Computer Science Journal of Moldova, vol.**25**, no.2(74), 2017

The Physical-mathematical Theory of Hyper-random Phenomena

Igor I. Gorban

Abstract

We give the survey of the researches dedicated to the statistical stability phenomenon and the physical-mathematical theory of hyper-random phenomena that takes into account the violation of statistical stability. It is presented the study technique of statistical stability, the results of the theoretical and experimental investigations of statistical stability of various processes, the mathematical apparatus of the theory of hyper-random phenomena, the particularities of mathematical statistics of hyper-random variables (including ones connected with the law of large numbers and the central limit theorem), and the explanation why the accuracy of actual measurements is limited. The description is constructed on the comparison of the theory of hyper-random phenomena with the probability theory.

Keywords: Phenomenon of statistical stability; Probability theory; Theory of hyper-random phenomena; Physical process; Violation of convergence.

1 Introduction

For description of mass physical phenomena in uncertainty conditions different mathematical and physical-mathematical theories are used.

Between these two types' theories there is essential difference: in mathematical theories the physical entity is ignored and in physical-mathematical ones it plays a key role. *Subject matter* of a mathematical theory is *abstract mathematical objects* and *subject matter* of a physical-mathematical theory is *physical phenomena* of the *real physical world.*

Classical probability theory based on Kolmogorov's axioms [1] is a typical example of mathematical theory, the *subject matter* of which is *abstract probability space.* Theories *exploiting uncertainty approach* and *approximate reasoning* (in particular, *imprecise probability theory* [2, 3], *interval analyses* [4], *interval probability theory* [5], *robust Bayesian analysis* [6, 7], *probability box theory* [8], *robust Neyman-Pearson theory* [9], *Huber's robust statistics* [10], etc.) are of *mathematical type* too (Table 1).

Table 1. Theories describing mass physical phenomena in uncertainty conditions

	Mathematical theories	**Physical-mathematical theories**
Probability approach	Probability theory as *mathematical discipline* (A.N. Kolmogorov)	Probability theory as *physical discipline* (D. Hilbert)
Uncertainty approach	*Mathematical theories* based on approximate reasoning	*Physical-mathematical theory of hyper-random phenomena*

Besides mathematical interpretation of the probability theory it is known the alternative one, the follower of which was David Hilbert. He and many other scientists regarded the probability theory as *physical discipline* [11]. Although the physical approach is less popular now among mathematicians but it is very popular among engineers and physicists. The *subject matter* of the physical-mathematical probability theory is *statistical stability of actual mass phenomena.*

The *probability theory* has the centuries-old development history. During this time it has established itself as the most powerful tool solving various statistical tasks. There is even opinion that any statistical problem can be effectively solved within the paradigm of the probability theory. However, as it turned out, it is not so.

Some conclusions of the probability theory do not accord to the experimental data. A typical example concerns the potential accuracy of measurement. If systematic error absents, according to the probability theory (*Cramer–Rao inequality* [12, 13]), with increasing of the number of measurement results of any physical quantity the error of the averaged estimator follows to zero. But every engineer or physicist knows that the actual measurement accuracy is always limited and to overcome the limit by the statistical averaging of the data is not possible.

Study of the causes of discrepancies between the theory and practice led to the understanding that the problem is related to the *unjustified idealization of the phenomenon of statistical stability.*

The modern probability theory regarded as a *physical-mathematical* one has mathematical and physical components [14]. The mathematical component is based on the A.N. Kolmogorov's classical axioms while the physical component is based on physical hypotheses, in particular the *hypothesis of ideal (perfect) statistical stability of actual events, variables, processes, and fields* assuming the *convergence of statistics* when the sample size goes to infinity.

The results of numerous experimental studies of various physical quantities and processes over long observation intervals have shown that the hypothesis of perfect statistical stability *is not confirmed experimentally.*

For relatively short temporal, spatial, or spatio-temporal observation intervals, an increase in data volume usually reduces the level of fluctuation in the statistics. However, when the volumes become very large, this tendency is no longer visible, and once a certain level is reached, the fluctuations remain practically unchanged or even grow. This indicates *a lack of convergence for real statistics (their inconsistency).*

If the volume of processing data is small, the violation of the convergence practically does not influence on the results, but if this volume is large, the influence is very significant.

The study of violations of statistical stability of physical processes and the development of an effective way for description of the actual world with taking into account such violations has resulted in the construction of the *physical-mathematical theory of hyper-random phenomena* (Table 1).

The *subject matter* of this theory as well as the *physical-mathematical probability theory* is *statistical stability of actual mass phenomena.* The *scope of study* of it is the *violation of statistical* stability among the characteristics and parameters of real physical phenomena.

The theory of hyper-random phenomena consists of physical and mathematical components. Mathematical component is based on the Kolmogorov's axioms and constructed on the scheme of the classical probability theory. However it accumulates knowledge obtained in the framework of a number of adjacent *mathematical theories exploiting approximate reasoning.*

Physical component of the theory is based on the hypotheses that essentially differ from the physical hypotheses of the physical-mathematical probability theory, in particular, on the *hypothesis of limited statistical stability* assuming the *absence of convergence of actual statistics.*

The theory of hyper-random phenomena began to develop at the end of the XX century. Quite a few scientific works concerning this theory are written. The publication list, in particular, includes eight monographs [14–21], two of which [14, 21] are written in English.

The *purpose* of this survey article is to *present main results concerning modern investigation of the phenomenon of statistical stability* and *to compare two approaches for its description* proposed by *probability theory* and *theory of hyper-random phenomena.*

In Sect. 2 we familiarize with the manifestations of the statistical stability phenomenon and two physical hypotheses: perfect and imperfect statistical stability.

Sect. 3 is devoted to description of Hilbert's sixth problem and its solution in the part of statistical stability proposed by the probability theory and the theory of hyper-random phenomena.

Sect. 4 presents the investigation technique of statistical stability on infinite and finite observation intervals as well as the results of theoretical researches of statistical stability of stochastic processes and the experimental investigations of statistical stability of actual processes of various physical nature.

Sect. 5 familiarizes with the *mathematical apparatus* used in the theory of hyper-random phenomenon for description of real physical events, quantities, processes, and fields in conditions of imperfect statistical stability. The mathematical apparatus is developed for hyper-random events, scalar and vector hyper-random variables, scalar, vector, stationary, and ergodic hyper-random functions, hyper-random differential equations, transformations of hyper-random variables and functions. The special part is devoted to mathematical statistics of hyper-random phenomena. We do not describe all these questions (they are presented in detail, in particular, in the monographs [14, 16, 17]). To obtain general representation about the developed mathematical approaches and main theoretical results we consider briefly only the *description of scalar hyper-random variables*, the particularities of mathematical statistics of hyper-random variables, notions of *generalized limit* and *convergence of sequences in the generalized sense*, *generalized law of large numbers*, and *generalized central limit theorem.*

Sect 6 concerns the *engineering* and *practical questions*. We describe here the *classic determinate–random measurement model* rested upon the probability theory and the *determinate–hyper-random one* based on the theory of hyper-random phenomena, compare these two models, and present estimation results of *potential measurement accuracy of physical quantities* calculated with using these models.

Note, from the issues presented in the article a special place occupies the questions concerning *experimental research of violation of statistical stability of actual processes*. The results of these investigations *give physical grounds* for correct using in practice not only hyper-random mathematical models but *other mathematical models based on approximate reasoning principles.*

2 The Physical Phenomenon of Statistical Stability

2.1 Manifestation of the Phenomenon of Statistical Stability

The statistical stability is manifested in stability of relative frequency of mass events. The first to draw attention to the phenomenon of statistical stability was the cloth merchant J. Graunt in 1662 [22]. Information about research on statistical stability is fragmentary for the period from the end of the XVII century to the end of the XIX century, e.g., by J. Bernoulli, S.D. Poisson, I.J. Bienayme, A.A. Cournot, L.A.J. Quetelet, J. Venn, etc. [23, 24].

Systematic study of statistical stability began at the end of the XIX century. In 1879, the German statistician W. Lexis made the first attempt to link the concept of statistical stability of the relative frequency with the dispersion [23]. At the turn of the century and in the early XX century, statistical stability was studied by C. Pearson, A.A. Chuprov, L. von Bortkiewicz, A.A. Markov, R.E. von Mises, and others [23, 24].

A lot of well known scientists led experimental investigations of the statistical stability of relative frequency of mass events. It is known, for example, that coin-tossing experiments were studied by P.S. de Laplace, G.L.L. de Buffon, K. Pearson, R.P. Feynman, A. de Morgan, W.S. Jevons, V.I. Romanovskiy, W. Feller, and others. At the first glance, this quite a trivial task per se was not presented for them.

A new stage of experimental research began in the late XX century. The necessity for additional studies is called due to the new applied tasks and the detection of a number of phenomena that can not be satisfactorily explained and described within the framework of the classical probability theory. The new tasks are, in particular, the *ultra-precise measurement* of physical quantities and *ultra-precise forecasting* of developments over large intervals of observation. To the relatively new phenomena can be led, for instance, an *unpredictable* measurement *progressive* (*drift*) *error* [25, 26], as well as a flicker noise [27], which is detected everywhere and *can not be suppressed by averaging the data.*

The phenomenon of statistical stability is manifested also in the stability of the average $y(t)$ of the process $x(t)$ and its sample mean $y_n = \frac{1}{n}\sum_{i=1}^{n} x_i$, where $x_1, \ldots, x_n$ are discrete samples of the process $x(t)$.

Interesting that this phenomenon occurs in case of averaging of the fluctuations that are of different types, in particular, of the *stochastic*, *determinate*, and *actual physical processes.*

Example 1. In Fig. 1**a** and Fig. 1**c** a realization of noise with uniform power spectral density (white noise) and a determinate periodical process are presented. In Fig. 1**b** and Fig. 1**d** the dependencies of the according averages on the averaging interval are shown. As can be seen from Fig. 1**b** and Fig. 1**d**, when the averaging interval increases, fluctuations in the sample mean decrease and the average value gradually stabilizes.

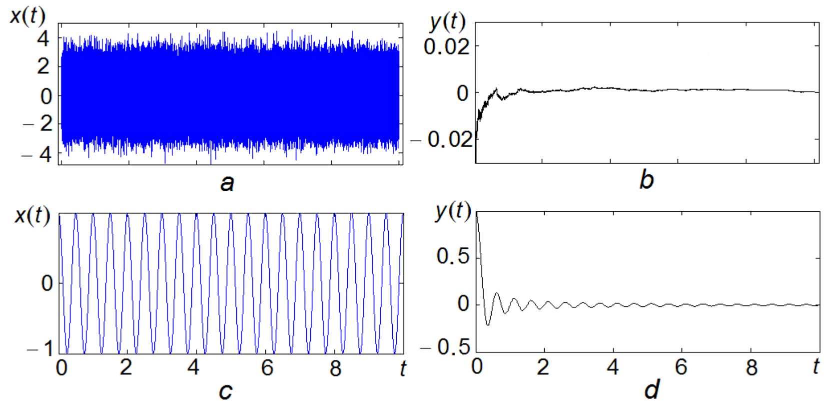


Figure 1. Realization of white Gaussian noise (**a**) and harmonic oscillation (**c**), together with the dependencies of the corresponding sample mean on the average interval (**b**, **d**)

Example 2. Fig. 2**a** and Fig. 2**b** show how the mains voltage in a city fluctuates quickly, while the average changes slowly. As the averaging interval increases from zero to one hour, the average voltage stabilizes (Fig. 2**b**).

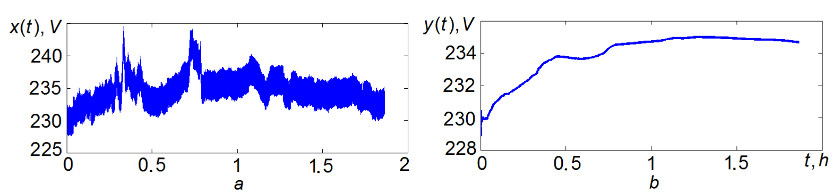


Figure 2. Dependence of the mains voltage (**a**) and the corresponding average (**b**) on time over 1.8 hours

The phenomenon of statistical stability is observed in calculation of other statistics too, in particular, the *sample standard deviation* $z_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - y_n)^2} \quad (n = 2, 3, \ldots)$.

2.2 The Hypothesis of Ideal Statistical Stability

Taking into account the statistical stability of the relative frequency of *actual physical events and actual statistics* it seems naturally to assume that if the number of the test n infinitely increases, the fluctuation level of the relative frequency $p_n(A)$ of *any actual event* A tends to zero, and also that in the unlimited increasing of the sample size n (increasing the observation time t) the fluctuation level of the sample mean y_n of any *random or real physical oscillation* $x(t)$ follows to zero too.

In other words, it is *possible to hypothesize* that there is a *convergence* of the sequence of the relative frequencies $p_1(A),\ p_2(A), \ldots$ of *any actual event* A to some *determinate value* $P(A)$ and there is a *convergence* of the sequence of averages $y_1,\ y_2, \ldots$ of *any stochastic or actual process* to *determinate value* m, viz. the limit of the relative frequency $\lim_{n\to\infty} p_n(A) = P(A)$, and the limit of the average $\lim_{n\to\infty} y_n = m$.

The modern probability theory is based on this *hypothesis of ideal* (*perfect*) *statistical stability* or, in other words, on the *assumption of convergence of statistics.*

The value $P(A)$ is interpreted in practice as the *probability* of the event A, and the value m is regarded as the *expectation* of the process $x(t)$.

2.3 The Hypothesis of Imperfect Statistical Stability

For many years it *was believed that the hypothesis of perfect statistical stability adequately reflects the reality.* Although some scholars (even the founder of axiomatic probability theory A.N. Kolmogorov [1, 28, 29] and such famous scientists as A.A. Markov [30], A.V. Skorokhod [31], E. Borel [32], V.N. Tutubalin [33], and others) noticed that, *in the real world, this hypothesis is valid only with certain reservations.*

Pay attention, the convergence of the relative frequency and other statistics is only a hypothesis. It does not follow from any experiments and any logical inferences. Not all processes, even of oscillatory type, have the property of perfect statistical stability.

Experimental studies of various processes of different physical nature over broad observation intervals show that the hypothesis of perfect statistical stability *is not confirmed*. The real world is continuously changing, and changes occur at all levels, including the statistical one. Statistical assessments formed on the basis of relatively small observation intervals are relatively stable. Their stability is manifested through a decrease in the fluctuation of statistical estimators when the volume of statistical data grows. This creates an *illusion* of perfect statistical stability. However, beyond a certain critical volume, the level of fluctuations remains practically unchanged (and sometimes even grows) when the amount of the data is increased. This indicates that the *statistical stability is not perfect*.

Example 3. Non-perfect statistical stability is illustrated in Fig. 3 [14] which presents mains voltage fluctuations over 2.5 days. Note, the fluctuation in Fig. 2**a** shows the beginning part of the fluctuation presented in Fig 3**a**. As can be seen from Fig. 3**b**, the sample average does not stabilize, even for very long averaging intervals.

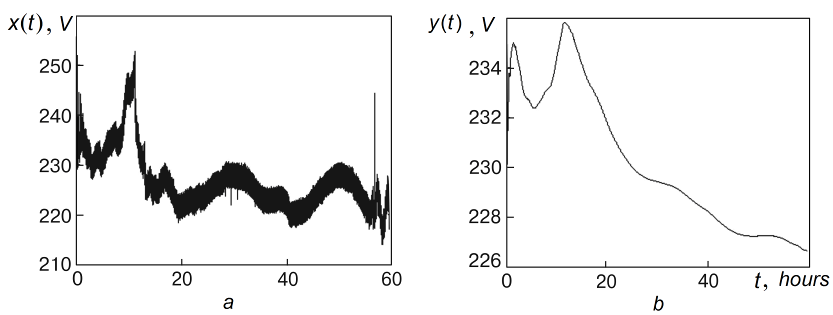


Figure 3. Dependence of the mains voltage (**a**) and the corresponding average (**b**) on time over 60 hours

3 Hilbert's Sixth Problem and Approaches for Its Solving

3.1 Description of the Phenomenon of Statistical Stability in the Framework of Probability Theory

Prior to the early twentieth century, probability theory was regarded as a *physical theory*, which described the phenomenon of statistical stability. Then at the beginning of the last century, the problem of axiomatizing probability theory was raised. In fact, David Hilbert formulated this as part of the problem of axiomatizing physics (the Hilbert's sixth problem) [11].

Many famous scientists tried to solve the problem and various approaches were proposed. Today, the most widely recognized approach is the set-theoretic one [1] developed by A.N. Kolmogorov in 1929 [34]. This approach has even been elevated to the rank of a standard [35].

The basic notion in Kolmogorov's probability theory is the notion of a *random event*. Random events are regarded as mathematical objects, described by means of a *probability space* defined as a triad $(\Omega, \Im, P)$, where Ω is the *space of elementary events* $\omega \in \Omega$, $\Im$ is a *σ-algebra of subsets of events* (*Borel field*), and P is a *probability measure on subsets of events.*

For any random event A the *probability* $P(A)$ is defined by the following three *axioms*:

1) the probability of any event A is a non-negative number, i.e. $P(A) \geq 0$;

2) for pairwise disjoint events A_1, $A_2, ...$ (both finite and *countable*), the probability of their union is the sum of the probabilities of the events, i.e. $P(\bigcup\limits_n A_n) = \sum\limits_n P(A_n)$;

3) the probability of the event Ω is equal unity (i.e. $P(\Omega) = 1$).

A *random variable* X is regarded as a measurable function defined on the space Ω of elementary random events ω, while a *random* (*stochastic*) *function* $X(t)$ is a function of an independent argument t, whose value is a random variable when this argument is fixed.

A *random phenomenon* is understood as a *mathematical object* (a random event, random variable, or stochastic function), which is exhaustively characterized by some specific probability distribution law. In particular, a random event is described exhaustively by the *probability*, a random variable X – by the *distribution function* $F(x) = P\{X < x\}$, where $P\{X < x\}$ is the probability of the inequality $X < x$, and a scalar random function $X(t)$ – by the *distribution function* $F(\vec{x};\vec{t}) = P\{X(t_1) < x_1, ..., X(t_L) < x_L\}$, where $\vec{x} = (x_1, ..., x_L)$ is the L-dimensional vector of values of the function $X(t)$ at times $t_1, ..., t_L$ represented by the L-dimensional vector $\vec{t} = (t_1, ..., t_L)$.

Note that a *phenomenon* or *mathematical model, not described by specific distribution law is not considered to be random.* This is an extremely important point that must be taken into account.

In probability theory, the probability of an event is a key concept. Note that, in Kolmogorov's definition, it is an *abstract mathematical concept.* Using a more visual statistical definition due to R. von Mises [36], the probability $P(A)$ of a random event A is interpreted as a limit of the relative frequency $p_N(A)$ of the event, when the experiments are carried out under identical statistical conditions and the number N of experiments tends to infinity. When N is small, the relative frequency $p_N(A)$ can fluctuate greatly, but with increasing N, it gradually stabilizes, and as $N \to \infty$, it tends to a definite limit $P(A)$.

All mathematical theories, including the version of probability theory based on Kolmogorov's axioms, are related to *abstract mathematical concepts* which *are not associated with the actual physical world.* In practice, these theories can be successfully applied if we admit certain *physical hypotheses* asserting the adequate description of real world objects by relevant mathematical models. For probability theory, such *physical hypotheses* are as follows [14]:

Hypothesis 1 *For mass phenomena occurring in the real world, the relative frequency of an event has the property of ideal (perfect) statistical stability, i.e., when the sample volume increases, the relative frequency converges to a constant value.*

Hypothesis 2 *Mass phenomena are adequately described by random models which are exhaustively characterized by distribution functions.*

It is often assumed that the hypothesis of perfect statistical stability is valid for any physical mass phenomena. In other words, a *stochastic concept of world structure is accepted*.

Kolmogorov's axioms with added Hypotheses 1 and 2 solve Hilbert's sixth problem in the part of axiomatizing of the probability theory as physical discipline.

3.2 Description of the Phenomenon of Statistical Stability in the Framework of Theory of Hyper-random Phenomena

In Sect. 2.3, attention was drawn to the fact that the experimental study of real physical phenomena over broad observation intervals does not confirm the hypothesis of perfect statistical stability (Hypothesis 1). For a correct application of the classical probability theory in this case, it is sufficient in principle to replace Hypothesis 1 by the following:

Hypothesis 1' *For real mass phenomena, the relative frequency of an event has the property of limited statistical stability, i.e., when the sample volume increases, the relative frequency does not converge to a constant value.*

The replacement of Hypothesis 1 by Hypothesis 1' leads to considerable mathematical difficulties due to the *violation of convergence*. There are different ways to overcome them. The development of one of these led to the *physical-mathematical theory of hyper-random phenomena* [14].

In classical probability theory, the basic mathematical entities are random events, random variables, and random functions. In the theory of hyper-random phenomena, the analogues of these basic entities are *hyper-random events*, *hyper-random variables*, and *hyper-random functions*, which are *sets* of *non-interconnected* random events, random variables, and stochastic functions, respectively, each regarded as a comprehensive whole.

A *hyper-random event* can be described by a tetrad $(\Omega, \Im, G, P_g)$, where Ω is a space of elementary events $\omega \in \Omega$, $\Im$ is a Borel field, G is a set of conditions $g \in G$, and P_g is a probability measure on subsets of events, depending on the condition g. Thus, the probability measure is defined for all subsets of events and all possible conditions $g \in G$. Note that the *measure for conditions* $g \in G$ *is not determined.*

Using a statistical approach, a hyper-random event A can be interpreted as an event whose relative frequency $p_N(A)$ *is not stabilized by growth of the number* N, *and which has no limit when* $N \to \infty$.

It is essential to understand that the hyper-random events, variables, and functions (*hyper-random phenomena*) are *many-valued objects exhaustively characterized by the sets of non-interconnected probability measures.* Hence,

- a hyper-random event is described exhaustively by the *collection of probabilities*;
- a hyper-random variable $X = \{X_g, \ \ g \in G\}$ is described exhaustively by the *collection of conditional distribution functions* $F(x/g)$ with conditions $g \in G$, forming the *many-valued distribution function* $\tilde{F}(x) = \{F(x/g), \ \ g \in G\}$[1], where $X_g = X/g$ is a random variable subject to the condition g, and the set G can be finite, countably infinite, or uncountable;
- a scalar hyper-random function $X(t) = \{X_g(t), \ \ g \in G\}$ is described exhaustively by the *collection of conditional multidimensional distribution functions* $F(\vec{x};\vec{t}/g)$ with conditions $g \in G$, forming the *many-valued distribution function* $\tilde{F}(\vec{x};\vec{t}) = \left\{F(\vec{x};\vec{t}/g), \ \ g \in G\right\}$, where $X_g(t) = X(t)/g$ is a random function subject to the condition g.

For correct use of the theory of hyper-random phenomena, one must also adopt the following hypothesis, in addition to Hypothesis 1'.

Hypothesis 2' *Mass phenomena are adequately described by hyper-random models which are exhaustively characterized by the sets of distribution functions.*

The assumption that these hypotheses are valid for a wide range of mass phenomena leads to *a world-building concept based on hyper-random principles.*

[1]A tilde under a letter indicates that the object described by the letter is many-valued.

So the *mathematical part of the theory of hyper-random phenomena* is based on the *classical axioms of probability theory*, and the *physical part* – on *Hypotheses 1' and 2'.*

Note, in contrast to the classical Kolmogorov's mathematical probability theory, the theory of hyper-random phenomena is *physical-mathematical* one. Its subject matter is phenomenon of statistical stability and the scope of research is adequate description of it by *hyper-random models* (*hyper-random phenomena*) taking into account the violation of statistical stability.

Since the mathematical part of the theory of hyper-random phenomena uses the system of mathematical axioms of probability theory, from the mathematical standpoint it is a *branch* of classical probability theory. But from the physical point of view, the theory of hyper-random phenomena is a new *physical theory* based on new physical hypotheses.

In general, *the theory of hyper-random phenomena* can be regarded as a *new physical-mathematical theory* constituting a *complete solution of Hilbert's sixth problem in the context of statistical stability.*

4 The Investigation of the Statistical Stability Violation

4.1 Formalization of the Statistical Stability Concept

Curiously enough is that the concept of statistical stability was not formalized until recent time. First of all note, a data statistically stable with respect to some statistics can be unstable with respect to other statistics. This means that the statistical stability is an attribute not only of a data, but also of the statistics. In addition, the level of statistical stability depends on the number of the data and on the sequence of this data.

It was proposed the number of parameters characterizing statistical stability violation. For the *random sequence* $X_1, \ldots, X_N$ most useful are parameters of statistical instability *with respect to the average* γ_N and *respect to the sample standard deviation* Γ_N described by the following expressions: $\gamma_N = \mathrm{E}\left[\bar{D}_{Y_N}\right]/\mathrm{E}\left[\bar{D}_{X_N}\right]$, $\Gamma_N = \mathrm{E}[\bar{D}_{Z_N}]/\mathrm{E}[\bar{D}_{X_N}]$, where $\bar{D}_{Y_N} = \frac{1}{N-1}\sum_{n=1}^{N}(Y_n - \bar{m}_{Y_N})^2$ is the sample variance of the fluctuations in the average $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ $(n = \overline{1,N})$, $\bar{m}_{Y_N} = \frac{1}{N}\sum_{n=1}^{N} Y_n$ is the sample mean of the average fluctuations, $\bar{D}_{Z_N} = \frac{1}{N-2}\sum_{n=2}^{N}(Z_n - \bar{m}_{Z_N})^2$ is the sample variance of the fluctuations in the sample standard deviation $Z_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - Y_n)^2}$ $(n = \overline{2,N})$, $\bar{m}_{Z_N} = \frac{1}{N-1}\sum_{n=2}^{N} Z_n$ is the average of the sample standard deviations, $\bar{D}_{X_N} = \frac{1}{N-1}\sum_{n=1}^{N}(X_i - Y_N)^2$ is the sample variance of the initial sequence.

The actual range of the parameters γ_N, Γ_N is $[0,\infty)$. The smaller the values of the parameters γ_N and Γ_N the more stable the sequence with respect to average and standard deviation respectively. Small values for large sample sizes N point to high statistical stability of the sequence, and large values point to statistical instability.

Random samples, for which the *parameters of statistical instability* γ_N and Γ_N do not follow to zero, are considered to be *statistically unstable with respect to the average and standard deviation* respectively.

Any measurement procedure consists in the comparison of the measurement result with some unit. For quantitative characterizing of the degree of instability, the *measurement units* are requested, the comparison with which would allow judging about the degree of instability in respect to the average and standard deviation. As the role of the *measurement unit*, a variable γ_{0N} can play, that is the parameter γ_N calculated for the *standard statistically stable sequence of uncorrelated samples of white Gaussian noise*.

The absolute level of statistical instability with respect to the average and standard deviation in units γ_{0N} characterize the parameters of the statistical instability h_N and H_N described by the following expressions: $h_N = \gamma_N/\gamma_{0N}$, $H_N = \Gamma_N/\gamma_{0N}$. The actual range of theses parameter is $[0,\infty)$. The measurement unit of them is the number

$h_{0N} = 1$ that does not depend on the sample size.

For solving of the practical tasks it is usually not important the behavior of statistics on the infinite observation interval, though it is laid in the basis for a formal definition of statistical stability. More important the behavior of statistics on the *actual observation interval*: the presence or absence of the *trends* indicating a violation of statistical stability. *If on the observation interval these trends are not tracing, the process can be considered as statistically stable, but otherwise as statistically unstable.*

Various statistics and processes, as a rule, have different *statistical stability intervals*. The concepts of the interval of statistical stability with respect to the average τ_{sm} and of the interval of statistical stability with respect to the standard deviation τ_{sd} can be formalized by the *statistical stability borders of the confidence intervals*.

For the parameters of the statistical instability γ_N and Γ_N the statistical stability upper border of the confidence interval is given by $\gamma_{0N}^{+} = \gamma_{0N} + \varepsilon\sigma_{\gamma_{0N}^{*}}$, where ε is the confidence parameter that determines the width of the confidence interval and $\sigma_{\gamma_{0N}^{*}}$ is the standard deviation of the variable $\gamma_{0N}^{*} = \bar{D}_{Y_N}/\mathrm{E}[\bar{D}_{X_N}]$ calculated for standard statistically stable sequence.

The *criteria of statistical stability violation with respect to the average and with respect to the standard deviation* (that determine the amounts of the *intervals of statistical stability* τ_{sm} and τ_{sd}) can be that the parameters γ_N and Γ_N go beyond the border γ_{0N}^{+} or the parameters h_N and H_N go beyond the border $h_{0N}^{+} = \gamma_{0N}^{+}/\gamma_{0N}$.

In practical work, due to the limited amount of data, instead of the parameters of statistical instability γ_N, h_N and Γ_N, H_N, we have to admit using of the appropriate estimates γ_N^{*}, h_N^{*} and Γ_N^{*}, H_N^{*}.

4.2 The Statistical Stability of Stochastic Processes

4.2.1 Dependence of the Statistical Stability on the Process's Spectrum

Studies show that the statistical stability of a stochastic sequence (process) with respect to the average and standard deviation is determined by its *spectrum.*

In particular, for the sequence $X_1, \ldots, X_N$ with zero expectation and power spectral density $S_{x_N}(k)$ the *parameter of statistical instability with respect to the average* γ_N when $N \to \infty$ is described by the following asymptotic formula:

$$\gamma_N = \frac{\sum_{k=2}^{N/2} \frac{1}{(k-1)^2} \left[\frac{\pi^2}{4} + (C + \ln(2\pi(k-1)))^2 \right] S_{x_N}(k)}{4\pi^2 \sum_{k=2}^{N/2} S_{x_N}(k)},$$

where k is the spectral sample number ($k = \overline{1, N}$), C is the *Euler–Mascheroni constant* ($C \approx 0.577216$).

4.2.2 Stochastic Processes Whose Spectrum is Described by a Power Function

In many cases, actual noise is well approximated by random processes whose power spectral density is described by a *power function* $1/f^{\beta}$ for various values of the shape parameter β, where f is frequency. Such noise sometimes is called a *color noise.* One thus speaks of violet, blue (cyan), white, pink, brown (red), and black noise that corresponds to $\beta = -2, -1, 0, 1, 2$, and > 2.

Flicker noise and *fractal* (*self-similar*) *processes* are other examples of the processes with power spectral density described by power functions.

Taking into account the prevalence of the processes with power spectral density described by a power function the research of their statistical stability was carried out.

Studies show that the *process with power spectral density described by a power function is statistically stable with respect to the average and with respect to standard deviation if* $\beta < 1$ *and statistically unstable if* $\beta \geq 1$.

Since the state of statistical stability of the process changes at the point $\beta = 1$, the process with this particular parameter value can be regarded as a *limiting unstable process with respect to average and standard deviation* (*limiting unstable in broad sense*).

Investigations show that *if* $\beta \leq 0$, *the process is more stable with respect to the average, than with respect to the standard deviation* ($\gamma_N < \Gamma_N$), and *if* $\beta > 0$, *on the contrary, it is less stable* ($\gamma_N > \Gamma_N$).

Summarizing these results it is possible to mark the following (see Fig. 4):

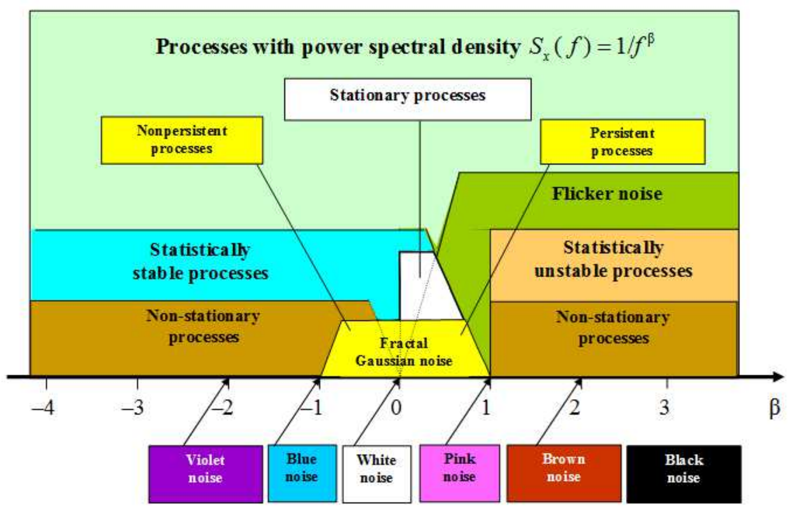


Figure 4. Processes with power spectral density described by a power function

- *statistically stable with respect to the average and standard deviation* (*stable in the broad sense*) are stationary processes, a part of the non-stationary processes, the so called fractal Gaussian noise, a part of the flicker noise, as well as violet, blue, and white noise;
- *statistically unstable with respect to the average and standard deviation* (*unstable in the broad sense*) are a part of the non-stationary processes, a part of the flicker noise, as well as pink, brown, and black noise.

4.2.3 Dependence of Statistical Stability on Other Process's Particularities

Investigation of dependence of statistical stability of the processes on the correlation of the samples shows that a *positive correlation* between the samples leads to a *decrease* in the statistical stability, and a *negative correlation,* to an *increase*.

Studies show [18] that violations of statistical stability occur not only in the case of low-frequency processes, but also for *narrowband stochastic processes* too.

Not only the non-stationary but *stationary in a narrow sense* stochastic processes can be statistically unstable in a broad sense. The statistically unstable processes, for example, are stationary stochastic processes, cut sets of which are described by distributions that do not have any moments or do not have moments higher than of the first order (processes described by Cauchy, Pareto, Fischer–Snedecor, Frechet, and et al. distributions).

Violation of statistical stability can have many causes. These include the *inflow into an open system of matter, energy, and* (*or*) *information* feeding non-equilibrium processes, various *nonlinear transformations*, *low-frequency linear filtering* of special type, etc. It is shown that, as the result of low-frequency filtration, broadband stationary and statistically stable noise can be transformed into a statistically unstable process.

4.3 The Results of Experimental Investigations of the Statistical Stability of Actual Processes of Various Physical Nature

To find out whether the *actual processes* are statistically stable or not, and if they are unstable, on the whole, but at what observation interval they can be considered as stable ones, various actual physical processes were studied over long observation intervals.

For instance, it is investigated the supply-line voltage. The active (effective) voltages were recorded in the computer memory and then analyzed. Recording sessions were conducted over two months, with

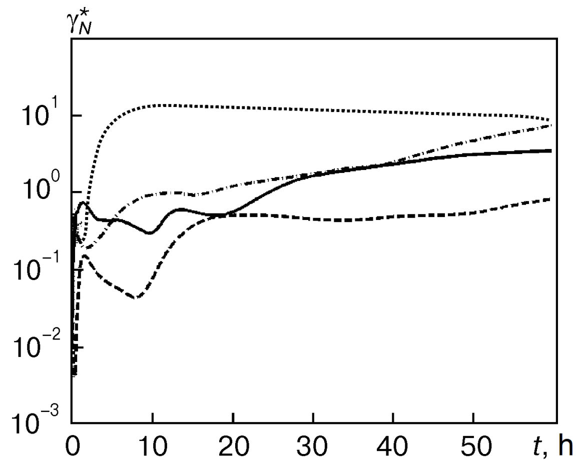


Figure 5. Variations in the estimate γ_N^* of the statistical instability parameter during a 60-hour observation of the mains voltage in four sessions

breaks of a few days. The duration of each session was about 60 hours. One of such records is presented in Fig. 3**a.** The estimates of the statistical instability parameter γ_N^* with respect to the average calculated for four sessions are shown in Fig. 5.

It follows from the figure that, for long observation times, the instability parameter does not show any tendency to fall to zero. Consequently, the *mains voltage is statistically unstable.* The *statistical stability interval with respect to average* τ_{sm} *of the mains voltage is approximately an hour.*

In the same mane it has been investigated statistical stability of a lot of various processes, in particular the Earth's magnetic field, the height and period of waves on the surface of the sea, the temperature and speed of sound in the ocean, the air temperature and atmospheric precipitation in different cities, exchange rates, the X-ray intensity of astrophysical objects, etc. [14, 17]. It has been found that all the processes have limited interval of statistical stability. Table 2 presents the estimation result of these intervals for some real processes.

All these estimates, except for the one in row 9, relate to statistical stability with respect to the average. The estimate in row 9 corresponds to statistical stability with respect to the standard deviation.

Table 2. Estimates of the statistical stability intervals for various real processes

No	Real process	Estimate of the statistical stability interval τ_s
1	Oscillations in the mains voltage	About 1 h
2	Currency rate oscillations	About 1 h
3	Height and period of sea surface waves	About half a day
4	Temperature and sound speed variations in the ocean	Ten hours
5	Radiation oscillations of astrophysical source Cygnus X-1	About a week
6	Variations of air temperature	Several weeks
7	Radiation oscillations of astrophysical source GRS 1915+105	About a month
8	Narrowband fluctuations of water temperature in the ocean with an average period from 2 to 10 hours	Several weeks
9	Radiation oscillations of pulsar PSR J1012+5307	Several months
10	Fluctuations in the wind speed in Chernobyl	Several months
11	Earth's magnetic field variations	Several months
12	Precipitation fluctuations	Many tens of years

It is important to note that *all the processes,* taken intentionally from different fields of knowledge, *are statistically unstable*. This allows us to suggest the following hypothesis: *all real physical phenomena are statistically unstable.* This physical hypothesis becomes the foundation for constructing of the mathematical part of the theory of hyper-random phenomena.

Note, the violation of statistical stability in the real world means that the *probability concept has no physical interpretation* [14, 17]. *Probability is thus a mathematical abstraction.*

5 The Mathematical Apparatus of the Theory of Hyper-random Phenomena

5.1 Scalar Hyper-random Variables

5.1.1 Conditional Characteristics

To describe the hyper-random variable $X = \{X_g,\ \ g \in G\}$, we use various probabilistic characteristics of the *conditional random variables* X_g ($g \in G$) such as the *conditional distribution functions* (Fig. 6) $F_{x/g}(x)$ and the *conditional probability density functions*[2] $f_{x/g}(x) = \frac{\mathrm{d}F_{x/g}(x)}{\mathrm{d}x}$.

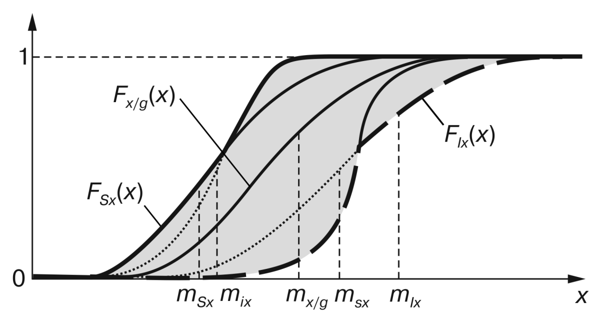


Figure 6. A set of conditional distribution functions $F_{x/g}(x)$ (thin lines) and the bounds of the distribution function $F_{Sx}(x)$, $F_{Ix}(x)$ (bold lines) of the hyper-random variable X

The most complete description of the hyper-random variable X gives its *many-valued distribution function* $\tilde{F}_x(x) = \{F_{x/g}(x),\ g \in G\}$.

[2]It is assumed here and below that all the above distribution functions are continuous or piecewise continuous.

A less complete description supplies the *conditional crude* and *central moments* of the *hyper-random variable* X (*conditional moments*), in particular, the *conditional expectation* $m_{x/g} = \mathrm{E}[X_g] = \int\limits_{-\infty}^{\infty} x\, f_{x/g}(x)\ \mathrm{d}x$, the *conditional variances* $D_{x/g} = \mathrm{E}[(X_g - m_{x/g})^2]$, the *conditional standard deviations* $\sigma_{x/g} = \sqrt{D_{x/g}}$, and others.

In this interpretation, the *expectation*, *variance*, and *standard deviation of the hyper-random variable* X are *many-valued values*, which is analytically described as follows: $\tilde{m}_x = \{m_{x/g},\quad g \in G\}$, $\tilde{D}_x = \{D_{x/g},\quad g \in G\}$, $\tilde{\sigma}_x = \{\sigma_{x/g},\quad g \in G\}$.

The *scalar hyper-random variables* X_1 and X_2 described by the respectively distribution functions $\tilde{F}_{x_1}(x) = \{F_{x_1/g}(x),\quad g \in G\}$ and $\tilde{F}_{x_2}(x) = \{F_{x_2/g}(x),\quad g \in G\}$ are said to be *equal in all conditions*, if under all conditions $g \in G$ for the same g their conditional distribution functions coincide: $F_{x_1/g}(x) = F_{x_2/g}(x)$.

5.1.2 Bounds of the Distribution Function and Their Moments

A general view of the hyper-random variable X is given by the *bounds of the distribution function* $F_{Sx}(x) = \sup\limits_{g \in G} F_{x/g}(x)$, $F_{Ix}(x) = \inf\limits_{g \in G} F_{x/g}(x)$ that are respectively the *upper* and *lower bounds* of probability that $X < x$ (see Fig. 6).

These *bounds can be considered as the distribution functions of some virtual random variables.* Between these bounds there is the *uncertainty area* (shaded area in Fig. 6).

The analogues of the probability density function of the *random variable* are the *probability densities functions of the bounds, viz.* $f_{Sx}(x) = \frac{\mathrm{d}\,F_{Sx}(x)}{\mathrm{d}\,x}$, $f_{Ix}(x) = \frac{\mathrm{d}\,F_{Ix}(x)}{\mathrm{d}\,x}$.

To describe a hyper-random variable, we may use the *moments of the bounds,* in particular, the expectations, variances, and standard deviations of the bounds, and so on.

The *expectations of the bounds* m_{Sx}, m_{Ix} of the hyper-random variable X are described by the formulas $m_{Sx} = \mathrm{E}_S\,[X] = \int\limits_{-\infty}^{\infty} x f_{Sx}(x)\mathrm{d}x$,

$m_{Ix} = \mathrm{E}_I\left[X\right] = \int\limits_{-\infty}^{\infty} x f_{Ix}(x)\mathrm{d}x$ (see Fig. 6).

For a *real* hyper-random variable X the *variances of the bounds* D_{Sx}, D_{Ix} are defined by $D_{Sx} = \mathrm{E}_S\left[(X - m_{Sx})^2\right]$, $D_{Ix} = \mathrm{E}_I\left[(X - m_{Ix})^2\right]$, and the *standard deviations of bounds* – by $\sigma_{Sx} = \sqrt{D_{Sx}}$, $\sigma_{Ix} = \sqrt{D_{Ix}}$.

The *scalar hyper-random variables* X_1 and X_2 described by the distribution functions $\tilde{F}_{x_1}(x)$ and $\tilde{F}_{x_2}(x)$ respectively, are said to be *equal* if their upper and lower bounds of the distribution coincide: $F_{Sx_1}(x) = F_{Sx_2}(x)$, $F_{Ix_1}(x) = F_{Ix_2}(x)$.

5.1.3 Bounds of the Moments

The *bounds of the moments* give a general view of the hyper-random variable X.

The upper and lower bounds of the expectation of the hyper-random variable X are the values $m_{sx} = \mathrm{E}_s[X] = \sup\limits_{g \in G} m_{x/g}$, $m_{ix} = \mathrm{E}_i[X] = \inf\limits_{g \in G} m_{x/g}$ (see Fig. 6).

The upper and lower bounds of the variance of the hyper-random variable X are the values $D_{sx} = \sup\limits_{g \in G} D_{x/g}$, $D_{ix} = \inf\limits_{g \in G} D_{x/g}$. The roots $\sigma_{sx} = \sqrt{D_{sx}}$, $\sigma_{ix} = \sqrt{D_{ix}}$ of these values are the *bounds of the standard deviation.*

In general, the operators $\mathrm{E}_s[\cdot]$, $\mathrm{E}_i[\cdot]$ do not coincide with the operators $\mathrm{E}_S[\cdot]$, $\mathrm{E}_I[\cdot]$, and the bounds of the expectation and variance m_{sx}, m_{ix}, D_{sx}, D_{ix} do not coincide with the expectations and variances of the bounds m_{Sx}, m_{Ix}, D_{Sx}, D_{Ix}.

5.2 Particularities of Statistics of Hyper-random Variables

5.2.1 A Hyper-random Sample

The concepts of mathematical statistics of the theory of hyper-random phenomena are based on the concepts of mathematical statistics of the probability theory.

The *entire assembly* (*general population*) *of the hyper-random variable* $X = \{X_g, \ g \in G\}$ is the *infinite set of all its determinate realizations* (*sample elements or components*) observed *under all conditions* $g \in G$. This set can be either countable or uncountable.

It implies from this definition that the general population of the hyper-random variable X is the union of the populations of all its random components $X_g, \ g \in G$.

The general population can be described by the *many-valued distribution function* $\tilde{F}_x(x)$ of the hyper-random variable X, the set of conditional distribution functions $F_{x/g}(x)$ $(g \in G)$, the upper and lower bounds of the distribution function $F_{Sx}(x)$, $F_{Ix}(x)$, the moments of the bounds, the bounds of the moments, and other characteristics.

A set of members of the general population

$$\vec{x} = (x_1, ..., x_N) = \{x_{1g}, \ldots, x_{Ng}, \ \ g \in G\} = \{\vec{x}_g, \ \ g \in G\}$$

of the hyper-random variable X obtained for a finite number N of experiments in different *fixed* or *non-fixed conditions* $g \in G$ is called the *sample of the population*, and its elements $x_1, ..., x_N$ or $x_{1g}, \ldots, x_{Ng}$ are called the *sampling values* or *realizations*.

Without specifying a condition g each sampling value x_n $(n = \overline{1, N})$ is a *set of determinate values* (*set of numbers*), and with specifying the condition g each sampling value x_{ng} is a *determinate value* (*number*).

Ones believe that the sample $x_1, \ldots, x_N$ belongs to the hyper-random variable $X = \{X_g, \ g \in G\}$ described by the conditional distribution functions $F_{x/g}(x)$, $g \in G$ if it is obtained from the general population described under condition g by the distribution function $F_{x/g}(x)$.

Infinite set of the samples $\vec{x} = (x_1, ..., x_N)$ of a volume N taken from a general population without specifying of a condition g forms N-dimensional hyper-random vector

$$\vec{X} = (X_1, ..., X_N) = \{X_{1g}, \ldots, X_{Ng}, \ \ g \in G\} = \left\{\vec{X}_g, \ \ g \in G\right\},$$

called *hyper-random sample* and the infinite set of samples $\vec{x}_g = (x_{1g}, \ldots, x_{Ng})$ of the volume N taken from this general population under condition g forms N-dimensional *random vector* (*random sample*) $\vec{X}_g = (X_{1g}, \ldots, X_{Ng})$.

Generally one believes that all elements of hyper-random vector are described by the same many-valued distribution function $\tilde{F}_x(x)$ and each component $X_{ng} \quad (n = \overline{1, N})$ of the random vector $\vec{X}_g$ corresponding to the specific condition g is described by the same single-valued distribution function $F_{x/g}(x)$ (or probability density function $f_{x/g}(x)$).

Ones usually assume that the components X_n of the hyper-random sample $\vec{X}$ are *mutually independent under all conditions*. Then the conditional distribution function $F_{\vec{x}/\vec{g}}(\vec{x})$ of the hyper-random sample $\vec{X}$ under conditions $g \in G$ factorizes: $F_{\vec{x}/g}(\vec{x}) = \prod_{n=1}^{N} F_{x/g}(x_n)$.

In the theory of hyper-random phenomena a *statistics* is *any function* of the *hyper-random sample* $\vec{X}$, *random sample* $\vec{X}_g$ under a fixed condition $g \in G$, *determinate many-valued sample* $\vec{x}$ or *determinate single-valued sample* $\vec{x}_g$ under a fixed condition $g \in G$.

5.2.2 Evaluations of Characteristics and Parameters of a Hyper-random Variable

Using the general population of a hyper-random variable theoretically it is possible to calculate various its *exact determinate characteristics and parameters*, such as the conditional distribution functions $F_{x/g}(x)$, bounds of distribution function $F_{Sx}(x)$, $F_{Ix}(x)$, conditional expectations $m_{x/g}$, expectations of bounds m_{Sx}, m_{Ix}, bounds of expectation m_{sx}, m_{ix}, conditional variances $D_{x/g}$, variances of bound D_{Sx}, D_{Ix}, bounds of variance D_{sx}, D_{ix}, and so on.

Using certain statistics of realizations of the hyper-random variable it is possible to calculate *approximate evaluations* of the same characteristics and parameters, in particular the evaluations of conditional distribution functions $F^*_{x/g}(x)$, bounds of distribution function $F^*_{Sx}(x)$, $F^*_{Ix}(x)$, conditional expectations $m^*_{x/g}$, expectations of bounds m^*_{Sx}, m^*_{Ix}, bounds of expectation m^*_{sx}, m^*_{ix}, conditional variances $D^*_{x/g}$, variances of bound D^*_{Sx}, D^*_{Ix}, bounds of variance D^*_{Sx}, D^*_{Ix}, and so on.

If the sample is *hyper-random*, then the evaluations are the *hyper-*

random estimators, if it is *determinate*, then the evaluations are *determinate estimates*.

The estimates can be made in several steps. First, samples $x_{1g}, \ldots, x_{Ng}$ are formed separately for each condition $g \in G$. Using samples $\vec{x}_g = (x_{1g}, \ldots, x_{Ng})$ for all $g \in G$, one then calculates the conditional characteristic and parameter estimates, in particular, estimates of the conditional distribution functions $F^*_{x/g}(x)$, estimates of the conditional expectations $m^*_{x/g}$, estimates of the conditional variances $D^*_{x/g}$, and others.

From the conditional distribution functions $F^*_{x/g}(x)$ for all $g \in G$, one can calculate estimates of the distribution function bounds: $F^*_{Sx}(x) = \sup_{g \in G} F^*_{x/g}(x)$, $F^*_{Ix}(x) = \inf_{g \in G} F^*_{x/g}(x)$, and estimates of the parameters describing these bounds: estimates m^*_{Sx}, m^*_{Ix} of the expectations of the bounds, estimates D^*_{Sx}, D^*_{Ix} of the variances of the bounds, and so forth.

Using estimates of the conditional variables, one can calculate estimates of the corresponding variable bounds, for example, estimates of the expectation bounds $m^*_{sx} = \sup_{g \in G} m^*_{x/g}$, $m^*_{ix} = \inf_{g \in G} m^*_{x/g}$, estimates of the variance bounds $D^*_{sx} = \sup_{g \in G} D^*_{x/g}$, $\quad D^*_{ix} = \inf_{g \in G} D^*_{x/g}$, etc.

When applying this technique, *certain difficulties can be expected in the first stage*, when the samples $\vec{x}_g$ for all $g \in G$ are formed, because at first glance, it is difficult to control and maintain the conditions g. The situation *is facilitated* by the facts that a lot of actual samples are possessed of *ergodic property* and the calculation of a number of characteristics *do not require information about the specific conditions under which the conditional characteristics have been obtained.*

Most important that, in the sample formation phase, all possible conditions g of the set G are represented, and for every fixed condition g in the sample $\vec{x}_g$, *only the data corresponding to this condition g is used.*

Typically, for actual phenomena occurring in the real world, in the case of a broad observation interval, the latter requirement can be easily provided, because, although the conditions often vary continuously, they vary sufficiently slowly, and it is possible to evaluate the maximum number of elements N_s for which the conditions can be treated as practically constant.

Therefore one can collect data on a broad observation interval (that is essentially larger than N_s) without taking care about what the statistical conditions are at any given time and in what way they alternate, and then one can separate the resulting data into a number of fragments containing N_s consistent elements. Using these fragments, which represent the variable under different statistical conditions g, one can then calculate the required estimates. The main requirement for this technique is to collect the data for all possible observation conditions in G.

Of course a number of questions arise. What are the conditions under which the hyper-random evaluations converge to the exact characteristics and parameters? What are types of these parameters and characteristics? What are their distribution laws? The *generalized law of large numbers* and the *generalized central limit theorem* help to obtain answer to these and other questions.

To understand this material, one should be familiar with some mathematical concepts, such as the *generalized limit* and the *convergence of sequences in the generalized sense.*

5.3 Generalized Limit and the Convergence of Sequences in the Generalized Sense

5.3.1 Generalized Limit

According to classical concepts, the *numerical sequence* $x_1, x_2, ..., x_n$ is considered as a *convergent sequence* if there is a limit $a = \lim_{n\to\infty} x_n$. If the limit exists, then it is *unique.* The sequence which has not the limit is considered as a *divergent sequence.*

From every infinite sequence one can form the set of *partial sequences* (*subsequences*) derived from the original sequence by discarding part of its members, while *maintaining the order of the remaining members.*

It is proved that when the sequence converges, all its partial sequences converge too. If the sequence diverges, then all its partial sequences do not necessary diverge. Some of them can converge to certain limits (*limit points*). The set of all limit points a_m, $m = 1, 2, \ldots$ of the sequence $x_1, x_2, ..., x_n$ also called *partial limits*, form the *spectrum of limit points* $\tilde{S}_x$.

The spectrum of limit points $\tilde{S}_x$ is a generalization of the limit concept on any sequence, including divergent. If the sequence converges, the spectrum of the limit points consists of a single element (number), and if it is divergent, it consists of a set of numbers. The spectrum of limit points can be described by the expression $\tilde{S}_x = \operatorname*{LIM}_{n\to\infty} x_n$, where, unlike the conventional limit $\lim\limits_{n\to\infty}$ the symbol of the *generalized limit* $\operatorname*{LIM}_{n\to\infty}$ is used.

This expression can be interpreted as the *convergence of the sequence to the spectrum of limit points*. The spectrum may be discrete, continuous, or mixed (discrete-continuous). If the spectrum forms a continuous interval, they say that the *sequence converges to the interval*.

A divergent sequence can be characterized by not only the spectrum of limit points, but also by a *set* (in general) *of the measures* described by the *many-valued* (in general) *distribution function of the limit points* $\tilde{F}_x(x) = \operatorname*{LIM}_{n\to\infty} \frac{m_n(x)}{n}$, where $m_n(x)$ is the number of terms of the sequence $x_1, x_2, ..., x_n$ that are less than x.

If the sequence converges in the usual sense to the number a, the distribution function of limit points is described by the unique distribution function $F_x(x)$ in the form of a unit step function at the point a (Fig. **7a**) (then the measure equals to one at the point a and zero at all other points).

If the sequence diverges (converges to the set of numbers (in the particular case converges to the interval)), the distribution function is either a single-valued non-decreasing function $F_x(x)$ that differs from the unit step function (Fig. **7b**), or a many-valued function $\tilde{F}_x(x)$ (Fig. **7c**). Note that the *special case of hyper-random variable is the interval variable*, the distribution function of which is described by a rectangle of unit height (Fig. **7d**).

Using the terminology of the theory of hyper-random phenomena,

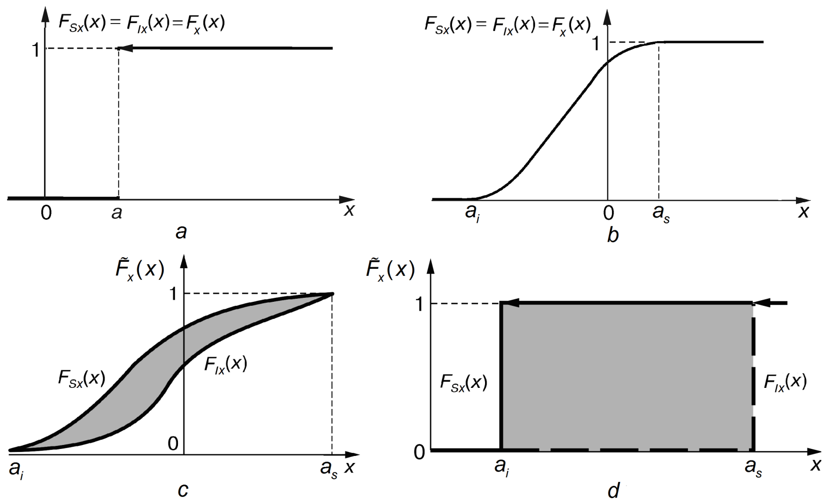


Figure 7. Single-valued $F_x(x)$ (**a**, **b**) and many-valued $\tilde{F}_x(x)$ (**c**, **d**) distribution functions of the limit points and their bounds $F_{Ix}(x)$, $F_{Sx}(x)$ for sequences converge to the number a (**a**) and to the interval $[a_i,\ a_s]$ (**b**–**d**)

we can say that the *spectrum of the limit points of a numerical sequence* can be

- a *number* (interpreted by the set of real numbers with the unit measure at the point $x = a$ and zero measure at all other points) (Fig. 7**a**),
- a *random variable* (Fig. 7**b**),
- a *hyper-random variable* (Fig. 7**c**) (in the degenerated case an *interval variable* (Fig. 7**d**)).

In other words, the numerical sequence may converge to a *number* or to a *set of numbers* (in the particular case to an *interval*). If it converges to a set of numbers, the spectrum of limit points may be either a *random variable* or a *hyper-random variable*.

5.3.2 Convergence of Sequences of Hyper-random Variables in the Generalized Sense

By analogy with the convergence of a *sequence of random variables*, in the theory of hyper-random phenomena the concept of *convergence* (*in generalized sense*) of a *sequence of hyper-random variables* is introduced. There is convergence *in distribution function, in mean-square, almost surely* (*with probability one*), *and in probability* (*in some measure*).

Consider the convergence of the sequence in the generalized sense in probability and in distribution function.

Suppose we have a sequence of hyper-random variables X $= \{X_1, ..., X_N\}$ and a hyper-random variable X, where $X_n = \{X_{ng},\ g \in G\}$ $(n = \overline{1, N})$ and $X = \{X_g,\ g \in G\}$. For all $X_1, ..., X_N$ and X, there are distribution functions $\tilde{F}_{x_1}(x) = \{F_{x_1/g}(x),\ g \in G\}, \ldots, \tilde{F}_{x_N}(x) = \{F_{x_N/g}(x),\ g \in G\}$ and $\tilde{F}_x(x) = \{F_{x/g}(x),\ g \in G\}$.

Then the sequence of hyper-random variables X *converges in the generalized sense* to the hyper-random variable X *in probability* $(P(|X_N - X| > \varepsilon) \to 0)$ if for all conditions $g \in G$ and $\varepsilon > 0$, when $N \to \infty$, $P(|X_{Ng} - X_g| > \varepsilon) \to 0$, i.e., for all $g \in G$, the random sequence $X_{1g}, ..., X_{Ng}$ converges in probability to the random variable X_g.

The sequence of hyper-random variables X *converges in the generalized sense* to the hyper-random variable X *in distribution* $(\tilde{F}_{x_N}(x) \to \tilde{F}_x(x))$ if for each point x, where $F_{x/g}(x)$ is continues, for all conditions $g \in G$, when $N \to \infty$, $F_{x_N/g}(x) \to F_{x/g}(x)$.

As in the case of the sequences of random variables, convergence in distribution is *weaker* than convergence in probability, i.e. the sequence of hyper-random variables that converges in probability converges in distribution too. The converse is not always true.

It follows from the definitions that, as well as a numerical sequence, the *hyper-random sequence* can converge to a *number* (determinate variable, the distribution function of which is a unit step function), to a *random variable* or to a *hyper-random variable.* It is obvious that a *random sequence* can also converge to a *number*, to a *random variable* or to a *hyper-random variable*.

5.4 Generalized Law of Large Numbers

Several variants of the law of large numbers for *random sequences* are known. Let us dwell on one of them formulated and proved by P.L. Chebyshev in 1867.

Chebyshev theorem. Let $X_1, \ldots, X_N$ be a sequence of pairwise independent random variables with expectations $m_1, \ldots, m_N$ and bounded variances. Then, when the sample size N goes to infinity, the average $m^*_{xN} = \frac{1}{N} \sum_{n=1}^{N} X_n$ of the sample values $X_1, \ldots, X_N$ tends in probability to the average $m_{xN} = \frac{1}{N} \sum_{n=1}^{N} m_{x_n}$ of the expectations $m_1, \ldots, m_N$: $\lim_{N \to \infty} P\{|m^*_{xN} - m_{xN}| > \varepsilon\} = 0$.

In typical for the probability theory interpretation the law of large numbers consists in that the average m^*_{xN} converges in probability to some number m_x that is a conventional limit of the average m_{xN} of the expectations $m_{x_1}, \ldots, m_{x_N}$.

The analysis of the proof of this assertion (which we will not present herein) shows that in the proof *it is not applied the assumption* that the average m^*_{xN} of the random samples and the average m_{xN} of the expectations have the conventional limits. This means that the sequences $\{m^*_{xN}\} = m^*_{x1}, \ldots, m^*_{xN}$ and $\{m_{xN}\} = m_{x1}, \ldots, m_{xN}$ *may not have limits in the conventional sense*, i.e. the sequences *may be divergent*.

But if they do not converge in the conventional sense, they *can converge in the generalized sense to the many-valued variables*: to random or hyper-random ones.

Hereafter, following the above mentioned agreement concerning designations of single-valued and many-valued variables and functions, the *single-valued* limits of the sequences $\{m^*_{xN}\}$ and $\{m_{xN}\}$ we shall denote by m^*_x and m_x, and a *many-valued* ones by the same manner but with tilde above: $\tilde{m}^*_x$ and $\tilde{m}_x$.

Whether the considered limits are single-valued or many-valued, according to the law of large numbers, when the sample size N increases, the *sample mean* m^*_{xN} *gradually approaches the average of the expectations* m_{xN}.

When $N \to \infty$, there are two possibilities:

Case 1 The variable m^*_{xN} converges to the *single-valued* average of the expectations m_x (number).

Case 2 The variable m^*_{xN}, becoming a many-valued variable $\tilde{m}^*_x$ in the limit, converges in the general sense to a *many-valued variable* $\tilde{m}_x$.

Case 1 is the idealized case considered in probability theory. In this case, the limit m_x of the average of the expectations is described by the distribution function $F_{m_x}(x)$, which is a unit step function at the point m_x. The distribution function $F_{m^*_{xN}}(x)$ of the sample mean m^*_{xN} tends to it when $N \to \infty$ (see Fig. 8**a**).

Case 2 is more realistic. Here the limit sample mean $\tilde{m}^*_x$ and the limit average of the expectations $\tilde{m}_x$ are described respectively by the many-valued spectra $\tilde{S}_{m^*_x}$ and $\tilde{S}_{m_x}$. In this case there may be two variants:

Case 2.1 The limit of the sample mean $\tilde{m}^*_x$ and the limit of the average expectations $\tilde{m}_x$ are *variables of random type.* Then the spectra $\tilde{S}_{m^*_x}$ and $\tilde{S}_{m_x}$ are characterized by the *single-valued distribution functions* $F_{m^*_x}(x)$ and $F_{m_x}(x)$ (see Fig. 8 **b**).

Case 2.2 The limit of the sample mean $\tilde{m}^*_x$ and the limit of the average expectations $\tilde{m}_x$ are *variables of hyper-random type.* Then the spectra $\tilde{S}_{m^*_x}$ and $\tilde{S}_{m_x}$ are characterized by the *many-valued distribution functions* $\tilde{F}_{m^*_x}(x)$ and $\tilde{F}_{m_x}(x)$ (see Fig. 8**c**).

Since the *convergence in distribution* of a sequence of *random variables* is weaker than the *convergence in probability*, in Case 2.1, the limit distribution function $F_{m^*_x}(x)$ coincides with the limit distribution function $F_{m_x}(x)$.

For *hyper-random variables*, convergence of the sequence *in distribution* is also weaker than convergence *in probability.* Therefore, in Case 2.2, the limit distribution function $\tilde{F}_{m^*_x}(x)$ coincides with the limit distribution function $\tilde{F}_{m_x}(x)$. In this case, the lower bound $F_{Im^*_x}(x)$ of the limit distribution function $\tilde{F}_{m^*_x}(x)$ coincides with the lower bound $F_{Im_x}(x)$ of the limit distribution function $\tilde{F}_{m_x}(x)$, and the upper bound

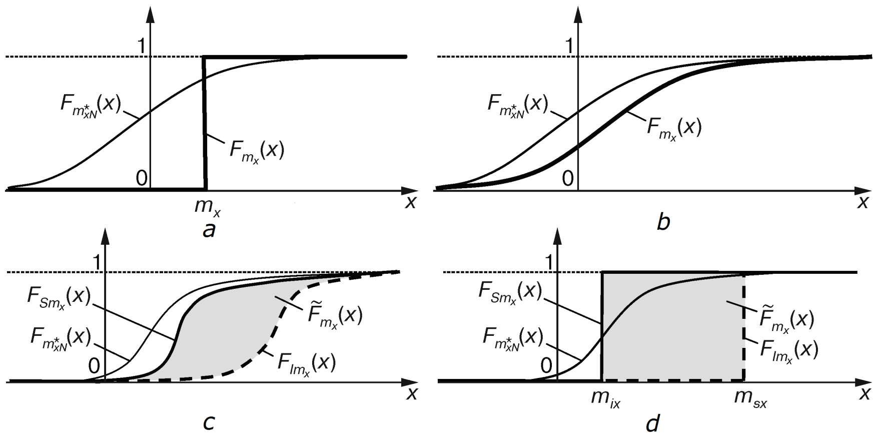


Figure 8. Formation of the limit distribution function $\tilde{F}_{m_x^*}(x)$ of the sample mean in the case of a random sequence: the limit sample mean and the limit average of expectations are a number (**a**), a random variable (**b**), and a hyper-random variable (**c**, **d**) (**c** is the general case and **d** is a special case)

$F_{Sm_x^*}(x)$ of the limit distribution function $\tilde{F}_{m_x^*}(x)$ coincides with the upper bound $F_{Sm_x}(x)$ of the limit distribution function $\tilde{F}_{m_x}(x)$.

The uncertainty area located between the specified bounds is shown in Fig. 8**c** by the shaded area. It is proved that, if the distribution function describing the spectrum of the sequence of averages of determinate values is many-valued, then the corresponding uncertainty area is continuous. So the *uncertainty area of the distribution function* $\tilde{F}_{m_x}(x)$ *is continuous.*

The interval in which the sample mean m_{xN}^* fluctuates when $N \to \infty$ is described by the lower bound m_{ix}^* when the function $F_{Sm_x^*}(x)$ begins to rise from zero and the upper bound m_{sx}^* when the function $F_{Im_x^*}(x)$ reaches unity. Naturally, these bounds coincide with the corresponding bounds m_{ix}, m_{sx} of the functions $F_{Sm_x}(x)$, $F_{Im_x}(x)$: $m_{ix}^* = m_{ix}$, $m_{sx}^* = m_{sx}$. These bounds can be either finite or infinite.

Note that Case 2.2 includes the special case when the limit sample mean $\tilde{m}_x^*$ and the limit average of the expectations $\tilde{m}_x$ are of *interval type* (Fig. 8**d**).

Systematizing the results of the present section, we may note that the sample mean m_{xN}^* of a random sample can converge to a *number* m_x (finite or infinite) or *fluctuate within a certain interval* $[m_{ix}, m_{sx}]$.

In the latter case, we shall say that there is *convergence of the sample mean to the interval.* Theoretically the limit of the sample mean $\tilde{m}_x^*$ and the limit average of the expectations $\tilde{m}_x$ can be *numbers*, *random variables*, *intervals*, or *hyper-random variables.* The spectra $\tilde{S}_{m_x^*}$ and $\tilde{S}_{m_x}$ can be *numbers* or *intervals.* The limit distribution functions $F_{m_x^*}(x)$ and $F_{m_x}(x)$ can be of *unit step type*, *single-valued functions*, or *many-valued functions with a continuous uncertainty area.*

Convergence of the sample mean to a number is not corroborated by the experiments and convergence to an interval is corroborated by a lot of them. We shall return to the question concerning the type of the limit distribution function after study of the generalized central limit theorem.

5.5 Generalized Central Limit Theorem

In the probability theory it is known the *central limit theorem.* There are many variants of it. One of them can be formulated with some simplification by the following manner.

Lindeberg-Feller theorem. Let $X_1, \ \ldots, X_N$ be, in general, a non-uniform random sample with mutually independent terms described by distribution functions $F_{x_n}(x)$ with expectations m_{x_n} and variances D_{x_n} $(n = \overline{1, N})$. We assume the so called *Lindeberg condition.* Then the distribution function $F_{m_{xN}^*}(x)$ of the sample mean m_{xN}^* *converges uniformly* to a *Gaussian distribution function* $F(x/m_{xN}, D_{xN}) = \Phi\left((x - m_{xN})/\sqrt{D_{xN}}\right)$ with expectation $m_{xN} = \frac{1}{N}\sum_{n=1}^{N} m_{x_n}$ and variance $D_{xN} = \frac{1}{N^2}\sum_{n=1}^{N} D_{x_n}$, viz.

$$\lim_{N\to\infty} F_{m_{xN}^*}(x) = \lim_{N\to\infty} F(x/m_{xN}, D_{xN}), \tag{1}$$

where $\Phi(x)$ is *Laplace function*.

According to this theorem, with increasing of the sample size the random variable m^*_{xN} becomes a *Gaussian random variable*.

Using the technique devised to obtain (1), a more general statement can be proved: if the conditions specified in Lindeberg-Feller theorem are satisfied, the difference between the distribution function $F_{m^*_{xN}}(x)$ of the sample mean m^*_{xN} and the Gaussian distribution function $F(x/m_{xN}, D_{xN})$ converges uniformly to zero

$$\lim_{N\to\infty}\left[F_{m^*_{xN}}(x) - F(x/m_{xN}, D_{xN})\right] = 0. \tag{2}$$

There is a significant difference between (1) and (2). The expression (1) implies that the sample mean m^*_{xN} has a single-valued limit distribution function $F_{m^*_x}(x)$ to which the distribution function $F_{m^*_{xN}}(x)$ tends when $N \to \infty$, and there is a *single-valued* Gaussian limit distribution function $F_{m_x}(x) = F(x/m_x, D_x)$ to which the distribution function $F(x/m_{xN}, D_{xN})$ tends, where m_x and D_x are the expectation and the variance of the limit distribution function, respectively.

The formula (2), on the other hand, allows the given limit distribution functions to be *many-valued*. The many-valuedness of the limit distribution function to which the function $F(x/m_{xN}, D_{xN})$ tends is stipulated by the many-valuedness of the expectation and (or) variance. Therefore, in the expression $\tilde{F}_{m_x}(x) = \tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ representing the limit distribution function of the average of the expectations, the many-valued parameters $\tilde{m}_x$ and $\tilde{D}_x$ appear. In general these parameters are hyper-random variables. Therefore the function $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ is a hyper-random function. It can be interpreted as a set of single-valued Gaussian distribution functions. Each of these is described by a single-valued expectation $m_x \in \tilde{m}_x$ and variance $D_x \in \tilde{D}_x$.

The relation $F_{m^*_{xN}}(x) \to \tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ follows from (2), implying that there is convergence in distribution of the sequence of determinate functions $F_{m^*_{xN}}(x)$ to the hyper-random function $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$. In other words, the many-valued limit distribution functions $\tilde{F}_{m^*_x}(x)$, $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ are described by identical sets of single-valued conditional distribution functions.

When $\tilde{m}_x = m_x$ and $\tilde{D}_x = D_x$ (i.e. the both parameters are numbers) and $D_x = 0$, the limit Gaussian distribution function $F_{m_x}(x) = F(x/m_x, D_x)$ is the unit step function shown in Fig. 8**a** by the bold line; and when $\tilde{m}_x = m_x$ and $\tilde{D}_x = D_x$ are numbers but $D_x \neq 0$, this distribution function is described by the single-valued Gaussian curve shown in Fig. 8**b** by the bold line.

When the limit expectation $\tilde{m}_x$, the limit variance $\tilde{D}_x$ or both these parameters are many-valued variables, the limit distribution function $\tilde{F}_{m_x}(x)$ is a many-valued function. In Fig. 8**c**, **d**, it is displayed by the shaded areas.

Note, the analogues results concerning the law of large numbers and the central limit theorem are generalized on hyper-random sequences.

5.6 Experimental Study of the Convergence of the Sample Mean

The theoretical research presented in Sects. 5.4 and 5.5 indicates that with increasing of the sample size the sample means are *not necessarily normalized* (i.e. they do not necessarily take on the Gaussian character) and *tend to a certain fixed value*. This result is quite different from the conclusion of the classical probability theory. It raises a very important question: how do the actual sample means behave?

To answer this question, we return to investigation of the mains voltage oscillations (see Fig. 3a) and present some results of additional experimental studies of the process.

Studies consisted in calculation and analysis of the estimates of the distribution functions of the voltage fluctuations $F_g^*(x)$ on adjacent observation intervals, each lasting about one hour ($g = \overline{1, 64}$) (Fig. 9**a**), and the estimate of the distribution function of the sample mean $F^*_{m^*_{xN}}(x)$ (Fig. 9**b**).

The curves of the distribution functions $F_g^*(x)$ corresponding to different values of the parameter g differ essentially from one another (primarily by their location) (see Fig. 9**a**), and this confirms the claimed *nonstationarity* of the oscillations.

The calculation results of the estimate of the sample mean distri-

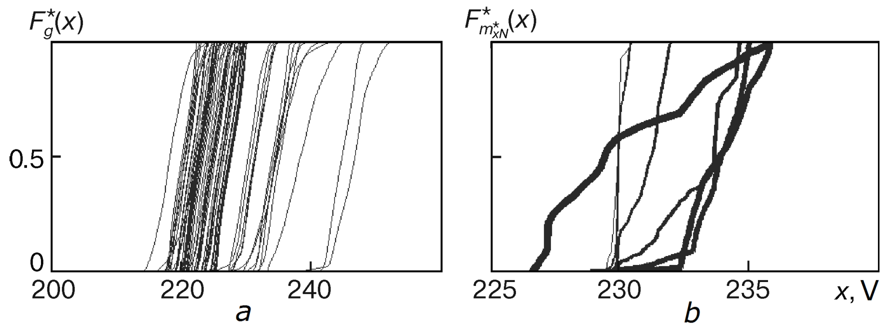


Figure 9. Estimates of the distribution functions of mains voltage oscillations on 64 contiguous observation intervals (**a**) and estimates of the distribution function of the voltage sample mean $F^*_{m^*_{xN}}(x)$ for sample sizes $N = 2^r$, $r = 8, 10, 12, 14, 16, 18, 20$ (**b**) (the line thickness increases with the value of the parameter r)

bution function $F^*_{m^*_{xN}}(x)$ for exponentially growing sample size (see Fig. 9**b**) show that $F^*_{m^*_{xN}}(x)$ *does not tend to a certain limit distribution function* $F_{m_x}(x)$, and the sample mean m^*_{xN} *does not tend to a certain limit value* m_x.

On the basis of the curves for the estimate of the distribution function of the sample mean $F^*_{m^*_{xN}}(x)$ for small values of the parameter r (8 and 10) (see Fig. 9 **b**), we may with some level of skepticism conclude that it is tending to a Gaussian distribution with decreasing variance, as probability theory would predict. However, for large values of r (starting from 10 to 20), the assumed trend is not confirmed.

When the sample size increases, the variance of the sample mean m^*_{xN} sometimes increases (for values of r from 8 to 14 and from 18 to 20) and sometimes decreases (for r from 14 to 18). In general, as one moves from small to large sample sizes, the variance does not manifest any tendency to go to zero, as would have been predicted by probability theory (see Fig. 8**a**), but in fact increases, even by a significant factor (the range of the sample mean increases approximately from $1V$ to $8V$).

It follows from these results that the *distribution function of the sample mean tends to a many-valued function* $\tilde{F}_{m_x}(x)$ *of general form* (see Fig. 8**c**).

Studies of the distribution functions of the sample means of a lot of processes show that when the data volume is large *there is not the aspiration* of the estimate $F^*_{m^*_{xN}}(x)$ of the distribution function of the sample mean to any specific distribution law, and more so to a Gaussian distribution with variance that tends to zero.

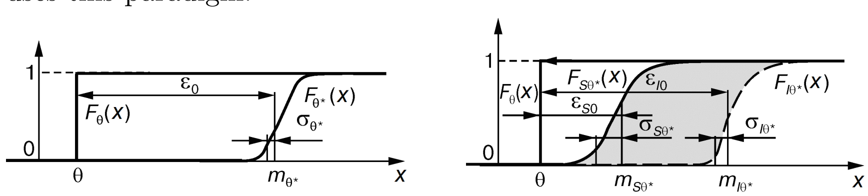
Thus, the experimental studies of the actual physical processes show that in case of a *small data volume* ones observe the *trends of normalization and stabilization of the sample means* and in case of a *large amount of data* such tendencies *are not fixed*.

The changing in the character of the behavior of the sample means can be explained by a violation of statistical stability of the actual processes on large observation intervals. These disorders lead to restriction of the accuracy of measurement and prediction of real physical quantities.

6 Accuracy and Measurement Models

6.1 Measurement Models

Any measurement is based on some models. It is usually suggested that the *measurand* (*measurement quantity*) *has a determinate character, while its estimator is random*. Modern classical measurement theory uses this paradigm.



a b

Figure 10. The classical determinate-random (**a**) and determinate–hyper-random (**b**) measurement models

When measuring a scalar quantity, the measurand θ can be represented by a unit step distribution function $F_\theta(x)$ and the measurement result Θ^* – by a distribution function $F_{\theta^*}(x)$ (Fig. 10**a**). Such a measurement model may naturally be referred to as *determinate–random*.

The foundations of this model were laid out by *Galileo Galilei*, who introduced the *concepts of systematic and random errors.* The systematic error is described by the bias of the estimator $\varepsilon_0 = m_{\theta^*} - \theta$ and the random error is often presented by the standard deviation σ_{θ^*} of the estimator Θ^*.

Modern metrology is based on the following hypotheses: the ideal value of a physical quantity is determinate, single-valued, and is not changed during the measurement time; the measure does not change its characteristics during the measurement; the statistical conditions are constant during the measurement time; and the result of a concrete measurement is unique.

All of these items, to put it mildly, not very reasonable. All actual physical objects and physical quantities describing them are subjected to change over time (except perhaps some universal constants). *Everything is changed: the object of measurement* (*measurand*), *the measure, and the measurement conditions.*

Any measurement is carried out not instantaneously, but over some time interval. Therefore the *measurement result is an average value representing over this interval the various states of the measuring object, the different states of the measure, and different measurement conditions.*

Of course, it is very convenient to represent the measurand by a determinate, unique, and unchanging value, and the measurement result – by a random variable. But this primitive model does not reflect many nuances of the real situation.

The theory of hyper-random phenomenon proposes different hyper-random mathematical models, taking into account some of them. *Determinate–hyper-random model* (Fig. 10**b**) describes, for instance, the measurand by a determinate model and the estimator by a hyper-random variable. In the figure, $F_{S\theta^*}(x)$ and $F_{I\theta^*}(x)$ are the upper and lower bounds of the distribution function of the hyper-random estimator Θ^*; $\varepsilon_{I0} = m_{S\theta^*} - \theta$ and $\varepsilon_{I0} = m_{I\theta^*} - \theta$ are the biases of the upper and lower bounds of the distribution function of the hyper-random estimator with respect to the measurand; $m_{S\theta^*}$, $m_{I\theta^*}$ are the expectations of the upper and lower bounds of the hyper-random estimator; and $\sigma_{S\theta^*}$, $\sigma_{I\theta^*}$ are the standard deviations of the appropriate bounds of the hyper-random estimator. The uncertainty area of the hyper-random estimator is shown by the shaded area.

6.2 Comparison of the measurement models

In the *determinate–random measurement model*, the error has a *random nature.* It is described by *systematic and random components*, and characterized by two parameters: the *bias* ε_0 and the *standard deviation of the estimator* σ_{θ^*} (Fig. 10**a**). In the *determinate–hyper-random measurement model*, the error has a *hyper-random nature.* It has an *uncertainty area* and is described by four parameters $\varepsilon_{S0}, \varepsilon_{I0}, \sigma_{S\theta^*}, \sigma_{I\theta^*}$ defining the location and size of the uncertainty area on the error axis (Fig. 10**b**).

Techniques of statistical measurement according to the comparing models are well known. They are described for instance in [14, 18]. Here we do not describe them and present only the calculation results for the parameters characterizing the mains voltage (Fig. 3**a**) at the end of 100-second and 60-hour observation intervals (Fig. 11).

The left side of the figure is obtained with using the classic determinate–random measurement model based on probability theory and the right side presents the parameters obtained with using the determinate–hyper-random measurement model based on the theory of hyper-random phenomena (except the parameter marked with a thin arrow).

For the 60-hour observation interval, the sample range and the range of the sample mean are obtained from the data of Figs. 3**a** and 9**b**. The confidence interval (THRP) marked with a bold arrow and the estimate (THRP) are calculated using the technique of the theory of hyper-random phenomena. The confidence interval (PT) marked by a thin arrow is calculated using the classic technique of the probability theory.

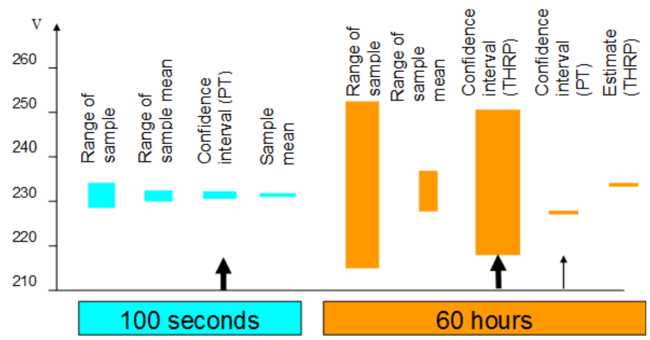


Figure 11. Calculation results for the parameters characterizing the city mains voltage over 100-second and 60-hour observation intervals, using calculation techniques based on probability theory (PT) and the theory of hyper-random phenomena (THRP)

The results shown in the figure for the 100-second and 60-hour observation intervals differ considerably. The parameters on the left side of the figure reflect the state of the electrical supply network under the *specific statistical conditions* that occurred during the relevant 100-second observation interval. The parameters on the right side (except for the one marked by a thin arrow) represent the state of the network for the *varying set of statistical conditions* that succeeded one another *unpredictably* during the relevant 60-hour observation period. The parameter marked by a thin arrow characterizes the state of the network for the *set of different but very specific statistical conditions* that succeeded each other over the same 60-hour period of observation.

For the *100-second observation interval, the most informative parameter is the confidence interval calculated using the classic technique of probability theory,* and for the *60-hour interval, it is the confidence interval calculated using the technique based on the theory of hyper-random phenomena* (in Fig. 11 these parameters are marked by two bold arrows).

For the 60-hour observation interval, the confidence interval, with width 50 mV and average value 229.4 V (calculated in accordance with probability theory and indicated in the figure by a thin arrow) is *not informative*, because it takes into consideration the *concrete sequence* of changes in the conditions which, in the next 60-hour observation interval, is likely to be something quite different. The confidence interval, with width 33 V and average 233.5 V (calculated using the theory of hyper-random phenomena and marked by a thick arrow) contains useful practical information about the average dynamics of the voltage changes in the power supply.

The loss of useful information in the first case and the fact that it is kept in the second arise because, when there are violations of statistical stability, the classic determinate–random measurement model reflects the real situation with considerable distortion, while the determinate–hyper-random measurement model is able to present it adequately. It follows from the above example that, ignoring the violation of statistical stability can lead to absurd results, and in particular, to an unjustifiable overstatement of measurement accuracy estimators by factors of hundreds or more.

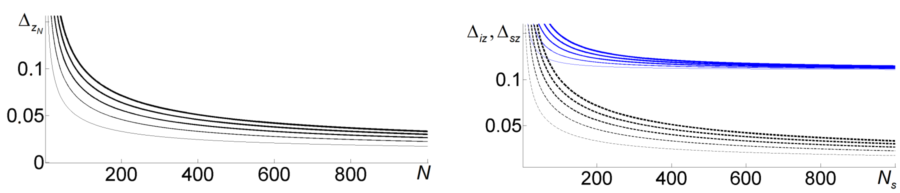
The conclusion is obvious: *when statistical stability is violated, the determinate–random measurement model and the measuring techniques based on it cannot be used.* In this case, *other models and measurement techniques must be used,* and in particular the *determinate–hyper-random measurement model and techniques based on it, which take into consideration the violations of statistical stability.*

6.3 Potential Measurement Accuracy

In case of determinate–random measurement model the error $\Delta_{z_N} = \sqrt{\varepsilon_0^2 + \sigma_{\theta_N^*}^2}$ is determined by the bias ε_0 and the standard deviation of the estimator $\sigma_{\theta_N^*}$. With increasing sample size N, theoretically this magnitude tends to the square of the bias ε_0^2. Let the estimator Θ_N^* be the average of the sample $(X_1^{,}..., X_N)$ and the sample elements are independent and have identical variance D_x. Then the variance of the estimator $\sigma_{\theta_N^*}^2 = D_x/N$ and the error is described by the expression $\Delta_{z_N}^{=} \sqrt{\varepsilon_0^2 + D_x/N}$. The dependence of the magnitude Δ_{z_N} on the defining parameters is shown in Fig. 12**a**.

It is clear from the figure that, when $N \to \infty$, the error tends to the bias ε_0 (the systematic error). If the bias ε_0 is *negligible*, the magnitude Δ_{z_N} is in inverse proportion to the root of the sample size N. It follows from this that, theoretically, by increasing N, the accuracy of the measurement can grow without limit, and as $N \to \infty$, it should become *infinitely large*.

Probability theory does not give a satisfactory explanation as to why, at low bias, an ultra-high measurement accuracy cannot be achieved by statistical processing of a large number of real data. The explanation of this effect gives the theory of hyper-random phenomena.



a b

Figure 12. Dependence of the error Δ_{z_N} (**a**) and the error bounds Δ_{iz}, Δ_{sz} (**b**) on the sample size N and the variance D_x accordingly for the determinate–random and the determinate–hyper-random measurement models. In case **a** $\varepsilon_0 = 0.01$ and in case **b** $\varepsilon_0 = \varepsilon_{S0} = 0.01$, $\Delta\varepsilon_0 = 0.1$. Thicker lines correspond to large values of the variance $D_x = 0.2;\ 0.4;\ 0.6;\ 0.8;\ 1$

Suppose the measurand θ is determinate and the estimator Θ^* is a hyper-random variable. The elements of the hyper-random sample $(X_1, \ldots, X_N)$ are independent. The statistical conditions change slowly and this allows us to divide the observation interval into G fragments of identical length corresponding to nearly constant statistical conditions. The elements of the sample are taken with uniform step. In any fragment, the number of samples is N_s.

The distribution law of the random elements $X_{1g}, \ldots, X_{N_s g}$ under the fixed condition g is fixed. Under different conditions g, the distribution laws of the elements are different, however all of them have the same variance D_x and differ from each other only in the expectation value. Then the error bounds can be written as $\Delta_{iz} = \sqrt{\varepsilon_i^2 + D_x/N_s}$, $\Delta_{sz} = \sqrt{\varepsilon_s^2 + D_x/N_s}$, where $\varepsilon_i^2 = \inf_{g \in G}[\varepsilon_{0/g}^2]$ and $\varepsilon_s^2 = \sup_{g \in G}[\varepsilon_{0/g}^2]$ are the lower and upper bounds of the square of the bias.

The dependence of the error bounds Δ_{iz}, Δ_{sz} on the defining parameters is shown in Fig. 12**b**. The dotted lines represent the lower error bounds and the solid ones the upper error bounds.

It is clear from the figure that, with increasing sample size N_s, the upper bound of the error Δ_{sz} tends to $\varepsilon_s = \varepsilon_0 + \Delta\varepsilon_0$ (ε_0 is the systematic error and $\Delta\varepsilon_0$ is the length of the uncertainty area).

Therefore, even if we make the unlikely assumption that the value N_s tends to infinity, the determining upper bound of the error Δ_{sz} will never be less than the value $\Delta\varepsilon_0 \neq 0$. When the bias ε_0 is negligible, the magnitude $\Delta_{sz} \to \Delta\varepsilon_0 \neq 0$.

So with the determinate–hyper-random measurement model, we can explain the *inability in practice to achieve infinitely high accuracy*, even with an unlimited amount of data.

7 Conclusions

Summing up the consideration of the issues it is drawn attention to the following key points.

1. *Statistical stability is a physical phenomenon* manifested in stability of *relative frequency of the actual mass events, sample means* and *other statistics*.

2. There are two theories describing statistical stability phenomenon: the *probability theory* and the *theory of hyper-random phenomena*. The *probability theory* is based on the assumption that the *phenomenon of statistical stability is perfect* (*statistics are converged* and *estimators are consistent*). The *theory of hyper-random phenomena* is based on the assumption that the *phenomenon of statistical stability is not perfect* (*statistics are not converged* and *estimators are not consistent*).

3. Numerous experimental studies of real phenomena of different physical nature indicate that *statistics obtained from actual samples do not demonstrate the tendency to convergence*. The trend towards convergence is observed only when the sample volume is small. In case of large sample volume such trend is not registered.

4. The violation of convergence of the relative frequency of actual events implies that the *probability, the basic concept of the probability theory is an abstract mathematical concept* that *does not have a physical interpretation*.

5. It is formulated and proved for *divergent sequences* the *generalized law of large numbers* and the *generalized central limit theorem*.

6. The results of the experimental studies conform the opinions of some scholars (including A.N. Kolmogorov, A.A. Markov, A.V. Skorokhod, E. Borel, and others) that the hypothesis of perfect statistical stability is valid in the actual world only in certain reservations. Apparently, the actual world really is obeyed to three *types of laws*: *determinate*, *statistically predicted* (*random, stochastic or otherwise probabilistic*), and *statistically unpredictable*.

7. For the small sample size the influence of statically unpredictable laws does not reflect essentially on the results of the measurement of physical quantities. This gives possibility to use the classical models and statistical methods of probability theory in a lot of important cases. *For the large sample size when the violation of statistical stability manifests itself clearly, the using of classical stochastic models leads to unacceptably large measurement errors*. Then the *hyper-random models have obvious advantages over the stochastic models*.

8. The hyper-random models, unlike the random ones theoretically can be used both in case of large and small observation intervals as in large and small samples. However, the hyper-random

models are more complicated. Therefore *for not very large sample sizes the stochastic models are preferred.* The using of the hyper-random models is justified when the stochastic models do not provide an adequate description of the reality.

9. The *limited accuracy of any statistical measurement of actual physical quantities* and the *limited accuracy of the temporal progress forecasting of actual events* can be explained by the presence of a statistically unpredictable laws.

10. The limited nature of statistical stability suggests that it may be necessary to review the postulates of a number of physical disciplines, in which the probability concept and convergence play a key role, in particular, *statistical mechanics, statistical physics, and quantum mechanics.* Taking into account statistical stability violations may lead to new scientific results that will be interesting for both theory and practice.

References

[1] A.N. Kolmogorov, *Fundamentals of Probability Theory*, Moscow: ONTI, 1974, 119p. (in Russian).

[2] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall, 1991, 706 p. ISBN: 0-412-28660-2.

[3] M. Beer, S. Ferson, and V. Kreinovich, "Imprecise Probabilities in Engineering Analysis," *Mechanical Systems and Signal Processing*, vol. 37, no. 1–2, pp. 4–29, 2013.

[4] R.E. Moor, *Interval Analyses*, Englewood Cliffs, NJ, United States: Prentice–Hall, 1966, 159 p.

[5] K. Weichselberger, "The theory of interval probability as a unifying concept for uncertainty," *International Journal of Approximate Reasoning*, vol. 24, no. 2–3, pp. 149–170, 2000. DOI: 10.1016/S0888-613X(00)00032-3.

[6] J.O. Berger, "An overview of robust Bayesian analysis (with discussion)," *Test*, vol. 3, pp. 5–124, 1994.

[7] P. Walley, "A bounded derivative model for prior ignorance about a real-valued parameter," *Scandinavian Journal of Statistics*, vol. 24, pp. 463–483, 1997.

[8] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz, "Constructing probability boxes and Dempster–Shafer structures," SAND, SAND report 2002–4015, 2003, 143 p.

[9] E.L. Lehmann, "The Neyman-Pearson theory after fifty years," In *Selected Works of E.L. Lehmann*, J. Rojo, Ed. Springer, 2012, pp. 1047–1060. ISBN: 978-1-4614-1412-4.

[10] P.J. Huber, E.M. Ronchett, *Robust Statistics*, Wiley, 2009, 380 p. ISBN: 978-0-470-12990-6.

[11] *Hilbert's Problems*, P.S. Aleksandrov, Ed. Moscow: Nauka, 1969, 240 p. (in Russian).

[12] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. 1, N.Y.: Wiley, 2004, 716 p. ISBN: 978-0-471-46382-5.

[13] I.I. Gorban, *Probability Theory and Mathematical Statistics for Scientists and Engineers*, Kiev: IMMSP NAS of Ukraine, 2003, 244 p. ISBN: 966-02-2664-0, DOI: 10.13140/2.1.5035.9366. (in Ukrainian).

[14] I. I. Gorban, *The Statistical Stability Phenomenon*, Springer, 2017, 362 p. ISBN: 978-3-319-43584-8. DOI: 10.1007/978-3-319-43585-5.

[15] I.I. Gorban, *Theory of Hyper-random Phenomena*, Kiev: IMMSP, NAS of Ukraine, 2007, 184 p. ISBN: 978-966-02-4367-5. DOI: 10.13140/2.1.2414.4967. (in Russian).

[16] I.I. Gorban, *The theory of Hyper-random Phenomena: Physical and Mathematical Basis*, Kiev: Naukova dumka, 2011, 318 p.

ISBN: 978-966-00-1093-2. DOI: 10.13140/2.1.2480.0328. (in Russian).

[17] I.I. Gorban, *The Phenomenon of Statistical Stability*, Kiev: Naukova dumka, 2014, 444 p. ISBN: 978-966-00-1422-0. (in Russian).

[18] I.I. Gorban, *Randomness and Hyper-randomness*, Kiev: Naukova dumka, 2016, 288 p. ISBN: 978-966-00-1561-6. DOI: 10.13140/RG.2.1.1545.5124. (in Russian).

[19] B.M. Uvarov, Yu.F. Zinkovskiy, *Design and Optimization of Mechanically Stable Radioelectronic Equipment with Hyper-random Characteristics*, Lugansk: LNPU, 2011, 180 p. (in Ukrainian).

[20] B.M. Uvarov, Yu.F. Zinkovskiy, *Optimization of Stability for Thermal Influences of Radioelectronic Equipment with Hyper-random Characteristics*, Lugansk: LNPU, 2011, 212 p. (in Ukrainian).

[21] I.I. Gorban, *Randomness and Hyper-randomness*, Springer, 2018, 268 p. ISBN: 978-3319607795.

[22] J. Graunt, *Natural and Political Observations Made Upon the Bills of Mortality*, Baltimore, 1939.

[23] O.B. Scheinin, "Probability Theory. Historical Review," 2009. [Online]. Available: http://www. sheynin.de. Accessed 21 June 2009. (in Russian).

[24] Y.V. Chaykovskiy, *About Random Nature*, Moscow: Centre for System Research, Institute of the History of Nature and Technique of the RAS, 2004, 280 p. (in Russian).

[25] A.G. Sergeev, V.V. Krokhin, *Metrology*, Moscow: Logos, 2001, 408 p. ISBN: 5-94010-039-2. (in Russian).

[26] P.E. Elyasberg, *Measuring Information. How Much Is Needed?*, Moscow: Nauka, 1983, 208 p. (in Russian).

[27] G.P. Zhigalskiy, "Nonequilibrium $1/f^{\gamma}$ noise in conducting films and contacts," *Physics–Uspekhy*, vol. 46, no. 5, pp. 449–471, 2003.

[28] A.N. Kolmogorov, "Probability theory," In: *Mathematics, its Content, Methods and Importance* (vol. 2), Moscow: USSR Academy of Sciences Publishing House, 1956, ch. 9, pp. 252–284. (in Russian).

[29] A.N. Kolmogorov, "About logical foundations of probability theory," In: *Probability theory and mathematical statistics*, Moscow: Nauka, 1986, pp. 467–471. (in Russian).

[30] A.A. Markov, *Calculus of Probability*, Moscow, 1924. (in Russian).

[31] V.I. Ivanenko, V.A. Labkovsky, *Uncertainty Problem in the Tasks of Decision Making*, Kiev: Naukova dumka, 1990, 135 p. (in Russian).

[32] E. Borel, "Sur les probabilités dénombrables et leurs applications arithmétiques," *Rend. Circ. Mat. Palermo*, vol. 26, pp. 247–271, 1909.

[33] V.N. Tutubalin, *Probability Theory*, Moscow: Moskow university, 1972, 230 p. (in Russian).

[34] A.N. Kolmogorov, "General measure theory and calculation of probability," in *Proceedings of Communist Academy. Mathematics*, 1929, pp. 8–21.

[35] *Statistics. Vocabulary and symbols. Part I: general statistical terms and terms used in probability*, ISO 3534–1:2006, 2006-10.

[36] R. Mises, *Mathematical Theory of Probability and Statistics*, London: Academic Press, 1964.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Igor I. Gorban
Received April 23, 2017

The Institute of Mathematical Machines and Systems Problem,
National academy of sciences of Ukraine,
Glushkova ave., 42, Kiev, Ukraine, 03680
E–mail: igor.gorban@yahoo.com