

# REFLECTIONS ON DUAL NATURE OF RISK. TOWARD A FORMALISM

Alexander Bochkov



JSC NIIAS, Moscow, Russia  
[a.bochkov@gmail.com](mailto:a.bochkov@gmail.com)

## Abstract

*We seem to know almost everything about risk and, at the same time, nothing. By focusing on the etymology of the word "risk", researchers have neglected its nature, causes and characteristics. At the same time, risk manifests itself differently in different situations and can be both a characteristic of a random event and a characteristic and measure of the quality of a process carried out over time. In the latter case, risk is inherent in the properties of a wave process, which requires the search for measures other than probabilistic ones to measure and assess it. This paper attempts to summarize the most characteristic different manifestations of risk and to propose a way of assessing risk that takes account of these differences. The paper can be seen as an invitation to debate the nature of risk and how its formalism should be constructed.*

**Keywords:** risk, random, Bayesian estimates, quality, anti-potential, difficulty of achieving, measure of disorderliness

## I. Introduction

Corpuscular-wave dualism (or quantum-wave dualism) is a property of nature consisting in the fact that material microscopic objects may under some conditions exhibit properties of classical waves and under other conditions - properties of classical particles. Typical examples of objects exhibiting dual corpuscular-wave behavior are electrons and light; the principle is also valid for larger objects, but, as a rule, the more massive the object, the less its wave properties are manifested (we are not talking here about the collective wave behavior of many particles, such as waves on the liquid surface). The idea of corpuscular-wave dualism was used during development of quantum mechanics to interpret phenomena observed in microcosm in terms of classical concepts. Quantum objects are neither classical waves, nor classical particles, exhibiting properties of the former or the latter only depending on conditions of experiments that are conducted on them. Corpuscular-wave dualism is unexplainable within classical physics and may be interpreted only in quantum mechanics.

Similarly, risk is understood by different researchers as either an event or a process. This difference in perception is partly dictated by the basic theoretical background of the researcher.

For example, specialists in the field of mathematical statistics and probability theory prefer to perceive risk as an event, because in this approach it is possible to apply the rich tools developed in this field of knowledge to describe this event in some space of states and to obtain a quantitative measure for subsequent comparison of different assessments. In this approach, risk is considered in terms of the probability of the event itself, and the expected damage from this event for the risk taker. Essentially, the expected damage (that is, what the subject is willing to risk) is reduced by a fraction proportional to the probability of the risk event. The most obvious example of this approach is insurance against frequent accidents (for example, traffic accidents, common diseases, failures of technical devices and mechanisms, etc.). In this case, the researcher has representative

statistics that he can process (analyze) and obtain the necessary distributions and analytical dependences for estimating. This also includes attempts to relate risk to uncertainty. This again introduces the concept of uncertainty, the rules for modeling it, and ways of determining the probability associated with risk. In this way, a probabilistic space is defined, in which risk is dealt with. At the same time, it is often argued that both concepts, uncertainty and probability, are basic and, therefore, no definitions can be given to them. Any such axiomatic attitude is based on set theory. It uses the basic concept of a set. There is no definition of a set (like any other basic concept in any theory). A set can be informally described as a set of objects having some common features (i.e., the concept of "set" here is expressed through the concept of "set").

Specialists, if I may say so, with basic engineering training, gravitating towards applied mathematical knowledge, consider risk as a process (in the limit - a wave process), which has points of maximum and minimum impact, at which, respectively, the phenomena of resonance (mutual strengthening or weakening) of different risk factors are possible. Under this approach, risk is seen as a measure of failure to achieve the goal set by the risk-taking subject, that is, it is an assessment of the quality of organization of a purposeful process of some activity. For a risk event to occur, a certain moment in time. As an example, we can take the task of removing a large vessel which happened to be on a shoal at low tide. It is extremely expensive to get it out at high tide, but it costs almost nothing if you know about tidal cycles and just wait for the right moment. Similar events occur infrequently, and to react to them adequately, one needs not analysis (because there is often almost nothing to analyze), but risk synthesis. That is, it is necessary to analyze all information known about the place of the expected occurrence of a risk event, as well as about the object that can be exposed to risk, and then perform risk synthesis and assess the possibility to achieve the goal (which will be the maximum risk avoidance). That is, risk, depending on the totality of factors and characteristics of the risk object (and risky subject) itself, manifests itself in two ways. And its analysis and assessment, respectively, should be carried out with this duality (wave-event duality of risk assessment), without substituting one assessment apparatus for another. The blind transfer of probability analysis to the field of risk synthesis leads to disastrous consequences (although, on a close planning horizon, it can provide estimates that are acceptable to the risk-taker).

The dual nature of risk leads to eight basic concepts that encompass the modern view of risk as an event assessment and a process assessment.

The first group is the concepts in which risk characterizes the event:

- **Risk as a relative value** (risk is defined as the ratio of the probability of an outcome in an exposed group to the probability of an outcome in an unexposed group).
- **Risk as a consequence of the occurrence of some random event** from a possible family of all events, or a set of possible damages in some stochastic situation and its probability (this concept covers the so-called frequency, statistical approach, most often applied to mass service systems, in insurance, reliability theory, etc.).
- **Risk as a criterion for choosing a decision in "games with nature"** when the response to the chosen decision is uncertain (this includes the so-called Wald's maximal utility (guaranteed result, minimal gain) or Savage's minimal regret (maximal loss), the Hurwitz criterion (coefficient of optimism)).
- **Risk as a Bayesian estimate** (here the probability is considered as the degree of confidence in the event, which can change when new information is collected, the risk in this case is the mathematical expectation of the variance of the posterior distribution).

The second group are concepts that describe risk as a characteristic of the process:

- **Risk, as the difficulty of achieving** the goal (risk is defined through a functional that describes the evolution of the system on a set of given trajectories, being a measure of the quality of the system in relation to the quality required to achieve the goal).
- **Risk as a measure of process quality assessment** (risk is a measure of the degree of mismatch between the real process and the reference process).

- **Risk as an anti-potential for development** (risks act as slower of the speed of reproduction of the entire system).
- **Risk as a measure of disorderliness** (risk is estimated as a minimum of the total inconsistency of expert evaluations (based on the equality of all participants of the examination) of the variants of system development, measured in the inversions of transitions, necessary to restore the lexicographic order of compared variants).

Let's look at each concept in more detail.

## II. Risk as a relative value

Risk in this concept is an objective-subjective feeling that under certain conditions something undesirable, dangerous event can happen [1]. Relative risk (RR) in medical statistics and epidemiology is the ratio of the risk of an event occurring in individuals exposed to a risk factor (p-exposed) to a control group (P non-exposed). Relative risk is often used in statistical analyses of paired outcomes when the outcome of interest has a relatively low probability. For example, in medical research to compare the risk of disease in patients receiving a treatment (or placebo) with patients receiving the treatment under study or to compare the risk of complications in patients receiving a drug with patients not receiving or receiving a placebo. A particular appeal of relative risk is the ease of calculation for uncomplicated cases.

Assuming a causal relationship between exposure and outcome, relative risk values can be interpreted as follows:

- $RR = 1$  means that the impact does not affect the result.
- $RR < 1$  means that the risk of the outcome is reduced by the exposure, which is a "protective factor."
- $RR > 1$  means that the risk of the outcome is increased by the exposure, which is a "risk factor."

As always, correlation does not mean causation; causation can be inverse, or they can both be caused by a common mixed variable. In regression models, risk is usually included as an indicator variable along with other factors that may affect risk. Relative risk is usually reported as the average of sample values of independent variables. In statistical modeling, approaches such as Poisson regression (for counting events per unit exposure) have an interpretation of relative risk: the estimated effect of the explanatory variable is multiplicative for rate and thus leads to relative risk. Logistic regression (for binary outcomes or counting successes in several trials) should be interpreted in terms of odds ratio: the effect of the explanatory variable is multiplicative of the odds and thus leads to an odds ratio. Relative risk can be interpreted in Bayesian terms as an a posteriori exposure relation (i.e., after observation of the disease) normalized by an a priori exposure relation.

## III. Risk as a consequence of a random event

One of the main, basic problems of mathematical statistics is the problem of restoring the unknown law of distribution of a random variable over a finite number of its realizations. In more detail, we consider a random variable  $\xi$  for which a set of its random realizations is known (from experiments)  $X_1, \dots, X_n$ . It is required to determine the distribution law as accurately as possible  $\xi$ , e.g., probability density  $p_\xi(x)$ .

A *random event* is a subset of the set of outcomes of a random experiment; when a random experiment is repeated many times, the frequency of occurrence of the event serves as an estimate of its probability. A random event that is never realized as a result of a random experiment is called impossible. A random event that is always realized by a random experiment is called credible. "...in homogeneous *mass operations*, the percentage of one or another kind of *event* important to us under given conditions is almost always about the same, only rarely deviating any

significant amount from some average figure. We can say that this average figure is a characteristic indicator of a given mass operation (under given, strictly defined conditions) ... So, what do we call the probability of events in each mass operation? A mass operation always consists of the repetition of many similar single operations. We are interested in a certain result of a single operation and, above all, in the number of such results in this or that mass operation. The percentage (or generally the proportion) of such "successful" results in each mass operation we call the probability of this result, which is important for us. We must always keep in mind that the question about the probability of this or that event (result) makes sense only in exactly defined conditions of our mass operation. Any essential change of these conditions entails, as a rule, change of the probability we are interested in". [2].

A set of random realizations  $X_1, \dots, X_n$  is called a sample, the number of realizations  $n$  - the volume of the sample. If the random realizations considered as random variables are independent, then the sample is called repeated, otherwise it is called unrepeated. For repeated sampling, and only for it, the joint probability density function of a population of random variables  $X_1, \dots, X_n$  (if it exists) has the form  $\prod_{i=1}^n p_{\xi}(x)$ . Further on, unless otherwise stated, it is assumed that  $X_1, \dots, X_n$  - repeated sampling. In mathematical statistics, any measurable function (scalar or vector) of sample terms  $X_1, \dots, X_n$  is usually called a statistic.

### 3.1. Risk in insurance

In insurance, we will call the risk the totality of the value of possible damage in some stochastic situation and its probability. The magnitude of the possible damage in a stochastic situation is obviously unknown before the implementation of this situation and therefore random. Thus, the theoretical and probabilistic analogue of the concept of damage, obviously, is the concept of a random variable. The totality of values of a random variable and their probabilities in the theory of probability is set by the distribution of a random variable. Thus, we would like to understand risk as a random variable. However, if risks are identified with random variables set on different probability spaces, the task of comparing such risks is fundamentally unsolvable and even meaningless, since their corresponding random variables as functions of elementary outcomes depend on arguments with different meaning. Therefore, in such situations, we must identify risks with distribution functions.

The mathematical theory of risk should be formally understood as a set of models and methods of probability theory applied to the analysis of random variables and their distributions. This interpretation is rather broad and is reduced to the fact that so interpreted risk theory should be identified with the discipline, which is assigned the name "applied probability theory" and which includes such an important and rich in results field as reliability theory.

Probability theory studies properties of mathematical models of random phenomena or processes. By randomness we will understand uncertainty that cannot be eliminated in principle. With the help of probability theory concepts and statements we can describe the very mechanisms of uncertainty manifestation, reveal regularities in manifestations of randomness. Here we will deal with the probability theory based on the system of axioms, which was proposed in 20-30s of XX century by Andrey Nikolayevich Kolmogorov.

Let us call a stochastic situation characterized by the following properties or conditions:

- *Unpredictability*: the outcome of the situation cannot be predicted in advance with absolute certainty.
- *Reproducibility*: it is at least theoretically possible to reproduce the situation in question as many times as desired under unchanged conditions.
- *Stability of frequencies*: whatever the event of interest related to the situation in question is, when this situation is repeatedly reproduced, the frequency of the event (i.e., the ratio of the number of cases in which the event in question was observed to the total number of reproductions

of the situation) fluctuates around a certain number, getting closer and closer to it as the number of reproductions of the situation increases.

The property of unpredictability is obvious. If the outcome of a situation is predictable unambiguously, then there is no need to involve the apparatus of probability theory at all.

The property of the reproducibility of a situation is key to be sure of the success of applying the apparatus of probability theory to its description. It is this property that is meant when they say that probability theory and mathematical statistics are aimed at studying mass phenomena. In connection with the condition of reproducibility, one should be very careful about attempts to apply probability theory to the analysis of unique phenomena or systems. For example, there have been numerous attempts to give a quantitative answer to the question of how likely it is that other planets inhabited by intelligent beings exist in the universe. However, so far there is insufficient evidence to believe that the existence of other planets and, even more so, the existence of intelligent life on them is a mass phenomenon. Therefore, the existing predictions are very controversial and therefore inadequate.

Finally, the property of frequency stability allows us to connect the mathematical definition of event probability with the intuitive notion of it as a certain understood limit on the frequency of an event in the unrestricted reproduction of the corresponding situation.

The individual claim (or insurance claim) equal to the total amount of funds paid by the insurer under some insurance contract, i.e. a random value taking a zero value if the insurer's payments under this insurance contract did not take place (no insurance event), and a non-zero value equal to the sum of all insurance payments under the contract if at least one insurance event took place, is usually considered an elementary risk component of the insurer. The conditional value of the claim value, provided that the claim is different from zero, is called the loss.

The existing literature on risk theory provides the following classification of risk models:

- *Individual risk model* (according to the terminology [3, 4, 5]) or dynamic insurance model (according to the terminology [6]) describes the situation in which a set of insurance objects (insurance portfolio) is considered, formed at one time, insurance premiums are collected at the moment of the portfolio formation, the term of all insurance contracts is the same, and during this period insurance events occur, leading to insurance payments – claims.

- *Collective risk model* (in the terminology of [3, 4, 5]) or dynamic insurance model (in the terminology of [6]), in which it is assumed that insurance contracts are concluded by the insurer at moments of time, forming some random process, each contract has its own duration, and during the time of this contract the insurance events may occur, leading to losses of the insurance company (insurer). Such a model can be considered both on a finite and infinite time interval. When considering a dynamic model, it is always assumed that there is some initial capital allocated by the insurer for a given insurance portfolio.

### 3.2. Risk in technology

Risk in engineering is associated with the occurrence of some random event  $\mathcal{A}$  which we will call a risk event from the possible family  $\mathcal{F}$  of all events describing the risk situation under consideration [7]. These events are usually distributed in some way over time and are accompanied by certain material or other costs - generally speaking, also incidental in magnitude.

Thus, risk is characterized by two quantities - the time  $T$  the time of occurrence of a risk event and the value  $X$  bringing them damage therefore under the risk we will understand a probabilistic model  $(\Omega, \mathcal{F}, \mathbf{P})$  on which a two-component random variable  $(T, X)$  the first component of which  $T$  - time of risk event occurrence  $\mathcal{A}$  counted from some fixed point in time, and the second  $X$  specifies the damage caused by this risk event. It should be kept in mind that the quantity  $T$  may depend on the moment  $t_0$  of the beginning of the counting. Therefore, when measuring the time to the occurrence of a sharp event, it should be measured from some natural starting point. In the theory of reliability such moment is the moment when the equipment is put into operation, in the models

of life insurance - the moment when a person is born, in the models of ecological risks the time is measured between successive moments of the relevant hazardous risk events, etc.

Note also that if the risk is investigated on a fixed time interval, the risk event may not occur during the time interval in question. As it has already been noted, we often must deal with a sequence of risk events. Such situations are studied as part of risk processes:

$$\{(S_n, X_n): n = 0, 1, \dots\},$$

where  $S_n$  - the non-decreasing sequence of moments of risk events, and  $X_n$  - the sequence of damages associated with them.

For any random variable, its basic measure is its bivariate distribution:

$$F(t, x) = \mathbf{P}\{T \leq t, X \leq x\},$$

concentrated, naturally, due to the nature of the phenomenon in the first quadrant of the plane  $\{t \geq 0, x \geq 0\}$ . In most real cases, however, information about the joint distribution of the time of occurrence of the sharp event and the amount of damage is rarely available and one has to be limited to the corresponding marginal distributions of the time of occurrence of the risk event:

$$F_T(t) = \mathbf{P}\{T \leq t\}$$

and the amount of damage

$$F_X(x) = \mathbf{P}\{X \leq x\}.$$

If the risk is considered on a fixed time interval, then instead of the time of  $T$  of risk event occurrence  $\mathcal{A}$  it is natural to consider its indicator  $1_{\{\mathcal{A}\}}$  and the damage is measured by its conditional distribution if the risk event occurs:

$$G(x; \mathcal{A}) = \mathbf{P}\{X \leq x | \mathcal{A}\}.$$

In this case, the unconditional value of the damage is represented by a distribution that is joint with the occurrence of a risk event, having a jump of zero (since, naturally, in the absence of a risk event, the value of the damage is zero):

$$F_X(x) = 1 - \mathbf{P}(\mathcal{A})(1 - G(x; \mathcal{A})).$$

In general, it is natural to measure risk by the distribution of the moment of the occurrence of a risk event  $F_T(t) = F(t) = F(t, \infty)$  and the conditional distribution of damage when it occurs:

$$G(x; t) = \mathbf{P}\{X \leq x | T = t\}.$$

Their joint distribution is expressed by the formula:

$$F(x, t) = \int_0^t G(x; u) f_T(u) du.$$

The simplest case is to assume the independence of these quantities. In this case the relations take place:

$$G(x; u) = G(x) = F_X(x) = F(\infty, x); \quad F(x, t) = G(x)F(t).$$

In many real-world situations, this assumption is perfectly acceptable, with the only note being that the value of future damage now is estimated by means of present damage, which is measured by the value:

$$\hat{X} = e^{-sT} X,$$

where  $s$  - is the inflation rate of the bank interest.

Indeed, to compensate for the damage  $X$  in time, it is enough to put in the bank the amount of  $\hat{X}$  at  $s$  interest. In this case the actual dependence of the future damage on time is just expressed in the form of the present value of the damage. Further on we will stick to the assumption of independence of the time of the risk event and the amount of the damage brought by it.

This approach allows a lot of analytical analysis of some specific characteristics and cases and is used in many other new research projects. It is done only within a theoretical framework and assumptions. Practical applications require sufficient statistical data to estimate the function  $F(x, t)$ .

#### IV. Risk as a criterion of choice in games with nature

In terms of "playing with nature," the decision-making problem can be formulated as follows: The decision-maker (DM) can choose one of  $m$  of possible variants of his decisions:

$X_1, X_2, \dots, X_M$ , and with respect to the conditions under which the possible choices will be realized, one can make  $n$  assumptions:  $Y_1, Y_2, \dots, Y_N$ . The estimates of each decision option under each condition are  $(X_m, Y_n)$  are known and are given in the form of a matrix of benefits of the DM:  $A = A(X_m, Y_n) = |A_{mn}|$ . It is assumed that there is no a priori information about the probabilities of this or that situation occurring  $Y_n$  is not available.

According to the methodology of game theory, alternatives are factors controlled by the DM, i.e., chosen by him at his discretion [8]. In addition to alternatives, there are non-controllable factors, which affect the outcome of the problem and which the DM cannot control (e.g., natural phenomena). To fully analyze and decide, the DM must have some information about the values of the uncontrollable factors.

The theory of operations research, based on the information of the DM divides [9] uncontrollable factors into three groups:

- *Fixed* - these are factoring whose values are known precisely, so, for example, the sale of shares occurs only when buyers know exactly the quotation of monetary units against the dollar and euro (in this case, the quotation and is an uncontrollable fixed factor).
- *Random* factors are random variables whose distribution functions are known precisely.
- *Uncertain* factors are deterministic or random variables for which only the range of possible values or the class of possible distribution laws is known.

Of the above, the groups of random and uncertain factors are of fundamental importance. The fixed factors inherently do not differ from all other parameters of the mathematical model, because their values are known and cannot be changed at will of the DM. Random and indeterminate factors also cannot be changed at will of the DM, but in addition, they take unknown values. With respect to random factors, the probability density function is known, which means that if, for example, a random factor takes some finite number of values  $y_1, \dots, y_k$  then the decision maker knows the probabilities  $p_1, \dots, p_k$  with which these values are taken. If a random factor is a continuous random value, the probability density is known. In both cases the criteria are replaced by their mathematical expectations  $p(x)$ . Even less is known about uncertainties. If the uncertainty is a deterministic quantity, then the range of its possible values is known  $Y$  i.e., it is only necessary to consider the inclusion of  $y \in Y$ . If the uncertainty is a random value, then we do not know the probability density, but a possible class of such densities.

Uncertainties can be roughly divided into the following five groups:

1. Uncertainties that have arisen due to actions on the part of persons who have their own goals, but who are not DMs. Uncertainties of this type are called [9] strategic uncertainties.
2. Uncertainties reflecting the vagueness of the DM's knowledge of their goals. This uncertainty is not an uncontrollable factor (in the strict sense), because the choice of the goal is at the disposal of the DM.
3. Uncertainties arising from insufficient knowledge of processes or values. Decision-making based on incomplete data can be understood as a conflict with nature.
4. Uncertainties arising in the process of collecting, processing and transmitting information. Approximated information can appear because of many reasons, which include, in particular: computational errors, errors in data transmission, limited accuracy of representation and processing of numbers, limitations on the accuracy of measurements. Already in manual calculations [10] one must deal with the rounding effect arising since only a finite number of decimal places are retained in the process of calculations. Direct application of interval methods [11] to computational processes makes it possible to enclose in intervals solutions of problems whose input data are known only that they lie in certain intervals. Rounding errors encountered in the process of calculations are also included in the resulting intervals. So-called interval arithmetic [12] was introduced to obtain two-sided approximations in which intervals rather than numbers are used and both input data and intermediate and, naturally, results are situated within certain intervals.
5. Special types of uncertainties arising in the control of mechanical systems. A special role here is assigned to control of motion of systems and observation processes under conditions

of incomplete information (in other words, under conditions of uncertainty). These problems represent a natural generalization of control problems with complete information and arise in many applied problems. Incompleteness of information in these problems is a consequence of several real factors.

It should be noted that when modeling most socio-economic, technical, organizational situations, as a rule, it is not possible to build a single scalar criterion describing the interests (and thus determining the behavior) of the parties involved in the simulated process. For example, the desire to increase output and improve its quality is supplemented by the requirement to reduce production costs and minimize environmental damage, that is, the quality of functioning of most socio-economic (and not only) systems is usually assessed by a set of criteria. The corresponding models are investigated within the framework of the theory of multicriteria problems - an actively developing direction of the theory of decision-making. In them, a single value of each criterion corresponds to each decision (alternative) to be taken. However, in real conditions this requirement is often not met. A typical situation is when with respect to some parameters of the system or external influences it is known only that they can change within certain limits, within which they can take any value, unpredictable in advance.

For the DM the problem is reduced to the following: a rectangular matrix is given  $m \times n$  matrix  $A = (a_{ij})$  and DM needs to choose a row  $i \in \{i_1, \dots, i_m\}$ , which is optimal according to some optimality criterion. In this case, DM does not know which of the possible "states of nature" (uncertainties, weakly formalizable threats)  $j \in \{j_1, \dots, j_n\}$  can be realized, i.e., when choosing a solution, DM must consider the possibility of any of these states. The DM does not have any probabilistic characteristics of the manifestation of outcomes  $a_{i1}, \dots, a_{in}$  because he has no experience in solving such problems. Nevertheless, since he must decide, the question arises about the criteria for choosing a decision under the conditions of weakly formalizable factors. This criterion must prescribe a precise algorithm that, for any decision problem under uncertainty, unambiguously specifies the action that can be characterized as "optimal, according to the selected criterion".

In the general case we consider a single-criteria problem with uncertainty

$$\langle X, Y, f_1(x, y) \rangle \#(1)$$

where the choice of solution (alternatives)  $x$  from the set  $X \subset \mathbb{R}^n$  is at the disposal of the DM. The goal of the DM is the choice  $x \in X$  for which the scalar criterion  $f_1(x, y)$  reaches the greatest possible value. At the same time, the decision maker must consider the effect of interferences, errors and other kinds of uncertainty  $y$ , of which we only know that they take a value from the given set  $Y \subset \mathbb{R}^m$ .

Criteria are chosen in such a way that they are well coordinated with the intuitive perceptions of the DM about the optimality (reasonableness).

Such criteria include:

- *Maximin criterion* (or Wald's maximin utility criterion [13]) - in each row of the matrix choose a minimum score. The optimal solution corresponds to such a solution, which corresponds to the maximum of this minimum, i.e.,  $\bar{F} = F(\bar{X}, Y) = \max_{1 \leq m \leq M} (\min_{1 \leq n \leq N} (A_{mn}))$ . This criterion is very cautious and focuses on worst-case conditions, only among which the best and now guaranteed result is found. The generally accepted approach to deciding in problem (1) is based on the maxim (guaranteed result or Wald's maximized utility) principle. As the unimproved solution  $x^* \in X$  here is chosen the maximal strategy defined by the first equality

$$\max_{x \in X} \min_{y \in Y} f_1(x, y) = \min_{y \in Y} f_1(x^*, y) = f_1^* \#(2)$$

The "substantive" meaning of the maxim's principle is that if the DM has chosen and used any arbitrary solution  $x \in X$ , then he "guarantees himself" the value of the criterion  $\min_{y \in Y} f_1(x, y) = f_1[x]$  при of any uncertainty  $y \in Y$ . This fact follows from the inequalities  $f_1[x] = \min_{y \in Y} f_1(x, y) \leq f_1(x, y), \forall y \in Y$ . It is natural for the decision maker to strive for the greatest such guarantee. It is realized on the solution  $x^* \in X$  of (2), since, according to (2), firstly,  $f_1(x^*, y) \geq f_1^*, \forall y \in Y$ , and second,  $f_1[x] = \min_{y \in Y} f_1(x, y) \leq f_1^* = \min_{y \in Y} f_1(x^*, y), \forall x \in X$ . Thus, the application of the principle of makismine leads to  $f_1^*$ - the largest (maximum) of all possible guarantees  $f_1[x]$ . The



meaning of this solution: by choosing and using  $x^*$ , the DM "guarantees" a criterion (outcome) value  $f_1(x^*, y)$  which, given the realization of any uncertainty  $y \in Y$  cannot become less than the guaranteed value  $f_1^*$ . Maximin guides the DM to a "catastrophe" - to the realization of the "worst" uncertainty for the DM. Usually such a realization is unlikely. That is why Savage suggested in 1951 the principle of minimax risk (minimax regret) as an improvement of the maximin criterion.

- *Minimax criterion* (or Savage's minimax regret criterion [14]). In each column of the matrix, the maximum score  $\bar{A}_n = \max_{1 \leq m \leq M} (A_{mn})$  and a new matrix is compiled, the elements of which are defined by the relation  $R_{mn} = \bar{A}_n - A_{mn}$ . This is the size of the regret that the strategy  $Y_n$  a non-optimal choice is made  $X_m$ . The value  $R_{mn}$  is called risk, which is understood as the difference between the maximum gain that would take place if it were reliably known that the most advantageous situation would occur  $\bar{Y}_n$  for the decision-maker, and the real gain in the choice of decision  $X_m$  under the conditions of  $Y_n$ . This new matrix is called the risk matrix. From the risk matrix, one chooses the decision in which the value of risk takes the smallest value in the most unfavorable situation, i.e.  $\bar{F} = F(\bar{X}, Y) = \min_{1 \leq m \leq M} (\max_{1 \leq n \leq N} (R_{mn}))$ . That is, the essence of this criterion is to minimize risk. Like the Wald criterion, Savage's criterion is very cautious. They differ in their different understanding of the worst situation: in the first case it is the minimum gain, in the second it is the maximum loss of gain compared to what could have been achieved under given conditions. To each uncertainty  $y^* \in Y$  let us assign a number  $\max_{x \in X} f_1(x, y^*)$ . Thus, DM defines for itself the maximal value of criterion under each possible uncertainty  $y^* \in Y$ . Then the decision-maker makes a difference between indicated maximal value of criterion  $f_1(x, y^*)$  and value of the same criterion at any solution  $x \in X$  that is

$$\max_{z \in X} f_1(z, y^*) - f_1(x, y^*) \quad \#(3)$$

where  $y^*$  - is a fixed uncertainty.

By doing so, the DM numerically assesses his "regret" that he is using  $x$  instead of  $\bar{x} = \arg \max_{x \in X} f_1(x, y^*)$ . Obviously, the "regret" will be zero if the solution chosen is  $\bar{x}$  under uncertainty  $y^*$ . The difference (3) characterizes the risk of the decision maker. The risk arises because the decision maker does not know exactly what uncertainty  $y^* \in Y$  may materialize. The DM then seeks to choose a solution  $x \in X$  in which the risk ("regret") would be as low as possible. To do this, we apply the maximization principle described above to the following problem  $\langle X, Y, -[\max_{z \in X} f_1(z, y) - f_1(x, y)] \rangle$ , which is reduced to  $\min_{x \in X} \max_{y \in Y} \Phi_1(x, y) = \max_{y \in Y} \Phi_1(x^0, y)$ , where the risk (regret) function  $\Phi_1(x, y) = \max_{z \in X} f_1(z, y) - f_1(x, y)$ . The value of the risk function on the concrete pair  $(x, y) \in X \in Y$  we further call the risk of the DM when he uses the alternative  $x \in X$  and the realization of uncertainty  $y \in Y$ . The DM evaluates its risk-the difference between the "best" (maximum) value of the criterion  $f_1$  and the value realized. It is natural for the DM to strive to minimize this risk. According to the above definition, the decision-maker, adhering to  $x^0$ , "secures for himself" the risk  $\Phi_1(x^0, y)$  which, given any uncertainty  $y \in Y$  cannot become greater than its guaranteed value.  $\Phi_1^0$ .

- *Pessimism-optimism index criterion* (Hurwitz criterion<sup>1</sup> [15]). We introduce a certain coefficient  $\alpha$  called the "optimism coefficient",  $0 < \alpha < 1$ . In each row of the winnings matrix we find the largest score  $\max_{1 \leq n \leq N} (A_{mn})$  and the smallest  $\min_{1 \leq n \leq N} (A_{mn})$ . They are multiplied by  $\alpha$  and  $(1 - \alpha)$  and then their sum is calculated. The optimal solution is the solution to which the maximum of this sum corresponds, i.e.

$$\bar{F} = F(\bar{X}, Y) = \max_{1 \leq m \leq M} \left( \alpha \times \max_{1 \leq n \leq N} (A_{mn}) + (1 - \alpha) \times \min_{1 \leq n \leq N} (A_{mn}) \right).$$

At  $(\alpha = 0)$  the Hurwitz criterion is transformed into the Wald criterion. This is a case of extreme "pessimism. At  $(\alpha = 1)$  (a case of extreme "optimism"), the person making the decision expects the most favorable situation to accompany it. The "optimism factor"  $\alpha$  is assigned subjectively, based on experience, intuition, etc. The more dangerous the situation, the more

<sup>1</sup> The Gurwitz criterion has nothing to do with risk analysis. Except for the subjective perception of "random" and "voluntary" risks.

cautious the approach to the choice of decision should be, and the lower value is assigned to the coefficient  $\alpha$ .

- *Laplace criterion.* Since the probabilities of occurrence of a given situation  $Y_n$  are unknown, they are equally probable. Then for each row of the winnings matrix the arithmetic mean of the evaluations is calculated. The optimal solution will correspond to such a solution to which the maximum value of this arithmetic mean corresponds, i.e.

$$\bar{F} = F(\bar{X}, Y) = \max_{1 \leq m \leq M} \left( \frac{1}{N} \sum_{n=1}^N A_{mn} \right)$$

A few remarks about uncertainties. According to Knight [16], the distinction between risk and uncertainty is a matter of knowledge about the future. Risk describes situations that are available to be "measured" in terms of probability, while uncertainty refers to situations in which the information is too imprecise to be generalized by means of probability. Uncertainty, then, is a consequence of a lack of knowledge about reality. Over time, what is currently considered uncertainty can turn into risks (e.g., long-term weather forecasts are currently characterized by great uncertainty, but medium-term forecasts can already characterize the risks of certain events). But, the harm and benefit of an action are determined by the totality of the circumstances. And often only the boundaries of changes are known about the uncertainties, and any statistical characteristics are either absent or unprofitable to obtain. Uncertainties in (1) are denoted by  $y$ , and their set by  $Y$ ; we may assume that the set  $Y$  is known to the DM a priori.

In the problem (1) the outcome for the DM is estimated by the value of the function  $f_1(x, y)$  defined on the product  $X \times Y$  and called the criterion for DM (it can be income for the seller, for the buyer - the amount of money saved, the amount of purchased goods, the time of goal achievement and other factors evaluating the outcome for DM). Although "non (ne) bis in idem" (you cannot do the same thing twice, that is, it should not be repeated, in Latin), once again describe the process of decision-making in the problem (1). It occurs as follows. The decision maker chooses and uses his solution  $x \in X$ . At the same time and independently of this choice some uncertainty is realized  $y \in Y$ . On the resulting pair of  $(x, y) \in X \times Y$  criterion  $f_1(x, y)$  takes a particular value  $f_1(x, y)$ , called the outcome for the DM. For certainty, let us assume that the DM seeks the greatest possible outcome (if, for example, the criterion  $f(x, y)$  evaluates the sum of losses or the cost of production, then in the problem (1.1.1) we should put  $f_1(x, y) = f(x, y)$ , for  $\max_{x \in X} | -f(x, y) | = - \min_{x \in X} f(x, y)$ ).

It follows from the above that risk assessment is only possible if there are alternative choices. If there is only a single choice, then the risk is automatically zero and the spread of payments is only a characteristic of the uncontrollable natural environment. However, the alternative is always present in the form of a refusal to decide. In some cases, the refusal to decide may give an optimum on the columns, and then there will be non-zero risks in the options at the expense of choosing the wrong decision. For example, it is better not to play in a casino than to play by sticking to a strategy. In chess, on the other hand, it makes sense to play even in the case of a single (forced) move. For example, if my opponent declares "check", there is nothing to close, and retreat is only possible for the only chess cell. The risk is also zero, since refusing to play is an automatic defeat.

Availability of probability estimates  $\sum_{n=1}^N p_n = 1$  to describe the state of the natural environment  $p_1 = p(Y_1), p_2 = p(Y_2), \dots, p_N = p(Y_N)$  allows to refuse from choosing the most unfavorable case when using the Savage criterion, and to write down the desired solution in the form:

$$\bar{F} = F(\bar{X}, Y) = \min_{1 \leq m \leq M} \left( \sum_{n=1}^N p_n \times \left( \max_{1 \leq n \leq N} A_{mn} - A_{mn} \right) \right)$$

This is the more correct formula. Only when, for any pair  $(X_m, Y_n)$  the payment is determined only by the size of the loss  $A_{mn} = B - C_{mn}$

$$\bar{F} = F(\bar{X}, Y) = \min_{1 \leq m \leq M} \left( \sum_{n=1}^M p_n \times (B - C_{mn}) \right) = B + \min_{1 \leq m \leq M} \left( \sum_{n=1}^M p_n \times C_{mn} \right).$$

And only when the level of loss at the optimum for the conditions  $Y_1, Y_2, \dots, Y_N$  does not depend on  $n$  and is equal to  $\bar{C}$ , then:

$$\bar{F} = F(\bar{X}, Y) = \min_{1 \leq m \leq M} \left( \sum_{n=1}^M p_n \times (B - C_{mn}) \right) = B - \bar{C} + \min_{1 \leq m \leq M} \left( \sum_{n=1}^M p_n \times C_{mn} \right).$$

Only in this case is the solution really determined by the value of the mathematical expectation of the loss. But adjusted for  $B$  и  $\bar{C}$ . The science of these corrections is contained in many papers. Usually, we take  $B$  и  $\bar{C}$  equal to zero. For example, in ecology, it costs nothing to improve "air" (no profit), and if no one gets sick, the optimal damage is taken as 0.

## V. Risk as a Bayesian evaluation

The Bayesian paradigm of statistical inference and non-competitive decision making is simple to state. It is essentially a probabilistic view of the world, which states that all uncertainty should be described only in terms of probability and its calculation, and that probability is personal or subjective. Why is the Bayesian paradigm relevant to reliability, risk and survivability analysis? From a philosophical point of view, the answer is obvious: the Bayesian paradigm is based on the logical structure of the calculus of probability. From a pragmatic point of view, we can say that in risk analysis we often deal with unique situations, so the notion of relative frequency is not always appropriate. Another argument is that in many cases there is no direct prior data, so any uncertainty estimates can only be based on raw information; the Bayesian paradigm allows for this. Finally, risk, reliability and survivability analyses are most reliable when experts play a key role; the Bayesian paradigm allows expert experience to be formally incorporated into the analysis by considering a priori probabilities.

In mathematical statistics and decision theory, a Bayesian decision estimate is a statistical estimate that minimizes the posterior expectation of the loss function (that is, the posterior expectation of the loss). In other words, it maximizes the posterior expectation of the utility function. In statistical decision theory it is shown that a detection system with a decision selection rule by the maximum posterior probability criterion minimizes the number of erroneous decisions. The sum of the number of false alarms and omissions in a sufficiently long sequence of decisions, that is, the probability of error of any kind, is minimal compared to a system using a rule with any other criterion. However, the rule does not establish any correlation between the number of false alarms and signal misses, so it should be applied when false alarms and signal misses are undesirable to the same extent. The effectiveness of the detection system is evaluated by their total number on some time interval. It is known that this criterion is used, for example, in communication systems.

When reconstructing the values of the parameters sought, the researcher usually has additional (a priori) information about the parameters, in addition to the information "inherent" in the sample  $X_1, \dots, X_n$ . So far, parameter estimates have been based only on the sample  $X_1, \dots, X_n$  and did not consider additional information about the parameters. Taking additional information into account will improve the reliability and accuracy of the estimation.

At present, there are two main methods of accounting for auxiliary information depending on its nature - Bayesian and minimax. Below, in addition to these two methods of accounting for a priori information, another method is outlined - a generalized maximum likelihood method [17].

In the framework of Bayesian theory, this estimate can be defined as a posteriori maximum estimate. A Bayesian estimate will be such an estimate, which minimizes the Bayesian risk among all other estimates. The risk (not Bayesian, but ordinary) will be equal to the mathematical expectation of the variance of the posterior distribution. Application of the Bayesian approach implies treating the sought parameter  $u$  as a random variable. Then the a priori information is

given in the form of a priori distribution  $p_{\text{app}}(u)$  of random variable  $u$ . In most applied problems, the sought parameter is a deterministic quantity, and therefore the artificially imposed Bayesian approach of treating deterministic quantities as random quantities causes numerous criticisms of the Bayesian method. It should be noted, however, that in many cases the Bayesian method successfully accounts for a priori information and produces reasonable estimates.

So, in the Bayesian approach  $u$  - is a random variable. Let us denote by  $p(x_1, \dots, x_n, u)$  the joint probability density of a set of random variables  $X_1, \dots, X_n, u$ .

We have  $p(x_1, \dots, x_n, u) = p(x_1, \dots, x_n | u)p_{\text{app}}(u) = p(u | x_1, \dots, x_n)p(x_1, \dots, x_n)$ , where  $p(u | x_1, \dots, x_n)$  - is the conditional probability density of a random variable  $u$  at fixed values  $x_1, \dots, x_n$ ,  $p(x_1, \dots, x_n)$  - is the marginal joint probability density of a set of random variables  $X_1, \dots, X_n$  which can be found by formula

$$p(x_1, \dots, x_n) = \int_U p(x_1, \dots, x_n, u) du = \int_U p(x_1, \dots, x_n | u)p_{\text{app}}(u) du.$$

From the above it follows the formula

$$p(u | x_1, \dots, x_n) = \frac{p(x_1, \dots, x_n | u)p_{\text{app}}(u)}{\int_U p(x_1, \dots, x_n | u)p_{\text{app}}(u) du}$$

Since there is one realization of a set of random variables  $X_1, \dots, X_n$  given by sample (1.1), then finally we obtain Bayes formula:

$$p(u | X_1, \dots, X_n) = L_n(u | X_1, \dots, X_n; u)p_{\text{app}}(u)/C,$$

where the normalization factor:

$$C = \int_U p(X_1, \dots, X_n | u)p_{\text{app}}(u) du.$$

does not depend on  $u$ .

The function  $p(u | x_1, \dots, x_n)$  is called the posterior probability density function.

The point Bayesian estimate (B-estimate)  $\hat{u}_B$  choose one of the characteristics of the posterior probability density function, such as the mode, the expectation, or the median of the posterior distribution. Clearly, in the general case, all three estimates will be different.

There is another way of obtaining a point Bayesian estimate, in which a positive loss function is specified  $W(|\hat{u} - u|)$  where  $\hat{u}$  - parameter estimate  $u$ . Based on the chosen loss function  $W(t)$  and the resulting posterior density  $p(u; x_1, \dots, x_n)$  we construct the functional

$$R(\hat{u}; X_1, \dots, X_n) = \int_U W(|\hat{u} - u|)p(u | X_1, \dots, X_n) du,$$

which is commonly referred to as *the posterior risk* of estimation and for a given sample  $X_1, \dots, X_n$ . Then the B-estimate is defined as a solution to an extreme problem:

$$\hat{u}_B = \arg \min_{\hat{u}} R(\hat{u}; X_1, \dots, X_n)$$

i.e., the B-estimate minimizes the posterior risk.

In determining the B-estimate according to the formula for  $\hat{u}_B$  there also remains the ambiguity of the Bayesian point estimate because of the possibility of choosing different loss functions  $W(t)$ . The first method usually takes as its B-estimate the mode of the posterior distribution (the most likely value of the posterior random variable  $u$ ). In the second method, the quadratic loss function is most often chosen  $W(t) = t^2$ . De Groot and Rao studied the type of B-estimates for different loss functions. They found that for  $W(t) = |t|$ ,  $W(t) = t^2$  и  $W(t) = \eta(t + \delta) - \eta(t - \delta)$ ,  $\delta > 0$  (rectangular window), the B-estimate coincides respectively with the median, expectation, and mode (for unimodal and symmetric with respect to the mode of the posterior density) of the posterior distribution. If posterior distribution is unimodal and symmetric about mode, then all B-estimates for any symmetrical convex loss function coincide. The theoretical foundations of the Bayes method are presented in the monographs by De Groot [18] and Zacks [19].

In mathematical statistics and decision theory, a Bayesian decision estimate is a statistical

estimate that minimises the posterior expectation of the loss function (that is, the posterior expectation of the loss). In other words, it maximises the posterior expectation of the utility function. In statistical decision theory, it is shown that a detection system with a decision selection rule using the maximum posterior probability criterion minimises the number of false alarms. The sum of the number of false alarms and omissions in a sufficiently long sequence of decisions, i.e. the probability of any type of error, is minimal compared to a system using a rule with any other criterion. However, the rule does not establish any correlation between the number of false alarms and the number of missed signals, so it should be used when false alarms and missed signals are equally undesirable. The effectiveness of the detection system is evaluated by their total number over a given time interval. It is well known that this criterion is used, for example, in communication systems. In the framework of Bayesian theory, this estimate can be defined as an a posteriori maximum estimate. A Bayesian estimate will be one that minimises the Bayesian risk among all other estimates. The (non-Bayesian, but ordinary) risk will be equal to the mathematical expectation of the variance of the posterior distribution.

## VI. Risk as a Difficulty of Achieving a Goal

This concept was formed and developed by the school of Russian mathematician Isaak Russman [20, 21]. Risk assesses the difficulty of obtaining a declared result  $d_k$  under the existing estimates of resource quality ( $\mu_k$ ) and the requirements for this quality ( $\varepsilon_k$ ) [22]. The notion of the difficulty of achieving a goal for a given quality and quality requirements of the resource and result derives from the considerations that it is more difficult to obtain a result of a certain quality the lower the quality of the resource ( $\mu_k$ ) or the higher the requirements for its quality ( $\varepsilon_k$ ).

The functioning of a reliable system is characterized by the preservation of its main characteristics within the established limits. The peculiarity of management in socio-economic systems is that in most cases it is focused not on the complete extinguishing of deviations (the performance of this task in modern conditions is extremely difficult), and to maintain fluctuations of output parameters within limits that do not threaten the system with loss of stability and destruction. In other words, this means that the actions of such a system are aimed at minimizing the deviations of its current state from some given ideal - the goal, which is a key aspect in the study of the properties and mechanisms of the behavior of control systems. In relation to the system the goal can be considered as a desired state of its outputs, i.e., some value of its target functions.

The system is considered in the process of achieving the goal, in the movement from its current state to some future result, the quantitative expression of which is  $A_{pl}$ . Suppose the time is given to achieve the goal  $t_{pl}$ . Let us also assume that there is a minimal speed  $V_{min}$  of moving to the goal in time and the maximal speed of  $V_{max}$ . It is most convenient to measure the result and the time needed to reach it in dimensionless values, for this purpose let us assume that  $A_{pl}$  и  $t_{pl}$  equal to one or 100%.

From general considerations, the difficulty  $d_k$  of the result must have the following basic properties:

- at  $\mu_k = \varepsilon_k$  be maximal, i.e., equal to one (indeed, the difficulty of obtaining a result is maximal at the lowest admissible value of quality);
- at  $\mu_k = 1$  и  $\mu_k > \varepsilon_k$  be minimal, i.e., equal to zero (at the highest possible value of quality regardless of the requirements (at  $\varepsilon_k < 1$ ) the difficulty should be minimal);
- at  $\mu_k > 0$  и  $\varepsilon_k = 0$  be minimal, i.e., equal to zero (obviously, if there are no requirements to the quality of a resource component, a  $\mu_k$  is greater than zero, then the difficulty of obtaining a result for this component must be minimal).

For these three conditions at  $\varepsilon_k < \mu_k$  a function of the form is valid:

$$d_k = \frac{\varepsilon_k(1 - \mu_k)}{\mu_k(1 - \varepsilon_k)}$$

We also believe that  $d_k = 0$  at  $\mu_k = \varepsilon_k = 0$  и  $d_k = 1$  at  $\mu_k = \varepsilon_k = 1$ .

Note, the resulting formula has a well-founded probabilistic interpretation.

In Fig. 1, the lines OD and OB correspond to the trajectories of the system with minimum and maximum velocities.

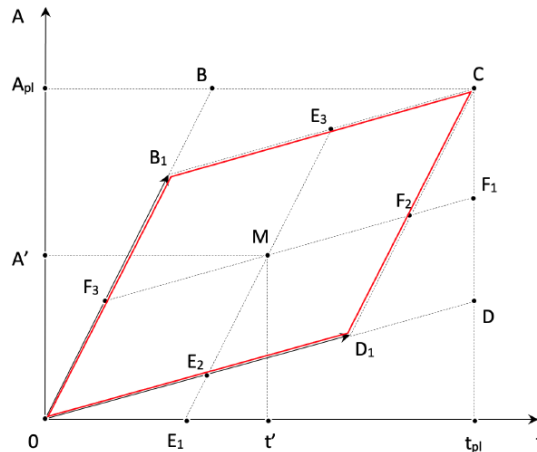


Fig. 1: Geometric interpretation of the system movement to the target

The broken line  $OD_1C$  is the boundary of the forbidden zone, and for any point  $M$  with coordinates  $(t', A')$ , describing the position of the system on an arbitrary trajectory of motion to the target within the limits of the parallelogram  $OB_1CD_1$ , the distance is taken as the value of the risk of not reaching the target:

$$r(M) = \max \left\{ \ln \frac{1}{1-d_1}, \ln \frac{1}{1-d_2} \right\},$$

$$\text{где } d_1 = \frac{\varepsilon_1(1-\mu_1)}{\mu_1(1-\varepsilon_1)}, d_2 = \frac{\varepsilon_2(1-\mu_2)}{\mu_2(1-\varepsilon_2)}, \varepsilon_1 = \frac{|E_1E_2|}{|E_1E_3|}, \mu_1 = \frac{|E_1M|}{|E_1E_3|}, \varepsilon_2 = \frac{|F_1F_2|}{|F_1F_3|}, \mu_2 = \frac{|F_1M|}{|F_1F_3|}$$

The Russman criterion contains *the idea of an optimal modification of a system* (a growing system - a reproducing growth system). This is the optimization of development by redistributing energies (resources).

## VII. Risk as a measure of process quality

If a certain quality standard of the object (process) is set, the risk of the object (process) is defined as a value proportional to the deviation of its current state from this quality standard. In fact, risk in this case is a measure of the quality of the object (process). The measure of risk itself is the threat of changes in the composition or properties of the object (process) or its environment, or the appearance of changes associated with the emergence of undesirable processes due to arbitrary influences [23, 24]. The measure of the threat of failure to achieve the goal is considered in this case as a variable, which is a function of the current state of the object (process): it increases when the assessed situation approaches a certain acceptable boundary, after which the object (process) cannot achieve the corresponding goals. The described approach assumes the availability of retrospective information on the realization of risks.

Mathematical formulation of the problem. The set of attributes  ${}^2X$  (risk factors), a set of admissible realizations of situations  $O = \{o_1, \dots, o_D\}$  (for example, the risk is realized or not realized), and there exists a target function  $o^*: X \rightarrow O$ , the values of which  $o_i = o^*(x_i)$  are known only on the finite subset of attributes  $\{x_1, \dots, x_\rho\} \subset X$ . The "sign-answer" pairs  $(x_i, y_i)$  will be called precedents. The set of pairs  $X_\ell = (x_i, o_i)_{i=1}^\rho = 1$  will constitute the training sample. You want to

<sup>2</sup> Traits can be binary (1/0, red/green), nominal (set of values), ordinal (set of ordered values) or quantitative.

use the sample  $X_\ell$  to restore the dependence  $o^*$  that is, we must construct the separating function  $\pi(o): X \rightarrow 0$ , which would approximate the target function  $o^*(x)$ , not only on the objects of the training sample, but also on the entire set of  $X$ . The separating function  $\pi(o)$  is called the choice logic function  $\pi: \pi(o) \rightarrow \{0,1\}$ , indicating that the situation  $o$  is selected into some subset  $\pi(o)$  ( $\pi(o) = 1$ ) or not ( $\pi(o) = 0$ ).

In the general case the choice functions may be arbitrary, but to use them correctly to describe the acts of choice, it is necessary to  $\pi(o)$  to impose several restrictions (the so-called axioms of choice).

**Axiom 1** (inheritance): if  $O' \subseteq O$ , then  $\pi(O') \supseteq (\pi(O) \cap O')$ , that is, if the choice is bounded, then both the "best of the best" objects belonging to  $(\pi(O) \cap O')$ , and those objects that are the best among those available in the restricted sample  $O' \subseteq O$ , but which would not be chosen if choice were available on all alternatives  $O$ .

**Axiom 2** (agreement):  $\prod_i \pi(O_i) \subseteq \pi(\cup_i O_i)$ , that is, if some object  $O$  has been chosen as the best in each of the sets  $O_i$ , then it should be chosen also when considering the whole set of sets  $\cup_i O_i$ .

**Axiom 3** (rejection):  $(\pi(O) \subseteq O' \subseteq O) \Rightarrow (\pi(O') = \pi(O))$  that is, if we discard any part of the "rejected" objects, then on the remaining subset of objects the result of selection will not change. The set of objects  $O = \{o_1, \dots, o_D\}$  which obeys all three listed axioms (inheritance, agreement, rejection) is called Pareto-optimal set.

**Axiom 4** (path-independent):  $\pi(X_1 \cup X_2) = \pi(\pi(X_1) \cup \pi(X_2))$ . This axiom is identical to the joint fulfillment of the axioms of inheritance and agreement. This axiom reflects the requirement to preserve the result of choice when implementing multi-step selection procedures. For example, the most systematically significant object is determined among the most systematically significant objects of the same type. Therefore, a set obeying the discard axiom and the Plott axiom is naturally a Pareto-optimal set.

So, we have two matrices of evaluated attributes  $p(l, k)$  и  $q(n, k)$  for situations with positive dynamics and situations with negative dynamics ( $k = 1, \dots, K$  - number of the attribute; the values  $l = 1, \dots, L$  - of the row of the situations with positive dynamics, and the values  $n = 1, \dots, N$  - of the row of the situations with negative dynamics). The axioms mentioned above are sufficient to adequately describe the structure of optimal solutions of choice problems. The accepted axiomatics shows that the constructed separating rule must be a monotone function with respect to the set of situations identified as regular (with positive dynamics). As a result, the resulting classifier of situations monotonically turns into a product of rules. This important property can be used in order not to retrain the classifier when new situations are received.

To set a partial order, you need the number of features at least  $\lceil \log_2 N \rceil + 1$  where  $N$  - is the number of situations. If there are fewer signs, there will necessarily be the same descriptions for at least a couple of situations.

General view of the separating rule:  $y = \sum_{v=1}^V x_{v,1} \cdot \dots \cdot x_{v,D_v}$ , where  $y$  - description score ( $y = 1$  for situations with "positive" dynamics for the DM,  $y = 0$  for situations with "negative" dynamics for DM);  $v$  - number of the group of variables;  $D_v$  - dimensionality (number of features in the group);  $x_{v,1}$  - value of the first characteristic in the group;  $x_{v,D_v}$  - value of the last trait in the group.

To identify the features that must be considered in the separating rule, a method for their analysis has been developed [24], using the Hamming metric<sup>3</sup>:

$$\rho(O_i, O_j) = \sum_{k=1}^K [O_{ik}(1 - O_{jk}) + (1 - O_{ik})O_{jk}]$$

The value of this metric - the distance between one-dimensional objects of the same type (rows, columns) is measured by the number of their non-matching pairs. The unit is added by the "exclusive OR" condition (addition modulo 2). Mismatching is interpreted as an error, and the

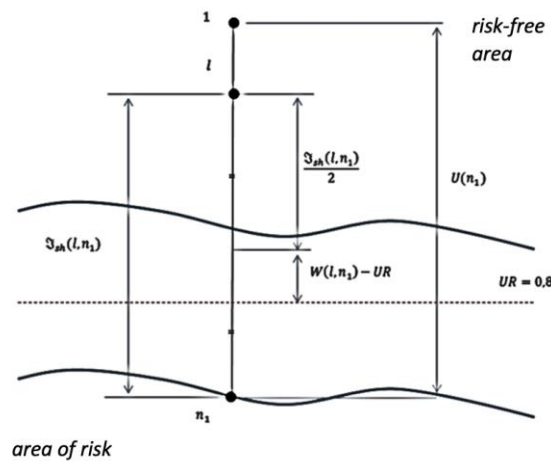
<sup>3</sup> Originally, the metric was formulated by Richard Hamming during his work at Bell Labs to determine the measure of difference between code combinations (binary vectors) in the vector space of code sequences, in which case the Hamming distance between two binary sequences (vectors) of length  $n$ . This is the formulation of Hamming distance in the NIST Dictionary of Algorithms and Data Structures.

closeness of objects is thus evaluated by the minimum number of "corrections" that one or the other, or both objects need to make for the objects to become identical, indistinguishable. Naturally, the equality is fulfilled:  $\rho(O_i, O_j) = \rho(O_j, O_i)$ . Since in the problem of risk identification we can limit ourselves to the natural class of monotone functions, we are not interested in all mismatches in a pair, but only in "ordered" ones, i.e., we can use the semi-Hemming metric:

$$\mu(O_i, O_j) = \sum_{k=1}^K O_{ik}(1 - O_{jk})$$

which reflects only the number of "successful" for the DM attributes  $\mu$  for which in the description of the first of the compared situations the value is "correct" (equal to 1), and in the second situation it is "wrong" (equal to 0). By combinations of "good" characteristics the situations "positive" for DM are separated from "negative", and by the metric  $\mu$  the size of the transition zone is specified. The semi-Hemming metric allows to eliminate inconsistent pairs of situations, which knowingly do not satisfy the conditions of strict inequalities, defining partial order.

Graphically, the solution algorithm is illustrated by the diagram in Fig. 2.



**Fig. 2:** Determining the measure of threat of non-achievement by an object (process) target state for situations with "positive" dynamics for the DM

Semi-Hemming matrix  $\mathfrak{S}_{sh}$  dimensions  $(L, N)$ :

$$\mathfrak{S}_{sh}(l, n) = \sum_k (p(l, k) - q(n, k)) \cdot b(p(l, k), q(n, k))$$

The penalty function  $b(p(l, k), q(n, k))$  Equals:

$$b(p(l, k), q(n, k)) = \begin{cases} 0, & \text{при } p(l, k) \leq q(n, k); \\ 1, & \text{при } p(l, k) > q(n, k). \end{cases}$$

In those cases where all penalties ( $k = 1, \dots, K$ ) are equal 0, we speak of the complete dominance of the situation with negative dynamics  $n$ . An unsolvable contradiction arises. Define "reserves" for situations with positive dynamics and situations with negative dynamics. For  $l = 1, \dots, L$  the value of the stockpile, respectively, is equal to:

$$Z(l) = \min_n \mathfrak{S}_{sh}(l, n).$$

The lowest of all  $Z(l)$  value  $ZZ = \min_l Z(l)$  determines the "separability threshold" of all situations with positive dynamics from all situations with negative dynamics. In principle, the "separability threshold" can be reached for more than one pair  $(l_1, n_1)$ .

For situations with "positive" for the DM dynamics for all  $n_1$  we calculate  $U(n_1) = \sum_k (1 - q(n_1, k))$ . These values account for the half-sum of the distance  $\mathfrak{S}_{sh}(l, n_1)$ , hence:

$$W(l, n_1) = UR + (1 - UR) \cdot \left( \frac{\mathfrak{S}_{sh}(l, n_1)}{2 \cdot (U(n_1) - \frac{\mathfrak{S}_{sh}(l, n_1)}{2})} \right) = UR + (1 - UR) \cdot \left( \frac{\mathfrak{S}_{sh}(l, n_1)}{2 \cdot (U(n_1) - \mathfrak{S}_{sh}(l, n_1))} \right).$$

The bottom line is a measure of risk:  $W(l) = \min_{n_1} W(l, n_1)$ .

Here  $UR$ - is the level of the acceptable risk boundary.



For situations with "negative" dynamics for the DM, similarly: for all  $l_1$  calculate  $V(l_1) = \sum_k(1 - p(l_1, k))$ . These values account for the half-sum of the distance  $\mathfrak{S}_{sh}(l_1, n)$ , therefore:

$$S(l_1, n) = UR - UR \left( \frac{\mathfrak{S}_{sh}(l_1, n)}{2 \cdot (V(l_1) - \frac{\mathfrak{S}_{sh}(l_1, n)}{2})} \right) = UR - UR \cdot \left( \frac{\mathfrak{S}_{sh}(l_1, n)}{2 \cdot (V(l_1) - \mathfrak{S}_{sh}(l_1, n))} \right).$$

As a result, the measure of risk, respectively:  $S(n) = \max_{l_1} S(l_1, n)$ .

### VIII. Risk as an Anti-Development Potential

The interpretation of risk as an anti-potential for development implies that it is considered in the context of some developing purposeful system [25]. Against the background of the global transformation processes that have begun in recent years, the relevance of solving the problems of effective management of structurally complex socio-technical systems has significantly increased. The notion of the value of objects (assets) included in such systems "presses" cost estimates of the significance of decisions made and becomes defining, but the definition of this "value" remains more of an art than a scientifically grounded methodology. Intuitively it is clear that with limited resources (of all kinds) it is necessary to strive to use these resources in the most rational way. However, to develop a rational solution, one must learn to evaluate the results of the system's purposeful activity, to compare them with the tasks set and the costs associated with this or that solution. To compare, respectively, one must learn to measure, that is, to have some quantitative measure that characterizes the result of the functioning of an individual object and the entire complex socio-technical system, as well as a "tool" that makes it possible to evaluate the result obtained.

Since the problem is to choose the best of the compared options for the growth and development of the system, we must first learn to measure the quality of the decisions made. The quality of any decision is fully manifested only in the process of its implementation (in the process of target functioning of the controlled object or system). Therefore, the most objective is to evaluate the quality of a decision by the effectiveness of its application. Thus, for a reasonable choice of the preferred solution it is necessary to measure the effectiveness of the target functioning of the managed object or system from its compared options [26] in conditions of existing uncertainty and risk. Comparison of development options and decision making directly depends on the competence of decision makers (DM), on their ability to comprehensively assess the risks associated with the functioning and development of the system. DM usually use analytical tools based on mathematical methods (from Kantorovich's "simplex method" [27] to modern methods of machine learning, neural networks [28,29,30], methods of support vectors [31], genetic algorithms [32], etc.) to provide a reasonable choice of DM.

There are several classes of decision-making tasks:

- Deterministic, which is characterized by an unambiguous connection between the decision and its outcome and is aimed at constructing a function of "progress", and the definition of stable parameters in which the optimum is achieved.
- Stochastic, in which each decision made can lead to one of a set of outcomes occurring with a certain probability, and using methods of simulation programming [33], game theory [34] and other methods of adaptive stochastic control [35] to choose the optimal strategy in the calculation on the average, statistical characteristics of random factors.
- In conditions of uncertainty, when the optimality criterion depends not only on the strategies of the operating party and fixed risk factors but also on uncertain factors of non-stochastic nature, and interval mathematics [36] or approximations in the form of fuzzy (fuzzy) sets [37, 38] are used for decision-making.

In the latter case, as a rule, methods of processing the opinions of independent experts are involved [39, 40]. Despite the widespread use of expert systems in practice for many DMs remains

unclear the fairness of using certain methods of analysis, especially when the results contradict "common sense" (in their understanding) [41], so developers must formulate and adhere to some principles, without which the automation of methods adopted in expert systems, becomes unacceptable.

Often expert evaluation procedures are based on the method of processing matrices of pairwise comparisons of different alternatives, known as the Saaty algorithm (or hierarchy analysis method) [42], which is quite widely used despite criticisms [43, 44, 45] and the lack of unambiguous solutions to several research questions.

First, if the dimensionality of the matrix of pairwise comparisons is large, the number of comparisons for each expert increases to  $N \times (N - 1) / 2$  where  $N$  - the number of alternatives under consideration. There are problems of "poor" filling of the matrix of comparisons by the experts and "insufficiency" of the qualitative scale used in the method.

Second, not all experts can compare in pairs all the proposed alternatives, so some matrices of pairwise comparisons will remain unevaluated (NA). This problem is partially solved by Saaty's development of the hierarchy analysis method to the analytic network method, but the latter contains several strong assumptions that impose restrictions on its application [46, 47].

Third, as a rule, there is no "benchmark" alternative, through which the remaining estimates are obtained by transformation  $A_{i,j} = A_{i,1} \times A_{j,1}^{-1}$ , which is used, for example, in combinatorial methods of missing data recovery.

Fourth, when generalizing the experts' opinions and moving to a general matrix of pairwise comparisons, values with a significant variation appear in the same cells, which leads to the need to work with the estimates given in the interval scale [48].

Finally, when the alternative for comparison is not the object itself, but some scalar way of determining the risks, then the problem of selecting objects is reduced to the assessment of the weights of the factors affecting the integral risk. As a result, there is a complex problem of analyzing the risks of objects, solved through the value of minimizing the integral risk [49].

The theory of complex systems (synergetic) uses nonlinear modeling and fractal analysis for forecasting. In the last decade such innovative directions as theoretical history, mathematical modeling of history based on a synergetic, holistic description of society as a non-linear developing system are actively developing.

Modern complex socio-technical systems are characterized by a distributed in space, a great variety of their constituent objects and the interaction of their various types, the heterogeneous structure of transport and technological chains, unique conditions of impact on individual objects and on the system of risks of different nature. If the stability of operation of such complex systems is understood as their fulfillment of the plan of their development with permissible deviations in the volume and time of task performance, their management is reduced to the minimization of unplanned losses in the event of abnormal situations and carrying out measures to prevent them, i.e., to the analysis, assessment and management of associated risks. The concept of management of such systems is to achieve an optimal balance between the value of the object, associated risks and performance indicators, on the basis of which economic goals are formed and the use of the object is ensured in such a way that it creates added value. In general, the optimal profit-oriented management consists in the ability to find a balance in the redistribution of available resources (material, human and informational) between "production activity" and "maintenance of the development potential".

The closest to the above is described by the models of interaction between the developing object and its environment – the model of self-improving developing systems by V. Glushkov. He introduced a new class of dynamic models based on nonlinear integrodifferential equations with prehistory [50]. He also developed approaches to modeling so-called "self-improving systems" and proved theorems on the existence and uniqueness of solutions describing their systems of equations [51]. However, it should be noted that the name "evolving" applied to the class of systems under consideration is not quite correct and contains some ambiguous ambiguity. The growth of the system may not be accompanied by its development (for example, improvement of

the science of creation and design, instructions of manufacturing or use of a product) and vice versa (for example, expectations of a quick practical return from basic science). Usually, growth and development combine with each other, there is a smooth or jumping change of proportions between them, and some "equilibrium" state with the external environment comes (or does not come).

Parallel to V. Glushkov's work, sectoral research also began on the scientific and technological revolution (STR) [52, 53], which laid the foundations for potential systems theory, and work based on biophysical and economic models [54], which proposed a model involving integrodifferential equations describing the production, adoption and forgetting of knowledge in production cycles due to the transition to a different scientific and technological basis. It shows the cyclical nature of capacity accumulation and the need to develop complex systems (health, education, industrial safety systems, ecology and other infrastructure projects by generations).

Understanding and development of the mentioned models led N. Zhigirev and A. Bochkov to the need to introduce a class of so-called "smart expansive systems". [55], which consists of three subsystems (Fig. 3).

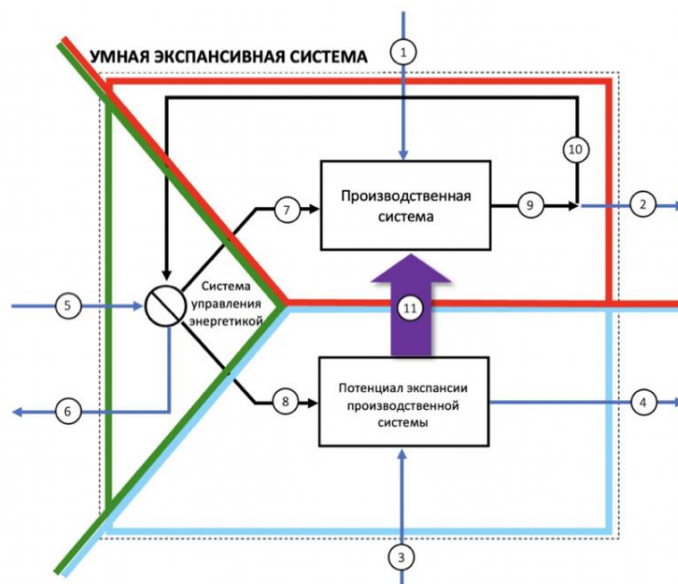


Fig. 3: Schematic diagram of a Smart Expansion System

Smart Expansive System (SES) is an open system, which can be growing or developing, or simultaneously both growing and developing (when, for example, objects of different "generations" are included in it). Sometimes the growth of a system is accompanied by so-called "flattening" (alignment at one level of development) and degradation in its development. SES openness is caused by the fact that it needs to effectively allocate the necessary resources from the external environment, possibly "cleaning them before using (including human resources), and to remove waste of vital activity.

The production subsystem is evaluated by the reproduction rate multiplier at a conditional minimum of development potential. Below this minimum (the critical mass of potential) the growth and development of the SES is impossible in principle. The potential catalytic function describing this multiplier is, in the limit, an asymptotic curve with saturation (like a logistic curve), although potential inhibition is also possible since the production subsystem occupies the space of the general SES. This behavior is analogous to the flow system of a "brusselator" (an intensively running conveyor belt), when the initial substrates flow out of it without time to react with the catalyst, not mot row for enough of it [56].

The production subsystem serves to measure the success of expansion, defined, for example, by the volume of useful products produced by the system.

The expansion potential subsystem of the production subsystem is designed to catalyze the management of forms produced and resources (sometimes measured by money, which has a dual structure - the cost of renewing matter and the cost of maintaining information in the broad sense of the word) produced in the production subsystem.

The energy management subsystem is a two-loop resource management system (financial and temporal) between production and contribution to infrastructure projects.

In Fig. 3, the externally directed flow of energies (5) is distributed by the regulating subsystem to the production subsystem (7) and is directed to the expansion potential subsystem to produce knowledge and improvement of technologies, "recipes" for the preparation of products (the so-called flow to the development of "infrastructure"). From the external environment the expansion potential subsystem also receives additional information about new knowledge, inventions and technologies (3) and has a catalytic effect on the production subsystem (11). The production subsystem, in turn, receives from the external environment a flow of "purified" semi-finished products (1) for further expansion. In the process of expansion inevitably there is a partial forgetting of information caused by various reasons, including physical death of carriers of original thought-forms (4), which causes weakening of expansion potential of the whole system.

From the production subsystem to the external environment there is an outflow (2) of products, unused semi-finished products, waste product assembly, etc. Cleared from (2), the energy flow coming by the results of labor into the production subsystem and the results of the sale of products on the market supports the functioning of the regulatory subsystem. In the latter, over time, dissipation of energies (6), not yet distributed across subsystems, is possible, capable of causing, under certain conditions, the collapse of the control system.

Let us dwell a little more on the peculiarities of the deterministic and stochastic approach to SES modeling.

For the deterministic case, the SES is described by a two-parameter model in terms of time (1) and in terms of energy distribution proportions (3).

The first equation describing the system is in some sense autonomous:

$$\frac{dX(t)}{dt} = \left( g \times \frac{\varphi(\beta)}{1 + \beta} - a \right) \times X(t) - b \times X^2(t), \quad (1)$$

where  $X(t)$  - the volume of the "production subsystem", measured by the number of products;  $a \times X(t)$  - the summand of the linear part - the maintenance of the production technology requires linear costs, in economics, for example, it is the cost of depreciation;  $g \times \frac{\varphi(\beta)}{1 + \beta} \times X(t)$  - the linear production function of the useful subsystem with the parameter  $\beta$ ;  $s = \frac{1}{1 + \beta}$  ( $0 \leq \beta \leq \infty$ ) - The proportion of the distribution of energy from the newly created forms  $s$  ( $s \in [0,1]$ ;  $(1 - s) = \frac{\beta}{1 + \beta}$  ( $0 \leq \beta \leq \infty$ );  $g$  - the scale factor of production losses, usually fulfilled by  $0 \leq g \leq 1$ ;  $\varphi(\beta)$  - a preset amplifier of mold production by reading "correct information" (assembly instructions) (information as a catalytic function);  $b \times X^2(t)$  - a quadratic term taking into account the limitedness of "semi-finished products" and the competition of finished "products" in the surrounding world.

The function  $\varphi(\beta)$  has the form of a logistic curve (Fig. 4), which in the general case is not necessary if the requirements of positive bounded monotonicity are met. This function can have discontinuities of the first kind. The final form of the function  $\varphi(\beta)$  with argument  $\beta$  is also determined by the degree of detail required in the calculation.

The segments on the abscissa axis  $[\beta_1; \beta_2]$ ,  $[\beta_2; \beta_L]$  и  $[\beta_R; \beta_5]$  at  $X_k(\beta) < 0$  are the degradation regions of the smart expansive system. Accordingly, the segments  $[0; \beta_1]$ ,  $[\beta_L; \beta^*]$  и  $[\beta^*; \beta_R]$  at  $X_k(\beta) > 0$  - the regions of its growth (development). And only from  $\beta_2$  the expansion of the system begins (on the segment  $[\beta_2; \beta^*]$  grows  $X_k(\beta)$  and on the other segments it only decreases). It does not make sense to search for a solution above the limit value  $\beta = \beta_5$  it does not make sense to search for a solution, although at the point  $\beta_R$  the system begins to actively degrade.

The type of the logistic curve is chosen in such a way that on the segment  $[0; \beta_1]$  the efficiency of the productive subsystem is extremely low (this is the area of low-skilled labor and some

potentially breakthrough ideas in science). The segment  $[\beta_L; \beta_R]$  corresponds to mass production with the use of available knowledge and skills. The optimum  $X_k(\beta^*)$  is located inside  $[\beta_L; \beta_R]$ , while at the same time there is an underdeveloped and under-demanded  $[\beta_L; \beta^*]$  science is insufficiently developed and demanded, and on  $[\beta^*; \beta_R]$  - science is "too much" and the results of scientific research simply do not have time to be implemented and mastered in the production subsystem.

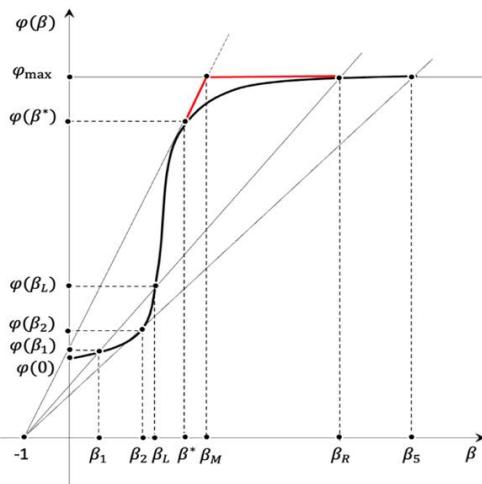


Fig. 4: Dependence of "development potential" on funds spent

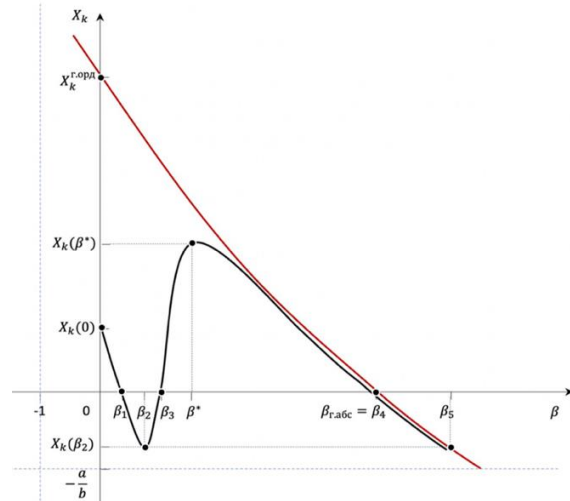


Fig. 5: Dependence graph  $X_k(\beta)$

Point  $\beta = 0$  in Fig. 4 corresponds to a situation where all resources are spent exclusively on the growth of the production subsystem. The potential of such a system is low because of constant losses, which can be avoided if there is the capacity to anticipate and deal with emerging risks.

Plot  $(0, \beta_1)$  shows that if the resources allocated to study and counteract threats and risks are low, the return on such research and activities is less than the resources allocated to them. Gathering information, research on internal and external threats at a low level does not allow to obtain an adequate assessment to improve the quality of decision-making in most cases of one way or another.

At the site  $(\beta_1, \beta^*)$  the contribution to the development potential begins to give a positive return, but only at the point  $\beta_L$  the so-called level of "self-repayment" of costs for the development of the "potential" of the system will be reached  $\varphi(\beta_L)$ . Therefore, it is reasonable to consider this point as a point of "critical" position. Decrease of potential  $\varphi(\beta)$  to the level of  $\varphi(\beta_L)$  is a threat that "by virtue of circumstances" it will be economically viable to implement a "survival strategy" - a strategy of complete rejection of expenditures to solve problems of anticipating and preempting threats and risks and ensuring reproduction only by building up low-efficient capacities in the production subsystem.  $\beta \rightarrow 0$ .

The optimum is reached at some point  $\beta^*$ . This point has a very definite meaning. Thus, if resources for capacity development are allocated "too much». ( $\beta > \beta^*$ ) then the resources  $(\beta - \beta^*)$  are incorrectly withdrawn from the current reproduction, and a situation occurs, when disproportionate efforts are spent on studying and counteracting numerous risks, which the developing system may never face. The optimum does not depend on the values of  $g, a, b$ . There may be cases where  $a$  is so large that even with an optimal solution the system does not develop but degrades. The condition of nondegeneracy of the solution is the presence of positive ordinates  $y X_k$  (Fig. 4).

In the general case on the segment  $[0, \beta_L]$  over-regulation and excessive formalization of the system control can only be harmful, and on the interval  $[\beta^*, \infty)$  there is a situation in which the cost of searching for an existing solution is so high that it is preferable to get it from scratch. The

final graph  $X_k(\beta)$  considering the graph compression  $\varphi(\beta)$  by  $g$  and shifting by the value  $a$  down vertically, and then dividing by the value of  $b$  is shown in Fig. 5.

The second equation describes the dynamics of the laws of the system's functioning, by which the  $X(t)$ :

$$\frac{dY(t)}{dt} = f \times (Y_0 - Y(t)) + h \times \frac{\beta \times \varphi(\beta)}{1 + \beta} \times X(t), \quad (2)$$

where  $Y_0$  - is some initial state. The regulator  $f$  describes the loss of knowledge, while the  $\left(h \times \frac{\beta \times \varphi(\beta)}{1 + \beta}\right) \times X(t)$  corresponds to the complexity of laws for new unknown control functions ("emergence") that are not available in subsystems.

The final solution is as follows:

$$Y_K = Y_0 + \frac{ha^2}{4 \times gbf} \times C, \quad (3)$$

where the parameter  $C = \left[\left(\frac{2}{A} \times \frac{\varphi(\beta)}{(1+\beta)} - 1\right)^2 - 1\right] \times \beta$  is searched as the maximum value from the above range. The maximum  $C$  is reached either at the edges of the segment or at one of the local minima. Thus, in Fig. 4, on the asymptote  $\beta \in [\beta_M, \beta_R]$  is satisfied by  $\varphi(\beta) = \varphi_{max}$ .

There are, however, unique systems whose value depends solely on the capacity of the producer. When the potential is high, and its value is not underestimated in relation to the "fair," i.e.  $g \gg 0$  - the second bracket is satisfied, even when  $x_{max} = 1$  and there is no competition from other producers  $b \cong 0$ .

The presence of two optimal solutions  $\beta^*$  (in terms of production quantity) and  $\beta^{opt}$  (intelligently) defined through the constant  $C_{opt}$ , suggests that the optimal one would be  $\beta^{opt} > \beta^*$  rejection of the shaft in favor of the maximum use of the expansion potential, that is - production with the "optimal" reserve of the "possible use" of the products produced (multifunctionality). Despite the schematic nature of the described model, it gives an idea that threats and risks can be considered as "anti-potentials" of development (i.e. they are retarders of the reproduction rate of the whole system). To model a real system, it is necessary to analyze the "raw" process data, and then synthesize them into a meaningful structure that explains the process under study.

The model of system growth considering the influence of random perturbations of system productivity on the rate of its reproduction considers that in the "quasi-linear section" of system expansion, not only the rate of expansion, but also the dispersion of the process is important. In this case "volatility" of the process itself plays a greater role than profitability of the "production subsystem". Despite the increase "on average" of the amount of product from each element of the system, nevertheless, each of the elements separately is characterized by a limited time of effective functioning. At the same time, the index of "population on mortality" under natural restrictions on the mathematical expectation is mainly influenced by the value of dispersion. That is why, for example, in economics, where processes with mathematical expectation values of the order of several percent are studied, it is the dispersion values that appear in the definitions of "risks."

Here it is extremely important to note that to estimate the values of mathematical expectation and dispersion, their quantitative calculation for the initial moment of time based on group estimates is made. Further, it is hypothesized that these estimates obtained for the group can be used to predict the motion trajectories of each element of the group individually. This is a very strong assumption, since it asserts, first, that the obtained estimates will remain constants for the whole prediction time, and second, it asserts that each element at any time behaves like some element at zero time. Such assumptions are valid only for ergodic processes.

But not all processes described by the model under consideration are ergodic. In systems consisting of elements of more than one type, the need to consider such "risks" increases significantly. And these "risks" themselves are much higher.

The described models can be complemented by the model of the influence of capital fluctuations on the growth of the system, which is connected not with the properties of the system

itself, but with the level of fluctuations in the parameters (influencing factors) of the external environment (fluctuations in the level of corruption, changes in tax legislation, etc.). The most probable values of the number of elements in such a model are always less than its average value. Some value is introduced as a threshold of criticality of the state. If the current value is lower than the critical value, the probability of bankruptcy increases sharply. It is important to note that as time increases, the critical values also increase. Moreover, if the amplitude of fluctuations of the distribution variance estimate is large and the mathematical expectation and the initial value are small enough, the probability of system degradation tends to unity.

Thus, on average, external fluctuations accelerate the growth of the system, but the price for such accelerated growth is an increased probability of its degeneration (decrease of mathematical expectation of its degeneration time). And since the expansion process is multifactorial, and the "prehistory" of behavior of such system (as it takes place, for example, in mass service systems), as a rule, there is no analysis based on statistics of past observation periods, but synthesis of risk of functioning of "smart expansion system" is necessary. For SES, it is more correct to speak of risk synthesis rather than analysis. Although the concept of "synthesis" is almost never used in relation to risk as opposed to "analysis", risk analysis is specific to systems where risk events are frequent enough to allow the well-developed apparatus of probability theory and mathematical statistics to be applied. This approach works in insurance, for example, in reliability theory, when we are dealing with flows of insured events, accidents or failures. But when it comes to safety in an era whose main characteristic is constant variability and variability, it can only be done through risk synthesis, by developing increasingly sophisticated automated advising systems - professional cues (PAFs) - or by replacing professionals with highly intelligent robotic systems. Risk becomes "synthetic" from "analytical" in this case.

As the analysis of integral assessments of the state of complex objects and systems used in systems research shows, generalized criteria (indices) of risks are widely used: additive (weighted arithmetic average) and multiplicative (weighted geometric average) forms:

- arithmetic (smoothing out the "outliers" of individual risk indicators)  $R_{ar} = \sum_{i=1}^M (\alpha_i \times r_i)$ ;
- geometric (reinforcing the negative "emissions" of NPR)  $R_{ge} = \prod_{i=1}^M r_i^{\alpha_i}$ ;
- geometric anti-risk  $1 - R_{\emptyset} = U_{\emptyset} = \prod_{i=1}^M (u_i)^{\alpha_i} = \prod_{i=1}^M (1 - r_i)^{\alpha_i}$ .

Weighting coefficients  $\alpha_i$  of partial evaluations  $r_i$  satisfy the condition:

$$\sum_{i=1}^M \alpha_i = 1; \alpha_i > 0 (i = 1, \dots, M). \quad (4)$$

Actual numbers  $r_i$  (partial risks) take values from the interval  $[0, 1]$ .

For UES, the most acceptable form of risk representation is geometric anti-risk, which satisfies the basic a priori requirements underlying the risk approach to constructing a nonlinear integral estimate  $R_{\emptyset}$ , namely:

1. *smoothness* - continuous dependence of the integral  $R$  and its derivatives from partial estimates:  $R(r_1, \dots, r_M)$ ;
2. *boundedness* - the boundaries of the interval of variation of the partial  $r_i$  and integral  $R$  estimates:

$$0 < R(r_1, \dots, r_M) < 1; \text{ при } 0 < r_1, r_2, \dots, r_M < 1. \quad (5)$$

3. *equivalence* - the same importance of private assessments  $r_i$  и  $r_j$ ;
4. *hierarchical unilevel* - only partial assessments are aggregated  $r_i$  which belong to the same level of the hierarchical structure;
5. *neutrality* - the integral assessment coincides with the private assessment, when the other takes the minimum value:

$$R(r_1, 0) = r_1; R(0, r_2) = r_2; R(0, 0) = 0; R(1, 1) = 1. \quad (6)$$

6. *uniformity*  $R(r_1 = r, \dots, r_M = r) = r$ .

Geometric anti-risk derives from the notion of "difficulty of reaching the goal" proposed by Isaak Russman (see above) and is a "top estimate" for the weighted average arithmetic and weighted average geometric risk. The geometric anti-risk satisfies, moreover, the so-called "fragility of good" theorem in the catastrophe theory, according to which "...for a system belonging to a particular part of the stability boundary, at small changes in parameters it is more likely to fall into the region of instability than into the region of stability. This is a manifestation of the general principle that all good (e.g., stability) is more fragile than bad" [57]. Risk analysis uses a similar principle of the limiting risk factor. Any system can be considered "good" if it satisfies a certain set of requirements but must be considered "bad" if at least one of them is not met. At the same time, all "good", such as the environmental safety of territories, is more fragile - it is easy to lose it, and difficult to restore.

A continuous function  $R(r_1, \dots, r_i, \dots, r_n)$ , satisfying the above conditions, has the following general form:

$$R(r_1, \dots, r_i, \dots, r_n) = 1 - \left\{ \prod_{i=1}^n (1 - r_i) \right\} \times g(r_1, \dots, r_i, \dots, r_n), \quad (7)$$

If in the special case  $g(r_1, \dots, r_i, \dots, r_n) \equiv 1$ , then, respectively:

$$R(r_1, \dots, r_i, \dots, r_n) = 1 - \left\{ \prod_{i=1}^n (1 - r_i) \right\}, \quad (8)$$

which gives an underestimation of the integral risk from the calculation that the flow of abnormal situations for the objects of the system is a mixture of ordinary events taken from homogeneous, but differing values of  $r_i$  ( $i = 1, \dots, n$ ) samples.

Since risks for real systems are usually dependent, we get

$$g(r_1, \dots, r_i, \dots, r_n) = 1 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n C_{ij} \times [r_i]^{\alpha_{ij}} \times [r_j]^{\beta_{ij}}, \quad (9)$$

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n C_{ij} \leq 1, C_{ij} \geq 0, \alpha_{ij} > 0, \beta_{ij} > 0, \quad (10)$$

where  $C_{ij}$  - coefficients of risk coherence of the  $i$ -th and  $j$ -th abnormal situation for the system objects;  $\alpha_{ij}$  и  $\beta_{ij}$  - positive coefficients of elasticity of substitution of the corresponding risks, which allow taking into account the facts of "substitution" of risks, mainly due to the fact that simultaneously the measures to reduce all risks cannot be carried out due to the limited time and resources of the DM.

The current values of partial risks  $r_i$  ( $i = 1, \dots, n$ ) included in (7) are values that change over time at different rates (for example, depending on the seasonal factor, the priorities of technological problems in some systems of the fuel and energy complex change significantly). Private risks  $r_i$  are built, as a rule, through the convolutions of the corresponding resource indicators - influencing factors, which have a natural or cost expression. These influence factors are measured in some own synthetic scales (e.g., the Saaty multiplicative scale of pairwise comparisons [46]), the mutual influence of which should also be studied, since they are generally non-linear and piecewise continuous. To obtain estimates of the influence factors, it is necessary to construct weighted scales. To solve this problem the so-called "vector compression method" was developed [58, 59, 60, 61].

## IX. Risk as a measure of disorderliness

It is obvious that when talking about risk, we cannot avoid talking about many accompanying circumstances during the development of a situation from less risky to riskier, up to the occurrence of the risky event itself. All these circumstances, as defined by the International Organization for Standardization (ISO), are combined in the concept of uncertainties affecting the achievement of the goal of any activity.



First, we are talking about the point in time when a risky event will happen, the place where it will happen, the time during which this event will last, the intensity of the impact of risk factors, the possibility (probability) of such an event in principle and the expected consequences after its realization. Since we are talking about activities, the results of which are affected by uncertainty, to assess a rare event, we must somehow estimate all the above circumstances and only after that, and only with a certain margin of error, speak about the coming risk event. If the circumstances of the activity periodically repeat, it is possible to determine the probability of the event by statistical methods based on the analysis of the frequency of this repeat, and the consequences (damage) can be estimated based on the mathematical expectation of losses from experience.

Both in the case of a statistically certain event and in the case of a rare event, it is important not only to know all the circumstances of the risk event, but also the order in which they occur and follow. Knowledge of this order can help assess the proximity of a risk event even without having actual information about the circumstances themselves. Here the apparatus of group theory and methods of group permutations come to the rescue. As an estimation of an order of circumstances (factors) of a risk event, as a rule, is made with attraction of expert opinions, the problem of collective choice inevitably arises - a problem of reduction of several individual expert opinions on an order of preference of compared objects (alternatives) in a uniform "group" preference. The complexity of collective choice consists in the necessity of processing the ratings of compared alternatives, set by different experts in private own scales.

Below is the author's original algorithm [61] for processing expert preferences in the collective choice problem, based on the notion of the total "error" of experts and measuring their contribution to the collective measure of their consistency.

In practice, the efficiency of decision-making requires the development and application of specialized algorithmic and methodological support. If a group of experts participates in the decision support process, the so-called collective (group) choice problem arises. The existing algorithms for solving collective choice problems [62, 63, 64] can be roughly divided into three classes. A representative of the first class is the Schulze method [65] (based on the proof of the Arrow theorem) with the selection of Pareto-optimal solutions (Schwartz exception) from the first ranking to the last, with the selection recalculating the criteria for the next step. The disadvantage of the method is a rather complicated algorithm of constant recalculation, which significantly complicates the practical use of the method.

A typical representative of the second class is the skating system [66], which has proven itself in the ballroom dance competitions. It is simple in computational calculations and is based on the so-called understandable majority principle. Unfortunately, in many ways, it is this simplicity that can lead to unstable decisions, and, therefore, the inability to distribute the final places among contestants in one round, or to recognize a draw between competitors [67, 68].

The third class consists of regression models, the type of nonlinear factor analysis and other methods of information compression [69, 70], in which the desired solution is constructed in the form of the problem of minimizing the accumulated errors. The difference between the methods of the third class is that they are not focused on the choice of the leader in the ratings, but are determined by the optimum, which is influenced by the entire volume of data.

The mentioned methods of solving the problems of collective choice in general are inherent to the problem of coordinating the experts' evaluations when comparing the evaluated objects.

In 1951 C. Arrow formulated [71] the theorem "On the impossibility of collective choice within the framework of the ordinality method", mathematically generalizing the Condorcet paradox [72]. The theorem asserts that within the framework of this approach there is no method for combining individual preferences for three or more alternatives that would satisfy some quite fair conditions (the axioms of choice) and would always give a logically consistent result.

When the uncertainty of the objects themselves is superimposed on the ambiguous opinions of the experts, some hierarchy is assumed in solving the choice problem. This is the case, for example, in the method of hierarchy analysis [46], when each of  $M$  of the experts has his/her own opinion, different from the others, concerning the weights of the objects in question  $N$  objects

through the coefficients of the preference matrix  $(S_{ij}^m = \frac{w_j^m}{w_i^m} (i = 1, \dots, N; j = 1, \dots, N; i \neq j; m = 1, \dots, M))$ . Usually, the weights are averaged and work with a generalized matrix  $S_{ij}$  which leads, as a rule, to a violation of the basic axioms of "right" choice (universality, completeness, monotone, lack of dictator, independence), proposed by W. Pareto [73, 74], R. Koch [75], C. Plott [76] and others. The rejection of one or another averaging procedure complicates the task of selection and leads, for example, to the need to solve the problem of "merging multidimensional scales" [59]. Experts need to reach consensus [55], at least with the accuracy of determining private ratings in the full order of objects, and then seek agreement in the weighting coefficients between neighboring nearest objects, setting a single scale, to obtain consistent solutions.

In the general case [61] we consider  $N$  comparison objects  $O_1, \dots, O_k, \dots, O_N$  whose indices are the first  $N$  members of the natural series  $E_{PIO} = \langle 1, \dots, k, \dots, N \rangle$  - correspond to the order in which the objects are presented for examination. In the examination of objects, the following people take part  $M$  experts  $E_1, \dots, E_m, \dots, E_M$ . Each of the experts  $E_m$  has his own idea of the order of objects  $g_m = \langle g_{m,1}, \dots, g_{m,n}, \dots, g_{m,N} \rangle$  which indexes increase with decreasing of some quality of objects from the expert's point of view. The value  $g_{m,1}$  corresponds to the index of object  $O_{k_1}$ , taking part in examination with maximal quality according to expert's opinion  $\exists_{m,}$  a  $g_{m,N}$  - the worst-quality object with the index  $O_{k_N}$ :

$$G = (g_{m,n})_{\substack{m=\overline{1,M} \\ n=\overline{1,N}}} = \begin{pmatrix} g_{1,1} & \dots & g_{1,N} \\ \dots & \ddots & \dots \\ g_{M,1} & \dots & g_{M,N} \end{pmatrix}.$$

Thereby  $g_m$  - it is a *permutation of object ratings* (PORs), the argument of which is the order of  $E_{POR} = \langle 1, \dots, n, \dots, N \rangle$ .

Places  $p_m = \langle p_{m,1}, \dots, p_{m,k}, \dots, p_{m,N} \rangle$  by values inverse to ABM  $g_m$  ( $p_m = g_m^{-1}$ ) are *permutations of object indices* (PIO) with argument  $E_{PIO}$ :

$$P = (p_{m,n})_{\substack{m=\overline{1,M} \\ n=\overline{1,N}}} = \begin{pmatrix} p_{1,1} = g_{1,1}^{-1} & \dots & p_{1,N} = g_{1,N}^{-1} \\ \dots & \ddots & \dots \\ p_{M,1} = g_{M,1}^{-1} & \dots & p_{M,N} = g_{M,N}^{-1} \end{pmatrix}.$$

It is necessary to find an optimum of the consistency measure and to restore the full collective order in preferences on the basis of private expert ratings, i.e., to compress all private POR rankings  $g_m (m = 1, \dots, M)$  in the form of a POR  $g_m^* = \langle g_1^*, \dots, g_N^* \rangle$  which would reduce the total inconsistency of expert evaluations  $g_{m,n} \rightarrow g_m^*$  (based on the equality of all expert participants), measured in the inversions of the transitions from  $g_{m,n} \kappa g_m^*$ , that is

$$K^* = \min_{g_m} K(g) = \min_{g_m} \left( \sum_{m=1}^M K_m(\langle g_1, \dots, g_N \rangle) \right),$$

where  $K_m(\langle g_1, \dots, g_N \rangle)$  - is the sum of inversions in the evaluations of  $m$  of the -one expert,  $K^*$  - is the marginal measure of inconsistency of the experts' opinions.

Finding an optimum in permutations of object ratings is equivalent to finding a permutation of the object index  $p^* : p^* = \langle p_1^*, \dots, p_N^* \rangle$ , since  $K(g_m^*) = K(p_m^*)$  where  $p^* = (g^*)^{-1}$  (the lengths of the inverse paths  $(E \rightarrow g)$  are the same as the forward paths  $(p = g^{-1} \rightarrow E)$  at any  $g$ ). This problem belongs to the class of integer programming problems (on the structure graph  $V(G, G \times G)$  of the POR graph arranged by levels of errors). Methods for solving such problems are sufficiently well developed [77,78], but none of them guarantees the uniqueness of the global minimum. A complete enumeration of all ABMs can provide some guarantee. Such a variant is possible for  $N \leq 10$ . For each  $g$  is considered  $K(P_m, g)$  и the sum of  $K(g)$ , and the current state of the set of global minima is "memorized". The subset  $g \in G$  for which  $K(g) = K^*$ , we call the set of global minima -  $G^K$ . Since  $M$  is odd, it, like the set of local minima, consists of isolated solutions (permutations) obeying the following rules.

Rule 1. If  $P_m g_l(E) < P_m g_{l+1}(E)$ , then the sum of inversions  $K(g_m)$  is increased by 1, and if  $P_m g_l(E) > P_m g_{l+1}(E)$ , the sum of inversions decreases by 1.

Rule 2. The decrease and increase of the sum depend on the number of rows in which the second condition  $M^2$  (Rule 1) dominates the first condition  $M^1$ . The ratio is

$M^1 + M^2 = M$ . Then the sum  $K(g)$  from the influence of  $s_i$  will decrease by exactly  $M^2 - M^1$  units (при  $M^2 > M/2$ ) or increase by  $M^1 - M^2$  units (при  $M^2 < M/2$ ).

**Rule 3.** "Cutoff condition. The POR  $g$  belongs to the set of local minima  $G^P$  if for all  $j = 1, \dots, N - 1$  the sum of errors only increases with rotation of neighboring columns by the symbol  $s_j$ . That is  $g \in G^P$  has neighboring vertices of the graph  $V$ , exceeding by sum the found local optimum  $g$  by at least one.

The search  $G^P$  makes sense at large  $N$ , but at small  $N$  it is also effective, because the decrease (increase) of some selected pair does not depend on the place where the pair stands, but only on the contents of the obtained inversions. Depending on the number of compared objects ( $N$ ) 2 variants of further actions are possible: either direct calculation, or consecutive iterations.

The questions concerning the comparison of the proposed method with other methods of information compression (for example, with the factor analysis, with the averaging method or with the Schulze method) are left out of the present discussion. The further development of this method implies its application in ranking determinations that allow equality of estimates of compared objects when determining weight coefficients of compared objects (like pairwise comparisons in the method of hierarchy analysis and solving problems of heterogeneous scales fusion).

## X. Conclusions

Risk is always a consequence of our ignorance, which generates uncertainties of various kinds in the decision-making process. Previously, uncertainty was assumed to be due only to the subject's lack of knowledge, while nature itself is dominated by a universal causal relationship of phenomena and events. However, many causal relationships remain unresolved due to the incompleteness of our knowledge. Science seeks to compensate for this deficiency through probabilistic laws, predicting the results of future events with the help of past events. If man had perfect knowledge of all causes and effects of natural phenomena, then everything would be precisely defined for him. So thought Laplace. But in human relations, as John von Neumann rightly believed, uncertainty arises from the ignorance of some people of the intentions, behavior, and actions of others. So, while the natural sciences have made remarkable progress in uncovering uncertainty in nature through laws based on the repeatability of the phenomena under study, the social and economic sciences have been much less successful. This is largely since the analysis of social processes must consider, along with objective conditions, such subjective factors as goals, interests and motives of people's activities, which are difficult to describe by probabilistic laws. The logic of circumstances is often stronger than the logic of intentions.

Based on previous research, science has developed a rational model of decision-making under uncertainty that describes the rational behavior of an individual or group, which often results in the successful achievement of a goal. In everyday life, we also make different decisions all the time, often without thinking about why some of them are successful and others are not. Experience shows that in the case of successful decisions, the goal is usually set and justified correctly, the possibility of achieving it is assessed intuitively correctly, and all reasoning is based on the logic of common sense. There is no doubt that intuition and life experience are quite sufficient for solving the simplest practical tasks of everyday and even managerial activity, which do not require precise calculations. However, personal experience, intuition, and common sense are insufficient for solving complex management problems in economics, social life, as well as in contemporary politics and other types of public activity. Careful analysis of the problem, precise calculations, and construction of mathematical models, including risk models, are necessary.

The most important requirement, which any rational decision must also satisfy, is that all alternatives for choosing a decision must be ordered by an appropriate preference relation, which has the properties of certainty, comparability, and transitivity. Although the methods of modern science make it possible to make increasingly accurate predictions and thus to overcome risks,

uncertainty remains an inevitable companion of human activity. Under these conditions, the problem of risk assessment and forecasting acquires relevance. Therefore, its solution should involve not only traditional probabilistic-statistical methods, but also new research methods that have emerged in the framework of synergetic, nonlinear dynamics and the theory of nonequilibrium systems, as well as expert methods.

**It is important to consider that the risk depends on the target function of the object.**

It can be forced (a farmer living next to a nuclear power plant, riding a bus as a passenger), it can be professional (working at a nuclear power plant, being a bus driver), it can be off system (nuclear power plants for the country, buses as part of transport). The farmer needs to be compensated for possible damage from the accident - for example, in the form of cheap electricity. A professional need to compensate for damage to health, but here the risk is voluntary - he can work in another safer place. With transport it is more complicated - under capitalism in firms a quality specialist is appointed whose main task is to ensure that within the allocated amounts (profit and profitability), for example, the bus is safe, but only within the "warranty" period. That is, it is important to him that everything "falls apart", but after this period. And then the responsibility lies with the professional.

It is the "risk synthesis" within the framework of the concepts presented above that is designed to solve such problems effectively.

## References

- [1] Dimitrov, B. The Axioms in My Understanding from Many Years of Experience. *Axioms* 2021, 10, 176. <https://doi.org/10.3390/axioms10030176>
- [2] Gnedenko B.V., Hinchin A.Y. Elementary Introduction to Probability Theory. Main Editorial Office of Physical and Mathematical Literature of "Nauka" Publisher, 1970. – 168 p.
- [3] Cramer H. Collective Risk Theory. - Stockholm: Skandia Jubilee Volume, 1955.
- [4] Bowers N.L., Gerber H.U., Hickman J.C., Jones D.A., Nesbitt C.J. Actuarial Mathematics. - Itasca, Illinois: The Society of Actuaries, 1986. Russian translation available: Bowers N., Gerber X., Jones D. Nesbitt S., Hickman J. Actuarial mathematics. - Moscow: Janus-K, 2001. – 644 p.
- [5] Panjer H. H., Willmot G. E. Insurance Risk Models. - Schaumburg, IL: The Society of Actuaries, 1992.
- [6] Rotar V. I., Beninet V. E. Introduction to the Mathematical Theory of Insurance // Review of Applied and Industrial Mathematics. Series: Financial and Insurance Mathematics. - 1994. - Ò.I. vol.5 pp.698-779.
- [7] Rykov V. V., Itkin V. Reliability of technical systems and technogenic risk: textbook. - MOSCOW: INFRA-M, 2017. - 192 p. - (Higher education).
- [8] Zhukovskiy V.I., Zhukovskaya L.V. Risk in multicriteria and conflict systems under uncertainty / ed. Ed. 2nd ed. Moscow: Publishing House LKI, 2010. - 272 p.
- [9] Germeyer Y.B. Games with non-opposite interests. Moscow: Nauka, 1976. 327 p.
- [10] Hemming R.W. Numerical methods. Moscow: Nauka, 1972. - 472 p.
- [11] Moore R.E. Interval Analysis. N.Y.: Prentice-Hall, 1966.
- [12] Shokin Y.I. Interval Analysis. Novosibirsk: Siberian Branch of Russian Academy of Sci., 1981 – 112 p.
- [13] Wald A. Statistical Decision Functions. N.Y.: Wiley, 1950.
- [14] Savage L.Y. The theory of statistical decision // J.American Statistic Association. 1951. № 46. pp.55-67.
- [15] Hurwicz L. Optimality criteria for decision making under ignorance, in Cowles Commission Discussion Paper, Statistics. 1951. № 370.
- [16] Frank H. Knight. The Meaning of Risk and Uncertainty. In: F.Knight. Risk, Uncertainty, and Profit. Boston: Houghton Mifflin Co. Translation by S. A. Afontsev.
- [17] Kryanev A. V., Lukin G. Mathematical Methods of Processing of Uncertain Data. -

MOSCOW: FIZMATLIT, 2003. - 216 p. - ISBN 5-9221-0412-8.

- [18] De Groot. M. Optimal statistical solutions. - Moscow: Mir, 1974. - 492 p.
- [19] Sachs S. Theory of Statistical Inference. - M.: The World, 1975. - 776 p.
- [20] Russman I. B., Bermant M. A. On the Problem of Quality Estimation. Journal of Economics and Mathematical Methods, No. 4, 1978, pp. 691-699.
- [21] Russman I. B., Gaidai A. A. Continuous control of goal attainment process. "Management of large systems". Collected works of the Institute of management problems of RAS, Issue 7, Moscow, 2004, pp. 106-113.
- [22] Memorial site of I.Russman. [https://www.adeptis.ru/russman/scientific\\_heritage.html](https://www.adeptis.ru/russman/scientific_heritage.html)
- [23] Bochkov, A.V. Hazard and Risk Assessment and Mitigation for Objects of Critical Infrastructure, pp. 57-135. In: Ram M., Davim J. (eds) Diagnostic Techniques in Industrial Engineering. Management and Industrial Engineering. Springer, Cham, DOI [https://doi.org/10.1007/978-3-319-65497-3\\_3](https://doi.org/10.1007/978-3-319-65497-3_3), Publisher Name: Springer, Cham. - 2017. ISBN 978-3-319-65496-6. - 247 p.
- [24] Bochkov Alexander Vladimirovich. Methodology of ensuring safe functioning and sustainability of the Unified system of gas supply in emergency situations: dissertation ... Doctor of Technical Sciences: 05.26.02 / Bochkov Alexander Vladimirovich; [Place of protection: LLC "Research Institute of natural gases and gas technologies - Gazprom VNIIGAZ"], 2019. - 385 p.
- [25] Zhigirev, N.; Bochkov, A.; Kuzmina, N.; Ridley, A. Introducing a Novel Method for Smart Expansive Systems' Operation Risk Synthesis. Mathematics 2022, 10, 427. <https://doi.org/10.3390/math10030427>
- [26] Petukhov G.B. Fundamentals of the theory of efficiency of purposeful processes. Part 1. Methodology, Methods, and Models. Moscow: Publishing house of the Ministry of Defense of the USSR, 1989. - 647 p.
- [27] Kantorovich L. V. Mathematical and Economic Works. Selected works. Moscow, Nauka, 2011. - 760 p.
- [28] Wasserman F. Neurocomputer Technique: Theory and Practice / Translated from English - Moscow: Mir, 1992. - 240 p.
- [29] Kohonen T. Self-Organizing Maps. Springer-Verlag, 1995.
- [30] Specht D. Probabilistic Neural Networks. Neural Networks, 1990, №1.
- [31] Vyugin V. Mathematical foundations of the theory of machine learning and prediction. - ICMNO, 2013. - 390 p.
- [32] Gladkov L. A., Kureichik V. V., Kureichik V. M. Genetic algorithms: Tutorial. - 2nd ed. - Moscow: Fizmatlit, 2006. - 320 p.
- [33] Taha Khemdi A. Operations research. Williams, 2016. - 912 p.
- [34] Germeyer Y. B. Introduction to the theory of operations research. Moscow - Nauka, 1971. - 383 p.
- [35] Saridis J. Self-organizing stochastic control systems. Moscow: Nauka, 1980. - 397 p.
- [36] Kalmykov S.A., Shokin Y.I., Yuldashev Z.H. Methods of interval analysis. Novosibirsk: Nauka, 1986.
- [37] Zadeh L. A. Fuzzy sets and their application in pattern recognition and cluster analysis // Classification and Cluster / Edited by J. Van Raisen. - M. Mir, 1980. - pp. 208-247.
- [38] Fuzzy sets and possibility theory. Recent advances. Ed. by R.R. Yager. - Moscow: Radio and Communications, 1986. - 408 p.
- [39] Larichev O. I. Theory and Methods of Decision-Making, and Chronicle of Events in Magic Countries. - 3rd ed. - M.: Logos, 2006. - 392 p.
- [40] Fodor J., Roubens M.: Fuzzy preference modelling and multicriteria decision support. (Kluwer Academic Publishers, Dordrecht, 1994).
- [41] Larichev O. I., Petrovsky A. B. Decision support systems. The current state and prospects for their development. //Progress of science and technology. Ser. of Technical Cybernetics. - Vol.21. Moscow: VINITI, 1987, pp. 131-164, [http://www.raai.org/library/papers/Larichev/Larichev\\_Petrovsky\\_1987.pdf](http://www.raai.org/library/papers/Larichev/Larichev_Petrovsky_1987.pdf)

- [42] Saati T. Decision Making. Method of hierarchy analysis // Translated from English by R.G. Vachnadze. Moscow: Radio and communications, 1993. - 278 p.
- [43] Nogin V. D. The simplified version of the method of hierarchy analysis based on nonlinear convolution of criteria // Zhurnal Vychisl. Vol. 44. No.7. pp. 1261-1270.
- [44] Podinovskiy V. V., Podinovskaya O. V. V. On the incorrectness of the method of hierarchy analysis // Problems of Control, 2011, № 1. - pp. 8-13.
- [45] Gusev S. S. Analysis of methods and approaches for solving the problems of multicriteria choice in conditions of uncertainty // Interactive Science. 2018. №1 (23). - pp. 69-75.
- [46] Saati T.L. Decision Making under Dependencies and Feedbacks: Analytical Networks / T.L. Saati. - M.: Librocom Book House, 2009. - 360 p.
- [47] Seredkin K.A. On the limits of applicability of the method of analytical networks in the problems of decision making in natural sciences / K.A. Seredkin // Artificial Intelligence and Decision Making. - 2018. - №2. - pp. 95-102.
- [48] Bochkov A. V., Zhigirev N. N., Ridley A. N. Method for Recovery of Priority Vector of Alternatives in Conditions of Uncertainty or Incompleteness of Expert Evaluations. Reliability. 2017. T. 17. № 3 (62). pp. 41-48.
- [49] Ridley A. N. Methodology of risk synthesis in systems management. In the book: Gagarin readings - 2019 Collection of abstracts of XLV International Youth Scientific Conference. Moscow Aviation Institute (National Research University), 2019.
- [50] Glushkov V.M., Ivanov V.V., Yanenko V.M. Modeling of developing systems. M.: Nauka, 1983. - 350p.
- [51] Glushkov V.M. Introduction to the Theory of Self-Enhancing Systems. Ed. 2, stereotypical. - MOSCOW. LENAND, 2022. - 112 p. (From the heritage of academician V.M. Glushkov; Science of the Artificial. No. 42). 52.
- [52] Automated control system for the scientific-production association (ACS "Ekstremum"): Technical Project / manuscript B.N. Onykiy, manuscript Yu.A. Erivanskiy, responsible executor. L.L. Semyonov, responsible executor. D. V. Mikhailov; the 17th Main Directorate. - M. Book 4: The Subsystem of Technical and Economic Production Planning. - 1974. - 113 p.
- [53] Automated control system of scientific-production association (ACS "Extremum"): Technical Project / hand in hand: B.N. Onykiy, Yu.A. Erivanskiy, responsible executor. A. Erivanskiy Yu. executor; 17 Main Department. - M. Kn. 2: Research and Development Control Subsystem. - 1974. - 124 p.
- [54] Zhigirev N.N. Human-machine procedures for resource allocation in developing systems [Text]: (05.13.06-automated control systems): Dissertation of Candidate of Technical Sciences / N.N. Zhigirev; Sci. Dissertation of doctoral candidate of engineering sciences / supervised by B.N. Onykiy - MIFI, 1987. - 134 p.
- [55] Zhigirev, N.; Bochkov, A.; Kuzmina, N.; Ridley, A. Introducing a Novel Method for Smart Expansive Systems' Operation Risk Synthesis. Mathematics 2022, 10, 427. <https://doi.org/10.3390/math10030427>
- [56] Prigogine I., Lefever R. Symmetry Breaking Instabilities in Dissipative Systems II, Journal of Chemical Physics, 48, 1968. pp. 1695-1700.
- [57] Arnold V. I. Theory of Catastrophes, Moscow: Nauka, 1990. -128 p. ISBN 5-02-014271-9
- [58] Bochkov, A. V., Zhigirev N. N., Ridley A. N. Method for restoring the vector of priorities of alternatives in conditions of uncertainty or incompleteness of expert evaluations. Reliability. 2017. T. 17. № 3 (62). pp.41-48.
- [59] Bochkov, A.V., Lesnykh, V.V., Zhigirev, N.N., Lavrukhin, Yu.N. Some methodical aspects of critical infrastructure protection // Safety Science, Volume 79, November 2015, pp. 229-242. <https://doi.org/10.1016/j.ssci.2015.06.008>
- [60] Bochkov, A.; Niias, J.; Ridley, A.; Kuzmina, N.; Zhigirev, N. Vector compression method to convert the incomplete matrix of pairwise comparisons in the analytic hierarchy process. In Proceedings of the International Symposium on the Analytic Hierarchy Process, Web Conference, December 3-6, 2020; <https://doi.org/10.13033/isahp.y2020.070>.



- [61] Bochkov A., Zhigirev N., Kuzminova A. "Inversion Method of Consistency Measurement Estimation Expert Opinions" // *Reliability: Theory & Applications*, vol. 17, no. 3 (69), 2022, pp. 242-252. doi:10.24412/1932-2321-2022-369-242-252
- [62] Kemeney J., Snell J. *Cybernetic modeling. Some applications*. Moscow: Soviet Radio, 1972. - 192 p.
- [63] Larichev O.I. *Theory and Methods of Decision-Making and Chronicle of Events in the Enchanted Lands*. 2nd edition revised and enlarged. Moscow: Logos, 2002. - 382 p. - ISBN 5-94010-180-1.
- [64] Kaplinsky A.I., Russman I.B., Umyvakin V.M. *Modeling and algorithmization of weakly formalized problems of the best system choice*. Voronezh: Publishing house of the All-Russian State University of Civil Engineering, 1991. - 168 p.
- [65] Markus Schulze, *The Schulze Method of Voting*. Computer Science and Game Theory. Cornell University, 2018. URL: <https://doi.org/10.48550/arXiv.1804.02973>.
- [66] Ralf Pickelmann, 1999. *Das Skating system*. URL: <http://www.tbw.de/rpcs/skating>.
- [67] Ferran Rovira, 1997. *Porqu e ganamos, porqu e perdemos?* TopDance, 15, 16. URL: <http://inicia.es/de/ballrun/skating.htm>.
- [68] Xavier Mora, *The Skating System*. 2nd edition. July 2001. URL: <https://mat.uab.cat/~xmora/escrutini/skating2en.pdf>.
- [69] Lisitsin D.V. *Methods of constructing regression models*. Novosibirsk: NSTU, 2011. - 77 p.
- [70] Kim O. J., Mueller C.W., Klecka W.R., et al. *Factor, discriminant and cluster analysis*. Ed. by I. S. Enyukov. - Moscow: Finances and Statistics, 1989. - 215 p.
- [71] Kenneth J. Arrow, 1951, 2nd ed. *Social Choice and Individual Values*, Yale University Press. ISBN 0-300-01364-7.
- [72] Condorcet J. *Esquisse d'un tableau historique des progres de l'esprit humain*. - Librocom, 2011. - 280 p. - (From the Heritage of World Philosophical Thought. Social Philosophy). ISBN 978-5-397-01568-4.
- [73] Nogin V. 73. D. *Multitude and the Pareto Principle* - SPb: Publishing and Printing Association of Higher Education Institutions, 2022, 2nd edition, revised and supplemented - 111 p.
- [74] Pareto V. *Textbook of political economy*. RIOR, 2018. - 592 p.
- [75] Koch R. *The 80/20 principle*. Exmo, 2012. - 443 p.
- [76] Ross M. Miller, Charles R. Plott, and Vernon L. Smith. *Intertemporal Competitive Equilibrium: An Empirical Study of Speculation*. *Economics. Quarterly Journal of Economics*. Volume 91. Issue 4, November 1977. pp. 599-624. URL: <https://doi.org/10.2307/1885884>
- [77] Coffman A., Henri-Laborder A. *Methods and models of operations research. Integer programming*. Textbook. - Moscow: Mir, 1977. - 432 p.
- [78] Schraever A. *Theory of linear and integer programming*. Monograph in two volumes. Translated from English: Mir, 1991. (360 p.) - 344 p.