

Typical problems of the theory of statistical hypothesis testing are considered. All these problems belong to the same object area and are formulated in a single system of axioms and assumptions using a common linguistic thesaurus. However, different approaches are used to solve each of these problems and a unique solution method is developed. In this regard, the work proposes a unified methodological approach for formulating and solving these problems. The mathematical basis of the approach is the theory of continuous linear programming (CLP), which generalizes the known mathematical apparatus of linear programming for the continuous case. The mathematical apparatus of CLP allows passing from a two-point description of the solution of the problem in the form $\{0; 1\}$ to a continuous one on the segment $[0; 1]$. Theorems justifying the solution of problems in terms of CLP are proved. The problems of testing a simple hypothesis against several equivalent or unequal alternatives are considered. To solve all these problems, a continuous function is introduced that specifies a randomized decision rule leading to continuous linear programming models. As a result, it becomes possible to expand the range of analytically solved problems of the theory of statistical hypothesis testing. In particular, the problem of making a decision based on the maximum power criterion with a fixed type I error probability, with a constraint on the average risk, the problem of testing a simple hypothesis against several alternatives for given type II error probabilities. The method for solving problems of statistical hypothesis testing for the case when more than one observed controlled parameter is used to identify the state is proposed

Keywords: statistical hypothesis testing, uniform method for solving problems, universal randomized criteria

DEVELOPMENT OF MODERN MODELS AND METHODS OF THE THEORY OF STATISTICAL HYPOTHESIS TESTING

L. Raskin

Doctor of Technical Sciences,
Professor, Head of Department*

E-mail: topology@ukr.net

O. Sira

Doctor of Technical Sciences, Professor*

E-mail: topology@ukr.net

*Department of Distributed Information

Systems and Cloud Technologies

National Technical University

«Kharkiv Polytechnic Institute»

Kyrpychova str., 2, Kharkiv, Ukraine, 61002

Received date 05.09.2020

Accepted date 05.10.2020

Published date 30.10.2020

Copyright © 2020, L. Raskin, O. Sira

This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0>)

1. Introduction

Methods for solving numerous problems of identifying the state of objects and making decisions were laid down as a result of the emergence and development of an important area of mathematical statistics – the theory of statistical hypothesis testing. As you know, a statistical hypothesis is any statement about the form or properties of the distribution of experimentally observed random variables. The mathematical theory of statistical hypothesis testing was created in the 30 s of the last century and was significantly developed in a number of fundamental works [1–3]. At the same time, the mathematical formulation and solution of various practical problems of statistical hypothesis testing do not have a single methodological basis and remain an art. This circumstance is an inevitable consequence of the insufficient methodological basis of the theory of statistical hypothesis testing, the development of which virtually no one was engaged in. The lack of a universal approach to solving various problems of this theory leads to the incompleteness of its mathematical framework. This seriously complicates the search and development of possible approaches to new challenges arising from the needs of practice. Thus, the problem of developing a general universal approach to solving problems of the theory of statistical hypothesis testing is urgent. Using

examples of specific problems of the theory of statistical hypothesis testing, we consider traditional technologies to solve them.

2. Literature review and problem statement

The works [1–3] consider the classical problem of testing a simple hypothesis H_0 against a simple alternative H_1 .

When solving such problems of testing hypotheses, it is necessary to construct a statistical criterion that allows making a decision on the degree of their agreement with the hypothesis H_0 on the basis of observation results. Usually, such a criterion is constructed using a critical region, when some function of observations (or observations themselves) falls into it, the hypothesis H_0 is rejected. The problem is to best select this region.

Let the observed parameter $X \in \Omega$ be a random variable with the distribution density $f_0(x/H_0)$, provided that the hypothesis H_0 is true, and $f_1(x/H_1)$, if the alternative hypothesis H_1 is true.

Let us introduce the critical region $\omega \subset \Omega$, when the observed parameter X falls into it, the hypothesis H_0 is rejected (i. e. the hypothesis H_1 is accepted). Then the probability of rejecting the hypothesis H_0 , when it is true, called the

criterion significance level (type I error probability), is determined by the expression:

$$\int_{\omega} f_0(x/H_0)dx = \alpha. \tag{1}$$

In this case, the probability of accepting hypothesis H_0 , when the hypothesis H_1 is true,

$$\beta = \int_{\Omega/\omega} f_1(x/H_1)dx = 1 - \int_{\omega} f_1(x/H_1)dx, \tag{2}$$

is called type II error probability.

Thus, the values of type I and II error probabilities depend on how the critical region ω is chosen. It is clear that the expansion of the region ω leads to an increase in α , but a decrease in β . On the other hand, narrowing the critical region results in opposite results. It is natural to set the problem of determining the critical region being the best in some chosen sense.

Usually, the problem of testing a simple hypothesis H_0 against a simple alternative H_1 comes down to finding a critical region ω for which the probability:

$$\mu(\omega) = 1 - \beta = \int_{\omega} f_1(x/H_1)dx, \tag{3}$$

of rejecting the hypothesis H_0 when the hypothesis H_1 is true (criterion power) takes the highest value, provided that type I error probability is equal to the set value α . This decision criterion is called the Neyman-Pearson criterion [1–3].

Along with the Neyman-Pearson criterion, the so-called Bayesian criteria based on average risk calculation are used in the statistical decision theory [4, 5]. Let it be known a priori that the hypotheses H_0 and H_1 are true with probabilities $P(H_0)$ and $P(H_1)$, respectively, $P(H_0) + P(H_1) = 1$. Then the average risk, depending on the choice of the critical area ω , is determined as follows:

$$R(\omega) = r_0 P(H_0) \int_{\omega} f_0(x/H_0)dx + r_1 P(H_1) \left[1 - \int_{\omega} f_1(x/H_1)dx \right], \tag{4}$$

where r_0, r_1 are some quantitative cost (risk) estimates of type I and II errors, respectively. Now the problem of making a decision according to the Bayesian average risk minimum criterion is to choose such a critical region for the observed parameter X at which the function (4) takes a minimum value.

A natural generalization of the given formulations of the classical problems of the theory of statistical hypothesis testing consists in rejecting the discrete nature of the criteria (acceptance or rejection of the hypothesis). A more general formulation of the hypothesis testing problem is as follows.

The decision rule $A(x)$ is introduced: if the observed parameter has the value x , then the decision on the validity of the hypothesis H_0 is rejected with the probability $A(x)$. Such a rule is called randomized. The function $A(x)$ obviously must satisfy the condition:

$$0 \leq A(x) \leq 1, \quad x \in \Omega. \tag{5}$$

Then the criterion power is written as:

$$\mu(A) = \int_{\Omega} f_1(x/H_1)A(x)dx, \tag{6}$$

the significance level is:

$$\int_{\Omega} f_0(x/H_0)A(x)dx = \alpha, \tag{7}$$

and the average risk function takes the form:

$$R(A) = r_1 P(H_1) + \int_{\Omega} r(x)A(x)dx, \tag{8}$$

where

$$r(x) = r_0 P(H_0) f_0(x/H_0) - r_1 P(H_1) f_1(x/H_1).$$

The relations obtained allow formulating the following typical problems of choosing the optimal randomized decision rule.

Problem 1. Find the function $A^*(x)$ that minimizes the functional (8) and satisfies the condition (5). In this case, the criterion determines the minimum average risk.

Problem 2. Find the function $A^*(x)$ that maximizes the functional (6) and satisfies the constraints (5), (7). In this case, the criterion power is maximized at a given level of significance.

Problem 3. Find the function $A^*(x)$ that maximizes the functional (6), satisfies the condition (5) and the relation:

$$\int_{\Omega} r(x)A(x)dx = R_0 - r_1 P(H_1),$$

where R_0 is the allowable level of average risk. In this case, the criterion power is maximized for a given risk.

Problems 1 and 2 are the usual two-alternative problems of testing a simple hypothesis in a traditional formulation [6]. However, the optimal decision rule is found here in a more general than usual class. Problem 3 is not considered by the classical theory of statistical decisions.

Attempts are known to formulate some problems for statistical hypothesis testing using a randomized criterion. For example, in [7] the problem of maximizing the criterion power provided that the type I error probability does not exceed a given one is formulated as follows. For a given sample of observations $X_{(n)}$, the function $\phi(X_{(n)})$ called critical is introduced, which determines the probability of rejecting the hypothesis H_0 . However, this work lacks any consideration as to how to find such a function. The number of examples of using randomized decision rules to formulate various problems of statistical hypothesis testing can be increased. Some of them propose an approximate solution of the problem, reduced by discretizing the decision space to the application of a non-randomized rule [8, 9], in the rest, the consideration of the problem is limited to correct formulation [10–13]. However, there is no general approach to solving the problems arising here. This circumstance is a consequence of the specific nature of mathematical models of problems of type (5)–(8). By nature, these are isoperimetric problems of the calculus of variations. But it is impossible to use this mathematical apparatus here, since the desired function enters into the integrands of the optimized functional and constraints linearly. The Euler equation does not vanish anywhere.

Thus, it is quite clear that the use of randomized decision rules allows a uniform formulation of traditional problems of statistical hypothesis testing, reducing each of them to solving the corresponding optimization problem. The need to improve the methodological basis of the theory of statistical hypothesis testing is emphasized in [14, 15]. At the same time, the problem of spreading the idea of randomization for solving more complex problems arising with several equivalent or unequal alternatives is important.

These circumstances initiate further research on the problem of statistical hypothesis testing.

3. The aim and objectives of the study

The aim of the study is to develop a general approach to the formulation of problems of statistical hypothesis testing using randomized decision rules. This allows making a reasonable choice of the method for solving optimization problems.

To achieve the aim, the following objectives were set:

- to develop a general mathematical model of the problems of the theory of statistical hypothesis testing using a randomized criterion;

- to develop a mathematical apparatus for optimizing the randomized criterion, adapted for solving traditional problems of statistical hypothesis testing;

- to develop mathematical models and methods for solving an extended set of problems of the theory of statistical hypothesis testing.

4. Development of a general approach to the formulation of problems of the theory of statistical hypothesis testing

Problems 1–3 formulated in terms of (5)–(8) fit into a uniform mathematical model defined as follows: find the function $y(x)$ that maximizes (minimizes) the functional:

$$F(y(x)) = \int_{r_1}^{r_2} g(x)y(x)dx, \quad (9)$$

and satisfying the conditions:

$$\int_{r_1}^{r_2} \eta_i(x)y(x)dx = b_i, \quad i = 1, 2, \dots, m, \quad (10)$$

$$0 \leq y(x) \leq d(x), \quad d(x) > 0, \quad x \in [r_1, r_2], \quad (11)$$

where $g(x)$, $\eta_i(x)$, $d(x)$ are known continuous functions on $[r_1, r_2]$, $r_k \in (0, \infty)$, $k = 1, 2$, $r_1 < r_2$.

A fundamental feature of the obtained mathematical models of the typical problems of the theory of statistical hypothesis testing is the need to find a continuous function that satisfies the system of integral constraints. Solving such a problem requires a special mathematical apparatus.

5. Development of a general method for solving problems of the theory of statistical hypothesis testing

The resulting problem (9)–(11) is a particular case of the general problem of continuous linear programming (CLP) [16].

CLP is a mathematical discipline dealing with the theory and methods of solving constrained optimization problems on a set of continuous functions. In this case, the optimized (objective) function of the problem and the constraints on it are described by integrals (Riemann or Stieltjes) that are linear with respect to the desired function of the continuous argument. Thus, CLP is a continuous generalization of conventional linear programming [17, 18]. To solve the CLP problem, a special mathematical apparatus has been developed [16], which provides an iterative solution. However, in some simple special cases, this solution is achieved directly.

Let us introduce the functions:

$$c(x) = g(x)d(x),$$

$$A(x) = \frac{y(x)}{d(x)},$$

$$a(x) = \eta(x)d(x), \quad m = 1, \quad x \in [r_1, r_2] = R.$$

Then the problem (9)–(11) is simplified to the form: find the function $A^*(x)$ that maximizes (minimizes) the functional:

$$L(A(x)) = \int_{r_1}^{r_2} c(x)A(x)dx, \quad (12)$$

and satisfies the conditions:

$$\int_{r_1}^{r_2} a(x)A(x)dx = b, \quad (13)$$

$$0 \leq A(x) \leq 1. \quad (14)$$

First, we solve an even simpler problem: find the function $A^*(x)$ that maximizes the functional (12) and satisfies only the condition (14). To solve this problem, we use theorem 1.

Theorem 1. The solution of the problem (12), (14) is:

$$A^*(x) = \begin{cases} 1, & c(x) < 0, \\ 0, & c(x) \geq 0, \quad x \in R. \end{cases} \quad (15)$$

Proof.

Let $A^*(x)$ be the function defined by (15), and $A(x)$ – an arbitrary function satisfying (14). We introduce the subset $\tau = \{x : c(x) < 0, \quad x \in R\}$, as well as $R/\tau = \{x : c(x) \geq 0, \quad x \in R\}$.

Then

$$\begin{aligned} \int_R c(x)A(x)dx &= \int_\tau c(x)A(x)dx + \int_{R/\tau} c(x)A(x)dx \geq \\ &\geq \int_\tau c(x)\max\{A^*(x), A(x)\}dx + \\ &+ \int_{R/\tau} c(x)\min\{A^*(x), A(x)\}dx = \int_R c(x)A^*(x)dx. \end{aligned} \quad (16)$$

From the validity of inequality (16) follows the optimality of the function $A^*(x)$. The theorem is proved.

Set the function $c(x)$ as follows:

$$c(x) = r_0 P(H_0) f_0\left(\frac{x}{H_0}\right) - r_1 P(H_1) f_1\left(\frac{x}{H_1}\right).$$

In this case, the functional (12) takes the form:

$$\int_{r_1}^{r_2} c(x)A(x)dx,$$

coinciding, up to a constant, with the average risk function (8), and the function $A(x)$ in accordance with (14), (15) takes the meaning of the rule for rejecting the hypothesis H_0 in problem 1. In accordance with theorem 1, the function $A^*(x)$ that minimizes the average risk is determined by the inequality (15). Thus, the average risk is minimized when $A^*(x)=1$, that is, if $c(x)<0$. In this case, the inequality holds:

$$r_0P(H_0)f_0\left(\frac{x}{H_0}\right)-r_1P(H_1)f_1\left(\frac{x}{H_1}\right)<0,$$

or

$$\frac{f_1\left(\frac{x}{H_1}\right)}{f_0\left(\frac{x}{H_0}\right)}>\frac{r_0P(H_0)}{r_1P(H_1)}. \tag{17}$$

The obtained solution coincides with the classical Bayesian decision rule, according to which the hypothesis H_0 is rejected if the inequality (17) is satisfied (the likelihood ratio exceeds a given threshold).

Let us now return to the original problem: find the function $A^*(x)$ that maximizes the functional (12) and satisfies the conditions (13), (14). The solution of the problem is determined by theorem 2.

Theorem 2.

Introduce the subset:

$$\tau_\lambda = \{x : c(x) > \lambda a(x), x \in R\}, \tag{18}$$

where λ is the root of the equation:

$$\int_{\tau_\lambda} a(x)dx = b. \tag{19}$$

Then the solution of the problem (12)–(14) is:

$$A^*(x) = \begin{cases} 1, & c(x) > \lambda a(x), \\ 0, & c(x) \leq \lambda a(x), \end{cases} \tag{20}$$

or

$$A^*(x) = \begin{cases} 1, & \text{if } x \in \tau_\lambda, \\ 0, & \text{if } x \notin \tau_\lambda. \end{cases}$$

Proof.

In accordance with (19) and (20), we have:

$$\int_R a(x)A^*(x)dx = \int_{\tau_\lambda} a(x)A^*(x)dx = b,$$

and hence the function $A^*(x)$ satisfies the condition (13). Next, we introduce:

$$R/\tau_\lambda = \{x : c(x) \leq \lambda a(x), x \in R\}.$$

Then

$$\begin{aligned} \int_R c(x)A(x)dx &= \int_{\tau_\lambda} c(x)A(x)dx + \int_{R/\tau_\lambda} c(x)A(x)dx = \\ &= \int_{\tau_\lambda} (c(x) - \lambda a(x))A(x)dx + \\ &+ \int_{R/\tau_\lambda} (c(x) - \lambda a(x))A(x)dx + \\ &+ \lambda \int_{\tau_\lambda} a(x)A(x)dx + \lambda \int_{R/\tau_\lambda} a(x)A(x)dx = \\ &= \int_{\tau_\lambda} (c(x) - \lambda a(x))A(x)dx + \\ &+ \int_{R/\tau_\lambda} (c(x) - \lambda a(x))A(x) + \lambda b \leq \\ &\leq \int_{\tau_\lambda} (c(x) - \lambda a(x))\max\{A^*(x), A(x)\}dx + \\ &+ \int_{R/\tau_\lambda} (c(x) - \lambda a(x))\min\{A^*(x), A(x)\}dx + \lambda b = \\ &= \int_{\tau_\lambda} (c(x) - \lambda a(x))A^*(x)dx + \\ &+ \int_{R/\tau_\lambda} (c(x) - a(x))A^*(x)dx + \lambda b = \\ &= \int_{\tau_\lambda} (c(x) - \lambda a(x))A^*(x)dx + \lambda b = \int_R c(x)A^*(x)dx. \end{aligned}$$

Thus, it is shown that the function $A^*(x)$ defined by the relation (20), satisfying the constraints (13), (14), maximizes (12). The theorem is proved.

Let us now give an interpretation of the obtained solution of the problem (12)–(14) in standard terms of the theory of statistical hypothesis testing. We set:

$$c(x) = f_1\left(\frac{x}{H_1}\right), \quad a(x) = f_0\left(\frac{x}{H_0}\right). \tag{21}$$

Then the relations (12) and (13) take the form of (6) and (7), respectively, and problem (12)–(14) will coincide with problem 2 (maximizing the criterion power at a given level of significance). Now, taking into account (21), we define the meaning of the decision rule (20). In these terms:

$$\tau_\lambda = \left\{x : f_1\left(\frac{x}{H_1}\right) > \lambda f_0\left(\frac{x}{H_0}\right)\right\} = \left\{x : \frac{f_1\left(\frac{x}{H_1}\right)}{f_0\left(\frac{x}{H_0}\right)} > \lambda\right\},$$

where λ is the root of the equation:

$$\int_{\tau_\lambda} f_0\left(\frac{x}{H_0}\right)dx = \int_{\frac{f_1\left(\frac{x}{H_1}\right)}{f_0\left(\frac{x}{H_0}\right)} > \lambda} f_0\left(\frac{x}{H_0}\right)dx = \alpha. \tag{22}$$

In accordance with (20)–(22), the hypothesis H_0 should be rejected if the likelihood ratio exceeds a given threshold. In this case, the maximum value of the criterion power is

achieved for a given level of significance. This decision coincides with the Neyman-Pearson decision.

Theorem 2 gives an optimal solution to problem 3. However, the relations in (12)–(14) must be specified as follows:

$$c(x) = f_1\left(\frac{x}{H_1}\right),$$

$$a(x) = r_0P(H_0)f_0\left(\frac{x}{H_0}\right) - r_1P(H_1)f_1\left(\frac{x}{H_1}\right) = r(x),$$

$$b = R_0 - r_1P(H_1).$$

Then the problem (12)–(14) is converted to the form: find the function $A^*(x)$ that maximizes the functional:

$$\int_{r_1}^{r_2} f_1\left(\frac{x}{H_1}\right) A(x) dx, \tag{23}$$

and satisfies the conditions:

$$\int_{r_1}^{r_2} r(x) A(x) dx = R_0 - r_1P(H_1), \tag{24}$$

$$0 \leq A(x) \leq 1, \tag{25}$$

which corresponds to (5), (6), (8). In this case, the solution of the problem (20), defined by theorem 2, is:

$$A^*(x) = \begin{cases} 1, & \frac{f_1\left(\frac{x}{H_1}\right)}{r(x)} > \lambda, \\ 0, & \text{otherwise,} \end{cases} \tag{26}$$

where λ is the solution relative to the equation λ :

$$\int_{\frac{f_1\left(\frac{x}{H_1}\right)}{r(x)}} r(x) dx = R_0 - r_1P(H_1). \tag{27}$$

Thus, it is shown that the traditional problems of the theory of statistical hypothesis testing can be formulated uniformly and solved by continuous linear programming methods.

6. Generalization of the proposed method for solving the problem of statistical hypothesis testing

The proposed mathematical apparatus made it possible to significantly expand the set of solved complex problems of the theory of statistical hypothesis testing. This, in particular, concerns the problem of testing one main hypothesis against several alternatives.

We formulate a mathematical model of such a problem. Let ξ be a random variable with a distribution density $f(\theta, x)$, $x \in \Omega$ depending on a discrete parameter, which can take one of the set of values $\Theta \in \{\theta_0, \theta_1, \theta_2, \dots, \theta_m\}$. Let us introduce a set of hypotheses $H_0, H_1, H_2, \dots, H_m$ regarding the equality of an unknown parameter θ to one of its possible values. We consider a priori probabilities of the implementation of the corresponding hypotheses as given:

$$\{P(H_i) = \text{Prob}\{\theta = \theta_i\}\}, \quad i = 0, 1, 2, \dots, m.$$

The problem is to find a rule according to which, on the basis of observation ξ , the main hypothesis H_0 is accepted or rejected, according to which the obtained measurement belongs to the distribution $f(\theta_0, x)$. The decision is made using the randomized decision rule $0 \leq B(x) \leq 1$, which sets the probability of rejecting the hypothesis H_0 when the observed value x is realized.

Acceptance (rejection) of the main hypothesis is accompanied by the possibility of errors. Moreover, if the hypothesis H_0 is rejected when it is true, then this is the type I error. On the other hand, if the hypothesis H_0 is accepted in a situation where the hypothesis H_i , $i = 1, 2, \dots, m$ is actually true, then this is the i -th type II error.

We introduce $f\left(\frac{x}{H_i}\right)$ – the distribution density of the values of the observed value x , provided that the hypothesis H_i , $i = 0, 1, 2, \dots, m$ is true. Then the conditional probability of the type I error, called, as before, the level of significance, is equal to

$$\alpha_0(B) = \int_{\Omega} f(x/H_0) B(x) dx,$$

and the conditional type II error probabilities are determined by the relation:

$$\beta_i(B) = \int_{\Omega} f(x/H_i)(1-B(x)) dx = 1 - \int_{\Omega} f(x/H_i) B(x) dx. \tag{28}$$

Taking into account a priori data on the possibility of implementing the hypotheses $H_0, H_1, H_2, \dots, H_m$, the unconditional type I error probability $\alpha(B)$ and the unconditional type II error probability $\beta(B)$ are calculated as follows:

$$\alpha(B) = P(H_0) \int_{\Omega} f(x/H_0) B(x) dx, \tag{29}$$

$$\beta(B) = \int_{\Omega} \sum_{i=1}^m P(H_i) f(x/H_i)(1-B(x)) dx = 1 - P(H_0) - \int_{\Omega} \sum_{i=1}^m P(H_i) f(x/H_i) B(x) dx. \tag{30}$$

By analogy with the two-alternative case, in a multi-alternative situation, the concept of criterion power is introduced as the probability of rejecting the hypothesis H_0 when any of the alternative hypotheses is true.

$$\mu(B) = \int_{\Omega} \sum_{i=1}^m P(H_i) f(x/H_i) B(x) dx = \sum_{i=1}^m P(H_i)(1-\beta_i). \tag{31}$$

Let us now specify a set of numbers r_i , $i = 0, 1, 2, \dots, m$, characterizing the loss (risk) associated with the occurrence of corresponding errors. Then the value of the average risk is determined by the following expression:

$$R(B) = r_0P(H_0) \int_{\Omega} f(x/H_0) B(x) dx + \sum_{i=1}^m r_iP(H_i) \int_{\Omega} f(x/H_i)(1-B(x)) dx = \sum_{i=1}^m r_iP(H_i) + \int_{\Omega} r(x) B(x) dx, \tag{32}$$

where

$$r(x) = r_0 P(H_0) f(x/H_0) + \sum_{i=1}^m r_i P(H_i) f(x/H_i). \quad (33)$$

The above relations make it possible to expand the set of problems of statistical hypothesis testing solved by continuous linear programming methods.

Problem 4. Making a decision according to the criterion of minimum average risk: find the function $B^*(x)$ that minimizes the functional:

$$L(B) = \int_{\Omega} r(x) B(x) dx, \quad (34)$$

provided

$$0 \leq B(x) \leq 1, x \in \Omega. \quad (35)$$

Problem 5. Making a decision according to the criterion of maximum power for a fixed unconditional type I error probability: find the function $B^*(x)$ that maximizes the functional:

$$\mu(B) = \int_{\Omega} \sum_{i=1}^m P(H_i) f(x/H_i) B(x) dx, \quad (36)$$

and satisfies the constraints:

$$P(H_0) \int_{\Omega} f(x/H_0) B(x) dx = \alpha, \quad (37)$$

$$0 \leq B(x) \leq 1, x \in \Omega. \quad (38)$$

Problem 6. Making a decision according to the criterion of maximum power with a limited average risk: find the function $B^*(x)$ that maximizes the functional:

$$\mu(B) = \int_{\Omega} \sum_{i=1}^m P(H_i) f(x/H_i) B(x) dx \quad (39)$$

provided

$$\int_{\Omega} r(x) B(x) dx = R_0 - \sum_{i=1}^m r_i P(H_i), \quad (40)$$

$$0 \leq B(x) \leq 1, x \in \Omega. \quad (41)$$

Problem 7. Testing the hypothesis H_0 against m simple alternatives $H_i, i=1, 2, \dots, m$, according to the criterion of the minimum type I error probability for given conditional type II error probabilities: find the function $B^*(x)$ that minimizes the functional:

$$\alpha_0(B) = \int_{\Omega} f_0(x/H_0) B(x) dx \quad (42)$$

and satisfies the constraints:

$$\int_{\Omega} f(x/H_i) B(x) dx \geq 1 - \beta_i, \quad i = 0, 1, 2, \dots, m, \quad (43)$$

$$0 \leq B(x) \leq 1, x \in \Omega. \quad (44)$$

The resulting optimization models belong to the class of continuous linear programming problems with a two-sided

constraint on the values of the plan function, the methods for solving which are considered in [11].

At the same time, problems 4, 5, and 6 are the simplest problems of continuous linear programming and do not differ in structure from problems 1, 2, 3 discussed above.

Problem 7 is more complex – it takes into account possible differences in the risk of different type II errors, which arise if the main hypothesis is accepted in a situation where $H_i, i=1, 2, \dots, m$ is actually true. This is important, since in practice it often matters to which of the alternative hypotheses an error is made. However, this problem also fits into the general mathematical model of canonical problems of continuous linear programming.

Finally, consider a natural generalization of the problem of testing the main hypothesis against several alternatives in the case when all alternatives are equivalent.

Let, as before, ξ be a random variable with the distribution density $\theta(x), x \in \Omega$, (or a multivariate density vector), depending on a discrete parameter that can take one of the set of values $\Theta \in \{\theta_0, \theta_1, \theta_2, \dots, \theta_m\}$.

We introduce a set of alternative hypotheses H_1, H_2, \dots, H_m regarding the equality of the unknown parameter θ to one of its possible values. Let us consider a priori probabilities of the implementation of the corresponding hypotheses as given:

$$\{P(H_i) = \text{Prob}\{\theta = \theta_i\}\}, \quad i = 1, 2, \dots, m.$$

The problem is to find a rule according to which, based on the observation x , one of m hypotheses $H_i, i=1, 2, \dots, m$ is accepted.

The decision is made using the vector randomized decision rule $R(x) = \{R_i(x)\}, i=1, 2, \dots, m$, the value of the component of which $R_i(x)$, when the observed value x is implemented, has the meaning of the probability of accepting the hypothesis H_i .

The introduced decision rule must satisfy the constraints:

$$\sum_{i=0}^m R_i(x) = 1, \quad R_i(x) \geq 0, \quad i = 1, 2, \dots, m.$$

Acceptance or rejection of the hypothesis $H_i, i=1, 2, \dots, m$ is accompanied by errors. Moreover, if the hypothesis H_i is rejected when it is true, then this is the i -th type I error. The number of such errors is m . On the other hand, if the hypothesis H_i is accepted in a situation where $H_j, i=1, 2, \dots, m, j \neq i$ is actually true, this is the type II error of type (ij) . In this situation, there may be $(m^2 - m)$ such errors.

Since $f_i(x/H_i)$ is the distribution density of the values of the parameter x , provided that the hypothesis $H_i, i=1, 2, \dots, m$ is true, the conditional probability of making a correct decision regarding the hypothesis H_i is:

$$\mu_i(R) = \int_{\Omega} f_i(x/H_i) R_i(x) dx, \quad i = 1, 2, \dots, m.$$

The conditional probability of the i -th type I error is calculated by the formula:

$$\alpha_i(R) = \int_{\Omega} f_i(x/H_i) (1 - R_i(x)) dx, \quad i = 1, 2, \dots, m,$$

and the conditional probability of the (ij) -th type II error is determined by the relation:

$$\beta_{ij}(R_i/H_j) = \int_{\Omega} f_j(x/H_j) R_i(x) dx,$$

$$i = 1, 2, \dots, m, \quad j = 1, 2, \dots, m, \quad j \neq i.$$

Let us now define the set g_{ij} , $i = 1, 2, \dots, m, j = 1, 2, \dots, m$ of numbers characterizing the risk associated with the occurrence of the corresponding type II errors. Then the average risk value is determined by the formula:

$$\begin{aligned} G(R) &= \sum_{i=1}^m \sum_{j=1}^m g_{ij} \beta_{ij}(R_i/H_j) = \\ &= \sum_{i=1}^m \sum_{j=1}^m g_{ij} \int_{\Omega} f_j(x/H_j) R_i(x) dx. \end{aligned}$$

Now the decision-making problem can be formulated as follows: find a set of functions $R^* = \{R_i^*(x)\}$ that minimizes the functional:

$$\begin{aligned} G(R) &= \sum_{i=1}^m \int_{\Omega} \sum_{j=1}^m g_{ij} f_j(x/H_j) R_i(x) dx = \\ &= \sum_{i=1}^m \int_{\Omega} c_i(x) R_i(x) dx, \end{aligned}$$

and satisfies the constraints:

$$\begin{aligned} \int_{\Omega} f_i(x/H_i) R_i(x) dx &\geq 1 - \alpha_i, \quad i = 1, 2, \dots, m, \\ \sum_{i=0}^m R_i(x) &= 1, \quad x \in \Omega; \\ R_i(x) &\geq 0, \quad c_i(x) = \sum_{j=1}^m g_{ij} f_j(x/H_j), \quad i = 1, 2, \dots, m. \end{aligned}$$

The obtained optimization model belongs to the class of distribution problems of continuous linear programming, the algorithms for solving which are also considered in [11].

Thus, it is shown that various problems of statistical hypothesis testing regarding the state of an object can be formulated uniformly and solved by continuous linear programming methods.

The completion of the considered cycle of problems is the extension of the described methods to the case when more than one controlled indicator is used to identify the state. The proposed method for solving this problem is as follows.

Let there be a set x_1, x_2, \dots, x_n of controlled indicators, for each of which the function $f\left(\frac{x_k}{H_i}\right)$ – distribution density of the values of the indicator x_k is given, provided that the hypothesis H_i , $k = 1, 2, \dots, n, i = 1, 2, \dots, m$ is true. The problem can be solved iteratively, for example, as follows.

At the first iteration, the indicator x_1 is selected and the problem of statistical hypothesis testing is solved in accordance with mathematical models (9)–(12), (13)–(14), (15)–(17), (18)–(20), (21)–(23), (24)–(27). As a result of solving the corresponding problem, we obtain the distribution of the probabilities of accepting the hypotheses H_i , $i = 1, 2, \dots, m$. This set is used as a priori for the next iteration. The procedure is repeated n times. The resulting set of probabilities $R_i(x_1, x_2, \dots, x_n)$ is a posteriori distribution of the probabilities of the object being in a set of possible states, $i = 1, 2, \dots, m$.

7. Discussion of the results of developing a general method for solving problems of statistical hypothesis testing

A review of the typical problems of statistical hypothesis testing is carried out, which revealed that known traditional methods for solving them are essentially and differently determined by the structure, nature and characteristics of each of these problems. The lack of a uniform technology, a common method for solving various problems of statistical hypothesis testing determines the theoretical and practical usefulness of developing a universal method for solving each of them in each particular case. The result obtained improves the methodological basis of the theory of statistical hypothesis testing, expanding the axiomatic foundation of this theory. The development of a uniform approach to the formulation of typical problems in the theory of statistical hypothesis testing makes it possible to significantly increase the list of such problems, enhancing the practical usefulness of the classical theory. The main result of the study is the development of a method of uniform formulation and solution of various problems of statistical hypothesis testing. The proposed method is based on the use of the mathematical apparatus of continuous linear programming (CLP). This method is a non-trivial generalization of the mathematical apparatus of linear programming by passing from a discrete space to a continuous one. At the same time, the range of solved optimization problems of the theory of statistical hypothesis testing is significantly expanded. These include the following problems. The problems of testing one main hypothesis against one or several alternative hypotheses, as well as problems of testing several equivalent alternative hypotheses are considered. In this case, solutions are sought not in a discrete, but in a more complete continuous class of mathematical descriptions. Theorems defining solutions to all typical problems of statistical hypothesis testing are formulated and proved. The use of the proposed randomized criteria and the mathematical apparatus of CLP significantly expands the range of problems in the theory of statistical hypothesis testing, which can be uniformly solved analytically.

A possible direction for further research is the development of the proposed method for an important case of uncertainty of initial data, when they are given indistinctly [19] or inaccurately [20]. In this case, the approaches proposed in [21, 22] may be useful.

8. Conclusions

1. Using universal randomized rules, a uniform formulation of models of all typical problems of statistical hypothesis testing in terms and on the basis of a modern optimization apparatus (CLP) is made.
2. In terms of continuous linear programming, the technology for the analytical solution of typical problems in the theory of statistical hypothesis testing is developed. In this case, the problems of the theory of statistical hypothesis testing are reduced to a uniform mathematical scheme for optimizing the integral linear functional in the presence of integral linear constraints.
3. The theory and methods of continuous linear programming make it possible to expand the class of analytically solved problems in the theory of statistical hypothesis

testing. In this case, randomized decision rules are used that ensure the transition from discrete two-alternative solutions to a continuous description of the solution in the form of functions that specify the probability of decision making.

The proposed method for solving problems of the theory of statistical hypothesis testing generalizes the known solution methods for the case when more than one independent indicator is used to identify the state of an object.

References

1. Kramer, G. (1975). *Matematicheskie metody statistiki*. Moscow: Mir, 648.
2. Zaks, Sh. (1975). *Teoriya statisticheskikh vyvodov*. Moscow: Mir, 776.
3. Leman, E. (1979). *Proverka statisticheskikh gipotez*. Moscow: Nauka, 408.
4. Rice, J. A. (2007). *Mathematical Statistics and Data Analysis*. Thomson Brooks/Cole, 9.3. Available at: https://portal.tpu.ru/SHARED/d/DIMMASSIKK/academics/Additional_chapters_of_mathematics/Rice%20J.A.%20Mathematical%20Statistics%20and%20Data%20Analysis.pdf
5. Lehmann, E. L., Romano, J. P. (2005). *Testing Statistical Hypotheses*. Springer. doi: <https://doi.org/10.1007/0-387-27605-x>
6. Shvedov, A. S. (2016). *Teoriya veroyatnostey i matematicheskaya statistika*. Moscow: Vysshaya shkola ekonomiki, 286.
7. Grauer, L. V., Arhipova, O. A. (2014). *Proverka statisticheskikh gipotez*. Sankt-Peterburg: CS center.
8. Ashkenazy, V. O. (2004). *Optimal'nye statisticheskie resheniya*. Tver': TGU, 126.
9. Moore, D., McCabe, G. (2003). *Introduction to the Practice of Statistics*. W.H. Freeman, 426.
10. Natan, A. A., Gorbachev, O. G., Guz, S. A. (2005). *Matematicheskaya statistika*. Moscow: MZ PRESS – MFTI, 160.
11. Young, G. A., Smith, R. L. (2005). *Essentials of statistical inference*. Cambridge University Press. doi: <https://doi.org/10.1017/cbo9780511755392>
12. Kobzar', A. I. (2012). *Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov*. Moscow: FIZMATLIT, 816.
13. Kel'bert, M. Ya., Suhov, Yu. M. (2007). *Veroyatnost' i statistika v primerah i zadachah. Vol. 1. Osnovnye ponyatiya teorii veroyatnostey i matematicheskoy statistiki*. Moscow: MTSNMO, 456.
14. *Hypothesis Testing. Introduction*. Available at: <http://www.randomservices.org/random/hypothesis/Introduction.html>
15. Berger, J. O. (2003). Could Fisher, Jeffreys and Neyman Have Agreed on Testing? *Statistical Science*, 18 (1), 1–32. doi: <https://doi.org/10.1214/ss/1056397485>
16. Raskin, L. G., Kirichenko, I. O., Seraya, O. V. (2014). *Prikladnoe kontinual'noe lineynoe programmirovaniye*. Kharkiv: Oberig, 292.
17. Dantsig, Dzh. (1966). *Lineynoe programmirovaniye, ego obobshcheniya i primeneniya*. Moscow: Progress, 581.
18. Yudin, D. B., Gol'shteyn, E. G. (1969). *Lineynoe programmirovaniye. Teoriya, metody i prilozheniya*. Moscow: Nauka, 424.
19. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8 (3), 338–353. doi: [https://doi.org/10.1016/s0019-9958\(65\)90241-x](https://doi.org/10.1016/s0019-9958(65)90241-x)
20. Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11 (5), 341–356. doi: <https://doi.org/10.1007/bf01001956>
21. Raskin, L., Sira, O. (2016). Fuzzy models of rough mathematics. *Eastern-European Journal of Enterprise Technologies*, 6 (4 (84)), 53–60. doi: <https://doi.org/10.15587/1729-4061.2016.86739>
22. Raskin, L., Sira, O. (2016). Method of solving fuzzy problems of mathematical programming. *Eastern-European Journal of Enterprise Technologies*, 5 (4 (83)), 23–28. doi: <https://doi.org/10.15587/1729-4061.2016.81292>