A Neyman-Pearson Proper Way to Universal Testing of Multiple Hypotheses Formed by Groups of Distributions

Evgueni A. Haroutunian and Aram O. Yesayan

Institute for Informatics and Automation Problems of NAS RA
e-mail: eghishe@sci.am, armfrance@yahoo.fr

Abstract

The asymptotically optimal Neyman-Pearson procedures of detection for models characterized by M discrete probability distributions arranged into K, $2 \leq K \leq M$ groups considered as hypotheses are investigated. The sequence of tests based on a growing number of observations is logarithmically asymptotically optimal (LAO) when a certain part of the given error probability exponents (reliabilities) provides positives values for all other reliabilities. LAO tests sequences for some models of objects, including cases, when rejection of decision may be permitted, and when part, or all given error probabilities decrease subexponentially with an increase in the of number of experiments, are desined. For all reliabilities of such tests single-letter characterizations are obtained. A simple case with three distributions and two hypotheses is considered.

Keywords: Statistical hypotheses testing, Families of hypotheses, Optimal detection, Test with no match detection, Neyman-Pearson approach, Neyman-Pearson Lemma, Principle of maximum of Kullback-Leibler distance, Error exponent.

1.   Introduction

This paper is devoted to the generalization of Neyman-Pearson criterion for some specific universal hypotheses testing problem pointed out in the title. In [8] and in the following papers [9], [10], Cox formulated a number of divers examples of problems for two families of hypotheses testing and developed a general modification of the Neyman-Pearson maximum-likelihood ratio procedures for solving such problems. In a series of papers and in disseration of F. Harmosi-nejad and all [27], two stage procedures were investigated for certain models of problems of hypotheses testing. The first stage in these actions executes detecting between families of distributions, and the second stage performs detection of certain distribution in the selected family. Investigation of the present paper can be considered as a more detailed analysis of this first-stage problems.

The asymptotically optimal testing of two hypotheses was investigated by Hoeffding in [32], also the concept of universal hypotheses testing was introduced there.

The hypotheses testing problems for two hypotheses were also studied by Borovkov [6], Levy [35], van Trees [40], Csiszár and Longo [12], Tusnady [39], Longo and Sgarro [36].

Neyman-Pearson criterion of multiple hypotheses testing for discrete random variables was explored in [25]. In publications [1], [24] and [26], many hypotheses logarithmically asymptotically optimal (LAO) testing for the models consisting of many independent objects was investigated. Following Birgé [3], we called the sequence of tests logarithmically asymptotically optimal (LAO), when for given values of some reliabilities (error probability exponents) the test ensures the best values for the rest of them. Haroutunian [18]-[20] investigated the problem of multiple hypotheses testing at the suggestion of R. L. Dobrushin.

Construction of LAO tests sequence is realized applying "Kullback-Leibler balls" around the hypothetic distributions in the space of distributions as sets for detection of corresponding hypotheses. This concept, introduced in [16]-[17] and applied in [20]-[31] and in the present paper, conforms to the idea of "r-divergent sequences" defined in [17] and used in other works.

Hypotheses testing with no-match decision was considered by Gutman [16]. In papers [28]-[30] the results of researches of characteristics of LAO hypotheses testing with possibility of rejection of decision for some models with one or multiple objects, with side information are presented.

Our study is based on information theoretic methods including the method of types. Applications of methods of information theory in mathematical statistics, in particular in hypotheses testing, are exposed in the monographs by Csiszár and Körner [11], Blahut [5], Cover and Thomas [7], Csiszár and Shields [13], Poor [37], Kullback [33] Haroutunian and all [31] , in paper of Blahut [4].

The structure of this paper is as follows. Section 2 contains definitions, notations and problem argument. In central Section 3 the construction of desired LAO tests is exposed for model with groups of distributions and with possibility of rejection of decision. In Section 4, the theorem of the Section 3 is reformulated for the case without rejection option. Section 5 is devoted to the models with some reliabilities equal to 0. In final Section 6, the testing for simplest model with three distributions and two hypotheses is discussed. Conclusion also contains some open problems.

2.   Problem Presentation

Let $\mathcal{P}(\mathcal{X})$ be the space of all probability distributions (PDs) on a finite alphabet $\mathcal{X}$. Let X be a random variable (RV) taking values in the set $\mathcal{X}$ with one of M possible PDs $G_m \in \mathcal{P}(\mathcal{X})$, $m = \overline{1,M}$. Let $\mathbf{x} = (x_1, x_2, ..., x_N)$, $x_n \in \mathcal{X}$, $n = \overline{1,N}$, be a vector of results of N independent observations of the RV X. Then the PD $G_m^N(\mathbf{x}) = \prod_{n=1}^{N} G_m(x_n)$ and $G_m^N \in \mathcal{P}(\mathcal{X}^N)$

The M different PDs are arranged into K, $2 \leq K \leq M$ different groups $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_K$, which we consider as K hypotheses (suppositions) H_k concerning the distributions of the studied object. We consider also an empty "group" $\mathcal{B}_{K+1}$. These groups are mutually disjoint, contain $|\mathcal{B}_1|, |\mathcal{B}_2|, ..., |\mathcal{B}_{K+1}|$ PDs such that

$$\sum_{k=1}^{K} |\mathcal{B}_k| = M, \quad |\mathcal{B}_{K+1}| = 0.$$

When $|\mathcal{B}_k| > 1$, the hypothesis H_k is composite [6], [14], [35]. In applications, the groups may be formed with some different values of parameters of a certain PD.

We study the hypothesis testing problem, which is that to decide, based on the observed sample $\mathbf{x}$, where this vector has originated from a source with a PD from a series $\mathcal{B}_k$, $k = \overline{1, K}$, or to accept $\mathcal{B}_{K+1}$, that is to reject to make any judgement. The procedure is the universal test (do not specializing individual PDs in groups), we denote it by Φ_N [34]. This problem may also be considered as specific task of detection for multiple composite hypotheses, also having the possibility to refuse any decision.

The test Φ_N can be defined by partition of the space $\mathcal{X}^N$ into $K + 1$ disjoint subsets $\mathcal{A}_1^N, \mathcal{A}_2^N, ..., \mathcal{A}_{K+1}^N$, where $\mathcal{A}_k^N$, $k = \overline{1, K}$, contains all vectors $\mathbf{x}$ for which the test adopts the hypothesis H_k, and $\mathcal{A}_{K+1}^N$ includes all vectors $\mathbf{x}$ for which the test refuses to take a certain answer.

We denote by Φ the infinite sequence of tests Φ_N. Let $\alpha_{l|k}(\Phi_N)$ for $l \neq k$, $l = \overline{1, K}$, $k = \overline{1, K}$ be the probability of the erroneous acceptance of the hypothesis H_l by the test Φ_N provided that the hypothesis H_k is true, we define (see [6], [35]):

$$\alpha_{l|k}^N = \alpha_{l|k}(\Phi_N) \triangleq \max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_l^N). \tag{1}$$

When we decline any decision, but the hypothesis H_k is true, we consider the following probability of error:

$$\alpha_{K+1|k}^N = \alpha_{K+1|k}(\Phi_N) \triangleq \max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_{K+1}^N), \; k = \overline{1, K}. \tag{2}$$

The probability of not accepting the true hypothesis H_k, we define in the following way:

$$\alpha_{k|k}^N = \alpha_{k|k}(\Phi_N) \triangleq \sum_{l \neq k, \, l = \overline{1, K+1}} \alpha_{l|k}^N = \max_{G_m \in \mathcal{B}_k} G_m^N(\overline{\mathcal{A}_k^N}), \; k = \overline{1, K}. \tag{3}$$

Note that our approach differs from the approaches in [6], where only $\alpha_{k|k}^N$ are studied, and in [38] where the $\alpha_{k|k}^N$ are not considered.

We study the corresponding reliabilities (error probabilities exponents) $E_{l|k}$ of the tests sequence Φ:

$$E_{l|k} = E_{l|k}(\Phi) \triangleq \lim_{N \to \infty} \left(-\frac{1}{N} \log \alpha_{l|k}^N \right), \; k = \overline{1, K}, \, l = \overline{1, K+1}. \tag{4}$$

All reliabilities are arranged in $(K + 1) \times K$ matrix. For instance, at $K = 3$ the matrix of reliabilities has the following form

$$\mathbf{E}(\Phi) = \begin{pmatrix} E_{1|1} & E_{2|1} & E_{3|1} & E_{4|1} \\ E_{1|2} & E_{2|2} & E_{3|2} & E_{4|2} \\ E_{1|3} & E_{2|3} & E_{3|3} & E_{4|3} \end{pmatrix}.$$

Definitions (3) and (4) imply that

$$E_{k|k} = \min_{l \neq k, \, l = \overline{1, K+1}} E_{l|k}, \; k = \overline{1, K}. \tag{5}$$

We call the tests sequence Φ^* logarithmically asymptotically optimal (LAO) for this model if for given positive values of certain K elements of the reliabilities matrix $\mathbf{E}(\Phi^*)$ the

procedure Φ^* provides maximal values for all other elements of it [3]. This criterion can be considered as a proper specification of the Neyman-Pearson approach to the universal test of multiple hypotheses in the sense of optimality of reliabilities. In certain publications, the LAO approach is referred to as the "exponential rate optimal" (ERO) [32], [16], [39].

In opposition to the criterion adopted by Gutman [16], we recognize the asymmetry in the importance of different hypotheses and consider unequal requirements to error probabilities, or reliabilities of their detection.

We use the following notions and notations:

Shannon entropy of PD P on alphabet $\mathcal{X}$:

$$H(P) = -\sum_x P(x) \log P(x),$$

divergence, Kullback-Leibler information, relative entropy, or "distance" of two PDs P_1 and P_2 on $\mathcal{X}$:

$$D(P_1||P_2) = \sum_x P_1(x) \log \frac{P_1(x)}{P_2(x)},$$

a new notion introduced in [22], divergence of three PDs P_1, P_2, P_3 on $\mathcal{X}$:

$$D(P_1||P_2||P_3) \triangleq \sum_x P_1(x) \log \frac{P_2(x)}{P_3(x)} = D(P_1||P_2) - D(P_1||P_3).$$

As was noted in introduction, our study applies the method of types, developed in information theory [7, 11, 13, 31]. The basic notion in this method is the notion of the type $Q_{\mathbf{x}}$ of the vector $\mathbf{x} \in \mathcal{X}^N$, which is equivalent to the statistical notion of the empirical distribution of the sample $\mathbf{x}$:

$$Q_{\mathbf{x}} = \{Q_{\mathbf{x}}(x) = N(x/\mathbf{x})/N, x \in \mathcal{X}\},$$

where $N(x/\mathbf{x})$ is the number of repetitions of the element x in the sample $\mathbf{x}$. We denote by $\mathcal{Q}(\mathcal{X}^N)$ the set of all possible types on $\mathcal{X}^N$. It is clear that $\mathcal{Q}(\mathcal{X}^N) \subset \mathcal{P}(\mathcal{X})$.

We will denote divergence by $D^N(Q||P)$ when $Q \in \mathcal{Q}(\mathcal{X}^N)$ and $P \in \mathcal{P}(\mathcal{X})$. Note that $D^N(Q||P) \to D(Q||P)$, when $N \to \infty$.

Let $\mathcal{T}_Q^N(\mathcal{X})$ be the family of all vectors $\mathbf{x}$ of the type Q. For $Q \notin \mathcal{Q}(\mathcal{X}^N)$, we have $\mathcal{T}_Q^N(\mathcal{X}) = \emptyset$. We will use the following estimates [11], [31]:

$$|\mathcal{Q}(\mathcal{X}^N)| \leq (N+1)^{|\mathcal{X}|}, \tag{6}$$

$$(N+1)^{-|\mathcal{X}|} \exp\{NH(Q)\} \leq |\mathcal{T}_Q^N(\mathcal{X})| \leq \exp\{NH(Q)\}. \tag{7}$$

We will denote for brevity:

$$\text{for } Q \in \mathcal{P}(\mathcal{X}), \ D(Q||\mathcal{B}_k) \triangleq \min_{G_m \in \mathcal{B}_k} D(Q||G_m), \tag{8}$$

$$\text{and for } Q \in \mathcal{Q}(\mathcal{X}^N), \ D^N(Q||\mathcal{B}_k) \triangleq \min_{G_m \in \mathcal{B}_k} D^N(Q||G_m), \tag{9}$$

$$\text{for } \mathcal{R}_l \subset \mathcal{P}(\mathcal{X}), \ D(\mathcal{R}_l||\mathcal{B}_k) \triangleq \min_{Q \in \mathcal{R}_l} D(Q||\mathcal{B}_k), \tag{10}$$

$$\text{and for } \mathcal{R}_l^N \subset \mathcal{Q}(\mathcal{X}^N), \ D^N(\mathcal{R}_l^N||\mathcal{B}_k) \triangleq \min_{Q \in \mathcal{R}_l^N} D^N(Q||\mathcal{B}_k). \tag{11}$$

In the following sections, we present ways of optimal tests construction for the considered models and investigate the corresponding error probabilities and reliabilities.

3.   Testing for Model with Rejection Option

To construct the desired LAO test corresponding to preliminary given strictly positive numbers $E_{1|1}, E_{2|2}, ..., E_{K|K}$ we define the following subsets of distributions:

$$\mathcal{R}_k^N \triangleq \{Q \in \mathcal{Q}(\mathcal{X}^N) : D^N(Q||\mathcal{B}_k) \le E_{k|k}\}, \;\; k = \overline{1, K}, \tag{12}$$

$$\mathcal{R}_{K+1}^N \triangleq \{Q \in \mathcal{Q}(\mathcal{X}^N) : D^N(Q||\mathcal{B}_k) > E_{k|k}, \;\; k = \overline{1, K}\}, \tag{13}$$

$$\mathcal{R}_k \triangleq \{Q \in \mathcal{P}(\mathcal{X}) : D(Q||\mathcal{B}_k) \le E_{k|k}\}, \;\; k = \overline{1, K}, \tag{14}$$

$$\mathcal{R}_{K+1} \triangleq \{Q \in \mathcal{P}(\mathcal{X}) : D(Q||\mathcal{B}_k) > E_{k|k}, \;\; k = \overline{1, K}\}, \tag{15}$$

It is clear that

$$\mathcal{R}_k^N \subset \mathcal{R}_k, \; k = \overline{1, K+1}. \tag{16}$$

Define also the following values of reliabilities:

$$E_{k|k}^* = E_{k|k}^*(E_{k|k}) \triangleq E_{k|k}, \;\; k = \overline{1, K}, \tag{17}$$

$$E_{l|k}^* = E_{l|k}^*(E_{l|l}) \triangleq D(\mathcal{R}_l||\mathcal{B}_k), \;\; k = \overline{1, K}, \;\; k \ne l, \;\; l = \overline{1, K}, \tag{18}$$

$$E_{K+1|k}^* = E_{K+1|k}^*(E_{1|1}, E_{2|2}, ..., E_{K|K}) \triangleq D(\mathcal{R}_{K+1}||\mathcal{B}_k) = E_{k|k}^*, \;\; k = \overline{1, K}. \tag{19}$$

Theorem 1: *If all distributions G_m, $m = \overline{1, M}$, are different in the sense that $D(G_{m'}||G_m) > 0$, $m' \ne m$, and the strictly positive numbers $E_{1|1}, E_{2|2}, ..., E_{K|K}$ are such that the following inequalities hold*

$$E_{1|1}^* < \min_{l=\overline{2,K}} D(\mathcal{R}_l||\mathcal{B}_1) \tag{20}$$

$$E_{k|k}^* < \min(\min_{l=\overline{1,k-1}} E_{l|k}^*(E_{l|l}), \min_{l=\overline{k+1,K}} D(\mathcal{R}_l||\mathcal{B}_k)) \tag{20'}$$

$$E_{K|K}^* < \min_{l=\overline{1,K-1}} E_{l|K}^*(E_{l|l}) \tag{20''}$$

then there exists an LAO sequence of tests, all elements of the reliability matrix $\mathbf{E}^* = \{E_{l|k}^*\}$ of which are defined in (17)-(19) and are strictly positive.

When at least one of the inequalities in (20) is violated, then at least one element of the matrix of reliabilities $\mathbf{E}^*$ is equal to 0. More than that, if we try to detect with such $E_{l|l}$ which for some $l \in [1; K+1]$ and $k \in [1; K]$ is greater than $D(\mathcal{R}_l||\mathcal{B}_k)$, then the test for all $N = 1, 2, ...$ will make an error with the probability 1.

Proof: Having a collection of numbers satisfying the conditions (20) we pass to the proof of the positive statement of the theorem, that is, to the construction of the test.

Consider a sequence of tests Φ^*, which is defined by partition of sample space $\mathcal{X}^N$ on the following $K + 1$ subsets:

$$\mathcal{A}_k^{N*} = \bigcup_{Q \in \mathcal{R}_k^N} \mathcal{T}_Q^N(\mathcal{X}), \;\; k = \overline{1, K},$$

$$\tag{21}$$

$$\mathcal{A}_{K+1}^{N*} = \mathcal{X}^N - \bigcup_{k=1}^{K} \mathcal{A}_k^{N*}, \;\; N = 1, 2, ...$$

This $\mathcal{A}_k^{N*} \subset \mathcal{X}^N$, because when $Q \notin \mathcal{Q}(\mathcal{X}^N)$, then $\mathcal{T}_Q^N(\mathcal{X})$ is empty. Let us prove that the collection of sets in (21) determines a test, namely, each $\mathbf{x}$ belongs to one and only to one subset $\mathcal{A}_k^{N*}$,

$$\mathcal{A}_k^{N*} \bigcap \mathcal{A}_r^{N*} = \emptyset, \ r \neq k, \ \text{ and } \ \sum_{k=1}^{K+1} \mathcal{A}_k^{N*} = \mathcal{X}^N.$$

Really, for $k = \overline{2, K}$, $r = \overline{1, K-1}$, for each $k > r$ let us consider arbitrary $\mathbf{x} \in \mathcal{A}_k^{N*}$. We see that in accordance with (21) and (12) there exists $\mathcal{T}_{Q_{\mathbf{x}}}^N(\mathcal{X}) \subset \mathcal{A}_k^{N*}$, such that $D^N(Q_{\mathbf{x}}||\mathcal{B}_k) \leq E_{k|k}$. As $k > r$ from conditions (20) it follows that $E_{r|r} < E_{k|r}^*(E_{k|k})$. From definition (18) and inequality $D^N(Q_{\mathbf{x}}||\mathcal{B}_k) \leq E_{k|k}$ we obtain $E_{r|r} < E_{k|r}^*(E_{k|k}) = \min_{Q \in \mathcal{R}_k^N} D(Q||\mathcal{B}_r) \leq D^N(Q_{\mathbf{x}}||\mathcal{B}_r)$.

Hence $Q_{\mathbf{x}} \notin \mathcal{R}_r^N$ and from (21) it follows that $\mathbf{x} \notin \mathcal{A}_r^{N*}$.

We can verify that $\mathcal{A}_{K+1}^{N*} \bigcap \mathcal{A}_k^{N*} = \emptyset$, $k = \overline{1, K}$, because if $\mathbf{x} \in \mathcal{A}_{K+1}^{N*}$, then by (15) for type $Q_{\mathbf{x}}$ the inequality $D^N(Q_{\mathbf{x}}||\mathcal{B}_k) > E_{k|k}$ is true for $k = \overline{1, K}$. According to the definition (21) of $\mathcal{A}_k^{N*}$, $k = \overline{1, K}$, we see that $\mathbf{x} \notin \mathcal{A}_k^{N*}$.

The sample $\mathbf{x}$ from $\mathcal{T}_Q^N(\mathcal{X}) \subset \mathcal{Q}(\mathcal{X}^N)$ has the following probability:

$$G_m^N(\mathbf{x}) = \prod_{n=1}^N G_m(x_n) = \prod_x G_m(x)^{N(x/\mathbf{x})} = \prod_x G_m(x)^{NQ_{\mathbf{x}}(x)}$$

$$= \exp\{N \sum_x (-Q_{\mathbf{x}}(x) \log \frac{Q_{\mathbf{x}}(x)}{G_m(x)} + Q_{\mathbf{x}}(x) \log Q_{\mathbf{x}}(x))\}$$

$$= \exp\{-N[D^N(Q_{\mathbf{x}}||G_m) + H(Q_{\mathbf{x}})]\}. \tag{22}$$

Now for $k = \overline{1, K}$, using (3), (21), (6), (7), (12) and (22) we can upper estimate $\alpha_{k|k}(\Phi_N^*)$ as follows:

$$\alpha_{k|k}(\Phi_N^*) = \max_{G_m \in \mathcal{B}_k} G_m^N(\overline{\mathcal{A}_k^{N*}}) = \max_{G_m \in \mathcal{B}_k} G_m^N(\bigcup_{Q \notin \mathcal{R}_k^N} \mathcal{T}_Q^N(\mathcal{X}))$$

$$\leq (N+1)^{|\mathcal{X}|} \max_{G_m \in \mathcal{B}_k} \max_{Q: D^N(Q||\mathcal{B}_k) > E_{k|k}} G_m^N(\mathcal{T}_Q^N(\mathcal{X}))$$

$$\leq (N+1)^{|\mathcal{X}|} \max_{Q: D^N(Q||\mathcal{B}_k) > E_{k|k}} \exp\{-ND^N(Q||\mathcal{B}_k)\}$$

$$\leq \exp\{-N[\inf_{Q: D^N(Q||\mathcal{B}_k) > E_{k|k}} D^N(Q||\mathcal{B}_k) - o_N(1)]\}$$

$$= \exp\{-N(E_{k|k} - o_N(1))\}, \tag{23}$$

where $o_N(1) \to 0$ with $N \to \infty$. From here $E_{k|k}^* \geq E_{k|k}$, $k = \overline{1, K}$.

Now let us prove the lower inequalities for $l = \overline{1, K}$, $k = \overline{1, K}$, $l \neq k$. From (1), (21), (7) and (22) we obtain

$$\alpha_{k|k}(\Phi_N^*) = \max_{G_m \in \mathcal{B}_k} G_m^N(\overline{\mathcal{A}_k^{N*}}) = \max_{G_m \in \mathcal{B}_k} G_m^N(\bigcup_{Q \notin \mathcal{R}_k^N} \mathcal{T}_Q^N(\mathcal{X}))$$

$$\geq \max_{G_m \in \mathcal{B}_k} \max_{Q \notin \mathcal{R}_k^N} G_m^N(\mathcal{T}_Q^N(\mathcal{X}))$$

$$\geq (N+1)^{-|\mathcal{X}|} \max_{G_m \in \mathcal{B}_k} \max_{Q: D^N(Q||\mathcal{B}_k) > E_{k|k}} \exp\{-ND^N(Q||G_m)\}$$

$$= \exp\{-N(\min_{G_m \in \mathcal{B}_k} \inf_{Q: D^N(Q||\mathcal{B}_k) > E_{k|k}} D^N(Q||G_m) + o(1))\}$$

$$= \exp\{-N(\inf_{Q:D^N(Q||\mathcal{B}_k)>E_{k|k}} D^N(Q||B_k) + o(1))\}$$

$$= \exp\{-N(E_{k|k} + o(1))\}. \tag{24}$$

(23) and (24) give us (17). We can obtain similar upper estimates $E^*_{l|k} \geq E_{l|k}$ for $l = \overline{1,K}$, $k = \overline{1,K}$, $l \neq k$. According to (1), (6), (7) and (10) we have

$$\alpha_{l|k}(\Phi^*_N) = \max_{G_m \in \mathcal{B}_k} G^N_m(\mathcal{A}^{N*}_l) = \max_{G_m \in \mathcal{B}_k} G^N_m(\bigcup_{Q \in \mathcal{R}^N_l} \mathcal{T}^N_Q(\mathcal{X}))$$

$$\leq (N+1)^{|\mathcal{X}|} \max_{G_m \in \mathcal{B}_k} \max_{Q \in \mathcal{R}^N_l} \exp\{-ND^N(Q||G_m)\}$$

$$= \exp\{-N(D^N(\mathcal{R}_l||\mathcal{B}_k) - o_N(1))\}. \tag{25}$$

Again for $l = \overline{1,K}, k = \overline{1,K}, l \neq k$, we lower estimate

$$\alpha_{l|k}(\Phi^*_N) = \max_{G_m \in \mathcal{B}_k} G^N_m(\mathcal{A}^{N*}_l) = \max_{G_m \in \mathcal{B}_k} G^N_m(\bigcup_{Q \in \mathcal{R}^N_l} \mathcal{T}^N_Q(\mathcal{X}))$$

$$\geq (N+1)^{-|\mathcal{X}|} \max_{G_m \in \mathcal{B}_k} \max_{Q \in \mathcal{R}^N_l} \exp\{-ND(Q||G_m)\}$$

$$= \exp\{-N(D(\mathcal{R}_l||\mathcal{B}_k) + o_N(1))\} = \exp\{-N(E_{l|k} + o_N(1))\}. \tag{26}$$

According to the definition (4), the reliability $E_{l|k}(\Phi^*)$ of the test sequence Φ^* is the limit inferiour $\varliminf_{N \to \infty}(-\frac{1}{N} \log \alpha_{l|k}(\Phi^*_N))$, taking into account (25), (26) and the continuity of the functional $D^N(Q||G_l)$, we obtain that $\lim_{N \to \infty}(-\frac{1}{N} \log \alpha_{l|k}(\Phi^*_N))$ exists and (18) is correct.

Similarly we can obtain upper and lower bounds for $\alpha_{K+1|k}(\Phi^*_N)$, $k = \overline{1,K}$. Applying the analogous resoning we get (19).

The proof of the first part of the theorem will be accomplished if we demonstrate that the sequence of tests Φ^* is LAO, that is, for every other sequence of tests Φ^{**} with the same reliabilities $E_{1|1}, ..., E_{K|K}$ for all $l = \overline{1, K+1}$, $l \neq k$, $k = \overline{1,K}$, inequalities $E_{l|k}(\Phi^{**}) \leq E_{l|k}(\Phi^*)$ hold. Suppose the contrary is the case, that is there exists sequence of tests Φ^{**} defined by the sets $\mathcal{D}^N_1, ..., \mathcal{D}^N_{K+1}$ such that

$$E_{l|k}(\Phi^{**}) > E_{l|k}(\Phi^*) \text{ for some } l \in [1, K+1], \ k \in [1, K], \ l \neq k. \tag{27}$$

For tests Φ^* and Φ^{**} the space $\mathcal{X}^N$ is decomposed into subsets $\mathcal{A}^{N*}_l$ and, respectively, into $\mathcal{D}^N_l$, $l = \overline{1, K+1}$, such that for $l = \overline{1,K}$ and N large enough.

$$\max_{G_m \in \mathcal{B}_l} G^N_m(\mathcal{A}^{N*}_l) = \max_{G_m \in \mathcal{B}_l} G^N_m(\mathcal{D}^N_l) = 1 - \exp\{-NE_{l|l}\} \tag{28}$$

and $\mathcal{A}^{N*}_l$ are constructed in (21) with sets of types $\mathcal{T}^N_Q(\mathcal{X})$ including almost all vectors $\mathbf{x}$ having positive probability $\max_{G_m \in \mathcal{B}_l} G_m$. From here for the set $\mathcal{A}^{N*}_l - \mathcal{D}^N_l \subset \overline{\mathcal{D}^{N*}_l}$ we have

$$\lim_{N \to \infty} \max_{G_m \in \mathcal{B}_l} G^N_m(\mathcal{A}^{N*}_l - \mathcal{D}^N_l) \leq \lim_{N \to \infty} \exp(-NE_{l|l}) = 0. \tag{29}$$

By (4), (1) and (27)

$$E_{l|k}(\Phi^{**}) = \varliminf_{N \to \infty} \left(-\frac{1}{N} \log \max_{G_m \in \mathcal{B}_k} G^N_m(\mathcal{D}^N_l)\right)$$

$$\geq \lim_{N \to \infty} \left(-\frac{1}{N} \log \max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_l^N) \right).$$

That is for N large enough

$$\max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{D}_l^N) \leq \max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_l^{N*}),$$

and hence

$$\max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_l^{N*} - \mathcal{D}_l^N) > 0,$$

which contradicts (29). So, we must conclude that (27) is not possible.

For the proof of the second part of Theorem 1, it is enough to note that if one of the conditions (20) is violated, then from (12)-(19) it follows that at least one of the elements $E_{l|k}$ is equal to 0. In case when $E_{l|l} > D(\mathcal{R}_l || \mathcal{B}_k)$ from (3), (21), (12), (22), (7) we have for all N

$$\alpha_{l|l}^N = \max_{G_m \in \mathcal{B}_l} G_m^N(\mathcal{A}_l^N) = \max_{G_m \in \mathcal{B}_l} G_m^N \left(\bigcup_{Q \in \mathcal{R}_k^N} \mathcal{T}_Q^N(\mathcal{X}) \right) \geq \max_{G_m \in \mathcal{B}_l} \max_{Q \in \mathcal{R}_l^N} G_m^N(\mathcal{T}_Q^N(\mathcal{X}))$$

$$= \exp\{-N \min_{G_m \in \mathcal{B}_l} \min_{Q \in \mathcal{R}_l^N} D^N(Q || G_m)\} = \exp\{-N \times 0\} = 1,$$

because for $G_m \in \mathcal{B}_l^N$ we have $\min_{G_m \in \mathcal{B}_l} \min_{Q \in \mathcal{R}_l^N} D^N(Q || G_m) = 0$. Theorem 1 is proved.

4.   Case without Rejection of Decision

Consider also the standard case when the decision is obligatory. Again we have M possible PDs $G_m \in \mathcal{P}(\mathcal{X})$, $m = \overline{1, M}$, which are placed in K groups $\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_K$, which we envisage as hypotheses H_k, $k = \overline{1, K}$. The unknown hypothesis must be detected on the base of sample $\mathbf{x} = (x_1, x_2, ..., x_N)$. The test Φ_N can be designed by dividing the sample space $\mathcal{X}^N$ into K subsets $\mathcal{A}_1^N, \mathcal{A}_2^N, ..., \mathcal{A}_K^N$ as acceptance regions for the hypotheses of the same number. The test is characterized by error probabilities.

$$\alpha_{l|k}^N = \alpha_{l|k}(\Phi_N) \triangleq \max_{G_m \in \mathcal{B}_k} G_m^N(\mathcal{A}_l^N), \ \ l \neq k, \ \ l, k = \overline{1, K}.$$

$$\alpha_{k|k}^N = \alpha_{k|k}(\Phi_N) \triangleq \sum_{l \neq k, \ l = \overline{1, K}} \alpha_{l|k}^N, \ \ k = \overline{1, K}.$$

The reliabilities are difined as in (4)

$$E_{l|k} = E_{l|k}(\Phi) \triangleq \lim_{N \to \infty} \left(-\frac{1}{N} \log \alpha_{l|k}^N \right), \ \ k, l = \overline{1, K}.$$

We shape the LAO sequence of tests Φ for preliminary given positive numbers $E_{1|1}, E_{2|2}, ..., E_{K-1|K-1}$ by the following regions of PDs

$$\mathcal{R}_k \triangleq \{Q \in \mathcal{P}(\mathcal{X}), D(Q || \mathcal{B}_k) \leq E_{k|k}\}, \ \ k = \overline{1, K-1},$$

$$\mathcal{R}_K \triangleq \{Q \in \mathcal{P}(\mathcal{X}), D(Q || \mathcal{B}_k) > E_{k|k}, \ \ k = \overline{1, K-1}\}.$$

Let the corresponding reliabilities be as in (17)-(19)

$$E_{k|k}^* = E_{k|k}^*(E_{k|k}) \triangleq E_{k|k}, \quad k = \overline{1, K-1}, \tag{30}$$

$$E_{l|k}^* = E_{l|k}^*(E_{l|l}) \triangleq D(\mathcal{R}_l||\mathcal{B}_k), \quad k = \overline{1, K}, \ k \neq l, \ l = \overline{1, K-1}, \tag{31}$$

$$E_{K|k}^* = E_{K|k}^*(E_{1|1}, E_{2|2}, ..., E_{K-1|K-1}) \triangleq D(\mathcal{R}_K||\mathcal{B}_k), \quad k = \overline{1, K}. \tag{32}$$

Theorem 2: *If all PDs G_m, $m = \overline{1, M}$, are different, $D(G_{m'}||G_m) > 0$, $m' \neq m$, and the strictly positive numbers $E_{1|1}, E_{2|2}, ..., E_{K-1|K-1}$ are such that the following inequlities hold*

$$E_{1|1}^* < \min_{l=\overline{2,K-1}} D(\mathcal{R}_l||\mathcal{B}_1)) \tag{33}$$

$$E_{k|k}^* < \min(\min_{l=\overline{1,k-1}} E_{l|k}^*(E_{l|l}^*), \min_{l=\overline{k+1,K-1}} D(\mathcal{R}_l||\mathcal{B}_k)) \, k = \overline{2, K-2}, \tag{33$'$}$$

$$E_{K-1|K-1}^* < \min_{l=\overline{1,K-2}} E_{l|K-1}^*(E_{l|l}^*) \tag{33$''$}$$

*then there exists an LAO sequence of tests, all elements of the reliabilities matrix **E** of which are defined in (30)-(32) and are strictly positive. When one of the inequalities in (33) is violated then at least one element of the matrix of reliabilities **E*** is equal to 0.*

5.   Some or All Given Reliabilities are Equal to Zero

The well-known Stein lemma [11] also called Chernoff-Stein lemma in [7] provides the estimate of the error probability for the case of two hypotheses. It asserts that when the error probability $\alpha_{1|1}^N$ is postulated as a constant, then the error probability $\alpha_{2|2}^N$ goes to 0 as $\exp\{-ND(P_1||P_2)\}$ as the number of observations N tends to infinity.

In this section, we present a generalization of Stein lemma in two directions. First, a more general model, when Section 2 consists of M PDs grouped in K hypotheses and the test has to detect an unknown hypothesis or reject any decision. And secondly, it is known that some error probabilities $\alpha_{l|l}^N$, or all K of them, tend to 0 when N goes to infinity as a function $\delta_{l|l}^N$, such that

$$\lim_{N \to \infty} \left(-\frac{1}{N} \log \delta_{l|l}^N\right) = 0. \tag{34}$$

In practice, $\delta_{l|l}^N$ can be constants or polynomials by N. The following theorem is a generalization of a result from [23] as an addition to Theorem 1.

Theorem 3: *When all distributions G_m, $m = \overline{1, M}$ are different in the sense that $D(G_{m'}||G_m) > 0$, $m' \neq m$ and given numbers $E_{k|k}$, $k = \overline{1, K}$ partly or all of them are equal to 0 and verify condition (34), then there exists an LAO test sequence Φ^*, the elements of reliabilities matrix $\mathbf{E}(\Phi^*) = \{E_{l|k}^*\}$ of which are defined by (14), (15), (17)-(19), if the conditions (20) hold. But in the case when the given $E_{l|l}$ is equal to zero, the formula (18) changes as follows:*

$$E_{l|k}^* = E_{l|k}^*(0) = D(\mathcal{B}_l||\mathcal{B}_k), \quad k = \overline{1, K}, \ k \neq l.$$

For the proof, it is enough to replace (12) by the following expressions:

$$\mathcal{R}_l^N \triangleq \{Q \in \mathcal{Q}(\mathcal{X}^N) : D^N(Q||\mathcal{B}_l) \leq -\frac{1}{N} \log \delta_{l|l}^N\}.$$

6.   Case of Three Distributions and Two Hypotheses

In this Section, we discuss a number of questions concerning the most simple model amongst the considered in this paper. At first we represent a generalization of the fundamental result of Neyman-Pearson lemma for the noted case of 3 PDs and two groups. There are given three distributions G_1, G_2, G_3 for a random variable X. These distributions are divided into two groups (hypotheses) such that the first hypothesis H_1 is the first distribution and the second hypothesis is the group of two other PDs

$$H_1 = (G_1), \quad H_2 = (G_2, G_3). \tag{35}$$

The statistician must accept or reject the first hypothesis on the base of the sample $\mathbf{x}$.

Theorem 4: (Neyman-Pearson lemma) *For a threshold $t > 1$, consider test Ψ_N^* defined by the region of acceptance $\mathcal{A}^{N*}$ for hypothesis H_1:*

$$\mathcal{A}^{N*} = \{\mathbf{x} : \frac{G_1^N(\mathbf{x})}{\max(G_2{}^N(\mathbf{x}); G_3{}^N(\mathbf{x}))} > t\}, \tag{36}$$

and acceptance region $\overline{\mathcal{A}^{N*}}$ for H_2. The corresponding error probabilities are

$$\alpha_{1|1}^{N*}(t) = \alpha_{2|1}^{N*}(t) = G_1^N(\overline{\mathcal{A}^{N*}})$$

$$\alpha_{2|2}^{N*}(t) = \alpha_{1|2}^{N*}(t) = \max(G_2^N(\mathcal{A}_1^{N*}); G_3^N(\mathcal{A}^{N*})).$$

Let $\mathcal{A}^N \subset \mathcal{X}^N$ be the decision region for H_1 of the another test Φ_N with error probabilities $\alpha_{1|1}^N$ and $\alpha_{2|2}^N$. If $\alpha_{1|1}^N \leq \alpha_{1|1}^{N*}$, then $\alpha_{2|2}^N \geq \alpha_{2|2}^{N*}$.

Proof: The numbers N and t are fixed, we can do not note them during proof. Let $\Psi_{\mathcal{A}^{N*}}$ and $\Psi_{\mathcal{A}^N}$ be indicator functions of the regions. It is not difficult to verify that for all $\mathbf{x} \in \mathcal{X}^N$,

$$(\Psi_{\mathcal{A}^{N*}}(\mathbf{x}) - \Psi_{\mathcal{A}^N}(\mathbf{x}))(G_1^N(\mathbf{x}) - t\max(G_2{}^N(\mathbf{x}); G_3{}^N(\mathbf{x}))) \geq 0.$$

Then

$$\sum_{\mathbf{x} \in \mathcal{X}^N} (\Psi_{\mathcal{A}^{N*}}(\mathbf{x})G_1^N(\mathbf{x}) - t\Psi_{\mathcal{A}^{N*}}(\mathbf{x})\max(G_2{}^N(\mathbf{x}); G_3{}^N(\mathbf{x}))$$

$$-\Psi_{\mathcal{A}^N}(\mathbf{x})G_1^N(\mathbf{x}) + t\Psi_{\mathcal{A}^N}(\mathbf{x})\max(G_2{}^N(\mathbf{x}); G_3{}^N(\mathbf{x})))$$

$$= \sum_{\mathbf{x} \in \mathcal{A}^{N*}} (G_1^N(\mathbf{x}) - t\max(G_2^N(\mathbf{x}); G_3^N(\mathbf{x}))) - \sum_{\mathbf{x} \in \mathcal{A}^N} (G_1^N(\mathbf{x}) - t\max(G_2^N(\mathbf{x}); G_3^N(\mathbf{x})))$$

$$= (1 - \alpha_{1|1}^*) - t\alpha_{2|2}^* - (1 - \alpha_{1|1}) + t\alpha_{2|2}$$

$$= (\alpha_{1|1} - \alpha_{1|1}^*) + t(\alpha_{2|2} - \alpha_{2|2}^*) \geq 0.$$

So from $\alpha_{1|1} \leq \alpha_{1|1}^*$ it follows that $\alpha_{2|2} \geq \alpha_{2|2}^*$.

Now we reformulate Theorem 2 for the model given in (35).

Theorem 5: *If PDs G_1, G_2, G_3 are different, the strictly positive number $E_{1|1}$ is such that*

$$E_{1|1} \leq \min(D(G_2||G_1), D(G_3||G_1)),$$

for

$$\mathcal{R}_1 \triangleq \{Q \in \mathcal{P}(\mathcal{X}), D(Q||G_1) \le E_{1|1}\},$$

$$\mathcal{R}_2 \triangleq \{Q \in \mathcal{P}(\mathcal{X}), D(Q||G_1) > E_{1|1}\},$$

we consider

$$E_{1|1}^* = E_{1|1} = E_{2|1}^*$$

$$E_{2|2}^* = E_{1|2}^* = \min_{Q \in \mathcal{R}_1} \min(D(Q||G_2), D(Q||G_3))$$

then for the hypotheses in (35) there exists an LAO sequence of tests with strictly positive reliabilities given above and with regions of decision for hypotheses H_k

$$\mathcal{A}_k^{N*} = \bigcup_{Q \in \mathcal{R}_k} \mathcal{T}_Q^N(\mathcal{X}), \ k = 1, 2.$$

In the paper [20] of 1990 Haroutunian noted that "the principle of maximum of likelihood is equivalent to the principle of maximum of Kullback-Leibler distance" and "the desired tests sequence is constructed by means of distances between the sample distribution and the hypothetical distributions". It is worth to note that this assertion is something in common with the following note in Cover and Thomas monograph of 1991 (p. 307) [7] concerning the test of two hypotheses (the next one with our adopted exposition). "In the above theorem (the Neyman-Pearson lemma), we have shown that the optimum test is a likelihood ratio test. We can rewrite the log-likelihood ratio as the difference between the relative entropy distances of the sample type to each of the two distributions. Hence the likelihood ratio test (in our notation)

$$\frac{G_1^N(\mathbf{x})}{G_2^N(\mathbf{x})} > t > 1$$

is equivalent to

$$D(Q_{\mathbf{x}}||G_2^N) - D(Q_{\mathbf{x}}||G_1^N) > \frac{\log t}{N}$$

or (with our new notation of divergence of three PDs)

$$D(Q_{\mathbf{x}}||G_1^N||G_2^N) > \frac{\log t}{N}.$$

It remains to add that for the case of simple model in (36) the likelihood ratio test is equivalent to the following condition specifying the region of detection of the first hypothesis in (36)

$$\min[D(Q_{\mathbf{x}}||G_1^N||G_2^N), D(Q_{\mathbf{x}}||G_1^N||G_3^N)] > \frac{\log t}{N}.$$

7. Conclusion

Here we offer some concluding remarks and open problems. In this paper, we have discussed error exponents trade-off of Neyman-Pearson suitable strategy of hypotheses testing for models with M known discrete probability distributions joined in K $(2 \le K \le M)$ clusters, considered as hypotheses.

We presented a single letter characterization of the error exponents of all possible pairs of hypotheses of tests for some cases. After a detailed proof of the point in question for a

general case with the possibility of decision rejection, analogical results are announced the case without a rejection option, the case when all or a part of the given reliabilities are equal to zero, and finally, for a particular case of three disributions and two hypotheses. The reasonings at the end of the previous section confirm the optimality of the tests considered in the paper based on the use of distances between the sample and hypotheses.

For further works it is deserving exploration of characteristics of testing for generalization and enlargement of models studied in this paper. Interesting is the case with multiple objects [cf. 21, 24, 26, 27]. Significant are arbitrarily varying models with a sequence of states known to the decision maker [cf. 12] and also the case when states are not known to the statistician [cf. 2, 25]. Important is the problem of hypothesis identification [cf. 1, 18, 23]. Bayesian framework of the problem, and sources with other than independent issues, for instance with Markov dependence must also be investigated [cf. 16].

Acknowledgement

References

[1] R. F. Ahlswede and E. A. Haroutunian, "On logarithmically asymptotically optimal testing of hypotheses and identification", *Lecture Notes in Computer Science, volume 4123, "General Theory of Information Transfer and Combinatorics", Springer*, pp. 462–478, 2006.

[2] R. F. Ahlswede, E. Aloyan E. A. Haroutunian, "On logarithmically asymptotically optimal hypothesis testing for arbitrarily varying source with side information", *Lecture Notes in Computer Science, volume 4123, "General Theory of Information Transfer and Combinatorics", Springer*, pp. 457-461, 2006.

[3] L. Birgé, "Vitesses maximals de décroissance des erreurs et tests optimaux associés". Z. Wahrsch. Verw. Gebiete, vol. 55, pp. 261–273, 1981.

[4] R. E. Blahut, "Hypotheses testing and information theory" *IEEE Trans. Inform Theory*, vol. 20, no 4, pp. 405-417, 1974.

[5] R. E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, 1987.

[6] A. A. Borovkov *Mathematical Statistics (in Russian)*. Nauka, Novosibirsk, 1997.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Second Edition, New York, Wiley, 2006.

[8] D. R. Cox, "Tests of separate families of hypotheses" In Proceeding 4th Berkley Simp. Math. Statist. Prob., University of Callifornia Press, Berkelly, pp. 105-123, 1961.

[9] D. R. Cox, "Furthe results on tests of separate families of hypotheses", Journal of the Royal Statist. Society: Serie B, vol. 24, no. 2, pp. 406-424, 1962.

[10] D. R. Cox, "A return to an old paper: Tests of separate families of hypotheses", Journal of the Royal Statist. Society: Serie B, vol. 75, no. 2, pp. 207-2015, 2013.

[11] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic press., New York, 1981.

[12] I. Csiszár and G. Longo, "On the error exponent for source coding and for testing simple statistical hypotheses", *Studia Sc. Math. Hungarica*, vol. 6, pp. 181–191, 1971.

[13] I. Csiszár and P. Shields, " Information theory and statistics: A tutorial", *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, 2004.

[14] M. Feder and N. Merhav, " Universal composite hypotheses testing: A competetive minimax approach," *IEEE Trans. Inform. Theory,* vol. 48, pp. 1504-1517, June 2002.

[15] F. W. Fu and S. Y. Shan, "Hypotheses testing for arbitrarily varying source with exponential type constraint", *IEEE Trans. on Inform. Theory*, vol. 14, no. 2, pp. 892-895, 1998.

[16] M. Gutman, "Asymptotically optimal classification for multiple tests with empirically observed statistics", *IEEE Trans. Inform. Theory*, vol. 35, no. 2, pp. 401–408, 1989.

[17] T. S. Han and K. Kobayashi , "Exponential-type error probabilities for multiterminal hypotheses testing", *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 2-13, 1989.

[18] E. A. Haroutunian, "Many statistical hypotheses: interdependence of optimal test's error probabilities exponents", (In Russian), Abstract of the report on the 3rd All-Union school-seminar, *"Program-algorithmical software for applied multi-variate statistical analysis"*, Tsakhkadzor, Part 2, pp. 177–178, 1988.

[19] E. A. Haroutunian, "On asymptotically optimal testing of many statistical hypotheses conserning Markov chain", (in Russian), *Isvestia Akademii Nauk Armenii, Mathematika*, vol. 23, no. 1, pp. 76-80, 1988.

[20] E. A. Haroutunian, "Logarithmically asymptotically optimal testing of multiple statistical hypotheses", *Problems of Control and Information Theory*, vol. 19(5-6), pp. 413–421, 1990.

[21] E. A. Haroutunian, "Reliability in multiple hypotheses testing and identification problems" *Nato Science Series III, Computer and System Sciences*, vol.198, IOS Press, pp. 189-201, 2003.

[22] E. A. Haroutunian, "On three hypotheses robust detection design under mismatch", *Proceedings of International Conference CSIT 2019*, pp. 125 – 128, Yerevan 2019.

[23] E. A. Haroutunian and P. M. Hakobyan, "On LAO testing of multiple hypotheses for pair of objects", *Mathematical Problems of Computer Sciences*, vol. 25, pp. 93-101, 2006.

[24] E. A. Haroutunian and P. M. Hakobyan, "Multiple hypotheses LAO testing for many independent objects", *International Journal "Scholarly Research Exchange"*, vol. 2009, pp. 1-6, 2009.

[25] E. A. Haroutunian and P. M. Hakobyan, "On Neyman-Pearson principle in multiple hypotheses testing", *Mathematical Problems of Computer Sciences*, vol. 40, pp. 34-37, 2013.

[26] E. A. Haroutunian and P. M. Hakobyan, "Multiple objects: Error exponents in hypotheses testing and identification", *Lecture Notes in Computer Science, volume 7777, "Ahlswede Festschrift", Springer* , pp. 313-345, 2013.

[27] E. A. Haroutunian, P. M. Hakobyan and F. Hormosi-nejad, "On two-stage LAO testing of multiple hypotheses for the pair of families of distributions", *Journal of Statistics and Econometrics Methods* , vol. 2, no. 2, pp. 127-156, 2013.

[28] E. A. Haroutunian, P. M. Hakobyan and A. O. Yessayan, "On multiple hypotheses LAO testing with rejection of decision for many independent objects", *Proceedings of International Conference CSIT*, pp. 117 – 120, 2011.

[29] E. A. Haroutunian, P. M. Hakobyan and A. O. Yessayan, "Many hypotheses LAO testing with rejection of decision for arbitrarily varying object", *Mathematical Problems of Computer Sciences*, vol. 35, pp. 77-85, 2011.

[30] E. A. Haroutunian, P. M. Hakobyan and A. O. Yessayan, "On multiple hypotheses LAO testing with liberty of rejection of decision for two independent objects", International Journal, *Information theories and applications*, vol. 25 , no. 1, pp. 38-46, 2018.

[31] E. A. Haroutunian, M. E. Haroutunian and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing", *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 2-3, 2008.

[32] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *The Annals of Mathematical Statistics*, vol. 36, pp. 369–401, 1965.

[33] S. Kullback, *Information Theory and Statistics*, New York, Wiley, 1959.

[34] E. Levitan and N. Merhav, "A competitive Neyman-Pearson approach to universal hypothesis testing with applications", *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2215-2229, 2002.

[35] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*, Springer, 2008.

[36] G. Longo and A. Sgarro, " The error exponent for the testing of simple statistical hypotheses: A combinatorial approach", *Journal of Combinatorics and Informational System Science*, vol. 5, no. 1, pp. 58-67, 1980.

[37] V. Poor, *An Introduction to Signal Deyection and Estimation*, Springer, 1994.

[38] E. Tuncel, "On error exponents in hypothesis testing', *IEEE Trans. Inform. Theory*, vol. 51, mo.8, pp. 2945-2950, 2005.

[39] G. Tusnády, "On asymptotically optimal tests," *Annals of Statatistics*, vol. 5, no. 2, pp. 385-393, 1977.

[40] H. van Trees, *Detection, Estimation and Modulation Theory*, pt.1 New York, Wiley, 1968.

Բաշխումների խմբերով կազմված բազմաթիվ վարկածների Նեյմանի-Պիրսոնի համընդհանուր տեստավորման հատուկ ուղի

Եվգենի Ա. Հարությունյան և Արամ Օ. Եսայան

ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտ
e-mail: eghishe@sci.am, armfrance@yahoo.fr

Ամփոփում

Հետազոտվել են Նեյմանի-Պիրսոնի ասիմպտոտորեն օպտիմալ հայտնաբերման ընթացակարգերը այն մոդելների համար, որոնք բնութագրվում են M դիսկրետ հավանականությունների բաշխումներով, որոնք խմբավորված են ըստ K դասերի, $2 \leq K \leq M$, որոնք դիտարկվում են որպես վարկածներ: Դիտարկումների քանակի աճի վրա հիմնված տեստերի հաջորդականությունը լոգարիթմորեն ասիմպտոտիկորեն օպտիմալ (LAO) է, երբ տրված ցուցիչների (հուսալիությունների) որոշակի մասը ապահովում ենbնացած բոլոր հուսալիությունների համար դրական արժեքներ ունենալը: Կառուցվել են օբյեկտների որոշ մոդելների համար LAO տեստերի այն դեպքերը, երբ թույլատրվում է հրաժարվել որոշում ընդունելուց և երբ սխալների հավանականությունների մի մասը կամ բոլորը փորձերի թվի աճի հետ մեկտեղ նվազում են ենթացուցչային օրենքով: Ստացվել են այդպիսի տեստերի բոլոր հուսալիությունների միատառ բնութագրումները: Դիտարկված է մի հասարակ դեպք, երբ բաշխումների թիվը երեք է, իսկ վարկածների թիվը՝ երկու:

Բանալի բառեր՝ վիճակագրական վարկածների ստուգում, վարկածների ընտանիքներ, օպտիմալ հայտնաբերում, Նեյման-Պիրսոնի մոտեցում, Նեյման-Պիրսոնի լեմմա, Կուլբակի-Լյոյբլերի հեռավորության մաքսիմումի սկզբունք, սխալի ցուցիչ:

Специальный путь Неймана-Пирсона к универсальной проверке многих гипотез, сформированных группами распределений

Евгений А. Арутунян и Арам О. Есаян

Институт проблем информатики и автоматизации НАН РА
e-mail: eghishe@sci.am, armfrance@yahoo.fr

Аннотация

Исследуются асимптотически оптимальные процедуры Неймана-Пирсона обнаружения для моделей, характеризующихся M дискретными распределениями вероятностей, сгруппированными в K групп, $2 \leq K \leq M$, рассматриваемых как гипотезы. Последовательность тестов, при возрастании числа наблюдений, является логарифмически асимптотически оптимальной (LAO), когда определенная

часть заданных экспонент (надежностей) обеспечивает положительные значения для всех других надежностей. Сконструированы LAO последовательности тестов для некоторых моделей, в том числе в случаях, когда разрешен отказ от принятия решения и когда часть или все заданные вероятности ошибок убывают субэкспоненциально с ростом количества экспериментов. Получены однобуквенные характеристики для всех надежностей таких тестов. Рассмотрен простой случай с тремя распределениями и двумя гипотезами.

Ключевые слова: проверка статистических гипотез, семейства гипотез, оптимальное обнаружение, подход Неймана-Пирсона, лемма Неймана-Пирсона, принцип максимума расстояния Кульбака-Лейблера, показатель ошибки.