Mathematical Problems of Computer Science 35, 33-36, 2011.

Error Probability Exponents and Achievable Region in Testing of Many Hypotheses for Two Independent Objects

Aram O. Yessayan, Evgueni A. Haroutunian and Parandzem M. Hakobyan

Institute for Informatics and Automation Problems of NAS of RA
e-mail: evhar@ipia.sci.am

Abstract

The model of many hypotheses testing for one objects was examined by E. Tuncel. In the present work it is supposed that L hypothetical probability distributions are known and two objects independently each from other follow to one of them. N-vectors of values of discrete independent random variables represent results of N observations for each object. Decisions concerning realized probability distributions of the objects must be made on the base of such samples. It is proved that defined region for vector of error probability exponents "reliabilities for two objects completly characterizes set of all achivable vectors.

1. Introduction

The applications of information theory in statistics were reflected in [1]-[5]. In monograph [3] the sequential methods for the multiple hypotheses identification problem was examined. The model error probabilities exponents of many hypotheses testing for one object was studied in [7]. Logarithmically asymptotically optimal (LAO) tests for multiple hypotheses were investigated in [6]. In papers [8], [9] and [10] the problems with many objects and multiple hypotheses were proposed and solved. The model with two objects which can not have the same probability distribution from three given was also examined in [8] and [11]. In present paper we develop the approach by Tuncel of error probabilities exponents for the model of two objects.

2. Problem Statement and Formulation of Result

Let X_1 and X_2 be random variables (RV) taking values in the finite set $\mathcal{X}$. Let $\mathcal{P}(\mathcal{X})$ be the space of all possible probability distributions (PD) on $\mathcal{X}$. There are given L PD $G_l = \{G_l(x), x \in \mathcal{X}\}, l = \overline{1, L}$ from $\mathcal{P}(\mathcal{X})$. In this paper we study the model consisting of 2 objects which independently follow to one of these L PD. Let $(\mathbf{x_1}, \mathbf{x_2}) = ((x_1^1, x_1^2), (x_2^1, x_2^2), ..., (x_N^1, x_N^2))$ be a sequence of results of N independent observations of the vector (X_1, X_2). The goal of the ststistician is to difine which pair of distributions corresponds to observed sample $(\mathbf{x_1}, \mathbf{x_2})$. The test is a procedure of making decision on the base of $(\mathbf{x}_1, \mathbf{x}_2)$. which we denote by φ_N. For each object the non-randomized test $\varphi_N(\mathbf{x}_i)$, $i = 1, 2$, can be defined by division of the sample space $\mathcal{X}^N$ on L disjoint subsets $\mathcal{A}_{l_i}^N = \{\mathbf{x}_i : \varphi_N(\mathbf{x}_i) = l_i\}$. $i = 1, 2$. $l_i = \overline{1, L}$. The set $\mathcal{A}_{l_i}^N$ consists of all vectors $\mathbf{x}_i$ for which the hypothesis G_{l_i} is adopted. We study the probabilities of the erroneous acceptance of hypothesis G_{l_i} provided that G_{m_i} is true for all pairs $l_i, m_i = \overline{1, L}$, $m_i \neq l_i$.

$$\alpha_{l_i|m_i}(\varphi_N^i) = G_{m_i}(\mathcal{A}_{l_i}^N),\ i = 1, 2. \tag{1}$$

The probability to reject G_{m_i}, when it is true, we define as follows

$$\alpha_{m_i|m_i}(\varphi_N^i) = \sum_{l_i \neq m_i} \alpha_{l_i|m_i}(\varphi_N^i) = G_{m_i}(\overline{\mathcal{A}_{m_i}^N}). \tag{2}$$

For each $i = 1, 2$ we denote by $\mathbf{E}_i = \{E_{l_i|m_i}\}$, $l_i \neq m_i$ the vector, elements of which correspond to the set of error exponents $-\frac{1}{N}\log G_{m_i}(\mathcal{A}_{l_i}^N)$.

Definition 1 [6] : *The set of error exponents indicated by vector* $\mathbf{E}_i$, $i = 1, 2$ *is called achievable if for all* $\varepsilon > 0$ *and large enough* N *there exists a decision scheme* $\{A_{l_i}^N\}$ *satisfying for* $m_i, l_i = \overline{1, L}$, $m_i \neq l_i$ *the following conditions:*

$$-\frac{1}{N}\log \alpha_{l_i|m_i}(\varphi_N^i) > E_{l_i|m_i}(\varphi^i) - \varepsilon.$$

The set of all achievable vectors is denoted by $\mathcal{R}_i$, $i = 1, 2$.

For each object let us define a regions $\mathcal{E}_i$:

$$\mathcal{E}_i = \{\mathbf{E}_i : \forall Q_i,\ \exists l_i,\ D(Q_i||G_{m_i}) > E_{l_i|m_i}(\varphi^i),\ \text{for all}\ m_i \neq l_i\},$$
$$i = 1, 2,\ m_i, l_i = \overline{1, L},\ \text{where}\ Q_i = \{Q_i(x^i),\ x^i \in \mathcal{X}\}$$

Theorem 1 [6]: *The following inclusion take place* $\mathcal{E}_i \subset \mathcal{R}_i$. *Conversely, if* $\mathbf{E}_i \in \mathcal{R}_i$, *then for any* $\delta > 0$, $\mathbf{E}_{i\delta} \in \mathcal{E}_i$, *where* $\mathbf{E}_{i\delta} \triangleq \{E_{l_i|m_i}(\varphi^i) - \delta\}$.

Let $\alpha_{l_1,l_2|m_1,m_2}(\varphi_N)$ be the probability of the erroneous acceptance of a pair hypotheses (G_{l_1}, G_{l_2}) when the pair (G_{m_1}, G_{m_2}) is true. where $(l_1, l_2) \neq (m_1, m_2)$, $m_i, l_i = \overline{1, L}$.
We denote by $\mathbf{E} = \{E_{l_1,l_2|m_1,m_2}\}$ the vector, elements of which correspond to the set of error exponents

Definition 2: *The set of error exponents indicated by the vector* $\mathbf{E}$ *is called achievable if for every* $\varepsilon > 0$ *there exists a decision scheme* $\{\mathcal{A}_{l_1,l_2}^N\}$, $l_1, l_2 = \overline{1, L}$ *satisfying*

$$-\frac{1}{N}\log \alpha_{l_1,l_2|m_1,m_2}(\varphi_N) > E_{l_1,l_2|m_1,m_2}(\varphi) - \varepsilon.$$

for $m_i, l_i = \overline{1, L}$, $i = 1, 2$ *and large enough* N. *The set of all achievable vectors is denoted by* $\mathcal{R}$.

For two objects let us define a region $\mathcal{E}$:

$$\mathcal{E} = \{\mathbf{E} : \forall Q_{1,2},\ \exists (l_1, l_2),\ D(Q_{1,2}||G_{m_1,m_2}) > E_{l_1,l_2|m_1,m_2}(\varphi),\ \forall (m_1, m_2) \neq (l_1, l_2),$$
$$m_i, l_i = \overline{1, L},\ i = 1, 2\}$$

where $Q_{1,2} = \{Q_1(x^1)Q_2(x^2),\ x^1, x^2 \in \mathcal{X}\}$ and $G_{m_1,m_2} = \{G_{m_1}(x^1)G_{m_2}(x^2),\ x^1, x^2 \in \mathcal{X}\}$

Theorem 2: $\mathcal{E} \subset \mathcal{R}$. *Conversely, if* $\mathbf{E} \in \mathcal{R}$, *then for any* $\delta > 0$, $\mathbf{E}_\delta \in \mathcal{E}$, *where*

$$\mathbf{E}_\delta \triangleq \{E_{l_1,l_2|m_1,m_2}(\varphi) - \delta\}.$$

3. Proof of Theorem 2.

From the independence of the objects it follows that.
if $l_i \neq m_i$, $i = 1,2$, then

$$\alpha_{l_1,l_2|m_1,m_2}(\varphi_N) = \alpha_{l_1|m_1}(\varphi_N^1)\alpha_{l_2|m_2}(\varphi_N^2). \tag{3}$$

if $m_i = l_i$, $l_{3-i} \neq m_{3-i}$, $i = 1,2$, then

$$\alpha_{l_1,l_2|m_1,m_2}(\varphi_N) = \alpha_{l_{3-i}|m_{3-i}}(\varphi_N^{3-i})(1 - \alpha_{m_i|m_i}(\varphi_N^i)). \tag{4}$$

Using (3) and (4) we can derive that, correspondingly,

$$E_{l_1,l_2|m_1,m_2}(\varphi) = E_{l_1|m_1}(\varphi^1) + E_{l_2|m_2}(\varphi^2), \tag{5}$$

$$E_{l_1,l_2|m_1,m_2}(\varphi) = E_{l_{3-i}|m_{3-i}}(\varphi^{3-i}). \tag{6}$$

Thus for every $\mathbf{E}_1 \in \mathcal{E}_1 \subset \mathcal{R}_1$ and every $\mathbf{E}_2 \in \mathcal{E}_2 \subset \mathcal{R}_2$ the components of $\mathbf{E}$ will be presented as follows: for each $m_2 = l_2$, $m_1 \neq l_1$, $m_i, l_i = \overline{1,L}$, $i = 1,2$ we obtain $\mathbf{E}_1$. By analogy for each $m_1 = l_1$, $m_2 \neq l_2$, $m_i, l_i = \overline{1,L}$, $i = 1,2$, we obtain $\mathbf{E}_2$. In other cases, if $m_i \neq l_i$, $m_i, l_i = \overline{1,L}$, $i = 1,2$, we obtain

$$E_{l_1,l_2|m_1,m_2}(\varphi) = E_{l_1|m_1}(\varphi^1) + E_{l_2|m_2}(\varphi^2),$$

where $E_{l_1|m_1}(\varphi^1) \in \mathbf{E}_1$ and $E_{l_2|m_2}(\varphi^2) \in \mathbf{E}_2$. Thus we obtain remaining elements of $\mathbf{E}$. So $\mathcal{E} \subset \mathcal{R}$. That is we get that the defined region for two objects is also achievable.

For the converse part, let $\mathbf{E} \in \mathcal{R}$. Then for every $\epsilon > 0$, we have that there exists some sequence of test φ_N and some number $N_0(\epsilon)$ such that

$$-\frac{1}{N}\log \alpha_{l_1,l_2|m_1,m_2}(\varphi_N) > E_{l_1,l_2|m_1,m_2}(\varphi) - \varepsilon. \tag{7}$$

$\forall (m_1, m_2) \neq (l_1, l_2)$ and $N > N_0(\epsilon)$. Now let us take any $\delta > 0$. If $\mathbf{E}_\delta \notin \mathcal{E}$, then there exists a distribution $Q_{1,2}$ such that

$$\forall (l_1, l_2)\, \exists (m_1, m_2) \neq (l_1, l_2) \text{ satisfying } D(Q_{1,2}||G_{m_1,m_2}) \leq E_{l_1,l_2|m_1,m_2}(\varphi) - \delta.$$

When $m_i = l_i$ and $m_{3-i} \neq l_{3-i}$, $i = 1,2$ we obtain

$$D(Q_1||G_{m_1}) + D(Q_2||G_{m_2}) \leq E_{l_{3-i}|m_{3-i}}(\varphi^i) - \delta$$

But it is impossible because for one object it is proven [6] that when in equality $D(Q_i||G_{m_i}) \leq E_{l_i|m_i}(\varphi^i) - \delta$ is assumed it gives contardiction. And when $m_i \neq l_i$, $i = 1,2$ we obtain

$$D(Q_1||G_{m_1}) + D(Q_2||G_{m_2}) \leq E_{l_1|m_1}(\varphi^1) + E_{l_2|m_2}(\varphi^2) - \delta.$$

But from here we can say that for $i = 1,2$ at least one of $D(Q_i||G_{m_i}) \leq E_{l_i|m_i}(\varphi^i) - \delta/2$ and according to [6] it gives also contradiction. So $\mathbf{E}_\delta \in \mathcal{E}$ and we get $\mathcal{E} = \mathcal{R}$.

References

[1] I. Csiszár and P. C. Shields. *Information Theory and Statistics: a tutorial. Foundations and Trends in Communications and Information Theory*, volume 1, no. 4. 2004.

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

[3] R. E. Bechhofer, J. Kiefer and M. Sobel, *Sequential Identification and Ranking Procedures*. The University of Chicago, Press, 1968.

[4] R. E. Blahut, "Hypotheses testing and informtion theory," *IEEE Transaction on Information Theory*, vol 20, pp. 405-417, 1974.

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

[6] E.Tuncel. "On error exponents in hypothesis testing", *IEEE Trans. Inf. Theory*. vol. 51, no. 8, pp. 2945-2950, 2005.

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

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

[9] E. Haroutunian, "Reliability in multiple hypotheses testing and identification". *Proceedings of NATO ASI, Yerevan, 2003, NATO Science Series III: Computer and System Sciences*, vol. 198, pp. 189-201. IOS Press, 2005.

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

[11] E. Haroutunian and A. Yessayan, "On optimal testing of three hypotheses for two dependent objects ", *Mathematical Problems of Computer Sciences*, vol. XXVI. pp. 89-94, 2006.

Երկու անկախ օբյեկտների վերաբերյալ բազմաթիվ վարկածների տեստավորման սխալների հավանականությունների ցուցիչները

Ա. Եսայան, Ե. Հարությունյան և Փ. Հակոբյան

Ամփոփում

Հոդվածում ենթադրվում է, որ L հավանականային բաշխումները հայտնի են, իսկ օբյեկտներից յուրաքանչյուրը անկախ մեկը մյուսից կարող են բաշխված լինել տրվածներից յուրաքանչյուրով: Օբյեկտների բաշխվածության վերաբերյալ որոշումները ընդունվում են երկու օբյեկտների N-ական անկախ դիտարկումների արդյունքների հիման վրա: Հոդվածում ապացուցվել է, որ անկախ օբյեկտների հուսալիությունների (սխալների հավանականությունների ցուցիչների) վեկտորն ամբողջությամբ բնութագրում է հասանելի կոչվող վեկտորների բազմությունը: