Mathematical Problems of Computer Science 49, 110–114, 2018.

On Multiple Hypotheses LAO Testing With Rejection of Decision for Two Dependent Objects

Evgueni Haroutunian, Aram Yesayan and Narine Harutyunyan

Institute for Informatics and Automation Problems of NAS RA
e-mail: eghishe@sci.am, armfrance@yahoo.fr, narineharutyunyan57@gmail.com

Abstract

Multiple statistical hypotheses testing with possibility of rejecting of decision is considered for model consisting of two dependent objects characterized by joint discrete probability distribution. The matrix of error probabilities exponents (reliabilities) of asymptotically optimal tests is studied.

Keywords: Multiple hypotheses testing, Optimal tests, Rejection option, Two object.

1. Introduction

This paper is devoted to the study of characteristics of logarithmically asymptotically optimal (LAO) hypotheses testing with possibility of rejection of decision for the model of two dependent objects with joint ptobability distributions (PDs). The correspondence presents a complement to problems studied in [1, 2], where the components of random vector characterizing two objects were independent, so it was possible to consider the test procedure with separate tests for two objects. In [2] two different models are studied: first, when rejection was allowed only for one of the objects and second, when rejection was allowed for both objects. The problem with analogous statement for one arbitrarily varying object with side information was examined in [3].

It is worth to recall the previous results on LAO testing of many hypotheses published in [4-6].

Methods and basic results of LAO testing are also presented in books [7-9]

2. Problem Formulation and Result

Let $\mathcal{P}(\mathcal{X})$ be a space of all probability distributions $G(x)$ on finite set $\mathcal{X}$. Let (X_1, X_2) be random vector taking values in the set $\mathcal{X} \times \mathcal{X}$ with one of M^2, $M \geq 2$ joint PDs $G_{m_1,m_2} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, $m_1, m_2 = \overline{1, M}$. Let $(\mathbf{x}^1, \mathbf{x}^2) \stackrel{\triangle}{=} ((x_1^1, x_1^2), ..., (x_n^1, x_n^2), ..., (x_N^1, x_N^2))$, $x_n^1, x_n^2 \in \mathcal{X}$, $n = \overline{1, N}$, be a vector of results of N independent observations of the vector (X_1, X_2), it is called a sample.

The statistician has to determine unknown PDs from the set of hypotheses: H_{m_1,m_2} : $G = G_{m_1,m_2}$, $m_1, m_2 = \overline{1, M}$ or withdraw to do any judgement using obtained sample.

We call this procedure a compound test and denote it by Φ_N.

The test Φ_N can be defined by the division of the space $\mathcal{X}^N \times \mathcal{X}^N$ into M^2+1 disjoint subsets, where $\mathcal{A}_{m_1,m_2}$, $m_1,m_2=\overline{1,M}$, contains all vectors $(\mathbf{x_1},\mathbf{x_2})$ for which the hypothesis H_{m_1,m_2} is adopted, and $\mathcal{A}_{M+1}$ contains all vectors for which we refuse to take a certain answer.

Let $\alpha_{l_1,l_2|m_1,m_2}(\Phi_N)$ be the probability of the erroneous acceptance of the hypothesis H_{l_1,l_2} by the test Φ_N provided that the hypothesis H_{m_1,m_2} is true, where $(m_1,m_2)\neq(l_1,l_2)$, $m_1,m_2,l_1,l_2=\overline{1,M}$,

$$\alpha_{l_1,l_2|m_1,m_2}(\Phi_N)=G^N_{m_1,m_2}(\mathcal{A}_{l_1,l_2}).$$

When the hypothesis H_{m_1,m_2} is true, but we decline the decision concerning to the hypotheses, the corresponding probability of error is:

$$\alpha_{M+1,M+1|m_1,m_2}(\Phi_N)=G^N_{m_1,m_2}(\mathcal{A}_{M+1}).$$

The probability not to accept a true hypotheses H_{m_1,m_2}, $m_1,m_2=\overline{1,M}$ is the following:

$$\alpha_{m_1,m_2|m_1,m_2}(\Phi_N)=\sum_{(l_1,l_2)\neq(m_1,m_2),\, l_1,l_2=\overline{1,M},\, (l_1,l_2)=(M+1,M+1)}\alpha_{l_1,l_2|m_1,m_2}(\Phi_N). \tag{1}$$

We study the corresponding reliabilities $E_{l_1,l_2|m_1,m_2}(\Phi)$ of the sequence of tests Φ,

$$E_{l_1,l_2|m_1,m_2}(\Phi)\triangleq\underline{\lim}_{N\to\infty}-\frac{1}{N}\log\alpha_{l_1,l_2|m_1,m_2}(\Phi_N),$$

$$m_1,m_2,l_1,l_2=\overline{1,M},(l_1,l_2)=(M+1,M+1). \tag{2}$$

Definitions (1) and (2) imply that

$$E_{m_1,m_2|m_1,m_2}(\Phi)=\min_{(l_1,l_2)\neq(m_1,m_2)}E_{l_1,l_2|m_1,m_2}(\Phi),$$

$$m_1,m_2,l_1,l_2=\overline{1,M},\ \ (l_1,l_2)=(M+1,M+1). \tag{3}$$

We call the test sequence Φ^* LAO for the model with two objects if for the given positive values of certain part of elements of the reliability matrix $\mathbf{E}(\Phi^*)$ the procedure Φ^* provides maximal values for all other elements of it.

For $M=2$ the matrix will be as follows:

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

With the given elements $E_{1,1|1,1}$, $E_{1,2|1,2}$, $E_{2,1|2,1}$, $E_{2,2|2,2}$ we define the regions of acceptance of the test.

In the general case of M hypotheses for given reliabilities $E_{1,1|1,1}$, $E_{1,2|1,2}$, $E_{2,1|2,1}$, ..., $E_{M,M|M,M}$ we define the following regions:

$$\mathcal{R}_{m_1,m_2}\triangleq\{Q:\ \ D(Q||G_{m_1,m_2})\leq E_{m_1,m_2|m_1,m_2}\},\ \ m_1,m_2=\overline{1,M}, \tag{4}$$

$$\mathcal{R}_{M+1,M+1} \triangleq \{Q: \quad D(Q||G_{m_1,m_2}) > E_{m_1,m_2|m_1,m_2}, \quad m_1,m_2 = \overline{1,M}\}, \tag{5}$$

$$E^*_{m_1,m_2|m_1,m_2} = E^*_{m_1,m_2|m_1,m_2}(E_{m_1,m_2|m_1,m_2}) \triangleq$$

$$\triangleq E_{m_1,m_2|m_1,m_2}, \quad m_1,m_2 = \overline{1,M}, \tag{6}$$

$$E^*_{l_1,l_2|m_1,m_2} = E^*_{l_1,l_2|m_1,m_2}(E_{l_1,l_2|l_1,l_2})$$

$$\triangleq \inf_{Q\in\mathcal{R}_{l_1,l_2}} D(Q||G_{m_1,m_2}), \;\; l_1,l_2,m_1,m_2 = \overline{1,M}, \;\; (m_1,m_2) \neq (l_1,l_2) \tag{7}$$

$$E^*_{M+1,M+1|m_1,m_2} = E^*_{M+1,M+1|m_1,m_2}(E_{1,1|1,1}, E_{1,2|1,2}, ..., E_{M,M|M,M})$$

$$\triangleq \inf_{Q\in\mathcal{R}_{M+1,M+1}} D(Q||G_{m_1,m_2}), \quad m_1,m_2 = \overline{1,M}. \tag{8}$$

Let us denote by $(m_1,m_2)^-$ the set of all pair indices in row of (m_1,m_2) varying from $(1,1)$ till previous of (m_1,m_2) and by $(m_1,m_2)^+$ the set of all pair indices in row of (m_1,m_2) varying from next of (m_1,m_2) till (M,M).

Theorem: *If all distributions* $G_{m_1,m_2} = \{G_{m_1,m_2}(x_1,x_2),\; x_1,x_2 \in \mathcal{X}\}$, $m_1,m_2 = \overline{1,M}$, *are different in the sense that* $D(G_{l_1,l_2}||G_{m_1,m_2}) > 0$, *and the positive numbers* $E_{1,1|1,1}, E_{2,2|2,2}, ..., E_{M,M|M,M}$ *are such that the following inequalities hold*

$$E_{1,1|1,1} < \min_{l_1,l_2=\overline{1,M},\,(l_1,l_2)\neq(1,1)} D(G_{l_1,l_2}||G_{1,1}), \tag{9}$$

$$E_{m_1,m_2|m_1,m_2} < \min[\min_{(l_1,l_2)\in(m_1,m_2)^-} E^*_{l_1,l_2|m_1,m_2}(E_{l_1,l_2|l_1,l_2}),$$

$$\min_{(l_1,l_2)\in(m_1,m_2)^+} D(G_{l_1,l_2}||G_{m_1,m_2})],$$

$$m_1,m_2 = \overline{1,M},\; (m_1,m_2)\neq(1,1),\; (m_1,m_2)\neq(M,M), \tag{10}$$

$$E_{M,M|M,M} < \min_{l_1,l_2=\overline{1,M},\,(l_1,l_2)\neq(M,M)} E^*_{l_1,l_2|M,M}(E_{l_1,l_2|l_1,l_2}), \tag{11}$$

then there exists a LAO sequence of tests, all elements of the reliability matrix of which $\mathbf{E}^* = \{E^*_{l_1,l_2|m_1,m_2}\}$ *are positive and are defined in* $(6)-(8)$ *.*

When one of the inequalities $(9)-(11)$ *is violated, then at least one element of the matrix* $(6)-(8)$ *is equal to* 0.

The proof of the theorem consists in presentation of the problem for two objects as a problem for one capacious object. If we renumerate as follows $(1,1)=1,(1,2)=2,...,(1,M)=M,(2,1)=M+1,...,(2,M)=2M,...,(M,M)=M^2$ and denote $(X_1,X_2)=Y$, $\mathcal{X}\times\mathcal{X}=\mathcal{Y}$ we will have problem of M^2 hypotheses testing for one object with possibility of decision rejection. So using this numeration we will have the corresponding error probabilities and reliabilities for $l=\overline{1,M^2+1}$, $,m=\overline{1,M^2}$, when we apply Theorem 2 of [3].

Generalization of the result is possible in many directions.

References

[1] 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 2011*, pp. 117 – 120, Yerevan 2011.

[2] E. Haroutunian, P. Hakobyan, A. Yessayan and N. Harutyunyan "On Multiple Hypotheses LAO Testing With Liberty of Rejection of Decision for Two Independent Objects",...

[3] E. Haroutunian, P. Hakobyan and A. Yessayan, "Many hypotheses LAO testing with rejection of decision for arbitrarily varying object", *Transactions of IIAP of NAS of RA and of YSU, Mathematical Problems of Computer Science*, vol. 35, pp. 77-85, 2011.

[4] 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.

[5] 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.

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

[7] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*, Second Edition, Cambridge University Press, 2011.

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

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

Submitted 10.10.2017, accepted 15.02.2018.

Բազմակի վարկածների որոշումից հրաժարմամբ ԼԱՕ ստուգման մասին երկու կախյալ օբյեկտների դեպքում

Ե. Հարությունյան, Ա. Եսայան և Ն. Հարությունյան

Ամփոփում

Համատեղ ընդհատ հավանականային բաշխմամբ բնութագրվող երկու կախյալ օբյեկտների նկատմամբ դիտարկվում է բազմակի վարկածների որոշումից հրաժարման հնարավորությամբ ստուգումը: Ուսումնասիրվել է սխալի հավանականությունների ասիմպտոտորեն օպտիմալ ցուցիչների (հուսալիությունների) մատրիցը:

О LAO тестировании многих гипотез с отказом от решения для двух зависимых объектов

Е. Арутюнян, А. Есаян и Н. Арутюнян

Аннотация

Для модели состоящей из двух зависимых объектов, характеризуемых совместным дискретным распределением вероятностей рассматривается тестирование многих статистических гипотез с возможностью отказа от решения. Изучена матрица асимптотически оптимальных экспонент вероятностей ошибок (надежностей).