Mathematical Problems of Computer Science 37, 25–34, 2012.

On Two-stage Logarithmically Asymptotically Optimal Testing of Multiple Hypotheses Concerning Distributions from the Pair of Families

Evgueni Haroutunian[1], Parandzem Hakobyan[1] and Farshin Hormozi nejad[2]

Institute for Informatics and Automation Problems of NAS of RA.[1]
Islamic Azad University, Ahvaz Branch, Iran.[2]
evhar@ipia.sci.am, par_h@ipia.sci.am, hormozi-nejad@iauahvaz.ac.ir

Abstract

Two-stage testing of multiple hypothesesfor a model with two given families of hypothetical probability distributions is considered. The matrix of reliabilities of logarithmically asymptotically optimal hypotheses testing by a pair of stages is studied and compared with the case of similar one-stage testing.

Keywords Logarithmically asymptotically optimal (LAO) test, multiple hypotheses testing, multistage tests, reliabilities matrix, error probability exponent.

Introduction

this paper the problems of hypotheses logarithmically asymptotically optimal (LAO) testg for a model consisting of two families of hypothetical distributions are studied. Hoeffding paper [9] and later Csiszár and Longo [4], Tusnady [10] and others studied asymptotically timal tests for two hypotheses. In paper [5] the problem of LAO testing of multiple statical hypotheses is solved. In paper [1] Ahlswede and Haroutunian, in [6] and in [7] some oblems on multiple hypotheses testing and identification for many objects are formulated. [8] multiple hypotheses LAO testing for many independent objects is investigated.

We examine multiple hypothesis LAO two-stage testing for an object characterized by a ir of disjoint families of PDs. Two-stage additive tests become popular in applications, pecially in the field of clinical trials, to achieve minimal economic expenditure.

Random variable (RV) X characterizing the studied object takes values in the finite set and $\mathcal{P}(\mathcal{X})$ is the space of all distributions on $\mathcal{X}$. Suppose S hypothetical probability stributions (PDs) of X are given, but they are divided into two disjoint families. The first mily includes R hypotheses $P_1, P_2, ..., P_R$ and the second one includes $S-R$ hypotheses $_{R+1}, P_{R+2}, ..., P_S$. The considered object is characterized by RV X following to one of ese S hypotheses. The statistician is trying to make a reliable decision about the correct stribution using the sample $\mathbf{x} = (x_1, \ldots, x_N)$ of results of N independent observations of e RV X.

Let $N(x|\mathbf{x})$ be the number of repetitions of the element $x \in \mathcal{X}$ in the vector $\mathbf{x} \in \mathcal{X}^N$, then

$$Q_{\mathbf{x}} \triangleq \left\{Q_{\mathbf{x}}(x) = \frac{N(x|\mathbf{x})}{N},\ x \in \mathcal{X}\right\},$$

is the PD, called in statistics *the empirical probability distribution* of the sample x, and in information theory – *the type* of x [2, 3].

Let $\mathcal{P}^N$ be the set of all possible types of samples from $\mathcal{X}^N$ and let $\mathcal{T}_Q^N$ be the set of all vectors x of the type $Q \in \mathcal{P}^N$. The entropy of RV X with PD Q and the divergence (Kullback-Leibler distance) of PDs P and Q, are defined [2, 3, 5] as follows:

$$\mathrm{H}_Q(X) \triangleq -\sum_{x\in\mathcal{X}} Q(x)\log Q(x),$$

$$\mathrm{D}(Q \parallel P) \triangleq \sum_{x\in\mathcal{X}} Q(x)\log\frac{Q(x)}{P(x)}.$$

Let us remind the following useful properties of types [2, 3].

$$|\mathcal{P}^N| \le (N+1)^{|\mathcal{X}|},$$

$$(N+1)^{-|\mathcal{X}|}.\exp\{N\mathrm{H}_Q(X)\} \le |\mathcal{T}_Q^N| \le \exp\{N\mathrm{H}_Q(X)\},$$

$$P^N(\mathbf{x}) = \exp\{-N(\mathrm{H}_Q(X) + \mathrm{D}(Q||P))\},\ \text{for}\ \mathbf{x} \in \mathcal{T}_Q^N.$$

On the base of N observations we denote the two-stage test by Φ^N, it may be composed by the pair of tests φ_1^N and φ_2^N for two consecutive stages, we write $\Phi^N = (\varphi_1^N, \varphi_2^N)$. The first stage for selection of a family of PDs is a non-randomized test $\varphi_1^N(\mathbf{x})$ based on the sample x. The next stage is for making a decision in the determined family of PDs, it is a non-randomized test $\varphi_2^N(\mathbf{x})$ based again on the same sample x and on the result of the test φ_1^N.

In Section 2, we consider the first stage of test for selecting one family of PDs and in Section 3 we construct the second stage of LAO test for accepting one PD. In Section 4, we compare reliabilities for the one-stage and the two-stage LAO hypotheses testing.

2. First Stage Test of Two Stages

Let us consider two sets of indices $\mathcal{D}_1 = \{\overline{1,R}\}$ and $\mathcal{D}_2 = \{\overline{R+1,S}\}$, then the pair of disjoint families of PDs $\mathcal{P}_1$ and $\mathcal{P}_2$ is:

$$\mathcal{P}_1 = \{P_s,\ s \in \mathcal{D}_1\},\quad \mathcal{P}_2 = \{P_s,\ s \in \mathcal{D}_2\}.$$

The first stage of decision making consists in using the sample x for the selection of one family of two of PDs by a test $\varphi_1^N(\mathbf{x})$, which can be defined by the division of the sample space $\mathcal{X}^N$ on the pair of disjoint subsets

$$\mathcal{A}_i^N \triangleq \{\mathbf{x} : \varphi_1^N(\mathbf{x}) = i\},\ \ i = 1, 2.$$

The set $\mathcal{A}_i^N$ consists of all vectors x for which i-th family $\mathcal{P}_i$ of PDs is adopted.

The test $\varphi_1^N(\mathbf{x})$ can have two kinds of errors for the pair of hypotheses $\mathcal{P}_i$, $i = 1, 2$. Let $\alpha'_{2|1}(\varphi_1^N)$ be the probability of the erroneous acceptance of the second family $\mathcal{P}_2$ provided that the first family $\mathcal{P}_1$ is true (that is the correct PD is in the first family) and $\alpha'_{1|2}(\varphi_1^N)$ be the probability of the erroneous acceptance of $\mathcal{P}_1$ provided that the second family $\mathcal{P}_2$ is true. We define

$$\alpha'_{2|1}(\varphi_1^N) \triangleq \alpha'_{1|1}(\varphi_1^N) \triangleq \max_{s:s\in\mathcal{D}_1} P_s^N(\mathcal{A}_2^N), \quad (1)$$

$$\alpha'_{1|2}(\varphi_1^N) \triangleq \alpha'_{2|2}(\varphi_1^N) \triangleq \max_{s:s\in\mathcal{D}_2} P_s^N(\mathcal{A}_1^N). \quad (2)$$

We have to consider reliabilities of the sequence of tests φ_1:

$$E'_{i|j}(\varphi_1) \triangleq \limsup_{N\to\infty}\left\{-\frac{1}{N}\log\alpha'_{i|j}(\varphi_1^N)\right\}, \quad i,j=1,2. \quad (3)$$

The reliability matrix for the first stage of the test is the following

$$\mathbf{E}'(\varphi_1) = \begin{bmatrix} E'_{1|1} & E'_{2|1} \\ E'_{1|2} & E'_{2|2} \end{bmatrix},$$

and it follows from (1)–(3) that there are only two different elements in it, namely

$$E'_{1|1} = E'_{2|1}, \quad E'_{1|2} = E'_{2|2}.$$

The test φ_1 is considered to be LAO if for the given value of $E'_{1|1}$ it provides the largest value to $E'_{2|2}$.

For the given $E'^{*}_{1|1}$ we can define LAO test φ_1^{*N} by division of $\mathcal{X}^N$ into two disjoint subsets

$$\mathcal{A}_1^{N*} = \bigcup_{Q_{\mathbf{x}}:\ \min\limits_{s:s\in\mathcal{D}_1} D(Q_{\mathbf{x}}\|P_s)\leq E'^{*}_{1|1}} \mathcal{T}_{Q_{\mathbf{x}}}^N, \qquad \mathcal{A}_2^{N*} = \mathcal{X}^N \setminus \mathcal{A}_1^{N*}.$$

We obtain the dependence $E'^{*}_{2|2}(E'^{*}_{1|1})$ applying the properties of types for the estimation of error probabilities. We estimate $\alpha'_{1|1}(\varphi_1^{*N})$ as follows:

$$\begin{aligned}
\alpha'_{1|1}(\varphi_1^{*N}) &= \max_{s:s\in\mathcal{D}_1} P_s^N\left(\mathcal{A}_2^{N*}\right) \\
&= \max_{s:s\in\mathcal{D}_1} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}\|P_l)>E'^{*}_{1|1}} \mathcal{T}_{Q_{\mathbf{x}}}^N\Big) \\
&< \max_{s:s\in\mathcal{D}_1} (N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}\|P_l)>E'^{*}_{1|1}} P_s^N\left(\mathcal{T}_{Q_{\mathbf{x}}}^N\right) \\
&\leq \max_{s:s\in\mathcal{D}_1} (N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}\|P_l)>E'^{*}_{1|1}} \exp\{-ND(Q_{\mathbf{x}}\|P_s)\} \\
&= \exp\Big\{-N[\min_{s:s\in\mathcal{D}_1} \inf_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}\|P_l)>E'^{*}_{1|1}} D(Q_{\mathbf{x}}\|P_s) - o_N(1)]\Big\} \\
&\leq \exp\left\{-N\{E'^{*}_{1|1} - o_N(1)\}\right\}.
\end{aligned}$$

We can estimate another error probability similarly:

$$\alpha'_{2|2}(\varphi_1^{*N}) = \max_{s:s\in\mathcal{D}_2} P_s^N\left(\mathcal{A}_1^{N*}\right)$$

$$= \max_{s:s\in\mathcal{D}_2} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} \mathcal{T}^N_{Q_{\mathbf{x}}}\Big)$$

$$\leq \max_{s:s\in\mathcal{D}_2} (N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} P_s^N\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right)$$

$$\leq \max_{s:s\in\mathcal{D}_2} (N+1)^{|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} \exp\{-ND(Q_{\mathbf{x}}||P_s)\}$$

$$= \exp\Big\{-N\big[\min_{s:s\in\mathcal{D}_2} \inf_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} D(Q_{\mathbf{x}}||P_s) - o_N(1)\big]\Big\}. \tag{4}$$

Now let us obtain the inverse inequality:

$$\alpha'_{2|2}(\varphi_1^{*N}) = \max_{s:s\in\mathcal{D}_2} P_s^N\left(\mathcal{A}_1^{N*}\right)$$

$$= \max_{s:s\in\mathcal{D}_2} P_s^N\Big(\bigcup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} \mathcal{T}^N_{Q_{\mathbf{x}}}\Big)$$

$$\geq \max_{s:s\in\mathcal{D}_2} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} P_s^N\left(\mathcal{T}^N_{Q_{\mathbf{x}}}\right)$$

$$\geq \max_{s:s\in\mathcal{D}_2} (N+1)^{-|\mathcal{X}|} \sup_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} \exp\{-ND(Q_{\mathbf{x}}||P_s)\}$$

$$= \exp\Big\{-N\big[\min_{s:s\in\mathcal{D}_2} \inf_{Q_{\mathbf{x}}:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q_{\mathbf{x}}||P_l)\leq E'^*_{1|1}} D(Q_{\mathbf{x}}||P_s) + o_N(1)\big]\Big\}. \tag{5}$$

According to the definition of the reliability $E'^*_{2|2}$ from (4) and (5) we conclude that

$$E'^*_{2|2}(E'^*_{1|1}) = \min_{s:s\in\mathcal{D}_2} \inf_{Q:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q||P_l)\leq E'^*_{1|1}} D(Q||P_s). \tag{6}$$

Theorem 1. *If all distributions P_s, $s=\overline{1,S}$, are different and $E'^*_{1|1}$ is such a positive number that the following inequality holds*

$$E'^*_{1|1} < \min_{s:s\in\mathcal{D}_2} \min_{l:l\in\mathcal{D}_1} D(P_s||P_l),$$

then there exists a LAO sequence of tests φ_1^* such that the reliability $E'^*_{2|2}(E'^*_{1|1})$ is positive and is defined in (6).

Corollary 1. If $R=1$, $S=2$ we have hypotheses P_1 and P_2, then we need only one-stage test and Theorem 1 in this case is equivalent to Hoeffding's Theorem [9], where for $E'^*_{1|1} < D(P_2||P_1)$,

$$E'^*_{2|2}(E'^*_{1|1}) = \inf_{Q:D(Q||P_1)\leq E'^*_{1|1}} D(Q||P_2).$$

From Theorem 1 the solution of the problem of LAO identification for the model with one family of S hypotheses can also be obtained.

The LAO statistical identification, which was considered in [1], [7], [9], gives the answer to the question: whether r-th PD occurred, or not. There are two error probabilities for each r: $\alpha_{l\neq r|s=r}$, $r,s=\overline{1,S}$, is the error probability that P_r is correct but it is rejected and $\alpha_{l=r|s\neq r}$ is the error probability that P_r is selected but it is not correct. The reliability approach to identification is to determine the optimal dependence of the reliability $E_{l=r|s\neq r}$ upon the given reliability $E_{l\neq r|s=r}$.

Corollary 2. *When we consider the sets $\mathcal{D}_1 = \{P_r\}$, $r = \overline{1,S}$, and $\mathcal{D}_2 = \{P_s : s \neq$, $s = \overline{1,S}\}$, Theorem 1 gives the result of [1], that is for $E_{l\neq r|s=r} < \min_{s:s\neq r} \mathrm{D}(P_s\|P_r)$, we have a solution of the problem of LAO identification:*

$$E_{l=r|s\neq r}(E_{l\neq r|s=r}) = \min_{s:s\neq r} \inf_{Q:\mathrm{D}(Q\|P_r)\leq E_{l\neq r|s=r}} \mathrm{D}(Q\|P_s).$$

. Second Stage Test of Two Stages

After selecting a family of PDs from the two, it is necessary to detect one PD in this family. If the first family of PDs is accepted, then we consider the test $\varphi_2^N(\mathbf{x})$ which can be defined by the division of the sample space $\mathcal{A}_1^N$ to R disjoint subsets

$$B_s^N \triangleq \{\mathbf{x} : \varphi_2^N(\mathbf{x}) = s\}, \quad s \in \mathcal{D}_1.$$

Let $\alpha''_{l|s}\left(\varphi_2^N\right)$ be the probability of the erroneous acceptance at the second stage of the test, in which PD P_l is accepted when P_s is true:

$$\alpha''_{l|s}\left(\varphi_2^N\right) \triangleq P_s^N\left(B_l^N\right), \; l \in \mathcal{D}_1, \; s = \overline{1,S}.$$

The probability to reject P_s, when it is true, is

$$\alpha''_{s|s}\left(\varphi_2^N\right) \triangleq P_s^N\left(\overline{B}_s^N\right) = \sum_{l=1,\, l\neq s}^{R} \alpha''_{l|s}\left(\varphi_2^N\right) + P_s(\mathcal{A}_2^N), \quad s \in \mathcal{D}_1. \tag{7}$$

Corresponding reliabilities for the second stage of the test, are defined as

$$E''_{l|s}(\varphi_2) \triangleq \limsup_{N\to\infty} \left\{-\frac{1}{N}\log \alpha''_{l|s}\left(\varphi_2^N\right)\right\}, \quad l \in \mathcal{D}_1, \; s = \overline{1,S}. \tag{8}$$

Using properties of types, we obtain the following equalities:

$$\lim_{N\to\infty} \left\{-\frac{1}{N}\log P_s^N(\mathcal{A}_2^N)\right\} = \inf_{Q:\, \min_{l:l\in\mathcal{D}_1} D(Q\|P_l) > E'_{1|1}} D(Q\|P_s) \triangleq E^I_{2|s}. \tag{9}$$

From (7)–(9) it follows that

$$E''_{s|s}(\varphi_2) = \min\left[\min_{r\in\mathcal{D}_1} E''_{r|s}(\varphi_2), E^I_{2|s}\right], \quad s \in \mathcal{D}_1.$$

If at the first stage the first family of PDs is accepted, the reliability matrix for the second stage of the test $\mathbf{E}''(\varphi_2)$ is the following:

$$\mathbf{E}''(\varphi_2) = \begin{bmatrix} E''_{1|1} & E''_{2|1} & \cdots & E''_{R|1} \\ E''_{1|2} & E''_{2|2} & \cdots & E''_{R|2} \\ \cdots & \cdots & \cdots & \cdots \\ E''_{1|R} & E''_{2|R} & \cdots & E''_{R|R} \\ E''_{1|R+1} & E''_{2|R+1} & \cdots & E''_{R|R+1} \\ E''_{1|R+2} & E''_{2|R+2} & \cdots & E''_{R|R+2} \\ \cdots & \cdots & \cdots & \cdots \\ E''_{1|S} & E''_{2|S} & \cdots & E''_{R|S} \end{bmatrix}.$$

In the following theorem we show the optimal dependence of reliabilities.

Theorem 2: *If in the first stage of test the first family of PDs is accepted, then for the given positive and finite values $E''_{1|1}, E''_{2|2}, \ldots, E''_{R-1|R-1}$ of the reliability matrix $\mathbf{E}''(\varphi_2)$, let us consider the regions:*

$$\mathcal{R}''_s = \left\{Q : \mathrm{D}(Q \parallel P_s) \leq E''_{s|s}\right\}, \quad s = \overline{1, R-1},$$

$$\mathcal{R}''_R = \left\{Q : \min_{l \in \mathcal{D}_1} \mathrm{D}(Q \| P_l) \leq E'^{*}_{1|1} \text{ and } \mathrm{D}(Q \parallel P_s) > E''_{s|s}, \quad s = \overline{1, R-1}\right\},$$

and the following values of elements of the future reliability matrix $\mathbf{E}''(\varphi_2^*)$ of the LAO test sequence:

$$E''^{*}_{s|s} = E''_{s|s}, \quad s = \overline{1, R-1},$$

$$E''^{*}_{l|s} = \inf_{Q \in \mathcal{R}''_l} \mathrm{D}(Q \parallel P_s), \quad l \in \mathcal{D}_1, \ s = \overline{1, S}, \ l \neq s.$$

$$E''^{*}_{R|R} = \min\left[\min_{l: l < R} E''^{*}_{l|R}, \ E^{I}_{2|R}\right].$$

When the following compatibility conditions are valid

$$E''_{1|1} < \min[\min_{s = \overline{2, R}} \mathrm{D}(P_s \parallel P_1), E^{I}_{2|1}],$$

$$E''_{s|s} < \min[\min_{1 \leq l < s} E''^{*}_{l|s}, \min_{s < l \leq R} \mathrm{D}(P_l \parallel P_s), E^{I}_{2|s}], \quad 2 \leq s \leq R-1,$$

then there exists a LAO sequence of tests φ_2^*, the elements of reliability matrix $\mathbf{E}''(\varphi_2^*)$ $\{E''^{*}_{l|s}\}$ of which are defined above and are positive.

When even one of the compatibility conditions is violated, then at least one element of the matrix $\mathbf{E}(\varphi_2^*)$ is equal to 0.

If the second family of PDs is accepted, then the test $\varphi_2^N(\mathbf{x})$ is a division of the sample space $\mathcal{A}_2^N$ to $S-R$ disjoint subsets B_s^N, $s \in \mathcal{D}_2$, such that

$$B_s^N \triangleq \{\mathbf{x} : \varphi_2^N(\mathbf{x}) = s\}, \quad s \in \mathcal{D}_2,$$

Let $\alpha''_{l|s}\left(\varphi_2^N\right)$ be the probability of the erroneous acceptance at the second stage of the test in which PD P_l is accepted when P_s is true. So

$$\alpha''_{l|s}\left(\varphi_2^N\right) \triangleq P_s^N\left(B_l^N\right), \ l \in \mathcal{D}_2, \ s = \overline{1, S}.$$

The probability to reject P_s, when it is true, is

$$\alpha''_{s|s}\left(\varphi_2^N\right) \triangleq P_s^N\left(\overline{B}_s^N\right) = \sum_{l=R+1}^{S} \alpha''_{l|s}\left(\varphi_2^N\right) + P_s(\mathcal{A}_1^N), \quad s \in \mathcal{D}_2. \qquad (1\ )$$

Corresponding reliabilities for the second stage of the test, are defined as

$$E''_{l|s}(\varphi_2) \triangleq \limsup_{N \to \infty}\left\{-\frac{1}{N}\log \alpha''_{l|s}\left(\varphi_2^N\right)\right\}, \quad l \in \mathcal{D}_2, \ s = \overline{1, S}. \qquad (1\ )$$

Using properties of types, we obtain the following equalities:

$$\lim_{N\to\infty}\left\{-\frac{1}{N}\log P_s^N(\mathcal{A}_1^N)\right\} = \inf_{Q:\ \min\limits_{l:l\in\mathcal{D}_1} D(Q\|P_l)\le E'^{*}_{1|1}} D(Q\|P_s) \triangleq E^I_{1|s}. \tag{12}$$

From (10)-(12) it follows that

$$E''_{s|s}(\varphi_2) = \min\left[\min_{l\in\mathcal{D}_2} E''_{l|s}(\varphi_2), E^I_{1|s}\right],\ \ s\in\mathcal{D}_2.$$

If at the first stage, the second family of PDs is accepted, the reliability matrix for the second stage of the test $\mathbf{E}''(\varphi_2)$ is the following:

$$\mathbf{E}''(\varphi_2) = \begin{bmatrix} E''_{R+1|1} & E''_{R+2|1} & \dots & E''_{S|1} \\ E''_{R+1|2} & E''_{R+2|2} & \dots & E''_{S|2} \\ \dots & \dots & \dots & \dots \\ E''_{R+1|R} & E''_{R+2|R} & \dots & E''_{S|R} \\ E''_{R+1|R+1} & E''_{R+2|R+1} & \dots & E''_{S|R+1} \\ E''_{R+1|R+2} & E''_{R+2|R+2} & \dots & E''_{S|R+2} \\ \dots & \dots & \dots & \dots \\ E''_{R+1|S} & E''_{R+2|S} & \dots & E''_{S|S} \end{bmatrix}.$$

Theorem 3: *If in the first stage of test, the second family of PDs is accepted, then for the given positive and finite values $E''_{R+1|R+1}, E''_{R+2|R+2}, \dots, E''_{S|S}$ of the reliability matrix $\mathbf{E}''(\varphi_2)$, let us consider the regions*

$$\mathcal{R}''_s = \left\{Q : \mathrm{D}(Q\parallel P_s)\le E''_{s|s}\right\},\ \ s=\overline{R+1,S-1},$$

$$\mathcal{R}''_R = \left\{Q:\ \min_{l:l\in\mathcal{D}_1}\mathrm{D}(Q\|P_l) > E'^{*}_{1|1}\ \textit{and}\ \mathrm{D}(Q\parallel P_s) > E''_{s|s},\ \ s=\overline{R+1,S-1}\right\},$$

and the following values of elements of the future reliability matrix $\mathbf{E}''(\varphi_2^*)$ of the LAO tests sequence:

$$E''^{*}_{s|s} = E''_{s|s},\ \ s=\overline{R+1,S-1},$$

$$E''^{*}_{l|s} = \inf_{Q\in\mathcal{R}''_l}\mathrm{D}(Q\parallel P_s),\ \ l\in\mathcal{D}_2,\ s=\overline{1,S}, l\neq s,$$

$$E''^{*}_{S|S} = \min\left[\min_{R<l<S} E''^{*}_{l|R},\ E^I_{1|S}\right].$$

When the following compatibility conditions are valid

$$E''_{R+1|R+1} < \min[\min_{s=\overline{R+2,S}}\mathrm{D}(P_s\parallel P_1), E^I_{1|R+1}],$$

$$E''_{s|s} < \min[\min_{R+1\le l<s} E''^{*}_{l|s}, \min_{s<l\le S}\mathrm{D}(P_l\parallel P_s), E^I_{1|s}],\ \ R+2\le s\le S-1,$$

then there exists a LAO sequence of tests φ_2^*, the elements of reliability matrix $\mathbf{E}''(\varphi_2^*) = \{E''^{*}_{l|s}\}$ of which are defined above and are positive.

When even one of the compatibility conditions is violated, then at least one element of the matrix $\mathbf{E}(\varphi_2^*)$ is equal to 0.

The proofs of Theorems 2 and 3 are similar to the proof of Theorem 1 from [5], where the method of types was used for the proof.

4. Reliabilities of Two-stage Test and Its Comparison to the One-stage Tes

We define a two-stage test by Φ. From the results of the first and the second stages w obtain the final reliabilities $E'''_{l|s}(\Phi)$, $l,s=\overline{1,S}$. Corresponding error probabilities are $\alpha'''_{l|}$ $l,s=\overline{1,S}$, it means that the hypothesis s is true, but we accepted hypothesis l. If l and are from the same set $\mathcal{D}_i$, $i=\overline{1,2}$, we do not have an error in the first stage, but the samp $\mathbf{x}$ is from $B_s^N \subset \mathcal{A}_i^{N*}$, that is $\alpha'''_{l|s}=P_s^N\left(B_l^N\right)=\alpha''_{l|s}$. When l and s are from different set $l\in\mathcal{D}_i$, $s\in\mathcal{D}_j$, $i\neq j$, $i,j=\overline{1,2}$, we can say that our wrong decision comes from the fir stage. It means that in the first stage the sample $\mathbf{x}$ belongs to $\mathcal{A}_i^{N*}$, and in the second stag $\mathbf{x}\in B_l^N\subset\mathcal{A}_i^{N*}$. That is $\alpha'''_{l|s}=P_s^N\left(B_l^N\right)=\alpha''_{l|s}$.
According to this discourse and the definition of the reliability we formulate the following

Theorem 4: *If all distributions P_s, $s=\overline{1,S}$, are different and positive numbers $E'^{*}_{1|1}$ ar $E''_{r|r}$, $r=\overline{1,R-1}\cup\overline{R+1,S-1}$ satisfy the compatibility conditions of Theorems 1, 2 and then*

$$E'''_{l|s}(\Phi)=E''_{l|s},\ \ l,s=\overline{1,S}.$$

Now we want to compare the reliabilities for two methods: of the one-stage test [5] an of the two-stage test. For comparison we will give the same diagonal elements $E_{s|s}=E'''_{s|}$ $s=\overline{1,S-1}$ of the reliability matrices of two-stage and one-stage cases.

For the one-stage test the elements of each column r, $r=\overline{1,S-1}$ are functions of diagonal element of the same column. For the two-stage test the element of columns $r=\overline{1,R-1}\cup\overline{R+1,S-1}$ are also functions of diagonal elements of the correspondir column, moreover, the values are the same for both cases, that is the elements of the colum $r=\overline{1,R-1}\cup\overline{R+1,S-1}$ of two matrices are equal.

The elements of the column R for the two-stage test we obtain by these formulas:
$E'''^{*}_{R|s}=\inf\limits_{Q\in\mathcal{R}'''_R}\mathrm{D}(Q\parallel P_s)$,
where $\mathcal{R}'''_R=\left\{Q:\ \min\limits_{l:l\in\mathcal{D}_1}\mathrm{D}(Q||P_l)\leq E'^{*}_{1|1}\text{ and }\mathrm{D}(Q\parallel P_s)>E'''_{s|s},\ \ s=\overline{1,R-1}\right\}$.

The elements of the column S for the one stage test are the following:
$E^{*}_{R|s}=\inf\limits_{Q\in\mathcal{R}_R}\mathrm{D}(Q\parallel P_s)$, where $\mathcal{R}_R=\{Q:\ D(Q\parallel P_R)\leq E_{R|R}\}$.

When we consider that $E_{s|s}=E'''_{s|s}$, $s=\overline{1,S-1}$ then we have two cases.

1. If $\mathcal{R}'''_R\subseteq\mathcal{R}_R$ or $\mathcal{R}_R$ consists of Q elements such that $\min\limits_{l:l\in\mathcal{D}_1}\mathrm{D}(Q||P_l)>E'^{*}_{1|1}$, then $E'''^{*}_{R|s}\geq E^{*}_{R|s}$.

2. If $\mathcal{R}'''_R\supseteq\mathcal{R}_R$, then $E'''^{*}_{R|s}\leq E^{*}_{R|s}$

For the elements of the column S we also have similar cases.

So, the elements of R-th and S-th columns of the two-stage reliabilities matrix can l smaller or greater than the corresponding elements of the one-stage matrix.

The procedure of testing for the two-stage test can be shorter when the first family distributions is accepted at the first stage. When in the first stage the second family accepted, then the procedures of the two-stage test and of the one-stage test are of an equ length.

5. Conclusion

We have shown that the number of the preliminarily given elements of the reliabilities matrix in two-stage and one-stage tests would be the same but the procedure of calculations for the first one would be shorter. Some element of the reliabilities matrix of the two-stage test can be even greater than corresponding elements of the one-stage test. So the customer has a possibility to use the method which is preferable.

In the next work we have an intention to consider the two-stage LAO test by the pair of samples $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X}^N$, $\mathbf{x}_1 = (x_1, x_2, \ldots, x_{N_1})$, $\mathbf{x}_1 \in \mathcal{X}^{N_1}$, $\mathbf{x}_2 = (x_{N_1+1}, x_{N_1+2}, \ldots, x_N)$, $\mathbf{x}_2 \in \mathcal{X}^{N_2}$, $N = N_1 + N_2$, $\mathcal{X}^N = \mathcal{X}^{N_1} \times \mathcal{X}^{N_2}$. The first stage of decision making consists in using sample $\mathbf{x}_1$ for selection of a family of PDs and after selecting a family of PDs, it is necessary to detect one PD in this family by using the sample $\mathbf{x}_2$.

References

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

[2] T.M Cover and J.A. Tomas. *Elements of Information Theory*, Second edition, Wiley, NewYork, 2006.

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

[4] 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.

[5] E.A. Haroutunian. "Logarithmically asymptotically optimal testing of multiple statistical hypotheses", *Problems of Control and Information Theory*, vol 19, no. 5-6, pp. 413-421, 1990.

[6] E.A. Haroutunian. "Reliability in multiple hypotheses testing and identification ", *Proceedings of the NATO-ASI Conference, vol. 198 of NATO Science Series III: Computer and Systems Sciences*, Yerevan, Armenia, pp. 189-201, IOS Press, 2008.

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

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

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

[10] G. Tusnady. "On asymptotically optimal tests", *Annals of Statistics*, vol. 5, no. 2, pp. 385-393, 1977.

Երկու ընտանիքներ կազմող բաշխումների վերաբերյալ բազմակի վարկածներ երկփուլ լոգարիթմորեն ասիմպտոտորեն օպտիմալ տեստավորում

Ե. Հարությունյան, Փ. Հակոբյան և Ֆ. Հորմոզի նեժադ

Ամփոփում

Դիտարկվել է բաշխումների երկու ընտանիքներից կազմված մոդելի վերաբերյա բազմակի վիճակագրական վարկածների ստուգման գործընթացը: Հետազոտվել ե երկու փուլերից յուրաքանչյուրի սխալների զույգերի ցուցիչների (հուսալիությունների փոխկապվածությունների մատրիցները: Համեմատվել են երկփուլ տեստի և միավո տեստի հուսալիությունները և գործընթացների երկարությունները:

О двух-этапном логарифмически асимптотически оптимальном тестировании многих гипотез относительно распределений из двух семейств

Е. Арутюнян, П. Акопян и Ф. Хормози нежад

Аннотация

Рассмотрено тестирование модели состояшей из двух семейств возможны распределений вероятностей. Матрица надежностей логорифмически асимпт тически оптимального тестирования в два этапа изучена и сравнена со случае аналогичного одноэтапного тестирования.