Mathematical Problems of Computer Science 23, 2004, 36–46.

On Multiple Hypotheses Testing by Informed Statistician for Arbitrarily Varying Object and Application to Source Coding*

Evgueni A. Haroutunian and Parandzem M. Hakobyan

Institue for Informatics and Automation Problems of NAS of RA
e-mails evhar@ipia.sci.am, par_h@ipia.sci.am

Abstract

The matrix of asymptotic interdependencies (reliability–reliability functions) of all possible pairs of the error probability exponents (reliabilities) in testing of multiple statistical hypotheses is studied for arbitrarily varying object with the current states sequence known to the statistician. The case of two hypotheses when state sequences are not known to the decision maker was studied by Fu and Shen, and when decision is founded on the known states sequence was considered by Ahlswede, Haroutunian and Aloyan.

In the same way as Fu and Shen we obtain from the main result rate-reliability and reliability-rate functions for arbitrarily varying source coding with side information. An illustrative example is presented.

1. Formulation of Results

The problem solved here is induced by the ideas of the paper of R. Ahlswede [1] concerning arbitrarily varying sources. It is natural generalization of the problem considered in [2], [3] in resolution of the problem proposed by R. L. Dobrushin [4] and devoted to discrete memoryless sources. The case of two hypotheses was considered in [5]. An arbitrarily varying object is a generalized model of the discrete memoryless one. Let $\mathcal{X}$ be a finite set of values of random variable X, and $\mathcal{S}$ is an alphabet of states. L conditional probability distributions on X depending on values s of states are known:

$$G_l = \{G_l(x|s),\ x \in \mathcal{X}, s \in \mathcal{S}\},\ \ l = \overline{1, L}.$$

The conditional probability distribution of the characteristic X of an object depends at any time instant n only on current state s_n. The statistician must select one among L alternative hypotheses $H_l : G = G_l,\ \ l = \overline{1, L}$. Let $\mathbf{x} = (x_1, ..., x_N)$ be a sequence of results of N observations of the object. We consider the model when the source of states produces arbitrary varying sequence $\mathbf{s} = (s_1, ..., s_N)$, which is connected with the statistician who must accept one of the hypotheses knowing sample $\mathbf{x}$ and the state sequence $\mathbf{s}$ of the same length

*The work was partially supported by INTAS, project 00–738.

N (as side information). It is supposed that for all l in each moment n the probabilities of observation x_n depends only on state s_n and is independent of the states and observations of other moments, that is

$$G_l^N(\mathbf{x}|\mathbf{s}) \triangleq \prod_{n=1}^{N} G_l(x_n|s_n), \quad l = \overline{1,L}. \tag{1}$$

The procedure of decision making is a non-randomized test $\varphi^N(\mathbf{x},\mathbf{s})$, it can be defined by division of the sample space $\mathcal{X}^N$ for each state sequence $\mathbf{s}$ on L disjoint subsets $\mathcal{A}_l^N(\mathbf{s}) = \{\mathbf{x} : \varphi^N(\mathbf{x},\mathbf{s}) = l\}, l = \overline{1,L}$. The set $\mathcal{A}_l^N(\mathbf{s})$ consists of all vectors $\mathbf{x}$ for which the hypothesis H_l is adopted knowing state sequence $\mathbf{s}$. We study for all pairs $l, m = \overline{1,L},\ m \neq l$, the probabilities of the erroneous acceptance of hypothesis H_l provided that H_m is true

$$\alpha_{m|l}^{(N)}(\varphi) = \max_{\mathbf{s}\in\mathcal{S}^N} G_m^N(\mathcal{A}_l^N(\mathbf{s})|\mathbf{s}). \tag{2}$$

The probability to reject H_m, when it is true, is also considered

$$\alpha_{m|m}^{(N)}(\varphi) = \sum_{l\neq m} \alpha_{m|l}^{(N)}(\varphi) = \max_{\mathbf{s}\in\mathcal{S}^N} G_m^N(\overline{\mathcal{A}_m^N(\mathbf{s})}|\mathbf{s}) = \max_{\mathbf{s}\in\mathcal{S}^N}(1 - G_m^N(\mathcal{A}_m^N(\mathbf{s})|\mathbf{s}). \tag{3}$$

Corresponding error probability exponents, called "reliabilities", are defined as

$$E_{m|l}(\varphi) \triangleq \overline{\lim_{N\to\infty}} - N^{-1} \log \alpha_{m|l}^{(N)}(\varphi), \ m, l = \overline{1,L}. \tag{4}$$

In the paper functions exp and log are considered at the base 2. It follows from (3) that

$$E_{m|m}(\varphi) = \min_{l\neq m} E_{m|l}(\varphi), \ m = \overline{1,L}. \tag{5}$$

The matrix $\mathbf{E} = \{E_{m|l}(\varphi)\}$ is called the reliability matrix of the tests sequence and is the object of our investigation.

Definition 1: *Following Birgé [6] we call the sequence of tests logarithmically asymptotically optimal (LAO) if for given positive values of $L-1$ diagonal elements of the matrix $\mathbf{E}$ the procedure provides maximal values for other elements of it.*

We exploit some combinatorial notions and facts [7], [8]. For $\mathbf{s} = (s_1, \ldots, s_N),\ \mathbf{s} \in \mathcal{S}^N$, let $N(s \mid \mathbf{s})$ be the number of occurrences of $s \in \mathcal{S}$ in the vector $\mathbf{s}$. The type (or empirical distribution) of $\mathbf{s}$ is the distribution $P_\mathbf{s} = \{P_\mathbf{s}(s), s \in \mathcal{S}\}$ defined by

$$P_\mathbf{s}(s) = \frac{1}{N} N(s \mid \mathbf{s}), \qquad s \in \mathcal{S}.$$

For a pair of sequences $\mathbf{x} \in \mathcal{X}^N$ and $\mathbf{s} \in \mathcal{S}^N$, let $N(x, s|\mathbf{x},\mathbf{s})$ be the number of occurences of $(x,s) \in \mathcal{X} \times \mathcal{S}$ in the pair of vectors $(\mathbf{x},\mathbf{s})$. The joint type of the pair $(\mathbf{x},\mathbf{s})$ is the distribution $Q_{\mathbf{x},\mathbf{s}} = \{Q_{\mathbf{x},\mathbf{s}}(x,s), x \in \mathcal{X},\ s \in \mathcal{S}\}$ defined by

$$Q_{\mathbf{x},\mathbf{s}}(x,s) = \frac{1}{N} N(x, s \mid \mathbf{x},\mathbf{s}), \qquad x \in \mathcal{X}, \quad s \in \mathcal{S}.$$

The conditional type of $\mathbf{x}$ for given $\mathbf{s}$ is the conditional distribution $Q_{\mathbf{x}|\mathbf{s}} = \{Q_{\mathbf{x}|\mathbf{s}}(x|s), x \in \mathcal{X},\ s \in \mathcal{S}\}$ defined by

$$Q_{\mathbf{x}|\mathbf{s}}(x|s) = \frac{Q_{\mathbf{x},\mathbf{s}}(x,s)}{P_\mathbf{s}(s)} = \frac{N(x,s|\mathbf{x},\mathbf{s})}{N(s \mid \mathbf{s})}, \qquad x \in \mathcal{X}, \quad s \in \mathcal{S}.$$

Let X and S are some random variables defined by probability distributions $P = \{P(s), s \in \mathcal{S}\}$ and $Q = \{Q(x|s), x \in \mathcal{X}, s \in \mathcal{S}\}$. The conditional entropy of X for given S is:

$$H_{P,Q}(X \mid S) = -\sum_{x,s} P(s)Q(x|s)\log Q(x|s).$$

The conditional divergences (Kullback-Leibler information) of the distribution $P \circ Q = =\{P(s)Q(x|s),\ x \in \mathcal{X},\ s \in \mathcal{S}\}$ with respect to the distribution $P \circ G_l=\{P(s)\,G_l(x|s),\ x \in \mathcal{X},\ s \in \mathcal{S}\}$ is

$$D(P \circ Q||P \circ G_l|P) = D(Q||G_l|P) = \sum_{x,s} P(s)Q(x|s)\log\frac{Q(x|s)}{G_l(x|s)},\ \ l = \overline{1, L}.$$

We denote by $\mathcal{P}^N(\mathcal{S})$ the set of all types on $\mathcal{S}$ for given N, by $\mathcal{P}(\mathcal{S})$ – the set of all possible probability distributions P on $\mathcal{S}$ and by $\mathcal{Q}^N(\mathcal{X}|\mathbf{s})$ – the set of all possible conditional types on $\mathcal{X}$ for given $\mathbf{s}$. Let $\mathcal{T}^N_{P_\mathbf{s},Q}(X \mid \mathbf{s})$ is the family of vectors $\mathbf{x}$ of the conditional type Q for given $\mathbf{s}$ of the type $P_\mathbf{s}$. It is known [7] that

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

$$(N+1)^{-|\mathcal{X}||\mathcal{S}|}\exp\{NH_{P_\mathbf{s},Q}(X|S)\} \leq |\ \mathcal{T}^N_{P_\mathbf{s},Q}(X \mid \mathbf{s})\ | \leq \exp\{NH_{P_\mathbf{s},Q}(X|S)\}. \tag{7}$$

Given positive numbers $E_{1|1}, ..., E_{L-1|L-1}$, let us define for each $P \in \mathcal{P}(\mathcal{S})$:

$$\mathcal{R}_l(P) \triangleq \{Q:\ D(Q||G_l|P) \leq E_{l|l}\},\ \ l = \overline{1, L-1}, \tag{8a}$$

$$\mathcal{R}_L(P) \triangleq \{Q:\ D(Q||G_l|P) > E_{l|l},\ \ l = \overline{1, L-1}\}, \tag{8b}$$

$$\mathcal{R}^{(N)}_l(P_\mathbf{s}) \triangleq R_l(P_\mathbf{s}) \bigcap \mathcal{Q}^N(\mathcal{X}|\mathbf{s}),\ \ l = \overline{1, L},\ \ \mathbf{s} \in \mathcal{S}^N. \tag{8c}$$

$$E^*_{l|l} = E^*_{l|l}(E_{l|l}) \triangleq E_{l|l},\ \ \ l = \overline{1, L-1}, \tag{9a}$$

$$E^*_{m|l} = E^*_{m|l}(E_{l|l}) \triangleq \inf_{P \in \mathcal{P}(S)}\ \inf_{Q \in \mathcal{R}_l(P)} D(Q||G_m|P),\ \ m = \overline{1, L},\ m \neq l,\ \ l = \overline{1, L-1}, \tag{9b}$$

$$E^*_{m|L} = E^*_{m|L}(E_{1|1}, E_{2|2}, ..., E_{L-1|L-1}) \triangleq \inf_{P \in \mathcal{P}(S)}\ \inf_{Q \in \mathcal{R}_L(P)} D(Q||G_m|P),\ \ \ m = \overline{1, L-1}, \tag{9c}$$

$$E^*_{L|L} = E^*_{L|L}(E_{1|1}, E_{2|2}, ..., E_{L-1|L-1}) \triangleq \min_{l: l=\overline{1, L-1}} E^*_{L|l}. \tag{9d}$$

The main result of the paper is formulated in

Theorem 1: *If all conditional distributions G_l, $l = \overline{1, L}$, are different in the sense that $D(G_l||G_m|P) > 0, l \neq m$, for every $P \in \mathcal{P}(\mathcal{S})$, and the positive numbers $E_{1|1}, E_{2|2}, ..., E_{L-1|L-1}$ are such that the following inequalities hold*

$$E_{1|1} < \min_{l=\overline{2,L}}\ \min_{P \in \mathcal{P}(S)} D(G_l||G_1|P),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \tag{10}$$

$$E_{m|m} < \min_{l=\overline{m+1,L}}\ \min_{P \in \mathcal{P}(S)} D(G_l||G_m|P),\ \ and\ \ E_{m|m} \leq \min_{l=\overline{1,m-1}} E^*_{m|l}(E_{l|l}),\ \ \ m = \overline{2, L-1},$$

then there exists a LAO sequence of tests, the reliability matrix of which $\mathbf{E}^* = \{E^*_{m|l}\}$ is defined in (9) and all elements of it are positive.

When one of the inequalities (10) is violated, then at least one element of the matrix $\mathbf{E}^*$ is equal to 0 .

Proof : For $\mathbf{s} \in \mathcal{S}^N$, $\mathbf{x} \in \mathcal{T}^N_{P_s,Q}(X|\mathbf{s})$ with fixed Q the conditional probability of $\mathbf{x}$ given $\mathbf{s}$ according to (1) can be presented as follows:

$$G^N_m(\mathbf{x} \mid \mathbf{s}) = \prod_{n=1}^{N} G_m(x_n|s_n) = \prod_{x,s} G_m(x|s)^{N(x,s|\mathbf{x},\mathbf{s})} = \prod_{x,s} G_m(x|s)^{NP_s(s)Q(x|s)} =$$

$$= \exp\{N \sum_{x,s}(-P_s(s)Q(x|s) \log \frac{Q(x|s)}{G_m(x|s)} + P_s(s)Q(x|s) \log Q(x|s))\} = \quad (11)$$

$$= \exp\{-N[D(Q \parallel G_m|P_s) + H_{P_s,Q}(X \mid S)]\}.$$

Let us consider the following sequence of tests φ^* defined for each $\mathbf{s} \in \mathcal{S}^N$ by the sets

$$B^{(N)}_l(\mathbf{s}) = \bigcup_{Q \in \mathcal{R}^{(N)}_l(P_s)} \mathcal{T}^N_{P_s,Q}(X|\mathbf{s}), \; l = \overline{1,L}. \quad (12)$$

We can show that each $\mathbf{x}$ is in one and only in one of $B^{(N)}_l(\mathbf{s})$, that is

$$B^{(N)}_l(\mathbf{s}) \bigcap B^{(N)}_m(\mathbf{s}) = \emptyset, \; l \neq m, \quad \text{and} \quad \bigcup_{l=1}^{L} B^{(N)}_l(\mathbf{s}) = \mathcal{X}^N.$$

Really, for $l = \overline{1, L-2}$, $m = \overline{2, L-1}$, for each $l < m$ and $\mathbf{s} \in \mathcal{S}^N$ let us consider arbitrary $\mathbf{x} \in B^{(N)}_l(\mathbf{s})$. It follows (from (8) and (12)) that there are $Q \in \mathcal{Q}^N(\mathcal{X}|\mathbf{s})$ such that $D(Q||G_l|P_s) \leq E_{l|l}$ and $\mathbf{x} \in \mathcal{T}^N_{P_s,Q}(X|\mathbf{s})$. From (8)–(10) we have $E_{m|m} < E^*_{m|l}(E_{l|l}) < D(Q||G_m|P_s)$. From definition of $B^{(N)}_m(\mathbf{s})$ we see that $\mathbf{x} \notin B^{(N)}_m(\mathbf{s})$. (8) and (12) show also that

$$B^{(N)}_L(\mathbf{s}) \bigcap B^{(N)}_l(\mathbf{s}) = \emptyset, \; l = \overline{1, L-1}.$$

Now, let us see that for $m = \overline{1, L-1}$, using (3), (6), (7), (8), (10), (11) and (12) we can estimate $\alpha^{(N)}_{m|m}(\varphi^*)$ as follows:

$$\alpha^{(N)}_{m|m}(\varphi^*) = \max_{\mathbf{s} \in \mathcal{S}^N} G^N_m(\overline{B^{(N)}_m(\mathbf{s})}|\mathbf{s}) = \max_{\mathbf{s} \in \mathcal{S}^N} G^N_m \left(\bigcup_{Q: D(Q||G_m|P_s) > E_{m|m}} \mathcal{T}^N_{P_s,Q}(X|\mathbf{s}) \middle| \mathbf{s} \right) \leq$$

$$\leq \max_{\mathbf{s} \in \mathcal{S}^N} (N+1)^{|\mathcal{X}||\mathcal{S}|} \sup_{Q: D(Q||G_m|P_s) > E_{m|m}} G_m(\mathcal{T}^N_{P_s,Q}(X|\mathbf{s})|\mathbf{s}) \leq$$

$$\leq (N+1)^{|\mathcal{X}||\mathcal{S}|} \sup_{P_s \in \mathcal{P}^N(S)} \sup_{Q: D(Q||G_m|P_s) > E_{m|m}} \exp\{-ND(Q||G_m|P_s)\} \leq$$

$$\leq \exp\{-N[\inf_{P \in \mathcal{P}(S)} \inf_{Q: D(Q||G_m|P) > E_{m|m}} D(Q||G_m|P) - o_N(1)]\} \leq \exp\{-N[E_{m|m} - o_N(1)]\},$$

where $o_N(1) \to 0$ with $N \to \infty$.

For $l = \overline{1, L-1}$, $m = \overline{1, L}$, $l \neq m$, we can obtain similar estimates:

$$\alpha^{(N)}_{m|l}(\varphi^*) = \max_{\mathbf{s} \in \mathcal{S}^N} G^N_m(B^{(N)}_l(\mathbf{s})|\mathbf{s}) = \max_{\mathbf{s} \in \mathcal{S}^N} G^N_m \left(\bigcup_{Q: D(Q||G_l|P_s) \leq E_{l|l}} \mathcal{T}^N_{P_s,Q}(X|\mathbf{s}) \middle| \mathbf{s} \right) \leq$$

$$\leq \max_{s\in\mathcal{S}^N}(N+1)^{|\mathcal{X}||\mathcal{S}|} \sup_{Q:D(Q||G_m|P_s)\leq E_{l|l}} G_m^N(\mathcal{T}_{P_s,Q}^N(X|\mathbf{s})|\mathbf{s}) \leq$$
$$\leq (N+1)^{|\mathcal{X}||\mathcal{S}|} \sup_{P_s\in\mathcal{P}^N(S)} \sup_{Q:D(Q||G_m|P_s)\leq E_{l|l}} \exp\{-ND(Q||G_m|P_s)\} = \tag{13}$$
$$= \exp\{-N(\inf_{P_s\in\mathcal{P}^N(S)} \inf_{Q:D(Q||G_m|P_s)\leq E_{l|l}} D(Q||G_m|P_s) - o_N(1))\}.$$

Now let us prove the inverse inequality

$$\alpha_{m|l}^{(N)}(\varphi^*) = \max_{s\in\mathcal{S}^N} G_m^N(\mathcal{B}_l^{(N)}(\mathbf{s})|\mathbf{s}) = \max_{s\in\mathcal{S}^N} G_m^N\left(\bigcup_{Q:D(Q||G_l|P_s)\leq E_{l|l}} \mathcal{T}_{P_s,Q}^N(X|\mathbf{s})\middle|\mathbf{s}\right) \geq$$
$$\geq \max_{s\in\mathcal{S}^N} \sup_{Q:D(Q||G_l|P_s)\leq E_{l|l}} G_m^N(\mathcal{T}_{P_s,Q}(X|\mathbf{s})|\mathbf{s}) \geq$$
$$\geq \sup_{P_s\in\mathcal{P}^N(S)} (N+1)^{-|\mathcal{X}||\mathcal{S}|} \sup_{Q:D(Q||G_l|P_s)\leq E_{l|l}} \exp\{-ND(Q||G_m|P_s)\} = \tag{14}$$
$$= \exp\{-N(\inf_{P_s\in\mathcal{P}^N(S)} \inf_{Q:D(Q||G_l|P_s)\leq E_{l|l}} D(Q||G_m|P_s) + o_N(1))\}.$$

According to the definition (4) the reliability $E_{m|l}(\varphi^*)$ of the test sequence φ^* is the limit superieure $\overline{\lim}_{N\to\infty} - N^{-1}\log\alpha_{m|l}^{(N)}(\varphi^*)$. Taking into account (13), (14) and the continuity of the functional $D(Q||G_l|P_s)$ we obtain that $\lim_{N\to\infty} -N^{-1}\log\alpha_{m|l}^{(N)}(\varphi^*)$ exists and in correspondence with (9b) equals to $E_{m|l}^*$. Thus $E_{m|l}(\varphi^*) = E_{m|l}^*$, $m=\overline{1,L}$, $l=\overline{1,l-1}$, $l\neq m$. Similarly we can obtain upper and lower bounds for $\alpha_{m|L}^{(N)}(\varphi^*)$, $m=\overline{1,L}$. Applying the same resonnement we get the reliability $E_{m|L}(\varphi^*) = E_{m|L}^*$. By the definitions (3) and (9d) $E_{L|L}(\varphi^*) = E_{L|L}^*$.

Thus

$$E_{m|l}(\varphi^*) = \lim_{N\to\infty} -N^{-1}\log\alpha_{m|l}^{(N)}(\varphi^*) = E_{m|l}^*, \quad m,l\in\overline{1,L}. \tag{15}$$

The proof of the first part of the theorem will be accomplished if we demonstrate that the sequence of the test φ^* is LAO, that is for given $E_{1|1},...,E_{L-1|L-1}$ and every sequence of tests φ for all $l,m\in\overline{1,L}$, $E_{m|l}(\varphi)\leq E_{m|l}^*$.

Let us consider any other sequence φ^{**} of tests which are defined for every $\mathbf{s}\in\mathcal{S}^N$ by the sets $\mathcal{D}_1^{(N)}(\mathbf{s}),...,\mathcal{D}_L^{(N)}(\mathbf{s})$ such that $E_{m|l}(\varphi^{**})\geq E_{m|l}^*$ for $m,l=\overline{1,L}$. This condition is equivalent to the inequality for N large enough

$$\alpha_{m|l}^{(N)}(\varphi^{**}) \leq \alpha_{m|l}^{(N)}(\varphi^*). \tag{16}$$

Let us examine the sets $D_l^{(N)}(\mathbf{s})\cap\mathcal{B}_l^{(N)}(\mathbf{s})$, $l=\overline{1,L-1}$. This intersection cannot be empty, because in that case

$$\alpha_{l|l}^{(N)}(\varphi^{**}) = \max_{s\in\mathcal{S}^N} G_l^N(\overline{\mathcal{D}}_l^{(N)}(\mathbf{s})|\mathbf{s} \geq \max_{s\in\mathcal{S}^N} G_l^N(\mathcal{B}_l^{(N)}(\mathbf{s})|\mathbf{s}) \geq$$
$$\geq \max_{P_s\in\mathcal{P}^N(S)} \max_{Q:D(Q||G_l|P_s)\leq E_{l|l}} G_l^N(\mathcal{T}_{P,Q}^N(X|S)|\mathbf{s}) \geq \exp\{-N(E_{l|l}+o_N(1))\}.$$

Let us show that $\mathcal{D}_l^{(N)}(\mathbf{s})\cap\mathcal{B}_m^{(N)}(\mathbf{s}) = \emptyset$, $m,l=\overline{1,L-1}$, $m\neq l$. If exists Q such that $D(Q||G_m|P_s)\leq E_{m|m}$ and $\mathcal{T}_{P_s,Q}^N(X|S)\in\mathcal{D}_l^{(N)}(\mathbf{s})$, then

$$\alpha_{m|l}^{(N)}(\varphi^{**}) = \max_{s\in\mathcal{S}^N} G_m^N(\mathcal{D}_l^{(N)}(\mathbf{s})|\mathbf{s}) > \max_{s\in\mathcal{S}^N} G_m^N(\mathcal{T}_{P_s,Q}^N(X|S)|\mathbf{s}) \geq \exp\{-N(E_{m|m}+o_N(1))\}.$$

When $\emptyset \neq \mathcal{D}_l^{(N)}(\mathbf{s}) \cap \mathcal{T}_{P_s,Q}^N(X|S) \neq \mathcal{T}_{P_s,Q}^N(X|S)$, we also obtain that

$$\alpha_{m|l}^{(N)}(\varphi^{**}) = \max_{\mathbf{s}\in S^N} G_l^N(\mathcal{D}_l^{(N)}(\mathbf{s})|\mathbf{s}) > \max_{\mathbf{s}\in S^N} G_l^N(\mathcal{D}_l^{(N)}(\mathbf{s}) \bigcap \mathcal{T}_{P_s,Q}^N(X|S)|\mathbf{s}) \geq \exp\{-N(E_{m|m}+o_N(1))\}.$$

Thus it follows that $E_{m|l}(\varphi^{**}) < E_{m|m}$, which contradicts to (5). Hence we obtain that $\mathcal{D}_l^{(N)}(\mathbf{s}) \cap \mathcal{B}_l^{(N)}(\mathbf{s}) = \mathcal{B}_l^{(N)}(\mathbf{s})$ for $l = \overline{1, L-1}$. The following intersection $D_l^{(N)}(\mathbf{s}) \cap \mathcal{B}_L^{(N)}(\mathbf{s})$ is empty too, because otherwise

$$\alpha_{m|L}^{(N)}(\varphi^{**}) \geq \alpha_{m|L}^{(N)}(\varphi^{*}),$$

which contradicts to (16), that $\mathcal{D}_l^{(N)}(\mathbf{s}) = \mathcal{B}_l^{(N)}(\mathbf{s})$, $l = \overline{1, L}$.

The proof of the second part of the Theorem 1 is simple. If one of the conditions (10) is violated, then from (8) and (9) it follows that at least one of the elements $E_{m|l}$ is equal to 0. For example, let in (10) the condition m is violated. If $E_{m|m} \geq \min\limits_{l=\overline{m+1,L}} \min\limits_{P\in\mathcal{P}(S)} D(G_l||G_m|P)$, then there are $l^* \in \overline{m+1, L}$ such that $E_{m|m} \geq \min\limits_{P\in\mathcal{P}(S)} D(G_l^*||G_m|P)$. From latter and (9b) we obtain that $E_{l^*|m}^* = 0$. From (5) we see that $E_{m|m} \leq \min\limits_{l=\overline{1,m-1}} E_{m|l}^*(E_{l|l})$. Theorem is proved.

It is interesting to examine the following consequences of the Theorem 1, which are generalizations of the Stein's Lemma for the case of multiple hypotheses for arbitrarily varying object and informed statistician. For simplicity we consider the case of three hypotheses.

Theorem 2 (Generalization of Stein's Lemma):
When $\alpha_{l|l}^{(N)}(\varphi) = \varepsilon_l$, $0 < \varepsilon_l < 1$, $l = 1, 2$, *then:*

$$\lim_{N\to\infty} \log \frac{1}{N} \alpha_{m|l}^{(N)}(\alpha_{l|l}^{(N)}(\varphi) = \varepsilon_l) = - \min_{P\in\mathcal{P}(S)} D(G_l||G_m|P), \ \ m \neq l, \ \ m = 1,2,3, \ \ l = 1,2,$$

$$\lim_{N\to\infty} \log \frac{1}{N} \alpha_{m|3}^{(N)}(\alpha_{1|1}^{(N)}(\varphi) = \varepsilon_1, \alpha_{2|2}^{(N)}(\varphi) = \varepsilon_2) = 0, \quad m = 1, 2,$$

$$\lim_{N\to\infty} \log \frac{1}{N} \alpha_{3|3}^{(N)}(\alpha_{1|1}^{(N)}(\varphi) = \varepsilon_1, \alpha_{2|2}^{(N)}(\varphi) = \varepsilon_2) =$$
$$= - \min[\min_{P\in\mathcal{P}(S)} D(G_l||G_m|P), m \neq l, l = 1,2, \ m = \overline{1,3,}].$$

Proof: For $\alpha_{l|l}^{(N)}(\varphi) = \varepsilon_l$, $0 < \varepsilon_l < 1$, $l = 1, 2$, the corresponding $E_{l|l} = 0$. When $E_{l|l} \to 0$ from (15), (8) and (9) we obtain the result of Theorem 2.

2. Application to Source Coding Problem

The tight connection of the hypothesis testing problem and the problem of estimating the optimum probability of incorrect decoding, when messages of length N from a discrete memoryless source are block coded, was emphasized by many authors, see, say, [[7],[9]- [12]]. In source coding problem the reliability function $E(R)$ expresses the dependence of the reliability (exponent of the optimal error probability) from the code rate R. Sometimes it is not less important to present the same dependence by the rate-reliability function $R(E)$, for $0 < E < \infty$. In [12] Fu and Shen examined application of their corresponding result on two hypotheses testing for arbitrarily varying source (AVS) to obtaining of functions $E(R)$ and $R(E)$ for the case when side information is absent.

In this section applying Theorem 1 (for $L = 2$, see also [5]) we obtain expressions of these functions for AVS with side information. Let the AVS is defined by conditional probability distribution $W = \{W(x|s),\ s \in \mathcal{S},\ x \in \mathcal{X}\}$.

Definition 2: *A code (f, g) is a pair of mappings: encoder $f : \mathcal{X}^N \to (1, 2, ..., M(N))$, and decoder $g : (1, 2, ..., M(N)) \times \mathcal{S}^N \to \mathcal{X}^N$.*

For given code (f, g) let us consider the set $\mathcal{A} \subseteq \mathcal{X}^N$ of vectors which are decoded correctly, $\mathcal{A} = \{\mathbf{x} :\ g(f(\mathbf{x}), \mathbf{s}) = \mathbf{x}\}$. Then the error probability of the code will be

$$e(f, g, N) = 1 - \min_{\mathbf{s}\in\mathcal{S}^N} W^N(\mathcal{A}|\mathbf{s}) = \max_{\mathbf{s}\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|\mathbf{s}).$$

The number $|\mathcal{A}|$ of elements of the finite set $\mathcal{A}$ is called the volume of the code and is denoted by $M(N)$.

Definition 3: *A number $R > 0$ is called E-achievable rate for reliability $E > 0$, if for every $\varepsilon > 0$ and sufficiently large N there exists a code (f, g) such that $e(f, g, N) \leq \exp\{-NE\}$, and $N^{-1} \log M(N) \leq R+\varepsilon$. We call the minimal E-achievable rate R the rate-reliability function and denote it by $R(E)$. It can be defined as follows:*

$$R(E) \triangleq \overline{\lim_{N\to\infty}} N^{-1} \log \min_{(f,g):\ e(f,g,N)\leq\exp\{-NE\}} M(N) =$$

$$= \overline{\lim_{N\to\infty}} N^{-1} \log \min_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{\mathbf{s}\in\mathcal{S}} W^N(\overline{\mathcal{A}}|\mathbf{s})\leq\exp\{-NE\}} |\mathcal{A}|.$$

Theorem 3: *In the presence of side information the rate-reliability function of* AVS *with conditional probability distribution W, for any $E > 0$ has the following single-letter presentation:*

$$R(E) = \max_{P\in\mathcal{P}(\mathcal{S})} \max_{Q:\ D(Q\|W|P)\leq E} H_{P,Q}(X|S).$$

Proof: When $L = 2$, we have $E_{1|1} = E_{1|2}$ and from Theorem 1 it follows that

$$E_{2|1}(E_{1|2}) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\ D(Q\|G_1|P)\leq E_{1|2}} D(Q \parallel G_2 \mid P). \tag{17}$$

According to the definition of $E_{2|1}(E_{1|2})$ we have

$$E_{2|1}(E_{1|2}) = \max_{\varphi:\ \alpha_{1|2}(\varphi)\leq\exp\{-NE_{1|2}\}} \overline{\lim_{N\to\infty}} - N^{-1} \log \alpha^N_{2|1}(\varphi) =$$

$$= \sup_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{\mathbf{s}\in\mathcal{S}^N} G^N_1(\overline{\mathcal{A}}|\mathbf{s})\leq\exp\{-NE_{1|2}\}} \overline{\lim_{N\to\infty}} - N^{-1} \log \max_{\mathbf{s}\in\mathcal{S}^N} G^N_2(\mathcal{A}|\mathbf{s}) = \tag{18}$$

$$= \overline{\lim_{N\to\infty}} - N^{-1} \log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{\mathbf{s}\in\mathcal{S}^N} G^N_1(\overline{\mathcal{A}}|\mathbf{s})\leq\exp\{-NE_{1|2}\}} \max_{\mathbf{s}\in\mathcal{S}^N} G^N_2(\mathcal{A}|\mathbf{s}).$$

From (17) and (18) we obtain that

$$E_{2|1}(E_{1|2}) = \overline{\lim_{N\to\infty}} - N^{-1} \log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{\mathbf{s}\in\mathcal{S}^N} G^N_1(\overline{\mathcal{A}}|\mathbf{s})\leq\exp\{-NE_{1|2}\}} \max_{\mathbf{s}\in\mathcal{S}^N} G^N_2(\mathcal{A}|\mathbf{s}) =$$

$$= \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\ D(Q\|G_1|P)\leq E_{1|2}} D(Q \parallel G_2 \mid P). \tag{19}$$

Let us define the probability distribution $P_0 \in \mathcal{P}(\mathcal{X})$ by $P_0(x) = 1/|\mathcal{X}|$, $x \in \mathcal{X}$ and let

$G_1 = W$, $G_2(\cdot|s) = P_0$ for every $s \in \mathcal{S}$. Then

$$\min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:D(Q\|G_1|P)\leq E} D(Q \parallel G_2 \mid P) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:D(Q\|W|P)\leq E} [\log|\mathcal{X}| - H_{P,Q}(X|S)] =$$

$$= \log|\mathcal{X}| - \max_{P\in\mathcal{P}(\mathcal{S})} \max_{Q:D(Q\|W|P)\leq E} H_{P,Q}(X|S), \tag{20}$$

and

$$\inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}^N} G_1^N(\overline{\mathcal{A}}|s)\leq\exp\{-NE\}} \max_{s\in\mathcal{S}^N} G_2^N(\mathcal{A}|s) = \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|s)\leq\exp\{-NE\}} \frac{|\mathcal{A}|}{|\mathcal{X}|^N}. \tag{21}$$

By definition of $R(E)$, from (19), (20) and (21) we obtain

$$R(E) = \overline{\lim_{N\to\infty}} N^{-1}\log \min_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}} W^N(\overline{\mathcal{A}}|s)\leq\exp\{-NE\}} |\mathcal{A}| =$$

$$= \log|\mathcal{X}| - \overline{\lim_{N\to\infty}} - N^{-1}\log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}^N} G_1^N(\overline{\mathcal{A}}|s)\leq\exp\{-NE\}} \max_{s\in\mathcal{S}^N} G_2^N(\mathcal{A}|s) =$$

$$= \max_{P\in\mathcal{P}(\mathcal{S})} \max_{Q:D(Q\|W|P)\leq E} H_{P,Q}(X|S).$$

Theorem is proved.

Corollary: *When $E \geq \min_{P\in\mathcal{P}(\mathcal{S})} D(P_0\|W|P)$, the rate-reliability function $R(E)$ is equal to $\log|\mathcal{X}|$.*

The minimal error probability among the codes with 2^{NR} codewords for AVS with side information is defined as

$$e(W^N, R) \triangleq \min_{\mathcal{A}\subseteq\mathcal{X}^N:\ |\mathcal{A}|\leq 2^{NR}} \max_{s\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|s).$$

Definition 4: *The reliability function $E(R)$ of the AVS for rate R is defined as follows:*

$$E(R) \triangleq \overline{\lim_{N\to\infty}} - N^{-1}\log e(W^N, R) = \overline{\lim_{N\to\infty}} - N^{-1}\log \min_{\mathcal{A}\subseteq\mathcal{X}^N:\ |\mathcal{A}|\leq 2^{NR}} \max_{s\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|s).$$

Theorem 4: *If $\max_{P\in\mathcal{P}(\mathcal{S})} H_{P,W}(X|S) < R \leq \log|\mathcal{X}|$ and conditional probability distribution W is given then the reliability function of AVS may be presented in the following form:*

$$E(R) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\ H_{P,Q}(X|S)\geq R} D(Q\|W|P).$$

When for $R \leq \max_{P\in\mathcal{P}(\mathcal{S})} H_{P,W}(X|S)$, then $E(R) = 0$.

Proof: $E_{2|1}(E_{1|2})$ is defined in (17), it can also be written as follows:

$$E_{2|1}(E_{1|2}) = \overline{\lim_{N\to\infty}} - N^{-1}\log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}^N} G_1^N(\mathcal{A}|s)\leq\exp\{-NE_{1|2}\}} \max_{s\in\mathcal{S}^N} G_2^N(\overline{\mathcal{A}}|s). \tag{22}$$

Let $G_1 = P_0$, $G_2 = W$ and $E_{1|2} = \log|\mathcal{X}| - R$. From this notations and (22) we have

$$E_{2|1}(\log|\mathcal{X}| - R) = \overline{\lim_{N\to\infty}} - N^{-1}\log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ \max_{s\in\mathcal{S}^N} P_0^N(\mathcal{A}|s)\leq\exp\{-N(\log|\mathcal{X}|-R)\}} \max_{s\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|s) =$$

$$= \overline{\lim_{N\to\infty}} - N^{-1}\log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ (|\mathcal{A}|/|\mathcal{X}|^N)\leq\exp\{-N(\log|\mathcal{X}|-R)\}} \max_{\mathbf{s}\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|\mathbf{s}) = \tag{23}$$

$$= \overline{\lim_{N\to\infty}} - N^{-1}\log \inf_{\mathcal{A}\subseteq\mathcal{X}^N:\ |\mathcal{A}|\leq 2^{NR}} \max_{\mathbf{s}\in\mathcal{S}^N} W^N(\overline{\mathcal{A}}|\mathbf{s}) = E(R).$$

Because $G_1 = P_0$ and $G_2 = W$ we see that

$$D(Q||P_0|P) = \sum_{x,s} P(s)Q(x|s)\log\frac{Q(x|s)}{P_0(x)} = \log|\mathcal{X}| - H_{P,Q}(X|S), \tag{24}$$

and

$$\min_{P\in\mathcal{P}(\mathcal{S})} D(W||P_0|P) = \sum_{x,s} P(s)W(x|s)\log\frac{W(x|s)}{P_0(x)} = \log|\mathcal{X}| - \max_{P\in\mathcal{P}(\mathcal{S})} H_{P,W}(X|S). \tag{25}$$

From (24) and our notations we obtain that

$$\min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\, D(Q||G_1|P)\leq E_{1|2}} D(Q||G_2|P) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\, D(Q||P_0|P)\leq\log|\mathcal{X}|-R} D(Q||W|P) =$$

$$= \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\, H_{P,Q}(X|S)\geq R} D(Q||W|P). \tag{26}$$

By the first part of Theorem 1, when $E_{1|2} < \min_{P\in\mathcal{P}(\mathcal{S})} D(G_2||G_1|P)$ then $E_{2|1}(E_{1|2}) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\, D(Q||G_1|P)\leq E_{1|2}} D(Q||G_2|P)$. According this fact, our notations, (23), (25) and (26) we have that when $\max_{P\in\mathcal{P}(\mathcal{S})} H_{P,W}(X|S) < R \leq \log|\mathcal{X}|$, then $E(R) = \min_{P\in\mathcal{P}(\mathcal{S})} \min_{Q:\, H_{P,Q}(X|S)\geq R} D(Q||W|P)$.

By the second part of Theorem 1, if $E_{1|2} \geq \min_{P\in\mathcal{P}(\mathcal{S})} D(G_2||G_1|P)$ then $E_{2|1}(E_{1|2}) = 0$. From here, (23), (25) and (26) it follows that when $R \leq \max_{P\in\mathcal{P}(\mathcal{S})} H_{P,W}(X|S)$ then $E(R) = 0$. Theorem is prove.

3. Illustrative example

Consider the set $\mathcal{X} = \{0,1\}$ and the set of states $\mathcal{S} = \{0,1\}$ each of two elements. Let conditional probability distributions on $\mathcal{X}$ for each value $s \in \mathcal{S}$ be given as follows:

$$G_1 = \begin{pmatrix} 0,70 & 0,30 \\ 0,20 & 0,75 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0,80 & 0,20 \\ 0,45 & 0.55 \end{pmatrix},$$

$$G_3 = \begin{pmatrix} 0,03 & 0,97 \\ 0,33 & 0,67 \end{pmatrix}, \quad G_4 = \begin{pmatrix} 0,40 & 0,60 \\ 0,30 & 0,70 \end{pmatrix}.$$

Results of calculations of $E^*_{m|2}(E_{2|2})$, for $m = 1,3,4$, and $E^*_{m|3}(E_{3|3})$, for $m = 1,2,4$, are presented on respective figers.

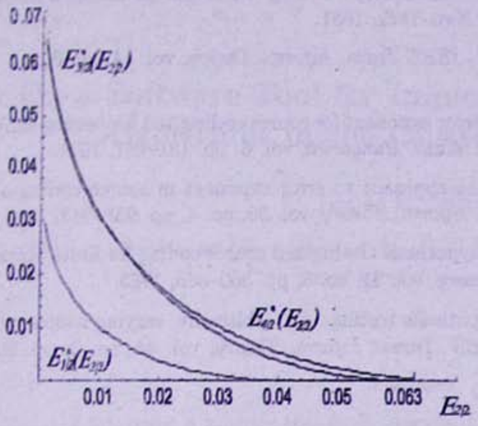

Fig. 1(a)

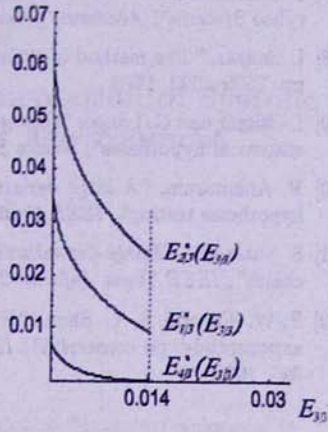

Fig. 1(b)

Here we have $\min[\min_{P\in\mathcal{P}(S)} D(G_3||G_2|P), \min_{P\in\mathcal{P}(S)} D(G_4||G_2|P)] = 0.63$ and $\min_{P\in\mathcal{P}(S)} D(G_4||G_3|P) = 0.014$. In Theorem 1, when one of the inequalities (10) is violated, then at least one element of the matrix $\mathbf{E}^*$ is equal to 0. In the example we see that when in (10) the second inequality is violated, that is $E_{2|2} \geq 0.063$, then three elements of the matrix $\mathbf{E}^*$ are equal to 0 (see Fig. 1(a)). When the third one is violated, i.e. $E_{3|3} \geq 0.014$, then one element is equal to 0 (see Fig. 1(b)).

References

[1] R. F. Ahlswede, "Coloring hypergraphs: a new approach to multi-user source coding" I, II, *J. Combin. Inform. and Syst. Sci.*, vol. 4, no. 1, pp. 76-115, 1979, vol. 5, no. 3, pp. 220–268, 1980.

[2] 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.

[3] 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.

[4] R. L. Dobrushin, Personal communication, 1987.

[5] R. F. Ahlswede, E. A. Haroutunian and E. V. Aloyan, "On logarithmically asymptotically optimal hypothesis testing for arbitrarily varying source with side information", presented for publication 2004.

[6] L. Birgé, "Vitesses maximals de dćroissence des erreurs et tests optimaux associeś". Z. Wahrsch. verw. Gebiete, vol. 55, pp. 261–273, 1981.

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

[8] I. Csiszár, "The method of types", *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2505–2523, 1998.

[9] 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.

[10] V. Anantaram, "A large deriations approach to error exponent in source coding and hypotheses testing", *IEEE Trans. Inform. Theory*, vol. 36, no. 4, pp. 938–943, 1990.

[11] S. Natarajan, "Large derivations hypothesis testing and source coding for finite Markov chains", *IEEE Trans. Inform. Theory*, vol. 31, no. 3, pp. 360–365, 1985.

[12] F.-W. Fu and S.-Y. Shen, "Hypothesis testing for arbitrarily varying source with exponential-type constraint", *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 892–895, 1998.

Տեղեկակացված վիճակագրի կողմից կամայականորեն փոփոխվող օբյեկտի նկատմամբ բազմակի վարկածների ստուգման և աղբյուրի կոդավորման համար կիրառության մասին

Ե. Ա. Հարությունյան և Փ. Մ. Հակոբյան

Ամփոփում

Ուսումնասիրվել է բազմակի վարկածների տեստավորման ընթացքում բոլոր հնարավոր զույգերի սխալների հավանականությունների հուսալիության ցուցիչների փոխկախվածությունների փոփոխվող օբյեկտի համար, որի վիճակները հայտնի են վիճակագրին: Երկու վարկածների դեպքը, երբ որոշում ընդունողին անհայտ է վիճակների հաջորդականությունը, քննարկվել է Ֆուի և Շենի կողմից, իսկ հայտնի վիճակներով տարբերակը դիտարկվել է Ալսվեդեի, Հարությունյանի և Ալոյանի կողմից: Ինչպես Ֆուն և Շենը, մենք նույնպես ստացել ենք կողմնակի ինֆորմացիայով կամայականորեն փոփոխվող աղբյուրի համար արագություն-հուսալիություն և հուսալիություն-արագություն ֆունկցիաները: Ներկայացված է պարզաբանող օրինակ: