CSIT Conference 2021, Yerevan, Armenia, September 27 - October 1

Optimal Neyman-Pearson Procedure of Detection with Side Information for Separate Groups of Arbitrarily Varying Markov Distributions

Evgueni Haroutunian
Institute for Informatics and Automation
Problems of NAS of RA
Yerevan, Armenia
email: eghishe@sci.am

Aram Yesayan
Institute for Informatics and Automation
Problems of NAS of RA,
French University in Armenia
Yerevan, Armenia
email: armfrance@yahoo.fr

***Abstract*—Our aim is to study the asymptotic performance of logarithmically asymptotically optimal (LAO) tests for $M \geq 2$ arbitrarily varying Markov chains grouped in $2 \leq K \leq M$ families using $N+1$ consecutive observations with side information, and also with possibility to reject any decision. The LAO rule for detection of such hypothesis is presented ensuring positiveness of all possible error probability exponents.**

***Keywords*— Markov chain, Markov types, Hypothesis testing, Family detection, Side information, Error exponents, Arbitrarily varying source.**

I. Introduction

A finite Markov chain is frequently a suitable probabilistic model for certain situations. This models are widely used to study behaviour of divers phenomenons in various physical systems, psychology, genetics, epidemiology, social sphere and industry.

This present paper is devoted to the certain problem of hypothesis testing which is the essential component of statistical inference. It is relevant to note that comprehensive, detailed exposition of progress of investigations in hypothesis testing can be found in works of Borovkov [6], Levy [26]. Problems of modern information theory and hypothesis testing are tightly interconnected. This fact is reflected in the books of Csiszár and Körner [8], Blahut [5], Cover and Thomas [7], Csizár and Shields [10]. In [24] multiple hypothesis testing groups of distribution is studied for case of independent identically distributed observations and in present paper we address to the case indicated in the title.

II. On Arbitrarily Varying Markov Chain

Let $\mathcal{X}$ be the alphabet of symbols x of the object and $\mathcal{S}$ be the finite set of object's states s and let $G \stackrel{\triangle}{=} \{G(x|u,s), x,u \in \mathcal{X}, s \in \mathcal{S}\}$ be a Markovian conditional transition probability distribution (PD) with strictly positive elements depending on states $s \in \mathcal{S}$. This states s changes independently each moment with known PD $P = \{P(s), s \in \mathcal{S}\}$. A state vector $\mathbf{s} \stackrel{\triangle}{=} (s_0, s_1, ..., s_N) \in \mathcal{S}^{N+1}$ is the realization of $N+1$ members of independently identically distributed (i.i.d.) sequence of RVs
$S_0, S_1, ..., S_n, ...$ with PD P. Object characterized by G, produces a sequence of random variables (RVs)
$X_0, X_1, ..., X_n, ...$ which form arbitrarily varying stationary Markov chain. If at an n-th moment the chain is in state S_n then

$$G(X_n = x | X_{n-1} = u, S_n = s) = G(x|u,s),\ x,u \in \mathcal{X},$$

$s \in \mathcal{S}$, $n = 1, 2, ...$
To conditional probability of Markov chain corresponds unique stationary PD

$$\tilde{G} \stackrel{\triangle}{=} \{\tilde{G}(X_0 = u|s) = \tilde{G}(u|s),\, u \in \mathcal{X},\, s \in \mathcal{S}\},$$

We consider the joint conditional PD on $\mathcal{X} \times \mathcal{X}$ for $s \in \mathcal{S}$

$$\tilde{G} \times G \stackrel{\triangle}{=} \{\tilde{G}(u|s)G(x|u,s),\, u,x \in \mathcal{X}, s \in \mathcal{S}\},$$

joint PD on $\mathcal{X} \times \mathcal{X} \times \mathcal{S}$

$$P \times \tilde{G} \times G \stackrel{\triangle}{=} \{P(s)\tilde{G}(u|s)G(x|u,s),\, u,x \in \mathcal{X}, s \in \mathcal{S}\},$$

marginal PD PQ on $\mathcal{X}$

$$P\tilde{G} \stackrel{\triangle}{=} \{P\tilde{G}(u) = \sum_{s\in\mathcal{S}} P(s)\tilde{G}(u|s),\, u \in \mathcal{X}\},$$

and conditional PD PG on $\mathcal{X}$ for $u \in \mathcal{X}$

$$PG \stackrel{\triangle}{=} \{PG(x|u) = \sum_{s\in\mathcal{S}} P(s)G(x|u,s),\, x,u \in \mathcal{X}\},$$

The conditional PD of the vector $\mathbf{x} = (x_0, x_1, ..., x_N) \in \mathcal{X}^{N+1}$ of the arbitrarily varying Markov chain with transition PD G and stationary PD $\tilde{G}$ with respect to corresponding states vector $\mathbf{s} = (s_0, s_1, ..., s_N) \in \mathcal{S}^{N+1}$ is defined as the following product:

$$\tilde{G} \times G^N(\mathbf{x}|\mathbf{s}) \stackrel{\triangle}{=} \tilde{G}(x_0|s_0) \prod_{n=1}^{N} G(x_n|x_{n-1}, s_n). \quad (1)$$

The conditional PD of a subset $\mathcal{A}^{N+1} \subset \mathcal{X}^{N+1}$ is

$$\tilde{G} \times G^N(\mathcal{A}^{N+1}|\mathbf{s}) \stackrel{\triangle}{=} \sum_{\mathbf{x}\in\mathcal{A}^{N+1}} \tilde{G} \times G^N(\mathbf{x}, \mathbf{s}),\ \mathbf{s} \in \mathcal{S}^{N+1}. \quad (2)$$

III. Problem Statement

We are given that there are M different Markov PDs G_m and corresponding stationary PDs $\tilde{G}_m$ as possible characteristics of the object:

$$\tilde{G}_m \times G_m = \{\tilde{G}_m(u|s)G_m(x|u,s), x, u \in \mathcal{X}, s \in \mathcal{S}\},$$

$m = \overline{1, M}$.

These PDs are arranged into K, $2 \leq K \leq M$, different groups, or families, $\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_K$, which are considered as K hypotheses H_k concerning the PD of the object. We consider also an empty "group" $\mathcal{B}_{K+1}$, these groups are disjoint, contains $|\mathcal{B}_1|, |\mathcal{B}_2|, ..., |\mathcal{B}_{K+1}|$ PDs such that

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

We study procedure of hypothesis testing (referred also as detection, or simply as test) which is the action of the statistician using observed vector $\mathbf{x}$ and states vector $\mathbf{s}$, either to accept decision concerning group $\mathcal{B}_k$ in which is the PD of the object, or accepting group $\mathcal{B}_{K+1}$ reject to do any judgement. Such procedure is universal test, we denote it by Φ_N, $N = 1, 2, ...$ When $K = M$, hypotheses are simple, hypotheses with $|\mathcal{B}_k| > 1$ are composite.

The test is the partition of the space $\mathcal{X}^{N+1}$ into $K + 1$ subsets $\mathcal{A}_{k,s}^{N+1}$, $k = \overline{1, K+1}$, $\mathbf{s} \in \mathcal{S}^{N+1}$, with dependence on known vector of states $\mathbf{s}$. With given $\mathbf{x}$ and $\mathbf{s}$ statistician can make corresponding decision.

We denote by Φ the infinite sequence of tests Φ_N. Taking decisions about the hypotheses statistician can commit the following errors. Let the hypothesis H_l is adopted but the correct is H_k, $l \neq k$, the probability of such error is defined (see [6]) as follows

$$\alpha_{l|k}^{N+1} = \alpha_{l|k}^{N+1}(\Phi_N) \stackrel{\triangle}{=}$$

$$\max_{\mathbf{s}\in\mathcal{S}^{N+1}} \max_{\tilde{G}_m \times G_m \in \mathcal{B}_k} \tilde{G}_m \times G_m^N(\mathcal{A}_l^{N+1}|\mathbf{s}) \quad (3)$$

The probability $\alpha_{k|k}^{N+1}$ to accept a hypothesis different from true hypothesis H_k we define in the following way:

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

$$\sum_{l\neq k,\, l=\overline{1,K+1}} \max_{\mathbf{s}\in\mathcal{S}^{N+1}} \max_{\tilde{G}_m \times G_m \in \mathcal{B}_k} \tilde{G}_m \times G_m^N(\mathcal{A}_l^{N+1}|\mathbf{s}) \quad (4)$$

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

$$E_{l|k} = E_{l|k}(\Phi) \stackrel{\triangle}{=} \liminf_{N\to\infty} \left(-\frac{1}{N+1} \log \alpha_{l|k}^{N+1}(\Phi_N)\right),$$

$$k = \overline{1, K},\ l = \overline{1, K+1}. \quad (5)$$

All reliabilities are arranged in $(K + 1) \times K$ matrix. For instance at $K = 4$ the matrix of reliabilities is the following:

$$\mathbf{E}(\Phi) = \begin{pmatrix} E_{1|1} & E_{2|1} & E_{3|1} & E_{4|1} & E_{5|1} \\ E_{1|2} & E_{2|2} & E_{3|2} & E_{4|2} & E_{5|2} \\ E_{1|3} & E_{2|3} & E_{3|3} & E_{4|3} & E_{5|3} \\ E_{1|4} & E_{2|4} & E_{3|4} & E_{4|4} & E_{5|4} \end{pmatrix}.$$

Definitions (6) and (7) imply that

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

Notes: 1) 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 [4]. 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) [25], [11].

2) Our approach differs from approaches in [6], where only $\alpha_{k|k}$ are considered.

3) In opposition to the criterion adopted by Gutman [11], we recognize the asymmetry in the importance of different hypotheses and consider unequal requirements to error probabilities, and reliabilities, of their detection.

IV. Information-theoretical Technical Tools

We use method of types and traditional notions and notation developed in information theory [7], [10], [23].

We specify the following definitions and properties of types for Markov chains, [4,23,34].

It is convenient don't take in consideration first element of the vector $\mathbf{s}$, so we consider $\mathbf{s} = (s_1, ..., s_N)$. Let $N(s|\mathbf{s})$ be the number of occurrences of a state $s \in \mathcal{S}$ in the N-vector $\mathbf{s}$. The type of the vector $\mathbf{s} \in \mathcal{S}^N$ is the PD

$$P_{\mathbf{s}} = \{P_{\mathbf{s}}^N(s) = \frac{1}{N} N(s|\mathbf{s}),\ s \in \mathcal{S}\}$$

For each $N = 1, 2, ...$, the second order type of vector $\mathbf{x} = (x_0, x_1, ..., x_N)$ from $\mathcal{X}^{N+1}$ is the PD $Q_{\mathbf{x}}^N(u, x)$ defined by the square matrix of $|\mathcal{X}|^2$ relative frequencies $N^{-1}N(u, x|\mathbf{x})$ of simultaneous appearance of pair (u, x) in neighbor positions in $\mathbf{x}$

$$Q_{\mathbf{x}}^N(u, x) \stackrel{\triangle}{=} \frac{1}{N} |(n : x_n = u, x_{n+1} = x,\ u, x \in \mathcal{X}^{N+1})|.$$

The second order joint type of the pair of vectors $\mathbf{x}$ and $\mathbf{s}$ is the PD

$$Q_{\mathbf{x},\mathbf{s}}^N(u, x, s) \stackrel{\triangle}{=} P_{\mathbf{s}}(s)\tilde{Q}_{\mathbf{x},\mathbf{s}}^N(u|s)Q_{\mathbf{x},\mathbf{s}}^N(x|u, s) =$$

$$\frac{1}{N} N(u, x, s|\mathbf{x}, \mathbf{s}),\ x, u \in \mathcal{X}, s \in \mathcal{S},$$

where $N(u, x, s|\mathbf{x}, \mathbf{s})$ is the number of n such that $(x_{n-1}, x_n) = (u, x)$ and $s_n = s$, $n = \overline{1, N}$, that is letters u and x appear in the vector $\mathbf{x}$ being neighbour and state s in vector $\mathbf{s}$ corresponds to x.

We use as conditional second order type of vector $\mathbf{x}$ with respect to vector $\mathbf{s}$ the conditional PD $\tilde{Q}_{\mathbf{x},\mathbf{s}}^N(u|s)$ defined by relation

$$\tilde{Q}_{\mathbf{x},\mathbf{s}}^N(u|s) \stackrel{\triangle}{=} N(u, s|\mathbf{x}, \mathbf{s})/N(s|\mathbf{s}),\ u \in \mathcal{X},\ s \in \mathcal{S},$$

where $N(u, s|\mathbf{x}, \mathbf{s})$ is the number of repetitions of u in $\mathbf{x}$ and s in $\mathbf{s}$ such that position of u precedes position of s.

There are also the conditional type of vector $\mathbf{x}$ with respect to vector $\mathbf{s}$ as conditional PD $Q_{\mathbf{x},\mathbf{s}}^N(x|u, s)$ defined by

$$Q_{\mathbf{x},\mathbf{s}}^N(x|u, s) \stackrel{\triangle}{=} N(u, s, x|\mathbf{x}, \mathbf{s})/N(u, s|\mathbf{x}, \mathbf{s}),\ u, x \in \mathcal{X},\ s \in \mathcal{S}.$$

Denote by $\tilde{Q} \times Q$ the following PD from $\mathcal{Q}(\mathcal{X}|\mathcal{S})$

$$\tilde{Q} \times Q = \{\tilde{Q}(u|s)Q(x|u, s), x, u \in \mathcal{X}, s \in \mathcal{S}\}.$$

We will use for $\mathbf{x} \in \mathcal{X}^{N+1}$, $\mathbf{s} \in \mathcal{S}^N$

$$\tilde{Q}_{\mathbf{x},\mathbf{s}} \times Q_{\mathbf{x},\mathbf{s}} = \{\tilde{Q}_{\mathbf{x},\mathbf{s}}(u|s)Q_{\mathbf{x},\mathbf{s}}(x|u, s) =$$

$$\frac{N(u, s, x|\mathbf{x}, \mathbf{s})}{N(s|\mathbf{s})},\ u, x \in \mathcal{X}, s \in \mathcal{S}\}$$

for the second order conditional type of $\mathbf{x}$ given $\mathbf{s}$. Let $\mathcal{Q}^N(X|\mathbf{s})$ be the set of all second order conditional types given $\mathbf{s}$. We denote by $\mathcal{T}_{\tilde{Q},Q}^N(X|\mathbf{s})$ the set of vectors $\mathbf{x}$ from $\mathcal{X}^{N+1}$ which have the second order conditional type given $\mathbf{s}$, such that $\tilde{Q}_{\mathbf{x},\mathbf{s}} \times Q_{\mathbf{x},\mathbf{s}} = \tilde{Q} \times Q$.

Let us define the conditional entropy of X with respect to $\mathbf{s}$

$$H_{\tilde{Q}\times Q|P_{\mathbf{s}}}(X|\mathbf{s}) =$$

$$-\sum_{s,u,x} P_{\mathbf{s}}(s)\tilde{Q}(u|s)Q(x|u, s) \log \tilde{Q}(u|s)Q(x|u, s).$$

We define also the conditional divergence $D(\tilde{Q} \times Q|\tilde{G}_m \times G_m|P)$ of the PD $P \times \tilde{Q} \times Q$ from the PD $P \times \tilde{G}_m \times G_m$ as follows

$$D(\tilde{Q} \times Q|\tilde{G}_m \times G_m|P) =$$

$$-\sum_{s,u,x} P(s)\tilde{Q}(u|s)Q(x|u, s) \log \frac{\tilde{Q}(u|s)Q(x|u, s)}{\tilde{G}_m(u|s)G_m(x|u, s)},$$

$m = \overline{1, M}$.

Similarly $D(\tilde{G}_m \times G_m|\tilde{G}_l \times G_l|P)$, $m, l = \overline{1, M}$, are defined.

We will denote for brevity:

$$\text{for } \tilde{Q} \times Q \in \mathcal{Q}(\mathcal{X}|\mathcal{S}),\ D(\tilde{Q} \times Q||\mathcal{B}_k|P) \stackrel{\triangle}{=}$$

$$\min_{\tilde{G}_m \times G_m \in \mathcal{B}_k} D(\tilde{Q} \times Q|\tilde{G}_m \times G_m|P),$$

$$\text{for } \tilde{Q} \times Q \in \mathcal{Q}^N(\mathcal{X}|\mathbf{s}),\ D^N(\tilde{Q} \times Q||\mathcal{B}_k|P_{\mathbf{s}}) \stackrel{\triangle}{=}$$

$$\min_{\tilde{G}_m \times G_m \in \mathcal{B}_k} D^N(\tilde{Q} \times Q|\tilde{G}_m \times G_m|P_{\mathbf{s}}),$$

$$\text{for } \mathcal{R}_l \subset \mathcal{Q}(\mathcal{X}|\mathcal{S}),\ D(\mathcal{R}_l||\mathcal{B}_k|P) \stackrel{\triangle}{=}$$

$$\min_{\tilde{Q} \times Q \in \mathcal{R}_l^N} D(\tilde{Q} \times Q||\mathcal{B}_k|P),$$

$$\text{for } \mathcal{R}_l^N \subset \mathcal{Q}^N(\mathcal{X}|\mathbf{s}),\ D^N(\mathcal{R}_l^N||\mathcal{B}_k|P_{\mathbf{s}}) \stackrel{\triangle}{=}$$

$$\min_{\tilde{Q} \times Q \in \mathcal{R}_l^N} D^N(\tilde{Q} \times Q||\mathcal{B}_k|P_{\mathbf{s}}).$$

It is elementary to verify that

$$|\mathcal{Q}^N(X|\mathbf{s})| \le (N + 1)^{|\mathcal{X}|^2}.$$

For every second order conditional type $\tilde{Q} \times Q$ of vectors $\mathbf{x} \in \mathcal{X}^{N+1}$ given state vector $\mathbf{s} \in \mathcal{S}$ the upper estimates holds

$$|\mathcal{T}_{\tilde{Q}\times Q}^N(X|\mathbf{s})| \le \exp\{NH_{\tilde{Q}\times Q,P_{\mathbf{s}}}(X|\mathbf{s}) + o(1)\}.$$

V. Formulation of Result

We consider the Markov chain with alphabet $\mathcal{X}$ of symbols x, and finite set $\mathcal{S}$ of states $\mathbf{s}$. The chain is characterized by one of M transition strictly positive PDs.

$$G_m \stackrel{\triangle}{=} \{G_m(x|u,s), x, u \in \mathcal{X}, s \in \mathcal{S}\},\ m = \overline{1,M},$$

such that at n-th moment

$$G_m(X_n = x|X_{n-1} = u, s = s_n) =$$

$$\{G_m(x|u,s), x, u \in \mathcal{X}, s \in \mathcal{S}, n = 1, 2, ..\}$$

and by unique stationary PD $\tilde{Q}_m$

$$\tilde{Q}_m \stackrel{\triangle}{=} \{\tilde{Q}_m(X_0 = u|s) = \tilde{Q}_m(u|s), u \in \mathcal{X}, s \in \mathcal{S}\}.$$

The conditional PD of a vector $\mathbf{x} = (x_0, x_1, ..., x_N) \in \mathcal{X}^{N+1}$ with respect to $\mathbf{s} = (s_0, s_1, ..., s_N)$ verify the corresponding equation (1) and for $\mathcal{A}^{N+1} \subset \mathcal{X}^{N+1}$ the corresponding sum in (2). As it was noted above, the PDs $G_1, ..., G_M$ are united into K, $2 \leq K \leq M$ families $\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_K$ which are hypotheses to be tested by statistician using observed vector of chain $\mathbf{x}$ and knowing states vector $\mathbf{s}$, which arises i.i.d arbitrary with PD $P = \{P(s),\ s \in \mathcal{S}\}$. There are also the empty "family" $\mathcal{B}_{K+1}$, corresponding to the case when any judgement is rejected. The set Φ_N^* will be constructed by generalization of the method proposed in [112,13] and applied for many different more general situation in [12, 14-24].

For constructing of the desirable LAO test ensuring preliminary given strictly positive reliabilities $E_{1|1}, E_{2|2}, ..., E_{K|K}$ we define the following subsets of PDs:

$$\mathcal{R}_k^N \stackrel{\triangle}{=} \{\tilde{Q} \times Q \in \mathcal{Q}^N(\mathcal{X}|\mathcal{S}) : D(\tilde{Q} \times Q||\mathcal{B}_k|P) \leq E_{k|k}\},$$

$k = \overline{1,K}$,

$\mathcal{R}_{K+1}^N \stackrel{\triangle}{=} \{\tilde{Q} \times Q \in \mathcal{Q}^N(\mathcal{X}|\mathcal{S}) : D(\tilde{Q} \times Q||\mathcal{B}_k|P) > E_{k|k},\ k = \overline{1,K}\}$.

Obviously, $\mathcal{R}_k^N \subset \mathcal{R}_k$, $k = \overline{1,K+1}$. We introduce the following values of elements of the reliabilites matrix:

$$E_{k|k}^* \stackrel{\triangle}{=} E_{k|k}^*(E_{k|k}) = E_{k|k},\ k = \overline{1,K}, \tag{6}$$

$$E_{l|k}^* = E_{l|k}^*(E_{l|l}) \stackrel{\triangle}{=} D(\mathcal{R}_l||\mathcal{B}_k|P),$$

$$k = \overline{1,K}, k \neq l,\ l = \overline{1,k}, \tag{7}$$

$$E_{K+1|k}^* = E_{K+1|k}^*(E_{1|1}, E_{2|2}, ..., E_{K|K}) \stackrel{\triangle}{=}$$

$$D(\mathcal{R}_{K+1}||\mathcal{B}_k|P),\ k = \overline{1,K}. \tag{8}$$

Theorem : *If all PDs $G_m,\ m = \overline{1,M}$ are different and the strictly positive numbers $E_{1|1}, E_{2|2}, ..., E_{K|K}$ are such that the following inequalities hold*

$$E_{1|1}^* < \min_{l=\overline{1,K}} D(\mathcal{R}_l||\mathcal{B}_1|P) \tag{9}$$

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

$$k = \overline{2,K-1} \tag{10}$$

$$E_{K|K}^* < \min_{l=\overline{1,K-1}} E_{l|K}^*(E_{l|l}), \tag{11}$$

then there exists an LAO sequence of tests all elements of the reliability matrix $E^* = (E_{l|k}^*)$ of which are defined in (6)-(8) and are strictly positive.

When at least one of inequalities (9)-(11) is violated, then at least one element of the reliabilities matrix E^* is equal to zero. Moreover, if we try to detect with $E_{l|l}$ which for $l \in [1; K+1]$ and $k \in [1; K]$ is greater than $D(\mathcal{R}_l||\mathcal{B}_k|P)$, then the tests for all $N = 1, 2, ...$ will make an error with the probability 1.

Proof: The theorem is a generalization of results from [20, 21, 24]. Accordingly our proof must be pertinent extension of the corresponding proofs applying tools from Section 4. Therefore we present here only the concise schematic exposition of that.

The proof of the positive statement of the theorem (of the existence of the desirable sequence of tests) consists in construction of the necessary test sequence Φ^* for each $N = 1, 2, ...$, each $\mathbf{s} \in \mathcal{S}^N$ as a partition of the space $\mathcal{X}^{N+1}$ on $K+1$ subsets $\mathcal{A}_{\mathbf{x},\mathbf{s}}^{N*}$ as follows

$$\mathcal{A}_{k,\mathbf{s}}^{N*} = \bigcup_{\tilde{Q} \times Q \in \mathcal{R}_k^N} \mathcal{T}_{\tilde{Q} \times Q}^{N+1}(\mathcal{X}|\mathbf{s}),\ k = \overline{1,K},$$

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

Note that $\mathcal{A}_{k,\mathbf{s}}^{N*} \subset \mathcal{X}^{N+1}$ for all $k = \overline{1,K+1}$, because $\tilde{Q} \times Q \notin \mathcal{Q}^{N+1}(X, \mathbf{s})$ we have $\mathcal{T}_{\tilde{Q} \times Q}^{N+1}(X|\mathbf{s}) = 0$. These $\mathcal{A}_{k,\mathbf{s}}^{N*}$ are the sets for acceptance of corresponding groups $\mathcal{B}_k$, $k = \overline{1,K+1}$, by obtained experimental pairs of vectors $\mathbf{s}$ and $\mathbf{x}$. Using in detail specified tools from Section 4 we show that sets $\mathcal{A}$ do not intersect and then prove that such test ensures that corresponding reliabilities presented in (6)-(8) are realized, and, finally, that this test is asymptotically optimal.

Verification of negative assertion of the Theorem is simple, when one of conditions (9)-(11) is violated then the corresponding error probability is equal to 1 and the reliability is equal to zero.

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] P. Bilingsly, ‘Statistical methods in Markov chains”, *Ann. math. statist.*, vol. 12, pp. 12-40, 1961.

[4] L. Birgé, “Vitesses maximals de décroissence des erreurs et tests optimaux associés”. Z. Wahrsch. Verw. Gebiete, vol. 55, pp. 261–273, 1981.

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

[9] I. Csiszár, “The method of types”, *IEEE Trans. Inform. Theory*, vol. 44, no.5, pp. 2505-2523, 1998.

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

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

[12] E. A. Haroutunian, “On asymptotically optimal test for Markov chains”, *The first World Congress of Bernoully Society*, Tashkent 1986, Vol. 1, p. 38, 1986.

[13] 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.

[14] 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.

[15] 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.

[16] E. A. Haroutunian and N. M. Grigoryan, “On arbitrarilty varying Markov source coding and hypothesis LAO testing by non-informed statistician”, *Proc. ISIT 2009, Seoul Korea*, pp. 981-985, 2009.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] E. A. Haroutunian P. M. Hakobyan and A. O. Yessayan, ‘On LAO testing of multiple hypotheses concerning Markov chain ”, *Mathematical Problems of Computer Sciences*, vol. 41, pp. 63-73, 2014.

[22] 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.

[23] 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.

[24] E. A. Haroutunian and A. O. Yessayan, ‘A Neyman-Pearson proper ways to universal testing of multiple hypotheses formed by groups of distributions ”, *Mathematical Problems of Computer Sciences*, vol. 54, pp. 18-33, 2020.

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

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

[27] S. Natarajan, “Large deviations, hypotheses testing, and source coding for finite Markov chains”, *IEEE Trans. Inform. Theory*, vol. 36, no.4, pp. 938-943, 1990. *Mathematical Problems of Computer Sciences*, vol. 29, pp. 89-96, 2007.