Serdica

Mathematical Journal

Сердика

Математическо списание

For further information on
Serdica Mathematical Journal
which is the new series of
Serdica Bulgaricae Mathematicae Publicationes
visit the website of the journal http://www.math.bas.bg/~serdica
or contact: Editorial Office
Serdica Mathematical Journal
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: serdica@math.bas.bg

Serdica Math. J. **37** (2011), 79–86

Serdica
Mathematical Journal

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

ON SPECIAL CASE OF MULTIPLE HYPOTHESES OPTIMAL TESTING FOR THREE DIFFERENTLY DISTRIBUTED RANDOM VARIABLES

Leader Navaei

Communicated by N. M. Yanev

Abstract. In this paper by using theory of large deviation techniques (LDT), the problem of hypotheses testing for three random variables having different distributions from three possible distributions is solved. Hypotheses identification for two objects having different distributions from two given probability distributions was examined by Ahlswewde and Haroutunian. We noticed Sanov's theorem and its applications in hypotheses testing.

1. Introduction. Haroutunian and Yessayan in [4] is solved the problem for the case of two objects having different distributions from three possible distributions without usage of LDT. In the present paper we introduce the proof of this theorem by using Sanov's theorem for the case of three random variables

2010 *Mathematics Subject Classification:* 62P30.
Key words: Optimal testing, error probability, three differently distributed, type.

having different distributions from three possible distributions.In the next Section we will express notations, basic concepts and theorem of Sanov and also in Section 3 we present the result and its proof.

2. Preliminaries. Assume that $\mathcal{X}$ is a finite set of the size $|\mathcal{X}|$. The set of all probability distributions (PDs) on $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$. For PD's, P and Q, $H(P)$ denotes entropy and $D(P\|Q)$ denotes information divergence (or the Kullback-Leibler distance)

$$H(P) \triangleq -\sum_{x\in\mathcal{X}} P(x)\log P(x), \quad D(P\|Q) \triangleq \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}.$$

In this paper we use exp-s and log-s at base 2. We also consider the standard conventions that $0\log 0 = 0$, $0\log\frac{0}{0} = 0$, $P\log\frac{P}{0} = \infty$ if $P > 0$. The type of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{X}^N$ is the empirical distribution given by $Q(x) \triangleq N^{-1}\cdot N(x|\mathbf{x})$ for all $x \in \mathcal{X}$, where $N(x|\mathbf{x})$ denotes the number of occurrences of x in $\mathbf{x}$. The subset of $\mathcal{P}(\mathcal{X})$ consisting of the possible types of sequences $\mathbf{x} \in \mathcal{X}^N$ is denoted by $\mathcal{P}_N(\mathcal{X})$. For $Q \in \mathcal{P}_N(\mathcal{X})$ the set of sequences of type class Q will be denoted by $\mathcal{T}_Q^N(X)$. The probability that N independent drawings from $P \in \mathcal{P}(\mathcal{X})$ give $\mathbf{x} \in \mathcal{X}^N$, is denoted by $P^N(\mathbf{x})$. If $\mathbf{x} \in \mathcal{T}_Q^N(X)$, then:

$$P^N(\mathbf{x}) \triangleq \prod_{x\in\mathcal{X}} P(x)^{NQ(x)} = \exp\left\{-N\left[H(Q) + D(Q\|P)\right]\right\}.$$

Lemma ([2, 3])**.**

a) The number of types of lengthN for sequences grows at most polynomially with N:

$$|\mathcal{P}_N(\mathcal{X})| < (N+1)^{|\mathcal{X}|}.$$

b) For any type $Q \in \mathcal{P}_N(\mathcal{X})$ we have:

$$(N+1)^{-|\mathcal{X}|}\exp\{NH(Q)\} \leqslant |\mathcal{T}_Q^N(X)| \leqslant \exp\{NH(Q)\}.$$

c) For any PD $P \in \mathcal{P}(\mathcal{X})$ we have:

$$\frac{P^N(\mathbf{x})}{Q^N(\mathbf{x})} = \exp\{-ND(Q\|P)\}, \quad \textit{if} \quad \mathbf{x} \in \mathcal{T}_Q^N(X),$$

and

$$(N+1)^{-|\mathcal{X}|}\exp\{-ND(Q\|P)\} \le P^N(\mathcal{T}_Q^N(X)) \le \exp\{-ND(Q\|P)\}.$$

Theorem 1 (Sanov's theorem [2, 3]). *Let $\mathcal{A}$ be a set of distributions from $\mathcal{P}$ such that its closure is equal to the closure of its interior, then for the empirical distribution $Q_{\mathbf{x}}$ of a vector $\mathbf{x}$ from a strictly positive distribution P on $\mathcal{X}$:*

$$\lim_{N\to\infty}\left(-\frac{1}{N}\log P^N(\mathbf{x}:Q_{\mathbf{x}}\in\mathcal{A})\right)=\inf_{Q_{\mathbf{x}}\in\mathcal{A}} D(Q_{\mathbf{x}}\|P).$$

3. Problem statement and formulation of results. Let Y_1,Y_2 and Y_3 be random variables (RV) taking values in the same finite set $\mathcal{Y}$ with one of L = 3 PDs. It is obvious that the RV's Y_1,Y_2 and Y_3 can have only different distributions from three given probability distributions PD's $P_i; i=\overline{1,3}$ from $\mathcal{P}(\mathcal{Y})$. Let$(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=((y_1^1,y_1^2,y_1^3),\ldots,(y_n^1,y_n^2,y_n^3),\ldots,(y_N^1,y_N^2,y_N^3)), y^i\in\mathcal{Y},\quad i=1,2,3,\quad n=\overline{1,N}$, be a sequence of results of N independent observations of the vector (Y_1,Y_2,Y_3). The goal of the statistician is to define which thriple of distributions corresponds to observed sample $(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})$. The test is a procedure of making decision on the base of $(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})$, which we denote by φ_N.

For this model the vector $(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})$ can have one of 6 joint probability distributions

$$P'_{l_1,l_2,l_3}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3}), l_1\neq l_2\neq l_3,\quad l_1,l_2,l_3=\overline{1,3},$$

where

$$P'_{l_1,l_2,l_3}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P'_{l_1}(\mathbf{y_1})P'_{l_2}(\mathbf{y_2})P'_{l_3}(\mathbf{y_3}),$$

We can take $(Y_1,Y_2,Y_3)=X,\mathcal{Y}\times\mathcal{Y}\times\mathcal{Y}=\mathcal{X}$ and $\mathbf{x}=(x_1,x_2,\ldots,x_N), x_n\in\mathcal{X},\mathbf{x}\in\mathcal{X}^N$, where $x_n=(y_n^1,y_n^2,y^3); n=\overline{1,N}$, then we will have 6 new hypotheses for one object.

$$P'_{1,2,3}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_1(\mathbf{x}),\quad P'_{1,3,2}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_2(\mathbf{x}),\quad P'_{2,1,3}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_3(\mathbf{x}),$$

$$P'_{2,3,1}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_4(\mathbf{x}),\quad P'_{3,1,2}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_5(\mathbf{x}),\quad P'_{3,2,1}(\mathbf{y_1},\mathbf{y_2},\mathbf{y_3})=P_6(\mathbf{x}),$$

By means of non-randomized test $\varphi_N(\mathbf{x})$ on the basis of a sample $\mathbf{x}$ of length N we must accept one of the hypotheses. For this aim we can divide the sample space $\mathcal{X}^N$ on 6 disjoint subsets

$$\mathcal{A}_m^N\triangleq\{\mathbf{x}:\varphi_N(\mathbf{x})=m\},\qquad m=\overline{1,6}.$$

The probability of the erroneous acceptance of hypotheses H_l provided that hypotheses H_m is true, for $m \neq l$ is denoted

$$\alpha^N_{m|l}(\varphi_N) \triangleq P^N_m(\mathcal{A}^N_l) = \sum_{\mathbf{x}\in\mathcal{A}^N_l} P^N_m(\mathbf{x}).$$

For $m = l$ we denote by $\alpha_{m|m}(\varphi_N)$ the probability to reject H_m when it is true:

$$\alpha^N_{m|m}(\varphi_N) \triangleq \sum_{l\neq m} \alpha^{(N)}_{m|l}(\varphi_N). \tag{1}$$

The matrix $\mathcal{A}(\varphi_N) \triangleq \left\{\alpha^N_{m|l}(\varphi_N)\right\}$ is called power of the test. We consider the rates of exponential decrees of the error probabilities and call them reliabilities

$$E_{m|l}(\varphi) \triangleq \overline{\lim_{N\to\infty}} -\frac{1}{N}\log\alpha^{(N)}_{m|l}(\varphi_N). \tag{2}$$

The matrix $E(\varphi) = \left\{E_{m|l}(\varphi)\right\}$ is called the reliability matrix of the tests sequences φ

$$E(\varphi) = \begin{bmatrix} E_{1|1} & \dots & E_{1|\ell} & \dots & E_{1|6} \\ \vdots & & \vdots & & \vdots \\ E_{m|1} & \dots & E_{m|\ell} & \dots & E_{m|6} \\ \vdots & & \vdots & & \vdots \\ E_{6|1} & \dots & E_{6|\ell} & \dots & E_{6|6} \end{bmatrix}.$$

The problem is to find the matrix $E(\varphi)$ with largest elements, which can be achieved by tests when a part of elements of the matrix $E(\varphi)$ is fixed. According to (1) and (2) we can derive that

$$E_{m|m} = \min_{l\neq m} E_{m|l}. \tag{3}$$

Definition. *The test sequence $\varphi^* = (\varphi_1, \varphi_2, \ldots)$ is called LAO if for given values of the elements $E_{1|1}, \ldots, E_{5|5}$ it provides maximal values for all other elements of $E(\varphi^*)$.*

Our aim is to define conditions on $E_{1|1}, \ldots, E_{5|5}$ under which there exists LAO sequence of tests φ^* and show how other elements $E_{m|l}(\varphi^*)$ of the matrix $E(\varphi^*)$ can be found from them.

Consider for a given positive and finite $E_{1|1}, \ldots, E_{5|5}$ the following family of regions:

$$\mathcal{R}_l \triangleq \left\{Q : D(Q\|P_l) \leqslant E_{l|l}\right\}, \quad l = \overline{1,5}, \tag{4}$$

$$\mathcal{R}_6 \triangleq \left\{Q : D(Q\|P_l) > E_{l|l}, \quad l = \overline{1,5}\right\}, \tag{5}$$

$$\mathcal{R}_l^N \triangleq \mathcal{R}_l \cap \mathcal{P}_N(\mathcal{X}), \qquad l = \overline{1,5}. \tag{6}$$

and the following numbers:

$$E^*_{l|l} = E^*_{l|l}(E_{l|l}) \triangleq E_{l|l}, \quad l = \overline{1,5}, \tag{7}$$

$$E^*_{m|l} = E^*_{m|l}(E_{l|l}) \triangleq \inf_{Q\in\mathcal{R}_l}(D(Q\|P_m)),\ m = \overline{1,6},\ m \neq l, \quad l = \overline{1,5}, \tag{8}$$

$$E^*_{m|6} = E^*_{m|M}(E_{1|1},\dots,E_{5|5}) \triangleq \inf_{Q\in\mathcal{R}_6}(D(Q\|P_m)),\ m = \overline{1,5}, \tag{9}$$

$$E^*_{6|6} = E^*_{6|6}(E_{1|1},\dots,E_{5|5}) \triangleq \min_{l=\overline{1,5}} E_{6|l}. \tag{10}$$

Now we explain application of Sanov's theorem in hypotheses testing.

With assumption $\mathcal{A} = \mathcal{R}_l$, $P = P_m$ in Sanov's theorem for conditions (4)–(6), (7)–(10) we have

$$\lim_{N\to\infty} -\frac{1}{N}\log \alpha^N_{m|l}(\varphi^*_N) = \lim_{N\to\infty} -\frac{1}{N}\log P^N_m(\mathcal{R}_l) = \inf_{Q\in\mathcal{R}_l} D(Q\|P_m). \tag{11}$$

We can use notation $y_1^N \approx y_2^N$, when $g(y_1^N) = g(y_2^N) + \epsilon_N$, where $\epsilon_N \to 0$, for $N \to \infty$.

Now using (11) we can write

$$E_{m|l}(\varphi^*) \approx \inf_{Q\in\mathcal{R}_l} D(Q\|P_m). \tag{12}$$

Therefore the value of:

$$\alpha_{m|l}(\varphi^*_N) \approx \exp(-N \inf_{Q\in\mathcal{R}_l} D(Q\|P_m)) \approx \exp(-N E_{m|l}(\varphi^*_N)). \tag{13}$$

In fact the error probability $\alpha_{m|l}(\varphi_N)$ still goes to zero with exponential rate $\inf\limits_{Q\in\mathcal{R}_l} D(Q\|P_m)$ for P_m not in the set of $\mathcal{R}_l$.

Theorem 2. *For fixed on finite set $\mathcal{X}$ family of distributions $P_1, \ldots, P_6$ the following two statements hold:If the positive finite numbers $E_{1|1}, \ldots, E_{5|5}$ satisfy conditions:*

$$
\begin{aligned}
&E_{1|1} < \min_{l=\overline{2,6}} D(P_l \| P_1) \quad , \\
&\vdots \\
&E_{m|m} < \min\left[\min_{l=\overline{1,m-1}} E^*_{m|l}(E_{l|l}), \min_{l=\overline{m+1,6}} D(P_l\|P_m)\right], \quad m=\overline{2,5}
\end{aligned} \tag{14}
$$

hence:

a) *There exists a LAO sequence of tests φ^*_N, the reliability matrix of which $E^* = \left\{E^*_{m|l}(\varphi^*)\right\}$ is defined in* (7)–(10), *and all elements $E^*_{m|m}$ of it are positive.*

b) *Even if one of conditions* (14) *is violated, then the reliability matrix of an arbitrary test necessarily has an element equal to zero,(the corresponding error probability does not tend exponentially to zero).*

P r o o f. At first we remark that $D(P_l\|P_m) > 0$, for $l \neq m$. That is all measures $P_l, \quad l = \overline{1,6}$ are distinct. Now we prow the sufficiency of the conditions (14). Consider the following sequence of tests φ^* given by the sets

$$
\mathcal{B}_l^N(\mathbf{x}) = \bigcup_{Q \in R_l^N} \mathcal{T}_Q^N(\mathbf{x}), \quad l = \overline{1,6}. \tag{15}
$$

The sets $\mathcal{B}_l^N(\mathbf{x}), \quad l = \overline{1,6}$, satisfies conditions to give test, by means:

$$
\mathcal{B}_l^N(\mathbf{x}) \bigcap \mathcal{B}_m^N(\mathbf{x}) = \emptyset, \quad l \neq m,
$$

and

$$
\bigcup_{l=1}^{6} \mathcal{B}_l^N(\mathbf{x}) = \mathcal{X}^N.
$$

Now let us show, that exponent $E_{m|m}(\varphi^*)$ for sequence of tests φ^* defined in (15) is not less than $E_{m|m}$. We know from Lemma that:

$$
|\mathcal{T}_Q^N(\mathbf{x})| \approx \exp\{NH(Q)\} \quad \textit{and} \quad P^N(\mathcal{T}_Q^N(\mathbf{x})) \approx \exp\{-N(D(Q\|P)\} \quad m = \overline{1,6}
$$

and also with (13) we have

$$
\alpha^N_{m|m}(\varphi^*) \approx \exp\{-NE_{m|m}\},
$$

and

$$\alpha^N_{m|l}(\varphi^*) \approx \exp\{-NE^*_{m|l}(E_{m|m})\} \quad l = \overline{1,5} \quad m = \overline{1,5} \quad m \neq l,$$

$$\alpha^N_{m|6}(\varphi^*) \approx \exp\{-NE^*_{m|6}(E_{1|1}, \cdots, E_{5|5})\} \quad l = 6, \quad m = \overline{1,5}.$$

And at last for $m = l = 6$ we have:

$$\alpha^N_{M|M}(\varphi^*) \approx \exp\{-NE^*_{6|6}(E_{1|1}, \cdots, E_{5|5})\}.$$

With using (14)we know that all $E^*_{m|l}$ are strictly positive. The proof of part (a) will be finished if one demonstrates that the sequence of the test φ^* is LAO, that is at given finite $E_{1|1}, \cdots, E_{5|5}$ for any other sequence of tests φ^{**}

$$E^*_{m|l}(\varphi^{**}) \leq E^*_{m|l}(\varphi^*), \quad m, l = \overline{1,6}.$$

For this purpose it is sufficient to see that the sequence of tests asymptotically does not became better if the sets $\mathcal{B}^N_m(\mathbf{x})$ will not be union of some number of whole types $\mathcal{T}^N_Q(\mathbf{x})$, in other words, if a test φ^{**} is defined, for example, by sets $\mathcal{G}^N_1, \cdots, \mathcal{G}^N_6$ and, in addition, Q is such that

$$0 < |\mathcal{G}^N_l(\mathbf{x}) \bigcap \mathcal{T}^N_Q(\mathbf{x}| \approx |\mathcal{T}^N_Q(\mathbf{x})|,$$

The test φ^{**} will not became worse if instead of the set $\mathcal{G}^N_l$ one takes $\mathcal{G}^N_l(\mathbf{x}) \bigcap \mathcal{T}^N_Q(\mathbf{x})$ correspondingly decreasing the sets, which had nonempty intersection with $\mathcal{T}^N_Q(\mathbf{x})$. And at last we prove the necessity of the condition (14).It is just now shown that if the sequence of the tests is LAO, then it can be given by sets of (15) form. But the non fulfilment of the conditions (14) is equivalent either to violation of (3) or to equality zero some of $E^*_{m|l}$ given in (14), and this again contradicts with (3) because $E_{m|m}, \quad m = \overline{1,5}$, must be positive.

REFERENCES

[1] R. F. Ahlswede, E. A. Haroutunian. On logarithmically asymptotically optimal testing of hypotheses and identification. Lecture Notes in Comput. Sci., vol. **4123**. General Theory of Information Transfer and Combinations, Springer, 2006, 462–478.

[2] I. Csiszar, G. Longo. On the error exponent for source coding and for testing simple statistical hypotheses. *Studia Sci. Math. Hungar.* **6** (1971), 181–191.

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

[4] E. Haroutunian, A. Yessayan. On hypotheses optimal testing for two differently distributed objects. *Mathematical Problems of Computer Sciences* **26** (2006), 91–96.

[5] L. Navaei. Application of LDT to many hypotheses optimal testing for Markov chain. *Mathematical Problems of Computer Sciences* **31** (2008), 73–78.

[6] E. A. Haroutunian, L. Navaei. On statistical identification of Markov chain distribution subject to the reliability criterian. *Mathematical Problems of Computer Sciences* **32** (2009), 65–70.

[7] L. Navaei. On many hypotheses logarithmically asymptotically optimal testing via the theory of large deviations. *Far East J. Math. Sci.* **25**, *2* (2007), 335–344.

[8] L. Navaei. Identification of Optimality of Multiple Hypotheses Testing. Lambert Academic Publishing, Germany, 2010.

Payame Noor University
19395-4697, Tehran
I. R. of IRAN
e-mail: `leadernavaei@yahoo.com`

Received April 8, 2011
Revised May 23, 2011