Australian Journal of Basic and Applied Sciences, 2(4): 1454-1462, 2008
ISSN 1991-8178

Application of Markov Processes for Many Hypotheses Optimal Testing via LDT

[1]Masoud Yarmohammadi and [2]Leader Navaei

[1]Payame Noor University (PNU), Faculty of Science, Tehran, Iran.
[2]Payame Noor University (PNU), Faculty of Science, Azarshahr-Iran.

Abstract: The problem of many hypotheses testing for a model consisting of L>2 hypotheses on Markov chain is studied. This problem was introduced by Dobrushin In march 1987 at the seminar in the Institute of Problems of Information Transmission of the Academy of Science of the USSR. We apply large deviations techniques (LDT) and the method of types to the empirical distributions of finite states of Markov chain. It is proved that this method of investigation in solving the problem of logarighmically asymptotically optimal (LAO) hypotheses testing is easier than with the procedure that was introduced by Haroutunian (1988) and gives identical results. The matrix of exponents $E=\{E_{l|m}\}, m, l=\overline{1,L}$ of probabilities for the optimal ф tests $E_{l|m}(\phi)=\lim_{N\to\infty} -\frac{1}{N}\log\alpha_{l|m}(\phi_N)$, is determined, where $\alpha^N_{l|m}(\phi_N)$ $l \neq m$ for is the probability to accept the hypothesis when the hypothesis m is true.

Key Words: Large Deviation Techniques (LDT), Markov chain, irreducible transition matrics, Logarithmically asymptotically optimal (LAO), hypotheses testing, reliability matrix, Exponents of probability

INTRODUCTION

Application of information-theoretical methods in mathematical statistics are reflected in the monographs by Blahut (1947), Kullback (1959), Csiszár and Körner (1981), Zeituoni and Gutman (1991), Csiszár and shields (2004).

Many papers have been devoted to the study of exponential decrease of the error probabilities $\alpha_1^{(N)}$ of the first kind and $\alpha_2^{(N)}$ of the second kind of the optimal tests for two simple statistical hypotheses as the sample size N goes to infinity. Similar problems for Markov dependence of experiments were investigated by Natarajan (1985) and Haroutunian (1988).

In the book of Csiszár and Shields (2004) for independent identically distributed observations different asymptotical aspects of two hypotheses testing are considered via theory of large deviations.

In this paper we solve the problem purposed by Dubrushin for the optimal tests to describe the matrix of exponents $E=\{E_{l|m}\}, m, l=\overline{1\,,\,L}$ of probabilities $\alpha^N_{l|m} \approx \exp(-NE_{l|m})$, where $\alpha^N_{l|m}$ for $l \neq m$ is the probability of accepting hypotheses l, when hypotheses m is ture, for finite state of Markov chain to LAO hypotheses testing by application of LDT. We solve this problem easier and shorter than procedure that was introduced by Haroutounian (1988). In the next Section we express notations, basic concepts of large deviation principle and the method of types. In Section 3 we present a theorem of LDT for Markov chain and results for hypotheses testing. Some remarks will be presented in section 4.

Preliminaries:

In this paper we use exp-s and log-s at base 2. We also consider the standard convention that :

0 log 0 = 0, $0\log\frac{0}{0}=0$, $\alpha\log\frac{\alpha}{0}=\infty$ if a >0.

In the sequel we present notation of measures of information and some identities and the ideas of the method of types.

Corresponding Author: Masoud Yarmohammadi, Payame Noor University (PNU), Faculty of Science, Tehran, Iran.

In this paper finite sets are considered, which are denoted by *x,u*...... The size of the set *x* is denoted by |*x*|. Random variable (*RVs*) with value for *x,u*..... are denoted by X, Y..... probability distributions (PDs) we denote by *Q,P,V,W,PV,POV,*...... . Let PD of RV X be P={P(x), x ε *x*}.

The conditional probability distributions of random variable Y for given value *x* is

$V=\{V(y|x),\ \ x\in\chi,\ \ y\in\mu\}$

The joint PD of RV X and Y is $PoV=\{PoV(x,y)=P(x)V(y|x),\ \ x\in\chi, y\in\mu\}$.

And the marginal PD of RV Y is $PV=\{PV(y)=\sum_{x\in\chi}P(x)V(y|x),\ \ y\in\mu\}$.

The entropy of RV Y with PD *PV* is

$$H_{P,V}(Y)=-\sum_{y\in\mu}PV(y)\log PV(y).$$

The conditional entropy of RV *Y* with PD *PV*,

$$H_{P,V}(Y|X)=-\sum_{x\in\chi,y\in\mu}P(x)V(y|x)\log V(y|x),$$

The informational divergence of PD *P* and PD *Q* onis *x* $D(P\|Q)=\sum_{x\in\chi}P(x)\log\frac{P(x)}{Q(x)}$,

And for informational conditional divergence of PD *PoV* and PD *PoW* on $\chi\times\mu$, where $W=\{W(y|x),$ $x\in\chi, y\in\mu\}$ is

$$D(PoV\|PoW)=D(V\|W|P)=$$

$$\sum_{x\in\chi,y\in\mu}P(x)V(y|x)\log\frac{V(y|x)}{W(y|x)}\quad .$$

Therefore:

$$D(PoV\|QoW)=D(P\|Q)+D(V\|W|P).$$

The type *P* for a sequence (or vector) $\overline{x}=(x_1,....,x_N)\in\chi^N$ is a PD $P=\{P(x)=N^{-1}N(x|\overline{x}),x\in\chi\}$, where $N(x|\overline{x})$ is the number of repetition of symbol x among . $x_1,....,x_N$

The idea of the method of type is to partition the sets of all *N*- length sequences into classes according to its empirical distributions (types).

The joint type of $\overline{x}\in\chi^N$ and$\in\mu^N$ is the PD $\{N(x,y|\overline{x},\overline{y})/N,\ x\in\chi\ , y\in\mu\}$ where $N(x,y|\overline{x},\overline{y})$ is the number of occurrences of symbols pair in the(x,y) pair of vectors . In other word, joint type is the$(\overline{x},\overline{y})$ type of the sequence ,(x_1,y_1),$(x_2,y_2)\ldots,(x_N,y_N)$ from$(\chi\times\mu)^N$. We say that the conditional type of $\overline{y}$ for given $\overline{x}$ is *PD V*= $\{V(y|x),\ \ x\in\chi,\ \ y\in\mu\}$ if $N(x,y|\overline{x},\overline{y})=N(x|\overline{x})V(y|x)$ for all $x\in\chi, y\in\mu$. The set of all *PD* on χ is denoted by $P(\chi)$ and the subset of $P(\chi)$ consisting of the possible type of sequences $x\in\chi^N$ is denoted by $P_N(\chi)$. The set of vectors $\overline{x}$ of type *P* is denoted by $T_P^N,(T_P^N=\phi$ for *PD* $P\notin P_N(\chi))$. The set of all sequences $\overline{y}\in\mu^N$ of conditional type V for given $\overline{x}\in T_P^N$ is denoted by $T_{P,V}^N(Y\mid X)$ and called V-shell of $\overline{x}$. The set of all possible V-shells $\overline{x}$ for of type *P* is denoted by$V_N(\mu,P)$.

The properties of types are formulated below (Csiszár and Körner (1981) and Csiszár (1998)):

$$|V_N(\mu,P)|<(N+1)^{|\chi||\mu|} \quad (1)$$

For any conditional type V and $\overline{x}\in T_P^N$,

$$\begin{aligned}&(N+1)^{-|\chi||\mu|}\exp\{NH_{P,V}(Y\mid X)\}<\\&T_{P,V}^N(Y\mid \overline{x})\le \exp\{NH_{P,V}(Y\mid X)\}.\end{aligned} \quad (2)$$

For any conditional type V and $\overline{x}\in T_P^N$,

$$\begin{aligned}&(N+1)^{-|\chi||\mu|}\exp\{-ND(V\,\|\,W\mid P)\}<\\&Q^N(T_{P,V}^N(Y\mid \overline{x}))\le \exp\{-ND(V\,\|\,W\mid P)\}.\end{aligned} \quad (3)$$

Let $\overline{y}\in T_{P,V}^N(Y\mid \overline{x})$, then

$$W^N(\overline{y}\mid \overline{x})=\exp\{-N(H_{P,V}(Y\mid X)+D(V\,\|\,W\mid P))\}. \quad (4)$$

3. Problem Statement and Formulation of Results:

Let $\overline{x}=(x_0,x_1,x_2,....,x_N)$, $x_n\in\chi=\{1,2,....,I\}$, $\overline{x}\in\chi^{N+1}, N=0,1,2,.....,$ be vectors of observed states of simple homogeneous stationary Markov chain with finite number *I* of states. The *L* hypotheses concern the irreducible matrices of the transition probabilities

$$P_l=\{P_l(j\mid i),\ i=\overline{1,I},\ \ j=\overline{1,I}\},\ \ l=\overline{1,L}.$$

The stationarity of the chain provides existence foe each $l=\overline{1,L}$ of the unique stationary distributions

$Q_l=\{Q_l(i),\, i=\overline{1,I}\},\, l=\overline{1,I}$, such that $\sum_i Q_l(i)P_l(j\mid i)=Q_l(j),\ \sum_i Q_l(i)=1,\ \ i=\overline{1,I},\quad j=\overline{1,I}.$

We define the joint distributions $Q_l oP_l$,

$$Q_l oP_l=\{Q_l oP_l(i,j)=Q_l(i)P_l(j\mid i),\ i=\overline{1,I},\ j=\overline{1,I}\ \}\ \ l=\overline{1,L}.$$

The second order type of vector $\overline{x}$ the square matrix of I^2 relative frequencies of the simultaneous appearance on the pairs of neighbor places of the states *i* and *j* is $\{N(i,j)N^{-1},\ i=\overline{1,I},\ j=\overline{1,I}\}$. It is clear that $\sum_{i,j}N(i,j)=N$. Denote T_{QoP}^N the set of vectors from χ^{N+1} which have the type such that for some joint probability distribution QoP $N(i,j)=NQ(i)P(j\mid i),\ \ i=\overline{1,I},\ j=\overline{1,I}$.

Note that if the vector $\overline{x}\in T_{QoP}^N$, then

$$\sum_j N(i,j)=NQ(i),\ \ i=\overline{1,I},$$

$$\sum_i N(i,j) = NQ'(i),\ j=\overline{1,I}.$$

For somewhat different *PD Q*, but in accordance with the definition of *N (i,j)*,

$$| NQ(i) - NQ'(i) | \le 1,\ i=\overline{1,I}.$$

Now for $N \to \infty$, the distribution *Q* coincides with *Q*` and may be taken as stationary for conditional probability distribution *P*:

$$\sum_i Q(i)P(j\,|\,i) = Q\ (j),\ j\in\chi\quad .$$

Note that for $l=\overline{1,L}$ the probability of $\bar{x}$ from T^N_{QoP} can be written as .$Q_l oP_l^N(\bar{x}) = Q_l(x_0)\prod_{i,j} P_l(j\,|\,i)^{NQ(i)P(j|i)}$

We shall use the following definition of the probability of the vector $\bar{x}\in\chi^{N+1}$ of the Markov chain with transition probabilities P_l and stationary distribution Q_l ,

$$Q_l oP_l^N(\bar{x}) \equiv Q_l(x_0)\prod_{n=1}^{N} P_l(x_n\,|\,x_{n-1}),\ l=\overline{1,I},$$
$$Q_l oP_l^N(A) \equiv \bigcup_{\bar{x}\in A} Q_l oP_l^N(\bar{x}),\ A\subset\chi^{N+1}.$$

By means of non-randomized test $\phi_N(x)$ on the bases of the trajectory $\bar{x}=(x_0,x_1,....,x_N)$ of the *N+1* observations, the test accepts one of the hypotheses $H_l,\ l=\overline{1,L}$.

Let us denoted $\alpha^{(N)}_{l/m}(\phi_N)$ the probability to accept the hypotheses in the H_l condition that the H_m, $m\neq l$, is true. For $l=m$ we denote $\alpha^{(N)}_{m/m}(\phi_N)$ the probability to reject the hypotheses H_m . It is clear that

$$\alpha^{(N)}_{m/m}(\phi_N) = \sum_{l\neq m}\alpha^{(N)}_{l/m}(\phi_N),\ m=\overline{1,L}. \tag{5}$$

This probability is called the error probability of the m-th kind of the test ϕ_N . To every trajectory $\bar{x}$ the determined test ϕ_N puts in correspondence one from *L* hypotheses. So the space χ^{N+1} is divided into parts

$$g_l^N = \{\bar{x}, \phi_N(\bar{x}) = l\},\ l=\overline{1,L},$$

and

$$\alpha_{l/m}(\phi_N) = Q_m oP_m(g_l),\ m,\ l=\overline{1,L}\ .$$

Denot $E_{l/m}(\phi) = \lim\limits_{N\to\infty} -\frac{1}{N}\log \alpha_{l/m}(\phi_N),$

$$m,\ l = \overline{1,L}. \qquad (6)$$

The matrix $E = \{E_{l/m},\ m = \overline{1,L}\ ,l = \overline{1,L}\ \}$ we call the reliability matrix of the tests sequence φ.
Note that from definitions (5) and (6) it follows that $E_{m/m} = \min\limits_{l \neq m} E_{l/m}$. (7)

Definition:

The test sequence $\phi^* = (\phi_1, \phi_2, ...)$ is called LAO if for a given family of $E_{1/1}, E_{2/2}, ..., E_{L-1/L-1}$, the reliability matrix contain in the diagonal these numbers and the remained $L^2 - L + 1$ components take the maximal possible values.
Let $P = \{P(j\,|\,i)\}$ be an irreducible matrix of transition probabilities of some stationary Markov chain with the set χ of states, and $Q = \{Q(i),\ i = \overline{1,I}\}$ be the corresponding stationary *PD*.
Let us define the decision rules by the sets

$$R_l \equiv \{QoP : D(QoP\,||\,QoP_l) \leq E_{l/l}\},\ \ l = \overline{1, L-1}, \qquad (8)$$
$$R_L \equiv \{QoP : D(QoP\,||\,QoP_l) > E_{l/l},\ \ l = \overline{1, L-1}\}.$$

And introduce the functions:

$$E^*_{l/l}(E_{l/l}) \equiv E_{l/l},\quad l = \overline{1, L-1},$$
$$E^*_{l/m}(E_{l/l}) \equiv \inf_{QoP \in R_l} D(QoP\,||\,QoP_m),\ \ m = \overline{1,L},\ l \neq m\ ,l = \overline{1, L-1},$$
$$E^*_{L/m}(E_{l/l}, ..., E_{L-l/L-l}) \equiv \inf_{QoP \in R_L} D(QoP\,||\,QoP_m),\ m = \overline{1, L-1}, \qquad (9)$$
$$E^*_{L/L}(E_{l/l}, ..., E_{L-l/L-l}) \equiv \min_{l=\overline{1,L-1}} E^*_{l/L}.$$

Theorem 1 :
Let $\chi = \{1, 2, ..., I\}$ be the set of states of the finite stationary Markov chain possessing an irreducible transition matrix $P(\chi)$ and *A* be a set of joint distributions QoP such that its closure is equal to the closure of its interior, then for the empirical distribution $QoP^N(\overline{x})$ of a vector $\overline{x}$ from a strictly positive distributions $Q_m oP_m$ on χ :

$$\lim_{N\to\infty} -\frac{1}{N}\log Q_m oP_m^N(\overline{x} : QoP(\overline{x}) \in A) = \inf_{QoP \in A} D(QoP\,||\,QoP_m).$$

Proof: Let $A_N \equiv A \cap P_N(\chi)$, by the upper bound (3) and (4),

$$Q_m oP_m^N(QoP(\bar{x}) \in A_N) = Q_m oP_m^N(\bigcup_{QoP \in A_N} T_{QoP}^N)$$

$$\sum_{QOP(\bar{x})} Q_m oP_m^N(T_{QoP(\bar{x})}^N) \leq \ = \sum_{QoP(\bar{x}) \in A_N} \exp\{-N(D(QoP \,||\, QoP_m) + o(1))\},$$

where $o(1) = \max(|N^{-1} \log Q_m(i)| : Q_m(i) > 0)$,

$$\leq (N+1)^{|x|^2} \exp\{-N \inf_{QoP(\bar{x}) \in A}(D(QoP \,||\, QoP_m) + o(1))\}.$$

And also, by lower bound of (3), (4),

$$\sum_{QoP(\bar{x}) \in A_N} Q_m oP_m^N(T_{QoP(\bar{x})}^N) \geq \sum_{QoP(\bar{x}) \in A_N} Q_m oP_m^N(T_{QoP(\bar{x})}^N) \geq$$
$$\geq (N+1)^{|x|^2} \exp\{-N \inf_{QoP(\bar{x}) \in A}(D(QoP \,||\, QoP_m) + o(1))\}.$$

Since $\lim_{N \to \infty} N^{-1} \log(N+1)^{|x|^2} = 0$, $N \to \infty,\ o(1) \to 0$, also $D(QoP \,||\, QoP_m)$ is continuous in QoP_m, the hypotheses on A implies that $\inf_{QoP(\bar{x}) \in A_N} D(QoP \,||\, QoP_m)$ is arbitrarily close to $\inf_{QoP(\bar{x}) \in A} D(QoP \,||\, QoP_m)$ if *N* is large. We know that for the set *A* relation $\bar{A} = \overline{A^o}$ is true, by using the charactric properties of infimum, the proof is complete.

Now we explain application of Theorem 1 in hypotheses testing.
With assumption A=R$_l$ in Theorem 1 and relations (8), (9) we have

$$\lim_{N \to \infty} -\frac{1}{N} \log \alpha_{l/m}^N(\phi_N^*) = \lim_{N \to \infty} -\frac{1}{N} \log QoP_m^N(R_l) = \inf_{QoP(\bar{x}) \in R_l} D(QoP \,||\, QoP_m) \tag{10}$$

Using the notation $y_1^N \approx y_2^N$ when $g(y_1^N) = g(y_2^N) + \varepsilon_N$, where $\varepsilon_N \to 0$, for $N \to \infty$,
Now using (10) we can write $E_{l/m}(\phi^*) \approx \inf_{QoP \in R_l} D(QoP \,||\, QoP_m)$ (11)
Therefore the value of:

$$\alpha_{l/m}(\phi_N^*) \approx \exp(-N \inf_{QoP \in R_l} D(QoP \,||\, QoP_m)) \approx \exp(-NE_{l/m}(\phi_N^*)) \tag{12}$$

In fact the error probability $\alpha_{l/m}(\phi_N)$ still goes to zero with exponential rate $\inf_{QoP \in R_l} D(QoP \,||\, QoP_m)$ for

$Q_m o P_m$ not in the set of R_l.

Theorem 2 : Let χ be fix finite set a family of different distributions $P_1,.....,P_L$ the following two statements hold: if the positive finite numbers $E_{1/1},.....,E_{L-1/L-1}$ satisfy conditions:

$$0 < E_{1/1} < \min[\inf_{Q_m} D(Q_m o P_m \,||\, Q_m o P_1),\ m = \overline{2,L}], \tag{13}$$

$$0 < E_{l/l} < \min[E^*_{l/m}(E_{m/m}),\ m = \overline{1,L-1},$$

$$\inf_{Q_m} D(Q_m o P_m \,||\, Q_m o P_l),\ m = \overline{l+1,L}],\ l = \overline{2,L-1},$$

then:

There exists a LAO sequence of tests ϕ^*_N, the reliability matrix of which $E^* = \{E^*_{m/l}(\varphi^*)\}$ is defined in (9), and all elements $E^*_{m/m}$ of it are positive.

Even if one of conditions (13) is violated, then the reliability matrix of on arbitrary test necessarily has an element equal to zero, (the corresponding error probability does not tend exponentially to zero).

Proof: First we remind that $D(Q_m o P_m \,||\, Q_m o P_l) > 0$, for $l \neq m$, because all measures P_l , $l = \overline{1,L}$, , are distinct. Now we prove the sufficiency of the conditions (13). Consider the following sequence of tests ϕ^* given by the sets

$$\beta_l^N = \bigcup_{QoP \in R_l^N} T^N_{QoP},\ l = \overline{1,L}. \tag{14}$$

The sets $\beta_l^N,\ l = \overline{1,L}\setminus$ satisfies conditions to give test, by means:

$$B_l^N \bigcap B_m^N = \phi \quad l \neq m,$$

and

$$\bigcup_{l=1}^{L} \beta_l^N = \chi^N.$$

Now we show that the exponent $E_{m/m}(\phi^*)$ for sequence of test ϕ^* defined in (14) is not less than $E_{m/m}$. We know from relations (2), (3) and (4),

$$|T^N_{QoP}| \approx \exp\{-N \sum_{i,j} Q(i) P(j \,|\, i) \log P(j \,|\, i)\},$$

And

$$Q_m o P_m^N (T^N_{QoP}) \approx \exp\{-N(D(QoP \,||\, QoP_m))\},\ m = \overline{1,L}$$

According to application of Theorem 1 in hypotheses testing we can write,

$$\alpha_{m/m}^{N}(\varphi^{*}) \approx \exp\{-NE_{m/m}\},$$

and

$$\alpha_{l/m}^{N}(\phi^{*}) \approx \exp\{-NE_{l/m}^{*}(E_{l/l})\},\ \ l=\overline{1,L-1},\ m=\overline{1,L},\ m\neq l$$
$$\alpha_{L/m}^{N}(\varphi^{*}) \approx \exp\{-NE_{L/m}^{*}(E_{l/l},....,E_{L-1/L-1})\},\ \ l=\overline{1,L}.$$

Using (13) and (8) and (9) we can see that all $E_{l/m}^{*}$ are strictly positive. The proof of part (a) will be finished if one demonstrate that the sequence of the test ϕ^{*} is LAO, that is, at given finite

$E_{l/l},....,E_{L-1/L-1}$ for any other sequence of test ϕ^{**}

$$E_{l/m}^{*}(\phi^{**}) \leq E_{l/m}^{*}(\phi^{*}),\ \ m,l=\overline{1,L}$$

For this purpose it is sufficie to see that the sequence of tests asymptotically does not became better if the sets β_m^N will not be union of some number of whole types T_{QoP}^N . In other words, if a test ϕ^{**} is defined, for example, by sets $g_1^N,...,g_L^N$ and, in addition, QoP is such that $0<|g_l^N \bigcap T_{QoP}^N| \approx |T_{QoP}^N|$, Since$G^N_I$ nonempty intersection with $T^N_{QoP}(\mathrm{x})$, then the test φ^{**} will not became worse if instead of the set G^N_I one takes $G_l^N \supset T_{QoP}^N(\mathrm{x})$. Finally we prove the necessity of the condition (13). It is just now shown that if the sequence of the tests is LAO, then it can be given by sets of (14) form. But the non fulfillment of the conditions (13) is equivalent either to violation of (7), or to equality to zero some of $E_{l/m}^{*}$ given in (13), and this again contradicts with (7) because $E_{m/m}$, $m=\overline{1,L-1}$ must be positive.

Remarks:

After the change of hypotheses enumeration the theorem remains valid with corresponding changes in conditions (13) .

Instead of the diagonal elements of the reliability matrix it is possible to give finite values to the one arbitrary element in each of the first $L-1$ columns. The assumption of equalities in (9) as equations, and taking into account the monotone decrease and continuity, when they are finite, of function, defined in (9), it is possible to find a unique $E_{l|l}$ for given $E_{m|l}$ $l=\overline{1,L-1}$, . It is not difficult to reestablish the compatibility conditions, in which matrix E has all positive elements. For further computation of the reliability matrix elements it will be necessary to replace the infinite values of the obtained by $E_{l|l}$ arbitrary large, but finite ones .

ACKNOWLEDGEMENT

We are grateful of Professor E. A Haroutunian for his very helpful comments which substantially improved the presentation of the paper.

REFERENCES

Blahut, R.E., 1947. "Hypotheses testing and information theory", IEEE Trans, Inform Theory, 20(4): 405-417.

Csiszár, I. and P. Shields, 2004. "Information Theory and Statistics," fundamentals and trends in communications and information theory", now publishers Inc.

Csiszár, I. and J. Körner, 1981, "Information theory, coding theorem for discrete memoryless systems". Academic press, New York.

Csiszár, I., 1993. " The method of types", IEEE trans. Inform. Theory, 44(6): 2505-2523.

Dembo, A. and O. Zeituoni, 1993. " large deviations techniques and applications", Jons. And Bartlet. Publishers. London.

Gutman, M., 1989. "Asymptotically optimal classification for multiple test with empirically observed statistics". IEEE Trans. Inform. Theory, 35(2): 401-408.

Haroutunian, E.A., 1988. "On asymptotically optimal testing of hypotheses concerning Markov chain (in Russian)", Izvestia Acad. Nauk Armenian SSR. Seria Mathem., 22(1): 76-80.

Natarajan, S., 1985. " Large deviations, hypotheses testing, and source coding for finite Markov chains". IEEE Trans. Inform. Theory, 31(3): 360-365.

Navaei, L., 2007. "On many hypotheses LAO testing via the theory of large deviations", Far East Journal of Mathematical Sciences, 25(2): 335-344.

Navaei, L. and M. Yarmohammadi, 2007. ”On the theory of large deviations and multiple hypothese LAO testing for many independent objects, Far east j. Math. Sci. (FJMS), 26(2): 345-356.

Kullback, S., 1959. " Information theory and statistics", Wiley, NewYork.

Zeituoni, O. and M. Gutman, 1991. " On universal hypotheses testing via large deviations", IEEE trans. Inform. Theory, 37(2): 285-291.