
ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

THEORY OF PROBABILITY AND MATHEMATICAL STATISTICS

УДК 519.2

ПОСЛЕДОВАТЕЛЬНЫЙ КРИТЕРИЙ ОТНОШЕНИЯ ВЕРОЯТНОСТЕЙ ДЛЯ ПРОВЕРКИ МНОГИХ ПРОСТЫХ ГИПОТЕЗ О ПАРАМЕТРАХ ВРЕМЕННЫХ РЯДОВ С ТРЕНДОМ

Т. Т. ТУ¹⁾, А. Ю. ХАРИН¹⁾

¹⁾Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Рассмотрена проблема последовательного тестирования многих простых гипотез о параметрах временных рядов с трендом. Для построения последовательного теста использованы два подхода, в том числе M -нарный последовательный критерий отношения вероятностей и матричный последовательный критерий отношения вероятностей. Даны достаточные условия конечных завершений теста и существования конечных моментов их времени остановки. Получены верхние оценки для среднего числа наблюдений. При подходящих порогах эти тесты могут принадлежать некоторым определенным классам статистических тестов. Приводятся результаты вычислительных экспериментов.

Ключевые слова: тестирование многих гипотез; M -нарный последовательный критерий отношения вероятностей; матричный последовательный критерий отношения вероятностей; временные ряды с трендом.

Образец цитирования:

Ту ТТ, Харин АЮ. Последовательный критерий отношения вероятностей для проверки многих простых гипотез о параметрах временных рядов с трендом. *Журнал Белорусского государственного университета. Математика. Информатика*. 2019;1:35–45.
<https://doi.org/10.33581/2520-6508-2019-1-35-45>

For citation:

Tu TT, Kharin AYu. Sequential probability ratio test for many simple hypotheses on parameters of time series with trend. *Journal of the Belarusian State University. Mathematics and Informatics*. 2019;1:35–45. Russian.
<https://doi.org/10.33581/2520-6508-2019-1-35-45>

Авторы:

Тон Тхат Ту – аспирант кафедры теории вероятностей и математической статистики факультета прикладной математики и информатики.

Алексей Юрьевич Харин – доктор физико-математических наук, доцент; заведующий кафедрой теории вероятностей и математической статистики факультета прикладной математики и информатики.

Authors:

Ton That Tu, postgraduate student at the department of probability theory and mathematical statistics, faculty of applied mathematics and computer science.
tthatu@gmail.com

Alexey Yu. Kharin, doctor of science (physics and mathematics), docent; head of the department of probability theory and mathematical statistics, faculty of applied mathematics and computer science.
kharinAY@bsu.by

SEQUENTIAL PROBABILITY RATIO TEST FOR MANY SIMPLE HYPOTHESES ON PARAMETERS OF TIME SERIES WITH TREND

T. T. TU^a, A. Yu. KHARIN^a

^aBelarusian State University, 4 Niezaliežnasci Avenue, Minsk 220030, Belarus
 Corresponding author: T. T. Tu (tthattu@gmail.com)

The problem of sequential test for many simple hypotheses on parameters of time series with trend is considered. Two approaches, including M -ary sequential probability ratio test and matrix sequential probability ratio test are used for constructing the sequential test. The sufficient conditions of finite terminations of the test and the existence of finite moments of their stopping times are given. The upper bounds for the average numbers of observations are obtained. With the thresholds chosen suitably, these tests can belong to some specified classes of statistical tests. Numerical examples are presented.

Key words: multiple hypothesis testing; M -ary sequential probability ratio test; matrix sequential probability ratio test; time series with trend.

Introduction

Sequential probability ratio test (SPRT) proposed by Wald [1] is an effective approach for testing two simple hypotheses, and it has many applications in practical problems because of the optimal properties [2]. The test characteristics of SPRT are well studied in the case of independent identically distributed observations [1; 3–5]. If the hypothetical model is distorted, robustness of sequential tests is studied in [6], and an approach to robust sequential test construction is developed in [7]. However, in practice there are several situations, where it is natural to consider more than two hypotheses. For example, in the clinical trials, deciding which of several possible medical treatments is the most effective as quick as possible is a multihypothesis sequential problem. Most researches on this problem are based on two approaches: (1) construct a recursive solution to the optimization problem to get the optimal sequential test in a Bayesian setting [8]; and (2) extend and generalize the SPRT to the case of more than two hypotheses [5; 9–11]. Among the generalized versions of SPRT, M -ary sequential probability ratio test (MSPRT) and matrix sequential probability ratio test (MaSPRT) seem to be much simpler to construct and implement. Optimal properties of these methods have been well studied in the case of independent identically distributed observations [5; 11]. In many applied problems, the observed data can come from more complicated resources, such as time series. Sequential test for time series with trend has also been studied by Kharin and Tu [12–14]. In this paper, we will use MSPRT and MaSPRT for sequentially testing parameters of time series with trend.

Mathematical model

Let x_1, x_2, \dots be the observations of time series with trend in the following form [15]:

$$x_t = \theta^T \psi(t) + \xi_t, \quad t \geq 1, \quad (1)$$

where $\psi(t) = (\psi_1(t), \psi_2(t), \dots, \psi_m(t))^T$, $t \geq 1$ are the vectors of basic functions of trend, $\theta = (\theta_1, \theta_2, \dots, \theta_m)^T \in \mathbb{R}^m$ is an unknown vector of coefficients, and $\{\xi_t, t \geq 1\}$ is a sequence of independent identically distributed random variables, $\xi \sim N(0, \sigma^2)$.

Consider M simple hypotheses:

$$\mathcal{H}_i : \theta = \theta^i, \quad i \in T, \quad (2)$$

where $\theta^i \in \mathbb{R}^m$, $i \in T$, are known vectors, $T = \{1, 2, \dots, M\}$ and $\theta^i \neq \theta^j$ if $i \neq j$.

Firstly, we consider the called M -ary SPRT [11] for testing the multiple hypotheses (2). Assume that the prior probabilities of the hypotheses are known. Put $\pi_i = P(\theta = \theta^i)$, $i \in T$, and for $n \geq 1$, $p_n = (p_n^1, p_n^2, \dots, p_n^M)^T$, where $p_n^j = P(\theta = \theta^j | x_1, x_2, \dots, x_n)$ is the posterior probability of the hypothesis \mathcal{H}_j given n observations x_1, x_2, \dots, x_n .

The stopping time N_a and the final decision d_a for the MSPRT $\delta_a = (N_a, d_a)$ can be described as follows:

$$N_a = \inf \left\{ n \geq 1, p_n^k > \frac{1}{1 + A_k} \text{ for at least one } k \right\}, \quad (3)$$

$$d_a = \mathcal{H}_m, m = \arg \max_{1 \leq j \leq M} p_{N_a}^j, \quad (4)$$

where $A_k, k \in T$, are specified constants, $A_k \in (0, 1]$.

Remark 1. Note that $\sum_{k=1}^M p_n^k = 1$ and $\frac{1}{1 + A_k} \geq \frac{1}{2}, \forall k \in T$. Thus, there is at most one index $k \in T$ such that $p_n^k > \frac{1}{1 + A_k}$.

When $M = 2$, the test δ_a can be rewritten as follows:

$$\left\{ \begin{array}{l} \text{accept } H_1, \quad \prod_{i=1}^n \frac{n_1(x_i; (\theta^2)^T \psi(i), \sigma^2)}{n_1(x_i; (\theta^1)^T \psi(i), \sigma^2)} < \frac{\pi_1 A_1}{\pi_2}, \\ \text{accept } H_2, \quad \prod_{i=1}^n \frac{n_1(x_i; (\theta^2)^T \psi(i), \sigma^2)}{n_1(x_i; (\theta^1)^T \psi(i), \sigma^2)} > \frac{\pi_1}{\pi_2 A_2}. \end{array} \right.$$

This is the Wald's sequential probability ratio test.

Let $B = \{b_{ij}\}_{M \times M}$ be a fixed square matrix of size M , whose elements are positive except for the diagonal elements which are zero. Next, we define the matrix SPRT δ_b by building $\frac{M(M-1)}{2}$ one-sided SPRTs between hypotheses \mathcal{H}_i and $\mathcal{H}_j, i, j \in T, j \neq i$, as follows [5]:

$$\text{stop at the first } n > 0 \text{ such that, for some } i, \Lambda_n(i, j) > b_{ij} \text{ for all } j \neq i, \quad (5)$$

and accept \mathcal{H}_i , where i is the unique index satisfying these inequalities, and

$$\Lambda_n(i, j) = \ln \left(\prod_{k=1}^n \frac{n_1(x_k; (\theta^i)^T \psi(k), \sigma^2)}{n_1(x_k; (\theta^j)^T \psi(k), \sigma^2)} \right).$$

For the test $\delta = (N, d)$ let $\alpha_{ij}(\delta) = P_i(d = j), i \neq j, i, j \in T$, be the error probabilities of the test δ , and $\bar{\alpha}_i(\delta), i \in T$, be the probabilities of accepting hypothesis \mathcal{H}_i incorrectly. Note that the probabilities of rejecting the hypothesis \mathcal{H}_i when it is true, $\alpha_i(\delta) = \sum_{j \neq i} \alpha_{ij}(\delta), i \in T$, are also of interest. In addition, we are interested in the weighted error probabilities defined as $\beta_j(\delta) = \sum_{i \in T} w_{ij} P_i(d = j)$, where $\{w_{ij}\}_{M \times M}$ is a given matrix of weights, all positive except for the diagonal elements w_{ii} which are zero. We introduce four classes of tests:

$$\mathbf{C}^0(\bar{\alpha}) = \left\{ \delta : P(\text{accept } \mathcal{H}_i \text{ incorrectly}) \leq \bar{\alpha}_i, i \in T \right\},$$

$$\mathbf{C}^1([\alpha_{ij}]) = \left\{ \delta : \alpha_{ij}(\delta) \leq \alpha_{ij}, i, j \in T, i \neq j \right\},$$

$$\mathbf{C}^2(\alpha) = \left\{ \delta : \alpha_i(\delta) \leq \alpha_i, i \in T \right\}, \quad \mathbf{C}^3(\beta) = \left\{ \delta : \beta_i(\delta) \leq \beta_i, i \in T \right\},$$

where $[\alpha_{ij}]$ is a matrix of given error probabilities that are positive numbers less than 1, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)^T$, $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_M)^T$ are two vectors of positive numbers less than 1, and $\beta = (\beta_1, \beta_2, \dots, \beta_M)^T$ is a vector of positive numbers.

Remark 2. There are some relations between four classes of tests defined above.

- i) $C^0(\alpha) = C^3(\alpha)$ if $w_{ij} = \pi_i \text{sign}(|i - j|)$, $i, j \in T$;
- ii) $C^1([\alpha_{ij}]) \subset C^0(\alpha)$ if $\alpha_{ij} \leq \frac{\alpha_j}{(M-1)\pi_i}$, $i, j \in T$, $i \neq j$;
- iii) $C^1([\alpha_{ij}]) \subset C^2(\alpha)$ if $\alpha_i \geq \sum_{j \neq i} \alpha_{ij}$;
- iv) $C^1([\alpha_{ij}]) \subset C^3(\beta)$ if $w_{ij} \leq \frac{\beta_j}{(M-1)\alpha_{ij}}$, $i, j \in T$, $i \neq j$.

Lemma 1 [16]. For positive semidefinite matrices A and B of the same order

$$0 \leq \text{tr}(AB) \leq \text{tr}(A)\text{tr}(B).$$

Lemma 2 [17]. If X is a non-negative, integer valued random variable, then

$$E(X) = \sum_{n=1}^{+\infty} P(X \geq n).$$

Lemma 3 [17]. Let $r > 0$, and suppose that X is a non-negative random variable. Then the following inequalities hold:

$$\sum_{n=1}^{+\infty} n^{r-1} P(X \geq n) \leq E(X^r) \leq 1 + \sum_{n=1}^{+\infty} n^{r-1} P(X \geq n),$$

and

$$E(|X|^r) < \infty \text{ if and only if } \sum_{n=1}^{+\infty} n^{r-1} P(X \geq n) < \infty.$$

Main results

M-ary sequential probability ratio test. Using Bayes' rule, the posterior probabilities can be rewritten as:

$$p_n^k = \frac{\pi_k \prod_{i=1}^n n_1(x_i; (\theta^k)^T \psi(i), \sigma^2)}{\sum_{j=1}^M \pi_j \prod_{i=1}^n n_1(x_i; (\theta^j)^T \psi(i), \sigma^2)}, \quad n \geq 1, k \in T,$$

where $n_1(x; \mu, \sigma^2)$ is the probability density function of the normal distribution $N(\mu, \sigma^2)$.

In practice, we can use the following recurrent formula for calculating the values of p_n^k :

$$p_{n+1}^k = P(\theta = \theta^k | x_1, x_2, \dots, x_n, x_{n+1}) = \frac{p_n^k n_1(x_{n+1}; (\theta^k)^T \psi(n+1), \sigma^2)}{\sum_{j=1}^M p_n^j n_1(x_{n+1}; (\theta^j)^T \psi(n+1), \sigma^2)}, \quad n \geq 0,$$

where $p_0^k = \pi_k$, $k \in T$.

Clearly, the condition $p_n^k > \frac{1}{1 + A_k}$ can be rewritten as follows:

$$\sum_{j \neq k} \frac{\pi_j}{\pi_k} \prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^T \psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^T \psi(i), \sigma^2)} < A_k.$$

Denote $\Gamma_{ij} = (\theta^i - \theta^j)(\theta^i - \theta^j)^T$, $i, j \in T$, and $H_n = \sum_{i=1}^n \psi(i)\psi^T(i)$, $n \geq 1$. The following theorem will give us a sufficient condition for the finite termination of the test (2)–(4).

Theorem 1. If $\text{tr}(\Gamma_{ij}H_n) \rightarrow +\infty$ as $n \rightarrow +\infty$ for all $i, j \in T, i \neq j$, then the test (2)–(4) will terminate finitely with probability 1.

Proof. Let $k \in T$ be a fixed value. We have:

$$\begin{aligned} P_k(N_a > n) &= P_k \left(\bigcap_{l=1}^n \bigcap_{l \in T} \left\{ \sum_{j \in T \setminus \{l\}} \frac{\pi_j}{\pi_l} \prod_{i=1}^l \frac{n_1(x_i; (\theta^j)^\top \Psi(i), \sigma^2)}{n_1(x_i; (\theta^l)^\top \Psi(i), \sigma^2)} \geq A_l \right\} \right) \leq \\ &\leq P_k \left(\bigcap_{l \in T} \left\{ \sum_{j \in T \setminus \{l\}} \frac{\pi_j}{\pi_l} \prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^\top \Psi(i), \sigma^2)}{n_1(x_i; (\theta^l)^\top \Psi(i), \sigma^2)} \geq A_l \right\} \right) \leq P_k \left(\sum_{j \in T \setminus \{k\}} \frac{\pi_j}{\pi_k} \prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^\top \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^\top \Psi(i), \sigma^2)} \geq A_k \right) \leq \\ &\leq P_k \left(\bigcup_{j \in T \setminus \{k\}} \left\{ \frac{\pi_j}{\pi_k} \prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^\top \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^\top \Psi(i), \sigma^2)} \geq \frac{A_k}{M-1} \right\} \right) \leq \sum_{j \in T \setminus \{k\}} P_k \left(\prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^\top \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^\top \Psi(i), \sigma^2)} \geq \frac{\pi_k}{\pi_j} \frac{A_k}{M-1} \right) = \\ &= \sum_{j \in T \setminus \{k\}} P_k(\Lambda_n(j, k) \geq \gamma(k, j)), \end{aligned}$$

where $\gamma(k, j) = \ln \left(\frac{\pi_k}{\pi_j} \frac{A_k}{M-1} \right)$.

Obviously, for $k, j \in T, k \neq j$,

$$\Lambda_n(j, k) = -\frac{1}{2\sigma^2} \left\{ 2 \sum_{i=1}^n x_i (\theta^k - \theta^j)^\top \Psi(i) + (\theta^j)^\top H_n \theta^j - (\theta^k)^\top H_n \theta^k \right\},$$

and under the hypothesis \mathcal{H}_k , statistic $\Lambda_n(j, k)$ has the normal distribution with the following parameters:

$$E^{(k)}(\Lambda_n(j, k)) = -\frac{1}{2\sigma^2} \left\{ 2(\theta^k - \theta^j)^\top H_n \theta^k + (\theta^j)^\top H_n \theta^j - (\theta^k)^\top H_n \theta^k \right\} = -\frac{\text{tr}(\Gamma_{jk} H_n)}{2\sigma^2},$$

$$D^{(k)}(\Lambda_n(j, k)) = \frac{1}{\sigma^2} \sum_{i=1}^n \left((\theta^k - \theta^j)^\top \Psi(i) \right)^2 = \frac{\text{tr}(\Gamma_{jk} H_n)}{\sigma^2}.$$

Under the conditions of this theorem we get:

$$P_k(\Lambda_n(j, k) \geq \gamma(k, j)) = 1 - \Phi \left(\frac{\gamma(k, j) - E^{(k)}(\Lambda_n(j, k))}{\sqrt{D^{(k)}(\Lambda_n(j, k))}} \right) \rightarrow 0, \text{ as } n \rightarrow +\infty, \forall j \in T \setminus \{k\},$$

which implies $\lim_{n \rightarrow +\infty} P_k(N_a > n) = 0$. This completes the proof.

Corollary 1. The conditional expectations of stopping time N_a satisfy the following inequalities:

$$E^{(k)}(N_a) \leq 1 + \sum_{j \in T \setminus \{k\}} \sum_{n=1}^{+\infty} \Phi \left(-\frac{2\sigma^2 \gamma(k, j) + \text{tr}(\Gamma_{jk} H_n)}{2\sigma \sqrt{\text{tr}(\Gamma_{jk} H_n)}} \right), k \in T.$$

Proof. This is directly derived from the proof of theorem 1 and lemma 2.

Remark 3. Under the theorem 1 conditions, we have

$$\sum_{i=1}^n \sum_{j=1}^m \Psi_j^2(i) \rightarrow +\infty, \text{ as } n \rightarrow +\infty.$$

Proof. This is directly derived from lemma 1 and the fact that $\text{tr}(H_n) = \sum_{i=1}^n \sum_{j=1}^m \Psi_j^2(i)$.

Theorem 2. *If there exist positive constants κ_{ij} , $i, j \in T$, $i \neq j$, such that $\frac{\text{tr}(\Gamma_{ij} H_n)}{n^{\kappa_{ij}}} \rightarrow +\infty$ as $n \rightarrow +\infty$, then the stopping time N_a has finite moments of all orders.*

Proof. Let $k \in T$ be a fixed value. From the proof of theorem 1 and Markov's inequality we have:

$$\begin{aligned} P_k(N_a > n) &\leq \sum_{j \in T \setminus \{k\}} P_k \left(\prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^T \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^T \Psi(i), \sigma^2)} > \frac{\pi_k A_k}{\pi_j M - 1} \right) \leq \\ &\leq \sum_{j \in T \setminus \{k\}} \sqrt{\frac{\pi_j(M-1)}{\pi_k A_k}} E^{(k)} \left(\sqrt{\prod_{i=1}^n \frac{n_1(x_i; (\theta^j)^T \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^T \Psi(i), \sigma^2)}} \right) = \\ &= \sum_{j \in T \setminus \{k\}} \sqrt{\frac{\pi_j(M-1)}{\pi_k A_k}} \prod_{i=1}^n E^{(k)} \left(\sqrt{\frac{n_1(x_i; (\theta^j)^T \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^T \Psi(i), \sigma^2)}} \right). \end{aligned}$$

On the other hand,

$$\begin{aligned} E^{(k)} \left(\sqrt{\frac{n_1(x_i; (\theta^j)^T \Psi(i), \sigma^2)}{n_1(x_i; (\theta^k)^T \Psi(i), \sigma^2)}} \right) &= E^{(k)} \exp \left\{ -\frac{2x_i(\theta^k - \theta^j)^T \Psi(i) + ((\theta^j)^T \Psi(i))^2 - ((\theta^k)^T \Psi(i))^2}{4\sigma^2} \right\} = \\ &= \frac{\exp \left\{ -\frac{1}{4\sigma^2} \left[((\theta^j)^T \Psi(i))^2 - ((\theta^k)^T \Psi(i))^2 \right] \right\}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{x(\theta^k - \theta^j)^T \Psi(i) + (x - (\theta^k)^T \Psi(i))^2}{2\sigma^2} \right\} dx = \\ &= \frac{\exp \left\{ -\frac{1}{4\sigma^2} \left[((\theta^j)^T \Psi(i))^2 - ((\theta^k)^T \Psi(i))^2 \right] \right\}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{x^2 - x(\theta^k + \theta^j)^T \Psi(i) + ((\theta^k)^T \Psi(i))^2}{2\sigma^2} \right\} dx = \\ &= \exp \left\{ -\frac{1}{4\sigma^2} \left[((\theta^j)^T \Psi(i))^2 - ((\theta^k)^T \Psi(i))^2 \right] \right\} \exp \left\{ -\frac{1}{2\sigma^2} \left[((\theta^k)^T \Psi(i))^2 - \frac{1}{4}((\theta^k + \theta^j)^T \Psi(i))^2 \right] \right\} = \\ &= \exp \left\{ -\frac{((\theta^j - \theta^k)^T \Psi(i))^2}{8\sigma^2} \right\}. \end{aligned}$$

Therefore,

$$P_k(N_a > n) \leq \sum_{j \in T \setminus \{k\}} \sqrt{\frac{\pi_j(M-1)}{\pi_k A_k}} \exp \left\{ -\frac{\text{tr}(\Gamma_{jk} H_n)}{8\sigma^2} \right\}. \tag{6}$$

The result is derived from lemma 3 and the last inequality above.

Remark 4. The results of theorem 1 can be derived directly from the inequality (6).

The relations between thresholds A_i , $i \in T$, of the test δ_a and its error probabilities are shown in theorem 4.2 [11]. This theorem is still valid for the model of general independent observations. Now we can restate this result with our notation above as follows.

Theorem 3 [11]. *If the test (3)–(4) terminates finitely with probability one, then the following inequalities hold:*

$$a) \bar{\alpha}_k(\delta_a) = \sum_{j \in T, j \neq k} \pi_j \alpha_{j,k}(\delta_a) \leq \pi_k A_k \text{ for all } k;$$

$$b) \bar{\alpha}(\delta_a) = \sum_k \bar{\alpha}_k(\delta_a) \leq \sum_k \pi_k A_k;$$

$$c) \text{ if, in addition, } A_1 = A_2 = \dots = A_M = A, \text{ then } \bar{\alpha}(\delta_a) \leq \frac{A}{1+A}.$$

Corollary 2. *Under the theorem 3 conditions the following inequality holds:*

$$\alpha(\delta_a) \leq \frac{A_{\max} M \pi_{\max}}{A_{\max} \pi_{\max} + \pi_{\min}},$$

where $\alpha(\delta_a) = \sum_{i=1}^M \alpha_i(\delta_a)$; $A_{\max} = \max\{A_i, i \in T\}$; $\pi_{\max} = \max\{\pi_i, i \in T\}$; $\pi_{\min} = \min\{\pi_i, i \in T\}$.

Proof. From the proof of theorem 4.2 in [11] we have $\alpha_{kk}(\delta_a) \geq \frac{\bar{\alpha}_k(\delta_a)}{\pi_k A_k}$, $k \in T$, which implies

$$1 - \alpha_k(\delta_a) \geq \frac{\pi_{\min}}{A_{\max} \pi_{\max}} \sum_{j \neq k} \alpha_{jk}(\delta_a), \quad k \in T.$$

Taking summation over k we get:

$$M - \alpha(\delta_a) \geq \frac{\pi_{\min}}{A_{\max} \pi_{\max}} \sum_{k=1}^M \sum_{j \neq k} \alpha_{jk}(\delta_a) = \frac{\pi_{\min}}{A_{\max} \pi_{\max}} \alpha(\delta_a).$$

This completes the proof.

Remark 5.

• If we choose the thresholds $A_k = \min\left\{1, \frac{\bar{\alpha}_k^0}{\pi_k}\right\}$, $k \in T$, then $\delta_a \in \mathbf{C}^0(\bar{\alpha}^0)$, where $\bar{\alpha}^0 = (\bar{\alpha}_1^0, \dots, \bar{\alpha}_M^0)^\top$.

• If we set the maximum of total probability $\bar{\alpha}(\delta_a)$ of an incorrect decision to be $\alpha_0 \in (0, 1)$ in advance, then we can select $A_k = \min\left\{1, \frac{\alpha_0}{M \pi_k}\right\}$, $k \in T$, or $A_k = \alpha_0$, $k \in T$.

• If we set the maximum of total probability $\alpha(\delta_a)$ of rejecting a hypothesis when it is true to be $\alpha_0 \in (0, M)$ in advance, then we can select $A_k = \min\left\{1, \frac{\alpha_0 \pi_{\min}}{\pi_{\max} (M - \alpha_0)}\right\}$, $k \in T$.

Matrix sequential probability ratio test. Denote $\tau_i = \inf\{n \in \mathbb{N} : \Lambda_n(i, j) > b_{ij}, \text{ for all } j \in T \setminus \{i\}\}$, $i \in T$.

Then, for the test $\delta_b = (N_b, d_b)$ the stopping time N_b and the final decision d_b can be rewritten as:

$$N_b = \min\{\tau_i, i \in T\}, \quad d_b = i \text{ if } N_b = \tau_i. \quad (7)$$

Theorem 4. *Under the theorem 1 conditions the test (7) will terminate finitely with probability one.*

Proof. For each $i \in T$ and $n \geq 1$, we have:

$$P_i(\tau_i > n) = P_i\left(\bigcap_{k=1}^n \bigcup_{j \in T \setminus \{i\}} \{\Lambda_k(i, j) \leq b_{ij}\}\right) \leq P_i\left(\bigcup_{j \in T \setminus \{i\}} \{\Lambda_n(i, j) \leq b_{ij}\}\right) \leq \sum_{j \in T \setminus \{i\}} P_i(\Lambda_n(i, j) \leq b_{ij}).$$

Under hypothesis \mathcal{H}_i , statistic $\Lambda_n(i, j)$ has the normal distribution with the following parameters:

$$E^{(i)}(\Lambda_n(i, j)) = -\frac{1}{2\sigma^2} \left\{ 2(\theta^j - \theta^i)^\top H_n \theta^i + (\theta^i)^\top H_n \theta^i - (\theta^j)^\top H_n \theta^j \right\} = \frac{\text{tr}(\Gamma_{ij} H_n)}{2\sigma^2},$$

$$D^{(i)}(\Lambda_n(i, j)) = \frac{1}{\sigma^2} \sum_{k=1}^n \left((\theta^i - \theta^j)^\top \Psi(k) \right)^2 = \frac{\text{tr}(\Gamma_{ij} H_n)}{\sigma^2}.$$

From that we get $P_i(\tau_i > n) \leq \sum_{j \in T \setminus \{i\}} \Phi \left(\frac{b_{ij} - E^{(i)}(\Lambda_n(i, j))}{\sqrt{D^{(i)}(\Lambda_n(i, j))}} \right) \rightarrow 0$, as $n \rightarrow +\infty$. Therefore, $P_i(N_b > n) \leq P_i(\tau_i > n) \rightarrow 0$ as $n \rightarrow +\infty$. This completes the proof.

Corollary 3. *The conditional expectations of stopping time N_b satisfy the following inequalities:*

$$E^{(i)}(N_b) \leq 1 + \sum_{j \in T \setminus \{i\}} \sum_{n=1}^{+\infty} \Phi \left(\frac{2\sigma^2 b_{ij} - \text{tr}(\Gamma_{ij} H_n)}{2\sigma \sqrt{\text{tr}(\Gamma_{ij} H_n)}} \right), i \in T.$$

Proof. This is directly derived from the proof of theorem 4 and lemma 2.

Theorem 5. *Under the theorem 2 conditions the stopping time N_b has finite moments of all orders.*

Proof. Denote $f(x) = \Phi(x) - \varphi(x)$, $x \in \mathbb{R}$. We have $f'(x) = (1+x)\varphi(x) \forall x \in \mathbb{R}$. Therefore, $\Phi(x) < \varphi(x) \forall x < -1$. Under the theorem conditions, we get:

$$\frac{1}{n^{\kappa_{ij}}} \left(\frac{b_{ij} - E^{(i)}(\Lambda_n(i, j))}{\sqrt{D^{(i)}(\Lambda_n(i, j))}} \right)^2 \rightarrow +\infty, \forall i, j \in T, i \neq j, \text{ as } n \rightarrow +\infty,$$

and there exists an index $n_0 \in \mathbb{N}$ such that

$$\frac{b_{ij} - E^{(i)}(\Lambda_n(i, j))}{\sqrt{D^{(i)}(\Lambda_n(i, j))}} < -1, \forall i, j \in T, i \neq j, n \geq n_0.$$

From the proof of theorem 4 we obtain:

$$P_i(N_b > n) \leq \sum_{j \in T \setminus \{i\}} \varphi \left(\frac{b_{ij} - E^{(i)}(\Lambda_n(i, j))}{\sqrt{D^{(i)}(\Lambda_n(i, j))}} \right), i \in T, n > n_0.$$

The rest part of proof is derived directly from lemma 3.

The following known results are very useful to choose the threshold matrix B so that the test δ_b can belong to one of the classes $\mathbf{C}^1([\alpha_{ij}])$, $\mathbf{C}^2(\alpha)$, or $\mathbf{C}^3(\beta)$ mentioned above.

Lemma 4 [5]. *The following assertions hold:*

i) $\alpha_{ij}(\delta_b) \leq e^{-b_{ij}}, i, j \in T, i \neq j;$

ii) $\alpha_i(\delta_b) \leq \sum_{j \neq i} e^{-b_{ij}}, i \in T;$

iii) $\beta_i(\delta_b) \leq \sum_{j \neq i} w_{ij} e^{-b_{ij}}, i \in T.$

Remark 6 [5]. We have the following implications:

i) $b_{ij} = \ln \left(\frac{1}{\alpha_{ji}} \right), i, j \in T, i \neq j$ implies $\delta_b \in \mathbf{C}^1([\alpha_{ij}]);$

ii) $b_{ij} = b_j = \ln \left(\frac{M-1}{\alpha_j} \right), i, j \in T, i \neq j$ implies $\delta_b \in \mathbf{C}^2(\alpha);$

iii) $b_{ij} = b_j = \ln \left(\sum_{k \neq j} \frac{w_{kj}}{\beta_j} \right), i, j \in T, i \neq j$ implies $\delta_b \in \mathbf{C}^3(\beta).$

Numerical examples

The model (1) is considered and the hypotheses (2) is tested with the following parameters:

$$M = 3, m = 4, \sigma = 10, \psi(t) = \left(1, \frac{t}{10}, \frac{t^2}{100}, \frac{1}{t}\right)^T,$$

$$\theta^0 = (1, 1, 1, 1)^T, \theta^1 = (2, 2, 1, 1)^T, \theta^3 = (3, 3, 1, 1)^T.$$

With these values of parameters it is easy to check the conditions $\text{tr}(\Gamma_{ij} H_n) \rightarrow +\infty$ as $n \rightarrow +\infty$ for all $i, j \in \{1, 2, 3\}$, $i \neq j$, e. g. the tests δ_a and δ_b terminate finitely with probability 1. Denote the Monte-Carlo estimate of a characteristic γ by $\hat{\gamma}$. The number of experiments used in Monte-Carlo method is 50 000.

For the test δ_a , from remark 5 we can use the thresholds $A_i = \min\left(1, \frac{\bar{\alpha}_i^0}{\pi_i}\right)$, $i \in T$, with different vectors $\bar{\alpha}^0 = (\bar{\alpha}_1^0, \dots, \bar{\alpha}_M^0)$ and the fixed prior probabilities $\pi = (0.2, 0.3, 0.5)$. In this case the test δ_a will be in class $C^0(\bar{\alpha}^0)$. The Monte-Carlo estimates of error probabilities $\alpha_{21}(\delta_a)$, $\alpha_{31}(\delta_a)$, and conditional average number of observations $t_1(\delta_a) = E(N_a | \mathcal{H}_1)$ are given in table 1, where $\hat{\alpha}_1(\delta_a) = \pi_2 \hat{\alpha}_{21}(\delta_a) + \pi_3 \hat{\alpha}_{31}(\delta_a)$ is an estimate of P (accept \mathcal{H}_1 incorrectly).

Table 1

Monte-Carlo estimates for the characteristics of the test δ_a

| $\bar{\alpha}^0$ | $\hat{\alpha}_{21}(\delta_a)$ | $\hat{\alpha}_{31}(\delta_a)$ | $\hat{\alpha}_1(\delta_a)$ | $\hat{t}_1(\delta_a)$ | $E^{(1)}(N_a) \leq$ | |
|--------------------|-------------------------------|-------------------------------|----------------------------|-----------------------|---------------------|----------|
| (0.1, 0.1, 0.1) | 0.20986 | 0.01258 | 0.06925 | 0.11148 | 27.97730 | 63.58791 |
| (0.05, 0.1, 0.1) | 0.10484 | 0.00222 | 0.03256 | 0.14548 | 33.41488 | 71.33009 |
| (0.01, 0.1, 0.1) | 0.01984 | 0.00006 | 0.00598 | 0.17368 | 43.38288 | 85.48740 |
| (0.05, 0.05, 0.1) | 0.11448 | 0.00176 | 0.03522 | 0.0765 | 34.59546 | 71.33009 |
| (0.05, 0.01, 0.1) | 0.11854 | 0.00224 | 0.03668 | 0.02822 | 35.42372 | 71.33009 |
| (0.05, 0.01, 0.05) | 0.12136 | 0.00212 | 0.03747 | 0.01452 | 35.90788 | 71.33009 |

In table 1, the inequality $\hat{\alpha}_1(\delta_a) \leq \bar{\alpha}_1^0$ is satisfied with all given values of vector $\bar{\alpha}^0$. With the same levels of $\bar{\alpha}_2^0, \bar{\alpha}_3^0$ the decrease of $\bar{\alpha}_1^0$ leads to the decrease of A_1 , and as a result the conditional average number of observations $t_1(\delta_a)$ increases. The changes in probability $\bar{\alpha}_1(\delta_a) = P$ (accept \mathcal{H}_1 incorrectly) and probability $\alpha_1(\delta_a)$ of rejecting hypothesis \mathcal{H}_1 when it is true are likely to be the opposite. Additionally, with the same levels of $\bar{\alpha}_1^0$, the value of $\hat{\alpha}_1(\delta)$ changes negligibly with respect to $\bar{\alpha}_2^0, \bar{\alpha}_3^0$. Using corollary 1, we can get the upper bounds for the conditional expected values of number of observations $E^{(k)}(N_a)$, $k \in T$. Because the dependence of the upper bound of $E^{(k)}(N_a)$ on the index k is expressed only by A_k and Γ_{kj} , $j \neq k$, this value will not change if we fix k -th element in vector $\bar{\alpha}^0$.

For the test δ_b we choose the matrix of thresholds B according to remark 6 as follows:

$$B = \{b_{ij}\}_{3 \times 3}, b_{ij} = b_j = \ln\left(\frac{M-1}{\alpha_j}\right), i, j \in \{1, 2, 3\}, i \neq j.$$

In this case the test δ_b will be in class $C^2(\alpha^0)$, where $\alpha^0 = (\alpha_1, \dots, \alpha_M)$ is a given vector of upper bounds for the error probabilities $\alpha_i(\delta_b)$, $i = \overline{1, 3}$. The Monte-Carlo estimates of error probabilities $\alpha_1(\delta_b)$, $\alpha_2(\delta_b)$ and conditional average number of observations $t_1(\delta_b) = E(N_b | \mathcal{H}_1)$, $t_2(\delta_b) = E(N_b | \mathcal{H}_2)$ are presented in table 2 with different vectors α^0 .

Table 2

Monte-Carlo estimates for the characteristics of the test δ_b

| α^0 | $\hat{\alpha}_1(\delta_b)$ | $\hat{\alpha}_2(\delta_b)$ | $\hat{\alpha}_3(\delta_b)$ | $\hat{t}_1(\delta_b)$ | $\hat{t}_2(\delta_b)$ | $\hat{t}_3(\delta_b)$ | $E^{(i)}(N_b) \leq$ |
|--------------------|----------------------------|----------------------------|----------------------------|-----------------------|-----------------------|-----------------------|---------------------|
| (0.1, 0.1, 0.1) | 0.021 22 | 0.073 32 | 0.020 56 | 44.480 80 | 55.164 92 | 44.588 76 | 74.509 47 |
| (0.05, 0.1, 0.1) | 0.011 92 | 0.069 42 | 0.020 16 | 44.790 36 | 57.141 22 | 44.553 26 | 74.509 47 |
| (0.01, 0.1, 0.1) | 0.002 72 | 0.072 38 | 0.016 42 | 45.040 42 | 61.639 28 | 44.714 20 | 74.509 47 |
| (0.05, 0.05, 0.1) | 0.010 94 | 0.034 28 | 0.019 18 | 49.036 96 | 58.239 30 | 48.827 16 | 78.331 66 |
| (0.05, 0.01, 0.1) | 0.011 00 | 0.006 66 | 0.020 26 | 57.040 46 | 59.004 56 | 56.735 46 | 85.742 48 |
| (0.05, 0.01, 0.05) | 0.010 88 | 0.006 22 | 0.010 70 | 57.090 78 | 60.861 86 | 57.018 48 | 88.113 51 |

In table 2 the inequalities $\hat{\alpha}_i(\delta_b) \leq \alpha_i^0$, $i = \overline{1, 3}$, are satisfied with all given values of vector α^0 . If we fix two elements in vector α^0 , the increase or decrease of the rest one leads to the change of conditional average number of observations under corresponding hypothesis in the opposite direction. Comparing with the results of the test δ_a in table 1, the test δ_b need more observations to get the final decision, but it seems to have much less error probabilities of rejecting a hypothesis when this hypothesis is true. Furthermore, we can use the results in corollary 3 to get the upper bounds for $E^{(k)}(N_b)$, $k \in T$. Note that from the expressions of the upper bounds for $E^{(k)}(N_b)$, $k \in T$, these values are independent of the index k , e. g. they do not change with respect to k -th element of vector α^0 .

Библиографические ссылки

1. Wald A. *Sequential analysis*. New York: John Wiley and Sons; 1947.
2. Айвазян С.А. Сравнение оптимальных свойств критериев Неймана – Пирсона и Вальда. *Теория вероятностей и ее применения*. 1959;4(1):86–93.
3. Харин А.Ю. *Робастность байесовских и последовательных статистических решающих правил*. Минск: БГУ; 2013.
4. Харин А.Ю. Об одном подходе к анализу последовательного критерия отношения правдоподобия для различения простых гипотез. *Вестник БГУ. Серия 1. Физика. Математика. Информатика*. 2002;1:92–96.
5. Tartakovsky A, Nikiforov I, Basseville M. *Sequential analysis: hypothesis testing and changepoint detection*. Boca Raton: CRC Press; 2015.
6. Kharin A. Performance and robustness evaluation in sequential hypotheses testing. *Communications in Statistics. Theory and Methods*. 2016;45(6):1693–1709. DOI: 10.1080/03610926.2014.944659.
7. Kharin A.Yu. Robustness of sequential testing of hypotheses on parameters of M-valued random sequences. *Journal of Mathematical Sciences*. 2013;189(6):924–931. DOI: 10.1007/s10958-013-1233-9.
8. Sobel M, Wald A. A sequential design procedure for choosing one of three hypotheses concerning the unknown mean of normal distribution. *Annals of Mathematical Statistics*. 1949;20:502–522.
9. Tartakovskii AG. Sequential testing of many simple hypotheses with independent observations. *Problems of Information Transmission*. 1989;24(4):299–309.
10. Armitage P. Sequential analysis with more than two alternative hypotheses and its relation to discriminant function analysis. *Journal of the Royal Statistical Society*. 1947;9:250–263.
11. Baum CW, Veeravalli VV. A sequential procedure for multihypothesis testing. *IEEE Transactions on Information Theory*. 1994;40(6):1994–2007. DOI: 10.1109/18.340472.
12. Харин А.Ю., Ту Т.Т. Последовательная статистическая проверка гипотез о параметрах временных рядов с трендом при пропусках наблюдений. *Известия Национальной академии наук Беларуси. Серия физико-математических наук*. 2016;3:38–46.
13. Харин А.Ю., Ту Т.Т. О вычислении вероятностей ошибок усеченного последовательного критерия отношения вероятностей. *Журнал Белорусского государственного университета. Математика. Информатика*. 2018;1:68–76.
14. Kharin A, Ton That Tu. Performance and robustness analysis of sequential hypotheses testing for time series with trend. *Austrian Journal of Statistics*. 2017;46(3–4):23–36. DOI: 10.17713/ajs.v46i3-4.668.
15. Андерсон Т. *Статистический анализ временных рядов*. Москва: Мир; 1976.
16. Coope ID. On matrix trace inequalities and related topics for products of Hermitian matrices. *Journal of Mathematical Analysis and Applications*. 1994;188:999–1001. DOI: 10.1006/jmaa.1994.1475.
17. Gut A. *Probability: A Graduate Course*. New York: Springer-Verlag; 2005.

References

1. Wald A. *Sequential analysis*. New York: John Wiley and Sons; 1947.
2. Aivazian SA. [A comparison of the optimal properties of the Neuman – Pearson and the Wald sequential probability ratio tests]. *Teoriya veroyatnostei i ee primeneniya*. 1959;4(1):86–93. Russian.
3. Kharin AY. *Robastnost' baiesovskikh i posledovatel'nykh statisticheskikh reshayushchikh pravil* [Robustness of Bayesian and sequential statistical decisions]. Minsk: Belarusian State University; 2013. Russian.

4. Kharin AY. [An approach to performance analysis of the sequential probability ratio test for simple hypotheses]. *Vestnik BGU. Seriya 1. Fizika. Matematika. Informatika*. 2002;1:92–96. Russian.
5. Tartakovsky A, Nikiforov I, Basseville M. *Sequential analysis: hypothesis testing and changepoint detection*. Boca Raton: CRC Press; 2015.
6. Kharin A. Performance and robustness evaluation in sequential hypotheses testing. *Communications in Statistics. Theory and Methods*. 2016;45(6):1693–1709. DOI: 10.1080/03610926.2014.944659.
7. Kharin AY. Robustness of sequential testing of hypotheses on parameters of M-valued random sequences. *Journal of Mathematical Sciences*. 2013;189(6):924–931. DOI: 10.1007/s10958-013-1233-9.
8. Sobel M, Wald A. A sequential design procedure for choosing one of three hypotheses concerning the unknown mean of normal distribution. *Annals of Mathematical Statistics*. 1949;20:502–522.
9. Tartakovskii AG. Sequential testing of many simple hypotheses with independent observations. *Problems of Information Transmission*. 1989;24(4):299–309.
10. Armitage P. Sequential analysis with more than two alternative hypotheses and its relation to discriminant function analysis. *Journal of the Royal Statistical Society*. 1947;9:250–263.
11. Baum CW, Veeravalli VV. A sequential procedure for multihypothesis testing. *IEEE Transactions on Information Theory*. 1994;40(6):1994–2007. DOI: 10.1109/18.340472.
12. Kharin AY, Tu TT. Sequential statistical hypotheses testing on parameters of time series with trend under missing values. *Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics Series*. 2016;3:38–46. Russian.
13. Kharin AY, Tu TT. On error probabilities calculation for the truncated sequential probability ratio test. *Journal of the Belarusian State University. Mathematics and Informatics*. 2018;1:68–76. Russian.
14. Kharin A, Ton That Tu. Performance and robustness analysis of sequential hypotheses testing for time series with trend. *Austrian Journal of Statistics*. 2017;46(3–4):23–36. DOI: 10.17713/ajs.v46i3-4.668.
15. Anderson T. *Statistical analysis of time series*. New York: John Wiley & Sons; 1971. 704 p.
Russian edition: Anderson T. *Statisticheskii analiz vremennykh ryadov*. Moscow: Mir; 1976.
16. Coope ID. On matrix trace inequalities and related topics for products of Hermitian matrices. *Journal of Mathematical Analysis and Applications*. 1994;188:999–1001. DOI: 10.1006/jmaa.1994.1475.
17. Gut A. *Probability: A Graduate Course*. New York: Springer-Verlag; 2005.

Received by editorial board 13.11.2018.