AUSTRIAN JOURNAL OF STATISTICS
Volume 34 (2005), Number 2, 153–162

Robust Sequential Testing of Hypotheses on Discrete Probability Distributions

Alexey Kharin and Dzmitry Kishylau
Belarusian State University, Minsk

Abstract: The paper is devoted to analysis of approximations of the models in sequential hypotheses testing and to construction of new robust sequential test under the distortions presented by mixtures of probability distributions. The previous results of the authors are extended to the case of arbitrary discrete probability distributions. The theory is illustrated by computer modelling.

Keywords: Robustness, Sequential Test, Distortion, Computer Modelling.

1 Introduction

In applications, especially in medicine (Whitehead, 1997), statistical quality control (Mason and Young, 2002), biology (Durbin, 1998) and finance (Lai, 2001), sequential methods of hypotheses testing (see Wald, 1947; Siegmund, 1985) are used quite often. The usage of this type of statistical procedures for such applications follows naturally from the data character (see, for example Bauer and Röhmel, 1995).

Sequential procedures are applied in practice to the data which does not follow exactly a considered hypothetical model (Huber, 1981). Usually a high percentage of the data fits a hypothetical model, and within this model assumptions the sequential test could be optimal. But a small part of data does not follow the hypothetical model — the hypothetical model is distorted (see Rieder, 1994). This leads often to loose of optimality of the test. That is why the robustness analysis for a sequential test used should be performed, and the robust tests need to be constructed to treat the data under distortions.

Some results on robustness analysis for sequential testing of hypotheses on data from continuous distributions are presented in (Quang, 1985). Our previous results are related to the special case of discrete model of observations (see Kharin, 2002). In this paper, we consider the case of arbitrary discrete probability distributions with finite sets of values.

2 Mathematical Model, Sequential Probability Ratio Test, Distortions

Let independent discrete random variables x_1, x_2, ... be defined on a measurable space $(\Omega, \mathcal{F})$, $x_t \in U = \{u_1, \dots, u_M\}$, $\forall t \in \mathbf{N}$, $U \subset \mathbf{R}^N$, $M < \infty$. Let these random variables be identically distributed:

$$P_\theta(u) = P\{x_t = u \mid \theta\},\ u \in U, \tag{1}$$

where $\theta \in \Theta = \{0, 1\}$ is an unobservable parameter. There are two simple hypotheses on the parameter θ:

$$H_0 :\ \theta = 0,\ \ H_1 :\ \theta = 1. \tag{2}$$

Denote the accumulated likelihood ratio statistic:

$$\Lambda_n = \Lambda_n(x_1, \ldots, x_n) = \sum_{t=1}^{n} \lambda_W(x_t),\ n \in \mathbf{N}, \tag{3}$$

where

$$\lambda_W(u) = \log \frac{P_1(u)}{P_0(u)},\ u \in U.$$

To test the hypotheses (2) after n ($n = 1, 2, \ldots$) observations, the decision

$$d_n = \mathbf{1}_{[C_+, +\infty)}(\Lambda_n) + 2 \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n) \tag{4}$$

is made according to the sequential probability ratio test (SPRT), (see Siegmund, 1985). By $\mathbf{1}_D(\cdot)$ we denote the indicator function of the set D. The decisions $d_n = 0$ and $d_n = 1$ mean stopping of the observation process and acceptance of the correspondent hypothesis. The decision $d_n = 2$ means that the $(n+1)$-th observation is to be made. The thresholds $C_-, C_+ \in \mathbf{R}$, $C_- < 0 < C_+$, are the parameters of the test.

Below we generalize the test (4) replacing the function $\lambda_W(\cdot)$ with another function $\lambda : U \to \mathbf{R}$. In this case, we indicate that the test (4) is based on the function $\lambda(\cdot)$.

Consider the model of distortions which will be analysed in Section 4. Let the hypothetical model be under contamination of the Tukey–Huber type, (see Huber, 1981). This means that instead of (1) the observations $x_1, x_2, \ldots$ come from a mixtured discrete probability distribution

$$\bar{P}_k(u) = \bar{P}(u; k) = (1 - \varepsilon_k) P_k(u) + \varepsilon_k \tilde{P}_k(u),\ u \in U,\ k = 0, 1, \tag{5}$$

where $\varepsilon_k \in [0, \varepsilon_{k+}]$, $k = 0, 1$, are unknown probabilities of contamination, $\tilde{P}_k(u)$ is an arbitrary contaminating probability distribution, $\tilde{P}_k(\cdot) \neq P_k(\cdot)$.

3 Evaluation of the Characteristics of the Test

In Kharin (2002) explicit expressions for the characteristics of the test (4) are given under additional assumption on the function $\lambda_W(\cdot)$:

$$\lambda_W(u) = m_u a,\ \forall u \in U,\ \ m_u \in \mathbf{Z},\ a \in \mathbf{R}_+. \tag{6}$$

Under this assumption the thresholds C_-, C_+ can be replaced with $a\lfloor C_-/a \rfloor$ and $a\lceil C_-/a \rceil$ respectively without changes in the test (4). As long as (6) is satisfied, assume that $\bar{C}_- = C_-/a$, $\bar{C}_+ = C_+/a \in \mathbf{Z}$. For $k = 0, 1$ introduce the notation:

$$Q^{(k)} = (q_{ij}^{(k)}),\ q_{ij}^{(k)} = \sum_{u \in U} \delta_{m_u, j-i} \bar{P}_k(u),\ \bar{C}_- < i, j < \bar{C}_+;$$

$$R^{(k)} = (r_{ij}^{(k)}),\ \bar{C}_- < i < \bar{C}_+,\ j = \bar{C}_-, \bar{C}_+,$$

$$r_{i\bar{C}_-}^{(k)} = \sum_{u \in U} \mathbf{1}(\bar{C}_- - i - m_u) \bar{P}_k(u),\ r_{i\bar{C}_+}^{(k)} = \sum_{u \in U} \mathbf{1}(m_u + i - \bar{C}_+) \bar{P}_k(u);$$

$$\pi^{(k)} = (\pi_i^{(k)}),\ \pi_i^{(k)} = \sum_{u\in U} \delta_{m_u,i}\bar{P}_k(u),\ \bar{C}_- < i < \bar{C}_+;$$

$$\pi_{\bar{C}_-}^{(k)} = \sum_{u\in U} \mathbf{1}(\bar{C}_- - m_u)\bar{P}_k(u),\ \pi_{\bar{C}_+}^{(k)} = \sum_{u\in U} \mathbf{1}(m_u - \bar{C}_+)\bar{P}_k(u),$$

where $\delta_{i,j}$ is the Kroneker symbol, $\mathbf{1}(\cdot)$ is the unit step function. Define the matrices $S^{(k)} = \mathbf{I}_{N-2} - Q^{(k)}$, $N = \bar{C}_+ - \bar{C}_- + 1$, and $B^{(k)} = \left(S^{(k)}\right)^{-1} R^{(k)}$, $k = 0, 1$. We denote the i-th column of the matrix $B^{(k)}$ as $B_{(i)}^{(k)}$. Let $\mathbf{1}_{N-2}$ be the vector column of size $(N-2)$ all elements of which are equal to 1.

The next result gives the explicit expressions for the error probabilities of type I and II denoted by α, β, and conditional expected sequence sizes (ESSs) $t^{(k)}$ (if the hypothesis H_k is true), $k = 0, 1$, for the test (4) under the assumption (6).

Theorem 1 (Kharin, 2002) *If under conditions (1), (2), (6) $|S^{(k)}| \neq 0$, then the characteristics of the test (4) have the following expressions:*

$$\begin{gathered} t^{(k)} = \left(\pi^{(k)}\right)' \left(S^{(k)}\right)^{-1} \mathbf{1}_{N-2} + 1,\ k = 0, 1, \\ \alpha = \left(\pi^{(0)}\right)' B_{(2)}^{(0)} + \pi_{\bar{C}_+}^{(0)},\ \beta = \left(\pi^{(1)}\right)' B_{(1)}^{(1)} + \pi_{\bar{C}_-}^{(1)}. \end{gathered} \tag{7}$$

If the assumption (6) is broken, we can use the so-called approximated test to get estimates of the conditional error probabilities and the conditional ESSs of the test (4). We construct a function $\lambda : U \to \mathbf{R}$ that satisfies (6) and approximates the function $\lambda_W(\cdot)$:

$$\max_{u\in U} |\lambda(u) - \lambda_W(u)| \leq \delta, \tag{8}$$

where $\delta > 0$ is a parameter of approximation. As it is shown in (Kharin, 2004), the function λ satisfying (8), (6) can always be constructed.

As $\lambda(\cdot)$ satisfies (6), Theorem 1 holds, and we get the conditional error probabilities and the ESSs of the test based on the function $\lambda(\cdot)$. With δ being small enough these values approximate the characteristics of the SPRT. Choosing $\lambda(\cdot)$ in a proper way we get upper and lower bounds of these characteristics.

Let $\underline{\alpha}$, $\underline{\beta}$, $\overline{\alpha}$, $\overline{\beta}$ be the error probabilities of the test (4) based on the functions $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot)$ respectively. Define random variables (stopping times)

$$\begin{gathered} \tau_0 = \inf\{n :\ \Lambda_n \leq C_-\},\ \tau_1 = \inf\{n :\ \Lambda_n \geq C_+\}, \\ \overline{\tau}_0 = \inf\{n :\ \overline{\Lambda}_n \leq C_-\},\ \overline{\tau}_1 = \inf\{n :\ \overline{\Lambda}_n \geq C_+\}, \\ \underline{\tau}_0 = \inf\{n :\ \underline{\Lambda}_n \leq C_-\},\ \underline{\tau}_1 = \inf\{n :\ \underline{\Lambda}_n \geq C_+\}, \end{gathered}$$

where

$$\overline{\Lambda}_n = \sum_{t=1}^{n} \overline{\lambda}(x_t),\ \underline{\Lambda}_n = \sum_{t=1}^{n} \underline{\lambda}(x_t).$$

Denote by $E_k\{\cdot\}$ the expectation w.r.t. the distribution P_k, $k = 0, 1$, by $o(1)$ the Landau symbol.

Theorem 2 *If the function $\lambda(\cdot)$ in the test (4), (3) is bounded by the functions $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot) : U \to \mathbf{R}$, that is*

$$\underline{\lambda}(u) \le \lambda(u) \le \overline{\lambda}(u),\ \forall u \in U, \tag{9}$$

then

$$\underline{\alpha} \le \alpha \le \overline{\alpha},\ \overline{\beta} \le \beta \le \underline{\beta}. \tag{10}$$

If, in addition, $\max_{u\in U}\left\{|\underline{\lambda}(u)-\lambda(u)|, |\overline{\lambda}(u)-\lambda(u)|\right\} = \delta \to 0$*, then*

$$t_-^{(k)} \le t^{(k)} \le t_+^{(k)},\ t_\pm^{(k)} = T_\pm^{(k)} + o(1),\ k = 0, 1, \tag{11}$$

where

$$\begin{array}{l} T_-^{(0)} = \overline{\alpha} E_0\left\{\overline{\tau}_1 \mid \overline{\tau}_1 < \overline{\tau}_0\right\} + (1-\underline{\alpha}) E_0\left\{\underline{\tau}_0 \mid \underline{\tau}_0 < \underline{\tau}_1\right\}, \\ T_+^{(0)} = \underline{\alpha} E_0\left\{\underline{\tau}_1 \mid \underline{\tau}_1 < \underline{\tau}_0\right\} + (1-\overline{\alpha}) E_0\left\{\overline{\tau}_0 \mid \overline{\tau}_0 < \overline{\tau}_1\right\}, \\ T_-^{(1)} = \underline{\beta} E_1\left\{\underline{\tau}_0 \mid \underline{\tau}_0 < \underline{\tau}_1\right\} + (1-\overline{\beta}) E_1\left\{\overline{\tau}_1 \mid \overline{\tau}_1 < \overline{\tau}_0\right\}, \\ T_+^{(1)} = \overline{\beta} E_1\left\{\overline{\tau}_0 \mid \overline{\tau}_0 < \overline{\tau}_1\right\} + (1-\underline{\beta}) E_1\left\{\underline{\tau}_1 \mid \underline{\tau}_1 < \underline{\tau}_0\right\}. \end{array} \tag{12}$$

Proof. From the sequential test construction it follows that $\tau_0 \ne \tau_1$. First prove the inequality $\alpha \le \overline{\alpha}$. By the definition, we have $\alpha = P_0\{\tau_1 < \tau_0\}$, $\overline{\alpha} = P_0\{\overline{\tau}_1 < \overline{\tau}_0\}$. As $\overline{\Lambda}_n \ge \Lambda_n$, we get $\tau_0 \le \overline{\tau}_0$ and $\tau_1 \ge \overline{\tau}_1$. Hence,

$$\alpha = P_0\{\tau_1 < \tau_0\} \le P_0\{\overline{\tau}_1 < \overline{\tau}_0\} = \overline{\alpha}.$$

The second inequality for α and inequalities (10) for β are proved in the same way.

From the total probability formula we get

$$t^{(0)} = E_0\left\{\tau_0 \mid \tau_0 < \tau_1\right\}(1-\alpha) + E_0\left\{\tau_1 \mid \tau_1 < \tau_0\right\}\alpha. \tag{13}$$

From the definition of the stopping times τ_k, $\overline{\tau}_k$, $\underline{\tau}_k$, $k = 0, 1$, it follows that

$$\underline{\tau}_0 \le \tau_0 \le \overline{\tau}_0,\ \overline{\tau}_1 \le \tau_1 \le \underline{\tau}_1,$$

and

$$\{\overline{\tau}_0 < \overline{\tau}_1\} \subset \{\tau_0 < \tau_1\} \subset \{\underline{\tau}_0 < \underline{\tau}_1\},\ \{\underline{\tau}_1 < \underline{\tau}_0\} \subset \{\tau_1 < \tau_0\} \subset \{\overline{\tau}_1 < \overline{\tau}_0\}. \tag{14}$$

From the definition of the conditional error probabilities we get

$$\begin{array}{l} P_0\{\{\overline{\tau}_0 > \overline{\tau}_1\} \cap \{\tau_0 < \tau_1\}\} = \overline{\alpha} - \alpha, \\ P_0\{\{\tau_0 > \tau_1\} \cap \{\underline{\tau}_0 < \underline{\tau}_1\}\} = \alpha - \underline{\alpha}. \end{array} \tag{15}$$

Let us analyse summands in (13):

$$\begin{array}{c} E_0\{\tau_0 \mid \tau_0 < \tau_1\}(1-\alpha) = \\ (1-\alpha)(E_0\left\{\tau_0 \mid \{\tau_0 < \tau_1\} \cap \{\overline{\tau}_0 < \overline{\tau}_1\}\right\} P_0\left\{\overline{\tau}_0 < \overline{\tau}_1 \mid \tau_0 < \tau_1\right\} + \\ E_0\left\{\tau_0 \mid \{\tau_0 < \tau_1\} \cap \{\overline{\tau}_0 > \overline{\tau}_1\}\right\} P_0\left\{\overline{\tau}_0 > \overline{\tau}_1 \mid \tau_0 < \tau_1\right\}. \end{array}$$

Using (14), (15) we get

$$\begin{array}{l} E_0\{\tau_0 \mid \tau_0 < \tau_1\}(1-\alpha) = \\ E_0\left\{\tau_0 \mid \overline{\tau}_0 < \overline{\tau}_1\right\}(1-\overline{\alpha}) + E_0\left\{\tau_0 \mid \{\tau_0 < \tau_1\} \cap \{\overline{\tau}_0 > \overline{\tau}_1\}\right\}(\overline{\alpha} - \alpha). \end{array} \tag{16}$$

Because of the finiteness of the integral (see Wald, 1947)

$$\int_{\tau_0<\tau_1} \tau_0 P_0(d\omega) < \infty,$$

and because of the "shrinking" probabilities (15) at $\delta \to 0$, we have from (16):

$$E_0\{\tau_0 \mid \tau_0 < \tau_1\}(1-\alpha) \le E_0\left\{\overline{\tau}_0 \mid \overline{\tau}_0 < \overline{\tau}_1\right\}(1-\overline{\alpha}) + o(1),\ \delta \to 0. \tag{17}$$

An upper bound for the second summand in (13) is constructed in the same way:

$$E_0\left\{\tau_1 \mid \tau_1 < \tau_0\right\}\alpha = \alpha(E_0\left\{\tau_1 \mid \{\tau_1 < \tau_0\} \cap \{\underline{\tau}_1 < \underline{\tau}_0\}\right\} P_0\left\{\underline{\tau}_1 < \underline{\tau}_0 \mid \tau_1 < \tau_0\right\} + E_0\left\{\tau_1 \mid \{\tau_1 < \tau_0\} \cap \{\underline{\tau}_1 > \underline{\tau}_0\}\right\} P_0\left\{\underline{\tau}_1 > \underline{\tau}_0 \mid \tau_1 < \tau_0\right\}) =$$

$$E_0\left\{\tau_1 \mid \underline{\tau}_1 < \underline{\tau}_0\right\}\underline{\alpha} + o(1),\ \delta \to 0.$$

Finally,

$$E_0\left\{\tau_1 \mid \tau_1 < \tau_0\right\}\alpha \le E_0\left\{\underline{\tau}_1 \mid \underline{\tau}_1 < \underline{\tau}_0\right\}\underline{\alpha} + o(1). \tag{18}$$

Considering (17) and (18) together, we come to the upper bound (11) for $t^{(0)}$.

The lower bound for $t^{(0)}$, and both bounds for $t^{(1)}$ are proved with the same scheme. ■

Theorem 3 *If the conditions of Theorem 2 hold, then the following inequalities are satisfied:*

$$\begin{gathered} t^{(0)} \le (1-\underline{\alpha})E_0\left\{\underline{\tau}_1 \mid \underline{\tau}_0 < \underline{\tau}_1\right\} + \overline{\alpha}E_0\left\{\overline{\tau}_0 \mid \overline{\tau}_1 < \overline{\tau}_0\right\}, \\ t^{(1)} \le (1-\overline{\beta})E_1\left\{\overline{\tau}_0 \mid \overline{\tau}_0 > \overline{\tau}_1\right\} + \underline{\beta}E_1\left\{\underline{\tau}_1 \mid \underline{\tau}_1 > \underline{\tau}_0\right\}. \end{gathered} \tag{19}$$

Proof. Analyse the first summand in (13):

$$\begin{gathered} E_0\left\{\tau_0 \mid \tau_0 < \tau_1\right\}(1-\alpha) = \tfrac{1}{1-\alpha} \cdot \textstyle\int_{\tau_0<\tau_1} \tau_0 P_0(d\omega) \cdot (1-\alpha) \le \\ \tfrac{1-\underline{\alpha}}{1-\underline{\alpha}} \cdot \textstyle\int_{\underline{\tau}_0<\underline{\tau}_1} \underline{\tau}_1 P_0(d\omega) = (1-\underline{\alpha}) \cdot E_0\left\{\underline{\tau}_1 \mid \underline{\tau}_0 < \underline{\tau}_1\right\}. \end{gathered} \tag{20}$$

Considering the second summand in (13), we get:

$$\begin{gathered} E_0\left\{\tau_1 \mid \tau_1 < \tau_0\right\}\alpha = \tfrac{1}{\alpha} \cdot \textstyle\int_{\tau_1<\tau_0} \tau_1 P_0(d\omega) \cdot \alpha \le \\ \tfrac{\overline{\alpha}}{\overline{\alpha}} \cdot \textstyle\int_{\overline{\tau}_1<\overline{\tau}_0} \overline{\tau}_0 P_0(d\omega) = \overline{\alpha} \cdot E_0\left\{\overline{\tau}_0 \mid \overline{\tau}_1 < \overline{\tau}_0\right\}. \end{gathered} \tag{21}$$

From (13), (20), (21) we get the first inequality of (19). The second inequality is proved in the same way. ■

The upper bounds (19) can be calculated using the theory of denumerable Markov chains (Kemeni et al., 1966). Although the result of Theorem 3 is not asymptotic, we recommend to use in practice the main terms (12) of the asymptotic expansions (11), because the inequalities in (19) are "rough".

The next result gives the explicit expressions for the conditional ESSs under assumption (6).

Theorem 4 *If under conditions (1), (2), (6)* $|S^{(k)}| \ne 0$, $k = 0, 1$, *then for the test (4)*

$$\begin{gathered} E_k\left\{\tau_0 \mid \tau_0 < \tau_1\right\} = \left(\pi^{(k)}_{\bar{C}_-} + (\pi^{(k)})'(S^{(k)})^{-1}R^{(k)}_{\bar{C}_-}\right)^{-1}(\pi^{(k)})'(S^{(k)})^{-2}R^{(k)}_{\bar{C}_-} + 1, \\ E_k\left\{\tau_1 \mid \tau_1 < \tau_0\right\} = \left(\pi^{(k)}_{\bar{C}_+} + (\pi^{(k)})'(S^{(k)})^{-1}R^{(k)}_{\bar{C}_+}\right)^{-1}(\pi^{(k)})'(S^{(k)})^{-2}R^{(k)}_{\bar{C}_+} + 1. \end{gathered} \tag{22}$$

Proof. Define a random sequence

$$\xi_n = \bar{C}_+ \cdot \mathbf{1}_{[C_+,+\infty)}(\Lambda_n) + \bar{C}_- \cdot \mathbf{1}_{(-\infty,C_-]}(\Lambda_n) + \frac{\Lambda_n}{a} \cdot \mathbf{1}_{(C_-,C_+)}(\Lambda_n) \in \mathbf{Z}, n \in \mathbf{N}.$$

Then ξ_n is a homogeneous Markov chain with N states, two of them, $\bar{C}_-$ and $\bar{C}_+$, being absorbing ones. If the hypothesis H_k, $k = 0, 1$, is true, the matrix of transition probabilities of ξ_n is given by

$$P^{(k)} = (p_{ij}^{(k)}) = \left(\begin{array}{c|c} \mathbf{I}_2 & 0_{2\times N-2} \\ \hline R^{(k)} & Q^{(k)} \end{array} \right),$$

where $\mathbf{I}_2$ is the (2×2)–identity matrix and $0_{2\times N-2}$ is $(2\times N-2)$–matrix with all elements equal to 0. The vector of initial probabilities of nonabsorbing states of ξ_n equals to $\pi^{(k)}$ and initial probabilities of absorbing states equal to $\pi^{(k)}_{\bar{C}_-}$, $\pi^{(k)}_{\bar{C}_+}$. The rest of the proof follows from the finite Markov chains theory (Kemeni and Snell, 1959). ■

Theorems 2, 4 give a possibility to evaluate lower and upper bounds of the characteristics of the test (4). We construct functions $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot)$ that approximate $\lambda_W(\cdot)$ and satisfy (6) and (9). Then we calculate characteristics of the tests based on $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot)$ using (7) and get upper and lower bounds via (10) and (12).

We also can use the results of Theorems 2, 4 to evaluate the accuracy of existing estimates obtained via approximation of $\lambda_W(\cdot)$. Suppose we have already the approximations $\tilde{\alpha}$, $\tilde{\beta}$ of the error probabilities obtained using the function $\tilde{\lambda}(\cdot)$ that satisfies (6) and approximates $\lambda_W(\cdot)$ with an accuracy δ. Considering (6), let us denote $m^* = \max\{m \in \mathbf{Z} : a/m \geq \delta\}$, $a_\delta = a/m^*$, and define the functions $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot)$:

$$\underline{\lambda}(u) = \begin{cases} \tilde{\lambda}(u), & \text{if } \tilde{\lambda}(u) \leq \lambda_W(u), \\ \tilde{\lambda}(u) - a_\delta, & \text{otherwise}, \end{cases} \quad \overline{\lambda}(u) = \begin{cases} \tilde{\lambda}(u), & \text{if } \tilde{\lambda}(u) \geq \lambda_W(u), \\ \tilde{\lambda}(u) + a_\delta, & \text{otherwise}. \end{cases} \tag{23}$$

Note that the functions (23) satisfy (6) with $a = a_\delta$.

Corollary 1 *If functions $\underline{\lambda}(\cdot)$, $\overline{\lambda}(\cdot)$ are defined by (23), then the following inequalities hold for the two tests of the type (4) based on these functions:*

$$|\tilde{\alpha} - \alpha| \leq \overline{\alpha} - \underline{\alpha}, \; |\tilde{\beta} - \beta| \leq \underline{\beta} - \overline{\beta}.$$

Proof. As $\underline{\lambda}(u) \leq \lambda_W(u) \leq \overline{\lambda}(u)$, and $\underline{\lambda}(u) \leq \tilde{\lambda}(u) \leq \overline{\lambda}(u)$, $\forall u \in U$, from Theorem 2 we get that $\underline{\alpha} \leq \alpha \leq \overline{\alpha}$ and $\underline{\alpha} \leq \tilde{\alpha} \leq \overline{\alpha}$. Hence, $|\tilde{\alpha} - \alpha| \leq \overline{\alpha} - \underline{\alpha}$. The second inequality is proved similarly. ■

Let us note that the asymptotics $\overline{\alpha} - \underline{\alpha} \to 0$, $\underline{\beta} - \overline{\beta} \to 0$ is reached by taking $\delta \to 0$.

The accuracy of the estimates of the conditional ESSs can be evaluated in the same way.

4 Minimax Robust Sequential Test

A family of modified sequential tests is proposed to robustify (4) in (Kharin, 2002):

$$d = \mathbf{1}_{[C_+,+\infty)}(\Lambda_n^g) + 2 \cdot \mathbf{1}_{(C_-,C_+)}(\Lambda_n^g), \tag{24}$$

$$\Lambda_n^g = \sum_{t=1}^{n} g(\lambda(x_t)),$$

where

$$g(z) = g_- \cdot \mathbf{1}_{(-\infty,g_-]}(z) + g_+ \cdot \mathbf{1}_{[g_+,+\infty)}(z) + z \cdot \mathbf{1}_{(g_-,g_+)}(z),$$

and $g_-, g_+ \in \mathbf{R}$, $g_- < g_+$, are parameters of (24).

The minimax robust test is defined as the solution of the extremal problem:

$$\begin{cases} w_0\pi_0\alpha(g(\cdot)) + w_1\pi_1\beta(g(\cdot)) \to \min_{g(\cdot)} \\ \pi_0 t^{(0)}(g(\cdot)) + \pi_1 t^{(1)}(g(\cdot)) \le C(\pi_0 t^{(0)} + \pi_1 t^{(1)}), \end{cases} \tag{25}$$

where $C > 1$ is a parameter, $w_0, w_1 \ge 0$ are losses caused by the errors of type I and II respectively, π_0, π_1 are known prior probabilities of the hypotheses, $\alpha(g(\cdot))$, $\beta(g(\cdot))$, $t^{(0)}(g(\cdot))$, $t^{(1)}(g(\cdot))$ are the error probabilities and conditional ESS of the test (24) under the least favorable contaminating distribution, that is given in (Kharin, 2002).

Note that the test (24) can be treated as the test (4) based on $\lambda(u) = g(\lambda_W(u))$. Hence, if $\lambda(\cdot)$ satisfies (6), Theorem 1 holds and the problem (25) can be solved numerically by iterating through all possible values of g_-, g_+.

If the assumption (6) doesn't hold, we can replace the characteristics $\alpha(g(\cdot))$, $\beta(g(\cdot))$, $t^{(0)}(g(\cdot))$, $t^{(1)}(g(\cdot))$, $t^{(0)}$, $t^{(1)}$ in (25) with their upper bounds using the results given by Theorem 2 and Corollary 1. The solution of the constructed extremal problem approximates the solution of the problem (25) according to the theory presented in Section 3, and can be found numerically.

5 Numerical Results

To illustrate the theoretical results we performed computer modelling. The case of the observed sequence of random vectors

$$x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}), x_{ij} \in N_4 = \{1, 2, 3, 4\},$$
$$M = 1024, U = N_4^5,$$

was considered.

The hypotheses H_0, H_1 were formed by the expressions:

$$P_0(u) = 1/4^5, \quad u \in U,$$
$$P_1\{x_{ij} = 1\} = 0.4,\ j = 1, 2, 3,$$
$$P_1\{x_{ij} = k\} = 0.2,\ j = 1, 2, 3,\ k \in \{2, 3, 4\},$$
$$P_1\{x_{ij} = k\} = 0.25,\ j = 4, 5,\ k \in N_4.$$

The thresholds C_- and C_+ for the test (4) were calculated using the Wald formulae:

$$C_+ = \log \frac{1-\beta_0}{\alpha_0}, C_- = \log \frac{\beta_0}{1-\alpha_0},$$

where α_0 and β_0 are the so-called "desired" conditional error probabilities (they are maximal possible values of the error probabilities of type I and II).

For a number of pairs of α_0, β_0 the estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{t}^{(0)}$, $\hat{t}^{(1)}$ of the error probabilities and the conditional ESSs were obtained using Monte-Carlo method. For each set of parameters $N = 1000000$ experiments were done. Also for each pair of α_0, β_0 the lower and upper bounds $\underline{\alpha}$, $\overline{\beta}$, $T_-^{(0)}$, $T_-^{(1)}$, $\overline{\alpha}$, $\underline{\beta}$, $T_+^{(0)}$, $T_+^{(1)}$, were calculated using (10) and (12), (22). The bounds were obtained using the approximation with the parameters $a = \delta = 0.005$. The results of computer modelling are presented in Tables 1 and 2. The cases where the Monte-Carlo estimate of the conditional ESS was out of the correspondent interval are marked with the bold font.

Table 1: Conditional error probabilities for the hypothetical model

α_0	β_0	$\underline{\alpha}$	$\hat{\alpha}$	$\overline{\alpha}$	$\overline{\beta}$	$\hat{\beta}$	$\underline{\beta}$
0.01	0.01	0.00630	0.00662	0.00749	0.00660	0.00731	0.00766
0.01	0.05	0.00646	0.00686	0.00763	0.03655	0.03960	0.04070
0.01	0.10	0.00671	0.00707	0.00781	0.07402	0.07706	0.07974
0.05	0.05	0.03527	0.03656	0.03868	0.03693	0.03907	0.04077
0.05	0.10	0.03673	0.03824	0.03949	0.07368	0.07632	0.07880
0.10	0.10	0.07418	0.07814	0.07938	0.07290	0.07614	0.07713

Table 2: Conditional ESSs for the hypothetical model

α_0	β_0	$T_-^{(0)}$	$\hat{t}^{(0)}$	$T_+^{(0)}$	$T_-^{(1)}$	$\hat{t}^{(1)}$	$T_+^{(1)}$
0.01	0.01	32.2909	32.5822	33.0935	30.0705	30.5502	30.9739
0.01	0.05	21.1276	21.3349	21.7386	28.3925	28.6475	28.9499
0.01	0.10	16.6312	16.7791	17.0242	26.4813	26.7070	26.8189
0.05	0.05	19.7907	19.9744	20.1483	18.7114	18.9015	18.9374
0.05	0.10	15.4720	15.5673	15.6722	17.2044	**17.1832**	17.2227
0.10	0.10	14.2885	14.2113	14.3149	13.2700	**13.1961**	13.3513

These results show that the lower and upper bounds of the characteristics of the test are precise enough. The upper bounds $\overline{\alpha}$, $\overline{\beta}$ of the error probabilities can be successfully used instead of Wald's estimates α_0 and β_0, when the thresholds C_-, C_+ are given. The main terms $T_\pm^{(k)}$ calculated according to (12) could also be used in practice.

Similar experiments were performed for the case where the hypothetical model (1) is distorted according to (5), $\pi_0 = \pi_1 = \frac{1}{2}$, $w_0 = w_1 = 1$, $C = 10$, and the minimax robust sequential test (24) with $g_- = -0.231158$, $g_+ = 0.241208$. The levels of contamination

were equal to $\varepsilon_0 = \varepsilon_{0+} = \varepsilon_1 = \varepsilon_{1+} = 0.1$. Contaminating distributions were chosen to be the least favorable ones (for construction see Kharin, 2002):

$$\tilde{P}_0\{x_{ij} = 1\} = 1,\ j = 1,2,3,$$
$$\tilde{P}_0\{x_{ij} = k\} = 0,\ j = 1,2,3,\ k \in \{2,3,4\},\ \tilde{P}_0\{x_{ij} = k\} = 0.25,\ j = 4,5,\ k \in N_4,$$

$$\tilde{P}_1\{x_{ij} = 2\} = 1,\ j = 1,2,3,$$
$$\tilde{P}_1\{x_{ij} = k\} = 0,\ j = 1,2,3,\ k \in \{1,3,4\},\ \tilde{P}_1\{x_{ij} = k\} = 0.25,\ j = 4,5,\ k \in N_4.$$

The results are presented in Tables 3 and 4. For comparison, the estimates of the SPRT (4) characteristics are also given in the tables. The results show the robustness of the test (24).

Table 3: Conditional error probabilities for the distorted model

α_0	β_0	$\hat{\alpha}$	$\underline{\alpha}(g)$	$\hat{\alpha}(g)$	$\overline{\alpha}(g)$	$\hat{\beta}$	$\overline{\beta}(g)$	$\hat{\beta}(g)$	$\underline{\beta}(g)$
0.01	0.01	0.24	0.0015	0.0034	0.0059	0.11	0.0049	0.0082	0.0185
0.01	0.05	0.19	0.0016	0.0035	0.0060	0.22	0.0301	0.0427	0.0722
0.01	0.10	0.18	0.0017	0.0037	0.0061	0.30	0.0662	0.0863	0.1298
0.05	0.05	0.31	0.0147	0.0246	0.0348	0.20	0.0309	0.0427	0.0709
0.05	0.10	0.28	0.0154	0.0249	0.0353	0.27	0.0675	0.0862	0.1272
0.10	0.10	0.36	0.0401	0.0581	0.0753	0.24	0.0690	0.0867	0.1253

Table 4: Conditional ESSs for the distorted model

α_0	β_0	$\hat{t}^{(0)}$	$T_-^{(0)}(g)$	$\hat{t}^{(0)}(g)$	$T_+^{(0)}(g)$	$\hat{t}^{(1)}$	$T_-^{(1)}(g)$	$\hat{t}^{(1)}(g)$	$T_+^{(1)}(g)$
0.01	0.01	60.2	197.5	225.0	244.7	50.7	241.7	263.4	304.9
0.01	0.05	38.6	130.1	147.6	160.0	40.1	235.0	246.7	270.3
0.01	0.10	30.2	101.0	114.2	123.4	33.5	222.7	228.0	240.5
0.05	0.05	28.0	127.3	139.8	146.5	25.7	153.7	158.9	170.5
0.05	0.10	21.1	98.4	107.0	111.5	21.3	143.7	144.3	148.7
0.10	0.10	16.9	92.9	98.0	99.2	16.0	108.8	**108.2**	110.1

Acknowledgement

The authors are grateful to the anonymous referee for the useful comments which have helped to present the results in a more correct and clear form.

References

P. Bauer and J. Röhmel. An adaptive method for establishing a dose response relationship. *Statistics in Medicine*, 14:1595–1607, 1995.

R. Durbin. *Biological Sequence Analysis*. Cambridge University Press, Cambridge, 1998.

P. Huber. *Robust Statistics*. Wiley, New York, 1981.

J.G. Kemeni and J.L. Snell. *Finite Markov Chains*. Wiley, New York, 1959.

J.G. Kemeni, J.L. Snell, and A.W. Knapp. *Denumerable Markov Chains*. Van Nostrand, Princeton, 1966.

A. Kharin. On robustifying of the sequential probability ratio test for a discrete model under "contaminations". *Austrian Journal of Statistics*, 31(4):267–277, 2002.

A. Kharin. On application of robust sequential discrimination under distortions. In G. Medvedev, editor, *Probability Theory, Mathematical Statistics and Applications*, pages 203–205. Belarusian State University, Minsk, 2004.

T.L. Lai. Sequential analysis: Some classical problems and new challenges. *Statistica Sinica*, 11:303–408, 2001.

R.L. Mason and J.C. Young. *Multivariate Statistical Process Control with Industrial Applications*. ASA-SIAM, Alexandria, 2002.

P.X. Quang. Robust sequential testing. *Annals of Statistics*, 13(2):638–649, 1985.

H. Rieder. *Robust Asymptotic Statistics*. Springer–Verlag, New York, 1994.

D. Siegmund. *Sequential Analysis. Tests and Confidence Intervals*. Springer–Verlag, New York, 1985.

A. Wald. *Sequential Analysis*. John Wiley and Sons, New York, 1947.

J. Whitehead. *The Design and Analysis of Sequential Clinical Trials*. John Wiley and Sons, New York, 1997.

Corresponding author's address:

Assoc. Prof. Dr. Alexey Kharin
Department of Probability Theory and Mathematical Statistics
Belarusian State University
Fr. Skoriny av. 4
220050 Minsk
Belarus

Tel. +375 17 2095129
Fax +375 17 2095054
E-mail: KharinAY@bsu.by