



# An Approach to Robustness Evaluation for Sequential Testing under Functional Distortions in $L_1$ -metric

Alexey Kharin

Belarusian State University

Sergey Chernov

Belarusian State University

---

## Abstract

The problem of sensitivity analysis for the sequential probability ratio test under functional distortions of the observation probability distribution is considered. For the situation where distorted densities of the log likelihood ratio statistic belong to  $\varepsilon$ -neighborhoods of hypothetical centers in the  $L_1$ -metric the least favorable distributions that maximize the conditional error probabilities are constructed. The instability coefficient is obtained to enable robustness evaluation for the sequential probability ratio test and its modification – trimmed sequential probability ratio test.

*Keywords:* sequential probability ratio test, error probability, distortion,  $L_1$ -metric, least favorable distribution, instability coefficient, robustness.

---

## 1. Introduction

The sequential approach to hypothesis testing (Wald 1947) is applied in various practical problems of statistical data analysis (Mukhopadhyay and de Silva 2009). If hypothetical suppositions are fulfilled, sequential tests require less observations at average in comparison with classical analogues based on the fixed number of observations, to provide the fixed small levels of error probabilities. However, in practice there are distortions in statistical data, i.e. the factual probability distribution of observations deviate from the hypothetical model (Kharin and Voloshko 2011). Therefore it is important to characterize the influence of the distortions on the error probabilities.

Similar problems of robustness analysis were investigated in Kharin (2002), Kharin and Kishylau (2005), Kharin (2013a) for discrete data under “contamination” (Huber and Ronchetti 2009). The problems of robustness analysis and of robust decision rules construction for case of composite hypotheses are investigated in Kharin (2008), Kharin (2011a) using the methodology of the asymptotic expansion construction for the characteristics w.r.t. the small parameter of distortion developed in Kharin and Shlyk (2009), Kharin (2005).

In Chernov and Kharin (2013) error probabilities of the sequential probability ratio test (SPRT) under functional distortions described by neighborhoods in the  $L_2$ -metric were studied.

In this paper we consider the case of continuous probability distribution of observations and analyze the influence of the distortions in the  $L_1$ -metric on the error probabilities of the SPRT. For a given maximal possible distance between the factual and the hypothetical probability distributions of the log likelihood ratio statistic the least favorable distributions (LFD) that maximize the conditional error probability of the SPRT are constructed. This maximal value of the error probability is required for the quantitative robustness analysis of sequential tests.

## 2. Mathematical Model

Consider the mathematical model from [Kharin and Chernov \(2011\)](#). Let  $x_1, x_2, \dots \in \mathbf{R}$  be independent and identically distributed random observations on a probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Let  $f(x, \theta)$  be the probability density function (p.d.f.) of  $x_i$ ,  $i \in \mathbf{N} = \{1, 2, \dots\}$ , with a parameter  $\theta \in \Theta = \{\theta_0, \theta_1\}$ ;  $F(x, \theta)$  be the cumulative distribution function that corresponds to  $f(x, \theta)$ .

There are two simple hypotheses concerning the unknown value of the parameter  $\theta$ :

$$\mathcal{H}_0 : \theta = \theta_0, \quad \mathcal{H}_1 : \theta = \theta_1. \quad (1)$$

Denote the accumulated log likelihood ratio test statistic:

$$\Lambda_n = \Lambda_n(x_1, \dots, x_n) = \sum_{k=1}^n \lambda_k, \quad (2)$$

where

$$\lambda_k = \lambda(x_k) = \ln \frac{f(x_k, \theta_1)}{f(x_k, \theta_0)} \quad (3)$$

is the logarithm of the likelihood ratio statistic calculated for the observation  $x_k$ ,  $k \in \mathbf{N}$ .

To test hypotheses (1) by observations  $x_1, x_2, \dots$  the SPRT ([Wald 1947](#)) can be used:

$$N = \min\{n \in \mathbf{N} : \Lambda_n \notin (C_-, C_+)\}, \quad (4)$$

$$d = \begin{cases} 0, & \Lambda_N \leq C_- \\ 1, & \Lambda_N \geq C_+ \end{cases}, \quad (5)$$

where  $N$  is the random stopping time; at this time point the decision  $d$  is made according to (5). In (4) the parameters  $C_-, C_+ \in \mathbf{R}$  are the test thresholds defined according to [Wald \(1947\)](#):

$$C_- = \ln \frac{\beta_0}{1 - \alpha_0}, \quad C_+ = \ln \frac{1 - \beta_0}{\alpha_0}, \quad (6)$$

where  $\alpha_0, \beta_0 \in (0, \frac{1}{2})$  are given maximal admissible values of probabilities of type I (to accept  $\mathcal{H}_1$  provided  $\mathcal{H}_0$  is true) and II (acceptance of  $\mathcal{H}_0$  provided the true hypothesis is  $\mathcal{H}_1$ ) errors respectively.

Let  $\alpha(f)$  and  $\beta(f)$  be the error probabilities of the test (4), (5) for the case where observations  $x_1, x_2, \dots$  have the probability density function  $f(\cdot)$ .

It is known that  $\alpha_0$  and  $\beta_0$  are only approximate values of the factual error probabilities  $\alpha(f)$  and  $\beta(f)$  of types I and II for the SPRT (4) – (6) (see [Wald 1947](#)) and can deviate from  $\alpha(f)$  and  $\beta(f)$  significantly ([Kharin 2013a](#)).

Without loss of generality, suppose that the hypothesis  $\mathcal{H}_0$  is true, so the value of the type I error probability  $\alpha$  is considered. To make formulation shorter, introduce the simplified notation:

$$F(x) = F(x, \theta_0), \quad f(x) = f(x, \theta_0), \quad F_\lambda(x) = \mathbf{P}_{\mathcal{H}_0}\{\lambda_1 \leq x\},$$

where  $\mathbf{P}_{\mathcal{H}_0}\{\cdot\}$  means the probability under the hypothesis  $\mathcal{H}_0$ . Let the probability density function  $p_\lambda(x)$  corresponds to the cumulative distribution function  $F_\lambda(x)$ .

### 3. Inequalities for Error Probabilities of the SPRT

Let  $x(\omega)$  and  $y(\omega)$  be random variables on some probability space  $(\Omega, \mathcal{F}, P)$  with some probability density functions  $a(x)$  and  $b(y)$  respectively; let also  $1_A(\cdot)$  be the indicator function of the set  $A$ .

**Lemma 1** *If the inequality  $\lambda(x(\omega)) \geq \lambda(y(\omega))$  is satisfied for every  $\omega \in \Omega$ , then the inequality*

$$\alpha(a) \geq \alpha(b)$$

*takes place.*

**Proof.** It follows from the Lemma condition that

$$\Lambda_n(a) = \sum_{k=1}^n \lambda(x_k) \geq \sum_{k=1}^n \lambda(y_k) = \Lambda_n(b). \quad (7)$$

From (5) we have

$$\alpha(a) = P_{\mathcal{H}_0} \{ \Lambda_N(a) \geq C_+ \},$$

where  $N$  is the random stopping time. Because of (7) we get the relation between the random events:

$$\{ \Lambda_N(a) \geq C_+ \} \supseteq \{ \Lambda_N(b) \geq C_+ \},$$

therefore,  $\alpha(a) \geq \alpha(b)$ . ■

**Lemma 2** *If the inequality  $\lambda(x) \geq \lambda(y)$  is satisfied for*

$$x \in M_{a>b} = \{z : a(z) > b(z)\}, \quad y \in M_{b \geq a} = \mathbf{R} \setminus M_{a>b},$$

*then the inequality  $\alpha(a) \geq \alpha(b)$  holds.*

**Proof.** From the norm conditions for  $a(\cdot)$ ,  $b(\cdot)$  we have:

$$\int_{M_{a>b}} a(x) dx + \int_{M_{b \geq a}} a(x) dx \equiv 1 \equiv \int_{M_{b \geq a}} b(x) dx + \int_{M_{a>b}} b(x) dx.$$

Using these equations denote

$$p = \int_{M_{a>b}} (a(x) - b(x)) dx = \int_{M_{b \geq a}} (b(x) - a(x)) dx \in [0, 1].$$

Note that if  $p = 0$ , then  $a(\cdot)$  and  $b(\cdot)$  coincide, if  $p = 1$ , they are orthogonal in the sense that  $a(x)b(x) = 0, \forall x$ .

Let  $\eta = \eta(\omega)$  be the Bernoulli random variable with the parameter value  $p$ :

$$P\{\eta = 1\} = p, \quad P\{\eta = 0\} = 1 - p;$$

$\xi = \xi(\omega)$ ,  $\xi^+ = \xi^+(\omega)$  and  $\xi^- = \xi^-(\omega)$  be random variables with the p.d.f.s

$$\begin{aligned} p_\xi(x) &= \frac{\min\{a(x), b(x)\}}{1 - p}, \\ p_{\xi^+}(x) &= \frac{1_{M_{a>b}}(x)(a(x) - b(x))}{p}, \\ p_{\xi^-}(x) &= \frac{1_{M_{b \geq a}}(x)(b(x) - a(x))}{p}, \end{aligned} \quad (8)$$

respectively, and  $\eta$ ,  $\xi$ ,  $\xi^+$ ,  $\xi^-$  be independent.

The norm condition is satisfied for the functions determined by (8):

$$\begin{aligned} \int_{-\infty}^{+\infty} p_{\xi}(x)dx &= \frac{1}{1-p} \left( \int_{M_{a>b}} b(x)dx + \int_{M_{b\geq a}} a(x)dx \right) = \\ &= \frac{1}{1-p} \left( 1 - \int_{M_{b\geq a}} b(x)dx + \int_{M_{b\geq a}} a(x)dx \right) = \\ &= \frac{1}{1-p} \left( 1 - \int_{M_{b\geq a}} (b(x) - a(x)) dx \right) = \frac{1}{1-p}(1-p) \equiv 1; \\ \int_{-\infty}^{+\infty} p_{\xi^+}(x)dx &= \frac{1}{p} \int_{M_{a>b}} (a(x) - b(x)) dx \equiv 1; \\ \int_{-\infty}^{+\infty} p_{\xi^-}(x)dx &= \frac{1}{p} \int_{M_{b\geq a}} (b(x) - a(x)) dx \equiv 1. \end{aligned}$$

The p.d.f.s  $p_{\xi^+}(\cdot)$  and  $p_{\xi^-}(\cdot)$  are orthogonal, and  $\xi^-(\omega) \geq \xi^+(\omega)$ ,  $\omega \in \Omega$ .

Construct random variables  $\xi_a = \xi_a(\omega)$ ,  $\xi_b = \xi_b(\omega)$  on  $(\Omega, \mathcal{F}, \mathbb{P})$ :

$$\xi_a(\omega) = (1 - \eta(\omega))\xi(\omega) + \eta(\omega)\xi^+(\omega), \quad \xi_b(\omega) = (1 - \eta(\omega))\xi(\omega) + \eta(\omega)\xi^-(\omega). \quad (9)$$

The p.d.f.s of random variables (9) can be found by (8):

$$\begin{aligned} p_{\xi_a}(x) &= p \cdot p_{\xi^+}(x) + (1-p) \cdot p_{\xi}(x) = \\ &= \frac{1-p}{1-p} \cdot \min\{a(x), b(x)\} + \frac{p}{p} \cdot 1_{M_{a>b}}(x) \cdot (a(x) - b(x)) = \\ &= \begin{cases} b(x) + a(x) - b(x), & \text{if } a(x) > b(x), \\ a(x) + 0, & \text{if } a(x) \leq b(x), \end{cases} \equiv a(x). \end{aligned} \quad (10)$$

Analogously we get

$$p_{\xi_b}(x) = p \cdot p_{\xi^-}(x) + (1-p) \cdot p_{\xi}(x) \equiv b(x). \quad (11)$$

From the construction of  $\xi^-$ ,  $\xi^+$  and the condition of this Lemma it follows that  $\lambda(\xi^+) \geq \lambda(\xi^-)$ .

Analyze now the two available cases using (9).

1. If  $\omega$ :  $\eta(\omega) = 1$ , then  $\xi_a(\omega) = \xi^+$ ,  $\xi_b(\omega) = \xi^-$ .

2. If  $\omega$ :  $\eta(\omega) = 0$ , then  $\xi_a(\omega) = \xi_b(\omega) = \xi(\omega)$ .

Combining these two results, we have  $\lambda(\xi_a) \geq \lambda(\xi_b)$ ,  $\forall \omega \in \Omega$ .

Finally, using Lemma 1 we get

$$\alpha(p_{\xi_a}) \geq \alpha(p_{\xi_b}),$$

that is equivalent to  $\alpha(a) \geq \alpha(b)$  because of (10), (11). ■

#### 4. Robustness Evaluation for SPRT

Let the hypothetical model described in Section 1 be not satisfied, so the log likelihoods  $\lambda_n = \lambda(x_n)$ ,  $n \in \mathbf{N}$ , are independent and identically distributed random variables with some p.d.f.  $\tilde{p}_{\lambda}(x)$ , that may deviate from the hypothetical p.d.f.  $p_{\lambda}(x)$ , but the distance between  $\tilde{p}_{\lambda}(x)$  and  $p_{\lambda}(x)$  in the  $L_1$ -metric does not exceed  $\varepsilon$ :

$$\rho_{L_1}(\tilde{p}_{\lambda}(\cdot), p_{\lambda}(\cdot)) = \int_{\mathbf{R}} |\tilde{p}_{\lambda}(x) - p_{\lambda}(x)| dx \leq \varepsilon, \quad (12)$$

where  $0 \leq \varepsilon \leq \varepsilon_0$ , and the maximal admissible deviation  $\varepsilon_0$  is a priori known.

Denote by  $L_1(p_\lambda, \varepsilon)$  the family of probability density functions  $\tilde{p}_\lambda(x)$  that satisfy the inequality (12) for the fixed value of  $\varepsilon$ . Let the cumulative probability distribution function  $\tilde{F}_\lambda(x)$  corresponds to the p.d.f.  $\tilde{p}_\lambda(x)$ . Let  $\alpha(\tilde{p}_\lambda, \varepsilon)$  be the type I error probability for the SPRT (4), (5), when the log likelihood (3) has the p.d.f.  $\tilde{p}_\lambda(\cdot) \in L_1(p_\lambda, \varepsilon)$ .

Let us construct the least favorable probability distribution of  $\lambda_n$ , i.e. the p.d.f. that maximizes the value of  $\alpha(\cdot, \varepsilon)$  within the set  $L_1(p_\lambda, \varepsilon)$ .

Consider the p.d.f.

$$\bar{p}_\lambda(x) = 1_{(g^-, +\infty)}(x)p_\lambda(x) + \frac{\varepsilon}{2}\delta(x - g^+), \tag{13}$$

where  $\delta(\cdot)$  is the Dirac  $\delta$ -function,

$$g^+ = C_+ - C_-, \quad F_\lambda(g^-) = \frac{\varepsilon}{2}.$$

**Lemma 3** *The function  $\bar{p}_\lambda(\cdot)$  belongs to  $L_1(p_\lambda, \varepsilon)$ .*

**Proof.** Find  $\rho_{L_1}(\bar{p}_\lambda(\cdot), p_\lambda(\cdot))$  using (13):

$$\begin{aligned} \int_{\mathbf{R}} |\bar{p}_\lambda(x) - p_\lambda(x)| dx &= \int_{(-\infty, g^-)} p_\lambda(x) dx + \int_{(g^-, +\infty)} |\bar{p}_\lambda(x) - p_\lambda(x)| dx = \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \cdot \int_{(g^-, +\infty)} \delta(x - g^+) dx = \varepsilon. \end{aligned} \tag{14}$$

Lemma is proved. ■

Now let us prove that if the random variables  $\{\lambda_n\}$  have the p.d.f.  $\bar{p}_\lambda(x)$ , then the type I error probability  $\alpha(\bar{p}_\lambda)$  is the highest value within the neighborhood  $L_1(p_\lambda, \varepsilon)$ .

**Theorem 1** *If the p.d.f.  $\tilde{p}_\lambda(\cdot)$  belongs to  $L_1(p_\lambda, \varepsilon)$ , then the following inequality holds:*

$$\alpha(\tilde{p}_\lambda, \varepsilon) \leq \alpha(\bar{p}_\lambda, \varepsilon). \tag{15}$$

**Proof.** Take any p.d.f.  $\tilde{p}_\lambda(\cdot) \in L_1(p_\lambda, \varepsilon)$ . Denote as in Lemma 2:

$$\begin{aligned} p &= \int_{(\tilde{p}_\lambda > p_\lambda)} (\tilde{p}_\lambda(x) - p_\lambda(x)) dx = \int_{(\tilde{p}_\lambda \leq p_\lambda)} (p_\lambda(x) - \tilde{p}_\lambda(x)) dx; \\ \varepsilon^-(q) &= \int_{(-\infty, g^-)} q(y) dy, \end{aligned}$$

where  $q(\cdot)$  is some arbitrary p.d.f.

Note that  $p \leq \varepsilon/2$  and construct the auxiliary p.d.f.s  $q_1(\cdot)$  and  $q_2(\cdot)$ :

$$\begin{aligned} q_1(x) &= 1_{(\tilde{p}_\lambda \leq p_\lambda)}(x)\tilde{p}_\lambda(x) + 1_{(\tilde{p}_\lambda > p_\lambda)}(x)p_\lambda(x) + p \cdot \delta(x - g^+) = \\ &= 1_{(\tilde{p}_\lambda < p_\lambda)}(x)\tilde{p}_\lambda(x) + 1_{(\tilde{p}_\lambda \geq p_\lambda)}(x)p_\lambda(x) + p \cdot \delta(x - g^+), \\ q_2(x) &= 1_{(g^-, +\infty) \cap (\tilde{p}_\lambda < p_\lambda)}(x)\tilde{p}_\lambda(x) + 1_{(g^-, +\infty) \cap (\tilde{p}_\lambda \geq p_\lambda)}(x)p_\lambda(x) + \\ &= \varepsilon^-(q_1)\delta(x - g^-) + p \cdot \delta(x - g^+). \end{aligned} \tag{16}$$

The p.d.f.  $q_1(x)$  is constructed from  $\tilde{p}_\lambda(x)$  by “transferring” of the probability measure equals to  $p$  from the set  $\{\tilde{p}_\lambda > p_\lambda\}$  to the point  $\{g^+\}$ . The p.d.f.  $q_2(x)$  is constructed from  $q_1(x)$  by “transferring” of the probability measure (equaled to  $\varepsilon^-(q_1) = \int_{(-\infty, g^-)} q_1(y) dy$ ) from the sets  $\{\tilde{p}_\lambda < p_\lambda\} \cap (-\infty, g^-)$  and  $\{\tilde{p}_\lambda \geq p_\lambda\} \cap (-\infty, g^-)$  to the point  $\{g^-\}$ .

Compare the four error probabilities  $\alpha(\tilde{p}_\lambda)$ ,  $\alpha(q_1)$ ,  $\alpha(q_2)$  and  $\alpha(\bar{p}_\lambda)$  using (16). Consider the sets, where the mentioned p.d.f.s differ from each other:

$$\{x : \tilde{p}_\lambda(x) < q_1(x)\} \subseteq \{g^+\}, \quad \{x : \tilde{p}_\lambda(x) > q_1(x)\} \subseteq \{x : \tilde{p}_\lambda(x) > p_\lambda(x)\} \setminus \{g^+\},$$

$$\begin{aligned} \{x : q_1(x) < q_2(x)\} &\subseteq \{g^-\}, \quad \{x : q_1(x) > q_2(x)\} \subseteq (g^-, +\infty), \\ \{x : q_2(x) < \bar{p}_\lambda(x)\} &\subseteq ((g^-, +\infty) \cap \{x : \tilde{p}_\lambda(x) < p_\lambda(x)\}) \cup \{g^+\}, \\ \{x : q_2(x) > \bar{p}_\lambda(x)\} &\subseteq \{g^-\}. \end{aligned}$$

According to Lemma 2 we have inequalities

$$\alpha(\tilde{p}_\lambda, \varepsilon) \leq \alpha(q_1, \varepsilon) \leq \alpha(q_2, \varepsilon) \leq \alpha(\bar{p}_\lambda, \varepsilon).$$

Therefore, the inequality (15) holds. ■

**Corollary 1** *The error probability  $\alpha(\bar{p}_\lambda, \varepsilon)$  is a monotone function w.r.t. the neighborhood size  $\varepsilon$ , and  $\forall \varepsilon \in [0, \varepsilon_0]$  the following inequality holds:*

$$\alpha(\bar{p}_\lambda, \varepsilon) \leq \alpha(\bar{p}_\lambda, \varepsilon_0).$$

**Proof** follows from the result of Lemma 2. ■

Calculate now the instability coefficient  $\kappa$  (Kharin 2013b) that characterizes the relative increment of the type I error probability for the SPRT under distortion (12) from the hypothetical version:

$$\kappa = \frac{\alpha^+ - \alpha^0}{\alpha^0} \geq 0,$$

where

$$\alpha^0 = \alpha(p_\lambda), \quad \alpha^+ = \sup_{\tilde{p}_\lambda \in L_1(p_\lambda, \varepsilon), \varepsilon \in [0, \varepsilon_0]} \alpha(\tilde{p}_\lambda, \varepsilon).$$

**Corollary 2** *The instability coefficient for the error type I probability of the SPRT is equal to*

$$\kappa = \frac{\alpha(\bar{p}_\lambda, \varepsilon_0) - \alpha(p_\lambda)}{\alpha(p_\lambda)} \geq 0.$$

**Proof.** The result follows from Lemma 3, Theorem 1 and Corollary 1. ■

## 5. Robustness Evaluation for Trimmed SPRT

To decrease the influence of distortions on the error probabilities of the test (4), (5) we construct the trimmed probability density function  $p_\lambda(x)$  for the log likelihood (3) following the idea of Kharin (2002):

$$p_\lambda^g(x) = 1_{(g^-, g^+)} p_\lambda(x) + \varepsilon^- \delta(x - g^-) + \varepsilon^+ \delta(x - g^+), \quad (17)$$

where  $g^-, g^+ \in \mathbf{R}$ ,  $g^- < g^+$ , are some trimming parameters for  $\lambda_n$ ;

$$\varepsilon^- = \varepsilon^-(p_\lambda) = F_\lambda(g^-), \quad \varepsilon^+ = \varepsilon^+(p_\lambda) = 1 - F_\lambda(g^+). \quad (18)$$

Note that the function  $p_\lambda^g(x)$  defined by (17) is some probability density function as it is nonnegative and the norm condition holds:

$$\begin{aligned} \int_{\mathbf{R}} p_\lambda^g(y) dy &= \int_{\mathbf{R}} 1_{(g^-, g^+)} p_\lambda(y) dy + \int_{\mathbf{R}} \varepsilon^- \delta(y - g^-) dy + \int_{\mathbf{R}} \varepsilon^+ \delta(y - g^+) dy = \\ &= \int_{(g^-, g^+)} p_\lambda(y) dy + \varepsilon^- + \varepsilon^+ = (F_\lambda(g^+) - F_\lambda(g^-)) + F_\lambda(g^-) + (1 - F_\lambda(g^+)) = 1. \end{aligned}$$

The sequential test (4) – (6) constructed using the test statistic with the trimmed probability density function (17) instead of  $\lambda(\cdot)$  will be called the trimmed SPRT. If  $g^- = -\infty$  and  $g^+ = +\infty$ , then the trimmed p.d.f.  $p_\lambda^g(\cdot)$  coincides with  $p_\lambda(\cdot)$ , i.e. we have no trimming.

Prove now that if the p.d.f.  $\tilde{p}_\lambda(\cdot)$  belongs to the  $\varepsilon$ -neighborhood in the  $L_1$ -metric of the function  $p_\lambda(\cdot)$ , then the trimmed p.d.f.  $\tilde{p}_\lambda^g(x)$  belongs to the  $\varepsilon$ -neighborhood of the function  $p_\lambda^g(\cdot)$  in the same metric.

**Lemma 4** *If  $\tilde{p}_\lambda \in L_1(p_\lambda, \varepsilon)$ , then  $\tilde{p}_\lambda^g \in L_1(p_\lambda^g, \varepsilon)$ .*

**Proof.** Using (17), (18) evaluate the distance:

$$\begin{aligned} \int_{\mathbf{R}} |\tilde{p}_\lambda^g(x) - p_\lambda^g(x)| dx &= \int_{(g^-, g^+)} |\tilde{p}_\lambda(x) - p_\lambda(x)| dx + \\ |\varepsilon^-(\tilde{s}) - \varepsilon^-(s)| \cdot \int_{\mathbf{R}} \delta(x - g^-) dx &+ |\varepsilon^+(\tilde{s}) - \varepsilon^+(s)| \cdot \int_{\mathbf{R}} \delta(x - g^+) dx = \\ \int_{(g^-, g^+)} |\tilde{p}_\lambda(x) - p_\lambda(x)| dx &+ |\varepsilon^-(\tilde{s}) - \varepsilon^-(s)| + |\varepsilon^+(\tilde{s}) - \varepsilon^+(s)| = \\ \int_{(g^-, g^+)} |\tilde{p}_\lambda(x) - p_\lambda(x)| dx &+ \left| \int_{(-\infty, g^-)} (\tilde{p}_\lambda(x) - p_\lambda(x)) dx \right| + \\ \left| \int_{(g^+, +\infty)} (\tilde{p}_\lambda(x) - p_\lambda(x)) dx \right| &\leq \int_{\mathbf{R}} |\tilde{p}_\lambda(x) - p_\lambda(x)| dx \leq \varepsilon, \end{aligned}$$

that proves the statement of the Lemma. ■

Let us find now the least favorable probability distribution for the fixed parameters of trimming  $g^-$  and  $g^+$ , that maximizes the value of  $\alpha(\cdot, \varepsilon)$  within  $L_1(p_\lambda^g, \varepsilon)$ . In other words, let us prove that if  $\bar{p}_\lambda(\cdot)$  corresponds to the LFD in  $L_1(p_\lambda, \varepsilon)$ , then  $\bar{p}_\lambda^g(\cdot)$  corresponds to the LFD in  $L_1(p_\lambda^g, \varepsilon)$ .

If  $\bar{p}_\lambda(\cdot)$  satisfies (13), then  $\bar{p}_\lambda^g(\cdot)$ , constructed according to (17), is determined by the equation

$$\bar{p}_\lambda^g(x) = 1_{(g^-, g^+)}(x) p_\lambda(x) + \left( \varepsilon^- - \frac{\varepsilon}{2} \right) \delta(x - g^-) + \left( \varepsilon^+ + \frac{\varepsilon}{2} \right) \delta(x - g^+). \quad (19)$$

**Theorem 2** *If the probability density function  $\tilde{p}_\lambda(\cdot)$  belongs to the set  $L_1(p_\lambda, \varepsilon)$ , then the following inequality holds:*

$$\alpha(\tilde{p}_\lambda^g, \varepsilon) \leq \alpha(\bar{p}_\lambda^g, \varepsilon).$$

**Proof.** The Theorem statement follows from Lemma 4 and Theorem 1. ■

**Corollary 3** *The error probability  $\alpha(\bar{p}_\lambda^g, \varepsilon)$  is a monotone function w.r.t. the variable  $\varepsilon$ , and for every  $\varepsilon$ ,  $0 \leq \varepsilon \leq \varepsilon_0$ , the following inequality takes place:*

$$\alpha(\bar{p}_\lambda^g, \varepsilon) \leq \alpha(\bar{p}_\lambda^g, \varepsilon_0).$$

**Proof.** The Corollary statement follows from Lemma 4 and Theorem 1. ■

Now calculate the instability coefficient (Kharin 2011b) for the type I error probability of the SPRT under distortion (12).

**Corollary 4** *The instability coefficient for the error type I probability of the trimmed SPRT is equal to*

$$\kappa = \frac{\alpha(\bar{p}_\lambda^g, \varepsilon_0) - \alpha(p_\lambda^g)}{\alpha(p_\lambda^g)} \geq 0.$$

**Proof** follows from Lemma 4, Theorem 2 and Corollary 3. ■

## 6. Conclusions

The least favorable probability distributions of the log likelihood ratio statistic are constructed in the paper for the distortions in the  $L_1$ -metric. The obtained results are useful for evaluation of the difference between hypothetical and actual error probabilities under functional distortions in observation distributions, adjusted in the mentioned metric.

The results for the error type II probabilities are obtained in the same way.

The instability coefficient characterizes robustness of the SPRT and of the trimmed SPRT quantitatively.

The research is partially supported by the ISTC Project B-1910.

## References

- Chernov S, Kharin A (2013). “Error Probabilities for Sequential Testing of Simple Hypotheses Under Functional Distortions in the  $L_2$ -Metric.” *Statistical Methods of Estimation and Hypotheses Testing (in Russian)*, **25**, 64–72.
- Huber P, Ronchetti E (2009). *Robust Statistics*. Wiley, New York.
- Kharin A (2002). “On Robustifying of the Sequential Probability Ratio Test for a Discrete Model Under “Contaminations”.” *Austrian Journal of Statistics*, **31**(4), 267–277.
- Kharin A (2005). “Robust Bayesian Prediction Under Distrtions of Prior and Conditional Distributions.” *Journal of Mathematical Sciences*, **126**(1), 992–997.
- Kharin A (2008). “Robustness Evaluation in Sequential Testing of Composite Hypotheses.” *Austrian Journal of Statistics*, **37**(1), 51–60.
- Kharin A (2011a). “Robustness Analysis for Bayesian Sequential Testing of Composite Hypotheses Under Simultaneous Distortions of Priors and Likelihoods.” *Austrian Journal of Statistics*, **40**(1), 65–73.
- Kharin A (2013a). “Robustness of Sequential Testing of Hypotheses on Parameters of M-valued Random Sequences.” *Journal of Mathematical Sciences*, **189**(6), 924–931.
- Kharin A, Chernov S (2011). “Evaluation of the Error Probabilities for the Sequential Probability Ratio Test.” *Proc. of the Belarusian State University (in Russian)*, (1), 96–100.
- Kharin A, Kishylau D (2005). “Robust Sequential Testing of Hypotheses on Discrete Probability Distributions.” *Austrian Journal of Statistics*, **34**(2), 153–162.
- Kharin A, Shlyk P (2009). “Robust Multivariate Bayesian Forecasting Under Functional Distortions in the Chi-square Metric.” *Journal of Statistical Planning and Inference*, **139**, 3842–3846.
- Kharin Y (2011b). “Robustness of the Mean Square Risk in Forecasting of Regression Time Series.” *Communications in Statistics – Theory and Methods*, **40**(16), 2893–2906.
- Kharin Y (2013b). “Robustness in Statistical Forecasting.” In C Becker et al (ed.), *Robustness and Complex Data Structures*, pp. 225–242. Springer, Berlin.
- Kharin Y, Voloshko V (2011). “Robust Estimation of AR Coefficients Under Simultaneously Influencing Outliers and Missing Values.” *Journal of Statistical Planning and Inference*, **141**(9), 3276–3288.
- Mukhopadhyay N, de Silva B (2009). *Sequential Methods and Their Applications*. Chapman and Hall / CRC, Boca Raton.
- Wald A (1947). *Sequential Analysis*. John Wiley and Sons, New York.



**Affiliation:**

Alexey Kharin  
Department of Probability Theory and Mathematical Statistics  
Belarusian State University  
Independence av. 4  
220030 Minsk, Belarus  
E-mail: [KharinAY@bsu.by](mailto:KharinAY@bsu.by)  
Telephone: +375 17 2095129  
Fax: +375 17 2095054