AUSTRIAN JOURNAL OF STATISTICS
Volume 40 (2011), Number 1 & 2, 65–73

Robustness Analysis for Bayesian Sequential Testing of Composite Hypotheses under Simultaneous Distortion of Priors and Likelihoods

Alexey Kharin
Belarusian State University, Minsk

Abstract: The Bayesian sequential test of composite hypotheses is considered. The situation of the simultaneously distorted prior probability distribution and likelihood is analyzed. The asymptotic expansions w.r.t. distortion parameters are constructed for the error type I and II probabilities as well as for the expected sample sizes.

Keywords: Sequential Test, Composite Hypothesis, Prior Distribution, Distortion, Robustness.

1 Introduction

In many applications, especially in medicine, engineering, finance, problems of statistical testing of composite hypotheses are typical (Ghosh and Sen, 1991; Jennison and Turnbull, 2000). Taking into consideration high costs of each observation, to solve these problems the sequential approach (Wald, 1947; Lai, 2001) is used. In the framework of this approach, the number of observations required to provide the prescribed performance is not fixed a priori, but is a random variable that depends on observations. This scheme of hypotheses testing needs a number of observations that is essentially less than what is required by the approach based on fixed sample sizes (see Aivazian, 1959).

The hypothetical probability model used in the sequential approach is often distorted in practice (Huber and Ronchetti, 2009; Kharin and Shlyk, 2009). There are two potential sources of distortions: the prior probability distribution of parameters and the conditional probability distribution of observations (the likelihood). Therefore, the problem of robustness analysis for sequential tests of composite hypotheses under simultaneous distortions seems to be an important one.

In Quang (1985) the robustness analysis is given for a special hypothetical model in case of simple hypotheses. An empirical study of robustness is performed in Pandit and Gudaganavar (2009) for the scale parameter of gamma and exponential distributions. In Kharin (2002b) an approach for quantitative robustness analysis of sequential tests for simple hypotheses is proposed. This approach is generalized in Kharin and Kishylau (2005). In case of composite hypotheses the approach is developed (see Kharin, 2008) under "contamination" (Huber and Ronchetti, 2009) of the likelihood. In (Kharin, 2010) the similar problem is solved in the situation where the distortion of the hypothetical model is caused by the distortion of the prior probability density function of parameters vector.

In this paper we generalize the results presented in Kharin (2008, 2010), and discuss the situation of simultaneous distortion of the hypothetical prior probability distribution and the hypothetical likelihood.

2 Mathematical Model

Let on a measurable space $(\Omega, \mathcal{F})$ a random sequence $x_1, x_2, \ldots \in \mathbf{R}$ be observed that corresponds to the n-dimensional conditional probability density function $p_n(x_1, \ldots, x_n \mid \theta)$, $n \in \mathbf{N}$, where $\theta \in \Theta \subseteq \mathbf{R}^k$ is an unknown value of the random parameters vector. The probability density function $p(\theta)$ of this vector in the Bayesian setting is supposed to be known. There are two composite hypotheses on the value of θ:

$$\mathcal{H}_0 : \ \theta \in \Theta_0 \, , \quad \mathcal{H}_1 : \ \theta \in \Theta_1 \, ; \qquad \Theta_0 \cup \Theta_1 = \Theta \, , \quad \Theta_0 \cap \Theta_1 = \emptyset \, . \tag{1}$$

Introduce the notation:

$$\mathbf{1}_S(s) = \begin{cases} 1, & s \in S, \\ 0, & s \notin S; \end{cases}$$

$$W_i = \int_{\Theta_i} p(\theta) d\theta; \ w_i(\theta) = \frac{1}{W_i} \cdot p(\theta) \cdot \mathbf{1}_{\Theta_i}(\theta), \ \theta \in \Theta \, , \qquad i = 0, 1.$$

Denote by

$$\Lambda_n = \Lambda_n(x_1, \ldots, x_n) = \log \frac{\int_\Theta w_1(\theta) p_n(x_1, \ldots, x_n \mid \theta) d\theta}{\int_\Theta w_0(\theta) p_n(x_1, \ldots, x_n \mid \theta) d\theta} \tag{2}$$

the logarithm of the generalized likelihood ratio statistic, that is calculated by n observations $x_1, \ldots, x_n$.

To test the hypotheses (1), the following parametric family of Bayesian sequential tests is used:

$$N = \min\{n \in \mathbf{N} : \ \Lambda_n \notin (C_-, C_+)\} \, , \tag{3}$$

$$d = \mathbf{1}_{[C_+, +\infty)}(\Lambda_N) \, , \tag{4}$$

where N is the random number of the observation that determines the stopping time, after that observation the decision d is made according to the decision rule (4). The decision $d = i$ means that the hypothesis $\mathcal{H}_i$ is accepted, $i = 0, 1$; $C_- < 0$, $C_+ > 0$ are parameters of the test (3), (4):

$$C_- = \log(\beta_0 / (1 - \alpha_0)) \, , \qquad C_+ = \log((1 - \beta_0)/\alpha_0) \, ,$$

where $\alpha_0, \beta_0 \in (0, \frac{1}{2})$ are some values close to maximal admissible levels of error type I and II probabilities (Wald, 1947). The actual values α, β of the error type I and II probabilities may deviate from α_0, β_0 (Kharin, 2002a).

For calculation of α, β and conditional mathematical expectations of the random variable N determined by (3), let us use a stochastic approximation of the statistic Λ_n, $n \in \mathbf{N}$. Let $m \in \mathbf{N}$ be a parameter of the approximation, $h = (C_+ - C_-)/m$. Let $p_{\Lambda_n}(u)$ be the probability density function of the statistic (2); $p_{\Lambda_{n+1}|\Lambda_n}(u \mid y)$ be the conditional probability density function, $n \in \mathbf{N}$; $[x]$ be the integer part of x (the smallest integer number, less or equal to x). Construct a discrete random sequence Z_n^m, $n = 0, 1, 2, \ldots$, with the state space $V = \{0, 1, \ldots, m+1\}$:

$$Z_n^m = \begin{cases} 0, & \text{if } Z_{n-1}^m = 0, \\ m+1, & \text{if } Z_{n-1}^m = m+1, \\ \left(\left[\frac{\Lambda_n - C_-}{h} \right] + 1 \right) \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n) + (m+1) \cdot \mathbf{1}_{[C_+, +\infty)}(\Lambda_n), & \text{otherwise}, \end{cases} \tag{5}$$

$n \in \mathbf{N}$, $Z_0^m = 0$. For this random sequence let us consider the $(m+2) \times (m+2)$-matrix of the conditional transition probabilities

$$P^{(n)}(\theta) = (p_{ij}^{(n)}(\theta)) = (\mathrm{P}\{Z_{n+1}^m = j \mid Z_n^m = i\}), \qquad i, j \in V, \quad n \in \mathbf{N}.$$

Let us approximate the random sequence Z_n^m by the Markov chain $z_n^m \in V$, $n \in \mathbf{N}$, with the same initial probability distribution (that corresponds to the observation number $n = 1$) and with the matrices of transition probabilities $P^{(n)}(\theta)$ at the stage n. After renumeration of the states $V := \{\{0\}, \{m+1\}, \{1\}, \ldots, \{m\}\}$ the matrix $P^{(n)}(\theta)$ is represented in the form:

$$P^{(n)}(\theta) = \left(\begin{array}{c|c} \mathbf{I}_2 & \mathbf{0}_{2\times m} \\ \hline R^{(n)}(\theta) & Q^{(n)}(\theta) \end{array} \right), \qquad \theta \in \Theta,$$

where $R^{(n)}(\theta)$ and $Q^{(n)}(\theta)$ are the blocks of the sizes $m \times 2$ and $m \times m$, respectively, $\mathbf{I}_k$ is the identity matrix of the size k, $\mathbf{0}_{(2\times m)}$ is the matrix of the size $(2 \times m)$, with all elements equal to 0. Let $\pi(\theta) = (\pi_i(\theta))$ be the vector of initial probabilities of the states $1, \ldots, m$ for the random sequence (5); $\pi_0(\theta)$, $\pi_{m+1}(\theta)$ be the initial probabilities of the absorbing states 0 and $m+1$; $\mathbf{1}_m$ be the vector of size m, with all components equal to 1. Denote:

$$S(\theta) = \mathbf{I}_m + \sum_{i=1}^{\infty} \prod_{j=1}^{i} Q^{(j)}(\theta);$$

$$B(\theta) = R^{(1)}(\theta) + \sum_{i=1}^{\infty} \prod_{j=1}^{i} Q^{(j)}(\theta) R^{(i+1)}(\theta).$$

Let $B_{(j)}(\theta)$ be the column number j of the matrix $B(\theta)$, $j = 1, 2$; $t_i = \mathrm{E}\{N \mid \theta \in \Theta_i\}$, $i = 0, 1$; $t = \mathrm{E}\{N\}$.

3 Simultaneous Distortions of Priors and Likelihoods

Let the hypothetical model described above be distorted, although the Bayesian sequential test (3), (4) is used. The test is constructed on the basis of the hypothetical probability density functions $p(\theta)$, $p_n(x_1, \ldots, x_n \mid \theta)$, but these probability density functions are simultaneously distorted. Actually, the parameters vector θ has the distorted probability density function

$$\bar{p}(\theta) = (1 - \varepsilon_\theta) \cdot p(\theta) + \varepsilon_\theta \cdot \tilde{p}(\theta), \qquad \theta \in \Theta, \tag{6}$$

where $\varepsilon_\theta \in [0, \frac{1}{2})$ is the probability of "contamination" w.r.t. the probability density of θ, and $\tilde{p}(\theta)$ is a "contaminating" probability density function that differs from $p(\theta)$. The distorted conditional probability density function of observations is also a mixture of the hypothetical $p_n(\cdot \mid \cdot)$ and the "contaminating" $\tilde{p}_n(\cdot \mid \cdot)$ probability density functions:

$$\bar{p}_n(x_1, \ldots, x_n \mid \theta) = (1 - \varepsilon_x) \cdot p_n(x_1, \ldots, x_n \mid \theta) + \varepsilon_x \cdot \tilde{p}_n(x_1, \ldots, x_n \mid \theta), \tag{7}$$

$$\theta \in \Theta,\ x_1, \ldots, x_n \in \mathbf{R}, \qquad n \in \mathbf{N},$$

where $\varepsilon_x \in [0, \frac{1}{2})$ can be interpreted as the probability of an "outlier" presence (see Huber and Ronchetti, 2009) w.r.t. the observations $x_1, x_2, \ldots$.

Let $\tilde{\pi}(\theta)$, $\tilde{\pi}_0(\theta)$, $\tilde{\pi}_{m+1}(\theta)$, $\tilde{Q}^{(n)}(\theta)$, $\tilde{R}^{(n)}(\theta)$ be the elements calculated analogously to $\pi(\theta)$, $\pi_0(\theta)$, $\pi_{m+1}(\theta)$, $Q^{(n)}(\theta)$, $R^{(n)}(\theta)$ by replacing the hypothetical p.d.f. $p_n(x_1, \ldots, x_n \mid \theta)$ with the "contaminating" p.d.f. $\tilde{p}_n(x_1, \ldots, x_n \mid \theta)$ in the probability distribution of the random sequence (5); $\Delta\pi_0(\theta) = \tilde{\pi}_0(\theta) - \pi_0(\theta)$, $\Delta\pi_1(\theta) = \tilde{\pi}_{m+1}(\theta) - \pi_{m+1}(\theta)$; $\bar{t}(\theta)$ and $\bar{\gamma}_{\mathcal{H}_i}(\theta)$, $i = 0, 1$, be the conditional mathematical expectation of the sample size and the conditional probability of acceptance of the hypothesis $\mathcal{H}_i$ respectively, provided the parameters vecor value is θ, for the distorted model (6), (7).

4 Asymptotic Analysis of Robustness

Introduce the notation:

$$\tilde{W}_i = \int_{\Theta_i} \tilde{p}(\theta) d\theta;$$

$$A(\theta) = \Big((\tilde{\pi}(\theta) - \pi(\theta))' S(\theta) +$$

$$(\pi(\theta))' \cdot \sum_{i=1}^{\infty} \sum_{j=1}^{i} \prod_{k=1}^{j-1} Q^{(k)}(\theta) (\tilde{Q}^{(j)}(\theta) - Q^{(j)}(\theta)) \prod_{k=j+1}^{i} Q^{(k)}(\theta) \Big) \cdot \mathbf{1}_m;$$

$$F_i(\theta) = \Delta\pi_i(\theta) + (\tilde{\pi}(\theta) - \pi(\theta))' B_{(i+1)}(\theta) + \tilde{R}^{(1)}(\theta) - R^{(1)}(\theta) +$$

$$\sum_{l=1}^{\infty} \Big(\sum_{j=1}^{l} \prod_{k=1}^{j-1} Q^{(k)} (\tilde{Q}^{(j)}(\theta) - Q^{(j)}(\theta)) \prod_{k=j+1}^{l} Q^{(k)}(\theta) R^{(l+1)}(\theta) +$$

$$\prod_{j=1}^{l} Q^{(j)}(\theta) (\tilde{R}^{(l+1)}(\theta) - R^{(l+1)}(\theta)) \Big), \qquad i = 0, 1.$$

Theorem 1 *Let the random sequence (2) satisfies the Markov property, $\forall \theta \in \Theta$, the probability density functions $p_{\Lambda_1}(u)$, $p_{\Lambda_{n+1}|\Lambda_n}(u \mid y)$ be differentiable functions w.r.t. the variable $u \in [C_-, C_+]$, and $\exists C \in (0, +\infty)$:*

$$\left| \frac{dp_{\Lambda_1}(u)}{du} \right| \le C, \quad \left| \frac{dp_{\Lambda_{n+1}|\Lambda_n}(u \mid y)}{du} \right| \le C, \qquad u, y \in [C_-, C_+], \quad n \in \mathbf{N}.$$

Then under simultaneous distortions (6), (7) the following asymptotic expansions hold for the error type I and II probabilities $\bar{\alpha}$, $\bar{\beta}$ at $\varepsilon_\theta \to 0$, $\varepsilon_x \to 0$, $h \to 0$:

$$\bar{\alpha} = \alpha + \varepsilon_x \cdot \frac{1}{W_0} \cdot \int_{\Theta_0} F_1(\theta) p(\theta) d\theta +$$

$$\varepsilon_\theta \Big(\frac{1}{W_0^2} \cdot \int_{\Theta_0} (\tilde{p}(\theta) - p(\theta)) d\theta \cdot \int_{\Theta_0} \gamma_{\mathcal{H}_1}(\theta) p(\theta) d\theta + \frac{1}{W_0} \cdot \int_{\Theta_0} \gamma_{\mathcal{H}_1}(\theta)(\tilde{p}(\theta) - p(\theta)) d\theta \Big) + \quad (8)$$

$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h)\,;$$

$$\bar{\beta} = \beta + \varepsilon_x \cdot \frac{1}{W_1} \cdot \int_{\Theta_1} F_0(\theta)p(\theta)d\theta +$$

$$\varepsilon_\theta\Big(\frac{1}{W_1^2}\cdot\int_{\Theta_1}(\tilde{p}(\theta)-p(\theta))d\theta\cdot\int_{\Theta_1}\gamma_{\mathcal{H}_0}(\theta)p(\theta)d\theta+\frac{1}{W_1}\cdot\int_{\Theta_1}\gamma_{\mathcal{H}_0}(\theta)(\tilde{p}(\theta)-p(\theta))d\theta\Big)+ \quad (9)$$

$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h)\,.$$

Proof. Using the results presented in Kharin (2008), if the conditions of Theorem 1 are satisfied, we get for the probability of acceptance of the hypothesis $\mathcal{H}_i$ under the distortion (6), (7):

$$\bar{\gamma}_{\mathcal{H}_i}(\theta) = \gamma_{\mathcal{H}_i}(\theta) + \varepsilon_x \cdot F_i(\theta) + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h)\,, \qquad \theta \in \Theta\,.$$

For the error type I probability, under the distortion we have

$$\bar{\alpha} = \int_{\Theta_0} \bar{\gamma}_{\mathcal{H}_1}(\theta) \cdot \frac{1}{\bar{W}_0} \cdot \bar{p}(\theta)d\theta\,,$$

where $\bar{W}_0 = \int_{\Theta_0} \bar{p}(\theta)d\theta$. Under the distortion (6) we have

$$\bar{W}_0 = (1-\varepsilon_\theta)\cdot W_0 + \varepsilon_\theta \cdot \int_{\Theta_0} \tilde{p}(\theta)d\theta = W_0 + \varepsilon_\theta \cdot \int_{\Theta_0} (\tilde{p}(\theta) - p(\theta))d\theta\,.$$

Therefore,

$$\frac{1}{\bar{W}_0} = \frac{1}{W_0}\cdot\left(1 + \frac{\varepsilon_\theta}{W_0}\cdot\int_{\Theta_0}(p(\theta)-\tilde{p}(\theta))d\theta + \mathcal{O}(\varepsilon_\theta^2)\right)\,. \quad (10)$$

From here we obtain

$$\bar{\alpha} = \int_{\Theta_0}(\gamma_{\mathcal{H}_1}(\theta) + \varepsilon_x \cdot F_1(\theta))\cdot\frac{1}{W_0}\cdot\left(1 + \frac{\varepsilon_\theta}{W_0}\cdot\int_{\Theta_0}(\tilde{p}(u)-p(u))du\right)\times$$

$$(p(\theta) + \varepsilon_x\cdot(\tilde{p}(\theta)-p(\theta)))\,d\theta + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h)\,.$$

By performing equivalent transformations we come to

$$\bar{\alpha} = \alpha + \varepsilon_x \cdot \frac{1}{W_0}\cdot\int_{\Theta_0} F_1(\theta)p(\theta)d\theta +$$

$$\varepsilon_\theta\Big(\frac{1}{W_0^2}\cdot\int_{\Theta_0}(\tilde{p}(\theta)-p(\theta))d\theta\cdot\int_{\Theta_0}\gamma_{\mathcal{H}_1}(\theta)p(\theta)d\theta + \frac{1}{W_0}\cdot\int_{\Theta_0}\gamma_{\mathcal{H}_1}(\theta)(\tilde{p}(\theta)-p(\theta))d\theta\Big)+$$

$$\varepsilon_x\varepsilon_\theta\cdot\Big(\frac{1}{W_0^2}\cdot\int_{\Theta_0}(\tilde{p}(\theta)-p(\theta))d\theta\cdot\int_{\Theta_0}F_1(\theta)p(\theta)d\theta + \frac{1}{W_0}\cdot\int_{\Theta_0}F_1(\theta)(\tilde{p}(\theta)-p(\theta))d\theta\Big)+$$

$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h)\,,$$

and, finally, get (8).

The result (9) for the error type II probability under distortions (6), (7) is obtained in a similar way. ■

The asymptotic expansions with the main terms of up to the second order w.r.t. ε_x, ε_θ can be constructed analogously. The mixed reminder term of the first orders is presented in the proof.

Corollary 1 *If the following equality holds under the conditions of Theorem 1:*

$$\int_{\Theta_0} \tilde{p}(\theta) d\theta = \int_{\Theta_0} p(\theta) d\theta \,, \tag{11}$$

then

$$\bar{\alpha} = \alpha + \varepsilon_x \cdot \frac{1}{W_0} \cdot \int_{\Theta_0} F_1(\theta) p(\theta) d\theta + \varepsilon_\theta \frac{1}{W_0} \cdot \int_{\Theta_0} \gamma_{\mathcal{H}_1}(\theta)(\tilde{p}(\theta) - p(\theta)) d\theta +$$
$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) \,;$$
$$\bar{\beta} = \beta + \varepsilon_x \cdot \frac{1}{W_1} \cdot \int_{\Theta_1} F_0(\theta) p(\theta) d\theta + \varepsilon_\theta \frac{1}{W_1} \cdot \int_{\Theta_1} \gamma_{\mathcal{H}_0}(\theta)(\tilde{p}(\theta) - p(\theta)) d\theta +$$
$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) \,.$$

Note that if the condition (11) is satisfied, then the distortion (6) does not change the prior probabilities of the hypotheses $\mathcal{H}_0$, $\mathcal{H}_1$, but just "re-distribute" the probability measure inside each of the sets Θ_0, Θ_1.

Corollary 2 *Under conditions of Theorem 1, if* $\varepsilon_x = 0$, *then*

$$\bar{\alpha} = \alpha + \frac{\varepsilon_\theta}{W_0} \cdot \int_{\Theta_0} ((\pi(\theta))' B_{(2)}(\theta) + \pi_{m+1}(\theta) - \alpha)(\tilde{p}(\theta) - p(\theta)) d\theta + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) \,,$$

$$\bar{\beta} = \beta + \frac{\varepsilon_\theta}{W_1} \cdot \int_{\Theta_1} ((\pi(\theta))' B_{(1)}(\theta) + \pi_0(\theta) - \beta)(\tilde{p}(\theta) - p(\theta)) d\theta + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) \,.$$

Note that Corollary 2 covers the situation where the distortion of the probability density function of observations is absent, and the hypothetical model is distorted via (6) only. These results correspond to those given in Kharin (2010) for this situation.

In case if $\varepsilon_\theta = 0$ the result of Theorem 1 meets one presented in Kharin (2008).

Denote by $\bar{t}_i$, $i = 0, 1$, the conditional mathematical expectation of the random number of observations N provided the hypothesis $\mathcal{H}_i$ is true, if the hypothetical model is distorted according to (6), (7).

Theorem 2 *Under the conditions of Theorem 1, the conditional expected sample sizes satisfy the asymptotic expansions:*

$$\bar{t}_i = t_i + \varepsilon_x \cdot \frac{1}{W_i} \cdot \int_{\Theta_i} A(\theta) p(\theta) d\theta + \varepsilon_\theta \cdot \Big(t_i \cdot (\tilde{W}_i - W_i) + \frac{1}{W_i} \cdot \int_{\Theta_i} t(\theta)(\tilde{p}(\theta) - p(\theta)) d\theta \Big) + \tag{12}$$
$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) \,, \qquad i = 0, 1 \,.$$

Proof. Under the conditions of Theorem 2 we have from Kharin (2008)

$$\bar{t}(\theta) = t(\theta) + \varepsilon_x \cdot A(\theta) + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h) .$$

The conditional expected sample sizes under distortions are

$$\bar{t}_i = \int_\Theta \bar{t}(\theta) \cdot \frac{1}{\bar{W}_i} \cdot \bar{p}(\theta) \cdot \mathbf{1}_{\Theta_i}(\theta) , \qquad i = 0, 1 . \tag{13}$$

Substituting (10) into (13), we get the asymptotic expansion

$$\bar{t}_i = \int_{\Theta_i} (t(\theta) + \varepsilon_x \cdot A(\theta)) \cdot \frac{1}{W_i} \cdot \Big(1 + \frac{\varepsilon_\theta}{W_i} \cdot \int_{\Theta_i} (\tilde{p}(u) - p(u))du\Big) \times$$
$$(p(\theta) + \varepsilon_\theta \cdot (\tilde{p}(\theta) - p(\theta)))d\theta + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) , \qquad i = 0, 1 .$$

After equivalent transformations we come to

$$\bar{t}_i = t_i + \varepsilon_x \cdot \frac{1}{W_i} \cdot \int_{\Theta_i} A(\theta)p(\theta)d\theta + \varepsilon_\theta \cdot \Big(t_i \cdot (\tilde{W}_i - W_i) + \frac{1}{W_i} \cdot \int_{\Theta_i} t(\theta)(\tilde{p}(\theta) - p(\theta))d\theta\Big) +$$
$$\varepsilon_x\varepsilon_\theta \cdot \frac{1}{W_i} \left((\tilde{W}_i - W_i) \cdot \int_{\Theta_i} A(\theta)p(\theta)d\theta + \int_{\Theta_i} A(\theta)(\tilde{p}(\theta) - p(\theta))d\theta\right) +$$
$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) , \qquad i = 0, 1 .$$

and get (12). ■

Corollary 3 *In the conditions of Corollary 1 the following asymptotic expansions hold:*

$$\bar{t}_i = t_i + \varepsilon_x \cdot \frac{1}{W_i} \cdot \int_{\Theta_i} A(\theta)p(\theta)d\theta + \varepsilon_\theta \cdot \frac{1}{W_i} \cdot \int_{\Theta_i} t(\theta)(\tilde{p}(\theta) - p(\theta))d\theta +$$
$$\mathcal{O}(\varepsilon_x^2) + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) , \qquad i = 0, 1 .$$

Corollary 4 *If the conditions of Corollary 2 are satisfied, then*

$$\bar{t}_i = t_i + \frac{\varepsilon_\theta}{W_i} \cdot \int_{\Theta_i} (1 + (\pi(\theta))' S(\theta) \mathbf{1}_m - t_i)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon_\theta^2) + \mathcal{O}(h) , \qquad i = 0, 1 .$$

The results of Corollary 4 correspond to those given in Kharin (2010) for the situation, where the distortion of the probability density function of observations is absent.

In case of $\varepsilon_\theta = 0$ the result of Theorem 2 turns to one presented in Kharin (2008).

Theorem 3 *If the conditions of Theorem 1 are satisfied, then the following asymptotic expansion holds for the expected sample size:*

$$\bar{t} = t + \varepsilon_x \cdot \int_\Theta A(\theta)p(\theta)d\theta + \varepsilon_\theta \cdot \int_\Theta t(\theta)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h) . \tag{14}$$

Proof. From the representation $\bar{t} = \int_\Theta \bar{t}(\theta)p(\theta)d\theta$ and (10) we have

$$\bar{t} = \int_\Theta (t(\theta) + \varepsilon_x \cdot A(\theta))(p(\theta) + \varepsilon_\theta \cdot (\tilde{p}(\theta) - p(\theta)))d\theta + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h) .$$

Performing equivalent transformations we get

$$\bar{t} = t + \varepsilon_x \cdot \int_\Theta A(\theta)p(\theta)d\theta + \varepsilon_\theta \cdot \int_\Theta t(\theta)(\tilde{p}(\theta) - p(\theta))d\theta +$$
$$\varepsilon_x \varepsilon_\theta \cdot \int_\Theta A(\theta)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h) ,$$

and come to (14). ■

Let us note that there is no terms of the order 2 and higher w.r.t. ε_θ in the expansion (14). This expansion could be also presented in the form

$$\bar{t} = (1 - \varepsilon_\theta) \cdot t + \varepsilon_\theta \cdot \tilde{t} + \varepsilon_x \cdot \Big((1 - \varepsilon_\theta) \cdot \int_\Theta A(\theta)p(\theta)d\theta +$$
$$\varepsilon_\theta \cdot \int_\Theta A(\theta)\tilde{p}(\theta)d\theta\Big) + \mathcal{O}(\varepsilon_x^2) + \mathcal{O}(h) ,$$

where $\tilde{t}$ means the expected sample size for the case where the hypothetical probability density function $p(\cdot)$ is replaced by the "contaminating" density $\tilde{p}(\cdot)$. This could be useful for interpretation of the influences of the distortions (6) and (7).

Corollary 5 *If the conditions of Corollary 2 are satisfied, then*

$$\bar{t} = t + \varepsilon_\theta \cdot \int_\Theta (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot (\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(h) .$$

The results of Corollary 5 correspond to those in Kharin (2010) for the situation, where the distortion of the probability density function of observations is absent.

In case of $\varepsilon_\theta = 0$ the result of Theorem 3 turns to one presented in Kharin (2008).

5 Conclusion

The Bayesian sequential statistical test of composite hypotheses is considered in the paper. The presented results give a possibility to evaluate the robustness under simultaneous distortion of two probability density functions that form the hypothetical model — the prior probability density function of parameters vector and the conditional probability density function of observations provided the parameters vector value is fixed. The asymptotic expansions are constructed for the error type I and II probabilities, and also for the conditional and total sample sizes w.r.t. the probabilities ε_θ, ε_x of "contamination".

It is proved that the deviations of the mentioned characteristics from the hypothetical values have the first order w.r.t. the "contamination" levels under the distortions. Special cases of distortions are analyzed.

The results are useful for construction of robust Bayesian sequential tests by the minimax criterion via the scheme used in Kharin (2002b) for simple hypotheses case.

References

Aivazian, S. A. (1959). Comparison of optimal properties of Neyman–Pearson and Wald tests. *Probability Theory and its Applications*, *4*, 86-93.

Ghosh, B. K., and Sen, P. K. (1991). *Handbook of Sequential Analysis*. New York, Basel, Hong Kong: Marcel Dekker.

Huber, P. J., and Ronchetti, E. M. (2009). *Robust Statistics*. New York: Wiley.

Jennison, C., and Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Boca Raton: Chapman & Hall / CRC.

Kharin, A. (2002a). An approach to performance analysis of the sequential probability ratio test for simple hypotheses. *Proceedings of the Belarusian State University*(4), 92-96.

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

Kharin, A. (2008). Robustness evaluation in sequential testing of composite hypotheses. *Austrian Journal of Statistics*, *37*, 51-60.

Kharin, A. (2010). Influence of distortions of the prior probability distribution to the sequential test characteristics for composite hypotheses. *Statistical Methods of Estimation and Hypotheses Testing*, *22*, 35-42.

Kharin, A., and Kishylau, D. (2005). Performance and robustness analysis for sequential testing of hypotheses on parameters of Markov chains (in Russian). *Proceedings of the National Academy of Sciences of Belarus*, *4*, 30-35.

Kharin, A., and Shlyk, P. (2009). Robust multivariate Bayesian forecasting under functional distortions in the chi-square metric. *Journal of Statistical Planning and Inference*, *139*, 3842-3846.

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

Pandit, P. V., and Gudaganavar, N. V. (2009). On robustness of a sequential test for scale parameter of gamma and exponential distributions. *Applied Mathematics*, *1*, 274-278.

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

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

Author's address:

Alexey Kharin
Department of Probability Theory and Mathematical Statistics
Belarusian State University
Independence avenue 4
220030 Minsk
Belarus

E-Mail: KharinAY@bsu.by