DOI 10.1007/s10958-022-06100-7
Journal of Mathematical Sciences, Vol. 267, No. 1, October, 2022

THE INFLUENCE OF DISTORTIONS OF THE PRIOR DISTRIBUTION ON THE CHARACTERISTICS OF THE SEQUENTIAL TEST FOR COMPOSITE HYPOTHESES

A. Yu. Kharin[1,2]

This study reports the influence of contamination-type distortions of the prior distribution density of the parameters on the conditional probabilities of erroneous decisions and the expected number of observations for sequential testing composite parametric hypotheses. Asymptotic expansions are constructed for the characteristics mentioned above that make it possible to estimate the deviations from the hypothetical values under distortions.

1. Introduction

Problems of statistical testing composite parametric hypotheses are relevant in many applied fields, e.g., engineering, medicine, and financial analysis. Since the cost of each observation may be high, the sequential approach introduced by A. Wald [1] is often used for solving these problems. Within this approach, the number of observations required to provide the desired accuracy of decisions (small error probabilities) is not fixed, but depends on available observations and is random. On average, this makes it possible to use fewer observations than are required by methods based on samples with fixed sizes to guarantee the prescribed accuracy [2–4].

In practice, the probability model often describes the phenomenon under study with some distortions [5]. Therefore the study of robustness (stability) of sequential statistical tests for parametric hypotheses is relevant. In [6-8] this problem is solved in the case where the tested hypotheses are simple. In [9], for the case of composite hypotheses, the robustness analysis of sequential tests was undertaken under outlier-type distortions [5] in observations.

In the present paper we study the influence of distortions of the prior distribution on the characteristics of the sequential test for testing composite hypotheses.

2. The mathematical model. Distortions of prior density

Consider a measurable space $(\Omega, \mathcal{F})$ with a random sequence $x_1, x_2, \ldots \in \mathbb{R}$ being observed having the n-dimensional probability density $p_{x_1,\ldots,x_n|\theta}(\cdot \mid \cdot) = p_n(\cdot \mid \cdot)$, $n \in \mathbb{N}$, $\theta \in \Theta \subseteq \mathbb{R}^k$, be the unknown value of the random vector of parameters whose distribution density $p(\theta)$ is assumed known. There are two composite hypotheses concerning the value of θ:

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

Let

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

be the indicator function,

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

[1] Belarusian State University, Minsk, Minsk, Belarus, e-mail: KharinAY@bsu.by
[2] Department of Probability Theory and Mathematical Statistics, Faculty of Applied Mathematics and Informatics, Belarusian State University, Minsk, Belarus

Translated from *Statisticheskie Metody Otsenivaniya i Proverki Gipotez*, Vol. 22, pp. 35–42, 2010
1072-3374/22/2671-01

Let Λ_n be the logarithm of the generalized likelihood ratio calculated from n observations,

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

In the paper, for testing hypotheses (1), the following parametric family of sequential tests is used:

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

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

where N is the random number of observation (stopping time), upon which a decision d is made in accordance with the rule (4). The decision $d = i$ means that the hypothesis $\mathcal{H}_i$ is accepted, $i = 0, 1$; $C_- < 0$ and $C_+ > 0$ are the parameters of the test (3), (4). To choose the values of C_-, C_+, the expressions $C_- = \ln(\beta_0/(1 - \alpha_0))$ and $C_+ = \ln((1 - \beta_0)/\alpha_0)$ can be used, where $\alpha_0, \beta_0 \in (0, \frac{1}{2})$ are quantities close to maximum admissible probabilities of errors of the first and second kinds [1]. Actual values α, β of error probabilities may differ from α_0, β_0 [10].

To calculate the values α, β and conditional expectations of the random variable N defined by (3), use the approximation for the statistic Λ_n, $n \in \mathbb{N}$. Let $m \in \mathbb{N}$ be the approximation parameter, $h = (C_+ - C_-)/m$. Let $p_{\Lambda_n}(u)$ be the probability density of statistic (2); $p_{\Lambda_{n+1}|\Lambda_n}(u \mid y)$ be the conditional density, $n \in \mathbb{N}$; $[x]$ be the integer part of x (the least integer no greater than x). Construct the discrete random sequence Z_n^m, $n = 0, 1, 2, \ldots$, taking values in the set $V = \{0, 1, \ldots, m+1\}$ so that $Z_0^m = 0$, and for $n \in \mathbb{N}$

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

For this sequence consider the $(m+2) \times (m+2)$ matrix of conditional probabilities

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

Approximate the random sequence Z_n^m by the Markov chain $z_n^m \in V$, $n \in \mathbb{N}$, with the same probability distribution for $n = 1$ (initial distribution) and matrices of transition probabilities $P^{(n)}(\theta)$ at the nth step. After the states are re-enumerated $V := \{\{0\}, \{m+1\}, \{1\}, \ldots, \{m\}\}$, the matrix $P^{(n)}(\theta)$ is representable as

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

where $R^{(n)}(\theta)$ and $Q^{(n)}(\theta)$ are correspondingly $(m \times 2)$ and $(m \times m)$ blocks, $\mathbf{I}_k$ is the unity matrix of size k, $\mathbf{0}_{(2\times m)}$ is the $(2 \times m)$ matrix whose elements are equal to 0. Let $\pi(\theta) = (\pi_i(\theta))$ be the vector of initial state probabilities of the sequence (5); $\pi_0(\theta)$, $\pi_{m+1}(\theta)$ be the initial probabilities of absorbing states 0 and $m+1$; $\mathbf{1}_m$ be the m-dimensional vector whose components are all 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 jth column of the matrix $B(\theta)$, $j=1,2$; $t_i = \mathrm{E}\{N \mid \theta \in \Theta_i\}$, $i=0,1$; $t=\mathrm{E}\{N\}$.

Assume that the hypothetical probability model described above is distorted, so that when the sequential test (3), (4) constructed on the basis of the hypothetical prior distribution density $p(\theta)$ is used, the parameter vector θ actually has the probability density

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

where $\varepsilon \in [0, \frac{1}{2})$ is the probability that contamination appears and $\tilde{p}(\theta)$ is the contaminating density different from $p(\theta)$.

3. Asymptotic expansions for the sequential test characteristics

Theorem 1. *Let the random sequence* (2) *be Markovian, for all* $\theta \in \Theta$, $n \in \mathbb{N}$, *the densities* $p_{\Lambda_1}(u)$, $p_{\Lambda_{n+1}|\Lambda_n}(u \mid y)$ *be differentiable with respect to* $u \in [C_-, C_+]$, *and there exists* $C \in (0,+\infty)$ *such that*

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

Then under the distortions (6), *as* $\varepsilon \to 0$, $h \to 0$, *the probabilities* $\bar{\alpha}$, $\bar{\beta}$ *of the errors of the first and second kinds admit the asymptotic expansions*

$$\bar{\alpha} = \alpha + \frac{\varepsilon}{W_0}\int\limits_{\Theta_0}((\pi(\theta))' B_{(2)}(\theta) + \pi_{m+1}(\theta) - \alpha)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon^2) + \mathcal{O}(h), \tag{7}$$

$$\bar{\beta} = \beta + \frac{\varepsilon}{W_1}\int\limits_{\Theta_1}((\pi(\theta))' B_{(1)}(\theta) + \pi_0(\theta) - \beta)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon^2) + \mathcal{O}(h). \tag{8}$$

Proof. The probability measure of the set Θ_0 under the distortions (6) is given by

$$\bar{W}_0 = \int\limits_{\Theta_0} \bar{p}(\theta)d\theta = (1-\varepsilon)\cdot W_0 + \varepsilon \cdot \int\limits_{\Theta_0}\tilde{p}(\theta)d\theta = W_0 + \varepsilon \cdot \int\limits_{\Theta_0}(\tilde{p}(\theta) - p(\theta))d\theta. \tag{9}$$

In [9], the following asymptotic expansion was obtained for the probability of the error of the first kind under no distortions ($h \to 0$):

$$\alpha = \int\limits_{\Theta}((\pi(\theta))' B_{(2)}(\theta) + \pi_{m+1}(\theta))w_0(\theta)d\theta + \mathcal{O}(h).$$

Using this result, (6) and (9), under the distortions we obtain

$$\bar{\alpha} = \frac{1}{\bar{W}_0}\cdot\int\limits_{\Theta_0}(p(\theta) + \varepsilon\cdot(\tilde{p}(\theta) - p(\theta)))((\pi(\theta))' B_{(2)}(\theta) + \pi_{m+1}(\theta))d\theta + \mathcal{O}(h) =$$

$$= \frac{1}{W_0 + \varepsilon\cdot\int\limits_{\Theta_0}(\tilde{p}(\theta) - p(\theta))d\theta}\cdot\left(\alpha\cdot W_0 + \varepsilon\cdot\int\limits_{\Theta_0}((\pi(\theta))' B_{(2)}(\theta) + \pi_{m+1}(\theta))(\tilde{p}(\theta) - p(\theta))d\theta\right) + \mathcal{O}(h). \tag{10}$$

Now make use of the asymptotic expansion ($\varepsilon \to 0$)

$$\frac{1}{W_0 + \varepsilon\cdot\int\limits_{\Theta_0}(\tilde{p}(\theta) - p(\theta))d\theta} = \frac{1}{W_0}\cdot\left(1 + \frac{\varepsilon}{W_0}\cdot\int\limits_{\Theta_0}(p(\theta) - \tilde{p}(\theta))d\theta + \mathcal{O}(\varepsilon^2)\right).$$

Substituting this expansion in (10), after some algebra we obtain (7). The validity of expansion (8) is proved in the same way.

Corollary 1. *If under the conditions of Theorem* 1 *the distortions* (6) *satisfy the additional condition*

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

then, as $h \to 0$, *we have*

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

$$\bar{\beta} = \beta + \frac{\varepsilon}{W_1} \cdot \int\limits_{\Theta_1} ((\pi(\theta))' B_{(1)}(\theta) + \pi_0(\theta))(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(h).$$

Proof. If in (9) we take account of (11), then we arrive at $\bar{W}_0 = W_0$. Substituting this in (10) we obtain the desired asymptotic relation for $\bar{\alpha}$. The result for $\bar{\beta}$ is proved in the same way.

Note that the meaning of the additional condition (11) consists in that under this condition the distortions (6) do not change the prior probabilities of hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ and only re-distribute the probability measure within each of the sets Θ_0, Θ_1.

Let $\bar{t}_i$, $i = 0, 1$, denote the conditional expectation of the random number N of observations under the hypothesis $\mathcal{H}_i$, given the model is distorted in accordance with (6).

Theorem 2. *Under the conditions of Theorem* 1, *under the distortions* (6), *the conditional expectations* $\bar{t}_i$ *admit the asymptotic expansions* $(\varepsilon \to 0,\ h \to 0)$:

$$\bar{t}_i = t_i + \frac{\varepsilon}{W_i} \cdot \int\limits_{\Theta_i} (1 + (\pi(\theta))' S(\theta) \mathbf{1}_m - t_i)(\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(\varepsilon^2) + \mathcal{O}(h),\ i = 0, 1. \tag{12}$$

Proof. Use the asymptotic expansions for the random number N of observations as $h \to 0$ [9] within the framework of the hypothetical model:

$$t_i = 1 + \int\limits_{\Theta} (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot w_i(\theta)d\theta + \mathcal{O}(h),\ i = 0, 1. \tag{13}$$

Under the distortions (6) we obtain

$$\bar{t}_i = 1 + \int\limits_{\Theta_i} (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot \frac{1}{W_i + \varepsilon \cdot \int\limits_{\Theta_i} (\tilde{p}(u) - p(u))du} \cdot \bar{p}(\theta)d\theta + \mathcal{O}(h) =$$

$$= 1 + \frac{1}{W_i} \cdot \left(1 + \frac{\varepsilon}{W_i} \cdot \int\limits_{\Theta_i} (p(\theta) - \tilde{p}(\theta))d\theta + \mathcal{O}(\varepsilon^2) \right) \times$$

$$\times \left(W_i(t_i - 1) + \varepsilon \cdot \int\limits_{\Theta_i} (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot (\tilde{p}(\theta) - p(\theta))d\theta \right) + \mathcal{O}(h).$$

Collecting similar terms we obtain (12).

Corollary 2. *If under the conditions of Theorem* 1 *the distortions* (6) *satisfy condition* (11), *then, as* $h \to 0$,

$$\bar{t}_i = t_i + \frac{\varepsilon}{W_i} \cdot \int\limits_{\Theta_i} (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot (\tilde{p}(\theta) - p(\theta))d\theta + \mathcal{O}(h),\ i = 0, 1. \tag{14}$$

Proof. Under the condition (11) we have the equalities $\bar{W}_i = W_i$, $i = 0, 1$. Taking this and expansions (13) into account, under the distortions (6) we arrive at (14).

Theorem 3. *Under the conditions of Theorem* 1, *under the distortions* (6) *of the prior density* $p(\theta)$, *as* $h \to 0$, *the unconditional mathematical expectation* $\bar{t}$ *of the random number* N *of observations admits the expansion*

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

Proof. Taking account of the expansion (see [9])

$$t = 1 + \int\limits_{\Theta} (\pi(\theta))' \cdot S(\theta) \cdot \mathbf{1}_m \cdot p(\theta) d\theta + \mathcal{O}(h),$$

which holds within the framework of the hypothetical model, and the distortions (6), we obtain

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

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

4. Conclusion

The presented results make it possible to estimate the influence of contamination-type distortions of the prior probability distribution of the parameter vector θ on the conditional probabilities of erroneous decisions and expectations of the random number of observations for the sequential test of composite hypotheses. It is demonstrated that under these distortions, the deviations of the characteristics under consideration have the first order of smallness with respect to the level ε of distortions. The asymptotic expansion with respect to ε of the mathematical expectation $\bar{t}$ does not contain terms of greater order of smallness than the first. The same property is inherent in conditional characteristics $\bar{\alpha}$, $\bar{\beta}$, $\bar{t}_i$, $i = 0, 1$, under an additional condition (11) which means that the distortions have no influence on the prior probabilities of hypotheses.

The obtained results make it possible to find the extreme contaminating distribution and to construct robust (stable) sequential tests by the principle of minimax of the risk of decision rule using the methods developed in [7] for the case of simple hypotheses.

REFERENCES

1. A. Wald, *Sequential Analysis*, Wiley, New York (1948).
2. B. K. Ghosh and P. K. Sen, *Handbook on Sequential Analysis*, Marcel Dekker, New York (1991).
3. S. A. Ayvazyan, "Comparison of the optimal properties of the Wald and Neyman-Pearson criteria," *Theor. Probab. Appl.*, **4**, No. 1, 83–89 (1959).
4. D. Siegmund, *Sequential Analysis. Tests and Confidence Intervals*, Springer, New York (1985).
5. P. J. Huber, *Robust Statistics*, John Wiley & Sons, New York (1981).
6. P. X. Quang, "Robust sequential testing," *Ann. Statist.*, **13**, No. 2, 638–649 (1985).
7. A. Kharin and D. Kishylau, "Robust sequential testing of hypotheses on discrete probability distributions," *Austrian J. Statist.*, **34**, No. 2, 153–162 (2005).
8. A. Yu. Kharin, "The robustness of sequential testing hypotheses concerning the parameters of random sequences," *Statist. Meth. Estim. Test. Hypot.*, **16**, 92–105 (2002).
9. A. Yu. Kharin, "On an approach to the analysis of a sequential likelihood ratio test for testing simple hypotheses," *Belarusian State Univ. Bull. Ser. Physics, Math., Informat.*, No. 1, 92–96 (2002).