Journal of Mathematics and System Science 8 (2018) 96-118
doi: 10.17265/2159-5291/2018.04.002

Constrained Bayesian Method for Testing the Directional Hypotheses

Kartlos J. Kachiashvili[1,2], N. K. Bansal[3] and I. A. Prangishvili[4]

1. Georgian Technical University, Faculty of Informatics and Control Systems, 77, st. Kostava, Tbilisi 0175, Georgia

2. Vekua Institute of Applied Mathematics of the Tbilisi State University, St. University, Tbilisi 0179, Georgia

3. Department of Mathematics, Statistics, and Computer Science, Marquette University, Milwaukee, WI 53201-1881, USA

4. Department of Computer Science, Swiss Federal Institute of Technology (ETHZ), Rämistrasse 101, Zürich 8092, Switzerland

Abstract: The paper discusses the generalization of constrained Bayesian method (CBM) for arbitrary loss functions and its application for testing the directional hypotheses. The problem is stated in terms of false and true discovery rates. One more criterion of estimation of directional hypotheses tests quality, the Type III errors rate, is considered. The ratio among discovery rates and the Type III errors rate in CBM is considered. The advantage of CBM in comparison with Bayes and frequentist methods is theoretically proved and demonstrated by an example.

Key words: CBM, Bayesian test, hypotheses testing, false discovery rate, type III error rate.

1. Introduction

Statistical hypothesis testing is one of the basic problems of the mathematical statistics theory and practice. Many different types of hypotheses have been considered in the literature. However directional hypotheses are comparatively new in comparison to traditional hypotheses. For parametrical models, this problem can be stated as H_0: $\theta = \theta_0$ vs. H_-: $\theta < \theta_0$ or H_+: $\theta > \theta_0$, where θ is the parameter of the model, θ_0 is known (see, for example, Ref. [1]).

The consideration of directional hypotheses started in the 50-ies of the last century. The earliest works considering this problem were by Lehmann [2-4] and Bahadur [5]. Interest in this problem has not decreased since (see, for example, Kaiser [6]; Leventhal & Huynh [7]; Finner [8]; Jones & Tukey [9] and Shaffer [10]; Bansal & Sheng [1]). For solving this problem, authors used traditional methods based on *p*-values, frequentist or Bayesian approaches and their modifications. A compact but exhaustive review of these works is given in Bansal & Sheng [1] where Bayesian decision theoretical methodology for testing the directional hypotheses was developed and compared with the frequentist method. In the same work, the decision theoretic methodology was used for testing multiple directional hypotheses. The cases of multiple experiments for directional hypotheses were also considered in Ref. [11, 12]. The choice of a loss function related to the Kullback-Leibler divergence in a general Bayesian framework for testing the directional hypotheses is considered in Ref. [13].

A new approach to the statistical hypotheses testing, called constrained Bayesian method (CBM), was developed by Kachiashvili et al. [14, 15], Kachiashvili & Mueed [16]. This method differs from the traditional Bayesian approach with a risk function split into two parts, reflecting risks for incorrect rejection and incorrect acceptance of hypotheses and stating the risk minimization problem as a constrained optimization problem when one of the risk components is restricted and the another one is minimized [15, 17]. Application of this method to different types of hypotheses (two and many simple, composite and multiple hypotheses) with parallel and

Corresponding author: Kartlos J. Kachiashvili, doctor of sciences, full professor, research fields: mathematical statistics, data analysis, mathematical modeling and simulation, new computer technologies development, system analysis, computing mathematics.

sequential experiments showed the advantage and uniqueness of the method in comparison with existing ones [18-21]. The uniqueness of the method consists in the emergence of the regions of impossibility of making a simple or any decision alongside with the regions of acceptance of tested hypotheses (like the sequential analysis method), which allows us based on this approach to develop both parallel and sequential method without any additional efforts. The advantage of the method is the optimality of made decisions with guaranteed reliability and minimality of necessary observations for given reliability (see, for example, Kachiashvili [18, 19]; Kachiashvili [20]; Kachiashvili [21]). CBM uses not only loss functions and priori probabilities for making decisions as the classical Bayesian rule does, but also a significance level as the frequentist method does. The combination of these opportunities improves the quality of made decisions in CBM in comparison with other methods. Taking into account the fact that CBM gives better results than other known methods for testing the traditional hypotheses, it is expected that it will give similar better results for testing the directional hypotheses as it, in addition to the classical Bayesian method, uses significance levels in appropriate restrictions.

In Section 2 the generalization of CBM for arbitrary loss functions is given. Application of CBM to the directional hypotheses and the investigation of the obtained decision rule are presented in Section 3. CBM for the normally distributed directional hypotheses is considered in Section 4. Computation results of a concrete example are given in Section 5. Some specific facts which take place in CBM are described in Section 6. Short conclusions are made in Section 7.

2. CBM for the General Loss Function

In the above mentioned works, CBM was introduced and investigated for the "0-1" loss function (see, for example, Kachiashvili [17]; Kachiashvili et al. [15]). Let us now consider the general case. Let the sample $x^T = (x_1, ..., x_n)$ be generated from $p(x;\theta)$ and the problem of interest is to test H_i: $\theta_i \in \Theta_i$, $i = 1, 2, ..., S$, where $\Theta_i \subset R^m$, $i = 1, 2, ..., S$, are disjoint subsets with $\cup\Theta_i = R^m$. The number of tested hypotheses is S. Let the prior on θ be denoted by $\sum_{i=1}^{S} \pi(\theta \mid H_i) p(H_i)$ where for $i = 1, 2, ..., S$, $p(H_i)$ is a priori probability of hypothesis H_i and $\pi(\theta \mid H_i)$ is a prior density with support Θ_i; $p(x \mid H_i)$ denotes the marginal density of x given H_i, i.e., $p(x \mid H_i) = \int_{\Theta_i} p(x \mid \theta)\, \pi(\theta \mid H_i) d\theta$ and $D = \{c\}$ is the set of solutions, where $d = \{d_1, \ldots, d_s\}$, it being so that

$$d_i = \begin{cases} 1, & if \quad hypothesis \quad H_i \quad is \quad accepted \quad , \\ 0, & otherwise; \end{cases}$$

$\delta(x) = \{\delta_1(x), \delta_2(x), \ldots, \delta_s(x)\}$ is the decision function that associates each observation vector x with a certain decision

$$x \xrightarrow{\delta(x)} d \in D$$

(notation: depending upon the choice of x, there is a possibility that $\delta_j(x) = 1$ for more than one j or $\delta_j(x) = 0$ for all $j = 1, ..., S$).

Γ_j is the region of acceptance of hypothesis H_j, i.e. $\Gamma_j = \{x : \delta_j(x) = 1\}$. It is obvious that $\delta(x)$ is completely determined by the Γ_j regions, i.e. $\delta(x) = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_s\}$. Let $L_1(H_i, \delta_j(x) = 1)$ and $L_2(H_i, \delta_j(x) = 0)$ be the losses of incorrectly accepted and incorrectly rejected hypotheses. Then one of possible formulations of CBM can be as follows: to minimize the averaged loss of incorrectly accepted hypotheses

$$r_\delta = \min_{\{\Gamma_j\}} \left\{ \sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} \int_{\Gamma_j} L_1(H_i, \delta_j(x)=1) p(x \mid H_i) dx \right\} \tag{1}$$

subject to the averaged loss of incorrectly rejected hypotheses

$$\sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} \int_{R^n - \Gamma_j} L_2(H_i, \delta_j(x)=0) p(x \mid H_i) dx$$

$$= \sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} \int_{R^n} L_2(H_i, \delta_j(x)=0) p(x \mid H_i) dx$$

$$- \sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} \int_{\Gamma_j} L_2(H_i, \delta_j(x)=0) p(x \mid H_i) dx \le r_1 \tag{2}$$

where r_1 is some real number determining the level of the averaged loss of incorrectly rejected hypotheses.

The kinds of functions in Eqs. (1) and (2) could be chosen differently depending on what type of restrictions is desired proceeding from the aim of the practical problem that must be solved [15, 17]. One of possible statements in Eqs. (1) and (2) minimizes the averaged risk caused by incorrectly accepted hypotheses with restriction of the averaged risk caused by incorrectly rejected hypotheses.

By solving problem in Eqs. (1) and (2), we have [15, 22]

$$\Gamma_j = \left\{ x : \sum_{i=1}^{S} L_1(H_i, \delta_j(x)=1) p(H_i) p(x \mid H_i) < \lambda \sum_{i=1}^{S} L_2(H_i, \delta_j(x)=0) p(H_i) p(x \mid H_i) \right\}$$

$$j = 1, \ldots, S \tag{3}$$

where Lagrange multiplier $\lambda (\lambda > 0)$ is defined so that in Eq. (2) the equality takes place.

Remark. When the losses are the following

$$L_1(H_i, \delta_j(x)=1) = \begin{cases} 0 & at \quad i = j, \\ 1 & at \quad i \ne j, \end{cases}, \quad L_2(H_i, \delta_j(x)=0) = \begin{cases} 0 & at \quad i \ne j, \\ 1 & at \quad i = j, \end{cases} \tag{4}$$

hypotheses acceptance regions in Eq. (3) coincide with the suitable regions of the appropriate task of CBM at loss "0-1" (see Task 1 in Kachiashvili et al. [15]).

Let us suppose that the losses are the same within the acceptance and rejection regions and introduce denotations $L_1(H_i, H_j)$ and $L_2(H_i, H_j)$ for incorrect acceptance of H_i when H_i is true and incorrect rejection of H_j in favour of H_i. Then decision making regions in Eq. (3) take the form

$$\Gamma_j = \left\{ x : \sum_{i=1}^{S} p(H_i) L_1(H_i, H_j) p(x \mid H_i) < \lambda \sum_{i=1}^{S} p(H_i) L_2(H_i, H_k) p(x \mid H_i); \right.$$

$$\left. \forall k : k \in (1, \ldots, j-1, j+1, \ldots, S) \right\}, \; j = 1, \ldots, S; \tag{5^1}$$

that is the same as

$$\Gamma_j = \left\{ x : \sum_{i=1}^{S} L_1(H_i, H_j) p(H_i \mid x) < \lambda \sum_{i=1}^{S} L_2(H_i, H_k) p(H_i \mid x); \right.$$

$$\left. \forall k : k \in (1, \ldots, j-1, j+1, \ldots, S) \right\}, \; j = 1, \ldots, S; \tag{5^2}$$

From Eq. (2) it is clear that the following condition must be fulfilled

$$r_1 < \sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} \int_{R^n} L_2(H_i, \delta_j(x)=0) p(x \mid H_i) dx \tag{6^1}$$

i.e. for losses $L_1(H_i, H_j)$ and $L_2(H_i, H_j)$,

$$r_1 < \sum_{i=1}^{S} p(H_i) \sum_{j=1}^{S} L_2(H_i, H_j) . \qquad (6^2)$$

Using the same denotations and introducing the general loss function $L(H_i, \delta(x))$ which determines the value of loss in the case when the sample has the probability distribution corresponding to hypothesis H_i, but, because of random errors, decision $\delta(x)$ is made, the Bayesian statement of S hypotheses testing is [23-25].

$$r_\delta^B = \min_{\{\delta(x)\}} \left\{ \sum_{i=1}^{S} p(H_i) \int_{R^n} L(H_i, \delta(x)) p(x \mid H_i) dx \right\}. \qquad (7)$$

In the general case, loss function $L(H_i, \delta(x))$ consists of two components:

$$L(H_i, \delta(x)) = \sum_{j=1}^{S} L_1(H_i, \delta_j(x) = 1) + \sum_{j=1}^{S} L_2(H_i, \delta_j(x) = 0), \qquad (8)$$

i.e. loss function $L(H_i, \delta(x))$ is the total loss of incorrectly accepted and incorrectly rejected hypotheses.

Taking into account Eq. (8), the solution of the problem Eq. (7) can be written down in the following (form Refs. [23, 25]:

$$\Gamma_j = \left\{ x : \sum_{i=1}^{S} L_1(H_i, \delta_j(x) = 1) p(H_i) p(x \mid H_i) < \sum_{i=1}^{S} L_2(H_i, \delta_j(x) = 0) p(H_i) p(x \mid H_i) \right\},$$

$$j = 1, \ldots, S, \qquad (9)$$

and, for losses $L_1(H_i, H_j)$ and $L_2(H_i, H_j)$, we have

$$\Gamma_j = \left\{ x : \sum_{i=1}^{S} L_1(H_i, H_j) p(H_i \mid x) < \sum_{i=1}^{S} L_2(H_i, H_k) p(H_i \mid x); \right.$$
$$\left. \forall k : k \in (1, \ldots, j-1, j+1, \ldots, S) \right\}, \; j = 1, \ldots, S . \qquad (10)$$

It is obvious that the difference between Eqs. (3) and (9), that is between Eqs. (5) and (10), consists in the Lagrange multiplier λ which cardinally changes the properties of decision-making regions.

Let us define the summary risk (SR) of making the incorrect decision at hypotheses testing as the weighed sum of probabilities of making incorrect decisions, i.e.

$$r_S(\Gamma) = \sum_{i=1}^{S} \sum_{j=1, j \neq i}^{S} L(H_i, H_j) p(H_i) \int_{\Gamma_j} p(x \mid H_i) dx . \qquad (11)$$

It is clear that, for given losses and probabilities, SR depends on the regions of making decisions. Let us denote by Γ^{CBM} and Γ^B the hypotheses acceptance regions in CBM and Bayes rule, respectively. Then SR for CBM and Bayes rules are $r_s(\Gamma^{CBM})$ and $r_s(\Gamma^B)$, respectively.

Theorem 1. For given losses and probabilities, SR of making incorrect decision in CBM is convex function of λ with maximum at $\lambda = 1$. At increasing or decreasing λ, SR decreases, and in the limit, i.e. at $\lambda \to \infty$ or $\lambda \to 0$, SR tends to zero.

Corollary 1. At the same conditions SR of making the incorrect decision in CBM is less or equal to SR of the Bayesian decision rule, i.e. $r_s(\Gamma^{CBM}) \leq r_s(\Gamma^B)$.

The justice of this corollary is obvious from theorem 1.

Providing proving consideration, similar to theorem 1, it is not difficult to be convinced that SR of making the incorrect decision in CBM is less or equal to SR of the frequentist decision rule, i.e. $r_s(\Gamma^{CBM}) \leq r_s(\Gamma^f)$, where $r_s(\Gamma^f)$ is SR for the frequentist method.

3. Consideration of Directional Hypotheses

Let us consider the directional hypotheses H_0: $\theta = \theta_0$ vs. H_-: $\theta < \theta_0$, or H_+: $\theta > \theta_0$. For testing these hypotheses, the loss functions that do not depend on x are used in Ref. [1]. Let us denote: Γ_0, Γ_- and Γ_+ are the regions of acceptance of the appropriate hypotheses.

In the considered case, decision-making region Eq. (5) becomes: the hypothesis H_- acceptance region

$$\Gamma_- = \{x : L_1(H_-,H_-)p(H_- \mid x) + L_1(H_0,H_-)p(H_0 \mid x) + L_1(H_+,H_-)p(H_+ \mid x) <$$

$$(k \equiv 0) \quad < \lambda\left[L_2(H_-,H_0)p(H_- \mid x) + L_2(H_0,H_0)p(H_0 \mid x) + L_2(H_+,H_0)p(H_+ \mid x)\right]$$

$$\& \, L_1(H_-,H_-)p(H_- \mid x) + L_1(H_0,H_-)p(H_0 \mid x) + L_1(H_+,H_-)p(H_+ \mid x) <$$

$$(k \equiv +) \quad < \lambda\left[L_2(H_-,H_+)p(H_- \mid x) + L_2(H_0,H_+)p(H_0 \mid x) + L_2(H_+,H_+)p(H_+ \mid x)\right] \tag{12}$$

similarly, for Γ_0, we have

$$\Gamma_0 = \{x : L_1(H_-,H_0)p(H_- \mid x) + L_1(H_0,H_0)p(H_0 \mid x) + L_1(H_+,H_0)p(H_+ \mid x) <$$

$$(k \equiv -) \quad < \lambda\left[L_2(H_-,H_-)p(H_- \mid x) + L_2(H_0,H_-)p(H_0 \mid x) + L_2(H_+,H_-)p(H_+ \mid x)\right]$$

$$\& \, L_1(H_-,H_0)p(H_- \mid x) + L_1(H_0,H_0)p(H_0 \mid x) + L_1(H_+,H_0)p(H_+ \mid x) <$$

$$(k \equiv +) \quad < \lambda\left[L_2(H_-,H_+)p(H_- \mid x) + L_2(H_0,H_+)p(H_0 \mid x) + L_2(H_+,H_+)p(H_+ \mid x)\right] \tag{13}$$

and, for Γ_+, we have

$$\Gamma_+ = \{x : L_1(H_-,H_+)p(H_- \mid x) + L_1(H_0,H_+)p(H_0 \mid x) + L_1(H_+,H_+)p(H_+ \mid x) <$$

$$(k \equiv 0) \quad < \lambda\left[L_2(H_-,H_0)p(H_- \mid x) + L_2(H_0,H_0)p(H_0 \mid x) + L_2(H_+,H_0)p(H_+ \mid x)\right]$$

$$\& \, L_1(H_-,H_+)p(H_- \mid x) + L_1(H_0,H_+)p(H_0 \mid x) + L_1(H_+,H_+)p(H_+ \mid x) <$$

$$(k \equiv -) \quad < \lambda\left[L_2(H_-,H_-)p(H_- \mid x) + L_2(H_0,H_-)p(H_0 \mid x) + L_2(H_+,H_-)p(H_+ \mid x)\right] \tag{14}$$

The following “0-K” loss function was used in Ref. [1].

$$L_1(H_-,H_-) = L_1(H_0,H_0) = L_1(H_+,H_+) = L_2(H_-,H_-) = L_2(H_0,H_0) = L_2(H_+,H_+) = 0$$

$$L_1(H_-,H_0) = L_1(H_+,H_0) = K_0, \quad L_2(H_-,H_0) = L_2(H_+,H_0) = K_0$$

$$L_1(H_0,H_-) = L_1(H_0,H_+) = L_1(H_-,H_+) = L_1(H_+,H_-) = K_1,$$

$$L_2(H_-,H_+) = L_2(H_+,H_-) = L_2(H_0,H_-) = L_2(H_0,H_+) = K_1 \tag{15}$$

Inputting these losses into decision-making regions Eqs. (12)-(14), we have: for Γ_-

$$(k \equiv 0) \quad \Gamma_- = \{x : K_1\left(p(H_0 \mid x) + p(H_+ \mid x)\right) < \lambda \cdot K_0\left(p(H_- \mid x) + p(H_+ \mid x)\right)$$

$$(k \equiv +) \quad \& K_1\big(p(H_0 \mid x) + p(H_+ \mid x)\big) < \lambda \cdot K_1\big(p(H_- \mid x) + p(H_0 \mid x)\big)\} \tag{16}$$

i.e.

$$(k \equiv 0)\, \Gamma_- = \{ x : \frac{p(H_0 \mid x) + p(H_+ \mid x)}{p(H_- \mid x) + p(H_+ \mid x)} = \frac{1 - p(H_- \mid x)}{1 - p(H_0 \mid x)} < \frac{K_0}{K_1}\lambda$$

$$(k \equiv +)\, \& \frac{p(H_0 \mid x) + p(H_+ \mid x)}{p(H_0 \mid x) + p(H_- \mid x)} = \frac{1 - p(H_- \mid x)}{1 - p(H_+ \mid x)} < \lambda \} \tag{17}$$

similarly, for Γ_0, we have

$$(k \equiv -) \quad \Gamma_0 = \{x : K_0\big(p(H_- \mid x) + p(H_+ \mid x)\big) < \lambda \cdot K_1\big(p(H_0 \mid x) + p(H_+ \mid x)\big)$$

$$(k \equiv +) \quad \& K_0\big(p(H_- \mid x) + p(H_+ \mid x)\big) < \lambda \cdot K_1\big(p(H_- \mid x) + p(H_0 \mid x)\big)\} \tag{18}$$

i.e.

$$(k \equiv -) \quad \Gamma_0 = \left\{ x : \frac{p(H_0 \mid x) + p(H_+ \mid x)}{p(H_- \mid x) + p(H_+ \mid x)} = \frac{1 - p(H_- \mid x)}{1 - p(H_0 \mid x)} > \frac{K_0}{K_1} \cdot \frac{1}{\lambda} \right.$$

$$(k \equiv +) \quad \left. \& \frac{p(H_- \mid x) + p(H_0 \mid x)}{p(H_- \mid x) + p(H_+ \mid x)} = \frac{1 - p(H_+ \mid x)}{1 - p(H_0 \mid x)} > \frac{K_0}{K_1} \cdot \frac{1}{\lambda} \right\} \tag{19}$$

and, finally

$$\Gamma_0 = \left\{ x : \min\left(\frac{1 - p(H_- \mid x)}{1 - p(H_0 \mid x)}, \frac{1 - p(H_+ \mid x)}{1 - p(H_0 \mid x)} \right) > \frac{K_0}{K_1} \cdot \frac{1}{\lambda} \right\} \tag{20}$$

for Γ_+, we obtain

$$\Gamma_+ = \{x : L_1(H_-, H_+)p(H_- \mid x) + L_1(H_0, H_+)p(H_0 \mid x) + L_1(H_+, H_+)p(H_+ \mid x)$$

$$(k \equiv 0) \quad < \lambda\big[L_2(H_-, H_0)p(H_- \mid x) + L_2(H_0, H_0)p(H_0 \mid x) + L_2(H_+, H_0)p(H_+ \mid x)\big]$$

$$\& L_1(H_-, H_+)p(H_- \mid x) + L_1(H_0, H_+)p(H_0 \mid x) + L_1(H_+, H_+)p(H_+ \mid x)$$

$$(k \equiv -) \quad < \lambda\big[L_2(H_-, H_-)p(H_- \mid x) + L_2(H_0, H_-)p(H_0 \mid x) + L_2(H_+, H_-)p(H_+ \mid x)\big] \tag{21}$$

i.e.

$$(k \equiv 0) \qquad \Gamma_+ = \{ x : \frac{p(H_0 \mid x) + p(H_- \mid x)}{p(H_- \mid x) + p(H_+ \mid x)} = \frac{1 - p(H_+ \mid x)}{1 - p(H_0 \mid x)} < \frac{K_0}{K_1}\lambda \ \&$$

$$(k \equiv -) \qquad \& \frac{p(H_0 \mid x) + p(H_- \mid x)}{p(H_0 \mid x) + p(H_+ \mid x)} = \frac{1 - p(H_+ \mid x)}{1 - p(H_- \mid x)} < \lambda \}. \tag{22}$$

Analyzing regions Eqs. (17), (20) and (22), we conclude that generally, for arbitrary $\lambda > 0$, in contradistinction to the classical cases, the following conditions take place: $\Gamma_i \cap \Gamma_j \neq \emptyset$, $i \neq j$, $i, j \in (-, 0, +)$ and $\Gamma_- \cup \Gamma_0 \cup \Gamma_+ \neq R^n$, i.e. in general, hypotheses acceptance regions intersect and the union of these regions does not coincide with the observation space. If more than one of conditions Eqs. (17), (20) and (22) or none of these conditions are fulfilled, then it is impossible to make a simple decision. In the first case more than one of the hypotheses are suspected to be true and, in the second case, it is impossible to make a

single decision. In such cases it is necessary to obtain one more observation and, on the basis of increased sample, to make a decision using condition Eqs. (17), (20) and (22) or to change r_1 in condition Eq. (2) upon fulfilling only one of conditions Eqs. (17), (20) and (22). When $\lambda = 1$, decision rules Eqs. (17), (20) and (22) completely coincide with the Bayesian decision rule given in Ref. [1].

In the case of loss functions in Eq. (15), condition Eq. (2) takes the form

$$K_1 p(H_-)\int_{\Gamma_-} p(x \mid H_-)dx + K_0 p(H_0)\int_{\Gamma_0} p(x \mid H_0)dx + K_1 p(H_+)\int_{\Gamma_+} p(x \mid H_+)dx$$

$$\geq p(H_-)\cdot K_1 + p(H_0)\cdot K_0 + p(H_+)\cdot K_1 - r_1 . \tag{23}$$

Hence it is clear that the following condition must always be satisfied

$$p(H_-)\cdot K_1 + p(H_0)\cdot K_0 + p(H_+)\cdot K_1 > r_1 .$$

Let us choose r_1 as follows

$$r_1 = p(H_-)\cdot K_1 \cdot \alpha_- + p(H_0)\cdot K_0 \cdot \alpha_0 + p(H_+)\cdot K_1 \cdot \alpha_+ ,$$

Where $0 \leq a_- \leq 1$, $0 \leq \alpha_0 \leq 1$ and $0 \leq \alpha_+ \leq 1$. Then, in the right side of (23), we have

$$p(H_-)\cdot K_1 \cdot (1-\alpha_-) + p(H_0)\cdot K_0 \cdot (1-\alpha_0) + p(H_+)\cdot K_1 \cdot (1-\alpha_+) . \tag{24}$$

Let us consider the following losses

$$L_1(H_i, H_j) = \begin{cases} 0 \quad at \quad i = j, \\ L_1(H_i, H_j) \quad at \quad i \neq j, \end{cases} \quad L_2(H_i, H_j) = \begin{cases} 0 \quad at \quad i \neq j, \\ L_2(H_i, H_i) \quad at \quad i = j. \end{cases} \tag{25}$$

It is clear that the "0-1" loss function is a private case of the step-wise loss (25).

For loss functions Eq. (25), Eq. (1) takes the form

$$r_\delta = \min_{\{\Gamma_-, \Gamma_0, \Gamma_+\}} \Bigg\{ p(H_-)\Big[L_1(H_-, H_0)\int_{\Gamma_0} p(x \mid H_-)dx + L_1(H_-, H_+)\int_{\Gamma_+} p(x \mid H_-)dx \Big]$$

$$+ p(H_0)\Big[L_1(H_0, H_-)\int_{\Gamma_-} p(x \mid H_0)dx + L_1(H_0, H_+)\int_{\Gamma_+} p(x \mid H_0)dx \Big]$$

$$+ p(H_+)\Big[L_1(H_+, H_-)\int_{\Gamma_-} p(x \mid H_+)dx + L_1(H_+, H_0)\int_{\Gamma_0} p(x \mid H_+)dx \Big] \Bigg\}, \tag{26}$$

and condition Eq. (2) transforms in the following form

$$p(H_-)L_2(H_-, H_-)\int_{\Gamma_-} p(x \mid H_-)dx + p(H_0)L_2(H_0, H_0)\int_{\Gamma_0} p(x \mid H_0)dx + p(H_+)L_2(H_+, H_+)\int_{\Gamma_+} p(x \mid H_+)dx$$

$$\geq p(H_-)L_2(H_-,H_-)+p(H_0)L_2(H_0,H_0)+p(H_+)L_2(H_+,H_+)-r_1 \,. \tag{27}$$

Stated problem Eq. (26), Eq. (27) can be written as

$$r_\delta = \min_{\{\Gamma_-,\Gamma_0,\Gamma_+\}} \{L_1(H_-,H_0)\cdot P(x\in\Gamma_0)\cdot P(H_- \mid x\in\Gamma_0)+L_1(H_-,H_+)\cdot P(x\in\Gamma_+)\cdot P(H_- \mid x\in\Gamma_+)$$

$$+L_1(H_0,H_-)\cdot P(x\in\Gamma_-)\cdot P(H_0 \mid x\in\Gamma_-)+L_1(H_0,H_+)\cdot P(x\in\Gamma_+)\cdot P(H_0 \mid x\in\Gamma_+)$$

$$+L_1(H_+,H_-)\cdot P(x\in\Gamma_-)\cdot P(H_+ \mid x\in\Gamma_-)+L_1(H_+,H_0)\cdot P(x\in\Gamma_0)\cdot P(H_+ \mid x\in\Gamma_0)\}, \tag{28}$$

at

$$L_2(H_-,H_-)\,p(x\in\Gamma_-)\,p(H_- \mid x\in\Gamma_-)+L_2(H_0,H_0)\,p(x\in\Gamma_0)\,p(H_0 \mid x\in\Gamma_0)$$

$$+L_2(H_+,H_+)\,p(x\in H_+)\,p(H_+ \mid x\in\Gamma_+)$$

$$\geq p(H_-)L_2(H_-,H_-)+p(H_0)L_2(H_0,H_0)+p(H_+)L_2(H_+,H_+)-r_1 \tag{29}$$

At the “0-1” loss function, Eqs. (28) and (29) take the form

$$r_\delta = \min_{\{\Gamma_-,\Gamma_0,\Gamma_+\}} \{P(x\in\Gamma_0)\cdot P(H_- \mid x\in\Gamma_0)+P(x\in\Gamma_+)\cdot P(H_- \mid x\in\Gamma_+)$$

$$+P(x\in\Gamma_-)\cdot P(H_0 \mid x\in\Gamma_-)+P(x\in\Gamma_+)\cdot P(H_0 \mid x\in\Gamma_+)$$

$$+P(x\in\Gamma_-)\cdot P(H_+ \mid x\in\Gamma_-)+P(x\in\Gamma_0)\cdot P(H_+ \mid x\in\Gamma_0)\} \tag{30}$$

at

$$p(x\in\Gamma_-)\,p(H_- \mid x\in\Gamma_-)+p(x\in\Gamma_0)\,p(H_0 \mid x\in\Gamma_0)+p(x\in H_+)\,p(H_+ \mid x\in\Gamma_+)\geq 1-r_1 \tag{31}$$

Let us rewrite Eqs. (30) and (31) in the following forms

$$r_\delta = \min_{\{\Gamma_-,\Gamma_0,\Gamma_+\}} \{P(H_-)\cdot\left[P(x\in\Gamma_0 \mid H_-)+P(x\in\Gamma_+ \mid H_-)\right]$$

$$+P(H_0)\cdot\left[P(x\in\Gamma_- \mid H_0)+P(x\in\Gamma_+ \mid H_0)\right]$$

$$+P(H_+)\cdot\left[P(x\in\Gamma_- \mid H_+)+P(x\in\Gamma_0 \mid H_+)\right]\} \tag{32}$$

and

$$p(x\in\Gamma_- \mid H_-)\,p(H_-)+p(x\in\Gamma_0 \mid H_0)\,p(H_0)+p(x\in H_+ \mid H_+)\,p(H_+)\geq 1-r_1 \,. \tag{33}$$

The results of Eqs. (32) and (33) can be stated in terms of positive false discovery rates (pFDR) for testing multiple hypotheses [26]. Let us call false discovery rates of the appropriate hypotheses the following

$$pFDR_-(\Gamma_-)=P(x\in\Gamma_0 \mid H_-)+P(x\in\Gamma_+ \mid H_-)\,,$$

$$pFDR_+(\Gamma_+)=P(x\in\Gamma_- \mid H_+)+P(x\in\Gamma_0 \mid H_+)\,, \tag{34}$$

$$pFDR_0(\Gamma_0) = P(x \in \Gamma_- \mid H_0) + P(x \in \Gamma_+ \mid H_0)$$

and true discovery rates of the appropriate hypotheses the following

$$TDR_-(\Gamma_-) = p(x \in \Gamma_- \mid H_-)\,,\ \ TDR_0(\Gamma_0) = p(x \in \Gamma_0 \mid H_0)\,,\ \ TDR_+(\Gamma_+) = p(x \in \Gamma_+ \mid H_+) \qquad (35)$$

Then Eqs. (32) and (33) will be written as follows

$$r_\delta = \min_{\{\Gamma_-,\Gamma_0,\Gamma_+\}} \left\{ P(H_-)\cdot pFDR_-(\Gamma_-) + P(H_0)\cdot pFDR_0(\Gamma_0) + P(H_+)\cdot pFDR_+(\Gamma_+) \right\} \qquad (36)$$

at

$$p(H_-)TDR_-(\Gamma_-) + p(H_0)TDR_0(\Gamma_0) + p(H_+)TDR_+(\Gamma_+) \ge 1 - r_1 \qquad (37)$$

For comparing the decision rules, let us consider a hierarchical structure of the prior on θ similarly to Ref. [1]. Let us introduce the first stage prior $p_- = p(H_-)$, $p_0 = p(H_0)$, $p_+ = p(H_+)$, with $p_- + p_0 + p_+ = 1$, and the second stage prior on θ as $\pi_-(\theta) = \pi(\theta \mid H_-)$, $\pi_0(\theta) = \pi(\theta \mid H_0)$ and $\pi_+(\theta) = \pi(\theta \mid H_+)$, where, $\pi_0(\theta) = I(\theta = \theta_0)$ $\pi_-(\cdot)$ and $\pi_+(\cdot)$ are the densities with supports in $(-\infty, 0)$ and $(0, +\infty)$, respectively. Then the prior on θ can be written as

$$\pi(\theta) = p_-\pi_-(\theta)I(\theta < 0) + p_0 I(\theta = \theta_0) + p_+\pi_+(\theta)I(\theta > 0)\,. \qquad (38)$$

For a fixed prior π, the decision rule can be compared by comparing the points in the space

$$S(\pi) = \left\{ pFDR_-(\Gamma_-),\, pFDR_0(\Gamma_0),\, pFDR_+(\Gamma_+) : \delta \in D^* \right\},$$

where D^* is the class of randomized decision rules. Let us consider a subclass of decision rules $D \in D^*$ such that $pFDR_0(\Gamma_0)$ is constant for all $\delta \in D$. Let us consider two different sets of priors: $p = \{p(H_-), p(H_0), p(H_+)\}$ and $P' = \{p'(H_-), p(H_0), p'(H_+)\}$ and suppose that the following relations take place $p(H_-) > p'(H_-)$ and thus $p(H_+) < p'(H_+)$. Then the following fact can be proved.

Theorem 2. If δ_{CBM} and δ'_{CBM} denote CBM rules within the class D under the priors p and p', respectively, then

$$pFDR_-^{\delta_{CBM}}(\Gamma_-) \le pFDR_-^{\delta'_{CBM}}(\Gamma_-)\,,\ \ TDR_-^{\delta_{CBM}}(\Gamma_-) \ge TDR_-^{\delta'_{CBM}}(\Gamma_-)$$

and

$$pFDR_+^{\delta_{CBM}}(\Gamma_+) \ge pFDR_+^{\delta'_{CBM}}(\Gamma_+)\,,\ \ TDR_+^{\delta_{CBM}}(\Gamma_+) \le TDR_+^{\delta'_{CBM}}(\Gamma_+)\,.$$

The proof of this theorem is similar to the proof of theorem 1 of Ref. [1], therefore its shortened version for only the false discovery rate, adapted to the considered case, is given in Appendix. The validity of this theorem is clearly demonstrated by the computation results shown in Figs. 1 and 4.

Corollary 1. If δ_{CBM} and δ'_{CBM} denote CBM rules under the priors p and p', respectively, then

$$P_{\delta_{CBM}}(x \notin \Gamma_0 \mid H_-) \ge P_{\delta'_{CBM}}(x \notin \Gamma_0 \mid H_-) \ \text{ and } \ P_{\delta_{CBM}}(x \notin \Gamma_0 \mid H_+) \le P_{\delta'_{CBM}}(x \notin \Gamma_0 \mid H_+)\,,$$

where

$$P_{\delta_{CBM}}(x \notin \Gamma_0 \mid H_-) = \int_{R^n - \Gamma_0(\delta_{CBM})} p(x \mid H_-)dx$$

and

$$P_{\delta_{CBM}}(x \notin \Gamma_0 \mid H_+) = \int_{R^n - \Gamma_0(\delta_{CBM})} p(x \mid H_+)dx\,.$$

The proof of the corollary directly follows from the proof of Theorem 2 (see Fig. 1).

When testing the directional hypotheses, some authors (see, for example Shaffer [10] and Jones & Tukey [9]) offer to use the Type III errors rate which is defined as

$$Type-III \quad error \quad rate = P(x \in \Gamma_- \mid H_0) + P(x \in \Gamma_+ \mid H_0)\,. \tag{39}$$

Somewhat different definition of this term is offered in Ref. [6]. In particular, the type III errors involve inferring incorrectly the direction of the effect. For example, when the population value of the tested parameter is actually more than the null value, getting a sample value that is so much below the null value that you reject the null and conclude that the population value is also below the null value. In the considered case this means:

$$Type-III \quad error \quad rate = P(x \in \Gamma_- \mid H_+) + P(x \in \Gamma_+ \mid H_-)\,. \tag{40}$$

Let us denote Type III error rate in Eq. (39) as ERR_{III}^T and Type III error rate in Eq. (40) as ERR_{III}^K.

Theorem 3. For the considered directional hypotheses, $ERR_{III}^T > ERR_{III}^K$ always takes place and, when $\min_{\{i,j \in (-,0,+)\ i \neq j\}} \mathrm{div}(H_i, H_j) \to \infty$, both error rates tend to zero.

From the comparison of expressions Eq. (34) with Eqs. (39) and (40), it is seen that $pFDR_0(\Gamma_0) = ERR_{III}^T$ and $pFDR_-(\Gamma_-) + pFDR_+(\Gamma_+) > ERR_{III}^K$.

The ratio between $pFDR_-(\Gamma_-) + pFDR_+(\Gamma_+)$ and $ERR_{III}^T + ERR_{III}^K$ can be arbitrary in general.

From the above given, it is clear that CBM is a data-dependent test (see Eqs. (30) and (31)) similarly to the Fisher's $p-$ value test, in addition to the fact that it also computes Type I and Type II error probabilities like the Neyman-Pearson's approach (see Eqs. (26) and (27)), and uses a posteriori probabilities like the Bayes test (see Eqs. (28) and (29)).

4. CBM for the Normally Distributed Directional Hypotheses

For illustration of the fact that the results of CBM are more promoted than the results of Bayes and frequentist methods when testing the directional hypotheses, let us consider the example given in Ref. [1] for showing some advantage of the Bayes rule in comparison with the frequentist one.

Let sample $X_1, X_2, \ldots, X_n$ be derived from $N(\theta, \sigma^2)$ with known σ^2, $p(x \mid H_-)$ and $p(x \mid H_+)$ be the truncated $N(0, \omega_0^{-1}\sigma^2)$ (ω_0 known) densities over $(-\infty, 0)$ and $(0, +\infty)$, respectively.

Due to the above-mentioned sample, the arithmetic mean is sufficient statistics. For determination of hypotheses acceptance regions in Eqs. (17), (20) and (22), the following ratios must be determined:

$$\frac{1-p(H_- \mid \bar{x})}{1-p(H_0 \mid \bar{x})},\ \frac{1-p(H_- \mid \bar{x})}{1-p(H_+ \mid \bar{x})}$$

and

$$\frac{1-p(H_+ \mid \bar{x})}{1-p(H_0 \mid \bar{x})}.$$

Taking into account the conditions of the stated problem, after routine computation, we have

$$\frac{1-p(H_-\,|\,\bar{x})}{1-p(H_0\,|\,\bar{x})}=\frac{p(H_0\,|\,\bar{x})+p(H_+\,|\,\bar{x})}{p(H_-\,|\,\bar{x})+p(H_+\,|\,\bar{x})}=\frac{p_0\sqrt{n+\omega_0}\cdot\exp\{-u^2/2\}+2\omega_0 p_+\Phi(u)}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]},$$

$$\frac{1-p(H_-\,|\,\bar{x})}{1-p(H_+\,|\,\bar{x})}=\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_+\Phi(u)}{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))},$$

$$\frac{1-p(H_+\,|\,\bar{x})}{1-p(H_0\,|\,\bar{x})}=\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]}, \tag{41}$$

where $\Phi(\cdot)$ is the standard normal distribution function and $u=n\bar{x}/\sigma\sqrt{n+\omega_0}$

Application of these ratios to hypotheses acceptance regions in Eqs. (17), (20) and (22) gives

$$\Gamma_0=\left\{\bar{x}:\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_+\Phi(u)}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]}>\frac{K_0}{K_1}\cdot\frac{1}{\lambda}\right.$$

$$\left.\&\,\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]}>\frac{K_0}{K_1}\cdot\frac{1}{\lambda}\right\}, \tag{42}$$

$$\Gamma_-=\left\{\bar{x}:\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_+\Phi(u)}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]}<\frac{K_0}{K_1}\lambda\right.$$

$$\left.\&\,\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_+\Phi(u)}{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))}<\lambda\right\}. \tag{43}$$

$$\Gamma_+=\left\{\bar{x}:\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))}{2\sqrt{\omega_0}\left[p_-(1-\Phi(u))+p_+\Phi(u)\right]}<\frac{K_0}{K_1}\lambda\right.$$

$$\left.\&\,\frac{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_+\Phi(u)}{p_0\sqrt{n+\omega_0}\,\exp\{-u^2/2\}+2\sqrt{\omega_0}\,p_-(1-\Phi(u))}>\frac{1}{\lambda}\right\}. \tag{44}$$

In Eqs. (42)-(44), the Lagrange multiplier λ is determined so that, in condition Eq. (26), the equality was provided.

Finally the decision rule in the considered case is the following: if $\bar{x}$ belongs to only one of the regions Γ_0, Γ_- or Γ_+ determined by Eqs. (42)-(44), then the appropriate hypothesis is accepted. Otherwise, i.e. if $\bar{x}$ belongs to more than one of the considered regions or it does not belong to any of them, a decision is not made. In the first case, it is impossible to make a single decision, because more than one hypothesis is suspected to be true and, in the second case, it is impossible to make a decision. For making a decision, it is necessary to change the restriction level r_1 in Eq. (26) or to add one more observation to the sample.

4.1 Determination of the Lagrange Multiplier

As mentioned above, the Lagrange multiplier λ is determined so that the equality was provided in condition Eq.

(23). For the solution of Eq. (23), the computation of the following integrals is necessary

$$\int_{\Gamma_0} p(\bar{x} \mid H_0) d\bar{x}, \ \int_{\Gamma_-} p(\bar{x} \mid H_-) d\bar{x}$$

and

$$\int_{\Gamma_+} p(\bar{x} \mid H_+) d\bar{x}. \tag{45}$$

The first integral can be easily computed by the Monte-Carlo method. It is necessary to generate the random variables $\bar{x}$ with distribution law $p(\bar{x} \mid H_0) = N(\bar{x} \mid 0, \sigma^2/n)$ N times and to check the condition $\bar{x} \in \Gamma_0$ (see Eq. (42)). Let the condition $\bar{x} \in \Gamma_0$ be fulfilled $N_1 \le N$ times. Then:

$$\int_{\Gamma_0} p(\bar{x} \mid H_0) d\bar{x} \approx N_1 / N.$$

For computation of the second integral of Eq. (45), we have to generate the random variables $\bar{x}$ with distribution law $p(\bar{x} \mid H_-)$ N times and to check the condition $\bar{x} \in \Gamma_-$ (see Eq. (43)). Let the condition $\bar{x} \in \Gamma_-$ be fulfilled $N_2 \le N$ times. Then:

$$\int_{\Gamma_-} p(\bar{x} \mid H_-) d\bar{x} \approx N_2 / N.$$

Taking into account the specificity of the considered case, for distribution law $p(\bar{x} \mid H_-)$, we have:

$$p(\bar{x} \mid H_-) = \int_{-\infty}^{0} \frac{\sqrt{n}}{\sqrt{2\pi} \cdot \sigma} \exp\left\{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right\} \cdot \frac{2\sqrt{\omega_0}}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left\{-\frac{\omega_0^2 \theta^2}{2\sigma^2}\right\} d\theta$$

$$= \frac{2\sqrt{\omega_0}}{\sqrt{n+\omega_0}} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot (1-\Phi(u)) \exp\left\{-\frac{\omega_0 u^2}{2n}\right\},$$

where $u = \dfrac{n\bar{x}}{\sigma\sqrt{n+\omega_0}}$. Therefore,

$$p(\bar{x} \mid H_-) = \frac{2\sqrt{\omega_0}}{\sqrt{n+\omega_0}} \cdot \frac{\sqrt{n}}{\sqrt{2\pi} \cdot \sigma} \cdot \left(1-\Phi\left(\frac{n\bar{x}}{\sigma\sqrt{n+\omega_0}}\right)\right) \exp\left\{-\frac{\omega_0 n^2 \bar{x}^2}{2n\sigma^2 (n+\omega_0)}\right\}.$$

Let us denote $\sigma_1^2 = \dfrac{\sigma^2(n+\omega_0)}{n\omega_0}$, then $\sigma_1 = \dfrac{\sigma\sqrt{n+\omega_0}}{\sqrt{n}\sqrt{\omega_0}}$, and

$$p(\bar{x} \mid H_-) = 2 \cdot \left(1-\Phi\left(\frac{\sqrt{n}\bar{x}}{\sigma_1\sqrt{\omega_0}}\right)\right) \cdot N(\bar{x} \mid 0, \sigma_1^2), \ \ \bar{x} \in (-\infty, 0),$$

where $N(\bar{x} \mid 0, \sigma^2)$ is the normal distribution function with mathematical equal to zero and variance equal to σ^2.

For getting a sample of $\bar{x}$ with pdf $p(\bar{x} \mid H_-)$, it is necessary to solve the equation:

$$\int_{-\infty}^{\bar{x}} 2 \cdot \left(1-\Phi\left(\frac{\sqrt{n}y}{\sigma_1\sqrt{\omega_0}}\right)\right) \cdot N(y \mid 0, \sigma_1^2) dy = u, \ \ \bar{x} \in (-\infty, 0),$$

where u is the uniformly distributed random variable from the interval [0, 1], i.e. $u \sim U[0, 1]$.

Conditional distribution density of $\bar{x}$ at validity of H_+ is

$$p(\bar{x} \mid H_+) = \int_0^{+\infty} \frac{\sqrt{n}}{\sqrt{2\pi}\cdot\sigma} \exp\left\{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right\} \cdot \frac{2\sqrt{\omega_0}}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left\{-\frac{\omega_0^2\theta^2}{2\sigma^2}\right\} d\theta$$

$$= 2\cdot\Phi\left(\frac{\sqrt{n}\bar{x}}{\sigma_1\sqrt{\omega_0}}\right)\cdot N\left(\bar{x} \mid 0,\sigma_1^2\right), \bar{x}\in(0,+\infty).$$

For getting a sample of $\bar{x}$ with pdf $p(\bar{x} \mid H_+)$, it is necessary to solve the equation:

$$\int_0^{\bar{x}} 2\cdot\Phi\left(\frac{\sqrt{n}y}{\sigma_1\sqrt{\omega_0}}\right)\cdot N\left(y \mid 0,\sigma_1^2\right)dy = u\,, \quad \bar{x}\in(0,+\infty),$$

where $u \sim U[0,1]$.

5. Computation Results

For the reasons noted in the beginning of Section 4, let us compute a concrete example with the initial data from Ref. [1]: priori probabilities $p = \{p_-, p_0, p_+\} = \{0.3975, 0.3975, 0.205\}$ and $p' = \{p'_-, p'_0, p'_+\} = \{0.205, 0.3975, 0.3975\}$; the values of the loss functions $K_0 = K_1 = 1$; probabilities in restriction (24) $\alpha_0 = \alpha_- = \alpha_+ = 0.05$; coefficient $\omega_0 = 1$; variance $\sigma^2 = 1$; sample size $n = 100$; the probabilities were computed by simulating 10,000 samples from the appropriate populations. Computation results are given in Table 1.

By the results of Table 1, the following graphs are constructed: dependences of the probabilities of impossibility of acceptance of H_0 hypothesis on the arithmetic mean of observation results (Fig. 1), dependences of the probability of acceptance of H_0 hypothesis on the arithmetic mean of observation results (Fig. 2), dependences of the probabilities of acceptance of H_- and H_+ hypotheses on the arithmetic mean of observation results (Figs. 3 and 4) and dependences of the probabilities of rejection of H_- and H_+ hypotheses on the arithmetic mean of observation results (Figs. 5 and 6). From these graphs, the rightness of the above-described theoretical results and the advantage of CBM in comparison with the Bayes rule and, accordingly, with the frequentist method is obvious.

Table 1 The results of testing directional hypotheses using CBM and Bayes rules.

Used method	Lagrange multiplier	Averaged probability on the left-side of (23)	Meth. expectation of the sample	Hypotheses acceptance probabilities			Hypotheses rejection probabilities (probabilities of impossibility of acceptance of Hypotheses)			Probability of impossibility of making a decision
	λ		$\bar{x}$	H_0	H_-	H_+	H_0	H_-	H_+	
CBM at p	8.7213	0.9658	-0.5	0	0.9865	0	0.9865 (1)	0 (0.0135)	1	0.0135
CBM at p'	6.5014	0.9573		0	0.9837	0	0.9837 (1)	0 (0.0163)	1	0.0163
Bayes at p	1			0.0007	0.9993	0	0.9993	0.0007	1	0
Bayes at p'	1			0.0023	0.9975	0	0.9977	0.0025	1	0
CBM at p	8.5937	0.9653	-0.4	0.0001	0.8942	0	0.8942 (0.9999)	0.0001 (0.1058)	1	0.1057
CBM at p'	7.5	0.9576		0.0007	0.853	0	0.853 (0.9993)	0.0007 (0.147)	1	0.1463
Bayes at p	1			0.0154	0.9834	0	0.9846	0.0166	1	0
Bayes at p'	1			0.0338	0.9643	0	0.9662	0.0357	1	0
CBM at p	8.75	0.9654	-0.3	0.0023	0.5924	0	0.5924 (0.9977)	0.0023 (0.4076)	1	0.4053
CBM at p'	6.1718	0.9580		0.0328	0.5459	0	0.5459 (0.9672)	0.0331 (0.4541)	0.9997 (1)	0.4213
Bayes at p	1			0.1231	0.8724	0	0.8769	0.1276	1	0
Bayes at p'	1			0.1917	0.7981	0	0.8083	0.2019	1	0
CBM at p	8.75	0.9652	-0.2	0.0353	0.2148	0	0.2148 (0.9647)	0.0368 (0.7852)	0.9985 (1)	0.7499
CBM at p'	7.1875	0.9580		0.1613	0.1771	0	0.1771 (0.8387)	0.1696 (0.8229)	0.9917 (1)	0.6616
Bayes at p	1			0.4246	0.567	0	0.5754	0.433	1	0
Bayes at p'	1			0.5513	0.4373	0.0001	0.4487	0.5627	0.9999	0
CBM at p	8.75	0.9662	-0.1	0.1916	0.036	0	0.036 (0.8084)	0.2187 (0.964)	0.9729 (1)	0.7724
CBM at p'	6.5014	0.9573		0.4775	0.0303	0.0002	0.0305 (0.5225)	0.541 (0.9697)	0.9365 (0.9998)	0.492
Bayes at p	1			0.7891	0.2043	0.0009	0.2109	0.7957	0.9991	0
Bayes at p'	1			0.874	0.1182	0.0025	0.126	0.8818	0.9975	0
CBM at p	8.75	0.9662	0	0.3915	0.0022	0.0016	0.0038 (0.6085)	0.5857 (0.9978)	0.8058 (0.9984)	0.6047
CBM at p'	7.0312	0.9579		0.534	0.0017	0.0037	0.0054 (0.466)	0.8532 (0.9983)	0.6808 (0.9963)	0.4606

Table 1 to be continued

Bayes at p	1			0.9493	0.033	0.0147	0.0507	0.967	0.9853	0
Bayes at p'	1			0.9492	0.0148	0.0331	0.0508	0.9852	0.9669	0
CBM at p	8.75	0.9652		0.3334	0	0.0227	0.0227 (0.6666)	0.888 (1)	0.4454 (0.9773)	0.6439
CBM at p'	6.5429	0.9574	0.1	0.2935	0.0001	0.0491	0.0492 (0.7065)	0.983 (0.9999)	0.3105 (0.9509)	0.6573
Bayes at p	1			0.8683	0.0023	0.1221	0.1317	0.9977	0.8779	0
Bayes at p'	1			0.7862	0.0006	0.2081	0.2138	0.9994	0.7919	0
CBM at p	8.75	0.9662		0.1166	0	0.1587	0.1587 (0.8834)	0.9874 (1)	0.1292 (0.8413)	0.7247
CBM at p'	6.2402	0.9574	0.2	0.0759	0	0.2633	0.2633 (0.9241)	0.9992 (1)	0.0767 (0.7367)	0.6608
Bayes at p	1			0.5595	0.0001	0.4296	0.4405	0.9999	0.5704	0
Bayes at p'	1			0.4247	0.0001	0.5666	0.5753	0.9999	0.4334	0
CBM at p	8.75	0.9652		0.0166	0	0.5095	0.5095 (0.9834)	0.9995 (1)	0.0171 (0.4905)	0.4739
CBM at p'	7.1875	0.9580	0.3	0.005	0	0.6214	0.6214 (0.995)	1	0.005 (0.3786)	0.3736
Bayes at p	1			0.1928	0	0.7994	0.8072	1	0.2006	0
Bayes at p'	1			0.1182	0	0.8792	0.8818	1	0.1208	0
CBM at p	8.75	0.9662		0.0005	0	0.8426	0.8426 (0.9995)	1	0.0005 (0.1574)	0.1569
CBM at p'	6.1718	0.9580	0.4	0	0	0.9126	0.9126 (1)	1	0 (0.0874)	0.0874
Bayes at p	1			0.0311	0	0.9668	0.9689	1	0.0332	0
Bayes at p'	1			0.0164	0	0.983	0.9836	1	0.017	0
CBM at p	8.75	0.9652		0	0	0.9775	0.9775 (1)	1	0 (0.0225)	0.0225
CBM at p'	7.5	0.9576	0.5	0	0	0.9896	0.9896 (1)	1	0 (0.0104)	0.0104
Bayes at p	1			0.0023	0	0.9976	0.9977	1	0.0024	0
Bayes at p'	1			0.0007	0	0.9992	0.9993	1	0.0008	0

Remark: The probabilities of impossibility of acceptance of hypotheses are given in the brackets of the columns of hypotheses rejection probabilities.

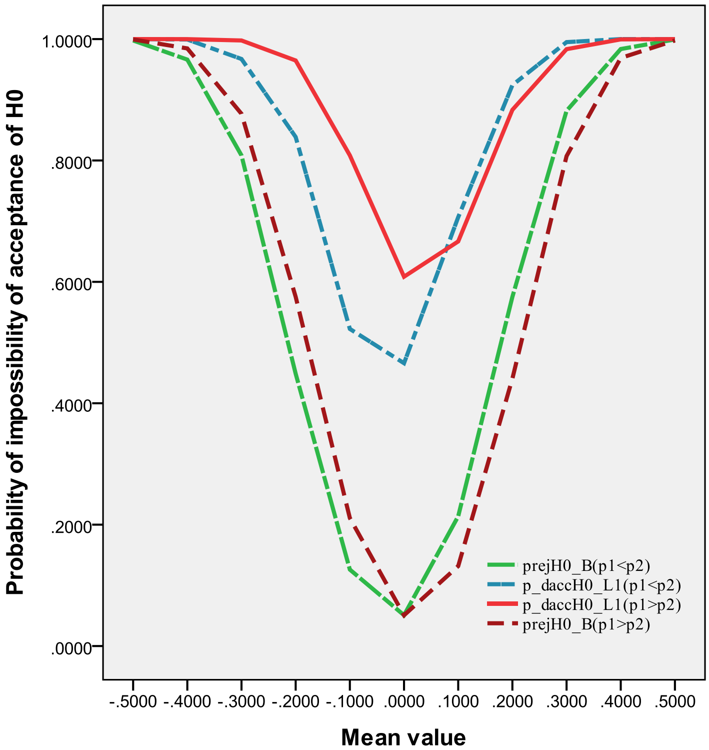

Fig. 1 Dependences of the probabilities of impossibility of acceptance of H_0 hypothesis on the arithmetic mean of observation results. Bayes rule; L1-CBM for losses (4); $P1 \equiv P(H_-)$ and $P2 \equiv P(H_+)$.

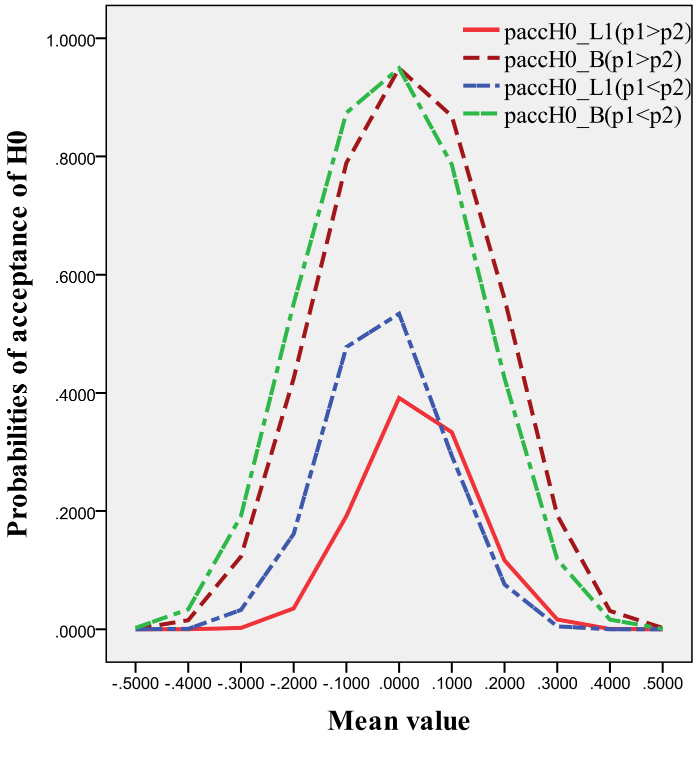

Fig. 2 Dependences of the probability of acceptance of H_0 hypothesis on the arithmetic mean of observation results.

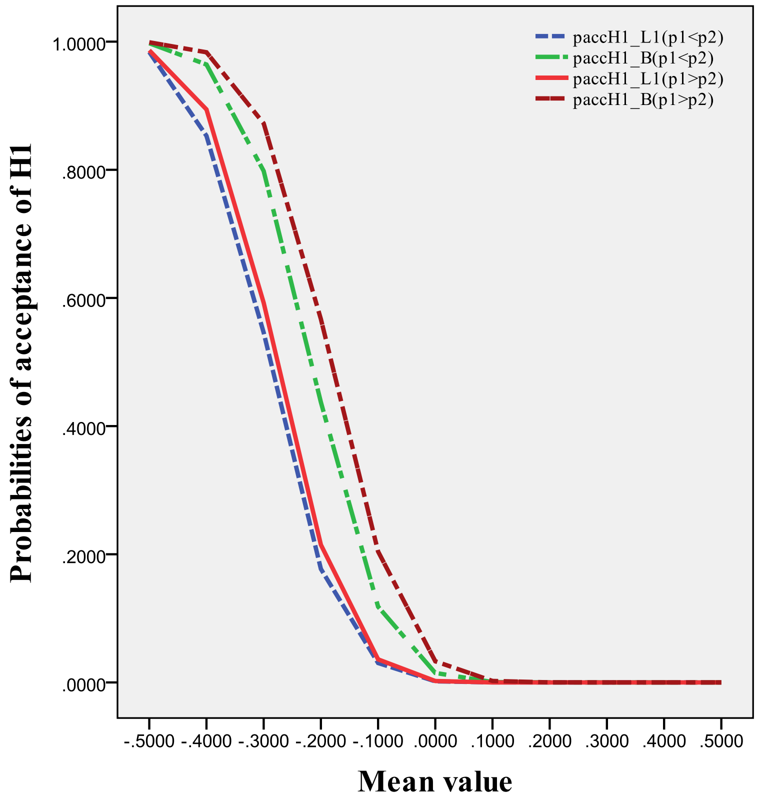

Fig. 3 Dependences of the probabilities of acceptance of H_- hypothesis on the arithmetic mean of observation results.

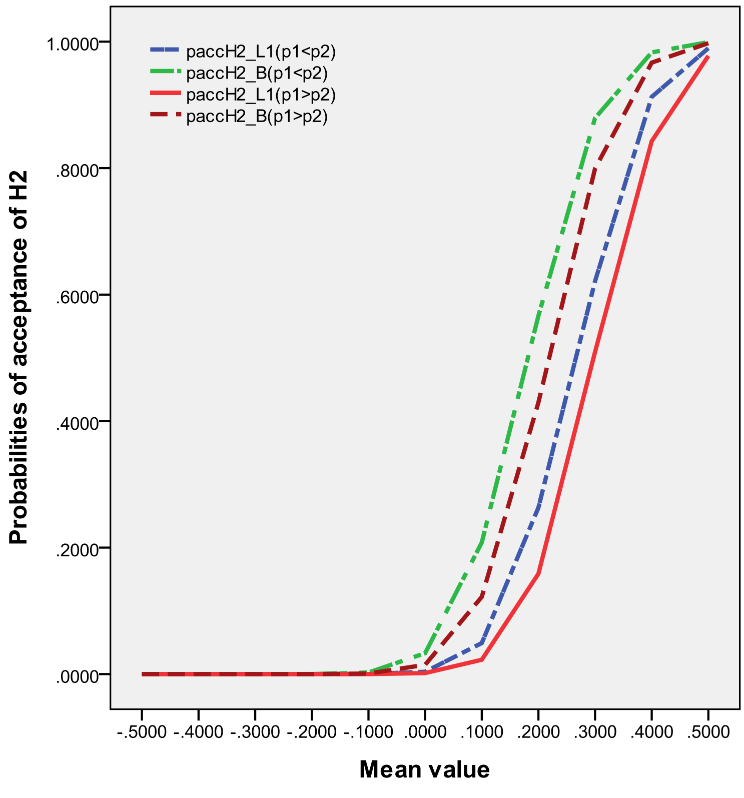

Fig. 4 Dependences of the probabilities of acceptance of H_+ hypothesis on the arithmetic mean of observation results.

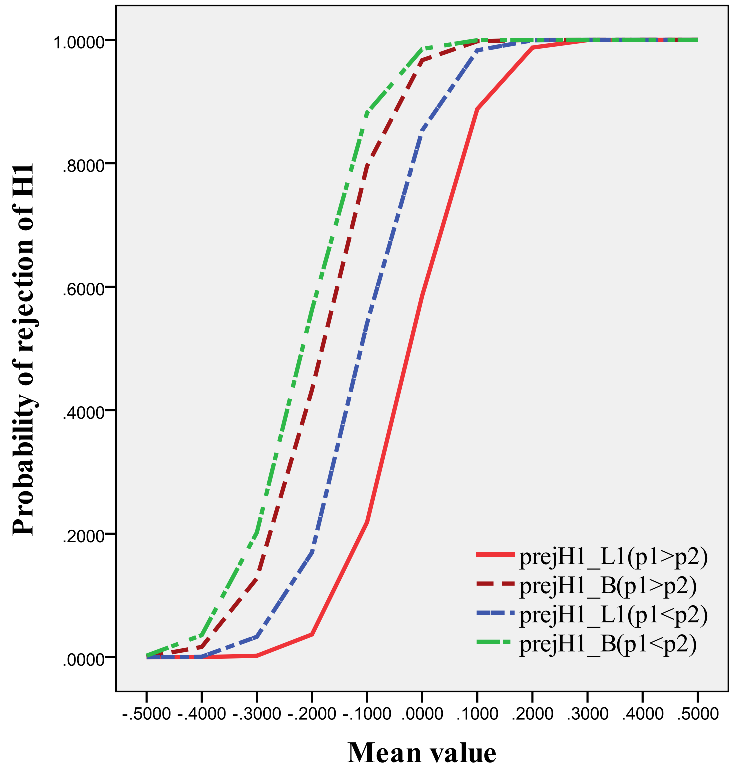

Fig. 5 Dependences of the probabilities of rejection of H_- hypotheses on the arithmetic mean of observation results.

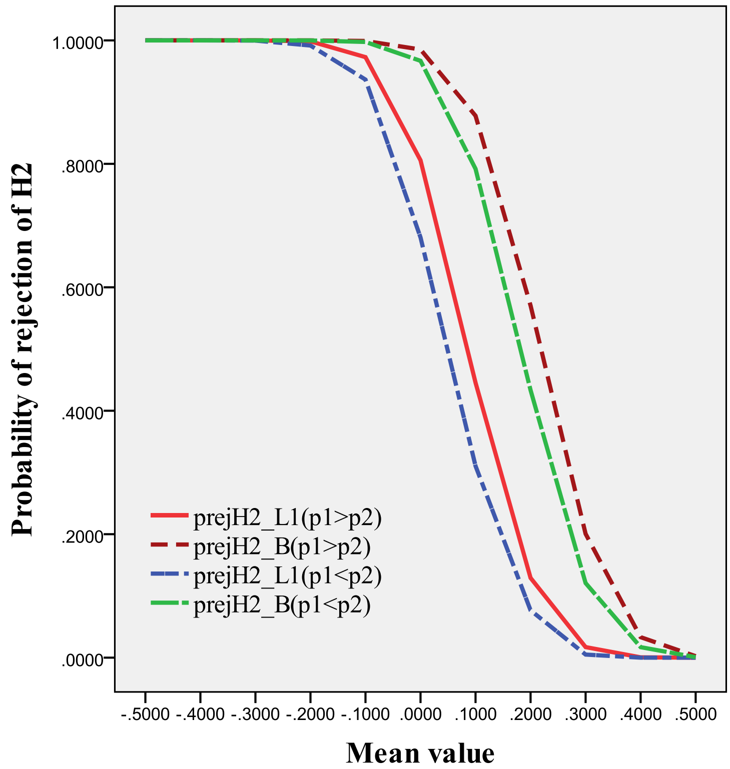

Fig. 6 Dependences of the probabilities of rejection of H_+ hypotheses on the arithmetic mean of observation results.

6. Discussion

CBM is more sensitive to the change of priori probabilities than the Bayes test because in CBM priori probabilities are multiplied by probabilities of significance levels, and hence the change in priori probabilities changes the restriction level in Eq. (2) more significantly and, accordingly, changes more significantly the decision-making regions.

From the specificity of the decision rule in CBM, the following relations take place among the computed probabilities:

(1) *(prob. of rejec. of H_0) = (prob. of rejec. of all hypotheses) + (prob. of rejec. of H_0 and H_+ and accep. of H_-) + (prob. of rejec. of H_0 and H_- and accep. of H_+)*;

(2) *(prob. of accep. of H_0) + (prob. of accep. of H_-) + (prob. of accep. of H_+) + (prob. of no making of decision) = 1*;

(3) (a) at absence of intersecting regions: *(prob. of accep. of H_0) + (prob. of rejec. of H_0) = 1*;

(b) at intersecting regions: *(prob. of accep. of H_0) + (prob. of rejec. of H_0) + (prob. of suspicion of more than one hypotheses to be true) = 1*;

(4) *(prob. of rejec. of H_- and H_+ and acceptance of H_0) = (prob. of $x \in \Gamma_0$) - (prob. of accep. of H_0 and H_- and rejection of H_+) - (prob. of accep. of H_0 and H_- and rejec. of H_-) - (prob. of accep. of all H_0, H_- and H_+).*

(5) Summary risk Eq. (11) can be computed using the appropriate computation results as follows:

$$SR = p(H_-)\cdot[(1-(prob.\ of\ accep\ of\ H_1))-(prob.\ of\ impos.\ of\ H_1)]$$
$$+\ p(H_0)\cdot[(1-(prob.\ of\ accep\ of\ H_0))-(prob.\ of\ impos.\ of\ H_0)]$$
$$+\ p(H_+)\cdot[(1-(prob.\ of\ accep\ of\ H_+))-(prob.\ of\ impos.\ of\ H_+)].$$

The dependences of SR on Lagrange multiplier are shown in Fig. 7. They clearly demonstrate the validity of theorem 1.

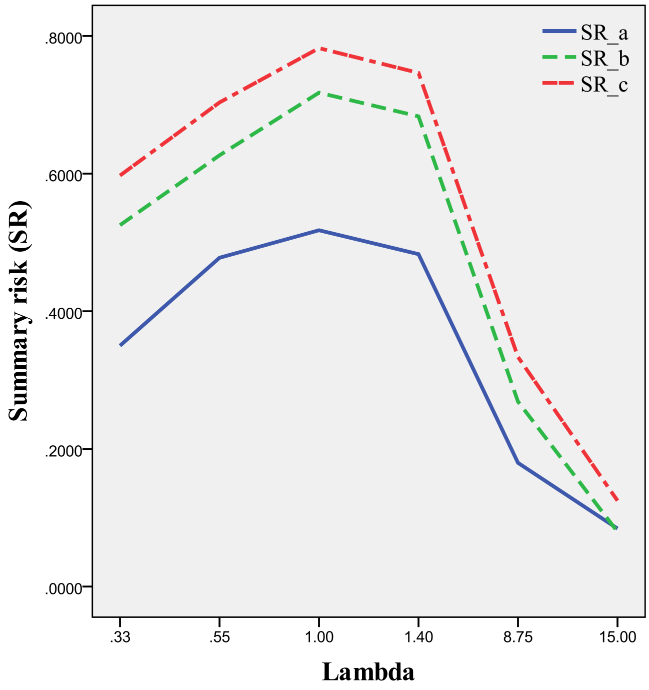

Fig. 7 Dependence of the summary risk (SR) on the Lagrange multiplier.

The graphs of summary risk (SR) are constructed by computed values of SR, using simulated samples at supposition of the validity of the following hypotheses:

a) $H_- : \bar{x} = -0.1$; $H_0 : \bar{x} = 0$; $H_+ : \bar{x} = 0.1$;

b) $H_- : \bar{x} = -0.2$; $H_0 : \bar{x} = 0$; $H_+ : \bar{x} = 0.2$;

c) $H_- : \bar{x} = -0.2$; $H_0 : \bar{x} = 0$; $H_+ : \bar{x} = 0.1$.

((6) Type III error rate Eq. (39) can be computed using the computation results by the following ratio:

ERR_{III}^{T} = (probab. of accep. of H_- | H_0 is true) + (probab. of accep. of H_+ | H_0 is true)

and Type III error rate Eq. (40) can be computed using the computation results as follows:

ERR_{III}^{K} = (probab. of accep. of H_- | H_+ is true) + (probab. of accep. of H_+ | H_- is true).

Appropriate computed results are shown in Fig. 8. They clearly demonstrate the validity of theorem 3.

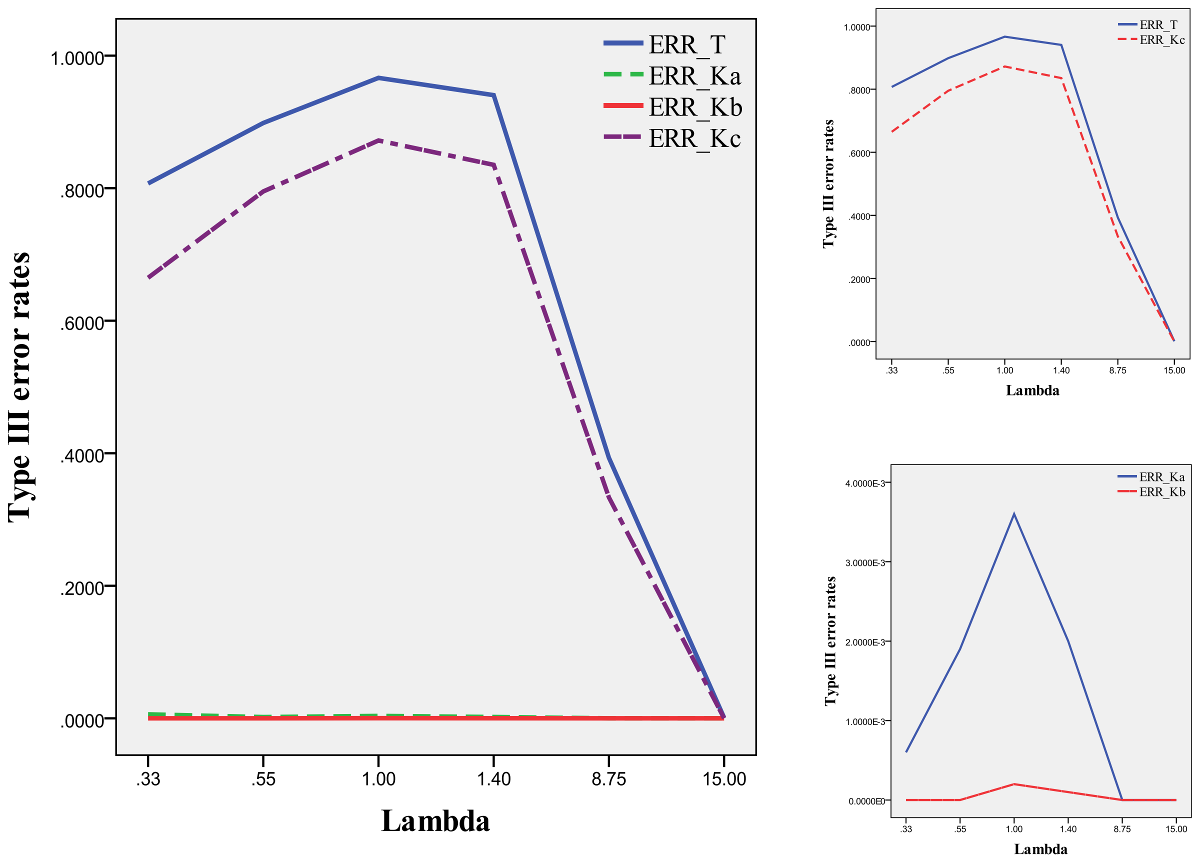

Fig. 8 Dependences of type III error rates on the Lagrange multiplier.

The graphs of type III error rates are constructed by computed values of ERRs, using the simulated samples at supposition of the validity of the following hypotheses:

a) $H_- : \bar{x} = -0.1$; $H_0 : \bar{x} = 0$; $H_+ : \bar{x} = 0.1$;

b) $H_- : \bar{x} = -0.2$; $H_0 : \bar{x} = 0$; $H_+ : \bar{x} = 0.2$;

c) $H_{-}:\bar{x}=-0.2$; $H_0:\bar{x}=0$; $H_{+}:\bar{x}=0.1$.

Remark: The values of different Type III error rates differ considerably. Therefore, the character of the change in the graphs of ERR_{III}^{K} for hypotheses (a) and (b) is not clear from the graph given on the left side of Fig. 8. For avoiding this inconvenience, the graphs of type III error rates are grouped depending on their values and presented in the two right graphs of Fig. 8 (ERR_{III}^{T} and ERR_{III}^{K}c) on the upper graph and ERR_{III}^{K}a) and ERR_{III}^{K}b) on the lower graph.

7. Conclusion

Generalization of CBM for arbitrary loss functions and its application for testing the directional hypotheses is offered in the paper. The advantage of CBM in comparison with Bayes and frequentist methods is theoretically proved and clearly demonstrated by a concrete computed example. The advantages of the use of CBM for testing the directional hypotheses are: (1) alongside with priori probabilities and loss functions, it uses the significance levels of hypotheses for sharpening the sensitivity concerning direction; (2) it makes decisions more carefully and with given reliability; (3) less values of SR and Type III error rates correspond to it. CBM allows making a decision with required reliability if the existing information is sufficient, otherwise it is necessary to increase the information or to reduce the required reliability of the made decision. CBM surpasses the Bayes and frequentist methods with guaranteed reliability of made decisions.

References

[1] Bansal, N. K., and Sheng, R. 2010. "Beyesian Decision Theoretic Approach to Hypothesis Problems with Skewed Alternatives." *J. Statistical Planning and Inference* 140: 2894-903.

[2] Lehmann, E. L. 1950. "Some Principles of the Theory of Testing Hypotheses." *The Annals of Mathematical Statistics* 20 (1): 1-26.

[3] Lehmann, E. L. 1957. "A Theory of Some Multiple Decision Problems I." *The Annals of Mathematical Statistics* 28: 1-25.

[4] Lehmann, E. L. 1957. "A Theory of Some Multiple Decision Problems II." *The Annals of Mathematical Statistics* 28: 547-72.

[5] Bahadur, R. R. 1952. "A Property of the *t*-Statistics." *Sankhya* 12: 79-88.

[6] Kaiser, H. F. 1960. "Directional Statistical Decisions." *Psychological Review* 67: 160-7.

[7] Leventhal, L., and Huynh, C. 1996. "Directional Decisions for Two-Tailed Tests: Power, Error Rates, and Sample Size." *Psychological Methods* 1: 278-92.

[8] Finner, H. 1999. "Stepwise Multiple Test Procedures and Control of Directional Errors." *The Annals of Statistics* 27 (1): 274-89.

[9] Jones, L. V., and Tukey, J. W. 2000. "A Sensible Formulation of the Significance Test." *Psychological Methods* 5 (4): 411-4.

[10] Shaffer, J. P. 2002. "Multiplicity, Directional (Type III) Errors, and the Null Hypothesis." *Psychological Methods* 7 (3): 356-69.

[11] Bansal, N. K., and Miescke, K. J. 2013. "A Bayesian Decision Theoretic Approach to Directional Multiple Hypotheses Problems." *J. of Multivariate Analysis* 120: 205-15.

[12] Bansal, N. K., Hamedani, G. G., and Maadooliat, M. 2015. "Testing Multiple Hypotheses with Skewed Alternatives." *Biometrics* 72 (2): 494-502.

[13] Bansal, N. K., Hamedani, G. G., and Sheng, R. 2012. "Bayesian Analysis of Hypothesis Testing Problems for General Population: A Kullback-Leibler Alternative." *J. of Statistical Planning and Inference* 142: 1991-8.

[14] Kachiashvili, G. K., Kachiashvili, K. J., and Mueed, A. 2012. "Specific Features of Regions of Acceptance of Hypotheses in Conditional Bayesian Problems of Statistical Hypotheses Testing." *Sankhya: The Indian J. of Statistics* 74 (1): 112-25.

[15] Kachiashvili, K. J., Hashmi, M. A., and Mueed A. 2012. "Sensitivity Analysis of Classical and Conditional Bayesian Problems of Many Hypotheses Testing." *Communications in Statistics—Theory and Methods* 41 (4): 591-605.

[16] Kachiashvili, K. J., and Mueed, A. 2013. "Conditional Bayesian Task of Testing Many Hypotheses." *Statistics* 47 (2): 274-93.

[17] Kachiashvili, K. J. 2011. "Investigation and Computation of Unconditional and Conditional Bayesian Problems of Hypothesis

Testing." *ARPN J. of Systems and Software* 1 (2): 47-59.

[18] Kachiashvili, K. J. 2014. "Comparison of Some Methods of Testing Statistical Hypotheses. Part I. Parallel Methods and Part II. Sequential Methods." *Int. J. of Statistics in Medical Research* 3: 174-97.

[19] Kachiashvili, K. J. 2014. "The Methods of Sequential Analysis of Bayesian Type for the Multiple Testing Problem." *Sequential Analysis* 33 (1): 23-38.

[20] Kachiashvili, K. J. 2015. "Constrained Bayesian Method for Testing Multiple Hypotheses in Sequential Experiments." *Sequential Analysis*: *Design Methods and Applications* 34 (2): 171-86.

[21] Kachiashvili, K. J. 2016. "Constrained Bayesian Method of Composite Hypotheses Testing: Singularities and Capabilities." *Int. J. of Statistics in Medical Research* 5 (3): 135-67.

[22] Kachiashvili K. J. 2018. *Constrained Bayesian Methods of Hypotheses Testing: A New Philosophy of Hypotheses Testing in Parallel and Sequential Experiments*. New York: Nova Science Publishers, Inc., p. 456.

[23] Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer.

[24] Kachiashvili, K. J. 1989. *Bayesian Algorithms of Many Hypothesis Testing*. Tbilisi: Ganatleba.

[25] Kachiashvili, K. J. 2003. "Generalization of Bayesian Rule of Many Simple Hypotheses Testing." *Int. J. of Information Technology & Decision Making* 2 (1): 41-70.

[26] Storey, J. D. 2003. "The Positive False Discovery Rate: A Bayesian Interpretation and the *q*-Value." *The Annals of Statistics* 31 (6): 2013-35.

[27] Kullback, S. 1978. *Information Theory and Statistics*. New York: Wiley & Sons.

Appendix

Proof of the Theorem 1. It is known that decision-making regions in the Bayesian rule satisfy conditions $\bigcup_{i=1}^{S}\Gamma_i^B = R^n$ and $\Gamma_i^B \bigcap \Gamma_j^B = \varnothing$, $i, j = 1, \ldots, S$, $i \neq j$. It is proved (see, for example, Kachiashvili et al. [14]; Kachiashvili & Mueed [16]) that in all tasks of CBM, when λ differs from 1, in observation space R^n, there appear sub-spaces of intersection of hypotheses acceptance regions or sub-spaces which do not belong to any region of acceptance of hypotheses. Both kinds of sub-spaces are the more than the more differs λ from 1 and, when $\lambda \to \infty$ or $\lambda \to 0$, their union coincides with observation space R^n, i.e. decision-making regions become empty (see hypotheses acceptance regions Eqs. (3) or (5) and (10)). In the first case, hypotheses acceptance regions are reduced by the intersection sub-region and, in the second case, hypotheses acceptance regions are reduced by the regions that do not belong to any region of acceptance of hypotheses. Thus in both cases (when $\lambda > 1$ and when $\lambda < 1$) hypotheses acceptance regions are reduced in comparison with the case of $\lambda = 1$, and in the limits ($\lambda \to \infty$ or $\lambda \to 0$) hypotheses acceptance regions become empty. Since in general in CBM $\lambda \neq 1$, the hypotheses acceptance regions are less than the regions when $\lambda = 1$. The hypotheses acceptance regions are the more reduced the more is the difference between λ and 1. Since SR of making the incorrect decision (11) is defined on these regions, to the reduced regions corresponds the reduced SR and vice versa. This proves the theorem.

Proof of the Theorem 2. Bayes Risk of a decision rule δ under the prior (38) is given by

$$r_\delta = p_- \cdot pFDR_-^{\delta_{CBM}} + p_0 \cdot pFDR_0^{\delta_{CBM}} + p_+ \cdot pFDR_+^{\delta_{CBM}} .$$

Thus, since δ_{CBM} and δ'_{CBM} are the constrained Bayesian rules under π and π', respectively,

$$p_- \cdot pFDR_-^{\delta_{CBM}} + p_0 \cdot pFDR_0^{\delta_{CBM}} + p_+ \cdot pFDR_+^{\delta_{CBM}} \le p_- \cdot pFDR_-^{\delta'_{CBM}} + p_0 \cdot pFDR_0^{\delta'_{CBM}} + p_+ \cdot pFDR_+^{\delta'_{CBM}} ,$$

$$p'_- \cdot pFDR_-^{\delta'_{CBM}} + p_0 \cdot pFDR_0^{\delta'_{CBM}} + p'_+ \cdot pFDR_+^{\delta'_{CBM}} \le p'_- \cdot pFDR_-^{\delta_{CBM}} + p_0 \cdot pFDR_0^{\delta_{CBM}} + p'_+ \cdot pFDR_+^{\delta_{CBM}} .$$

Now, since $pFRD_0$ is constant within the class D, and since δ_{CBM} and δ'_{CBM} belong to the class D,

$$p_- \cdot pFDR_-^{\delta_{CBM}} + p_+ \cdot pFDR_+^{\delta_{CBM}} \le p_- \cdot pFDR_-^{\delta'_{CBM}} + p_+ \cdot pFDR_+^{\delta'_{CBM}}$$

$$p'_- \cdot pFDR_-^{\delta'_{CBM}} + p'_+ \cdot pFDR_+^{\delta'_{CBM}} \le p'_- \cdot pFDR_-^{\delta_{CBM}} + p'_+ \cdot pFDR_+^{\delta_{CBM}}$$

which implies that

$$p_- \cdot \left(pFDR_-^{\delta_{CBM}} - pFDR_-^{\delta'_{CBM}}\right) + p_+ \cdot \left(pFDR_+^{\delta_{CBM}} - pFDR_+^{\delta'_{CBM}}\right) \le 0$$

$$p'_- \cdot \left(pFDR_-^{\delta_{CBM}} - pFDR_-^{\delta'_{CBM}}\right) + p'_+ \cdot \left(pFDR_+^{\delta_{CBM}} - pFDR_+^{\delta'_{CBM}}\right) \ge 0 . \quad \text{(A.1)}$$

Now, if we will denote $\left(pFDR_-^{\delta_{CBM}} - pFDR_-^{\delta'_{CBM}}\right)$ by x and $\left(pFDR_+^{\delta_{CBM}} - pFDR_+^{\delta'_{CBM}}\right)$ by y and will consider system of Eq. (A.1) relatively x and y, we will easily be convinced in the validity of the theorem.

Proof of Theorem 3. If we recall the character of considered directional hypotheses and the fact that the increasing divergence among hypotheses entails a decrease in the probabilities of errors of the first and the second types at hypotheses testing, and, in the limit, when $\min_{\{i,j \in (-,0,+)\ i \neq j\}} \text{div}(H_i, H_j) \to \infty$, there takes place $\alpha \to 0$ and $\beta \to 0$ [27], we will be convinced in the validity of the theorem.