

# Constrained Bayesian Method of Composite Hypotheses Testing: Singularities and Capabilities

K.J. Kachiashvili\*

Georgian Technical University, 77, st. Kostava, Tbilisi, 0175, I. Vekua Institute of Applied Mathematics of the Tbilisi State University, 2, st. University, Tbilisi, 0179, Georgia

**Abstract:** The paper deals with the constrained Bayesian Method (CBM) for testing composite hypotheses. It is shown that, similarly to the cases when CBM is optimal for testing simple and multiple hypotheses in parallel and sequential experiments, it keeps the optimal properties at testing composite hypotheses. In particular, it easily, without special efforts, overcomes the Lindley's paradox arising when testing a simple hypothesis versus a composite one. The CBM is compared with Bayesian test in the classical case and when the a priori probabilities are chosen in a special manner for overcoming the Lindley's paradox. Superiority of CBM against these tests is demonstrated by simulation. The justice of the theoretical judgment is supported by many computation results of different characteristics of the considered methods.

**Keywords:** CBM, Bayesian test, Composite hypotheses, Lindley's paradox, Hypotheses testing.

## 1. INTRODUCTION

In [1] was offered Constrained Bayesian Methods (CBM) of statistical hypotheses testing. The CBM have all positive characteristics of well-known classical approaches. It is a data-dependent measure like Fisher's test for making a decision, it uses a posteriori probabilities like the Jeffreys test and computes error probabilities Type I and Type II like the Neyman-Pearson's approach does [2]. The parallel and sequential methods for testing many (more than or equal to two) and multiple simple hypotheses are considered in the works [2, 3-6, 7-9]. Also there are given the results of comparison of CBM with the Fisher, Jeffreys, Neyman-Pearson, Berger (parallel and sequential), Wald (sequential), Bonferroni and step-up and step-down methods of multiple hypotheses testing, which shows the unique properties of this method. In particular, it gives more logical and reliable decisions, using information contained in the sample more completely than the existing methods. Continuing the development and examination of CBM, below we offer its application for testing composite hypotheses.

A lot of works are dedicated to the problem of testing composite hypotheses. In particular, a compact and exhaustive review of the history of development of the theory of hypotheses testing (including composite ones) is given in the paper [10]. In this case Neyman and Pearson replace the density at a single parameter value with the maximum of the density over all parameters in that hypothesis. The maximum likelihood

ratio test rejects the null hypothesis for large values of  $\sup_{\theta \in \Theta_A} f_{\theta}(x) / \sup_{\theta \in \Theta_0} f_{\theta}(x)$ . This is a useful method for finding good testing procedures. Stein's method [11, 12] integrates the density over  $\theta$  using special measures to obtain the density of the maximum invariant statistic, which can then be used to analyze the problems; for example, to find the uniformly most powerful invariant test. The Bayesian approach to hypothesis testing uses the integrated likelihood ratio  $\int_{\Omega_A} f_{\theta}(x) \gamma_A(\theta) d\theta / \int_{\Omega_0} f_{\theta}(x) \gamma_0(\theta) d\theta$  for making a decision.

Here  $\gamma_A$  is the density of the parameter  $\theta$  at the alternative supposition,  $\gamma_0$  is the density of the parameter at the null supposition. In the Bayesian statement, the a priori probability that the null hypothesis is true  $\pi_0$  is necessary. Hence it is clear that, in the Bayesian approach, the problems of proper choice of the densities  $\gamma_0$  and  $\gamma_A$  and the a priori probability arise. For overcoming this problem, in [13] was recommended "reference" priors  $\gamma_A$  and  $\gamma_0$  in many common problems. They must be carefully chosen so that the priors would not overwhelm the data. Other approaches include developing priors using imaginary prior data, and using part of the data as a training sample [14, 15]. The conditioning strategy proposed in [16] for the simple versus simple hypothesis was generalized to testing a precise null hypothesis versus a composite alternative hypothesis in [17]. In the latter paper, it is shown that the conditional frequentist method of testing a precise hypothesis can be made virtually equivalent to Bayesian testing. The method is adapted to sequential testing involving precise and composite hypotheses in [18]. Testing of a composite null hypothesis versus a composite alternative when both have a related invariance structure is considered in [19]. There were developed conditional frequentist tests that allowed reporting the data-dependent error probabilities. It was shown that the resulting tests were also Bayesian tests,

\*Address correspondence to this author at the Georgian Technical University, 77, st. Kostava, Tbilisi, 0175, I. Vekua Institute of Applied Mathematics of the Tbilisi State University, 2, st. University, Tbilisi, 0179, Georgia; Tel: (+99532) 223 7247; E-mail: k.kachiashvili@gtu.edu.ge, kkachiashvili@gmail.com

because the reported frequentist error probabilities were also the posterior probabilities of the hypotheses under definite choices of the prior distribution. The new procedures are illustrated in a variety of applications to model selection and testing of multivariate hypothesis. The paper considered the case when both hypotheses were composite, the situation that arises most frequently in practice. The Bayesian test for simple versus two-sided hypotheses in the multivariate case was developed in [20]. The Bayesian and the p-value approaches were compared. A better approximation was obtained by the Bayesian approach, because the p-value is in the range of the Bayesian measures of evidence. Testing of simple versus composite alternative hypotheses concerning the mean vector of the asymptotic normal distribution was given in [21]. The numerical algorithm used to compute the inverses were considered. There were compared different weighting strategies and local asymptotic powers of different test procedures. As a result, it was inferred that no test was uniformly dominated by another, and the ranking of the powers varied significantly with the direction of the alternative, and it is different if the alternative was fixed or local. For simple and composite null hypotheses, the likelihood ratio test (LR test), Wald's test, and Rao's score test were derived and turned out to have simple representations in [22]. The asymptotic distribution of the corresponding test statistics under the null hypothesis was stated, and, asymptotic optimality of the LR test in the case of a simple null hypothesis was shown. A power study was performed to measure and compare the quality of the tests for both simple and composite null hypotheses.

The problem of testing simple versus composite hypotheses using the Bayesian approach was considered in [23]. The method of choosing the prior distribution of hypotheses with consideration of the simple basic hypotheses versus the composite alternative hypothesis for avoiding the so-called Lindley's paradox was offered. The essence of the paradox is in the following: for any fixed prior probability of the basic hypothesis, the posterior probability of this hypothesis can be made as close to one as desired, *whatever the data* for a sufficiently large prior value of some parameter of the considered distribution at the alternative hypothesis. Similar behavior was founded in [24, 25] when a uniform distribution over a finite interval was chosen for the parameter under investigation at the alternative hypothesis (though the same type of a result is hold with any other distributional assumption). Namely, the posterior probability of the basic hypothesis tends to one as the size of the interval increases. The method offered in [23] was concretized for the cases of normal and uniform distributions of the parameter of mathematical expectation under a composite alternative hypothesis.

Below we apply CBM to composite hypotheses testing for the cases considered in [23]. The results of investigation show high quality of CBM in these cases as well. In particular, it preserves all positive properties that it has at testing simple and multiple hypotheses. Moreover it naturally overcomes the problem of the Lindley's paradox arising in classical methods based on the likelihood ratio. These facts are shown theoretically and on the basis of computation of the concrete examples taken from work [23]. The Application of CBM to composite hypotheses testing is described in Section 2. The Lindley's paradox in classical hypothesis testing and CBM is considered in Section 3. Computation results of some examples are presented in Section 4. The discussion of the problem and conclusions made are given in Sections 5 and 6, respectively.

## 2. CBM FOR TESTING COMPOSITE HYPOTHESES

Let us consider the problem of testing composite hypotheses. We assume that the data are represented by  $x$ , which has density  $f_{\theta}(x)$  for parameter  $\theta$ . Let us test the hypotheses

$$H_i : \theta \in \Theta_i, \quad (1)$$

where  $\Theta_i$ ,  $i = 1, \dots, S$  ( $S \geq 2$ ), are the disjoint subsets of the parameter space, i.e.  $\Theta_i \cap \Theta_j = \emptyset$ ,  $i \neq j$ .

For testing hypotheses (1) there are different ways of action. As was mentioned above, one of possible ways was offered by Neyman and Pearson. They replaced the density at a single parameter value with the maximum of the density over all parameters in that hypothesis, i.e. instead of density  $f_{\theta}(x)$ , the conditional density  $f_{\theta}(x | H_i) = \sup_{\theta \in \Theta_i} f_{\theta}(x)$  is considered. Stein's method integrates the density over  $\theta$  using special measures to obtain the density of the maximum invariant statistic, which can then be used to analyze the problems. Here, instead of density  $f_{\theta}(x)$ , the integrated density  $f_{\theta}(x | H_i) = \int_{\Theta_i} f_{\theta}(x) \gamma_i(\theta) d\theta$  over the subset  $\Theta_i$ , where the density  $\gamma_i$  of the parameter conditional on it being in the hypothesis  $H_i$  is used. As was mentioned in [10], the Neyman and Pearson's method is an extremely useful method for finding good testing procedures and the Stein's method allows us to find the uniformly most powerful invariant test, though there arises the problem of finding the proper densities  $\gamma_i$ ,  $i = 1, \dots, S$ . When testing composite hypotheses (1) using CBM, we can use one of the possible decision rules of different statements of CBM depending on the final aim [1, 4, 9]. In these decision rules, one must use densities  $f_{\theta}(x | H_i)$ ,  $i = 1, \dots, S$ , conditioned in

accordance with Neyman-Pearson or Stein’s methods depending on the existed information and desired properties of decisions made. The integrated densities are used in the examples considered below with the aim to study the CBM on the Lindley’s paradox problem.

**3. LINDLEY’S PARADOX AND CBM**

The likelihood ratio tests are based on the relations of the densities  $f_{\theta}(x | H_i)$ , and it is logically clear that, the bigger is the difference among these densities, the bigger is evidence for making a proper decision on the basis of the observation results. When the integrated density  $f_{\theta}(x | H_i) = \int_{\Theta_i} f_{\theta}(x)\gamma_i(\theta)d\theta$  is used for testing composite hypotheses, unfortunately, it is not possible to make a proper decision for distanced densities for all the possible densities  $\gamma_i, i = 1, \dots, S$ . For example, when testing two hypotheses, null and alternative, the integrated likelihood ratio is  $\int_{\Omega_A} f_{\theta}(x)\gamma_A(\theta)d\theta / \int_{\Omega_0} f_{\theta}(x)\gamma_0(\theta)d\theta$ . The Bayesian approach to hypothesis testing uses this ratio, and the posterior probability that the null hypothesis is true is calculated as

$$p(H_0 | D) = 1 / (1 + posterior odds), \tag{2}$$

where  $posterior\ odds = (1 - \pi_0) \int_{\Omega_A} f_{\theta}(x)\gamma_A(\theta)d\theta / \pi_0 \int_{\Omega_0} f_{\theta}(x)\gamma_0(\theta)d\theta$ ,  $\pi_0$  is a priori probability that the null hypothesis is true.

It is obvious that the posterior odds depend heavily on  $\pi_0$  and the individual priors  $\gamma_A$  and  $\gamma_0$ . It is known that, as the prior  $\gamma_A$  becomes increasingly flat, the posterior odds usually approach 0 [10, 23]. In [13] was recommended “reference” priors  $\gamma_A$  and  $\gamma_0$  so that the priors did not overwhelm the data. Other approaches recommend to use part of the data as a training sample. In [14] was offered averaging over training samples for choosing so-called “intrinsic” priors. A number of methods for finding the priors were reviewed in [15]. For choosing  $\pi_0$ , in [13] was proposed to take  $\pi_0 = \frac{1}{2}$ , which means that the posterior odds are equal to the Bayes factor.

The Bayesian rule of testing hypotheses computes the posterior probability of the null hypothesis on the basis of some data  $D = \{x_1, x_2, \dots, x_n\}$  and, if it is close to one, accepts the null hypothesis, otherwise it accepts the alternative hypothesis. It is known that, when both null and alternative hypotheses have the

same dimension, i.e. when both are simple or both are composite, the solution is easily obtained, but when the null hypothesis is simple and the alternative composite, the posterior probability of the null hypothesis may be rather misleading. For showing this fact, the following examples were considered in [23]. The first of them consists in the following. Let  $D = \{x_1, x_2, \dots, x_n\}$  be a random sample of the random variable  $\xi$  with normal distribution. The null and alternative hypotheses are

$$H_0 : \xi \sim N(x | \mu, \sigma^2) \text{ and } H_A : \xi \sim N(x | \cdot, \sigma^2),$$

$$p(\mu | H_A) = N(\mu | \mu_1, \sigma_1^2),$$

with a priori probabilities  $p(H_0) = p$  and  $p(H_1) = 1 - p$ .

The arithmetic mean of the observation  $\bar{x}$  is sufficient and the posterior probability of the null hypothesis is

$$p(H_0 | D) = p(H_0 | \bar{x}) = \left( 1 + \frac{1-p}{p} \frac{p(\bar{x} | H_A)}{p(\bar{x} | H_0)} \right)^{-1}, \tag{3}$$

where

$$p(\bar{x} | H_0) = N(\bar{x} | \mu_0, \sigma^2 / n),$$

$$p(\bar{x} | H_A) = \int N(\bar{x} | \mu, \sigma^2 / n) N(\mu | \mu_1, \sigma_1^2) d\mu = N(\bar{x} | \mu_1, \sigma_1^2 + \sigma^2 / n)$$

The final form of (3) is

$$p(H_0 | D) = \left[ \frac{1 + \frac{1-p}{p} \left( \frac{\sigma^2 / n}{(\sigma_1^2 + \sigma^2 / n)} \right)^{1/2}}{\exp \left\{ \frac{(\bar{x} - \mu_0)^2}{2(\sigma^2 / n)} \right\}} \right]^{-1} \cdot \left[ \frac{1}{\exp \left\{ \frac{(\bar{x} - \mu_1)^2}{2(\sigma_1^2 + \sigma^2 / n)} \right\}} \right] \tag{4}$$

It is easy to see from (4) that, for any fixed  $\bar{x}$  and  $p$ , the right hand side of (4) tends to one as  $\sigma_1^2$  increases, i.e. for sufficiently large prior variance  $\sigma_1^2$ , independently of the data values and a priori probability  $p$ , the null hypothesis will always be accepted. This fact is called the Lindley’s statistical paradox. Similar behavior was mentioned in [24, 25] at uniform distribution of  $p(\mu | H_A)$  when the size of the interval increases. For avoiding this misleading, in [23] it was suggested not to choose the fixed value of  $p$ , but to choose it depending on the form of  $p(\mu | H_A)$ . In particular, we must choose the prior probability which maximizes the amount of missing information about  $\theta$ , which gives

$$\frac{1-p}{p} = \left( \frac{\sigma_1^2 + \sigma^2/n}{\sigma^2/n} \right)^{1/2}$$

Application of this methodology to considered case (3) gives the following posterior for null

$$p(H_0 | D) = \left[ 1 + \frac{\exp\left\{(\bar{x} - \mu_0)^2 / (2\sigma^2/n)\right\}}{\exp\left\{(\bar{x} - \mu_1)^2 / (2\sigma_1^2 + 2\sigma^2/n)\right\}} \right]^{-1} \quad (5)$$

which does not tend to one as  $\sigma_1^2$  increases.

In the second example, at alternative hypothesis  $\mu \in ]-A, A[$  and  $p(\mu | H_A) = \frac{1}{2A}$ , the posterior probability of  $H_0$  is [23]

$$p(H_0 | D) = \left[ 1 + \frac{1-p}{p} \frac{1/(2A)}{(n/2\pi\sigma^2)^{1/2} \exp\left\{-n(\bar{x} - \mu_0)^2 / 2\sigma^2\right\}} \right]^{-1} \quad (6)$$

which, for fixed  $p$ , tends to one as  $A$  increases.

In this case the proper choice of a priori probability in accordance with the above-mentioned methodology gives the following result

$$\frac{1-p}{p} = \left( \frac{n}{2\pi e\sigma^2} \right)^{1/2} 2A \quad (7)$$

and, substituting it into (6), we get

$$p(H_0 | D) = \left[ 1 + \exp\left\{ \frac{1}{2} \left[ \left( \frac{n}{\sigma^2} \right) (\bar{x} - \mu_0)^2 - 1 \right] \right\} \right]^{-1}, \quad (8)$$

which does not depend on the arbitrary constant  $A$ .

Thus, the method of choosing the proper priors for overcoming the Lindley's paradox was offered in [23]. But let us imagine the situation when the priors are known on the basis of previous experience or they are defined on the basis of preliminary investigation of the problem. The existing knowledge is very important and it is not reasonable to ignore it as it can lead to misleading conclusions.

Let us apply one of CBM, for example, task 1 with restrictions on the averaged probability of acceptance of true hypotheses for testing hypotheses [1, 4, 9]. For two hypotheses, it has the following form

$$p(H_0) \int_{\Gamma_0} p(\bar{x} | H_0) d\bar{x} + p(H_A) \int_{\Gamma_A} p(\bar{x} | H_A) d\bar{x} \Rightarrow \min_{\{\Gamma_0, \Gamma_A\}}, \quad (9)$$

at

$$p(H_0) \int_{\Gamma_0} p(\bar{x} | H_0) d\bar{x} + p(H_A) \int_{\Gamma_A} p(\bar{x} | H_A) d\bar{x} \geq 1 - \alpha, \quad (10)$$

where  $\Gamma_0$  and  $\Gamma_A$  are the acceptance regions of hypotheses  $H_0$  and  $H_A$ , respectively, and

$$\Gamma_0 = \left\{ \bar{x} : \frac{p(\bar{x} | H_A)}{p(\bar{x} | H_0)} < \lambda \frac{p}{1-p} \right\},$$

$$\Gamma_A = \left\{ \bar{x} : \frac{p(\bar{x} | H_0)}{p(\bar{x} | H_A)} < \lambda \frac{1-p}{p} \right\}.$$

For the first example considered above, we have

$$\Gamma_0 = \left\{ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \frac{\exp\left\{(\bar{x} - \mu_0)^2 / (2\sigma^2/n)\right\}}{\exp\left\{(\bar{x} - \mu_1)^2 / (2\sigma_1^2 + 2\sigma^2/n)\right\}} < \lambda \right\}, \quad (11)$$

$$\Gamma_A = \left\{ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \frac{\exp\left\{(\bar{x} - \mu_0)^2 / (2\sigma^2/n)\right\}}{\exp\left\{(\bar{x} - \mu_1)^2 / (2\sigma_1^2 + 2\sigma^2/n)\right\}} > 1/\lambda \right\}. \quad (12)$$

It is evident that, for any fixed  $\bar{x}$  and  $p$ , the left-hand-side member of inequalities of  $\Gamma_0$  and  $\Gamma_A$  tend to zero as  $\sigma_1^2$  increases. If  $\lambda$  is fixed (as it is in the classical Bayesian statement when  $\lambda = 1$ ), for a large prior variance  $\sigma_1^2$ , inequality in (11) will always be satisfied and that in (12) will never be satisfied, i.e. hypothesis  $H_0$  will always be accepted.

In the considered case, Lagrange multiplier  $\lambda$  must be chosen so that condition (10) was satisfied. This means that the average power of the test will always be equal to the chosen level  $1 - \alpha$ . The value of  $\lambda$  compensates the influence of  $\sigma_1^2$  on the conditions in (11) and (12), and its value depends on the value of  $\alpha$ . When  $\sigma_1^2$  increases the value of  $\lambda$  decreases for satisfying condition (10). Therefore, the sizes of hypotheses acceptance regions  $\Gamma_0$  and  $\Gamma_A$  will be kept the same for the preservation of condition (10) and the situation like the Lindley's paradox will never arise.

For the given values of  $\mu_0, \mu_1, \sigma^2, \sigma_1^2$  and  $\alpha$ , the value of  $\lambda$  will be defined from condition (10). That means that, to the concrete values of  $\mu_0, \mu_1, \sigma^2, \sigma_1^2$  and  $\alpha$ , the concrete value of  $\lambda$  corresponds. In this condition, it is evident that, when  $\bar{x} \rightarrow \mu_0$ , the

probability of satisfying of the inequality in (11) increases but the same probability in (12) decreases, i.e. the probability that hypothesis  $H_0$  will be accepted increases, and the probability that  $H_A$  will be accepted decreases independently of the value of  $\sigma_1^2$ . On the other hand, when  $\bar{x} \rightarrow \mu_1$ , there is diametrically opposite circumstance: the probability of satisfying the inequality in (12) increases but that in (11) decreases. This means that the probability of acceptance of  $H_A$  increases, and the probability of acceptance of  $H_0$  decreases.

Similarly to [23], let us denote  $\gamma_x$  and  $\gamma_1$

$$\bar{x} = \mu_0 + \gamma_x \cdot \sigma / \sqrt{n},$$

$$\mu_1 = \mu_0 + \gamma_1 \cdot \sigma_1,$$

so that  $\gamma_x$  and  $\gamma_1$  are the measures in standard unites, how far the sample and the prior mean are from the null hypothesis, respectively.

Substitution of them into (11) and (12) yields

$$\Gamma_0 = \left\{ \left[ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ \frac{1}{2} \left( \gamma_x^2 - \frac{(\gamma_x \cdot \sigma / \sqrt{n} - \gamma_1 \sigma_1)^2}{\sigma_1^2 + \sigma^2/n} \right) \right\} \right] < \lambda \right\}, \quad (13)$$

$$\Gamma_A = \left\{ \left[ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ \frac{1}{2} \left( \gamma_x^2 - \frac{(\gamma_x \cdot \sigma / \sqrt{n} - \gamma_1 \sigma_1)^2}{\sigma_1^2 + \sigma^2/n} \right) \right\} \right] > 1/\lambda \right\}. \quad (14)$$

When either  $\sigma_1^2$  or  $n$  increases, the expression in the power of exponents tends to  $\exp \left\{ \frac{1}{2} (\gamma_x^2 - \gamma_1^2) \right\}$  and (13) and (14) could be rewritten as follows

$$\Gamma_0 = \left\{ \bar{x} : \frac{1-p}{p} (\varepsilon)^{1/2} \exp \left\{ \frac{1}{2} (\gamma_x^2 - \gamma_1^2) \right\} < \lambda \right\}, \quad (15)$$

$$\Gamma_A = \left\{ \bar{x} : \frac{1-p}{p} (\varepsilon)^{1/2} \exp \left\{ \frac{1}{2} (\gamma_x^2 - \gamma_1^2) \right\} > 1/\lambda \right\}. \quad (16)$$

where  $\varepsilon = \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n}$  is a very small value. Form here it is evident that, for fixed  $\sigma^2$ ,  $\sigma_1^2$  and  $n$ , the only

important features are the distances in standard unites  $\gamma_x$  and  $\gamma_1$  from the sample and from the prior mean to the null hypothesis. Under the null hypothesis,  $\gamma_x$  will be moderate,  $(\gamma_x^2 - \gamma_1^2)$  will not be too large, therefore the fulfillment of the inequality defining the region  $\Gamma_0$  is more probable, i.e. it is more probable to accept  $H_0$ . Under the alternative hypothesis,  $\gamma_x$  will increase as  $\sqrt{n}$ ,  $(\gamma_x^2 - \gamma_1^2)$  increases as  $n$ , exponent in (16) will increase and the probability of fulfilling the inequality defined the region  $\Gamma_A$  (see (16)) increases, i.e. acceptance of  $H_A$  is more probable.

Analyzing (13) and (14), we arrive at the following conclusion. Let us suppose that  $\bar{x} \rightarrow \mu_0$ . Then (13) and (14) tend to the following expressions

$$\Gamma_0 = \left\{ \left[ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ -\frac{(\mu_1 - \mu_0)^2}{2(\sigma_1^2 + \sigma^2/n)} \right\} < \lambda \right] \right\}, \quad (17)$$

$$\Gamma_A = \left\{ \left[ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ -\frac{(\mu_1 - \mu_0)^2}{2(\sigma_1^2 + \sigma^2/n)} \right\} > 1/\lambda \right] \right\}. \quad (18)$$

Hence it is evident that, when for fixed  $\sigma^2$ ,  $\sigma_1^2$  and  $n$  we have  $(\mu_1 - \mu_0)^2 \rightarrow \infty$ ,  $P(\bar{x} \in \Gamma_0) \rightarrow 1$  and  $P(\bar{x} \in \Gamma_A) \rightarrow 0$ .

Let us suppose that  $\bar{x} \rightarrow \mu_1$ . Then (13) and (14) tend to the following expressions

$$\Gamma_0 = \left\{ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ \frac{(\mu_1 - \mu_0)^2 \cdot n}{2\sigma^2} \right\} < \lambda \right\}, \quad (17)$$

$$\Gamma_A = \left\{ \bar{x} : \frac{1-p}{p} \left( \frac{\sigma^2/n}{\sigma_1^2 + \sigma^2/n} \right)^{1/2} \exp \left\{ \frac{(\mu_1 - \mu_0)^2 \cdot n}{2\sigma^2} \right\} > 1/\lambda \right\}. \quad (18)$$

When for fixed  $\sigma^2$ ,  $\sigma_1^2$  and  $n$  we have  $(\mu_1 - \mu_0)^2 \rightarrow \infty$ , the following is true  $P(\bar{x} \in \Gamma_0) \rightarrow 0$  and  $P(\bar{x} \in \Gamma_A) \rightarrow 1$ .

Let us consider the second example when, with the alternative hypothesis, the mathematical expectation is uniformly distributed in the finite interval. In this case, the decision-making regions have the following forms

$$\Gamma_0 = \left\{ \bar{x} : \frac{1-p}{P} \frac{1/(2A)}{(n/2\pi\sigma^2)^{1/2} \exp\{-n(\bar{x}-\mu_0)^2/2\sigma^2\}} < \lambda \right\}, \quad (19)$$

$$\Gamma_A = \left\{ \bar{x} : \frac{1-p}{P} \frac{1/(2A)}{(n/2\pi\sigma^2)^{1/2} \exp\{-n(\bar{x}-\mu_0)^2/2\sigma^2\}} > 1/\lambda \right\}. \quad (20)$$

Similarly, for the given  $\mu_0$ ,  $\sigma^2$  and  $n$ , in the classical Bayesian rule  $\lambda = 1$  and, when  $A$  increases, the inequality in (19) will always be satisfied, while the inequality in (20), on the contrary, will not be satisfied. That means that, for fixed  $p$ , the probability of inequality in (19) tends to one as  $A$  increases and the probability of inequality in (20) tends to zero as  $A$  increases. This is the same that, when  $A$  increases the null hypothesis will always be accepted.

In the constrained Bayesian task, for given  $\alpha$ , when  $A$  increases the value of  $\lambda$  decreases in such a way that condition (10) was satisfied. For given  $p$ ,  $\mu_0$ ,  $\sigma^2$ ,  $n$ ,  $A$  and  $\alpha$ , the value of  $\lambda$  is defined from condition (10), and, the smaller is the difference between  $\bar{x}$  and  $\mu_0$ , i.e. the smaller is  $(\bar{x} - \mu_0)^2$ , the bigger is the probability of satisfying the inequality in (20). This means that the probability of acceptance of  $H_0$  increases. On the other hand, the bigger is the divergence between  $\bar{x}$  and  $\mu_0$ , i.e. the bigger is  $(\bar{x} - \mu_0)^2$ , the bigger is the probability of inequality in (20) and the smaller is the inequality in (19). This means that the probability of acceptance of  $H_A$  increases.

#### 4. COMPUTATION RESULTS

Let us present the computation results of the considered examples for demonstration of the theoretically investigated properties of the Bayes method with arbitrary and a priori probabilities offered by Bernardo and CBM.

The following notations are used below in addition to the above-introduced ones:

$r$  - risk function computed by ratio (9),

$\alpha_0$  - probability of incorrect rejection of the basic hypothesis,

$\alpha_1$  - probability of incorrect rejection of the alternative hypothesis,

$\beta_0$  - probability of incorrect acceptance of the basic hypothesis,

$\beta_1$  - probability of incorrect acceptance of the alternative hypothesis;

at  $\lambda > 1$

$p_{01}$  - probability of making no decision when the observed value belongs to the intersection area of the regions of acceptance of tested hypotheses at validity of  $H_0$ ,

$p_{11}$  - probability of making no decision when the observed value belongs to the intersection area of the regions of acceptance of tested hypotheses at validity of  $H_1$ ;

at  $\lambda < 1$

$p_{02}$  - probability of making no decision when the observed value does not belong to the regions of acceptance of tested hypotheses at validity of  $H_0$ ,

$p_{12}$  - probability of making no decision when the observed value does not belong to the regions of acceptance of tested hypotheses at validity of  $H_1$ .

The results given below were obtained by simulation of 10,000 experiments in which by  $\bar{x}_0$  and  $\bar{x}_1$  computed at validity of hypotheses  $H_0 : \xi \sim N(x | \mu, \sigma^2)$  and  $H_A : \xi \sim N(x | \cdot, \sigma^2)$ ,  $p(\mu | H_A) = N(\mu | \mu_1, \sigma_1^2)$ , respectively, with sample sizes equal to  $n$  and a priori probability  $p$ , decisions were made and the appropriate error probabilities were computed.

**Example 1.** The results of computation of Lagrange multiplier, risk function and error probabilities depending on the variance of the alternative hypothesis in CBM are given in Table 1. Figures 1a, b and c were constructed by these results. All computations for this paper are realized in MATLAB by programs developed by the author of the paper. The codes of these programs are given in Appendix.

From the computation results, the following is evident. When the variance of alternative hypotheses increases, the Lagrange multiplier in the decision rule changes: in the beginning it increases, achieves its maximum and then decreases. For the considered example, it increased from 0.4615 for  $\sigma_1^2 = 1$  up to the maximum value 9.7657 for  $\sigma_1^2 = 50$  and then decreased to 1.0898 for  $\sigma_1^2 = 500$ . The change of  $\sigma_1^2$  from 1 to 500 caused the change of the value of the probability of rejection of the null hypothesis  $\alpha_0$  from 0.0857 practically to zero (0.0004) for  $\sigma_1^2 = 100$ , after

**Table 1: Computation Results of Lagrange Multiplier, Risk Function and Error Probabilities Depending on the Variance of the Alternative Hypothesis**

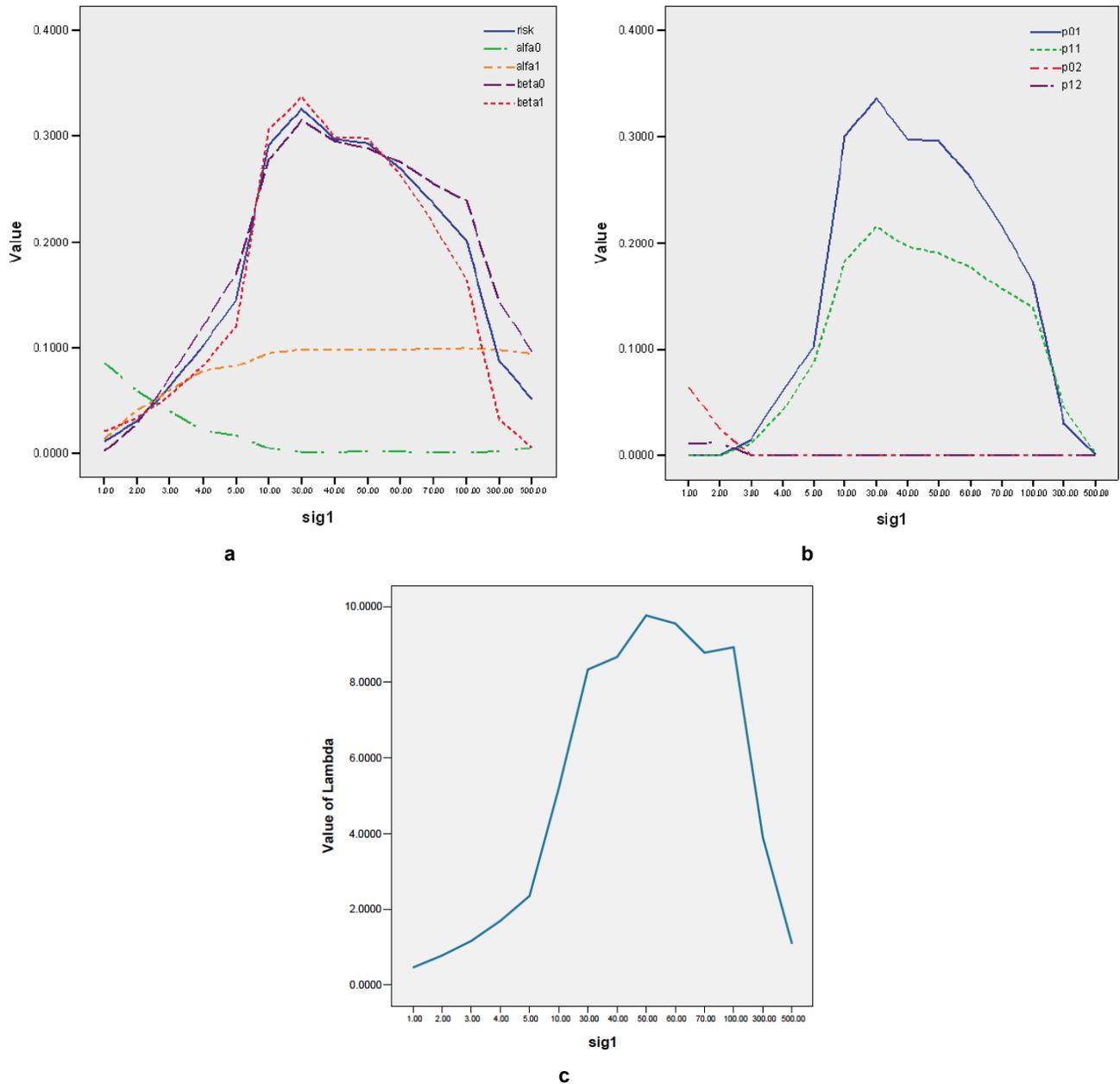
Initial values $m = 10,000$ $n = 5, \sigma^2 = 1$ $\mu_0 = 1, \mu_1 = 3$			Computed values									
$\sigma_1^2$	$p$	$\alpha$	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$	$p_{12}$
1	0.5	0.05	0.4615	0.0118	0.0857	0.0143	0.0024	0.0211	0	0	0.0646	0.0119
2			0.7793	0.0307	0.0594	0.0406	0.0278	0.0335	0	0	0.0259	0.0128
3			1.1656	0.0636	0.0399	0.0601	0.0720	0.0553	0.0154	0.0119	0	0
4			1.6952	0.1016	0.0220	0.0780	0.1203	0.0828	0.0608	0.0423	0	0
5			2.3528	0.1444	0.0174	0.0826	0.1693	0.1195	0.1021	0.0867	0	0
10			5.1938	0.2916	0.0052	0.0948	0.2776	0.3056	0.3004	0.1828	0	0
30		8.3372	0.3261	0.0015	0.0985	0.3147	0.3375	0.3360	0.2162	0	0	
40		8.6671	0.2971	0.0014	0.0986	0.2952	0.2990	0.2976	0.1966	0	0	
50		9.7657	0.2934	0.0018	0.0982	0.2888	0.2979	0.2961	0.1906	0	0	
60		9.5523	0.2693	0.0017	0.0983	0.2752	0.2635	0.2618	0.1769	0	0	
70		8.7808	0.2364	0.0010	0.0990	0.2554	0.2173	0.2163	0.1564	0	0	
100		8.9242	0.2017	0.0004	0.0996	0.2392	0.1642	0.1638	0.1396	0	0	
300		3.9063	0.0874	0.0021	0.0979	0.1428	0.0320	0.0299	0.0449	0	0	
500		1.0898	0.0514	0.0054	0.0946	0.0965	0.0062	0.0008	0.0019	0	0	

that it increased a little up to 0.0054 for  $\sigma_1^2 = 500$ . In this case the value of the probability of rejection of the alternative hypothesis  $\alpha_1$  increased from 0.0143 for  $\sigma_1^2 = 1$  up to 0.0996 for  $\sigma_1^2 = 100$  and then it decreased to 0.0946 for  $\sigma_1^2 = 500$ . The probabilities of incorrect acceptance of basic and alternative hypotheses  $\beta_0$  and  $\beta_1$  in the beginning increased from 0.0024 and 0.0211 up to 0.3147 and 0.3375, respectively, for  $\sigma_1^2 = 30$  and then they decreased to 0.0965 and 0.0062, respectively, for  $\sigma_1^2 = 500$ . When  $\lambda > 1$ , the probabilities  $p_{01}$  and  $p_{11}$  of making no decision when the observed value belongs to the intersection area of the regions of acceptance of testing hypotheses at validity of  $H_0$  and  $H_1$ , change similarly from 0.0154 and 0.0119, respectively, for  $\sigma_1^2 = 3$  up to 0.3360 and 0.2162 for  $\sigma_1^2 = 30$ , and then they decrease to 0.0008 and 0.0019 for  $\sigma_1^2 = 500$ . When  $\lambda < 1$ , the probability  $p_{02}$  of making no decision when the observed value does not belong to the regions of acceptance of testing hypotheses at validity of  $H_0$  decreases from 0.0646 for  $\sigma_1^2 = 1$  to 0.0259 for  $\sigma_1^2 = 2$  and the probability  $p_{12}$  of making no decision when the observed value does not

belong to the regions of acceptance of testing hypotheses at validity of  $H_1$  increases from 0.0119 for  $\sigma_1^2 = 1$  up to 0.0128 for  $\sigma_1^2 = 2$ .

In accordance with expression (9), the value of the risk function is within the interval between  $\beta_0$  and  $\beta_1$ .

As was mentioned above, for simulation there were used 10,000 random numbers distributed by appropriate distribution laws. Let us explain the essence of the computed values of the considered probabilities from the point of view of decision-making. Let us consider the case  $\sigma_1^2 = 500$ . In this case, for  $\lambda = 1.0898$ , the decisions about the validity of tested hypotheses were made 10,000 times. Among them, the null hypothesis was incorrectly rejected 54 times, while the alternative hypothesis – 946 times. The null hypothesis was incorrectly accepted 965 times, while the alternative hypothesis – 62 times. No decision was made 27 times, because the information contained in the sample was not sufficient for making a simple decision with given reliability, among them 8 times because both hypotheses were supposed to be true and 19 times both hypotheses were supposed to be false. It seems that these results are quite good for so noised alternative hypothesis.



**Figure 1:** Dependences of Lagrange multiplier, risk function and error probabilities on the variance of the alternative hypothesis in the constrained Bayesian test.

The results of computation of Lagrange multiplier, risk function and error probabilities depending on the sample size are given in Table 2. The dependences of the considered probabilities on the sample size  $n$  are shown in Figure 2.

Hence the following is obvious. When the sample size  $n$  increases from 1 to 50, the error probabilities of type II for both hypotheses, i.e. the probabilities of incorrect acceptance of both hypotheses and the risk function which is the averaged value of these probabilities, decrease. They tend to zero (see Figure 2a). The probabilities of type I for both hypotheses, i.e. the probabilities of incorrect rejection of both

hypotheses change in such a manner that their sum weighed by a priori probabilities was not more than  $\alpha$ , i.e. restriction (10) in the statement of the constrained Bayesian task was satisfied. The changes have the following character. The probability of rejection of the basic hypothesis decreases and the probability of rejection of the alternative hypothesis increases when the sample size  $n$  increases from 1 up to 10. After that, i.e. when  $n$  increases from 10 up to 50, the probability of rejection of the basic hypothesis increases and the probability of rejection of the alternative hypothesis decreases (see Figure 2a). These two probabilities are equal to each other when  $n = 37$ . The described situation logically can be explained completely from the

**Table 2: Computation Results of Lagrange Multiplier, Risk Function and Error Probabilities Depending on the Sample Size**

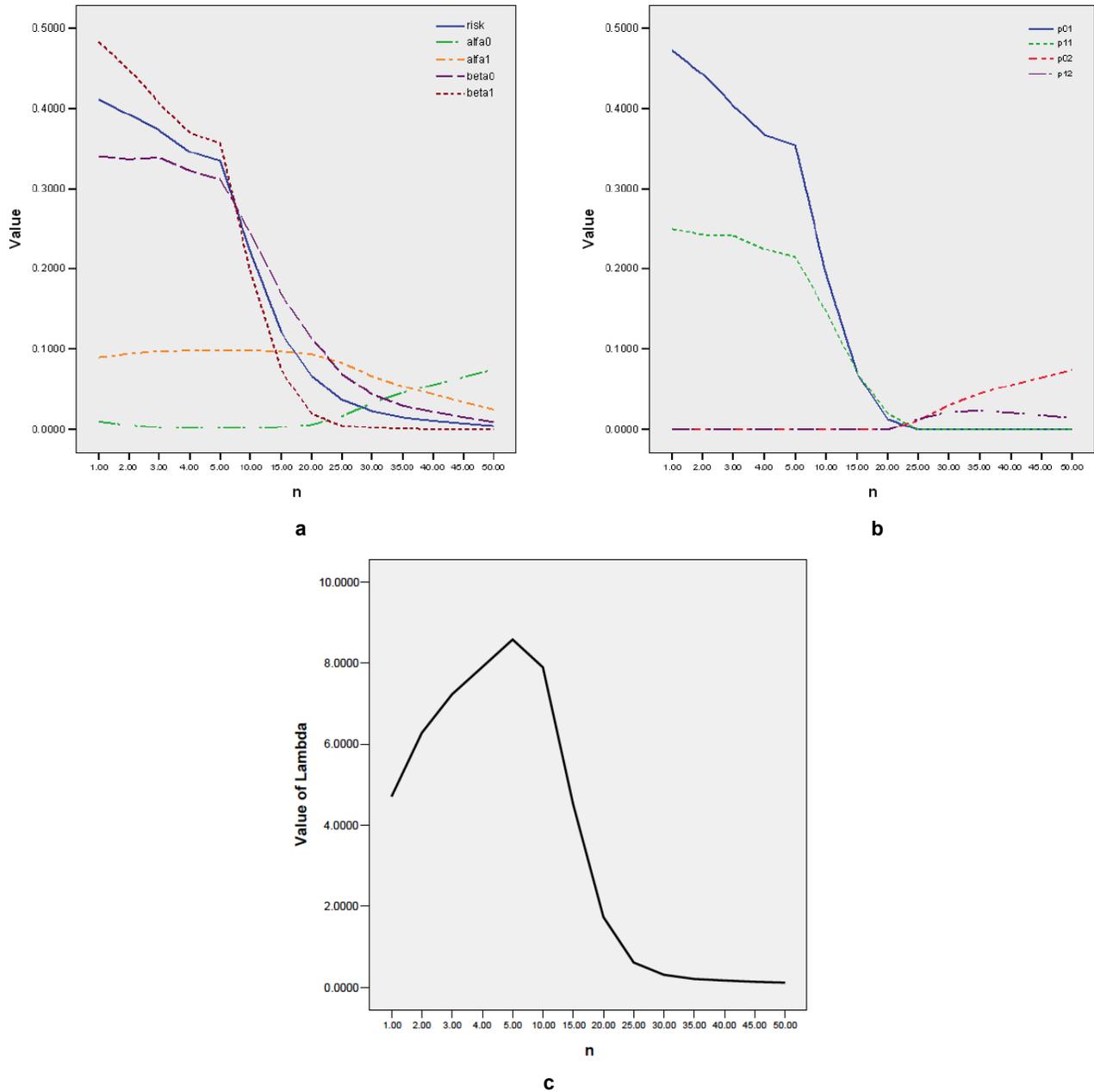
Initial values $m = 10,000$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, \sigma_1^2 = 30$ $\alpha = 0.05$			Computed values									
$n$	$P$	$p_{opt}$	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$	$p_{12}$
1	0.5	0.1523*	4.7098	0.4113	0.0106	0.0894	0.3400	0.4825	0.4719	0.2506	0	0
			1.4938	0.2644	0.3283	0	0.1322	1	0.6717	0.1322	0	0
2			6.2814	0.3924	0.0061	0.0939	0.3364	0.4485	0.4424	0.2425	0	0
3			7.2328	0.3725	0.0029	0.0971	0.3386	0.4065	0.4036	0.2415	0	0
4			7.9070	0.3458	0.0023	0.0977	0.3223	0.3693	0.3670	0.2246	0	0
5		0.0753*	8.5832	0.3346	0.0025	0.0975	0.3120	0.3571	0.3546	0.2145	0	0
			1.0686	0.1236	0.6075	0.0046	0.0550	0.9657	0.3582	0.0504	0	0
10			7.8974	0.2207	0.0022	0.0978	0.2452	0.1961	0.1939	0.1474	0	0
15			4.5156	0.1211	0.0033	0.0967	0.1682	0.0740	0.0707	0.0715	0	0
20			1.7333	0.0669	0.0067	0.0933	0.1138	0.0201	0.0134	0.0205	0	0
25			0.6143	0.0372	0.0171	0.0829	0.0692	0.0053	0	0	0.0118	0.0137
30			0.3100	0.0233	0.0336	0.0664	0.0438	0.0028	0	0	0.0308	0.0226
35			0.2074	0.0156	0.0463	0.0537	0.0298	0.0014	0	0	0.0449	0.0239
40			0.1699	0.0115	0.0560	0.0440	0.0225	0.0004	0	0	0.0556	0.0215
45			0.1380	0.0083	0.0656	0.0344	0.0160	0.0006	0	0	0.0650	0.0184
50			0.1169	0.0052	0.0752	0.0248	0.0098	0.0007	0	0	0.0745	0.0150

\*these two lines of computation results are obtained for optimal (in accordance with Bernardo for overcoming the Lindley's paradox) values of a priori probability at appropriate sample sizes.

point of view of the information theory. In particular, the more is the information about the studied phenomenon, the more reliable must be the decision made in relation with this phenomenon. In the considered case, the increase of the information, i.e. the increase of sample size  $n$ , causes the decrease of probabilities of type II errors and, on the basis of the condition of the considered problem, type I errors stay in the frame of chosen restriction (10). To the small or big values of the variance of the alternative hypothesis corresponds the increase of information about tested hypotheses. Therefore, to the small or big values of the variance of alternative hypothesis correspond small values of type II errors, and the weighted sum of type I error probabilities satisfy restriction (10) (see Figure 1a).

Constrained Bayesian methods of hypotheses testing have the following specificity [1, 9]: when the Lagrange multiplier  $\lambda$  differs from the one in the observation space there appears the region of no acceptance of the simple hypothesis and the region of supposition of the validity of both hypotheses. In the considered case,  $p_{01}$  and  $p_{11}$  are the probabilities of no acceptance of basic and alternative hypotheses,

respectively, when they are true, and  $p_{02}$  and  $p_{12}$  are the probabilities of supposition of the validity of both hypotheses when basic and alternative hypotheses are true, respectively. When  $n$  changes from 1 up to 50, the value of  $\lambda$  changes. In the beginning it is more than one and after  $n = 22$  it becomes less than one (see Figure 2c). Therefore, in the computed values (see Table 2), up to  $n = 25$   $p_{02}$  and  $p_{12}$  are equal to zero, after that  $p_{01}$  and  $p_{11}$  become equal to zero. The fact that  $n$  increases corresponds to the increase of the existing information about tested hypotheses. To the increasing information corresponds the increase in the information distance between the tested hypotheses in the considered case. The increase in the information distance between the tested hypotheses logically causes the decrease in the intersection of hypotheses acceptance regions because of restriction (10). The decrease in the intersection of hypotheses acceptance regions causes the decrease in the probabilities  $p_{01}$  and  $p_{11}$ . When the information distance increases, after some its value (which is determined by restriction (10)), the regions of acceptance of tested hypotheses are not intersected,



**Figure 2:** Dependences of Lagrange multiplier, risk function and error probabilities on the sample size in the constrained Bayesian test.

and the following increase in the distance causes the appearance of the region which does not belong to the hypotheses acceptance regions. The bigger is the distance, the bigger is the latter region, i.e. the bigger is the sum of probabilities  $p_{02}$  and  $p_{12}$ . The changes of probabilities  $p_{01}$ ,  $p_{11}$ ,  $p_{02}$  and  $p_{12}$  given in Figure 2b correspond completely to this logic.

If there is no other objective reason, the choice of a priori probability  $p$  in accordance with Bernardo for only overcoming the Lindley's paradox, on the one hand, unjustifiably significantly increases the probability of incorrect rejection of the basic hypothesis ( $\alpha_0$ ) and

the probability of incorrect acceptance of the alternative hypothesis ( $\beta_1$ ) and, on the other hand, unjustifiably significantly decreases the probability of incorrect rejection of the alternative hypothesis ( $\alpha_1$ ) and the probability of incorrect acceptance of the basic hypothesis ( $\beta_0$ ) (see  $p_{opt}$  in lines of Table 2 at  $n = 1$  and  $n = 5$ ).

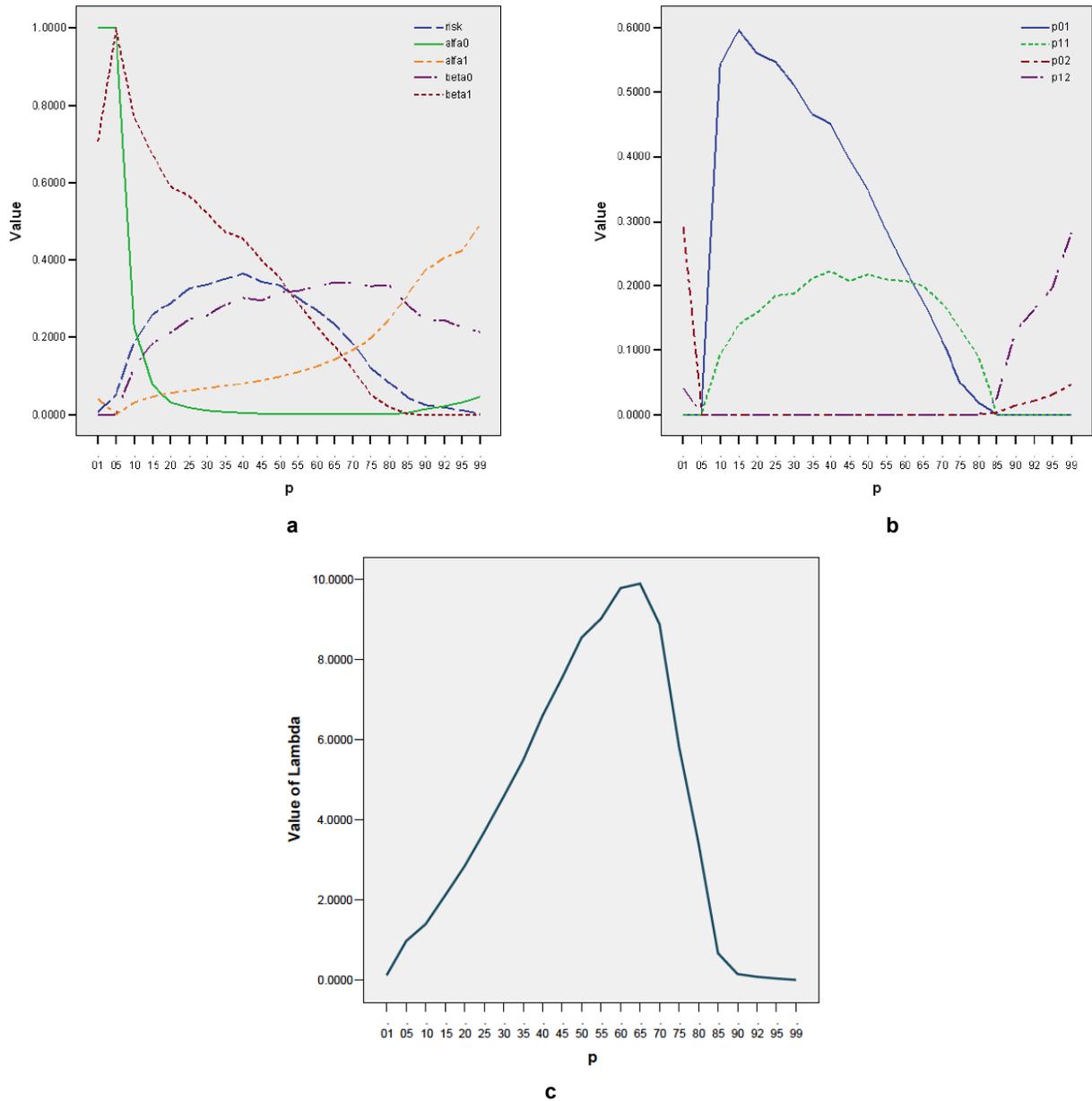
The results of computation of Lagrange multiplier, risk function and error probabilities depending on the a priori probability are given in Table 3. Figures 3a, b and c are constructed by these results.

**Table 3: Computation Results of Lagrange Multiplier, Risk Function and Error Probabilities Depending on the a Priori Probability**

Initial values $m = 10,000$ $n = 5, \alpha = 0.05$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, \sigma_1^2 = 30$	Computed values									
	$p$	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$
0.01	0.1235	0.0071	1	0.0404	0	0.7070	0	0	0.2930	0.0404
0.05	0.9767	0.0500	1	0	0	1	0	0	0	0
0.1	1.3996	0.1873	0.2247	0.0306	0.1228	0.7677	0.5430	0.0922	0	0
0.15	2.1136	0.2581	0.0749	0.0456	0.1851	0.6717	0.5968	0.1395	0	0
0.2	2.8459	0.2876	0.0312	0.0547	0.2116	0.5918	0.5606	0.1569	0	0
0.25	3.6939	0.3249	0.0176	0.0608	0.2450	0.5647	0.5471	0.1842	0	0
0.30	4.5781	0.3347	0.0106	0.0669	0.2543	0.5224	0.5118	0.1874	0	0
0.35	5.4853	0.3497	0.0069	0.0732	0.2838	0.4720	0.4651	0.2106	0	0
0.4	6.6015	0.3631	0.0044	0.0804	0.3017	0.4553	0.4509	0.2213	0	0
0.45	7.5412	0.3415	0.0026	0.0888	0.2950	0.3984	0.3958	0.2062	0	0
0.5	8.5460	0.3329	0.0019	0.0981	0.3148	0.3510	0.3491	0.2167	0	0
0.55	9.0180	0.3012	0.0010	0.1099	0.3186	0.2869	0.2859	0.2087	0	0
0.6	9.7802	0.2682	0.0008	0.1238	0.3303	0.2268	0.2260	0.2065	0	0
0.65	9.8921	0.2337	0.0008	0.1414	0.3404	0.1763	0.1755	0.1990	0	0
0.7	8.8764	0.1828	0.0006	0.1653	0.3382	0.1162	0.1156	0.1729	0	0
0.75	5.8290	0.1203	0.0008	0.1976	0.3310	0.0501	0.0493	0.1334	0	0
0.8	3.4200	0.0821	0.0012	0.2452	0.3336	0.0192	0.0180	0.0884	0	0
0.85	0.6715	0.0437	0.0045	0.3078	0.2829	0.0015	0	0	0.0030	0.0249
0.9	0.1527	0.0243	0.0142	0.3722	0.2424	0.0001	0	0	0.0141	0.1298
0.9247	0.0835	0.0183	0.0211	0.4050	0.2426	0	0	0	0.0211	0.1624
0.95	0.0417	0.0113	0.0304	0.4224	0.2259	0	0	0	0.0304	0.1965
0.99	0.0059	0.0021	0.0455	0.4949	0.2125	0	0	0	0.0455	0.2824

Hence the following is clear. With the increasing a priori probability of basic hypothesis  $p$  and accordingly the decreasing a priori probability of alternative hypothesis  $(1 - p)$ , the probability of acceptance of basic hypothesis  $\beta_0$ , when it is not true, increases up to  $p = 0.65$  and then it decreases. At the same time, the probability of acceptance of the alternative hypothesis  $\beta_1$ , when it is not true, decreases (see Figure 3a). The risk function is sum of these two probabilities, weighed by a priori probabilities (see (9)). Therefore, the form of the change of the risk function up to  $p = 0.5$  has the form of  $\beta_1$  and then, it has the form of  $\beta_0$  (see Figure 3a).

As was investigated in [1], the change of a priori probability affects the decision made. This influence is not so significant as in the classical Bayesian test, but still it is present in the constrained test too, especially, when  $p$  is very small ( $< 0.1$ ) or very big ( $> 0.8$ ) (see Table 3 and Figure 3a). When  $p$  changes from 0.1 up to 0.99, the probability of rejection of the basic hypothesis  $\alpha_0$  when it is true decreases to  $p = 0.8$  and then it increases. At the same time, the probability of rejection of the alternative hypothesis  $\alpha_1$  when it is true increases from zero up to 0.4949 so that restriction (10) in the statement of the task was fulfilled. (Remark: here must be taken into account the fact that  $\alpha_0$  and  $\alpha_1$  from Table 3 contain the probabilities of no making the



**Figure 3:** Dependences of Lagrange multiplier, risk function and error probabilities on the a priori probability in the constrained Bayesian test.

decision, i.e. probabilities  $p_{01}$ ,  $p_{11}$ ,  $p_{02}$  and  $p_{12}$  and, after their exclusion, restriction (10) is fulfilled).

The Lagrange multiplier  $\lambda$  changes depending on  $p$  and, when  $p$  changes from 0.1 up to 0.99,  $\lambda$  increases from 0.1235 up to 9.8921 for  $p = 0.65$ , and then, it decreases to 0.0059 (see Figure 3c). The probabilities of no-acceptance of hypotheses  $p_{01}$ ,  $p_{11}$ ,  $p_{02}$  and  $p_{12}$  change depending on  $\lambda$ . When  $\lambda < 1$ , to which correspond  $p < 0.1$  and  $p > 0.8$ ,  $p_{01}$  and  $p_{11}$  are equal to zero. When  $0.1 \leq p \leq 0.8$ ,  $p_{01}$  and  $p_{11}$  increase in the beginning and then, they decrease. The probabilities  $p_{02}$  and  $p_{12}$  are equal to zero when

$0.1 \leq p \leq 0.8$ , i.e. when  $\lambda > 1$ . They decrease on the left of this interval and they increase on the right of this interval.

Let us present the simulation results for the considered example obtained by the classical Bayesian method and with Bernardo's correction. Computation results of the risk function and error probabilities depending on the variance of the alternative hypothesis and on the sample size are given in Tables 4 and 5, respectively, for both methods. Appropriate graphs are shown in Figures 4 and 5, respectively.

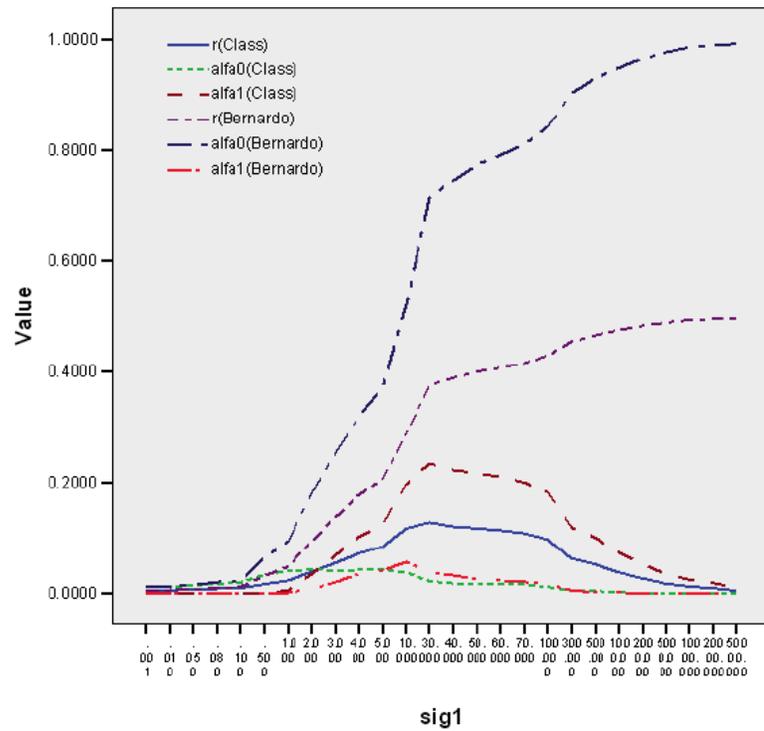
From these results, the following is evident. The property of (4) that, for the sufficiently large prior

**Table 4: Computation Results of the Risk Function and Error Probabilities Depending on the Variance of the Alternative Hypothesis**

Initial values $m = 10,000, n = 5$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, p = 0.5$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$ <b>ia</b>
0.001	0.0060	0.0119	0	0.0060	0.0120	0
0.01	0.0063	0.0126	0	0.0069	0.0138	0
0.05	0.0076	0.0151	0	0.0077	0.0154	0
0.08	0.0092	0.0183	0	0.0105	0.0210	0
0.1	0.0098	0.0197	0	0.0117	0.0234	0
0.5	0.0172	0.0338	0.0007	0.0321	0.0642	0.0001
1	0.0236	0.0409	0.0063	0.0484	0.0949	0.0019
2	0.0387	0.0430	0.0344	0.0935	0.1789	0.0080
3	0.0554	0.0421	0.0687	0.1371	0.2536	0.0206
4	0.0722	0.0425	0.1019	0.1772	0.3189	0.0354
5	0.0843	0.0438	0.1247	0.2062	0.3710	0.0415
10	0.1176	0.0376	0.1977	0.2895	0.5219	0.0570
30	0.1286	0.0225	0.2347	0.3756	0.7122	0.0389
40	0.1207	0.0184	0.2230	0.3892	0.7456	0.0328
50	0.1177	0.0178	0.2176	0.3997	0.7737	0.0257
60	0.1139	0.0165	0.2112	0.4072	0.7908	0.0236
70	0.1086	0.0165	0.2007	0.4158	0.8102	0.0213
100	0.0978	0.0124	0.1831	0.4287	0.8408	0.0166
300	0.0629	0.0061	0.1197	0.4543	0.9021	0.0065
500	0.0528	0.0052	0.1004	0.4654	0.9287	0.0021
1000	0.0383	0.0025	0.0741	0.4747	0.9470	0.0025
2000	0.0273	0.0016	0.0529	0.4828	0.9648	0.0004
5000	0.0180	0.0014	0.0346	0.4883	0.9765	0.0001
10000	0.0130	0.0008	0.0252	0.4930	0.9859	0.0001
20000	0.0098	0.0005	0.0192	0.4947	0.9894	0
50000	0.0055	0.0004	0.0107	0.4959	0.9918	0

variance  $\sigma_1^2$  independently of the observation results and a priori probability  $p$ , the null hypothesis will always be accepted is theoretically right. But, when the decision is made on the basis of real data (or on the basis of modeled results, as it is in the considered case), the situation is changed. It is true that the coefficient in (4) tends to zero at increasing  $\sigma_1^2$ . When the variance  $\sigma_1^2$  is very big, the arithmetic mean of observation results  $\bar{x}$  (for a small number of observations  $n$ ), at the validity of hypothesis  $H_1$

significantly differs from the mathematical expectations  $\mu_0$  and  $\mu_1$ , and the value of the exponent in the denominator is significantly smaller in comparison with the value of the exponent in the numerator. The result of their division is so big that the result of its multiplication by the coefficient becomes more than one and its inverse result becomes less than one. Therefore, in such cases, hypothesis  $H_1$  is accepted (see the columns  $\alpha_0 = \beta_1$  and  $\alpha_1 = \beta_0$  in Table 4 giving the results of modeling for the classical Bayesian method). For this reason, for the real data the classical



**Figure 4:** Dependences of the risk function and error probabilities on the variance of the alternative hypothesis in the Bayesian and Bernardo tests.

**Table 5: Computation Results of the Risk Function and Error Probabilities Depending on the Sample Size**

Initial values $m = 10,000, p = 0.5$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, \sigma_1^2 = 5$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$n$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$
1	0.2248	0.1013	0.3484	0.2808	0.3683	0.1933
2	0.1774	0.0751	0.2797	0.2512	0.3669	0.1356
3	0.1389	0.0551	0.2226	0.2304	0.3664	0.0945
4	0.1106	0.0522	0.1689	0.2132	0.3624	0.0639
5	0.0871	0.0441	0.1300	0.2024	0.3596	0.0453
10	0.0287	0.0276	0.0297	0.1833	0.3599	0.0068
15	0.0143	0.0223	0.0063	0.1826	0.3641	0.0011
20	0.0107	0.0203	0.0011	0.1815	0.3628	0.0001
25	0.0088	0.0172	0.0004	0.1844	0.3687	0
30	0.0079	0.0158	0	0.1888	0.3775	0
35	0.0072	0.0143	0	0.1829	0.3658	0
40	0.0066	0.0133	0	0.1870	0.3741	0
45	0.0061	0.0122	0	0.1856	0.3712	0
50	0.0052	0.0104	0	0.1905	0.3810	0

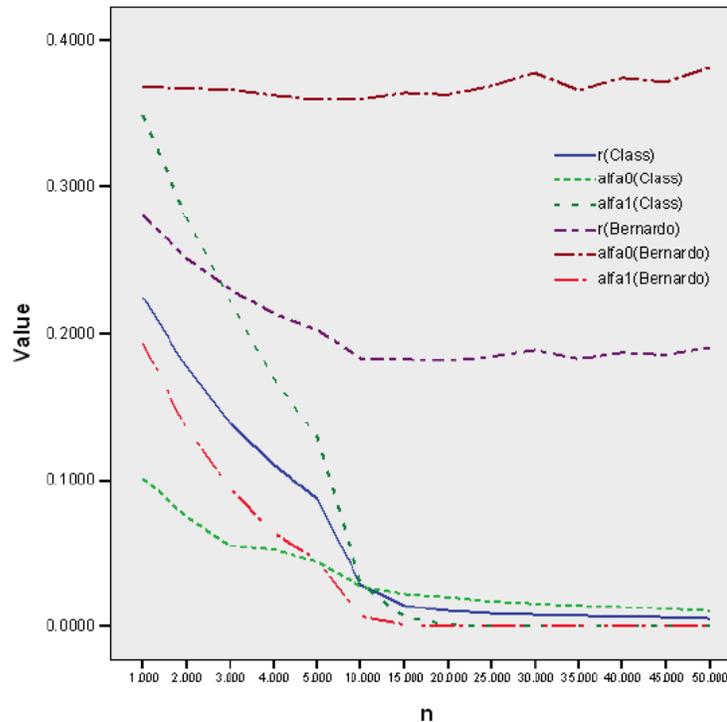


Figure 5: Dependences of the Lagrange multiplier, the risk function and error probabilities on the sample size in the Bayesian and Bernardo tests.

Bayesian method gives considerably better results in comparison with the Bernardo’s modification (see Table 4 and Figure 4). In the latter method, the probabilities of rejection and acceptance of basic and alternative hypotheses, respectively, converge to one at increasing  $\sigma_1^2$ , which is incorrect. The advantage of the classical Bayesian method in comparison with the Bernardo’s correction depending on the sample size is evident from the data given in Table 5 and Figure 5.

The computation results of the risk function and error probabilities depending on the a priori probability in the Bayesian test are given in Table 6. The appropriate graphs are shown in Figure 6. Hence it is evident that, when  $p$  increases,  $\alpha_0$  (that is the same as  $\beta_1$ ) decreases, and  $\alpha_1$  (that is the same as  $\beta_0$ ) increases, but, the bigger is  $\sigma_1^2$ , the faster are these changes, i.e. to the bigger value of the variance  $\sigma_1^2$  (which defines indefiniteness in the alternative hypothesis  $H_1$ ) correspond the less value of the rejection probability of the basic hypothesis  $H_0$  and the big value of the rejection probability of the alternative hypothesis  $H_1$  for one and the same  $p$ . In other words, for the given value of  $p$ , when the value of variance  $\sigma_1^2$  increases, the frequency of rejection of the basic hypothesis  $H_0$  decreases, i.e. it will be more frequently accepted and the frequency of rejection of the alternative hypothesis  $H_1$  increases, i.e. it will be

more frequently rejected. This corresponds completely to the noted above theoretical property of this test.

The choice of a priori probability  $p$  in accordance with the Bernardo’s condition does not change the natural motion of the investigated probabilities (see the cases  $p = 0.1640$  for  $\sigma_1^2 = 5$  and  $p = 0.0753$  for  $\sigma_1^2 = 30$ ). It only avoids the Lindley’s paradox effect for fixed  $\bar{x}$ .

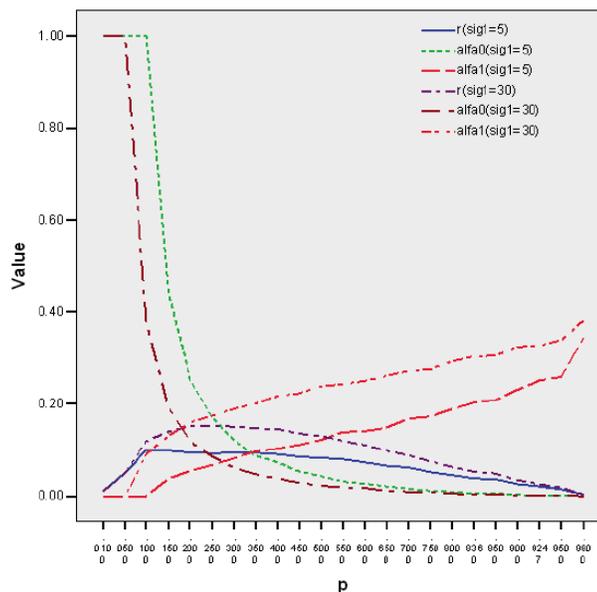
Dependences of the risk function and error probabilities on the variance of the alternative hypothesis in the considered tests are shown in Figure 7. As was mentioned above, for classical and Bernardo methods there are the following ratios:  $\alpha_0 = \beta_1$  and  $\alpha_1 = \beta_0$ . In CBM they differ from each other. Hence, the optimality of CBM in comparison with the considered tests is obvious.

**Example 2.** For saving in room, we bring similar computation results for the second example with minimal explanations, because the behavior and ratios of the results obtained by the considered tests are similar to the results of the first example.

On the basis of the results given in Tables 13, 14 and Figures 14, 15, the following is evident. The probability of rejection of the basic hypothesis when it is true (that is the same, the probability of acceptance of the alternative hypothesis when it is false) at increasing variance of the basic hypothesis tends to

**Table 6: Computation Results of the Risk Function and Error Probabilities Depending on the a Priori Probability in the Bayesian and Bernardo Tests**

A priori probability	Initial values $m = 10,000, n = 5$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, \sigma_1^2 = 5$			Initial values $m = 10,000, n = 5$ $\mu_0 = 1, \mu_1 = 3$ $\sigma^2 = 1, \sigma_1^2 = 30$		
	$p$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$
0.01	0.0100	1	0	0.0100	1	0
0.05	0.0500	1	0	0.0500	1	0
0.0753	0.0753	1	0	0.0912	0.7178	0.0402
0.1	0.1000	1	0	0.1193	0.3791	0.0904
0.15	0.0983	0.4422	0.0376	0.1402	0.1908	0.1313
0.1640	0.0959	0.3651	0.0431	0.1455	0.1670	0.1413
0.2	0.0941	0.2533	0.0543	0.1516	0.1190	0.1598
0.25	0.0928	0.1708	0.0668	0.1522	0.0855	0.1744
0.30	0.0935	0.1211	0.0817	0.1508	0.0608	0.1894
0.35	0.0937	0.0879	0.0968	0.1470	0.0472	0.2007
0.4	0.0906	0.0721	0.1029	0.1454	0.0374	0.2174
0.45	0.0850	0.0533	0.1110	0.1357	0.0289	0.2231
0.5	0.0824	0.0425	0.1223	0.1304	0.0220	0.2388
0.55	0.0796	0.0312	0.1388	0.1198	0.0186	0.2436
0.6	0.0727	0.0273	0.1408	0.1104	0.0174	0.2500
0.65	0.0653	0.0206	0.1484	0.0985	0.0107	0.2615
0.7	0.0616	0.0161	0.1678	0.0882	0.0095	0.2718
0.75	0.0517	0.0113	0.1727	0.0748	0.0079	0.2756
0.8	0.0452	0.0093	0.1887	0.0629	0.0055	0.2925
0.85	0.0364	0.0061	0.2082	0.0486	0.0034	0.3046
0.9	0.0256	0.0027	0.2314	0.0343	0.0020	0.3247
0.95	0.0145	0.0016	0.2589	0.0182	0.0013	0.3394
0.99	0.0035	0.0001	0.3435	0.0039	0.0001	0.3815



**Figure 6: Dependences of the risk function and error probabilities on the a priori probability in the Bayesian test.**

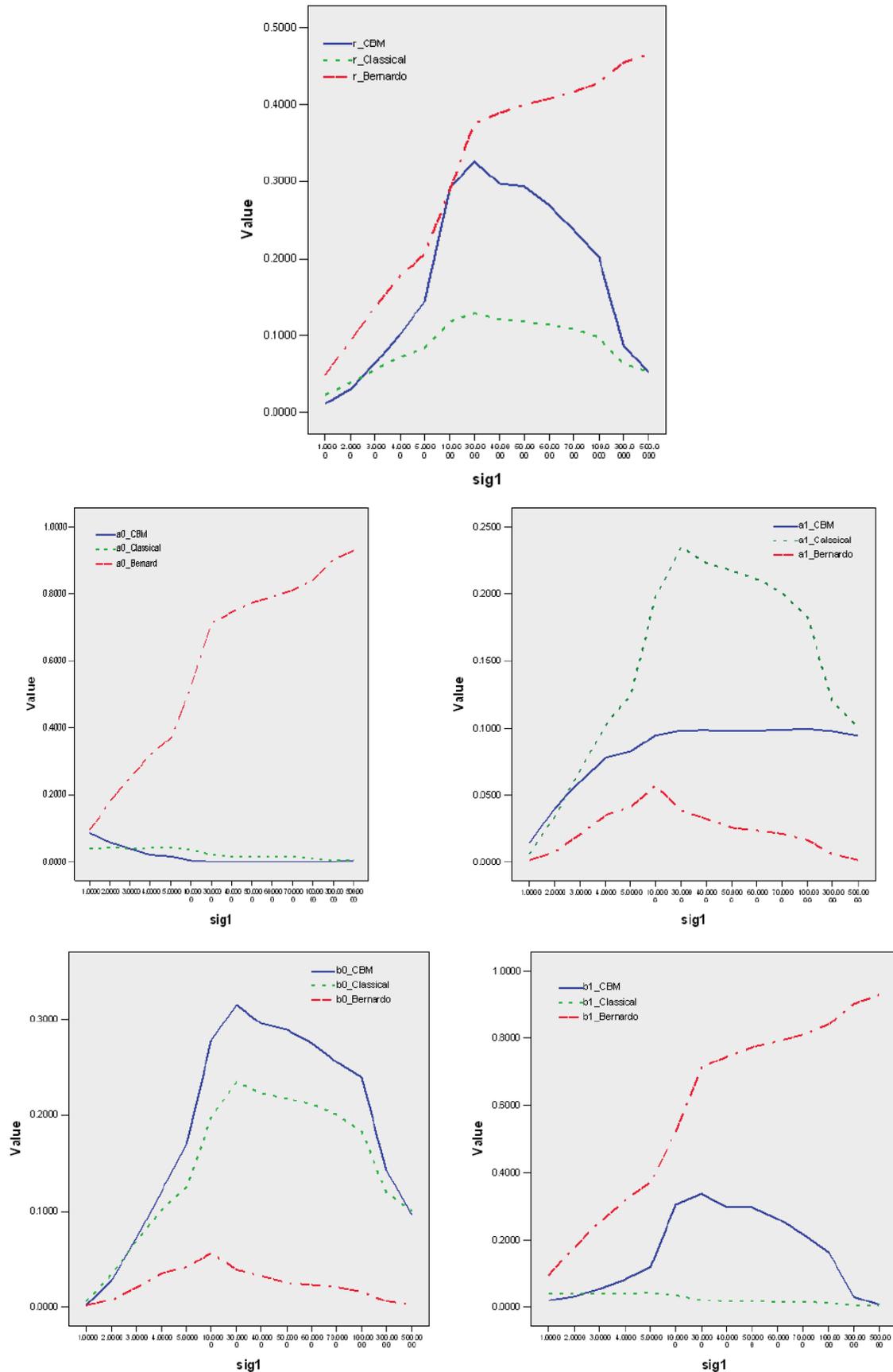
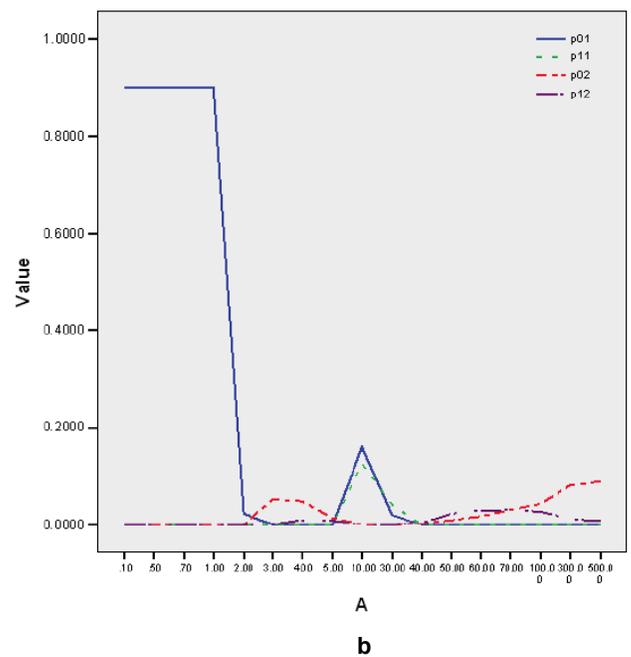
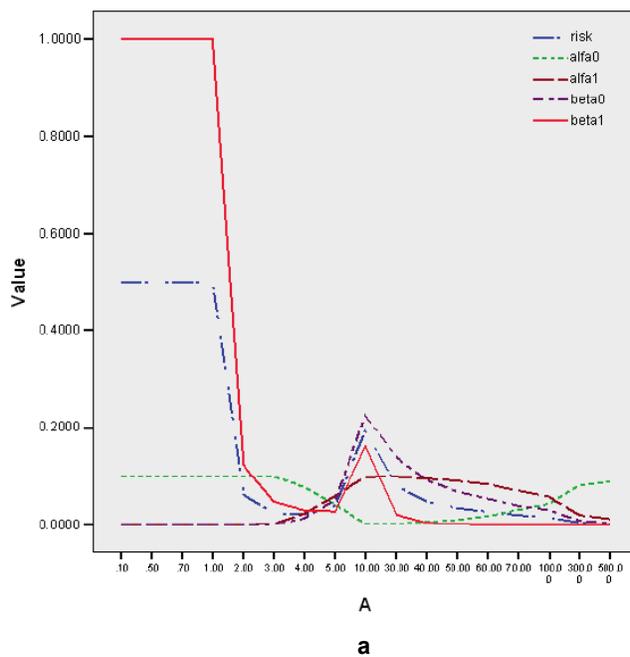


Figure 7: Dependences of the risk function and error probabilities on the variance of the alternative hypothesis in the considered tests for example 1.

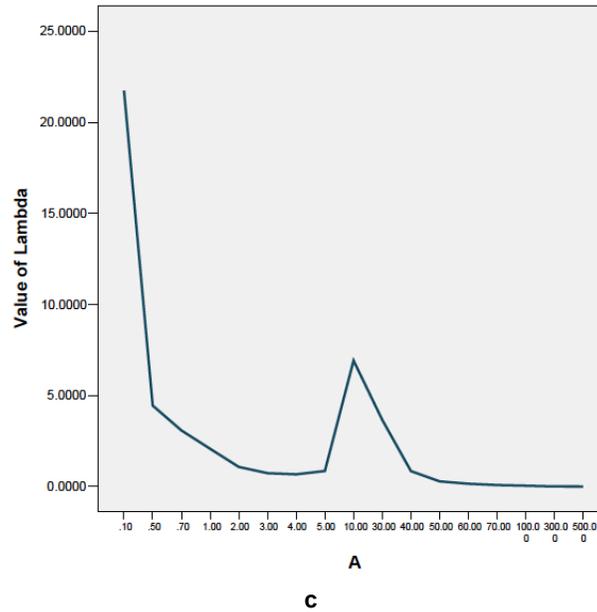
**Table 7: Computation Results of the Lagrange Multiplier, the Risk Function and Error Probabilities of CBM Depending on the Length of the Domain of Definition of Uniformly Distributed Random Variable at the Alternative Hypothesis**

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, \sigma^2 = 1$ $p = 0.5$			Computed values									
A	xe0	xe1	$\lambda$	r	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$	$p_{12}$
0.1	2.8096	0.0109	21.7270	0.5000	0.1000	0	0	1	0.9000	0	0	0
0.5	3.9502	-0.2002	4.4443	0.5000	0.1000	0	0	1	0.9000	0	0	0
0.7	3.2897	0.1467	3.1860	0.5000	0.1000	0	0	1	0.9000	0	0	0
1	3.5571	-0.0549	2.0753	0.5000	0.1000	0	0	1	0.9000	0	0	0
2	2.7300	0.7892	1.0807	0.0613	0.1000	0	0	0.1226	0.0226	0	0	0
3	3.6723	0.3437	0.7362	0.0239	0.0986	0.0014	0.0006	0.0471	0	0	0.0515	0.0008
4	3.9171	0.4262	0.6715	0.0212	0.0773	0.0227	0.0130	0.0294	0	0	0.0479	0.0097
5	3.1847	-1.5526	0.8535	0.0392	0.0414	0.0586	0.0508	0.0276	0	0	0.0138	0.0078
10	3.4541	-4.4365	6.9195	0.1925	0.0019	0.0981	0.2237	0.1613	0.1594	0.1256	0	0
30	2.7012	5.1465	3.6624	0.0800	0.0015	0.0985	0.1391	0.0209	0.0194	0.0406	0	0
40	3.3332	11.9239	0.8495	0.0472	0.0056	0.0944	0.0909	0.0035	0	0	0.0021	0.0035
50	3.4695	4.3492	0.2840	0.0342	0.0100	0.0900	0.0676	0.0009	0	0	0.0091	0.0224
60	2.0492	-17.8571	0.1546	0.0265	0.0175	0.0825	0.0528	0.0002	0	0	0.0173	0.0297
70	3.1209	5.5814	0.0827	0.0195	0.0307	0.0693	0.0388	0.0001	0	0	0.0306	0.0305
100	2.9632	35.1740	0.0459	0.0150	0.0430	0.0570	0.0300	0	0	0	0.0430	0.0270
300	2.9835	19.2266	0.0085	0.0042	0.0803	0.0197	0.0083	0	0	0	0.0803	0.0114
500	3.3755	-23.0923	0.0047	0.0019	0.0884	0.0116	0.0039	0	0	0	0.0884	0.0077

**Note:** The averaged values of  $n$  simulated random variables distributed in accordance with the hypotheses  $H_0$  and  $H_A$  are given in the columns of xe0 and xe1, respectively, i.e.  $xe0 \sim p(\bar{x} | H_0)$  and  $xe1 \sim p(\bar{x} | H_A)$  take place.



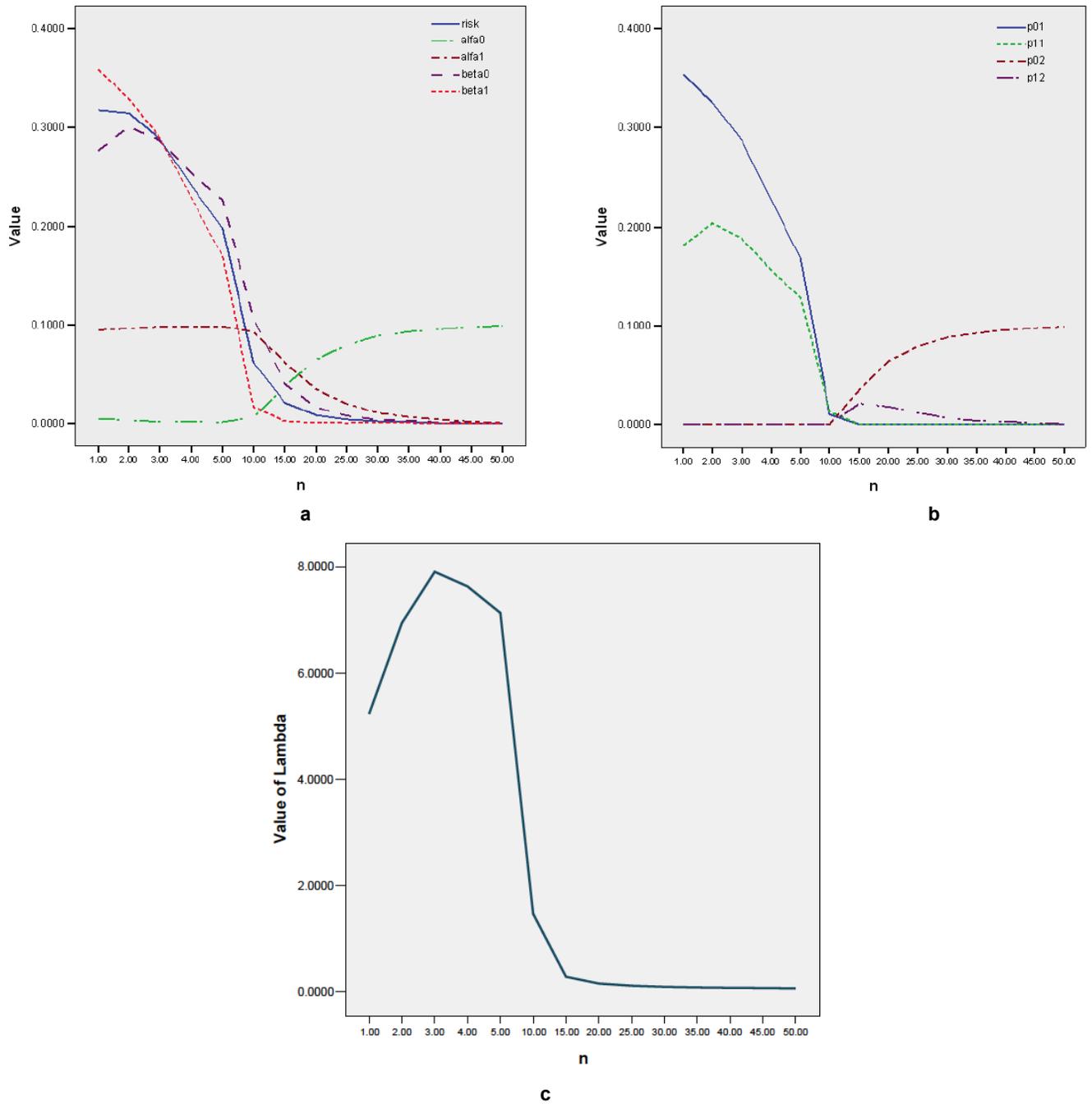
(Figure 8). Continued.



**Figure 8:** Dependences of the Lagrange multiplier, the risk function and error probabilities on the length of the domain of definition of uniformly distributed random variable at the alternative hypothesis in the constrained Bayesian test.

**Table 8: Computation Results of the Lagrange Multiplier, the Risk Function and Error Probabilities of CBM Depending on the Sample Size**

Initial values $m = 10,000, A = 10$ $\mu_0 = 3, \sigma^2 = 1$ $p = 0.5, \alpha = 0.05$			Computed values									
$n$	$xe0$	$xe1$	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$	$p_{12}$
1	4.5576	7.3393	5.2301	0.3178	0.0051	0.0949	0.2768	0.3588	0.3537	0.1819	0	0
2	2.0512	0.1734	6.9418	0.3146	0.0035	0.0965	0.3005	0.3288	0.3253	0.2040	0	0
3	3.2742	-3.9028	7.9040	0.2880	0.0018	0.0982	0.2868	0.2891	0.2873	0.1886	0	0
4	4.3315	3.1191	7.6294	0.2416	0.0017	0.0983	0.2541	0.2291	0.2274	0.1558	0	0
5	2.6525	5.0844	7.1296	0.1984	0.0015	0.0985	0.2270	0.1698	0.1683	0.1285	0	0
10	2.9451	-0.3409	1.4727	0.0620	0.0067	0.0933	0.1067	0.0172	0.0105	0.0134	0	0
15	3.1862	-0.5697	0.2892	0.0216	0.0378	0.0622	0.0406	0.0025	0	0	0.0353	0.0216
20	2.9242	-1.7879	0.1604	0.0087	0.0650	0.0350	0.0167	0.0008	0	0	0.0642	0.0183
25	3.0555	-0.0751	0.1186	0.0043	0.0798	0.0202	0.0081	0.0006	0	0	0.0792	0.0121
30	2.9433	-0.3194	0.0978	0.0024	0.0890	0.0110	0.0040	0.0008	0	0	0.0882	0.0070
35	3.0924	-0.1571	0.0866	0.0019	0.0932	0.0068	0.0030	0.0007	0	0	0.0925	0.0038
40	2.9752	0.0476	0.0791	0.00055	0.0959	0.0041	0.0009	0.0002	0	0	0.0957	0.0032
45	3.3062	0.0921	0.0746	0.0004	0.0980	0.0020	0.0005	0.0003	0	0	0.0977	0.0015
50	2.9466	-0.7556	0.0685	0.0002	0.0992	0.0008	0.0002	0.0002	0	0	0.0990	0.0006



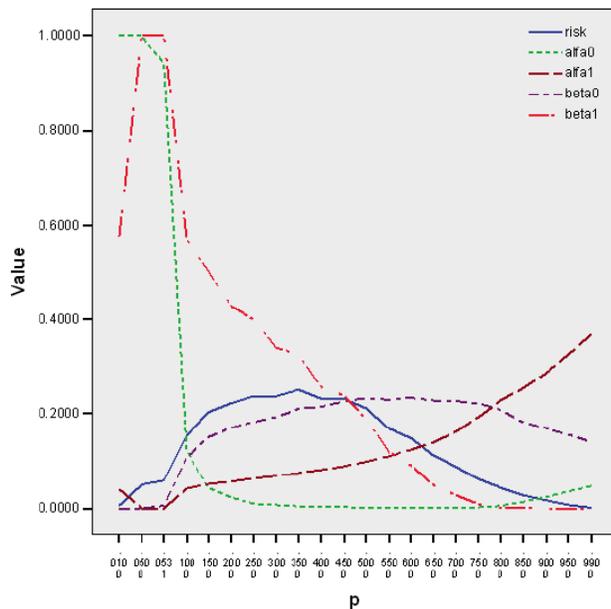
**Figure 9:** Dependences of the Lagrange multiplier, the risk function and error probabilities on the sample size in the constrained Bayesian test.

one. At the same time, the probability of rejection of the alternative hypothesis when it is true (that is the same, the probability of acceptance of the basic hypothesis when it is false) tends to zero. The first of these probabilities for the Bernardo's correction practically does not change and the second one tends to one. It is not difficult to be convinced that these behaviors are not correct because, with increasing indefiniteness in relation to the basic hypothesis, the probability of its rejection must tend to one and the probability of its acceptance must tend to zero, as they do in the classical Bayesian method. The behavior of CBM

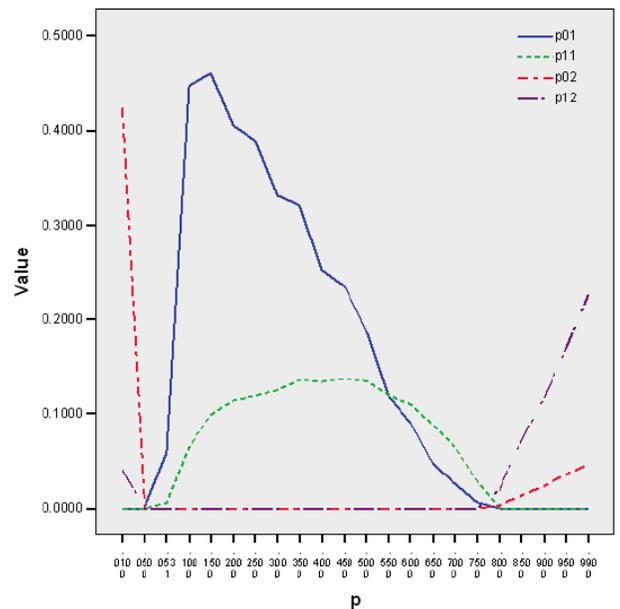
coincides with the classical Bayesian test in principle, but it is more thoughtful, because with the increasing  $\sigma^2$ , the indefiniteness in relation to both hypotheses increases, and it becomes more and more difficult to make a correct decision in relation to the tested hypotheses. In accordance with that, the regions of acceptance of both hypotheses increase (see  $\beta_0$  and  $\beta_1$ ) and consequently the region of their intersection increases too. Therefore, the probabilities of supposition for both hypotheses  $p_{01}$  and  $p_{11}$  to be true

**Table 9: Computation Results of the Lagrange Multiplier, the Risk Function and Error Probabilities of CBM Depending on the a Priori Probability**

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, \sigma^2 = 1$ $A = 10, \alpha = 0.05$			Computed values									
$p$	$xe0$	$xe1$	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$P_{01}$	$P_{11}$	$P_{02}$	$P_{12}$
0.01	2.7537	-0.4090	0.1544	0.0058	1	0.0404	0	0.5761	0	0	0.4239	0.0404
0.05	2.3094	3.3013	0.9767	0.0500	1	0	0	1	0	0	0	0
0.0531	2.9630	2.5257	1.0021	0.0585	0.9415	0	0.0057	1	0.0585	0.0057	0	0
0.1	2.9099	2.3337	1.6789	0.1529	0.1221	0.0420	0.1066	0.5695	0.4474	0.0646	0	0
0.15	2.8971	0.8313	2.5025	0.2031	0.0427	0.0513	0.1502	0.5028	0.4601	0.0989	0	0
0.2	3.1834	-2.1573	3.2837	0.2220	0.0224	0.0569	0.1706	0.4275	0.4051	0.1137	0	0
0.25	3.8357	2.3950	4.1382	0.2364	0.0101	0.0633	0.1820	0.3994	0.3893	0.1187	0	0
0.30	3.0008	3.7126	4.8407	0.2365	0.0073	0.0683	0.1930	0.3381	0.3308	0.1247	0	0
0.35	2.6914	-4.2711	5.9248	0.2503	0.0045	0.0745	0.2101	0.3250	0.3205	0.1356	0	0
0.4	2.8161	-0.1680	6.1932	0.2311	0.0032	0.0812	0.2145	0.2560	0.2528	0.1333	0	0
0.45	2.1548	-0.8875	7.2070	0.2313	0.0027	0.0887	0.2258	0.2380	0.2353	0.1371	0	0
0.5	3.3688	1.1759	7.3491	0.2113	0.0018	0.0982	0.2326	0.1900	0.1882	0.1344	0	0
0.55	2.0364	-2.2779	6.7140	0.1685	0.0010	0.1099	0.2291	0.1190	0.1180	0.1192	0	0
0.6	3.1860	-3.1922	6.7302	0.1481	0.0010	0.1235	0.2332	0.0913	0.0903	0.1097	0	0
0.65	3.6665	-4.0843	4.7608	0.1112	0.0017	0.1397	0.2275	0.0486	0.0469	0.0878	0	0
0.7	3.6420	0.0720	3.4596	0.0877	0.0023	0.1613	0.2264	0.0282	0.0259	0.0651	0	0
0.75	3.6826	-4.9689	1.7012	0.0618	0.0025	0.1925	0.2211	0.0087	0.0062	0.0286	0	0
0.8	3.2409	1.8051	0.6450	0.0433	0.0055	0.2280	0.2080	0.0021	0	0	0.0034	0.0200
0.85	2.6527	1.7469	0.2297	0.0276	0.0141	0.2534	0.1811	0.0005	0	0	0.0136	0.0723
0.9	3.3631	-0.0057	0.0828	0.0171	0.0238	0.2858	0.1690	0.0002	0	0	0.0236	0.1168
0.95	2.9828	2.3266	0.0258	0.0078	0.0356	0.3236	0.1558	0	0	0	0.0356	0.1678
0.99	3.3919	-1.5773	0.0039	0.0014	0.0468	0.3664	0.1399	0	0	0	0.0468	0.2265

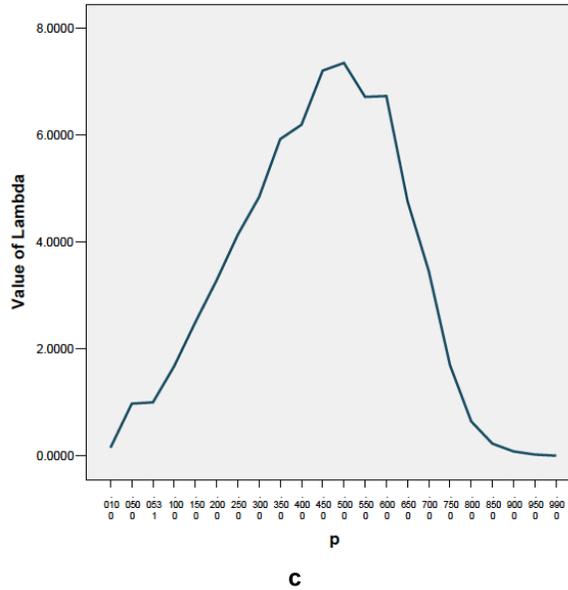


a



b

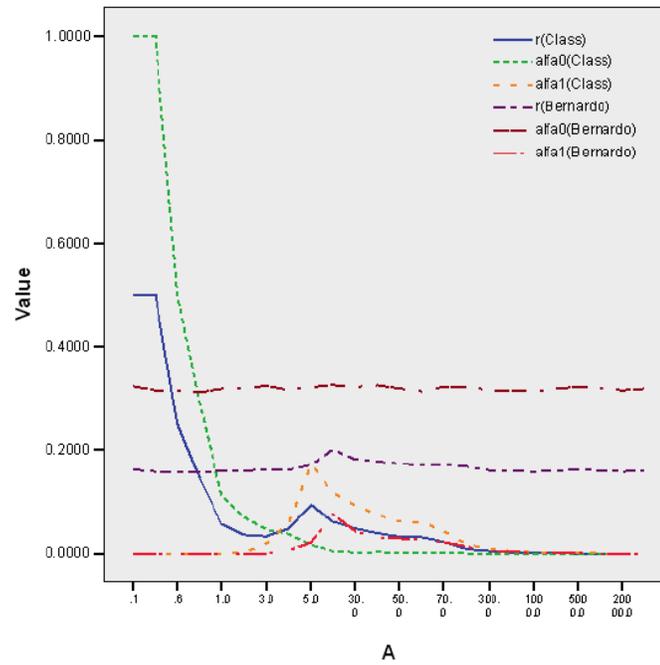
(Figure 10). Continued.



**Figure 10:** Dependences of the Lagrange multiplier, the risk function and error probabilities on the a priori probability in the constrained Bayesian test.

**Table 10: Computation Results of the Risk Function and Error Probabilities in the Bayesian and Bernardo Tests Depending on the Length of the Domain of Definition of Uniformly Distributed Random Variable at the Alternative Hypothesis**

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, \sigma^2 = 1$ $p = 0.5$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$A$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$
0.1	0.5000	1	0	0.1615	0.3229	0
0.5	0.5000	1	0	0.1576	0.3152	0
0.7	0.2493	0.4985	0	0.1559	0.3118	0
1	0.1442	0.2883	0	0.1590	0.3179	0
2	0.0558	0.1117	0	0.1592	0.3183	0
3	0.0352	0.0693	0.0010	0.1621	0.3241	0.0001
4	0.0323	0.0450	0.0195	0.1607	0.3160	0.0054
5	0.0468	0.0372	0.0563	0.1710	0.3213	0.0208
10	0.0938	0.0161	0.1715	0.2003	0.3254	0.0751
30	0.0610	0.0044	0.1176	0.1819	0.3218	0.0420
40	0.0474	0.0022	0.0927	0.1779	0.3251	0.0306
50	0.0385	0.0026	0.0744	0.1734	0.3187	0.0282
60	0.0319	0.0024	0.0614	0.1696	0.3129	0.0262
70	0.0304	0.0016	0.0593	0.1717	0.3221	0.0213
100	0.0217	0.0008	0.0425	0.1678	0.3220	0.0135
300	0.0085	0.0002	0.0169	0.1600	0.3159	0.0040
500	0.0048	0.0001	0.0095	0.1589	0.3147	0.0032
1000	0.0022	0	0.0043	0.1578	0.3138	0.0018
2000	0.0011	0	0.0022	0.1592	0.3177	0.0007
5000	0.0007	0	0.0014	0.1612	0.3220	0.0004
10000	0.0002	0	0.0004	0.1601	0.3201	0.0001
20000	0.00015	0	0.0003	0.1576	0.3150	0.0001
50000	0.00005	0	0.0001	0.1591	0.3182	0



**Figure 11:** Dependences of the risk function and error probabilities on the interval of definition of uniformly distributed random variable in the Bayesian and Bernardo tests.

**Table 11: Computation Results of the Risk Function and Error Probabilities in the Bayesian and Bernardo Tests Depending on the Sample Size**

Initial values $m = 10,000$ $\mu_0 = 3, \sigma^2 = 1$ $p = 0.5, A = 10$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$
1	0.1214	0.0413	0.2015	0.2117	0.3188	0.1046
2	0.1219	0.0254	0.2184	0.2132	0.3235	0.1028
3	0.1152	0.0218	0.2086	0.2080	0.3199	0.0960
4	0.1055	0.0200	0.1911	0.2016	0.3210	0.0822
5	0.0965	0.0172	0.1758	0.1983	0.3215	0.0751
10	0.0543	0.0107	0.0979	0.1785	0.3219	0.0351
15	0.0325	0.0092	0.0557	0.1683	0.3172	0.0194
20	0.0185	0.0077	0.0293	0.1643	0.3174	0.0112
25	0.0106	0.0068	0.0143	0.1589	0.3147	0.0031
30	0.0071	0.0050	0.0092	0.1596	0.3170	0.0022
35	0.0050	0.0059	0.0041	0.1608	0.3204	0.0012
40	0.0032	0.0044	0.0020	0.1615	0.3226	0.0003
45	0.0035	0.0060	0.0010	0.1568	0.3134	0.0001
50	0.0027	0.0053	0.0002	0.1587	0.3174	0

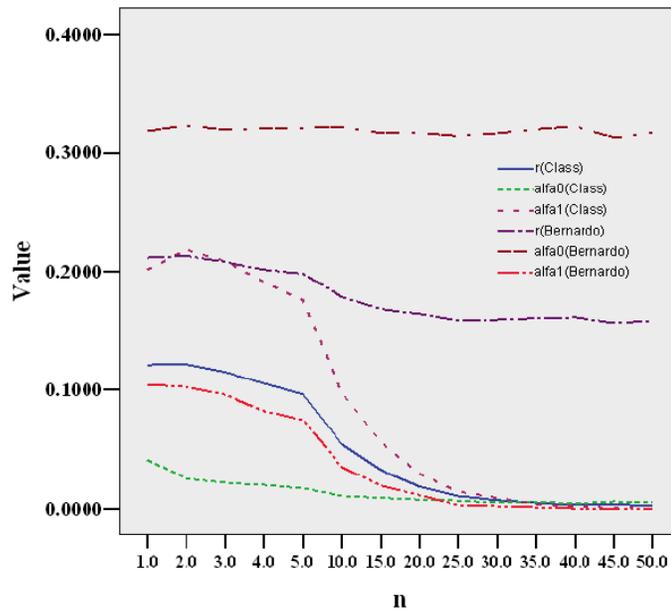


Figure 12: Dependences of the risk function and error probabilities on the sample size in the Bayesian and Bernardo tests.

Table 12: Computation Results of the Risk Function and Error Probabilities in the Bayesian and Bernardo Tests Depending on the a Priori Probability

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, \sigma^2 = 1$ $A = 10$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$
0.01	0.0100	1	0	0.0738	0.3160	0.0714
0.05	0.0500	1	0	0.0847	0.3106	0.0728
0.0531	0.0534	0.9735	0.0018	0.0850	0.3057	0.0726
0.1	0.1000	0.2435	0.0841	0.0961	0.3161	0.0717
0.15	0.1137	0.1261	0.1115	0.1090	0.3167	0.0723
0.2	0.1157	0.0866	0.1230	0.1189	0.3083	0.0716
0.25	0.1166	0.0579	0.1362	0.1350	0.3208	0.0731
0.30	0.1144	0.0408	0.1460	0.1510	0.3208	0.0783
0.35	0.1096	0.0298	0.1526	0.1587	0.3195	0.0721
0.4	0.1085	0.0259	0.1635	0.1747	0.3263	0.0736
0.45	0.1047	0.0190	0.1748	0.1822	0.3198	0.0696
0.5	0.0932	0.0141	0.1723	0.1962	0.3167	0.0757
0.55	0.0922	0.0142	0.1876	0.2048	0.3121	0.0737
0.6	0.0786	0.0097	0.1819	0.2163	0.3129	0.0714
0.65	0.0722	0.0081	0.1913	0.2317	0.3182	0.0711
0.7	0.0626	0.0054	0.1959	0.2463	0.3214	0.0711
0.75	0.0554	0.0056	0.2048	0.2564	0.3173	0.0736
0.8	0.0461	0.0036	0.2163	0.2666	0.3146	0.0747
0.85	0.0353	0.0025	0.2213	0.2814	0.3179	0.0749
0.9	0.0245	0.0015	0.2316	0.2961	0.3204	0.0771
0.95	0.0130	0.0007	0.2461	0.3063	0.3186	0.0734
0.99	0.0032	0.0004	0.2840	0.3141	0.3165	0.0734

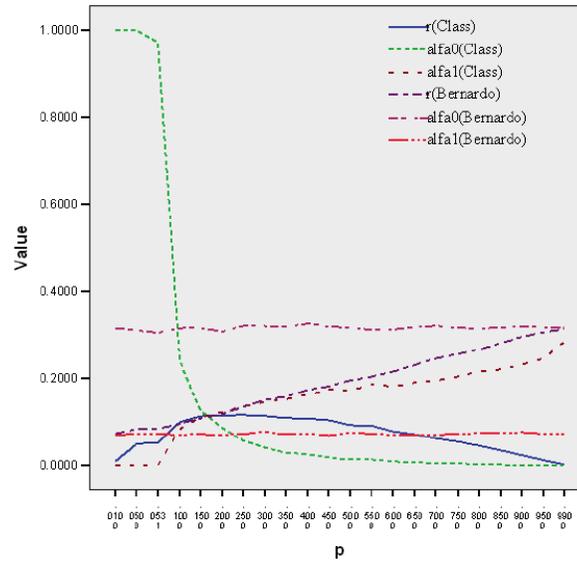


Figure 13: Dependences of the risk function and error probabilities on the a priori probability in the Bayesian and Bernardo tests.

Table 13: Computation Results of the Risk Function and Error Probabilities in the Bayesian and Bernardo Tests Depending on the Variance of Basic Hypothesis

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, p = 0.5$ $A = 10, \alpha = 0.05$	Computed values for classical Bayesian method			Computed values for Bernardo's correction		
	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$	$r$	$\alpha_0 = \beta_1$	$\alpha_1 = \beta_0$
0.0001	0.0014	0	0.0027	0.1600	0.3189	0.0011
0.001	0.0040	0.0004	0.0077	0.1623	0.3221	0.0026
0.01	0.0124	0.0011	0.0237	0.1608	0.3146	0.0070
0.05	0.0251	0.0031	0.0471	0.1668	0.3191	0.0144
0.08	0.0312	0.0029	0.0595	0.1713	0.3229	0.0197
0.1	0.0359	0.0043	0.0675	0.1693	0.3152	0.0234
0.5	0.0695	0.0100	0.1290	0.1829	0.3137	0.0521
1	0.0958	0.0152	0.1763	0.1939	0.3118	0.0759
2	0.1287	0.0246	0.2328	0.2089	0.3179	0.0998
3	0.1481	0.0286	0.2675	0.2221	0.3139	0.1303
4	0.1720	0.0371	0.3069	0.2303	0.3135	0.1470
5	0.1904	0.0417	0.3391	0.2400	0.3183	0.1616
10	0.2453	0.0636	0.4271	0.2770	0.3241	0.2299
30	0.3691	0.1252	0.6130	0.3606	0.3160	0.4051
40	0.4067	0.1552	0.6582	0.3901	0.3213	0.4588
50	0.4357	0.1746	0.6968	0.4172	0.3254	0.5089
60	0.4543	0.1945	0.7141	0.4414	0.3182	0.5646
70	0.4731	0.2177	0.7284	0.4634	0.3132	0.6135
100	0.5199	0.2816	0.7582	0.5071	0.3124	0.7017
300	0.5551	0.7993	0.3109	0.6361	0.3070	0.9652
500	0.5000	1	0	0.6601	0.3218	0.9983
1000	0.5000	1	0	0.6625	0.3251	1
2000	0.5000	1	0	0.6593	0.3187	1
5000	0.5000	1	0	0.6564	0.3129	1
10000	0.5000	1	0	0.6611	0.3221	1
20000	0.5000	1	0	0.6610	0.3220	1
50000	0.5000	1	0	0.6580	0.3159	1

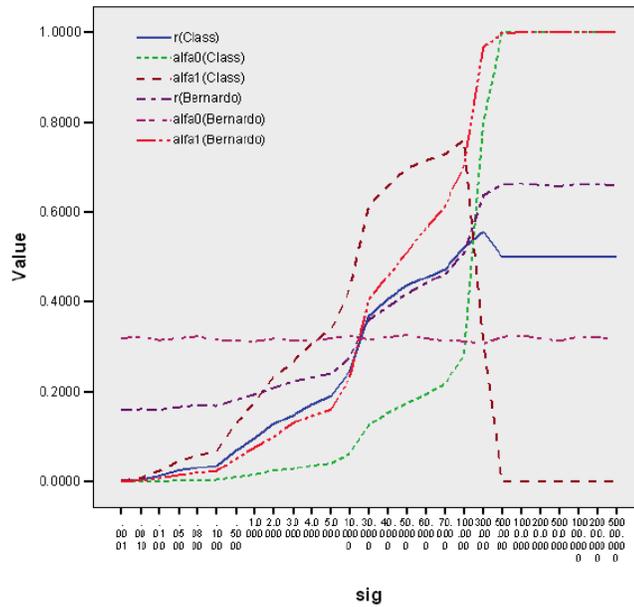
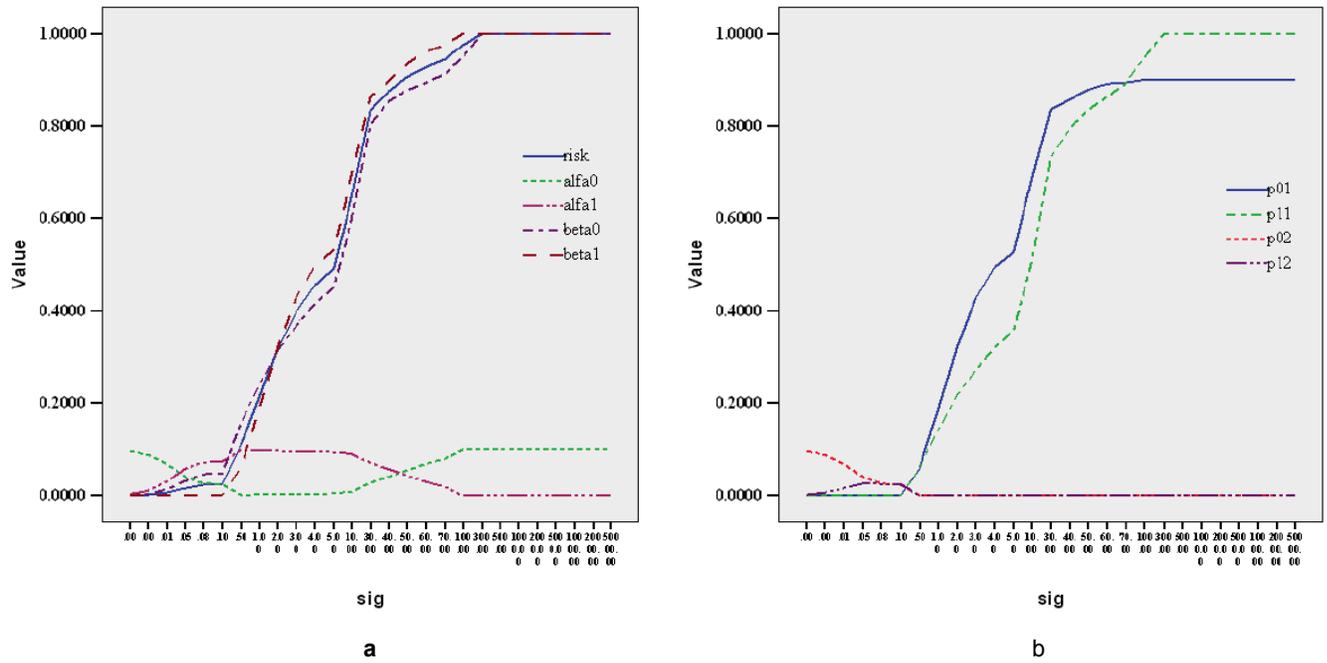


Figure 14: Dependences of the risk function and error probabilities in the Bayesian and Bernardo tests depending on the variance of the basic hypothesis.

Table 14: Computation Results of the Lagrange Multiplier, the Risk Function and Error Probabilities of CBM Depending on the Variance of the Basic Hypothesis

Initial values $m = 10,000, n = 5$ $\mu_0 = 3, p = 0.5$ $A = 10, \alpha = 0.05$			Computed values for CBM									
$\sigma^2$	xe0	xe1	$\lambda$	$r$	$\alpha_0$	$\alpha_1$	$\beta_0$	$\beta_1$	$p_{01}$	$p_{11}$	$p_{02}$	$p_{12}$
0.0001	2.9978	-1.1466	0.0022	0.0007	0.0964	0.0036	0.0014	0	0	0	0.0964	0.0022
0.001	2.9874	4.7805	0.0076	0.0019	0.0893	0.0107	0.0039	0	0	0	0.0893	0.0068
0.01	3.0010	-2.7428	0.0290	0.0070	0.0693	0.0307	0.0140	0	0	0	0.0693	0.0167
0.05	2.9201	0.5121	0.0982	0.0164	0.0404	0.0596	0.0326	0.0001	0	0	0.0403	0.0270
0.08	3.1208	0.5918	0.1832	0.0238	0.0270	0.0730	0.0471	0.0005	0	0	0.0265	0.0259
0.1	3.1015	4.2024	0.2312	0.0254	0.0252	0.0748	0.0495	0.0012	0	0	0.0240	0.0253
0.5	2.4304	3.7244	4.3946	0.1101	0.0015	0.0985	0.1586	0.0616	0.0601	0.0601	0	0
1	2.2747	1.3500	7.5685	0.2122	0.0025	0.0975	0.2377	0.1868	0.1843	0.1402	0	0
2	3.7352	-1.8508	7.8123	0.3181	0.0028	0.0972	0.3140	0.3221	0.3193	0.2168	0	0
3	3.8066	-0.8789	7.5968	0.3982	0.0032	0.0968	0.3669	0.4294	0.4262	0.2701	0	0
4	2.4517	-1.2185	7.1248	0.4562	0.0031	0.0969	0.4152	0.4972	0.4941	0.3183	0	0
5	4.5725	0.1347	6.5919	0.4920	0.0054	0.0946	0.4515	0.5325	0.5271	0.3569	0	0
10	2.5069	-2.7955	5.2387	0.6473	0.0087	0.0913	0.5992	0.6954	0.6867	0.5079	0	0
30	1.7302	-2.7131	3.2087	0.8337	0.0286	0.0714	0.8039	0.8634	0.8348	0.7325	0	0
40	-0.3488	0.0214	2.7992	0.8764	0.0421	0.0579	0.8538	0.8990	0.8569	0.7959	0	0
50	1.1060	-3.0332	2.5140	0.9063	0.0561	0.0439	0.8778	0.9349	0.8788	0.8339	0	0
60	9.6382	2.0775	2.3004	0.9274	0.0700	0.0300	0.8935	0.9612	0.8912	0.8635	0	0
70	1.2111	5.7723	2.1313	0.9437	0.0808	0.0192	0.9128	0.9746	0.8938	0.8936	0	0
100	-0.5579	-1.8288	2.1325	0.9746	0.1000	0	0.9492	1	0.9000	0.9492	0	0
300	-6.2150	-1.0295	3.6622	1	0.1000	0	1	1	0.9000	1	0	0
500	-5.9096	1.3254	4.9059	1	0.1000	0	1	1	0.9000	1	0	0
1000	7.3755	-1.5548	6.9129	1	0.1000	0	1	1	0.9000	1	0	0
2000	-20.1732	-2.4341	9.7693	1	0.1000	0	1	1	0.9000	1	0	0
5000	38.5150	-1.4164	15.1238	1	0.1000	0	1	1	0.9000	1	0	0
10000	16.3573	2.0443	21.7477	1	0.1000	0	1	1	0.9000	1	0	0
20000	46.7002	2.0201	29.2970	1	0.1000	0	1	1	0.9000	1	0	0
50000	-90.0415	-1.1883	50.7818	1	0.1000	0	1	1	0.9000	1	0	0



**Figure 15:** Dependences of the risk function and error probabilities of CBM depending on the variance of basic hypothesis.

**a)** Probabilities of incorrect making decision. **b)** Probabilities of no making decision.

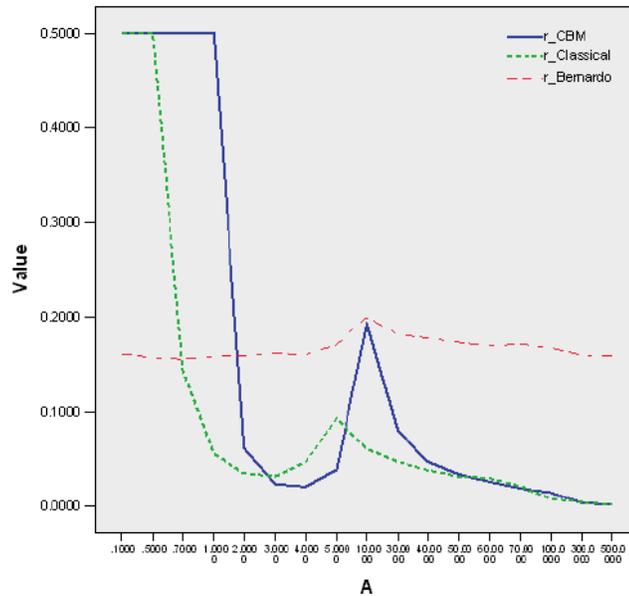
increase. Because  $\alpha_0 \rightarrow 0.1$  and  $\alpha_1 \rightarrow 0$ , the following takes place  $p_{01} \rightarrow 0.9$  and  $p_{11} \rightarrow 1$ . At the same time, the risk function  $r \rightarrow 1$ . Because of increasing regions of acceptance of both hypotheses, the regions which do not belong to hypotheses acceptance regions become empty and hence the appropriate probabilities  $p_{02}$  and  $p_{12}$  become equal to zero.

The graphs of dependences of the risk function and error probabilities of the considered methods on the length of the domain of definition of the uniformly distributed random variable with the alternative

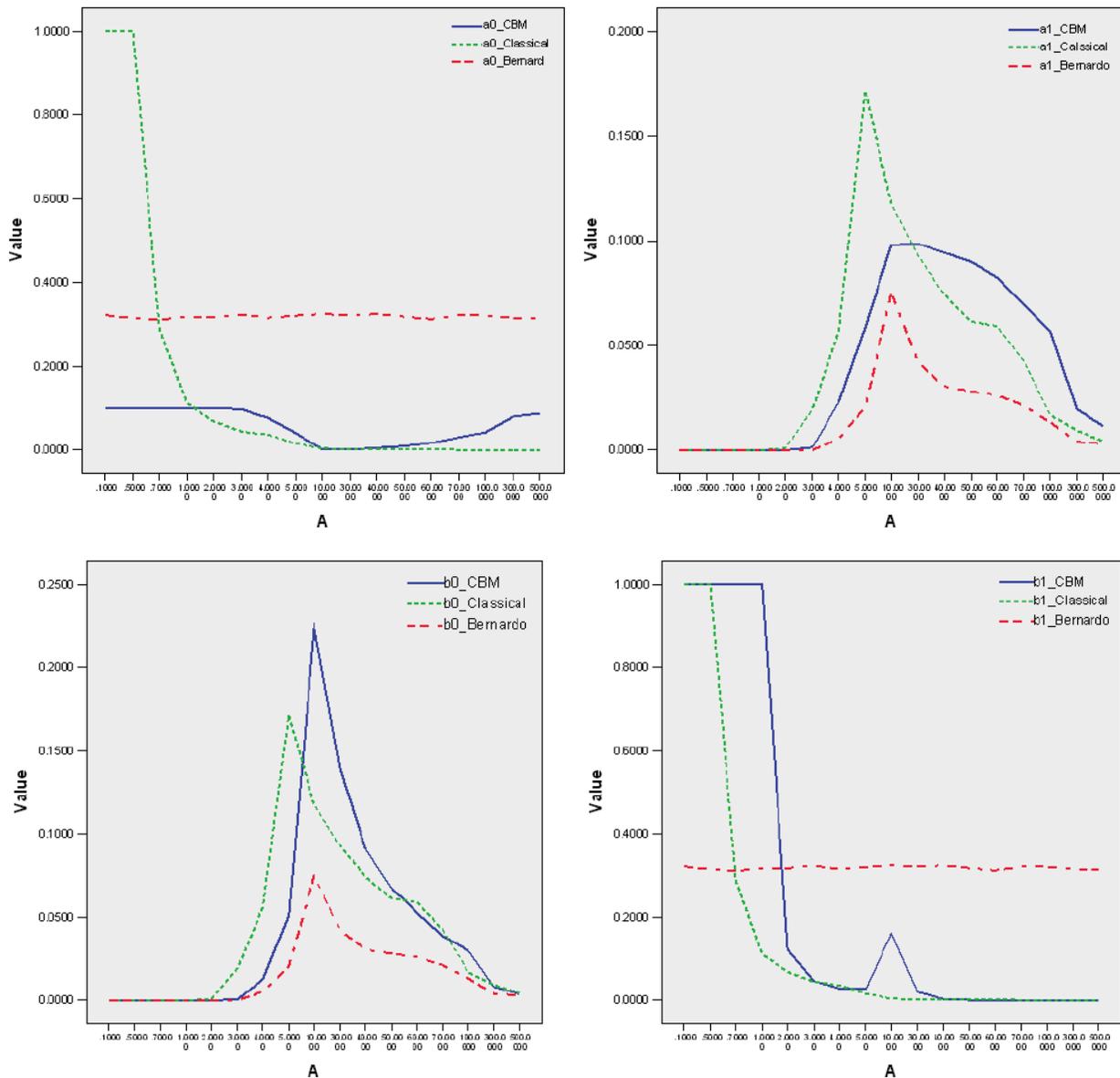
hypothesis are given below. The optimality of CBM in comparison with two other tests is obvious (similarly to the first example given in Figure 7).

### 5. DISCUSSION

The computation results show that the choice of a priori probabilities on the basis of a preliminarily stated aim does not make sense, because it deteriorates the results of the test. A priori probabilities must be chosen on the basis of objective information [14, 26], and, the more objectively they are chosen, the better is for the



(Figure 16). Continued.



**Figure 16:** Dependences of the risk function and error probabilities on the length of the domain of definition of the uniformly distributed random variable with the alternative hypothesis in the considered tests for example 2.

final results. The choice of a priori probabilities on the basis of the sample which is used for making a decision is not a good solution of the problem either, because they are subjective probabilities which are determined by a concrete sample, i.e. by the information contained in the concrete sample. The purpose of the a priori probabilities in the Bayesian statement is to decrease the subjective factor (which is contained in a concrete sample as it is determined by concrete (subjective) values of the random component) on the basis of previous knowledge and experience for neutralization of the concrete values of the random component. If there is no objective information for

determination of the a priori probabilities of tested hypotheses, it is better to choose equal values of these probabilities for all hypotheses.

**6. CONCLUSION**

The optimality of CBM for testing composite hypotheses is shown. In particular, there is shown that the risk of the Lindley's paradox does not exist for it. Superiority of CBM in comparison with the Bayesian test with classical and Bernardo's choice of the a priori probabilities of tested hypotheses is shown with theoretical consideration that is supported by computation results of many characteristics of different

practical examples. Necessity in testing composite hypotheses arises very often at solving many practical problems. In particular, they are very important in many medical and biological applications. It is shown that, CBM keeps the optimal properties at testing composite hypotheses and, therefore it gives great opportunities for making reliable decisions in theory and practice.

## ACKNOWLEDGEMENTS

The author is grateful to the Editor, members of Editorial committee, and to the anonymous referee for deep, insightful, and encouraging comments that helped us tremendously.

This research was supported by Shota Rustaveli National Science Foundation of Georgia grant # 216680" that I forgot in previous case.

## APPENDIX

The code of the program for computation of Example 1

% Example for normal versus normal

```
%
%
% clear all
% format short

% m - Number of generated random values for
% computation integrals;
% n - sample size, i.e. dimension of the normal vector;
% a - restriction level alfa;
% LamL, LamU - Lower and upper values of Lambda;
% Lam - Lambda corresponding to the solution of
% restriction equation;
% miu0, miu1 - mathematical expectations for H0 and
% H1;
% sig, sig1 - variances for H0 and H1;
% p - a priori probability of H0;

m=10000;
n=5;

a=0.05;

LamL=0.0001;
LamU=1000;

miu0=1;
miu1=3;

sig=1;
sig1=30;
```

```
p=0.01;

for i1=1:30 % generating random numbers for
improving generator's quality
cn(i1)=randn; % normally distributed random
numbers

end

ind=0;

% computation of the integrals in restriction

l=0;

while abs(l-(1-a))>0.00001 % cycle for solution of
restriction equation, i.e. finding Lam

Lam=(LamL+LamU)/2;

l0=0;
l1=0;
lb0=0;
lb1=0;
lp01=0;
lp02=0;
lp11=0;
lp12=0;

ind=ind+1;

for j=1:m

xe0=0;
xe1=0;

for i=1:n
x0(i)=miu0+sqrt(sig)*randn;
x1(i)=miu1+sqrt(sig1+sig/n)*randn;
xe0=xe0+x0(i);
xe1=xe1+x1(i);
end

xe0=xe0/n; % mathematical expectation of normal
vector at H0;
xe1=xe1/n; % mathematical expectation of normal
vector at H1;

% computation for acceptance region of H0;

cc=((1-p)/p)*((sig/n)/(sig1+sig/n))^0.5;

d0=(xe0-miu0)^2/(2*sig/n)-(xe0-
miu1)^2/(2*sig1+2*sig/n);

if d0>100
```

```

    d0=100;
end
if d0<-100
    d0=-100;
end
g0=cc*exp(d0);
% computation for acceptance region of H1;
d1=(xe1-miu0)^2/(2*sig/n)-(xe1-
miu1)^2/(2*sig1+2*sig/n);
if d1>100
    d1=100;
end
if d1<-100
    d1=-100;
end
g1=cc*exp(d1);
% computation of integrals in restriction;
if g0<Lam
    I0=I0+1; % I0 and I1 integrals in restriction for H0
    and H1
end
if g1>1/Lam
    I1=I1+1;
end
% Computation of error probabilities
if g1<Lam
    lb0=lb0+1;
end
if g0>1/Lam
    lb1=lb1+1;
end
if 1/Lam<=g0&g0<=Lam
    lp01=lp01+1;
end
if Lam<=g0&g0<=1/Lam
    lp02=lp02+1;
end
if 1/Lam<=g1&g1<=Lam
    lp11=lp11+1;
end
    if Lam<=g1&g1<=1/Lam
        lp12=lp12+1;
    end
    % End of computation of error probabilities
end % for j=1:m
I=p*I0/m+(1-p)*I1/m;
if I>(1-a)
    LamU=Lam;
else
    LamL=Lam;
end
ind
end % while
    alfa0=1-I0/m;
    alfa1=1-I1/m;
    beta0=lb0/m;
    beta1=lb1/m;
    p01=lp01/m;
    p02=lp02/m;
    p11=lp11/m;
    p12=lp12/m;
    r=(1-p)*beta0+p*beta1
Lam
I
alfa0
alfa1
beta0
beta1
p01
p02
p11
p12
The code of the program for computation of Example 2
% Example for normal versus uniform
%
%
clear all
format short
% m - Number of generated random values for

```

```

computation integrals;
% n - sample size, i.e. dimension of the normal vector;
% a - restriction level alfa;
% LamL, LamU - Lower and upper values of Lambda;
% Lam - Lambda corresponding to the solution of
restriction equation;
% miu0, miu1 - mathematical expectations for H0 and
H1;
% sig, sig1 - variances for H0 and H1;
% p – a priori probability of H0;

m=10000;
n=5;

a=0.05;

LamL=0.0001;
LamU=1000;

miu0=3;
A=10;

sig=50000;

p=0.5;

for i1=1:30 % generating random numbers for
improving generator's quality
    cu(i1)=rand; % uniformly distributed random
    numbers
end

% computation of the integrals in restriction

I=0;

while abs(I-(1-a))>0.00001 % cycle for solution of
restriction equation, i.e. finding Lam

Lam=(LamL+LamU)/2;

I0=0;
I1=0;
Ib0=0;
Ib1=0;
Ip01=0;
Ip02=0;
Ip11=0;
Ip12=0;

for j=1:m

    xe0=0;
    xe1=0;

    for i=1:n

```

```

x0(i)=0;
x1(i)=0;
end

for i=1:n
x0(i)=miu0+sqrt(sig)*randn;
x1(i)=-A+2*A*rand;
xe0=xe0+x0(i);
xe1=xe1+x1(i);
end

xe0=xe0/n; % mathematical expectation of normal
vector at H0;
xe1=xe1/n; % mathematical expectation of uniform
vector at H1;

% computation for acceptance region of H0;

pp=(1-p)/p;
a1=1/(2*A);
a2=1/(n/(2*pi*sig))^0.5;
g0=pp*a1*a2*exp(n*(xe0-miu0)^2/(2*sig));

% computation for acceptance region of H1;

g1=pp*a1*a2*exp(n*(xe1-miu0)^2/(2*sig));

% computation of integrals in restriction;

if g0<Lam
    I0=I0+1; % I0 and I1 integrals in restriction for H0
    and H1
end

if g1>1/Lam
    I1=I1+1;
end

% Computation of error probabilities

if g1<Lam
    Ib0=Ib0+1;
end

if g0>1/Lam
    Ib1=Ib1+1;
end

if 1/Lam<=g0&g0<=Lam
    Ip01=Ip01+1;
end

if Lam<=g0&g0<=1/Lam
    Ip02=Ip02+1;
end

```

```
if 1/Lam<=g1&g1<=Lam
  lp11=lp11+1;
end
```

```
if Lam<=g1&g1<=1/Lam
  lp12=lp12+1;
end
```

```
% End of computation of error probabilities
```

```
end % for j=1:m
```

```
l=p*I0/m+(1-p)*l1/m;
```

```
if l>(1-a)
  LamU=Lam;
else
  LamL=Lam;
end
```

```
end % while
```

```
alfa0=1-l0/m;
alfa1=1-l1/m;
```

```
beta0=l0/m;
beta1=l1/m;
```

```
p01=lp01/m;
p02=lp02/m;
```

```
p11=lp11/m;
p12=lp12/m;
```

```
r=(1-p)*beta0+p*beta1
```

```
Lam
l
alfa0
alfa1
beta0
beta1
p01
p02
p11
p12
xe0
xe1
```

## REFERENCES

- [1] Kachiashvili KJ, Hashmi MA, Mueed A. Sensitivity Analysis of Classical and Conditional Bayesian Problems of Many Hypotheses Testing. *Communications in Statistics—Theory and Methods* 2012; 41: 591-605. <http://dx.doi.org/10.1080/03610926.2010.510255>

- [2] Kachiashvili KJ. Comparison of Some Methods of Testing Statistical Hypotheses. (Part I. Parallel Methods and Part II. Sequential Methods). *International Journal of Statistics in Medical Research* 2014; 3: 174-197. <http://dx.doi.org/10.6000/1929-6029.2014.03.02.11>
- [3] Kachiashvili KJ. Generalization of Bayesian Rule of Many Simple Hypotheses Testing. *International Journal of Information Technology and Decision Making* 2003; 2: 41-70. <http://dx.doi.org/10.1142/S0219622003000525>
- [4] Kachiashvili KJ. Investigation and Computation of Unconditional and Conditional Bayesian Problems of Hypothesis Testing. *ARPN Journal of Systems and Software* 2011; 1: 47-59.
- [5] Kachiashvili KJ. The Methods of Sequential Analysis of Bayesian Type for the Multiple Testing Problem. *Sequential Analysis* 2014; 33: 23-38. <http://dx.doi.org/10.1080/07474946.2013.843318>
- [6] Kachiashvili KJ, Hashmi MA. About Using Sequential Analysis Approach for Testing Many Hypotheses. *Bulletin of the Georgian Academy of Sciences* 2010; 4: 20-25.
- [7] Kachiashvili KJ, Hashmi MA, Mueed A. Quasi-optimal Bayesian procedures of many hypotheses testing. *Journal of Applied Statistics* 2013; 40: 103-122. <http://dx.doi.org/10.1080/02664763.2012.734797>
- [8] Kachiashvili GK, Kachiashvili KJ, Mueed A. Specific Features of Regions of Acceptance of Hypotheses in Conditional Bayesian Problems of Statistical Hypotheses Testing. *Sankhya* 2012; 74: 112-125. <http://dx.doi.org/10.1007/s13171-012-0014-8>
- [9] Kachiashvili KJ, Mueed A. Conditional Bayesian Task of Testing Many Hypotheses. *Statistics* 2013; 47: 274-293. <http://dx.doi.org/10.1080/02331888.2011.602681>
- [10] Marden JI. Hypothesis Testing: From p Values to Bayes Factors, *Am Stat Assoc* 2000; 95: 1316-1320. <http://dx.doi.org/10.2307/2669779>
- [11] Andersson S. Distributions of Maximal Invariants Using Quotient Measures. *Ann Statist* 1982; 10: 955-961. <http://dx.doi.org/10.1214/aos/1176345885>
- [12] Wijsman RA. Cross-Sections of Orbits and Their Application to Densities of Maximal Invariants, *Proceedings of the 5<sup>th</sup> Berkeley Symposium on Mathematical Statistics and Probability*, 1; 1967; pp. 389-400.
- [13] Jeffreys H. *Theory of Probability*, Oxford: Clarendon Press 1939.
- [14] Berger JO, Pericchi LR. The Intrinsic Bayes Factor for Model Selection and Prediction, *Am Stat Assoc* 1996; 91: 109-122. <http://dx.doi.org/10.1080/01621459.1996.10476668>
- [15] Kass RE, Wasserman L. The Selection of Prior Distributions by Formal Rules, *Am Stat Assoc* 1996; 91: 1343-1370. <http://dx.doi.org/10.1080/01621459.1996.10477003>
- [16] Berger JO, Brown LD, Wolpert RL. A unified conditional frequentist and Bayesian test for fixed and sequential simple hypothesis testing. *Ann Statist* 1994; 22: 1787-1907. <http://dx.doi.org/10.1214/aos/1176325757>
- [17] Berger JO, Boukai B, Wang Y. Unified Frequentist and Bayesian Testing of a Precise Hypothesis. *Statistical Science* 1997; 12: 133-148.
- [18] Berger JO, Boukai B, Wang Y. Simultaneous Bayesian-frequentist sequential testing of nested hypotheses. *Biometrika* 1999; 86: 79-92. <http://dx.doi.org/10.1093/biomet/86.1.79>
- [19] Dass SC, Berger JO. Unified Conditional Frequentist and Bayesian Testing of Composite Hypotheses. *Scand J of Stat* 2003; 30: 193-210. <http://dx.doi.org/10.1111/1467-9469.00326>
- [20] Gomez-Villegas MA, Main P, Sanz L. A Bayesian Analysis For The Multivariate Point Null Testing Problem. *Statistics* 2009; 43: 379-391. <http://dx.doi.org/10.1080/02331880802505173>

- [21] Duchesne P, Franco Ch. Multivariate hypothesis testing using generalized and  $\{2\}$ -inverses – with applications. *Statistics* 2014; 1-22.
- [22] Bedbur S, Beutner E, Kamps U. Multivariate testing and model-checking for generalized order statistics with applications. *Statistics* 2013; 1-114.
- [23] Bernardo JM. A Bayesian analysis of classical hypothesis testing, Universidad de Valencia 1980; 605-617.
- [24] Bartlett MS. A comment on D.V. Lindley's statistical paradox. *Biometrika* 1957; 44: 533-534.  
<http://dx.doi.org/10.1093/biomet/44.3-4.533>
- [25] Demster AP. Model searching and estimation in the logic of inference. In Godambe VP, Sprott DA, editors. *Foundations of Statistical Inference*. Toronto: Holt, Rinehart and Winston 1971; pp. 56-81.
- [26] Moreno E, Giron FJ. On the frequentist and Bayesian approaches to hypothesis testing. University of Granada and University of Malaga 2006; SORT 30: 3-28.

---

Received on 09-11-2015

Accepted on 05-04-2016

Published on 16-08-2016

<http://dx.doi.org/10.6000/1929-6029.2016.05.03.1>

© 2016 K.J. Kachiashvili; Licensee Lifescience Global.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.