DOI 10.1007/s10958-015-2300-1
Journal of Mathematical Sciences, Vol. 206, No. 2, April, 2015

ON ASYMPTOTICALLY EFFICIENT STATISTICAL INFERENCE ON A SIGNAL PARAMETER

M. S. Ermakov* UDC 519

We consider the problems of confidence estimation and hypotheses testing of the parameter of a signal observed in a Gaussian white noise. For these problems, we point out lower bounds of asymptotic efficiency in the zone of moderate deviation probabilities. These lower bounds are versions of the local asymptotic minimax Hajek–Le Cam lower bound in estimation and the lower bound for the Pitman efficiency in hypotheses testing. The lower bounds are obtained both for logarithmic and sharp asymptotics of moderate deviation probabilities. Bibliography: 23 titles.

1. Introduction

In the normal approximation zone, lower bounds of asymptotically efficient statistical inference were comprehensively studied. The local asymptotic minimax Hajek–Le Cam lower bound [1–6] for estimation and the Pitman efficiency [4–7] for hypotheses testing are natural measures of efficiency in parametric statistical inference. In the large deviation zone, the Bahadur efficiencies [2, 6–11] are the most widespread measures of asymptotic efficiency of tests and estimators. The goal of this paper is to study lower bounds of asymptotic efficiency in the zone of moderate deviation probabilities for the problem of statistical inference on the value of the parameter of a signal observed in a Gaussian white noise. Thus, for this problem, we fill the gap between asymptotic efficiencies given by the normal approximation and Bahadur asymptotic efficiencies.

For statistical inference on the parameter of the distribution of an independent sample, this problem was considered in [12–16]. The goal of this paper is to obtain similar results for the problem of statistical inference on the signal parameter. Lower bounds of asymptotic efficiency are given both for logarithmic and sharp asymptotics of moderate deviation probabilities of tests and estimators. The problem of asymptotic efficiency in statistical inference on the signal parameter both for the zone of normal approximation and the zone of large deviation probabilities was studied in a large number of papers (see [2, 3, 17–20] and references therein).

The asymptotic equivalence of various statistical models and the model of a signal in a Gaussian white noise is a very popular topic of research [4, 5, 21, 22]. These results show a tight relation between these models and the model of a signal in a Gaussian white noise. From this viewpoint, the paper helps one to compare different results on moderate deviation probabilities of tests and estimators for various models.

Usually, coverage errors of confidence sets have small values. Type I error probabilities are small in hypotheses testing. These problems are prime examples of application of large and moderate deviation probabilities in statistics. In particular, lower bounds of asymptotic efficiency of estimators in moderate deviation zone admit natural interpretation as lower bounds of asymptotic efficiency in confidence estimation [13, 14, 23].

Lower bounds of asymptotic efficiency in the problem of signal detection are easily deduced from the Neyman–Pearson lemma and are given for completeness. In the case of a one-dimensional parameter, the proof of lower bounds in estimation is based on lower bounds for hypotheses testing. The proof of local asymptotic minimax lower bounds for estimation of a

*Institute of Problems of Mechanical Engineering of RAS, St.Petersburg State University, St.Petersburg, Russia, e-mail: erm2512@mail.ru.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 420, 2013, pp. 70–87. Original article submitted October 7, 2013.

multidimensional parameter is obtained by a modification of the proof of similar results for an independent sample [14].

We use the following notation. Denote by C and c positive constants. For any $x \in R^1$ denote by $[x]$ the integer part of x. For any event A denote by $\chi(A)$ the indicator of this event. The limits of integration are the same throughout the paper. For this reason, we omit the limits of integration. We write $\int$ instead of $\int\limits_0^1$. For any function $f \in L_2(0,1)$ denote

$$\|f\|^2 = \int f^2(t)\, dt.$$

Define the function of standard normal distribution,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\{-t^2/2\}\, dt, \quad x \in R^1.$$

We omit the index of the true value of parameter θ_0 and write $\mathbf{E}[\cdot] = \mathbf{E}_{\theta_0}[\cdot]$ and $\mathbf{P}(\cdot) = \mathbf{P}_{\theta_0}(\cdot)$.

2. Lower bounds of asymptotic efficiency

2.1. Lower bounds of asymptotic efficiency for logarithmic asymptotics of moderate deviation probabilities of tests and estimators. Let us observe a realization of a random process $Y_\epsilon(t), t \in (0,1), \epsilon > 0$, defined by the stochastic differential equation

$$dY_\epsilon(t) = S(t,\theta)\ dt + \epsilon dw(t). \tag{2.1}$$

Here $S \in L_2(0,1)$ is a signal and $dw(t)$ is a Gaussian white noise. Parameter θ is unknown, $\theta \in \Theta$, and Θ is an open set in R^d.

Assume that $S(t,\theta)$ is differentiable in θ in $L_2(0,1)$ at a point θ_0, i.e., there exists a function $S_\theta(t,\theta_0)$ such that

$$\int (S(t,\theta) - S(t,\theta_0) - (\theta-\theta_0)' S_\theta(t,\theta_0))^2\, dt = o(|\theta-\theta_0|^2). \tag{2.2}$$

Here $(\theta-\theta_0)'S_\theta(t,\theta_0)$ is the inner product of $\theta-\theta_0$ and $S_\theta(t,\theta_0)$.

The Fisher information matrix equals

$$I(\theta) = \int S_\theta(t,\theta) S'_\theta(t,\theta)\, dt. \tag{2.3}$$

We make the following assumption.

A1. Relation (2.2) holds at the point $\theta_0 \in \Theta$. The Fisher information matrix $I(\theta_0)$ is positive definite.

For logarithmic asymptotics, the problem on lower bounds of efficiency for large and moderate deviation probabilities of tests and estimators is usually reduced to the one-dimensional one. Thus, in this section, we assume that $d = 1$.

Consider the problem of testing the hypothesis $H_0 : \theta = \theta_0$ versus $H_\epsilon : \theta = \theta_\epsilon := \theta_0 + u_\epsilon$, where $u_\epsilon > 0$, $u_\epsilon \to 0$, and $\epsilon^{-1}u_\epsilon \to \infty$ as $\epsilon \to 0$.

For any test K_ϵ, denote by $\alpha(K_\epsilon)$ and $\beta(K_\epsilon)$, respectively, its type I and type II error probabilities.

Define the test statistics

$$T = I^{-1/2}(\theta_0) \int S_\theta(t,\theta_0)\, dY_\epsilon(t). \tag{2.4}$$

Theorem 2.1. *Assume that condition* A1 *holds. Then for any family of tests* K_ϵ *such that* $\alpha(K_\epsilon) < c < 1$ *and* $\beta(K_\epsilon) < c < 1$,

$$\limsup_{\epsilon \to 0} (\epsilon^{-1} u_\epsilon I^{1/2}(\theta_0))^{-1} (|2 \log \alpha(K_\epsilon)|^{1/2} + |2 \log \beta(K_\epsilon)|^{1/2}) \le 1. \tag{2.5}$$

The lower bound in (2.5) *is attained at the family of tests generated by the test statistics* T.

Theorem 2.2. *Assume that condition* A1 *holds. Then*

$$\lim_{\epsilon \to 0} \sup_{\theta = \theta_0, \theta_0 + 2u_\epsilon} \epsilon^2 u_\epsilon^{-2} I^{-1}(\theta_0) \log \mathbf{P}_\theta(|\widehat{\theta}_\epsilon - \theta| \ge u_\epsilon) \ge -\frac{1}{2} \tag{2.6}$$

for any estimator $\widehat{\theta}_\epsilon$.

2.2. Lower bounds of efficiency for sharp asymptotics of moderate deviation probabilities of tests and estimators. The case of one-dimensional parameter. Fix λ, $0 < \lambda \le 1$.

The results are proved in the zone $u_\epsilon = o(\epsilon^{\frac{2}{2+\lambda}})$ under the following additional assumption.

A2. The following relations hold:

$$\int (S(t,\theta) - S(t,\theta_0) - (\theta - \theta_0)' S_\theta(t,\theta_0))^2 \, dt = O(|\theta - \theta_0|^{2+\lambda}) \tag{2.7}$$

and

$$\int (S(t,\theta) - S(t,\theta_0))^2 dt - (\theta - \theta_0)' I(\theta_0)(\theta - \theta_0) = O(|\theta - \theta_0|^{2+\lambda}). \tag{2.8}$$

In the case of a multidimensional parameter, the lower bound in hypotheses testing essentially depends on the geometry of sets of hypotheses and alternatives. We only consider the case of a one-dimensional parameter. Usually, in this case, the problem is reduced to the problem of testing a simple hypothesis versus a simple alternative. Consider the problem of testing the hypothesis $H_0 : \theta = \theta_0$ versus alternatives $H_\epsilon : \theta = \theta_\epsilon := \theta_0 + u_\epsilon$. We additionally assume that $\epsilon^{-2} u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$.

Theorem 2.3. *Assume that conditions* A1 *and* A2 *hold. Let* $\epsilon^{-1} u_\epsilon \to \infty$ *and* $\epsilon^{-2} u_\epsilon^{2+\lambda} \to 0$ *as* $\epsilon \to 0$. *For any family of tests* K_ϵ *such that* $\alpha_\epsilon := \alpha(K_\epsilon) < c < 1$,

$$\beta(K_\epsilon) \ge \Phi(x_{\alpha_\epsilon} - \epsilon^{-1} u_\epsilon I^{1/2}(\theta_0))(1 + o(1)), \tag{2.9}$$

where x_{α_ϵ} *is determined by the equation* $\alpha_\epsilon = \Phi(x_{\alpha_\epsilon})$.

The lower bound (2.9) *is attained at the tests* L_ϵ *generated by the test statistic* T.

If equality is attained in (2.9), *then*

$$\lim_{\epsilon \to 0} \alpha_\epsilon^{-1} \mathbf{E}_{\theta_0}[|K_\epsilon - L_\epsilon|] = 0 \tag{2.10}$$

and

$$\lim_{\epsilon \to 0} (\Phi(x_{\alpha_\epsilon} - \epsilon^{-1} u_\epsilon I^{1/2}(\theta_0)))^{-1} \mathbf{E}_{\theta_\epsilon}[|K_\epsilon - L_\epsilon|] = 0 \tag{2.11}$$

for the family of tests L_ϵ *with* $\alpha_\epsilon = \alpha(L_\epsilon)$.

Remark. For $u_\epsilon = \epsilon u, u > 0$, the lower bound (2.9) becomes the lower bound for the Pitman efficiency.

Define the statistic

$$T_0 = I^{-1/2}(\theta_0) \int S_\theta(t,\theta_0) \, dw(t).$$

Theorem 2.4. *Let $d = 1$ and let conditions* A1 *and* A2 *be satisfied. Let $\epsilon^{-1}u_\epsilon \to \infty$ and $\epsilon^{-2}u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$. Then, for any estimator $\widehat{\theta}_\epsilon$,*

$$\liminf_{\epsilon\to 0} \sup_{|\theta-\theta_0|<C_\epsilon u_\epsilon} \frac{\mathbf{P}_\theta(|\widehat{\theta}_\epsilon - \theta| > u_\epsilon)}{2\Phi(-\epsilon^{-1}I^{1/2}(\theta_0)u_\epsilon)} \ge 1 \tag{2.12}$$

for any family of constants $C_\epsilon \to \infty$ as $\epsilon \to 0$.

If equality is attained in (2.12) *for $C_\epsilon \to \infty$ and $\epsilon^{-2}C_\epsilon^{2+\lambda}u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$, then*

$$\lim_{\epsilon\to 0}\Big(\Phi(-\epsilon^{-1}I^{1/2}(\theta_0)u_\epsilon)\Big)^{-1}\mathbf{E}_{\theta_\epsilon}\Big[|\chi(|\widehat{\theta}_\epsilon-\theta_\epsilon|>u_\epsilon)-\chi(|I^{-1/2}(\theta_0)T_0-(\theta_\epsilon-\theta_0)|>u_\epsilon)|\Big]=0 \tag{2.13}$$

for any family of parameters $\theta_\epsilon, |\theta_\epsilon - \theta_0| < C_\epsilon u_\epsilon$.

Theorems 2.1, 2.2, 2.3, and 2.4 are versions of Theorems 2.2, 2.5, 2.3, and 2.7 established in [13] for problems of statistical inference on a parameter of the distribution of an independent sample.

2.3. Lower bound of efficiency for the sharp asymptotic of confidence estimation of a multidimensional parameter. For a multidimensional parameter, we can derive a version of Theorem 2.4 under some additional assumptions.

We say that a set $\Omega \subset R^d$ is centrally symmetric if $x \in \Omega$ implies that $-x \in \Omega$. Denote by $\partial\Omega$ the boundary of Ω.

We make the following assumptions.

A3. For any $v \in R^d$,

$$v'I(\theta)v - v'I(\theta_0)v = O(|v|^2|\theta - \theta_0|^\lambda). \tag{2.14}$$

A4. The set Ω is bounded, convex, and centrally symmetric. The boundary $\partial\Omega$ is a C^2-manifold. The principal curvatures at each point of $\partial\Omega$ are negative.

Theorem 2.5. *Assume the conditions* A1–A3 *hold for all $\theta_0 \in \Theta$. Let the set Ω satisfy condition* A4*. Let Θ_0 be a bounded open set such that $\partial\Theta_0 \subset \Theta$. Let $\epsilon^{-1}u_\epsilon \to \infty$ and $\epsilon^{-2}u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$. Then*

$$\liminf_{\epsilon\to 0} \inf_{\theta_0\in\Theta_0} \sup_{|\theta-\theta_0|<C_\epsilon u_\epsilon} \frac{\mathbf{P}_\theta(I^{1/2}(\theta_0)(\widehat{\theta}_\epsilon - \theta) \notin u_\epsilon\Omega)}{\mathbf{P}(\zeta \notin \epsilon^{-1}u_\epsilon\Omega)} \ge 1 \tag{2.15}$$

for any estimator $\widehat{\theta}_\epsilon$ with $C_\epsilon \to \infty$ as $\epsilon \to 0$. Here ζ is a Gaussian random vector with identity covariance matrix and $\mathbf{E}[\zeta] = 0$.

Theorems 2.4 and 2.5 can be considered as lower bounds of asymptotically efficient confidence estimation. In confidence estimation, the covariance matrices of estimators are often unknown. Then the confidence sets are defined on the base of pivotal statistics. Pivotal statistics are widely applied in hypotheses testing as well. For such a setup, one can modify Theorems 2.4 and 2.5. The general approach to such a setup is given in Theorem 2.2 of [14]. We do not discuss this problem in detail in order not to overload the paper.

3. Proofs of Theorems 2.1, 2.2, 2.3, and 2.4

Proofs of Theorems 2.1 and 2.3 are based on straightforward application of the Neyman–Pearson lemma and analysis of the asymptotic distribution of the logarithm of the likelihood ratio. Theorems 2.2 and 2.4 are deduced from Theorems 2.1 and 2.3, respectively, using the same reasoning as in [13, Theorems 2.3 and 2.7]. In particular, (2.13) follows from (2.10) and (2.11). For this reason, proofs of Theorems 2.2 and 2.4 are omitted. In [13], similar results were obtained for the problem of statistical inference on a parameter of the distribution of an independent sample.

We prove Theorem 2.3. The proof of Theorem 2.1 is similar. It only suffices to replace $O(|u_\epsilon|^{2+\lambda})$ by $o(|u_\epsilon|^2)$ in all the estimates. The correctness of such a replacement follows from Assumption A1.

Assume that the hypothesis is valid. Then the logarithm of the likelihood ratio equals (see [2,3])

$$\begin{aligned} L(\theta_0 + u_\epsilon, \theta_0) &:= \epsilon^{-2}\int (S(t,\theta_\epsilon) - S(t,\theta_0))\, dY_\epsilon(t) - (2\epsilon^2)^{-1}(\|S(t,\theta_\epsilon)\|^2 - \|S(t,\theta_0)\|^2) \\ &= \epsilon^{-1}\int (S(t,\theta_\epsilon) - S(t,\theta_0)) dw(t) - (2\epsilon^2)^{-1}\|S(t,\theta_\epsilon) - S(t,\theta_0)\|^2. \end{aligned} \tag{3.1}$$

Therefore, the test statistics can be defined as follows:

$$T_1 = \xi(\theta_\epsilon, \theta_0) = \epsilon^{-1}\int (S(t,\theta_\epsilon) - S(t,\theta_0))\, dY_\epsilon(t). \tag{3.2}$$

By condition A2, to prove the asymptotic efficiency of the tests L_ϵ, it suffices to estimate the difference of stochastic parts of T_1 and $(\theta_\epsilon - \theta_0)I^{1/2}(\theta_0)T$ defined by the statistics

$$T_{1\epsilon} = \epsilon^{-1}\int (S(t,\theta_\epsilon) - S(t,\theta_0))\, dw(t)$$

and $\epsilon^{-1}I^{1/2}(\theta_0)(\theta_\epsilon - \theta_0)T_0$, respectively.

Denote

$$\rho^2(\theta_\epsilon, \theta_0) = \|S(t,\theta_\epsilon) - S(t,\theta_0)\|^2.$$

Straightforward calculations show that

$$\mathbf{E}_{\theta_0}[T_{1\epsilon}] = 0, \tag{3.3}$$

and it follows from (2.8) that

$$\mathbf{E}_{\theta_0}[T_{1\epsilon}^2] = \epsilon^{-2}\rho^2(\theta_\epsilon, \theta_0) = \epsilon^{-2}u_\epsilon^2 I(\theta_0) + O(\epsilon^{-2}u_\epsilon^{2+\lambda}). \tag{3.4}$$

For the alternative, we get the relations

$$\begin{aligned} \mathbf{E}_{\theta_\epsilon}[T_{1\epsilon}] &= \epsilon^{-1}\mathbf{E}_{\theta_0}[\xi(\theta_\epsilon,\theta_0)\exp\{\epsilon^{-1}\xi(\theta_\epsilon,\theta_0) - (2\epsilon^2)^{-1}\rho^2(\theta_\epsilon,\theta_0)\}] \\ &= \epsilon^{-2}\rho^2(\theta_\epsilon,\theta_0) = \epsilon^{-2}(u_\epsilon^2 I(\theta_0) + O(u_\epsilon^{2+\lambda})) \end{aligned} \tag{3.5}$$

and

$$\mathbf{Var}_{\theta_\epsilon}[T_{1\epsilon}^2] = \epsilon^{-2}\rho^2(\theta_\epsilon,\theta_0) = \epsilon^{-2}u_\epsilon^2 I(\theta_0) + O(u_\epsilon^{2+\lambda}). \tag{3.6}$$

The lower bound (2.9) follows from (3.1)-(3.6)

The proof of asymptotic efficiency of the test statistics T is based on the following lemma.

Lemma 3.1. *Let* $\vec{\eta}_\epsilon = (\eta_{1\epsilon}, \eta_{2\epsilon})'$ *be Gaussian random vectors such that* $\mathbf{E}[\eta_{1\epsilon}] = 0$, $\mathbf{E}[\eta_{2\epsilon}] = 0$, $\mathbf{E}[\xi_{1\epsilon}^2] = 1$, $\mathbf{E}[\xi_{2\epsilon}^2] = O(|u_\epsilon|^\lambda)$, *and* $\mathbf{E}[\eta_{1\epsilon}\eta_{2\epsilon}] = O(|u_\epsilon|^\lambda)$. *Then*

$$\mathbf{P}(\eta_{1\epsilon} > \epsilon^{-1}u_\epsilon) = \mathbf{P}(\eta_{1\epsilon} + \eta_{2\epsilon} > \epsilon^{-1}u_\epsilon)(1 + o(1)). \tag{3.7}$$

Proof. Denote by A_ϵ the covariance matrix of the random vector $\vec{\eta}_\epsilon$. Let ζ_1 and ζ_2 be independent random variables having the standard normal distribution. Define the random vector $\vec{\zeta} = (\zeta_1, \zeta_2)$. Denote $\vec{\omega}_\epsilon = (\omega_{1\epsilon}, \omega_{2\epsilon})' = A_\epsilon^{1/2}\vec{\zeta}$. Then

$$\mathbf{P}(\eta_{1\epsilon} + \eta_{2\epsilon} > \epsilon^{-1}u_\epsilon) = \mathbf{P}(\omega_{1\epsilon} + \omega_{2\epsilon} > \epsilon^{-1}u_\epsilon). \tag{3.8}$$

A straightforward calculation using the identity $A_\epsilon^{1/2}A_\epsilon^{1/2} = A_\epsilon$ shows that entries of the matrix $A_\epsilon^{1/2} = \{a_{\epsilon,ij}\}_{i,j=1}^2$ have the following orders: $a_{\epsilon,22} = O(|u_\epsilon|^{\lambda/2})$ and $a_{\epsilon,12} = O(|u_\epsilon|^{\lambda/2})$. Hence, using (3.8), we get (3.7).

Thus, it remains to verify that the normalized random variables

$$\eta_{1\epsilon} = u_\epsilon^{-1}\xi(\theta_\epsilon, \theta_0) \quad \text{and} \quad \eta_{2\epsilon} = u_\epsilon^{-1}(\xi(\theta_\epsilon, \theta_0) - u_\epsilon \tau)$$

satisfy the assumptions of Lemma 3.1 in the case of the hypothesis and alternative. Here

$$\tau = \tau_{\theta_0} = \int S_\theta(t, \theta_0)\, dw(t).$$

Assume that the hypothesis is valid.

By (3.4),

$$\mathbf{E}\eta_{1\epsilon}^2 = I + O(|u_\epsilon|^\lambda). \tag{3.9}$$

We note that

$$\mathbf{E}[\eta_{1\epsilon}\eta_{2\epsilon}] = u_\epsilon^{-2}\rho^2(\theta_\epsilon, \theta_0) - u_\epsilon^{-1}\int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt. \tag{3.10}$$

By (2.7),

$$\begin{aligned} O(u_\epsilon^{2+\lambda}) &= \|S(t, \theta_\epsilon) - S(t, \theta_0) - u_\epsilon S_\theta(t, \theta_0)\|^2 \\ &= \rho^2(\theta_\epsilon, \theta_0) - 2u_\epsilon \int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt + u_\epsilon^2 I(\theta_0). \end{aligned} \tag{3.11}$$

Hence, by (2.8),

$$u_\epsilon \int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt = u_\epsilon^2 I(\theta_0) + O(|u_\epsilon|^{2+\lambda}). \tag{3.12}$$

It follows from (2.8), (3.10), and (3.12) that

$$\mathbf{E}[\eta_{1\epsilon}\eta_{2\epsilon}] = O(|u_\epsilon|^\lambda). \tag{3.13}$$

By (3.9) and (3.13), the assumptions of Lemma 3.1 are satisfied if the hypothesis is valid.

Assume that the alternative is valid. Straightforward calculations using (3.12) show that

$$\begin{aligned} \mathbf{E}_{\theta_\epsilon}[\tau] &= \mathbf{E}_{\theta_0}[\tau \exp\{\epsilon^{-1}\xi(\theta_\epsilon, \theta_0) - (2\epsilon^2)^{-1}\rho^2(\theta_\epsilon, \theta_0)\}] \\ &= \epsilon^{-1}\int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt = \epsilon^{-1}u_\epsilon I(\theta_0) + O(\epsilon^{-1}|u_\epsilon|^{1+\lambda}), \end{aligned} \tag{3.14}$$

and, arguing similarly, we conclude that

$$\mathbf{E}_{\theta_\epsilon}[\xi(\theta_0, \theta_\epsilon)] = \epsilon^{-1}\rho^2(\theta_\epsilon, \theta_0) = \epsilon^{-1}u_\epsilon^2 I(\theta_0) + O(\epsilon^{-1}|u_\epsilon|^{2+\lambda}). \tag{3.15}$$

The same reasoning shows that

$$\begin{aligned} \mathbf{E}_{\theta_\epsilon}[\xi^2(\theta_\epsilon, \theta_0)] &= \rho^2(\theta_\epsilon, \theta_0) + \epsilon^{-2}\rho^4(\theta_\epsilon, \theta_0) \\ &= u_\epsilon^2 I(\theta_0) + \epsilon^{-2}u_\epsilon^4 I^2(\theta_0) + O(|u_\epsilon|^{2+\lambda} + \epsilon^{-2}|u_\epsilon|^{4+\lambda}), \end{aligned} \tag{3.16}$$

$$\begin{aligned} u_\epsilon \mathbf{E}_{\theta_\epsilon}[\xi(\theta_\epsilon, \theta_0)\tau] &= u_\epsilon \int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt(1 + \epsilon^{-2}\rho^2(\theta_\epsilon, \theta_0)) \\ &= u_\epsilon^2 I(\theta_0) + \epsilon^{-2}u_\epsilon^4 I^2(\theta_0) + O(|u_\epsilon|^{2+\lambda} + \epsilon^{-2}|u_\epsilon|^{4+\lambda}), \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} u_\epsilon^2 \mathbf{E}_{\theta_\epsilon}[\tau^2] &= u_\epsilon^2 I(\theta_0) + \epsilon^{-2}u_\epsilon^2 \left(\int (S(t, \theta_\epsilon) - S(t, \theta_0))S_\theta(t, \theta_0)\, dt\right)^2 \\ &= u_\epsilon^2 I(\theta_0) + \epsilon^{-2}u_\epsilon^4 I^2(\theta_0) + O(|u_\epsilon|^{2+\lambda} + \epsilon^{-2}|u_\epsilon|^{4+\lambda}). \end{aligned} \tag{3.18}$$

By (3.16)–(3.18),

$$\mathbf{E}_{\theta_\epsilon}[\eta_{2\epsilon}^2] = O(|u_\epsilon|^\lambda + \epsilon^{-2}|u_\epsilon|^{2+\lambda})$$

and

$$\mathbf{E}_{\theta_\epsilon}[\eta_{1\epsilon}\eta_{2\epsilon}] = O(|u_\epsilon|^\lambda + \epsilon^{-2}|u_\epsilon|^{2+\lambda}).$$

This implies that the assumptions of Lemma 3.1 are satisfied in the case of the alternative. □

Our proof of Theorem 2.4 is based on the following version of Theorem 2.3. In this version, we treat the problem of testing the hypothesis $H_0 : \theta = \theta_0 + C_1 u_\epsilon$ versus the alternatives $H_{1\epsilon} : \theta = \theta_0 + C_2 u_\epsilon$.

Lemma 3.2. *Assume that conditions* A1 *and* A2 *hold. Then for any family of tests K_ϵ such that $\alpha_\epsilon := \alpha(K_\epsilon) < c < 1$,*

$$\beta(K_\epsilon) \ge \Phi(x_{\alpha_\epsilon} - \epsilon^{-1}(C_2 - C_1)u_\epsilon I^{1/2}(\theta_0))(1 + o(1)), \tag{3.19}$$

where x_{α_ϵ} is determined by the equation $\alpha_\epsilon = \Phi(x_{\alpha_\epsilon})$.

The lower bound (3.19) *is attained at the tests L_ϵ generated by the tests statistics T.*

If equality is attained in (3.19), *then*

$$\lim_{\epsilon\to 0} \alpha_\epsilon^{-1}\mathbf{E}_{\theta_0}[|K_\epsilon - L_\epsilon|] = 0 \tag{3.20}$$

and

$$\lim_{\epsilon\to 0}(\Phi(x_{\alpha_\epsilon} - \epsilon^{-1}(C_2 - C_1)u_\epsilon I^{1/2}(\theta_0)))^{-1}\mathbf{E}_{\theta_\epsilon}[|K_\epsilon - L_\epsilon|] = 0 \tag{3.21}$$

for any family of tests L_ϵ such that $\alpha_\epsilon := \alpha(L_\epsilon)$.

The remaining reasoning in the proof of Theorem 2.4 is identical to the proof of Theorem 2.7 of [13] and is omitted.

4. Proof of Theorem 2.5

In Theorem 2.1 of [14], a version of Theorem 2.5 has been proved for confidence estimation of a parameter of the distribution of an independent sample. The proof of Theorem 2.5 is a revised version of the proof of this theorem.

In what follows, we assume that $\theta_0 = 0$.

We divide the proof into the following steps.

1. The Bayes approach. We refer to the fact that the Bayes risk does not exceed the minimax one and reduce the problem to the problem of calculation of asymptotics of Bayes risks. We define a uniform Bayes a priori distribution on the lattice Λ_ϵ in the cube $K_{v_\epsilon} = (-v_\epsilon, v_\epsilon)^d$, where $v_\epsilon = C_\epsilon u_\epsilon$, $C_\epsilon \to \infty$, and $\epsilon^{-2}(C_\epsilon u_\epsilon)^{2+\lambda} \to 0$. The lattice spacing equals $\delta_{1\epsilon} = c_{1\epsilon}\epsilon^2 u_\epsilon^{-1}$ with $c_{1\epsilon} \to 0$ and $c_{1\epsilon}^{-3}\epsilon^{-2}u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$. Denote $l_\epsilon = [v_\epsilon/\delta_{1\epsilon}]$.

2. We split the cube K_{v_ϵ} into small cubes,

$$\Gamma_{i\epsilon} = x_{\epsilon i} + (-c_{2\epsilon}\epsilon^2 u_\epsilon^{-1}, c_{2\epsilon}\epsilon^2 u_\epsilon^{-1}]^d, \quad 1 \le i \le m_n,$$

with $c_{2\epsilon} \to 0$, $c_{2\epsilon}c_{1\epsilon}^{-1} \to \infty$, and $c_{1\epsilon}^{-3}\epsilon^{-2}b_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$.

Using the fact that a normalized a posteriori Bayes risk tends to a constant in probability as $\epsilon \to 0$, we study the asymptotics of a posteriori Bayes risks independently for each event $W_{i\epsilon} : \tau \in \epsilon^{-1}\Gamma_{\epsilon i}$.

3. To narrow down the set of parameters for a posteriori Bayes risk minimization, we split the lattice Λ_ϵ into subsets $\Lambda_{ie}, 1 \le i \le m_{2i\epsilon}$. Each set Λ_{ie} is a lattice in the union of a finite

number of very narrow parallelepipeds K_{ije}. The problem of minimization of a posteriori Bayes risk is solved independently for each set Λ_{ie}, and the results are added:

$$\begin{aligned}\inf_{\widehat{\theta}_\epsilon}\sup_{\theta\in K_{v_\epsilon}}\mathbf{P}_\theta(\widehat{\theta}_\epsilon-\theta\notin u_\epsilon\Omega) &\geq \inf_{\widehat{\theta}_\epsilon}(2l_\epsilon)^{-d}\sum_{i=1}^{m_n}\sum_{\theta\in\Lambda_\epsilon}\mathbf{P}_\theta(\widehat{\theta}_\epsilon-\theta\notin u_\epsilon\Omega, W_{i\epsilon})\\ &\geq (2l_\epsilon)^{-d}\sum_{i=1}^{m_\epsilon}\sum_{e=1}^{m_{2i\epsilon}}\inf_{\widehat{\theta}_\epsilon}\sum_{\theta\in\Lambda_{ie}}\mathbf{P}_\theta(\widehat{\theta}_\epsilon-\theta\notin u_\epsilon\Omega, W_{i\epsilon}).\end{aligned} \tag{4.1}$$

4. To estimate the accuracy of a linear approximation of the stochastic part of the logarithm of the likelihood ratio, we prove the following inequalities (see also Lemma 4.2 and, for comparison, (3.4) and Lemma 5.3 of [14]). For any $\theta_j, \theta_k \in \Lambda_\epsilon \cap K_{ije}$ and $\kappa > 0$,

$$\begin{aligned}&\mathbf{P}(\epsilon^{-1}|\xi(\theta_j,\theta_k)-(\theta_k-\theta_j)'\tau_{\theta_j}-\rho_\epsilon'\tau_{\theta_j}|>\kappa, W_{i\epsilon})\\ &\qquad\leq C\int\limits_{\Gamma_{\epsilon i}}\exp\left\{-\frac{|t|^2}{2\epsilon^2\|S_\theta(t,0)\|^2}\right\}dt\exp\left\{-c\kappa^2|\theta_k-\theta_j|^{-2-\lambda}\epsilon^{-2}\right\},\end{aligned} \tag{4.2}$$

where

$$\rho_\epsilon = \rho_\epsilon(\theta_j,\theta_k) = \epsilon^2\|S_\theta(t,\theta_j)\|^{-2}\int S_\theta(t,\theta_j)(S(t,\theta_k)-S(t,\theta_j)-(\theta_k-\theta_j)'S_\theta(t,\theta_j))\ dt.$$

Since $\tau\in\epsilon^{-1}\Gamma_{i\epsilon}$, it is easy to show that

$$\rho_\epsilon'\int S_\theta(t,0)dw(t) < \delta_\epsilon \to 0 \tag{4.3}$$

as $\epsilon\to 0$.

5. Estimates similar to (4.2) and (4.3) and the "chaining" method allow us to apply to terms of the right-hand side of (4.1) the technique of the proof of the multidimensional local asymptotic minimax lower bound [2] based on the same reasoning as in [14].

For clarity, we define parallelepipeds K_{ij} in the case where $x_{i\epsilon}$ is parallel to the first ort e_1 of the coordinate system. Consider the subspace Π_1 orthogonal to e_1. Define in the lattice $\Lambda_\epsilon\cap\Pi_1$ a sublattice $\Lambda_i^1=\{\theta_{ij}\}_{1\leq j\leq m_{1i\epsilon}}$ with spacing $2c_{3\epsilon}\delta_{1\epsilon}$, where $c_{3\epsilon}$ is such that $c_{3\epsilon}/c_{2\epsilon}\to\infty$, $c_{3\epsilon}\delta_{1\epsilon}=o(\epsilon^2u_\epsilon^{-1})$, and $c_{3\epsilon}^3c_{1\epsilon}^{-3}\epsilon^{-2}u_\epsilon^{2+\lambda}\to 0$ as $\epsilon\to 0$.

Set

$$K_{ij}=K(\theta_{ij})=\left\{x:x=\lambda x_{i\epsilon}+u+\theta_{ij},\ u=\{u_k\}_{k=1}^d,\ u\perp x_{ni}, |u_k|\leq c_{3n}\delta_{1n},\ \lambda\in R^1,\ u\in R^d\right\}\cap K_{v_\epsilon},\quad 1\leq j\leq m_{1i\epsilon}.$$

We define the sets Λ_{ie} for the most simple geometry in which the distance from the set $\partial\Omega$ to zero is attained only at two points. Each set Λ_{ie} consists of subsets $K(\theta_{ij})\cap\Lambda_\epsilon$ such that

$$\begin{aligned}\theta_{ij}\in\Theta_{ie}=\Theta_i(k_1,\dots,k_{d-d_1}) = \{\theta:\theta=\theta_{ij}&+(-1)^{t_2}2k_2c_{3\epsilon}\delta_{1\epsilon}e_2\\ &+\dots+(-1)^{t_d}2k_dc_{3\epsilon}\delta_{1\epsilon}e_d;\ t_2,\dots t_d=0,1\},\end{aligned}$$

where $k_2,\dots,k_d$ are fixed for each Λ_{ie} and $0\leq k_2,\dots,k_d < C_{1\epsilon}$ with $C_{1\epsilon}c_{3\epsilon}c_{1\epsilon}\to\infty$ and $\epsilon^{-2}C_{1\epsilon}^3c_{3\epsilon}^3c_{1\epsilon}^3u_\epsilon^{2+\lambda}\to 0$ as $\epsilon\to 0$.

Denote $K_{ie}=\bigcup\limits_{\theta\in\Theta_{ie}}K(\theta)$.

For an arbitrary geometry of $\partial\Omega$, the definition of the sets Λ_{ie} is more complicated, and, moreover, the indexation becomes cumbersome (see [14]). However, the reasoning remains almost unchanged.

Fix $\delta > 0$. For all $\theta \in \Lambda_{ie}$ define the events $A_i(0,\theta,\delta) : \epsilon^{-1}(\xi(0,\theta) - \theta'\tau) > \delta$. Denote $A_{ie} = \bigcap_{\theta\in\Theta_{ie}} A_i(0,\theta,\delta)$. Denote by B_{ie} the event additional to A_{ie}.

Arguing similarly to (3.8) and (3.9) in [14], we conclude that

$$
\begin{aligned}
&\inf_{\widehat{\theta}_\epsilon} \sum_{\theta\in\Lambda_{ie}} \mathbf{P}_\theta(\widehat{\theta}_\epsilon - \theta \notin u_\epsilon\Omega, W_{i\epsilon}) \\
&\quad \ge \inf_{\widehat{\theta}_\epsilon} \sum_{\theta\in\Lambda_{ie}} \mathbf{E}\Big[\chi(\widehat{\theta}_\epsilon - \theta \notin u_\epsilon\Omega) \exp\{\epsilon^{-1}\tau - (2\epsilon^2)^{-1}\rho^2(\theta_0,\theta_\epsilon)\}, \quad W_{i\epsilon}, A_i(\theta,0,\kappa)\Big] \\
&\quad \ge \mathbf{E}\Bigg[\inf_t \sum_{\theta\in\Lambda_{ie}} \chi(t-\theta\notin u_\epsilon\Omega) \exp\left\{\epsilon^{-1}\theta'\tau - (2\epsilon^2)^{-1}\theta' I\theta + o(1)\right\}, \quad W_{i\epsilon}, A_{ie}\Bigg] = R_\epsilon
\end{aligned}
\tag{4.4}
$$

if $\delta = \delta_\epsilon$ tends to zero sufficiently slowly as $\epsilon \to 0$.

Denote $\Delta_\epsilon = \exp\{\tau'\tau/2\}$ and $y = y_\theta = \epsilon^{-1}\theta - \tau$. Since $\epsilon^{-2}u_\epsilon\delta_{1\epsilon} \to 0$ and $\epsilon^{-2}u_\epsilon^{2+\lambda} \to 0$ as $\epsilon \to 0$, we see that

$$
\begin{aligned}
(2l_\epsilon)^{-d}R_\epsilon &\ge (2l_\epsilon)^{-d}\mathbf{E}\Bigg[\Delta_\epsilon \inf_t \sum_{\theta\in\Lambda_{ie}} \chi(t-y_\theta-\tau\notin\epsilon^{-1}u_\epsilon\Omega)\exp\left\{-\frac{1}{2}y_\theta' I y_\theta\right\}, W_{i\epsilon}, A_{ie}\Bigg](1+o(1)) \\
&= (2v_\epsilon)^{-d}\mathbf{E}\Bigg[\Delta_\epsilon \inf_t \int\limits_{\epsilon^{-1}K_{ie}-\psi_\epsilon} \chi(t-y\notin\epsilon^{-1}u_\epsilon\Omega)\exp\left\{-\frac{1}{2}y' I y\right\} dy, W_{i\epsilon}, A_{ie}\Bigg](1+o(1)) \\
&:= (2v_\epsilon)^{-d} I_{ie\epsilon}(1+o(1)).
\end{aligned}
\tag{4.5}
$$

For $\kappa \in (0,1)$ denote

$$
K_{i\kappa}(\theta_{ij}) = \left\{x : x = \lambda x_{i\epsilon} + u + \theta_{ij},\ u = \{u_k\}_1^d, |u_k| \le (c_{3\epsilon} - Cc_{2\epsilon})\delta_{1\epsilon},\ u \perp x_{i\epsilon}, \lambda \in R^1\right\} \cap K_{(1-\kappa)v_\epsilon}
$$

and

$$
K_{ie\kappa} = \bigcup_{\theta\in\Theta_{ie}} K_{i\kappa}(\theta).
$$

Here $u \perp x_{i\epsilon}$ indicates that the vectors u and $x_{i\epsilon}$ are orthogonal.

If $\tau \in \epsilon^{-1}\Gamma_{i\epsilon}$, then $\epsilon^{-1}K_{ie\kappa} \subset \epsilon^{-1}K_{ie} - \tau$, and, therefore,

$$
I_{ie\epsilon} \ge U_{ie\epsilon}\overline{J}_{ie\epsilon}(1+o(1)), \tag{4.6}
$$

where

$$
U_{ie\epsilon} = \mathbf{E}\left[\Delta_\epsilon, W_{i\epsilon}, A_{ie}\right]
$$

and

$$
\overline{J}_{ie\epsilon} := \inf_t J_{ie\epsilon}(t) := \inf_t \int\limits_{\epsilon^{-1}K_{ie\kappa}} \chi(t-y\notin\epsilon^{-1}u_\epsilon\Omega)\exp\left\{-\frac{1}{2}y' I y\right\} dy.
$$

By Lemma 3.1 of [14],

$$
\overline{J}_{ie\epsilon} = J_{ie\epsilon}(0). \tag{4.7}
$$

We note that

$$
\mathbf{E}\left[\Delta_\epsilon, W_{i\epsilon}\right] = \mathrm{mes}(\Gamma_{i\epsilon})(1+o(1)). \tag{4.8}
$$

Thus, to prove Theorem 2.5, it only remains to prove that

$$
U_{2ie\epsilon} := \mathbf{E}\left[\Delta_\epsilon, W_{i\epsilon}, B_{ie}\right] = \exp\{\epsilon^{-2}|x_{i\epsilon}|^2/2\}\mathbf{P}(W_{i\epsilon}, B_{ie}) = o(\mathrm{mes}(\Gamma_{i\epsilon})). \tag{4.9}
$$

Then Theorem 2.5 follows from (4.1) and (4.4)–(4.9).

Thus, it remains to estimate $\mathbf{P}(W_{i\epsilon}, B_{ie})$. For this estimation, we implement the "chaining" method.

To simplify notation, we assume that $l_\epsilon = 2^m$. Fix $\theta \in \Theta_{ie}$. Define sets Ψ_j, $j = 0, 1, 2, \dots, m$, by induction. We put $\Psi_0 = \{\theta\}$. Set

$$\Psi_j = \{\theta : \theta = \theta_{j-1} \pm v_\epsilon 2^{-j} x_{i\epsilon}, \theta_{j-1} \in \Psi_{j-1}\}, \quad 1 \le j \le m.$$

We put $\Psi_{m+1} = \Lambda_{ie\epsilon} \setminus \bigcup_{j=1}^{m} \Psi_j$. For each $\theta_j \in \Psi_j$ denote by θ_{j-1} the nearest to it point $\theta \in \Psi_{j-1}$ that lies between zero and θ_j.

We have the equalities

$$S(\theta_j, 0) = \xi(\theta_j, 0) - (\theta_j - \theta)'\tau = S_1(\theta_j, \theta_{j-1}) + S(\theta_{j-1}, 0) + S_2(\theta_j, \theta_{j-1}), \tag{4.10}$$

where

$$S_1(\theta_j, \theta_{j-1}) = \xi(\theta_j, \theta_{j-1}) - (\theta_j - \theta_{j-1})'\tau_{\theta_{j-1}} \tag{4.11}$$

and

$$S_2(\theta_j, \theta_{j-1}) = (\theta_j - \theta_{j-1})'(\tau_{\theta_{j-1}} - \tau). \tag{4.12}$$

Then

$$\mathbf{P}(W_{i\epsilon}, B_{ie}) \le C \left(\sum_{\theta_0 \in \Theta_{ie}} \left(V_0(\theta_0) + \sum_{\theta \in \Lambda_{1ile}(\theta_0)} (V_1(\theta) + V_2(\theta)) \right) \right), \tag{4.13}$$

where $\Lambda_{1ie}(\theta_0) = \Lambda_{ie}(\theta_0) \setminus \Theta_{ie}$,

$$V_0(\theta_0) = \mathbf{P}(|S(0, \theta_0) > \delta/4, W_{i\epsilon}),$$

and

$$V_s(\theta_j) = \mathbf{P}(j^2 |S_s(\theta_j, \theta_{j-1})| > \delta/4, W_{i\epsilon}), \quad s = 1, 2.$$

Lemma 4.1. *The estimate*

$$V_0(\theta) < C \exp\big\{ - C|\theta|^{-2-\lambda} \delta^2 \epsilon^{-2} \big\} \mathbf{P}(W_{i\epsilon}) \tag{4.14}$$

holds. For $\theta_j \in \Psi_j$,

$$V_s(\theta_j) < C \exp\big\{ - C|\theta_j - \theta_{j-1}|^{-2} u_\epsilon^\lambda \delta^2 j^{-4} \epsilon^{-2} \big\} \mathbf{P}(W_{i\epsilon}) \tag{4.15}$$

for $s = 1, 2$.

Substituting (4.14) and (4.15) into (4.13), we get (4.9).

The proofs of (4.14) and (4.15) are akin to the proofs of (5.6) and (5.7) in [14] and are based on the same estimates of Lemmas 5.4–5.8 of [14]. The proofs of these lemmas for the setup of this paper do not differ from those in [14]. We omit the proofs of versions of Lemmas 5.4–5.6 of [14] for our setup. We only give versions of Lemmas 5.7 and 5.8 of [14] and their proofs.

Denote $\overline{h} = \theta_j - \theta_{j-1}, h = \theta_j$, and $h_1 = \theta_{j-1}$.

Lemma 4.2. *For any $u \in R^d$,*

$$\mathbf{E}[(u'(\tau - \tau_h))^2] = O(|u|^2 |h|^\lambda). \tag{4.16}$$

Lemma 4.3. *Let $v \perp \overline{h}, v \in R^d$. Then*

$$\mathbf{E}[(\overline{h}'(\tau_{h_1} - \tau))(v'\tau)] = O(|v||\overline{h}||h_1|^{\lambda/2}). \tag{4.17}$$

If $v \parallel \overline{h}$, then

$$\mathbf{E}[(\overline{h}'(\tau_{h_1} - \tau))(v'\tau)] = O(|v||\overline{h}||h_1|^\lambda). \tag{4.18}$$

Proof of Lemma 4.2. Using (2.7), we conclude that

$$J(h,u) := \mathbf{E}[(\xi(h,h+u)-\xi(0,u))^2] \le C(\mathbf{E}[(\xi(\theta_0,h+u)-(h+u)'\tau)^2] \\ +\mathbf{E}[(\xi(\theta_0,h+u)-u'\tau)^2]+\mathbf{E}[(\xi(\theta_0,h)-h'\tau)^2]) \le C(|h+u|^{2+\lambda}+|h|^{2+\lambda}+|u|^{2+\lambda}). \tag{4.19}$$

At the same time,

$$\begin{aligned}\mathbf{E}[(u'(\tau-\tau_h))^2] &\le C(E[(\xi(h,h+u)-u'\tau_h-\xi(0,u)+u'\tau)^2]+J(h,u))) \\ &\le C(\mathbf{E}[(\xi(h,h+u)-u'\tau_h)^2]+\mathbf{E}[(\xi(0,u))^2-u'\tau)^2]+J(h,u)) \\ &\le C(|h+u|^{2+\lambda}+|h|^{2+\lambda}+|u|^{2+\lambda}).\end{aligned} \tag{4.20}$$

Putting $|u|=C|h|$, we get (4.16). □

Proof of Lemma 4.3. Applying the Cauchy inequality and Lemma 4.2, we see that

$$|\mathbf{E}[(\overline{h}'(\tau_{h_1}-\tau))(v'\tau)]| \le (\mathbf{E}[((\overline{h}'(\tau_{h_1}-\tau)))^2])^{1/2}(\mathbf{E}[(v'\tau)^2])^{1/2} = O(|v||\overline{h}||h_1|^{\lambda/2}). \tag{4.21}$$

Let us prove (4.18). Note that

$$O(|v|^2|h|^\lambda) = \mathbf{E}[(v'(\tau-\tau_h))^2] = v'I(0)v+v'I(h)v-2\mathbf{E}[(v'\tau)(v'\tau_h)]. \tag{4.22}$$

Hence, using (2.8), we get the relation

$$\mathbf{E}[(v'\tau)(v'\tau_h)] = v'I(0)v+O(|v|^2|h|^\lambda). \tag{4.23}$$

Hence,

$$\mathbf{E}[(\overline{h}'(\tau_{h_1}-\tau))(v'\tau)] = C(\mathbf{E}[(h'\tau_{h_1})(h\tau)]-\mathbf{E}[(h'\tau)(h\tau)]) = O(|v||\overline{h}||h_1|^\lambda). \tag{4.24}$$

□

This research was supported by the RFBR (projects 11-01-00577 and 11-01-00769).

Translated by M. S. Ermakov.

REFERENCES

1. J. Hajek, “Local asymptotic minimax and admissibility in estimation,” in: *Proc. Sixth Berkeley Symp. Math. Statist. Probab.,* California Univ. Press, Berkeley, **1** (1972), pp. 175–194.
2. I. A. Ibragimov and R. Z. Hasminskii, *Statistical Estimation: Asymptotic Theory,* Springer (1981).
3. Yu. A. Kutoyants, *Identification of Dynamical System With Small Noise,* Springer (1994).
4. L. Le Cam, “Limits of experiments,” in: *Proc. Sixth Berkeley Symp. Math. Statist. Probab.,* California Univ. Press, Berkeley, **1** (1972), pp. 245–261.
5. H. Strasser, *Mathematical Theory of Statistics,* W. de Gruyter, Berlin (1985).
6. A. W. van der Vaart, *Asymptotic Statistics,* Cambridge Univ. Press (1998).
7. R. E. Blahut, “Hypothesis testing and information theory,” *IEEE Trans. Inform. Theory,* **20**, 405–415 (1974).
8. R. R. Bahadur, “Asymptotic efficiency of tests and estimates,” *Sankhyā,* **22**, 229–252 (1960).
9. J. Bishwal, *Parameter Estimation of Stochastic Differential Equations,* Springer (2008).
10. T. Chiyonobu, “Hypothesis testing for signal detection problem and large deviations,” *Nagoya Math. J.,* **162**, 187–203 (2003).
11. A. Puhalskii and V. Spokoiny, “On large-deviation efficiency in statistical inference,” *Bernoulli,* **4**, 203–272 (1998).

12. A. A. Borovkov and A. A. Mogulskii, *Large Deviations and the Testing of Statistical Hypotheses* [in Russian], Tr. Inst. Mat. Sib. Otd. RAS, Novosibirsk (1992).
13. M. S. Ermakov, "Asymptotically efficient statistical inference for moderate deviation probabilities," *Teor. Veroyatn. Primen.,* **48**, 676–700 (2003).
14. M. S. Ermakov, "The sharp lower bound of asymptotic efficiency of estimators in the zone of moderate deviation probabilities," *Electronic J. Statist.,* **6**, 2150–2184 (2012).
15. M. Radavičius, "From asymptotic efficiency in minimax sense to Bahadur efficiency," in: V. Sazonov and T. Shervashidze (eds.), *New Trends Probab. Statist.,* VSP/Mokslas, Vilnius (1991), pp. 629–635.
16. W. C. M. Kallenberg, "Intermediate efficiency, theory and examples," *Ann. Statist.,* **11**, 170–182 (1983).
17. M. V. Burnashev, "On maximum likelihood estimator of signal parameter in Gaussian white noise," *Probl. Pered. Inform.,* **11**, 55–69 (1975).
18. F. Q. Gao and F. J. Zhao, "Moderate deviation and hypotheses testing for signal detection problem," *Sci. China. Math.,* **55**, 2273–2284 (2012).
19. S. Ihara and Y. Sakuma, "Signal detection in white Gaussian channel," in: *Proc. Seventh Japan–Russian Symp. Probab. Theory, Math. Stat.,* World Scientific, Singapore (1996), pp. 147–156.
20. R. S. Liptzer and A. N. Shiriaev, *Statistics of Random Processes,* Springer (2005).
21. L. D. Brown, A. V. Carter, M. G. Low, and C. H. Zhang, "Equivalence theory for density estimation, Poisson processes, and Gaussian white noise with drift," *Ann. Statist.,* **32**, 2074–2097 (2004).
22. G. K. Golubev, M. Nussbaum, and H. H. Zhou, "Asymptotic equivalence of spectral density estimation and Gaussian white noise," *Ann. Statist.,* **38**, 181–214 (2010).
23. J. Wolfowitz, "Asymptotic efficiency of the maximum likelihood estimator," *Teor. Veroyatn. Primen.,* **10**, 267–281 (1965).