Theory of Stochastic Processes
Vol.13 (29), no.1-2, 2007, pp.294-307

ALEXANDER VERETENNIKOV

ON ASYMPTOTIC INFORMATION INTEGRAL INEQUALITIES

Asymptotical versions of Bayesian Cramér – Rao inequalities are discussed.

1. Introduction

Apparently, the Cramér–Rao inequality for unbiased estimators and its asymptotic version are among the most basic results of point estimation theory in modern textbooks in Statistics. This is despite the existence of super-efficiency in a wider class of estimators which makes the use of this inequality in its conventional form at least doubtful. A logically correct approach which is free from the obstacle mentioned above can be based on *LAN* – local asymptotic normality – and *minimax* concepts, which provide improved methods of using the idea of this inequality (see, e.g., [2, 6]). Unfortunately, corresponding monographs are hardly suitable for courses at the undergraduate level. At the same time, there is a nice, relatively easy, and also logically correct way of teaching point estimation theory, based on a certain *integral version* of the famous inequality. The point is that these versions do not require unbiasedness, at least not the usual one. In this way, either Maximum Likelihood Estimator, or Bayesian one under mild assumptions are asymptotically efficient in the integral Mean Square sense. In the other words, there is no super-efficiency phenomenon here. These integral inequalities are in use in practical applications, for example, in signal processing, see [12] et al.

Some history of integral versions of Cramér–Rao's lower bounds can be found, e.g., in the paper [8]. Apparently, the first integral inequality of this sort in a rather general form, although without a precise set of assumptions was established by Schützenberger [9, 10]; the first of these references is cited in the paper [8]. It should be admitted that both sources (i.e., [9, 10]) are not easily available nowadays, however, there is a useful link on

Invited lecture.
2000 *Mathematics Subject Classifications.* 62F15; 60E15.
Key words and phrases. Cramér – Rao inequality, integral asymptotic case.

the publication list, http://www-igm.univ-mlv.fr/˜herstel/Mps. Anyway, in about a decade, this idea of integration with respect to a parameter was re-discovered by Van Trees, see [12], and since that time it is called usually the Van Trees inequality. Probably, the name Schützenberger would be more appropriate here. But in this respect it is rather instructive to remind that an *earlier* version of the classical Cramér–Rao inequality has been established, actually, by Fréchet. This fact was known and seemed to be recognized in old days. Because of this variety of names, – and who knows if there is no other papers which remain just unknown, – we use here a general term, *information inequality*. In any case, in the paper [10], two other names are mentioned, but the author did not manage to trace their papers on the subject.

Let us return to *asymptotic* integral information inequalities, our goal in this paper. Indeed, they are most important for the construction of a complete asymptotic theory of estimation. We will discuss mainly the lower bound (6), due to Borovkov and Sakhanenko. For some other results in this direction see also [1, 7].

Let us say a few words about assumptions, for, in fact, the aim of this paper is new assumptions for an existing inequality (6). The most standard assumption in this theory is a so-called "weak unbiasedness condition", – see (2) below, – which is used in most of the works on the subject, see [8, 5], et al. Notice here that this term "weak unbiasedness" should not confuse the reader: it has practically nothing to do with *unbiasedness*. The paper [3] (see also [2, §30]) introduced an amazing idea of replacing this already weak condition by a very unusual and even weaker technical assumption for the limiting inequality: namely, the "prior density", – even though it could be not a density, – has to be *Riemann integrable* if Θ is bounded, or *directly Riemann integrable* if it is not (see [4, Chapter 11]). We notice that the words "directly integrable" is not used in [2], however, the calculus in the proof, apparently, exploits exactly the latter. There were some generalizations, see [8, 11], and references therein. However, the author did not see any discussion of this surprising Riemann integrability condition anywhere but in the textbook [2]. So, the question arises, could one drop even this remarkable condition? Surely, Riemann integrability as an assumption in the theory based on Lebesgue integration looks strange. The aim of this note is to show that some *asymptotical* versions of this integral inequality are valid without any Riemann integrability at all.

The structure of the paper is as follows. Firstly, we present a new version of the Theorem 5 [3], – that is, a rather similar, although not fully identical result with Riemann integrability, – with a proof. Then we present one more version of this inequality without Riemann integrability at all. In the proofs, we use the approach developed in the proof of the Theorem 30.5 [2], and combine it with some further hints. Notice that in the proof of the

Theorem 30.5 [2] there are some minor inaccuracies, which apparently can be corrected; see the remarks 4 and 5 below. This does not influence on our deep appreciation of the result by Borovkov and Sakhanenko, however, corrections, being unthankful, are not the aim of this paper. Following [2], the general idea of dropping weak unbiasedness is to use approximations, with a level, say, ϵ, adjusted to the sample size n. Finally, let us mention that we will consider here bounded intervals in R^1 as parametric sets; generalizations to unbounded parametric sets and to any finite dimension are possible. The idea of this work gradually arose from teaching theoretical statistics in the University of Leeds, and earlier in Moscow State University (cf. [13]).

2. Setting and results: $\Theta = (a, b)$

We consider a family of densities $(f(x\,|\,\theta),\ \theta \in \bar\Theta)$, with a parametric set $\Theta = (a, b)$, $-\infty < a < b < \infty$, independent samples $X = (X_1, \dots, X_n)$, $n \to \infty$, Fisher's information $I(\theta) := E_\theta \left(\partial \ln L_n(X \mid \theta)/\partial\theta\right)^2$, where L_n is a conditional density for the sample X given θ, and the "prior density" $q(\theta)$, $\theta \in \bar\Theta$, which is a Borel measurable function. We assume that the function I is also Borel, but *do not* require that it is necessarily continuous. All these are standard assumptions which are not reminded in the sequel. We just notice that considering $\bar\Theta$ is theoretically rather convenient, see [6].

In addition to these standard requirements, we assume in all Theorems and Proposition below, as in [2, §30], the following:

(A1)

$$0 < J := \int_a^b \frac{q(t)}{I(t)}\,dt < \infty, \qquad \int_a^b \sqrt{I(t)}\,dt < \infty. \tag{1}$$

The *weak unbiasedness* condition which we mentioned above, – and which we do not use below, – reads,

$$q(b)E_b(\theta^* - b) - q(a)E_a(\theta^* - a) = 0, \quad q \in C[a, b]. \tag{2}$$

In turn, a simple sufficient condition for (2) is

$$q(a) = q(b) = 0, \quad q \in C[a, b]. \tag{3}$$

On the other hand, (2) could be guaranteed also by ordinary unbiasedness at the two limiting points, i.e. at a and b. Instead of (2), the Theorem 5 [3] assumes (1), and, in addition, requires that

($\mathbf{A}_{BS}$)

$$\text{the function } q \text{ is Riemann integrable on } [a, b]. \tag{4}$$

We believe that some other minor condition should be added to (A_{BS}) in order to be sufficient for the inequality (6), see the Remark 5 below. Instead, we will use here a slightly different version of the latter assumption,

(A2)

$$\text{the function } q/I \text{ is Riemann integrable on } [a,b], \quad \text{and} \quad \inf_{a\le t\le b} q(t) > 0. \qquad (5)$$

First of all, we propose a new version of the Theorem 5 from [3] and [2, §30], under the assumption (A2) instead of (A_{BS}).

Theorem 1. *Let (A1) and (A2) hold true. Then*

$$\liminf_{n\to\infty} n\, E(\theta^* - \theta)^2 = \liminf_{n\to\infty} n \int E_\theta(\theta^* - \theta)^2\, q(\theta)\, d\theta \ge J. \qquad (6)$$

Now, we are going to show that *in some cases* it is possible to relax the assumption on Riemann integrability further, even drop it completely, see the Theorem 2 below. The following new assumption which describes one of these cases will be used,

(A3) there exists $C > 0$ such that

$$C^{-1} \le q(t)/I(t) \le C, \quad a \le t \le b, \qquad \text{and} \quad \inf_{a\le t\le b} I(t) > 0.$$

Theorem 2. *Let the assumptions (A1) and (A3) be satisfied. Then (6) holds true.*

The "price" of dropping the assumption on Riemann integrability is a new restriction on the "prior", (A3); at least, in certain cases this restriction could be considered as mild. Certain "mixture" or the two assumptions, (A2) and (A3) is possible, – that is, for example, one could require (A2) on one part of Θ, and (A3) on the complementary part, – but we do not pursue this here. Let us state a useful technical lemma which will be applied in the proof of the Theorem 2; however, it may have some independent interest, too: see the Corollary 1 below.

Lemma 1. *Let the assumption (A1) be satisfied, let $q \ge 0$, and let there be a sequence $0 \le q_m \uparrow q$ (a.e.) as $m \to \infty$, such that for any estimator θ^*,*

$$\liminf_{n\to\infty} n \int E_\theta(\theta^* - \theta)^2 \tilde q_m(\theta)\, d\theta \ge \int \frac{\tilde q_m(t)}{I(t)}\, dt,$$

where

$$\tilde q_m(t) := \frac{q_m(t)}{\int q_m(\theta)\, d\theta}.$$

Then (6) holds true with the prior q*.*

Corollary 1. *Let* $q \geq 0$*, and*

$$0 \leq q_m(t) \uparrow q(t), \quad m \to \infty, \quad a.e.$$

If every function q_m *is Riemann integrable and satisfies (A1), then (6) holds true for this* q*.*

Some independent interest here might arise because the class of such functions q that possess monotonic approximation from below by Riemann integrable ones is wider than Riemann integrable, although it is yet more narrow in compare to $L_1[a, b]$.

Remark 1. Is it tempting to formulate an analogue of the Lemma 1 for a *decreasing* sequence approximating q, in order to extend the Theorem 2 to all Lebesgue's integrable densities. However, this idea does not seem to help much, because it only works, apparently, if the convergence is *uniform*; and as such, it could be re-formulated as a uniform *increasing* convergence, too, just by a multiplication by an appropriate constant close to one. Hence, this new Lemma, which could be, indeed, stated, follows from, and is actually strictly weaker then the Lemma 1 above.

Remark 2. In certain works, the notion of "Fisher's integral information" is used,

$$\bar{I}(n) := E(\partial \ln L_n(X, \theta)/\partial\theta)^2; \tag{7}$$

a similar definition via the second derivative could be applied, too. Then, e.g., under the assumption of weak unbiasedness (3), the following Cramér–Rao integral inequality can be established,

$$E(\theta^* - \theta)^2 \geq \bar{I}(n)^{-1} = \frac{1}{n \int q(\theta) I(\theta)\, d\theta + \int (q')^2/q\, d\theta}\,. \tag{8}$$

In the limit this gives,

$$\liminf_n nE(\theta^* - \theta)^2 \geq \frac{1}{\int q(\theta) I(\theta)\, d\theta}\,. \tag{9}$$

Due to Jensen's inequality applied to the strictly convex function ($t \mapsto 1/t$, $t > 0$), we have,

$$J = \int q(\theta) \frac{1}{I(\theta)}\, d\theta \geq \frac{1}{\int q(\theta) I(\theta)\, d\theta} =: (\bar{I})^{-1}.$$

This fact is, of course, well known and mentioned in the literature. Naturally, in most of cases the latter inequality is strict. So, the integral asymptotic inequality with the Borovkov–Sakhanenko bound J, - as in [3] or in the Theorem 1 above, – is strictly stronger, and, therefore, is clearly more reasonable than the lower bound with $\bar{I}$, of course, if J is finite. As shown in [2, Example 30.1], and as it may be proved for MLE's under mild conditions, the bounds with J is often attainable, see [2]. In the latter Example from [2], of course, $(\bar{I})^{-1} < J$, hence, the use of $(\bar{I})^{-1}$ is not reasonable, because this bound is not attainable. We are not speaking of the case $\bar{I} = 0$, nor of non-differentiable densities. The question about optimal finite lower bounds apparently was not studied in the literature.

We will be using the following result, see, e.g., [2, §30, Theorem 4]. Notice that the inequality (10) with $h_\epsilon = q$, leads to (9), while $h_\epsilon = q/I$ provides (6), – of course, each one under the condition that q, or correspondingly q/I, satisfies the assumptions of the Proposition.

Proposition 1. *Let h_ϵ be absolutely continuous function, – that is, possess a representation as an integral of some Lebesgue integrable function, – satisfy the equalities*

$$h_\epsilon(a) = h_\epsilon(b) = 0,$$

and

$$supp(h_\epsilon) \subset supp(q),$$

and let the second condition in (A1) be satisfied. (The first one is not exploited here because we do not use J.) Then

$$n \int E_\theta(\theta^* - \theta)^2 \, q(\theta) \, d\theta \geq \frac{\left(\int_a^b h_\epsilon(t) \, dt\right)^2}{\int_a^b I(t) h_\epsilon(t)^2/q(t) \, dt + (1/n) \int_a^b (h'_\epsilon(t))^2/q(t) \, dt}. \tag{10}$$

For convenience of reading let us remind the main idea of the proof, that is an application of the Cauchy inequality to the identity,

$$E\left((\theta^*(X) - \theta) \frac{(f(X \mid \theta) h_\epsilon(\theta))'_\theta}{f(X, \theta)}\right) = \int h_\epsilon(t) \, dt = E \, \frac{h_\epsilon(\theta)}{q(\theta)} \, ,$$

which, in turn, follows from

$$E\left(\theta^*(X) \frac{(f(X \mid \theta) h_\epsilon(\theta))'_\theta}{f(X, \theta)}\right) = \int \theta^*(x) \left(\int (f(x \mid t) h_\epsilon(t))'_t \, dt\right) dx = 0,$$

due to $h_\epsilon(a) = h_\epsilon(b) = 0$, and

$$-E\left(\theta \, \frac{(f(X \mid \theta) h_\epsilon(\theta))'_\theta}{f(X, \theta)}\right) = -\int \left(\int t \, (f(x \mid t) h_\epsilon(t))'_t \, dt\right) dx$$

$$= \int \left(\int (f(x \mid t) h_\epsilon(t)) \, dt\right) dx = \int \left(\int (f(x \mid t) h_\epsilon(t)) \, dx\right) dt$$

$$= \int \int \frac{(f(x \mid t) h_\epsilon(t))}{q(t)} \, q(t) \, dx \, dt = E \, \frac{h_\epsilon(\theta)}{q(\theta)} \, ,$$

where $h_\epsilon(a) = h_\epsilon(b) = 0$ has been also used explicitly. We skip further details which can be found in [2].

Remark 3. Notice that in in [2, Theorem 30.4], one may wish to use a function h_ϵ which is not necessarily non-negative; in this case, one should require $|h_\epsilon| \le h$ instead of $h_\epsilon \le h$. We do not need either of these in the Proposition above, because we do not aim to get an inequality with h in this assertion.

3. Proof of Theorem 1

1. Let us denote,

$$q_- := \inf_{a\le t\le b} q(t) > 0,$$

see the assumption (A2). For the function $\int_a^b (h'_\epsilon(t))^2/q(t)\,dt$, the following notation will be used,

$$H(\epsilon) := \int_a^b (h'_\epsilon(t))^2/q(t)\,dt.$$

Let

$$h_0(t) := q(t)/I(t),$$

$$\bar h_\epsilon(t) := \min_{|u|\le\epsilon} \frac{q(t+u)}{I(t+u)}, \quad \tilde h_\epsilon(t) := \bar h_\epsilon(t) \wedge \frac{q_-}{\epsilon}, \qquad a \le t \le b,$$

and

$$h_\epsilon(t) := \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \tilde h_\epsilon(v)\,dv. \tag{11}$$

With this definition, we clearly have

$$\tilde h_\epsilon(t) \le h_0(t), \quad \& \quad 0 \le h_\epsilon(t) \le h_0(t). \tag{12}$$

Now, the function h_ϵ defined in (11) is absolutely continuous and differentiable almost everywhere, with

$$|h'_\epsilon(t)| \le \frac{Cq(t)}{\epsilon} \wedge \frac{q(t)}{I(t)},$$

and $h_\epsilon(a) = h_\epsilon(b) = 0$, for any $\epsilon > 0$; of course, we define q outside $[a, b]$ as identical zero. Due to the assumption (A2), the function $H(\epsilon)$ is finite, and, moreover,

$$H(\epsilon) \le \frac{C}{\epsilon^2}\int \frac{q^2(t)}{q(t)}\,dt \le \frac{C}{\epsilon^2}.$$

2. Let us show that

$$\tilde h_\epsilon(t) \to h_0(t), \quad \epsilon \downarrow 0 \quad (a.e.). \tag{13}$$

For that, due to the Lebesgue dominated convergence theorem, it suffices to show that

$$\int_a^b (h_0(t) - \tilde{h}_\epsilon(t))\, dt \downarrow 0, \quad \epsilon \downarrow 0. \tag{14}$$

This follows similarly to [2, Proof of Theorem 30.5], where this hint is applied to the function q. We have, by virtue of Riemann integrability condition and the theorem about Darboux integral sums from the Calculus,

$$\sum_k \bar{h}_\delta(2k\delta)\, 2\delta \to \int_a^b h_0(t)\, dt, \quad \delta \to 0,$$

$$\sum_k \bar{h}_\delta((2k+1)\delta)\, 2\delta \to \int_a^b h_0(t)\, dt, \quad \delta \to 0.$$

Estimate the difference,

$$0 \le \sum_k (\bar{h}_\delta(2k\delta) - \tilde{h}_\delta(2k\delta))\, 2\delta$$

$$\le 2\delta \sum_k \bar{h}_\delta(2k\delta)\, 1(\bar{h}_\delta(2k\delta) > q_-/(2\delta)).$$

However, since h_0 is Riemann integrable, in must be bounded on $[a, b]$, and so is $\bar{h}_\delta \le h_0$. Since $\inf_{t\in[a,b]} q(t) > 0$, then it follows from (A2) that $\tilde{h}_\delta \equiv \bar{h}_\delta$ as δ is small enough. Then, of course,

$$1(\bar{h}_\delta(2k\delta) > q_-/(2\delta)) = 0.$$

Therefore the sum $\sum_k \bar{h}_\delta(2k\delta)\, 1(\bar{h}_\delta(2k\delta) > q_-/(2\delta))$ equals zero if δ is small enough. So,

$$0 \le \sum_k (\bar{h}_\delta(2k\delta) - \tilde{h}_\delta(2k\delta))\, 2\delta \to 0, \qquad \delta \to 0.$$

Similarly,

$$0 \le \sum_k (\bar{h}_\delta((2k+1)\delta) - \tilde{h}_\delta((2k+1)\delta))\, 2\delta \to 0, \qquad \delta \to 0.$$

Hence,

$$\int_a^b \tilde{h}_\epsilon(t)\, dt \ge \left(\sum_k \tilde{h}_{2\epsilon}(4k\epsilon)\, 2\epsilon + \sum_k \tilde{h}_{2\epsilon}((4k+2)\epsilon)\, 2\epsilon \right) \to \int_a^b h_0(t) dt, \ \epsilon \to 0. \tag{15}$$

Since $\int \tilde{h}_\epsilon \le \int h$, the latter convergence implies (14); hence, (13) holds true *almost everywhere* for $a \le t \le b$.

3. Notice that h_ϵ satisfies the assumptions of the Proposition 1, being differentiable and since it vanishes at a and b. So, we get, with $\epsilon = (Cn)^{-1/3}$,

$$n\,E(\theta^* - \theta)^2 \ge \frac{\left(\int_a^b h_\epsilon(t)\,dt\right)^2}{\int_a^b I(t)h_\epsilon(t)^2/q(t)\,dt + n^{-1/3}}. \tag{16}$$

Hence, to complete the proof, it suffices to establish

$$\int_a^b h_\epsilon(t)\,dt \to \int_a^b h_0(t)\,dt, \tag{17}$$

and

$$\int_a^b I(t)h_\epsilon(t)^2/q(t)\,dt \to \int_a^b q(t)/I(t)\,dt, \quad \epsilon \to 0. \tag{18}$$

4. We have,

$$0 \le \int_a^b \left(h_0(t) - h_\epsilon(t)\right)\,dt$$

$$= \int_a^b \left(h_0(t) - \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \tilde{h}_\epsilon(v)\,dv\right)\,dt$$

$$= \int_a^b \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \left(h_0(t) - \tilde{h}_\epsilon(v)\right)\,dv\,dt$$

$$= \int_a^b \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \left(h_0(t) - \tilde{h}_\epsilon(t)\right)\,dv\,dt$$

$$+ \int_a^b \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \left(\tilde{h}_\epsilon(t) - \tilde{h}_\epsilon(v)\right)\,dv\,dt.$$

Here,

$$\int_a^b \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \left(h_0(t) - \tilde{h}_\epsilon(t)\right)\,dv\,dt = \int_a^b \left(h_0(t) - \tilde{h}_\epsilon(t)\right)\,dt \to 0, \quad \epsilon \to 0,$$

due to (14). On the other hand side,

$$\int_a^b \frac{1}{2\epsilon}\int_{t-\epsilon}^{t+\epsilon} \left(\tilde{h}_\epsilon(t) - \tilde{h}_\epsilon(v)\right)\,dv\,dt =$$

$$= \int \tilde{h}_\epsilon(t)\,dt - \int \tilde{h}_\epsilon(v)\left(\frac{1}{2\epsilon}\int_{v-\epsilon}^{v+\epsilon} 1\,dt\right)\,dv = 0.$$

Thus, indeed, (17) holds true.

5. Further, by virtue of (12) and (17, we also have,

$$0 \le \int \frac{I(t)}{q(t)}(h_0^2(t) - h_\epsilon^2(t))\,dt$$

$$= \int h_0^{-1}(h_0(t) - h_\epsilon(t))(h_0(t) + h_\epsilon(t))\,dt$$

$$\le 2\int h_0^{-1}(h_0(t) - h_\epsilon(t))2h_0(t)\,dt$$

$$= 2\int (h_0(t) - h_\epsilon(t))\,dt \to 0, \quad \epsilon \to 0.$$

Whence, from (16), (17) and (18) the desired inequality (6) follows. Q.E.D.

Remark 4. Here is a simple example where the sum $\sum_k q_\delta(2k\delta)\,2\delta$ may decrease when δ decreases in [2]. Let $a = 0$, $b = 5$, $\delta_1 = 1$, $\delta_2 = 3/2$, and $q(t) = 1(1 < t < 5) \times (1/4)$. Then, clearly, $\delta_1 < \delta_2$, but the sum with δ_2 is not less than that with δ_1: indeed,

$$\sum_k q_{\delta_1}(2k\delta_1)\,2\delta_1 = 0 < 3/4 = \sum_k q_{\delta_2}(2k\delta_2)\,2\delta_2.$$

This does not change the rest of the proof in [2, Theorem 30.5].
Remark 5. More important is that for the construction used in [2], – which is slightly different from ours, – the inequality $h_\epsilon \le h_0$ may fail; there is simply no reason why it could be guaranteed. This inequality which asserts the applicability of Lebesgue's dominated convergence theorem, might be replaced by $h_\epsilon \le Ch_0$ *under additional suitable conditions* on q and/or I, for example, under $0 < \inf_{a\le t\le b} I(t) \le \sup_{a\le t\le b} I(t) < \infty$ (cf. (A2)). However, as was explained above, we do not pursue this goal.

4. Proof of Lemma 1

We have, by the assumptions,

$$\liminf_{n\to\infty} \int E_\theta(\theta^* - \theta)^2 \tilde q_m(\theta)\,d\theta \ge \int \frac{\tilde q_m}{I}(\theta)\,d\theta.$$

From the monotone convergence theorem it follows,

$$\kappa_m := \int q_m \uparrow 1.$$

Again by the monotone convergence theorem,

$$\int \frac{q_m}{I}\,dt \uparrow \int \frac{q}{I}\,dt = J.$$

Thus, we get,

$$\liminf_{n\to\infty} n\int E_\theta(\theta^*-\theta)^2 q(\theta)\,d\theta \ge \liminf_{n} n\int E_\theta(\theta^*-\theta)^2 q_m(\theta)\,d\theta$$

$$= \kappa_m \liminf_{n} n\int E_\theta(\theta^*-\theta)^2 \tilde q_m(\theta)\,d\theta \ge \kappa_m \int \frac{\tilde q_m(\theta)}{I(\theta)}\,d\theta = \int \frac{q_m(\theta)}{I(\theta)}\,d\theta.$$

Since the left hand side here does not depend on m, we deduce that

$$\liminf_{n\to\infty} n\int E_\theta(\theta^*-\theta)^2 q(\theta)\,d\theta \ge \limsup_{m\to\infty} \int \frac{q_m(\theta)}{I(\theta)}\,d\theta = J.$$

Q.E.D.

5. Proof of Theorem 2

1. We will construct approximations q_m suitable for using the Lemma 1. Let

$$q_m(t) := q(t)\,1(a+1/m<t<b-1/m), \quad m\ge 1.$$

Then $0\le q_m(t)\uparrow q(t)$, $m\to\infty$. Denote

$$\kappa_m = \int q_m, \qquad \& \qquad \tilde q_m = q_m/\kappa_m.$$

To prove the Theorem, it suffices to show that for every m,

$$\liminf_{n\to\infty} n\int E_\theta(\theta^*-\theta)^2 \tilde q_m(\theta)\,d\theta \ge \int \frac{\tilde q_m(t)}{I(t)}\,dt. \tag{19}$$

Denote

$$S_m := \mathrm{supp}\,(q_m) = [a+1/m, b-1/m],$$

and

$$h_{0,m}(t) := \frac{\tilde q_m(t)}{I(t)},$$

and consider the following continuous piece-wise linear function $\varphi=\varphi_{\epsilon,m}$, with $\epsilon<(b-a-1/m)/2$ and $m>2/(b-a)$,

$$\varphi_{\epsilon,m}(t) = \begin{cases} a+\frac{1/m+\epsilon}{1/m}\,(t-a), & a\le t\le a+1/m,\\ +\epsilon\,\frac{a+b}{b-a-2/m}+t\left(1-\frac{2\epsilon}{b-a-2/m}\right), & a+1/m\le t\le b-1/m,\\ b+\frac{1/m+\epsilon}{1/m}(t-b), & b-1/m\le t\le b. \end{cases}$$

Notice that

$$\varphi(a)=a, \quad \varphi(a+1/m)=a+1/m+\epsilon, \quad \varphi(b-1/m)=b-1/m-\epsilon, \quad \varphi(b)=b,$$

$$0 < C^{-1} \le \varphi'_{\epsilon,m} \le C < \infty, \quad \sup_t \left|\varphi'_{\epsilon,m}(t) - 1\right| \to 0, \quad \epsilon \to 0,$$
$$\sup_t |\varphi_{\epsilon,m}(t) - t| \to 0, \quad \epsilon \to 0,$$

and

$$\tilde{q}_m(a + 1/m-) = \tilde{q}_m(b - 1/m+) = 0.$$

In particular, it follows,

$$\sup_v \left|1 - \frac{1}{2\epsilon}\int_{\varphi^{-1}_{\epsilon,m}(v-\epsilon)}^{\varphi^{-1}_{\epsilon,m}(v+\epsilon)} dt\right| \to 0, \quad \epsilon \to 0. \tag{20}$$

Let

$$h_{\epsilon,m}(t) := \frac{1}{2\epsilon}\int_{\varphi_{\epsilon,m}(t)-\epsilon}^{\varphi_{\epsilon,m}(t)+\epsilon} h_{0,m}(v)\, dv.$$

Then there exists C such that for every ϵ small enough, and every m large enough,

$$|h'_{\epsilon,m}(t)| \le C/\epsilon. \tag{21}$$

By virtue of the Proposition 1 with this function as h_ϵ, we get the following,

$$n \int E_\theta(\theta^* - \theta)^2 \tilde{q}_m(\theta)\, d\theta \ge \frac{\left(\int_a^b h_{\epsilon,m}(t)\, dt\right)^2}{\int_a^b I(t) h^2_{\epsilon,m}(t)/q(t)\, dt + (1/n) \int_a^b (h'_{\epsilon,m}(t))^2/q(t)\, dt}. \tag{22}$$

2. Let us show that

$$\int_a^b h_{\epsilon,m}(t)\, dt \to \int_a^b h_{0,m}(t)\, dt, \tag{23}$$

and

$$\int_{a+1/m}^{b-1/m} I(t) h^2_{\epsilon,m}(t)/\tilde{q}_m(t)\, dt \to \int_{a+1/m}^{b-1/m} \tilde{q}_m(t)/I(t)\, dt, \quad \epsilon \to 0. \tag{24}$$

To show (23), we simply notice that

$$\int h_{\epsilon,m}(t)\, dt = \int dt\, \frac{1}{2\epsilon}\int_{\varphi_{\epsilon,m}(t)-\epsilon}^{\varphi_{\epsilon,m}(t)+\epsilon} h_{0,m}(v)\, dv$$

$$= \int h_{0,m}(v)\, dv\, \frac{1}{2\epsilon}\int_{\varphi^{-1}_{\epsilon,m}(v-\epsilon)}^{\varphi^{-1}_{\epsilon,m}(v+\epsilon)} dt \to \int h_{0,m}\, dt, \quad \epsilon \to 0,$$

due to the Lebesgue dominated convergence theorem and (20). To show (24), we notice that

$$\int \frac{I(t)}{q_m(t)} \left(h_{\epsilon,m}(t)^2 - h_{0,m}(t)^2\right) dt$$

$$= \int \frac{I(t)}{q_m(t)} \left(h_{\epsilon,m}(t) - h_{0,m}(t)\right)\left(h_{\epsilon,m}(t) + h_{0,m}(t)\right) dt.$$

Since the terms $\frac{I(t)}{q_m(t)}$ and $(h_{\epsilon,m}(t) + h_{0,m}(t))$ are uniformly bounded on S_m, it suffices to establish

$$\int |h_{\epsilon,m}(t) - h_{0,m}(t)| \, dt \to 0, \quad \epsilon \to 0. \tag{25}$$

Let $\delta > 0$ be any positive value, and let us approximate the function $h_{0,m}$ in $L_1[a,b]$ by some continuous function $h^\delta_{0,m}$ so that

$$\int |h_{0,m} - h^\delta_{0,m}| < \delta.$$

Then, denoting

$$h^\delta_{\epsilon,m}(t) = \frac{1}{2\epsilon} \int_{\varphi_{\epsilon,m}(t)-\epsilon}^{\varphi_{\epsilon,m}(t)+\epsilon} h^\delta_{0,m}(v) \, dv,$$

we get,

$$\begin{aligned}
\int |h_{\epsilon,m} - h^\delta_{\epsilon,m}|(t) \, dt &= \int \left| \frac{1}{2\epsilon} \int_{\varphi_{\epsilon,m}(t)-\epsilon}^{\varphi_{\epsilon,m}(t)+\epsilon} \Big(h_{0,m}(v) - h^\delta_{0,m}(v) \Big) \, dv \right| \, dt \\
&\le \int \frac{1}{2\epsilon} \int_{\varphi_{\epsilon,m}(t)-\epsilon}^{\varphi_{\epsilon,m}(t)+\epsilon} \Big| h_{0,m}(v) - h^\delta_{0,m}(v) \Big| \, dv \, dt \\
&= \int dv \Big| h_{0,m}(v) - h^\delta_{0,m}(v) \Big| \frac{1}{2\epsilon} \int_{\varphi^{-1}_{\epsilon,m}(v-\epsilon)}^{\varphi^{-1}_{\epsilon,m}(v+\epsilon)} dv \, dt \le C\delta.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int |h_{\epsilon,m}(t) - h_{0,m}(t)| \, dt \\
&\le \int \Big| h_{\epsilon,m}(t) - h^\delta_{\epsilon,m}(t) \Big| \, dt + \int \Big| h_{0,m}(t) - h^\delta_{0,m}(t) \Big| \, dt + \int \Big| h^\delta_{\epsilon,m}(t) - h^\delta_{\epsilon,m}(t) \Big| \, dt \\
&\le C\delta + \int \Big| h^\delta_{\epsilon,m}(t) - h^\delta_{0,m}(t) \Big| \, dt.
\end{aligned}$$

For every fixed $\delta > 0$, the latter integral tends to zero as $\epsilon \to 0$, because the function $h^\delta_{0,m}$ is uniformly continuous, and, hence, $\sup_x (h^\delta_{\epsilon,m} - h^\delta_{0,m})(x) \to 0$, $\epsilon \to 0$. Therefore, for every $\delta > 0$,

$$\limsup_{\epsilon \to 0} \int |h_{\epsilon,m} - h^\delta_{\epsilon,m}|(t) \, dt \le C\delta;$$

however, the left hand side $\int |h_{\epsilon,m}(t) - h_{0,m}(t)| \, dt$ does not depend on δ, hence, (25) holds true, which implies (24). From (23), (24) and (22) we deduce (19), which, finally, implies the desired asymptotic inequality (6) by virtue of the Lemma 1. Q.E.D.

6. Acknowledgements

The author thanks the grants by Leeds University for conferences on teaching in higher education, and RFBR 05-01-00449 (Russia) for support. The author is grateful to Mrs Hélène Schützenberger who has kindly provided a copy of the paper [10] by her father.

Bibliography

1. Bobrovsky, B. Z., Zakai, M. *A lower bound on the estimation error for certain diffusion processes,* IEEE Trans. Inform. Theory. **IT-22**, (1976), 45–52.
2. Borovkov, A. A. *Mathematical statistics,* Gordon and Breach Science Publishers, Amsterdam, (1998).
3. Borovkov, A. A., Sakhanenko, A. I. *Estimates for averaged quadratic risk,* (Russian) Probab. Math. Statist. **1**, (1980), no. 2, 185-195 (1981).
4. Feller, W. *An introduction to probability theory and its applications*, Vol. 2., John Wiley & Sons, New York-London-Sydney, (1966).
5. Gill, R. D., Levit, B. Y. *Applications of the Van Trees inequality: a Bayesian Cramér-Rao bound,* Bernoulli, **1**, (1995), no. 1-2, 59–79.
6. Ibragimov, I. A.; Hasminskii, R. Z. *Statistical estimation. Asymptotic theory,* Springer, New York - Berlin, (1981).
7. Nazin, A. V. *Information approach to optimization problems and adaptive control of discrete stochastic systems,* DrSc Thesis, Moscow, Inst. Control Sci., 1995.
8. Prakasa Rao, B. L. S. *On Cramér-Rao type integral inequalities,* Calcutta Statist. Assoc. Bull. **40**, (1990/91), no. 157-160, 183–205.
9. Schützenberger, M. P. *A generalization of the Frechet–Cramér inequality to the case of Bayes estimation,* Bull. Amer. Math. Soc., **63**, (1957), 142.
10. Schützenberger, M. P. *A propos de l'inégalité de Fréchet – Cramér,* Publ. Inst. Statist. Univ. Paris 7, no. 3/4 (1958), 3–6.
11. Shemyakin, A. E. *Rao–Cramér type integral inequalities for estimates of a vector parameter,* Theory Probab. Appl., **32**, (1987), 426–434; translation from Teor. Veroyatn. Primen. 32(3), (1987), 469–477.
12. Van Trees, H. *Detection, Estimation and Modulation Theory,* Vol. I., Wiley, New York, (1968).
13. Veretennikov, A. Yu. *Parametric and non-parametric estimation for Markov chains,* Lecture notes, Moscow, Publ. Centre Appl. Research, Mech. & Math. Faculty, Moscow State Univ., (2000).
14. Weiss, A. J., Weinstein, E. *A lower bound on the mean-square error in random parameter estimation,* IEEE Trans. Inf. Theory, **31**, (1985), 680–682.

School of Mathematics, University of Leeds, LS2 9JT, Leeds, UK, & Institute of Information Transmission Problems, Moscow, 127994, GSP-4, Russia.

E-mail address: a.veretennikov@leeds.ac.uk