ON LARGE DEVIATION PROBABILITIES FOR THE MAXIMUM LIKELIHOOD ESTIMATORS

K. DZHAPARIDZE and E. VALKEILA*

Centre for Mathematics and Computer Science,
P.O. Box 4079, 1009 AB Amsterdam, the Netherlands
Computer Centre, University of Helsinki,
Teollisuuskatu 23, SF-00510 Helsinki, Finland

ABSTRACT

This paper is devoted to estimating large deviation probabilities of the maximum likelihood estimator from the "true" value of an unknown parameter. In contrast with (Ibragimov and Has'minskii, 1981, Theorem I.5.1) we consider a sequence of parametric families of filtered experiments (cf. (Valkeila and Vostrikova, 1987)). This allows us to verify the conditions of this theorem in so-called "predictable" terms — in terms of the Hellinger processes $h(\alpha; P, P')$ of order $\alpha \in (0, 1)$, associated with two probability measures P and P' in a filtered space as, e.g., in (Jacod and Shiryaev, 1987).

1. INTRODUCTION

1.1. Let
$$(\Omega^n, F^n, \mathcal{F}^n, \{P_\theta^n, \theta \in \Theta\}, Q^n), \qquad n = 1, 2, \ldots$$
be a sequence of parametric families of filtered experiments with a measure Q^n dominating a family $\{P_\theta^n, \theta \in \Theta\}$ which is indexed by a certain set Θ in $\mathbf{R}^d$. A filtration $\mathcal{F}^n = (F_t^n)_{t \geqslant 0}$, satisfying the usual assumptions with respect to Q^n, is supposed to be such that $\vee_t F_t^n = F_\infty^n = F^n$ and $F_0^n = \{\emptyset, \Omega^n\}$, for simplicity.

1.2. For a $\mathcal{F}^n$ stopping time T restrict the measures P_θ^n and Q^n to the sub–σ–field F_T^n of F^n. Consider the density process $z_t^n(\theta) = dP_\theta^n/dQ^n|F_t^n$, and for each θ take a version with paths in the space $\mathbf{D}$ of right–continuous functions with left–hand limits. The density process $z^n(\theta)$ is supposed to be continuous in θ.
For a fixed time t define the maximum likelihood estimator θ_t^n as the value of θ in Θ that renders $z_t^n(\theta)$ as large as possible: $z_t^n(\theta_t^n) = \sup_{\theta \in \Theta} z_t^n(\theta)$.
It is supposed that a solution to this equation exists P_θ^n – a.s., for all θ (this is certainly true if Θ is compact), and therefore, as is shown in (Sieders and Dzhaparidze, 1987), it is measurable.
Our result on properties of the maximum likelihood estimator θ_t^n is of asymptotical nature, i.e. it is valid for n and R large enough and t fixed, where $n \to \infty$ describes

---

the approach to the "limiting experiment", and R the normalized deviation from the "true" value θ (see 3.2. below).

1.3. Following (Ibragimov and Has'minskii, 1981, Theorem I.5.1) to characterize the rate of convergence (cf. also (Sieders and Dzhaparidze, 1987)), for each fixed n we consider a function g_n on $[0, \infty[$, monotonically increasing to infinity and such that $R^L \exp\{-g_n(R)\} \to 0$ as n and R tend to infinity, for each L.

1.4. As in the works mentioned above, conditions are sought for establishing the following inequality concerning large deviation probabilities:

$$P_\theta^n\{|\varphi_n^{-1}(\theta_t^n - \theta)| > N\} \leqslant B \exp\{-bg_n(N)\}$$

valid uniformly in $\theta \in \Theta$, for n and N large enough and t fixed, with certain positive constants b and B and a certain normalizing sequence of positive definite matrices φ_n, $n = 1, 2, \ldots$, which usually depends on θ, i.e. $\varphi_n = \varphi_n(\theta)$, $n = 1, 2, \ldots$

Unlike these works, however, our conditions are expressed in so–called "predictable" terms, namely in terms of the Hellinger processes of order $\alpha \in (0, 1)$, associated to our sequence of parametric families of experiments in the same fashion as, e. g., in (Jacod and Shiryaev, 1987, Ch. IV) (see Section 2 below for more details). In this sense the results presented here can be viewed as an extention of the results in (Valkeila and Vostrikova, 1987), as well as the ideas and results in the works mentioned above, especially (Ibragimov and Has'minskii, 1981).

1.5. The present paper is organized as follows. In Section 2 we restrict our attention to the case of binary experiments involving two probability measures P and P' in order to associate to it the Hellinger processes and related, so–called "divergency" process, and to present in 2.7 a necessary result (Dzhaparidze and Valkeila, 1989), concerning the Hellinger–type distances between P and P'. Next, following (Valkeila and Vostrikova, 1987), we present in 2.6 a simple, useful estimate for the Hellinger integral associated to P and P'. We end this Section by giving more upper bounds for the Hellinger integrals.

In Section 3 we formulate and prove our main theorem concerning large deviation probabilities mentioned in Subsection 1.4 above. Based on the results mentioned in the preceding section, the proof is essentially reduced to verifying the conditions of a version by (Sieders and Dzhaparidze, 1987) of Theorem I.5.1 in (Ibragimov and Has'minskii, 1981).

We refer to (Kutoyants, 1984) and (Ibragimov and Has'minskii, 1981) for various applications. As noted above, we shall formulate our results for a general filtered experiments. We hope to give applications of our results elsewhere.

2. BINARY EXPERIMENTS

2.1. Consider a binary experiment $(\Omega, F, \mathcal{F}, P, P')$ and let a measure Q dominate both P and P'. As in 1.2 define the density processes z and z' by

$$z_T = dP_T/dQ_T \quad \text{and} \quad z'_T = dP'_T/dQ_T.$$

2.2. As usual, z^c and Δz will denote the continuous part and jumps of the process z. We associate to them the quadratic characteristic $\langle z^c \rangle$ and the jump measure μ^z with the $(Q, \mathcal{F})$-compensator $\nu^{z,Q}$. According to (Jacod and Shiryaev, 1987, Theorem IV.1.33), μ^z and its compensator only charges the set

$$\{(\omega, t, x) : z_{t-}(\omega) > 0, z'_{t-}(\omega) > 0, -z_{t-}(\omega) \leqslant x \leqslant z'_{t-}(\omega)\}.$$

For bravity, we use the following notations $X = zz'$ and $Y(\alpha) = z^\alpha z'^{1-\alpha}$ with $\alpha \in (0,1)$. Put also $\lambda = \lambda(x) = 1 + x/z_-$ and $\lambda' = \lambda'(x) = 1 - x/z'_-$.

2.3. We present here briefly certain known facts concerning the Hellinger integrals and Hellinger processes; for more details see (Jacod and Shiryaev, 1987, Section IV.1).
Recall that

$$H(\alpha; P, P') = \mathbf{E}_Q(z^\alpha z'^{1-\alpha}) = \mathbf{E}_Q Y(\alpha)$$

is called the Hellinger integral of order $\alpha \in (0,1)$. Obviously, it is free of a particular choise of a dominating measure Q.
By definition, the Hellinger process $h(\alpha) = h(\alpha; P, P')$ is a predictable increasing process for which $Y(\alpha) + Y(\alpha)_- \circ h(\alpha)$ is a $(Q, \mathcal{F})$-martingale. Recall that not only $H(\alpha; P, P')$ is free of a particular choise of a measure Q, but also the process $h(\alpha; P, P')$. Hence, for a $\mathcal{F}$-stopping time T we have $H(\alpha; P_T, P'_T) = = 1 - \mathbf{E}_Q(Y(\alpha)_- \circ h(\alpha; P_T, P'_T))$. Besides, the following multiplicative decomposition takes place (Jacod and Shiryaev, 1987, Section V.4):

$$Y(\alpha) = N(\alpha)\mathcal{E}(-h(\alpha)) \tag{2.1}$$

with a certain $(Q, \mathcal{F})$-supermartingale $N(\alpha)$. As usual, $\mathcal{E}$ denotes the Dolean's exponential. Hence

$$H(\alpha; P_T, P'_T) = \mathbf{E}_Q Y(\alpha) = \mathbf{E}_Q\{N_T(\alpha)\mathcal{E}_T(-h(\alpha))\}. \tag{2.2}$$

2.4. Recall also the following explicit expression for $h(\alpha)$: if $Q = (P + P')/2$, then by (Jacod and Shiryaev, 1987, Section IV.1, Corollary 1.35)

$$h(\alpha) = 2\alpha(1 - \alpha)X_-^{-2} \circ \langle z^c \rangle + \psi_\alpha(\lambda, \lambda') \star \nu^{z,Q},$$

where $\psi_\alpha(x, y) = \alpha x + (1 - \alpha)y - x^\alpha y^{1-\alpha}$ (see 2.2 for notations).

2.5. As in (Dzhaparidze and Valkeila, 1989) (where more details can be found), for $p \geqslant 2$ we introduce the following divergency process $k(p) = k(p; P, P')$ related to the discontinuous part of z only: $k(p) = |(\lambda)^{1/p} - (\lambda')^{1/p}|^p \star \nu^{z,Q}$. It exists since

$k(p) \leqslant 2h(1/2)$, and it is free of a measure Q. Note that in case of an even $p > 2$ the process $k(p)$ is related to the Hellinger processes $h(\alpha)$ as follows:

$$k(p) = -\sum_{j=1}^{p-1} \binom{p}{j} (-1)^j h(j/p).$$

2.6. Here and in the next section we will evaluate from above the Hellinger integral $H(\alpha; P_T, P'_T)$ in terms of the Hellinger process $h(\alpha; P, P')$. First we present an inequality due to (Valkeila and Vostrikova, 1987), and give a simple alternative proof for an arbitrary $\alpha \in (0, 1)$.

LEMMA 2.1.   *For any $\mathcal{F}$-stopping time T and constant $c > 0$, and each $\alpha \in (0, 1)$*

$$H(\alpha; P_T, P'_T) \leqslant e^{-c} + P^\alpha \{ (h_T(\alpha; P, P') < c) \}. \tag{2.3}$$

Proof. Clearly, $\mathcal{E}(-h(\alpha)) \leqslant e^{-c} + \mathcal{E}(-h(\alpha)) I\{\mathcal{E}(-h(\alpha)) > e^{-c}\}$ since $x - 1 \leqslant \leqslant xI(x \geqslant 1)$ for $x \geqslant 0$. Hence, by (2.1) and (2.2)

$$H(\alpha; P_T, P'_T) \leqslant e^{-c} \mathbf{E}_Q N_T(\alpha) + \mathbf{E}_Q(Y_T(\alpha) I\{\mathcal{E}(-h_T(\alpha; P, P')) > e^{-c}\}) \leqslant$$
$$\leqslant e^{-c} + \mathbf{E}_Q(Y_T(\alpha) I\{h_T(\alpha; P, P') < c\}),$$

as $\mathbf{E}_Q N_T(\alpha) \leqslant 1$ and $\mathcal{E}(-h(\alpha)) \leqslant \exp\{-h(\alpha)\}$; see (Jacod and Shiryaev, 1987, Section V.4). Finally, apply Hölder's inequality with exponents $1/\alpha$ and $1/(1 - \alpha)$ to the second term on the right-hand side of the last inequality: this gives the desired estimate, since $\mathbf{E}_Q z'_T \leqslant 1$ and

$$\mathbf{E}_Q(Y_T(\alpha) I\{h_T(\alpha; P, P') < c\}) = \mathbf{E}_P((z'_T/z_T)^{1-\alpha} I\{h_T(\alpha; P, P') < c\}) \leqslant$$
$$\leqslant \mathbf{E}_Q(z'_T)^{1-\alpha} P^\alpha \{h(\alpha; P_T, P'_T) < c\}.$$

2.7. As an alternative to Lemma 2.1 we give below Corollary 2.1 to the following

LEMMA 2.2.   *For any $\alpha \in]0, 1[$ and stopping time T we have*

$$H(\alpha; P_T, P'_T) \leqslant \left(\mathbf{E}_P \mathcal{E}^{1/\alpha}(-h_T(\alpha; P, P')) \right)^{\alpha/2}. \tag{2.4}$$

Remark 2.1.   Compare with a result from (Kabanov et al., 1980): for any stopping time T and any pair $\alpha < \beta$ we have

$$H(\beta; P_T, P'_T) \leqslant \{ \mathbf{E}_P(\mathcal{E}_T(-h(\alpha; P, P')))^{(1-\beta)/(\beta-\alpha)} \}^{(\beta-\alpha)/(1-\alpha)}$$

(see (Jacod and Shiryaev, 1987), p. 278).

Proof. By (2.1)

$$Y(\alpha) = z^{\alpha/2} \sqrt{\mathcal{E}(-h(\alpha))} z'^{(1-\alpha)/2} \sqrt{N(\alpha)}.$$

Apply now Hölder's inequality to the right–hand side of the above inequality with exponents $p = 2/\alpha$ and $q = 2/(2 - \alpha)$ to get

$$H(\alpha; P_T, P'_T) \leqslant (\mathbf{E}_P(\mathcal{E}_T(-h(\alpha)))^{1/\alpha})^{\alpha/2}(\mathbf{E}_Q(N_T(\alpha)^{1/(2-\alpha)} z'^{(1-\alpha)/(2-\alpha)}_T))^{(2-\alpha)/\alpha}.$$

It remains to verify that

$$\mathbf{E}_Q(N_T(\alpha)^{1/(2-\alpha)} z'^{(1-\alpha)/(2-\alpha)}_T) \leqslant 1. \tag{2.5}$$

To see this apply again Hölder's inequality to the left–hand side of the above inequality with exponents $p = (2 - \alpha)/(1 - \alpha)$ and $q = 2 - \alpha$. Since $\mathbf{E}_Q z'_T = 1$ and $\mathbf{E}_Q N_T(\alpha) \leqslant 1$, this gives (2.5).

COROLLARY 2.1.   *For any stopping time T and constant c we have*

$$H(\alpha; P_T, P'_T) \leqslant e^{-c/2} + P^{\alpha/2}(h_T(\alpha; P, P') < c). \tag{2.6}$$

Proof.   Apply the inequality $\mathcal{E}^{1/\alpha}(-h(\alpha)) \leqslant e^{-h(\alpha)/\alpha}$, along with

$$\mathbf{E}e^{-X} \leqslant e^{-c} + \mathbf{P}(X < c) \tag{2.7}$$

and the inequality $(a + b)^\alpha \leqslant a^\alpha + b^\alpha$ to get (2.6) from (2.4).

2.8.   In (Dzhaparidze and Valkeila, 1989) the following Hellinger-type distances between stopped measures P_T and P'_T are studied:

$$\rho_p(P_T, P'_T) = \{\mathbf{E}_Q|(z_T)^{1/p} - (z'_T)^{1/p}|^p\}^{1/p},$$

where $p \geqslant 2$. It has been shown that in the simplest case of the usual Hellinger distance ρ_2 we have

$$\rho_2(P_T, P'_T) \leqslant \sqrt{8\mathbf{E}_P\{h_T(P, P')\}},$$

with $h(P, P') = h(1/2; P, P')$, and, generally, the following assertion holds:

LEMMA 2.3.   *For any even integer $p \geqslant 2$ there is a constant $C_p > 0$ such that*

$$\rho_p^p(P_T, P'_T) \leqslant C_p \mathbf{E}_P\{h_T(P, P')^{p/2} + k_T(p; P, P')\}.$$

Proof.   See (Dzhaparidze and Valkeila, 1989).
As it is shown in the same paper, for deterministic processes h and $k(p)$ the following inequalities hold:
for any $p > 2$ and any fixed t we have that

$$\rho_p^p(P_t, P'_t) \leqslant C_p\{h_t^{p/2} + k_t(p)\} \tag{2.8}$$

and

$$\rho_p^p(P_t, P'_t) \geqslant c(p; h(t))\{h_t^{p/2} + k(p)\}, \tag{2.9}$$

where the constant in (2.9) is $c(p; h^n(t)) = c_p(\mathcal{E}_t(-h^n))^{2p}$, and so $c(p; h^n(t)) \to$ $\to c_p > 0$ as $h^n(t) \to 0$.

2.9. Finally, we mention the following relation between the Hellinger–type distances:

$$\rho_2(P, P') \leqslant C(p)\rho_p(P, P') \qquad (2.10)$$

valid for $p \geqslant 2$ (see (Dzhaparidze and Valkeila, 1989)).

3. LARGE DEVIATION PROBABILITIES

3.1. Consider a sequence of parametric families of experiments, indexed by a set Θ in $\mathbf{R}^d$ as in 1.1, and relate to it the density process as in 1.2.
As is said in 1.4, our statement will be uniform in $\theta \in \Theta$. It will be valid for a certain normalizing sequence of positive definite matrices depending usually on θ, i.e. $\varphi_n = \varphi_n(\theta)$, $n = 1, 2, \ldots$.
Denote $U_n(\theta) = \varphi_n^{-1}(\theta)(\Theta - \theta)$ and $\Gamma_{n,R}(\theta) = \overline{U_n(\theta)} \cap \{u : R \leqslant |u| \leqslant R + 1\}$.

3.2. For all n and R large enough, and t fixed, consider the following three conditions concerning the Hellinger processes $h_t(P_\theta^n, P_{\theta'}^n)$ and the divergency processes $k_t(P_\theta^n, P_{\theta'}^n)$:
Condition I. Given Θ in $\mathbf{R}^d$, there corresponds to it a function g_n on $[0, \infty[$, introduced in 1.3, such that

$$P_\theta^n\{h_t^n(\theta, \theta + \varphi_n u) < Cg_n(|u|)\} \leqslant \exp\{-cg_n(R)\} \qquad (3.1)$$

for each $\theta \in \Theta$ and $u \in \Gamma_{n,R}(\theta)$, where c and C are certain positive constants.
The following is an alternative to Condition I.
Condition I'. Given Θ in $\mathbf{R}^d$, there corresponds to it a function g_n on $[0, \infty[$, introduced in 1.3, such that

$$\mathbf{E}_\theta^n \exp\{-h_t^n(\theta, \theta + \varphi_n u)\} \leqslant \exp\{-cg_n(R)\}, \qquad (3.2)$$

for each $\theta \in \Theta$ and $u \in \Gamma_{n,R}(\theta)$, where c is a certain positive constant.

LEMMA 3.1. *We have*

$$\text{Condition I} \Rightarrow \text{Condition I}'$$

and if the function g_n in (3.2) is such that $g_n(R+1)/g_n(R) \leqslant \gamma$ for some $\gamma > 0$ for all R large enough, then

$$\text{Condition I}' \Rightarrow \text{Condition I}.$$

Proof. Assume first that (3.1) holds. Then apply (2.7) with $c = Cg_n(|u|)$ to guarantee Condition I with $g_n'(R) = (c \wedge C)g_n(R) - \log 2$.
Assume that (3.2) holds with some c'. For a non–negative random variable X and a non-negative constant a we have

$$\mathbf{P}(X < a) \leqslant e^a \mathbf{E}e^{-X}.$$

Taking here $a = Cg_n(|u|)$ with some $C < c'/2\gamma$ we get (3.1) with $c = c'/2$.
Condition II. Given Θ in $\mathbf{R}^d$, there corresponds to it numbers p and γ, $p \geqslant \gamma > d$, p is even, and a polynomial $\mathrm{pol}_\Theta(R)$ in R such that for $|u| < \delta$

$$\mathbf{E}_\theta^n \{ h_t^n(\theta, \theta + \varphi_n u)^{p/2} + k_t^n(p; \theta, \theta + \varphi_n u) \} \leqslant C(\theta)|u|^\gamma \qquad (3.3)$$

for each $\theta \in \Theta$, where $C(\theta)$ is a certain positive function of θ admitting the following estimate for each $\theta \in \Theta$ and $u \in \Gamma_{n,R}(\theta)$: $C(\theta + \varphi_n u) \leqslant \mathrm{pol}_\Theta(R)$.

3.3. THEOREM 3.1. *Under Conditions* I *and* II *or* I' *and* II, *there exist positive constants* b *and* B *such that for all* n *and* N *large enough*

$$\sup_{\theta \in \Theta} P_\theta^n \{ |\varphi_n^{-1}(\theta_t^n - \theta)| > N \} \leqslant B \exp\{-bg_n(N)\}. \qquad (3.4)$$

Moreover, the constant b *can be made arbitrary close from below to* $(\gamma - d)/(\gamma - d + pd)$ *by choosing* B *large enough.*

Proof. Denote $Z_t^n(\theta, \theta + \varphi_n u) = z_t^n(\theta + \varphi_n u)/z_t^n(\theta)$ with $\theta \in \Theta$ and $\theta + \varphi_n u \in \Theta$. To prove (3.4) it suffices, according to (Sieders and Dzhaparidze, 1987), to check the following two conditions for n and R large enough:

(i) $$P_\theta^n \{ Z_t^n(\theta, \theta + \varphi_n u) \geqslant \varepsilon \} \leqslant \eta \exp\{-g_n(R)\}$$

for each $\theta \in \Theta$ and each $u \in \Gamma_{n,R}(\theta)$, where $0 < \varepsilon < 1$ and $\eta > 0$ are some constants, and

(ii) $$\rho_p^p(P_{t,\theta+\varphi_n u}^n, P_{t,\theta+\varphi_n v}^n) \leqslant |u - v|^\gamma \mathrm{pol}_\Theta(R)$$

for each $\theta \in \Theta$ and each $u, v \in \Gamma_{n,R}(\theta)$ such that $|u - v| < \delta$.
First we will check condition (i). Assume that Condition I holds. Fix an $\varepsilon \in (0,1)$ and note that by Markov's inequality and the definition of the Hellinger integral $(H(1/2; P, P') = H(P, P'))$ we have

$$P_\theta^n \{ Z_t^n(\theta, \theta + \varphi_n u) \geqslant \varepsilon \} \leqslant \varepsilon^{-1/2} H(P_{t,\theta}^n, P_{t,\theta+\varphi_n u}^n).$$

Apply now Lemma 2.1 with $\alpha = 1/2$ and $c \doteq g_n(u)$. Then by the last inequality and Condition I we get (i) with $\eta = 2\varepsilon^{-1/2}$.
To deduce from the same inequality that Condition I' implies (i), we extend it by using first (2.4) with $\alpha = 1/2$ then the inequality $\mathcal{E}(-h) \leqslant \exp(-h)$.
We show now that Condition II implies (ii). By Lemma 2.3 we have that

$$\rho_p^p(P_{t,\theta+\varphi_n u}^n, P_{t,\theta+\varphi_n v}^n) \leqslant C_p \mathbf{E}_{\theta+\varphi_n u}^n \{ h_t(\theta + \varphi_n u, \theta + \varphi_n v)^{p/2} +$$

$$+ k_t(p; \theta + \varphi_n u, \theta + \varphi_n v) \}.$$

Due to Condition II the right-hand side of this inequality is bounded by

$$C(\theta + \varphi_n u)|u - v|^\gamma \leqslant \mathrm{pol}_\Theta(R)|u - v|^\gamma.$$

Hence (ii) is satisfied. The assertion conserning the constant b is proved in (Sieders and Dzhaparidze, 1987).

3.4. We specify the conditions I, I' and II in the case of deterministic processes h^n and $k^n(p)$. Note meanwhile that by the inequality $\log(EX) \geqslant E\log(X)$ it is easily seen that (3.2) implies

$$\mathbf{E}_\theta^n(h_t^n(\theta, \theta + \varphi_n u)) \geqslant cg_n(R) \tag{3.5}$$

for each $\theta \in \Theta$ and $u \in \Gamma_{n,R}(\theta)$. The inverse implication is not true in general, except for deterministic processes h^n. Therefore introduce the following condition:
Condition I(d). Let the Hellinger processes h^n be deterministic satisfying

$$h_t^n(\theta, \theta + \varphi_n u) \geqslant g_n(R)$$

for each $\theta \in \Theta$ and $u \in \Gamma_{n,R}(\theta)$, with c and $g_n(R)$ as in Condition I'.
Condition II is specified as follows.
Condition II(d). Let h^n and k^n be deterministic processes. Given Θ in $\mathbf{R}^d$, there corresponds to it numbers p and γ, $p \geqslant \gamma > d$ and a polynomial $\mathrm{pol}_\Theta(R)$ in R such that for $|u| < \delta$

$$h_t^n(\theta, \theta + \varphi_n u) \leqslant C(\theta)|u|^{2\gamma/p}$$

and

$$k_t^n(p; \theta, \theta + \varphi_n u) \leqslant C(\theta)|u|^\gamma$$

for each $\theta \in \Theta$, where $C(\theta)$ is as in Condition II.

COROLLARY 3.1. *Suppose that the processes $h_t^n(\theta, \theta + \varphi_n u)$ and $k_t^n(p; \theta, \theta + \varphi_n u)$ are deterministic, satisfying Conditions* I(d) *and* II(d). *Then the conclusion of Theorem 3.1 holds.*

REFERENCES

Dzhaparidze, K. and Valkeila, E. (1989). On Hellinger type distances for filtered experiments. *Probab. Theory Rel. Fields* (to appear).

Ibragimov, I.A. and Has'minskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory.* Springer, Berlin–Heidelberg etc.

Jacod, J. and Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin–Heidelberg etc.

Kabanov, Yu.M., Liptser, R.Sh. and Shiryaev, A.N. (1980). Some limit theorems for simple point processes (a martingale approach). *Stochastics* **3**, 203–216.

Kutoyants, Yu.A. (1984). *Parameter Estimation for Stochastic Processes.* Heldermann, Berlin.

Sieders, A. and Dzhaparidze, K. (1987). A large deviation result for parameter estimators and its application to nonlinear regression analysis. *Ann. Statist.* **15**, 1031–1049.

Valkeila, E. and Vostrikova, L. (1987). On predictable criteria of (c_n)–consistency of estimates. *Theory Probab. Appl.* **32**, 477–489.