Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

R.J. Chitashvili, N.L. Lazrieva, T.A. Toronjadze

Asymptotic theory of M-estimators in general statistical models
Part I: On asymptotic behaviour of estimators under model disturbance

Department of Operations Research, Statistics, and System Theory Report BS-R9019 June

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Asymptotic Theory of M-Estimators in General Statistical Models

Part I: On asymptotic behaviour of estimators under model disturbance

R.J. Chitashvili, N.L. Lazrieva, T.A. Toronjadze
Department of Probability Theory and Mathematical Statistics
Tbilisi Mathematical Institute, Academy of Sciences of the Georgian SSR
150 a, Agmashenebeli Avenue, Tbilisi 380012, USSR

The asymptotic behaviour of the solutions to the equation for M-estimators is considered under model disturbance in a general scheme of statistical models.

1980 Mathematics Subject Classification: 60G44, 62F12.
Key Words & Phrases: semimartingales, statistical estimation.

Note: To appear in Transactions of Tbilisi Mathematical Institute, 1989. This work was completed while the first author was visiting the Centre for Mathematics and Computer Science, Amsterdam.

The paper deals with asymptotics of parameter estimators in a general scheme of statistical models defined on a given probability space with a filtration. On studying the asymptotic behaviour of estimators, defined as solutions certain equations (M-estimators), we systematically utilize the representation of likelihood ratios in the form of exponential martingales, useful for application of martingale limit theorems.

The paper is organized as follows: in § 1 main objects and notions are introduced; in § 2 a theorem is proved concerning asymptotic properties of M-estimators; in § 3 similar problems for alternative (perturbed) models are studied; in § 4 several examples of particular schemes of models (independent observations, diffusion processes, point processes) are given. Further, § 5 is devoted to the investigation of contigual alternatives. In § 6 the global asymptotic behaviour of M-estimators is studied. And, finally, in the last section (§ 7) some references and short historical notes are given.

Throughout this paper we use the conventional notations of martingale theory (see, e.g., [1,2,3]).

1. Model of an experiment, L-transformations, regularity conditions, regular M-estimators

1.1. Consider a sequence of statistical models [1)]

$$\mathscr{E} = (\mathscr{E}_n)_{n\geq 1} = (\Omega^n, \mathscr{F}^n, F^n, P_\theta^n, P^n)_{n\geq 1},\ \theta\in\Theta\subset R^1,$$

where $(\Omega^n, \mathscr{F}^n, F^n, P^n)$ is a stochastic basis for each $n\geq 1$, i.e. a probability space with a filtration $F^n = (\mathscr{F}_t^n)_{0\leq t\leq T}$ satisfying the usual conditions. Here $(\mathscr{F}_t^n)_{0\leq t\leq T}$ is a right-continuous increasing family of σ-subalgebras of $\mathscr{F}^n$ augmented by sets of zero P^n-measure; Θ is an open subset of R^1; the measures P_θ^n and P^n are equivalent for all $n\geq 1$ and $\theta\in\Theta$.

Let $P_\theta^n(t) = P_\theta^n \mid \mathscr{F}_t^n$ and $P^n(t) = P^n \mid \mathscr{F}_t^n$ be restrictions of the measures P_θ^n and P^n to the σ-algebra $\mathscr{F}_t^n$ and let $\rho_\theta^n = (\rho_\theta^n(t))_{0\leq t\leq T}$ denote the likelihood ratio process

$$\rho_\theta^n(t) = \frac{dP_\theta^n(t)}{dP^n(t)}.$$

1) For brevity, a sequence $\mathscr{E} = (\mathscr{E}_n)_{n\geq 1}$ will be sometimes called an experiment or a scheme of experiments.

Report BS-R9019
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

For convenience we assume $\rho_\theta^n(0) = 1$.

As is known, there exists a local P^n-martingale $M_\theta^n = (M_\theta^n(t))_{0<t<T}$ such that ρ_θ^n, satisfying Dolean's equation

$$d\rho_\theta^n(t) = \rho_\theta^n(t-)dM_\theta^n(t),\ \rho_\theta^n(0) = 1, \tag{1.1}$$

can be represented as an exponential martingale

$$\rho_\theta^n = \mathcal{E}(M_\theta^n) := \exp(M_\theta^n - \frac{1}{2}<M_\theta^{n,c}>)\Pi(1+\Delta M_\theta^n)e^{-\Delta M_t^n}, \tag{1.2}$$

where M^c is a continuous part of M, $<M>$ its square characteristic, and $\Delta M = M - M_-$ its jump.

1.2. We shall need the martingale transformation formula under an absolutely continuous change of a measure. Let $(\Omega,\mathcal{F},F,P)$ be some stochastic bases, and Q some probability on $\mathcal{F}$ absolutely continuous w.r.t. $P(Q<<P)$. If $m \in \mathcal{M}_{loc}(F,P)$ [1] and M is a P-martingale such that $dQ/dP = \mathcal{E}(M)$, then by the general Girsanov theorem [1] the process (L-transformation)

$$L(m,M) := m - <m^c,M^c> - \Sigma\frac{\Delta m \Delta M}{1+\Delta M} \tag{1.3}$$

is a Q-local martingale ($L(m,M) \in \mathcal{M}_{loc}(F,Q)$).

1.3. We call an experiment $\mathcal{E}$ regular if the following set of conditions (R) holds:

(R1) For any $n \geq 1$ and $\theta \in \Theta$ we have $M_\theta^n \in \mathcal{M}^2(F^n,P^n)$ (where as usual, $\mathcal{M}^2(F,P)$ is a class of square integrable martingales w.r.t. a flow F and a measure P).

(R2) For any $n \geq 1$ the mapping $M^n : \Theta \to \mathcal{M}^2(F^n,P^n)$ is twice continuously differentiable w.r.t. θ in the sense of the norm $\|\cdot\|$, and the second derivative $\ddot{M}_\theta^n$ for all $t \in [0,T]$ is continuous w.r.t. θ P^n-a.s. (If $m \in \mathcal{M}^2(F,P)$, then $\|M\| = E^{\frac{1}{2}} <M>_T = E^{\frac{1}{2}} M_T^2$. Everywhere below the dot means the derivative w.r.t. the parameter θ, e.g. $\dot{M}_\theta^n = \frac{\partial}{\partial\theta}M_\theta^n$, $\dot{L}(m_\theta^n,M_\theta^n) = \frac{\partial}{\partial\theta}L(m_\theta^n,M_\theta^n)$ etc.)

(R3) For all $n \geq 1$ the following series are termwise continuously differentiable:

$$\Sigma(ln(1+\Delta M_\theta^n) - \Delta M_\theta^n),\ \Sigma\frac{\Delta \dot{M}_\theta^n \Delta M_\theta^n}{1+\Delta M_\theta^n};$$

besides

$$E_\theta^n \Sigma(\frac{\Delta \dot{M}_\theta^n}{1+\Delta M_\theta^n})^2 < \infty,\ E_\theta^n \Sigma(\frac{\Delta \ddot{M}_\theta^n}{1+\Delta M_\theta^n})^2 < \infty.$$

(R4) Fisher's information

$$I_\theta^n := E_\theta^n <L(\dot{M}_\theta^n,M_\theta^n)>_T$$

is finite, positive and θ-continuous for all $n \geq 1$.

We present now a number of direct implications of the regularity conditions.

PROPOSITION 1.1. *Under conditions* (R)

a) *The log likelihood* $\ln\rho_\theta^n$ *is continuously differentiable w.r.t.* θ *for all* t, P^n*-a.s. for all* ω, *and*

$$\frac{\partial}{\partial\theta}\ln\rho_\theta^n = L(\dot{M}_\theta^n,M_\theta^n). \tag{1.4}$$

b) *At each* t *the transformations* $L(\dot{M}_\theta^n,M_\theta^n)$ *and* $L(\ddot{M}_\theta^n,M_\theta^n)$ *are continuous w.r.t.* θ P^n*-a.s. for all* ω,

1) $\mathcal{M}_{loc}(F,P)$ is a class of local martingales w.r.t. a filtration F and a measure P.

$$L(\dot{M}_\theta^n, M_\theta^n) \in \mathfrak{M}^2(F^n, P_\theta^n),$$
$$L(\ddot{M}_\theta^n, M_\theta^n) \in \mathfrak{M}^2(F^n, P^n).$$

c)

$$\dot{L}(\dot{M}_\theta^n, M_\theta^n) = L(\ddot{M}_\theta^n, M_\theta^n) - [L(\dot{M}_\theta^n, M_\theta^n)], \tag{1.5}$$

where [.] is the square variation.

Proof. The formal differentiation of

$$\ln\rho_\theta^n = M_\theta^n - \frac{1}{2} <M_\theta^{n,c}> + \Sigma(ln(1+\Delta M_\theta^n) - \Delta M_\theta^n)$$

yields

$$\frac{\partial}{\partial\theta}\ln\rho_\theta^n = \dot{M}_\theta^n - \frac{1}{2}\frac{\partial}{\partial\theta} <M_\theta^{n,c}> - \Sigma\frac{\Delta\dot{M}_\theta^n \Delta M_\theta^n}{1+\Delta M_\theta^n}. \tag{1.6}$$

Now assertion a) follows directly from the estimation

$$\Sigma\left|\frac{\Delta\dot{M}_\theta^n \Delta M_\theta^n}{1+\Delta M_\theta^n}\right| \leq [\Sigma(\frac{\Delta\dot{M}_\theta^n}{1+\Delta M_\theta^n})^2 \Sigma(\Delta M_\theta^n)^2]^{\frac{1}{2}} < \infty,$$

provided

$$\frac{\partial}{\partial\theta} <M_\theta> = 2 <\dot{M}_\theta, M_\theta> \tag{1.7}$$

for any $M_\theta \in \mathfrak{M}^2(F,P)$. The last equality is easily proved. Indeed, by the definition of the square characteristic the processes $m_\theta = M_\theta^2 - <M_\theta, M_\theta>$ and $2\dot{M}_\theta M_\theta - 2<\dot{M}_\theta, M_\theta>$ are martingales as well as the process $\dot{m}_\theta = 2\dot{M}_\theta M_\theta - \frac{\partial}{\partial\theta}<M_\theta>$. Hence, (1.7) follows from the uniqueness of a mutual square characteristic.

By the general Girsanov theorem [2] the continuous component and the jump of a purely discontinuous component of the L-transformation $L(m,M)$ have the form

$$L^c(m,M) = m^c - <m^c, M^c>, \quad \Delta L^d(m,M) = \frac{\Delta m}{1+\Delta M} \tag{1.8}$$

with $<L^c(m,M)> = <m^c>$. Therefore assertion b) follows directly from conditions (R2) and (R3).

Finally, by the formal differentiation of the expression $L(\dot{M}_\theta^n, M_\theta^n)$ we obtain assertion c):

$$\frac{\partial}{\partial\theta}L(\dot{M}_\theta^n, M_\theta^n) = \ddot{M}_\theta^n - <\ddot{M}_\theta^{n,c}, M_\theta^{n,c}> - \Sigma\frac{\Delta\ddot{M}_\theta^n \Delta M_\theta^n}{1+\Delta M_\theta^n} - <\dot{M}_\theta^{n,c}, \dot{M}_\theta^{n,c}> - \Sigma(\frac{\Delta\dot{M}_\theta^n}{1+\Delta M_\theta^n})^2$$
$$= L(\ddot{M}_\theta^n, M_\theta^n) - [L(\dot{M}_\theta^n, M_\theta^n)]. \quad \square$$

Note that the finiteness of the Fisher information in condition (R4) is ensured by conditions (R2) and (R3), since by (1.8)

$$I_\theta^n = E_\theta^n <L(\dot{M}_\theta^n, M_\theta^n)>_T = E_\theta^n <\dot{M}_\theta^{n,c}>_T + E_\theta^n \sum_{s \leq T}(\frac{\Delta\dot{M}_\theta^n(s)}{1+\Delta M_\theta^n(s)})^2. \tag{1.9}$$

1.4. In view of (1.4), the maximum likelihood equation (MLE) may be written in terms of an L-transformation, namely in the following form:

$$L_T(\dot{M}_\theta^n, M_\theta^n) = 0. \tag{1.10}$$

This form suggests a natural extension: define the estimator as a solution to the equation

$$L_T(m_\theta^n, M_\theta^n) = 0 \tag{1.11}$$

where $m_\theta^n, n \geq 1$ is some sequence of θ-dependent P^n-martingales. Similarly to the classical schemes (see the examples below) such estimators will be called M-estimators.

Under regularity of the experiment we can naturally restrict ourselves to M-estimators corresponding to the martingale m_θ^n (which will be sometimes called the martingale defining the M-estimator) with properties similar to those of $\dot{M}_\theta^n$.

A sequence of martingales defining the M-estimator will be called regular if the following set of conditions holds (R_m):

$(R_m 1)$ For any $n \geq 1$ and $\theta \in \Theta$ we have $m_\theta^n \in \mathfrak{M}^2(F^n, P^n)$.

$(R_m 2)$ For any $n \geq 1$ the mapping $m^n : \Theta \to \mathfrak{M}^2(F^n, P^n)$
is continuously differentiable in the sense of the norm $\|\cdot\|$ and for all $t \geq 0$ $\dot{m}_\theta^n$ is continuous P^n-a.s.

$(R_m 3)$ For all $n \geq 1$ the following series are termwise continously differentiable:

$$\Sigma \frac{\Delta m_\theta^n \Delta M_\theta^n}{1+\Delta M_\theta^n};$$

besides

$$E_\theta^n \Sigma\left(\frac{\Delta m_\theta^n}{1+\Delta M_\theta^n}\right)^2 < \infty, \quad E_\theta^n \Sigma\left(\frac{\Delta \dot{m}_\theta^n}{1+\Delta M_\theta^n}\right)^2 < \infty.$$

$(R_m 4)$ The quantity

$$I_\theta^n(m) := E_\theta^n < L(\dot{M}_\theta^n, M_\theta^n), L(m_\theta^n, M_\theta^n) >_T$$

is finite, positive and θ-continuous for all $n \geq 1$.

PROPOSITION 1.2. *Under conditions (R_m) we have*

b_m) *For all t the L-transformations $L(m_\theta^n, M_\theta^n)$ and $L(\dot{m}_\theta^n, M_\theta^n)$ are continuous in θ, P^n-a.s. and*

$$L(m_\theta^n, M_\theta^n), L(\dot{m}_\theta^n, M_\theta^n) \in \mathfrak{M}^2(F^n, P_\theta^n)$$

$$c_m) \qquad \dot{L}(m_\theta^n, M_\theta^n) = L(\dot{m}_\theta^n, M_\theta^n) - [L(\dot{M}_\theta^n, M_\theta^n), L(m_\theta^n, M_\theta^n)] \tag{1.12}$$

The proof is similar to that of Proposition 1.1. Note that the quantity $I_\theta^n(m)$ is finite due to conditions $R_m 1)$ and $R_m 2)$, and it can be written in the following explicit form

$$I_\theta^n(m) = E_\theta^n < m_\theta^{n,c}, \dot{M}_\theta^{n,c} >_T + E_\theta^n \sum_{s \leq T} \frac{\Delta m_\theta^n(s) \Delta \dot{M}_\theta^n(s)}{1+\Delta M_\theta^n(s)}. \tag{1.13}$$

A class of regular martingales defining M-estimator is denoted by $\mathfrak{M}_R$.

1.5. In the assertions to the forthcoming sections concerning asymptotic properties of the solutions to (1.10) and (1.11) it is convenient to express the conditions imposed on the sequences of martingales $(M_\theta^n)_{n \geq 1}$ (defining the model) and $(m_\theta^n)_{n \geq 1}$ (defining the estimators) in terms of L-transformations or the corresponding predictable characteristics. In this connection it is useful to present formulas giving explicit expressions for the characteristics of L-transformations by means of the characteristics of the martingales M_θ^n and m_θ^n. From (1.8) we can obtain the following facts for P-local martingales m, m', M, M':

$$1^\circ. \qquad [L(m,M), L(m',M')] = < m^c, m'^c > + \Sigma \frac{\Delta m \Delta m'}{(1+\Delta M)(1+\Delta M')}. \tag{1.14}$$

2°. If ν^L denotes the P-compensator of the jump measure of the process $L(m,M)$, then (for any integrable function ψ) we have

$$\int_0^T\int_E \psi(s,x)\nu^L(ds,dx) = (\Sigma\psi(s,\frac{\Delta m(s)}{1+\Delta M(s)})I_{\{\frac{\Delta m(s)}{1+\Delta M(s)}\neq 0\}})^{p,P} \tag{1.15}$$

where $\{\cdot\}^{p,P}$ denotes a dual predictable projection w.r.t. P.
3°. If P-martingales m and M allow integral representations

$$M_t = \int_0^t\int_E g(s,x)(\mu-\nu)(ds,dx)+\int_0^t g(s)dn_s, \tag{1.16}$$

$$m_t = \int_0^t\int_E \psi(s,x)(\mu-\nu)ds,dx)+\int_0^t \psi(s)dn_s,$$

where n is a continuous P-martingale, μ an integer-valued random measure on the product $[0,T]\times E$ and ν its P-compensator, then for the measure $\tilde{P}$ with $d\tilde{P}/dP = \mathcal{E}(M)$ we have

$$L_t(m,M) = \int_0^t\int_E \phi(s,x)(\mu-\tilde{\nu})(ds,dx)+\int_0^t \psi(s)(dn_s-g(s)d<n>_s), \tag{1.17}$$

where $\tilde{\nu}$ is the $\tilde{P}$-compensator of the measure μ and $\phi(s,x)$ has the form

$$\phi(s,x) = \frac{\psi(s,x)-\hat{\psi}(s)}{1+g(s,x)-\hat{g}(s)} + \frac{\hat{\psi}(s)}{1-\hat{g}(s)} \tag{1.18}$$

with

$$\hat{\psi}(s) = \int_E \psi(s,x)\nu(\{s\},dx),\ \ \hat{g}(s) = \int_E g(s,x)\nu(\{s\},dx).$$

Indeed the validity of (1.17) easily follows from the fact that, on the one hand,

$$\Delta L^d(m,M) = \frac{\psi(s,\beta_s)I_D-\hat{\psi}(s)}{1+g(s,\beta_s)I_D-\hat{g}(s)}$$

where

$$D = \{(s,\omega):\mu(\{s\},E) = 1\}$$

with Diraque's measure $\mu(\{s\},dx) = \epsilon_{\beta_s}(dx)$ and, on the other hand,

$$\Delta(\int_0\int_E \phi(s,x)(\mu-\tilde{\nu})(ds,dx)) = \phi(s,\beta_s)I_D-\tilde{\hat{\phi}}(s), \tag{1.19}$$

where

$$\tilde{\hat{\phi}}(s) = \int_E \phi(s,x)\tilde{\nu}(\{s\},dx) = \frac{\hat{\psi}(s)}{1-\hat{g}(s)}. \tag{1.20}$$

Now, substituting (1.18) and (1.20) into (1.19) we obtain that

$$\Delta L^d(m,M) = \Delta[\phi*(\mu-\tilde{\nu})] \tag{1.21}$$

and, hence, the purely discontinuous part of the martingale $L(m,M)$ is equal to $\phi*(\mu-\tilde{\nu})$. □

4°. If m' is a martingale allowing the integral representation (1.16) with $\psi'(s,x),\psi'(s)$ instead of $\psi(s,x),\psi(s)$, then

$$<L(m,M),L(m',M)> = \phi\phi'*\nu-\Sigma\frac{\hat{\phi}\hat{\phi}'}{1-\hat{g}}(1-a)+\int g\psi' d<n>,$$

where $a_s = \nu(\{s\},E)$ and ϕ' is defined by (1.18) with ψ' instead of ψ.

5°. Let

$$P \sim Q \sim \tilde{P}, Q = \mathcal{E}(M)\cdot P, \quad \tilde{P} = \mathcal{E}(\tilde{M})\cdot P, \quad M,\tilde{M} \in \mathfrak{M}^2(F,P).$$

Then if m is a P-martingale the transformation of m into a $\tilde{P}$-martingale can be carried out in two ways: either directly using L-transformation, or transforming it first into a Q-martingale and then into a $\tilde{P}$-martingale. This is expressed by the following transitivity property of the L-transformation:

$$L(L(m,M),L(\tilde{M}-M,M)) = L(m,\tilde{M}). \tag{1.22}$$

Besides the expression

$$L(m,M,\tilde{M}) := L(m,M) - \langle L(m,M),L(\tilde{M}-M,M)\rangle \tag{1.23}$$

is a $\tilde{P}$-martingale and

$$L(m,M) = L(m,\tilde{M}) + [L(m,M),L(M-\tilde{M},\tilde{M})]. \tag{1.24}$$

Note also that by the uniqueness of the canonical decomposition (of the $\tilde{P}$-special semimartingale $L(m,M)$) we have [1)]

$$[L(m,M),L(M-\tilde{M},M)]^{p,\tilde{P}} = \langle L(m,M),L(\tilde{M}-M,M)\rangle^{Q}.$$

REMARK 1. If m is a P-martingale, then w.r.t. a measure Q it is a semimartingale and hence has, generally speaking, a non-unique decomposition. The L-transformation presents the Q-martingale component in one of such decompositions. Another decomposition is obtained by Girsanov's transformation

$$G(m,M) = m - \langle m,M\rangle.$$

(G-transformation defines a Q-martingale in a canonical decomposition of a P-martingale m). Note that both L and G-transformations are one-to-one, hence for every $m \in \mathfrak{M}(F,P)$ there exists a unique $\tilde{m} \in \mathfrak{M}(F,P)$ such that

$$G(m,M) = L(\tilde{m},M).$$

Analogously to (1.11), one might think of the estimational equation of type

$$G(m_\theta,M_\theta) = 0 \tag{1.25}$$

(or some other estimational equation based on a one-to-one transformation, different from L and G), but this cannot enreach the class of all M-estimators: it is easily seen that this class remains unchanged. The special role attached in this paper to L-transformation becomes transparent below. Here we only note the fact that, as it can be seen from Assertion IV of Theorem 2.1, the asymptotic efficiency of an M-estimator $\hat{\theta}$ is given as the square of the (asymptotic) correlation coefficient

$$\mathit{eff}\,\hat{\theta} = P^n_\theta - \lim_{n\to\infty} \frac{[L(m^n_\theta,M^n_\theta),L(\dot{M}^n_\theta,M^n_\theta)]^2_T}{\langle L(m^n_\theta,M^n_\theta)\rangle_T \langle L(\dot{M}^n_\theta,M^n_\theta)\rangle_T}$$

where m^n_θ is martingale envolved in the equation (1.11) defining $\hat{\theta}$.

As for the equation (1.25), it corresponds to the classical method of moments, as it easily seen by taking into consideration that the left-hand side of (1.25) presents nothing less then the martingale m_θ centered (conditionally) w.r.t. P_θ.

1) $\langle\cdot\rangle^Q$ denotes the square characteristic w.r.t. Q.

2. Asymptotic properties of M-estimators. Ergodic models and estimators

2.1. As is known the classical proof of the assertions concerning the asymptotic behaviour of M-estimators as solutions to (1.11) is carried out in two steps: firstly, the required asymptotic properties are established for the left-hand side of (1.11); secondly, the asymptotic properties of the estimators (considered as implicit functions) are obtained by linearization.

Following the same scheme, we will utilize in the second part of the proof the following sufficiently general assertion.

Let for every $\theta\in\Theta$ sequences of probability measures $\{Q_\theta^n\}_{n\geq 1}$ $(Q_\theta^n \sim P^n)$ and random variables $\{L_n(\theta)\}_{n\geq 1}$ be given, on a measurable space $(\Omega^n, \mathcal{F}_T^n)$ as well as a sequence of positive numbers $\{\phi_n(\theta)\}_{n\geq 1}$.

Lemma 2.1. *Let the following conditions hold:*

a) *For every* $\theta\in\Theta$ $\lim_{n\to\infty} \phi_n(\theta) = 0$;

b) *For every* $n\geq 1$ *the mapping* $\theta \to L_n(\theta)$ *is continuously differentiable in* θ, P^n*-a.s.*

c) *For every* $\theta\in\Theta$ *there exists a unique point* $\theta' = \theta'(\theta)$ *such that*

$$Q_\theta^n - \lim_{n\to\infty} \phi_n^2(\theta) L_n(\theta') = 0;$$

(here and elsewhere below $Q_\theta^n - \lim_{n\to\infty} \xi_n = \eta$ *denotes the convergence in* Q_θ^n*-probability:* $Q_\theta^n\{|\xi_n - \eta| > \rho\} \to 0, \forall \rho > 0.$*)*

d) $$Q_\theta^n - \lim_{n\to\infty} \phi_n^2(\theta) \dot{L}_n(\theta') = -\gamma(\theta)$$

where $\gamma(\theta)$ *is positive for every* $\theta\in\Theta$.

e) $$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}} Q_\theta^n\{\sup_{z:|\theta'-z|<\epsilon} \phi_n^2(\theta)|\dot{L}_n(z) - \dot{L}_n(\theta')| > \rho\} = 0$$

for any positive ρ.

Then for every $\theta\in\Theta$ *there exists a sequence of random variables* $\hat{\theta} = (\hat{\theta}_n)_{n\geq 1}$ *such that*

I. $\lim_{n\to\infty} Q_\theta^n\{L_n(\hat{\theta}_n) = 0\} = 1$;

II. $Q_\theta^n - \lim_{n\to\infty} \hat{\theta}_n = \theta'$;

III. *if* $\tilde{\theta}_n$ *is another sequence of random variables with properties I and II, then*

$$\lim_{n\to\infty} Q_\theta^n\{\hat{\theta}_n = \tilde{\theta}_n\} = 1$$

and, finally,

IV. *if the sequence of the distributions of* $\phi_n(\theta)L_n(\theta')$ *w.r.t.* Q_θ^n, *denoted here by*

$$\mathcal{L}_{Q_\theta^n} - \phi_n(\theta)L_n(\theta'),\ n = 1,2,...,$$

converges weakly to some distribution Φ, *i.e.*

$$\mathcal{L}_{Q_\theta^n} - \phi_n(\theta)L_n(\theta') \Rightarrow \Phi,$$

then

$$\mathcal{L}_{Q_\theta^n} - \gamma(\theta)[\phi_n^{-1}(\theta)(\hat{\theta}_n - \theta')] \Rightarrow \Phi. \tag{2.1}$$

Proof. Since we follow the classical Dugue-Cramér scheme for proving this kind of theorems (see, e.g., Le Breton [5]), we present here its outline only.

1. Using Tailor's formula we obtain the decomposition

$$\phi_n^2(\theta)L_n(z) = \phi_n^2(\theta)L_n(\theta') - (z-\theta')\gamma(\theta) + (z-\theta')\delta_n(z), \tag{2.2}$$

where the remainder term $\delta_n(z)$ is small in the sense that

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} Q_\theta^n\{\sup_{z:|z-\theta'|<\epsilon} |\delta_n(z)|>\rho\} = 0, \ \forall\rho>0.$$

2. By (2.2) a sequence $\Omega(n,\epsilon)$ of the subsets of Ω^n is constructed according to the formula

$$\Omega(n,\epsilon) = \{\omega\in\Omega^n: |\phi_n^2(\theta)L_n(\theta')| \leq \frac{\gamma(\theta)\epsilon}{2}, \sup_{z:|z-\theta'|<\epsilon} |\delta_n(z)| < \frac{\gamma(\theta)}{2}\}$$

and it is shown that

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} Q_\theta^n\{\Omega(n,\epsilon)\} = 1$$

and that for any $\omega\in\Omega(n,\epsilon)$ a unique solution θ_n^* of the equation

$$L_n(z) = 0$$

exists in the interval $|z-\theta'|<\epsilon$.

3. A set $\Omega_n = \bigcup_k \Omega(n,\frac{1}{k})$ and a measurable space

$$(\Omega_n\cap\Omega^n, \mathcal{F}_T^n\cap\Omega_n) = (\tilde{\Omega}_n, \tilde{\mathcal{F}}_T^n)$$

are considered.

Obviously, for any $\omega\in\tilde{\Omega}_n$ there exists $\kappa(\omega)>0$ such that in the interval $|z-\theta'|<1/\kappa(\omega)$ the equation $L_n(z) = 0$ has a unique $\tilde{\mathcal{F}}_T^n$-measurable[1)] solution θ_n^*. The desired solution $\{\hat{\theta}_n\}, n\geq 1$ is constructed as follows:

$$\hat{\theta}_n = \begin{cases} \theta_n^* & \omega\in\tilde{\Omega}_n \\ arbitr, & \omega\in\tilde{\Omega}^c. \end{cases}$$

4. Properties I-III of the sequence $\{\hat{\theta}_n\}, n\geq 1$ easily follow from the form of $\tilde{\Omega}_n$. We will briefly remark on Property IV.

By the decomposition (2.2) we have

$$|\phi_n(\theta)(L_n(\hat{\theta}_n)-L_n(\theta'))-\gamma(\theta)\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta')| \leq \phi_n^{-1}(\theta)|\hat{\theta}_n-\theta'||\delta_n(\hat{\theta}_n)|. \tag{2.3}$$

Denote $X_n = \phi_n(\theta)(L_n(\hat{\theta}_n)-L_n(\theta'))$, $Y_n = \gamma(\theta)\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta')$ and $z_n = \gamma(\theta)^{-1}|\delta_n(\hat{\theta}_n)|$. Then inequality (2.3) will take the form

$$|X_n - Y_n| \leq z_n|Y_n|. \tag{2.4}$$

It is known (see [6], 2, § 4, Part I) that if X_n converges to X in distribution ($X_n \xrightarrow{\mathcal{D}} X$) and (2.4) is satisfied, then $Y_n \xrightarrow{\mathcal{D}} X$. □

REMARK 2.1. We will need the following generalization of problem 2 from § 4, Part 1 of [6]. Let $X_n \xrightarrow{\mathcal{D}} X$ and let the following inequality

$$|X_n - Y_n| \leq Z_n|X_n| + U_n \tag{2.5}$$

be satisfied with $Z_n \xrightarrow{P} O$ and $U_n \xrightarrow{P} O$. Then, obviously $Y_n \xrightarrow{\mathcal{D}} X$. Now, let

$$|X_n - Y_n| \leq Z_n|Y_n| + Z_n\beta_n, \tag{2.6}$$

1) the existence of a measurable solution is the well-known fact in the function theory (see, e.g., [5])

with $Z_n \xrightarrow{P} O$ and $Z_n\beta_n \xrightarrow{P} O$. Then $Y_n \xrightarrow{\mathfrak{D}} X$. Indeed, for any numbers a,b,c and $\epsilon<\frac{1}{2}$ the inequality

$$|a-b|<\epsilon|a|+c$$

implies the inequality $|a-b|<2\epsilon|b|+\frac{1}{1-\epsilon}c$. □

REMARK 2.2. Unlike [5] we assume here the scheme of series, and we do not restrict ourselves by the special case of $\theta' = \theta$.

2.2. Specifying the function $L_n(\theta)$, $n\geq 1$ as the values of the process $L(m_\theta^n, M_\theta^n)$ at $t = T$ (with a martingale $m_\theta^n \in \mathfrak{M}_R$), replacing Q_θ^n by P_θ^n and setting $\theta'(\theta) = \theta$, the assertion of Lemma 2.1 may be reformulated as Theorem 2.1 given below.

We will specify here also the normalizing sequence as follows:

$$\phi_n(\theta) = (I_\theta^n)^{-\frac{1}{2}}.$$

THEOREM 2.1. *Let for every $\theta\in\Theta$ the following conditions hold:*

a) $\lim_{n\to\infty} \phi_n(\theta) = 0;$

b1) $P_\theta^n - \lim_{n\to\infty} \phi_n^2(\theta)<L(m_\theta^n,M_\theta^n)>_T = \Gamma(\theta);$

b2) $P_\theta^n - \lim_{n\to\infty} \phi_n^2(\theta)[L(m_\theta^n,M_\theta^n),L(\dot{M}_\theta^n,M_\theta^n)]_T = \gamma(\theta)$

where $\Gamma(\theta)$ and $\gamma(\theta)$ are positive for every $\theta\in\Theta$;

c) $P_\theta^n - \lim_{n\to\infty} \phi_n^2(\theta)L_T(\dot{m}_\theta^n,M_\theta^n) = 0;$

d) $P_\theta^n - \lim_{n\to\infty} \int_0^T \int_{|x|>\epsilon} x^2\nu_n(ds,dx) = 0,$ *for $\epsilon>0$,*

where ν_n is the compensator w.r.t. P_θ^n of the jump measure of the process $\phi_n(\theta)L(m_\theta^n,M_\theta^n)$;

e) $\lim_{\epsilon\to 0}\lim_{n\to\infty} P_\theta^n\{\sup_{z:|z-\theta|<\epsilon} \phi_n^2(\theta)|\dot{L}_T(m_\theta^n,M_\theta^n)-\dot{L}_T(m_z^n,M_z^n)|>\rho\} = 0$

for any $\rho>0$.

Then for any $\theta\in\Theta$ there exists an M-estimator $(\hat{\theta}_n),n\geq 1$ with Properties I-III given in Lemma 2.1 (with $Q_\theta^n = P_\theta^n,\theta'(\theta) = \theta$), such that

$$\mathcal{L}_{P_\theta^n} - \phi_n^{-1}(\theta)(\hat{\theta}_n-\theta)\Rightarrow N(0,\frac{\Gamma(\theta)}{\gamma^2(\theta)}).$$

PROOF. It can be easily seen that

$$\dot{L}(m_\theta^n,M_\theta^n) = L(\dot{m}_\theta^n,M_\theta^n)-[L(m_\theta^n,M_\theta^n),L(\dot{M}_\theta^n,M_\theta^n)],$$

and that by virtue of Conditions b2) and c) of the theorem this ensures Condition d) of Lemma 2.1. Further, Condition b1) of the theorem implies Condition c) of Lemma 2.1 and, finally, by virtue of the central limit theorem for martingales ([3]) Conditions b1) and a) lead to the convergence of the distributions $\mathcal{L}_{P_\theta^n} - \phi_n(\theta)L_T(m_\theta^n,M_\theta^n)$ to the normal law $N(0,\Gamma(\theta))$. □

2.3. In view of the facts in 1.5, the conditions of Theorem 2.1 can be expressed, if necessary, in terms of the characteristics of the martingales M_θ^n and m_θ^n. Moreover, the sufficient conditions expressed in this terms are even simplefied. For example, due to (1.15) Condition d) is equivalent to

d′) $P_\theta^n - \lim \phi_n^2(\theta)(\sum_{s<T}(\frac{\Delta m_\theta^n(s)}{1+\Delta M_\theta^n(s)})^2 I\{|\frac{\Delta m_\theta^n(s)}{1+\Delta M_\theta^n(s)}|>\epsilon\phi_n^{-1}(\theta)\})^{p,P_\theta^n} = 0.$

Further the following condition c′) is sufficient for c) to hold:

c′) $P_\theta^n - \lim_{n\to\infty} \phi_n^4(\theta)[<\dot{m}_\theta^{n,c}>_T+(\sum_{s<T}(\frac{\Delta\dot{m}_\theta^n(s)}{1+\Delta M_\theta^n(s)})^2)^{p,P_\theta^n}] = 0.$

This fact easily follows from (1.14) and Lenglart's inequality [3].

If we assume that the martingales M_θ^n and m_θ^n are differentiable to the third and second order, respectively, then condition e) will follow from the set e′) of stronger conditions:

e1′) $\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\, \phi_n^2(\theta)E_\theta^n\{\sup_{z:|\theta-z|\leq\epsilon}|[L(m_y^n,M_z^n),L(M_\theta^n-M_y^n,M_\theta^n)]_T|\} = 0;$

e2′) $\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\, \phi_n^2(\theta)E_\theta^n\{\sup_{z:|z-\theta|\leq\epsilon}|[L(m_z^n,M_z^n),L(\dot{M}_y^n,M_y^n)]_T-[L(m_\theta^n,M_\theta^n),L(\dot{M}_\theta^n,M_\theta^n)]_T|\} = 0;$

e3′) $\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\, \epsilon^2\phi_n^4(\theta)\sup_{z:|z-\theta|\leq\epsilon} E_\theta^n<L(\dot{m}_y^n,M_\theta^n)>_T = 0.$

Indeed, by the definition of the L-transformation we have

$$\dot{L}_T(m_\theta^n,M_\theta^n)-\dot{L}_T(m_y^n,M_y^n) = L_T(\dot{m}_\theta^n-\dot{m}_y^n,M_\theta^n)+[L(\dot{m}_y^n,M_y^n),L(M_\theta^n-M_y^n,M_\theta^n)]_T \qquad (2.7)$$
$$+([L(m_y^n,M_y^n),\ L(\dot{M}_y^n,M_y^n)]_T-[L(m_\theta^n,M_\theta^n),L(\dot{M}_\theta^n,M_\theta^n)]_T).$$

Now applying successively Chebyshev's inequality, linearity of the L-transformation and the Cauchi-Buniakovsky inequality we get

$$P_\theta^n\{\sup_{z:|z-\theta|\leq\epsilon}\phi_n^2(\theta)|L_T(\dot{m}_\theta^n-\dot{m}_z^n,M_\theta^n)|>\rho\}\leq\frac{\phi_n^4(\theta)}{\rho^2}E_\theta^n\{\sup_{z:|z-\theta|\leq\epsilon}|L_T(\dot{m}_\theta^n-\dot{m}_z^n,M_\theta^n)|^2\}$$
$$=\frac{\phi_n^4(\theta)}{\rho^2}E_\theta^n\{\sup_{z:|z-\theta|\leq\epsilon}(L_T(\int_\theta^z\ddot{m}_u^n du,M_\theta^n))^2\}\leq\frac{\phi_n^4(\theta)}{\rho^2}\epsilon E_\theta^n\int_{\theta-\epsilon}^{\theta+\epsilon}L_T^2(\ddot{m}_z^n,M_\theta^n)dz$$
$$\leq\frac{\phi_n^4(\theta)}{\rho^2}\epsilon^2\sup_{z:|\theta-z|\leq\epsilon}E_\theta^n<L(\ddot{m}_z^n,M_\theta^n)>_T.$$

Evidently, by virtue of (2.7) conditions e′) imply e). □

If martingales M_θ^n and m_θ^n admit the integral representation (1.16) with corresponding coefficients $g_n(s,x,\theta)$, $g_n(s,\theta),\phi_n(s,x,\theta),\phi_n(s,\theta),\theta\in\Theta$ w.r.t. some integer-valued random measure μ with the P^n-compensator ν_n, then Lindeberg's condition d) can be written in the following equivalent form:

$$d'')\quad P_\theta^n-\lim_{n\to\infty}\phi_n^2(\theta)[\int_0^T\int_E(\frac{\psi_n(s,x,\theta)-\hat{\psi}_n(s,\theta)}{1+g_n(s,x,\theta)-\hat{g}_n(s,\theta)})^2 I\{|\frac{\psi_n(s,x,\theta)-\hat{\psi}_n(s,\theta)}{1+g_n(s,x,\theta)-\hat{g}_n(s,\theta)}|>\epsilon\phi_n^{-1}(\theta)\}\nu_\theta^n(ds,dx)$$
$$+\sum_{s<T}(\frac{\hat{\psi}_n(s,\theta)}{1-\hat{g}_n(s,\theta)})^2 I\{|\frac{\hat{\psi}_n(s,\theta)}{1-\hat{g}_n(s,\theta)}|>\epsilon\phi_n^{-1}(\theta)\}(1-a_\theta^n(s))] = 0,$$

where ν_θ^n is the compensator of the measure μ w.r.t. the measure P_θ^n and $a_\theta^n(s) = \nu_\theta^n(\{s\},E)$.

2.4. Conditions b) and c) of Theorem 2.1 are of ergodic nature. Their validity usually follows from the ergodicity of underlaying processes (see the examples of particular models of statistical experiments below). In this connection it is convenient to introduce the ergodicity concept for an experiment.

An experiment is called *ergodic* if the condition

$$P_\theta^n-\lim_{n\to\infty}\phi_n^2(\theta)<L(\dot{M}_\theta^n,M_\theta^n)>_T = 1.$$

holds.

A sequence of martingales $(m_\theta^n),n\geq 1$ is said to be *ergodically related* to an experiment if Conditions b) and c) are satisfied.

Of course, if the experiment is ergodic, then the sequence of martingales $\{\dot{M}_\theta^n\},n\geq 1$, defining the MLE, is ergodically related to this model.

We shall say below that an experiment is *strongly regular* if in addition to Conditions (R1)-(R4) the following condition is satisfied: $\forall\rho>0$

$$(\text{R5}) \lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}} P_\theta^n\{\sup_{z:|z-\theta|<\epsilon}\phi_n^2(\theta)|\dot{L}_T(\dot{M}_\theta^n,M_\theta^n)-\dot{L}_T(\dot{M}_z^n,M_z^n)|>\rho\} = 0.$$

A sequence of martingales (m_θ^n) will be called *strongly regular* if Condition e) of Theorem 2.1 is satisfied.

We will say that an experiment $\mathcal{E} = (\mathcal{E}_n), n\geq 1$ satisfies Lindeberg's condition if

$$P_\theta^n - \lim_{n\to\infty}\phi_n^2(\theta)\int_0^T\int_{|x|>\phi_n^{-1}(\theta)\epsilon} x^2\nu_n^{L(\dot{M}_\theta^n,M_\theta^n)}(ds,dx) = 0,\ \forall\epsilon\in(0,1],$$

where $\nu_n^{L(\dot{M}_\theta^n,M_\theta^n)}$ is the compensator of the jump measure of the process $L(\dot{M}_\theta^n,M_\theta^n)$.

We shall say that a sequence of martingales $m_\theta = \{m_\theta^n\}, n\geq 1$ satisfies Lindeberg's condition if condition d) is satisfied. We shall restrict below our attention to strongly regular ergodic experiments, satisfying Lindeberg's condition. A class of such experiments will be denoted by $\mathcal{C}$. If $\mathcal{E} = (\mathcal{E}_n)_{n\geq 1}\in\mathcal{C}$, then by $\mathfrak{M}(\mathcal{C})$ we denote the class of strongly regular sequences of martingales defining M-estimators, which are ergodically related to $\mathcal{C}$ and satisfy Lindeberg's condition.

3. Asymptotic behaviour of M-estimator under alternatives

3.1. Let $\tilde{\mathcal{E}} = (\tilde{\mathcal{E}}_n)_{n\geq 1}\in\mathcal{C}$ be an experiment, alternative to $\mathcal{E}$, and let $\tilde{M}_\theta = \{(\tilde{M}_\theta^n)_{n\geq 1},\theta\in\Theta\}$ be the family of martingales defining the densities $\{d\tilde{P}_\theta^n(t)/dP^n(t) = \mathcal{E}_t(\tilde{M}_\theta^n), 0\leq t\leq T, n\geq 1, \theta\in\Theta\}$.

We shall investigate the properties (under the alternative measure $\tilde{P}_\theta^n$) of the M-estimator which solves (1.11) with a martingale $m_\theta = (m_\theta^n)_{n\geq 1}\in\mathfrak{M}(\mathcal{E})$. To this end we will need certain additional conditions (see Theorem 3.1 below) which specify the behaviour of the sequence of martingales $m_\theta = (m_\theta^n)_{n\geq 1}$ under the alternative measure $\tilde{P}_\theta^n$ (the natural requirement $m_\theta\in\mathfrak{M}(\tilde{\mathcal{E}})$ solely may be insufficient) and also establish the required relation between the experiments $\mathcal{E}$ and $\tilde{\mathcal{E}}$.

Theorem 3.1. *Let the following conditions hold:*

$\tilde{a}$) *for each $\theta\in\Theta$ we have $\phi_n(\theta)$ and $\tilde{\phi}_n(\theta)$ related as follows*

$$\lim_{n\to\infty}\phi_n(\theta) = 0, \lim\frac{\tilde{\phi}_n(\theta)}{\phi_n(\theta)} = c(\theta)$$

with $0<c(\theta)<\infty$, *besides*

$$\tilde{\phi}_n^{-2}(\theta) = \tilde{E}_\theta^n <L(\dot{\tilde{M}}_\theta^n,\tilde{M}_\theta^n)>_T;$$

$\tilde{c}$) *for each $\theta,\theta'\in\Theta$ there exists a function of two arguments $\tilde{\Delta}(\theta,\theta')$ such that*

$$\tilde{P}_\theta^n - \lim_{n\to\infty}\phi_n^2(\theta)L_T(m_\theta^n,M_{\theta'}^n) = \tilde{\Delta}(\theta,\theta')$$

and for every $\theta\in\Theta$ the equation (w.r.t. θ')

$$\tilde{\Delta}(\theta,\theta') = 0 \tag{3.1}$$

is uniquely solved by $\theta' = b(\theta)$;

$\tilde{d}$) $$\tilde{P}_\theta^n - \lim_{n\to\infty}\phi_n^2(\theta)\dot{L}_T(m_{b(\theta)}^n,M_{b(\theta)}^n) = -\gamma(\theta)$$

where $\gamma(\theta)$ is positive for every $\theta\in\Theta$;

$\tilde{e}$) $$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\tilde{P}_\theta^n\{\sup_{z:|b(\theta)-z|<\epsilon}\phi_n^2(\theta)|\dot{L}_T(m_{b(\theta)}^n,M_{b(\theta)}^n)-\dot{L}_T(m_z^n,M_z^n)|>\rho\} = 0$$

for any $\rho>0$.

Then all assertions of Lemma 2.1 hold with $\tilde{P}_\theta^n$ instead of $Q_\theta^n, L_T(m_\theta^n,M_\theta^n)$ instead of $L_n(\theta)$ and $b(\theta)$ instead of θ'.

This theorem is an immediate consequence of Lemma 2.1. □

3.2. Introduce the following conditions c′) and d′).

c′) For all $\theta,\theta'\in\Theta$ let

c1′) $$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta) L_T(m^n_{\theta'},\tilde{M}^n_\theta) = 0;$$

c2′) there exists a function $\tilde{\Delta}(\theta,\theta')$ such that

$$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta)[L(m^n_{\theta'},M^n_{\theta'}),L(M^n_{\theta'}-\tilde{M}^n_\theta,\tilde{M}^n_\theta)]_T = \tilde{\Delta}(\theta,\theta')$$

and the equation

$$\tilde{\Delta}(\theta,\theta') = 0$$

has the unique solution $b(\theta)$.

d′). For each $\theta\in\Theta$

d1′) $$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta) L_T(\dot{m}^n_{b(\theta)},\tilde{M}^n_\theta) = 0;$$

d2′) $$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta)[L(m^n_{b(\theta)},M^n_{b(\theta)}),L(\dot{M}^n_{b(\theta)},M^n_{b(\theta)})]_T = \gamma_1(\theta),$$

$$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta)[L(\dot{m}^n_{b(\theta)},M^n_{b(\theta)}),L(M^n_{b(\theta)}-\tilde{M}^n_\theta,\tilde{M}^n_\theta)]_T = \gamma_2(\theta)$$

with

$$\gamma(\theta) = \gamma_1(\theta)+\gamma_2(\theta)>0.$$

COROLLARY 3.1. *Let conditions $\tilde{a}$), $\tilde{c}'$), $\tilde{d}'$), $\tilde{e}$) be satisfied. Then all assertions of Theorem 3.1 are valid.*

The proof immediately follows from the decomposition (1.24) and the formula (1.12).

REMARK 3.1. We may apply the decomposition (1.23) to $L_T(m^n_{\theta'},M^n_{\theta'})$, provided there exists the mutual square characteristic envolved in (1.23). As for $\tilde{c}$), it is satisfied if, for instance,

$$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta) L_T(m^n_\theta,M^n_\theta,\tilde{M}^n_\theta) = 0, \tag{3.2}$$

$$\tilde{P}^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta) < L(m^n_{\theta'},M^n_{\theta'}),L(\tilde{M}^n_\theta-M^n_{\theta'},M^n_{\theta'}) >_T = \tilde{\Delta}(\theta,\theta'),$$

where $\tilde{\Delta}(\theta,\theta')$ is the function envolved in Theorem 3.1, Condition c).

Using (1.23) we can analogously obtain sufficient conditions for d) to hold.

Further, as in Theorem 1.1, conditions a) - e) can be expressed directly in terms of martingales m^n_θ and M^n_θ or corresponding integrand functions in their integral representation (when these representations are known of course).

REMARK 3.2. i) The passage from the experiment $\mathfrak{E}$ to its alternative $\tilde{\mathfrak{E}}$ leads to a drift $b(\theta)$ in the parameter estimation, caused by the fact that the process $L_n(\theta) = L(m^n_\theta,M^n_\theta)$ is not a martingale w.r.t. the measure $\tilde{P}^n_\theta$. Now, in Theorem 3.1, Condition $\tilde{c}$) the shift $b(\theta)$ of the parameter θ is determined which forces the normalized non-martingale component of the semimartingale L_n to vanish asymptotically.

ii) In contrast with Section 2 (see Theorem 2.1) here we do not mention central limit theorems (CLT) of any kind for estimators, because they might be true only in certain special situations; it can be seen (see Lemma 2.1, Assertion V) that the main problem here consists in establishing the CLT for the process $L_n(b(\theta))$. It is hard to investigate the weak convergence of the second summand in the formula

$$L(m^n_{b(\theta)},M^n_{b(\theta)}) = L(m^n_{b(\theta)},\tilde{M}^n_\theta)+[L(m^n_{b(\theta)},M^n_{b(\theta)}),L(M^n_{b(\theta)}-\tilde{M}^n_\theta,\tilde{M}^n_\theta)].$$

In the simple examples given below the problem is solved due to the specific character of the model.
iii) Comparing (2.1) with the expression $\gamma(\theta)$ in Theorem 3.1, Condition $\tilde{d}$), we see that the scale parameter in the distribution of the estimator $\hat{\theta}_n$ differs under the alternative: we get the additional term $\gamma_2(\theta)$ depending on the derivative $\dot{m}^n_\theta$ of the martingale m^n_θ.

4. Examples

4.1. Before presenting the examples we recall the standard device of reducing an experiment with a fixed family of distributions and with an infinitely increasing sample size, to a scheme of series. Such experiment is defined by the set

$$(\Omega,\mathcal{F},F = (\mathcal{F}_t)_{t\geq 0},\ P_\theta,P) \text{ with } P_\theta \overset{loc}{\sim} P \text{ i.e. } P_\theta(t)\sim P(t) \text{ for } all\ t\geq 0.$$

By $M_\theta = (M_\theta(t),t\geq 0)$ we will denote martingales, defining the densities.
The corresponding scheme of series has the form

$$\mathcal{E}_n = (\Omega,\mathcal{F},F^n = (\mathcal{F}^n_t := \mathcal{F}_{nt}),0\leq t\leq T,P^n_\theta = P_\theta|\mathcal{F}^n,P^n = P|\mathcal{F}^n),$$

with the family of martingales $M^n_\theta = (M^n_\theta(t) := M_\theta(nt)),0\leq t\leq T$. It is useful in such schemes to express the conditions of Theorems 2.1 and 3.1 directly in terms of martingales $M_\theta = (M_\theta(t)),t\geq 0,\theta\in\Theta$. For example, if $\phi_t^{-2}(\theta) = E_\theta\langle L(M_\theta,M_\theta)\rangle_t$ is Fisher's information quantity up to time t, then the first two of the conditions in Theorem 2.1, for instance, take the following form:

a) $\lim\limits_{t\to\infty}\phi_t(\theta) = 0,$

(b1) $P_\theta - \lim\limits_{t\to\infty}\phi_t^2(\theta)\langle L(m_\theta,M_\theta)\rangle_t = \Gamma(\theta)<\infty.$

Example 1. *Independent Identically Distributed Observations.*
Let μ and μ_θ be probability measures defined on some measurable space $(X,\mathcal{B})$ and let $\mu_\theta\sim\mu$, $f(x,\theta) = \dfrac{d\mu_\theta}{d\mu}(x)$. Put $T = 1$.
Consider a sequence of experiments $\mathcal{E}_n,n\geq 1$, with

$$\Omega^n = X^n,\ \mathcal{F}^n = \mathcal{B}^n,\ F^n = \{\mathcal{F}^n_t := \mathcal{B}^{[nt]},0\leq t\leq 1\},\ P^n_\theta = \overbrace{\mu_\theta\times..\times\mu_\theta}^{n},\ P^n = \overbrace{\mu\times..\times\mu}^{n}$$

($[\cdot]$ denotes an integer part). It can be easily seen that for $\omega = (x_1,...,x_n)\in X^n$

$$P^n_\theta(t) = \prod_{i=1}^{[nt]}f(x_i,\theta),\ M^n_\theta(t) = \sum_{i=1}^{[nt]}(f(x_i,\theta)-1).$$

Besides, if $f(x,\theta)$ is twice continuously differentiable w.r.t. θ in $L_2(X,\mathcal{B},\mu)$, the second derivative $f(x,\theta)$ is θ continuous $\mu-$ a.s. for all x and

$$0<I_\theta = \int(\frac{\dot{f}(x,d\theta)}{f(x,\theta)})^2f(x,\theta)\mu(dx)<\infty,\int(\frac{\ddot{f}(x,\theta)}{f(x,\theta)})^2f(x,\theta)\mu(dx)<\infty,$$

then the experiment is regular.
In this scheme one often considers only the subclass $\mathcal{M}^0_R\subset\mathcal{M}_R$, consisting of martingales m^n_θ of the form

$$m^n_\theta = \sum_{i=1}^{[n.]}\psi(x_i,\theta),\ \psi\in\Psi(f),$$

where $\Psi(f)$ is a class of continuously differentiable in $L_2(X,\mathcal{B},\mu)$ functions with

$$\int\psi(x,\theta)\mu(dx) = 0,\ \int\left(\frac{\psi(x,\theta)}{f(x,\theta)}\right)^2 f(x,\theta)\mu(dx)<\infty,$$

$$\int\left(\frac{\dot\psi(x,\theta)}{f(x,\theta)}\right)^2 f(x,\theta)\mu(dx)<\infty,\quad 0<\int\frac{\psi(x,\theta)\dot f(x,\theta)}{f^2(x,\theta)}f(x,\theta)\mu(dx)<\infty.$$

The equation (1.11) takes the form

$$L_1(m_\theta^n,M_\theta^n) = \sum_{i=1}^n\frac{\psi(x_i,\theta)}{f(x_i,\theta)} = 0. \tag{4.1}$$

In Theorem 2.1 Conditions a) -d) with $\phi_n^{-2}(\theta) = nI_\theta$,

$$\Gamma(\theta) = \frac{1}{I_\theta}\int\frac{\psi^2(x,\theta)}{f^2(x,\theta)}f(x,\theta)\mu(dx),$$

$$\gamma(\theta) = \frac{1}{I_\theta}\int\frac{\psi(x,\theta)\dot f(x,\theta)}{f^2(x,\theta)}f(x,\theta)\mu(dx),$$

are verified by the law of large numbers and CLT. If, in addition Condition e) is required, then Assertion IV of Theorem 2.1 coincides with the classical result

$$\mathcal{L}_{P_\theta} - \sqrt{n}(\tilde\theta_n-\theta)\Rightarrow N(0,\frac{\Gamma(\theta)}{\gamma^2(\theta)I_\theta}).$$

Consider now an alternative regular experiment corresponding to the density $\tilde f(x,\theta)$ and let $\psi\in\Psi(f)\cap\Psi(\tilde f)$.

The conditions in Section 3 involve the following quantities, written in the explicit form:

$$\Delta(\theta,\theta') = \int\frac{\psi(x,\theta')}{f(x,\theta')}(f(x,\theta')-\tilde f(x,\theta))\mu(dx),$$

$$\gamma_1(\theta) = \frac{1}{I_\theta}\int\frac{\psi(x,b(\theta))\dot f(x,b(\theta))}{f^2(x,b(\theta))}\tilde f(x,\theta)\mu(dx),$$

$$\gamma_2(\theta) = \frac{1}{I_\theta}\int\frac{\dot\psi(x,b(\theta))(f(x,b(\theta))-\tilde f(x,\theta))}{f(x,b(\theta))}\mu(dx).$$

Here we have the following central limit theorem (cf. Remark 3.2, (ii)):

$$\mathcal{L}_{\tilde P_\theta} - \sqrt{n}(\hat\theta_n-b(\theta))\Rightarrow N(0,\frac{\gamma^2(\theta)\int(\psi(x,b(\theta))^2(f(x,b(\theta))^{-2}\tilde f(x,\theta)\mu(dx)}{I_\theta^2})$$

where $\gamma(\theta) = \gamma_2(\theta)+\gamma_2(\theta)$; it is deduced from Lemma 2.1, Assertion IV - the CLT w.r.t. the measure $\tilde P_\theta$ for the sums

$$L_1(m_{b(\theta)}^n,M_{b(\theta)}^n) = \sum_{i=1}^n\frac{\psi(x_i,b(\theta))}{f(x_i,b(\theta))}.$$

Example 2. *Estimation of a drift parameter of a diffusion process.*
Suppose that for every $n\geq 1$ one observes the process ξ_n with a differential

$$d\xi_n(t) = a_n(t,\xi_n,\theta)dt+dW_n(t),\ 0\leq t\leq T,$$

and one needs to estimate the unknown parameter θ ($a_n(\cdot,\cdot,\theta)$ is a non-anticipating functional).

To cover this problem by a general scheme of experiments, set $\Omega^n = C_{[0,T]},\mathcal{F}^n = \mathcal{B}(C_{[0,T]})$, $F^n = \{\mathcal{F}_t^n := \sigma(x:x_s,s\leq t),0\leq t\leq T\},P_\theta^n$ a distribution of the process ξ_n with a given θ,P^n a Wiener measure. Assume for all $n\geq 1$ and $\theta\in\Theta$ that

$$P^n_\theta(\int_0^T a_n^2(s,x,\theta)ds<\infty) = P^n(\int_0^T a_n^2(s,x\theta)ds<\infty) = 1.$$

Then $P^n_\theta \sim P^n$ and the local density takes the form

$$\rho^n_\theta(t) = \exp(\int_0^t a_n(s,x,\theta)ds - \frac{1}{2}\int_0^t a_n^2(s,x,\theta)ds) = \mathcal{E}_t(M^n_\theta)$$

where

$$M^n_\theta(t) = \int_0^t a_n(s,x,\theta)dx_s$$

is a local P^n-martingale.

Further, assume that the following condition is satisfied: for every n the function $a_n(s,x,.)$ is twice differentiable w.r.t. θ in the sense of the norm $\|\cdot\|$ where $\|f\|^2 = E^n\int_0^T f^2(s)ds$, with the θ-continuous second derivative for all t P^n - a.s. Then $\dot{M}^n_\theta(t) = \int_0^t \dot{a}_n(s,x,\theta)dx_s$. So the Fisher information takes the form

$$\phi_n^{-2}(\theta) = I_n(\theta) = E^n_\theta\int_0^T \dot{a}^2(s,x,\theta)ds.$$

Let $0<I_n(\theta)<\infty$. Under these conditions the experiment is obviously regular and the likelihood equation takes the form

$$L_T(\dot{M}^n_\theta, M^n_\theta) = \int_0^T \dot{a}_n(s,x,\theta)(dx_s - a_n(s,,x,\theta)ds) = 0.$$

In view of the fact that any P^n-martingale has an integral representation w.r.t. a Wiener process, we consider martingales defining M-estimators of the form

$$m^n_\theta(t) = \int_0^t \psi_n(s,x,\theta)dx_s,$$

where $\psi_n(\cdot,\cdot,\theta)$ is a non-anticipating functional and the equation (1.11) giving the M-estimator can be written as

$$\int_0^t \psi_n(s,x,\theta)(dx_s - a_n(s,x,\theta)ds) = 0. \tag{4.2}$$

The regularity of such M-estimator means, for all n, that the function ψ is continuously differentiable w.r.t. θ in the sense of the norm $\|\cdot\|$, with the θ-continuous derivative $\dot{\psi}(t,x,\theta)$ for all $t\in[0,T]$ P^n-a.s. Besides, the values $E^n_\theta\int_0^T \psi_n^2(s,x,\theta)ds$, $E^n_\theta\int_0^T \dot{\psi}_n^2(s,x,\theta)ds$ are finite and positive.

Assume that the following conditions hold: for all $\theta\in\Theta$ and for $\forall\rho>0$

1) $\lim_{n\to\infty}\phi_n(\theta) = 0$;
2) $P^n_\theta - \lim_{n\to\infty}\phi_n^2(\theta)\int_0^T \psi_n^2(s,x,\theta)ds = \Gamma(\theta)$;
3) $P^n_\theta - \lim_{n\to\infty}\phi_n^2(\theta)\int_0^T \psi_n(s,x,\theta)\dot{a}_n(s,x,\theta)ds = \gamma(\theta)$
   where $\Gamma(\theta)$ and $\gamma(\theta)$ are positive numbers;
4) $P^n_\theta - \lim_{n\to\infty}\phi_n^4(\theta)\int_0^T \dot{\psi}_n^2(s,x,\theta)ds = 0$;

5) $\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}\phi_n^4(\theta)E_\theta^n\int_0^T\sup_{z:|\theta-z|<\epsilon}\ddot{\psi}_n^2(s,x,z)ds = 0,$

$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}P_\theta^n\{\phi_n^2(\theta)\sup_{z:|z-\theta|<\epsilon}|\int_0^T\dot{\psi}_n(s,x,z)(a_n(s,x,z)-a_n(s,x,\theta))ds|>\rho\} = 0,$

$\lim_{\epsilon\to 0}\overline{\lim_{n\to\infty}}P_\theta^n\{\phi_n^2(\theta)\sup_{z:|z-\theta|<\epsilon}|\int_0^T(\psi_n(s,x,z)\dot{a}_n(s,x,z)-\psi_n(s,x,\theta)\dot{a}_n(s,x,\theta))ds|>\rho\} = 0.$

It is easily seen that all conditions of Theorem 2.1 are satisfied and, hence, all assertions of Lemma 2.1 are valid and, besides,

$$\mathcal{L}_{p_\theta^n}\{\phi_n^{-1}(\theta)(\tilde{\theta}_n^\psi-\theta)\}\Rightarrow N(0,\frac{\Gamma(\theta)}{\gamma^2(\theta)}).$$

Note, for instance, that conditions 4) and 5) imply conditions c) and e) of Theorem 2.1, respectively.

Now let $\tilde{P}_\theta^n$ be the distribution of the process ξ_n with the differential

$$d\xi_n(t) = \tilde{a}_n(t,\xi_n,\theta)dt+dW_n(t),\ 0\leq t\leq T,$$

and suppose that the corresponding experiment is regular and *ergodically related* to $\mathcal{E}_n$ in the sense that condition a) is satisfied in Theorem 3.1. We are interested in the asymptotic properties w.r.t. $\tilde{P}_\theta^n$ of the *M*-estimator obtained from (4.2). We obtain sufficient conditions for validity of Theorem 3.1 in the same manner as above. In particular, the drift equation takes the form

$$\tilde{P}_\theta^n - \lim_{n\to\infty}\phi_n^2(\theta)\int_0^T\psi_n(s,x,\theta')(a_n(s,x,\theta')-\tilde{a}_n(s,x,\theta))ds = \Delta(\theta,\theta') = 0.$$

It can be easily seen (cf. [5]) that the following elementary conditions are, in turn, sufficient for condition e) of Theorem 3.1 to hold; functionals $U_n^K, V_n, \tilde{V}_n^K$ exist such that[1)]

$$|\psi_n^{(k)}(s,x,\theta)|^2\leq U_n^k(s,x),\quad k = 0,1,2;\ |a_n(s,x,\theta)|^2\leq V_n(s,x),$$

$$|\tilde{a}_n^{(k)}(s,x,\theta)|^2\leq \tilde{V}_n^k(s,x),\quad k = 0,1;\ \lim_{n\to\infty}\tilde{E}_\theta^n\int_0^T U_n^{(k)}(s,x)ds<\infty,\quad k = 0,1,2,$$

$$\lim_{n\to\infty}\tilde{E}_\theta^n\int_0^T V_n(s,x)ds<\infty,\ \lim_{n\to\infty}E_\theta^n\int_0^T U_n^{(k)}(s,x)ds<\infty,\quad k = 0,1.$$

EXAMPLE 3. *Small diffusion ([7],[8]).*

Let the MLE $\hat{\theta}_T^\epsilon$ be constructed by means of observations on ξ with the differential

$$d\xi_t = a(\xi_t,\theta)dt+\epsilon dW_t, \xi_0 = 0,\ 0\leq t\leq T,$$

Assuming $\epsilon\to 0$, we study its asymptotic properties w.r.t. P^ϵ, where P^ϵ is the distribution of ξ with the differential

$$d\xi_t = b(\xi_t)dt+\epsilon dW_t,\ \xi_0 = 0.$$

Evidently, this problem is equivalent to the problem of constructing the *MLE* $\hat{\theta}_T^\epsilon$ by means of observations on η with the differential

$$d\eta_t = a_\epsilon(\eta_t,\theta)dt+dW_t,\ \eta_0 = 0,$$

and then studying its asymptotic properties w.r.t. the distribution $\tilde{P}^\epsilon$ of η with the differential

$$d\eta_t = b_\epsilon(\eta_t)dt+dW_t,\ \eta_0 = 0,$$

1) We denote by $f^{(k)}(x,\theta)$ the k-th derivative of a function f w.r.t. θ.

where

$$\eta_t = \xi_t/\epsilon,\ a_\epsilon(Y,\theta) = a(\epsilon Y,\theta)/\epsilon,\ b_\epsilon(Y) = b(\epsilon Y)/\epsilon.$$

In this case

$$L_T(\dot{M}^\epsilon_\theta, M^\epsilon_\theta) = \int_0^T \dot{a}_\epsilon(Y_s,\theta)(dY_s - a_\epsilon(Y_s,\theta)ds).$$

Following [8] we assume that the function $a(x,\theta)$ has three bounded derivatives w.r.t. x and θ, and $b(x)$ has a bounded derivative w.r.t. $x \in R^1$. Introduce the notation $x_0 = (x_0(t)), 0 \leqslant t \leqslant T$ for the solution of the deterministic equation

$$\frac{dx_0(t)}{dt} = b(x_0(t)),\ x_0(0) = 0.$$

Denote

$$G(\theta) = \int_0^T (a(\theta,x_0(t)) - b(x_0(t)))dt,$$

$$I(\theta) = \int_0^T [\dot{a}^2(\theta,x_0(s)) + \ddot{a}(\theta,x_0(s))(a(\theta,x_0(s)) - b(x_0(s))]ds,$$

$$\zeta(\theta) = I^{-1}(\theta)\{\int_0^T\{\int_0^T \dot{a}(\theta,x_0(s))dW_s - \int_0^T[\frac{\partial}{\partial x}\dot{a}(\theta,x_0(s))[a(\theta,x_0(s))$$

$$-b(x_0(s))]x_0^{(1)}(s) + \dot{a}(\theta,x_0(s))[\frac{\partial}{\partial x}\dot{a}(\theta,x_0(s)) - \frac{\partial}{\partial x}b(x_0(s))]x_0^{(1)}(s)]ds$$

$$\theta^* = \arg\min G(\theta),$$

where $x_0^{(1)} = (x_0^{(1)}(t)), 0 \leqslant t \leqslant T$ is the solution of the equation

$$dx_0^{(1)}(t) = \frac{\partial}{\partial x}b(x_0(t))x_0^{(1)}(t)dt + dW_t,\ x_0^{(1)}(0) = 0,\ 0 \leqslant t \leqslant T.$$

Then all conditions of Theorem 3.1 are satisfied, provided the equation (of drift)

$$\tilde{\Delta}(\theta,\theta') = \int_0^T \dot{a}(x_0(s),\theta')(b(x_0(s)) - a(x_0(s),\theta'))ds = 0$$

has the unique solution and

$$\int_0^T \ddot{a}(x_0(s),\theta'(\theta))(b(x_0(s)) - a(x_0(s),\theta'(\theta))ds - \int_0^T \dot{a}^2(x_0(s),\theta'(\theta))ds = \tilde{\Gamma}(\theta) < 0.$$

Then by Theorem 3.1

$$\tilde{P}^\epsilon_\theta - \lim_{\epsilon \to 0} \hat{\theta}^\epsilon_T = \theta'(\theta),$$

$$\lim_{\epsilon \to 0} \mathfrak{L}_{\tilde{P}_\theta^\epsilon} - \epsilon^{-1}(\hat{\theta}^\epsilon_T - \theta'(\theta)) = \lim_{\epsilon \to 0} \mathfrak{L}_{\tilde{P}_\theta^\epsilon} - \frac{\epsilon L_T(\dot{M}^\epsilon_{\theta'(\theta)}, M^\epsilon_{\theta'(\theta)})}{-\tilde{\Gamma}(\theta)}$$

It can be easily seen that

$$\lim_{\epsilon \to 0} \mathfrak{L}_{\tilde{P}_\theta^\epsilon} - \frac{\epsilon \mathfrak{L}_T(\dot{M}^\epsilon_{\theta'(\theta)}, M^\epsilon_{\theta'(\theta)})}{-\tilde{\Gamma}(\theta)} = \mathfrak{L}(\zeta(\theta'(\theta))).$$

But under our assumptions $\theta'(\theta) = \theta^* = \mathrm{argmin} G(\theta)$. Hence

$$\tilde{P}^\epsilon_\theta - \lim_{\epsilon \to 0} \hat{\theta}^\epsilon_T = \theta^*,$$

$$\lim_{\epsilon \to 0} \mathcal{L}_{\tilde{P}^\epsilon_\theta} - \epsilon^{-1}(\hat{\theta}^\epsilon_T - \theta^*) = \mathcal{L}_p - \zeta(\theta^*) := \mathcal{L}(\zeta(\theta^*)),$$

that is the same result as in [8].

EXAMPLE 4. *Multivariate point process.*
These processes are defined (see e.g. [3]) by a given sequence $(T_k, \xi_k), k \geq 1$ of random variables taking values in $(R_+ \times E)$ where E is a certain Luzin space with $T_k < T_{k+1}$ being the jump sizes of the ξ_k process. The distributions of the process are defined by a given sequence of regular conditional distributions

$$G_k(B) = P(T_{k+1}, \xi_{k+1}) \in B \mid T_i, \xi_i, i \leq K), \quad B \in \mathcal{B}(R_+ \times E).$$

The corresponding model of the experiment (in a scheme of series) is defined by the objects:

(i) $(\Omega^n, \mathcal{F}^n, F^n) = (D_{[0,T]}, \mathcal{B}(D_{[0,T]}), \mathcal{B}^n_t = \sigma(x : x_s, s \leq t)),$

where $D[0,T]$ is Skorokhod's space of functions which take values in E;

(ii) $\mu((0,t] \times B) = \sum_{s \leq t} I\{\Delta x_s \in B\}, \quad B \in \mathcal{B}(E),$

that is the integer-valued jump measure of a coordinate process x;

(iii) the measure P^n_θ which is characterized by the compensator of the measure μ, given by

$$\nu^n_\theta(dt, dx) = \sum_{k > 0} I\{T_k < t \leq T_{k+1}\} \frac{G^{n,k}_\theta(dt, dx)}{G^{nk}_\theta([t, \infty) \times E)},$$

where $G^{n,k}_\theta$ is the conditional distribution of (T_{k+1}, ξ_{k+1}) and $T_k = \infty\{t > T_{k-1} : \Delta x_t \neq 0\}$, $\xi_k = \Delta x_{T_k}$.

For all $\theta \in \Theta$ the assumption $P^n_\theta \sim P^n$ is equivalent to $\nu^n_\theta \sim \nu^n$ where ν^n_θ and ν are the compensators of μ w.r.t. P^n_θ and P^n respectively (see, e.g., [3]).

The martingales M^n_θ have the integral representation (for simplicity we consider here the case $\nu^n_\theta(\{t\}, E) = 0$):

$$M^n_\theta(t) = \int_0^t \int_E (f^n(s,x,\theta) - 1)(\mu - \nu^n)(ds, dx), \quad f^n(s,x,\theta) := \frac{d\nu^n_\theta}{d\nu^n}(s,x).$$

As for the martingales defining M-estimators, we consider the martingales

$$m^n_\theta(t) = \int_0^t \int_E \psi^n(s,x,\theta)(\mu - \nu^n)(ds, dx).$$

It can be easily calculated, for instance, that

$$L_t(\dot{M}^n_\theta, M^n_\theta) = \int_0^t \int_E \frac{\dot{f}^n(s,x,\theta)}{f^n(s,x,\theta)} (\mu - \nu^n_\theta)(ds, dx),$$

$$L_t(m^n_\theta, M^n_\theta) = \int_0^t \int_E \frac{\psi^n(s,x,\theta)}{f^n(s,x,\theta)} (\mu - \nu^n_\theta)ds, dx),$$

$$<L(m^n_\theta, M^n_\theta)>_t = \int_0^t \int_E \left(\frac{\psi^n(s,x,\theta)}{f^n(s,x,\theta)}\right)^2 \nu^n_\theta(ds, dx),$$

$$[L(m^n_\theta, M^n_\theta), L(\dot{M}^n_\theta, M^n_\theta)]_t = \int_0^t \int_E \dot{f}^n(s,x,\theta) \psi^n(s,x,\theta) \mu(ds, dx)$$

$$= \sum_{k \geq 1} I\{T_k \leq t\} \frac{\dot{f}^n(T_k,\xi_k,\theta)\psi^n(T_k,\xi_k,\theta)}{f^2(T_k,\xi_k,\theta)}.$$

The regularity of the experiment means that the function $f^n(s,x,\theta)$ is twice differentiable continuously w.r.t. θ in the Hilbert space

$$L_2(\Omega^n \times R_+ \times E, \tilde{\mathcal{P}}, M^{P^n}_{\nu})$$

where $M^{P^n}_{\nu}$ is Dolean's measure of the compensator ν^n (restriction of the measure $\nu^n(dt,dx)P^n(dw)$ to a predictable σ-algebra $\tilde{\mathcal{P}}$) so that the integrability condition can be written as

$$E^n \int_0^T \int_E f_n^2(s,x,\theta)\nu^n(ds,dx) < \infty, E^n_\theta \int_0^T \int_E \left(\frac{\dot{f}_n(s,x,\theta)}{f(s,x,\theta)}\right)^2 \nu^n_\theta(ds,dx) < \infty$$

etc.

Similarly, the regularity conditions for the martingale m^n_θ can be explicitly expressed in terms of differentiability w.r.t. the parameter of the function ψ^n.

An alternative experiment can be defined by derivatives $\tilde{f}_n(s,x,\theta) = \frac{d\tilde{\nu}^n_\theta}{d\nu^n}(s,x)$ of the corresponding compensators $\tilde{\nu}^n_\theta$.

5. Contigual alternatives

5.1. Consider a sequence of alternative experiments

$$\tilde{\mathcal{E}}_n = (\Omega^n, \mathcal{F}^n, \mathbf{F}^n, \tilde{P}^n_\theta, P^n),\ n \geq 1, \theta \in \Theta \subset R^1,$$

such that for all $n \geq 1$ and $\theta \in \Theta$, $d\tilde{P}^n_\theta / dP^n = \mathcal{E}(\tilde{M}^n_\theta)$, and the martingales M^n_θ and $\tilde{M}^n_\theta$ corresponding to the basic and alternative experiments respectively, are related as follows:

$$L(\tilde{M}^n_\theta, M^n_\theta) = L(M^n_\theta, M^n_\theta) + \phi_n(\theta) L(N^n_\theta, M^n_\theta), \tag{5.1}$$

where N^n_θ are some P^n-martingales.

Based on the properties of L-transformations, we can represent (5.1) in the following equivalent form:

$$\tilde{M}^{n,c}_\theta = M^{n,c}_\theta + \phi_n(\theta) N^{n,c}_\theta, \tag{5.2}$$
$$\Delta\tilde{M}^n_\theta = \Delta M^n_\theta + \phi_n(\theta)\Delta N^n_\theta.$$

Remark 5.1. Treating independent identically distributed (i.i.d.) observations with the density $f(x,\theta), x \in R^1, \theta \in \Theta$, P. Huber [9] have introduced the sequence of alternative experiments $\tilde{\mathcal{E}}_n$ generated by i.i.d. observations with the density $\tilde{f}_n(x,\theta) = f(x,\theta) + \frac{c}{\sqrt{n}} h(x,\theta)$ where $\int h(x,\theta)\mu(dx) = 0$. (so called 'contamination model'). In this case

$$\frac{d\tilde{P}^n_\theta}{dP^n} = \prod_{i \leq n} \tilde{f}(x_i,\theta) = \mathcal{E}(\tilde{M}^n_\theta)$$

with $\tilde{M}^n_\theta = \sum_{i \leq n} (\tilde{f}(x_i,\theta) - 1)$. Therefore

$$L(\tilde{M}^n_\theta, M^n_\theta) - L(M^n_\theta, M^n_\theta) = \sum_{i \leq n} \frac{\tilde{f}(x_i,\theta) - f(x_i,\theta)}{f(x_i,\theta)} = \frac{c}{\sqrt{n}} \sum_{i \leq n} \frac{h(x_i,\theta)}{f(x_i,\theta)}$$

$$= \frac{c}{\sqrt{n}} L(N^n_\theta, M^n_\theta), \tag{5.3}$$

where $N_\theta^n = \sum_{i\leq n} h(x_i,\theta)$. Hence the sequence of experiments $\tilde{\mathcal{E}}_n$ with property (5.1) may be viewed as a natural extension of the 'contamination model'.

5.2. It is well known that in the scheme of i.i.d. observations mentioned above, the condition $\int(\frac{h(x,\theta)}{f(x,\theta)})^2 f(x,\theta)\mu(dx)<\infty$ implies the contiguity of the sequence of measures $(\tilde{P}_\theta^n)$ w.r.t. the sequence P_θ^n $((\tilde{P}_\theta^n)\lhd(P_\theta^n))$.

In the lemma given below we present sufficient conditions for contiguity w.r.t. (P_θ^n) of the sequence of measures $(\tilde{P}_\theta^n)$ satisfying condition (5.1).

LEMMA 5.1. *Suppose* $L(N_\theta^n,M_\theta^n)\in\mathfrak{M}^2(F,P_\theta^n)$ *and let*

$$\lim_{N\to\infty}\lim_{n\to\infty}\tilde{P}_\theta^n\{\phi_n^2(\theta)<L(N_\theta^n,M_\theta^n)>_T\geq N\} = 0. \tag{5.4}$$

Then $(\tilde{P}_\theta^n)\lhd(P_\theta^n)$.

PROOF. The following necessary and sufficient conditions for $(\tilde{P}_\theta^n)\lhd(P_\theta^n)$ are wellknown (see, e.g. [4]):

$$\lim_{N\to\infty}\lim_{n\to\infty}\tilde{P}_\theta^n\{h_T(\tfrac{1}{2};\tilde{P}_\theta^n,P_\theta^n)>N\} = 0, \tag{5.5}$$

$$\lim_{N\to\infty}\lim_{n\to\infty}\tilde{P}_\theta^n\{\sup_{s\leq T}\alpha_n(s)>N\} = 0, \tag{5.6}$$

where $h(\frac{1}{2},\tilde{P}_\theta^n,P_\theta^n)$ is the Hellinger process of order 1/2, and $\alpha_n(s) = \tilde{\rho}_\theta^n(t)/\rho_\theta^n(t-)$ where

$$\tilde{\rho}_\theta^n = \frac{d\tilde{P}_\theta^n}{dP_\theta^n} = \mathcal{E}(L(\tilde{M}_\theta^n-M_\theta^n,M_\theta^n)) = \mathcal{E}(\phi_n(\theta)L(N_\theta^n,M_\theta^n)).$$

It can be easily seen that

$$h(\tfrac{1}{2};\tilde{P}_\theta^n,P_\theta^n)\leq\tfrac{1}{2}\phi_n^2(\theta)<L(N_\theta^n,M_\theta^n)>. \tag{5.7}$$

Indeed

$$h(\tfrac{1}{2};\tilde{P}_\theta^n,P_\theta^n) = \tfrac{1}{8}\phi_n^2(\theta)<L^c(N_\theta^n,M_\theta^n)>+\tfrac{1}{2}(\Sigma(1-\sqrt{1+\phi_n(\theta)\Delta L(N_\theta^n,M_\theta^n)})^2)^{p,P_\theta^n}$$

But since $(1-\sqrt{1+x})^2 = \frac{x^2}{(1+\sqrt{1+x})^2}\leq x^2$ for $x\geq-1$, we have

$$h(\tfrac{1}{2};\tilde{P}_\theta^n,P_\theta^n)\leq\tfrac{1}{8}\phi_n^2(\theta)<L^c(N_\theta^n,M_\theta^n)>+\tfrac{1}{2}(\Sigma\phi_n^2(\theta)(\Delta L(N_\theta^n,M_\theta^n))^2)^{p,P_\theta^n}$$

$$\leq\tfrac{1}{2}\phi_n^2(\theta)<L(N_\theta^n,M_\theta^n)>.$$

Further

$$\tilde{P}_\theta^n(\sup_{s\leq T}\alpha_n(s)>N) = \tilde{P}_\theta^n(\sup_{s\leq T}(1+\phi_n(\theta)\Delta L(N_\theta^n,M_\theta^n))>N)$$
$$\leq\tilde{P}_\theta^n(\sup_{s\leq T}\phi_n^2(\theta)(\Delta L(N_\theta^n,M_\theta^n))^2>N-1)\leq\tilde{P}_\theta^n(\sum_{s\leq T}\phi_n^2(\theta)(\Delta L(N_\theta^n,M_\theta^n))^2>N-1).$$

By virtue of Lenglart's inequality for any $\eta>0$

$$\tilde{P}_\theta^n(\sum_{s\leq T}\phi_n^2(\theta)(\Delta L(N_\theta^n,M_\theta^n))^2>N)\leq\frac{\eta}{N}+\tilde{P}_\theta^n(\sum_{s\leq T}(\phi_n^2(\theta)(\Delta L(N_\theta^n,M_\theta^n))^2)^{p,P_\theta^n}>\eta). \tag{5.8}$$

Now (5.5) and (5.6) follow from (5.4) and (5.8).

5.3. In Theorem 5.1 below we present certain asymptotic properties of M-estimators under contigual alternatives. As above, consider regular experiments and martingales defining the estimators.

THEOREM 5.1. *Let conditions of Theorem 2.1 hold as well as condition* $\tilde{e}$) *of Theorem 3.1 with* $b(\theta) = \theta$. *Besides, let condition (5.4) be satisfied.*

Then all assertions of Lemma 2.1 are valid with

$$L_n(\theta) = L_T(m_\theta^n, M_\theta^n),\quad Q_\theta^n = \tilde{P}_\theta^n,\ \theta' = \theta,$$

and also

$$\lim_{n\to\infty} \mathcal{L}_{\tilde{P}_\theta^n}\{\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta)-\gamma^{-1}(\theta)\phi_n^2(\theta)\langle L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)\rangle_T\} = N(0,\frac{\Gamma(\theta)}{\gamma^2(\theta)}).$$

PROOF. By Lemma 5.1 $(\tilde{P}_\theta^n)\lhd(P_\theta^n)$, hence, conditions a) -b) of Lemma 2.1 are satisfied with

$$L_n(\theta) = L_T(m_\theta^n, M_\theta^n),\quad Q_\theta^n = \tilde{P}_\theta^n,\ \theta' = \theta.$$

Consequently assertions I-III of Lemma 2.1 are valid.

Further, as in the proof of Lemma 2.1, we can easily obtain the following inequality

$$|U_n - V_n| \leq Z_n|V_n| + Z_n\beta_n, \tag{5.9}$$

where

$$U_n = \phi_n(\theta)L_T(m_{\hat{\theta}_n}^n, M_{\hat{\theta}_n}^n)-(X^n-\phi_n^2(\theta)\langle L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)\rangle_T),$$

with

$$X^n = \phi_n(\theta)L_T(m_\theta^n,M_\theta^n),\ V_n = \gamma(\theta)\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta)-\phi_n^2(\theta)\langle L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)\rangle_T,$$

$$Z_n = \gamma^{-1}(\theta)\delta_n(\hat{\theta}_n)$$

($\delta_n(\theta)$ is defined in the course of proving Lemma 2.1) and

$$\beta_n = \phi_n^2(\theta)\langle L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)\rangle_T.$$

If now

$$\tilde{P}_\theta^n - \lim Z_n\beta_n = 0, \tag{5.10}$$

then by virtue of Remark 2.1 the limiting distributions (w.r.t. the measure $\tilde{P}_\theta^n$) of U^n and V^n coincide and since by Lemma 2.1, Assertion I,

$$\tilde{P}_\theta^n - \lim_{n\to\infty} \phi_n(\theta)L(m_{\hat{\theta}_n}^n, M_{\hat{\theta}_n}^n) = 0,$$

it suffices to determine the limiting distribution of the expression

$$X^n - \phi_n^2(\theta)\langle L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)\rangle_T.$$

To this end we shall use the following assertion (see [4]): for any $n\geq 1$ let X^n be a semimartingale w.r.t. measures P^n and $\tilde{P}^n$ with the characteristics (B^n,C^n,ν^n) and $(\tilde{B}^n,\tilde{C}^n,\tilde{\nu}^n)$ and the modified second characteristics $C^{1,n}$ and $\tilde{C}^{1,n}$, respectively.

Besides, let M be a Gaussian martingale defined on the space $(\Omega,\mathcal{F},F,P)$, with $M_0 = 0$ and $C_t = E_p(M_t)^2$, i.e. a semimartingale with characteristics $(0,C,0)$. Let D be some subset of R_+. Introduce the conditions:

(1) $I_{\{|x|>\epsilon\}} * \nu_t^n \xrightarrow{P^n} 0$ for every $t\in D, \epsilon>0$,

(2) $C_t^{1,n} \xrightarrow{P^n} C_t$ for every t.

PROPOSITION 5.1. *Suppose that $(\tilde{P}_t^n) \triangleleft (P_t^n)$ for all $t \in R$. Then if conditions (1) and (2) are satisfied we have*

$$X^n - \tilde{B}^n \xrightarrow{\mathcal{L}(D \mid \tilde{P}^n)} M,$$

where $\xrightarrow{\mathcal{L}(D \mid \tilde{P}^n)}$ denotes the weak convergence of the distributions on D.

Note that under conditions (1) and (2) we have (see [4]):

$$X^n - B^n \xrightarrow{\mathcal{L}(D \mid P^n)} M$$

We can apply this proposition to the case when

$$P^n = P_\theta^n, \ \tilde{P}^n = \tilde{P}_\theta^n,$$

$$X^n = \phi_n L(m_\theta^n, M_\theta^n), \ M = \sqrt{\Gamma(\theta)}\, W,$$

where W is a Wiener process.

The process X^n w.r.t. the measure P_θ^n is a square integrable martingale with the triplet

$$(B^n = -I_{\{|x|>1\}} * \nu_{X^n}^{P_\theta^n}, \ C_n = \phi_n^2(\theta) < m_\theta^{nc} >, \ \nu_{X^n}^{P_\theta^n}).$$

By conditions b) and d) of Theorem 2.1 it is evident that conditions (1) and (2) of Proposition 5.1 are satisfied. Hence

$$X^n - B^n \xrightarrow{\mathcal{L}(\{T\} \mid P_\theta^n)} \sqrt{\Gamma(\theta)}\, W.$$

Therefore the assertion of Proposition 5.1 implies that

$$X^n - \tilde{B}^n \xrightarrow{\mathcal{L}(\{T\} \mid \tilde{P}_\theta^n)} \sqrt{\Gamma(\theta)}\, W, \tag{5.11}$$

where $\tilde{B}^n$ is the first characteristic in the triplet of the semimartingale X^n w.r.t. the measure $\tilde{P}_\theta^n$. Now we can show that

$$\tilde{P}_\theta^n - \lim_{n \to \infty} | \tilde{B}_T^n - \phi_n^2(\theta) < L(m_\theta^n, M_\theta^n) L(N_\theta^n, M_\theta^n) >_T | = 0. \tag{5.12}$$

Define the semimartingale $\overline{X}^n$ by the relation

$$\overline{X}^n = X^n - \Sigma \Delta X^n I\{|\Delta X^n| \geq 1\}. \tag{5.13}$$

Since $\overline{X}^n$ is a special semimartingale the unique decomposition

$$\overline{X}^n = \overline{M}^n + \overline{A}^n \tag{5.14}$$

takes place with a predictable $\overline{A}^n$.

On the other hand, (5.13) leads to

$$\overline{X}^n = X^n - xI\{|x| \geq 1\} * (\mu_{X^n} - \nu_{X^n}^{P_\theta^n}) - xI_{\{|x|>1\}} * \nu_{X^n}^{P_\theta^n},$$

and, consequently,

$$\overline{M}^n = X^n - xI\{|x| \geq 1\} * (\mu_{X^n} - \nu_{X^n}^{P_\theta^n}). \tag{5.15}$$

Further, applying the traansformation formulas for the triplet of a semimartingale under absolutely continuous change of a measure we have

$$\tilde{B}^n = B^n - \phi_n(\theta) < \overline{M}^n, L(N_\theta^n, M_\theta^n) >^{P, P_\theta^n}$$

Hence,

$$\tilde{B}^n - \phi_n^2(\theta)<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)> \; = \; B^n + \phi_n(\theta)<X^n - \overline{M}^n, L(N_\theta^n,M_\theta^n)>. \tag{5.16}$$

By the Lindeberg condition and contiguity $(\tilde{P}_\theta^n) \lhd (P_\theta^n)$ we have

$$\tilde{P}_\theta^n - \lim_{n\to\infty} B_T^n = 0. \tag{5.17}$$

Thus to prove (5.12) we have only to show that

$$\tilde{P}_\theta^n - \lim_{n\to\infty} \phi_n(\theta)<X^n - \overline{M}^n, L(N_\theta^n,M_\theta^n)>_T = 0. \tag{5.18}$$

From (5.15) we have

$$X^n - \overline{M}^n \; = \; xI_{\{|x|\geq 1\}} \star (\mu_{X^n} - \nu_{X^n}^{P_\theta^n}).$$

Therefore

$$<X^n - \overline{M}^n> \; = \; x^2 I\{|x|\geq 1\} \star \nu_{X^n}^{P_\theta^n} - \Sigma(\textstyle\int xI\{|x|\geq 1\}\nu_{X^n}^{P_\theta^n}(\{s\},dx))^2.$$

Again, by the Lindeberg condition and contiguity of $(\tilde{P}_\theta^n)$ w.r.t. (P_θ^n) we have

$$\tilde{P}_\theta^n - \lim_{n\to\infty} <X^n - \overline{M}^n>_T = 0. \tag{5.19}$$

Further by the Kunita-Watanabe inequality we have

$$\phi_n^2(\theta)<X^n - \overline{M}^n, L(N_\theta^n,M_\theta^n)>_T^2 \leq <X^n - M^n>_T \phi_n^2(\theta)<L(N_\theta^n,M_\theta^n)>_T.$$

Now (5.12) follows from (5.17) and (5.19).

Hence by (5.11) and (5.12) we have

$$X^n - \phi_n^2(\theta)<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)> \overset{\mathcal{L}(\{T\}|\tilde{P}_\theta^n)}{\rightarrow} \sqrt{\Gamma(\theta)}\,W. \tag{5.20}$$

To complete the proof of the theorem it remains to show that

$$\tilde{P}_\theta^n - \lim_{n\to\infty} Z_n\beta_n \; = 0. \tag{5.21}$$

Since $\tilde{P}_\theta^n - \lim_{n\to\infty} Z_n \; = 0$, (5.21) follows from the boundedness of β_n in probability $\tilde{P}_\theta^n$.

We have

$$\phi_n^2(\theta)|<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)>_T| \leq \frac{1}{2}\phi_n^2(\theta)(<L(m_\theta^n,M_\theta^n)>_T + <L(N_\theta^n,M_\theta^n)>_T).$$

Now the required 'boundedness' follows from (5.4) and condition b) of Theorem 2.1. □

REMARK 5.2. Taking into consideration the notions introduced at the end of 2.4, Theorem 5.1 can be restated in the following way:

THEOREM 5.2. *Let $\mathfrak{S} \in \mathcal{C}_*(m_\theta^n) \in \mathfrak{M}(\mathfrak{S}) \cap \mathfrak{M}(\tilde{\mathfrak{S}})$ and let condition (5.4) be satisfied. Then assertions of Theorem 5.1 are valid.*

Under the additional ergodicity condition that there exists a deterministic 'second order shift', the estimator is asymptotically normal.

COROLLARY 5.1. *If under the conditions of Theorem 5.1 there exists a deterministic limit*

$$\beta(\theta) := \tilde{P}_\theta^n - \lim_{n\to\infty} \gamma^{-1}(\theta)\phi_n^2(\theta)<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)>_T$$

then

$$\mathcal{L}_{\tilde{P}_\theta}\{\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta)\}\Rightarrow N(\beta(\theta),\frac{\Gamma(\theta)}{\gamma^2(\theta)}).$$

5.4. Suppose that martingales m_n^θ and M_n^θ allow the integral representation (1.10) and the martingale N_θ^n also has the integral representation

$$N_\theta^n(t) = \int_0^t G_n(s,\theta)dn_s+\int_0^t\int_E q_n(s,x,\theta)(\mu-\nu)(ds,dx).$$

Then $L(N_\theta^n,M_\theta^n)$ can be represented as (for simplicity the arguments are omitted)

$$L(N_\theta^n,M_\theta^n) = G_n\cdot(n-q_n<n>)+e_n\star(\mu-\nu_\theta^n),$$

where ν_θ^n is the P_θ^n-compensator of the measure μ and

$$l_n = \frac{q_n-\hat{q}_n}{1+g_n-\hat{g}_n}-\frac{\hat{q}_n}{1-\hat{g}_n}. \tag{5.23}$$

Consequently,

$$<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)> = \psi_nG_n\cdot<n>+(\Phi_n-\hat{\Phi}^{n,\theta})(l_n-\hat{l}^{n,\theta})\star\tilde{\nu}-\Sigma\hat{\phi}^{n,\theta}\hat{l}^{n,\theta}(1-\tilde{a}_\theta^n),$$

where

$$\hat{\phi}^{n,\theta} = \int_E\phi_n(s,x)\nu_\theta^n(\{s\},dx),\quad \hat{l}^{n,\theta} = \int_E \ln(s,x)\nu_\theta^n(\{s\},dx).$$

But taking into account an easily verified equality $\frac{1-a_\theta^n}{1-\hat{g}_n} = 1-\tilde{a}_\theta^n$ and equalities (1.18) and (1.23) we can, finally, derive a formula which establishes the limiting 'second order shift'

$$\begin{aligned}\beta_n(\theta) &= \phi_n^2(\theta)<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)>_T\\ &= \phi_n^2(\theta)[\psi_nG_n\cdot n+\frac{(\psi_n-\hat{\psi}_n)(g_n-\hat{g}_n)}{1+g_n-\hat{g}_n}\star\nu^n-\Sigma\frac{\hat{\psi}_n\hat{g}_n}{1-\hat{g}_n}(1-a^n)].\end{aligned}$$

5.5. Example 1. Independent identically distributed observations.
Consider a sequence of alternative regular experiments $\tilde{\mathcal{E}}_n$ corresponding to densities $\tilde{f}(x,\theta) = f(x,\theta)+\frac{c}{\sqrt{nI(\theta)}}h(x,\theta)$ with $\int h(x,\theta)\mu(dx) = 0$, $\int(\int(h/f)^2(x)\mu(dx)<\infty$. As we know $(\tilde{P}_\theta^n)\triangleleft(P_\theta^n)$ for all $\theta\in\Theta$ (the conditions of Lemma 5.1 are trivially satisfied; cf. [4]).

In this case the representation (5.1) takes place with the martingale $N_\theta^n(t) = \sum_{i=1}^{[nt]}h(x_i,\theta), 0\leqslant t\leqslant 1$ and the conditions of Corollary 5.1 are trivially satisfied. Indeed,

$$\begin{aligned}\tilde{P}_\theta^n-\lim\phi_n^2(\theta)<L(m_\theta^n,M_\theta^n),L(N_\theta^n,M_\theta^n)>_1 &= \tilde{P}_\theta^n-\lim_{n\to\infty}\frac{1}{nI_\theta}n\int\frac{\psi(x,\theta)h(x,\theta)}{f^2(x,\theta)}f(x,\theta)\mu(dx)\\ &= \frac{1}{I_\theta}\int\frac{\psi(x,\theta)h(x,\theta)}{f(x,\theta)}\mu(dx).\end{aligned}$$

Therefore by the same Corollary 5.1

$$\mathcal{L}_{\tilde{P}_\theta}-\lim_{n\to\infty}\sqrt{n}(\hat{\theta}_n-\theta) = N(\frac{\gamma^{-1}(\theta)}{I_\theta^{\frac{3}{2}}}\int\frac{\psi(x,\theta)h(x,\theta)}{f(x,\theta)}\mu(dx),\frac{\Gamma(\theta)}{I(\theta)\gamma^2(\theta)}).$$

Example 2. Estimation of a drift parameter of a diffusion type process.
Suppose that $\tilde{a}_n(s,x,\theta) = a_n(s,x,\theta)+\phi_n(\theta)b_n(s,x,\theta),b_n$ are such that

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \tilde{P}^n_\theta \{\phi_n^2(\theta)\int_0^T b_n^2(s,x,\theta)ds>N\} = 0.$$

Then $(\tilde{P}^n_\theta)\lhd(P^n_\theta)$ and by formula (5.24)

$$\lim_{n\to\infty} \mathcal{L}_{\tilde{P}_\theta^n}(\phi_n^{-1}(\theta)(\hat{\theta}_n-\theta)-\gamma^{-1}(\theta)\phi_n^2(\theta)\int_0^T \psi_n(s,x,\theta)b_n(s,x,\theta)ds) = N(0,\frac{\Gamma(\theta)}{\gamma^2(\theta)}).$$

Example 3. Multivariate point process.
The drift is calculated here directly by formula (5.24).

6. Global asymptotic behaviour of solutions to an equation for M-estimator

6.1. The theorems presented above are of a local character; in particular, Theorem 2.1 does by no means guarantee the existence of a consistent solution to (1.11). It only states that for every fixed $\theta = \theta_0$ a solution $\{\hat{\theta}_n\},n\geq 1$ to equation (1.11) exists such that

$$P^n_{\theta_0} - \lim_{n\to\infty} \hat{\theta}_n = \theta_0.$$

In this connection the question arises whether consistent solutions to (1.11) can be constructed and moreover, whether one can guarantee the consistency of all solutions to (1.11)?

Below we give sufficient conditions for the positive answer to this question. In the course of investigating asymptotic properties of the solutions to (1.11) both for a fixed model (Theorem 2.1) and under an alternative (Theorems 3.1 and 5.1) we have used Lemma 2.1. We have specified the function $L_n(\theta)$ as the L-transformation $L_n(\theta) = L_T(m^n_\theta,M^n_\theta)$ and the measures Q^n_θ and the functions $\theta'(\theta)$ as $Q^n_\theta = P^n_\theta$, $\theta'(\theta) = \theta$ and $Q^n_\theta = \tilde{P}^n_\theta$, $\theta'(\theta) = b(\theta)$, respectively. On stipulating conditions a) and e) in Lemma 2.1 we rely upon the indicated structure of the process $L(m^n_\theta,M^n_\theta)$.

To achieve the aim of this section we have to strengthen only condition c) of Lemma 2.1. We assume below that $\Theta = (a,b)$ is an interval from R_1 and for convenience, without loosing generality, we set $a = -\infty$ and $b = +\infty$.

6.2. For every $\theta\in\Theta$ consider the set

$$S_\theta = \begin{cases} \hat{\theta} = \{\hat{\theta}_n\}_{n\geq 1}: \text{for every } n\geq 1\ \hat{\theta}_n \text{ is a random variable and} \\ Q^n_\theta - \lim_{n\to\infty} \phi_n^2(\theta)L_n(\hat{\theta}_n) = 0. \end{cases}$$

Evidently, the set S_θ includes the set S_{sol} of all solutions to the equation

$$L_n(\theta) = 0.$$

Theorem 6.1. *Let the following conditions be satisfied*
(sup$C$) *there exists a θ'-continuous function $\Delta(\theta,\theta')$ of two arguments $\theta,\theta'\in\Theta$ such that*
(A) *the equation $\Delta(\theta,\theta') = 0$ has the unique solution $\theta' = \theta'(\theta)$;*
(B) *for any $c, 0<c<\infty$ and $\rho>0$*

$$\lim_{n\to\infty} Q^n_\theta(\sup_{|\theta'|\leq c} |\phi_n^2(\theta)L_n(\theta')-\Delta(\theta,\theta')|>\rho) = 0.$$

Then
I. *the following alternative takes place: if $\hat{\theta}\in S_\theta$, then either*

$$Q^n_\theta - \lim_{n\to\infty} \hat{\theta}_n = \theta'(\theta), \tag{6.1}$$

or

$$\overline{\lim_{n\to\infty}} Q_\theta^n\{|\hat{\theta}_n|>c\}>0, \tag{6.2}$$

for any $c, 0<c<\infty$;

II. *if, in addition, the following condition holds:*

$$(C^+) \quad \lim_{|\theta'|\to\infty} |\Delta(\theta,\theta')| = c(\theta)>0,$$

and for all $\rho>0$

$$\lim_{n\to\infty} Q_\theta^n(\sup_{\theta'\in\Theta}|\phi_n^2(\theta)L_n(\theta')-\Delta(\theta,\theta')|>\rho) = 0,$$

then representation (6.1) takes place.

PROOF. Let

$$\hat{\theta} = (\hat{\theta}_n)_{n\geq 1}\in S_\theta$$

and assume that (6.2) is not satisfied. Then a number $c_0>0$ can be found such that

$$\lim_{n\to\infty} Q_\theta^n\{|\hat{\theta}_n|>c_0\} = 0.$$

We have

$$\begin{aligned} Q_\theta^n\{|\phi_n^2(\theta)L(\hat{\theta}_n)-\Delta(\theta,\hat{\theta}_n)|>\rho\} &\leq Q_\theta^n\{|\hat{\theta}_n|>c_0\} \\ &+Q_\theta^n\{|\phi_n^2(\theta)L_n(\hat{\theta}_n)-\Delta(\theta,\hat{\theta}_n)|>\rho, |\hat{\theta}_n|\leq c_0\} \\ &\leq Q_\theta^n\{|\hat{\theta}_n|>c_0\}+Q_\theta^n\{\sup_{|\theta'|\leq c_0}|\phi_n^2(\theta)L_n(\theta')-\Delta(\theta',\theta)|>\rho\}\to 0. \end{aligned}$$

On the other hand,

$$Q_\theta^n - \lim_{n\to\infty}\phi_n^2(\theta)L(\hat{\theta}_n) = 0$$

and, hence,

$$Q_\theta^n - \lim_{n\to\infty}\Delta(\theta,\hat{\theta}_n) = 0. \tag{6.3}$$

Assume now that (6.1) is not satisfied. Then $\epsilon>0$ can be found such that

$$\overline{\lim_{n\to\infty}} Q_\theta^n\{|\hat{\theta}_n-\theta|>\epsilon\}>0.$$

By condition (A) for every $\epsilon>0$ we have

$$\Delta(\epsilon) = \inf_{\theta':|\theta'-\theta|\geq\epsilon, |\theta'|\leq c_0} |\Delta(\theta,\theta')|>0.$$

Therefore

$$\begin{aligned} \overline{\lim_{n\to\infty}} Q_\theta^n\{|\Delta(\theta,\hat{\theta}_n)|>\Delta(\epsilon)\} &\geq \overline{\lim_{n\to\infty}} Q_\theta^n\{|\Delta(\theta,\hat{\theta}_n)|>\Delta(\epsilon), |\hat{\theta}_n|\leq c_0\} \\ &\geq \overline{\lim_{n\to\infty}} Q_\theta^n\{|\hat{\theta}_n-\theta|>\epsilon, |\hat{\theta}_n|\leq c_0\}>0, \end{aligned}$$

which contradicts (6.3).

Taking into consideration that under condition (C^+)

$$\inf_{\theta':\theta'\in\Theta\setminus(\theta-\epsilon,\theta+\epsilon)} |\Delta(\theta,\theta')|>0$$

and argueing as above, we easily arrive at the last assertion of the theorem. □

COROLLARY 6.1. I. *If conditions a), b), d), e) of Lemma 2.1 and conditions (A) and (B) are satisfied for any* $\theta\in\Theta$, *then there exists the solution* $\hat{\theta} = (\hat{\theta}_n)_{n\geq 1}\in S_{sol}$ *such that*

$$Q_\theta^n - \lim_{n\to\infty}\hat{\theta}_n = \theta'(\theta) \tag{6.4}$$

for any $\theta\in\Theta$.
II. Besides, if condition (C^+) *is satisfied for any* $\theta\in\Theta$, *then any sequence* $\hat{\theta} = (\hat{\theta}_n)_{n\geq 1}\in S_{sol}$ *has the same property (6.4).*

PROOF. It suffices to construct a solution $\hat{\theta} = (\hat{\theta}_n)_{n\geq 1}$ for which (6.2) does not take place for any $\theta\in\Theta$. For any $n\geq 1$ and $\epsilon>0$ one can choose θ_n^* such that

$$|\theta_n^*|\leq \operatorname*{ess\,inf}_{\hat{\theta}\in S_{sol}} |\hat{\theta}_n|+\epsilon \quad Q_\theta - a.s.$$

By Lemma 2.1 for any $\theta\in\Theta$ there exists an estimator $\hat{\theta}_n = \hat{\theta}_n(\theta)$ such that

$$Q_\theta^n\{L_n(\hat{\theta}_n(\theta)) = 0\}\underset{n\to\infty}{\to} 1 \tag{6.5}$$

and

$$Q_\theta^n - \lim_{n\to\infty}\hat{\theta}_n(\theta) = \theta'(\theta). \tag{6.6}$$

For every $c, 0<c<\infty$ we have

$$\overline{\lim_{n\to\infty}}\, Q_\theta^n\{|\theta_n^*|\geq c\}\leq \overline{\lim_{n\to\infty}}\, Q_\theta^n\{|\theta_n^*|\geq c, L_n(\hat{\theta}_n(\theta))\neq 0\}+\overline{\lim_{n\to\infty}}\, Q_\theta^n\{|\theta_n^*|\geq c, L_n(\hat{\theta}_n(\theta)) = 0\}\leq$$
$$\leq \overline{\lim_{n\to\infty}}\, Q_\theta^n\{L_n(\hat{\theta}_n(\theta))\neq 0\}+\overline{\lim_{n\to\infty}}\, Q_\theta^n\{|\hat{\theta}_n(\theta)|+\epsilon\geq c\}.$$

The first term on the right-hand side of the last relation tends to zero by (6.5), and the second one by (6.6). □

6.3. As in Theorems 2.1 and 3.1 we can stipulate conditions sufficient for (sup$C$) of Theorem 6.1, for instance, in the case of a fixed model, i.e. $L_n(\theta) = L_T(m_\theta^n, M_\theta^n)$ where $(m_\theta^n)\in\mathfrak{M}(\mathfrak{S}), \theta'(\theta) = \theta$, $Q_\theta^n = P_\theta^n$.

Note that for every $n\geq 1$ and $z\in\Theta$ the process $L_n(z)$ is a semimartingale w.r.t. the measure P_θ^n, so that sufficient conditions for uniform ergodicity w.r.t. the parameter z (condition (sup$C$)) can be naturally formulated as certain restrictions imposed on the component with finite variation and on the martingale component of the semimartingale $L_n(z)$. We shall give two types of such conditions corresponding to two decompositions (1.23) and (1.24).

THEOREM 6.2. *Let conditions of Theorem 2.1 hold, as well as conditions* (B'): *for any* $c>0, \rho>0$,

(B'_1) $\overline{\lim_{n\to\infty}}\,\phi_n^4(\theta)\int_{|\theta-z|<c} E_\theta^n<L(m_z^n, M_\theta^n)>_T dz = 0;$

(B'_2) $\overline{\lim_{n\to\infty}}\, P_\theta^n\{\sup_{|z-\theta|<c}|\phi_n^2(\theta)[L(m_z^n, M_z^n), L(M_z^n - M_\theta^n, M_\theta^n)]_T - \Delta(\theta, z)|>\rho\} = 0.$

Then there exists the consistent M-estimator $\hat{\theta} = \{\hat{\theta}_n\}, n\geq 1$ *such that*

$$L_{P_\theta^n} - \phi_n^{-1}(\theta)(\hat{\theta}_n - \theta)\to N(0, \Gamma(\theta)/\gamma^2(\theta)).$$

PROOF. By the decomposition (1.24) we have

$$L(m_z^n, M_z^n) = L(m_z^n, M_\theta^n)+[L(m_z^n, M_z^n), L(M_z^n - M_\theta^n, M_\theta^n)].$$

Evidently, it suffices to prove that condition (B) implies

$$\overline{\lim_{n\to\infty}} P_\theta^n(\sup_{|z-\theta|<c} |\phi_n^2(\theta) L_T(m_z^n, M_\theta^n)| > \rho) = 0.$$

Indeed, argueing as in the course of establishing conditions sufficient for Theorem 2.1, we arrive at the inequality

$$P_\theta^n(\sup_{|z-\theta|<c} |\phi_n^2(\theta) L_T(m_z^n, M_\theta^n)| > \rho) \leq \frac{2}{\rho^2}\phi_n^4(\theta) c \int_{|z-\theta|<c} <L(\dot{m}_z^n, M_\theta^n)>_T dz$$
$$+\frac{2}{\rho_2}\phi_n^4(\theta) E_\theta^n <L(m_\theta^n, M_\theta^n)>_T.$$

The second term in the right-hand side of (6.7) tends to zero by condition b) of Theorem 2.1. The assertion of the theorem follows now from Corollary 6.1. □

REMARK 6.1. Typically, it is hard to verify the uniform ergodicity (as condition (B'_2) involves the unpredictable process $[L(m_z^n, M_z^n), L(M_z^n - M_\theta^n, M_\theta^n)]$). The decomposition (1.23) allows us to present slightly different conditions (B''): for any $c>0, \rho>0$

(B''_1) $\overline{\lim_{n\to\infty}} \phi_n^4(\theta) \int_{|z-\theta|<c} E_\theta^n <\dot{L}(m_z^n, M_z^n, M_\theta^n)>_T dz = 0;$

(B''_2) $\overline{\lim_{n\to\infty}} P_\theta^n(\sup_{|z-\theta|<c} |\phi_n^2(\theta) <L(m_z^n, M_z^n), L(M_\theta - M_z, M_z)>_T - \Delta(\theta, z)| > \rho) = 0,$

where $\dot{L}(m_z^n, M_z^n, M_\theta^n)$ is the derivative w.r.t. z of the martingale component in the decomposition (1.23) with $m = m_z^n, M = M_z^n, \tilde{M} = M_\theta^n, P = P_\theta^n$.

6.4. We will check conditions (B') and (B'') in the case of independent identically distributed observations. Using the notations of § 4, Example 1, we have

$$[L(m_z^n, M_z^n), L(M_z^n - M_\theta^n)]_t = \sum_{i=1}^{[nt]} \phi(x_i, \theta, z),$$

where

$$\phi(x, \theta, z) = \frac{\psi(x,z)}{f(x,z)} \frac{f(x,z) - f(x,\theta)}{f(x,\theta)},$$

$$E_\theta^n <L(\dot{m}_z^n, M_\theta^n)>_1 = n \int \frac{\dot{\psi}^2(x,z)}{f(x,\theta)} \mu(dx),$$

$$\dot{L}_t(m_z^n, M_z^n, M_\theta^n) = \sum_{i=1}^{[nt]} (\psi(x_i, z) - \int \psi(x,z) f(x,\theta) \mu(dx))$$

with

$$\psi(x,z) = \frac{\dot{\psi}(x,z) f(x,z) - \psi(x,z) \dot{f}(x,z)}{f^2(x,z)},$$

and

$$\Delta(\theta, t) = \int \Phi(x, \theta, z) f(x, \theta) \mu(dx) = \frac{1}{nI(\theta)} <L(m_z^n, M_\theta^n), \mathcal{L}(M_\theta^n - M_z^n, M_z^n)>_T.$$

Hence, condition (B''_2) is trivially satisfied and therefore conditions (B') and (B''), respectively, take the form: for any $c>0, \rho>0$,

(B'_1) $\int_{|z-\theta|<c} \int \frac{\dot{\psi}^2(x,z)}{f(x,\theta)} \mu(dx) dz < \infty,$

(B'_2) $\overline{\lim_{n\to\infty}} P_\theta(\sup_{|z-\theta|<c} |\frac{1}{n}\sum_{i=1}^{[nt]} \phi(x_i, \theta, z) - \Delta(\theta, z)| > \rho) = 0$

$(B'_3)\ \int_{|z-\theta|<c} \int (\psi(x,z)-\int\psi(x,z)f(x,\theta)\mu(dx))^2 f(x,\theta)\mu(dx)dz<\infty$

One can avoid conditions (B') by taking into consideration the sufficient condition for uniform ergodicity proposed in [10]: this leads, specifically, to the condition

$(B'_{2'})\ \int \sup_{z:|z-\theta|<c} |\Phi(x,\theta,z)| f(x,\theta)\mu(dx)<\infty$

for each $c>0$, which is sufficient for (B'_2).

7. Comments and historical remarks

7.1. The general scheme of statistical experiments is considered in Ibragimov and Has'minskii [11]. The main object studied in this book, is a normalized likelihood ratio

$$Z_\theta^n(u) = \frac{\rho_{\theta+\phi_n u}^n(T)}{\rho_\theta^n(T)}$$

The method for investigating asymptotic properties of parameter estimators developed originally by Le Cam [12] and Hajek [13], is based on a certain behaviour of this ratio or some related parameter dependent random variables (conditions of the local asymptotic normality (LAN) type, Helder type w.r.t. the parameter, etc.) (see [27] for a survey of asymptotic theory of statistical decisions).

However, on considering particular schemes of models (i.i.d. observations, diffusion processes, point processes [5, 14, 7, 15, 16, 17, 18]), one needs conditions imposed on local characteristics of the processes under study (on marginal densities, drift coefficients and intensities of jumps, respectively), which guarantee the desired behaviour of $Z_\theta^n(u)$.

Treating separately each special scheme, one quickly notices certain common features of reasonings: each time the problem is reduced to imposing useful conditions (which are wide enough) on the characteristics of martingales M_θ^n (envolved in the exponential representation of likelihood ratio processes $\rho_\theta^n(t), 0\leq t\leq T$). Moreover, the methods of investigating each particular scheme can be classified in terms developed within special branches of martingale theory (such as the branches treating functional limit theorems, theorems on absolute continuity and contiguity, etc.)

In the present paper the asymptotic theory of estimation in viewed in this light (in the spirit of the earlier works [19]-[21]). In order to mitigate technical difficulties we restrict the considerations to the 'smooth' case. The regularity conditions imposed on the scheme of models can be considered as a generalization of the conditions of the classical Dugue-Cramer [22] method.

7.2. Our approach leads naturally to the consideration of Eq. (1.11) for defining general type of estimators, the so called M-estimators. Such estimators were treated by Huber [9] in studying robustness under model disturbance. We do not treat the questions risen in this field. Observe nevertheless that the asymptotic behaviour of M-estimators under model disturbance, investigated here, lies on the basis for the consequent study of construction problems of the robust estimators (see, e.g., [13], [24] for robust estimators).

Expressing the main equation (1.11) in terms of L-transformations, we are able to formulate elegantly our theorems. In the applications, however (see the examples in § 4,5,6), we give integral representations of the martingales M_θ^n defining the models.

Conclusions about the asymptotic behaviour of M-estimators take a rather complete character at least under contigual alternatives (§ 5) (in the general scheme of the so called 'contamination model' due to Huber [9]; for particular schemes see, e.g. [25,26]).]

7.3. We devote § 6 to the global asymptotic behaviour of M-estimators. This is motivated by the following observation. Under the undoubted influence of Cramer's works, the statements concerning the asymptotic behaviour of the estimators defined as solutions to certain equations, in particular, to the likelihood equations, are of a local character (cf. § 2-5 of the present paper) in the sense that for eached fixed parameter value the existance of a solution to the equation under consideration is stated which tends to a given parameter value (in other words, the identifiability of a parameter is proved rather then the existence of a consistent estimator). The consistency of the MLE is usually proved by the LAN technique where MLE is understood as the maximizor of the likelihood ratio.

From the practical point of view, however, it seems interesting to investigate the consistency of (all) solutions of (1.11).

In particular cases these problems were treated by WALD, LE CAM, KENDALL and STUART, etc. PERLMAN [10] presents a short historical review of the problem, as well as sufficiently wide conditions for the consistency of all solutions of the likelihood equation based on i.i.d. observations.

The sufficient conditions presented in this paper (of uniform ergodicity) are slightly more restrictive in the classical scheme then those of Perlman. This is caused, actually, by the fact that they are established for the general equation (1.11) rather then for its particular case - the likelihood equation.

REFERENCES

1. R.J. ELLIOTT (1982). *Stochastic Calculus and Applications.* New York, Springer.
2. J. JACOD (1979). Calcul stochastique et problemes de martingales. *Lect. Notes Math., 714, Springer Verlag.*
3. R. SH. LIPTSER, A.N. SHIRYAYEV (1989). *Theory of Martingales.* Dordrecht, Kluwer Acad. Publishers.
4. J. JACOD, A.N. SHIRYAYEV (1987). *Limit Theorems for Stochastic Processes.* New York, Springer.
5. A. LE BRETON (1976). *Dissertation,* Grenoble.
6. P. BILLINGSLEY (1968). *Convergence of Probability Measures.* New York, Wiley.
7. YU.A. KUTOYANTS (1984). *Parameter Estimation for Stochastic Processes.* Berlin, Hedermann Verlag.
8. YU.A. KUTOYANTS (1986). On parametric model choice for a diffusion process. *Abstr. of Commun. XX school in Probab. Th. and Math. Stat.* Tbilisi. (Russian)
9. P. HUBER (1984). *Robust Statistics.* New York, Wiley.
10. H.D. PERLMAN (1983). *The Limiting Behaviour of Multiple Roots of the Likelihood Equation,* In: Recent Advances in Statistics. Papers in Honor of H. Chernoff. N.Y. Academic Press, 339-370.
11. I.A. IBRAGIMOV, R.Z. KHAS'MINSKII (1981). *Statistical Estimation: Asymptotic Theory.* New York, Springer.
12. L. LE CAM (1960). Locally asymptotically normal families of distributions. *Univ. Calif. Publ. Statist., 3,* 27-98.
13. J. HAJEK (1970). A characterization of limiting distribution of regular estimates. *Z. Wahrscheinlichkeitstheorie verw. Geb., 14,* 323-330.
14. A. LE BRETON (1977). Parameter estimation in a linear stochastic differential equation. *Trans. of the Seventh Prague Conf. on Inform. Theory* (Prague, 1974), v. A, Dordrecht, Reidel, 353-366.
15. YU.N. LIN'KOV (1982). Asymptotic properties of statistical estimators and tests for Markov processes. *Theory Probab. and Math. Statist.* N25, 83-98.
16. A.F. TARASKIN (1984). On the behaviour of the likehood ratio of semimartingales. *Theory Probab. Appl.,* v. 29, 452-464.
17. L. VOSTRIKOVA (1985). On a weak convergence of likelihood ratio processes of general statistical models. In: *Intern. Vilnius Conf. on Probab. Th. and Math. Statist., Abstr. of Commun., v. 3,* Vilnius, p. 331-332.
18 K. DZHAPARIDZE (1986). On LAN for counting processes. Report MS-8606, *Centre for Mathematics and Computer Sciences,* Amsterdam.
19. N.L. LAZRIEVA, T.A. TORONJADZE (1986). Asymptotic properties of maximal likelihood estimator in a general scheme of statistical models. *Soobsh. AN SSSR, v. 123,* 25-28 (Russian).
20. N.L. LAZRIEVA, T.A. TORONJADZE (1986). Ito-Ventzel's formula for semimartingales, asymptotic properties of MLE and recursive estimation. In: *Stochastic Differential Systems. Proceedings of the*

Fifth IFIP Working Conf. (Aisenach). Lect. Notes Math.

21. N.L. Lazrieva, T.A. Toronjadze (1986). Joint asymptotic distribution of the maximal likehood estimators and *M*-estimators. In: *Proceedings of the Fifth Japan-USSR Symp. on Prob. Th.* (Kyoto). Lect. Notes. Math.
22. H. Cramer (1946). *Mathematical Methods of Statistics.* Prinston U.P.
23. P.J. Bickel. Quelques aspects de la statistiques robuste. *Lect. Notes. Math. B. 876,* 1-72.
24. P.J. Bickel (1982). On adaptive estimation. *Ann. Statist., v. 10,* 647-671.
25. G.V. Srogovitch (1985). Nonparametric estimation of linear regression coefficients. *Reports in Applied Math. VC Acad. of Science of USSR,* Moscow (Russian).
26. R.J. Beran (1978). An efficient and robust adaptive estimator of location. *Ann. Statist., v. 6,* 292-313.
27. H. Strasser (1985). *Mathematical Theory of Statistics.* Berlin, de Gruyter.