Publications mathématiques et informatiques de Rennes

L. Vostrikova

Divergence Processes and Weak Convergence of Likelihood Ratio Processes

Publications de l'Institut de recherche mathématiques de Rennes, 1989-1990, fascicule 1 « Probabilités », , p. 134-146

<http://www.numdam.org/item?id=PSMIR_1989-1990___1_134_0>

DIVERGENCE PROCESSES AND WEAK CONVERGENCE OF LIKELIHOOD RATIO PROCESSES

L. VOSTRIKOVA
I.R.M.A.R.
U.F.R. Mathématiques
Campus de Beaulieu
35042 RENNES CEDEX

Abstract

We use divergence processes introduced by Valkeila and Dzhaparidze [1] to generalize our criteria of weak convergence of likelihood ratio processes for general statistical parametric models, given in [16].

Key words : general statistical parametric model, general parametric experiment, Gaussian statistical model, statistical model generated by processes with independent increments, likelihood ratio processes, Hellinger processes, divergence processes.

1. Introduction

Let $\mathcal{E}^n = \{(\Omega^n, F^n, \mathbb{F}^n), (P^n_\theta)_{\theta\in\Theta}\}$ be a sequence of general statistical parametric models (or general parametric experiments) where $(\Omega^n, F^n, \mathbb{F}^n)$ is a measurable filtered space (Ω^n, F^n) with a right-continuous filtration $\mathbb{F}^n = (F^n_t)_{t\geq 0}$, $\vee_{t\geq 0} F^n_t = F^n$, $(P^n_\theta)_{\theta\in\Theta}$ is a family of probability measures on (Ω^n, F^n) which depend on the parameter θ belonging to some closed convex set $\Theta \subseteq R^m (m \geq 1)$ depending may be on n and containing $\{0\}$.

Assuming that P^n_θ is absolutely continuous with respect to some σ-finite measure μ^n, $P^n_\theta \ll \mu^n$, for each $\theta \in \Theta$, and denoting by $P^{n,t}_\theta, P^{n,t}_0, \mu^{n,t}$ the restrictions of the measures P^n_θ, P^n_0, μ^n to the σ-algebra F^n_t, we introduce the likelihood ratio process $Z^n = (z^n_t(\theta))_{t\geq 0,\theta\in\Theta}$ of $\mathcal{E}^n$ with

$$z^n_t(\theta) = \frac{dP^{n,t}_\theta / d\mu^{n,t}}{dP^{n,t}_0 / d\mu^{n,t}}$$

where by convention $0/0 = 0$, $a/0 = \infty$ if $a > 0$.

Let $\mathcal{E} = \{(\Omega, F, \mathbb{F}), (P_\theta)_{\theta\in R^m}\}$ be a general parametric model generated by a process with independent increments, with locally equivalent measures $P_\theta \overset{\text{loc}}{\sim} P_{\theta'}$ for all $\theta, \theta' \in R^m$. We denote by $Z = (z_t(\theta))_{t\geq 0,\theta\in R^m}$ the likelihood ratio process of $\mathcal{E}$ with

$$z_t(\theta) = dP^t_\theta / dP^t_0.$$

We recall (see JACOD [4]) that this model is such that for each $k \geq 1$ the k-dimensional process $(\log Z(\theta_1), \log Z(\theta_2), \dots, \log Z(\theta_k))$ with $\log Z(\theta_i) = (\log z_t(\theta_i))_{t\geq 0}$ and $\theta_i \in R^m, 1 \leq i \leq k$, is a process with independent increments under each $P_{\theta'}$, $\theta' \in R^m$. We recall also that the model $\mathcal{E}$ is called continuous Gaussian if in addition $\log Z(\theta)$ is a continuous (a.s.) in t Gaussian process for each $\theta \in R^m$ under each $P_{\theta'}$, $\theta' \in R^m$.

The first results about weak convergence of the likelihood ratio processes were obtained for binary statistical models (i.e. models with only two measures) when the limiting model is continuous Gaussian, by GREENWOOD and SHIRYAYEV [2] in discrete time, by KORDZAHIA [7] and VOSTRIKOVA [15] in continuous time.

Then, these results were generalized for a limiting model $\mathcal{E}$ generated by a process with independent increments : in binary cas by MEMIN [12], for the models with a finite number of measures by JACOD [5].

One can see that the number of the measures in $\mathcal{E}^n$ is infinite, in general. In [16] we gave the conditions for having weak convergence of Z^n to Z when $\mathcal{E}$ is continuous Gaussian.

The aim of this paper is to extend our results of [16] to the cas of a limiting model generated by a process with independent increments. The main results are formulated in Theorems 4.2, 4.3. In corollaries 4.4, 4.5, 4.6 we consider the important particular cas when the limiting model is continuous Gaussian. In corollary 4.7 we give results for a Poisson limiting case.

We begin by recalling the necessary facts on Hellinger processes and divergence processes.

2. Basic facts on Hellinger processes and divergence processes

Let $(\Omega, F, \mathbb{F})$ be a measurable filtered space, i.e. a measurable space (Ω, F) with a filtration $\mathbb{F} = (F_t)_{t\geq 0}$ such that $F_t \subseteq F_{t'}$, if $t \leq t'$, $F_{t+} = F_t$ for each $t \geq 0$, $\vee_{t\geq 0} F_t = F$.

Let $(P_\theta)_{\theta\in\Theta}$ be probability measures on (Ω, F) dominated by a probability measure Q, $\Theta = \{\theta_0, \theta_1, \dots, \theta_k\}$, $k \geq 1$. We denote by $P_\theta^t = P_\theta / F_t$, $Q^t = Q/F_t$ the restrictions of the measures P_θ and Q to the σ-algebra F_t and we introduce the likelihood process $V(\theta) = (V_t(\theta))_{t\geq 0}$ with

$$V_t(\theta) = dP_\theta^t / dQ^t.$$

We will consider below a cadlag version of this process.

We recall the definition of the Hellinger process h^α of order α of the measures $(P_{\theta_0}, P_{\theta_1}, \dots, P_{\theta_k})$ where α is a multi-index, $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)$ with $0 < \alpha_i < 1$ for $0 \leq i \leq k$ and $\sum_{i=0}^{k} \alpha_i = 1$.

Definition 2.1. *(see* JACOD *[4]). The Hellinger process $h^\alpha = (h_t^\alpha)_{t\geq 0}$ of order α of the measures $(P_{\theta_0}, P_{\theta_1}, \dots, P_{\theta_k})$ is a process which satisfies the following conditions :*

1) $h_0^\alpha = 0$,

2) it is increasing and predictable,

3) the process $M = (M_t, F_t)_{t\geq 0}$ with

$$M_t = \prod_{i=0}^{k} V_t^{\alpha_i}(\theta_i) + \int_0^t \prod_{i=0}^{k} V_{s-}^{\alpha_i}(\theta_i) dh_s^\alpha$$

is Q-martingale.

From JACOD [4] we know that the Hellinger process h^α of order α exists and that it is unique on the set

$$\Gamma = \{(\omega, t) : \prod_{i=0}^{k} V_{t-}(\theta_i) > 0\} \cup [0].$$

It can be calculated by the formula :

$$\begin{aligned} h_t^\alpha &= \frac{1}{2}\sum_{i=0}^{k} \alpha_i \frac{1}{V_-^2(\theta_i)} \circ \langle V(\theta_i)\rangle \\ &- \frac{1}{2}\sum_{i=0}^{k}\sum_{j=0}^{k} \alpha_i\alpha_j \frac{1}{V_-(\theta_i)V_-(\theta_j)} \circ \langle V(\theta_i), V(\theta_j)\rangle + \phi_\alpha(\overline{y}) * \nu_{\overline{V}} \end{aligned}$$

where $\langle V(\theta_i)\rangle$, $\langle V(\theta_i), V(\theta_j)\rangle$ are the angle brackets of $V(\theta_i)$ and $(V(\theta_i), V(\theta_j))$, $\nu_{\overline{V}}$ is the compensator of the jump-measure of $\overline{V} = (V(\theta_0), V(\theta_1), \dots, V(\theta_k))$ with respect to $(\mathbb{F}, Q)$, $\circ$ and $*$ are two types of Lebesgue-Stieltjes integrals, the first one with respect to an increasing process, the second one with respect to a random measure, and

$$\phi_\alpha(\overline{y}) = \sum_{i=0}^{k} \alpha_i y_i - \prod_{i=0}^{k} (y_i)^{\alpha_i}$$

with $\overline{y} = (y_0, y_1, \dots, y_k)$, $y_i = 1 + x_i/V_-(\theta_i)$, $0 \le i \le k$.

Remark 1 : If $k = 1$ we obtain the usual definition of the Hellinger process of order α, $0 < \alpha < 1$ for two measures (see JACOD, SHIRYAYEV [6]).

Example 1 : Let μ_θ^j be a probability measure on the measurable space (Ω_j, F_j), $j \ge 1$. Let

$$\Omega = \prod_{j=1}^{\infty} \Omega_j, \; F = \prod_{j=1}^{\infty} F_j, \; F_t = \prod_{j=1}^{[t]} F_j$$

and

$$P_\theta = \prod_{j=1}^{\infty} \mu_\theta^j.$$

Then one of the versions of h^α is given by :

$$h_t^\alpha = \sum_{j=1}^{[t]} \{1 - H^\alpha(\mu_{\theta_0}^j, \mu_{\theta_1}^j, \dots, \mu_{\theta_k}^j)\}$$

where $H^\alpha(\cdot, \cdots, \cdot)$ is the Hellinger integral of order α of the corresponding measures. □

Example 2 : Let (Ω, F) be the measurable space of piece-wise constant right-continuous functions $X = (X_t)_{t\ge 0}$ with jumps ΔX_t of size 0 or 1. Let $F_t = \sigma\{X_s | s \le t\}$ for every $t \ge 0$ and $F = \vee_{t\ge 0} F_t$. If P_θ is the measure corresponding to a Poisson process with intensity $\lambda_t(\theta) \ge 0$, then one of the versions of the Hellinger process h^α is given by :

$$h_t^\alpha = \int_0^t \left\{ \sum_{i=0}^{k} \alpha_i \lambda_s(\theta_i) - \prod_{i=0}^{k} \lambda_s^{\alpha_i}(\theta_i) \right\} ds.$$

□

Example 3 : Let (Ω, F) be the measurable space of continuous functions $X = (X_t)_{t\ge 0}$ with $F_t = \sigma\{X_s | s \le t\}$ and $F = \vee_{t\ge 0} F_t$.

Let P_θ, $\theta \in \Theta$, be the unique probability measure corresponding to the diffusion process

$$dX_t = a_t(X,\theta) + dW_t,\ X_0 = 0,$$

where $W = (W_t)_{t\geq 0}$ is a Wiener process, $a_t(X,\theta)$ is a non-anticipating functional, $\int_0^t a_s^2(X,\theta)ds < \infty$ for each $\theta \in \Theta$, $X \in \Omega, t \geq 0$. Then one of the versions of the Hellinger process h^α is given by

$$h_t^\alpha = \frac{1}{2}\sum_{i=1}^k \alpha_i \int_0^t a_s^2(X,\theta_i)ds - \frac{1}{2}\sum_{i=1}^k\sum_{j=1}^k \alpha_i\alpha_j \int_0^t a_s(X,\theta_i)a_s(X,\theta_j)ds$$

□

We begin now to discuss the notion of divergence process introduced by DZHAPARIDZE and VALKEILA [1].

Suppose that we are given the measurable filtered space $(\Omega, F, \mathbb{F})$ equipped with two probability measures P and $\tilde{P}$. We denote as before by $V = (V_t)_{t\geq 0}$ and $\tilde{V} = (\tilde{V}_t)_{t\geq 0}$ the cadlag versions of the likelihood processes of P and $\tilde{P}$ with respect to a majorating measure Q.

We know that the Hellinger process of order $\frac{1}{2}$ of the probability measures P and $\tilde{P}$ at $t > 0$ behaves like the Hellinger distance squared $\rho_2^2(P^t, \tilde{P}^t) = E_Q(\sqrt{V_t} - \sqrt{\tilde{V}_t})^2$ (see VALKEILA, VOSTRIKOVA [14]). We know also that if the dimension m of the parameter space of θ is big, the smoothness conditions in θ in terms of the Hellinger distance could be very restrictive. So, it leads to consider the following distances :

$$\rho_p(P^t, \tilde{P}^t) = \{E_Q|V_t^{1/p} - \tilde{V}_t^{1/p}|^p\}^{1/p},\ p \geq 1,$$

used by many autors (see, for example, IBRAGIMOV, HASMINSKIJ [3], KUTOYANTS [8], [9], [10], LIESE and VAJDA [11], VOSTRIKOVA [16]) and called some times p-divergence. Note that for $p = 1$ we obtain the variation distance, for $p = 2$ we have the Hellinger distance.

This distance has three advantages :

a) choosing p big enough we can obtain a distance with a very smooth behaviour,

b) we do not need to ask the existence of hight order moments of $V_t, \tilde{V}_t$,

c) this distance does not depend on the majorating measure Q.

But working with the distance $\rho_p(P^t, \tilde{P}^t)$ is not so easy. To simplify this, DZHAPARIDZE and VALKEILA [1] have introduced divergence processes of order p.

Definition 2.2. *The predictable increasing process $k^p = (k_t^p, F_t)_{t\geq 0}$ with $k_0^p = 0$ and*

$$k_t^p = \left|(1+\frac{x}{V_-})^{1/p} - (1+\frac{y}{\tilde{V}_-})^{1/p}\right|^p * \nu_{V,\tilde{V}}$$

where $\nu_{V,\tilde{V}}$ is the compensator of the jump-measure of $(V,\tilde{V})$ with respect to $(\mathbb{F}, Q)$, is called the divergence process of order $p, p \geq 1$, of the measures $P, \tilde{P}$.

Note that, like the Hellinger process of order $\alpha, 0 < \alpha < 1$, the divergence process of order $p \geq 1$ is defined and is unique only on the set $\Gamma = \{(\omega, t) : V_{t-} \cdot \tilde{V}_{t-} > 0\} \cup [0]$.

Example 4 : In the situation of example 1 when $P = \prod_{i=1}^{\infty} \mu_i$, $\tilde{P} = \prod_{i=1}^{\infty} \tilde{\mu}_i$ we have that a version of the divergence process is given by :

$$k_t^p = \sum_{i=1}^{[t]} \rho_p^p(\mu_i, \tilde{\mu}_i).$$

□

Example 5 : In the situation of example 2 where P and $\tilde{P}$ correspond to Poisson processes with intensities $\lambda_s \geq 0$ and $\tilde{\lambda}_s \geq 0$ we have

$$k_t^p = \int_0^t |\lambda_s^{1/p} - \tilde{\lambda}_s^{1/p}|^p ds.$$

□

Remark 2 : One can see that $k_t^p = 0$ for all $t \geq 0$ and $p \geq 1$ in example 3 because V and $\tilde{V}$ have no jump.

3. Skorohod space $D(R^+, C_{\text{loc}}(R^m))$

We consider the space $D(R^+, C_{\text{loc}}(R^m))$ of the right-continuous functions $z = (z_t(\theta))_{t \geq 0, \theta \in R^m}$ with left-hand limits and with values in the space $C_{\text{loc}}(R^m)$ of the continuous functions endowed with the locally uniform distance :

$$d_{\text{loc}}(x, y) = \sum_{\ell=1}^{\infty} \frac{1}{2^\ell} \cdot \frac{\max_{||\theta|| \leq \ell} |x(\theta) - y(\theta)|}{1 + \max_{||\theta|| \leq \ell} |x(\theta) - y(\theta)|}$$

where $|| \cdot ||$ in the Euclidean norm in R^m. We consider also the subspace D_0 of the space $D(R^+, C_{\text{loc}}(R^m))$ which contains the functions $z = (z_t(\theta))_{t \geq 0, \theta \in R^m}$ of $D(R^+, C_{\text{loc}}(R^m))$ such that for each $N > 0$

$$\gamma_L^N(z) = \sup_{1/N \leq t \leq N} \sup_{||\theta|| \geq L} z_t(\theta)$$

tend to zero as $L \to \infty$.

Lemma 3.1. *The function $z = (z_t(\theta))_{t \geq 0, \theta \in R^m}$ belongs to $D(R^+, C_{loc}(R^m))$ iff*
a) for each $\theta \in R^m$, $z(\theta) = (z_t(\theta))_{t \geq 0}$ belongs to $D(R^+, R)$
b) for each $\ell > 0$ and $N > 0$

$$\kappa_h^{\ell,N}(z) = \sup_{0 \leq t \leq N} \sup_{||\theta - \theta'|| \leq h, ||\theta|| \leq \ell, ||\theta'|| \leq \ell} |z_t(\theta) - z_t(\theta')|$$

tends to zero as $h \to 0$.

Proof : We set $K_\ell = \{||\theta|| \leq \ell\}$ and we denote by $z^\ell = (z_t(\theta))_{0 \leq t < N, \theta \in K_\ell}$ the restrictions of z to $D([0, N[, C(K_\ell))$. First we note that $z \in D(R^+, C_{\text{loc}}(R^m))$ iff $z^\ell \in D([0, N[, C(K_\ell))$ for each $\ell \geq 1$ and $N > 0$.

Suppose that the conditions a) and b) are satisfied. We have to show that the modulus of continuity $w_h^{\ell,N}(z^\ell)$ associated with $D([0, N[, C(K_\ell))$ tends to zero as $h \to 0$. We recall that

$$w_h^{\ell,N}(z) = \inf_{T_h} \max_{0\le i\le n} \sup_{t,t'\in[t_i,t_{i+1}[} \max_{\theta\in K_\ell} |z_t(\theta) - z_{t'}(\theta)| \tag{1}$$

where inf is taken over all subdivisions $T_h = \{0 = t_0 < t_1 < t_2 < \cdots < t_{n+1} = N\}$ of the interval $[0, N[$ where $t_{i+1} - t_i > h$ for $0 \le i \le n-1$.

For that we cover the set K_ℓ by a finite number $N_{h_1} = N_{h_1}(\ell)$ of balls of radius h_1 with centers at $\theta_j, 1 \le j \le N_{h_1}$, $\theta_j \in K_\ell$. Since for each $t', t'' \in [0, N[$ we have

$$\sup_{\|\theta\|\le\ell} |z_{t'}(\theta) - z_{t''}(\theta)| \le \sup_{1\le j\le N_{h_1}} |z_{t'}(\theta_j) - z_{t''}(\theta_j)| + 2\kappa_{h_1}^{\ell,N}(z),$$

we obtain that there exists $h, 0 < h \le \min(h_1, h_2)$ such that

$$w_h^{\ell,N}(z) \le \max_{1\le j\le N_{h_1}} w_{h_2}^N(z(\theta_j)) + 2\kappa_{h_1}^{\ell,N}(z) \tag{2}$$

where $z(\theta_j) = (z_t(\theta_j))_{t\ge0}$ and $w_{h_2}^N(\cdot)$ is the modulus of continuity in $D([0,N[,R)$ defined like in (1) with omiting $\max\limits_{\theta\in K_\ell}$.

Taking $\lim\limits_{h_1\to0} \lim\limits_{h_2\to0} \lim\limits_{h\to0}$ we have that $w_h^{\ell,N}(z) \to 0$ as $h \to 0$, hence, $z \in D([0,N[,C(K_\ell))$.

If $z \in D([0,N[,C(K_\ell))$ for each $\ell \ge 1$, $w_h^{\ell,N}(z) \to 0$ as $h \to 0$, hence, $w_h^N(z(\theta)) \to 0$ as $h \to 0$ for each $\theta \in K_\ell$ and each $\ell \ge 1$, and we have the condition a).

Suppose that the condition b) is not satisfied. Then there are sequences $\{t_n\}$, $\{\theta_n\}$ and $\{\theta'_n\}$, $0 \le t_n \le N, \|\theta_n\| \le \ell, \|\theta'_n\| \le \ell$ such that $\|\theta_n - \theta'_n\| \to 0$ as $n \to \infty$ and

$$|z_{t_n}(\theta_n) - z_{t_n}(\theta'_n)| \not\to 0 \tag{3}$$

as $n \to \infty$.

We can suppose further that $t_n \to t_0$ and either $t_n \ge t_0$ or $t_n < t_0$ for all n.

If $t_n \ge t_0$ we have

$$|z_{t_n}(\theta_n) - z_{t_n}(\theta'_n)| \le 2 \sup_{\|\theta\|\le\ell} |z_{t_n}(\theta) - z_{t_0}(\theta)| + |z_{t_0}(\theta_n) - z_{t_0}(\theta'_n)| \tag{4}$$

and the left-hand side tends to zero by the uniform right-continuity of z and by the continuity of $z_{t_0}(\cdot)$ with respect to θ, which contradicts (3).

If $t_n < t_0$ we have the analog of (4) with z_{t_0} replaced by z_{t_0-}. Since $z_{t_0-}(\cdot)$ is also continuous in θ we have a contradiction. □

In [16] we obtained the following criterion for weak convergence in $D(R_+, C_{\text{loc}}(R^m))$.

Theorem 3.2. *(see [16], p. 281) Suppose that the finite dimensional distributions of Z^n weakly converge to the ones of Z and that for every $\epsilon > 0$ and $N > 0$*

a) $\overline{\lim\limits_{h\to0}} \sup\limits_{n\ge1} P^n(w_h^N(Z^n(\theta)) \ge \epsilon) = 0, \quad \forall\theta \in R^m,$

b) $\overline{\lim\limits_{h\to0}} \sup\limits_{n\ge1} P^n(\kappa_h^{\ell,N}(Z^n) \ge \epsilon) = 0, \quad \forall\ell \ge 1.$

Then there is weak convergence

$$Z^n \xrightarrow{\mathcal{L}(P^n)} Z$$

in the Skorohod space $D(R^+, C_{loc}(R^m))$ with respect to P^n. If in addition

$$\overline{\lim_{L\to\infty}} \sup_{n\ge1} P^n(\gamma_L^N(Z^n) \ge \epsilon) = 0, \quad \forall\epsilon > 0, \quad \forall N > 0,$$

then Z^n and Z belong to D_0.

4. Weak convergence of the likelihood ratio processes

Suppose that we are given a sequence of general statistical parametric models $\mathcal{E}^n = \{(\Omega^n, F^n, \mathbb{F}^n), (P^n_\theta)_{\theta\in\Theta}\}$ and let $\mathcal{E} = \{(\Omega, F, \mathbb{F}), (P_\theta)_{\theta\in R^m}\}$ be a limiting statistical model with locally equivalent measures which generated by a process with independent increments.

We denote by $Z^n(\theta) = (z^n_t(\theta))_{t\geq 0}$ and $Z(\theta) = (z_t(\theta))_{t\geq 0}$ the cadlag versions of the likelihood ratio processes of P^n_θ, P^n_0 and P_θ, P_0 respectively.

We denote by $h^{n,\alpha}(\overline{\theta})$ and $h^\alpha(\overline{\theta})$ the Hellinger processes of order $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)$ of the measures $P^n_{\theta_0}, P^n_{\theta_1}, \dots, P^n_{\theta_k}$ and $P_{\theta_0}, P_{\theta_1}, \dots, P_{\theta_k}$ respectively, $\overline{\theta} = (\theta_0, \theta_1, \dots, \theta_k)$ with $\theta_i \in \Theta, 0 \leq i \leq k$.

We denote also by $H^{n,\alpha}_t(\overline{\theta})$ and $H^\alpha_t(\overline{\theta})$ the Hellinger integrals of order α of the restrictions of the measures $P^n_{\theta_0}, P^n_{\theta_1}, \dots, P^n_{\theta_k}$ and $P_{\theta_0}, P_{\theta_1}, \dots, P_{\theta_k}$ to the σ-algebras F^n_t and F_t respectively.

In the following theorem we recall the conditions for finite dimensional convergence of Z^n to Z obtained by JACOD [5]. We set

$$\mathcal{A}_k = \{\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k) : 0 < \alpha_i < 1, 0 \leq i \leq k, \ \sum_{i=0}^{k} \alpha_i = 1\}.$$

Theorem 4.1. *Let for each $k \geq 1$, $\alpha \in \mathcal{A}_k$ and $\overline{\theta} \in \Theta^k$ the following conditions be satisfied :*

1) $H^{n,\alpha}_0(\overline{\theta}) \to H^\alpha_0(\overline{\theta})$,

2) $h^{n,\alpha}(\overline{\theta}) \xrightarrow{\mathcal{L}(P^n_{\theta_i})} h^\alpha(\overline{\theta})$,

as $n \to \infty$ in the Skorohod space $D(R^+, R)$ with respect to each measure $P^n_{\theta_i}$, $0 \leq i \leq k$.

Then we have weak convergence

$$(Z^n(\theta_0),\ Z^n(\theta_1), \dots, Z^n(\theta_k)) \xrightarrow{\mathcal{L}(P^n_{\theta^*})} (Z(\theta_0),\ Z(\theta_1), \dots, Z(\theta_k))$$

in $D(R^+, R^{k+1})$ with respect to any fixed probability measure $P^n_{\theta^*}$, $\theta^* \in \Theta$.

Proof : From JACOD [5], Theorem 4.32, p. 61, we have the result with respect to P^n_0.

We show that the measures $(P^{n,t}_{\theta^*})$ are contiguous with respect to $(P^{n,t}_0)$ for each $\theta^* \in \Theta$. For this we note that

$$Z^n(\theta^*) \xrightarrow{\mathcal{L}(P^n_0)} Z(\theta^*)$$

in the Skorohod space $D(R^+, R)$. We can suppose that t is a point of the continuity of $Z(\theta^*)$ (if not we take any point of the continuity of $Z(\theta^*)$ greater than t). So,

$$\mathcal{L}(z^n_t(\theta^*)|P^n_0) \to \mathcal{L}(z_t(\theta^*)|P_0).$$

Since the limit model is locally equivalent we have $E_0 z_t(\theta^*) = 1$, which implies the contiguity $(P^{n,t}_{\theta^*}) \triangleleft (P^{n,t}_0)$ (see JACOD, SHIRYAYEV [6], lemma 1.10, p. 252).

Since this theorem is true for P^n_0 we have

$$(Z^n(\theta^*), Z^n(\theta_0), \dots, Z^n(\theta_k)) \xrightarrow{\mathcal{L}(P^n_0)} (Z(\theta^*), Z(\theta_0), \dots, Z(\theta_k))$$

for each $k \geq 1$ and $\theta_i \in \Theta$, $0 \leq i \leq k$. Then from the extension of LE CAM'S lemma (see JACOD, SHIRYAYEV [6], Theorem 3.3, p. 564) we have the result with respect to $P^n_{\theta^*}$. □

Remark 3 : If $h^\alpha(\overline{\theta})$ is continuous with respect to t for each $\alpha \in \mathcal{A}_k$ and $\overline{\theta} \in \Theta^k$ we can replace 2) by

$$2')h_t^{n,\alpha}(\overline{\theta}) \xrightarrow{P^n_{\theta_i}} h_t^\alpha(\overline{\theta}), \quad \forall t > 0.$$

In fact, $h^\alpha(\overline{\theta})$ is a deterministic increasing function (see JACOD [4], Theorem 5.25, p. 21) and by the result of MC LEISH [13], lemma 1, for each $t > 0$

$$\sup_{0 \leq s \leq t} |h_s^{n,\alpha}(\overline{\theta}) - h_s^\alpha(\overline{\theta})| \xrightarrow{P^n_{\theta_i}} 0.$$

□

Remark 4 : If the model $\mathcal{E}$ is continuous Gaussian we obtain from JACOD [5], Theorem 5.3, p. 64, that for 2′) is necessary and sufficient that

$$2'')h_t^{n,\alpha}(\theta,\theta') \xrightarrow{P^n_\theta} h_t^\alpha(\theta,\theta'), \quad \forall \theta,\theta' \in \Theta, \quad \forall t > 0,$$

for three values of α only : $\alpha = \frac{1}{2}, \alpha = a$ with $0 < a < \frac{1}{2}$ and $\alpha = 1 - a$. □

We denote by $Z^n = (z_t^n(\theta))_{t \geq 0, \theta \in R^m}$ and $Z = (z_t(\theta))_{t \geq 0, \theta \in R^m}$ the likelihood ratio processes of $\mathcal{E}^n$ and $\mathcal{E}$ (the first one is extended from Θ to R^m if necessary, in a way which preserws the modulus of continuity, see [15]). To obtain a criterion for weak convergence in $D(R^+, C_{\text{loc}}(R^m))$ we verify the conditions of Theorem 3.2. Using essentially the same proof as in [16], Theorem II.3.1, p. 292, we can obtain the following result.

Theorem 4.2. *Suppose that the conditions of Theorem 4.1 are satisfied. Assume that there exist constants $p \geq 1, 0 < \gamma < 1, \beta > m, r \geq 0, c \geq 0$ such that for every $t \geq 0$*

$$1) \sup_{n \geq 1} \sup_{||\theta|| \leq L, ||\theta'|| \leq L} \frac{\rho_p^p(P_{\theta'}^{n,t}, P_\theta^{n,t})}{||\theta' - \theta||^\beta} \leq cL^r < \infty.$$

Then there are processes $\tilde{Z}^n$ and $\tilde{Z}$ with paths in $D(R^+, C_{loc}(R^m))$ such that $\tilde{Z}^n(\theta) = Z^n(\theta), \tilde{Z}(\theta) = Z(\theta)$ for each $\theta \in R^m$ (P_0^n - a.s. and P_0 - a.s.) and

$$\tilde{Z}^n \xrightarrow{\mathcal{L}(P^n_{\theta_*})} \tilde{Z}. \tag{5}$$

If in addition we have for each $t > 0$ that

$$2) \overline{\lim_{L \to \infty}} \sup_{n \geq 1} \int_L^\infty y^{m-1} \left[\sup_{||\theta|| \geq y} H_t^{n,\gamma}(\theta, 0) \right]^{(\beta - m)/\beta} dy = 0$$

then $\tilde{Z}^n$ and $\tilde{Z}$ have paths in D_0 (P_0^n - a.s. and P_0 - a.s.).

Remark 5 : The condition 1) provides smoothness of the trajectories of Z^n with respect to θ and gives the possibility to estimate the modulus of continuity $\kappa_h^{\ell,N}$. The condition 2) provides the convergence of $\gamma_L^N(Z^n)$ and $\gamma_L^N(Z)$ to zero as $L \to \infty$. □

Proof : Using the proof of Theorem II.3.1 in [16] we obtain the result with respect to P_0^n.

From the proof of Theorem 4.1 we know that $(P_{\theta^*}^{n,t}) \lhd (P_0^{n,t})$ for each $t > 0$. Using the Skorohod representation theorem we can show that

$$(\tilde{Z}^n, \tilde{Z}^n(\theta^*)) \xrightarrow{\mathcal{L}(P_0^n)} (\tilde{Z}, \tilde{Z}(\theta^*))$$

in $D(R^+, C_{\text{loc}}(R^{m+1}))$, where $\tilde{Z}^n(\theta^*) = (\tilde{z}_t^n(\theta^*))_{t\geq 0}, \tilde{Z}(\theta^*) = (\tilde{z}_t(\theta^*))_{t\geq 0}$. Then using the proof of Theorem 3.3 of JACOD, SHIRYAYEV [6] we obtain the extention of the third LE CAM'S lemma for our case also and, hence, (5).

In the same manner as in [16] we show that $\tilde{Z}^n$ and $\tilde{Z}$ have their paths in D_0 (P_0^n - a.s. and P_0 - a.s.). □

In the following theorem we express the condtions 1) and 2) in terms of the Hellinger processes and the divergence processes. We denote by $k^{n,p}(\theta', \theta)$ the divergence process of order p and by $h^{n,1/2}(\theta', \theta)$ the Hellinger process of order $\frac{1}{2}$ of the measures $P_{\theta'}^n, P_\theta^n$. By $\mathcal{E}(\cdot)$ we denote the Doleans exponential function.

Theorem 4.3. *Suppose that the conditions of Theorem 4.1 are satisfied. Assume that there exist constants $p \geq 2, \beta > m, r \geq 0, c \geq 0$ such that for each $t > 0$*

1) $$\sup_n \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{\rho_p^p(P_{\theta'}^{n,0}, P_\theta^{n,0})}{\|\theta' - \theta\|^\beta} \leq cL^r < \infty,$$

2) $$\sup_n \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{E_\theta^n k_t^{n,p}(\theta', \theta)}{\|\theta' - \theta\|^\beta} \leq cL^r < \infty,$$

3) $$\sup_n \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{E_\theta^n [h_t^{n,1/2}(\theta', \theta)]^{p/2}}{\|\theta' - \theta\|^\beta} \leq cL^r < \infty,$$

4) $$\lim_{L\to\infty} \sup_n \int_L^\infty y^{m-1} \left\{ \sup_{\|\theta\|\geq y} E_0^n \mathcal{E}(-h^{n,1/2}(\theta, 0))_t \right\}^{\frac{\beta-m}{2\beta}} dy = 0.$$

Then the conclusions of Theorem 4.2 hold.

Proof : From the conditions 1), 2), 3) of this theorem and from the following inequality of DZHAPARIDZE, VALKEILA [1] :

$$\rho_p^p(P_{\theta'}^{n,t}, P_\theta^{n,t}) \leq \rho_p^p(P_{\theta'}^{n,0}, P_\theta^{n,0}) + c_p E_\theta^n [h_t^{n,1/2}(\theta', \theta)]^{p/2} + c_p E_\theta^n k_t^{n,p}(\theta', \theta)$$

where $c_p > 0$ is some constant, we obtain the condition 1) of Theorem 4.2.

From the condition 4) of this theorem and from the inequality of Kabanov (see JACOD, SHIRYAYEV [6], p. 278) we get

$$H_t^{n,3/4}(\theta, 0) \leq \{E_0^n \mathcal{E}(-h^{n,1/2}(\theta, 0))_t\}^{1/2}$$

which gives condition 2) of Theorem 4.2. □

Remark 6 : If the model $\mathcal{E}^n$ is generated by a processes with independent increments, the processes $k^{n,p}$ and $h^{n,\alpha}$ have deterministic versions (see JACOD [4]). Hence, we can omit E_θ^n in conditions 2), 3), 4) of Theorem 4.3. If the initial filtrations F_0^n and F_0 are trivial, we can omit condition 1). □

Consider now the important particular case of a Gaussian continuous limit model. Let $\mathcal{E}$ be the model generated by the m-dimensional process $X = (X_t)_{t\geq 0}$ such that

$$dX_t = \theta dt + d\eta_t, X_0 = 0,$$

where $\eta = (\eta_t)_{t\geq 0}$ is an m-dimensional continuous a.s. Gaussian process with independent increments and with a covariance matrix C_t. It is well known that in this case

$$z_t(\theta) = \exp\{{}^T\theta\eta_t - \frac{1}{2}{}^T\theta C_t\theta\} \tag{6}$$

and as in example 3 we get

$$h_t^\alpha(\theta,\theta') = \frac{1}{2}\alpha(1-\alpha)^T(\theta-\theta')C_t(\theta-\theta').$$

Suppose that the probability measures P_θ^n in $\mathcal{E}^n$ have the following structure :

$$P_\theta^n = \prod_{j=1}^n \mu_\theta^{n,j}$$

where $\mu_\theta^{n,j}$ are probability measures. Then from examples 1, 4 and Theorem 4.3 we have the following result.

Corollary 4.4. *Suppose that for each $t > 0$, $\theta, \theta' \in \Theta$, $\alpha = \frac{1}{2}$, $\alpha = a$ with $0 < a < \frac{1}{2}$ and $\alpha = 1 - a$ as $n \to \infty$*

1) $\displaystyle\sum_{j=1}^{[nt]}\{1 - H^\alpha(\mu_\theta^{n,j}, \mu_{\theta'}^{n,j})\} \to \frac{1}{2}\alpha(1-\alpha)^T(\theta-\theta')C_t(\theta-\theta'),$

2) $\displaystyle\sup_{n\geq 1}\ \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{\sum_{j=1}^{[nt]}\rho_p^p(\mu_\theta^{n,j}, \mu_{\theta'}^{n,j})}{\|\theta-\theta'\|^\beta} \leq cL^r < \infty,$

3) $\displaystyle\sup_{n\geq 1}\ \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{\left[\sum_{j=1}^{[nt]}\rho_2^2(\mu_\theta^{n,j}, \mu_{\theta'}^{n,j})\right]^{p/2}}{\|\theta-\theta'\|^\beta} \leq cL^r < \infty,$

4) $\displaystyle\overline{\lim_{L\to\infty}}\ \sup_{n\geq 1}\int_L^\infty y^{m-1}\left\{\sup_{\|\theta\|\geq y}\prod_{j=1}^{[nt]} H^{1/2}(\mu_\theta^{n,j}, \mu_0^{n,j})\right\}^{\frac{\beta-m}{2\beta}} dy = 0,$

where $\beta > m$, $c \geq 0$, $r \geq 0$ and $p \geq 2$. Then the conclusions of Theorem 4.2 hold with $Z = (z_t(\theta))_{t\geq 0, \theta\in R^m}$ defined by (6).

Suppose that $\mathcal{E}^n$ is a model generated by a Poisson process starting from zero with intensities $\lambda_t^n(\theta) \geq 0$. Then from Theorem 4.3 and examples 2, 5 we can obtain the following :

Corollary 4.5. *Suppose that for each $t > 0, \theta, \theta' \in \Theta, \alpha = \frac{1}{2}, \alpha = a$ with $0 < a < \frac{1}{2}$ and $\alpha = 1 - a$ as $n \to \infty$*

1) $\int_0^t \{\alpha\lambda_s^n(\theta) + (1-\alpha)\lambda_s^n(\theta') - (\lambda_s^n(\theta))^\alpha(\lambda_s^n(\theta'))^{1-\alpha}\}ds \to \frac{1}{2}\alpha(1-\alpha)^T(\theta-\theta')C_t(\theta-\theta'),$

2) $$\sup_{n\geq 1} \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{\int_0^t \left|(\lambda_s^n(\theta))^{1/p} - (\lambda_s^n(\theta'))^{1/p}\right|^p ds}{\|\theta-\theta'\|^\beta} \leq cL^r < \infty,$$

3) $$\sup_{n\geq 1} \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{\left[\int_0^t (\sqrt{\lambda_s^n(\theta)} - \sqrt{\lambda_s^n(\theta')})^2 ds\right]^{p/2}}{\|\theta-\theta'\|^\beta} \leq cL^r < \infty,$$

4) $$\overline{\lim_{L\to\infty}} \sup_{n\geq 1} \int_L^\infty y^{m-1} \left\{\sup_{\|\theta\|\geq y} \exp\left\{-\int_0^t (\sqrt{\lambda_s^n(\theta)} - \sqrt{\lambda_s^n(0)})^2 ds\right\}\right\}^{\frac{\beta-m}{2\beta}} dy = 0,$$

where $\beta > m$, $c \geq 0, r \geq 0$ and $p \geq 2$. Then the conclusions of Theorem 4.2 hold with $Z = (z_t(\theta))_{t\geq 0, \theta\in R^m}$ defined by (6).

Suppose that $\mathcal{E}^n$ is a model generated by the processes $X^n = (X_t^n)_{t\geq 0}$ satisfing

$$dX_t^n = a_t^n(X,\theta)dt + dW_t, X_0 = 0,$$

where $a^n(X,\theta)$ are non-anticipating functionals with $\int_0^t (a_s^n(X,\theta))^2 dx < \infty$ for each $t > 0, X \in \Omega, \theta \in \Theta$, $W = (W_t)_{t\geq 0}$ is the Wiener process. Then from example 3 and remark 2 we get the following :

Corollary 4.6. *Suppose that for each $t > 0$ and $\theta, \theta' \in \Theta$ we have*

1) $\int_0^t (a_s^n(X,\theta) - a_s^n(X,\theta'))^2 ds \xrightarrow{P_\theta^n} {}^T(\theta-\theta')C_t(\theta-\theta'),$

2) $$\sup_{n\geq 1} \sup_{\|\theta\|\leq L, \|\theta'\|\leq L} \frac{E_\theta^n\left[\int_0^t (a_s^n(X,\theta) - a_s^n(X,\theta'))^2 ds\right]^{p/2}}{\|\theta-\theta'\|^\beta} \leq cL^r < \infty,$$

3) $$\overline{\lim_{L\to\infty}} \sup_{n\geq 1} \int_L^\infty y^{m-1} \left\{\sup_{\|\theta\|\geq y} E_0^n \exp\left\{-\int_0^t (a_s^n(X,\theta) - a_s^n(X,0))^2 ds\right\}\right\}^{\frac{\beta-m}{2\beta}} dy = 0$$

where $p \geq 2$, $\beta > m$, $c \geq 0$, $r \geq 0$. Then the conclusions of Theorem 4.2 hold with $Z = (z_t(\theta))_{t\geq 0, \theta\in R^m}$ defined by (6).

Consider now the case of Poisson limit model. Let the model $\mathcal{E}$ be of the form

$$z_t(\theta) = \exp\{\theta N_t - (e^\theta - 1)t\} \tag{7}$$

where $N = (N_s)_{s\geq 0}$ is a Poisson process with intensity 1.

Suppose that the probability measures P_θ^n in $\mathcal{E}^n$ have the following structure : $P_\theta^n = \prod_{j=1}^n \mu_\theta^{n,j}$ where $\mu_\theta^{n,j}$ are probability measures.

Corollary 4.7. *Suppose that for each $k \geq 1$, $\alpha \in \mathcal{A}_k$, $\theta_i \in \Theta$, $0 \leq i \leq k$ and $t > 0$*

$$1)\ \sum_{j=1}^{[nt]}\{1 - H^\alpha(\mu_{\theta_0}^{n,j}, \mu_{\theta_1}^{n,j}, \ldots, \mu_{\theta_k}^{n,j})\} \rightarrow \{\sum_{i=0}^{k}\alpha_i \exp(\theta_i) - \exp(\sum_{i=0}^{k}\alpha_i\theta_i)\}t$$

as $n \rightarrow \infty$. Suppose also the conditions 2), 3), 4) of corollary 4.4. Then the conclusions of Theorem 4.2 hold with $Z = (z_t(\theta))_{t\geq 0, \theta \in R^m}$ defined by (7).

References

[1] K. Dzhaparidze, E. Valkeila. *On the Hellinger type distances for filtered experiments.* Probab. Th. Rel. Fields 85, 105-117 (1990).

[2] P.E. Greenwood, A.N. Shiryayev. *Contiguity and the statistical invariance principle.* Stochastics Monographs I, Gordon and Breach, New York, 1985.

[3] I.A. Ibragimov, R.Z. Hasminski. *Asymptotic theory of estimations.* Moscow, "Nauka", 1979 (in Russian).

[4] J. Jacod. *Filtered statistical models and Hellinger processes.* Stochastic Processes and their Applications 32, 3-45 (1989).

[5] J. Jacod. *Convergence of filtered statistical models and Hellinger processes.* Stochastic Processes and their Applications 32, 47-68 (1989).

[6] J. Jacod, A.N. Shiryayev. *Limit theorems for stochastic processes.* Springer-Verlag, Berlin Heildelberg, 1987.

[7] N.E. Kordzahia. *Weak convergence of likelihood ratio process.* Uspehi mat. nauk, 39 (4), 167 (1984).

[8] Ju. A. Kutoyants. *Parameter estimation for stochastic processes.* Erevan, A.N. ARSSR, 1980 (in Russian)

[9] Ju. A. Kutoyants. *Parameter estimation for stochastic processes.* Heldermann Verlag, Berlin, 1984.

[10] Ju. A. Kutoyants. *Identification of dynamic systems with small noise (to appear).*

[11] F. Liese, I. Vajda. *Convex statistical distances.* Leipzig : Teubner, 1987.

[12] J. Mémin. *Théorèmes limite fonctionnels pour les processus de vraisemblance (cadre asymptotiquement non gaussien).* Publications de l'I.R.M.A.R., Rennes 1985.

[13] D.L. Mc Leish. *An extended martingale principle.*
Ann. Probab. 6(1), 144-150 (1978).

[14] E. Valkeila, L. Vostrikova. *An integral representation for the Hellinger distance.*
Math. Scand. 58, 239-257 (1986).

[15] L. Vostrikova. *Functional limit theorems for the likelihood ratio processes.*
Annales Universitatis Scientiarum Budapestinenses de Rolando Eötvös Nominatae, vi, 145-182 (1985).

[16] L. Vostrikova. *On the weak convergence of likelihood ratio processes of general statistical parametric models.*
Stochastics 23, 277-298 (1988).