STRONG CONVERGENCE OF STATISTICAL EXPERIMENTS AND HELLINGER PROCESSES.

E. Mordecki

Steklov Mathematical Institute, Moscow, Russia,
Universidad de la Republica, Montevideo, Uruguay

Abstract.

Convergence of experiments plays a central role in the verification of the optimality of diferent statistical procedures. Since the introduction of Hellinguer processes, several works were devoted to the study of convergence of statistical experiments in terms of Hellinger processes. The main result of this paper is the proof of strong convergence (or Δ-convergence) of statistical experiments to a Gaussian or Mixed Gaussian experiment when uniform convergence of Hellinger processes takes place.

0. Introduction and notations.

Convergence of experiments plays a central role in the verification of the optimality of different statistical procedures (Le Cam,1986; Ibraguimov and Hazminskii, 1981).

Since the introduction of Hellinger processes by (Liptser and Shiryaev, 1983), several works were devoted to the study of convergence of statistical experiments (expresed as the convergence of the corresponding likelihood processes) in terms of Hellinger processes (see (Jacod and Shiryaev, 1987) and (Jacod, 1989b) and the references therein).

In this paper we prove strong convergence of experiments (or Δ-convergence) when uniform convergence of Hellinger processes takes place in case the limit is a Gaussian or Mixed Gaussian experiment.

The notion of "filtered experiment" is considered, and it is shown how the introduction of a filtration in the usual scheme of statistical experiment allows to employ the tecniques related to the theory of weak convergence of stochastic processes.

0.1. For each $n \geq 0$ consider a statistical experiment

$$\mathcal{E}^n = (\Omega^n, \mathcal{F}^n; P_\theta^n \colon \theta \in \Theta^n)$$

(see Definition 7.1 in (Strasser, 1985)) with $0 \in \Theta^n \subset \mathbb{R}^d$ and Θ^n increasing with n, endowed with a filtration (i.e. an increasing and right continuous family of sub σ-algebras of $\mathcal{F}^n$

$$\mathbb{F}^n = (\mathcal{F}_t^n)_{0 \leq t \leq 1} \quad \text{with} \quad \mathcal{F}_1^n = \mathcal{F}^n$$

(see Definition I.1.2 in (Jacod and Shiryaev, 1987)).

The pair $(\mathcal{E}^n, \mathbb{F}^n)$ will be called statistical experiment with filtration, or filtered experiment.

With each filtered experiment we associate:

a) The likelihood process with basis $0 \in \Theta^n$

$$Z_\theta^n(\omega) = \frac{dP_\theta^n}{dP_0^n} \qquad \theta \in \Theta^n$$

where the quotient denotes the likelihood ratio of the two probability measures (see Definition 1.3 in (Strasser, 1985)).

b) The density process with basis $0 \in \Theta^n$, denoted by $(Z_\theta^n(\omega)_t)_{0\le t\le 1, \theta\in\Theta^n}$ and defined by

$$Z_\theta^n(\omega)_t = \frac{d(P_\theta^n|\mathcal{F}_t^n)}{d(P_0^n|\mathcal{F}_t^n)} \qquad \theta \in \Theta^n,\ t \in [0,1],$$

where $P_\theta^n|\mathcal{F}_t^n$ denotes the restriction of the measure P_θ^n to the σ-algebra $\mathcal{F}_t^n$. Observe that $Z_\theta^n(\omega)_1 = Z_\theta^n(\omega)$.

c) The Hellinger process of order $\alpha \in \mathcal{A}^n$ denoted by

$$h^n(\alpha) = (h^n(\alpha)_t)_{0\le t\le 1}$$

where

$$\mathcal{A}^n = \{\alpha = (\alpha^\theta)_{\theta\in I}, I \text{ is a finite subset of } \Theta^n, \alpha^\theta > 0, \sum_{\theta\in I} \alpha^\theta = 1\}$$

When the set I has two points, $I = \{\theta, \zeta\}$ then $\alpha = (a, 1-a)$ with $a \in (0,1)$, the Hellinger process will be called binary and denoted by $H^{n,\theta,\zeta}(a)$ or $H^n(a,\theta,\zeta)$, and $H^{n,\theta,\zeta}$ when $a = 1/2$, (see (Jacod, 1989a)).

For further reference define for $\Theta \subset \mathbb{R}^d$

$$\mathcal{A} = \{\alpha = (\alpha^\theta)_{\theta\in I}, I \text{ is a finite subset of } \Theta, \alpha^\theta > 0, \sum_{\theta\in I} \alpha^\theta = 1\}$$

0.2. The rest of the paper is organized as follows:

Section 1 contains the Theorem on strong convergence of a sequence of experiments to a Gaussian limit experiment. Section 2 contains some generalizations of results of (Jacod, 1989b), that is weak convergence to an experiment generated by a process with conditionally independent increments. Section 3 is devoted to the study of strong convergence to a Mixed Gaussian experiment.

1. STRONG CONVERGENCE TO A GAUSSIAN EXPERIMENT.

1.1. Filtered Gaussian experiment.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P_0)$ be a stochastic basis where

$$(\Omega, \mathcal{F}, \mathbb{F}) = (W^d, \mathcal{B}(W^d), (\mathcal{B}_t(W^d))_{0\le t\le 1})$$

with $W^d = \mathbb{C}([0,1] \to \mathbb{R}^d)$ is the metric space of continuous functions $w : [0,1] \to \mathbb{R}^d$ endowed with the metric corresponding to the norm

$$\|w\| = \sup_{0\le t\le 1} \max_{1\le i\le d} |w_i(t)|$$

For each $t \in [0,1]$ denote by $\mathcal{B}^o_t(W^d)$ the σ-algebra generated by the maps $w \to w(s)$ with $o \le s \le t$,

$$\mathcal{B}_t(W^d) = \cap_{t<u\le 1}\mathcal{B}^o_u(W^d) \quad \text{for } 0 \le t < 1$$
$$\mathcal{B}(W^d) = \mathcal{B}_1(W^d) = \vee_{0\le t\le 1}\mathcal{B}^o_t(W^d)$$

Note that $\mathcal{B}(W^d)$ coincides with the Borel σ-algebra of W^d (see 1§ 4 in (Ikeda and Watanabe, 1981)).

Assume that the canonical process $M = (M_t)_{0\le t\le 1}$ defined by $M(\omega)_t = \omega_t$ is a d-dimensional continuous Gaussian martingale on $(\Omega, \mathcal{F}, \mathbb{F}, P_0)$ with deterministic compensator $\langle M, M\rangle = (\langle M^i, M^j\rangle)_{i,j=1,\dots,d}$, that is $\langle M, M\rangle$ is a continuous function with values in the space M^d of simmetric positive definite matrices.

For $\theta \in \mathbb{R}^d$ define the continuous function

$$Z_\theta : \Omega \to \mathbb{R}$$
$$Z_\theta(\omega) = \exp\{\theta'\omega_1 - \theta'\langle M, M\rangle_1\theta\} \quad \omega \in \Omega$$

As $E_{P_0}(Z_\theta(\omega)) = 1$ for each $\theta \in \mathbb{R}^d$ the relation

$$dP_\theta = Z_\theta(\omega)dP_0$$

defines a probability measure, and the Gaussian experiment will be denoted, for $\Theta \subset \mathbb{R}^d$ by

$$G = (\Omega, \mathcal{F}; P_\theta : \theta \in \Theta)$$

The binary Hellinger process of order $a \in [0,1]$ is (see (Mordecki,1994))

$$H(a, \theta, \zeta) = \frac{a(1-a)}{2}(\theta - \zeta)'\langle M, M\rangle(\theta - \zeta).$$

Consider now a sequence of experiments with filtration as in **0.1** with $\Theta^n \uparrow \Theta$, and assume that $P^n_\theta|\mathcal{F}^n_0$ does not depend on $\theta \in \Theta$.

Let K be a compact subset of Θ. For simplicity suppose that $0 \in K$ and that $\{e_i\}_{i=1,\dots,d} \subset K$ where $\{e_i\}_{i=1,\dots,d}$ is an ortonormal basis of $\mathbb{R}^d$. Denote also $e_0 = 0$.

1.2.Theorem. *Assume that*

$$\sup_{\theta\in K}|H^n(a,\theta,\theta_0)_1 - H(a,\theta,\theta_0)_1| \xrightarrow{P_0^n} 0 \quad \textit{when} \quad n\to+\infty$$

for all $a\in(0,1)$ *and* $\theta_0\in K$.

Then, the experiments $(\mathcal{E}^n)$, $n\geq 1$, *are strong convergent to the Gaussian experiment* G *in* K, *that is*

$$\mathcal{E}^n \xrightarrow{\Delta(K)} G.$$

In order to prove Theorem 1.2 we need some preliminary results. The following Lemma is based on the fact that G is an homogeneous experiment ($P_\theta\sim P_\zeta$ for all $\theta,\zeta\in\Theta$).

1.3.Lemma. *Let be given a sequence of statistical experiments* $(\mathcal{E}^n)$, $n\geq 1$, *and* G *as above. Then* $\mathcal{E}^n$ *weakly converges to* G ($\mathcal{E}^n \xrightarrow{w}\to G$) *if and only if*

$$\mathcal{L}\left((Z_\theta^n)_{\theta\in I}|P_0^n\right) \xrightarrow{w} \mathcal{L}\left((Z_\theta)_{\theta\in I}|P_0\right) \tag{1.1}$$

for all finite sets $I\in\Theta$.

Proof. According to Theorem 60.3 in (Strasser, 1985) we only have to prove (1.1)$\Rightarrow$ $\mathcal{E}^n \xrightarrow{w}\to G$.

We know that $Z_\theta > 0 P_0$-a.s. and $E_{P_0}(Z_\theta)=1$ for any $\theta\in\Theta$. So (1.1) and Theorems V.1.12 and V.1.15 in (Jacod and Shiryaev, 1987)give that the sequences (P_θ^n) and (P_0^n) are mutually contiguous, that is $(P_\theta^n)\triangleleft\triangleright(P_0^n)$ for any $\theta\in\Theta$. From this follos that $(P_\theta^n)\triangleleft\triangleright(P_\zeta^n)$ for any $\theta,\zeta\in\Theta$ and in consequence, the experiments $(\mathcal{E}^n)$ are contiguous. Theorem 61.6 in (Strasser, 1985) gives the result.

The following result is a simple consequence of Theorem 5.3 of (Jacod, 1989b).

1.4.Proposition. *Let* $(\mathcal{E}^n)$ *and* G *be given as above. Assume that*

$$H^n(a,\theta,\zeta)_1 \xrightarrow{P_0^n} H(a,\theta,\zeta)_1$$

for all θ,ζ *in* Θ *and* $a\in(0,1)$. *Then*

$$\mathcal{E}^n \xrightarrow{w}\to G.$$

Proof. By (ii) in the proof of Theorem 5.3 in (Jacod, 1989b) with $D=\{1\}$ and $\theta=0$ we obtain (1.1) and Lemma 1.3 concludes the proof.

Proof of Theorem 1.2. As $(\Omega,\mathcal{F})$ in the limit experiment is a separable and complete metric space with its Borel σ-algebra, we can apply the Corollary in § 4 of (Spokoiny and Shiryaev, 1993). This means verify the conditions of the uniform λ-convergence of experiments in K, that means to find for each $n\geq 1$ a $\mathcal{F}^n/\mathcal{F}$-measurable random element

$$\lambda^n : \Omega^n \to \Omega,$$

such that

a) $\mathcal{L}(\lambda^n|P_0^n) \xrightarrow{w} \mathcal{L}(\lambda|P_0)\equiv P_0$

b') $\sup_{\theta\in K} P_0^n(|Z_\theta^n(\omega) - Z_\theta(\lambda^n(\omega))|>\varepsilon)\longrightarrow 0.$

In order to define

$$\lambda^n = (\lambda_1^n, \ldots, \lambda_d^n)$$

consider

$$\lambda_i^n = \log Z_{e_i}^n + \frac{1}{2}\langle M^i, M^i\rangle \qquad i = 1, \ldots, d.$$

Let then $(\theta_n)_{n\geq 0}$ be an arbitrary sequence in K such that $\theta_n \to \theta_1 \in K$. Consider now the sequence of experiments with finite parameter space $S = \{0, 1, \ldots, d, d+1\}$

$$\mathcal{E}_S^n = (\Omega^n, \mathcal{F}^n, P^{n,i} : i \in S)$$

where

$$P^{n,i} = P_{e_i}^n \quad i = 0, 1, \ldots, d \quad \text{and} \quad P^{n,d+1} = P_{\theta_n}^n,$$

and the experiment

$$G_S = (\Omega, \mathcal{F}, P^i : i \in S)$$

with

$$P^i = P_{e_i} \quad i = 0, 1, \ldots, d \quad P^{d+1} = P_{\theta_1}.$$

Let us verify that

$$\mathcal{E}_S^n \xrightarrow{w} G_S \tag{1.2}$$

According to Proposition 1.4 it is necessary to check the convergence of the correponding binary Hellinger processes.

From our hypotesis, for $a \in (0, 1)$

$$H^n(a, e_i, e_j)_1 \xrightarrow{P_0^n} H(a, e_i, e_j)_1$$

Furthermore

$$\begin{aligned} |H^n(a, e_i, \theta_n) - H(a, e_i, \theta_1)| &\leq |H^n(a, e_i, \theta_n) - H(a, e_i, \theta_n)| \\ &\quad + |H(a, e_i, \theta_n) - H(a, e_i, \theta_1)| \\ &\leq \sup_\theta |H^n(a, e_i, \theta) - H(a, e_i, \theta)| \\ &\quad + |H(a, e_i, \theta_n) - H(a, e_i, \theta_1)| \xrightarrow{P_0^n} 0 \end{aligned}$$

by hypotesis and the continuity of $H(a, e_i, \theta)$ with respect to $\theta \in K$, and (1.2) is proved.

As a consequence

$$\mathcal{L}(Z_{e_1}^n, Z_{e_2}^n, \ldots, Z_{e_d}^n, Z_{\theta_n}^n | P_0^n) \xrightarrow{w} \mathcal{L}(Z_{e_1}, Z_{e_2}, \ldots, Z_{e_d}, Z_{\theta_1} | P_0)$$

Applying the Continuous Mapping Theorem (Theorem 5.1 in (Billingsley, 1968)), follows that

$$\mathcal{L}(\lambda^n, Z_{\theta_n}^n | P_0^n) \xrightarrow{w} \mathcal{L}(\lambda, Z_{\theta_1} | P_0) \tag{1.3}$$

giving a).

Let us suppose now that b) is false. Then we can find a sequence (θ_n) converging to $\theta_1 \in K$, and an $\varepsilon_0 > 0$, such that

$$P_0^n(|Z_{\theta_n}^n - Z_{\theta_n}(\lambda^n)| > \varepsilon_0) \geq \varepsilon_0 \quad \forall n \tag{1.4}$$

But in view of (1.3), the generalization of the Continuous Mapping Theorem (Theorem 5.5 in (Billingsley, 1968)) gives that

$$\mathcal{L}(Z^n_{\theta_n}, Z_{\theta_n}(\lambda^n)|P^n_0) \xrightarrow{w} \mathcal{L}(Z_{\theta_1}, Z_{\theta_1}(\lambda)|P_0)$$

From this

$$Z^n_{\theta_n} - Z_{\theta_n}(\lambda^n) \xrightarrow{P^n_0} 0$$

that contradicts (1.4) and the Theorem is proved.

2. WEAK CONVERGENCE TO AN EXPERIMENT GENERATED BY A PROCESS WITH CONDITIONALLY INDEPENDENT INCREMENTS.

2.1. Experiment generated by a process with conditionally independent increments.

Consider the stochastic basis

$$\mathcal{B} = (\Omega \otimes \mathbb{D}(\mathbb{R}^d), \mathcal{F} \otimes \mathcal{D}(\mathbb{R}^d), (\cap_{u>t}\mathcal{F} \otimes \mathcal{D}_u(\mathbb{R}^d))_{0\leq t\leq 1}, P)$$

where $(\mathbb{D}(\mathbb{R}^d), \mathcal{D}(\mathbb{R}^d), (\mathcal{D}_t(\mathbb{R}^d))_{0\leq t\leq 1})$ is the Skorokhod space of cadlag functions (see Definition VI.1.1 in (Jacod and Shiryaev, 1987)) and $(\Omega, \mathcal{F}, P_0)$ is a probability space.

Suppose that we observe a process $X = (X_t)_{0\leq t\leq 1}$ defined on $\mathcal{B}$ with values in $\mathbb{R}^d$, such that for each $\theta \in \Theta \subset \mathbb{R}^d$ X is a semimartingale with triplet of predictable characteristics $(B^\theta, C^\theta, \nu^\theta)$ under the measure $P_\theta(d\omega, d\gamma) = P_0(d\omega)Q_\theta(\omega, d\gamma)$ with Q_θ a Markov kernel such that for each fixed $\omega \in \Omega$ $Q_\theta(\omega, d\gamma)$ is the law of a process with independent increments. It can be concluded that X is a process with $\mathcal{F}$-conditionally independent increments and that the characteristics are $\mathcal{F}$-measurable and depend only in $\omega \in \Omega$. For simplicity, assume that

(1) X is quasi-left continuous for each $\theta \in \Theta$, this amounts to say, that

$$\nu^\theta(\{t\} \times \mathbb{R}^d) = 0 \quad \forall\theta \in \Theta, \quad \forall t \in [0, 1].$$

(2) The measures $(P_\theta)_{\theta\in\Theta}$ are mutually equivalent, that is $P_\theta \sim P_\zeta$ for all $\theta, \zeta \in \Theta$. As a consequence

$$C^\theta = C = c.A \quad \forall\theta \in \Theta$$

where A is an increasing and predictable process.

(3) $P_\theta|(\mathcal{F} \otimes \mathcal{D}_0(\mathbb{R}^d))$ does not depend on $\theta \in \Theta$.

In this case the Hellinger process of order $\alpha \in \mathcal{A}$ is given by (see (Jacod, 1990) or (Mordecki, 1994))

$$h(\alpha) = \frac{1}{2}\sum_{\theta\in I} \alpha^\theta\beta^\theta c\beta^\theta.A - \frac{1}{2}\sum_{\theta,\zeta\in I} \alpha^\theta\alpha^\zeta\beta^\theta c\beta^\zeta.A + \varphi_\alpha((Y^\theta)_{\theta\in I}) * \bar{\nu} \tag{2.1}$$

with β^θ and Y^θ $\mathcal{F}$-measurable processes given by Girsanov's Theorem (III.3.4 in (Jacod and Shiryaev, 1987)) applied to the measures P_θ and P_0. We deduce that

$h(\alpha)_t$ is $\mathcal{F}$-measurable, and it is not difficult to see, in view of (1) that it can be choosen to be continuous.

Then denoting

$$\mathcal{E} = (\Omega \otimes \mathbb{D}(\mathbb{R}^d), \mathcal{F} \otimes \mathcal{D}(\mathbb{R}^d); P_\theta : \theta \in \Theta),$$

$$\mathcal{F}_t = \cap_{u>t} \mathcal{F} \otimes \mathcal{D}_u(\mathbb{R}^d)\ 0 \le t < 1\ \mathcal{F}_1 = \mathcal{F} \otimes \mathcal{D}(\mathbb{R}^d)$$

$(\mathcal{E}, (\mathcal{F}_t)_{0 \le t \le 1})$ is a filtered experiment generated by a process with $\mathcal{F}$-conditionally independent increments.

Consider now a sequence of filtered experiments $(\mathcal{E}^n, \mathbb{F}^n)$, $n \ge 1$, as in **0.1** with $(\Omega^n, \mathcal{F}^n, P_0^n) = (\Omega, \mathcal{F}, P_0)$ a fixed probability space.

In order to formulate our results, we introduce the following condition (see (Feigin, 1985) or VII.5.37 in (Jacod and Shiryaev, 1987)).

2.2.Nesting condition. *Assume there exists a sequence of real numbers* (α_n) *decreasing to* 0*, such that*

(1) $\mathcal{F}^n_{\alpha_n} \subset \mathcal{F}^{n+1}_{\alpha_{n+1}}$,

(2) $\mathcal{F} = \bigvee_n \mathcal{F}^n_{\alpha_n}$.

2.3.Theorem. *Assume that the nesting condition 2.2 holds, and*

$$h^n(\alpha)_t \xrightarrow{P_0} h(\alpha)_t \quad \forall t \in [0,1], \forall \alpha \in \mathcal{A}. \tag{2.2}$$

Then

$$\mathcal{E}^n \xrightarrow{w} \mathcal{E}.$$

In case the limit experiment $(\mathcal{E}, \mathbb{F})$ has continuous density processes (see b) in **0.1**) the convergence of the binary hellinger process is a sufficient condition to obtain weak convergence.

This is the case, if $\nu^\theta = 0 \forall \theta \in \Theta$, as a consequence of Theorem III.3.35 in (Jacod and Shiryaev, 1987).

2.4.Corollary. *Assume that the nesting condition 2.2 holds, that the limit experiment* $(\mathcal{E}, \mathbb{F})$ *has continuous density processes* P_0*-a.s.,and*

$$H^n(a, \theta\zeta)_t \xrightarrow{P_0} H(a, \theta, \zeta)_t \qquad \forall t \in [0,1],\ \forall a \in (0,1)\ \forall \zeta, \theta \in \Theta.$$

Then

$$\mathcal{E}^n \xrightarrow{w} \mathcal{E}.$$

Proof of Theorem 2.3. As the limit experiment $\mathcal{E}$ is homogeneous (see condition (2) in **2.1**) as in Lemma 1.3 we deduce that it is enough to prove that

$$\mathcal{L}\left((Z^n_\theta)_{\theta \in I} | P_0\right) \xrightarrow{w} \mathcal{L}\left((Z_\theta)_{\theta \in I} | P_0\right)$$

for all finite sets $I \subset \Theta$. In fact, it is possible to show a slightly stronger result, namely

$$\mathcal{L}\left((Z^n_\theta)_{\theta \in I}, F | P_0\right) \xrightarrow{w} \mathcal{L}\left((Z_\theta)_{\theta \in I}, F | P_0\right) \tag{2.3}$$

for all finite sets $I \subset \Theta$ and all $\mathcal{F}$-measurable random variables F that we can suppose to be positive and buonded, for instance $0 \leq F \leq 1$.

The proof of (2.3) is presented in several steps. Denote

$$K_t^n = \mathbb{E}(-h^n(\alpha)_t), \quad K_t = \mathbb{E}(-h^n(\alpha)_t)\ t \in [0,1].$$

the Doleans exponent of the Hellinger process.

(i) First of all, let us see that for each $\alpha \in \mathcal{A}$

$$K_t^n \xrightarrow{P_0} K_t \tag{2.4}$$

We know by (2.1) that $h(\alpha)$ is a continuous process.

Applying Theorem VI.3.37 in (Jacod and Shiryaev, 1987) we obtain the weak convergence

$$h^n(\alpha) \xrightarrow{w} h(\alpha)$$

(because the Hellinger processes are increasing) and taking into account that the limit is continuous we obtain

$$\sup_{0\leq s\leq 1} |\Delta h_s^n| \leq w(h^n, \delta) \xrightarrow{P_0} 0$$

where $w(h^n, \delta)$ denotes the modulus of continuity of the Hellinger process

$$w(h^n, \delta) = \sup_{|s-t|<\delta} |h^n(\alpha)_t - h^n(\alpha)_s| \qquad 0 < \delta \leq 1$$

By definition

$$\mathbb{E}(-h^n(\alpha)_t) = \exp\{-h^n(\alpha)_t + \sum_{s\leq t}(\log[1 - \Delta h^n(\alpha)_s] - \Delta h^n(\alpha)_s)\}$$

Let us see now that

$$I_t^n = \sum_{s\leq t}(\log[1 - \Delta h^n(\alpha)_s] - \Delta h^n(\alpha)_s) \xrightarrow{P_0} 0 \quad \text{when} \quad n \to +\infty.$$

See that

$$0 \leq x \leq 1/2 \ \Rightarrow\ |\log(1-x) - x| \leq \frac{x^2}{2}$$

Now

$$\begin{aligned} P_0(I_t^n > \varepsilon) &\leq P_0(I_t^n > \varepsilon,\ \sup_{0\leq s\leq t} |\Delta h^n(\alpha)_s| > 1/2) + \\ &\quad P_0(I_t^n > \varepsilon,\ \sup_{0\leq s\leq t} |\Delta h^n(\alpha)_s| \leq 1/2) \\ &\leq P_0(\sup_{0\leq s\leq t} |\Delta h^n(\alpha)_s| > 1/2) + P_0(\sum_{0\leq s\leq t} |\Delta h^n(\alpha)_s|^2 > 2\varepsilon) \longrightarrow 0 \end{aligned}$$

taking into account that

$$\sum_{0\leq s\leq t} |\Delta h^n(\alpha)_s|^2 \leq \sup_{0\leq s\leq t} |\Delta h^n(\alpha)_s|.h^n(\alpha)_t \xrightarrow{P_0} 0$$

and (2.4) is proved.

(ii) Denote $A = \{K_1 \geq a\}$ with $a > 0$. For a finite set $I = \{\theta_1, \ldots, \theta_n\} \subset \Theta$, $\sum_{i=0}^{p} \beta_i < 1, \quad \beta_i \geq 0$ denote

$$Z^n(\beta) = \prod_{i=1}^{p} Z^n(\theta_i)^{\beta_i} \qquad Z(\beta) = \prod_{i=1}^{p} Z(\theta_i)^{\beta_i}$$

$$\begin{aligned} v^n &= E(F^{\beta_0} Z^n(\beta)) \qquad v = E(F^{\beta_0} Z(\beta)) \\ v_a^n &= E(\mathbb{I}_A F^{\beta_0} Z^n(\beta)) \quad v_a = E(\mathbb{I}_A F^{\beta_0} Z(\beta)) \end{aligned}$$

Applying Theorem 5.16 in (Strasser, 1985), (2.3) is equivalent to the convergence of the Mellin transforms

$$v^n \longrightarrow v \tag{2.5}$$

Applying now Hölder inequality with $\rho = 1 - \beta_0 - \sum_{i=1}^{p} \beta_i > 0$

$$|v_a^n - v^n| = E(\mathbb{I}_{A^c} F^{\beta_0} Z^n(\beta)) \leq [P(K_1 < a)]^{\rho}$$

and the same is true for $|v_a - v|$. Taking into account, that $K_1 > 0$ P_0-a.s. we conclude that (2.5) is true if and only if

$$v_a^n \longrightarrow v_a \quad \text{for any} \quad a > 0. \tag{2.6}$$

(iii) Let us now define the following predictable times

$$R^n = \inf\{t \geq 0 \colon K_1^n \leq a\}$$

From (2.2) follows, that

$$P_0(R^n \leq 1, K_1 \geq a) \underset{n}{\longrightarrow} 0.$$

We deduce the existence of another sequence (S^n) of stopping times, such that

$$\begin{aligned} &S^n < R^n \quad \text{when} \quad R^n > 0 \\ &P_0(S^n \leq 1, K_1 \geq a) \underset{n}{\longrightarrow} 0 \end{aligned} \tag{2.7}$$

(iv) Let us prove (2.6)

$$\begin{aligned} v_a^n - v_a &= E(\mathbb{I}_{A \cap \{S^n \leq 1\}} F^{\beta_0} Z^n(\beta)) + E(\mathbb{I}_{A \cap \{S^n > 1\}} F^{\beta_0} Z^n(\beta)) - v_a \\ &= E(\mathbb{I}_{A \cap \{S^n \leq 1\}} F^{\beta_0} Z^n(\beta)) \\ &+ E\left(\frac{Z^n(\beta)_{S^n \wedge 1}}{K^n_{S^n \wedge 1}} (\mathbb{I}_{A \cap \{S^n > 1\}} F^{\beta_0} K^n_{S^n \wedge 1} - E(\mathbb{I}_A F^{\beta_0} K_1 / \mathcal{F}^n_{\alpha_n})) \right) \\ &+ E\left(\frac{Z^n(\beta)_{S^n \wedge 1}}{K^n_{S^n \wedge 1}} E(\mathbb{I}_A F^{\beta_0} K / \mathcal{F}^n_{\alpha_n}) \right) - E(\mathbb{I}_A F^{\beta_0} K_1) \end{aligned}$$

As K_1 is $\mathcal{F}$-measurable in the limit

$$v_a = E\left(\mathbb{I}_A F^{\beta_0}\frac{Z(\beta)}{K_1}K_1\right) = E(\mathbb{I}_A F^{\beta_0} K_1 E\left(\frac{Z(\beta)}{K}/\mathcal{F}\right)) = E(\mathbb{I}_A F^{\beta_0} K_1).$$

Then

$$v_a^n - v_a = H_1^n + H_n^2 + H_3^n$$

with

$$\begin{aligned}
H_1^n &= E(\mathbb{I}_{A\cap\{S^n\le 1\}} F^{\beta_0} Z^n(\beta)) \\
H_2^n &= E\left(\frac{Z^n(\beta)_{S^n\wedge 1}}{K^n_{S^n\wedge 1}}(\mathbb{I}_{A\cap\{S^n>1\}} F^{\beta_0} K^n_{S^n\wedge 1} - E(\mathbb{I}_A F^{\beta_0} K_1/\mathcal{F}^n_{\alpha_n}))\right) \\
H_3^n &= E\left(\frac{Z^n(\beta)_{S^n\wedge 1}}{K^n_{S^n\wedge 1}} E(\mathbb{I}_A F^{\beta_0} K_1/\mathcal{F}^n_{\alpha_n})\right) - E(\mathbb{I}_A F^{\beta_0} K_1).
\end{aligned}$$

(v) Let us see that $H_1^n \to 0$.
Hölder inequality with $\rho = 1 - \sum_{i=1}^p \beta_i$ gives

$$H_1^n \le P(S^n \le 1, K_1 > a)^{1/\rho} \longrightarrow 0$$

(vi) Let us see, that $H_2^n \to 0$.
Denote

$$W^n = K^n_{S^n\wedge 1}(\mathbb{I}_{A\cap\{S^n>1\}} F^{\beta_0} K^n_{S^n\wedge 1} - E(\mathbb{I}_A F^{\beta_0} K_1/\mathcal{F}^n_{\alpha_n}))$$

By (2.7), (2.4) and (2) in the Nesting Condition 2.2 we obtain that

$$W^n \xrightarrow{P_0} 0$$

Furthermore

$$|W^n| \le 2$$

As $(K^n_{S^n\wedge 1})^{-1} \le 2/a$ and applying Hölder inequality with $\rho = 1 - \sum_{i=1}^p \beta_i$ we obtain that

$$H_2^n \le \frac{2}{a} E(Z^n(\beta)_{S_n\vee 1}) W_n) \le \frac{2}{a} E([W^n]^\rho)^{1/\rho} \to 0$$

(vii) Let us see that $H_3^n \to 0$.
Denote

$$M^n = (M_t^n)_{0\le t\le 1}, \qquad M_t^n = \frac{Z^n(\beta)_{t\wedge S^n}}{K^n_{t\wedge S^n}}$$

$$Q^n = \frac{1}{p+1}\left(\sum_{i=1}^p P^n_{\theta_i} + P_0\right), \quad z^n_{\theta,t} = \frac{d(P^n_\theta|\mathcal{F}^n_t)}{d(Q^n|\mathcal{F}^n_t)}, \quad z^n_{0,t} = \frac{d(P_0|\mathcal{F}^n_t)}{d(Q^n|\mathcal{F}^n_t)},$$

and consider

$$Y^n(\beta) = \prod (z^n_{\theta_i,t})^{\beta_i} (z^n_{0,t})^{1-\sum\beta_i}$$

See, that

$$z_0^n M^n = \frac{z_0^n}{K^n_{S^n\wedge 1}} \prod_{i=1}^p \left(\frac{z^n_{\theta_i}}{z^n_0}\right)^{\beta_i} = \frac{(z_0^n)^{1-\sum\beta_i}\prod (z^n_{\theta_i})^{\beta_i}}{K^n_{S^n\wedge 1}} = \frac{Y^n(\beta)}{K^n_{S^n\wedge 1}}$$

Now, as

$$0 \le Y^n(\beta) \le p, \qquad \frac{a}{2} \le K^n_{S^n \wedge t} \le 1,$$

and

$$\frac{Y^n(\beta)}{K^n_{S^n \wedge 1}} \text{ is a } Q\text{-local martingale}$$

(see 2.10 c) in (Jacod, 1989a)), it follows that

$$z^n_0 M^n \text{ is a } Q\text{-martingale}$$

By Theorem III.3.8 a) in (Jacod and Shiryaev, 1987) we obtain, that

$$M^n \text{ is a } P_0\text{-martingale}$$

On the other side

$$H^n_3 = E(\mathbb{I}_A F^{\beta_0} K_1 (1 - M^n_{\alpha_n})) \le E|1 - M^n_{\alpha_n}|$$

In order to prove that

$$E|1 - M^n_{\alpha_n}| \underset{n}{\longrightarrow} 0 \tag{2.8}$$

introduce the new measure $\tilde{P}^n$ by

$$d\tilde{P}^n = M^n_{\alpha_n} dP_0$$

Note, that

$$E_{P_0}|1 - M^n_{\alpha_n}| \le E_{P_0}|1 - M^n_{\alpha_n}| + E_{\tilde{P}^n}|1 - M^n_{\alpha_n}| = \|P - \tilde{P}^n\|.$$

By Theorem V.4.31 in (Jacod and Shiryaev, 1987) (2.8) follows from

$$H^n = H^n(1/2, \tilde{P}^n, P)_{\alpha_n} \xrightarrow{P_0} 0$$

The binary experiment $(\Omega, \mathcal{F}, \mathbb{F}^n = (\mathcal{F}_t)_{0 \le t \le \alpha_n}; \tilde{P}^n, P)$ has a Hellinger process of order $1/2$ characterized by

$$\frac{\sqrt{M^n}}{\mathbb{E}(-H^n)} \text{ is a local martingale in } \mathbb{E}(-H^n) > 0.$$

Denoting

$$\frac{\beta}{2} = (\frac{\beta_i}{2})_{1 \le i \le p} \quad \alpha = (1 - \sum_{i=1}^{p} \beta_i, \beta_1, \ldots, \beta_p) \quad \frac{\alpha}{2} = (1 - \sum_{i=1}^{p} \frac{\beta_i}{2}, \frac{\beta_1}{2}, \ldots, \frac{\beta_p}{2})$$

we have

$$\frac{\sqrt{M^n}}{\mathbb{E}(-H^n)} = \sqrt{\frac{Z^n(\beta)}{\mathbb{E}(-h^n(\alpha))}} \frac{1}{\mathbb{E}(-H^n)} = \frac{Z^n(\frac{\beta}{2})}{\sqrt{\mathbb{E}(-h^n(\alpha))}} \frac{1}{\mathbb{E}(-H^n)}$$

Knowing that

$$\frac{Z^n(\frac{\beta}{2})}{\mathbb{E}(-h^n(\frac{\alpha}{2}))} \text{ is a } P_0\text{-martingale}$$

the unicity in the descomposition of positive semimartingales (Theorem 1 in 2 § 5 of (Liptser and Shiryaev, 1986)) gives

$$\mathbb{E}(-H^n) = \frac{\mathbb{E}(-h^n(\frac{\alpha}{2}))}{\mathbb{E}(-h^n(\alpha))}$$

From (2.4) and the fact, that $\mathbb{E}(-h(\alpha))$ is a continuous process it follows that

$$\mathbb{E}(-H^n)_{\alpha_n} \xrightarrow{P_0} 1 \qquad n \to +\infty$$

As

$$H^n_{\alpha_n} = \int_0^{\alpha_n} \frac{d\mathbb{E}(-H^n)}{\mathbb{E}(-H^n)}$$

we obtain that

$$H^n_{\alpha_n} \xrightarrow{P_0} 0 \qquad n \to +\infty$$

and (2.8) is proved.

3. STRONG CONVERGENCE TO A MIXED GAUSSIAN EXPERIMENT.

3.1. Mixed Gaussian experiment.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P_0)$ be a stochastic basis where

$$(\Omega, \mathcal{F}) = (V^d \times W^d, \mathcal{B}(V^d) \otimes \mathcal{B}(W^d))$$
$$\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}, \quad \mathcal{F}_t = \cap_{u>t}\mathcal{B}(V^d) \otimes \mathcal{B}_u(W^d), \quad \mathcal{F}_1 = \mathcal{B}(V^d) \otimes \mathcal{B}(W^d).$$

with $(W^d, \mathcal{B}(W^d), (\mathcal{B}_t(W^d))_{0\le t\le 1})$ as in 1.1 and $V^d = \mathbb{C}([0,1] \to M^d)$ is the space of continuous functions with values in the space of simmetric positive definite matrices M^d, and $\mathcal{B}(V^d)$ is the Borel σ-algebra under the metric

$$\|v\| = \sup_{0\le t\le 1} \max_{1\le i,j\le d} |v_{ij}(t)|.$$

Assume that the process $M = (M_t(v,w))_{0\le t\le 1}$ is for each fixed $v \in V^d$ a d-dimensional continuous Gaussian martingale under $P_0'(v, dw)$ with compensator $\langle M, M\rangle(v)_t = v_t$, where $P_0'(v, dw)$ is a Markov kernel and

$$P_0(dv, dw) = P_0''(dv)P_0'(v, dw)$$

for some probability measure $P''(dv)$ on $(V^d, \mathcal{B}(V^d))$. In other words, M is a $\mathcal{B}(V^d)$-conditionally Gaussian martingale.

For $\theta \in \mathbb{R}^d$ define the continuous function

$$Z_\theta \colon \Omega \to \mathbb{R}$$
$$Z_\theta(\omega) = \exp\{\theta' w_1 - \theta' v_1 \theta\} \quad \omega = (v, w) \in \Omega$$

As $E_{P_0}(Z_\theta(\omega)) = 1$ for each $\theta \in \mathbb{R}^d$ the relation

$$dP_\theta = Z_\theta(\omega) dP_0$$

defines a probability measure, and the Mixed Gaussian experiment will be denoted, for $\Theta \subset \mathbb{R}^d$ by

$$MG = (\Omega, \mathcal{F}; P_\theta : \theta \in \Theta)$$

The binary Hellinger process of order $a \in [0,1]$ is (see (Mordecki,1994))

$$H(a, \theta, \zeta) = \frac{a(1-a)}{2} (\theta - \zeta)' \langle M, M \rangle (\theta - \zeta).$$

Consider now a sequence of experiments with filtration as in **0.1** with $\Theta^n \uparrow \Theta$, with $(\Omega^n, \mathcal{F}^n, P_0^n) = (\Omega, \mathcal{F}, P_0)$ a fixed probability space, and assume that $P_\theta^n | \mathcal{F}_0^n$ does not depend on $\theta \in \Theta$.

Let K be a compact subset of Θ as in Theorem 1.2.

3.2.Theorem. *Assume that*

$$\sup_{\theta \in K} |H^n(a, \theta, \theta_0)_t - H(a, \theta, \theta_0)_t| \xrightarrow{P_0} 0 \quad \textit{when} \quad n \to +\infty$$

for all $a \in (0,1)$, $t \in [0,1]$ and $\theta_0 \in K$.

Then, the experiments $(\mathcal{E}^n)$, $n \geq 1$, are strong convergent to the Mixed Gaussian experiment MG in K, that is

$$\mathcal{E}^n \xrightarrow{\Delta(K)} MG.$$

In order to prove this Theorem, we will need the following simple result.

3.3.Lemma. *Consider (X^n, F^n), $n \geq 1$, and (X, F) random elements defined on a probability space $(\Omega, \mathcal{F}, P)$ taking values in $\mathbb{R}^d \times \mathbb{R}^p$. Let $\mathcal{G} \in \mathcal{F}$. Assume that*

- *$(X^n, G) \xrightarrow{w} (X, G)$ for any G $\mathcal{G}$-measurable.*
- *F^n and F are calg-measurable.*
- *$F^n \xrightarrow{P} F$.*

Then

$$(X^n, F^n) \xrightarrow{w} (X, F).$$

Proof. Let $\lambda \in \mathbb{R}^d$ and $\mu \in \mathbb{R}^p$.

$$|Ee^{i(\lambda X^n + \mu F^n)} - Ee^{i(\lambda X + \mu F)}| \leq$$
$$E|e^{i\mu F^n} - e^{i\mu F}| + |Ee^{i\mu F}(e^{i\lambda X^n} - e^{i\lambda X}| \xrightarrow{P} 0$$

because $F^n \xrightarrow{P} F \Rightarrow E|e^{i\mu F^n} - e^{i\mu F}| \to 0$, and taking $G = e^{i\mu F}$ the second term goes to zero.

Proof of Theorem 3.2. As the proof is similar to that of Theorem 1.2 we indicate the main differences. In this case $\lambda^n = (\gamma^n, \Gamma^n)$, and are definde by

$$\gamma^n = (\gamma_1^n, \ldots, \gamma_d^n), \quad n \geq 1, \qquad \Gamma^n = (\Gamma_{i,j}^n)_{1 \leq i,j \leq d}, \quad n \geq 1$$
$$\gamma_i^n = \log Z_{e_i}^n + \frac{1}{2} H^{n, e_i, 0} \qquad i = 1, \ldots, d.$$
$$\Gamma_{i,j}^n = 4H^{n, e_i + e_j, 0} - 4H^{n, e_i, 0} - 4H^{n, e_j, 0}$$

With the same definition of $\mathcal{E}_S^n$ and MG_S, and verifying the convergence of the binary Hellinger processes, Corollary 2.4 gives

$$\mathcal{L}(Z_{e_1}^n, Z_{e_2}^n, \ldots, Z_{e_d}^n, Z_{\theta_n}^n, F|P_0) \xrightarrow{w} \mathcal{L}(Z_{e_1}, Z_{e_2}, \ldots, Z_{e_d}, Z_{\theta_1}, F|P_0)$$

for any $\mathcal{F}$-measurable random variable F.

As $H^n(a, \theta, \theta_0) \xrightarrow{P_0} 0$, Lemma 3.3 gives

$$\mathcal{L}(\lambda^n, Z_{\theta_n}^n|P_0) \xrightarrow{w} \mathcal{L}(\lambda, Z_{\theta_1}|P_0)$$

and the proof is concluded as in the proof of Theorem 1.2.

References

P.D. Feigin, *Stable convergence of semimartingales*, Stoch. Proc. and Appl. **19** (1985), 125–134.

Ikeda N. and Watanabe S., *Stochastic diffeential equations and diffusion processes*, North-Holland, Amsterdam-Oxford-New York, 1981.

J. Jacod, *Filtered statistical models and Hellinger processes*, Stochastic Process. Appl. **32** (1989), 3–45.

J. Jacod, *Convergence of filtered statistical models and Hellinger processes*, Stochastic Processes Appl. **32** (1989), 47–68.

J. Jacod, *Sur le processus de vraisemblance partielle*, Ann. Inst. Henri Poincaré **26** (1990), 299–329.

Jacod, J. and Shiryaev, A. N., *Limit theorems for stochastic processes*, Springer, Berlin Heidelberg New York, 1987.

L. Le Cam, *Asymptotic methods in Statistical Decision Theory*, Springer, Berlin Heildeberg, 1987.

R.S. Liptser, A.N. Shiryaev, *On the problem of "predictable" criteria of contiguity*, Proc. 4th Japan–URSS symp. Lecture notes in mathematics 1021, Springer, Berlin Heidelberg New York, 1983, pp. 384–418.

Strasser H., *Mathematical theory of statistics*, De Grutier, Berlin, 1985.

Mordecki E., *Asymptotic mixed normality and Hellinger processes*, Stochastics and stochastic reports (1994); To appear.

Mordecki E., *Convergence of statistical experiments and Hellinger processes*, PhD. Thesis., Steklov Matematical Institute; In russian, 1994.

Ibraguimov I. and Hasminskii R., *Asymptotical statistical estimation*, Nauka, Moscow., 1979.

Liptser R. S. and Shiryaev A.N., *Theory of Martingales*, Nauka, Moscow, 1986.

Spokoiny V. G. and Shiryaev A.N., *On the concept of λ-convergence of statistical experiments.*, Publications of the Steklov Mathematical Institute **202** (1993), 282–286; In russian.

Billingsley P., *Convergence of probability measures*, 1968.