ASYMPTOTIC MIXED NORMALITY AND HELLINGER PROCESSES

Ernesto Mordecki

Steklov Mathematical Institute, Vavilov 42, 117966 Moscow GSP–1 Russia
Universidad de la Republica, Montevideo, Uruguay

Abstract: We characterize the stable convergence of a sequence of density processes corresponding to binary filtered experiments to the exponential of a mixture of gaussian processes, in terms of the convergence of their respective Hellinger processes, allowing random Hellinger processes in the limit. Within the applications we obtain necessary and sufficient conditions for the log–likelihood function of a supercritical branching process to be asymptotically mixed normal, in terms of the offspring distributions.

Keywords: Asymptotic mixed normality, Hellinger processes, Binary filtered experiments, Asymptotic mixed normality for supercritical branching processes.

1.Introduction.

The subject of the present paper is the characterization of the behaviour of the density processes corresponding to a binary filtered experiment, in case the limit is the exponential of a proccess with conditionally independent increments; the so called "local asymptotic mixed normal case".

Our characterization is done in terms of the "Hellinger processes" $h(\alpha)$, a sort of predictable characteristics for density processes introduced by Liptser and Shiryaev [10] with the purpose of studying contiguity of probability measures corresponding to stochastic processes.

The main difference with previous results is that we allow random Hellinger processes in the limit.

A similar characterization (in the discrete time case) when the $h(\alpha)$ are deterministic (asymptotic normal case) were obtained by Greenwood and Shiryaev [4], where also is presented a slightly weaker version of the sufficient condition of our case. Other relevant results concerning weak convergence of density processes in terms of Hellinger processes, were obtained by Vostrikova in the continuous time case when $h(\alpha)$ is deterministic (under a mild additional assumption), giving also finite dimensional convergence relative to the parameter space [16], and functional convergence in parameter and time [17].

See also Jacod [6] for the general parameter space case, Höpfner, Jacod and Ladelli [5] for Markov statistical models, and Jacod and Shiryaev [7] for a general overview.

Typeset by $\mathcal{AMS}$-TEX

Sufficient conditions in terms of the Hellinger processes and the characteristics of the observed processes were recently obtained by Kramkov [10] with general parameter space and quasi–left continuous limit.

For others results about convergence of density processes see Taraskin [15] who covers a very wide class of limit processes, whose finite dimensional distributions are mixtures of infinitely divisively laws (including the LAMN case); under conditions on the log–likelihood; Luschy [11], for a similar approach; where LAMN property (according to Basawa–Scott [1]) is established and some efficient results are obtained.

The paper is organized as follows:

In section 2 we state the main results. Section 3 is devoted to the study of two applications. The first one can be stated as "LAMN under loss of information" and the second concerns supercritical branching processes (or Galton – Watson processes) where necessary and sufficient conditions are given for stably convergence of log-likelihood processes to a mixed gaussian limit to hold, in terms of the Hellinger integrals of the offspring laws. In section 4 we present the proofs of the main results.

For general concepts about semimartingales and their characteristics we refer to Jacod and Shiryaev [7] chapters I and II.

We will use the following notation:

$$X.Z_t = \int_0^t X_s dZ_s \qquad Y * \nu_t = \int_0^t \int_{\mathbb{R}^d} Y(\omega, s, x)\nu(\omega, ds, dx)$$

to denote the stochastic integral of a predictable process X with respect to a semimartingale Z, and the integral of a random function Y with respect to a random measure ν in the interval $[0, t]$.

$\mathbb{I}_A$ denotes the indicator of the set A ($\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ if $x \notin A$), $P \sim Q$ that P and Q are mutually absolutely continuous, and $P^n \triangleleft Q^n$ that the sequences of probability measures P^n is contiguous to the sequence Q^n (for all sequences A^n such that $Q^n(A^n) \to 0$ we have $P^n(A^n) \to 0$).

We will denote the $\mathcal{F}$–stably convergence in $\mathbb{D}(\mathbb{R}^d)$ by

$$X^n \xrightarrow{d} X \quad (\mathcal{F} - \text{stably})$$

defined by

$$E(Yf(X^n)) \to E(Yf(X)) \quad \text{as} \quad n \to +\infty$$

for all:

Y bounded and $\mathcal{F}$–measurable random variables

$f : \mathbb{D}(\mathbb{R}^d) \to \mathbb{R}$ continuous and bounded functions.

2. Main results.

(E) *Let be given a sequence of binary filtered experiments*

$$\mathcal{E}^n = (\Omega, \mathcal{F}, \mathbb{F}^n; P, \tilde{P}^n)$$

where $(\Omega, \mathcal{F}, P)$ *is a probability space,* $\mathbb{F}^n = (\mathcal{F}_t^n : t \geq 0)$ *is a filtration (i.e. an increasing and right continuous family of sub* σ*–algebras of* $\mathcal{F}$*) that satisfies* $\mathcal{F} = \bigvee\{\mathcal{F}_t^n : t \geq 0\}$ *for each* n*, and* $\tilde{P}^n$ *a sequence of probability measures defined on*

$(\Omega, \mathcal{F})$. For each $t \geq 0$ we denote P_t^n and $\tilde{P}_t^n$ the restrictions of P and $\tilde{P}^n$ to $\mathcal{F}_t^n$ respectively.

For simplicity we will always assume

(a) $P_0^n = \tilde{P}_0^n$

(b) *P_t^n and $\tilde{P}_t^n$ are mutually absolutely continuous for each $t \in \mathbb{R}^+$ ($P \overset{loc}{\sim} \tilde{P}$)*

(This restriction is not essential in obtaining the results; and allow to gain in simplicity and comprehension in what follows).

With each $\mathcal{E}^n$ we associate the following processes:

(1) *$Z^n = (Z_t^n)_{t \geq 0}$ defined by $Z_t^n \equiv \frac{d\tilde{P}_t^n}{dP_t^n}$, the density process or likelihood process.*

(2) *$Y^n = logZ^n$, the log–density or log–likelihood process (under our assumptions it is always well defined), whose modified characteristics will be denoted by $(B^n, \tilde{C}^n, \nu^n)$.*

(3) *$h^n(\alpha)$, with $\alpha \in (0,1)$ the Hellinger process of order α, defined by the conditions $h^n(\alpha)_0 = 0$ and*

$$M^n(\alpha)_t = (Z_t^n)^{1-\alpha} + \int_0^t (Z_{s-}^n)^{1-\alpha} dh^n(\alpha)_s$$

is a P–martingale (see [7] chapter IV).

(4) *$I^n(a) = e^x |1 - e^{-x}| \mathbb{I}_{[0,\frac{1}{a}] \cup [a,\infty)}(e^{-x}) * \nu^n \quad with \quad a \in (1, \infty)$*

(5) *$i_\beta(P, \tilde{P}^n) = e^x \mathbb{I}_{\{e^{-x} \leq \beta\}} * \nu^n \quad with \quad \beta \in (0,1)$*

The limit experiment is constructed as follows:

(L) $$\mathcal{E} = (\Omega \times \mathbb{C}, \mathcal{F} \otimes \mathcal{C}, \{\cap_{s>t}(\mathcal{F} \otimes \mathcal{C}_s)\}_{t \geq 0}; Q, \tilde{Q})$$

where $(C, \mathcal{C}, (\mathcal{C}_t)_{t \geq 0}, P_W)$ is the canonical Weiner space of continuous functions $\gamma = (\gamma_t)_{t \geq 0}$, P_W is the Wiener measure, $Q(d\omega, d\gamma) = P(d\omega) P_W(d\gamma)$, and $\tilde{Q}(d\omega, d\gamma)$ defined by

$$\frac{d(\tilde{Q}|\bar{\mathcal{F}}_t)}{d(Q|\bar{\mathcal{F}}_t)}(\omega, \gamma) = Z_t(\omega, \gamma),$$

where $\bar{\mathcal{F}}_t = \bigcap_{s>t}(\mathcal{F} \otimes \mathcal{C}_s)$, $Z_t(\omega, \gamma) = exp\{W_{C_t(\omega)}(\gamma) - \frac{1}{2} C_t(\omega)\}$, $W = (W_t(\gamma))_{t \geq 0}$ is the Wiener process, $C = (C_t(\omega))_{t \geq 0}$ is a continuous increasing process, that play the role of change of time, such that $C_0(\omega) = 0$ and $C_t(\omega)$ is $\mathcal{F}_t$–measurable for each $t \geq 0$.

Theorem 1. *Assume* **(E)** *and* **(L)**. *Let D be a dense subset in $\mathbb{R}^+$. The following conditions are equivalent:*

(1) *$Z^n \xrightarrow{d} Z \quad (\mathcal{F} - stably)$*

(2) *$h^n(\alpha)_t \xrightarrow{P} \frac{\alpha(1-\alpha)}{2} C_t \quad \forall t \in D \quad \forall \alpha \in (0,1)$*

(3) *$h^n(\alpha)_t \xrightarrow{P} \frac{\alpha(1-\alpha)}{2} C_t \quad \forall t \in D;$*
$for\ \alpha = 1/2, \alpha = \beta, \alpha = 1 - \beta,\ for\ some\ \beta \in (0, 1/2)$

(4) *$h^n(1/2)_t \xrightarrow{P} \frac{1}{8} C_t$ and $I^n(1+\epsilon)_t \xrightarrow{P} 0 \quad \forall t \in D, \quad \forall \varepsilon > 0$*

Remarks:

(1) Theorem 1 is a generalization of Theorems 1.12 and 1.64 in chapter X of [7], where the "process" C is deterministic.

(2) Observe that the condition

$$I^n(a)_t \xrightarrow{P} 0 \quad \forall t \geq 0 \quad \forall a > 1$$

is a "Lindberg Condition" in the sense that it controls the jumps of Y^n.

3. Applications.

In order to see how Theorem 1 works, we will examine two related examples. The first one consist in the problem of "loss of information" formulated in this way: Take a sequence $\mathcal{E}^n = (\Omega, \mathcal{F}, \mathbb{F}^n, P, \tilde{P}^n)$ generated by independent identically distributed random variables, whose log–density processes are asymptotically normal, and suppose we are not able to observe all the sample but only up to certain random times. Under the condition of convergence in probability of the normalized stopping times, and taking a convenient subsequence of $(P, \tilde{P}^n)$, we obtain asymptotic mixed normality for the density processes of the restricted experiment. The second one is a particular case of the first, where we consider a supercritical branching process as the restricted experiment.

3.1.Asymptotic mixed normality under loss of information.

Consider for each $n \in \mathbb{N}$ the binary filtered experiment

$$\mathcal{E}^n = (\Omega, \mathcal{F}, \mathbb{F}^n; P, \tilde{P}^n)$$

consisting in the observation of the first n variables of a sequence $X_1, \ldots, X_n, \ldots$ of i.i.d. random variables defined on a measurable space $(\Omega, \mathcal{F})$

That is

$$P = \overset{\infty}{\underset{i=1}{\otimes}} \mu \quad \tilde{P}^n = \overset{\infty}{\underset{i=1}{\otimes}} \tilde{\mu}^n$$

where μ and $\tilde{\mu}^n$ are the possible distributions of X_1; $\mathbb{F}^n = (\mathcal{F}^n_t)_{t\geq 0}$ with $\mathcal{F}^n_t = \mathcal{F}_{nt}$ and $\mathcal{F}_t = \sigma - \{X_i : 0 \leq i \leq t\}$

In order to construct the "restricted experiments" $\mathcal{G}^n$ obtained by loss of information consider $(T_n : n \in \mathbb{N})$ a sequence of stopping times satisfying:

(a) $\exists (a_n) \subset \mathbb{N}$ strictly increasing, and a random variable $\xi(\omega)$ on $(\Omega, \mathcal{F})$ such that

$$\frac{T_n}{a_n} \xrightarrow{P} \xi(\omega) \qquad \text{as } n \to +\infty$$

(b) $\exists (b_n) \subset \mathbb{N}$ strictly increasing such that: $T_n \geq b_n \quad \forall n$

Remark: Condition (b) ensures the "nesting condition" in [3] to hold.

Then consider

$$\mathcal{G}^n = (\Omega, \mathcal{F}, \mathbb{G}^n; P, \tilde{P}^{a_n})$$

with $\mathbb{G}^n = \{\mathcal{G}^n_t : 0 \leq t \leq 1\}$ where $\mathcal{G}^n_t = \mathcal{F}_{[T_n t]}$.

Denote by $Z^n(\mathcal{E})$, $h^n(\alpha, \mathcal{E})$, $Z^n(\mathcal{G})$, $h^n(\alpha, \mathcal{G})$ the density and Hellinger processes of the experiments $\mathcal{E}^n$ and $\mathcal{G}^n$ respectively.

Now we can state the following

Theorem 2. *Suppose $\mathcal{E}^n$ satisfies*

$$Z^n(\mathcal{E}) \xrightarrow{\mathcal{L}(P)} exp\{W_t - \frac{1}{2}t\}$$

Then $\mathcal{G}^n$ satisfies

$$Z^n(\mathcal{G}) \xrightarrow{d} exp\{W_{\xi(\omega)t} - \frac{1}{2}\xi(\omega)t\} \quad \mathcal{F} - stably$$

Proof. Chose $\alpha \in (0,1)$ and $t \in \mathbb{R}^+$. In view of Theorem 1 with $C_t = t$ we know:

$$h^n(\alpha, \mathcal{E}) \xrightarrow{P} \frac{\alpha(1-\alpha)}{2}t$$

Furthermore

$$h^n(\alpha, \mathcal{E}) = \sum_{i=1}^{[nt]}(1 - H(\alpha, \mu, \tilde{\mu}^n)) = [nt](1 - H(\alpha, \mu, \tilde{\mu}^n))$$

(see [7] IV.1.64). As the Hellinger processes is defined trough the condition of certain process to be a martingale and this property is stable under stopping, we obtain:

$$h^n(\alpha, \mathcal{G})_t = \sum_{i=1}^{[T_n t]}(1 - H(\alpha, \mu, \tilde{\mu}^{a_n})) = [T_n t](1 - H(\alpha, \mu, \tilde{\mu}^{a_n}))$$

Then as $[nt](1 - H(\alpha, \mu, \tilde{\mu}^n)) \to \frac{\alpha(1-\alpha)}{2}t$ we get

$$[T_n t](1 - H(\alpha, \mu, \tilde{\mu}^{a_n})) = \frac{[T_n t]}{a_n} a_n(1 - H(\alpha, \mu, \tilde{\mu}^{a_n})) \xrightarrow{P} \frac{\alpha(1-\alpha)}{2}\xi(\omega)t$$

and Theorem 1 gives the result. □

3.2. Asymptotic mixed normality for supercritical branching processes.

In this section we show that the experiments corresponding to supercritical branching processes can be fitted in our first application, obtaining necessary and sufficient conditions for asymptotic mixed normality for the density processes in terms of the Hellinger integrals corresponding to the offspring laws, giving light into a problem stated in Le Cam and Yang [9] in reference to a paper of Davies [2]. See also Basawa and Scott [1] for LAMN models and in particular results about branching processes.

Denote the by $p = (p_i)$ a probability measure on the positive integers $\{1, 2, \dots\}$ called the offspring distribution. Denote $m = \sum_{i=1}^{+\infty} ip_i$. We assume $1 < m < +\infty$, corresponding to the supercritical case.

Consider $M_0, \dots, M_n$ random variables defined on a probability space $(\Omega, \mathcal{F}, P)$; M_n stands for the number of individuals of the n^{th} generation and $M_0 = 1$. If the n^{th} generation contains M_n individuals, each of them independently of the "history" an the other individuals transform itself into i individuals with probability p_i .

In what follows we will need the following results:

Proposition. *Let $M_0, \dots, M_n$ be a supercritical branching process defined on a probability space $(\Omega, \mathcal{F}, P)$ with $m = E(M_1) \in (1, +\infty)$. Then there exist an a.s. nonegative random variable $\xi(\omega)$ on $(\Omega, \mathcal{F}, P)$ with $E(\xi(\omega)) \leq 1$ such that*

(1) $\frac{M_n}{m^n} \to \xi(\omega) \quad P - a.s.$
(2) $\frac{M_0 + \dots + M_n}{1 + m + \dots + m^n} \to \xi(\omega) \quad P - a.s.$

Proof. (1) is a consequence of the Doob's Theorem applied to the martingale $\frac{M_n}{m^n}$. (2) is a consequence of the Kronecker's Lemma (see for instance [14] ch. 4.3) $\square$

Let us now consider the experiments $\mathcal{E}^n = (\Omega, \mathcal{F}, \mathbb{F}^n; P, \tilde{P}^n)$ constructed as in 3.1 with $\mu = p$ and $\tilde{\mu}^n = \tilde{p}^n$

In order to construct the restricted experiments $\mathcal{G}^n$ consider the random walk

$$R(0) = 1 \quad R(1) = 1 + X_1 \quad \dots \quad R(n) = 1 + X_1 + \cdots + X_n$$

Define T_n by the following relations:

$$T(0) = 1 \quad T(1) = R(T(0)) \quad \dots \quad T(n) = R(T(n-1))$$

and let us observe that if

$$M_0 = T(0) = 1 \quad M_1 = T(1) - T(0) \quad \dots \quad M_n = T(n) - T(n-1)$$

then $(M_0, M_1, \dots, M_n)$ is a branching process with offspring law p or $\tilde{p}^n$.

Now by (2) in the Proposition as $T_n = M_0 + M_1 + \cdots + M_n$ taking $a_n = [1 + m + \cdots + m^n]$ we obtain

$$\frac{T_n}{a_n} \xrightarrow{P} \xi(\omega)$$

So considering

$$\mathcal{G}^n = (\Omega, \mathcal{F}, \mathbb{G}^n; P, \tilde{P}^{a_n})$$

as in 3.1 we obtain the following result:

Theorem 3. *The following conditions are equivalent:*

(1) $Z^n(\mathcal{G}) \xrightarrow{d} exp\{W_{\xi(\omega)t} - \frac{1}{2}\xi(\omega)t\}$ $(\mathcal{F} - stably)$
(2) $H(\alpha, p, \tilde{p}^{a_n}) = 1 - \frac{\alpha(1-\alpha)}{2}\frac{m-1}{m^{n+1}} + o(\frac{1}{m^{n+1}})$

Proof. As in Application 3.1

$$h^n(\alpha, \mathcal{G}) = [T_n t](1 - H(\alpha, p, \tilde{p}^{a_n})) \xrightarrow{P} \frac{\alpha(1-\alpha)}{2}\xi(\omega)t$$

if and only if (2) holds $\square$

Remark: In refernece to the works of Davies [2] and Le Cam and Yang [9] we can observe that in Theorem 4.1 of Davies [2] it is not stated whether the normalization "$\delta_n(\theta)$" is changed and how in the restricted experiments $\mathcal{G}^n$, though this fact becames clear in the applications.

In the last example of section 3 given by Le Cam and Yang [9] the measures "p^n" and "q^n" are not changed when constructing the "bigger" experiments $\mathcal{E}^n$ (this is equivalent in the set up of Davies to have the same rates "$\delta_n(\theta)$" for both experiments) obtaining that if the bigger experiments are contiguous, their restrictions tends to the totally uninformative experiment.

In Theorem 3 seems clear that to construct the bigger experiments $\mathcal{E}^n$ form the restricted ones; it would be necessary to enlarge the sequences of probability measures in such a way that the measures in $\mathcal{G}^n$ result a subsequence of the measures in $\mathcal{E}^n$ with the adequate rate of mutual approximation as n tends to infinity.

4. Proofs.

4.1 We begin with some preliminary results.

Proposition 1. *Denote by $\mathcal{E}$ any of the experiments considered in* **(E)**, *by Z and $Y = logZ$ the corresponding density and log–density processes respectively. Y^c denotes the continuous martingale part of Y, $(B, C =< Y^c, Y^c >, \nu^Y)$, $(\tilde{C})$ its characteristics (modified second characteristic) under P.*

Then:

(1) $h(\alpha, P, \tilde{P}) = \frac{\alpha(1-\alpha)}{2} < Y^c, Y^c > + \varphi_\alpha(e^x) * \nu^Y$
where $\varphi_\alpha(u) = \alpha + (1-\alpha)u - u^{1-\alpha}$.

(2) $B = -[\frac{1}{2} < Y^c, Y^c > + (e^x - 1 - h(x)) * \nu^Y]$

(3) $\tilde{C} =< Y^c, Y^c > + h^2(x) * \nu^Y - \sum_{s \leq \cdot} (\Delta B_s)^2$

(4) $i_\beta(\tilde{P}, P) = \mathbb{I}_{[0,\beta]}(e^x) * \nu^Y$

Proof of (1). First we apply Ito's formula to $Y = logZ$.

Denote by M the local martingale $M = (Z_-)^{-1}.Z$, by M^c and μ^M its continuous martingale part and jump measure respectively.

We obtain:

$$Y = M - \frac{1}{2} < M^c, M^c > + (log(1+x) - x) * \mu^M$$

so

$$< Y^c, Y^c > =< M^c, M^c >$$

and

$$\Delta Y = log(1 + \Delta M)$$

So for any real function $g(x)$:

$$g(x) * \mu^Y = g(log(1+x)) * \mu^M$$

and this for $g(x) = 1 + x - e^x$ gives:

$$(1 + x - e^x) * \mu^Y = (log(1+x) - x) * \mu^M$$

so

$$Y = M - \frac{1}{2} < Y^c, Y^c > + (1 + x - e^x) * \mu^Y \tag{4.1}$$

setting $\gamma(x) = e^x - 1 - x$ we obtain that

$$Y + \frac{1}{2} < Y^c, Y^c > + \gamma(x) * \nu^Y$$

is a local martingale.

Now apply Ito's formula to $f(Y) = exp\{(1-\alpha)Y\}$:

$$\begin{aligned} f(Y) &= 1 + (1-\alpha) f(Y_-).Y + \frac{1}{2}(1-\alpha)^2 f(Y_-). < Y^c, Y^c > \\ &\quad + \sum_{s \leq \cdot} [f(Y_s) - f(Y_{s-}) - (1-\alpha) f(Y_{s-}) \Delta Y_s] \\ &= 1 + (1-\alpha) f(Y_-).[Y + \frac{1}{2} < Y^c, Y^c > + \gamma(x) * \nu^Y] \\ &\quad - f(Y_-).[\frac{\alpha(1-\alpha)}{2} < Y^c, Y^c > + \varphi_\alpha(e^x) * \nu^Y] \end{aligned}$$

As the first integral is a local martingale; the uniqueness of $h(\alpha)$ gives the result. $\square$

Proof of (2). Consider

$$Y(h) = Y - \check{Y}$$

with

$$\check{Y}(h) = \sum_{s \leq \cdot} \Delta Y_s - h(\Delta Y_s) = (x - h(x)) * \mu^Y$$

In view of (4.1)

$$\begin{aligned} Y(h) &= M - \frac{1}{2} < Y^c, Y^c > + (1 + h(x) - e^x) * \mu^Y \\ &= M + (1 + h(x) - e^x) * (\mu^Y - \nu^Y) \\ &- \frac{1}{2} < Y^c, Y^c > + (1 + h(x) - e^x) * \nu^Y \end{aligned}$$

And as $Y(h)$ has buonded jumps, and the first two summands are local martingales, the uniqueness of the decomposition of a special semimartingale gives (1). $\square$

Proof of (3). (3) is a general fact, see for instance II.2.17 in [7].

Proof of (4). Let us see that

$$\Delta Y = \Delta log Z = -\Delta log(1/Z) = -\Delta \tilde{Y}$$

if $\tilde{Y} = log\tilde{Z} = log(1/Z)$; $\tilde{Z}_t = \frac{dP_t}{d\tilde{P}_t}$ so for any real function $g(x)$,

$$g(x) * \mu^Y = g(-x) * \mu^{\tilde{Y}}$$

and this relation carries over the compensators with respect to $\tilde{P}$:

$$g(x) * \nu^{Y,\tilde{P}} = g(-x) * \nu^{\tilde{Y},\tilde{P}}$$

Now to obtain the relation between $\nu^{Y,\tilde{P}}$ and $\nu^{\tilde{Y},\tilde{P}}$ we will apply Girsanov's Theorem for random measures (see [7] chapter III)

$$\frac{Z}{Z_-} = \frac{e^Y}{e^{Y_-}} = e^x * \mu^Y(dt, dx) \quad a.s.$$

This gives

$$M_\mu^P(\frac{Z}{Z_-}/\tilde{\mathcal{P}}) = e^x$$

so

$$\nu^{Y,\tilde{P}} = e^x \nu^{Y,P}$$

Then

$$\begin{aligned} i_\beta(\tilde{P}, P) &= e^x \mathbb{I}_{\{e^{-x} \leq \beta\}} * \nu^{\tilde{Y},\tilde{P}} \\ &= e^{-x} \mathbb{I}_{\{e^x \leq \beta\}} * \nu^{Y,\tilde{P}} \\ &= \mathbb{I}_{\{e^x \leq \beta\}} * \nu^{Y,P} \end{aligned}$$

□

Let us consider the following conditions:

$[\beta - D] \quad B^n_t \xrightarrow{P} B_t \quad \forall t \in D$

$[sup\beta - D] \quad \sup_{s\le t} |B^n_t - B_t| \xrightarrow{P} 0 \quad \forall t \in D$

$[\gamma - D] \quad \tilde{C}^n_t \xrightarrow{P} C_t \quad \forall t \in D$

$[\delta - D] \quad \mathbb{I}_{\{|x|>\varepsilon\}} * \nu^n_t \xrightarrow{P} 0 \quad \forall t \in D \quad \forall \varepsilon > 0$

Lemma 1. *Assume* **(E)**.

(1) *Assume* $[\delta - \mathbb{R}^+]$. *Then*

$$\sup_{s\le t} |\Delta B^n_s| \xrightarrow{P} 0 \tag{4.2}$$

(2) *Assume* $[\delta - \mathbb{R}^+]$ *and* $(Var(B^n_t)|P^n_t)$ *is tight. Then*

$$\sum_{s\le t} (\Delta B^n_s)^2 \xrightarrow{P} 0$$

Proof of Lemma 1. (1) follows from

$$|\Delta B^n_s| = h(x) * \nu^n(\{s\}, dx) \le \varepsilon + \Delta \mathbb{I}_{\{|x|>\varepsilon\}} * \nu^n_s \le \varepsilon + \mathbb{I}_{\{|x|>\varepsilon\}} * \nu^n_t$$

(assuming $h(x) = x$ in the set $\{|x| < \varepsilon\}$ and $|h(x)| \le 1$). So

$$\sup_{s\le t} |\Delta B^n_s| \le \varepsilon + \mathbb{I}_{\{|x|>\varepsilon\}} * \nu^n_t$$

obtaining (4.2) by $[\delta - \mathbb{R}^+]$.

(2) is a consequence of (1) if we observe that

$$\sum_{s\le t} (\Delta B^n_s)^2 \le (\sup_{s\le t} |\Delta B^n_s|) Var(B^n_t)$$

□

Lemma 2. *Assume* **(E)**.

(1) *Suppose the following conditions hold, for some fixed* $t \in \mathbb{R}^+$.

$$\mathbb{I}_{\{|x|>\varepsilon\}} * \nu^n_t \xrightarrow{P} 0 \tag{4.3}$$

$$\lim_{a\to\infty} \limsup_n P(1_{\{x>a\}} * \nu^n_t > \eta) = 0 \quad \forall \eta > 0 \tag{4.4}$$

$$\exists \varepsilon > 0 \text{ such that } (1_{\{|x|<\varepsilon\}} x^2 * \nu^n_t | P^n_t) \text{ is tight} \tag{4.5}$$

Then

$$\frac{1}{8} h^n(\alpha)_t - \frac{\alpha(1-\alpha)}{2} h^n(1/2)_t \xrightarrow{P} 0$$

(2) *And if furthermore*

$$(Var(B^n_t)|P^n_t) \text{ is tight} \tag{4.6}$$

then:

$$h^n(\alpha)_t - \frac{\alpha(1-\alpha)}{2} \tilde{C}^n_t \xrightarrow{P} 0$$

Proof of (1).

$$\frac{1}{8}h^n(\alpha)_t - \frac{\alpha(1-\alpha)}{2}h^n(1/2)_t = [\frac{1}{8}\varphi_\alpha(e^x) - \frac{\alpha(1-\alpha)}{2}\varphi_{\frac{1}{2}}(e^x)] * \nu_t^n$$

After computations

$$|(\frac{1}{8}\varphi_\alpha(e^x) - \frac{\alpha(1-\alpha)}{2}\varphi_{\frac{1}{2}}(e^x))| \leq e^x\mathbb{I}_{\{x>a\}} + K_a\mathbb{I}_{\{\varepsilon\leq|x|\leq a\}} + x^3e^\varepsilon\mathbb{I}_{\{|x|<\varepsilon\}} + \mathbb{I}_{\{x<-a\}}$$

So let $\delta > 0$ be given: Chose a and n_0 such that

$$P(e^x\mathbb{I}_{\{x>a\}} * \nu_t^n > \delta/4) < \delta/4 \quad \forall n \geq n_0$$

as $1_{\{|x|<\varepsilon\}}x^2 * \nu_t^{Y^n}$ is tight for some ε and $e^\varepsilon(x^3)\mathbb{I}_{\{|x|<\varepsilon\}} * \nu_t^n \leq \varepsilon e^\varepsilon(x^2)\mathbb{I}_{\{|x|<\varepsilon\}} * \nu_t^n$ we can chose ε_0 and n_1 such that

$$P(x^3e^\varepsilon\mathbb{I}_{\{|x|<\varepsilon\}} * \nu_t^n > \delta/4) < \delta/4 \quad \forall n \geq n_1$$

As

$$(\mathbb{I}_{\{\varepsilon\leq|x|\leq a\}} + \mathbb{I}_{\{x<-a\}}) * \nu_t^n \xrightarrow{P} 0$$

(1) follows.

Proof of (2).

$$h^n(\alpha)_t - \frac{\alpha(1-\alpha)}{2}\tilde{C}_t^n = [\varphi_\alpha(e^x) - \frac{\alpha(1-\alpha)}{2}h(x)^2] * \nu_t^n + \sum_{s\leq t}(\Delta B_s^n)^2$$
$$= J_1^n + J_2^n$$

Now $J_1^n \xrightarrow{P} 0$ as in (1) because the integrands are continuous and have the same behaviour at $x = +\infty$, $x = -\infty$ and $x = 0$.

Finally $J_2^n \xrightarrow{P} 0$ by (2) in Lemma 1. □

4.2. Proof of Theorem 1.

Proofs of (3) ⇒ (1).

According to Theorem VIII.5.42 in [7] (see also [3]) and Lemma 1 in [12] (that gives when D is dense in $\mathbb{R}^+$, $[\beta - D] \Rightarrow [sup\beta - D]$) it is enough to verify $[\beta - D]$, $[\gamma - D]$ and $[\delta - D]$. Let us then choose $t \in D$.

Proof of $[\delta - D]$

We know that

$$h^n(\beta)_t + h^n(1-\beta)_t - 8\beta(1-\beta)h^n(1/2)_t \xrightarrow{P} 0$$

This means

$$[\varphi_\beta + \varphi_{1-\beta} - 8\beta(1-\beta)\varphi_{\frac{1}{2}}](e^x) * \nu_t^n \xrightarrow{P} 0$$

Denote $\lambda_\beta(e^x) = [\varphi_\beta + \varphi_{1-\beta} - 8\beta(1-\beta)\varphi_{\frac{1}{2}}](e^x)$ and $\eta = 1 - 4\beta(1-\beta)$; $\quad 0 < \eta < 1$, we obtain

$$\lambda_\beta(e^x) = e^{\frac{1}{2}x}[2\eta ch(x/2) - 2ch(\frac{\eta^{1/2}}{2}x) + 2(1-\eta)]$$

by derivation becomes clear that

$$\lambda_\beta(e^x) \geq 0 \quad \text{and} \quad \lambda_\beta(e^x) = 0 \text{ iff } x = 0$$

Furthermore

$$\lambda_\beta(e^x) \sim \eta e^x \quad \text{if} \quad x \to +\infty$$
$$\lambda_\beta(e^x) \sim \eta \quad \text{if} \quad x \to -\infty$$

And we obtain the following two facts: for any $\varepsilon > 0$

$$\mathbb{I}_{\{x>\varepsilon\}} e^x * \nu_t^n \xrightarrow{P} 0, \qquad \mathbb{I}_{\{x<-\varepsilon\}} * \nu_t^n \xrightarrow{P} 0 \tag{4.7}$$

which clearly implies $[\delta - D]$.

Proof of $[\gamma - D]$

As $h^n(\alpha)_t \xrightarrow{P} \frac{\alpha(1-\alpha)}{2} C_t$, (2) in of Lemma 2 will give the result if we verify (4.5) and (4.6).

(We have just verified (4.3) and (4.4)). Let us now see (4.5).

As $\varphi_{1/2} \sim \frac{1}{8}x^2$ if $x \to 0$ the tightness is a consequence of $1_{\{|x|<\varepsilon\}} x^2 * \nu_t^{Y^n} \prec 8h^n(1/2)_t$ (where $A \prec B$ stands for $B - A$ is an increasing process)

Similarly (4.6) follows from the fact

$$Var(B_t^n) \prec K_0 h^n(1/2)_t \tag{4.8}$$

for some K_0, and this is a consequence of the existence of K_1 such that

$$|e^x - 1 - h(x)| \leq K_1 \varphi_{\frac{1}{2}}(e^x)$$

Proof of $[\beta - D]$

From (2) in Proposition 1 we obtain:

$$-2B_t^n = < Y^n, Y^n >_t + 2(e^x - 1 - h(x)) * \nu_t^n$$

As $C - 2B = 0$ in the limit, it is enough to prove $C_t^n - 2B_t^n \xrightarrow{P} 0$

Now

$$\begin{aligned} C_t^n - 2B_t^n &= [2(e^x - 1 - h(x)) - h(x)^2] * \nu_t^n + \sum_{s \leq t} (\Delta B_s^n)^2 \\ &= \mathbb{I}_{\{|x| \geq \varepsilon\}} [2(e^x - 1 - h(x)) - h(x)^2] * \nu_t^n \\ &+ \mathbb{I}_{\{|x|<\varepsilon\}} [2(e^x - 1 - h(x)) - h(x)^2] * \nu_t^n \\ &+ \sum_{s \leq t} (\Delta B_s^n)^2 \\ &= K_1^n + K_2^n + K_3^n \end{aligned}$$

$K_1^n \xrightarrow{P} 0$ as a consequence of (4.7)

In what respects K_2^n as $|2(e^x - 1 - h(x)) - h(x)^2| \leq x^3 e^\varepsilon$ in the set $\{|x| < \varepsilon\}$ we obtain $K_2^n \xrightarrow{P} 0$ by the same argument as in Lemma 2 (1).

$K_3^n \xrightarrow{P} 0$ by (2) in Lemma 1. $\square$

Proof of (1) ⇒ (2) in Theorem 1.

Let us first verify the hypothesis of Theorem 2 in [13] for Y^n and Y.

As for any $t \in \mathbb{R}^+$ $Z_t > 0$ a.s. and $E(Z_t) = 1$ we obtain

$$(P_t^n) \lhd (\tilde{P}_t^n) \lhd (P_t^n) \tag{4.9}$$

From Theorem V.2.3 in [7] applied to $(\tilde{P}_t^n, P_t^n)$, we obtain $(h^n(1/2, P^n, \tilde{P}^n)_t | P_t^n)$ is tight, and $(Var(B_t^n)|P_t^n)$ is tight follows from (4.8). In consequence Theorem 2 in [13] gives

$$\tilde{C}_t^n \xrightarrow{P} C_t \quad \forall t \geq 0$$

Let us now verify conditions of Lemma 2:

(4.3) is a consequence of the fact that Y is continuous.

(4.5) is a consequence of

$$\mathbb{I}_{\{|x|<\varepsilon\}} x^2 * \nu_t^n \prec K_\varepsilon h^n(1/2)_t$$

Let us verify (4.4). The contiguity $(\tilde{P}_t^n) \lhd (P_t^n)$ gives, by Theorem V.2.3 in [7]

$$\lim_{a\to\infty} \limsup_n \tilde{P}^n(\mathbb{I}_{\{x>a\}} e^x * \nu_t^n > \eta) = 0 \quad \forall \eta > 0 \tag{4.10}$$

but in (4.4) instead of $\tilde{P}^n$ we need P. The possibility of changing $\tilde{P}^n$ into P follows from the fact $(P_t^n) \lhd (\tilde{P}_t^n)$ by the following argument:

Denote

$$F(n,a) = \{\mathbb{I}_{\{x>a\}} e^x * \nu^{Y^n} > \eta\}$$

and observe that $F(n,a) \subset F(n,b)$ if $a < b$. From (4.10) we know that there exist two sequences $a_k \nearrow \infty$ and (n_k) subsequence of $\mathbb{N}$ such that

$$\tilde{P}^{n_k}(F(n_k, a_k)) \leq \frac{1}{k}$$

Then suppose

$$\lim_{a\to\infty} \limsup_n P(F(n,a)) = \theta > 0$$

We have

$$\limsup_n P(F(n,a)) \geq \theta$$

for all a, so exists $n_k^{'} \geq n_k$ such that $P(F(n_k^{'}, a_k)) \geq \theta$ for all k.

Since $\tilde{P}^{n_k}(F(n_k^{'}, a_k)) \leq \frac{1}{k}$ this contradicts $(P_t^n) \lhd (\tilde{P}_t^n)$ and (4.4) follows, so (4.8) and (4.4) gives

$$h^n(\alpha)_t \xrightarrow{P} \frac{\alpha(1-\alpha)}{2} C_t \quad \forall t \geq 0$$

Proof of (2) ⇔ (4).

Finally we have to prove:

Under

$$h^n(1/2)_t \xrightarrow{P} \frac{1}{8} C_t \tag{4.11}$$

$$I_t^n(1+\varepsilon) \xrightarrow{P} 0 \quad \forall \varepsilon \geq 0 \quad \Leftrightarrow \quad h^n(\alpha)_t \xrightarrow{P} \frac{\alpha(1-\alpha)}{2} C_t \quad \alpha \in (0,1)$$

Proof of ($\Rightarrow$)

$$I_t^n(1+\varepsilon) = [(1-e^x)\mathbb{I}_{\{e^x<1/(1+\varepsilon)\}} + (e^x-1)\mathbb{I}_{\{e^x>1+\varepsilon\}}] * \nu_t^n \xrightarrow{P} 0$$

clearly implies (4.3) and (4.4) in Lemma 2. As $h^n(1/2)_t$ converges in probability it is tight, giving (4.5). Then (4.11) and (1) in Lemma 2 give the result.

Proof of ($\Leftarrow$)

As in the proof of Theorem 1 we obtain (4.7) and

$$I_t^n(1+\varepsilon) \leq [\mathbb{I}_{\{x<-log(1+\varepsilon)\}} + e^x \mathbb{I}_{\{x>log(1+\varepsilon)\}}] * \nu_t^n$$

concludes the proof of (2) $\Leftrightarrow$ (4). The proof of Therem 1 is concluded observing that (2) $\Rightarrow$ (3) is trivial $\square$

Aknowledgements:. The author is in debt with A.N. Shiryaev, D.O. Kramkov and A.A. Guschin for helpful discussion.

References

1. I.V. Basawa, D.J. Scott, *Asymptotic optimal inference for nonergodic models*, Lecture notes in statistics 17, Berlin Heidelberg New York, Springer, 1983.
2. R.B. Davies, *Asymptotic inference when the amount of information is random*, Proc. Berkeley Conference in Honor of J. Neyman and J. Kiefer Vol 2, L. Le Cam, R. A. Olsen, Wadsworth Inc., 1985, pp. 841–864.
3. P.D. Feigin, *Stable convergence of semimartingales*, Stoch. Proc. and Appl. **19** (1985), 125–134.
4. P. Greenwood, A.N. Shiryaev, *Contiguity and the statistical invariance principle*, Gordon and Breach, London, 1985.
5. R. Höpfner, J. Jacod, L. Ladelli, *Local asymptotic normality and mixed normality for Markov statistical models*, Probab. Th. Rel. Fields **86** (1990), 105–129.
6. J. Jacod, *Convergence of filtered statistical models and Hellinger processes*, Stochastic Processes Appl. **32** (1989), 47–68.
7. Jacod, J. and Shiryaev, A. N., *Limit theorems for stochastic processes*, Springer, Berlin Heidelberg New York, 1987.
8. D.O. Kramkov, *Convergence of filtered experiments to the experiment generated by a semimartingale*, Lect. Notes in Math. Proceedings of the Soviet–Finish Simpos. on Probab. and Math. Stat., Springer, Berlin Heildeberg New York, 1992 (to appear).
9. L. Le Cam, and G.L. Yang, *Distinguished statistics, loss of information and a theorem of Robert B. Davies*, Statistical decision theory and related topics IV, (S.S. Gupta, J.O. Berger, eds.), vol. 2, Springer, New York, 1988, pp. 163-175.
10. 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.
11. H. Luschy, *Local asymptotic normality for semimartingales*, Probab. Th. Rel. Fields **92** (1992), 151–176.
12. D. L. Mc Leish, *An extended martingale principle*, Ann. Probab. **6** (1978), 144–150.
13. E. Mordecki, *Stable convergence of semimartingales: reciprocals*, Russian Math. Surveys (1993) (to appear).
14. A.N. Shiryaev, *Probability*, Springer, Berlin Heidelberg New York, 1985.
15. A.F. Taraskin, *On the behaviour of the likelihood ratio of semimartingales*, Theory Probab. Appl. **29** (1984), 452–464.
16. L. Yu. Vostirkova, *Functional limit theorems for the likelihood ratio processes*, Ann. Univ. Sci. Budapest R. Eotvos Nom 6 (1987), 145–182.
17. L.Yu. Vostrikova, *On the weak convergence of likelihood ratio processes of general statistical parametric models*, Stochastics **23** (1988), 277–298.