Convergence Results for Multivariate Martingales

Irene Crimaldi* - Luca Pratelli†

January 18, 2003

Abstract

The Central Limit Theorem for multivariate martingales is a very useful tool in applications, for example in order to obtain asymptotic results in Likelihood Theory for stochastic processes. In this paper we present a version of this result which is more general than others published so far. Indeed, we not only eliminate some superfluous assumptions but we also formulate the statement in terms of the convergence of suitable conditional expectations.

1 Introduction

The Central Limit Theorem for multivariate martingales is a very useful tool in applications, for example in order to obtain asymptotic results in Likelihood Theory for stochastic processes. In this paper we present a version of this result which is more general than others published so far. Indeed, we not only eliminate some superfluous assumptions but we also formulate the statement in terms of the convergence of suitable conditional expectations.

The structure of the paper is the following:

- In section 2 we give some notations.
- In section 3 we formulate our main Theorem in terms of stable convergence.
- In section 4 we give some recalls and lemmas that we need for the proof of the main Theorem.
- In section 5 we prove the main Theorem.
- In section 6 we compare our result with previously published limit theorems for martingales (see [7], [9]).
- In section 7 we introduce the notion of convergence with respect to a conditioning system and we illustrate an extension of the main Theorem given in section 3. More precisely, we get, under the same assumptions given in section 3, the convergence in L^1 of suitable conditional expectations. Referring to [3] for details, we observe that, in our case, this type of convergence implies the stable convergence.
- In section 8 we give some complements. In particular, we give a weak law of large numbers for martingales, under less restrictive assumptions than those used so far in literature.
- Finally, in section 9 we apply our results to the Likelihood Theory for stochastic processes in order to illustrate an example of situations in which results of this kind can be useful.

*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa

†Gruppo insegnamento Matematiche, Accademia Navale, Viale Italia 72, I-57100 Livorno

2 Notation

If a is a real matrix, we denote by $|a|$ the sum of the absolute values of its entries. With this notation, if a is equal to the row-column product of two matrices b, c, we have $|a| \leq |b|\,|c|$. Moreover, we denote by a' the transpose of a. A d-dimensional real vector x is identified with a column matrix: so the symbol $|x|$ denotes the sum of the absolute values of its entries. Hence, we have

$$|xx'| = |x|^2. \tag{1}$$

If U is a random variable with values in the space of positive semi-definite $d \times d$-matrices, the family of Gaussian distributions

$$\big(\mathcal{N}(0, U(\omega))\big)_{\omega \in \Omega}$$

is a Markovian kernel with respect to the couple

$$(\Omega, \mathcal{A}), \qquad (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)).$$

We denote this kernel by $\mathcal{N}(0, U)$ and we call it the *Gaussian kernel* associated with the random variable U. Moreover, given an event A, we denote by P_A the probability measure $P(\cdot|A)$ and by $P_A\mathcal{N}(0, U)$ the probability measure on $\mathbb{R}^d$ defined by

$$P_A\mathcal{N}(0, U)(B) = \int P_A(\mathrm{d}\omega)\, \mathcal{N}\big(0, U(\omega)\big)(B)$$

for each B in $\mathcal{B}(\mathbb{R}^d)$.

3 Main theorem

Theorem 3.1 *On a probability space $(\Omega, \mathcal{A}, P)$, endowed with a filtration $\mathcal{F}$ satisfying the usual conditions, let M be a (right-continuous with limits from the left) d-dimensional martingale such that $M_0 = 0$. Let us denote by Q the process $[M, M]$ (with values in the space of positive semi-definite $d \times d$-matrices). Further let $(a_t)_{t \geq 0}$ be a family of $d \times d$-matrices. Let us suppose that, as t goes to $+\infty$, the following relations hold:*

(a) $|a_t| \longrightarrow 0$.

(b) $\mathrm{E}\big[\sup_{0 \leq s \leq t} |a_t \Delta_s M|\big] \longrightarrow 0$.

(c) $a_t Q_t a_t' \longrightarrow U$ *in probability (where U is a random variable with values in the space of positive semi-definite $d \times d$-matrices).*

Then the random vector $a_t M_t$ converges $\mathcal{A}$–stably to the Gaussian kernel $\mathcal{N}(0, U)$.

A definition of the concept of stable convergence can be found in [1], [4], [8] and [11].

4 Preliminaries

In order to prove the main theorem, we shall need some preliminary notions and lemmas. Let us start with the following proposition.

Proposition 4.1 *Under the same hypotheses as in Theorem 3.1, let S, T be two bounded stopping times, with $S \leq T$, and let $(S_{n,j})_{n\geq 1, j\geq 0}$ be a double-sequence of stopping times, with $S \leq S_{n,j} \leq T$ and satisfying the following properties:*

(a) For each n, the sequence $(S_{n,j})_{j\geq 0}$ is increasing and we have

$$S_{n,0} = S, \qquad \bigcup_{j\geq 0}\{S_{n,j} = T\} = \Omega. \tag{2}$$

(b) The sequence $\left(\sup_{j\geq 1}(S_{n,j} - S_{n,j-1})\right)_{n\geq 1}$ converges pointwise to zero.

Then the sequence

$$\left(\sum_{j\geq 1}(M_{S_{n,j}} - M_{S_{n,j-1}})\,(M_{S_{n,j}} - M_{S_{n,j-1}})'\right)_{n\geq 1}$$

converges in probability to $Q_T - Q_S$.

Remark 4.2 With the same assumptions as in the previous proposition, because of (2), the difference $M_T - M_S$ can be put in the form

$$M_T - M_S = \sum_{j\geq 1}(M_{S_{n,j}} - M_{S_{n,j-1}}), \tag{3}$$

i.e. it can be expressed as the pointwise limit of the finite sums

$$\sum_{j=1}^{k}(M_{S_{n,j}} - M_{S_{n,j-1}}).$$

Similarly, each sum of the form

$$\sum_{j\geq 1}(M_{S_{n,j}} - M_{S_{n,j-1}})\,(M_{S_{n,j}} - M_{S_{n,j-1}})'$$

can be expressed as the pointwise limit of its partial sums.

Moreover, we shall need the following "vectorial version" of the result obtained in [12].

Proposition 4.3 *On the probability space $(\Omega, \mathcal{A}, P)$ let $(X_{n,j})_{n\geq 1, 1\leq j\leq k_n}$ be a triangular array of d-dimensional real random vectors, such that, for each n, the finite sequence $(X_{n,j})_{1\leq j\leq k_n}$ is a martingale difference array with respect to a given filtration $(\mathcal{G}_{n,j})_{j\geq 0}$. Let us set*

$$X_n^* = \sup_{1\leq j\leq k_n}|X_{n,j}|, \qquad U_n = \sum_{j=1}^{k_n} X_{n,j}X_{n,j}'.$$

Moreover, let us denote by $\mathcal{H}_j$ the algebra $\liminf_n \mathcal{G}_{n,j}$ and by $\mathcal{U}$ the σ-field generated by the algebra $\bigcup_j \mathcal{H}_j$. Let us suppose that (X_n^*) converges to zero in L^1 and that (U_n) converges in probability to a random variable U (with values in the space of positive semi-definite $d \times d$-matrices), which is measurable with *respect to $\mathcal{U}$.*

Then the sequence $\left(\sum_{j=1}^{k_n} X_{n,j}\right)_{n\geq 1}$ converges $\mathcal{U}$–stably to the Gaussian kernel $\mathcal{N}(0,U)$.

The following lemmas will be useful.

Lemma 4.4 *With the same assumptions as in Theorem 3.1, let us denote by M^* the real process defined by $M^*_t = \sup_{0\leq s\leq t}|M_s|$. Further, for each real number t, let us denote by T_t the first entrance time of the real process $|a_t|^{1/2}M^* \vee |a_t|\,|Q|$ into $]1,\infty[$. Finally, let us set*

$$S_t = T_t \wedge t. \tag{4}$$

Then the following statements hold:

(a) S_t is a stopping time which is bounded by the constant t.

(b) S_t goes to $+\infty$.

(c) $|a_t\, M_{S_t}|$ converges to zero in L^1.

(d) $|a_t\, Q_{S_t} a_t'|$ converges to zero in probability.

Remark 4.5 For each t, let M^i_t be the ith entry of the random vector M_t and $Q^{i,j}_t$ the (i,j)th entry of the random matrix Q_t. Then the jump of the real process $Q^{i,j}$ at time s is given by $\Delta_s Q^{i,j} = (\Delta_s M^i)\,(\Delta_s M^j)$ and so the jump of the process Q at time s is given by

$$\Delta_s Q = (\Delta_s M)\,(\Delta_s M)'. \tag{5}$$

Proof of Lemma 4.4. Since the two processes M^*, $|Q|$ are adapted and right-continuous, T_t is a stopping time. So, because of (4), we get statement (a).

By assumption (a) in Theorem 3.1 and the fact that the two processes M^*, $|Q|$ have locally bounded trajectories, we obtain statement (b).

In order to prove statement (c), it suffices to observe that we have

$$M_{S_t} = M_{S_t-} + \Delta_{S_t} M, \qquad |a_t|^{1/2}\,|M_{S_t-}| \leq 1,$$

and so

$$\begin{aligned}
|a_t M_{S_t}| &\leq |a_t M_{S_t-}| + |a_t \Delta_{S_t} M| \\
&\leq |a_t|\,|M_{S_t-}| + \sup\nolimits_{0\leq s\leq t} |a_t \Delta_s M| \\
&\leq |a_t|^{1/2} + \sup\nolimits_{0\leq s\leq t} |a_t \Delta_s M|.
\end{aligned}$$

Similarly, in order to prove statement (d), it suffices to observe that we have

$$Q_{S_t} = Q_{S_t-} + \Delta_{S_t} Q, \qquad |a_t|\,|Q_{S_t-}| \leq 1,$$

and so, because of (5),

$$\begin{aligned}
|a_t Q_{S_t} a_t'| &\leq |a_t Q_{S_t-} a_t'| + |a_t (\Delta_{S_t} Q) a_t'| \\
&\leq |a_t|\, |Q_{S_t-}|\, |a_t| + \sup_{0\leq s\leq t} |a_t (\Delta_s Q) a_t'| \\
&\leq |a_t| + \sup_{0\leq s\leq t} |a_t (\Delta_s M)(\Delta_s M)' a_t'| \\
&= |a_t| + \sup_{0\leq s\leq t} |a_t \Delta_s M|^2
\end{aligned}$$

(where the last equality follows from (1)).

Lemma 4.6 *Under the same assumptions as in Theorem 3.1, let t be a positive number and S_t the stopping time defined in the previous lemma. Then, for each strictly positive number ϵ, there exist an increasing sequence $(T_j)_{j\geq 0}$ of stopping times, with $S_t \leq T_j \leq t$, and a strictly positive integer k, such that the array $(X_j)_{1\leq j\leq k}$ defined by*

$$X_j = a_t (M_{T_j} - M_{T_{j-1}}) \quad \textit{for} \quad 1 \leq j \leq k \tag{6}$$

(which is a martingale difference array with respect to the filtration $(\mathcal{F}_{T_j})_{j\geq 0}$) has the following properties:

$$P\left\{ \left| a_t (M_t - M_{S_t}) - \textstyle\sum_{j=1}^k X_j \right| > \epsilon \right\} < \epsilon, \tag{7}$$

$$P\left\{ \left| a_t (Q_t - Q_{S_t}) a_t' - \textstyle\sum_{j=1}^k X_j X_j' \right| > \epsilon \right\} < \epsilon, \tag{8}$$

$$\sup_{1\leq j\leq k} |X_j| \leq \epsilon + \sup_{0\leq s\leq t} |a_t \Delta_s M|. \tag{9}$$

Proof. Let us fix $\epsilon > 0$. For each integer n, let us define (by induction) the increasing sequence $(S_{n,j})_{j\geq 0}$ of stopping times, with $S_t \leq S_{n,j} \leq t$, setting

$$S_{n,0} = S_t, \qquad S_{n,j} = t \wedge (S_{n,j-1} + n^{-1}) \wedge U_{n,j},$$

where

$$U_{n,j}(\omega) = \inf\left\{ s \in \mathbb{R} : \; s > S_{n,j-1}(\omega), \; \left| a_t \big(M_s(\omega) - M_{S_{n,j-1}}(\omega) \big) \right| > \epsilon \right\}. \tag{10}$$

Then, it is easy to verify that conditions (a), (b) in Proposition 4.1 (where S, T are replaced by S_t, t) are satisfied. Thus, because of Proposition 4.1 and Remark 4.2, it is possible to find a pair (n, k) of strictly positive integers such that, setting $T_j = S_{n,j}$ and defining X_j by relation (6), conditions (7), (8) are satisfied. Moreover, condition (9) holds because of relation (10).

Lemma 4.7 *On the space $(\Omega, \mathcal{A})$ let $\mathcal{F}$ be a right-continuous filtration, and let us set $\mathcal{F}_\infty = \bigvee_t \mathcal{F}_t$. Further, let $(T_n)_{n\geq 1}$ be a sequence of finite stopping times going to $+\infty$ and let us denote by $\mathcal{H}$ the algebra $\liminf_n \mathcal{F}_{T_n}$.*

Then the σ-field generated by $\mathcal{H}$ coincides with the whole σ-field $\mathcal{F}_\infty$.

Proof. The sequence $(T_n)_{n\geq 1}$ is bounded from below by the increasing sequence

$$(T_n \wedge T_{n+1} \wedge \cdots)_{n\geq 1},$$

which goes to $+\infty$ and which is still a sequence of stopping times (because of the right-continuity of the filtration $\mathcal{F}$). Thus, without loss of generality, we may suppose that the sequence $(T_n)_{n\geq 1}$ itself is increasing. In this case, the algebra $\mathcal{H}$ coincides with the union of the σ-fields $\mathcal{F}_{T_n}$. Therefore, if we take a positive real number t and an element A in $\mathcal{F}_t$, we have, for each n,

$$A \cap \{t < T_n\} \in \mathcal{F}_{T_n} \subset \mathcal{H}.$$

Hence we find

$$A = \bigcup_n (A \cap \{t < T_n\}) \in \sigma(\mathcal{H}),$$

where $\sigma(\mathcal{H})$ denotes the σ-field generated by $\mathcal{H}$. Since A is arbitrary, we get the inclusion $\mathcal{F}_t \subset \sigma(\mathcal{H})$. Finally, since t is arbitrary, we get the inclusion $\mathcal{F}_\infty \subset \sigma(\mathcal{H})$, and so we can conclude.

5 Proof of Theorem 3.1

We are now able to prove Theorem 3.1.

Let us fix an increasing sequence $(t_n)_{n\geq 1}$ of real positive numbers going to $+\infty$. It is enough to prove that the sequence $\left(a_{t_n} M_{t_n}\right)_{n\geq 1}$ converges $\mathcal{F}_\infty$–stably to the Gaussian kernel $\mathcal{N}(0, U)$. To this end, let us start observing that, if we denote by S_n the stopping time S_{t_n} defined in Lemma 4.4, this lemma shows that S_n goes to $+\infty$ and the following two limits hold in probability:

$$\left|a_{t_n} M_{S_n}\right| \longrightarrow 0, \tag{11}$$

$$\left|a_{t_n} Q_{S_n} a'_{t_n}\right| \longrightarrow 0. \tag{12}$$

Therefore it suffices to prove that the sequence $\left(a_{t_n}(M_{t_n} - M_{S_n})\right)_{n\geq 1}$ converges $\mathcal{F}_\infty$–stably to the kernel $\mathcal{N}(0, U)$. To this end, let us apply Lemma 4.6 with t_n instead of t and with n^{-1} instead of ϵ. Thus we obtain a double sequence $(T_{n,j})_{n\geq 1, j\geq 0}$ of stopping times and a sequence $(k_n)_{n\geq 1}$ of strictly positive integers such that, for each n, the sequence $(T_{n,j})_{j\geq 0}$ is an increasing sequence of stopping times with $S_n \leq T_{n,j} \leq t_n$ and, setting

$$X_{n,j} = a_{t_n}\left(M_{T_{n,j}} - M_{T_{n,\,j-1}}\right) \quad \text{for} \quad 1 \leq j \leq k_n,$$

the following properties hold:

$$P\left\{\left|a_{t_n}(M_{t_n} - M_{S_n}) - \textstyle\sum_{j=1}^{k_n} X_{n,j}\right| > n^{-1}\right\} < n^{-1}, \tag{13}$$

$$P\left\{\left|a_{t_n}(Q_{t_n} - Q_{S_n})a'_{t_n} - \textstyle\sum_{j=1}^{k_n} X_{n,j} X'_{n,j}\right| > n^{-1}\right\} < n^{-1}, \tag{14}$$

$$\sup_{1\leq j\leq k_n} |X_{n,j}| \leq n^{-1} + \sup_{0\leq s\leq t_n} |a_{t_n} \Delta_s M|. \tag{15}$$

Then, setting $\mathcal{G}_{n,j} = \mathcal{F}_{T_{n,j}}$, the triangular array $(X_{n,j})_{n\geq 1,\, 1\leq j\leq k_n}$, satisfies (with respect to $(\mathcal{G}_{n,j})_{n\geq 1,\, j\geq 0}$) the assumptions of Proposition 4.3. More precisely: from relation (15) we deduce, by assumption (b) in Theorem 3.1, that the sequence

$$\big(\sup_{1\leq j\leq k_n} |X_{n,j}|\big)_{n\geq 1}$$

converges in L^1 to zero; from relations (12) and (14) we deduce, by assumption (c) in Theorem 3.1, that the sequence

$$\big(\textstyle\sum_{j=1}^{k_n} X_{n,j} X'_{n,j}\big)_{n\geq 1}$$

converges in probability to U. Finally, since U is measurable with respect to $\mathcal{F}_\infty$, the condition of measurability required for U in Proposition 4.3 is obviously verified: indeed, for each j, the sequence $(T_{n,j})_{n\geq 1}$ (which is bounded from below by $(S_n)_{n\geq 1}$) goes to $+\infty$, and so, by Lemma 4.7, the σ-field generated by the algebra

$$\mathcal{H}_j = \liminf_n \mathcal{F}_{T_{n,j}}$$

(and, consequently, the σ-field $\mathcal{U}$) coincides with $\mathcal{F}_\infty$.

Thus, applying Proposition 4.3, we get that the sequence

$$\big(\textstyle\sum_{j=1}^{k_n} X_{n,j}\big)_{n\geq 1} \tag{16}$$

converges $\mathcal{F}_\infty$–stably to the kernel $\mathcal{N}(0, U)$. Then, in order to conclude, it suffices to observe that, thanks to (13), the sequence $\big(a_{t_n}(M_{t_n} - M_{S_n})\big)_{n\geq 1}$ differs from the sequence (16) up to a sequence converging in probability to zero.

6 Comparison with previous results

Let us start with the comparison between Theorem 3.1 and the following theorem proved in [9] (see also Theorem A.7.7 in [10]):

Theorem 6.1 *On a probability space $(\Omega, \mathcal{A}, P)$, endowed with a filtration $\mathcal{F}$ satisfying the usual conditions, let M be a (right-continuous with limits from the left) d-dimensional martingale such that $M_0 = 0$ and $M^i_t \in L^2$ for each t and i. Let us denote by Q the process $[M, M]$ (with values in the space of positive semi-definite $d\times d$-matrices). Further let $(a_t)_{t\geq 0}$ be a family of invertible $d\times d$-matrices. Let us suppose that, as t goes to $+\infty$, the following relations hold:*

(a) $|a_t| \longrightarrow 0$.

(b) $\sum_i |a^{i,j}_t|\, \mathrm{E}\big[\sup_{0\leq s\leq t} |\Delta_s M^j|\big] \longrightarrow 0$ *for each* j.

(c) $a_t Q_t a'_t \longrightarrow U$ *in probability (where U is a random variable with values in the space of positive semi-definite $d\times d$-matrices).*

(d) $a_t \mathrm{E}[M_t M'_t] a'_t \longrightarrow \Sigma$ *where Σ is a positive definite $d\times d$-matrix.*

Then the random vector $a_t M_t$ converges $\mathcal{A}$–stably to the Gaussian kernel $\mathcal{N}(0, U)$.

In Theorem 3.1 we do not need the following assumptions:

$$M_t^i \in L^2, \quad \det a_t \neq 0.$$

Moreover, we do not need assumption (d), and assumption (b) in Theorem 3.1 is weaker than assumption (b) in Theorem 6.1. Indeed, we have:

$$\begin{aligned}\sup_{0\leq s\leq t} |a_t \Delta M_s| &\leq \sup_{0\leq s\leq t} \textstyle\sum_i \sum_j |a_t^{i,j}|\,|\Delta M_s^j| \\ &\leq \textstyle\sum_j \sum_i |a_t^{i,j}| \sup_{0\leq s\leq t} |\Delta M_s^j|.\end{aligned}$$

Finally, let us do a comparison between Theorem 3.1 and the following theorem proved in [7] (see Theorem 12.6):

Theorem 6.2 *On a probability space $(\Omega, \mathcal{A}, P)$, endowed with a filtration $\mathcal{F}$ satisfying the usual conditions, let M be a (right-continuous with limits from the left) d-dimensional martingale such that $M_0 = 0$ and $M_t^i \in L^2$ for each t and i. Let us denote by Q the process $[M, M]$ (with values in the space of positive semi-definite $d \times d$-matrices). Further let $(a_t)_{t\geq 0}$ be a family of diagonal $d \times d$-matrices with $a_t^{i,i} > 0$. Let us suppose that, as t goes to $+\infty$, the following relations hold:*

(a) $a_t^{i,i} \downarrow 0$ for each i.

(b) $a_t^{i,i} \sup_{0\leq s\leq t} |\Delta_s M^i| \longrightarrow 0$ in probability for each i.

(c) $a_t Q_t a_t \longrightarrow U$ in probability (where U is a random variable with values in the space of positive semi-definite $d \times d$-matrices).

(d) $a_t \mathrm{E}[M_t M_t'] a_t \longrightarrow \Sigma$ where Σ is a positive definite $d \times d$-matrix.

Then the random vector $a_t M_t$ converges $\mathcal{A}$–stably to the Gaussian kernel $\mathcal{N}(0, U)$.

In Theorem 3.1 the matrices a_t are not necessarily diagonal and their entries are not necessarily decreasing. Moreover, assumptions (b) and (d) in Theorem 6.2 imply assumption (b) in Theorem 3.1. Indeed, we have:

$$\sup_{0\leq s\leq t} |a_t \Delta M_s| \leq \textstyle\sum_i a_t^{i,i} \sup_{0\leq s\leq t} |\Delta M_s^i|$$

and so assumption (b) in Theorem 6.2 implies that, as t goes to $+\infty$,

$$\sup_{0\leq s\leq t} |a_t \Delta M_s|$$

converges in probability to zero. On the other hand, assumption (d) in Theorem 6.2 implies that, for each i,

$$\big(\, (a_t^{i,i})^2\, \mathrm{E}\big[(M_t^i)^2\big]\, \big)_t$$

is bounded. Since we have $\mathrm{E}[(M_t^i)^2] = \mathrm{E}[Q_t^{i,i}]$ and

$$\sup_{0\leq s\leq t} |\Delta M_s^i|^2 \leq \textstyle\sum_{0\leq s\leq t} |\Delta M_s^i|^2 \leq Q_t^{i,i},$$

we get that, for each i,

$$\big((a_t^{i,i})^2\, \mathrm{E}\big[\sup_{0\leq s\leq t} |\Delta M_s^i|^2\big]\big)_t$$

is bounded, that is

$$\left(a_t^{i,i}\ \sup_{0\le s\le t}|\Delta M_s^i|\right)_t$$

is bounded in L^2 and so it is uniformly integrable. Thus the assumption (b) in Theorem 3.1 is satisfied.

7 An extension of the main theorem

Let us start with the following definitions. (We refer to [3] for more details and for a systematic study of these concepts.)

Definition 7.1 We call *conditioning system* a family $\mathcal{G} = (\mathcal{G}_t)_{t\ge 0}$ of sub-σ-fields of $\mathcal{A}$.

Definition 7.2 Let $(X_t)_{t\ge 0}$ be a family of d-dimensional real random vectors, $\mathcal{G}$ a conditioning system and N a Markovian kernel with respect to the pair $\big((\Omega,\mathcal{A}),\, (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\big)$. We say that $(X_t)_t$ *converges with respect to* $\mathcal{G}$ *to* N if, for each bounded continuous function f on $\mathbb{R}^d$, the conditional expectation $\mathrm{E}[f(X_t)|\mathcal{G}_t]$ converges in L^1 to the random variable Nf defined by $Nf(\omega) = \int N(\omega, \mathrm{d}x) f(x)$. If the kernel N is constant, i.e it is reduced to a distribution μ, we say that $(X_t)_t$ converges with respect to $\mathcal{G}$ to μ.

We can reformulate Theorem 3.1 as the following:

Theorem 7.3 *Under the same assumptions and with the same notation as in Theorem 3.1, let (S_t) be a family of stopping times with properties (a)-(d) in Lemma 4.4.*

Then the random vector $a_t M_t$ converges with respect to the conditioning system $(\mathcal{F}_{S_t})_t$ to the Gaussian kernel $\mathcal{N}(0, U)$.

The proof of this theorem is exactly the same of Theorem 3.1, but we have to apply the following result instead of Proposition 4.3. (We refer to [3] for the proof of the following theorem.)

Theorem 7.4 *On the probability space $(\Omega, \mathcal{A}, P)$ let $(X_{n,j})_{n\ge 1,\, 1\le j\le k_n}$ be a triangular array of d-dimensional real random vectors, such that, for each n, the finite sequence $(X_{n,j})_{1\le j\le k_n}$ is a martingale difference array with respect to a given filtration $(\mathcal{G}_{n,j})_{j\ge 0}$. Let us set*

$$X_n^* = \sup_{1\le j\le k_n}|X_{n,j}|, \qquad U_n = \sum_{j=1}^{k_n} X_{n,j}X_{n,j}'$$

and, for each j, let us denote by $\mathcal{H}_j$ the algebra $\liminf_n \mathcal{G}_{n,j}$. Moreover, let us suppose that (X_n^*) converges to zero in L^1 and that (U_n) converges in probability to a random variable U (with values in the space of positive semi-definite $d\times d$-matrices), which is measurable with respect to the σ-field generated by a certain $\mathcal{H}_j$.

Then the sequence $\left(\sum_{j=1}^{k_n} X_{n,j}\right)_{n\geq 1}$ *converges with respect to the conditioning system* $(\mathcal{G}_{n,j})_n$ *to the Gaussian kernel* $\mathcal{N}(0,U)$.

Indeed, under the assumptions of Theorem 7.3, if we use the same notations used in the proof of Theorem 3.1, we can say that the triangular array $(X_{n,j})$ and the σ-fields $\mathcal{G}_{n,j}$ satisfy the assumptions of Theorem 7.4. In particular, as we have already noticed, the random variable U is measurable with respect to $\mathcal{F}_\infty$ and, for each j, the σ-field generated by the algebra

$$\mathcal{H}_j = \liminf_n \mathcal{F}_{T_{n,j}}$$

coincides with $\mathcal{F}_\infty$. Thus, applying Theorem 7.4, we get that the sequence

$$\left(\sum_{j=1}^{k_n} X_{n,j}\right)_{n\geq 1} \tag{17}$$

converges to the kernel $\mathcal{N}(0,U)$ with respect to the conditioning system $(\mathcal{G}_{n,0})$, which coincides with $(\mathcal{F}_{S_n})$.

Remark 7.5 Under the assumptions of Theorem 7.3, let us use the same notations used in the proof of Theorem 3.1. Since U is measurable with respect to $\mathcal{F}_\infty$ and $\mathcal{F}_\infty$ coincides with the σ-field generated by the algebra $\liminf_n \mathcal{F}_{S_n}$, it can be shown (see [3]) that the convergence with respect to $(\mathcal{F}_{S_n})$ to $\mathcal{N}(0,U)$ implies the $\mathcal{F}_\infty$–stable convergence. Thus, Theorem 7.3 is really stronger than Theorem 3.1.

8 Some complements

From Theorem 3.1 we get the following corollary:

Corollary 8.1 *With the same assumptions and notation as in Theorem 3.1, let us set* $A = \{\det U > 0\}$. *Moreover, let us suppose* $P(A) > 0$ *and denote by* P_A *the probability measure* $P(\,\cdot\,|A)$. *Then, under* P_A, *we have the following* $\mathcal{A}$*–stable convergences:*

(a) $U^{-\frac{1}{2}} a_t M_t \longrightarrow \mathcal{N}(0, I_d)$.

(b) $\left(a_t Q_t a_t'\right)^{-\frac{1}{2}} a_t M_t \longrightarrow \mathcal{N}(0, I_d)$.

(c) $M_t' Q_t^{-1} M_t \longrightarrow \chi^2(d)$.

Proof. The convergences (a), (b) and (c) are immediate consequences of Theorem 3.1 and of elementary properties of stable convergence. In particular, for (b) and (c) let us observe that on the set A the matrix $a_t Q_t a_t'$ is positive definite for t large enough.

It is worthwhile to notice that the previous convergences hold also in the sense of the convergence with respect to the conditioning system $(\mathcal{F}_{S_t})$ defined in section 7.

Moreover, from Theorem 3.1 we deduce the following "weak law of large numbers" for martingales:

Corollary 8.2 *With the same assumptions and notation as in the previous corollary, the random vector $Q_t^{-1}M_t$ converges to zero in probability under P_A.*

Proof. It suffices to observe that

$$Q_t^{-1}M_t = a_t'(a_tQ_ta_t')^{-1}a_tM_t.$$

We conclude this section observing that, under a suitable assumption on the martingale M, the family (S_t) of stopping times with properties (a)-(d) in Lemma 4.4 may be taken deterministic.

Proposition 8.3 *With the same assumptions and notation as in Lemma 4.4, let us suppose that, for each t, the random variable M_t^* is integrable. Further, for each real number t, let us denote by $f(t)$ the first entrance time of $s \mapsto |a_t|^{1/2}\mathrm{E}[M_s^*]$ into $]1,\infty[$. Finally, let us set*

$$g(t) = f(t) \wedge t. \tag{18}$$

Then the following statements hold:

(a) g is right-continuous with limits from the left and bounded by the constant t.

(b) $g(t)$ goes to $+\infty$.

(c) $\sup_{0\le s\le g(t)} |a_t\, M_s|$ (and so $|a_t\, M_{g(t)}|$) converges to zero in L^1.

(d) $|a_t\, Q_{g(t)}a_t'|$ converges to zero in probability.

Proof. Statement (a) follows from the monotone convergence theorem and relation (18). By assumption (a) in Theorem 3.1, we obtain statement (b). In order to prove statement (c), it suffices to observe that we have

$$M_s = M_{s-} + \Delta_s M, \qquad |a_t|^{1/2}\,\mathrm{E}\big[M^*_{g(t)-}\big] \le 1,$$

and so

$$\begin{aligned}\mathrm{E}\big[\,\sup_{0\le s\le g(t)}|a_tM_s|\,\big] &\le \mathrm{E}\big[\,\sup_{0\le s\le g(t)}|a_tM_{s-}|\,\big] + \mathrm{E}\big[\,\sup_{0\le s\le g(t)}|a_t\Delta_sM|\,\big]\\ &\le |a_t|\,\mathrm{E}\big[M^*_{g(t)-}\big] + \mathrm{E}\big[\,\sup_{0\le s\le t}|a_t\Delta_sM|\,\big]\\ &\le |a_t|^{1/2} + \mathrm{E}\big[\,\sup_{0\le s\le t}|a_t\Delta_sM|\,\big].\end{aligned}$$

Finally, in order to obtain statement (d), we observe that, thanks to the Burkholder-Davis-Gundy inequality, if we set $N = a_tM$, we have:

$$\mathrm{E}\big[\,\sup_{0\le s\le g(t)}\big(\textstyle\sum_i (N_s^i)^2\big)^{1/2}\,\big] \sim \mathrm{E}\big[\,\big(\textstyle\sum_i [N^i,N^i]_{g(t)}\big)^{1/2}\,\big].$$

Moreover, the following inequalities hold:

$$\big(\textstyle\sum_i (N_s^i)^2\big)^{1/2} \le \textstyle\sum_i |N_s^i| = |N_s|, \qquad \textstyle\sum_{i,j}|\,[N^i,N^j]_{g(t)}\,| \le c(d)\textstyle\sum_i [N^i,N^i]_{g(t)}$$

(where $c(d)$ is a constant which depends on d). Thus, since

$$\textstyle\sum_{i,j}|\,[N^i,N^j]_{g(t)}\,| = |a_tQ_{g(t)}a_t'|,$$

we may conclude.

9 Application to likelihood theory

Let $\big(\Omega, \mathcal{A}, \{\mathcal{F}_t\}, \{P_\theta : \theta \in \Theta\}\big)$ be a filtered statistical model, where $\{\mathcal{F}_t\}$ is a filtration with usual conditions and Θ is an open subset of $\mathbb{R}^d$. Let us assume that the probability measures P_θ are locally equivalent to a probability measure P on $\mathcal{A}$. This means that, if we denote by P_θ^t and P^t the restrictions of P_θ and P to $\mathcal{F}_t$ respectively, we have

$$P_\theta^t \sim P^t$$

for each $t \geq 0$ and $\theta \in \Theta$.

For every θ in Θ, the stochastic process $\big(L_t(\theta)\big)_{t\geq 0}$ defined by

$$L_t(\theta) = \frac{\mathrm{d}P_\theta^t}{\mathrm{d}P^t}$$

is a martingale under P. (We take the version which is right-continuous with limits from the left.) The map

$$\theta \longmapsto L_t(\theta)$$

is called the *likelihood function* corresponding to the experiment with index t.

Let us set $\ell_t(\theta) = \log\big(L_t(\theta)\big)$. If, for each t, the likelihood function is continuously differentiable and the set of gradients $\{\nabla L_t(\theta) : \theta \in \Theta\}$ is locally dominated integrable under P, then, for each θ, the $\mathbb{R}^d$-valued stochastic process $\big(\dot{\ell}_t(\theta)\big)_{t\geq 0}$ defined by

$$[\dot{\ell}_t(\theta)]^i = \frac{\partial \ell_t(\theta)}{\partial \theta^i} = \frac{1}{L_t(\theta)}\,\frac{\partial L_t(\theta)}{\partial \theta^i}$$

is a martingale under P_θ (see [2]) and it is called the *score function.* In the sequel, we take the version of $\big(\dot{\ell}_t(\theta)\big)_t$ which is right-continuous with limits from the left.

In a similar way as it is done in [2] and [10] (but using Theorem 3.1), we find the following asymptotic result for the score function:

Theorem 9.1 *Let us denote by $Q(\theta)$ the process $[\dot{\ell}(\theta), \dot{\ell}(\theta)]$ (with values in the space of positive semi-definite $d \times d$-matrices). Further let $\big(a_t(\theta)\big)_{t\geq 0}$ be a family of $d \times d$-matrices. Let us suppose that, as t goes to $+\infty$, the following relations hold for each θ in Θ:*

(a) $|a_t(\theta)| \longrightarrow 0$.

(b) $\mathrm{E}_\theta\big[\sup_{0\leq s\leq t} |a_t(\theta)\Delta_s\dot{\ell}(\theta)|\big] \longrightarrow 0$, *where* $\mathrm{E}_\theta[\,\cdot\,]$ *denotes the expectation under* P_θ.

(c) $a_t(\theta)Q_t(\theta)a_t(\theta)' \longrightarrow U(\theta)$ *in probability under P_θ (where $U(\theta)$ is a random variable with values in the space of positive semi-definite $d \times d$-matrices).*

Then under the probability measure $P_\theta\big(\,\cdot\,|\{\det U(\theta) > 0\}\big)$, we have the following $\mathcal{A}$–stable convergences:

(a) $U(\theta)^{-\frac{1}{2}}a_t(\theta)\dot{\ell}_t(\theta) \longrightarrow \mathcal{N}(0, I_d)$.

(b) $\left(a_t(\theta)Q_t(\theta)a_t(\theta)'\right)^{-\frac{1}{2}}a_t(\theta)\dot{\ell}_t(\theta) \longrightarrow \mathcal{N}(0, I_d)$.

(c) $\dot{\ell}_t(\theta)'Q_t(\theta)^{-1}\dot{\ell}_t(\theta) \longrightarrow \chi^2(d)$.

This theorem is a consequence of Theorem 3.1 and Corollary 8.1.

Moreover, if we assume that, for t sufficiently large, there exists a *maximum likelihood estimator*, that is a solution $\widehat{\theta}_t$ of the likelihood equation

$$\dot{\ell}_t(\theta) = 0,$$

we have the following asymptotic result:

Theorem 9.2 *Let us assume the same hypotheses as in the previous theorem. Moreover, let us suppose that $\theta \mapsto \dot{\ell}_t(\theta)$ is continuously differentiable and set*

$$j_t(\theta) = -\ddot{\ell}_t(\theta).$$

Finally, let us denote by $j_t(\theta^{(1)}, \dots, \theta^{(d)})$ the matrix obtained by replacing θ by $\theta^{(i)}$ in the ith row of $j_t(\theta)$.

If, for each vector $(\theta_t^{(1)}, \dots, \theta_t^{(d)})$ of convex combinations of θ and $\widehat{\theta}_t$,

$$a_t(\theta)j_t(\theta_t^{(1)}, \dots, \theta_t^{(d)})a_t(\theta)' \longrightarrow V(\theta)$$

in probability under P_θ, where $V(\theta)$ is a random variable with values in the space of positive semi-definite $d\times d$-matrices with $\det V(\theta) > 0$ on the set $\{\det U(\theta) > 0\}$, then we have the following $\mathcal{A}$–stable convergence under $P_\theta(\,\cdot\,|\{\det U(\theta) > 0\})$:

$$\left(a_tQ_t(\theta)a_t'\right)^{-\frac{1}{2}}a_t(\theta)j_t(\theta)(\widehat{\theta}_t - \theta) \longrightarrow \mathcal{N}(0, I_d).$$

Proof. By Taylor expansion, we find

$$\dot{\ell}_t(\theta) = j_t(\theta_t^{(1)}, \dots, \theta_t^{(d)})(\widehat{\theta}_t - \theta),$$

where $\theta_t^{(i)}$ is a convex combination of θ and $\widehat{\theta}_t$. Hence the two random vectors

$$\left(a_tQ_t(\theta)a_t'\right)^{-\frac{1}{2}}a_t(\theta)j_t(\theta)(\widehat{\theta}_t - \theta) \qquad \text{and} \qquad \left(a_tQ_t(\theta)a_t'\right)^{-\frac{1}{2}}a_t(\theta)\dot{\ell}_t(\theta)$$

differs up to the vector

$$\left(a_tQ_t(\theta)a_t'\right)^{-\frac{1}{2}}a_t(\theta)\left(j_t(\theta) - j_t(\theta_t^{(1)}, \dots, \theta_t^{(d)})\right)a_t(\theta)'(a_t(\theta)')^{-1}(\widehat{\theta}_t - \theta).$$

Thus, in order to conclude, it suffices to observe that

$$(a_t(\theta)')^{-1}(\widehat{\theta}_t - \theta) = (a_t(\theta)')^{-1}j_t(\theta_t^{(1)}, \dots, \theta_t^{(d)})^{-1}a_t(\theta)^{-1}a_t(\theta)\dot{\ell}_t(\theta).$$

These results (and other similar to these, see [2] and [10]) are useful to construct confidence regions in order to estimate the parameter θ. We conclude underlying that, thanks to our generalization of Theorem 6.1 and Theorem 6.2, we have been able to eliminate the assumption

$$a_t(\theta)\mathrm{E}_\theta\left[\,\dot{\ell}_t(\theta)\dot{\ell}_t(\theta)'\,\right]a_t(\theta)' \longrightarrow \Sigma(\theta)$$

(where $\Sigma(\theta)$ is a positive definite $d \times d$-matrix) that appears in the asymptotic results published so far in Likelihood Theory. In applications, this corresponds to an enlargement of the statistical model that can be taken.

References

[1] Aldous D.J. - Eagleson G.K., *On Mixing and Stability of Limit Theorems*, Ann. Probab., 6 (1978), 325-333.

[2] Barndorff-Nielsen O.E. - Sørensen M., *A Review of Some Aspects of Asymptotic Likelihood Theory for Stochastic Processes*, International Statistical Review (1994), 62, 1, 133-165.

[3] Crimaldi I. - Pratelli L., *Convergence with respect to a Conditioning System*, working paper.

[4] Crimaldi I., *Convergenza stabile e applicazioni*, graduation thesis (1998), Department of Mathematics, University of Pisa.

[5] Fegin P. D., *Stable Convergence of Semimartingales*, Stochastic Processes and their Applications, 19 (1985), 125-134.

[6] Hall P. - Heyde C. C., *Martingale Limit Theory and its Application*, Academic Press (1980).

[7] Heyde C. C., *Quasi-Likelihood and its Application*, Springer Series in Statistics, Springer-Verlag (1997).

[8] Jacod J. - Memin J., *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité*, Sem. de Prob. XV (1981), Lecture Notes in Math. 850, Springer 529-546.

[9] Küchler U. - Sørensen M., *A Note on Limit Theorems for Multivariate Martingales*, Bernoulli 5 (3), 1999, 483-493.

[10] Küchler U. - Sørensen M., *Exponential Families of Stochastic Processes*, Springer Series in Statistics, Springer-Verlag (1997).

[11] Letta G., *Convergence stable et applications*, Atti Sem. Mat. Fis. Univ. Modena, supplemento al vol. XLVI (1998), 191-211.

[12] Letta G. - Pratelli L., *Convergence stable vers un noyau gaussien*, Rendic. Acc. Naz. Sci. detta dei XL, Mem. di Mat. e Appl., 114, vol. XX (1996), 205-213.