Martingales

[These notes are based on the books:
* Probability with Martingales (1991) by David Williams

* Martingale Limit Theory and Its Application (1980) by Hall and Heyde

and • C.B. Phillips Lecture Notes]

In this course of Time Series Analysis, Martingale theory will be used to obtain asymptotic results (LLN and CLT).

The word martingale comes from Martingale

"... chausses, a medieval form of culottes that can be opened from behind."

Think on the relation of this meaning with the word martingale used in probability.

In probability the name Martingale is due to Ville (1939), "Etude Critique de la Notion de Collectif", Gauthier-Villars, Paris, and later on used by Levy (1937), Doob (1940).

Apart from the books I mentioned before there are two other books that contain excellent chapters on martingales:

* Stochastic Processes (1953) by Doob

and

* Real Analysis and Probability (1972) by Ash.

<u>Def</u>: A filtered space is the quadruple

$$(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$$

where

* $(\Omega, \mathcal{F}, P)$ is a probability space

* $\{\mathcal{F}_n : n \geq 0\}$ is a FILTRATION, that is, an increasing family of sub σ-algebras of $\mathcal{F}$:

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \ldots \subseteq \mathcal{F}.$$

We define

$$\mathcal{F}_\infty \equiv \sigma\left(\bigcup_n \mathcal{F}_n\right) \subseteq \mathcal{F}.$$

<u>Intuition</u>: The information about w in Ω available to us at time $\underline{n}$ consists precisely of the values of $Z(w)$ for all $\mathcal{F}_n$ measurable functions Z. Usually, $\{\mathcal{F}_n\}$ is the <u>natural filtration</u>

$$\mathcal{F}_n = \sigma(W_0, W_1, \ldots, W_n)$$

of some stochastic process $W = (W_n : n \in \mathbb{Z}^+)$ and then the information about w which we have at time n consists of the values

$$W_0(w), W_1(w), \ldots, W_n(w).$$

<u>Def</u> A process $Y = (Y_n ; n \geq 0)$ is called <u>adapted</u> (to the filtration $\{\mathcal{F}_n\}$) if for each n, Y_n is $\mathcal{F}_n$-measurable

<u>Intuition</u>: If Y is adapted, the value $Y_n(w)$ is known to us at time n. Usually, $\mathcal{F}_n = \sigma(W_0, W_1, \ldots W_n)$ and $Y_n = f_n(W_0, W_1, \ldots W_n)$ for some $\mathcal{B}^{n+1}$-measurable function f_n on $\mathbb{R}^{n+1}$.

Def: A process Y is called a martingale (relative to $(\{F_n\}, P)$) if

(i) Y is adapted
(ii) $E(|Y_n|) < \infty, \ \forall n$
(iii) $\boxed{E[Y_n | F_{n-1}] = Y_{n-1}}$, a.s $(n \geq 1)$.

A supermartingale (relative to $(\{F_n\}, P)$) is defined similarly, except that (iii) is replaced by

$$E[Y_n | F_{n-1}] \leq Y_{n-1} \quad \text{a.s} \quad (n \geq 1)$$

A submartingale is defined with (iii) replaced by

$$E[Y_n | F_{n-1}] \geq Y_{n-1} \quad \text{a.s} \quad (n \geq 1)$$

A supermartingale "decreases on average"; a submartingale "increases on average".

Some Remarks

* A process Y for which $Y_0 \in L^1(\Omega, F_0, P)$ is a martingale (sub, super) iff the process $Y - Y_0 = (Y_n - Y_0 : n \in \mathbb{Z}^+)$ has the same property. So we only need to consider processes which are null at 0.

* $E[Y_n | F_m] = E[Y_n | F_{n-1} | F_m] \leq E[Y_{n-1} | F_m] \leq \ldots$

$$\leq Y_m, \quad \text{a.s,}$$

if Y_n is a supermartingale.

Some Examples of martingales

Example 1 (Sums of independent zero-mean RVs)

$$\boxed{S_n \equiv Y_1 + Y_2 + \dots + Y_n}$$

with

$Y_1, Y_2, \dots$ a sequence of independent RVs with $E[|Y_k|] < \infty \ \forall k$, and $E(Y_k) = 0, \ \forall k$.

Prove S_n is a martingale w.r.t $F_n \equiv \sigma(Y_1, Y_2, \dots Y_n)$

Example 2 (Products of non-negative independent RV's of mean 1)

$$\boxed{M_n = Y_1 Y_2 \dots Y_n}$$

show that M_n is a martingale w.r.t $F_n \equiv \sigma(Y_1, Y_2, \dots Y_n)$

Example 3: (Accumulating data about a random variable)

Let $\{F_n\}$ be our filtration, and let $\xi \in L^1(\Omega, F, P)$.

Define $M_n \equiv E(\xi | F_n)$

Show that M is a martingale

Example 4 - Define $\sigma_n^2 = \sigma_0^2 \prod_1^n (\alpha + \beta Z_s^2)$; $\alpha + \beta = 1$, $\alpha, \beta > 0$

Show that $\{\sigma_n^2, F_n\}$ is a martingale

with $F_n \equiv \sigma(\sigma_0^2, Z_1, Z_2 \dots Z_n)$, and

$\{Z_s\}_1^\infty \equiv$ iid $N(0,1)$ independent of σ_0^2

Some Inequalities for martingales (MG)

Theorem (Doob, maximal inequality for MG)

If $(Y_n, \mathcal{F}_n) \equiv$ sub MG. Then $\forall \lambda \in \mathbb{R}$ we have

$$\lambda P(\max_{i \le n} Y_i > \lambda) \le E[Y_n 1(\max_{i \le n} Y_i > \lambda)]$$

indicator function

Proof: Define $F = \{\max_{i \le n} Y_i > \lambda\} = \bigcup_{i=1}^{n} \{Y_i > \lambda, \max_{j<i} Y_j < \lambda\}$

$$= \bigcup_{i=1}^{n} F_i$$

where the F_i are disjoint and $F_i \in \mathcal{F}_i$. Thus

$$\lambda P(F) = \sum_{i=1}^{n} \lambda P(F_i) = \sum_{i=1}^{n} \lambda P(F_i) = \sum_{i=1}^{n} \lambda E(1_{F_i})$$

$$\le \sum_{i=1}^{n} E(Y_i 1_{F_i}) \le$$

(as $Y_i > \lambda$ on F_i)

$$\le \sum_{i=1}^{n} E(E(Y_n | \mathcal{F}_i) 1_{F_i})$$

(as $E(Y_n | \mathcal{F}_i) \ge Y_i$ by sub MG property)

$$= \sum_{i=1}^{n} E(E(Y_n 1_{F_i} | \mathcal{F}_i))$$

$$= \sum_{i=1}^{n} E(Y_n 1_{F_i})$$

$$= E(Y_n (\textstyle\sum_1^n 1_{F_i}))$$

$$= E(Y_n 1_F)$$

giving

$$\lambda P(F) \le E(Y_n 1_F) \quad \text{Q.E.D}$$

<u>Corollary</u>: If $(Y_n, F_n) \equiv MG$, $p \geq 1$, $\lambda > 0$.

Then
$$\lambda^p P\left(\max_{i \leq n} |Y_i| > \lambda\right) < E|Y_n|^p$$

[This is obtained applying the Thm to the submartingale $|Y_i|$]

<u>Remark</u>: Set $p=2$ in this corollary and we get the Kolmogorov's inequality for iid r.v.s

Recall: If $X_1, X_2, \ldots$, indep $E(X_j)=0$,
$E(X_j^2) < \infty$

$$S_k = X_1 + X_2 + \cdots + X_k$$

Then

(1) $$P\left(\max_{1 \leq k \leq n} |S_k| \geq \alpha\right) \leq \frac{var(S_n)}{\alpha^2}$$

and then we also have Tchebychef's inequality

(2) $$P(|S_n| > \alpha) \leq \frac{var(S_n)}{\alpha^2}$$

(1) is sharper than (2) because

$$|S_n| > \alpha \Rightarrow \max_{k \leq n} |S_k| > \alpha \text{ so } P(|S_n| > \alpha) \leq P\left(\max_{k \leq n} |S_k| > \alpha\right)$$

Theorem (Doob, MG convergence Theorem)

If $Y_n \equiv$ MG with natural filtration and

$$(*) \quad \sup_n E|Y_n| < \infty$$

Then $\exists$ r.v Y_∞ with $E|Y| < \infty$ s.t

$$Y_n \longrightarrow Y_\infty \quad a.s$$

Let $\xi_1, \xi_2, \ldots$ be a sequence of independent random variable with $P(\xi_i = 0) = P(\xi_i = 2) = \frac{1}{2}$.

Then $Y = (Y_n, F_n^\xi)$, with $Y_n = \prod_{i=1}^n \xi_i$ and $F_n^\xi = \sigma(w: \xi_1, \ldots \xi_n)$ is a MG with $EY_n = 1$ and $Y_n \to Y_\infty = 0$ (a.s). At the same time is clear that $E|Y_n - Y_\infty| = 1$ and therefore

$Y_n \not\xrightarrow{L^1} Y_\infty$. So condition $(*)$ doesn't in general guarantee the convergence of X_n to X_∞ in L^1 sense. In order to get convergence in L^1 sense we need to strength $(*)$ to uniform integrability.

<u>Def</u>: A family $\{\xi_n\}_{n \geq 1}$ of random variables is said to be uniformly integrable if

$$\sup_n E[|\xi_n| \mathbb{1}_{(|\xi_n| > c)}] \longrightarrow 0 \quad c \to \infty.$$

It is clear that if ξ_n, $n \geq 1$, satisfy $|\xi_n| \leq \eta$, $E\eta < \infty$, then the family $\{\xi_n\}_{n \geq 1}$ is U. integrable

Theorem: Let $Y = \{Y_n, \mathcal{F}_n\}$ be a uniformly integrable MG (that is, the family $\{Y_n\}$ is uniformly integrable). Then there is a random variable Y_∞ with $E|Y_\infty| < \infty$, such that as $n \to \infty$

$$Y_n \to Y_\infty \quad (a.s)$$
$$Y_n \xrightarrow{L_1} Y_\infty.$$

Theorem (MG convergence theorem for L_2 MG)

If $Y_n \equiv$ MG with the natural filtration and

$$\sup_n E Y_n^2 < \infty$$

Then $\exists$ r.v Y_∞ with $EY^2 < \infty$ s.t

$$(*)\ Y_n \longrightarrow Y_\infty \quad a.s$$
$$Y_n \xrightarrow{L^2} Y_\infty$$

Proof of $(*)$

By the Cauchy criterion Y_n cgt a.s iff $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\left(|Y_{n+k} - Y_n| > \varepsilon \text{ for some } k \geq 1\right) = 0$$

i.e.

$$\lim_{n\to\infty} \lim_{m\to\infty} P\left(|Y_{n+k} - Y_n| > \varepsilon \text{ for some } k \quad 1 \leq k \leq m\right) = 0$$

Clearly this condition holds iff

$$(1) \quad \lim_{n\to\infty} \lim_{m\to\infty} P\left(\max_{1\leq k\leq m} |Y_{m+k} - Y_n| > \varepsilon\right) = 0$$

Now note that

$$Y_{n+k} - Y_n = \sum_{n+1}^{n+k} u_s \quad \left(\begin{array}{c} \text{with } u_s \text{ a m.d.s} \\ E[u_s \mid F_{s-1}] = 0 \end{array} \right)$$

is a MG so that by the maximal inequality

$$(2) \quad P\left(\max_{1 \le k \le m} |Y_{n+k} - Y_n| > \varepsilon \right) < E(Y_{n+m} - Y_n)^2 / \varepsilon^2$$

But

$$E(Y_{n+m} - Y_n)^2 = E(Y_{n+m}^2) - 2E(Y_{m+n} Y_n) + E(Y_n^2)$$

$$= E(Y_{n+m}^2) - 2E(Y_n E(Y_{m+n} \mid F_n)) + E(Y_n^2)$$

$$= E(Y_{n+m}^2) - E(Y_n^2)$$

So

$$E(Y_{n+m}^2) = E(Y_n^2) + E(Y_{n+m} - Y_n)^2$$

Thus, $\{E(Y_n^2)\}_{n=1}^{\infty}$ is an increasing sequence which is bounded above by assumption (i.e $\sup E(Y_n^2) < \infty$). Hence $E(Y_n^2)$ is cgt and we must have

$$E(Y_{n+m} - Y_n)^2 \longrightarrow 0 \quad \text{as } n, m \to \infty$$

as $E(Y_{n+m} - Y_n)^2$ is the difference between two sequences which have the same limit

(1) now follows by bounding it by the maximal inequality (2)

This proves the result that $\{Y_n\}$ cgt a.s. The limit r.v Y must satisfy $EY^2 < \infty$ since $\sup_n E(Y_n^2) < \infty$.

Applications of the MG convergence Theorem

Example 1 : Explosive AR(1) :

$X_t = \theta X_{t-1} + u_t \quad , \quad |\theta| > 1$

$(u_t, F_t) \equiv mds(0, \sigma^2) , \quad F_t = \sigma(u_1, \dots, u_t)$

(i) $\left(\frac{1}{\theta^t} X_t , F_t\right) \equiv MG$

$$\frac{1}{\theta^t} X_t = \frac{1}{\theta^t} u_t + \frac{1}{\theta^{t-1}} u_{t-1} + \dots + \frac{1}{\theta} u_1$$

$$= \frac{1}{\theta^t} \sum_{j=0}^{t-1} \theta^j u_{t-j}$$

and

$$E\left(\frac{1}{\theta^t} X_t \,\middle|\, F_{t-1}\right) = \frac{1}{\theta^{t-1}} u_{t-1} + \dots + \frac{1}{\theta} u_1 =$$
$$= \frac{1}{\theta^{t-1}} X_{t-1}$$

(ii) $$\sup_t E\left(\frac{1}{\theta^t} X_t\right)^2 = \sup_t \left(\frac{1}{\theta^{2t}} \sum_{j=0}^{t-1} \theta^{2j} \sigma^2\right)$$
$$= \sup_t \left[\sigma^2 \frac{1}{\theta^{2t}} \left(\frac{\theta^{2t}-1}{\theta^2-1}\right)\right] < \infty$$

(iii) Hence by the MG convergence theorem

$$\frac{1}{\theta^t} X_t \xrightarrow[a.s]{} Y(\theta)$$ some r.v. depends on θ

Note If $\{u_t\} \equiv iid\ N(0, \sigma^2)$ then X_t is normal for all t and we deduce $\frac{1}{\theta^t} X_t \xrightarrow[a.s]{} Y(\theta) \equiv N\left(0, \frac{\sigma^2}{\theta^2-1}\right)$

Note: What we have obtained is

$$\sum_1^\infty \frac{1}{\theta^t} u_t \text{ converges a.s}$$

Example 2 (conditional variances)

$$X_n = \sigma_n Z_n \quad n \in N$$

Z_n iid $N(0,1)$

$$\sigma_n^2 = \alpha + \beta X_{n-1}^2 \qquad \alpha, \beta > 0 \quad \alpha+\beta = 1$$

solving it $\quad \sigma_n^2 = \sigma_0^2 \prod_1^n (\alpha + \beta Z_s^2)$

$E(\sigma_n^2 \mid F_{n-1}) = \sigma_{n-1}^2 \qquad$ since $E(\alpha + \beta Z_n^2) = \alpha + \beta = 1$

So $\{\sigma_n^2\} \equiv MG$

Now $E(\sigma_n^2) = E(\sigma_0^2) \ \forall n$ so that

$$\sup E(\sigma_n^2) = E(\sigma_0^2) < \infty$$

Hence

$$\{\sigma_n^2\} \text{ cgt a.s}$$

But

$$\begin{aligned}\sigma_n^2 - \sigma_{n-1}^2 &= \sigma_{n-1}^2 \{\alpha + \beta Z_n^2 - 1\} \\ &= \sigma_{n-1}^2 \{\alpha + \beta Z_n^2 - (\alpha+\beta)\} \\ &= \beta \sigma_{n-1}^2 Z_n^2\end{aligned}$$

Since $\{\sigma_n^2\}$ converges a.s we have

$$\sigma_n^2 - \sigma_{n-1}^2 = \beta \sigma_{n-1}^2 Z_n^2 \longrightarrow 0 \quad \text{a.s}$$

If $\beta \neq 0$ this ensures that

$$\sigma_n^2 \xrightarrow[a.s]{} 0$$

Thus $\{\sigma_n^2\}$ converges a.s. to zero

Example 3 : (SLLN for mds)

$\{\varepsilon_t\} \equiv \text{mds}\ (0, \sigma_t^2) \qquad \sigma_t^2 = E(\varepsilon_t^2)$

$\sup_t \sigma_t^2 = K < \infty$

Consider

$$S_n = \Sigma_1^n \varepsilon_t / t \ , \quad \{S_n\} \equiv MG$$

$$E(S_n^2) = \Sigma_1^n E(\varepsilon_t^2)/t^2 = \Sigma_1^n \frac{\sigma_t^2}{t^2}$$

$$\sup E(S_n^2) \leq \left(\Sigma_1^\infty \frac{1}{t^2}\right)\left(\sup_t \sigma_t^2\right) = \frac{\pi^2}{6} K$$

Hence

(*) $\Sigma_1^n \frac{\varepsilon_t}{t}$ is convergent a.s (by the MG C Theorem)

Lemma (Kronecker Lema, see proof in Williams' book page 117)

Let x_n be a seq. of real nos with $\Sigma_1^\infty x_n$ cgt. Let $\{b_n\}$ be a monotone sequence (i.e $b_n \nearrow \infty$; e.g. $b_n = n$) increasing to ∞.

Then $\frac{1}{b_n} \Sigma_1^n b_i x_i \longrightarrow 0$

Take $b_n = n$ in (*) and since (*) is cgt a.s we have

$$\frac{1}{n} \Sigma_1^n t \frac{\varepsilon_t}{t} = \boxed{\frac{1}{n} \Sigma_1^n \varepsilon_t \xrightarrow[a.s]{} 0}$$

[SLLN for mds $(0, \sigma_t^2)$ with $K = \sup \sigma_t^2 < \infty$]

Remarks

1) The Kronecker Lemma used above is very important in SLLN proofs. It shows how to convert a convergent series into a sequence that $\xrightarrow[a.s]{} 0$. Since the MGCT establishes a.s. cgt series so easily it is ideally suited for setting up an application of the Kronecker lemma to establish a SLLN

2) Example 3 is a special case of the following:

Theorem

If (i) $\{Y_n\}$ is a seq. of L_2 r.v.'s

(ii) $\sum_1^\infty E(Y_n^2) < \infty$

Then
$$\sum_1^\infty (Y_n - E(Y_n / \mathcal{F}_{n-1})) \text{ is cgt a.s.}$$

where $\mathcal{F}_n = \sigma(Y_1, .. Y_n)$

Proof:
Set $Y_i^* = Y_i - E(Y_i / \mathcal{F}_{i-1})$, $S_n = \sum_1^n Y_i^*$ and note that S_n is MG since $Y_i^* \equiv$ mds by construction.

Now
$$E(S_n^2) = \sum_1^n E(Y_i^*)^2 \quad \text{(mds is orthogonal)}$$
and
$$E(Y_i^{*2}) = E(Y_i^2) - E(E(Y_i / \mathcal{F}_{i-1}))^2$$
$$\leq E(Y_i^2)$$
moreover
$$\sup E(S_n^2) \leq \sup_n \sum_1^n E(Y_i^2) = \sum_1^\infty E(Y_i^2) < \infty$$
by hypothesis. The result follows by the MGCT.

Corollary (SLLN for mds)

Let $\{X_n\}$ be a seq. of r.v.'s satisfying

$$(*) \quad \sum_{n=1}^{\infty} \frac{E(X_n^2)}{n^2} < \infty$$

then

$$\frac{1}{n} \sum_{j=1}^{n} \left(X_j - E(X_j \mid F_{j-1})\right) \xrightarrow{a.s} 0$$

Proof: Define $Y_n = X_n / n$. Then Y_n is in L_2, and

$$\sum_1^{\infty} E(Y_n^2) < \infty$$

Hence, by the theorem

$$\sum_1^{\infty} Y_n^* = \sum_1^{\infty} \left(\frac{X_n - E(X_n \mid F_{n-1})}{n}\right) \text{ is cgt a.s.}$$

Then by the Kronecker lemma

$$\frac{1}{n} \sum_1^n j \left(\frac{X_j - E(X_j \mid F_{j-1})}{j}\right) = \frac{1}{n} \sum_1^n \left(X_j - E(X_j \mid F_{j-1})\right) \xrightarrow[a.s]{} 0$$

Remark: The above corollary nearly gives us a general SLLN for sequences $\{X_n\}$ that satisfy $(*)$. We can often use alternative methods like mixing conditions to establish that $\frac{1}{n} \sum_1^n E(X_j / F_{j-1})$ is arbitrary small. Then the result give us a SLLN for weakly dependent sequences $\frac{1}{n} \sum_1^n X_j \xrightarrow[a.s]{} 0$

Doob's decomposition (Shiryayev's book pg 454)

Theorem: Let $Y = (Y_n, F_n)$ be a submartingale.

Then there are a martingale $m = (m_n, F_n)$ and a predictable increasing sequence $A = (A_n, F_{n-1})$ such that, for every $n \geq 0$, Doob's decomposition

$$Y_n = m_n + A_n \quad (P\text{-a.s})$$

holds. A decomposition of this kind is unique.

Proof: Let us put $m_0 = Y_0$, $A_0 = 0$ and

$$m_n = m_0 + \sum_{j=0}^{n-1} [Y_{j+1} - E(Y_{j+1} | F_j)]$$

$$A_n = \sum_{j=0}^{n-1} [E(Y_{j+1} | F_j) - Y_j]$$

It is evident that m and A defined in this way have the required properties.
Now you have to prove that this decomposition is unique (very easy).

Def: A predictable (Fn-1 measurable) increasing sequence $A = (A_n, F_{n-1})$ appearing in the Doob decomposition is called a **compensator** (of the submartingale Y)

The Doob decomposition plays a key role in the study of square integrable martingales

$M = (M_n, F_n)$ i.e martingale for which $E M_n^2 < \infty$, $n \geq 0$;

this depends on the observation that the stochastic sequence $M_n^2 = (M^2, F_n)$ is a submartingale.

From the previous theorem, there are a martingale

$$m = (m_n, F_n)$$

and a predictable increasing sequence

$$\langle M \rangle = (\langle M \rangle_n, F_{n-1}) \text{ s.t.}$$

$$M_n^2 = m_n + \langle M \rangle_n$$

$\langle M \rangle$ is called the QUADRATIC variation of M and in many aspects determines its structure and properties

$$\langle M \rangle_n = \sum_{j=1}^{n} E[(\Delta M_j)^2 | F_{j-1}]$$

and for all $l \leq k$

$$E[(M_k - M_l)^2 | F_l] = E[M_k^2 - M_l^2 | F_l] =$$

$$= E[\langle M \rangle_k - \langle M \rangle_l | F_l]$$

In particular if $M_0 = 0$ (P-a.s) then

$$E M_k^2 = E \langle M \rangle_k$$

For instance if $M_0 = 0$ and $M_n = \xi_1 + \dots + \xi_n$ where (ξ_n) is a sequence of independent r.v with $E\xi_1 = 0$ and $E\xi_1^2 < \infty$ the quadratic variation

$$\langle M \rangle_n = E M_n^2 = V(\xi_1) + \dots + V(\xi_n)$$

is not random and coincides with the variance

If $X = (X_n, F_n)$ and $Y = (Y_n, F_n)$ are square integrable martingales, we put

$$\langle X; Y \rangle_n = \frac{1}{4}\left[\langle X+Y \rangle_n - \langle X-Y \rangle_n\right]$$

It is easily verified that $(X_n Y_n - \langle X, Y \rangle_n, F_n)$ is a martingale and therefore, for $l \leq k$

$$E\left[(X_k - X_l)(Y_k - Y_l) \mid F_l\right] = E\left[\langle X, Y \rangle_k - \langle X, Y \rangle_l \mid F_l\right]$$

In the case $X_n = \xi_1 + \dots + \xi_n$ and $Y_n = \eta_1 + \dots + \eta_n$ where (ξ_n) and (η_n) are sequences of independent r.v with $E(\xi_i) = E(\eta_i) = 0$ and $E\xi_i^2 < \infty$, $E\eta_i^2 < \infty$, the variable $\langle X, Y \rangle_n$ is given by

$$\langle X, Y \rangle_n = \sum_{i=1}^{n} \mathrm{cov}(\xi_i, \eta_i)$$

The sequence $\langle X, Y \rangle = (\langle X, Y \rangle_n, F_{n-1})$ is often called the mutual variation of the martingales X and Y.

What about CLT for martingales?

To investigate the convergence in distribution of the sequence $\{S_n\}$, sum of random variables, is convenient to work with triangular arrays

$$\begin{bmatrix} \xi_{11} & & & \\ \xi_{21} & \xi_{22} & & \\ \xi_{31} & \xi_{32} & \xi_{33} & \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{n1} & \xi_{n2} & \xi_{n3} \cdots & \xi_{nn} \\ \vdots & & & \end{bmatrix}$$

We assume that for each n, $\xi_{n1}, \dots, \xi_{nn}$ are independent. We set $S_n = \xi_{n1} + \xi_{n2} + \dots + \xi_{nn}$

If we were interested in sequences $\{a_n^{-1}(S_n - b_n)\}$ where S_n is the sum of independent random variables $\xi_1, \dots, \xi_n$ we may construct an appropriate triangular array; take

$$\xi_{ni} = \frac{\xi_i}{a_n} - \frac{b_n}{n a_n}$$

then $$S_n = \sum_{i=1}^{n} \xi_{ni} = \frac{S_n - b_n}{a_n}$$

Thus the triangular array scheme includes the models you have previously studied

MG Arrays (MGA's)

This is a triangular sequence of the form

$$\{S_{ni}, F_{ni}\}_{i=1}^{k_n} \qquad k_n \nearrow \infty, \ n \nearrow \infty$$

which for each n (each row) the sequence is a MG.
Frequently $k_n = n$ and the array is triangular.

Example: Let $\{S_n, F_n\}_1^\infty \equiv$ MG with $E(S_n) = 0$

Set

$$k_n = n, \quad F_{ni} = F_i, \quad S_{ni} = \left(\frac{1}{\sigma_n}\right) S_i$$

where

$$\sigma_n^2 = var(S_n) = E(S_n^2)$$

Then

$$\left\{\{S_{ni}, F_{ni}\}_{i=1}^{n}\right\}_{n=1}^{\infty} = \left\{\left\{\frac{1}{\sigma_n} S_i, F_i\right\}_{i=1}^{n}\right\}_{n=1}^{\infty} \equiv MGA$$

MG difference arrays (MGDA's)

Let $X_{nj} = S_{nj} - S_{nj-1}$ in the above example.
Then for each n X_{nj} is a mds as $S_{nj} \equiv$ MG.

Then $\left\{\{X_{nj}, F_{nj}\}_{j=1}^{k_n}\right\}_{n=1}^{\infty} \equiv$ MGDA

MG conditional variances

$$v_{nj}^2 = E(X_{nj}^2 \mid F_{nj-1})$$ condit. variances of MGDA

$$V_{ni}^2 = \sum_{j=1}^{i} E(X_{nj}^2 \mid F_{nj-1})$$
$$= \sum_{j=1}^{i} v_{nj}^2$$
$$= \text{MG condit. variance}$$

Theorem (CLT for MGDA's)

If (a) $\sum_1^n v_{ni}^2 \xrightarrow[p]{} \sigma^2$

(b) $\sum_{i=1}^n E\left[X_{ni}^2 1(|X_{ni}|>\varepsilon) \mid F_{ni-1}\right] \xrightarrow{p} 0 \quad \forall \varepsilon > 0$

Then
$$\sum_{i=1}^n X_{ni} \Rightarrow N(0, \sigma^2)$$

Proof (see Hall & Heyde pg 58 Thm 3.2)

Remark 1: Condition (a) requires the convergence of the condit. variance of the MG. This is analogous to the convergence of

$$\lim_{n\to\infty} n^{-1} E(S_n^2)$$

i.e variance stability of the standardized sum $\left(\frac{1}{\sqrt{n}}\right) S_n$.

Remark 2: Condition (b) is a conditional form of "Lindberg" condition. Recall the conventional Lindberg condition

$$(*) \sum_{j=1}^{k_n} E\left(X_{nj}^2 \, 1\left(|X_{nj}| > \eta\right)\right) \longrightarrow 0$$

Now for iid sequences & setting $k_n = n$ we have

$$X_{nj} = \frac{1}{\sqrt{n}} X_j \quad , \quad X_j \equiv X \quad \forall_j$$

and then $(*)$ is

$$\sum_{j=1}^{n} \left(\frac{1}{\sqrt{n}}\right)^2 E\left(X^2 \, 1\left(|X| > \eta\sqrt{n}\right)\right) \longrightarrow 0$$

or simply

$$E\left(X^2 \, 1\left(|X| > \eta\sqrt{n}\right)\right) \longrightarrow 0$$

which holds if $E(X^2) < \infty$ (finite variance)

Remark 3 Note, of course that if the X_{nj} that appear in the conditional Lindberg condition (b) are i.n.i.d, the independence ensures that (b) is equivalent to $(*)$, the conventional Lindberg condition.

<u>Corollary</u> (CLT for MG's with stat & ergodic differences)

If

(i) $(S_n, F_n) \equiv$ MG with $E(S_n) = 0$

(ii) $X_n = S_n - S_{n-1}$ stat & ergodic

(iii) $E(X_n^2) < \infty$

Then $\frac{1}{\sqrt{n}} S_n = \frac{1}{\sqrt{n}} \sum_1^n X_t \Rightarrow N(0, E(X_1^2))$

<u>Proof</u> Set up to MGA

$$\{\{S_{ni}, F_{ni}\}_{i=1}^n\}_{n=1}^\infty = \{\{\frac{1}{\delta_n} S_i, F_i\}_{i=1}^n\}_{n=1}^\infty$$

where $\delta_n^2 = var(S_n) = E(S_n^2) = \sum_1^n E(X_t^2) = nE(X_1^2)$

Then, with $X_{ni} = (1/\delta_n) X_i$ we have

$$v_{ni}^2 = E(X_{ni}^2 | F_{n\,i-1}) = \frac{1}{\delta_n^2} E(X_i^2 | F_{i-1})$$

$$= \frac{1}{\delta_n^2} E(X_1^2 | F_0)$$

$$= \frac{1}{\delta_n^2} E(X_1^2)$$

$$= \frac{1}{n}$$

where $F_0 = (\phi, \Omega)$ is the trivial σ field.

Thus $\sum_1^n v_{ni}^2 = 1 \xrightarrow{P} 1$

satisfying (a) of the Theorem

Next observe that

$$\sum_{i=1}^{n} E\left[X_{ni}^2 1(|X_{ni}|>\varepsilon) \middle| F_{n,i-1}\right] = \sum_{i=1}^{n} \frac{1}{\delta_n^2} E\left[X_i^2 1(|X_i|>\varepsilon\delta_n) \middle| F_{i-1}\right]$$

$$= \frac{1}{n E(X_1^2)} \, n \, E\left[X_1^2 1(|X_1| > \varepsilon\sqrt{n}(E(X_1^2))^{\frac{1}{2}}\right]$$

$$\longrightarrow 0 \qquad \forall \varepsilon > 0$$

satisfying (b) of the CLT for MGDA's

Remark:

(i) This Corollary provides us with a simple version of the "Lindeberg Levy" Theorem for mds's rather than iid sequences. It forms the backbone of the general MG approach to CLT's. It was originally proved by Billingsley (1961) "The Lindeberg-Levy theorem for martingales" Proc. Amer. Math. Soc. 12

(ii) Note that the iid assumption of Lindberg-Levy is replaced by the requirement

X_n = mds & be stat & ergodic

The finite variance condition of Lindberg-Levy is retained $(E(X_1^2)<\infty)$