Chapter 1

1.1 Martingales

p.s. for a probability space.
r.v. for a random variable.

Definition 1.1

For a p.s. $(\Omega, \mathcal{F}_0, P)$ a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a r.v. $X \in \mathcal{F}_0$ with $E|X| < \infty$. We define the conditional expectation of X given $\mathcal{F}$, $E(X|\mathcal{F})$ to be any r.v. Y that has
a. $Y \in \mathcal{F}$.
b. $\int_A XdP = \int_A YdP$ for all $A \in \mathcal{F}$. and Y is said to be a $version$ of $E(X|\mathcal{F})$.

♣

Lemma 1.1

If Y satisfies (a),(b) above, then it is integrable.

♡

Proof

We know

$$\int_{\{Y>0\}} YdP = \int_{\{Y>0\}} XdP < \infty \int_{\{Y<0\}} YdP = \int_{\{Y<0\}} XdP < \infty$$

and hence $\int |Y|dP$ finite.

Lemma 1.2

If Y' also satisfies (a),(b) in Def.1.1., then $Y = Y'$ a.s.

♡

Proof

Assume $E_n = \{Y' - Y > n^{-1}\}, F_n = \{Y - Y' > n^{-1}\}, n \in \mathbb{N}$, then we know

$$n^{-1}P(E_n) \leq \int_{E_n} (Y - Y')dP = \int_{E_n} YdP - \int_{E_n} Y'dP = 0$$

and hence $P(E_n) = 0$ for any $n \in \mathbb{N}$, similarly, we know $P(F_n) = 0$ for any $n \in \mathbb{N}$, therefore, $Y = Y'$ a.s.

Theorem 1.1

If $X_1 = X_2$ on $B \in \mathcal{F}$ then $E(X_1|\mathcal{F}) = E(X_2|\mathcal{F})$ a.s. on B.

♡

Proof

For any $E \subset B$, we have

$$0 = \int_{\{E(X_1|\mathcal{F}) - E(X_2|\mathcal{F}) > n^{-1}\} \cap E} (X_1 - X_2)dP \geq n^{-1}P(\{E(X_1|\mathcal{F}) - E(X_2|\mathcal{F}) > n^{-1}\} \cap E)$$

and the rest is similar.

Theorem 1.2

$E(X|\mathcal{F})$ exists.

♡

Proof

Define $\nu(E) = \int_E XdP$ for $E \in \mathcal{F}$ and we know $\nu \ll P$ and hence there exists $Y \in \mathcal{F}$ such that

$$\int_E YdP = \nu(E) = \int_E XdP$$

for all $E \in \mathcal{F}$ by Radon-Nikodym's Theorem.

Example 1.1 a. If $X \in \mathcal{F}$, then $E(X|\mathcal{F}) = X$.

b. If X is independent to $\mathcal{F}$, i.e. $P(\{X \in B\} \cap A) = P(X \in B)P(A)$, then X is independent to χ_A for any $A \in \mathcal{F}$ and hence $E(X|\mathcal{F}) = EX$.

c. Suppose $\Omega_1, \Omega_2, \cdots$ is a finite or infinite partition of Ω into disjoint sets, with $P(\Omega_i) > 0, i \geq 1$ and then let $\mathcal{F} = \sigma(\Omega_1, \Omega_2, \cdots)$ and then

$$E(X|\mathcal{F}) = \frac{E(X;\Omega_i)}{P(\Omega_i)} \quad \text{on } \Omega_i$$

d. Suppose X, Y have joint density $f(x, y)$ i.e.,

$$P((X,Y) \in B) = \int_B f(x,y)dxdy \quad \text{for } B \in \mathcal{R}^2$$

then if $E|g(X)| < \infty$, then $E(g(X)|Y) = h(Y)$, where

$$h(y) = \int g(x)f(x,y)dx / \int f(x,y)dx$$

on $\{(x,y), \int f(x,y)dx > 0\}$, and hence a.s.

e. Suppose X and Y are independent, let ϕ be a function with $E|\phi(X,Y)| < \infty$ and let $g(x) = E(\phi(x,Y))$, then $E(\phi(X,Y)|X) = g(X)$.

Proof

c. By the $\pi - \lambda$ theorem, it suffices to show that

$$\int_A XdP = \int_A YdP$$

for any $A \in \{\bigcup_{1\leq i \leq n} \Omega_i\}$ where Y was defined as above.

d. Firstly, we recall any simple function $\phi \geq 0$ will cause $\int \phi(x,y)dy$ is measurable since $\int \phi(x,y)dy = \nu(E_y)$ when $\phi = \chi_E$ and then we know for any $g \geq 0$, $\int g(x)f(x,y)dy$ is measurable and then $\int g(x)f(x,y)dy$ is measurable for general g, then we will know $h(Y) \in \sigma(Y)$.

Consider $A \in \sigma(Y)$, where $A = \{Y \in B\}$, then

$$E(h(Y);A) = \int_{Y \in B} h(y)f(x,y)dxdy = \int_B \int h(y)f(x,y)dxdy = \int_B \int g(x)f(x,y)dxdy = E(g(X);A)$$

and the conclusion goes.

e. We know $g(X) \in \sigma(X)$ and then for any $A = \{X \in B\}$, we will know

$$E(g(X);A) = \int_B g(x)dx = \int_B \int \phi(x,y)dydx = E(\phi(X,Y);A)$$

and hence $E(\phi(X,Y)|X) = g(X)$.

Definition 1.2

Denote

$$P(A|\mathcal{G}) = E(1_A|\mathcal{G})$$
$$P(A|B) = P(A \cap B)/P(B)$$

and $E(X|Y) = E(X|\sigma(Y))$.

♣

Theorem 1.3

For the first two parts, we assume $E|X|, E|Y| < \infty$.

(a) $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.

(b) If $X \leq Y$ *then* $E(X|\mathcal{F}) \leq E(Y|\mathcal{F})$.

(c) If $X_n \geq 0$ *and* $X_n \uparrow X$ *with* $EX < \infty$ *then* $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$.

♡

Theorem 1.4

If ϕ *is convew and* $E|X|, E|\phi(X)| < \infty$ *then*

$$\phi(E(X|\mathcal{F})) \leq E(\phi(X)|\mathcal{F})$$

♡

Proof

Let $S = \{(a, b) : a, b \in \mathbb{Q}, ax + b \leq \phi(x) \text{ for all } x\}$, then $\phi(x) = \sup\{ax + b : (a, b) \in S\}$. And we know

$$E(\phi(X)|\mathcal{F}) \geq aE(X|\mathcal{F}) + b$$

and hence $E(\phi(X)|\mathcal{F}) \geq \phi(E(X|\mathcal{F}))$.

Theorem 1.5

Conditional expectation is a contraction in L^p, $p \geq 1$.

Proof

By Theorem 1.5., we have $|E|(X|\mathcal{F})|^p \leq E(|X|^p|\mathcal{F})$, then we know

$$E(|E(X|\mathcal{F})|^p) \leq E(E(|X|^p|\mathcal{F})) = E|X|^p$$

Theorem 1.6

If $\mathcal{F} \subset \mathcal{G}$ and $E(X|\mathcal{G}) \in \mathcal{F}$, then $E(X|\mathcal{F}) = E(X|\mathcal{G})$.

Theorem 1.7

If $\mathcal{F}_1 \subset \mathcal{F}_2$ then
(i) $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_1)$
(ii) $E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1)$

Proof

For $A \in \mathcal{F}_1$, we know

$$\int_A E(E(X|\mathcal{F}_1)|\mathcal{F}_2)dP = \int_A E(X|\mathcal{F}_1)dP = \int_A XdP$$
$$\int_A E(E(X|\mathcal{F}_2)|\mathcal{F}_1)dP = \int_A E(X|\mathcal{F}_2)dP = \int_A XDP$$

therefore, the equalities go.

Theorem 1.8

If $X \in \mathcal{F}$ and $E|Y|, E|XY| < \infty$ then

$$E(XY|\mathcal{F}) = XE(Y|\mathcal{F})$$

Proof

For any $X, Y \geq 0$, assume $\phi_n \uparrow X$ simple, then we know $\phi_n Y \uparrow XY$ and then

$$\int_A E(\chi_B Y|\mathcal{F}) = \int_A \chi_B Y dP = \int_{AB} Y dP = \int_{AB} E(Y|\mathcal{F})dP = \int_A \chi_B E(Y|\mathcal{F})$$

for any $A, B \in \mathcal{F}$ and hence $E(\chi_B Y|\mathcal{F}) = \chi_B E(Y|\mathcal{F})$ for any $B \in \mathcal{F}$, therefore, we know $E(\phi_n Y|\mathcal{F}) = \phi_n E(Y|\mathcal{F})$. By theorem 1.3 we know $E(\phi_n T|\mathcal{F}) \uparrow E(XY|\mathcal{F})$ and hence $E(XY|\mathcal{F}) = XE(Y|\mathcal{F})$, so for any $X \in \mathcal{F}, E|Y| < \infty, E|XY| < \infty$, we can consider the positive and negative parts and the conclusion goes.

Theorem 1.9

Suppose $EX^2 < \infty$. $E(X|\mathcal{F})$ is the variable $Y \in \mathcal{F}$ that minimizes the "mean square error" $E(X - Y)^2$.

Proof

If $Z \in L^2(\mathcal{F})$, then

$$ZE(X|\mathcal{F}) = E(ZX|\mathcal{F})$$

then we know

$$E(ZE(X|\mathcal{F})) = E(E(ZX|\mathcal{F})) = E(ZX)$$

and hence $E(Z(X - E(X|\mathcal{F}))) = 0$ for any $Z \in L^2(\mathcal{F})$.

If $Z = E(X|\mathcal{F}) - Y$, then

$$E(X-Y)^2 = E(X - E(X|\mathcal{F}) + Z)^2 = E(X - E(X|\mathcal{F}))^2 + EZ^2$$

and hence $E(X-Y)^2$ are minimal when $Y = E(X|\mathcal{F})$.

Definition 1.3

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $\mathcal{G}$ a σ-algebra contained by $\mathcal{F}$. $\mu : \Omega \times \mathcal{S} \to [0,1]$ is said to be a regular conditional distribution for X given $\mathcal{G}$ if

a. For each A, $\omega \to \mu(\omega, A)$ is a version of $P(X \in A|\mathcal{G})$.

b. For a.e. ω, $A \to \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$.

When $S = \Omega$ and X is the identity map, μ is called a regular condition probability. ♣

Proposition 1.1

Suppose X and Y have a joint density $f(x,y) > 0$. If

$$\mu(y, A) = \int_A f(x,y)dx \Big/ \int f(x,y)dx$$

then $\mu(Y(\omega), A)$ is a r.c.d for X given $\sigma(Y)$. ♠

Proof

Here we know $X : (\Omega . \mathcal{F}) \to (\mathbb{R}, \mathcal{R})$, so we should check:

a. $\mu(Y(\omega), A) = \int_A f(x, Y(\omega))dx / \int f(x, Y(\omega))dx$ is a version of $P(X \in A|Y)$.

b. For a.e. ω, $\mu_{Y(\omega)}(A) = \mu(Y(\omega), A)$ is a probability measure on $(\mathbb{R}, \mathcal{R})$.

To see the first claim, consider

$$\begin{aligned}
\int_{Y \in B} P(X \in A|Y)dP = \int_{Y \in B} \chi_{X \in A} dP &= \int_B \int_A f(x,y)dxdy \\
&= \int_A \int_B f(x,y)dydx \\
&= \int_B \int_A f(x,y)dxdy
\end{aligned}$$

$$= \int_B \int \int_A f(x,y)dx \Big/ \int f(x,y)dx f(x,y)dxdy = \int_{Y \in B} \mu(Y(\omega), A)dP$$

and the second claim is trivial.

Theorem 1.10

Let $\mu(\omega, A)$ be a r.c.d for X given $\mathcal{F}$. If $f : (S, \mathcal{S}) \to (\mathbb{R}, \mathcal{R})$ has $E|f(X)| < \infty$ then

$$E(f(x)|\mathcal{F}) = \int \mu(\omega, dx) f(x) \quad a.s.$$

♡

Proof

Consider $f = \chi_A$ for some A mrb in $\mathcal{R}$, then $\int \mu(\omega, dx)f(x) = \mu(\omega, A) = P(X \in A|\mathcal{F})$ and hence the equality holds for all simple functions, then the problem goes.

Here we skip some properties of regular conditional distribution and continue to martingale.

Definition 1.4

$\mathcal{F}_n$ is a filtration, i.e. an increasing sequence of σ-fields. A sequence X_n is said to be adapted to $\mathcal{F}_n$ if $X_n \in \mathcal{F}_n$ for all n. If X_n is sequence with

a. $E|X_n| < \infty$.

b. X_n is adapted to $\mathcal{F}_n$.

c. $E(X_{n+1}|\mathcal{F}_n) = X_n$ for all n then X is said to be a martingale (resp to $\mathcal{F}_n$). If we replace the equality into $\leq$ or $\geq$, then X is said to be a supermartingale or submartingale. ♣

Example 1.2 (Random walk)Let $\xi_1, \xi_2, \cdots$ be independent and id.d, $S_n = S_0 + \sum\limits_{i=1}^{n} \xi_i$ where S_0 is a constant. $\mathcal{F}_n = \sigma(\xi_1, \cdots, \xi_n)$ and take $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

a. If $\mu = E\xi_i = 0$ then $S_n, n \geq 0$ is a martingale with respect to $\mathcal{F}_n$.

b. $\mu = E\xi_i = 0$ and $\sigma^2 = \text{var}(\xi_i) < \infty$, then $S_n^2 - n\sigma^2$ is a martingale.

Proof

a. Notice $E|S_n| < \infty, n \geq 0$, for any $A \in \mathcal{F}_n$, then notice

$$E(S_{n+1}|\mathcal{F}_n) = E(\xi_{n+1}|\mathcal{F}_n) + S_n = E\xi_{n+1} + S_n = S_n$$

b. Notice that $E|S_n - n\sigma^2| < \infty$, and

$$E(S_{n+1}^2 - (n+1)\sigma^2|\mathcal{F}) = S_n^2 - (n+1)\sigma^2 + \sigma^2 = S_n^2 - n\sigma^2$$

Example 1.3 Let $Y_1, Y_2, \cdots$ be nonnegative i.i.d r.v.s with $EY_m = 1$. If $\mathcal{F}_n = \sigma(Y_1, \cdots, Y_n)$, then $M_n = \prod_{m \leq n} Y_m$ defines a martingale.

Then assume $\phi(\theta) = Ee^{\theta\xi_i}, Y_i = e^{\theta\xi_i}/\phi(\theta)$, then we know $M_n = e^{\theta S_n}/\phi(\theta)^n$.

Theorem 1.11

If X_n is a (super-/sub-)martingale then for $n > m$, $E(X_n|\mathcal{F}_m) \leq (\geq / =) X_m$. ♡

Proof Notice

$$E(X_{m+k}|\mathcal{F}_m) = E(E(X_{m+k}|\mathcal{F}_{m+k-1})|\mathcal{F}_m) \leq E(X_{m+k-1}|\mathcal{F}_m)$$

the rest proof is similar.

Theorem 1.12

If X_n is a martingale w.r.t. $\mathcal{F}_n$ and ϕ is a convex function with $E|\phi(X_n)| < \infty$ for all n then $\phi(X_n)$ is a sub-martingale w.r.t. $\mathcal{F}_n$. Consequently, if $p \geq 1$ and $E|X_n|^p < \infty$ for all n, then $|X_n|^p$ is a submartingale w.r.t. $\mathcal{F}_n$. ♡

Proof Notice

$$E(\phi(X_{n+1})|\mathcal{F}_n) \geq \phi(E(X_{n+1})|\mathcal{F}_n) = \phi(X_n)$$

and the problem goes.

Theorem 1.13

If X_n is a submartingale w.r.t. $\mathcal{F}_n$ and ϕ is an increasing convex function with $E|\phi(X_n)| < \infty$ for all n, then $\phi(X_n)$ is a submartingale w.r.t. $\mathcal{F}_n$. Consequently

a. If X_n is a submartingale then $(X_n - a)^+$ is a submartingale.

b. If X_n is a supermartingale then $\min(X_n, a)$ is a supermartingale. ♡

Proof Notice

$$E(\phi(X_{n+1})|\mathcal{F}_n) \geq \phi(E(X_{n+1})|\mathcal{F}_n) \geq \phi(X_n)$$

then (a) is easy to be checked and hence (b) is correct.

Definition 1.5

Let $\mathcal{F}_n, n \geq 0$ be a filtration. $H_n, n \geq 1$ is said to be a predictable sequence if $H_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$. ♣

Definition 1.6

We denote

$$(H \cdot X)_n = \sum_{m=1}^{n} H_m(X_m - X_{m-1})$$

♣

Theorem 1.14

Let $X_n, n \geq 0$ be a supermartingale. If $H_n \geq 0$ is predictable and each H_n is bounded then $(H \cdot X)_n$ is a supermartingale.

Proof Consider

$$E((H \cdot X)_{n+1}|\mathcal{F}_n) = E(\sum_{m=1}^{n+1} H_m(X_m - X_{m-1})|\mathcal{F}_n) = (H \cdot X)_n + E(X_{n+1}|\mathcal{F}_n) - X_n \leq (H \cdot X)_n$$

Definition 1.7

A r.v. N is said to be a stopping time if $\{N = n\}$ in $\mathcal{F}_n$ for all $n > \infty$.

Theorem 1.15

If N is a stoppiing time and X_n is a supermartingale, then $X_{N \wedge n}$ is a supermartingale.

Proof Consider

$$E(X_{N\wedge n+1}|\mathcal{F}_n) = E(X_{n+1}\chi_{N\geq n+1} + \sum_{k=0}^{n} X_k\chi_{N=k}|\mathcal{F}_n) \leq \chi_{N\geq n+1}X_n + \sum_{k=0}^{n} X_k\chi_{N=k} = X_{N\wedge n}$$

Definition 1.8

Suppose $X_n, n \geq 0$ is a submartingale. Let $a < b, N_0 = -1$ and for $k \geq 1$ let

$$N_{2k-1} = \inf\{m > N_{2k-2}, X_m \leq a\}$$
$$N_{2k} = \inf\{m > N_{2k-1}, X_m \geq b\}$$

The N_j are stopping times so

$$H_m = \begin{cases} 1 & \text{if } N_{2k-1} < m \leq N_{2k} \text{ for some } k \\ 0 & \text{otherwise} \end{cases}$$

defines a predictable sequence.

Proof

Notice

$$\{N_{2k-1} = n\} = \bigcup_{0\leq m\leq n-1} \{N_{2k-2} = m\} \cap (\bigcap_{n-1-m\geq k\geq 0} \{X_{m+k} > a\}) \cap \{X_n \leq a\}$$

and

$$\{N_{2k} = n\} = \bigcup_{0\leq m\leq n-1} \{N_{2k-1} = m\} \cap (\bigcap_{n-1-m\geq k\geq 0} \{X_{m+k} < b\}) \cap \{X_n \geq b\}$$

and hence N_{2k-1}, N_{2k} are stopping times by induction.

And notice

$$\{N_{2k-1} < m \leq N_{2k} \text{ for some } k\} = \bigcup_{k\geq 0}\{N_{2k-1} \leq m-1\} \cap \{N_{2k} \geq m\} \in \mathcal{F}_{m-1}$$

and hence H_m is predictable.

Theorem 1.16

(Upcoming inequality) If $X_m, m \geq 0$, is a submartingale then

$$(b-a)EU_n \leq E(X_n - a)^+ - E(X_0 - a)^+$$

where $U_n = \sup\{k, N_{2k} \leq n\}$.

Proof Here we assume $Y_m = a + (X_m - a)^+$ and we have

$$(b-a)U_n \leq (H \cdot Y)_n$$

let $K_m = 1 - H_m$ and then we know that $(K \cdot X)_n$ is a submartingale and then

$$E(K \cdot X)_n \geq E(K \cdot X)_0 = 0$$

so we know

$$E(H \cdot Y)_n \leq E(Y_n - Y_0) = E(X_n - a)^+ - E(X_0 - a)^+$$

since $Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n$

Theorem 1.17

(Martingale convergence theorem) If X_n is a submartingale with $\sup EX_n^+ < \infty$ then as $n \to \infty$, X_n converges a.s. to a limit X with $E|X| < \infty$. ♡

Proof We know $(X - a)^+ \leq X^+|a|$, then we know

$$EU_n \leq (|a| + EX_n^+)/(b - a)$$

so $\sup X_n^+$ will imply than $EU < \infty$ where $U = \lim U_n$ and hence for all rational a, b, we know

$$P(\{\liminf X_n < a < b < \limsup X_n\}) = 0$$

and hence $\lim X_n$ exists a.s. and $EX^+ \leq \liminf EX_n^+ < \ infty$ and hence $X < \infty$ a.s. and notice

$$EX_n^- = EX_n^+ - EX_n \leq EX_n^+ - EX_0$$

and hence $EX^- \leq \liminf EX_n^- \leq \liminf EX_n^+ - EX_0 < \infty$ therefore $E|X| < \infty$.

Theorem 1.18

If $X_n \geq 0$ is a supermartingale then as $n \to \infty$, $X_n \to X$ a.s. and $EX \leq EX_0$. ♡

Proof Let $Y_n = -X_n$ and hence a submartingale with $EY_n^+ = 0$, then we know $X_n \to X$ a.s. and we also have

$$EX \leq \liminf EX_n^+ \leq EX_0$$

Proposition 1.2

The theorem 1.18. provide a method to show that a.s. convergence does not guarantee convergence in L^1. ♠

Proof Let S_n be a symmetric simple random walk with $S_0 = 1$ and $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$, let $N = \inf\{n : S_n = 0\}$ and $X_n = S_{N \wedge n}$. Then we know X_n nonnegative and $EX_n = EX_0 = 1$ since X_n is a martingale, then we know $X_n \to X$ where X is some r.v. and hence $X = 0$, because there is no way to converge to others and hence X_n do not converge to X in L^1.

Proposition 1.3

Convergence in probability do not guarantee convergence a.s. ♠

Proof Let $X_0 = 0$ and $P(X_k = 1|X_{k-1} = 0) = P(X_k = -1|X_{k-1} = 0) = \frac{1}{2k}, P(X_k = 0|X_{k-1} = 0) = 1 - \frac{1}{k}$ and $P(X_k = kX_{k-1}|X_{k-1} \neq 0) = \frac{1}{k}, P(X_k = 0|X_{k-1} \neq 0) = 1 - \frac{1}{k}$, then we know $X_k \to 0$ in probability, but $P(X_k = 0, k \geq K)$ and it picks discrete values and hence X_k can not converge to 0 a.s.