Probability theory II Facts

Taeyoung Chang

Contents

1 Conditional Expectation

- Projection Thm for Hilbert Space

    - If E is a Hilbert space and $M \subset E$ is closed and convex, then for any $y \in E$,
    $\exists$ a unique $w \in M$ s.t. $\|y - w\| = d(y, M) := \inf\{\|y - v\| : v \in M\}$.
    Denote it as $w = proj_M y$ i.e. w is a projection of y onto M .

    - If E is a Hilbert space and $M \subset E$ is a closed vector subspace, then for any $y \in E$,

        i. $\exists$ a unique decomposition $y = w + v$ with $w = proj_M y \in M$ and $v \in M^{\perp}$
        ii. For $w \in M$, $w = proj_M y \Leftrightarrow \langle y - w, z \rangle = 0$ $\forall z \in M$

* $\mathcal{L}^2 := \{$ Random Variable $X : E(X^2) = \int X^2 \, dP < \infty\}$

$\surd$ If $X \in \mathcal{L}^2$ then $E|X| < \infty$ i.e. every element of $\mathcal{L}^2$ is integrable.

    ★ Trick : $|X| \le X^2 + \frac{1}{4}$

$\surd$ $\mathcal{L}^2$ is a vector space

    ★ Trick : inequality $(aX + bY)^2 \le 2(a^2 X^2 + b^2 Y^2)$

- $\mathcal{L}^2$ is a Hilbert space with inner product $\langle X, Y \rangle = E(XY)$

    ★ Trick : Cauchy seq. having a subseq. converging to a point converges to the point.

- Lemma for proving $\mathcal{L}^2$ is a complete normed space.

    - If $\{X_n\} \underset{seq}{\subset} \mathcal{L}^2$ and $\|X_n - X_{n+1}\| \le 2^{-n}$ $\forall n \in \mathbb{N}$ then $\exists X \in \mathcal{L}^2$ s.t. $X_n \to X$ $a.s.$ and
    $\|X_n - X\| \to 0$ i.e. $X_n \to X$ in $\mathcal{L}^2$.
        ★ Lemma : If a random varaible Z satisfies $Z \ge 0$ and $E(Z) < \infty$ then $Z < \infty$ $a.s.$

* For $X \in \mathcal{L}^2$, $\mathcal{L}^2(X) := \{g(X) \,|\, g : \mathbb{R} \to \mathbb{R}$ is a Borel function, $E[(g(X))^2] < \infty\}$

$\surd$ For $X \in \mathcal{L}^2$, $\mathcal{L}^2(X)$ is a vector subspace of $\mathcal{L}^2$.

- For $X \in \mathcal{L}^2$, $\mathcal{L}^2(X)$ is a closed vector subspace of $\mathcal{L}^2$ so that $\mathcal{L}^2(X)$ is also a Hilbert space.

* Geometric definition for conditional expectation

    - For $X, Y \in \mathcal{L}^2$, define $E[Y|X] = Proj_{\mathcal{L}^2(X)} Y$
    - $E[Y|X] = g(X)$ $a.s.$ for some Borel function g
    - $\|Y - E[Y|X]\| = \min_{h(X) \in \mathcal{L}^2(X)} \|Y - h(X)\|$
    i.e. $E[(Y - E[Y|X])^2] \le E[(Y - h(X))^2]$ $\forall h(X) \in \mathcal{L}^2$
    - For $g(X) \in \mathcal{L}^2(X)$, $g(X) = E[Y|X] \Leftrightarrow \langle Y - g(X), h(X) \rangle = 0$ $\forall h(X) \in \mathcal{L}^2$
    $\Leftrightarrow E[(Y - g(X))h(X)] = 0$ $\forall h(X) \in \mathcal{L}^2$

- Elementary properties of conditional expectation from geometric definition

    - If $X, Y, Z \in \mathcal{L}^2$ then the followings are true.
        i. $E[c|X] = c \, a.s.$ $\forall c \in \mathbb{R}$
        ii. $E[\alpha Y + \beta Z|X] = \alpha E[Y|X] + \beta E[Z|X]$ $\forall \alpha, \beta \in \mathbb{R}$
        iii. $E[Y|X] = E[Y]$ if X and Y are independent.

iv. $E[g(X)Y|X] = g(X)E[Y|X]$ if g satisfies $g(X) \in \mathcal{L}^2(X)$ and $\sup_x |g(x)| < \infty$

v. $E[E[Y|X]] = E[Y]$

$\sqrt{}$ In fact, the additional assumption about boundedness of g in (iv) is not necessary. We will see later.

- Extending the definition from $\mathcal{L}^2$ to all integrable functions

$$E[\{Y - E[Y|X]\}I(X \in A)] = 0 \quad \forall A \in \mathcal{B}(\mathbb{R}) \quad \because I(X \in A) \in \mathcal{L}^2(X)$$

$$\int_{(X \in A)} Y \, dP = \int_{(X \in A)} E[Y|X] \, dP \quad \forall A \in \mathcal{B}(\mathbb{R})$$

$$\int_B Y \, dP = \int_B E[Y|X] \, dP \quad \forall B \in \sigma(X)$$

  − $E[Y|X] \in \sigma(X)$ and $\int_B Y \, dP = \int_B E[Y|X] \, dP \quad \forall B \in \sigma(X)$. Such r.v. is unique in the sense that if any r.v. Z satisfies $Z \in \sigma(X)$ and $\int_B Y \, dP = \int_B Z \, dP \quad \forall B \in \sigma(X)$ then $Z = E[Y|X]$ a.s. provided $E|Y| < \infty$

  − From the theory on $\mathcal{L}^2$ space, we get geometric understanding about conditional expectation. But now, from the equation above, we can guess that definition for conditional expectation may be extended to all integrable random variables.

- Proof for the uniqueness mentioned above

  − $(\Omega, \mathcal{F}, P)$: a prob. space. $Y \in \mathcal{F}$ and $E|Y| < \infty$. $\mathcal{G} \subset \mathcal{F}$ is a sub σ-field. If X is a random variable satisfying (a) $X \in \mathcal{G}$ (b) $\int_A Y \, dP = \int_A X \, dP \quad \forall A \in \mathcal{G}$ then

    i. X is integrable

    ii. Such X is unique in the sense that if there is another X' then $X = X'$ a.s.
      ★ Trick : For any r.v. Z, $(Z > 0) = \bigcup_{\varepsilon > 0}(Z \geq \varepsilon) = \bigcup_{n \in \mathbb{N}}(Z > \frac{1}{n})$
      ★ Lemma : For any $\mathcal{F}$-measurable and integrable X and Y,
        if $\int_A X \, dP = \int_A Y \, dP \quad \forall A \in \mathcal{F}$ then $X = Y$ a.s.

- Radon-Nikodym Thm

  − If μ, ν are σ−finite measures on $(\Omega, \mathcal{F})$ and $\nu \ll \mu$ ($\mu(A) = 0 \Rightarrow \nu(A) = 0 \quad \forall A \in \mathcal{F}$) then $\exists$ a $\mathcal{F}$-measurable nonnegative function g s.t. $\nu(A) = \int_A g \, d\mu \quad \forall A \in \mathcal{F}$. The function g is unique in the sense that if h is another such function then $g = h \ \mu - a.e.$

* Definition of conditional expectation

  − $(\Omega, \mathcal{F}_0, P)$: a prob. space. $\mathcal{F} \subset \mathcal{F}_0$: a sub σ-field.
  X is a random variable s.t. $X \geq 0$, $X \in \mathcal{F}_0$ and $E|X| < \infty$. Then $\exists$ a unique r.v. Y s.t. $Y \geq 0$, $Y \in \mathcal{F}$ and $\int_A X \, dP = \int_A Y \, dP \quad \forall A \in \mathcal{F}$. Such Y is unique in the sense that if another Y' exists then $Y = Y'$ a.s.

  − $Y = E[X|\mathcal{F}]$ is said to be conditional expectation of X given $\mathcal{F}$

    ★ Applying Radon Nikodym thm to measures $P|_{\mathcal{F}}$ and Q on $(\Omega, \mathcal{F})$ where Q is defined by $Q(A) = \int_A X \, dP \quad \forall A \in \mathcal{F}$. Note that $Q \ll P|_{\mathcal{F}}$ and Q is a finite measure.

  − We can extend the definition to general integrable r.v. X
  $Y = E[X|\mathcal{F}]$ is a unique random varaible s.t. $Y \in \mathcal{F}$ and $\int_A X \, dP = \int_A Y \, dP \quad \forall A \in \mathcal{F}$. $E[X|\mathcal{F}]$ is also integrable and the uniqueness is in the sense of a.s. equivalence relation. $Y = E[X|\mathcal{F}]$ can be derived by $Y = Y_1 - Y_2$ where $Y_1 = E[X^+|\mathcal{F}]$ and $Y_2 = E[X^-|\mathcal{F}]$

* Conditional expectation given a random variable

  – X : integrable r.v. For a random variable Y, define $E[X|Y] := E[X|\sigma(Y)]$

  $\sqrt{}$ Y need not be integrable.

  $\sqrt{}$ Since $E[X|Y] \in \sigma(Y)$, $E[X|Y] = g(Y)$ for some Borel function g. This coincides with the definition of conditional expectation in $\mathcal{L}^2$ space.

* Conditional probability

  – For $A \in \mathcal{F}_0$ and a sub σ-field $\mathcal{F} \subset \mathcal{F}_0$, define $P(A|\mathcal{F}) := E[I_A|\mathcal{F}]$

  – For $A, B \in \mathcal{F}_0$, define $P(A|B) = P(A \cap B)/P(B)$

• Elementary properties of conditional expectation

  – $(\Omega, \mathcal{F}_0, P)$: a prob. space. $\mathcal{F} \subset \mathcal{F}_0$: a sub σ-field. X, Y : integrable random variables

    i. $E[c|\mathcal{F}] = c$

    ii. $E[\psi(X)|X] = \psi(X)$ given $E|\psi(X)| < \infty$

    iii. If $\mathcal{F}$ is a trivial σ-field i.e. $\mathcal{F} = \{\Omega, \phi\}$ then $E[X|\mathcal{F}] = E[X]$

    iv. $\Omega = \bigcup_{i=1}^{\infty} \Omega_i$ is a partition of Ω with $\Omega_i \in \mathcal{F}_0$ and $P(\Omega_i) > 0$ $\forall i \in \mathbb{N}$
    $\mathcal{F} = \sigma\{\Omega_1, \Omega_2, \cdots\} = \{\bigcup_{j \in \kappa} \Omega_j : \kappa \subset \mathbb{N}\}$ ($\mathcal{F}$ is a σ-field).   Then we have

    $$E[X|\mathcal{F}] = \sum_{i=1}^{\infty} a_i I_{\Omega_i} \quad with \quad a_i = \frac{E[X I_{\Omega_i}]}{P(\Omega_i)}$$

    $\sqrt{}$ For $A \in \mathcal{F}_0$, $P(A|\mathcal{F}) = P(A|\Omega_i) I_{\Omega_i}$

    ★ Lemma : If $Z \in \mathcal{F}$ for such $\mathcal{F}$, then we can write $Y = \sum_{i=1}^{\infty} c_i I_{\Omega_i}$ where $c_i \in \mathbb{R}$

    v. $E[aX + bY|\mathcal{F}] = aE[X|\mathcal{F}] + bE[Y|\mathcal{F}]$ $\forall a, b \in \mathbb{R}$

    vi. $X \geq 0 \Rightarrow E[X|\mathcal{F}] \geq 0$ $a.s.$

    ★ Lemma : If $Z > 0$ on A with $P(A) > 0$ then $\int_A Z \, dP > 0$

    vii. $X \leq Y \Rightarrow E[X|\mathcal{F}] \leq E[Y|\mathcal{F}]$ $a.s.$

    viii. $\left|E[X|\mathcal{F}]\right| \leq E\left[|X| \,\big|\, \mathcal{F}\right]$

    □ $|X| \leq M$ for some $M > 0$ $\Rightarrow$ $\left|E[X|\mathcal{F}]\right| \leq M$ $a.s.$

    ix. $E[|X| \,|\, \mathcal{F}] = 0$ $\Rightarrow$ $X = 0$ $a.s.$

    x. $E[E[X|\mathcal{F}]] = E[X]$

• X, Y : integrable r.v's where $X \perp\!\!\!\perp Y$. $\psi : \mathbb{R}^2 \to \mathbb{R}$ Borel measurable s.t. $E|\psi(X,Y)| < \infty$
Define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = E[\psi(x,Y)]$ $\forall x \in \mathbb{R}$.   Then $E[\psi(X,Y)|X] = g(X)$

  $\sqrt{}$ $g(x) = E[\psi(x,Y)] = \int \psi(x,Y) \, dP = \int_{\mathbb{R}} \psi(x,y) dPY^{-1}(y) = \int_{\mathbb{R}} \psi_x(y) \, d\mu_Y(y)$ $\forall x \in \mathbb{R}$
  By Fubini thm in real analysis course, it is shown that g is Borel measurable & integrable.

• Conditional expectation and convergence

  – $(\Omega, \mathcal{F}_0, P)$: a probability space. $\mathcal{F} \subset \mathcal{F}_0$: a sub σ-field

    i. (MCT) If $X_n \geq 0$ and $X_n \nearrow X$ $a.s.$ with $E(X) < \infty$ then $E[X_n|\mathcal{F}] \nearrow E[X|\mathcal{F}]$ $a.s.$

    □ If $Y_n \searrow Y$ $a.s.$ with $E|Y_1|$, $E|Y| < \infty$ then $E[Y_n|\mathcal{F}] \searrow E[Y|\mathcal{F}]$ $a.s.$

    ii. (DCT) If $|X_n| \leq Y$, $E(Y) < \infty$ and $X_n \to X$ $a.s.$ then $E[X_n|\mathcal{F}] \to E[X|\mathcal{F}]$ $a.s.$

iii. (Fatou's lemma) If $X_n \geq 0$ and $X_n \to X$ $a.s.$ with $E(X_n) < \infty$, $E(X) < \infty$ then
$E[X|\mathcal{F}] \leq \liminf E[X_n|\mathcal{F}]$

iv. (Continuity from below) $\{B_n\} \underset{seq}{\subseteq} \mathcal{F}_0$ s.t. $B_n \subset B_{n+1}$ $\forall\, n \in \mathbb{N}$. $B := \bigcup_n B_n$
Then $P(B_n|\mathcal{F}) \nearrow P(B|\mathcal{F})$

v. (Countable additivity) If $\{C_n\} \underset{seq}{\subseteq} \mathcal{F}_0$ is mutually disjoint then $P(\bigcup_n C_n|\mathcal{F}) = \sum_n P(C_n|\mathcal{F})$

- Essential inequalities

  i. (Markov) $\quad P(|X| \geq c\,|\mathcal{F}) \leq \frac{1}{c} E\big[|X|\big|\mathcal{F}\big] \quad \forall\, c > 0$

  ii. (Jensen) If $\phi : \mathbb{R} \to \mathbb{R}$ is convex then $\phi(E[X|\mathcal{F}]) \leq E[\phi(X)|\mathcal{F}]$ $a.s.$

  ★ Trick : For each $x \in \mathbb{R}$ and convex function $\phi : \mathbb{R} \to \mathbb{R}$, we have
  $\phi(x) = \sup\{ax + b : (a,b) \in S\}$ where $S = \{(a,b) \in \mathbb{R}^2 : ax + b \leq \phi(x) \,\forall\, x \in \mathbb{R}\}$

  iii. (Cauchy-Schwarz) For $X, Y \in \mathcal{L}^2$, we have $E^2[XY|\mathcal{F}] \leq E[X^2|\mathcal{F}]E[Y^2|\mathcal{F}]$ $a.s.$

- Smoothing property of conditional expectation

  i. If $X \in \mathcal{F}$, $E|Y| < \infty$, and $E|XY| < \infty$ then $E[XY|\mathcal{F}] = X E[Y|\mathcal{F}]$ $a.s.$

  $\surd$ $E|X| < \infty$ assumption is not required.

  □ If $X \in \mathcal{F}$ and $E|X| < \infty$ then $E[X|\mathcal{F}] = X$ $a.s.$

  ii. If $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_0$ are sub σ-fields and $E|X| < \infty$ then

  (a) $E[E(X|\mathcal{F}_1)|\mathcal{F}_2] = E[X|\mathcal{F}_1]$
  (b) $E[E(X|\mathcal{F}_2)|\mathcal{F}_1] = E[X|\mathcal{F}_1]$

  ★ Lemma : If $\mathcal{F}_1 \subset \mathcal{F}_2$ then $Y \in \mathcal{F}_1 \Rightarrow Y \in \mathcal{F}_2$

  $\surd$ In short, "the smaller wins". In view of information, it is similar to projection onto
  vector subspaces $S_1 \subset S_2 \subset S$ where $Proj_{S_1} Proj_{S_2} = Proj_{S_2} Proj_{S_1} = Proj_{S_1}$

- Def. of conditional expectation by Radon-Nikodym derivative agrees with def. in $\mathcal{L}^2$ space.

  – If $E(X^2) < \infty$ then for $\mathcal{C} = \{Y : Y \in \mathcal{F}, E(Y^2) < \infty\}$,
  $E[\{X - E[X|\mathcal{F}]\}^2] = \inf_{Y \in \mathcal{C}} E[\{X - Y\}^2]$ and $E[X|\mathcal{F}] = \arg\min_{Y \in \mathcal{C}} E[\{X - Y\}^2]$

  ★ Lemma : If $X \in \mathcal{L}^2$ then $E[X|\mathcal{F}] \in \mathcal{L}^2$

* Independence of a random variable and a σ-field

  – A random variable X and a σ-field $\mathcal{F}$ are said to be independent if $\sigma(X)$ and $\mathcal{F}$ are
  independent

- If an integrable random variable X and a σ-field $\mathcal{F}$ are independent then $E[X|\mathcal{F}] = E[X]$

□ Two extreme cases of conditional expectations w.r.t information

  – Perfect information : If $X \in \mathcal{F}$ then $E[X|\mathcal{F}] = X$
  – No information : If $X \perp\!\!\!\perp \mathcal{F}$ then $E[X|\mathcal{F}] = E[X]$

* Conditional variance

$$Var(X|\mathcal{F}) := E[\{X - E[X|\mathcal{F}]\}^2|\mathcal{F}] = E[X^2|\mathcal{F}] - E^2[X|\mathcal{F}]$$

Conditional variance is defined for $X \in \mathcal{L}^2$

2 Martingales

* Definition needed for martingales

    – Given a probability space $(\Omega, \mathcal{F}, P)$, increasing sequence of sub σ-fields $\{\mathcal{F}_n\}_{n=0}^{\infty}$ is called a filtration.

    – A random sequence $\{X_n\}_{n=0}^{\infty}$ is said to be adapted to $\{\mathcal{F}_n\}$ if $X_n \in \mathcal{F}_n \quad \forall n \in \mathbb{N} \cup \{0\}$

* Definition of martingale and their cousins

    – $\{X_n\}_{n=0}^{\infty}$: a random sequence. $\{\mathcal{F}_n\}_{n=0}^{\infty}$: a filtration. Assume $E|X_n| < \infty \quad \forall n \in \mathbb{N} \cup \{0\}$ and $\{X_n\}$ is adapted to $\{\mathcal{F}_n\}$. Then $\{X_n\}$ is said to be a martingale (w.r.t $\{\mathcal{F}_n\}$) if $E[X_{n+1}|\mathcal{F}_n] = X_n \quad \forall n \in \mathbb{N} \cup \{0\}$

    – $\{X_n\}$ is said to be a submartingale (w.r.t $\{\mathcal{F}_n\}$) if $E[X_{n+1}|\mathcal{F}_n] \geq X_n \quad \forall n \in \mathbb{N} \cup \{0\}$

    – $\{X_n\}$ is said to be a supermartingale (w.r.t $\{\mathcal{F}_n\}$) if $E[X_{n+1}|\mathcal{F}_n] \leq X_n \quad \forall n \in \mathbb{N} \cup \{0\}$

    $\sqrt{}$ These are abbreviated to 'mtg', 'submtg', 'supermtg' respectively.

• Examples of martingales

    i. $\{\xi_n\}_n$ $i.i.d$ with $E(\xi_1) = 0$. $X_0 = 0$. $X_n = \xi_1 + \cdots + \xi_n$ and $\mathcal{F}_0 = \{\phi, \Omega\}$.
    $\mathcal{F}_n = \sigma(\xi_1, \cdots, \xi_n)$. Then $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$

    ★ Trick : $E[Z]$ is finite $\Leftrightarrow Z$ is integrable. ($\because$ the definition of expectation)

    ii. Adding assumption $Var(\xi_1) = \sigma^2 < \infty$ to i. above.
    Then $\{X_n - n\sigma^2\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$

    iii. $\{\varepsilon_n\}_n$ $i.i.d \sim (0,1)$. $X_0 = 0$. $X_{n+1} = X_n + h(X_n)\varepsilon_{n+1}$ with $h : \mathbb{R} \to \mathbb{R}$ Borel function s.t.
    $E|h(X_n)| < \infty \quad \forall n \in \mathbb{N} \cup \{0\}$ and $\mathcal{F}_0 = \{\phi, \Omega\}$. $\mathcal{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$
    Then $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$

    iv. $\{\varepsilon_n\}_n$ $i.i.d \sim (0,1)$. $Y_0 = 0$. $Y_{n+1} = \phi(Y_n)\varepsilon_{n+1}$ with $\phi(y) = w + \alpha y^2$ $(w > 0, 0 \leq \alpha < 1)$
    and $E[\phi(Y_n)] < \infty \quad \forall n \in \mathbb{N}$. and $\mathcal{F}_0 = \{\phi, \Omega\}$. $\mathcal{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$.
    Let $X_0 = 0$. $X_n = Y_1 + \cdots Y_n$. Then $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$

    $\sqrt{}$ Such $\{Y_n\}$ is called as ARCH (autoregressive conditional heteroskedasticity) process

• Elementary facts about Martingales

    – Every martingale is a submartingale and a supermartingale

    – If $\{X_n\}$ is a submartingale then $\{-X_n\}$ is a supermartingale

    $\sqrt{}$ We develop theory about martingales often assuming submartingale since every martingale is submartingale and every supermartingale is negative version of submartingale

    – If $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$ then $E[X_n|\mathcal{F}_m] = X_m$ whenever $n \geq m$

    – If $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$ then $E[X_n|\mathcal{F}_m] \geq X_m$ whenever $n \geq m$

    – If $\{X_n\}$ is a supermartingale w.r.t $\{\mathcal{F}_n\}$ then $E[X_n|\mathcal{F}_m] \leq X_m$ whenever $n \geq m$

    – If $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$ then $\{E[X_n]\}$ is constant.

    – If $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$ then $\{E[X_n]\}$ is increasing.

    – If $\{X_n\}$ is a supermartingale w.r.t $\{\mathcal{F}_n\}$ then $\{E[X_n]\}$ is decreasing.

- Convex transformation of martingale

  – If $\{X_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$ and $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function s.t. $E|\phi(X_n)| < \infty \quad \forall n \in \mathbb{N}$ then $\{\phi(X_n)\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$

  – If $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$ and $\phi : \mathbb{R} \to \mathbb{R}$ is a convex and increasing function s.t. $E|\phi(X_n)| < \infty \quad \forall n \in \mathbb{N}$ then $\{\phi(X_n)\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$

  – If $\{X_n\}$ is a supermartingale w.r.t $\{\mathcal{F}_n\}$ and $\phi : \mathbb{R} \to \mathbb{R}$ is a concave and increasing function s.t. $E|\phi(X_n)| < \infty \quad \forall n \in \mathbb{N}$ then $\{\phi(X_n)\}$ is a supermartingale w.r.t $\{\mathcal{F}_n\}$

  (Ex) If $\{X_n\}$ is a martingale and $E[|X_n|^p] < \infty$ for some $p \geq 1$, then $\{|X_n|^p\}$ is a submartingale

  (Ex) If $\{X_n\}$ is a submartingale then for any $a \in \mathbb{R}$, $\{(X_n - a)^+\}$ is a submartingale

  (Ex) If $\{X_n\}$ is a supermartingale then for any $a \in \mathbb{R}$, $\{X_n \wedge a\}$ is a supermartingale

  (Ex) If $\{X_n\}$ is a submartingale then $\{X_n^+\}$ is a submartingale and $\{X_n^-\}$ is a supermartingale

* Predicatable sequence and a process using it

  – For a filtration $\{\mathcal{F}_n\}_{n=0}^{\infty}$, a random sequence $\{H_n\}_{n=1}^{\infty}$ is said to be a predicatable sequence (w.r.t $\{\mathcal{F}_n\}$) if $H_n \in \mathcal{F}_{n-1} \quad \forall n \in \mathbb{N}$

    $\sqrt{}$ A letter H stands for a 'height'

  – Suppose $\{X_n\}$ is adapted to $\{\mathcal{F}_n\}$. For a predicatable sequence $\{H_n\}$ (w.r.t $\{\mathcal{F}_n\}$), we define a process $\{(H \cdot X)_n\}$ by

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

    $\sqrt{}$ Note that $\{(H \cdot X)_n\}$ is adapted to $\{\mathcal{F}_n\}$

    $\sqrt{}$ The definition above can be extended from $\{(H \cdot X)_n\}_{n \in \mathbb{N}}$ to $\{(H \cdot X)_n\}_{n \in \mathbb{N} \cup \{0\}}$ with additionally defining $(H \cdot X)_0 = 0$. Obviously $(H \cdot X)_0 \in \mathcal{F}_0$. For the following theorems using this process, we can regard it as $\{(H \cdot X)_n\}_{n \in \mathbb{N} \cup \{0\}}$

- Elementary facts about martingale transform with predicatable sequence

  – Let $\{X_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=1}^{\infty}$ be a random sequence and $\{H_n\}$ is a predicatable sequence w.r.t. a filtration $\{\mathcal{F}_n\}_{n=0}^{\infty}$. Assume $E|X_n H_n| < \infty$, $E|X_{n-1} H_n| < \infty \quad \forall n \in \mathbb{N}$

    i. If $\{X_n\}$ is a martingale (w.r.t $\{\mathcal{F}_n\}$) then $\{(H \cdot X)_n\}$ is also a martingale

    ii. If $\{X_n\}$ is a submartingale (w.r.t $\{\mathcal{F}_n\}$) and $H_n \geq 0$ then $\{(H \cdot X)_n\}$ is also a submartingale

    iii. If $\{X_n\}$ is a supermartingale (w.r.t $\{\mathcal{F}_n\}$) and $H_n \geq 0$ then $\{(H \cdot X)_n\}$ is also a supermartingale

    $\sqrt{}$ The condition "$E|X_n H_n| < \infty$, $E|X_{n-1} H_n| < \infty \quad \forall n \in \mathbb{N}$" can be replaced with "For each $n \in \mathbb{N}$, H_n is bounded".

* Stopping time

  – A (extended) random variable N taking values of $\mathbb{N} \cup \{0, \infty\}$ is said to be a stopping time (w.r.t a filtration $\{\mathcal{F}_n\}$) if an event $(N = n) \in \mathcal{F}_n \quad \forall n \in \mathbb{N}$

  $$(N \leq n) = \bigcup_{j=0}^{n}(N = j) \in \mathcal{F}_n \qquad\qquad (N > n) = (N \leq n)^C \in \mathcal{F}_n$$

  $$(N < n) = \bigcup_{j=0}^{n-1}(N = j) \in \mathcal{F}_{n-1} \qquad\qquad (N \geq n) = (N < n)^C \in \mathcal{F}_{n-1}$$

- $(N \geq n)$ is a $\mathcal{F}_{n-1}$-measurable event. $I(N \geq n)$ is $\mathcal{F}_{n-1}$-measurable random variable. Hence, $\{I(N \geq n)\}_n$ is a predictable sequence given N is a stopping time.
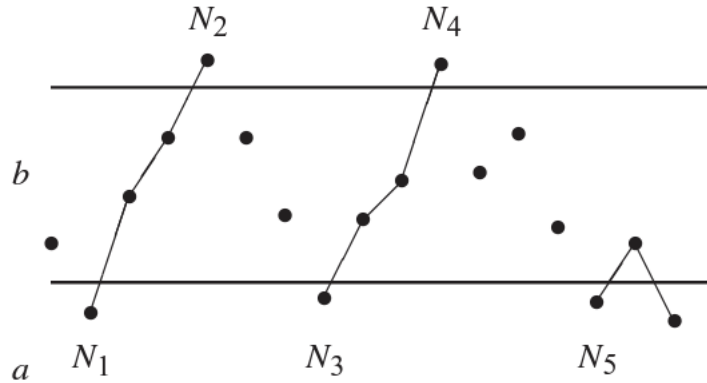
- **Martingale stopped by stopping time**

  - Let $\{X_n\}$ be a random sequence adapted to $\{\mathcal{F}_n\}$. Let N be a stopping time w.r.t $\{\mathcal{F}_n\}$ and put $H_n = I(N \geq n) \quad \forall n \in \mathbb{N}$. Then $(H \cdot X)_n = X_{N \wedge n} - X_0$.

  - The process $\{X_{N \wedge n}\}_n$ is said to be a martingale stopped by stopping time N, provided $\{X_n\}$ is a martingale.

    ★ If $\{X_n\}$ and $\{Y_n\}$ are martingales (w.r.t. $\{\mathcal{F}_n\}$) then $\{X_n + Y_n\}$ is also a martingale. The same holds for submartingales and supermartingales too.

  - If $\{X_n\}$ is a martingale and N is a stopping time then $\{X_{N \wedge n}\}$ is martingale.

  - If $\{X_n\}$ is a submartingale and N is a stopping time then $\{X_{N \wedge n}\}$ is submartingale.

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

- **Stopping time and Upcrossing**

  - Suppose $\{X_n\}_{n=0}^{\infty}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$. Let $a < b$. Define N_j's as below :

$$N_1 = \inf\{m \geq 0 \ : \ X_m \leq a\} \qquad N_2 = \inf\{m > N_1 \ : \ X_m \geq b\}$$
$$N_3 = \inf\{m > N_2 \ : \ X_m \leq a\} \qquad N_4 = \inf\{m > N_3 \ : \ X_m \geq b\}$$
$$\vdots \qquad\qquad\qquad\qquad \vdots$$
$$N_{2k-1} = \inf\{m > N_{2k-2} \ : \ X_m \leq a\} \qquad N_{2k} = \inf\{m > N_{2k-1} \ : \ X_m \geq b\}$$
$$\vdots \qquad\qquad\qquad\qquad \vdots$$

  - Every N_j for $j \in \mathbb{N}$ is stopping time w.r.t $\{\mathcal{F}_n\}$. $N_1 < N_2 < N_3 \cdots$ provided all N_j's are finite. (It is possible that $N_j = \infty$ provided it has a form of $\inf(\phi)$)

  - 'Upcrossing' is a case where the submartingale $\{X_n\}$ crosses from below a to above b.

  - $\mathcal{U}_n := \sup\{k \ : \ N_{2k} \leq n\}$ is the number of upcommings completed by time n

- **Upcrossing inequality**

  - Suppose $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$. If stopping time N_j and the number of upcrossings $\mathcal{U}_n$ are defined as above then

$$(b - a)E[\mathcal{U}_n] \leq E[(X_n - a)^+] - E[(X_0 - a)^+]$$

- Submartingale convergence theorem

  - If $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\}$ with $\sup_n E(X_n^+) < \infty$ then $X_n \to X$ $a.s.$ for some integrable random variable X

    - ★ Trick : If $X_n \to X$ $a.s.$ then $X_n^+ \to X^+$ $a.s.$ and $X_n^- \to X^-$ $a.s.$
    - ★ Lemma : If the number of upcrossings of $[a, b]$ by submartingale $\{X_n\}$ is finite for any $a, b \in \mathbb{Q}$, then $\lim_n X_n$ exists. i.e. X_n converges to some r.v. almost surely.

  - ☐ If $\{X_n\}$ is a nonnegative supermartingale w.r.t $\{\mathcal{F}_n\}$ then $X_n \to X$ $a.s.$ for some integrable random variable X s.t. $E(X) \leq E(X_0)$

- Example of martingale which converges almost surely but not in L^1

  - $\{\xi_n\}_n$ $i.i.d$ with $P(\xi_1 = 1) = P(\xi_1 = -1) = 1/2$. Let $S_0 = 1$, $S_n = S_{n-1} + \xi_n$ and $\mathcal{F}_0 = \{\phi, \Omega\}$, $\mathcal{F}_n = \sigma(\xi_1, \cdots, \xi_n)$. Then $\{S_n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$
    Let $N = \inf\{n \in \mathbb{N} : S_n = 0\}$. Then N is a stopping time.
    $X_n := S_{N \wedge n}$ so that $X_n = S_n$ if $n < N$ and $X_n = 0$ if $n \geq N$. $\{X_n\}$ is a nonnegative integer valued martingale w.r.t $\{\mathcal{F}_n\}$. $X_n \to 0$ $a.s.$ but $X_n \nrightarrow 0$ in $\mathcal{L}^1$.

- If $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ is a negative submartingale w.r.t $\{\mathcal{F}_n\}_{n \in \mathbb{N} \cup \{0\}}$ then so is $\{X_n\}_{n \in \mathbb{N} \cup \{0, \infty\}}$ w.r.t $\{\mathcal{F}_n\}_{n \in \mathbb{N} \cup \{0, \infty\}}$ where $X_\infty = \lim_n X_n$ and $\mathcal{F}_\infty = \sigma\left(\bigcup_{n=0}^\infty \mathcal{F}_n\right)$

  - If $\{X_n\}_{n \in \mathbb{N}}$ is a martingale w.r.t $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ and $X_n \to X_\infty$ $a.s.$ then $X_\infty \in \mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$

- Doob's decomposition

  - Any submartingale $\{X_n\}$ can be written as $X_n = M_n + A_n$ where $\{M_n\}$ is a martingale and $\{A_n\}$ is a predictable increasing sequence with $A_0 = 0$. Also, this expression is unique in the sense that if $X_n = M_n' + A_n'$ is another expression then $M_n = M_n'$ and $A_n = A_n'$ $a.s.$

  - $\sqrt{}$ The exact form of M_n, A_n for given X_n is $A_n = A_{n-1} + E[X_n | \mathcal{F}_{n-1}] - X_{n-1}$ $\forall n \in \mathbb{N}$ and $M_n = X_n - A_n$ $\forall n \in \mathbb{N} \cup \{0\}$ (Since $A_0 = 0$, $A_n = \sum_{k=1}^n (E[X_k | \mathcal{F}_{k-1}] - X_{k-1})$)

- * Almost sure convergence on a restricted event

  - Define "$X_n \to X$ $a.s.$ on B" for measurable set B as
    $P((X_n \to X) \cap B) = P(B)$ or equivalently, $P((X_n \nrightarrow X) \cap B) = 0$

  - ★ Lemma : $X_n \to X$ $a.s.$ on B $\Rightarrow$ $X_n \to X$ $a.s.$ on A whenever $A \subset B$

  - ★ Lemma : $X_n \to X$ $a.s.$ on A_k $\forall k \in \mathbb{N}$ $\Rightarrow$ $X_n \to X$ $a.s.$ on $\bigcup_{k \in \mathbb{N}} A_k$

- Martingales with bounded increments either converge or oscillate between ∞ and $-\infty$

  - Let $\{X_n\}$ be a martingale with $|X_n - X_{n-1}| \leq M < \infty$ $\forall n \in \mathbb{N}$ for some $M > 0$.
    Define disjoint subsets $C, D \subset \Omega$ by

$$C = (\lim_n X_n \text{ exists and } -\infty < \lim_n X_n < \infty)$$

$$D = (\limsup X_n = \infty \text{ and } \liminf X_n = -\infty)$$

  Then $P(C \cup D) = 1$

- Conditional Borel-Cantelli second lemma

– Let $\{\mathcal{F}_n\}_{n \in N \cup \{0\}}$ be a filtration with $\mathcal{F}_0 = \{\phi, \Omega\}$. If $A_n \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}$ then

$$(A_n \ i.o.) = \Big(\sum_{n=1}^{\infty} P(A_n | \mathcal{F}_{n-1}) = \infty \Big) \ a.s.$$

★ Define "$A = B \ a.s.$" for measurable sets A and B by $P(A \Delta B) = 0$ where $A \Delta B$ denotes the symmetric difference of two sets.

★ $\sum_{k=1}^{n} I_{A_k}$ is a submartingale whose martingale compoenent of Doob's decomposition is

$$\sum_{k=1}^{n} I_{A_k} - \sum_{k=1}^{n} \Big(E\Big[\sum_{j=1}^{k} I_{A_j} | \mathcal{F}_{k-1} \Big] - \sum_{j=1}^{k-1} I_{A_j} \Big) = \sum_{k=1}^{n} I_{A_k} - \sum_{k=1}^{n} P(A_k | \mathcal{F}_{k-1})$$

and this is the martingale we exploit in the proof of conditional B-C 2nd lemma

★ Trick : $(A_n \ i.o.) = \big(\sum_{n=1}^{\infty} I_{A_n} = \infty \big)$

√ Given $\{A_n\}$ is independent, by setting $\mathcal{F}_n = \sigma(A_1, \cdots, A_n)$, conditional Borel-Cantelli second lemma implies original Borel-Cantelli second lemma which is given by

$$\sum_n P(A_n) = \infty \ \Rightarrow \ P(A_n \ i.o.) = 1$$

* Branching process (Galton-Watson process)

– Let $\{\xi_i^n\}_{i \in \mathbb{N},\, n \in \mathbb{N}}$ be i.i.d nonnegative integer-valued random variables.
Define a Galton-Watson process $\{Z_n\}_{n \in \mathbb{N} \cup \{0\}}$ as below :

$$Z_0 = 1$$
$$Z_{n+1} = \begin{cases} \xi_1^{n+1} + \cdots + \xi_{Z_n}^{n+1} = \sum_{j=1}^{Z_n} \xi_j^{n+1} & \text{if} \quad Z_n > 0 \\ 0 & \text{if} \quad Z_n = 0 \end{cases}$$

√ The idea behind the definitions is that Z_n is the population in the n-th generation and each member of the n-th generation gives birth independently to an identically distributed number of offspring.

– $P(\xi_1^1 = k) \quad \forall\, k \in \mathbb{N} \cup \{0\}$ is called the offspring distribution. $\mu = E(\xi_1^1)$ is the expected number of offspring per individual.

• Properties of the branching process

– Let $\mathcal{F}_n = \sigma(\{\xi_i^m : i \in \mathbb{N},\, 1 \le m \le n\}) \quad \forall\, n \in \mathbb{N}$, $\mathcal{F}_0 = \{\phi, \Omega\}$. If $\mu = E(\xi_1^1) \in (0, \infty)$ then $\{Z_n / \mu^n\}$ is a martingale w.r.t $\{\mathcal{F}_n\}$ and $E(Z_n) = \mu^n \quad \forall\, n \in \mathbb{N}$

– If $\mu = E(\xi_1^1) \in (0, 1)$ then $Z_n = 0$ for large enough n's $a.s.$ i.e. the species goes extinct.

• Inequality for bounded stopping time

– If $\{X_n\}$ is a submartingale and N is a stopping time with $P(N \le K) = 1$ for some $K \in \mathbb{N}$ then

$$E(X_0) \le E(X_N) \le E(X_K)$$

√ Since $\{X_n\}$ is a submartingale, $E(X_0) \le E(X_j) \le E(X_K)$ whenever $0 \le j \le K$. This thm tells us that similar inequality still holds true when the index is random.

□ If $\{X_n\}$ is a martingale and N is a stopping time with $P(N \leq K) = 1$ for some $K \in \mathbb{N}$ then
$$E(X_0) = E(X_N) = E(X_K)$$

- Doob's inequality

  - Let $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ be a submartingale. Take $n \in \mathbb{N}$ and define $\overline{X}_n = \max_{0 \leq m \leq n} X_m$. Let $\lambda > 0$ and define an event $A = (\overline{X}_n \geq \lambda)$. Then the inequality below holds true.
  $$\lambda P(A) \leq E[X_n I_A] \leq E[X_n^+ I_A] \leq E[X_n^+]$$

  □ Let $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ be a supermartingale. Take $n \in \mathbb{N}$ and define $\overline{X}_n = \max_{0 \leq m \leq n} X_m$. Let $\lambda > 0$ and define an event $A = (\overline{X}_n \geq \lambda)$. Then the inequality below holds true.
  $$\lambda P(A) \leq E[X_0] - E[X_n I_{A^c}] \leq E[X_0] + E[X_n^-]$$

  √ Note that $P(A)$ involves $\max_{0 \leq m \leq n}$ term while $E[X_n^+]$ or $E[X_n^-]$ only depends on n

- Doob's L^p maximal inequality

  - If $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ is a nonnegative submartingale, then for $1 < p < \infty$ and $\overline{X}_n = \max_{0 \leq m \leq n} X_m$, the inequality below holds true.
  $$E(\overline{X}_n^p) \leq \left(\frac{p}{p-1}\right)^p E[X_n^p]$$

  □ If $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ is a martingale then for $1 < p < \infty$ and $|\overline{X}_n| = \max_{0 \leq m \leq n} |X_m|$, the inequality below holds true.
  $$E|\overline{X}_n|^p \leq \left(\frac{p}{p-1}\right)^p E|X_n|^p$$

  ★ Lemma : If $X \geq 0$ then $E(X) = \int_0^\infty P(X > t)\, dt$

- L^p convergence thm

  - If $\{X_n\}$ is a martingale with $\sup_n E|X_n|^p < \infty$ for some $p > 1$ then $X_n \to X$ a.s. and $X_n \to X$ in L^p for some integrable r.v. X

  √ For a martingale convergence thm, the condition was $\sup_n E(X_n^+) < \infty$

  ★ Trick : $a, b \in \mathbb{R}$ and $p \geq 1 \Rightarrow |a + b|^p \leq 2^p(|a|^p + |b|^p)$

* σ-field generated by a stopping time

  - Let τ be a stopping time w.r.t. a filtration $\{\mathcal{F}_n\}$. Then we define $\mathcal{F}_\tau$ as the following :
  $$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap (\tau = n) \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}\}$$

  √ Note that $\mathcal{F}_\tau$ is indeed a σ-field.
  √ τ is $\mathcal{F}_\tau$-measurable
  √ If $\{X_n\}$ is adapted to $\{\mathcal{F}_n\}$ then X_τ is $\mathcal{F}_\tau$-measurable

- Bounded optional stopping thm

  - Let $\{X_n\}$ be a submartingale . Let σ and τ be two bounded stopping times s.t. $\sigma \leq \tau \leq B$ a.s. for some $B \in \mathbb{N}$. Then $E[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma$ a.s.

$\checkmark$ $X_\tau = \sum_{n=0}^B X_n I(\tau = n)$ is well-defined and integrable.

$\checkmark$ By defining property of submartingale, $E[X_m | \mathcal{F}_n] \geq X_n \quad \forall \, m \geq n$. The thm tells us that this property is preserved even when indices are stopping times if they are bounded.

    ★ Trick : For a random variable X and a σ-field $\mathcal{F}$,
        i. $(X \leq a) \in \mathcal{F} \quad \forall \, a \in \mathbb{R} \Rightarrow (X \in A) \in \mathcal{F} \quad \forall \, A \in \mathcal{B}(\mathbb{R})$
        ii. For $S \in \mathcal{F}$, $(X \leq a) \cap S \in \mathcal{F} \quad \forall \, a \in \mathbb{R} \Rightarrow (X \in A) \cap S \in \mathcal{F} \quad \forall \, A \in \mathcal{B}(\mathbb{R})$

    ★ Lemma : For any $\mathcal{F}$-measurable and integrable X and Y ,
        i. If $\int_A X \, dP = \int_A Y \, dP \quad \forall A \in \mathcal{F}$ then $X = Y$ a.s.
        ii. If $\int_A X \, dP \leq \int_A Y \, dP \quad \forall A \in \mathcal{F}$ then $X \leq Y$ a.s.

    ★ Lemma : $\{X_n\}$ is a submartingale w.r.t $\{\mathcal{F}_n\} \Rightarrow \int_A X_n \, dP \leq \int_A X_{n+1} \, dP \quad \forall \, A \in \mathcal{F}_n$

* Uniform integrability

    − A collection of r.v.'s $\{X_t : t \in T\}$ is said to be uniformly integrable if

$$\lim_{a \to \infty} \sup_{t \in T} \int_{|X_t| \geq a} |X_t| \, dP = \lim_{a \to \infty} \sup_{t \in T} E|X_t| I(|X_t| \geq a) = 0$$

$\checkmark$ Denote it as $\{X_t\}_{t \in T}$ $u.i.$

$\checkmark$ A uniformly integrable family is well-controlled in the sense that if $\{X_t\}_{t \in T}$ $u.i.$ then $\exists \, M > 0$ s.t. $\sup_{t \in T} E|X_t| \leq M + 1 < \infty$

$\checkmark$ If $\{X_t\}_{t \in T}$ is uniformly integrable then each X_t is integrable .

• If $\{X_t\}_{t \in T}$ is dominated by a nonnegative integrable r.v. X i.e. $|X_t| \leq X$ a.s. $\quad \forall \, t \in T$ then $\{X_t\}_{t \in T}$ is uniformly integrable.

    ★ Lemma: If X is integrable then $\int_{|X| \geq a} |X| \, dP = E|X| I(|X| \geq a) \to 0$ as $a \to \infty$

• Equivalent condition for uniform integrability

    − $\{X_t\}_{t \in T}$ is uniformly integrable iff both of two conditions below are satified.
        i. $\sup_t E|X_t| < \infty$
        ii. $\forall \varepsilon > 0, \exists \delta > 0$ s.t. $\sup_t \int_A |X_t| \, dP < \varepsilon$ whenever $A \in \mathcal{F}$ and $P(A) < \delta$

• Elementary properties of uniform integrable family

    − If $\{X_n\}_{n \in \mathbb{N}}$ and $\{Y_n\}_{n \in \mathbb{N}}$ are both uniformly integrable then $\{X_n + Y_n\}_{n \in \mathbb{N}}$ $u.i.$
    − If $|X_n| \leq |Y_n| \quad \forall n \in \mathbb{N}$ and $\{Y_n\}_{n \in \mathbb{N}}$ is uniformly integrable then $\{X_n\}_{n \in \mathbb{N}}$ $u.i.$

• Vitali's lemma

    − For $p \geq 1$, if $\{X_n\} \underset{seq}{\subset} L^p$ and $X_n \xrightarrow{P} X$ then the followings are equivalent.
        i. $\{X_n{}^p\}_{n \in \mathbb{N}}$ is uniformly integrable.
        ii. $X \in L^p$ and $X_n \to X$ in L^p
        iii. $E|X_n|^p \to E|X|^p < \infty$
        ★ Lemma : For a r.v. Z, continuity set $\{z \in \mathbb{R} : P(Z = z) = 0\}$ is dense in $\mathbb{R}$

• If $\{X_n\}_{n \in \mathbb{N}}$ is uniformly integrable and $X_n \xrightarrow{D} X$ then $E|X_n| \to E|X|$ and $E(X_n) \to E(X)$

★ Lemma : If $Y_n \to Y$ in L^1 then $E|Y_n| \to E|Y|$ and $E(Y_n) \to E(Y)$

* Regular martingale and closable martingale

    – Let $\{X_n\}_{n \in \mathbb{N} \cup \{0\}}$ be a martingale.

        i. $\{X_n\}$ is said to be regular if $\exists\, X \in L^1$ s.t. $X_n = E[X|\mathcal{F}_n]$ $a.s.$ $\quad \forall n \in \mathbb{N}$

        ii. $\{X_n\}$ is said to be closable if $\exists\, X_\infty \in L^1$ s.t. $X_n \to X_\infty$ $a.s.$, $X_\infty \in \mathcal{F}_\infty$ where
$\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$ and $E[X_\infty|\mathcal{F}_n] = X_n$ $\quad \forall n \in \mathbb{N}$ so that
$\{X_n\}_{n \in \mathbb{N} \cup \{0,\infty\}}$ is a martingale w.r.t $\{\mathcal{F}_n\}_{n \in \mathbb{N} \cup \{0,\infty\}}$

        $\sqrt{}$ Every closable martingale is regular.

• For a martingale $\{X_n\}_{n \in \mathbb{N}}$, the followings are equivalent.

    i. $\{X_n\}$ is regular.

    ii. $\{X_n\}$ is uniformly integrable and converges $a.s.$

    iii. $\{X_n\}$ converges in L^1

    iv. $\{X_n\}$ is closable.

□ For a martingale $\{X_n\}_{n \in \mathbb{N}}$ w.r.t $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$

    – If $X_n \to X$ in L^1 then $X_n \to X$ $a.s.$ and $X_n = E[X|\mathcal{F}_n]$ $\quad \forall n \in \mathbb{N}$

    – If $\{X_n\}$ is uniformly integrable then $X_n \to X$ $a.s.$ for some integrable r.v. X and
$X_n = E[X|\mathcal{F}_n]$ $\quad \forall n \in \mathbb{N}$

    – If $X_n = E[X|\mathcal{F}_n]$ for some integrable r.v. X then $\{X_n\}$ is uniformly integrable and
$\exists$ integrable r.v. $X_\infty \in \mathcal{F}_\infty$ s.t. $E[X_\infty|\mathcal{F}_n] = X_n$ $\quad \forall n \in \mathbb{N}$ and $X_n \to X_\infty$ $a.s.$ and in L^1.

• Levy's thm

    – If $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is a filtration and $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$ then for an integrable r.v. X,
$E[X|\mathcal{F}_n] \to E[X|\mathcal{F}_\infty]$ $a.s.$ and in L^1.

• Conditional DCT (generalized version)

    – Let $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ be a filtration and $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$. If $X_n \to X$ $a.s.$ and $|X_n| \leq Z$ for
some integrable r.v. Z , then $E[X_n|\mathcal{F}_n] \to E[X|\mathcal{F}_\infty]$ $a.s.$

* Potential

    – A nonnegative supermartingale $\{X_n\}$ is said to be potential if $E(X_n) \to 0$

    $\sqrt{}$ If $\{X_n\}$ is potential then $\{X_n\}$ is uniformly integrable and $X_n \to 0$ $a.s.$

• Riesz decomposition

    – Let $\{X_n\}$ be a uniformly integrable nonnegative supermartingale. Then we can express
X_n as $X_n = M_n + V_n$ where $\{M_n\}$ is uniformly integrable martingale and $\{V_n\}$ is potential.
Furthermore, such decomposition is unique.

• If $\{X_n\}$ is uniformly integrable submartingale, then for any stopping time N, stopped process
$\{X_{N \wedge n}\}$ is also uniformly integrable submartingale.

    ★ Lemma : If $X_n \to X$ $a.s.$ then $X_n^+ \to X^+$ $a.s.$ and $X_n^- \to X^-$ $a.s.$

- Inequality for unbounded stopping time

    - If $\{X_n\}$ is uniformly integrable submartingale then for any stopping time N, we have $E(X_0) \leq E(X_N) \leq E(X_\infty)$ where $X_n \to X_\infty$ a.s.

    □ If $\{X_n\}$ is uniformly integrable martingale then for any stopping time N, we have $E(X_0) = E(X_N) = E(X_\infty)$ where $X_n \to X_\infty$ a.s.

- Optional stopping thm

    - If $L \leq M$ are stopping times and $\{X_n\}$ is uniformly integrable submartingale then $E[X_L] \leq E[X_M]$ and $X_L \leq E[X_M|\mathcal{F}_L]$ a.s.

- Suppose $\{X_n\}$ is a submartingale and $E\big[|X_{n+1} - X_n|\big|\mathcal{F}_n\big] \leq B$ a.s. $\forall n \in \mathbb{N}$. If N is a stopping time with $E(N) < \infty$ then $\{X_{N\wedge n}\}$ is unifomly integrable and $E(X_0) \leq E(X_N)$

    √ Note that $E(N) < \infty$ condition implies that N is almost surely finite.

    ★ Lemma : $E|X| < \infty \Leftrightarrow \sum_n P(|X| \geq n) < \infty$

- If $\{X_n\}$ is a nonnegative supermartingale and N is a stopping time then $E(X_0) \geq E(X_N)$

- Comment for X_N with stopping time N and (sub)martingale $\{X_n\}$

    - $X_N = \sum_{n=0}^{\infty} X_n I(N = n)$
    - Note that N can take value of $N = \infty$. Thus, for X_N to make sense, N should be almost surely bounded or X_∞ is well-defined.
    - If X_∞ is well-defined such that $X_n \to X_\infty$ a.s. then $X_{N\wedge n} \to X_N$ a.s.
    - How can we figure out integrability of X_N ?

        i. If N is bounded a.s.
            - $N \leq K$ a.s. for some $K \in \mathbb{N}$. Hence $E|X_N| \leq \sum_{n=0}^{K} E|X_n| < \infty$
        ii. If $\{X_n\}$ is uniformly integrable submartingale
            - $X_n \to X_\infty$ a.s. $\Rightarrow X_{N\wedge n} \to X_N$ a.s. Since $\{X_{N\wedge n}\}$ is also uniformly integrable submartingale, by Vitali lemma, $X_N \in L^1$ i.e. X_N is integrable.
        iii. If $\{X_n\}$ is nonengative supermartingale
            - $X_n \to X_\infty$ a.s. $\Rightarrow X_{N\wedge n} \to X_N$ a.s. By inequality for bounded stopping time, $E[X_{N\wedge n}] \leq E[X_0]$ and using Fatou's lemma, we have $0 \leq E[X_N] \leq E[X_0] < \infty$

- Asymmetric simple random walk

    - Let $\{\xi_i\}_{i\in\mathbb{N}}$ be i.i.d. random seq. s.t. $P(\xi_1 = 1) = p$ and $P(\xi_1 = -1) = q = 1 - p$ for some $0 < p < 1$. $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$ and $\mathcal{F}_n = \sigma(\xi_1, \cdots \xi_n)$ $\forall n \in \mathbb{N}$

        i. For $\psi : \mathbb{R} \to \mathbb{R}$ defined by $\psi(x) = \left(\frac{1-p}{p}\right)^x$, $\{\psi(S_n)\}$ is a martingale.

        ii. Define $T_m = \inf\{n \in \mathbb{N} \cup \{0\} : S_n = m\}$ for $m \in \mathbb{Z}$ where $\inf(\phi)$ is interpreted as ∞ For any $a, b \in \mathbb{Z}$ s.t. $a < 0 < b$, we have

        $$P(T_a < T_b) = \frac{\psi(b) - \psi(0)}{\psi(b) - \psi(a)}$$

        $p > q \Rightarrow T_b < \infty$ a.s. , $p < q \Rightarrow T_a < \infty$ a.s. , $p = q \Rightarrow T_a, T_b < \infty$ a.s.

iii. If $p > q$ ($p > \frac{1}{2}$) then

$$P(\inf_n S_n \leq a) = P(T_a < \infty) = \left(\frac{1-p}{p}\right)^{-a} \quad \forall\, a < 0$$

$$E[T_b] = \frac{b}{2p-1} \quad \forall b > 0$$

★ Trick : For $|r| < 1$, $\sum_{n=1}^{\infty} n r^n$ converges to a finite number $r/(1-r)^2$

- Square integrable Martingales

  - Suppose $\{X_n\}$ is a martingale with $X_0 = 0$ and $E[X_n^2] < \infty \quad \forall\, n \in \mathbb{N}$
  - Since X_n^2 is a submartingale, we can find $X_n^2 = M_n + A_n$ which is a Doob's decomposition for X_n^2. Here, A_n is called the increasing process associated with X_n

  $$A_n = \sum_{k=1}^{n} E[X_k^2 | \mathcal{F}_{k-1}] - X_{k-1}^2 = \sum_{k=1}^{n} E[(X_k - X_{k-1})^2 | \mathcal{F}_{k-1}] \quad \text{with} \quad A_0 = 0$$

  - Since A_n is increasing, $A_\infty = \lim_n A_n = \sup_n A_n$ exists.
  - A_n can be though of as a path by path measurement of the variance at time n and A_∞ can be viewed as the total variance in the path.
  - A_n is integrable and $E[X_n^2] = E[A_n] \quad \forall\, n \in \mathbb{N}$
  - Two results about boundedness

  $$E[\sup_n X_n^2] \leq 4 E[A_\infty] \quad \text{and} \quad E[\sup_n |X_n|] \leq 3 E[\sqrt{A_\infty}]$$

  - Two results about convergence
    i. X_n converges to a finite limit $a.s.$ on $(A_\infty < \infty)$
       ★ Trick : If N is a stopping time , then $X_{N \wedge n}$ is a square integrable martingale and $A_{N \wedge n}$ is the increasing process associated with $X_{N \wedge n}$
    ii. If $f : \mathbb{R} \to \mathbb{R}$ be an increasing function with $f(t) \geq 1 \,\forall\, t \in \mathbb{R}$ and $\int_0^\infty 1/f^2(t)\, dt < \infty$ then $X_n/f(A_n) \to 0$ $a.s.$ on $(A_\infty = \infty)$
       ★ Trick : For such function f , we have $f(t) \nearrow \infty$ as $t \to \infty$
       ★ Lemma : [Kronecker's lemma] If $0 < a_n \nearrow \infty$ and $\sum_{n=1}^{\infty} \frac{1}{a_n} x_n$ converges then $\frac{1}{a_n} \sum_{j=1}^{n} x_j \to 0$

- Conditional Borel-Cantelli second lemma (2nd version)

  - Suppose $B_n \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}$ and $\mathcal{F}_0 = \{\phi, \Omega\}$. Let $p_n = P(B_n | \mathcal{F}_{n-1}) \quad \forall\, n \in \mathbb{N}$. Then ,

  $$\frac{\sum_{m=1}^{n} I_{B_m}}{\sum_{m=1}^{n} p_m} \to 1 \;\; a.s. \quad \text{on} \quad \left(\sum_n p_n = \infty\right)$$

√ This extends the original Borel-Cantelli second lemma .

- Buckholder's inequality

  - Suppose $\{S_i\}_{0 \leq i \leq n}$ is a martingale with $S_0 = 0$. Let $X_i = S_i - S_{i-1}$ be a 'martingale difference' . Then for $1 < p < \infty$, $\exists\, c_1$ and c_2 depending only on p such that

  $$c_1 \cdot E \big| \sum_{i=1}^{n} X_i^2 \big|^{\frac{p}{2}} \leq E|S_n|^p \leq c_2 \cdot E \big| \sum_{i=1}^{n} X_i^2 \big|^{\frac{p}{2}}$$

  In fact, $c_1 = (18p^{1/2}q)^{-p}$ and $c_2 = (18pq^{1/2})p$ where q is Holder conjugate of p

  $\sqrt{}$ We can handle martingale with squared sum of martingale difference sequences.

☐ Marcinkiewicz-Zygmund inequality

  - Suppose X_i's are i.i.d. random variables with $E[X_1] = 0$ and $E|X_1|^P < \infty$ for some $p > 1$. Let $S_n = X_1 + \cdots + X_n$ (S_n is a martingale w.r.t. $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$) . Then we have

  $$\|S_n\|_p \leq 18q^{1/2}p\sqrt{n}\|X_1\|_p$$

  with q being Holder conjugate of p

- Uniform integrability of moments of sample mean

  - If $\{X_n\}$ is a seq. of i.i.d. random variables with mean 0 and variance 1 and $E|X_1| < \infty$ for some $r > 1$ then

  $$\left\{ \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \right|^r \right\}_{n \in \mathbb{N}}$$

  is uniformly integrable.

  $\sqrt{}$ Combining this thm with CLT and Vitalli lemma, we can obtain that under the same assumption,

  $$E \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \right|^r \to E|Z|^r$$

  where $Z \sim N(0,1)$

3   Infinitely Divisible Distribution

* Definition of infinitely divisible distributions

  - A distribution F is called inifinitely divisble if one of these equivalent conditions are satisfied.

    i. For each $n \in \mathbb{N}$, $\exists$ a distribution F_n s.t. $F = F_n * F_n * \cdots * F_n$ where $*$ denotes the convolution

    ii. There is a probability space $(\Omega, \mathcal{F}, P)$, a random variable X , and a double array $\{X_{nk} : k = 1, \cdots n\}$ such that

    $$X \sim F \quad \text{and} \quad X_{nk} \overset{i.i.d.}{\sim} F_n \quad \text{with} \quad X \overset{D}{=} X_{n1} + \cdots + X_{nn} \qquad \forall\, n \in \mathbb{N}$$

    iii. For each $n \in \mathbb{N}$, $\exists$ a distribution F_n s.t.

    $$\psi_F = (\psi_{F_n})^n$$

    where ψ_G denotes the characteristic function of the distribution G

– We will say a characteristic function ψ is infinitely divisible if it is a characteristic function of inifinite divisble distribution F

- For any characteristic function ψ , we have

$$1 - |\psi(2t)|^2 \leq 4(1 - |\psi(t)|^2) \quad \forall \, t$$

- If ψ is inifinitely divisble characteristic function then $\psi(t) \neq 0$ for any $t \in \mathbb{R}$

- Canonical representation of infinitely divisble characteristic function

– Let μ be a finite measure in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Define ψ as

$$\psi(t) = \exp\left(\int_{-\infty}^{\infty} \frac{e^{itx} - 1 - itx}{x} \, d\mu(x) \right)$$

Then ψ is a characteristic function of an infinite divisible distribution with mean 0 and variance $\mu(\mathbb{R})$.

– This is called canonical representation of ψ and μ is called a canonical measure.

- condition $\mathcal{R}$ of a row-wise independent double array

– A row-wise independent double array $\{X_{nk} : 1 \leq k \leq r_n\}_n$ is said to satisfy condition $\mathcal{R}$ if

  i. $E(X_{nk}) = 0, E[X_{nk}^2] = \sigma_{nk}^2 < \infty, s_n^2 = \sum_{k=1}^{r_n} \sigma_{nk}^2 > 0$
  ii. $\sup_n s_n^2 < \infty$
  iii. $\max_{1 \leq k \leq r_n} \sigma_{nk}^2 \to 0$

- If X and Y are independent and $E[(X + Y)^2] < \infty$ then $E[X^2] + E[Y^2] < \infty$

- Every inifinite divisble distribution F with mean 0 and finite variance σ^2 is the limit distribution of $S_n = X_{n1} + \cdots + X_{nr_n}$ for some row-wise independent double array $\{X_{nk}\}$ satisfying condition $\mathcal{R}$. In fact, if double array $\{X_{nk} : k = 1, \cdots, n\}_n$ is associated with infinite divisble distribution F then it satisfies stronger condition than condition $\mathcal{R}$

- Infinite divisible distribution has a canonical representation .

– If F is the limit distribution of $S_n = X_{n1} + \cdots X_{nn}$ for some row-wise independent double array $\{X_{nk} : k = 1, \cdots, n\}$ satisfying condition $\mathcal{R}$ then F has the characteristic function ψ with canonical representation

– By the theorem above, if F is an infinite divisble distribution with mean 0 and finite variance σ^2 then it has canonical representation.