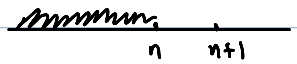
Markov Chain (Review)

$X_n(\omega)$, $n = 0, 1, \cdots$



$X_0(\omega), \cdots X_n(\omega)$

$X_{n+1}(\omega)$?

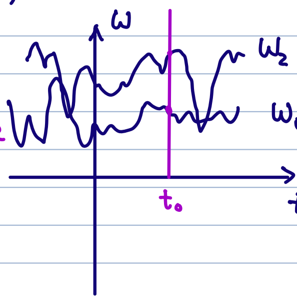
$X(t, \omega)$ $\quad t \in R^+ \cup \{0\}$

$X(t_0, \underline{\underline{\omega}})$

Random Variable

$X(t, \omega_i)$

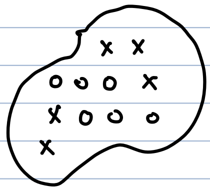
function



$$\Pr\Big(\underset{\downarrow \atop \text{future}}{X_{n+1}(\omega)} \Big| \underbrace{X_0 \cdots X_n}_{\text{history}}\Big) = \Pr\Big(X_{n+1}(\omega) \Big| \underset{\downarrow \atop \text{present}}{X_n(\omega)}\Big)$$

Example A

(HW1)



- Urn has R red balls, B black balls.
- Randomly pick one ball, put back this ball and additional c balls of the same color and d balls of the other color
- Repeat ....

$n = 0, 1, 2$

$\{X_n\}_{n \geq 0}$ = # of red balls after n-th pick

$\triangle$ $\{Y_n\}_{n \geq 0}$ = # of black balls after n-th pick

Either is "equivalent" for analysis

$\vee$ $\{Z_n\}_{n \geq 0}$ = # of total balls after n-th pick $\rightarrow$ $R + B + n(c+d)$

$\triangle\triangle$ $\{W_n\}_{n \geq 0}$ $\quad W_n = \begin{cases} 1 & \text{if n-th pick is black} \\ 0 & \text{otherwise} \end{cases}$

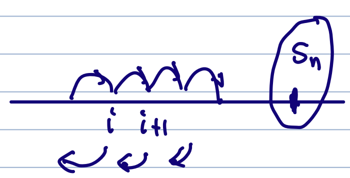
Claim $\{X_n\}$ is M. C.

$\{W_n\}$ is NOT M.C.

Example B $\quad$ M.C. $\Longleftarrow$ non M.C. $\quad \{X_n\} \Rightarrow \vec{V}_n = \begin{pmatrix} X_n \\ \vdots \\ X_1 \end{pmatrix}$

(HW1) $S_n = \sum_{i=1}^{n} X_i \qquad X_0 = 0 \qquad \Pr(X_i = -1) = p = 1 - \Pr(X_i = 1)$

$Y_n = \max(S_i, i \leq n)$ ~ running maximum of simple random walk.



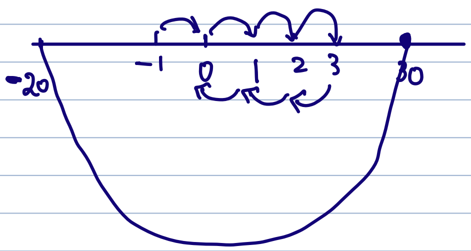
claim: Y_n is NOT M.C. (Why)?

Q: Some enlargement of Y_n to creat a M. C. ?

(Y_n, S_n) is a M.C. ?

Example C: M.C. $\overset{?}{\Longrightarrow}$ non M.C.

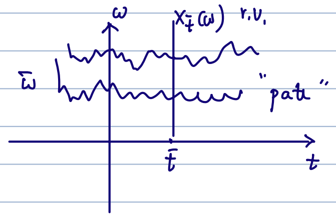
(HW1) Simple random walk   Collapse of states, say 30 & -20.



Martingale $\qquad\qquad\qquad\qquad$ $\{X_n\}_{n\geq 0}$ — Markov Chain

A class of stochastic process indexed by A.

$$\{X_t\}_{t\in A}$$



E.g. $A = \mathbb{N}\cup\{0\}$ $\quad$ "time"

$A = \{\text{location, time, magnitude}\}$

Fix $A = \mathbb{N}\cup\{0\}$ discrete-time stochastic processes

"independence" in both Martingale and M.C.

<u>Martingale</u>: $\{X_n\}_{n\geq 1}^{\infty}$, $\qquad\qquad\qquad$ $\{Y_n\}_{n\geq 1}^{\infty}$

"fair" $\qquad$ Stochastic process of interest $\qquad$ Stochastic process that provides "information"

$$E[X_{n+1}\mid \underbrace{Y_1\cdots Y_n}_{\text{information up to } n,\ \text{"now"}}] \overset{?}{=} X_n$$

$E[X\mid Y]$ — Conditional expectation of X, given Y.

<u>Example</u>: $\begin{cases} X_1 \sim \text{Poisson}(\lambda_1) \\ X_2 \sim \text{Poisson}(\lambda_2) \\ X_1 \perp\!\!\!\perp X_2 \end{cases}$ $\Rightarrow$ $X_1 + X_2 \sim \text{Poisson}(\lambda_1+\lambda_2)$

$X_1 \perp\!\!\!\perp X_2$ indep. $X_i \sim \text{Poisson}(\lambda_i)$

$P(X_1 = m \mid X_1+X_2 = n)$, $\quad 0\leq m\leq n$

$$= \frac{P(X_1=m,\ X_1+X_2=n)}{\Pr(X_1+X_2=n)} = \frac{\overbrace{P(X_1=m)}^{\text{Poi}(\lambda_1)}\cdot \overbrace{\Pr(X_2=n-m)}^{\text{Poi}(\lambda_2)}}{\underbrace{\Pr(X_1+X_2=n)}_{\text{Poisson}(\lambda_1+\lambda_2)}}$$

$$= \binom{n}{m}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^m\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-m}$$

$$\sim \text{Biominal}\left(\frac{\lambda_1}{\lambda_1+\lambda_2},\ n\right)$$

$$\Rightarrow \underline{E[X_1\mid X_1+X_2=n] = n\,\frac{\lambda_1}{\lambda_1+\lambda_2}}$$

$E[X_1\mid X_1+X_2]$ is a r.v. taking values at $\dfrac{n\lambda_1}{\lambda_1+\lambda_2}$ with prob $(X_1+X_2=n)$

In general if X and Y are all discrete r.v.'s, say $X = 1,\dots m$, $Y = 1,\cdots n$

- If $X(\omega)$ is a r.v taking values with $\underline{p_i} = \Pr(X=i)$
- If $Y(\omega)$ is a r.v. taking values with $p_j = \Pr(Y=j)$

$Y: \Omega\to R$, $\quad X: \Omega\to R$ $\qquad$ Then $E[X\mid Y]$ is a r.v taking values at $E[X\mid Y=y_j]$ with prob. $P(Y=y_j)$

$$\begin{array}{ccccc} & & Y & & E[X\mid Y] \\ \omega_1 & \longrightarrow & y_1 & \longrightarrow & E[X\mid Y=y_1] \\ \vdots & & \vdots & & \vdots \\ \omega_m & \longrightarrow & y_m & \longrightarrow & E[X\mid Y=y_m] \end{array}$$

$$Y: \Omega \to R \longrightarrow R$$

$E[X|Y]$

- $E[X|Y]$ is a function of Y, is a r.v.
- $E[X|Y]$ is an estimate of X, given information from Y.
  is the "best estimate" of X, given Y. under L^2

Q: Are there any other norms for which $E[X|Y]$ is optimal ?

$$\boxed{\underset{g}{\arg\min}\, E[f(X,Y)] = E[X|Y]}$$ f: loss function

2006

necessary & sufficient condition: f - Bregman function

IEEE INFORMATION THEORY (2007-2008)

φ: convex function $\quad f_\varphi(x,y) = \varphi(x) - \varphi(y) - \varphi'(y)(x-y)$

$\varphi = x^2 \quad f_\varphi = (x-y)^2$

Properties of conditional expectation.

1) $E[g(X)|Y] \geq 0$ if $g(\cdot) \geq 0$
2) $E[a_1 g(X_1) + a_2 g(X_2) | Y] = a_1 E[g(X_1)|Y] + a_2 E[g(X_2)|Y]$
3) $E[g(X)|Y] = E[g(X)]$ if $X \perp\!\!\!\perp Y$
4) $E[g(X)|X] = g(X)$
5) $E[E[g(X)|Y]] = E[g(X)]$

Defn: $\{X_1 \cdots X_n \cdots\}$, $\{Y_1 \cdots Y_n \cdots\}$ stochastic processes

$\{X_n\}_{n\geq 1}^{\infty}$ is a martingale w.r.t $\{Y_n\}_{n\geq 1}^{\infty}$ if

a) $E[X_{n+1} | Y_1 \cdots Y_n] = X_n$ ("Fairness")

b) $E[|X_n|] < \infty$ for any n.

Example: ① simple symmetric random walk

$\{X_i\}$. i.i.d $\quad Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}$

$S_n = \sum_{i=0}^{n} X_i \quad S_0 = 0$

$\{S_n\}_{n=0}^{\infty}$ is a martingale w.r.t $\{X_n\}_{n=0}^{\infty}$ (or equivalently w.r.t $\{S_n\}_{n=0}^{\infty}$)

Proof: First, $E[|S_n|] \leq n < \infty$ for any n.

Second. $E[S_{n+1} | X_1 \cdots X_n] \stackrel{?}{=} S_n$

$$\begin{aligned} \text{L.H.S} &= E[\underline{S_{n+1} - S_n} + S_n | X_1 \cdots X_n] \\ &= E[X_{n+1} | X_1 \cdots X_n] + E[S_n | X_1 \cdots X_n] \\ &= \underset{0}{\underset{\parallel}{E[X_{n+1}]}} + S_n = \text{R.H.S.} \end{aligned}$$

Generally, ① $\{X_i\}$ i.i.d. $E[X_i]=0$ $E[|X_i|]<\infty$

$S_n = \sum_{i=0}^{n} X_i$, $S_0=0$ $\{S_n\}_{n\geq 0}^{\infty}$ is a martingale w.r.t $\{X_n\}_{n\geq 0}^{\infty}$

② $\{X_i\}$ i.i.d $E[X_i] = C_i$ $E[|X_i|]<\infty$

$Y_i = X_i - E[X_i]$

$\tilde{S}_n = \sum_{i=0}^{n} Y_i$ $S_0=0$ $\{\tilde{S}_n\}$ is a martingale w.r.t $\{X_n\}_{n\geq 0}^{\infty}$

Q: martingale vs M.C ?

Example: ① both martingale and M.C. ?

② martigale but not M.C. ?

③ M.C but not martingale ?

④ Neiter

Why: martingale ?

Example 1: Simple random walk

$S_n = \sum_{i=1}^{n} X_i$ $S_0 = X_0 = 0$, $P_r(X_i=1) = p = 1 - P_r(X_i=-1)$

Game Strategy $T_a = \inf\{ n>0, S_n = a)$ for $a>0$

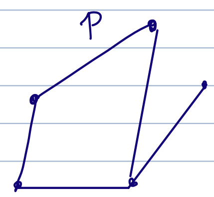
$T_{-b} = \inf\{ n>0$ $S_n = -b)$ for $b>0$

Questions. $P(T_a < T_{-b}) = ?$ (MGF can get some answer)

$E[T_a \wedge T_{-b}] = ?$

We will study martingale technique of optional sampling theorem

Example 2: $G(n,p)$ random graph n vertices



Chromatic number: the minimal number of color needed to color the vertices, so that any two vertices that are connected by an edge use different colors?

We will use martingale techniques of martingale inequality to derive the concentration inequality for the chromatic number around its mean.

Martingale: $\{X_n\}_{n\geq 1}$ w.r.t $\{Y_n\}_{n\geq 1}^{\infty}$

$\downarrow$ information

① $E[|X_n|] < \infty$ for any n.

② $E[X_{n+1} | Y_1 \cdots Y_n] = X_n$ (☆) martingale property (☆)

$\Downarrow$

L.H.S. is a function of $Y_1 \cdots Y_n$, implying X_n is a function of $Y_1 \cdots Y_n$

$X_n = E[X_n | Y_1 \cdots Y_n]$

$\Downarrow$

(☆☆) $E[X_{n+1} - X_n | Y_1 \cdots Y_n] = 0$ assuming X_n being a function of $Y_1 \cdots Y_n$

$\Updownarrow$

(☆)

Example: ① symmetric simple random walk $S_n = \sum_{i=1}^{n} X_i$ is a martingale

$\Downarrow$ generalizing

i.i.d $\{X_i\}$ $E[X_i] = \mu_i$ $\tilde{S}_n = \sum \tilde{X}_i$, $\tilde{X}_i = X_i - E[X_i]$

② $\{X_i\}$ i.i.d. $E[X_i] = 0$ (W.L.O.G.), $Var(X_i) = \sigma^2$

$\{S_n^2 - n\sigma^2\}$ is a martingale.

③ Simple random walk $S_n = \sum_{i=1}^{n} X_i$ $P_r(X_i = 1) = p = 1 - P_r(X_i = -1)$

If $p > \frac{1}{2}$, $Z_n = \left(\frac{1-p}{p}\right)^{S_n}$ is a martingale.

④ Doob's martingale

X is a r.v. $\{Y_1 \cdots Y_n\}$ is a random process. $E[|X|] < \infty$

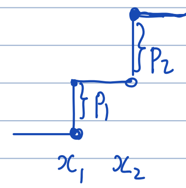
$X_1 = E[X | Y_1]$

$X_2 = E[X | Y_1, Y_2]$

$\vdots$

$X_n = E[X | Y_1 \cdots Y_n]$

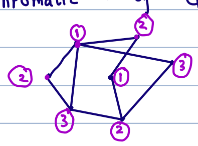
Claim: $\{X_i\}_{i\geq 1}^{\infty}$ is a martingale, w.r.t $\{Y_n\}_{n\geq 1}$



$x_1 < x_2 < \cdots$

Chromatic # of $G(n,p)$. $= Z_n$

Chromatic # $= 1$



Chromatic # $= 3$

For any fixed n, and p,

there are $m = \binom{n}{2}$ possible edges, $Y_1 \cdots Y_m$

$Y_i = \begin{cases} 0 & \text{edge } i \text{ not present} \\ 1 & \text{edge is is present} \end{cases}$

$\left.\begin{array}{l} X_1 = E[Z_n | Y_1] \\ \vdots \\ X_m = E[Z_n | Y_1 \cdots Y_m] = Z_n \end{array}\right\}$ Edge exposure martingale

---

Let us now show that Doob's construction from X & $\{Y_1 \cdots Y_n\}$ is a martingale. In particular

(a) $E[X_{n+1} | Y_1 \cdots Y_n] = X_n$

$\parallel$ L.H.S $\parallel$ R.H.S

$E\Big[\underline{E[X | Y_1 \cdots Y_{n+1}]} \Big| Y_1 \cdots Y_n\Big] = E[X | Y_1 \cdots Y_n]$

$\hookrightarrow$ tower property of conditional expectation

"Consistency property of conditional expectation"

(Remark: non-linear expectation theory)

?

(b) $E[|X_n|] < \infty$ (need Jensen's inequality)

$\hookrightarrow \infty > E[|X|] = E\Big[E[|X| | Y_1 \cdots Y_n]\Big] \overset{\text{Jensen}}{\geq} E\Big[\big|E[X | Y_1 \cdots Y_n]\big|\Big] = E[|X_n|]$

$\phi(x) = |x|$

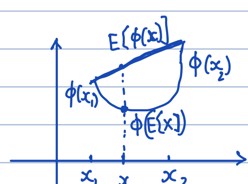
Doob's Construction

Recall: (Jensen's inequality)

Suppose $\phi: R \to R$ is convex, then

1) $E[\phi(X)] \geq \phi(E[X])$

2) $E[\phi(X) | Y_1 \cdots Y_n] \geq \phi(E[X | Y_1 \cdots Y_n])$



$E[X]$ for $X \sim \{(x_i, p_i)\}_{i\geq 1}$

$E[X] = \sum p_i x_i$

$X \sim f_X(\cdot)$ $E[X] = \int x f_X(x) dx$

$\left(F_X(a) = \int_{-\infty}^{a} f_X(x) dx \right)$

$= \int x \, dF_X(x)$

$G(x)$ is non-decreasing

$G(x) = G_{Ac}(x) + G_{jump}(x)$

$E[h(X)] = \int_{-\infty}^{\infty} h(x) \, dG_{Ac}(x)$

$$+\int_{-\infty}^{\infty} h(x)\, d\, G_{jump}(x)$$

$E[|X|]<\infty \qquad X = X^{+} - X^{-}$

$E[X]$ is well defined $\Leftrightarrow$ $E[X^{+}]<\infty$ and $E[X^{-}]<\infty$ or one of them is finite

Property of martingale $\{X_n\}_{n\geq 1}^{\infty}$

$(*)$

$$E[X_1] = \cdots E[X_n] \quad \forall n. \quad \text{(Constant expectation)}$$

Proof: $E[E[X_{n+1} | Y_1 \cdots Y_n]] = E[X_{n+1}]$

$\parallel$ by defn.

$E[X_n]$

Remark: $(\text{✡}) \Rightarrow (*)$

$(*) \not\Rightarrow (\text{✡})$

Counter Example: $Pr(X_n = 1) = \frac{1}{2} = Pr(X_n = -\frac{1}{2})$ i.i.d

$E[X_n] = 0 \qquad E[X_n | X_1 \cdots X_{n-1}] = E[X_n] = 0$

$\neq X_{n-1}$

---

martingale theory $<$ Concept of martingale / Concept of stopping time

τ: Markov time / Stopping time w.r.t information provided by $\{Y_1 \cdots Y_n \cdots\}$

① τ is a r.v taking values at $1, \cdots 2 \cdots$

② $\{\tau = n\}$ $(\{\tau \leq n\}, \{\tau \geq n\})$ can be determined by $\{Y_1 \cdots Y_n\}$.

Example: Simple random walk $\tau_a = \inf\{n \geq 0 \quad S_n \geq a\}$ is stopping time — first hitting time

Counter-example: $\tau_a = \sup\{n \geq 0, \ S_n \geq a\}$ last exit time

is NOT a stopping time

martingale + stopping time w.r.t $\{\mathcal{F}_n\}$

$\{X_n\}$ τ ↳ information up to n.

$\parallel$ represented

$\{X_1 \cdots X_n\}$ or $\{Y_1 \cdots Y_n\}$

optional sampling theorem

idea: $\{X_n\}$ – martingale

$\Downarrow$

$E[X_1] \cdots = E[X_n]$

$\nearrow \tau$ simple random walk $\tau_a = \inf\{n \geq 0, S_n \geq a\}$, $X_0 = 0$ $a > 0$.

$\Longrightarrow$ $E[X_\tau] = ?$ under what conditions $E[X_\tau] \overset{?}{=} E[X_i]$ for martingale $\{X_n\}$ & τ ?

E.g. $E[S_{\tau_a}] = E[a] \neq E[X_i]$ in general

What is the problem?

1) X_τ is different from X_n, τ is a stopping time

E.g. X_i i.i.d $S_\gamma = \sum_{i=1}^{\gamma} X_i$, γ is a Poisson (λ), $\perp\!\!\!\perp X_i$

Compound poisson

Wald's identity $\Longrightarrow$ $E[S_\gamma] = E[\gamma]E[X_i]$

different from $S_n = \sum_{i=1}^{n} X_i$

$E[S_n] = nE[X_i]$

idea: $P(\tau < \infty) \overset{?}{=} 1$

$E[\tau] \overset{?}{<} \infty$ NOT equivalent (exercise)

Why?

Suppose we have OST holds for some $\{X_n\}$ & τ. i.e. $E[X_\tau] = E[X_i]$

$\downarrow$ martingale $\downarrow$ stopping time

E.g. Simple symmetric random walk

$\left(P_r(X_i = 1) = P_r(X_i = -1) = \frac{1}{2}, \quad S_n = \sum_{i=1}^{n} X_i \quad S_0 = X_0 = 0 \right)$

$\downarrow$ martingale

$\tau_{-b} = \inf\{n \geq 0, X_n \leq -b\}$ $b, a > 0$

$\tau_a = \inf\{n \geq 0 \quad X_n \geq a\}$

$\tau = \inf\{\tau_a, \tau_b\} = \tau_a \wedge \tau_{-b}$. $\longrightarrow$ stopping time

$P_r(\tau_{-b} < \tau_a) = P_r(\tau = \tau_{-b})$ – ruin prob.

Suppose OST holds $\Rightarrow$ $E[S_\tau] = E[S_i] = 0$

$\Downarrow$

$(-b)P_r(\tau = \tau_{-b}) + a\underbrace{P_r(\tau = \tau_a)}_{1 - P_r(\tau = \tau_{-b})} = 0$

$P_r(\tau = \tau_{-b}) = \frac{a}{a+b}$.

$E[\tau] = ?$ $S_n^2 - n$ is a martingale

Suppose OST holds $E[S_\tau^2 - \tau] = E[S_0^2 - 0] = 0$

$\Downarrow$

$E[\tau] = E[S_\tau^2] = (-b)^2 P_r(\tau = \tau_{-b}) + a^2 P_r(\tau = \tau_a)$

$= b^2 \frac{a}{a+b} + a^2 \frac{b}{a+b} = ab$.

X_τ $\tau = 1, 2, \ldots n, \ldots$

$\boxed{\{X_{\tau \wedge n}\}_{n=1}^{\infty}}$ $\tau \wedge n$ $= \min(n, \tau)$ $\{X_n\}_{n \geq 1}^{\infty}$ martingale

Claim. $\downarrow$ stopped martingale

If we accept the "fact" that $\{X_{\tau \wedge n}\}$ is a martingale $\Rightarrow E[X_{\tau \wedge n}] = E[X_n]$

$\lim_{n \to \infty} E[X_{\tau \wedge n}] \overset{?}{=} E[\lim_{n \to \infty} X_{\tau \wedge n}]$

$\parallel$ OST $\Downarrow$

$E[X_n] \overset{?}{=} E[X_\tau]$

Thm (OST version 1)

Suppose $\{X_n\}$ is a martingale and τ is a stopping time (w.r.t $\tilde{\mathcal{F}}$)

If 1) $P(\tau<\infty)=1$ $\quad (P(\tau=\infty)=0)$

2) $E(|X_\tau|)<\infty$

3) $\lim_{n\to\infty} E[X_n 1_{\{\tau>n\}}]=0$

Then $E[X_\tau]=E[X_0]$.

Intermediate Step of $E[X_{\tau\wedge n}]$ – as a stopped-martingale.

$\mathcal{F}_k=\{Y_1\cdots Y_k\}$

Thm: If $\{X_n\}$ is a martingale, τ is a stopping time (w.r.t $\tilde{\mathcal{F}}$)

Then $\{X_{\tau\wedge n}\}$ is a (stopped)-martingale. In particular $E[X_{\tau\wedge n}]=E[X_n]$. $(*)$

Proof of $(*)$. Recall $\tau\wedge n = \min(\tau,n)=\begin{cases}\tau & \tau<n\\ n & \tau\geq n\end{cases}$

LHS $E[X_{\tau\wedge n}] = E[X_\tau 1_{\{\tau<n\}}] + E[X_\tau 1_{\{\tau\geq n\}}]$

?

$= E[X_\tau 1_{\{\tau<n\}}] \qquad E[X_n 1_{\{\tau\geq n\}}]$

$=\sum_{k=1}^{n-1} E[X_k 1_{\{\tau=k\}}]$

helps me to understand condition 3)

RHS $E[X_n] = E[X_n 1_{\{\tau<n\}}] + E[X_n 1_{\{\tau\geq n\}}]$

?

$=\sum_{k=1}^{n-1} E[X_n 1_{\{\tau=k\}}]$

To see why (?) holds,

$$E[X_n 1_{\{\tau=k\}}] \overset{?}{=} E[X_k 1_{\{\tau=k\}}] \quad \checkmark$$

$$= E\big[E[X_n 1_{\{\tau=k\}} \mid Y_1\cdots Y_k]\big] \qquad = E\big[E[X_k 1_{\{\tau=k\}} \mid Y_1\cdots Y_k]\big]$$

$$E\big[1_{\{\tau=k\}} E[X_n \mid Y_1\cdots Y_k]\big] \overset{?}{=} E\big[1_{\{\tau=k\}} E[X_k \mid Y_1\cdots Y_k]\big]$$

Lemma $\checkmark$

Lemma: $\{X_n\}$ is a martingale w.r.t $\{Y_1\cdots Y_n\}$ $\iff$ $E[X_n \mid Y_1,\cdots Y_{n-1}]=X_{n-1}$

To see this,

$$E[E[X_n|Y_1\cdots Y_{n-1}]|Y_1\cdots Y_{n-2}]$$

Tower property $\parallel$

$$= E[X_{n-1}|Y_1\cdots Y_{n-2}] = X_{n-2}$$

$$E[X_n|Y_1\cdots Y_{n-2}]$$

$$E[X_n|Y_1\cdots Y_k] = X_k$$

$k<n.$

$\{X_n\}$ – martingale

$\{\tau\}$ – stopping time

$(Y_1 \cdots Y_n \cdots)$ — information (filtration) → OST, → martingale inequality

OST $\quad E[X_\tau] \overset{?}{=} E[X_i]$ (interchange of $\lim$ & E)

$$\{X_n\} \longrightarrow \{X_{\tau\wedge n}\} \longrightarrow \{X_\tau\}$$

$\{X_n\}$: martingale; $\{X_{\tau\wedge n}\}$: stopped martingale

"key step" $\quad E[X_n | Y_1 \cdots Y_k] \qquad E[X_k | Y_1 \cdots Y_k] = X_k$

X_k is a function of $Y_1 \cdots Y_k$

using equivalent defn of martingale

Conditioning & $E[X|Y]$ properties

Classical prob. "independence" i.i.d identical and independence

In the martingale world, "Conditional equivalence"

Recall

OST (version 1)

1) $P(\tau < \infty) = 1$

2) $E[|X_\tau|] < \infty$

3) $\lim_{n\to\infty} E[X_n 1_{\{\tau \geq n\}}] = 0 \quad (*)$ last time

$\Rightarrow E[X_\tau] = E[X_1] = E[X_n] \quad n \geq 1$

Proof: $\quad E[X_\tau] = E[X_\tau 1_{\{\tau < n\}}] + E[X_\tau 1_{\{\tau \geq n\}}] \qquad \forall n$

$$\Rightarrow \lim_{n\to\infty} E[X_\tau] = \underline{\lim_{n\to\infty} E[X_\tau 1_{\{\tau<n\}}]} + \boxed{\lim_{n\to\infty} E[X_\tau 1_{\{\tau \geq n\}}]} \overset{?}{\to} 0$$

$$= \underline{\lim_{n\to\infty} E[X_{\tau\wedge n} 1_{\{\tau<n\}}]}_{\Delta\Delta}$$

$$\lim_{n\to\infty} E[X_n] = \underline{\lim_{n\to\infty} E[X_n 1_{\{\tau<n\}}]}_{\Delta} + \lim_{n\to\infty} E[X_n 1_{\{\tau \geq n\}}]$$

by 3)

b/c $X_{\tau\wedge n}$ is a stopped martingale $(\Delta) = (\Delta\Delta)$

It is therefore sufficient to $\quad \lim_{n\to\infty} E[X_\tau 1_{\{\tau\geq n\}}] = 0 \quad (\ast)$

Apply Lemma below by $W := X_\tau$

Lemma: W be a random variable $E|W| < \infty$.

And $P(\tau < \infty) = 1$ Then

$$\lim_{n\to\infty} E[W 1_{\{\tau \geq n\}}] = 0$$

To see this lemma:

$$\infty > E[|W|] = \lim_{n\to\infty} \sum_{k=1}^{n} E[|W| 1_{\{\tau = k\}}]$$

$$\Rightarrow \lim_{n\to\infty} E[|W| 1_{\{\tau > n\}}] \to 0$$

$$\Longrightarrow \lim E[W 1_{\{\tau > n\}}] \to 0$$

Remark: When applying O.S.T. (previous examples from last lecture), checking Conditions 1) 2) 3) are the key.

---

martingale inequality $\begin{cases} \text{martingale} \Rightarrow \text{super, sub-martingale} \\ \text{stopping time} \end{cases}$

Defn: $\{X_n\}_{n\geq 1}^{\infty}$ and $\{Y_n\}_{n\geq 1}$ are two stochastic processes. We say that $\{X_n\}_{n\geq 1}^{\infty}$ is a super, (sub) martingale w.r.t $\{Y_n\}_{n\geq 1}^{\infty}$ if

1) X_n is a function of $Y_1 \cdots Y_n$

2) $E[X_{n+1} | Y_1 \cdots Y_n] \overset{(\geq)}{\leq} X_n$

3) $E[\min(0, X_n)] > -\infty, \forall n.$ $\left(E[\max(0, X_n)] < \infty \ \forall n\right).$

Construction of sub (super)-martingale is based on the

<u>Prop.</u> $\{X_n\}_{n\geq 1}^{\infty}$ - martingale w.r.t $\{Y_n\}$. Assume $\phi(\cdot)$ is a convex function.
Then $\{\phi(X_n)\}_{n\geq 1}$ is a sub-martingale, assuming $E[\max(0, \phi(X_n))] < \infty$. $\forall n$.

(Proof directly follows from Jensen's)

Example: $\phi(x) = |x|$, $\phi(x) = x^2$, $\phi(x) = e^x$ are all convex.
$\{|X_n|\}_{n\geq 1}$ $\{X_n^2\}_{n\geq 1}$ $\{e^{X_n}\}_{n\geq 1}$ are all sub-martingale, if $\{X_n\}_{n\geq 1}$ is martingale,
subject to appropriate integrability conditions.

Remark: 1) Super $X_n \longleftrightarrow$ sub $-X_n$

2) martingale is both super and sub-martingale

3) All the existing properties established so far apply to super, sub-martingale,
with "=" replaced by "$\leq$" or "$\geq$"
$\qquad\qquad\qquad\qquad\qquad\qquad$ Super $\qquad$ Sub.

---

martingale inequality:

Recall:

Markov inequality: X is a r.v. $X \geq 0$. Then $Pr(X \geq \lambda) \leq \frac{E[X]}{\lambda}$. for any $\lambda > 0$

Proof. $P_r(X \geq \lambda) \overset{?}{\leq} \frac{E[X]}{\lambda} \Leftrightarrow E[X] \overset{?}{\geq} \lambda Pr(X \geq \lambda)$

$E[X] \geq E[X 1_{\{X\geq\lambda\}}] \geq E[\lambda 1_{\{X\geq\lambda\}}] = \lambda E[1_{\{X\geq\lambda\}}]$

Maximal Inequality: $\{X_n\}_{n\geq 1}^{\infty}$ is a sub-martingale. $X_i \geq 0$ Then for any given $n > 0$

$$Pr\left(\max_{1\leq k\leq n} X_k > \lambda\right) \leq \frac{E[X_n]}{\lambda} \qquad \forall \lambda > 0.$$

Hint: designing appropriate stopping time τ and use Markov inequality.

Martingale inequality $\longleftrightarrow$ Classical inequality in Prob.

Maximal Inequality $\longleftrightarrow$ Markov inequality

$\{X_n\}$ sub-martingale, $X_i \geq 0$

$$P\left(\max_{1\leq k\leq n} X_k > \lambda\right) \leq \frac{E[X_n]}{\lambda}$$

X, $X \geq 0 \Rightarrow P(X \geq \lambda) \leq \frac{E[X]}{\lambda}$

$$\lambda \Pr(X \geq \lambda) = \underbrace{E[1_{\{X\geq\lambda\}} X]}$$

$$\underbrace{\lambda \Pr\left(\max_{1\leq k\leq n} X_k > \lambda\right)}_{\substack{\| \\ E\left[\lambda 1_{\{\max_{1\leq k\leq n} X_k > \lambda\}}\right]}} \leq E[X_n]$$

$\geq$ some appropriate stopping time $E[X_{\tau\wedge n}]$

$$\geq E\left[X_{\tau\wedge n} 1_{\{\max X_k > \lambda\}}\right]$$

$\frac{1}{1-\alpha} t^{\alpha-1}$

Proof define stopping time τ such that

$$\tau = \begin{cases} \min\{k, X_k \geq \lambda\} & \text{if for some } k = 1, \cdots n \\ n & \text{otherwise} \end{cases}$$

a) In the case of there is no such $k \leq n$ for which $X_k \geq \lambda$

$$\lambda \Pr\left(\max_{0\leq k\leq n} X_k \geq \lambda\right) = 0 \leq E[X_n] \quad \checkmark$$

b) If there exists some $k_0 \leq n$, for which $X_{k_0} \geq \lambda$

$$E[X_n] \geq E[X_{\tau\wedge n}] \geq E\left[X_{\tau\wedge n} 1_{\{\max X_k \geq \lambda\}}\right] \geq \lambda E\left[1_{\{\max_{1\leq k\leq n} X_k \geq \lambda\}}\right] = \lambda \Pr\left(\max_{1\leq k\leq n} X_k > \lambda\right)$$

$\checkmark$

Azuma Inequality $\longleftrightarrow$ Chernoff's bound $\quad \lambda > 0 \quad$ #

If $\{S_n\}$ is any martingale

Such that $|S_i - S_{i-1}| \leq 1$

C.L.T. (Review)

$\{X_i\}$ i.i.d. $E[X_i] = 0$ $\quad \underline{Var(X_i) = 1}$

$S_n = \sum_{i=1}^{n} X_i \qquad \frac{S_n}{\quad} \xrightarrow{n\to\infty} N(0, 1)$

and $E[S_n]=0$. Then

$$\Pr\left(\frac{S_n}{\sqrt{n}} > \lambda\right) \le e^{-\frac{\lambda^2}{2}},$$

for $\forall \lambda > 0$, $\forall n$.

$$\frac{\quad}{\sqrt{n}} \to N(0,1)$$

$$\lim_{n\to\infty} \Pr\left(\frac{S_n}{\sqrt{n}} > \lambda\right) = \frac{1}{\sqrt{2\pi}} \int_\lambda^\infty e^{-\frac{x^2}{2}} dx$$

?? $< e^{-\frac{\lambda^2}{2}}$

Note A

Remark: martingale difference

$X_i \triangleq S_i - S_{i-1}$

$E[X_i]=0$, $|X_i| \le 1$

Chernoff has stronger Conditions vs Azuma b/c $\{X_i\}$ i.i.d implies $\{S_n\}$ martingale.

Chernoff

Suppose $\underline{\{X_i\} \text{ i.i.d}}$ $E[X_i]=0$, $\underline{|X_i| \le 1}$. $S_n = \sum_{i=1}^n X_i$

$$\Pr\left(\frac{S_n}{\sqrt{n}} > \lambda\right) \le e^{-\frac{\lambda^2}{2}}, \quad \forall n.$$

Note A

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} dx = \frac{1}{2}$$

$$\Uparrow$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{x^2}{2}} dx$$

$$\frac{1}{\sqrt{2\pi}} \int_\lambda^\infty e^{-\frac{x^2}{2}} dx \le \begin{cases} \le \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} dx = \frac{1}{2} < e^{-\frac{\lambda^2}{2}} & \lambda < 1 \\ \frac{1}{\sqrt{2\pi}} \int_\lambda^\infty x e^{-\frac{x^2}{2}} dx \le e^{-\frac{\lambda^2}{2}} & \lambda > 1 \end{cases}$$

Proof of Chernoff's bound.

$$\Pr\left(\frac{S_n}{\sqrt{n}} > \lambda\right) \overset{?}{\leq} e^{-\frac{\lambda^2}{2}} \quad \forall \lambda > 0$$

$$\Updownarrow$$

$$\Pr\left(S_n \geq \sqrt{n}\,\lambda\right)$$

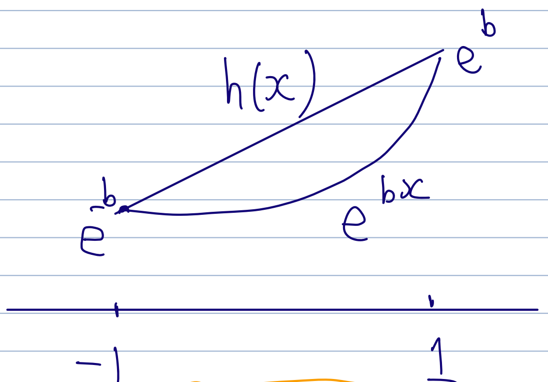
Consider $\Pr(S_n \geq a) \quad \forall a > 0, \ \forall b > 0$

$$= \Pr\left(e^{bS_n} \geq e^{ba}\right)$$

Markov Inequality $\leq \dfrac{E[e^{bS_n}]}{e^{ba}}$ $\quad$ X_i i.i.d, $S_n = \sum_{i=1}^{n} X_i$ $\quad$ $\dfrac{\left(E[e^{bX_1}]\right)^n}{e^{ba}}$

Note: linearization trick

$$E[e^{bX_1}] \leq E[h(X_1)] = h(E[X_1]) = h(0)$$



$$h(x) = \frac{e^{b}+e^{-b}}{2} + \frac{e^{b}-e^{-b}}{2}\,x$$

$E[X_1] = 0$

$|X_1| \leq 1$

$$\leq \frac{\left(h\,E[X_1]\right)^n}{e^{ba}} = \frac{\left(h(0)\right)^n}{e^{ba}}$$

$$= \frac{\left(\frac{e^{-b}+e^{b}}{2}\right)^n}{e^{ba}} \leq \frac{e^{+\frac{b^2 n}{2}}}{e^{ba}} \quad \forall b, a > 0$$

[Note: $\frac{e^{b}+e^{-b}}{2} \leq e^{+\frac{b^2}{2}}$ (by Taylor expansion)]

Since b is arbitrage

$$\Pr(S_n > a) \leq \min_{b>0} e^{\frac{b^2 n}{2} - ba} = e^{-\frac{a^2}{2n}}$$

$$\underset{a=\sqrt{n}\lambda}{=\!=} e^{-\frac{\lambda^2}{2}}.$$

Proof of Azuma with $X_i = S_i - S_{i-1}$

where $\{S_n\}$ - martingale, (i.e. X_i are not necessarily i.i.d.)

Chernoff's bound

- X_i i.i.d.
- $E[X_i] = 0$ $\qquad S_n = \sum_{i=1}^{n} X_i$
- $|X_i| \le 1$

Then $\Pr\left(\frac{S_n}{\sqrt{n}} > \lambda\right) \le e^{-\frac{\lambda^2}{2}}, \quad \lambda > 0$

Proof: $\Pr(S_n > a) = \Pr(e^{bS_n} > e^{ba})$

$$\le \frac{E[e^{bS_n}]}{e^{ba}} \quad ✓$$

i.i.d

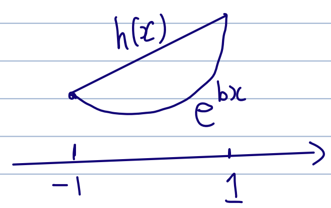
$$= \frac{(E[e^{bX_i}])^n}{e^{ba}}$$

$E[X_i]=0$

$|X_i| \le 1$

$$\le \frac{(h(0))^n}{e^{ba}} \qquad (*)$$

with $h(x) = \frac{e^b + e^{-b}}{2} + \frac{e^b - e^{-b}}{2} x$



Azuma's inequality

- $\{S_i\}$ martingale
- $|S_i - S_{i-1}| \le 1$ $\qquad S_0 = 0$

Then $\Pr\left(\frac{S_n}{\sqrt{n}} \ge \lambda\right) \le e^{-\frac{\lambda^2}{2}} \quad \forall \lambda > 0$

$$\underbrace{E[S_n \mid S_1 \cdots S_{n-1}] = S_{n-1}}_{\text{martingale}} \Leftrightarrow E[\underbrace{S_n - S_{n-1}}_{X_n} \mid S_1 \cdots S_{n-1}]$$

Proof

$$\Pr(S_n > a) \le \frac{E[e^{bS_n}]}{e^{ba}} \qquad S_n = X_n + S_{n-1}$$

$$= \frac{E[E[e^{bS_{n-1} + X_n} \mid S_1 \cdots S_{n-1}]]}{e^{ba}}$$

$$= \frac{E[e^{bS_{n-1}} E[e^{bX_n} \mid S_1 \cdots S_{n-1}]]}{e^{ba}}$$

Now our goal is

$$E[e^{bX_n} \mid S_1 \cdots S_{n-1}]$$

(Note $|X_n| = |S_n - S_{n-1}| < 1$

$E[X_n] = 0$)

mimicking Chernoff's bound

$$\le E[h(X_n) \mid S_1 \cdots S_{n-1}]$$

$$= h(E[X_n \mid S_1 \cdots S_{n-1}])$$

martingale property

$$= h(0)$$

Now repeating $E[e^{bS_n} E[e^{bX_n} \mid S_1 \cdots S_{n-1}]]$

$$\le (h(0))^n \qquad (*)$$

#

Applications of martingale inequalities / Concentration inequality

Chromatic # $\underbrace{f(G)}$ $\quad G(n,p)$

Recall: for any n & p, there are at most $\binom{n}{2} = m$ edges.

$I_1 \cdots I_m$

Constructing an edge-exposure martingale

$$\left.\begin{aligned} Z_1 &= E[\, f(G) \mid I_1] \\ &\vdots \\ Z_m &= E[\, f(G) \mid I_1 \cdots I_m] = f(G) \end{aligned}\right\} \quad \{Z_i\} \text{ martingale}$$

Clearly $|Z_i - Z_{i-1}| \le 1$

$E[Z_i] \neq 0$

$X_i \triangleq Z_i - E[Z_i]$

1) $\{X_i\}$ is martingale

2) $|X_i - X_{i-1}| \le 1$

3) $E[X_i] = 0$

Goal: $\underbrace{\Pr\big(f(G) - E[f(G)] \big) > \lambda}$

Concentration

Applying Azuma

$$\Pr\left(\frac{X_m}{\sqrt{m}} > \frac{\lambda}{\sqrt{m}} \right) = \Pr\left(\frac{Z_m - E[Z_m]}{\sqrt{m}} > \frac{\lambda}{\sqrt{m}} \right)$$

$$\boxed{= \Pr\left(\frac{f(G) - E[f(G)]}{\sqrt{m}} > \frac{\lambda}{\sqrt{m}} \right)}$$

$\lambda^2 \qquad \lambda^2$

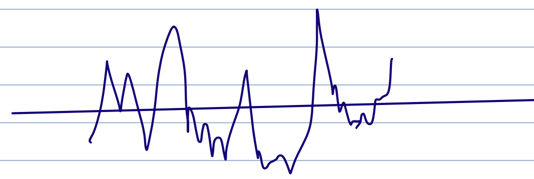
$$\le e^{-\frac{\cdots}{2m}} = e^{-\frac{\cdots}{n(n-1)}}$$

#

General martingale theory $\Rightarrow$ Classical examples $\begin{cases} \text{random walk} \\ \Downarrow \\ \text{B.M.} \checkmark \\ \text{Poisson Process} \\ \Downarrow \\ \text{point process} \checkmark \end{cases}$

<u>Brownian motion</u> : $(W_t)_{t\ge0}$, $(B_t)_{t\ge0}$

(Winer Process)



- Biological origin
- (1900) L. Bachelier
- (1905) A. Einstein
- Wiener.

- Rescaled random walk

$$Pr(X_i=1)=\frac{1}{2}=Pr(X_i=-1) \quad \text{i.i.d}$$

$$Y_k=\sum_{i=1}^{k} X_i$$

For any fixed $t>0$, take a fixed n, $[nt]$ = largest integer $\le nt$

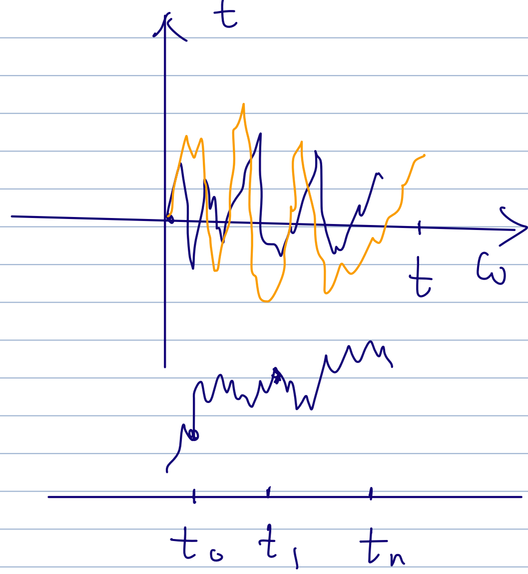
$$W^{(n)}=Y_{[nt]}/\sqrt{n}=\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]} X_i, \quad W(t)=\lim_{n\to\infty}\frac{1}{\sqrt{n}}Y_{[nt]}\to N(0,t)$$

B.M. $(\Omega, \mathcal{F}, \mathbb{P})$ Given. For $\omega\in\Omega$, there is a continuous function $B(t)$ $(t\ge0)$ such that $B(0)=0$ and $B(t,\omega)$ depends on ω. $B(t,\omega)$ is called B.M if $\forall\, 0=t_0<\cdots t_n$

$B(t,\omega)$

a) independent increment

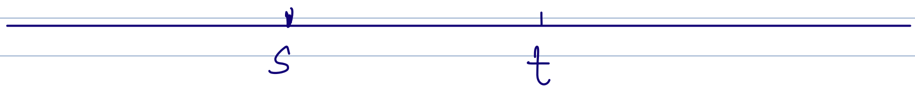
$$B(t_i) - B(t_{i-1}) \perp\!\!\!\perp B(t_{i+1}) - B(t_i)$$



b) stationary increment

$$B(t_i) - B(t_{i-1}) \sim \underline{N(0, t_i - t_{i-1})}$$

i.e. $B(t) - B(s) \sim N(0, t-s)$



Note: If without continuous path, Poisson (λt) has stationary and independent increment.

Review: $B(t)$, $t \geq 0$, $B_0 = 0$

1) Independent increment $\quad B(t) - B(s) \perp\!\!\!\perp B(u) - B(t)$ $\quad s < t < u$

2) Stationary increment $\quad B(t) - B(s) \sim \underline{N(0, t-s)}$

B.M is a rescaled random walk

For every t

$$W^{(n)} = \frac{1}{\sqrt{n}} Y_{[nt]} \qquad Y_{[nt]} = \sum_{i=1}^{[nt]} X_i \qquad Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}$$

<u>Proposition</u> $\quad W^{(n)} \xrightarrow{n \to \infty} N(0,t)$ in distribution.

<u>Proof</u>: By moment generating function

(Hw to check)

$$E\left[e^{\lambda \frac{1}{\sqrt{n}} Y_{[nt]}}\right] \xrightarrow{n \to \infty} E\left[e^{\lambda N(0,t)}\right] = \underline{e^{\frac{1}{2}\lambda^2 t}}, \qquad \lambda > 0$$

$\|$ i.i.d for simplicity $[nt] = nt$

$$E\left[e^{\lambda \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} X_i}\right]$$

$$\| $$

$$\prod_{i=1}^{nt} E\left[e^{\lambda \frac{1}{\sqrt{n}} X_i}\right] \underset{\text{Def of } X_i}{=} \prod_{i=1}^{nt} \left[\frac{1}{2} e^{\frac{\lambda}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{\lambda}{\sqrt{n}}}\right]$$

$$= \underline{\left(\frac{1}{2} e^{\frac{\lambda}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{\lambda}{\sqrt{n}}}\right)^{nt}}$$

It suffices to $\quad \underset{\downarrow n\to\infty \atop \infty}{nt} \ln \underbrace{\left(\frac{1}{2} e^{\frac{\lambda}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{\lambda}{\sqrt{n}}}\right)}_{\Downarrow \text{ as } n \to \infty \atop 0} \overset{?}{\longrightarrow} \frac{1}{2}\lambda^2 t$

This can be verifies by L'Hopitâl rule.

$E[X|Y]$ Review: evaluation of X, given y_i

$Y = y_1 \cdots y_n$ $\quad$ r.v, taking values at $E[X|Y=y_i]$, with $P_r(Y=y_i)$.

To define $E[X|Y]$ in general, we need σ-algebra and "measurability".

Intuition: Flip a coin twice

$$X_1 = \begin{cases} 1 & \text{if the first flip head} \\ 0 & \text{otherwise} \end{cases} \qquad A_1 = (HH, HT) \quad A_0 = (TH, TT)$$

$$X_2 = \begin{cases} 1 & \text{if the second flip is head} \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{A}_1 = (TH, HH) \quad \tilde{A}_0 = (HT, TT)$$

$$Z = X_1 + X_2 = \begin{cases} 2 & (HH) \\ 1 & (HT, TH) \\ 0 & (TT) \end{cases}$$

Information in Z is NOT rich enough in determining X_1 or X_2.

σ-algebra. $\mathcal{F}$

Defn: σ-algebra $\mathcal{F}$ is a collection of subsets of Ω, with the conditions

a) $A \in \mathcal{F} \Leftrightarrow A^c \in \mathcal{F}$

b) $\phi \in \mathcal{F} \Leftrightarrow \Omega \in \mathcal{F}$

c) $A_i \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

a) + c) $\Rightarrow A^i \in \mathcal{F}$, $\bigcap_{i=1}^{\infty} A_i^c \in \mathcal{F}$.

In particular, σ-algebra generated by random variable $\sigma(X)$

Example: flip a coin twice $\quad \Omega =$ collection of all samples

$$X_1 = \begin{cases} 1 & \text{if the first flip is head} \\ 0 & \text{otherwise} \end{cases}$$

$A_1 = (HH, HT)$

$A_2 = (TH, TT) = A_1^c$

$$\sigma(X_1) = \phi, \Omega, A_1, A_2, A_1 \cup A_2$$
$$= \{(HH, HT), (TH, TT), \phi, \Omega\}$$

$X = x_1 \cdots x_n$

$A_i = X^{-1}(x_i)$

In general: $\sigma(X) =$ from A_i following rules, such as

$$\{A_i, A_i^c, \cup A_i, \cap A_i \cdots\}$$

Now, $\sigma(Z) = \{\phi, (HH)\ (TT), (HT, TH), \Omega\}$

$\phi(X_1) = \{A_1 = (HH, HT)\ \ \phi\ \ \Omega\}$

$A_0 = (TH, TT)$

Observing $\sigma(Z) \neq \sigma(X_1) \neq \sigma(X_2)$, moreover, knowing $X_i \not\Rightarrow$ predict X_j or Z
know $Z \not\Rightarrow$ predict X_1, X_2.

Defn: Let X be a r.v. on Ω.

Let $\mathcal{G}$ be a given σ-algebra on some subsets of Ω.

We say that X is $\mathcal{G}$-measurable, if for any $A \in \sigma(X)$, we have $A \in \mathcal{G}$.

Def $E[X|Y]$ in the language of σ-algebra

Let X be a r.v on some prob. space $(\Omega, \mathcal{F}, \mathbb{P})$, on which $E[|X|] < \infty$.

Let $\mathcal{A}$ be a sub-σ-algebra of $\mathcal{F}$. (That is, $\forall A \in \mathcal{A}$, we have $A \in \mathcal{F}$).

Then $E[X|\mathcal{A}]$ is a r.v satisfying

1) $E[X|\mathcal{A}]$ is measurable w.r.t $\mathcal{A}$. (recall $E[X|Y]$ is a function of Y)

2) $E[X 1_A] = E[E[X|\mathcal{A}] 1_A]$ for all $A \in \mathcal{A}$.

"Translation" "Equivalent to"

<1> Y is a r.v. such that Y is measurable w.r.t $\mathcal{A}$

<2> $\int_A Y \, dP(\omega) = \int_A X \, dP(\omega)$ for any $A \in \mathcal{A}$

$\Leftrightarrow E[Y 1_A] = E[X 1_A]$.

Then Y is called conditional expectation of X w.r.t $\mathcal{A}$,
& written as $E[X|\mathcal{A}]$.

Q: Existence? Yes, need Radon-Nikodyn derivatives
Uniqueness? Yes.

Proof of uniqueness.

Suppose Y_1 and Y_2 are two r.v's satisfying <1> and <2>,

We need to show $Y_1 = Y_2$ in prob. $P(A) = 0$ on $A = \{\omega : Y_1(\omega) \neq Y_2(\omega)\}$.

proof by contradiction. Suppose $P(A) \neq 0$.

$A_1 \subset A, \quad A_1 = \{\omega : Y_1 - Y_2 > 0\} \quad A_1 \in \mathcal{A}.$

$$\int_{A_1} Y_1 \, dp(\omega) = \int_{A_1} Y_2 \, dp \iff \int_{\underbrace{A_1}_{P(A_1) > 0}} \underbrace{Y_1 - Y_2}_{> 0} \, dp = 0 \text{ , impossible.}$$

Properties of $E[X|\mathcal{A}]$

1) $E[aX_1 + bX_2 | \mathcal{A}] = E[aX_1 | \mathcal{A}] + E[bX_2 | \mathcal{A}]$

2) $E[X|\mathcal{A}] = X$ for every $\mathcal{A}$-measurable X

3) $E[XZ|\mathcal{A}] = XE[Z|\mathcal{A}]$ if X is $\mathcal{A}$-measurable

4) $E[X] = E[E[X|\mathcal{A}]]$

5) $E[X|\mathcal{H}] = E[E[X|\mathcal{G}]|\mathcal{H}]$ if $\mathcal{H} \subset \mathcal{G}$. (i.e. if $A \in \mathcal{H}, \Rightarrow A \in \mathcal{G}$)

$E[X|\mathcal{A}]$ for σ-algebra $\mathcal{A}$.

Properties of $E[X|\mathcal{A}]$

1) $E[aX_1+bX_2|\mathcal{A}] = E[aX_1|\mathcal{A}] + E[bX_2|\mathcal{A}]$

2) $E[X|\mathcal{A}] = X$ for every $\mathcal{A}$-measurable X

3) $E[XZ|\mathcal{A}] = XE[Z|\mathcal{A}]$ if X is $\mathcal{A}$-measurable

$X \quad \mathcal{H}\subset\mathcal{G} \to Y = E[X|\mathcal{G}] \to E[Y|\mathcal{H}]$

4) $E[X] = E[E[X|\mathcal{A}]]$ $\qquad \longrightarrow E[X|\mathcal{H}]$

(*) 5) $E[X|\mathcal{H}] = E[E[X|\mathcal{G}]|\mathcal{H}]$ if $\mathcal{H}\subset\mathcal{G}$. (i.e. if $A\in\mathcal{H} \Rightarrow A\in\mathcal{G}$)

Proof of (*). Define $Z = E[X|\mathcal{G}] \overset{\text{Def}}{\Longleftrightarrow} \begin{cases} Z \text{ is } \mathcal{G} \text{ measurable} \\ \forall A\in\mathcal{G} \quad \int_A Z\,dP = \int_A X\,dP \end{cases}$ $\Rightarrow$ $\forall A\in\mathcal{H}\subset\mathcal{G}$ (2)' $\int_A Z\,dP = \int_A X\,dP$

Define $Y = E[X|\mathcal{H}] \Longleftrightarrow \begin{cases} Y \text{ is } \mathcal{H}\text{-measurable} \quad (1) \\ \forall A\in\mathcal{H}, \quad \int_A Y\,dP = \int_A X\,dP. \quad (2) \end{cases}$

We need to show that

$Y = E[Z|\mathcal{H}] \Longleftrightarrow \begin{cases} Y \text{ is } \mathcal{H}\text{-measurable} \checkmark \text{ from (1)} \\ \forall A\in\mathcal{H}, \quad \int_A Z\,dP = \int_A Y\,dP \checkmark \text{ (from (2) \& (2)')} \end{cases}$

Back to B.M. $\{B(t)\}_{t\geq 0}$

1) stationary & independent increment. $\quad B(t+s)-B(s) \sim N(0,t)$

2) Continuous path

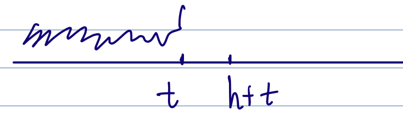
If we drop 2), another process is Poisson (λt).

$0 \qquad t$

B.M. is both a Markov process and martingale.

Thm 1: B.M. is a Markov process

Clearly: $P\left(B_{t+h} \,\middle|\, \sigma(B_s, s\le t)\right)$



$$= P\Big(\underbrace{B_{t+h} - B_t}_{N(0,h)\ \perp\!\!\!\perp\ \sigma(B_s,\, s\le t)} + B_t \,\Big|\, \sigma(B_s, s\le t)\Big)$$

$$= P\left(B_{t+h} - B_t + B_t \,\middle|\, \sigma(B_t)\right)$$

Thm 2: ① B.M. is a martingale w.r.t $\mathcal{F}_t = \sigma(B_s, s\le t)$

Recall the analogous martingale in simple random walk?

② ⊛ $\{B^2(t) - t\}_{t\ge 0}$ is a martingale

③ $\left\{e^{\lambda B(t) - \frac{1}{2}\lambda^2 t}\right\}_{t\ge 0}$ is a martingale (exponential martingale) ← critical martingale

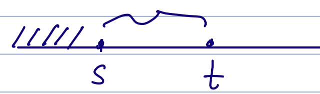
(Exercise)

e.g. Novikov condition

One can easily check $E\left[\,|B^2(t) - t|\,\right] < \infty$ for $\forall t$ fixed.

Proof ② critical to check martingale property.

$$E\left[B^2(t) - t \,\middle|\, \sigma(B_s, s\le t)\right]$$



$$= E\left[\left(B(t) - B(s) + B(s)\right)^2 - t \,\middle|\, \sigma(B_s, s\le t)\right]$$

$$= E\left[(B(t)-B(s))^2 + 2(B(t)-B(s))B(s) + B^2(s) - t \,\middle|\, \sigma(B_s, s\le t)\right]$$

$$= E\left[(B(t)-B(s))^2 \,\middle|\, \sigma(B_s, s\le t)\right] + 2E\left[(B(t)-B(s))B(s) \,\middle|\, \sigma(B_s, s\le t)\right]$$

$$= E\left[(N(0, t-s))^2 \right] + 2B(s)\,E\big[\underbrace{B(t)-B(s)}_{=0}\big] \qquad + B^2(s) - t$$

$$= t - s + 0 + B^2(s) - t$$

$$= B^2(s) - s \qquad \#$$

$B(t)$ $\xrightarrow{\text{SDE}}$ Construction of many more martingales

martingale

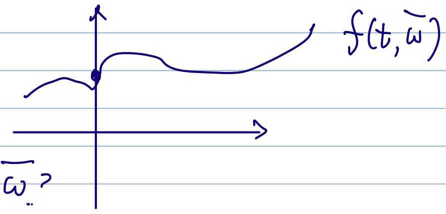
$f(t)$ $\qquad \int_a^b f(t)\,dt \qquad \int_a^b dF(t) \qquad \int_a^b f(t)\,dg(t)$

function

"White noise" $\int_a^b f(t,\omega)\, dB(t,\omega)$

$\hookrightarrow$ stochastic process

"path view" $\int_a^b f(t,\bar\omega)\, dB(t,\bar\omega)$ for every $\bar\omega$?



Starting point:

quadratic variation of B.M.

Recall: (1) symmetric random walk $Y_n = \sum_{i=1}^{n} X_i$ $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$ i.i.d

$$E[Y_n] = 0$$

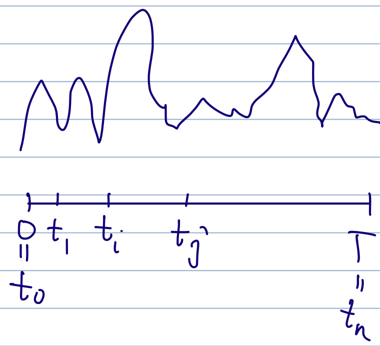
$$Var(Y_n) = \sum_{i=1}^{n} Var(X_i) = n.$$

$$[Y\ Y](\omega) = \sum_{j=1}^{n} (Y_j - Y_{j-1})^2 = n. \text{ for every } \omega.$$

$\uparrow$ quadratic variation

(2) Def of quadratic variation of function $f(t)$, $t \in [0,T]$

$$[f\ f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left(f(t_{j+1}) - f(t_j)\right)^2$$



$\Pi = (t_0, \cdots t_n)$ is a partition of $[0,T]$

$$\|\Pi\| = \max_{n-1 \le i \le 0} (t_{i+1} - t_i)$$

Thm: $\underbrace{[B\ B](T)}_{\text{is a r.v. with } T \text{ fixed.}} = T.$

$$\Leftrightarrow \begin{cases} Var\left([B\ B](T)\right) = 0 & (1) \\ E\left([B\ B](T)\right) = T & (2) \end{cases}$$

Proof: To see this, take any partition Π of $[0,T]$

$$Q_\Pi := \sum_{j=0}^{n-1} \left(B(t_{j+1}) - B(t_j)\right)^2$$

$$E[Q_\Pi] = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = T - 0 = T$$

Focus on $Var(Q_\Pi)$, suffices to analysis $Var\left(B(t_{j+1}) - B(t_j)\right)^2$

$$Var\left(B(t_{j+1})-B(t_j)\right)^2 = E\left[\left\{\left(B(t_{j+1})-B(t_j)\right)^2-(t_{j+1}-t_j)\right\}^2\right]$$

$$= E\left[\underbrace{\left(B(t_{j+1})-B(t_j)\right)^4}_{E[(N(0,\,t_{j+1}-t_j))^4]} - 2(t_{j+1}-t_j)\,\underbrace{E\left[\left(B(t_{j+1})-B(t_j)\right)^2\right]}_{=\,t_{j+1}-t_j} + (t_{j+1}-t_j)^2\right]$$

$$= 3(t_{j+1}-t_j)^2 - (t_{j+1}-t_j)^2$$

$$= 2(t_{j+1}-t_j)^2$$

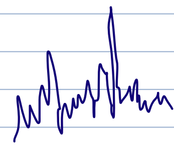
$$Var(Q_\Pi) = \sum_{j=0}^{n-1} 2(t_{j+1}-t_j)^2$$

$$\leq 2\|\Pi\| \underbrace{\sum_{j=0}^{n-1}(t_{j+1}-t_j)}_{T} \xrightarrow{\|\Pi\|\to 0} 0.$$

#.

Lemma: If a function $f: [0,T] \to \mathbb{R}$ has a continuous derivative (i.e. f' is continuous)

Then $[f\ f](T) = 0$ (Trivial to prove)

Corollary (of Lemma) B.M. has no smooth path anywhere. $\to$ 

B.M.

Next, stochastic integration: $[B\ B](T)$

$$\int_a^b f(t,\omega)\, dB(t,\omega)$$

$$\mathcal{H} \subset \mathcal{G} \quad \begin{cases} X \to \underbrace{E[X|\mathcal{G}]}_{Y} \to E[Y|\mathcal{H}] \\ X \longrightarrow E[X|\mathcal{H}] \end{cases}$$

$$E[Y|\mathcal{H}] = E[X|\mathcal{H}]$$

Last time

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}] \quad \text{when } \mathcal{H} \subset \mathcal{G} \qquad (*)$$

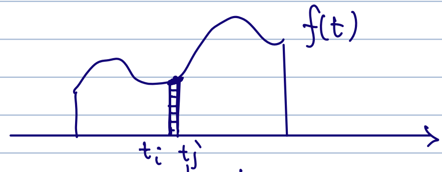
$$\mathcal{H} \subset \mathcal{G} \quad \begin{cases} X \to \underbrace{E[X|\mathcal{H}]}_{Z \text{ — is } \mathcal{H}\text{-measurable}} \to E[Z|\mathcal{G}] = E[E[X|\mathcal{H}]|\mathcal{G}] \\ X \longrightarrow E[X|\mathcal{H}] = Z \end{cases}$$

$B(t)$ —
- Markov
- Several martingales $f(B(t))$
- $[B\ B][T] \to$ Non-smooth path of B.M.

---

building block of stochastic integral

$\int_a^b f(t)\,dt$



$f[t_i\ t_j] \sim f(\xi)(t_j - t_i)$

$\xi \in (t_i\ t_j)$

↓ Calculus

↳ linearization (e.g. Signature (*))

↳ universal non-linearity

step 1 → $\int_a^b c\,dt$ — Constant

step 2 → $\sum_i \int_{a_i}^{b_i} C_i\,dt$ — piecewise-constant

step 3 → $\int_a^b f(t)\,dt$ — Continuous function

$$\int_0^T t\,dt = \frac{1}{2}T^2$$

$\int_a^b B(t,\omega)\,dt$ defined by each ω.

Step 1 $\int_0^T c\ dB(t,\omega)$?

Finance

$$\int_a^b f(t,\omega)\ \boxed{dB(t,\omega)}$$ — changes in "price"

model for some stocks

$f(t,\omega)$ → # of shares

$$= c\int_0^T dB(t,\omega)$$

$$dB(t,\omega) = B(t+h,\omega) - B(t,\omega)$$

$$= c\,(B(T,\omega) - B(0,\omega)) = c\,B(T,\omega)$$

Step 2 Simple process (vs piecewise constant)

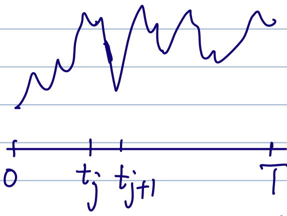
Fixed ω: $f(t,\omega)$ function → $e_i(t_i\ t_j)(\omega)$ — piecewise constant for fixed ω.

Step 3 define appropriate metric to measure the approximation error ($\Rightarrow$ limit).

Example $$\int_0^T B(t,\omega)\, dB(t,\omega) \longrightarrow \int_0^T t\,dt = \tfrac{1}{2}T^2$$

$$\overset{?}{=} \tfrac{1}{2}B^2(T,\omega)$$

$$\Phi(t_j^*) \;\underset{\text{calculus}}{\overset{\text{mimicking}}{=\!=}}\; \lim_{\|\Pi\|\to 0} \sum_{j=1}^n B(t_j^*)\left[B(t_j)-B(t_{j-1})\right]$$



Π = partition of $[0,T]$

$t_j^* \in [t_{j-1}, t_j)$

$f(t_{j-1}) = f(t_j^*)$

t_{j-1} $\quad$ t_j

Two test cases

$t_j^* = t_{j-1}$

Ito's integral

$$\Phi_1 = \lim_{\|\Pi\|\to 0} \sum_{j=1}^n B(t_{j-1})\left[B(t_j)-B(t_{j-1})\right]$$

$E[\Phi_1] = 0$ b/c

$$E\left[B(t_{j-1})\left(B(t_j)-B(t_{j-1})\right)\right]$$
$$= E\left[B(t_{j-1})\right]E\left[B(t_j)-B(t_{j-1})\right]$$
$$\| $$
$$0$$

$t_j^* = t_j$

$$\Phi_2 = \lim_{\|\Pi\|\to 0} \sum_{j=1}^n B(t_j)\left[B(t_j)-B(t_{j-1})\right]$$

$$E\left[\Phi_2 - \Phi_1\right] = \lim_{\|\Pi\|\to 0} E\left(\left[B(t_j)-B(t_{j-1})\right]^2\right) = T.$$

S-integral is used mostly to model in finance about "inside trading" + filtration enlargement

---

Ito's integral $\int_0^T f(t,\omega)\, dB(t,\omega)$

Key $f(t,\omega)$ should always be adapted w.r.t $\sigma(B(s), s\le t)$

If we do take Ito's view

$$\int_0^T B(t)\, dB(t,\omega)$$

$$\Phi_1 = \lim_{\|\Pi\|\to 0} \sum_{j=1}^n B(t_{j-1})\left[B(t_j)-B(t_{j-1})\right] \overset{?}{=} \tfrac{1}{2}B^2(T) - \tfrac{1}{2}T$$

Warning $\Rightarrow$ $\displaystyle \lim_{\|\Pi\|\to} \sum_{j=1}^n \left[B(t_j)-B(t_{j-1})\right]^2$
$$\|$$
$$T$$

Ito's correction term from $[B\ B](T) = T$.

<u>Note:</u>

$$2\sum_{j=1}^{n} B(t_{j-1})\left[B(t_j) - B(t_{j-1})\right] + \underbrace{\sum_{j=1}^{n}\left[B(t_j) - B(t_{j-1})\right]^2}_{T}$$

$$= \underbrace{\sum_{j=1}^{n} B^2(t_j) - B^2(t_{j-1})}_{\shortparallel \atop B^2(T)}$$

$$\int_0^T B(t,\omega)\, dB(t,\omega) \longleftrightarrow \int_0^T t\, dt = \frac{1}{2}T^2$$

$$\Pi = (t_0, \cdots t_n) = [0,T] \qquad \|\Pi\| = \max_{t_j}(t_{j+1} - t_j)$$

$0 \quad t_j \quad t_{j+1} \quad T$

Itô — adaptiveness

$$\phi_n(t,\omega) = \sum B(t_j)(\omega)\chi_{(t_j\ t_{j+1}]}(t) \qquad \chi_{(t_j\ t_{j+1}]}(t) = \begin{cases} 1 & t \in (t_j\ t_{j+1}] \\ 0 & 0 \end{cases}$$

Simple process $\longleftrightarrow$ Step function (piece-wise linear)

$\sum B(t_{j+1}^*)(\omega)$ for $t_j^* \in (t_j\ t_{j+1}]$

Step 1: Using $\Phi_n(t,\omega)$ as an approximation of $B(t,\omega)$ on $[0,T]$.

$$\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} B(t_j)[B(t_{j+1}) - B(t_j)] = \lim_{\|\Pi\|\to 0} \frac{1}{2}B^2(T) - \frac{1}{2}T = \frac{1}{2}B^2(T) - \frac{1}{2}T. \checkmark$$

Step 2: $\Phi_n(t,\omega)$ is an approximation of $B(t,\omega)$, in what sense ?

Prop: $E\left[\int_0^T |\phi_n(t,\omega) - B(t,\omega)|^2 dt\right] \Rightarrow 0$ as $\|\Pi\| \to 0$.

Proof:
$$= E\left[\sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (B(t_j) - B(t))^2\, dt\right]$$

$$= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} E(B(t_j) - B(t))^2\, dt$$

$$= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} (t - t_j)\, dt = \sum_{j=0}^{n-1} \frac{1}{2}(t_{j+1} - t_j)^2 \leq \|\Pi\| \frac{1}{2} \underbrace{\sum (t_{j+1} - t_j)}_{\frac{1}{2}\cdot T} \to 0$$

as $\|\Pi\| \to 0$.

Step 3:
Ito's isometry.

$$E\int_0^T |\phi_n(t,\omega) - B(t,\omega)|^2 dt = E\left[\int_0^T (\phi_n(t,\omega) - B(t,\omega))\, dB(t)\right]^2$$

$$\| $$

$$\text{distance}\left(\int_0^T \phi_n(t,\omega)\, dB(t) \quad \int B(t,\omega)\, dB(t,\omega)\right)$$

Step 2:
$\phi_n(t,\omega)$ is "an" approximation of $B(t,\omega)$

General recipe: $\int_0 f(t,\omega)\,dB(t,\omega) = \lim_{\|\Pi\|\to 0} \int_0 \phi_n(t,\omega)\,dB(t,\omega)$

$$\Uparrow$$

$f(t,\omega)$

$\uparrow$

Step2 Approximation of f by simple process

$\phi_n \longrightarrow \int_0^T \phi_n(t,\omega)\,dB(t,\omega)$

$\vdots$

$\phi_1 \longrightarrow \int_0^T \phi_1(t,\omega)\,dB(t,\omega)$

proposition

$\boxed{Q1}$: What class of f can ensure such sequence of ϕ_n approximation? ✓

$\boxed{Q2}$: Would such $\lim_{\|\Pi\|\to 0} \int_0^T \phi_n(t,\omega)\,dB(t,\omega)$ exist? How does it depend on choice of $\phi_n(t,\omega)$? (Ito's isometry!) ✓

In order for the general recipe to work, we need to have

Ito's isometry: For all $f(t,\omega)\in\mathcal{H}$, where $\mathcal{H} = \left\{ f(t,\omega):\ E\left[\int_s^T f^2(t,\omega)\,dt\right] < \infty \text{ and } \begin{array}{l} f(t,\omega) \text{ is adaptive} \\ \text{to } \mathcal{F}_t^B \end{array} \right\}$

(✱✱) $$E\left[\int_s^T f(t,\omega)\,dB(t)\right]^2 = E\left[\int_s^T f^2(t,\omega)\,dt\right]$$

Proof of (✱✱) is to establish it first on simple process.

E.g. $B(t,\omega)$ is adaptive
$B(2t,\omega)$ is NOT
$B(\frac{t}{2},\omega)$ is.

Remark: Simple process is a generalization of step function to stochastic settings,

such that $\phi_\Pi(t,\omega) = \sum e_j(\omega)\,\chi_{(t_j,\,t_{j+1}]}(t)$

Where $\left\{\begin{array}{l} \chi_{(t_j\,t_{j+1}]}(t) = \begin{cases} 1 & t\in(t_j\ t_{j+1}] \\ 0 & 0 \end{cases} \\ e_{t_j}(\omega) \text{ is } \mathcal{F}_{t_j}^B\text{-measurable} \end{array}\right.$

For $f(t,\omega)\in\mathcal{H}$, $\int_0^T f(t,\omega)\,dB(t) \stackrel{\Delta}{=} \lim_{\|\Pi\|\to 0} \int_0^T \phi_\Pi(t,\omega)\,dB(t,\omega)$ (Ito's integral)

if $E\int_0^T |\phi_\Pi - f|^2\,dt \to 0$ as $\|\Pi\|\to 0$.

Example (as before)

Correction term

$$\int_0^T B(t,\omega)\,dB(t,\omega) \stackrel{\Delta}{=} \lim_{\|\Pi\|\to 0} B(t_j)\big(B(t_{j+1})-B(t_j)\big) = \tfrac{1}{2}B^2(T) - \tfrac{1}{2}T.$$

Compared to $\int_0^T t\,dt = \frac{1}{2}T^2$

$$\text{where } \varphi_\Pi = \sum_{j=0}^{n-1} e_{t_j}\,\chi_{(t_j,\,t_{j+1}]}(t)$$

$$e_{t_j} = B(t_j)$$

Oksendal (textbook chapter, post in bcourses)

Properties of Ito's integral:

1) $\int_S^T (c_1 f + c_2 g)(t,\omega)\,dB(t) = c_1 \int_S^T f(t,\omega)\,dB(t,\omega) + c_2 \int_S^T g(t,\omega)\,dB(t,\omega)$

2) $\int_S^T f(t,\omega)\,dB(t,\omega) = \int_S^u f(t,\omega)\,dB(t,\omega) + \int_u^T f(t,\omega)\,dB(t,\omega)$

3) $I(t) \stackrel{\Delta}{=} \int_0^t f(t,\omega)\,dB(t,\omega)$

Then $I(t)_{t\geq 0}$ is a martingale with zero mean and variance $E\left[\int_0^t f^2(t,\omega)\,dt\right]$. (Statement of Ito's isometry).