§ Chapter 3. Conditioning, Martingales, Stopping times

Recall: $(\Omega, \mathcal{F}, P)$, X is $\mathcal{F}$-measurable

$$\Leftrightarrow \{\omega \in \Omega : X(\omega) \leq c\} \in \mathcal{F} \quad \forall c \in \mathbb{R}$$

Recall Conditional probability

Take $B \in \mathcal{F}$, $P(B) > 0$, Then the probability of of an event A given B is $\quad P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Example: $A = \{\text{heads on first throw}\} = \{HH, HT\}$

$B = \{\text{heads at least once}\} = \{HH, HT, TH\}$

$$P(A) = P(\{HH, HT\}) = P(\{HH\}) + P(\{HT\}) = \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2}$$

$$P(A \cap B) = P(\{HH, HT\}) = \tfrac{1}{2}$$

$$P(B) = P(\{HH, HT, TH\}) = \tfrac{3}{4}$$

$$\Rightarrow P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{3}$$

Two questions: <1> How to condition on a σ-algebra $\mathcal{G}$ rather than
(motivation) a specific event B (i.e. $P(A|\mathcal{G})$)

<2> How to condition on events for which $P(B) = 0$

Extensions:

<1> Assume $\mathcal{G} = \sigma(B_1, B_2, \cdots B_N)$ $N \in \mathbb{N}$. $B_i \cap B_j = \phi$ $\forall i \neq j$

$\bigcup_{i=1}^{N} B_i = \Omega$

N can be extended to ∞

$$1_{B_i}(\omega) = \begin{cases} 1, & \omega \in B_i \\ 0, & \omega \notin B_i \end{cases}$$

Define $P(A|\mathcal{G})(\omega) := \sum_{i=1}^{N} P(A|B_i) 1_{B_i}(\omega)$

Ex. $\omega \in B_6$ $\quad 1_{B_i}(\omega) = \begin{cases} 1 & i=6 \\ 0 & \text{otherwise} \end{cases}$ $\Rightarrow P(A|\mathcal{G})(\omega) = P(A|B_6)$

More generally, take a r.v. X, then we define the conditional expectation

$$E[X|\mathcal{G}](\omega) = \sum_{i=1}^{\infty} \mathbb{E}[X_i|B_i]\cdot 1_{B_i}(\omega)$$

Example: coin-toss. $B_1 = \{HT, HH\}$, $B_2 = \{TH, TT\}$

$$\mathcal{G} = \sigma(B_1, B_2) = \{\Omega, \phi, B_1, B_2\}$$

$X \triangleq$ number of heads

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|B_1]\cdot 1_{B_1} + \mathbb{E}[X|B_2]\cdot 1_{B_2}$$
$$= \frac{1+2}{2} 1_{B_1} + \frac{0+1}{2} 1_{B_2}$$
$$= \frac{3}{2} 1_{B_1} + \frac{1}{2} 1_{B_2}$$

<2> Def: (sub-σ-algebra) if $\mathcal{G}$ is a σ-algebra, and

$$A \in \mathcal{G} \Rightarrow A \in \mathcal{F} \quad \forall A \in \mathcal{G}$$

Def 3.0.1 (Conditional Expectation) Let X be a ($\mathcal{F}$-measurable) r.v. $\mathcal{G} \subseteq \mathcal{F}$ is a sub-σ-algebra. We define $E[X|\mathcal{G}]$, the conditional expectation of X given $\mathcal{G}$, to be a r.v. satisfying

only information contained in $\mathcal{G}$ is used to compute conditional expectation

① $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable (not $\mathcal{F}$-measurable in general)

② $\forall G \in \mathcal{G}$, $E[\underbrace{X 1_G}_{\text{r.v.}}] = E[\underbrace{E[X|\mathcal{G}]}_{\text{r.v.}}\cdot 1_G]$ (averaging identity)

F-measurable, not $\mathcal{G}$-measurable.

in general, this will not hold for events $F \in \mathcal{F}/\mathcal{G}$

Remark:

- $\mathcal{G} \subseteq \mathcal{F}$: $\Rightarrow$ $\mathcal{G}$ contains less information than $\mathcal{F}$.

▷ $\mathbb{E}[X|\mathcal{G}]$ is the "best" approximation of X on the subsets of $\mathcal{G}$.

☆▷. Conditional expectation is unique. For any $\mathcal{G}$-measurable r.v. Y satisfies $E[X1_G] = \mathbb{E}[Y1_G]$ $\forall G \in \mathcal{G}$, we have

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

⇒ we often "guess" $E[X|\mathcal{G}]$, and then check (1) & (2) of the definition

▷ Plug in $G = \Omega$ into ②

$$E[X] = E[X 1_\Omega] = E[E[X|\mathcal{G}]1_\Omega] = E[E[X|\mathcal{G}]]$$

⇒ law of total probability

▷ assume $\mathcal{G} = \sigma(Y)$ for another r.v. Y on $(\Omega, \mathcal{F}, P)$

$$E[X|Y] := E[X|\mathcal{G}] = g(Y)$$

lemma: A $\sigma(Y)$-measurable r.v. Z can be written as a function of Y. i.e. $g: \mathbb{R} \to \mathbb{R}$ $Z = g(Y)$

Property 3.0.2

① linearity: If X, Y are r.v and $\lambda \in \mathbb{R}$, then

$$\mathbb{E}[X + \lambda Y|\mathcal{G}] = E[X|\mathcal{G}] + \lambda \mathbb{E}[Y|\mathcal{G}]$$

② Positivity: If $X \le Y$, then $\mathbb{E}[X|\mathcal{G}] \le E[Y|\mathcal{G}]$

③ "Take-out-what's-known"

If X is $\mathcal{G}$-measurable and Y is any r.V. then

$$\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$$

④ Tower Property: If $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ are σ-algebras, then

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

(the smaller σ-algebra always wins)

Remark:

$\mathcal{G} \subseteq \mathcal{F}$, X is $\mathcal{F}$-measurable, Y is $\mathcal{G}$-measurable, then

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY|\mathcal{G}]] = \mathbb{E}[Y\,\mathbb{E}[X|\mathcal{G}]]$$

Property 3.0.1 If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X|\mathcal{G}] = X\,\underbrace{\mathbb{E}[1|\mathcal{G}]}_{}$ $= X$

If X is independent of $\mathcal{G}$, that is

$$P(\{X \le t\} \cap G) = P(X \le t) \cdot P(G) \quad \forall t \in \mathbb{R}, G \in \mathcal{G}$$

then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$

In particular. If $\mathcal{G} = \sigma(Y)$ and X and Y are independent, then

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X]$$

Lemma 3.0.1 (Independence Lemma). Suppose X, Y are two random variables such that X is independent of $\mathcal{G}$ and Y is $\mathcal{G}$-measurable. Then if $g = g(x,y)$ is any function of two variables we have

$$\mathbb{E}[g(X,Y)|\mathcal{G}] = h(Y)$$

where $h = h(y)$ is the function defined by

$$h(y) := \mathbb{E}[g(X,y)]$$

3.1. Martingales

Example: Assume today is time t, and you want to invest until $T \geq t$. Your portfolio has price $(X_s)_{t \leq s \leq T}$

If X is a martingale, then "on average", you will not be able to lose or win money on your investment in the future

$\rightarrow$ $E[\underbrace{X_T}_{\text{future value of the portfolio}} | \mathcal{F}_t] = X_t \quad (T \geq t)$

"Martingales are fair games"

3.2. Adapted processes & filtration

Def 3.2.1: A filtration is a family of σ-algebra $(\mathcal{F}_t)_{t \geq 0}$ such that whenever $s \leq t$ we have $\mathcal{F}_s \subseteq \mathcal{F}_t$

Def 3.2.3. Given a stochastic process X, the filtration generated by X is the family of σ-algebra $(\mathcal{F}_t^X)_{t \geq 0}$ so that

$\forall t \quad (X_{s_1}, X_{s_2}, \dots) \; s_1 \leq s_2 \leq \dots \leq t$ is $\mathcal{F}_t^X$-measurable

i.e. $\mathcal{F}_t^X$ is generated by all events that can be observed using $\{X_s : s \leq t\}$

Def 3.2.4. (Adapted process). A stochastic process X is said to be adapted to $(\mathcal{F}_t)_{t \geq 0}$ if for all $t \geq 0$, $\underbrace{X_t \text{ is } \mathcal{F}_t\text{-measure}}$

$\{X_t \leq c_i\} \in \mathcal{F}_t \quad \forall c_i \in \mathbb{R}$

(definition of measurabl

Remark:

X is always adapted to $(\mathcal{F}_t^*)_{t\geq 0}$

(recall that a r.v. y is $\sigma(y)$-measurable)

Def 3.3.1 (Martingale) A stochastic process $\mathcal{M}$ is a martingale (MG)

w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ if

① $\mathcal{M}$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$

② $\forall s \leq t$, we have $E[M_t | \mathcal{F}_s] = M_s$ (fair game)

Def 3.3.2 **Sub-Martingale:** → "earn money on average"

cond ① & $E[M_t | \mathcal{F}_s] \geq M_s \quad \forall t \geq s$

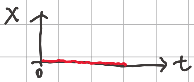
Super-Martingale: → "losing money on average"

cond ① & $E[M_t | \mathcal{F}_s] \leq M_s \quad \forall t \geq s$

Example:

- Martingale

$M_t = 0 \quad \forall t \geq 0$



$E[M_t | \mathcal{F}_s] = M_s \quad \forall t \geq s$

$\{0 \leq c_i\} = \begin{cases} \emptyset & c_i < 0 \\ \Omega & c_i \geq 0 \end{cases}$

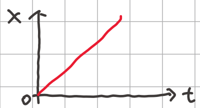
↑

$\{M_t \leq c_i\} \in \mathcal{F}_t \quad \forall c_i \in \mathbb{R}$

$\Rightarrow$ $\mathcal{M}$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$

- sub-martingale

$M_t = t \quad \forall t \geq 0$



$E[M_t | \mathcal{F}_s] = t \geq s = M_s \quad \forall t \geq s$

- Super-martingale

$M_t = -t$

Example Let $T>0$ and $(\mathcal{F}_t)_{0\le t\le T}$, X is $\mathcal{F}_T$-measurable

Then $X_t := E[X|\mathcal{F}_t]$ is a martingale w.r.t. $(\mathcal{F}_t)_{0\le t\le T}$

Proof: $X_t = E[X|\mathcal{F}_t] \Rightarrow X_t$ is adapted to $(\mathcal{F}_t)_{t\ge 0}$

$\forall s\ge t$ $\quad E[X_s|\mathcal{F}_t] = E[E[X|\mathcal{F}_s]|\mathcal{F}_t]$

$$= E[X|\mathcal{F}_t] = X_t$$

therefore X_t is a martingale w.r.t. $(\mathcal{F}_t)_{0\le t\le T}$

3.4. Martingale property of random walks (discrete times)

In Lecture 1, we defined $S_k = \sum_{n=1}^{k} \xi_n$

Take $\mathcal{F}_k = \mathcal{F}_k^S$

<1> S is $(\mathcal{F}_k^S)_{k\ge 0}$-adapted by definition

<2> Take $k, l \in \mathbb{N}$ and $k\le l$

$$E[S_l|\mathcal{F}_k] = E[S_l - S_k + S_k | \mathcal{F}_k]$$

S_k is adapted to $\mathcal{F}_k$

(linearity) $= E[S_l - S_k|\mathcal{F}_k] + E[S_k|\mathcal{F}_k]$

$$= E\Big[\sum_{n=k+1}^{l} \xi_n \Big| \mathcal{F}_k\Big] + S_k$$

(linearity) $= \sum_{n=k+1}^{l} E[\xi_n|\mathcal{F}_k] + S_k$

$$= \sum_{n=k+1}^{l} E[\xi_n] + S_k$$

$(\xi_{k+1}, \cdots \xi_l)$ is independent of $(\xi_1, \cdots, \xi_k)$

$$= S_k$$

3.5. Martingale property of Brownian motion

Theorem 3.5.1 Let W be a Brownian motion, $\mathcal{F}_t = \mathcal{F}_t^W$, then

W is a MG w.r.t. $(\mathcal{F}_t)_{t\geq 0}$

proof: Recall independence of increments

$$\forall\ 0 \leq t_1 \leq \cdots \leq t_n$$

$W_{t_1}, W_{t_2} - W_{t_1}, \cdots W_{t_n} - W_{t_{n-1}}$ are independent r.v.s

In particular, $W_t - W_s$ $(t \geq s)$ is independent of W_r. $\forall r \leq s$

$\Rightarrow W_t - W_s$ is independent of $\mathcal{F}_s^W$

Thus
$$\begin{aligned} E[W_t \mid \mathcal{F}_s] &= E[W_t - W_s + W_s \mid \mathcal{F}_s] \\ &= E[W_t - W_s \mid \mathcal{F}_s] + E[W_s \mid \mathcal{F}_s] \\ &= \underbrace{E[W_t - W_s]}_{\sim N(0,\sigma^2(t-s))} + W_s \\ &= 0 + W_s = W_s \end{aligned}$$

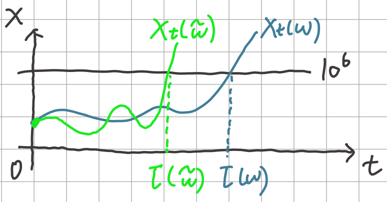
3.6. Stopping times

Imagine you want to decide when to stop investing depending on your current & historical walk. Your stopping rule is given by a random time, which depends on the information available to you at a specific time.

Def 3.6.1 (stopping time). A stopping time is a function

$\tau: \Omega \to [0, \infty)$, such that $\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0$

example:



Prop. 3.6.1 Let X be a continuous adapted process, let $a \in \mathbb{R}$

$\tau := \underset{t}{\operatorname{argmin}} \{ t \geq 0 : X_t = a \}$. Thus τ is a stopping time.

It is called the (first) hitting time of level a.

Theorem 3.6.1. (Doob's optional sampling theorem)

$\tau(\omega) \leq C$ $\forall \omega \in \Omega$

Let M be a martingale, τ be a bounded stopping time,

then the stopped process $M_t^{\tau} := M_{\tau \wedge t}$ is also a martingale.

Consequently, $E[M_t^{\tau}] = E[M_{\tau \wedge t}] = E[M_0] = E[M_t]$ $\forall t \geq 0$

(the proof is omitted)

Remark:

if instead of assuming τ is bounded, we assume M^{τ} is bounded. the Doob's optional sampling theorm is still valid.

One failed case for this theorem, for example, Let W be a standard Brownian motion (that is $W_0 = 0$ and $Var(W_t) = t$) let τ be the first hitting time of W to 1. Then obviously $E[W_{\tau}] = 1 \neq 0 = E[W_0]$. This is because $E[\tau] = \infty$ (or $\exists \omega \in \Omega$ $\tau(\omega) = \infty$) and W_{τ} is not bounded either. Thus, the Doob's optional sampling theorem is not valid.