Lec 2: Conditional Expectation and Martingales

1) A simple motivating example
2) Conditioning and Prediction
3) Classical conditional probability
4) Abstract conditional expectation
5) Basic properties
6) Conditional expectation and L^2 projection

Martingales
1) Filtrations and semi-martingales.
2) Martingale Transforms, Doob Martingales
3) Martingale decompositions
4) Stopping times and optional stopping
5) Maximal inequalities
6) Martingale convergence
7) Zero-One Law
8) Option pricing and Black-Scholes.

Let $X \sim Bin(100, 1/2)$ represent outcomes of coin toss. Player 2 is supposed to guess the value of X. What should his guess be?

Obviously 50. But in what sense is this the best guess. Let's suppose your guess is g, and you want to minimize the L^2 norm between X and g.

$$\mathbb{E}[|X-g|^2] = \|X-g\|_2^2$$

$$= \mathbb{E}[(X - EX + EX - g)^2]$$

$$= \overbrace{Var(X)}^{\sigma^2} + (\mu - g)^2 = f(g)$$

$f(g)$ takes its minimum value at:

Ex: What if I define some "reasonable" distance function $d(X,g)$. What happens then? Start with the p norms and graduate to $d(X,g) = \varphi(X-g)$ where φ is some convex function.

Now suppose I give you "EXTRA INFORMATION"? Say Y is the # of heads tosses in the first 10 coin tosses, what should your guess be?

lets call our guess g How do we minimize

$$E((X - g)^2 | Y)$$

<u>POLL</u>

What should our guess be for X?

A	B	C
$50 - Y$	$45 + Y$	50

Well, the remaining 90 tosses are independent of the first 10. So we should guess $g = 45 + Y$

We will call $E[X|Y] = g$.
Our "best guess" for X given Y is the conditional expectation

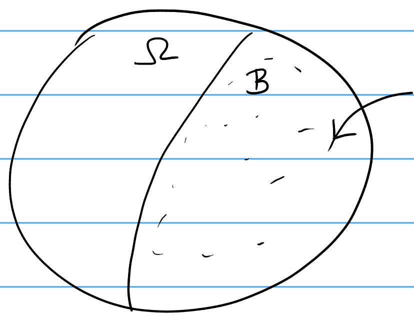
Ex (from Le Gall)

let $\Omega = \{1,2,\ldots,6\}$ $P(\omega) = 1/6$.

$$Y(\omega) = \begin{cases} 1 & \omega \text{ is odd} \\ 0 & \text{otherwise} \end{cases}$$

let $X(\omega) = \omega$

$$E[X|Y] = \begin{cases} \sum_{i \in 1,3,5} i \, P(X = i \mid Y \in 1,3,5) = \frac{1+3+5}{3} = 3 \\ 4 \end{cases}$$



$$E[X|B] = \sum_{y \in B} P(X = y \mid B)\, y$$

Classical conditioning

For 2 events A and B, $\mathbb{P}(B) \neq 0$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

So for a discrete random variable $X \in \{a_0, a_1, \dots\}$

$$\mathbb{E}[X|B] = \sum_{i=0}^{\infty} a_i \, \mathbb{P}(X = a_i | B)$$

$$= \sum_{i=0}^{\infty} a_i \, \frac{\mathbb{P}(X = a_i \cap B)}{\mathbb{P}(B)}$$

If X is continuous

$$\mathbb{E}[X|B] = \frac{1}{\mathbb{P}(B)} \int X \mathbb{1}_B \, d\mathbb{P}$$

☆ Can you prove the above ?

So certainly, we want some notion of conditional expectation that tallies with the above.

Now suppose Y is discrete. Then what should $E[X|Y]$ be? Let $Y \in \{b_0, b_1, \dots\}$

$$E[X \mid Y = b_i] = \sum_j a_j \, P(X = a_j \mid Y = b_i)$$

$$= f(b_i)$$

So $E[X|Y] = f(Y)$ some function of Y.

BUT if Y is continuous, then $\mathbb{P}(Y = z) = 0$ for any z, then you cannot divide by $\mathbb{P}(Y = z)$

POLL

Similarly, what should $E[X \mid \mathbb{1}_B]$ be?

A

$$\frac{1}{\mathbb{P}(B^c)} \int X \mathbb{1}_{B^c} \, d\mathbb{P} + \frac{1}{\mathbb{P}(B)} \int X \mathbb{1}_B \, d\mathbb{P}$$

B

$$\mathbb{1}_B \, E[X \mid B] + \mathbb{1}_{B^c} E[X \mid B^c]$$

1) We defined conditional prob on events.
2) Conditional expectation on events
3) " " on discrete rvs.

We learned $E[X | Y] = f(Y)$ a random variable.

What we have learned is that we can condition on r.v.s ; thus we can condition on the σ algebra associated with Y itself. So our abstract version of conditional expectation will condition on σ-algebras.

<u>Prop</u> Let $X \in L^1$, Y discrete. Then

$$E\big[\,|E[X|Y]|\,\big] \leq E\big[|X|\big]$$

That is $E[X|Y] \in L^1(\Omega, \sigma(Y), \mathbb{P})$

Pf: $$E\big[\,|E[X|Y]|\,\big] = \sum_{i=0}^{\infty} P(Y=b_i)\, |E[X| Y=b_i]|$$

$$= \sum_{i=0}^{\infty} P(\cancel{Y=b_i}) \frac{|E[X 1_{\{Y=b_i\}}]|}{P(\cancel{Y=b_i})} \leq \sum_{i=0}^{\infty} E[|X| 1_{\{Y=b_i\}}]$$

and the rest follows from MCT or Fubini.

Prob

Let $Z: \Omega \to \mathbb{R}$ be bounded $\sigma(Y)$ measurable, then

$$E[ZX] = E[Z\, E[X|Y]]$$

Pf:

Let $\xi(Y) = Z$.

$$E[\xi(Y)\, E[X|Y]]$$

$$= \sum_{i=0}^{\infty} \xi(b_i)\, E[X|Y=b_i]\, P(Y=b_i)$$

$$= \sum_{i=0}^{\infty} \xi(b_i)\, E[X 1_{\{Y=b_i\}}] = E\left[\sum_{i=0}^{\infty} \xi(b_i) X 1_{\{Y=b_i\}}\right]$$

Using $E[X|Y] \in L^1$ and Fubini / DOM.

$$= E[\xi(Y) X]$$ — ☆ Important.

∎

How should I think about this?

Choose $\xi(Y) = 1_{[b_k-\varepsilon, b_k+\varepsilon]}(Y)$ and can reconstruct $E[X|Y]$ using this property (if b_k is not an accumulation pt of Y)

$$E[\xi\, E[X|Y]] = E[X 1_{\{b_k-\varepsilon, b_k+\varepsilon\}}(Y)]$$

Consequence:

If $\sigma(Y) = \sigma(Y')$ then $E[X|Y] = E[X|Y']$ a.s.

$E[Z\, E[X|Y']] = E[ZX] = E[Z\, E[X|Y]]$ from prev.

choose $Z = 1_{E[X|Y'] > E[X|Y]}$ and apply a standard argument.

<u>Remark</u>: $E[X|Y]$ only depends on $\sigma(Y)$!

So how does one define $E[X|Y]$ for general Y?

Note, $E[X|1_B] = 1_B\, E[X|B] + 1_{B^c}\, E[X|B^c]$

If Y simple $Y = \sum a_i 1_{B_i}$

$\Rightarrow E[X|Y] = \sum_{i=1}^{n} E[X|B_i]\, 1_{B_i}$

Can one somehow take a limit here? People usually take the reverse approach nowadays. So we will revisit this.

Space $(\Omega, \mathcal{F}, P)$. Subalgebra: $\mathcal{G} \subset \mathcal{F}$

Prop 8.1 (from Khoshnevisan): If $X \in L^1(P)$ then there exists an a.s.-unique random variable $E[X|\mathcal{G}] \in L^1(P)$ st

1) $\mathcal{G}$ measurable
2) Defining property: $\forall \xi$ that is $\mathcal{G}$-meas.

$$E\left[\xi \, E[X|\mathcal{G}]\right] = E[\xi X]$$

☆ Ex: If X is $\mathcal{G}$ measurable then $X = E[X|\mathcal{G}]$ a.s.

Def: $E[X|Y] := E[X|\sigma(Y)]$

Pf: Will define a new meas. space with a signed measure $(\Omega, \mathcal{G}, \nu)$:

$\mu(\emptyset) = 0$

μ countably additive

μ cannot take both $+\infty$ and $-\infty$

LeGall's version of signed meas. at least seems a bit stronger.

$$\nu(A) = \int X 1_A \, dP$$

Restrict P to $\mathcal{G}$ and thus $\nu \ll P$ (definition of integral)

If $P(A) = 0$,

$$|\nu(A)| \leq \int |X| 1_A \, dP = 0$$

Similarly, one checks that ν is finite.

Radon-Nikodym theorem says there exists an $\tilde{X}$ st for any g that is $\mathcal{G}$-measurable and integrable,

$$\int Z \, d\nu = \int Z \tilde{X} \, dP$$

Choose $Z = 1_{\{\tilde{X} > 0\}}$ and the finiteness of ν shows $\tilde{X} \in L^1(\Omega, \mathcal{G}, P)$

Using simple functions and bounded convergence we can also get

$$\int Z \, d\nu = \int Z X \, dP$$

$$\Rightarrow \int Z \, d\nu = \int Z \tilde{X} \, dP = \int Z X \, dP \quad (☆1)$$

We call $\tilde{X} =: E[X|G]$ and

$\star 1$ is the defining property.

Remarks: $E[X|G]$ is a random variable

$E[X|Y]$ is an r.v. that is a function of Y

Lem: (Exercise) $X \geq 0$ a.s $\Rightarrow E[X|G] \geq 0$ a.s.

Theorem. $(\Omega, \mathcal{F}, P)$ $\mathcal{G} \subset \mathcal{F}$

1) Basic properties

a) $E[E[X|\mathcal{G}]] = E[X]$ $(Z=1)$

b) $E[X | \mathcal{F}] = X$ ($\mathcal{F}$ ← information) $(Z = 1_{\{X > E[X|\mathcal{F}]\}})$

c) $E[X | \{\phi, \Omega\}] = E[X]$ ($\{\phi, \Omega\}$ ← trivial σ-algebra) (Constant fn, and choose $Z=1$)

2) Linearity

If $X_1, X_2, \dots, X_n \in L^1(P)$ and
$a_1, a_2, \dots, a_n \in \mathbb{R}$ a.s.

$$E\left[\sum_{j=1}^{n} a_j X_j \;\middle|\; \mathcal{G}\right] = \sum_{j=1}^{n} a_j E[X_j | \mathcal{G}]$$

Pf: $E[Z E[(aX_1 + bX_2) | \mathcal{G}]] = E[Z(aX_1 + bX_2)]$

$= a E[Z E[X_1|\mathcal{G}]] + b E[Z E[X_2 | \mathcal{G}]]$

3) Jensen: If φ is convex and $\varphi(X) \in L^1(P)$

$$\varphi(E[X \mid \mathcal{G}]) \leq E[\varphi(X) \mid \mathcal{G}]$$

(In particular $\varphi(x) = |x|$)

Pf: $\varphi(x) = \sup_{a,b} \{ax+b : ay+b \leq \varphi(y) \ \forall y\}$

In fact $\varphi(x) = \sup_{a,b \in \mathbb{Q}^2} \{ax+b : ay+b \leq \varphi(y) \ \forall y\}$

Call $E_\varphi = \{(a,b) \in \mathbb{Q}^2 : ay+b \leq \varphi(y) \ \forall y\}$

One inequality is obvious. The other follows from the existence of the supporting line

$$\varphi(x) \geq \varphi(x_0) + (x-x_0)c_{x_0}$$

Fix a,b st for any $ay+b \leq \varphi(y) \ \forall y$. Then

$\varphi(X(\omega)) \geq aX(\omega)+b$. For any fixed $Z \in C_b$

$$E[Z\,E[\varphi(X) \mid \mathcal{G}]] \geq E[Z(aX(\omega)+b)]$$

$$\geq aE[ZX] + b = aE[Z\,E[X \mid \mathcal{G}]] + b$$

$$= E[Z(aE[X \mid \mathcal{G}]+b)] \quad \forall a,b \in \mathbb{Q}^2 \text{ st } ay+b \leq \varphi(y) \ \forall y$$

One needs to write $\sup E[Z(aE[X \mid \mathcal{G}]+b)]$

$$= E[Z \sup(aE[X \mid \mathcal{G}]+b)]$$

$$= E[Z\,\varphi(E[X \mid \mathcal{G}])]$$

But how does one pass the sup inside?

Instead we use le Gall: since $\varphi(X) \geq aX+b$

$$E[\varphi(X) \mid G] \geq aE[X \mid G] + b \quad \text{a.s.}$$

This is true $\forall\ (a,b) \in E_\varphi$, a countable set.

$$\Rightarrow E[\varphi(X) \mid G] \geq \sup_{a,b \in E_\varphi} aE[X \mid G] + b = \varphi(E[X \mid G]) \quad \text{a.s.}$$

We want to prove the MCT but this is hard to do since we can have $X_n \in L^1$ $X_n \uparrow X$ but $X \notin L^1$. So $E[X_n \mid G]$ is well defined, but $E[X \mid G]$ is not defined.

So let us define $E[X \mid G]$ $\forall\ X \geq 0$, not just $X \in L^1$.

Def: Let $X \geq 0$ $E[X \mid G] := \lim_{n \to \infty} E[X \wedge n \mid G]$. ★2

$X \wedge n$ is bounded and $E[X \wedge n \mid G] \geq E[X \wedge m \mid G]$ a.s. $\forall\ n \geq m$. So $E[X \wedge n \mid G]$ is a.s. increasing, and so the limit in ★2 exists.

MCT. Suppose $X_n \geq 0$ and $X_n \uparrow X$ a.s. Then

$$\lim_{n \to \infty} E[X_n \mid G] = E[X \mid G]$$

4) Fatou, BCT, DCT

5) Conditional Hölder

$$|E[XY \mid \mathcal{G}]| \leq E\left[|X|^p \mid \mathcal{G}\right]^{1/p} E\left[|Y|^q \mid \mathcal{G}\right]^{1/q}$$

6) Conditional Minkowski holds (triangle inequality for L^p norms)

Does it tally with the classical notion of conditional expectation? (Use the defining property and go back to basics)

let B be a set s.t. $P(B) \neq 0$. Then

$$E[X \mid 1_B] = E[X \mid \underbrace{\{B, B^c, \Omega\}}_{G}]$$

POLL

What is the set of G measurable fns?

A

$\{1_B, 1_{B^c}, 1_\Omega\}$

B

$a\,1_B + b\,1_{B^c}$, $a,b \in \mathbb{R}$

Thus

$$E[X \mid 1_B] = E[X \mid \sigma(1_B)] = p\,1_B + q\,1_{B^c}$$

The defining property says for
$\xi = a 1_B + c 1_{B^c}$

$$E[\xi\, E[X \mid 1_B]] = a p\, \mathbb{P}(B) + b q\, \mathbb{P}(B^c)$$

$$= E[a X 1_B + b X 1_{B^c}]$$

$$\Rightarrow p = \frac{E[X 1_B]}{\mathbb{P}(B)} \qquad q = \frac{E[X 1_{B^c}]}{\mathbb{P}(B^c)}$$

☆

Ex: $E[X \mid Y] = \sum 1_{\{Y = a_i\}} E[X \mid Y = a_i]$

L^2-projection

Prop 8.4: $X \in L^2(P)$, $\mathcal{G} \subset \mathcal{F}$. Then for every r.v. Y that is $\mathcal{G}$-meas.

$$E\left[\left(X - E[X|\mathcal{G}]\right)^2\right] \leq E\left[(X-Y)^2\right]$$

$$\| \quad \|_2 \leq \| \quad \|_2$$

Let $H = L^2(\Omega, \mathcal{F}, P)$

$M = L^2(\Omega, \mathcal{G}, P)$ $\quad M \subset H$

Let $\Pi : H \longrightarrow M$ be the L^2 orthogonal projection operator. It satisfies

1) $\Pi^2 = \Pi$

2) $(x - \Pi x, y) = 0 \quad \forall x \in H, y \in M$

Then $\Pi f = E[f|\mathcal{G}]$

$$(x, y) = (\Pi x, y)$$

$$\int XY \, dP = \int \Pi X \, Y \, dP$$

Ex: Show that the L^2 orthogonal projection operator onto a subspace is unique (upto a.s. equivalence)

By the defining property, this shows $\Pi X = E[X|\mathcal{G}]$ a.s.

Theorem 8.5 (Tower Property)

Let $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$ be subalgebras and suppose $X \in L^1(P)$

$$E\left[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1 \right] = E[X \mid \mathcal{G}_1]$$

↑ information interpretation

Pf:

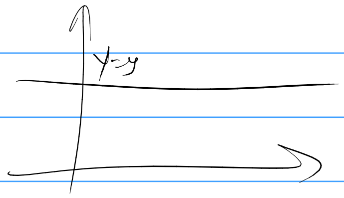
Take Z $\mathcal{G}_1$ meas. Then it's definitely $\mathcal{G}_2$ meas as well.

LHS: $E\left[Z \, E\left[E[X \mid \mathcal{G}_2] \mid \mathcal{G}_1 \right]\right] = E\left[Z \, E[X \mid \mathcal{G}_2] \right] = E[ZX]$

RHS: $E\left[Z \, E[X \mid \mathcal{G}_1] \right] = E[ZX]$

Let X, Y be RVs w/ joint density $f(x,y)$ s.t. $\int_{\mathbb{R}} f(x,y)\,dx > 0\ \forall y$, $E|X| < \infty$. What is $E(X|Y)$? Given $X \in L^1(\Omega)$

Intuitively



given Y, best guess for X is its expected value on the line corresp. to the value of Y.

$$E(X|Y) = h(Y) \quad \text{where} \quad h(y) = \int x\, f(x,y)\,dx \Big/ \int f(x,y)\,dx$$

Pf: $h(Y)$ is $\sigma(Y)$-m'ble. $\forall A \in \sigma(Y)$. $h \in L^1$ by previous.

$\sigma(Y)$ - smallest σ-alg. s.t. $Y: \Omega \to \mathbb{R}$ is m'ble : $\sigma(Y) = Y^{-1}(\mathcal{B})$

$\Rightarrow \exists B \in \mathcal{B}$ s.t. $A = Y^{-1}(B)$ i.e. $(x,y) \in A$ iff $y \in B$.

$$E[1_A\, h(Y)] = \iint 1_B(y) h(y) f(x,y)\,dx\,dy \overset{?}{=} E[X 1_A]$$

$$= \iint 1_B(y)\, x\, f(x,y)\,dx\,dy \overset{\text{Fubini}}{=} \int 1_B(y) \int x\, f(x,y)\,dx\,dy$$

$$= \int 1_B(y)\, h(y) \int f(x,y)\,dx\,dy = \iint 1_B(y) h(y)\, f(x,y)\,dx\,dy$$

which is what we wanted to show.

Prop: If X and Y are rvs, and let Y be $\mathcal{G}$ meas. If X, Y are bounded or $YX \in L^1$ then

$$E[YX \mid \mathcal{G}] = Y E[X \mid \mathcal{G}] \quad a.s.$$

Pf. Let Z be $\mathcal{G}$ meas and bdd. Assume $X, Y \geq 0$. Let $Y_n = Y \wedge n$

$$E[Z\, E[Y_n X \mid \mathcal{G}]] = E[Z Y_n X]$$

since ZY_n is bounded.

$$E[Z Y_n E[X \mid \mathcal{G}]] = E[Z Y_n X]$$

$\Rightarrow$ $E[Y_n X \mid \mathcal{G}] = Y_n E[X \mid \mathcal{G}]$ and using conditional MCT.

$$\lim_{n \to \infty} E[Y_n X \mid \mathcal{G}] = E[YX \mid \mathcal{G}] = Y E[X \mid \mathcal{G}]$$

The integrable case follows from $X = X^+ - X^-$

Prop: $\mathcal{G}_1, \mathcal{G}_2 \subset \Sigma$ are sub σ-fields. They are independent iff $\forall\ B \in \mathcal{G}_2$ we have

$$\mathbb{E}[1_B \mid \mathcal{G}_1] = \mathbb{P}(B)$$

Pf: ($\Rightarrow$) Let $\mathcal{G}_1, \mathcal{G}_2$ be indep and Z be bdd and $\mathcal{G}_2$ meas.

$$\mathbb{E}[Z\, 1_B] \overset{\text{indep.}}{=} E[Z] P(B) \overset{P(B) \text{ is a const}}{=} E[Z\, P(B)]$$

$$= E[Z\, E[1_B \mid \mathcal{G}]] \qquad \text{By prev. } P(B) = E[1_B \mid \mathcal{G}_1] \text{ a.s.}$$

$\Leftarrow$ If $A \in \mathcal{G}_1$ $\quad P(A \cap B) = E[1_A 1_B] = E[1_A E[1_B | \mathcal{G}_1]]$

$= P(B) E(1_A)$

$\boxtimes$

Martingales

Example/motivation:

Let $X_0, X_1, X_2, \ldots$ be a s/ce of RVs on Ω

Z - any RV on Ω

We start on day zero & gain new information every day & observe the RV X_n on day n.

The information we have on day n can be represented by a σ-alg $\mathcal{F}_n$.

We'll have $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$

e.g. if the only info we gain on day n is the value of X_n, we'll have $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

E.g.: X_n - value of a stock on day n.

Say we're on day n. What is the best guess for X_{n+1}?

i.e. $E(X_{n+1} \mid$ info we have on day $n)$

$E(X_{n+1} \mid \mathcal{F}_n)$.

assuming no insider info, dividends, etc,

best guess for $E(X_{n+1} \mid \mathcal{F}_n)$ should be X_n.

($\mathcal{F}_n$ doesn't need to be $\mathcal{F}_n = \sigma(X_0, X_1, X_2, \ldots, X_n)$.

e.g. maybe we're also keeping track of other stocks daily,

say $Y_0, Y_1, Y_2, \ldots$

$Z_0, Z_1, Z_2, \ldots$

so $\mathcal{F}_n = \sigma(X_0, \ldots, X_n, Y_0, \ldots, Y_n, Z_0, \ldots, Z_n)$.)

Defn: A stochastic process is a collection of RVs on $(\Omega, \mathcal{F}, P)$

If $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots \subseteq \mathcal{F}$ are sub σ-algebras, then $\{\mathcal{F}_n\}_{n\geq 1}^{\infty}$ is called a filtration.

The process $\{X_n\}_{n\geq 1}^{\infty}$ is adapted to the filtration $\{\mathcal{F}_n\}_{n\geq 1}^{\infty}$ if X_n is $\mathcal{F}_n$-m'ble $\forall n \geq 1$.

(think of it as, if have the info $\mathcal{F}_n$, then know the value of X_n).

Ex: $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Then $\{X_n\}_{n\geq1}^\infty$ is always adapted to $\{\mathcal{F}_n\}_{n\geq1}^\infty$

Def: A stochastic process $X = \{X_n\}_{n\geq1}^\infty$ is a martingale wrt the filtration $\mathcal{F} = \{\mathcal{F}_n\}_{n\geq1}^\infty$ if

1) X_n is adapted to $\mathcal{F}$
2) $X_n \in L^1(P)$ $\forall n \geq 1$
3) $E[X_{n+1} | \mathcal{F}_n] = X_n$ a.s. $\forall n \geq 1$.

X is a submartingale if

3') $E[X_{n+1} | \mathcal{F}_n] \geq X_n$ a.s. $\forall n \geq 1$

X is a supermartingale if

3'') $E[X_{n+1} | \mathcal{F}_n] \leq X_n$ a.s. $\forall n \geq 1$

Rmk: Martingale "the best guess for X_{n+1} given present info is X_n."
Super vs sub: If a supermartingale agrees w/ a "similar" martingale now, then in the past it would tend to be larger

Rmk: martingale $\Leftrightarrow$ superm & subm

Ex: Simple RW. $X_1, X_2, \ldots$ iid $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$
(example of fair game)

$S_k = X_1 + \cdots + X_k$ – cumulative fortune after k steps.

$$E[X_k | X_1, \ldots, X_{k-1}] \overset{\text{indep'ce}}{=} E X_k = 0$$

so $E[S_k | X_1, \ldots, X_{k-1}] = S_{k-1}$ so $S = \{S_n\}_{n\geq1}^\infty$ is a martingale wrt the filtration $\mathcal{F} = \{\mathcal{F}_n\}_{n\geq1}^\infty$ w/ $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$.

Rmk: Only needed indep'ce & $E X_i = 0$ $\forall i$, to get S martingale wrt $\mathcal{F}$. No need for identically distributed

Ex: If $\mathcal{F}$ is a filtration & $Y \in L^1(P)$, $X_n := E[Y | \mathcal{F}_n]$ is a martingale (these are called Doob martingales).

These are called "closed" by le Gall.

Def: A stoch. pr. $\{A_i\}_{i\geq 1}^{\infty}$ is previsible wrt filtration $\mathcal{F}=\{\mathcal{F}_n\}_{n\geq 1}^{\infty}$ if A_n is $\mathcal{F}_{n-1}$ m'ble $\forall n\geq 1$, where $\mathcal{F}_0=\{\emptyset, \Omega\}$.

Rmk: Previsible is also called predictable.

Ex: $X_1, X_2, X_3, \ldots$ indep. $EX_i=0$ (fair game).

$S_n = X_1+\dots+X_n$ — winnings after n'th round.

Of course S_n - martingale

What if we change the bets every day?

If we bet A_n on day n, our overall winnings will be

$$Y_n = Y_{n-1} + A_n X_n = Y_{n-1} + A_n(S_n - S_{n-1})$$
$$= \underset{\substack{\uparrow \\ \text{some starting amount}}}{y_0} + \sum_{i=1}^{n} A_i(S_i - S_{i-1})$$

How much we bet can only be based on info we had up to that point so $A_n \in \mathcal{F}_{n-1}$, i.e. $A=\{A_n\}_{n\geq 1}^{\infty}$ is previsible.

More generally

Martingale transforms

Let $\{S_n\}_{n=1}^{\infty}$ be a martingale wrt filtration $\mathcal{F}=\{\mathcal{F}_i\}_{i\geq 1}^{\infty}$.

Let $S_0=0$, y_0 - const & $A=\{A_n\}_{n\geq 1}^{\infty}$ previsible pr wrt $\mathcal{F}$.

Define pr Y by $Y_n := y_0 + \sum_{j\geq 1}^{n} A_j(S_j - S_{j-1})$ $\forall n \geq 0$.

Exercise: Y is a martingale

Y is called a martingale transform of S.

Le Gall writes if $(A_n)_n$ previsible, $(X_n)_n$ is a Martingale,

$$(A \bullet X)_n = \sum_{i=1}^{n} A_i(X_i - X_{i-1}) \text{ "="} \int A\, dX$$

Prop: if (X_n) is a super(sub)martingale and (H_n) previsible and nonnegative Then $(H \bullet X)_n$ is a super (sub) martingale.

Example: $\Omega = \{-1,+1\}^{\mathbb{Z}_+}$, $P = \mu^{\otimes \mathbb{Z}_+}$ (product meas.)
let $Y_n(\omega) = \omega_n$ represent the outcome of the game.
If its fair $E[Y_n(\omega)] = 0$.

Then $S_n = \sum_{i=1}^{n} Y_i$ is a martingale.

If $H_n = f(Y_1, \dots, Y_{n-1})$ is a previsible "bet" then if you win you gain $H_n(S_n - S_{n-1}) = H_n$ and if you lose you gain $H_n(S_n - S_{n-1}) = -H_n$. Thus your net winnings after n games is $(H \bullet S)_n$.

Another transform:

Prop: let $\varphi: \mathbb{R} \to \mathbb{R}$ be convex, and let X_n be $\mathcal{F}_n$ adapted.

1) If (X_n) is a martingale, and $E|\varphi(X_n)| < \infty$ then $\varphi(X_n)$ is a submartingale.

2) If $(X_n)_{n \in \mathbb{Z}_+}$ is a submartingale and φ is increasing then $\varphi(X_n)$ is a submartingale.

Pf: 1) $E[\varphi(X_{n+1}) \mid \mathcal{F}_n] \geq \varphi(E[X_{n+1} \mid \mathcal{F}_n]) = \varphi(X_n)$ (★1)

↑ This needs the L^1 condition to be defined. If $\varphi: \mathbb{R} \to \mathbb{R}_+$ as in Le Gall, then can drop $E|\varphi(X_n)| < \infty$. Recall that we defined $E[X \mid \mathcal{F}] := \lim_{n\to\infty} E[X \wedge n \mid \mathcal{F}]$ if $X \geq 0$.

2) Similarly, if X_n is only a submartingale, the last step has $E[X_{n+1} \mid \mathcal{F}_n] \geq X_n$ then we get (★1).

Properties of martingales

Lemma: If X is a subm. wrt filtration $\mathcal{F}$, then it is a subm. wrt the filtration generated by X itself, i.e.

$$\mathbb{E}[X_{n+1} \mid X_1, \dots, X_n] \geq X_n \text{ a.s.}$$

Pf: Know $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad \forall n$.

X_n – $\mathcal{F}_n$-m'ble $\forall n$ & $\mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_{n-1} \subseteq \mathcal{F}_n \Rightarrow \sigma(X_1, \dots, X_n) \subseteq \mathcal{F}_n$.

$$\mathbb{E}\left[\mathbb{E}[X_{n+1}|\mathcal{F}_n] \,|\, X_1,\dots,X_n\right] \geq \mathbb{E}[X_n|X_1,\dots,X_n]$$
$$\mathbb{E}[X_{n+1}|X_1,\dots,X_n] \geq X_n$$
$\triangle$.

Lemma: If X is a martingale & φ is convex s.t. $\varphi(X_n) \in L^1(P)\ \forall n$ then $\varphi(X)$ is a subm. If X subm & φ - non-decr. convex, w/ $\varphi(X_n) \in L^1(P)\ \forall n$, then $\varphi(X)$ is also subm.

Pf: Conditional Jensen

$$\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \geq \varphi\,\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \varphi(X_n). \quad \text{if } X \text{ martingale}$$

If X - subm & φ - non decr, then

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$$

so $\varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) \geq \varphi(X_n)$ $\triangle$

Cor: If X - martingale, then X^+, $|X|^p$, e^X are subm. if they stay in $L^1(P)$ $\forall n$.

If X - subm., then X^+, e^X also subm if in $L^1(P)\ \forall n$.

If X - subm, $|X|$ might not be.

e.g. $X_k := -\frac{1}{k}$.

Thm: (Doob's decomposition thm).

Any submartingale X can be written as

$X_n = Y_n + Z_n$ where Y is a martingale & Z is a non-neg. previsible a.s. increasing process w/ $Z_n \in L^1(P)\ \forall n$

Pf: Set $X_0 := 0$ & $d_j = X_j - X_{j-1}$ so $X_n = d_1 + \dots + d_n$.

Suppose have Y_n, Z_n, want to find out Y_{n+1}, Z_{n+1}.

Let $e_{n+1} = Y_{n+1} - Y_n$, $f_{n+1} = Z_{n+1} - Z_n$. Have $d_{n+1} = e_{n+1} + f_{n+1}$

Must have Z_{n+1} is $\mathcal{F}_n$-m'ble & $\mathbb{E}(e_{n+1}|\mathcal{F}_n) = 0$

$$\mathbb{E}(d_{n+1}|\mathcal{F}_n) = \mathbb{E}(e_{n+1}|\mathcal{F}_n) + \mathbb{E}(f_{n+1}|\mathcal{F}_n)$$
$$\text{need } = 0 + f_{n+1}$$

So should take $f_{n+1} = \mathbb{E}(d_{n+1}|\mathcal{F}_n)$

$$e_{n+1} = d_{n+1} - \mathbb{E}(d_{n+1}|\mathcal{F}_n).$$

Thus $Z_n = \sum_{k=1}^{n} E(d_k | \mathcal{F}_{k-1})$

$Y_n = X_n - Z_n$

Hw: Check these work. △

Another decomp. thm

Def: $\{X_i\}_{i=1}^{\infty}$ is bdd in $L^r(P)$ if $\sup_n \|X_n\|_r < \infty$.

Thm: (Krickeberg's decomposition)
If X is a subm. bdd in $L^1(P)$ then can write it as
$X_n = Y_n - Z_n$ w/ Y_n - mart., Z_n - non-neg. superm.

Pf: Set $Y_n = \lim_{m\to\infty} E[X_m | \mathcal{F}_n]$. Does the limit exist?
Yes since it is a.s. increasing (X is a subm.)

Let $m \geq n$. $E[X_{m+1} | \mathcal{F}_n] \geq E[E[X_{m+1}|\mathcal{F}_m] | \mathcal{F}_n] \geq E(X_m | \mathcal{F}_n) \geq X_n$
(use towering property) So $Y_n \geq X_n \Rightarrow Z_n = Y_n - X_n$ is non-neg.

Check Y_n - martingale. Certainly adapted.

$E Y_n = \lim_{m\to\infty} E[E[X_m | \mathcal{F}_n]] = \lim_m E X_m \overset{\text{subm}}{=} \sup_m E X_m \overset{X_n \text{ bdd in } L^1(P)}{<} \infty$

$E[Y_{n+1} | \mathcal{F}_n] = E[\lim_{m\to\infty} E[X_{m+1} | \mathcal{F}_{n+1}] | \mathcal{F}_n] \underset{\text{cond'l MCT}}{=} \lim_{m\to\infty} E[E[X_{m+1}|\mathcal{F}_{n+1}]|\mathcal{F}_n]$

$= \lim_{m\to\infty} E[X_{m+1} | \mathcal{F}_n] = Y_n$

So Y_n - mart. $\Rightarrow Z_n$ - superm. △

Rmk: Y & Z are L^1-bdd.

Pf: $Y_n = X_n + Z_n$ $\qquad Z_n \geq 0$

$E|Y_n| \leq \underbrace{\sup_k E|X_k|}_{\text{indep of } n,\text{ finite}} + E Z_n$.

$|E Z_n| = |E(Y_n - X_n)| = |E Y_1 - E X_n| \leq |E Y_1| + \underbrace{\sup_k E|X_k|}_{\text{indep of } n,\text{ finite}}$. △

Doob + Krickeberg: submartingales are bounded above and below by martingales.

Thm: ∀ bdd L^1-bdd subm is bdd above & below by martingales.

Pf: Doob — ∀ subm bdd below by mat.
Krickeberg — ∀ L^1-bdd subm bdd above by mat. △

Stopping times

Defn: A stopping time wrt a filtration $\mathcal{F}$ is a RV $T: \Omega \to \mathbb{N} \cup \{\infty\}$ s.t. $\{T = k\} \in \mathcal{F}_k$ ∀ $k \in \mathbb{N}$

(equiv. $\{T \leq k\} \in \mathcal{F}_k$ ∀ $k \in \mathbb{N}$)

$\{T = +\infty\} = \bigcap_{n \in \mathbb{Z}_+} \{T \leq n\}^c$ Each of these are $\mathcal{F}_n$ meas.

$\Rightarrow \{T = +\infty\} \in \mathcal{F}_\infty$

Ex 1) $T = n$ a.s is a stopping time since $\{T = n\} \Delta \Omega = \phi$

2) If $A \in \mathcal{B}(\mathbb{R})$, $T_A = \inf\{n : X_n \in A\}$ is a stopping time.

$\{T_A = k\} = \{X_1 \notin A, \dots, X_{k-1} \notin A, X_k \in A\} \in \mathcal{F}_k$

Actually, every stopping time can be expressed as a 1st visit time

If T - stopping time, define

$X_j(\omega) := \mathbb{1}_{\{T = j\}}(\omega)$.

Then T is the 1st time the Stoch. pr. $\{X_j\}_{j \geq 1}^{\infty}$ visits $A = \{1\}$

This may not be that useful in general.

3) $L_A = \sup\{n : X_n \in A\}$ is not in general.

$\{L_A = k\} = \{X_k \in A, X_{k+1} \notin A, \dots\}$

4) S, T stopping times $\{\max(S,T)=k\} = \{S=k, T\leq k\} \cup \{T=k, S\leq k\} \in \mathcal{F}_k$

Hwk: Check that if $\{T_i\}_{i=1}^n$ are stopping times then so are $T_1 + \dots + T_n$, $\min_{1\leq i\leq n} T_i$, $\max_{1\leq i\leq n} T_i$

Stopped σ-algebras

$$\mathcal{F}_T := \{A \in \mathcal{F} \mid A \cap \{T \leq k\} \subseteq \mathcal{F}_k \ \forall k \in \mathbb{N}\}.$$

(note this is not a random σ-alg, you don't choose $T=k$ & pick $\mathcal{F}_k$)

Hwk: $\mathcal{F}_T$ - σ-alg

Lemma: 1) If S, T - stopping times st $P(S<\infty) = P(T<\infty) = 1$, & $S \leq T$ a.s., then $\mathcal{F}_S \subseteq \mathcal{F}_T$.

2) $E[Y | \mathcal{F}_T] \mathbb{1}_{T=n} = E[Y | \mathcal{F}_n] \mathbb{1}_{T=n} \quad \forall Y \in L^1(P), \ \forall n \geq 1.$

Pf 1) If $A \in \mathcal{F}_S$, then $A \cap \{T=n\} = \bigcup_{k=0}^{\infty} (A \cap \{S=k\}) \cap \{T=n\}$

$= \bigcup_{k=0}^{n} (A \cap \{S=k\}) \cap \{T=n\} \in \mathcal{F}_n$ since $S \leq T$

$\Rightarrow A \in \mathcal{F}_T$

Skip 2)

Prop If S, T stopping times, $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$

$\mathcal{F}_{S\wedge T} \subset \mathcal{F}_S \cap \mathcal{F}_T$ from 1)

If $A \in \mathcal{F}_S \cap \mathcal{F}_T$ $\quad A \cap \{S \wedge T > n\} = A \cap (\{S>n, T>n\})$

$= \underbrace{(A \cap \{S>n\})}_{\in \mathcal{F}_n} \cup \underbrace{(A \cap \{T>n\})}_{\in \mathcal{F}_n}$

$\Rightarrow A \in \mathcal{F}_{S\wedge T}$

Rem: The whole point of $\mathcal{F}_T$ is to find a σ algebra st $X_T := X_{T(\omega)}(\omega)$ is meas. wrt to it. One option is to consider $\sigma(T)$, but if $T = \text{cos}$, then $\sigma(T)$ is trivial.

Prop Let (Y_n) be $\mathcal{F}_n$ adapted and let T be a stopping time. Then Y_T is well defined on $\{T < \infty\}$ and is meas. wrt $\mathcal{F}_T$.

Pf: $\{Y_T \in B\} \cap \{T = n\} = \{Y_n \in B\} \cap \{T = n\}$
$\in \mathcal{F}_n$ $\in \mathcal{F}_n$

Thm: (Optional stopping, easy) Let $(X_n)_{n \in \mathbb{Z}}$ be a martingale (or super) Let T be a stopping time, then $(X_{n \wedge T})$ is a martingale (or super)

a) If $T \leq K$ a.s. $E[X_T] = E[X_0]$

b) If $|X_n| \leq K$ a.s., and $T < \infty$ a.s., $E[X_T] = E[X_0]$

Pf: Let $H_n = \mathbb{1}_{\{T \geq n\}}$ $\mathcal{F}_{n-1}$ meas.

$$X_{n \wedge T} = \sum_{i=1}^{n \wedge T} X_i - X_{i-1} + X_0 = X_0 + \sum_{i=1}^{n} \mathbb{1}_{\{i \leq T\}} (X_i - X_{i-1})$$

$$= X_0 + (H \bullet X)_n$$

For a) $\lim_{n \to \infty} E[X_{n \wedge T}] = E[X_{K \wedge T}] = E[X_0]$

b) $\lim_{n \to \infty} E[X_{n \wedge T}] \overset{\text{bdd}}{=} E[X_T] = E[X_0]$

Optional stopping theorem

Consider a fair game. E.g. $X_1, X_2, \ldots$ indep, $EX_i = 0 \ \forall i$,
$S_n = X_1 + \cdots + X_n$.
$ES_n = 0$, so if playing for n steps your expected winnings are 0.
What if you devise some strategy for when you stop instead of stopping after n steps? You have to stop at a stopping time (can't use info from the future).
Can you come up w/ a strategy T s.t. $ES_T > 0$?
No (saw w/ random walks, $ES_N = EX_1 EN = 0 \ \forall$ stopping time N)

Thm: (Doob's optional stopping thm).
Suppose S, T are a.s. bdd stopping times s.t. $S \le T$ a.s.
If X is a subm, then $E[X_T | \mathcal{F}_S] \geq X_S$ a.s.
supm $\leq$
martingale $=$

pf: Let $k \in \mathbb{N}$ be s.t. $S \le T \le k$ a.s. (k – const).
X – subm.
Write $X_n = d_1 + d_2 + \cdots + d_n$ (i.e. $d_k := X_k - X_{k-1}, \ \forall k \in \mathbb{N}$ w/ $X_0 = 0$).
X – subm. $\Rightarrow E[d_j | \mathcal{F}_{j-1}] \geq 0$ a.s. $\forall j$

$$X_T = \sum_{j=1}^{k} d_j \mathbb{1}_{j \le T} \quad \text{ditto for } X_S.$$

ETS $E[X_T - X_S; A] \geq 0 \quad \forall \ A \subseteq \mathcal{F}_S$.

$$E[X_T - X_S; A] = \sum_{j=1}^{k} E[d_j \mathbb{1}_{S < j \le T}; A]$$

$$= \sum_{j=1}^{k} E[d_j \mathbb{1}_{\{S < j \le T\} \cap A}]$$

$$= \sum_{j=1}^{k} \mathbb{E}\left[\mathbb{E}\left[d_j \mathbb{1}_{\{S<j\le T\}\cap A} \mid \mathcal{F}_{j-1}\right]\right]$$

$\{S<j\}\cap A \cap \{T<j\}^c \cap A$, each $\in \mathcal{F}_{j-1}$

$$= \sum_{j=1}^{k} \mathbb{E}\left[\mathbb{E}[d_j \mid \mathcal{F}_{j-1}]\, \mathbb{1}_{\{S<j\le T\}\cap A}\right] \geq 0. \quad \triangle$$

Cor: If T is a stopping time wrt $\mathcal{F}$ & X submart. (supermartingale), then so is $\{X_{T\wedge n}\}_{n\ge1}^{\infty}$ wrt $\{\mathcal{F}_{T\wedge n}\}_{n\ge1}^{\infty}$.

Pf: apply prev thm w/ $T = T\wedge n+1$ & $S = T\wedge n$ $\triangle$

Read Section 8.4 to see how to obtain the Random walk results from earlier via martingale theory. Especially Thm 8.34.

St. Petersberg paradox

Recall $\Omega = \{-1,+1\}^{\mathbb{Z}_+}$, $P = \mu^{\otimes \mathbb{Z}_+}$ (product meas.)

let $Y_n(\omega) = \omega_n$ represent the outcome of the game.

If its fair $E[Y_n(\omega)] = 0$.

Let H_0 be initial bet, $H_n = 2H_{n-1}$ (double)

Let $T = \inf\{n : Y_n = 1\}$ (quits)

Net winnings after n games is $(H \bullet S)_n$.

We care about $(H \bullet S)_T = W_T$

$$W_T = \sum_{i=0}^{T-1} H_i Y_i + H_T Y_T \quad \text{on } T<\infty$$

$$= -H_0 \sum_{i=1}^{T-1} 2^i + H_0 2^T$$

$$= H_0 \left[-\left(\frac{2^T - 1}{2-1}\right) + 2^T \right] = H_0$$

If $P(T<\infty) = 1$ we have ensured $E[W_T] = H_0$

$\neq E[W_0] = E[Y_0 H_0] = 0$

Gambler's ruin

Consider the previous example, but now, let $H_n = 1$. Let $S_0 = k$ represent your initial fortune, and if $S_n = M$, then you bankrupt Mr. House.

Let $T = \inf\{n : S_n = 0 \text{ or } M\}$

Let $A = \{S_T = 0\}$ this is $\mathcal{F}_T$ meas.

$S_n - k$ is a martingale. $|S_{n\wedge T} - k| \leq k + M$, so stopping theorem gives

$$E[S_T - k] = E[S_0 - k] = 0$$

$$\Rightarrow \quad -k\,P(S_T = 0) + M\,(1 - P(S_T = 0)) = 0$$

$$\Rightarrow \quad P(S_T = 0) = \frac{M-k}{M}$$

Need to show $P(S_T = 0) + P(S_T = M) = 1$

But this is equivalent to showing $P(T < \infty) = 1$ (which was needed for the stopping theorem anyway)

How to show? $P(0 < S_n < M) = P\left(0 < \frac{S_n}{\sqrt{n}} < \frac{M}{\sqrt{n}}\right)$

$$\leq P\left(0 \leq \frac{S_n}{\sqrt{n}} \leq \epsilon\right) \longrightarrow \int_0^\epsilon \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx \leq C\epsilon$$

But

$$P(T = +\infty) = P\left(\bigcap_n \{0 < S_n < M\}\right) \leq P(0 < S_n < M)$$

Ballot theorem

Given A receives α votes & B receives β, what's the chance that A always strictly leads B in the vote counting process?

This can be proved in two different ways.

1) Using the reflection principle
2) Using "backward" martingales.

1)

Assume all vote counting sequences are equally likely.
$n = \alpha + \beta$. Let

$$X_i \in \begin{cases} +1 & \text{if vote for A} \\ -1 & \text{if vote for B} \end{cases}$$

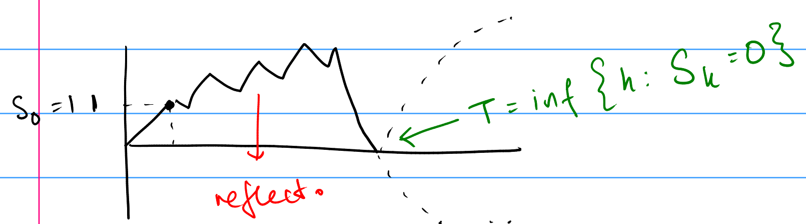
X_i represents the i^{th} vote.
Let $S_k = \sum_{i=1}^{k} X_i$

Find sequences $\vec{X}_n = (X_1, \dots X_n)$ st $S_k > 0 \quad \forall k = 1, \dots, n$.
$(S_k)_{k=1}^{n}$ is clearly a SRW. We need $\alpha > \beta$

$$\# \{ \vec{X}_n : S_0 = 0, S_k > 0 \quad \forall k, S_n = \alpha - \beta \}$$

$$= \# \{ \vec{X}_n : S_1 = 1, S_k > 0 \quad k = 2, \dots, n, \quad S_n = \alpha - \beta \}$$

$$= \# \{ \vec{X}_{n-1} \in \{\pm 1\}^{n-1} : S_0 = 1, S_k > 0, k = 1, \dots, n-1, \quad S_{n-1} = \alpha - \beta \}$$



Consider the set

$\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta\}$. If $T \le n$, that is, the trajectory hits the y-axis then the trajectory is not good; i.e.,

$\vec{X}_{n-1} \notin \{\vec{X}_{n-1} \in \{\pm 1\}^{n-1} : S_0 = 1, S_k > 0, k = 1, \dots, n-1, S_{n-1} = \alpha - \beta\}$

But for trajectories in this set, we can replace the initial part of the trajectory with each $(X_1, \dots X_T)$ replaced by $(-X_1, \dots, -X_T)$

Then it's easy to see that the reflected trajectory starts at $S_0 = -1$ and ends up at $S_{n-1} = \alpha - \beta$. In fact

$$\#\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta\} = \#\{\vec{X}_{n-1} : S_0 = -1, S_{n-1} = \alpha - \beta\}$$

In $n-1$ steps of ± 1, we have to end up at $\alpha - \beta + 1$

$\#\{\text{Up steps}\} - \#\text{down steps} = \alpha - \beta + 1$

$$a - (n-1-a) = \alpha - \beta + 1 \Rightarrow 2a = n - 1 + \alpha - \beta + 1$$

$$a = \frac{\alpha + \beta + \alpha - \beta}{2} = \alpha.$$

So $\#\{\vec{X}_{n-1} : S_0 = -1, S_{n-1} = \alpha - \beta\} = \binom{n-1}{\alpha}$

$$\#\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta\} = \binom{n-1}{\alpha - 1}$$

↑ one upstep is already taken

$$\#\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta, T > n\} = \#\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta\}$$
$$- \#\{\vec{X}_{n-1} : S_0 = 1, S_{n-1} = \alpha - \beta, T \le n\}$$

$$= \binom{n-1}{\alpha-1} - \binom{n-1}{\alpha} = \frac{(n-1)!}{(\alpha-1)!\,\beta!} - \frac{(n-1)!}{\alpha!\,(\beta-1)!}.$$

$$= \frac{(n-1)!}{(\alpha-1)!\,(\beta-1)!}\left[\frac{\alpha-\beta}{\alpha\beta}\right] = \frac{(n-1)!}{\alpha!\,\beta!}(\alpha-\beta)$$

$$\#\{\vec{X}_n,\ S_0=0,\ S_n=\alpha-\beta\} = \binom{n}{\alpha} = \frac{n!}{\alpha!\,\beta!}.$$

Taking the ratio gives $\frac{\alpha-\beta}{\alpha+\beta}$.

2)

<u>Backward Martingales</u>

Suppose $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ $\quad \mathcal{F}_n = \sigma(S_n, S_{n+1}, \dots)$

Then S_n is $\mathcal{F}_n$ measurable. Let $\tilde{S}_n = \frac{S_n}{n}$

$$E[\tilde{S}_n \mid \mathcal{F}_{n+1}] = E\left[\frac{S_{n+1}}{n} - \frac{X_n}{n} \,\middle|\, \mathcal{F}_{n+1}\right] = \frac{S_{n+1}}{n} - E\left[\frac{X_n}{n} \,\middle|\, \mathcal{F}_{n+1}\right]$$

— ☆1

$E[X_i \mid \mathcal{F}_{n+1}] = E[X_j \mid \mathcal{F}_{n+1}]$ $\quad i,j \le n+1$ by symmetry.

$$\Rightarrow \sum_{i=1}^{n+1} E[X_i \mid \mathcal{F}_{n+1}] = S_{n+1} \Rightarrow E[X_n \mid \mathcal{F}_{n+1}] = \frac{S_{n+1}}{n+1}$$

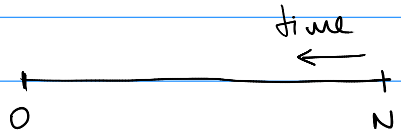
$$☆1 = \frac{S_{n+1}}{n} - \frac{S_{n+1}}{n+1} = \tilde{S}_{n+1}$$

Thus $\tilde{S}_n$ is a backwards martingale satisfying

$$E[\tilde{S}_n \mid \mathcal{F}_{n+1}] = \tilde{S}_{n+1}$$

The stopping Theorem applies to backward martingales.

Let N be some fixed horizon, and let $0 \le T \le N$ be st $\{T \ge k\}$ is $\mathcal{F}_k$ measurable.



As before $T \vee k$ is a stopping time,

$$E[\tilde{S}_T] = \lim_{k \downarrow 0} E\left[\tilde{S}_{T \vee k}\right] = \tilde{E}[S_N]$$

Remark: There is no centering here to worry about in $\tilde{S}_n$

Let us apply this to Ballots.

Let $X_i \in \{0,2\}$, where 2 represents a vote for B & 0 for A.

$\bar{X}_i = -(X_i - 1) \in \{-1,1\}$ $\quad \{\bar{S}_0 = 0, \bar{S}_k > 0, k = 1, \dots, N, \bar{S}_N = \alpha - \beta\}$

$= \{S_0 = 0, -(S_k - k) > 0, k = 1, \dots N, S_N = (N - (\alpha - \beta))\}$ $\qquad \alpha + \beta - \alpha + \beta = 2\beta$

$= \{S_0 = 0, k > S_k, k = 1, \dots N, S_N = 2\beta\}$

$= G$

Let $T = \sup\{0 \le k \le N : S_k \ge k\}$ and $T = 1$ if it doesn't happen.

$\tilde{S}_T = 1$ on G^c. $\quad S_{T+1} < T+1 \quad$ Then

$S_T \le S_{T+1} \le T$ since S_{j+1} is integer valued and $X_i \ge 0$.

$\Rightarrow S_T = T \Rightarrow \hat{S}_T = 1$.

On G, $T = 1$, and $S_1 < 1$ (since G happens)

$\Rightarrow S_1 = 0$ (since $X_1 \in \{0,2\}$)

Therefore $S_T = 0$ on G.

Then $\tilde{S}_T = 1_G$

$$E[1_G] = E[\tilde{S}_T] = E[\tilde{S}_N] = \frac{S_N}{N} = \frac{\alpha - \beta}{\alpha + \beta}$$

Crazy, huh? A more general version appears in Durrett.

Thm: (Doob's Maximal inequalities)

If X - subm, then $\forall \lambda > 0$, $\forall n \geq 1$

1) $(*)$ $\quad \lambda P\left(\max_{1 \leq j \leq n} X_j \geq \lambda\right) \leq E\left[X_n ; \max_{1 \leq j \leq n} X_j \geq \lambda\right] \leq E[X_n^+]$

2) $\quad \lambda P\left(\min_{1 \leq j \leq n} X_j \leq -\lambda\right) \leq E[X_n^+] - EX_1$

Pf: 1) Let $T = \inf\{j : X_j \geq \lambda\}$. T is a stopping time

$\{T \leq n\} = \max_{1 \leq j \leq n} X_j \geq \lambda$.

$$\text{LHS of } (*) = \lambda P(T \leq n) = \sum_{j=1}^{n} \lambda P(T = j)$$
$$\leq \sum_{j=1}^{n} E[X_j ; T = j)$$

X - subm $\Rightarrow$ $\quad E[X_j ; T=j] \leq E\left[E[X_n | \mathcal{F}_j] ; T=j\right]$

$\{T=j\} \subseteq \mathcal{F}_j \quad \Rightarrow \quad = E[X_n ; T=j]$

So LHS of $(*)$ $\leq \sum_{j=1}^{n} E[X_n ; T=j] = E[X_n ; T \leq n] \leq EX_n^+$

2) $\tau := \inf\{1 \le j \le n : X_j \le -\lambda\}$. - stopping time.
($\inf \emptyset = \infty$)

So $\{\min_{1\le j\le n} X_j \le -\lambda\} = \{\tau \le n\}$.

On $\tau \le n$ have $X_\tau \le -\lambda$ so

$E[X_\tau ; \tau \le n] \le -\lambda P(\tau \le n)$

also $E[X_n ; \tau > n] \le E[X_n^+]$

add: $E X_{\tau\wedge n} \le -\lambda P(\tau \le n) + E[X_n^+]$

$1 \le \tau \wedge n$ a.s. $\Rightarrow$ by Optional stopping

$$X_1 \le E[X_{\tau\wedge n} | \mathcal{F}_1]$$

$$E X_1 \le E \text{ (〃)} = E X_{\tau\wedge n}$$

△

Thm) Doob's Martingale cvge thm

Let X be a subm. Suppose

i) X is bdd in $L^1(P)$

or ii) $X \le 0$ a.s.

Then $\lim_{n\to\infty} X_n$ exists a.s. & is finite a.s.

Pf) Like LLN, 1st do L^2 case, ~~then use truncation~~ ~~for L^1~~

non-neg, L^2-bdd case

1st let's show X_n is Cauchy, so cv in L^2.

$X \ge 0$, L^2-bdd, so

$$\|X_{n+k} - X_n\|_2^2 = \|X_{n+k}\|_2^2 + \|X_n\|_2^2 - 2E[X_{n+k}X_n]$$

$$-2E[X_{n+k}X_n] = -2E[E(X_{n+k}|\mathcal{F}_n)X_n]$$

$$\le \|X_{n+k}\|_2^2 - \|X_n\|_2^2$$

X^2 subm $\Rightarrow$ $\|X_n\|_2 \nearrow \sup_n \|X_n\|_2 < \infty$ since L^2 bdd.

$\Rightarrow \{X_n\}_{n\ge 1}^{\infty}$ Cauchy seq in L^2 $\Rightarrow$ cv in L^2.

Let $X_\infty =$ L^2 limit of X_n

To show $X_n \to X_\infty$ a.s., use Borel-Cantelli

Need $\sum_{k\geq 1}^{\infty} P(|X_\infty - X_n| > \varepsilon) < \infty$. Might not have st. Work w/ a subseq 1st.

$\exists n_k \uparrow \infty$ s.t. $\|X_\infty - X_{n_k}\|_2 \leq 2^{-k}$.

Then $\sum_{k\geq 1}^{\infty} P(|X_\infty - X_{n_k}| \geq \varepsilon) \leq \frac{1}{\varepsilon}\sum 4^{-k} < \infty$ so $X_{n_k} \to X_\infty$ a.s.

Let's show X_j w/ $n_k \leq j \leq n_{k+1}$ are close to X_{n_k}.

Consider $\max_{n_k \leq j \leq n_{k+1}} |X_j - X_{n_k}|$.

Use Doob's max'l ineq. Apply to $X_j - X_{n_k} = X_{n_k+\ell} - X_{n_k}$

$\{X_{n_k+\ell} - X_{n_k}\}_{\ell\geq 0}^{\infty}$ submartingale $\Rightarrow$ by Doob's ineq.

(combine both parts)

$$P\left(\max_{n_k \leq j \leq n_{k+1}} |X_j - X_{n_k}| > \varepsilon\right) \leq \frac{2}{\varepsilon} E|X_{n_{k+1}} - X_{n_k}| - \frac{1}{\varepsilon}E0$$

$$\leq \frac{2}{\varepsilon}\|X_{n_{k+1}} - X_{n_k}\|_2 \leq \frac{2}{\varepsilon}\left(\|X_{n_{k+1}} - X_\infty\|_2 + \|X_{n_k} - X_\infty\|_2\right)$$

$$\leq \frac{2}{\varepsilon}\left(2^{-k} + 2^{-k-1}\right) = \frac{3}{\varepsilon}2^{-k}$$

Since this is summable argue here by Borel-Cantelli

$$\max_{n_k \leq j \leq n_{k+1}} |X_j - X_{n_k}| \xrightarrow[k\to\infty]{} 0 \quad \text{a.s.}$$

Combined w/ $X_{n_k} \to X_\infty$ a.s. get $X_n \to X_\infty$ a.s.

$X_\infty \in L^2(P) \Rightarrow X_\infty$ is a.s. finite.

<u>non-positive case</u>

If $X_n \leq 0$ subm, $\Rightarrow e^{X_n}$ bdd non-neg subm $\Rightarrow$ by prev. case $e^{X_n} \to e^{X_\infty}$ a.s. w/ e^{X_∞} finite a.s.

<u>non-neg L^1-bdd case</u>

$X_n \geq 0$, subm, bdd in $L^1(P) \Rightarrow$ by Krickeberg decomp.

$X_n = Y_n - Z_n$, Y_n - martingale, Z_n - non-neg superm.

Know pf of Krickeberg $Y_n \geq 0$.

Both Y_n, Z_n – L^1-bdd

$-Y_n$ – nonpos. subm ⇒ by prev. step $-Y_n$ cv a.s. to a.s. finite limit.

$-Z_n$ – 〃

<u>L^1-bdd case</u>

$X_n = Y_n - Z_n$ Y L^1-bdd m, Z – non-neg L^1-bdd super-

as in prev. case Z_n cv a.s. to a.s. finite limit.

$Y_n = Y_n^+ - Y_n^-$, Y_n^+, Y_n^- – non-neg L^1-bdd subm ⇒ by prev. done. △

Suggest reading applications of Martingales

e.g. Pf of Kolmogorov's 0-1 law

Lévy's Borel–Cantelli

Khintchine's law of iterated logarithms

$\{X_i\}_{i=1}^{\infty}$ iid $S_n = X_1 + \dots + X_n$

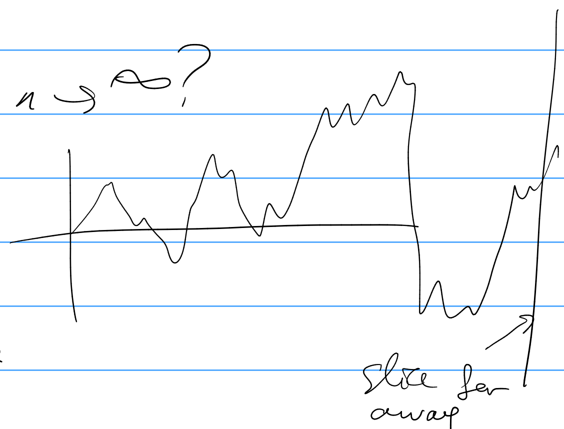
mean μ

Var σ^2

How does S_n behave as $n \to \infty$?

$\frac{S_n}{n} \xrightarrow{a.s.} \mu$

$\frac{S_n - n\mu}{\sqrt{n}} \Rightarrow N(0,1)$.



So typically S_n of order $\sqrt{n}$ away from the mean

Q: How large does it get?

Khintchine's Law of the iterated logarithm

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = 1 \text{ a.s.} \quad (\text{for liminf get } -1)$$

<u>Pf in book.</u>

Another big application - Stock market
Option pricing
Black - Scholes

Rademacher's thm

I - any interval $\subseteq \mathbb{R}$

<u>Def:</u> A fn $f: I \to \mathbb{R}$ is Lipschitz if $\exists$ const $A > 0$ s.t. $\forall x, y \in \mathbb{R}$

$$|f(x) - f(y)| \le A|x-y|.$$

The optimal A is called the Lipschitz const of f.
(Can define for f.)

Lipschitz $\Rightarrow$ Cts.

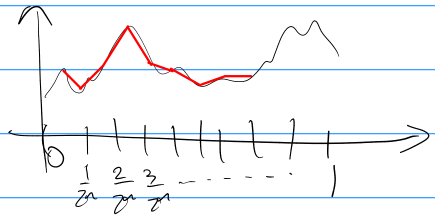
diff'ble w/ cts derivative on cpct set $\Rightarrow$ Lipschitz

<u>converse</u>

In pf: if f has deriv, f'=g, and g int'l, then can write f(x)-f(0)=\int_0^x g(t)dt. Let's try to write f as such integral.If g is cts, then FundThmCalc tells us f diff'le. We won't have g diff'le, but only int'le. Lebesgue differentiation thm gives that integrals of L1 functions are almost everywhere differentiable. What's a good candidate for g? Take approximation of g by finite difference quotients.

<u>Thm</u> (Rademacher)
If $f: I \to \mathbb{R}$ is Lipschitz where I is an interval, then f is diff'le almost everywhere.

<u>Pf:</u> Let $I = [0,1]$, P - the Leb m're on $([0,1], \mathcal{B}([0,1]))$
P - prob m're.



Divide $[0,1]$ into 2^n parts & consider the slopes on each:

$$\frac{f((j+1)2^{-n}) - f(j2^{-n})}{(j+1)2^{-n} - j2^{-n}} \quad j \in \{0,1,\dots,2^n-1\}.$$

Let $\mathcal{F}_n^0 = \{ (j2^{-n}, (j+1)2^{-n}],\ j \in \{0, \dots, 2^n-1\}\}$ - dyadic intervals

$\mathcal{F}_n = \sigma(\mathcal{F}_n^0)$

$\mathcal{F} = \{\mathcal{F}_n\}_{n \ge 1}^{\infty}$ - dyadic filtration

Given $Q \in \mathcal{F}_n^0$, let $\ell(Q)$ be the left endpt
$r(Q)$ — " — right — " —.

Define $$X_n(\omega) := \sum_{Q \in \mathcal{F}_n^0} \frac{f(r(Q)) - f(\ell(Q))}{r(Q) - \ell(Q)} \mathbb{1}_{Q}(\omega) \quad \forall \omega \in (0,1].$$

<u>Rmk:</u> notice that $\forall \omega \in (0,1]$ exactly one summand is nonzero.

Rmk 2: Given $\omega \in [0,1]$, $\{X_n(\omega)\}_n$ is a seq of difference quotients approximating what would be $f'(\omega)$.

Claim: X is a martingale wrt $\mathcal{F}$.

Pf: NTS $E[X_n | \mathcal{F}_{n-1}] = X_{n-1}$.

Write X_n into $\mathcal{F}^0_{n-1}$

$$X_n = \sum_{J \in \mathcal{F}^0_{n-1}} \sum_{\substack{Q \in \mathcal{F}^0_n \\ Q \subseteq J}} \frac{f(r(Q)) - f(l(Q))}{2^{-n}} \mathbb{1}_Q(\omega)$$

For $Q \subseteq J$ as in the sum $E[\mathbb{1}_Q | \mathcal{F}_{n-1}] = \frac{1}{2}\mathbb{1}_J$

so

$$E[X_n | \mathcal{F}_{n-1}] = \frac{1}{2} \sum_{J \in \mathcal{F}^0_{n-1}} \sum_{\substack{Q \in \mathcal{F}^0_n \\ Q \subseteq J}} \frac{f(r(Q)) - f(l(Q))}{2^{-n}} \mathbb{1}_J(\omega) = X_{n-1}.$$

So X - Martingale wrt $\mathcal{F}$.

X - bdd $\Rightarrow$ by conv thm $\exists$ RV X_∞ s.t.

$X_n \to X_\infty$ P almost surely & in $L^1(P)$ $\triangle$ claim

Claim: $f(x) - f(0) = \int_0^x X_\infty(\omega)\,d\omega \quad \forall x \in [0,1]$.

Pf: Given $x \in [0,1]$, $\exists!\, J \in \mathcal{F}^0_n$ s.t. $x \in J$.

Then $|f(x) - f(r(J))| \leq A|x - r(J)| \leq \frac{A}{2^n}$ ①

(A ↑ Lipschitz const of f)

$$f(r(J)) - f(0) = \sum_{\substack{L \in \mathcal{F}^0_n \\ L \leq J}} f(r(L)) - f(l(L))$$ ②

$$= \sum_{\substack{L \in \mathcal{F}^0_n \\ L \leq J}} \int_L X_n(\omega)\,d\omega = \int_0^{r(J)} X_n(\omega)\,d\omega$$

Now $\left| \int_0^{r(J)} X_n(\omega)\,d\omega - \int_0^x X_n(\omega)\,d\omega \right| \leq \int_x^{x+2^{-n}} |X_n(\omega)|\,d\omega \xrightarrow[\text{by DCT}]{n \to \infty} 0$ ③

So combining ①, ②, ③ get $\lim_{n\to\infty} f(x) - f(0) - \int_0^x X_n(\omega)\,d\omega = 0$.

Now $\left| \int_0^x X_n(w) - X_\infty(w)\, dw \right| \le \int_0^x |X_n(w) - X_\infty(w)|\, dw \le \int_0^1 |X_n(w) - X_\infty(w)|\, dw \to 0$

so $f(x) - f(0) = \int_0^x X_\infty(w)\, dw \quad \forall x \in (0,1]$. $\triangle$ Claim.

If X_∞ were cts, could use the fundamental thm of calc to say f is diff'ble everywhere and $f' = X_\infty$.
Only know $X_\infty \in L^1(P)$.
Generalization of FTC, Lebesgue's differentiation thm (see pf in section 86.4) says
If $\int_0^1 |X_\infty(w)|\, dw < \infty$, then $\lim_{\delta \downarrow 0} \frac{1}{\delta} \int_w^{w+\delta} X_\infty(y)\, dy = X_\infty(y)$ almost every $w \in (0,1]$

i.e. $\frac{f(x+\delta) - f(x)}{\delta} \to X_\infty(x)$ //

$\triangle$

Random patterns

Let $X_1, X_2, \ldots$ be iid $P(X_i = 1) = p \in (0,1)$.
$P(X_i = 0) = q = 1 - p$.

$X_1, X_2, X_3, \ldots$ random seq of 0's & 1's.

Q: How long do you need to wait for the 1st 0?
$N = \inf\{n : X_1, \ldots, X_n \text{ contains a } 0\}$.
$EN = ?$

More generally, fix any deterministic pattern.
e.g. $P = 011001$
$N_P := \inf\{k : X_1, \ldots, X_k \text{ contains the pattern } P\}$
$EN_P = ?$

Easy to check $N_P < \infty$ a.s. $\forall P$.
i.e. $\forall$ pattern appears a.s.

E.g. $P(N_0 = j) = p^{j-1} q$ so $EN_0 = \sum_{j=1}^\infty j p^{j-1} q = \frac{1}{q}$

Hard to replicate even for N_{01}.

Let $Y_n := \frac{1}{q} \mathbb{1}_{\{X_1=0\}} + \dots + \frac{1}{q} \mathbb{1}_{\{X_n=0\}} = \frac{\#\,0\text{s up to } n}{q}$

$\mathcal{F}_n := \sigma(X_1, \dots, X_n)$.

$E[Y_{n+1} | \mathcal{F}_n] = Y_n + \frac{1}{q} P(X_{n+1}=0 | \mathcal{F}_n) = Y_n + 1$

So $\{Y_n - n\}_{n\geq 1}^{\infty}$ - mean zero martingale

$\Rightarrow$ by the optional stopping thm

$E(Y_{N\wedge n} - N\wedge n) = 0$.

So $E Y_{N\wedge n} = E N\wedge n$.

$N < \infty$ a.s. & both $\nearrow$ increasing

so can apply MCT, $n \to \infty$ get $E Y_N = E N$.

$Y_N = \frac{1}{q}$ a.s. so $EN = \frac{1}{q}$

This generalizes.

E.g. $E N_{\underbrace{010}_{T}}$ - ? $\quad Z_n := \frac{1}{qpq} \mathbb{1}_{(X_1,X_2,X_3)=T} + \frac{1}{qpq} \mathbb{1}_{(X_2,X_3,X_4)=T} + \dots + \frac{1}{qpq} \mathbb{1}_{(X_{n-2},X_{n-1},X_n)=T}$

$+ \frac{1}{qp} \mathbb{1}_{(X_{n-1},X_n)=(01)} + \frac{1}{q} \mathbb{1}_{X_n=0}$.

$$E[Z_{n+1} - Z_n | \mathcal{F}_n] = E\Big[\frac{1}{qpq} \mathbb{1}_{(X_{n-1},X_n,X_{n+1})=(010)} + \frac{1}{qp} \mathbb{1}_{(X_n,X_{n+1})=(01)} + \frac{1}{q} \mathbb{1}_{X_{n+1}=0}$$
$$- \frac{1}{qp} \mathbb{1}_{(X_{n-1},X_n)=(01)} - \frac{1}{q} \mathbb{1}_{X_n=0} \Big| \mathcal{F}_n\Big]$$

$$= \frac{1}{qpq} \mathbb{1}_{(X_{n-1},X_n)=(01)} \underbrace{E[\mathbb{1}_{X_{n+1}=0} | \mathcal{F}_n]}_{P(X_{n+1}=0)=q}$$
$$+ \frac{1}{qp} \mathbb{1}_{X_n=0} \underbrace{E[\mathbb{1}_{X_{n+1}=1} | \mathcal{F}_n]}_{P(X_{n+1}=1)=p}$$
$$+ \frac{1}{q} \underbrace{E[\mathbb{1}_{X_{n+1}=0} | \mathcal{F}_n]}_{P(X_{n+1}=0)=q} - \frac{1}{qp} \mathbb{1}_{(X_{n-1},X_n)=(01)} - \frac{1}{q} \mathbb{1}_{X_n=0}$$

$= 1$.

So $(Z_n - n)$ - martingale w/ mean 0.

$\Rightarrow \quad EN_{010} = EZ_N = \frac{1}{qpq} + \frac{1}{q}.$

Similarly $EN_{001} = \frac{1}{qqp}$: $W_n = \sum_{i=3}^{n} \frac{1}{qqp} 1_{(X_{i-2,i-1,i}) = 001} + \frac{1}{qq} 1_{(X_{n-1}, X_n) = (00)} + \frac{1}{q} 1_{X_n = 0}.$

Q: Which pattern comes 1st, 010 or 011?

$P(010 \text{ before } 001) = ?$

$Z_n - W_n$ — Mean 0 martingale

Let $T = \inf\{k : X_1, \ldots, X_k \text{ contains } 010 \text{ or } 011\}$.

Optional stopping & MCT $E[Z_T - W_T] = 0.$

$$Z_T - W_T = \begin{cases} \frac{1}{qpq} + \frac{1}{qp} + \frac{1}{q} - \frac{1}{q} & \text{if } 010 \text{ comes 1st} \\ -\left(\frac{1}{qqp} + \frac{1}{qq} + \frac{1}{q} - \frac{1}{q}\right) & \text{if } 001 \text{ comes 1st} \end{cases}$$

So $E[T_T - W_T] = \underbrace{P(010 \text{ 1st})}_{S} \cdot \left(\frac{1}{qpq} + \frac{1}{qp}\right) - \underbrace{P(011 \text{ 1st})}_{1-S} \left(\frac{1}{qpp} + \frac{1}{qq}\right) = 0$

$$P(010 \text{ before } 011) = S = \frac{\frac{1}{qpp} + \frac{1}{qq}}{\frac{1}{qpp} + \frac{1}{qq} + \frac{1}{qpq} + \frac{1}{qp}} = \frac{\frac{1}{pp} + \frac{1}{q}}{\frac{1}{pp} + \frac{1}{q} + \frac{1}{pq} + \frac{1}{p}} = \frac{q + p^2}{q + p^2 + p + pq}$$

$$= \frac{p^2 - p + 1}{\quad}$$