18.175: Lecture 26

More on martingales

Scott Sheffield

MIT

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Recall: conditional expectation

- Say we're given a probability space $(\Omega, \mathcal{F}_0, P)$ and a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a random variable X measurable w.r.t. $\mathcal{F}_0$, with $E|X| < \infty$. The **conditional expectation of** X **given** $\mathcal{F}$ is a new random variable, which we can denote by $Y = E(X|\mathcal{F})$.

Recall: conditional expectation

- Say we're given a probability space $(\Omega, \mathcal{F}_0, P)$ and a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a random variable X measurable w.r.t. $\mathcal{F}_0$, with $E|X| < \infty$. The **conditional expectation of X given $\mathcal{F}$** is a new random variable, which we can denote by $Y = E(X|\mathcal{F})$.
- We require that Y is $\mathcal{F}$ measurable and that for all A in $\mathcal{F}$, we have $\int_A X dP = \int_A Y dP$.

Recall: conditional expectation

- Say we're given a probability space $(\Omega, \mathcal{F}_0, P)$ and a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a random variable X measurable w.r.t. $\mathcal{F}_0$, with $E|X| < \infty$. The **conditional expectation of X given $\mathcal{F}$** is a new random variable, which we can denote by $Y = E(X|\mathcal{F})$.
- We require that Y is $\mathcal{F}$ measurable and that for all A in $\mathcal{F}$, we have $\int_A XdP = \int_A YdP$.
- Any Y satisfying these properties is called a **version** of $E(X|\mathcal{F})$.

Recall: conditional expectation

- Say we're given a probability space $(\Omega, \mathcal{F}_0, P)$ and a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a random variable X measurable w.r.t. $\mathcal{F}_0$, with $E|X| < \infty$. The **conditional expectation of X given $\mathcal{F}$** is a new random variable, which we can denote by $Y = E(X|\mathcal{F})$.
- We require that Y is $\mathcal{F}$ measurable and that for all A in $\mathcal{F}$, we have $\int_A X dP = \int_A Y dP$.
- Any Y satisfying these properties is called a **version** of $E(X|\mathcal{F})$.
- **Theorem:** Up to redefinition on a measure zero set, the random variable $E(X|\mathcal{F})$ exists and is unique.

Recall: conditional expectation

- Say we're given a probability space $(\Omega, \mathcal{F}_0, P)$ and a σ-field $\mathcal{F} \subset \mathcal{F}_0$ and a random variable X measurable w.r.t. $\mathcal{F}_0$, with $E|X| < \infty$. The **conditional expectation of X given $\mathcal{F}$** is a new random variable, which we can denote by $Y = E(X|\mathcal{F})$.
- We require that Y is $\mathcal{F}$ measurable and that for all A in $\mathcal{F}$, we have $\int_A X dP = \int_A Y dP$.
- Any Y satisfying these properties is called a **version** of $E(X|\mathcal{F})$.
- **Theorem:** Up to redefinition on a measure zero set, the random variable $E(X|\mathcal{F})$ exists and is unique.
- This follows from Radon-Nikodym theorem.

Conditional expectation observations

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

Conditional expectation observations

- Linearity: $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \le Y$ then $E(E|\mathcal{F}) \le E(Y|\mathcal{F})$.

Conditional expectation observations

- Linearity: $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \leq Y$ then $E(E|\mathcal{F}) \leq E(Y|\mathcal{F})$.
- If $X_n \geq 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).

Conditional expectation observations

- Linearity: $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \leq Y$ then $E(E|\mathcal{F}) \leq E(Y|\mathcal{F})$.
- If $X_n \geq 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).
- If $\mathcal{F}_1 \subset \mathcal{F}_2$ then

Conditional expectation observations

- Linearity: $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \leq Y$ then $E(E|\mathcal{F}) \leq E(Y|\mathcal{F})$.
- If $X_n \geq 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).
- If $\mathcal{F}_1 \subset \mathcal{F}_2$ then
  - $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_1)$.

Conditional expectation observations

- Linearity: $E(aX+Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \leq Y$ then $E(E|\mathcal{F}) \leq E(Y|\mathcal{F})$.
- If $X_n \geq 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).
- If $\mathcal{F}_1 \subset \mathcal{F}_2$ then
  - $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_1)$.
  - $E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1)$.

Conditional expectation observations

- Linearity: $E(aX + Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \leq Y$ then $E(E|\mathcal{F}) \leq E(Y|\mathcal{F})$.
- If $X_n \geq 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).
- If $\mathcal{F}_1 \subset \mathcal{F}_2$ then
  - $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_1)$.
  - $E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1)$.
- Second is kind of interesting: says, after I learn $\mathcal{F}_1$, my best guess of what my best guess for X will be after learning $\mathcal{F}_2$ is simply my current best guess for X.

Conditional expectation observations

- Linearity: $E(aX+Y|\mathcal{F}) = aE(X|\mathcal{F}) + E(Y|\mathcal{F})$.
- If $X \le Y$ then $E(E|\mathcal{F}) \le E(Y|\mathcal{F})$.
- If $X_n \ge 0$ and $X_n \uparrow X$ with $EX < \infty$, then $E(X_n|\mathcal{F}) \uparrow E(X|\mathcal{F})$ (by dominated convergence).
- If $\mathcal{F}_1 \subset \mathcal{F}_2$ then
  - $E(E(X|\mathcal{F}_1)|\mathcal{F}_2) = E(X|\mathcal{F}_1)$.
  - $E(E(X|\mathcal{F}_2)|\mathcal{F}_1) = E(X|\mathcal{F}_1)$.
- Second is kind of interesting: says, after I learn $\mathcal{F}_1$, my best guess of what my best guess for X will be after learning $\mathcal{F}_2$ is simply my current best guess for X.
- Deduce that $E(X|\mathcal{F}_i)$ is a martingale if $\mathcal{F}_i$ is an increasing sequence of σ-algebras and $E(|X|) < \infty$.

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Regular conditional probability

- Consider probability space $(\Omega, \mathcal{F}, P)$, a measurable map $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $\mathcal{G} \subset \mathcal{F}$ a σ-field. Then $\mu : \Omega \times \mathcal{S} \to [0, 1]$ is a **regular conditional distribution for** X **given** $\mathcal{G}$ if

Regular conditional probability

- Consider probability space $(\Omega, \mathcal{F}, P)$, a measurable map $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $\mathcal{G} \subset \mathcal{F}$ a σ-field. Then $\mu : \Omega \times \mathcal{S} \to [0, 1]$ is a **regular conditional distribution for** X **given** $\mathcal{G}$ if
  - For each A, $\omega \to \mu(\omega, A)$ is a version of $P(X \in A | \mathcal{G})$.

Regular conditional probability

- Consider probability space $(\Omega, \mathcal{F}, P)$, a measurable map $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $\mathcal{G} \subset \mathcal{F}$ a σ-field. Then $\mu : \Omega \times \mathcal{S} \to [0,1]$ is a **regular conditional distribution for** X **given** $\mathcal{G}$ if
  - For each A, $\omega \to \mu(\omega, A)$ is a version of $P(X \in A|\mathcal{G})$.
  - For a.e. ω, $A \to \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$.

Regular conditional probability

- Consider probability space $(\Omega, \mathcal{F}, P)$, a measurable map $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $\mathcal{G} \subset \mathcal{F}$ a σ-field. Then $\mu : \Omega \times \mathcal{S} \to [0, 1]$ is a **regular conditional distribution for** X **given** $\mathcal{G}$ if
  - For each A, $\omega \to \mu(\omega, A)$ is a version of $P(X \in A|\mathcal{G})$.
  - For a.e. ω, $A \to \mu(\omega, A)$ is a probability measure on $(S, \mathcal{S})$.
- **Theorem:** Regular conditional probabilities exist if $(S, \mathcal{S})$ is nice.

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Outline

Martingales

- Let $\mathcal{F}_n$ be increasing sequence of σ-fields (called a **filtration**).

Martingales

- Let $\mathcal{F}_n$ be increasing sequence of σ-fields (called a **filtration**).
- A sequence X_n is **adapted** to $\mathcal{F}_n$ if $X_n \in \mathcal{F}_n$ for all n. If X_n is an adapted sequence (with $E|X_n| < \infty$) then it is called a **martingale** if
$$E(X_{n+1}|\mathcal{F}_n) = X_n$$
for all n. It's a **supermartingale** (resp., **submartingale**) if same thing holds with $=$ replaced by $\leq$ (resp., $\geq$).

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \leq X_m$.

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \leq X_m$.
- **Proof idea:** Follows if $n = m + 1$ by definition; take $n = m + k$ and use induction on k.

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \le X_m$.
- **Proof idea:** Follows if $n = m + 1$ by definition; take $n = m + k$ and use induction on k.
- Similar result holds for submartingales. Also, if X_n is a martingale and $n > m$ then $E(X_n|\mathcal{F}_m) = X_m$.

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \leq X_m$.
- **Proof idea:** Follows if $n = m + 1$ by definition; take $n = m + k$ and use induction on k.
- Similar result holds for submartingales. Also, if X_n is a martingale and $n > m$ then $E(X_n|\mathcal{F}_m) = X_m$.
- **Claim:** if X_n is a martingale w.r.t. $\mathcal{F}_n$ and ϕ is convex with $E|\phi(X_n)| < \infty$ then $\phi(X_n)$ is a submartingale.

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \leq X_m$.
- **Proof idea:** Follows if $n = m + 1$ by definition; take $n = m + k$ and use induction on k.
- Similar result holds for submartingales. Also, if X_n is a martingale and $n > m$ then $E(X_n|\mathcal{F}_m) = X_m$.
- **Claim:** if X_n is a martingale w.r.t. $\mathcal{F}_n$ and ϕ is convex with $E|\phi(X_n)| < \infty$ then $\phi(X_n)$ is a submartingale.
- **Proof idea:** Immediate from Jensen's inequality and martingale definition.

Martingale observations

- **Claim:** If X_n is a supermartingale then for $n > m$ we have $E(X_n|\mathcal{F}_m) \le X_m$.
- **Proof idea:** Follows if $n = m + 1$ by definition; take $n = m + k$ and use induction on k.
- Similar result holds for submartingales. Also, if X_n is a martingale and $n > m$ then $E(X_n|\mathcal{F}_m) = X_m$.
- **Claim:** if X_n is a martingale w.r.t. $\mathcal{F}_n$ and ϕ is convex with $E|\phi(X_n)| < \infty$ then $\phi(X_n)$ is a submartingale.
- **Proof idea:** Immediate from Jensen's inequality and martingale definition.
- Example: take $\phi(x) = \max\{x, 0\}$.

Predictable sequence

- Call H_n **predictable** if each $H + n$ is $\mathcal{F}_{n-1}$ measurable.

Predictable sequence

- Call H_n **predictable** if each $H + n$ is $\mathcal{F}_{n-1}$ measurable.
- Maybe H_n represents amount of shares of asset investor has at nth stage.

Predictable sequence

- Call H_n **predictable** if each $H+n$ is $\mathcal{F}_{n-1}$ measurable.
- Maybe H_n represents amount of shares of asset investor has at nth stage.
- Write $(H \cdot X)_n = \sum_{m=1}^n H_m(X_m - X_{m-1})$.

Predictable sequence

- Call H_n **predictable** if each $H+n$ is $\mathcal{F}_{n-1}$ measurable.
- Maybe H_n represents amount of shares of asset investor has at nth stage.
- Write $(H \cdot X)_n = \sum_{m=1}^n H_m(X_m - X_{m-1})$.
- **Observe:** If X_n is a supermartingale and the $H_n \geq 0$ are bounded, then $(H \cdot X)_n$ is a supermartingale.

Predictable sequence

- Call H_n **predictable** if each $H+n$ is $\mathcal{F}_{n-1}$ measurable.
- Maybe H_n represents amount of shares of asset investor has at nth stage.
- Write $(H \cdot X)_n = \sum_{m=1}^n H_m(X_m - X_{m-1})$.
- **Observe:** If X_n is a supermartingale and the $H_n \geq 0$ are bounded, then $(H \cdot X)_n$ is a supermartingale.
- Example: take $H_n = 1_{N \geq n}$ for stopping time N.

Two big results

- **Optional stopping theorem:** Can't make money in expectation by timing sale of asset whose price is non-negative martingale.

Two big results

- **Optional stopping theorem:** Can't make money in expectation by timing sale of asset whose price is non-negative martingale.
- **Proof:** Just a special case of statement about $(H \cdot X)$.

Two big results

- **Optional stopping theorem:** Can't make money in expectation by timing sale of asset whose price is non-negative martingale.
- **Proof:** Just a special case of statement about $(H \cdot X)$.
- **Martingale convergence:** A non-negative martingale almost surely has a limit.

Two big results

- **Optional stopping theorem:** Can't make money in expectation by timing sale of asset whose price is non-negative martingale.
- **Proof:** Just a special case of statement about $(H \cdot X)$.
- **Martingale convergence:** A non-negative martingale almost surely has a limit.
- **Idea of proof:** Count upcrossings (times martingale crosses a fixed interval) and devise gambling strategy that makes lots of money if the number of these is not a.s. finite.

Problems

- How many primary candidates ever get above twenty percent in expected probability of victory? (Asked by Aldous.)

Problems

- How many primary candidates ever get above twenty percent in expected probability of victory? (Asked by Aldous.)
- Compute probability of having conditional probability reach a before b.

Wald

- **Wald's equation:** Let X_i be i.i.d. with $E|X_i| < \infty$. If N is a stopping time with $EN < \infty$ then $ES_N = EX_1 EN$.

Wald

- **Wald's equation:** Let X_i be i.i.d. with $E|X_i| < \infty$. If N is a stopping time with $EN < \infty$ then $ES_N = EX_1 EN$.
- **Wald's second equation:** Let X_i be i.i.d. with $E|X_i| = 0$ and $EX_i^2 = \sigma^2 < \infty$. If N is a stopping time with $EN < \infty$ then $ES_N = \sigma^2 EN$.

Wald applications to SRW

- $S_0 = a \in \mathbb{Z}$ and at each time step S_j independently changes by ± 1 according to a fair coin toss. Fix $A \in \mathbb{Z}$ and let $N = \inf\{k : S_k \in \{0, A\}\}$. What is $\mathbb{E}S_N$?

Wald applications to SRW

- $S_0 = a \in \mathbb{Z}$ and at each time step S_j independently changes by ± 1 according to a fair coin toss. Fix $A \in \mathbb{Z}$ and let $N = \inf\{k : S_k \in \{0, A\}\}$. What is $\mathbb{E}S_N$?
- What is $\mathbb{E}N$?

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Outline

Conditional expectation

Regular conditional probabilities

Martingales

Arcsin law, other SRW stories

Reflection principle

- How many walks from $(0, x)$ to (n, y) that don't cross the horizontal axis?

Reflection principle

- How many walks from $(0, x)$ to (n, y) that don't cross the horizontal axis?
- Try counting walks that *do* cross by giving bijection to walks from $(0, -x)$ to (n, y).

Ballot Theorem

- Suppose that in election candidate A gets α votes and B gets $\beta < \alpha$ votes. What's probability that A is ahead throughout the counting?

Ballot Theorem

- Suppose that in election candidate A gets α votes and B gets $\beta < \alpha$ votes. What's probability that A is ahead throughout the counting?
- Answer: $(\alpha - \beta)/(\alpha + \beta)$. Can be proved using reflection principle.

Arcsin theorem

- Theorem for last hitting time.

Arcsin theorem

- Theorem for last hitting time.
- Theorem for amount of positive positive time.