18.175: Lecture 16

Conditional expectation, random walks, martingales

Scott Sheffield

MIT

Outline

Conditional expectation

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

Outline

Conditional expectation

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

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 $Y = E(X|\mathcal{F})$.

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 $\int_A X dP = \int_A Y dP$.

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 $E(X|\mathcal{F})$.

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 $E(X|\mathcal{F})$.
- Is it possible that there exists more than one version of $E(X|\mathcal{F})$ (which would mean that in some sense the conditional expectation is not canonically defined)?

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 $E(X|\mathcal{F})$.
- Is it possible that there exists more than one version of $E(X|\mathcal{F})$ (which would mean that in some sense the conditional expectation is not canonically defined)?
- Is there some sense in which $E(X|\mathcal{F})$ always exists and is always uniquely defined (maybe up to set of measure zero)?

Conditional expectation

- **Claim:** Assuming $Y = E(X|\mathcal{F})$ as above, and $E|X| < \infty$, we have $E|Y| \leq E|X|$. In particular, Y is integrable.

Conditional expectation

- **Claim:** Assuming $Y = E(X|\mathcal{F})$ as above, and $E|X| < \infty$, we have $E|Y| \leq E|X|$. In particular, Y is integrable.
- **Proof:** let $A = \{Y > 0\} \in \mathcal{F}$ and observe: $\int_A YdP = \int_A XdP \leq \int_A |X|dP$. By similar argument, $\int_{A^c} -YdP \leq \int_{A^c} |X|dP$.

Conditional expectation

- **Claim:** Assuming $Y = E(X|\mathcal{F})$ as above, and $E|X| < \infty$, we have $E|Y| \leq E|X|$. In particular, Y is integrable.
- **Proof:** let $A = \{Y > 0\} \in \mathcal{F}$ and observe: $\int_A YdP = \int_A XdP \leq \int_A |X|dP$. By similar argument, $\int_{A^c} -YdP \leq \int_{A^c} |X|dP$.
- **Uniqueness of** Y: Suppose Y' is $\mathcal{F}$-measurable and satisfies $\int_A Y'dP = \int_A XdP = \int_A YdP$ for all $A \in \mathcal{F}$. Then consider the set $Y - Y' \geq \epsilon\}$. Integrating over that gives zero. Must hold for any ϵ. Conclude that $Y = Y'$ almost everywhere.

Radon-Nikodym theorem

- Let μ and ν be σ-finite measures on $(\Omega, \mathcal{F})$. Say $\nu << \mu$ (or ν is **absolutely continuous w.r.t.** μ if $\mu(A) = 0$ implies $\nu(A) = 0$.

Radon-Nikodym theorem

- Let μ and ν be σ-finite measures on $(\Omega, \mathcal{F})$. Say $\nu << \mu$ (or ν is **absolutely continuous w.r.t.** μ if $\mu(A) = 0$ implies $\nu(A) = 0$.
- Recall **Radon-Nikodym theorem:** If μ and ν are σ-finite measures on $(\Omega, \mathcal{F})$ and ν is absolutely continuous w.r.t. μ, then there exists a measurable $f : \Omega \to [0, \infty)$ such that $\nu(A) = \int_A f d\mu$.

Radon-Nikodym theorem

- Let μ and ν be σ-finite measures on $(\Omega, \mathcal{F})$. Say $\nu << \mu$ (or ν is **absolutely continuous w.r.t.** μ if $\mu(A) = 0$ implies $\nu(A) = 0$.
- Recall **Radon-Nikodym theorem:** If μ and ν are σ-finite measures on $(\Omega, \mathcal{F})$ and ν is absolutely continuous w.r.t. μ, then there exists a measurable $f : \Omega \to [0, \infty)$ such that $\nu(A) = \int_A f d\mu$.
- Observe: this theorem implies existence of conditional expectation.

Outline

Conditional expectation

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

Outline

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.
- **Martingale convergence:** A non-negative martingale almost surely has a limit.

Outline

Conditional expectation

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

Outline

Exchangeable events

- Start with measure space $(S, \mathcal{S}, \mu)$. Let $\Omega = \{(\omega_1, \omega_2, \ldots) : \omega_i \in S\}$, let $\mathcal{F}$ be product σ-algebra and P the product probability measure.

Exchangeable events

- Start with measure space $(S, \mathcal{S}, \mu)$. Let $\Omega = \{(\omega_1, \omega_2, \ldots) : \omega_i \in S\}$, let $\mathcal{F}$ be product σ-algebra and P the product probability measure.
- **Finite permutation** of $\mathbb{N}$ is one-to-one map from $\mathbb{N}$ to itself that fixes all but finitely many points.

Exchangeable events

- Start with measure space $(S, \mathcal{S}, \mu)$. Let $\Omega = \{(\omega_1, \omega_2, \ldots) : \omega_i \in S\}$, let $\mathcal{F}$ be product σ-algebra and P the product probability measure.
- **Finite permutation** of $\mathbb{N}$ is one-to-one map from $\mathbb{N}$ to itself that fixes all but finitely many points.
- Event $A \in \mathcal{F}$ is permutable if it is invariant under any finite permutation of the ω_i.

Exchangeable events

- Start with measure space $(S, \mathcal{S}, \mu)$. Let $\Omega = \{(\omega_1, \omega_2, \ldots) : \omega_i \in S\}$, let $\mathcal{F}$ be product σ-algebra and P the product probability measure.
- **Finite permutation** of $\mathbb{N}$ is one-to-one map from $\mathbb{N}$ to itself that fixes all but finitely many points.
- Event $A \in \mathcal{F}$ is permutable if it is invariant under any finite permutation of the ω_i.
- Let $\mathcal{E}$ be the σ-field of permutable events.

Exchangeable events

- Start with measure space $(S, \mathcal{S}, \mu)$. Let $\Omega = \{(\omega_1, \omega_2, \ldots) : \omega_i \in S\}$, let $\mathcal{F}$ be product σ-algebra and P the product probability measure.
- **Finite permutation** of $\mathbb{N}$ is one-to-one map from $\mathbb{N}$ to itself that fixes all but finitely many points.
- Event $A \in \mathcal{F}$ is permutable if it is invariant under any finite permutation of the ω_i.
- Let $\mathcal{E}$ be the σ-field of permutable events.
- This is related to the tail σ-algebra we introduced earlier in the course. Bigger or smaller?

Hewitt-Savage 0-1 law

- If $X_1, X_2, \ldots$ are i.i.d. and $A \in \mathcal{A}$ then $P(A) \in \{0, 1\}$.

Hewitt-Savage 0-1 law

- If $X_1, X_2, \ldots$ are i.i.d. and $A \in \mathcal{A}$ then $P(A) \in \{0, 1\}$.
- **Idea of proof:** Try to show A is independent of itself, i.e., that $P(A) = P(A \cap A) = P(A)P(A)$. Start with measure theoretic fact that we can approximate A by a set A_n in σ-algebra generated by $X_1, \ldots X_n$, so that symmetric difference of A and A_n has very small probability. Note that A_n is independent of event A'_n that A_n holds when $X_1, \ldots, X_n$ and $X_{n_1}, \ldots, X_{2n}$ are swapped. Symmetric difference between A and A'_n is also small, so A is independent of itself up to this small error. Then make error arbitrarily small.

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^{n} X_i$ is a **random walk** on $\mathbb{R}^n$.

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:
  - $S_n = 0$ for all n

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:
  - $S_n = 0$ for all n
  - $S_n \to \infty$

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:
  - $S_n = 0$ for all n
  - $S_n \to \infty$
  - $S_n \to -\infty$

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:
  - $S_n = 0$ for all n
  - $S_n \to \infty$
  - $S_n \to -\infty$
  - $-\infty = \liminf S_n < \limsup S_n = \infty$

Application of Hewitt-Savage:

- If X_i are i.i.d. in $\mathbb{R}^n$ then $S_n = \sum_{i=1}^n X_i$ is a **random walk** on $\mathbb{R}^n$.
- **Theorem:** if S_n is a random walk on $\mathbb{R}$ then one of the following occurs with probability one:
  - $S_n = 0$ for all n
  - $S_n \to \infty$
  - $S_n \to -\infty$
  - $-\infty = \liminf S_n < \limsup S_n = \infty$
- **Idea of proof:** Hewitt-Savage implies the $\limsup S_n$ and $\liminf S_n$ are almost sure constants in $[-\infty, \infty]$. Note that if X_1 is not a.s. constant, then both values would depend on X_1 if they were not in $\pm\infty$

Outline

Conditional expectation

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

Outline

Stopping time definition

- Say that T is a **stopping time** if the event that $T = n$ is in $\mathcal{F}_n$ for $i \leq n$.

Stopping time definition

- Say that T is a **stopping time** if the event that $T = n$ is in $\mathcal{F}_n$ for $i \leq n$.
- In finance applications, T might be the time one sells a stock. Then this states that the decision to sell at time n depends only on prices up to time n, not on (as yet unknown) future prices.

Stopping time examples

- Let $A_1, \ldots$ be i.i.d. random variables equal to -1 with probability .5 and 1 with probability .5 and let $X_0 = 0$ and $X_n = \sum_{i=1}^n A_i$ for $n \geq 0$.

Stopping time examples

- Let $A_1, \ldots$ be i.i.d. random variables equal to -1 with probability .5 and 1 with probability .5 and let $X_0 = 0$ and $X_n = \sum_{i=1}^n A_i$ for $n \geq 0$.
- Which of the following is a stopping time?
  1. The smallest T for which $|X_T| = 50$
  2. The smallest T for which $X_T \in \{-10, 100\}$
  3. The smallest T for which $X_T = 0$.
  4. The T at which the X_n sequence achieves the value 17 for the 9th time.
  5. The value of $T \in \{0, 1, 2, \ldots, 100\}$ for which X_T is largest.
  6. The largest $T \in \{0, 1, 2, \ldots, 100\}$ for which $X_T = 0$.

Stopping time examples

- Let $A_1, \ldots$ be i.i.d. random variables equal to -1 with probability .5 and 1 with probability .5 and let $X_0 = 0$ and $X_n = \sum_{i=1}^n A_i$ for $n \geq 0$.
- Which of the following is a stopping time?
  1. The smallest T for which $|X_T| = 50$
  2. The smallest T for which $X_T \in \{-10, 100\}$
  3. The smallest T for which $X_T = 0$.
  4. The T at which the X_n sequence achieves the value 17 for the 9th time.
  5. The value of $T \in \{0, 1, 2, \ldots, 100\}$ for which X_T is largest.
  6. The largest $T \in \{0, 1, 2, \ldots, 100\}$ for which $X_T = 0$.

Stopping time examples

- Let $A_1, \ldots$ be i.i.d. random variables equal to -1 with probability .5 and 1 with probability .5 and let $X_0 = 0$ and $X_n = \sum_{i=1}^n A_i$ for $n \geq 0$.
- Which of the following is a stopping time?
  1. The smallest T for which $|X_T| = 50$
  2. The smallest T for which $X_T \in \{-10, 100\}$
  3. The smallest T for which $X_T = 0$.
  4. The T at which the X_n sequence achieves the value 17 for the 9th time.
  5. The value of $T \in \{0, 1, 2, \ldots, 100\}$ for which X_T is largest.
  6. The largest $T \in \{0, 1, 2, \ldots, 100\}$ for which $X_T = 0$.
- Answer: first four, not last two.

Stopping time theorems

- **Theorem:** Let $X_1, X_2, \ldots$ be i.i.d. and N a stopping time with $N < \infty$.

Stopping time theorems

- **Theorem:** Let $X_1, X_2, \ldots$ be i.i.d. and N a stopping time with $N < \infty$.
- Conditioned on stopping time $N < \infty$, conditional law of $\{X_{N+n}, n \geq 1\}$ is independent of $\mathcal{F}_n$ and has same law as original sequence.

Stopping time theorems

- **Theorem:** Let $X_1, X_2, \dots$ be i.i.d. and N a stopping time with $N < \infty$.
- Conditioned on stopping time $N < \infty$, conditional law of $\{X_{N+n}, n \geq 1\}$ is independent of $\mathcal{F}_n$ and has same law as original sequence.
- **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 $ES_N = EX_1 EN$.

Stopping time theorems

- **Theorem:** Let $X_1, X_2, \ldots$ be i.i.d. and N a stopping time with $N < \infty$.
- Conditioned on stopping time $N < \infty$, conditional law of $\{X_{N+n}, n \geq 1\}$ is independent of $\mathcal{F}_n$ and has same law as original sequence.
- **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

- **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

Martingales

Random walks

Stopping times

Arcsin law, other SRW stories

Outline

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 a head 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 a head 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.