Probability, Markov Chains, and Stochastic Processes

Eren Kizildag, Zied Ben Chaouch

MIT, Cambridge

September 10, 2017

Contents

0.1 Abstract

We are doing a Probability and Stochastic Processes review. This will take all Summer 2017 and continue during the Fall 2017 semester. We will be covering the following textbooks (in depth reading is in bold):

- **David Williams, *Probability With Martingales*, 1991 [13]**
- Rick Durrett, *Probability: Theory and Example*, 2010 [5]
- Robert Ash & Catherine Doleans-Dade, *Probability With Measure Theory*, 2008 [1]
- Patrick Billingsley, *Probability And Measure*, 1995 [3]
- Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes*, 2001 [8]
- Geoffrey Grimmett & David Stirzaker, *One Thousand Exercises in Probability*, 2001 [7]
- Gerald Folland, *Real analysis: modern techniques and their applications*, 1999 [6]
- Pierre Bremaud, *Point Processes and Queues*, 1981 [4]
- **David Levin, Yuval Peres, Elizabeth Wilmer, *Markov Chains and Mixing Times*, 2008 [10]**
- **Peter Morters & Yuval Peres, *Brownian Motion*, 2010 [11]**
- **Bernt ∅ksendal, *Stochastic Differential Equations*, 2013 [12]**
- Alice Guionnet, *MIT 18.176 Introduction to Stochastic Analysis*, 2013 [9]
- Martin Bazant, *MIT 18.366 Random Walks and Diffusion*, 2006 [2]

0.2 Syllabus

0.2.1 Probability Review:

Books Considered:

–[**W**] David Williams, *Probability With Martingales*, 1991

–[**D**] Rick Durrett, *Probability: Theory and Example*, 2010

–[**A**] Robert Ash & Catherine Doleans-Dade, *Probability With Measure Theory*, 2008

–[**B**] Patrick Billingsley, *Probability And Measure*, 1995

–[**B**] Pierre Bremaud, *Point Processes and Queues*, 1981

Calendar:

- Week 1 (May 29):
  - [**W**] Chapters 1-5: Measure Spaces, Events, Random Variables, Independence and Integration.
  - [**D**] 2.3 Borel-Cantelli
- Week 2 (June 5):
  - [**W**] Chapters 6-8: Expectation, Strong Law and Product Measure.
  - [**D**] 2.4.7 Glivenko-Cantelli (SLLN) Theorem
  - [**D**] 2.5 Convergence of Random Series
  - [**D**] 2.2.3 Truncation
- Week 3 (June 12):
  - [**W**] Chapters 9-11: Conditional Expectation, Martingales and Convergence Theorem.
- Week 4 (June 19):
  - [**W**] Chapters 12-13: Martingales Bounded in L^2 and Uniform Integrability.
  - [**D**] Chapter 5: Martingales
  - [**Br**] Chapter 1: Martingales
- Week 5 (June 26):
  - [**W**] Chapter 14: Uniformly Integrable Martingales.
  - [**D**] Chapter 5: Martingales
  - [**Br**] Chapter 1: Martingales
- Week 6 (July 3):
  - [**W**] Chapters 16-18: Characteristic Functions, Weak Convergence and Central Limit Theorem.
  - [**D**] 3.4.4 Berry-Esseen Inequality
  - [**D**] 3.3.4 Polya's Criterion (characteristic functions)
  - [**D**] 3.9 Limit theorems in $\mathbb{R}^d$
- Week 7 (July 10):
  - [**W**] Chapters 0 & 15: Branching Processes and Applications.

Additional Notes

Whenever deemed necessary, we will look at Ash and Billingsley. Priority will be on Ash.

0.2.2 Markov Chains

Books Considered:

–**[P]** David Levin, Yuval Peres, Elizabeth Wilmer, *Markov Chains and Mixing Times*, 2008

–**[D]** Rick Durrett, *Probability: Theory and Example*, 2010

Calendar:

- Week 8 (July 17):
  - **[P]** Chapters 1 & 2: Markov Chains, Classical Examples.
  - **[D]** Chapter 4: Random Walks
- Week 9 (July 24):
  - **[P]** Chapters 4 & 5: Mixing and Coupling.
- Week 10 (July 31):
  - **[P]** Chapters 6 & 7: Strong Stationary Times and Lower Bounds.
- Week 11 (August 7):
  - **[P]** Chapters 9-11: Networks, Hitting Times and Cover Times.
- Week 12 (August 14):
  - **[P]** Chapters 12 & 3: Eigenvalues and Metropolis/Glauber Dynamics.
- Week 13 (August 21):
  - **[P]** Chapters 13 & 17: Eigenfunctions/Comparison of chains and Martingales on Evolving Sets.
- Week 14 (August 28):
  - **[P]** Chapters 18 & 20: Cutoff and Continuous Time Chains.
- Week 15 (September 4):
  - **[P]** Chapter 21: Countable States.

Additional Notes

In Peres, we will skip Chapters:

- $\boxed{\text{8: Shuffling}} \to \boxed{\text{16: Shuffling Genes}}$
- $\boxed{\text{14: Path Coupling}} \to \boxed{\text{22: Coupling from the Past}}$ & $\boxed{\text{15: Ising Model}}$
- $\boxed{\text{19: Lamplighter}}$

0.2.3 Brownian Motion & Stochastic Calculus

Books Considered:

–**[P]** Peter Morters & Yuval Peres, *Brownian Motion*, 2010

–**[O]** Bernt ∅ksendal, *Stochastic Differential Equations*, 2013

–**[B]** Pierre Bremaud, *Point Processes and Queues*, 1981

–**[G]** Alice Guionnet, *MIT 18.176 Introduction to Stochastic Analysis*, 2013

–**[Ba]** Martin Bazant, *MIT 18.366 Random Walks and Diffusion*, 2006

Calendar:

- Unit 1:
  - **[P]** Chapter 1: Brownian Motion as a Random Function.
- Unit 2:
  - **[P]** Chapter 2: Brownian Motion as a Strong Markov Process.
- Unit 3:
  - **[P]** Chapter 3: Harmonic Functions, Transience and Recurrence.
- Unit 4:
  - **[P]** Chapter 4: Haussdorff Dimension.
- Unit 5:
  - **[P]** Chapter 5: Brownian Motion and Random Walk.
- Unit 6:
  - **[P]** Chapter 6: Brownian Local Time.
- Unit 7:
  - **[O]** Chapters 1-4: Introduction, Preliminaries and Ito Integrals.
- Unit 8:
  - **[O]** Chapters 4 & 5: Ito Formula, Martingale Representation Theorem and Stochastic Differential Equations.
- Unit 9:
  - **[O]** Chapter 6: Filtering.
- Unit 10:
  - **[O]** Chapter 7: Diffusions.
- Unit 11:
  - **[O]** Chapter 8: Topics in Diffusion Theory.
- Unit 12:
  - **[O]** Chapter 9: Boundary Value Problems.

- Unit 13:
  - **[O]** Chapter 10: Optimal Stopping.
- Unit 14:
  - **[O]** Chapter 11: Stochastic Control.
- Unit 15:
  - **[O]** Chapter 12: Mathematical Finance.
- Unit 16:
  - **[P]** Chapter 7: Stochastic Integrals.
- Unit 17:
  - **[P]** Chapter 8: Potential Theory of Brownian Motion.
- Unit 18:
  - **[P]** Chapter 9: (Self) Intersections of Brownian Paths.
- Unit 19:
  - **[P]** Chapter 10: Exceptional Sets For Brownian Motion.

Additional Notes

Try to get to Bremaud book and Guionnet's notes as well. In Peres, we will skip Chapter 11: Stochastic Loewner Evolution and Planar Brownian Motion.

Chapter 1

Week 1: Measure Spaces, Events and Random Variables

1.1 Day 1: Measure Spaces

<u>**Main Reference(s):**</u> David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Set Operations $\cap, \cup, (.)^c, \setminus$ and $A\Delta B = (A \cup B) \setminus (A \cap B)$ (symmetric difference)

Algebra Collection Σ_0 of subsets of set S, S.T. stable under finitely many sets operations:

(i) $S \in \Sigma_0$

(ii) $F \in \Sigma_0 \implies F^c \in \Sigma_0$

(iii) $F, G \in \Sigma_0 \implies F \cup G \in \Sigma_0$

σ-algebra Collection Σ of subsets of set S, S.T. stable under finitely many sets operations:

(i) $S \in \Sigma$

(ii) $F \in \Sigma \implies F^c \in \Sigma$

(iii) $F_n \in \Sigma \implies \bigcup_{n=1}^{\infty} F_n \in \Sigma$

π-system Family of subsets of set S that is stable under finite $\bigcap$

Measurable Space (S, Σ) and **Measure Space** $(S, \Sigma, \mathbb{P})$

σ-Algebra Generated by C: C=class of subsets of $S \Rightarrow \sigma(C) = \bigcap\limits_{\substack{\Sigma_i \supset C \text{ is a} \\ \sigma\text{-algebra on } S}} \Sigma_i$

$\Rightarrow \sigma(C) =$ smallest σ-algebra on S containing C.

Borel σ-algebra $\mathscr{B}(S) = \sigma(\text{open})$ (S =topological space)

Countable Additivity Just right: more powerful than finite additivity but does not lead to contradictions (uncountable additivity: $\mathbb{P}(\bigcup_{[0,1]}\{x\}) = 0$ or 1?)

Measure $\mu : \Sigma \to [0, \infty]$ with **Countable additivity:** $\mathbb{P}\left(\bigcup_{n=1}^{\infty} \overbrace{A_n}^{disjoint}\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

Finite Meas $\mu(S){<}\infty$; **Prob Meas** $\mu(S){=}1$; **σ-Finite Meas** $\exists$ partition (S_n) with $\mu(S_n){<}\infty$

Construction of Lebesgue Meas. Define it on simple sets $(a, b]$ (guarantees no overlap): get an algebra, extend via Caratheodory, argue the uniqueness.

Lebesgue σ-algebra Completion of $\mathscr{B}(\mathbb{R})$ with nullset.

MAIN THEOREMS:

Vitali Sets On $I=[0,1]$: $x \sim y \Leftrightarrow (x-y) \in \mathbb{Q}$. Let $S/\sim$=set of equivalence classes determined by $\sim$: these classes are nonempty (axiom of choice) and we can choose one element of each equivalence class.

Axiom of Choice $\Rightarrow \exists f:(S/\sim) \to S = \cup_{A\in(S/\sim)} A$ (choose 1 element of each equiv. class).

$V(S) := \{f(A) : A \in (S/\sim)\}$= set formed by picking exactly 1 element of each equiv class.

$V(S)$ is not meas. $(\because) A_q = \{a+q \pmod 1 \mid a \in A\}$ $(q \in \mathbb{Q})$ we have a countably family of pairwise disjoint sets whose union is [0,1]. $\mu([0,1]) = \sum \mu(A_q) = \sum_{i=1}^{\infty} \mu(A) = \infty \times \mu(A)$.

Banach-Tarski Paradox $\exists F \subset S^2$ (Sphere in $\mathbb{R}^3$) S.T. $S^2 = \cup_{i=1}^{k} \tau_i^{(k)} F$ $\forall 3 \le k < \infty$ $(\tau - i^{(k)}$= rotation). F must be non-meas., otherwise length(F)=0.

Characterization of open subsets of $\mathbb{R}$ $G \subset \mathbb{R}$ open $\Leftrightarrow G = \bigcup_i O_i$, $i \in I$ countable, O_i disjoint open interval.

Generators of $\mathscr{B}(\mathbb{R})$, π-systems $\mathscr{B}(S) = \sigma(\text{open}) = \sigma(\pi(\mathbb{R}))$, with $\pi(\mathbb{R}) = \{(-\infty, x] : x \in \mathbb{R}\}$

Uniqueness of Extension $\mathbb{P}_1 = \mathbb{P}_2$ (general: $\mu(\Omega) < \infty$) on a π-system $\Rightarrow$ agree on $\sigma(\pi)$. Application, from CDF's to probability laws with $(-\infty, x]$.

Caratheodory's Extension Theorem Extend a pre-measure on an algebra, to a measure $\mu(\cdot)$ on σ(algebra). Used together with previous theorem.

Union Bound $\mu\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} \mu(A_i)$

Inclusion/Exclusion Formula $A_1, \ldots, A_n \in \Sigma$: $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2)$ and

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j}^{n} \mu(A_i \cap A_j) + \sum_{i<j<k}^{n} \mu(A_i \cap A_j \cap A_k) - .. + (-1)^{n-1} \mu(A_i \cap .. \cap A_n)$$

Nullsets $\bigcup_{\mathbb{N}}(\text{nullsets}) = \text{nullset}$

Baire Category Theorem $X = \bigcup_{\mathbb{N}} F_n$ complete metric space (F_n closed) $\Rightarrow \exists F_n$ S.T. $F_n^{\circ} \neq \varnothing$

Monotone Convergence of Measures $A_n \in \Sigma$:

$\star A_n \nearrow A \implies \mu(A_n) \nearrow \mu(A)$

$\star A_n \searrow A$ & $\exists k$ S.T. $\mu(A_k) < \infty \implies \mu(A_n) \searrow \mu(A)$

Careful! $(\star\ 2)$ Fails if $\mu(A_n) = \infty\ \forall n$

Example: $A_n = [n, \infty) \to \varnothing \implies \lim_n \mu(A_n) = \infty \neq 0 = \mu(\varnothing) = \mu\left(\lim_n A_n\right)$

1.2 Day 2: Events

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Almost sure A statement S holds almost surely (a.s.) if $\mathbb{P}(S) = 1$.

Liminf of Sequences $\liminf x_n = \lim_n \inf_{k\geq n} x_k = \sup_{n\geq 1} \inf_{k\geq n} x_k$ exist $(\because) \inf_{k\geq n} x_k$=monotonic

Limsup of Sequences $\limsup x_n = \lim_n \sup_{k\geq n} x_k = \inf_{n\geq 1} \sup_{k\geq n} x_k$ exist $(\because) \sup_{k\geq n} x_k$=monotonic

Liminf of Sets $\liminf A_n = \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k$, $\omega \in$ all A_n's except possibly a finitely many of A_n's

Limsup of Sets $\limsup A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k$, $\omega \in$ infinitely many A_n's.

MAIN THEOREMS:

a.s. Intersection $F_n \in \mathscr{F}$ and $\mathbb{P}(F_n) = 1 \implies \mathbb{P}(\bigcap_n F_n) = 1$

Sandwich Limits $\inf_{n\geq 1} x_n \leq \inf_{n\geq k} x_n \leq \liminf_n x_n \leq \limsup_n x_n \leq \sup_{n\geq k} x_n \leq \sup_{n\geq 1} x_n$.

$z > \limsup_n x_n \implies x_n < z$ eventually (n big enough)

$z < \limsup_n x_n \Rightarrow x_n > z$ i.o.(infinitely often/infinitely many n).

Sandwich Limits of sets $\bigcap_{k\geq 1} A_k \subset \bigcup_{n\geq i} \bigcap_{k\geq n} A_k \subset \liminf_n A_n \subset \limsup_n A_n \subset \bigcap_{n\geq i} \bigcup_{k\geq n} A_k \subset \bigcup_{k\geq 1} A_k$.

Fatou Lemmas Limsup/liminf on LHS for sets, on RHS for sequences.

1. $\mathbb{P}(\limsup_n E_n) \geq \limsup_n \mathbb{P}(E_n)$. Need $\mathbb{P}(\cdot)$ finite (prob. measures OK).
2. $\mathbb{P}(\liminf_n E_n) \leq \liminf_n \mathbb{P}(E_n)$. All measures.

Proof: *(click)*

Sketch: cf. Williams + By Exercise 2.9 $\mathbb{1}_{\limsup_n E_n} = \limsup_n \mathbb{1}_{E_n}$

$\Rightarrow \mathbb{P}(\limsup_n E_n) = \int \limsup_n \mathbb{1}_{E_n} d\mathbb{P} \geq \limsup_n \int 1_{E_n} d\mathbb{P} = \limsup_n \mathbb{P}(E_n).$ $\square$

Borel-Cantelli 1 $\sum_n \mathbb{P}(A_n) < \infty \implies \mathbb{P}(A_n \text{ i.o}) := \mathbb{P}(\limsup_n A_n) = 0$

Proof: *(click)*

$\mathbb{P}(\limsup_n A_n) \leq \mathbb{P}\left(\bigcup_{n\geq k} A_n\right) \leq \sum_{k=n}^{\infty} \mathbb{P}(A_k) \to 0$, as $n \to \infty$ $\square$

Convergence of Events $\{A_n\}$ be a seq of sets: $\lim_{n\to\infty} A_n = A \iff \lim_{n\to\infty} I_{A_n}(\omega) = I_A(\omega)\ \forall\omega$.

Proof: *(click)*

$(\Rightarrow)$ **Assume that** $\lim_{n\to\infty} \mathbf{A_n} = \mathbf{A}$. Consider the two following cases:

(i) $\underline{\omega \in A}$: $\liminf_{n\to\infty} A_n = A \implies \omega$ belongs to all but finitely many of the sets A_n

$\implies I_{A_n}(\omega) = 1$ for all but finitely many $n \implies \lim_{n\to\infty} I_{A_n}(\omega) = 1 = I_A(\omega)$.

(ii) $\underline{\omega \notin A}$: $\limsup_{n\to\infty} A_n = A \implies \omega$ belongs to at most finitely many of the sets A_n

$\implies I_{A_n}(\omega) = 0$ for all but finitely many $n \implies \lim_{n\to\infty} I_{A_n}(\omega) = 0 = I_A(\omega)$

$(\Leftarrow)$ **Assume that** $\lim_{n\to\infty} \mathbf{I_{A_n}}(\omega) = \mathbf{I_A}(\omega)\ \forall\omega$**.** Consider two cases:

(i) $\underline{\lim_{n\to\infty} A_n = B \text{ exists for some set } B \neq A}$:

$\implies$ either $B \setminus A \neq \varnothing$ or $A \setminus B \neq \varnothing \implies \lim_{n\to\infty} A_n = B$ & $\lim_{n\to\infty} I_{A_n}(\omega) = I_A(\omega)$.

– If $B \setminus A \neq \varnothing$: $\omega \in B \setminus A \xRightarrow[\text{above}]{\text{Result}} \lim_{n\to\infty} I_{A_n}(\omega) = I_B(\omega) = 1$

But, $\omega \notin A \implies I_A(\omega) = 0 \implies \lim_{n\to\infty} I_{A_n}(\omega) = I_A(\omega) = 0$ $\boxed{\Rightarrow\Leftarrow}$

– If $A \setminus B \neq \varnothing$: $\omega \in A \setminus B \xRightarrow[\text{above}]{\text{Result}} \lim_{n\to\infty} I_{A_n}(\omega) = I_B(\omega) = 0$

But $\omega \in A \implies I_A(\omega) = 1 \implies \lim_{n\to\infty} I_{A_n}(\omega) = I_A(\omega) = 1$ $\boxed{\Rightarrow\Leftarrow}$

(ii) $\underline{\text{The limit } \lim_{n\to\infty} A_n \text{ does not exist:}}$

$\implies \liminf_{n\to\infty} A_n \subsetneq \limsup_{n\to\infty} A_n \implies \exists\omega \in \infty$ many A_n's, but also $\omega \notin$ to ∞ many A_n's.

$$\implies I_{A_n}(\omega) = \begin{cases} 0, & \text{for } \infty \text{ many choices of } n, \\ 1, & \text{for } \infty \text{ many choices of } n. \end{cases}$$

$\implies I_{A_n}(\omega)$ does not converge $\implies$ the condition $\lim_{n\to\infty} I_{A_n}(\omega) = I_A(\omega)\ \forall\omega$, cannot hold. $\square$

Continuity of Prob. Measures $\{A_n\}$ a sequence of events: $A_n \to A \implies \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$

Proof: *(click)*

Dr Can: As "$A_n \to A \Leftrightarrow \mathbb{1}_{A_n} \to \mathbb{1}_A$" and $\mathbb{1}_{A_n} \leq 1 \in L^1(\Omega, \mathscr{F}, \mathbb{P}) \xRightarrow{\text{DCT}} \mathbb{P}(A_n) = \mathbb{E}[\mathbb{1}_{A_n}] \to \mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$.

ZBC: $B_n = \bigcap_{k=n}^{\infty} A_k \nearrow A$ and $C_n = \bigcup_{k=n}^{\infty} A_k \searrow A$. By Monotone Convergence of Measures:

$\limsup_{n\to\infty} \mathbb{P}(A_n) \leq \lim_{n\to\infty} \mathbb{P}(C_n) = \mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(B_n) \leq \liminf_{n\to\infty} \mathbb{P}(A_n)$ so all are equal. $\square$

1.3 Day 3: Random Variables

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Measurable Function $h : (X, \Sigma) \to (\mathbb{R}, \mathscr{B})$ measurable if

$$f^{-1}(B) := \{\omega \in \Omega : f(\omega) \in B\} \in \Sigma, \ \forall B \in \mathscr{B}$$

Compactification Use $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$

Borel function $]\ \Sigma = \sigma(S) = \mathscr{B}(S)$, S topological

Random Variable $X : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$ is $(\mathscr{F}, \mathscr{B})$measurable

σ-algebra generated by rv's $\sigma(X_i : i \in I) = \sigma\left(\{\omega \in \Omega : X_i(\omega) \in B\} : i \in I, B \in \mathscr{B}\right) \subset \mathscr{F}$. Smallest $\sigma-$algebra preserving measurability of X_i's. $\sigma(X_n : n \in \mathbb{N}) = \sigma\underbrace{\left(\bigcup_n \sigma(X_k : k \le n)\right)}_{\pi \text{ system}}$

Law of a R.V. $X : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$, $\mathbb{P} : \mathscr{F} \to [0, 1]$, define $\mathbb{P}_X : \mathscr{B} \to [0, 1]$ as $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$. CDF uniquely defines $\mathbb{P}_X$ (argue by uniqueness in $\pi-$systems)

MAIN THEOREMS:

Properties of Inverse Map and Measurability

- $h^{-1}(\bigcup_\alpha A_\alpha) = \bigcup_\alpha h^{-1}(A_\alpha)$ and $h^{-1}(A^c) = \left(h^{-1}(A)\right)^c$
- If $\mathscr{B} = \sigma(C)$, then suffice to check measurability of $h^{-1}(G)$, $\forall G \in C$

Proof: *(click)*

$A = \{G \in \mathscr{B} : f^{-1}(G) \in \Sigma\}$. $C \subseteq A \subseteq \mathscr{B} \Rightarrow \mathscr{B} = \sigma(C) \subseteq \sigma(A) \subseteq \sigma(\mathscr{B}) = \mathscr{B}$.

Conclude by showing A is a $\sigma-$algebra. $\square$

- $h : S \to \mathbb{R}$ continuous $\Rightarrow$ h Borel (S topological). Application: take $C = \pi(\mathbb{R})$.
- $h : (\mathbb{R}, \mathscr{B}) \to (\mathbb{R}, \mathscr{B})$ monotonic $\Rightarrow$ h measurable.
- $h_1, h_2 : (S, \Sigma) \to (\mathbb{R}, \mathscr{B})$ measurable $\Rightarrow$ $h_1 + h_2$, $h_1 \cdot h_2$, and λh_1 ($\lambda \in \mathbb{R}$) measurable.
- $h_1 : (S_1, \Sigma_1) \to (S_2, \Sigma_2)$ & $h_2 : (S_2, \Sigma_2) \to (S_3, \Sigma_3) \Rightarrow h_2 \circ h_1 : (S_1, \Sigma_1) \to (S_3, \Sigma_3)$ measurable.
- $h_n : (S, \Sigma) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ measurable $\Rightarrow$ $\inf h_n$, $\liminf_n h_n$, $\limsup_n h_n$ $(\Sigma, \mathscr{B})$ measurable.
- h_n:$(S, \Sigma) \to (\mathbb{R}, \mathscr{B})$ measurable $\Rightarrow A := \{\omega \in \Omega : \lim_n h_n(\omega)$ exists in $\mathbb{R}\}$ measurable($\in\Sigma$)

Proof: *(click)*

$h^+ = \limsup_n h_n$, and $h_- = \liminf_n h_n$, so:

$A = \{h^+(x) < \infty\} \bigcap \{h_-(x) > -\infty\} \bigcap \{h^+(x) = h_-(x)\}$

$= \left(\bigcup_n \{h^+(x) < n\}\right) \bigcap \left(\bigcup_n \{h_-(x) > -n\}\right) \bigcap \left((h^+ - h_-)^{-1}(\{0\})\right)$

Can also use a trick: take the union via $\mathbb{Q}$ Trick: $\square$

CDF $F : \mathbb{R} \to [0, 1]$ is a CDF $\Leftrightarrow$

- $\star$ $F(-\infty) = 0$, $F(+\infty) = 1$
- $\star$ F monotonically increasing (note: so # of discontinuities of $f \leq \aleph_0 = |\mathbb{N}|$)
- $\star$ F right continuous: $\lim\limits_{t \searrow t_0} F(t) = F(t_0)$ (filled dot $\bullet$ always on right, empty dot $\circ$ on left)

Chapter 2

Week 2: Random Variables, Independence and Integration

2.1 Day 1: Skorokhod representation. Doob-Dynkin Lemma. Monotone Class Theorem.

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

σ-algebra Generated by a r.v. Intuition: if event $A \in \sigma(X) \Leftrightarrow$ I can decide if A happened or not based on the result of the experiment I am doing on X.

Observing the value of $X(\omega) \iff$ observing the value of $\mathbb{1}_A(\omega)\ \forall A \in \sigma(X)$

Observing the values of $X_i(\omega),\ i \in I \iff$ observing the value of $\mathbb{1}_A(\omega)\ \forall A \in \sigma(X_i,\ i \in I)$

Ex: flip coin, X_i = outcome of i^{th} toss. $X_1 \in \sigma(X_2)$ but $X_2 \notin \sigma(X_1)$

***d*-system** S a set, $\mathcal{D}$ = collection of subsets of S. $\mathcal{D}$ = d-system on $S \iff$

(a) $S \in \mathcal{D}$

(b) $A, B \in \mathcal{D}$: $A \subset B \implies B \backslash A \in \mathcal{D}$

(c) $A_n \in \mathcal{D}$: $A_n \nearrow A \implies A \in \mathcal{D}$

Dynkinization $\mathcal{C}$ = class of subsets of S: $d(\mathcal{C}) = \cap\{\text{all d-systems containing } \mathcal{C}\}$.

$d(\mathcal{C})$ = smallest d-system containing $\mathcal{C}$ & $d(\mathcal{C}) \subset \sigma(\mathcal{C})$

MAIN THEOREMS:

Skorokhod Representation Construct r.v. with given CDF $F(\cdot)$.

Note you/computer can generate any r.v. Y by applying F_Y^{-1} to a random $U[0,1]$ r.v.

Proof: *(click)*

$F : \mathbb{R} \to [0,1]$ a CDF, $(\Omega, \mathscr{F}, \mathbb{P}) = ([0,1], \mathscr{B}([0,1]), \text{Leb})$ $(\Rightarrow \mathbb{P}(\omega \leq t) = t$ for $t \in [0,1])$

$X^- := \inf\{z : F(z) \geq \omega\}$ and $X^+ := \inf\{z : F(z) > \omega\}$.

Goal: $\omega \leq F(c) \Leftrightarrow X^-(\omega) \leq c$, $F(c) = \mathbb{P}(\omega \leq F(c)) = \mathbb{P}(X^-(\omega) \leq c) = F_{X^-}(c)$.

$(\because) z > X^-(\omega) \Rightarrow F(z) \geq \omega$ ($z > \inf$, hence can't be a lower bd on set).

So $F(X^-(\omega)) = \lim_{z \searrow X^-(\omega)} F(z) \geq \omega$ (right-continuity).

Note $\mathbb{P}(X^+ = X^-) = 1$ $(\because)$use: $\{X^- \neq X^+\} \overset{\text{Trick:}}{=\!=} \bigcup_{q \in \mathbb{Q}} \{X^- < q < X^+\}$. $\square$

Handy: $\sigma(Y) = Y^{-1}(\mathscr{B}) := (\{\omega : Y(\omega) \in B\}\ \forall B \in \mathscr{B})$

& $\pi(Y) = Y^{-1}(\pi(\mathbb{R})) := (\{\omega : Y(\omega) \leq x\}\ \forall x \in \mathbb{R})$

Doob-Dynkin $X, Y : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$: X is $\sigma(Y)$-measurable $\Leftrightarrow X = f(Y)$, f $(\mathscr{B}, \mathscr{B})$-meas.

Note $X, \{Y_i : i \in I\} : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$: X is $\sigma(Y_i : i \in I)$-measurable

$\Leftrightarrow \exists$ countable seq $(i_1, i_2, \ldots, i_n, \ldots)$ S.T. $X = f(Y_{i_1}, Y_{i_2}, \ldots, Y_{i_n}, \ldots)$, f $(\mathscr{B}(\mathbb{R}^{\mathbb{N}}), \mathscr{B})$-meas.

Careful! ! I finite: stop at n & use $f \in \mathscr{B}^n$. I uncountable: $\mathscr{B}(\mathbb{R}^I) \gg \prod_{i \in I} \mathscr{B}(\mathbb{R})$ so need $f \in \prod_{i \in I} \mathscr{B}(\mathbb{R})$.

Proof: *(click)*

$\mathcal{H} := \{X : \exists f : (\mathbb{R}, \mathscr{B}) \to (\mathbb{R}, \mathscr{B})$ bounded, S.T. $X = f(Y)\}$. Goal: $\mathcal{H}$ is a Monotone Class.

Step 1: $\mathcal{H} =$ V-Space

Step 2: $X(\omega) = 1 \in \mathcal{H}$

Step 3: $X_n \in \mathcal{H}$, $0 \leq X_n \nearrow X$ & X bdd $\Rightarrow X \in \mathcal{H}$. 1,2,3 $\Rightarrow$ apply Monotone Class Thm

$(\because) X_n = f_n(Y)$, take $f = \limsup_n f_n$, show $X = f(Y)$ (need: argue f measurable, bdd, etc.).

$f(Y(\omega) = \limsup_n f_n(Y(\omega)) = \lim_n f_n(Y(\omega)) = X(\omega)$, 1st eq from def, 2nd lim $\exists$, 3rd by assertion.

Step 4: $\forall F \in \sigma(Y)$, $1_F \in \mathcal{H}$.

$(\because) \sigma(Y) = Y^{-1}(\mathscr{B}) \Rightarrow F \in \sigma(Y) \Rightarrow F = Y^{-1}B$, $\exists B \in \mathscr{B}$. Show $1_F(\omega) = 1_B(Y(\omega))$.

$\square$

π & d-systems S = set, Σ = collection of subsets S: $\Sigma = \sigma$-algebra $\Leftrightarrow \Sigma = \pi$-syst AND d-syst.

Dynkin's Lemma $\mathcal{I} = \pi$-system $\Rightarrow d(\mathcal{I}) = \sigma(\mathcal{I})$

If a d-system $\mathcal{D}$ contains a π-system $\mathcal{I}$: $\mathcal{D}$ contains the σ-algebra $\sigma(\mathcal{I})$ generated by $\mathcal{I}$.

Proof: *(click)*

Goal: $d(\mathcal{I}) = \pi$-system (use above prop to conclude).

Let $\mathcal{D}_1 := \{A \in d(\mathcal{I}) : B \cap A \in d(\mathcal{I}),\ \forall B \in \mathcal{I}\}$ & $\mathcal{D}_2 := \{A \in d(\mathcal{I}) : B \cap A \in d(\mathcal{I}),\ \forall B \in d(\mathcal{I})\}$

Show: $\mathcal{D}_1 \supset \mathcal{I}$, $\mathcal{D}_1 = d$-system $\Rightarrow \mathcal{D}_1 = d(\mathcal{D}_1) = d(\mathcal{I})$ (as $\mathcal{D}_1 \subset \mathcal{I}$ by def)

Show: $\mathcal{D}_2 \supset \mathcal{I}$ (use $\mathcal{D}_1 = \mathcal{I}$ result), $\mathcal{D}_2 = d$-system $\Rightarrow \mathcal{D}_2 = d(\mathcal{D}_2) = d(\mathcal{I})$ (as $\mathcal{D}_2 \subset \mathcal{I}$ by def)

$\Rightarrow d(\mathcal{I}) = \pi$-system. $\square$

Monotone-Class $\mathcal{P}$ a property, $\mathcal{H} = \{f : S \to \mathbb{R} : f$ bdd and satisfying $\mathcal{P}\}$, $\mathcal{I}$ a π-system.

$$(\star) = \begin{cases} (i) & f, g \in \mathcal{H}, \alpha \in \mathbb{R} \implies f + g \;\&\; \alpha f \in \mathcal{H} \\ (ii) & 1_S \in \mathcal{H} \text{ (cst function 1)} \\ (iii) & 0 \leq f_n \nearrow f, f_n \in \mathcal{H}, f \text{ bounded } \implies f \in \mathcal{H} \\ (iv) & A \in \mathcal{I} \implies \mathbb{1}_A \in \mathcal{H} \end{cases}$$

THEN: "$\mathcal{H}$ satisfies $(\star) \implies g \in \mathcal{H},\ \forall g \in \sigma(\mathcal{I})$ bounded $(g : S \to \mathbb{R})$", i.e., $\boxed{\mathrm{b}\sigma(I) \subset \mathcal{H}}$.

<u>**Idea:**</u> approx any bdd funct $S \to \mathbb{R}$ by an $\nearrow$ seq of simple functs (lin. combo of $\mathbb{1}$ functs)

Note : The natural setup is as follows. Given a property $\mathcal{P}$, we want to verify this is fulfilled by any b$\sigma(I)$-function. We define $\mathcal{H}$ as the set of all $b\sigma(I)$-functions obeying $\mathcal{P}$, then try to show $(\star)$ is also satisfied to reach to the conclusion.

Proof: *(click)*

$D := \{A \in \sigma(\mathcal{I}) : 1_A \in \mathcal{H}\}$. Show $D \supseteq \mathcal{I}$ (by def) & D=d-system (use all $\star$)

$\Rightarrow D = d(D) \supseteq d(\mathcal{I}) = \sigma(\mathcal{I})$. Take g_n simple, $0 \leq g_n \nearrow g$ and use $\star(i, iii)$, so $g \in \mathcal{H}$. $\square$

Uniqueness Lemma S a set, $\mathcal{I} = \pi$-system on S, $\Sigma = \sigma(\mathcal{I})$, μ_1, μ_2 measures on (S, Σ).

$\mu_1 = \mu_2$ on $\mathcal{I}$(with $\mu_1(S) = \mu_2(S) < \infty) \implies \mu_1 = \mu_2$ on Σ.

Proof: *(click)*

$\mathcal{D} = \{F \in \Sigma : \mu_1(F) = \mu_2(F)\}$: Show $\mathcal{D} = d$-system on S, $D \supset \mathcal{I}$ (by def)

$\implies \mathcal{D} = d(\mathcal{D}) \supset d(\mathcal{I}) = \sigma(\mathcal{I}) = \Sigma$ (by Dynkin) $\implies \mathcal{D} = \Sigma$. $\square$

2.2 Day 2: Independence & Borel-Cantelli Lemma

Main Reference(s): – David Williams, *Probability With Martingales*, 1991 [13] Chap.4

– Rick Durrett, *Probability: Theory and Example*, 2010 [5] Sec. 2.3

KEYWORDS: Given $(\Omega, \mathscr{F}, \mathbb{P})$ triplet

Indep. σ-alg. $\mathscr{F}_1, \mathscr{F}_2, \dots \subset \mathscr{F}$, if $\forall\ A_{i_1} \in \mathscr{F}_{i_1}, \dots, A_{i_n} \in \mathscr{F}_{i_n}$, $\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_n}) = \prod_{k=1}^n \mathbb{P}(A_{i_k})$.

Careful! here done for finite collection; infinite needs continuity argument

Indep. of r.v. $X_1, X_2, \dots \perp\!\!\!\perp. \Leftrightarrow \sigma(X_1), \sigma(X_2), \dots$ are $\perp\!\!\!\perp$.

Indep. of events $E_1, E_2, \dots$ indep. if $\sigma(1_{E_1}), \sigma(1_{E_2}), \dots$ are $\perp\!\!\!\perp$.

Careful! $\sigma(1_{E_i}) = \{\varnothing, \Omega, E_i, E_i^c\}$, so complements also $\perp\!\!\!\perp$.

MAIN THEOREMS:

Indep. π-sys. $\Rightarrow$ Indep. σ-alg. $\mathcal{I} \perp\!\!\!\perp \mathcal{J}$ (π-sys.) $\Rightarrow \sigma(\mathcal{I}) \perp\!\!\!\perp \sigma(\mathcal{J})$.

Borel-Cantelli 1 $\sum_n \mathbb{P}(A_n) < \infty \implies \mathbb{P}(A_n \text{ i.o}) := \mathbb{P}(\limsup_n A_n) = 0$

Proof: *(click)*

$\mathbb{P}(\limsup_n A_n) \leq \mathbb{P}\left(\bigcup_{n \geq k} A_n\right) \leq \sum_{k=n}^\infty \mathbb{P}(A_k) \to 0$, as $n \to \infty$ $\square$

Borel-Cantelli 2 $\sum_n \mathbb{P}(A_n) = \infty$ & $A_n \perp\!\!\!\perp \implies \mathbb{P}(A_n \text{ i.o}) := \mathbb{P}(\limsup_n A_n) = 1$

Proof: *(click)*

$$\mathbb{P}\left((\limsup_n A_n)^c\right) = \mathbb{P}\left(\liminf_n A_n^c\right) = \mathbb{P}\left(\cup_{n\geq 1} \cap_{k \geq n} A_k^c\right) \leq \sum_{n=1}^\infty \mathbb{P}\left(\cap_{k\geq n} A_k^c\right) \overset{A_n \perp\!\!\!\perp}{=\!=} \sum_{n=1}^\infty \prod_{k=n}^\infty \mathbb{P}\left(A_k^c\right) = \sum_{n=1}^\infty 0$$

$$(\because) \prod_{k \geq n} \mathbb{P}(A_k^c) = \prod_{k\geq n}(1 - \mathbb{P}(A_k)) \leq \underbrace{e^{-\sum_{k\geq n} \mathbb{P}(A_k)}}_{=e^{-\infty}} = 0 \text{ since } 1 - x \geq e^{-x}\ \forall x \geq 0.$$ $\square$

Kolmogorov's Law of Iterated Logarithm $X_1, X_2, \dots$ i.i.d. r.v.s with $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$.

Let $S_n = X_1 + X_2 + \ldots + X_n$. Then: $\begin{cases} \limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = +1, & \text{a.s.,} \\ \liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = -1, & \text{a.s.} \end{cases}$

Note See Section 4.7 for proof when $X_n \sim \mathscr{N}(0,1)$.

Strassen's Law of Iterated Logarithm $X_1, X_2, \ldots$ i.i.d. r.v.s with mean 0 and variance 1. Let $S_n = X_1+X_2+\ldots+X_n$ and define the linear interpolation $S_t = (t-n)S_{n+1}+(n+1-t)S_n$, $\forall t \in [n, n+1)$. Define:

$\star Z_n(t,\omega) = \frac{S_{nt}(\omega)}{\sqrt{2n \log\log n}}$, $\forall t \in [0,1]$

$\star K(\omega) = \{f : t \mapsto f(t,\omega) \mid \exists$ a seq $n_1(\omega), n_2(\omega) \ldots \in \mathbb{N}$ S.T.$Z_n(t,\omega) \to f(t,\omega)$ unif in $t \in [0,1]\}$

$\star K = \left\{ f : t \mapsto f(t,\omega) \mid f \in C[0,1] \text{ S.T. } f(t) = \int_0^t h(s)\, ds \text{ where } \int_0^1 h(s)^2\, ds \leq 1 \right\}$

THEN: $\quad \mathbb{P}(K(\omega) = K) = 1$, and $\begin{cases} \sup\{f(1) \mid f \in K\} = +1, \\ \inf\{f(1) \mid f \in K\} = -1. \end{cases}$,

$\implies$ **Almost all paths have the same limiting shapes.**

But $\{f \in K \mid f(1) = 1\} = \{f(t) = t\}$ only, big values of S occur when the whole (rescaled to Z) path looks like the line $f(t) = t$.

$\implies$ **Almost all (rescaled) paths Z look i.o. to $\mathbf{f(t) = t}$ and i.o. to $\mathbf{f(t) = -t}$.**

Durrett 2.3.7/Williams E4.6 $\{X_n$ i.i.d $\}$, $S_n = X_1 + \ldots + X_n$:

$\mathbb{E}[|X_i|] = \infty \implies \mathbb{P}(|X_n| \geq n \text{ i.o. }) = 1$ and $\mathbb{P}\left(\lim_n \frac{S_n}{n} \text{ exists} \in (-\infty, +\infty)\right) = 0$

So $\mathbb{E}[|\mathbf{X}|] < \infty$ is necessary for the SLLN

Proof: *(click)*

First: $\infty = \mathbb{E}[|X|] = \int_0^{+\infty} \mathbb{P}\left(|X| > x\right)\, dx \leq \sum_{n=0}^{\infty} \mathbb{P}\left(|X| > n\right)$ & $\{X_i\}$ i.i.d $\implies \mathbb{P}\left(|X_n| \geq n \text{ i.o. }\right) = 1$.

Second: Let $C = \{\omega : \lim_n \frac{S_n}{n} \text{ exists} \in (-\infty, +\infty)\}$. GOAL: $\mathbb{P}(C) = 0$.

$\frac{S_n}{n} - \frac{S_{n+1}}{n+1} = \frac{S_n}{n(n+1)} - \frac{X_{n+1}}{n+1} \Rightarrow$ on $C \cap \{\omega : |X_n| \geq n \text{ i.o. }\}$: $\frac{S_n}{n(n+1)} \to 0$ and $|\frac{S_n}{n} - \frac{S_{n+1}}{n+1}| > \frac{2}{3}$ i.o. ,

therefore $\omega \notin C$, hence $C \cap \{\omega : |X_n| \geq n \text{ i.o. }\} = \varnothing$. But $\mathbb{P}\left(|X_n| \geq n \text{ i.o. }\right) = 1 \implies \mathbb{P}(C) = 0$. $\square$

Random Cesaro Means $\{X_n \perp\!\!\!\perp\}$, $c \in \mathbb{R}$, $S_n = X_1 + \ldots + X_n$: $X_n \xrightarrow{\text{a.s.}} c \implies \frac{S_n}{n} \xrightarrow{\text{a.s.}} c$.

Proof: *(click)*

For simplicity, assume that $X_n \geq c$. $X_n \xrightarrow{\text{a.s.}} c$, so for a.e. ω, $\forall \varepsilon > 0$, $\exists N(\omega)$, $X_n - c \leq \varepsilon$, $\forall n \geq N(\omega)$.

Now fix $\omega \in \Omega$: $\frac{S_n(\omega)}{n} = \frac{X_1(\omega) + \ldots + X_N(\omega)}{n} + \frac{n-N}{n} \frac{X_{N+1}(\omega) + \ldots + X_n(\omega)}{n-N}$, so

$\frac{S_n(\omega)}{n} = \frac{X_1(\omega) + \ldots + X_N(\omega)}{n} + \frac{n-N}{n} \frac{(X_{N+1}(\omega) - c) + \ldots + (X_n(\omega) - c) + (n-N) \times c}{n-N}$, so

$\frac{S_n(\omega)}{n} \leq \underbrace{\frac{X_1(\omega) + \ldots + X_N(\omega)}{n}}_{\to 0 \text{ as } n \to \infty} + \underbrace{\frac{n-N}{n} \frac{(n-N) \times (c+\varepsilon)}{n-N}}_{\to c \text{ as } n \to \infty} \longrightarrow c$

Also, $X_n \geq c \Rightarrow S_n \geq c$. Hence $\lim_n S_n(\omega) = c$ for almost all ω, i.e., $S_n \xrightarrow{\text{a.s.}} c$.

We can generalize the proof by looking at $c - \varepsilon \leq X_n \leq c + \varepsilon$ when $n \geq N(\omega)$. $\square$

EXERCISES:

Exercise:

Let X_n be nonnegative r.v.s (not necessarily independent) S.T. $0 < \mathbb{P}(X_n \geq B) \leq \frac{1}{B^2}$, $\forall\ B \geq 0$, $\forall\ n$.

(a) Fix some $B > 0$. Show that $\mathbb{P}\Big(\limsup_{n\to\infty} \frac{X_n}{n} \geq B\Big) = 0$.

(b) Show that, with probability 1, $\lim_{n\to\infty}(X_n/n)$ exists and is equal to zero.

Solution: *(click)*

(a) We have

$$\mathbb{P}\left(\frac{X_n}{n} \geq \frac{B}{2}\right) = \mathbb{P}\left(X_n \geq n\frac{B}{2}\right) \leq \frac{4}{B^2 n^2}, \quad \forall n.$$

Therefore,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{X_n}{n} \geq \frac{B}{2}\right) \leq \sum_{n=1}^{\infty} \frac{4}{B^2 n^2} < \infty.$$

By the BC Lemma, the event $\{X_n/n \geq B/2\}$ can only happen a finite number of times (w.p. 1), so

$$\mathbb{P}\Big(\limsup_{n\to\infty} \frac{X_n}{n} > \frac{B}{2}\Big) = 0.$$

The strict inequality is crucial here: e.g., $\frac{X_n}{n} = \frac{B}{2} - \frac{1}{n} \implies \limsup_n \frac{X_n}{n} = \frac{B}{2}$, therefore $\mathbb{P}(\limsup_{n\to\infty} \frac{X_n}{n} > \frac{B}{2}) = \mathbb{P}(\limsup_{n\to\infty} \frac{B}{2} > \frac{B}{2}) = 0$ would fail for "$\leq$".

Since $0 \leq \mathbb{P}\Big(\limsup_{n\to\infty} \frac{X_n}{n} \geq B\Big) \leq \mathbb{P}\Big(\limsup_{n\to\infty} \frac{X_n}{n} > \frac{B}{2}\Big) = 0$, we get $\mathbb{P}\Big(\limsup_{n\to\infty} \frac{X_n}{n} \geq B\Big) = 0$.

(b) We define $A_k = \Big\{\omega : \limsup_{n\to\infty} \frac{X_n(\omega)}{n} \leq \frac{1}{k}\Big\}$:
Since X_n are nonnegative, we observe that

$$A = \bigcap_{k=1}^{\infty} A_k = \Big\{\omega : \lim_{n\to\infty} \frac{X_n(\omega)}{n} = 0\Big\}.$$

Using the result from part (a), we have

$$\mathbb{P}(A_k^c) = \mathbb{P}\left(\Big\{\omega : \limsup_{n\to\infty} \frac{X_n(\omega)}{n} > \frac{1}{k}\Big\}\right) = 0, \quad \forall k.$$

Therefore,

$$\mathbb{P}(A^c) = \mathbb{P}\Big(\bigcup_{k=1}^{\infty} A_k^c\Big) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k^c) = 0 \implies \mathbb{P}(A) = 1.$$

Exercise:

Let A_n be a sequence of independent events with $\mathbb{P}(A_n) < 1$ for all n, and $\mathbb{P}(\cup_n A_n) = 1$.

(a) Show that $\mathbb{P}(A_n \text{ i.o.}) = 1$.

(b) Give an example to show that the conclusion of part (a) does not necessarily hold if we remove the assumption that $\mathbb{P}(A_n) < 1$ for all n.

Solution: *(click)*

We are given that

$$\mathbb{P}\big((\cup_n A_n)^c\big) = 0,$$

or

$$\mathbb{P}(\cap_n A_n^c) = 0.$$

Using the multiplication theorem (part (e) of Theorem 1 in the notes for Lecture 3), and independence, we conclude that

$$\prod_n (1 - \mathbb{P}(A_n)) = 0.$$

We will then use the following fact.

Claim: if a sequence $\{x_i\}$, with $0 \le x_i < 1$, satisfies $\prod_i (1 - x_i) = 0$ x_i in $[0, 1)$, then $\sum_i x_i = \infty$.

Proof of the claim: We can assume that after some integer N, every x_i is below $1/2$; else, the conclusion follows trivially. Note that for every $x \in [0, 1/2]$, we have $\log(1 - x) \ge -2x$. To see this, observe that both the left-hand and right-hand side are 0 at $y = 0$, but the right-hand side has a smaller derivative throughout $[0, 1/2]$.

Since $\prod_{i=N}^{\infty} (1 - x_i) = 0$, we take logarithms to obtain

$$\sum_{i=N}^{\infty} -2x_i \le \sum_{i=N}^{\infty} \log(1 - x_i) = -\infty,$$

which establishes the claimed result.

Now we apply the above claim to obtain

$$\sum_n \mathbb{P}(A_n) = \infty.$$

Using the Borel-Cantelli lemma, we conclude that A_n occurs infinitely often, with probability 1.

For part (b), consider the sequence of events where $A_1 = \Omega$, and A_n is empty for $n > 1$. We have $\mathbb{P}(\cup_n A_n) = 1$, and the events are independent. However, $\mathbb{P}(A_n \text{ i.o.}) = 0$.

Exercise:

(Durrett [5] 2.3.13) Let $\{X_n\}$ be a sequence of independent random variables. Show that $\sup_n X_n = \infty$, a.s., if and only if $\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) < \infty$ for some c.

Remark: The problem indicates that X_i are random variables (as opposed to extended-valued random variables), which means that their range is the real numbers. On the other hand, we can't say the same about $\sup_i X_i$, since it can take the value of $+\infty$.

Solution: *(click)*

Suppose $\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) = \infty$ for all c. For any c, the probability of $\sup_i X_i < c$ must be 0 since the event $X_i > c$ occurs infinitely often with probability 1 by the Borel-Cantelli lemma. It follows that

$$\begin{aligned}\mathbb{P}\left(\sup_i X_i < \infty\right) &= \mathbb{P}\left(\bigcup_{n=0}^{\infty}\left\{\sup_i X_i < n\right\}\right)\\ &\leq \sum_n \mathbb{P}\left(\sup_i X_i < n\right)\\ &\leq \sum_n 0 = 0\end{aligned}$$

so $\sup_i X_i$ must equal $+\infty$ with probability 1.

Suppose that for some c, $\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) < \infty$. By the Borel-Cantelli lemma, this means that with probability 1, the number of times $X_i > c$ is finite. If $X_i > c$ occurs finitely many times, then $\sup_i X_i$ is finite, and it follows that $\sup_i X_i$ is finite with probability 1.

Exercise:

Let X_n be independent, identically distributed (i.i.d.) random variables, defined on the same probability space. Each X_n is exponentially distributed that are exponentially distributed, with PDF $f_X(x) = e^{-x}$, $x \geq 0$, so that $\mathbb{P}(X \geq x) = e^{-x}$, for all $x \geq 0$. Let c be a positive constant, and consider the event A that "$X_n \geq c \log n$ for infinitely many values of n."

Find a necessary and sufficient condition on c for $\mathbb{P}(A)$ to be equal to 1.

Solution: *(click)*

First, note that if X_n has PDF $f_X(x) = e^{-x}$, its CDF is $F_X(x) = 1 - e^{-x}$ and $\mathbb{P}(X_n \geq x) = e^{-x}$. Defining $A_n = \{X_n \geq c \log n\}$, a necessary and sufficient condition for the statement "A_n occurs infinitely often" is, by independence of A_n and the second Borel-Cantelli lemma, $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$. However, $\mathbb{P}(A_n) = e^{-c \log n} = \dfrac{1}{n^c}$. Therefore the series $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \frac{1}{n^c}$ diverges if and only if $c \leq 1$.

Exercise:

Let X be a nonnegative r.v., with $\mathbb{E}[X] < \infty$. Let $A_1, A_2, \ldots$ be an arbitrary sequence of events such that $\mathbb{P}(A_n) \leq 2^{-n}$, for every n, and let I_n be the indicator r.v. of A_n. Show that $\lim_{n\to\infty} \mathbb{E}[X\mathbb{1}_n] = 0$.

Solution: *(click)*

We have that $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$. By the BC lemma, and w.p. 1, only finitely many of the events A_n will occur. In particular, the seq. of r.v.s $\mathbb{1}_n$ converges to 0 a.s. It follows that $X\mathbb{1}_n$ also converges to 0 a.s. The r.v.s $X\mathbb{1}_n$ are bounded above by the integrable r.v. X. Conclude with the dominated convergence thm.

Exercise:

(Durrett [5] 2.3.11) Suppose that the events A_n satisfy $\mathbb{P}(A_n) \to 0$ and $\sum_{n=1}^{\infty} \mathbb{P}(A_n^c \cap A_{n+1}) < \infty$. Show that $\mathbb{P}(A_n \text{ i.o.}) = 0$.

Solution: *(click)*

Define the set

$$A = \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

We wish to show $\mathbb{P}(A) = 0$. Now, $A \subseteq \cup_{m=n}^{\infty} A_m$ for all m, and by monotonicity of the measure, $\mathbb{P}(A) \leq \mathbb{P}(\cup_{m=n}^{\infty} A_m)$, for all n. In addition,

$$\begin{aligned} \bigcup_{m=n}^{\infty} A_m &= A_n \cup (A_{n+1} \setminus A_n) \cup (A_{n+2} \setminus A_{n+1}) \cup \cdots \\ &= A_n \cup (A_{n+1} \cap A_n^c) \cup (A_{n+2} \cap A_{n+1}^c) \cup \cdots, \end{aligned}$$

and these are disjoint sets. Therefore by the union bound, and countable additivity,

$$\mathbb{P}(A) \leq \mathbb{P}\left(\bigcup_{m=n}^{\infty} A_m\right) = \mathbb{P}(A_n) + \sum_{m=n}^{\infty} \mathbb{P}(A_{m+1} \cap A_n^c).$$

This holds for all n, and therefore it holds in the limit as n goes to infinity. But the limit of the final expression is zero, since $\mathbb{P}(A_n) \to 0$, and since $\sum_{n=1}^{\infty} \mathbb{P}(A_n^c \cap A_{n+1}) < \infty$.

Second solution outline We observe that if A_n occurs infinitely often, then either (a) $A_n^c \cap A_{n+1}$ occurs infinitely often, or (b) there is some k such that A_n occurs for every $n \geq k$. Alternative (a) has zero probability (by applying the Borel-Cantelli Lemma to the sequence $A_n^c \cap A_{n+1}$. Alternative (b) is the event $\cup_{k=1}^{\infty} B_k$, where $B_k = \cap_{n\geq k}^{\infty} A_n$. Note that $\mathbb{P}(B_k) \leq \mathbb{P}(A_m)$ for every $m \geq k$. Since $\mathbb{P}(A_n) \to 0$, it follows that $\mathbb{P}(B_k) = 0$ for every k, from which it follows (using the union bound) that $\mathbb{P}(\cup_{k=1}^{\infty} B_k) = 0$.

Exercise:

(Williams [13] p.41) Let $\{X_n\}$ be a sequence of i.i.d Exp(1) r.v.s: then $\mathbb{P}(X_n > x) = e^{-x},\ x \geq 0$.

Show that: $\mathbb{P}(X_n > \alpha \log n \text{ for infinitely many } n) = \begin{cases} 0, & \text{if } \alpha > 1, \\ 1, & \text{if } \alpha \leq 1. \end{cases}$ and $\limsup_{n\to\infty} \dfrac{X_n}{\log n} = 1$ a.s.

Note Similarly, we can prove that:

$\mathbb{P}(X_n > \log n + \alpha \log\log n, \text{ i.o.}) = \begin{cases} 0, & \text{if } \alpha > 1, \\ 1, & \text{if } \alpha \leq 1. \end{cases}$ and

$\mathbb{P}(X_n > \log n + \log\log n + \alpha \log\log\log n, \text{ i.o.}) = \begin{cases} 0, & \text{if } \alpha > 1, \\ 1, & \text{if } \alpha \leq 1. \end{cases}$ etc.

Solution: *(click)*

$\mathbb{P}(X_n > x) = e^{-x},\ x \geq 0$ implies that $\mathbb{P}(X_n > \alpha \log n) = n^{-\alpha}$, for $\alpha > 0$. Therefore:

$\sum_{n=1}^{\infty} \mathbb{P}(X_n > \alpha \log n) = \sum_{n=1}^{\infty} n^{-\alpha} : \begin{cases} < \infty, & \text{if } \alpha > 1 \quad \Rightarrow \text{ apply BC. 1}, \\ = \infty, & \text{if } \alpha \leq 1\ (+\perp\!\!\!\perp) \Rightarrow \text{ apply BC. 2}. \end{cases}$, as desired Now $\mathbb{P}\left(\limsup_{n\to\infty} \dfrac{X_n}{\log n} \geq 1\right) \geq \mathbb{P}(X_n > \log n, \text{ i.o.}) = 1 \implies \mathbb{P}\left(\limsup_{n\to\infty} \dfrac{X_n}{\log n} \geq 1\right) = 1$. We now show that it is actually equal to 1: $\forall k \in \mathbb{N},\ \mathbb{P}\left(\limsup_{n\to\infty} \dfrac{X_n}{\log n} > 1 + \dfrac{2}{k}\right) \leq \mathbb{P}\left(X_n > \left(1 + \dfrac{1}{k}\right) \log n, \text{ i.o.}\right) = 0$.

$$\Rightarrow \mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{\log n} > 1\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} \left\{\limsup_{n\to\infty} \frac{X_n}{\log n} > 1 + \frac{2}{k}\right\}\right) \leq \sum_{k=1}^{\infty} \mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{\log n} > 1 + \frac{2}{k}\right) = 0$$

$$\Rightarrow \mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{\log n} = 1\right) = 1 \implies \limsup_{n\to\infty} \frac{X_n}{\log n} = 1 \text{ a.s.}$$

The higher order approximations are obtained in a similar way.

Exercise:

(Williams [13] E4.7) Let $\{X_n \perp\!\!\!\perp\}_{n=1}^{\infty}$ S.T. $X_n = \begin{cases} n^2 - 1, & \text{w.p. } \frac{1}{n^2}, \\ -1, & \text{w.p. } 1 - \frac{1}{n^2}. \end{cases}$

Show that $\mathbb{E}[X_n] = 0$ & $S_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{\text{a.s.}} -1$

Note **In the absence of identicality of distribution, SLLN does not hold.**

Solution: *(click)*

$\sum_{n=1}^{\infty} \mathbb{P}(X_n > -1) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. By BC. 1, $\mathbb{P}(X_n > -1 \text{ i.o.}) = 0 \Rightarrow$ For a.e. ω, $\exists N(\omega)$, $X_n \leq -1,\ \forall n \geq N(\omega)$ So $X_n(\omega) = -1,\ \forall n \geq N(\omega)$ and $X_n \xrightarrow{\text{a.s.}} -1$. Now: $S_n(\omega) = \dfrac{X_1(\omega) + \ldots + X_N(\omega)}{n} + \dfrac{n-N}{n} \dfrac{X_{N+1}(\omega) + \ldots + X_n(\omega)}{n - N}$, so $S_n(\omega) = \underbrace{\dfrac{X_1(\omega) + \ldots + X_N(\omega)}{n}}_{\to 0 \text{ as } n\to\infty} + \underbrace{\dfrac{n-N}{n} \dfrac{-1 \times (n-N)}{n-N}}_{\to -1 \text{ as } n\to\infty}$

Exercise 6. Let $\{X_n\}$ be a sequence of independent non-negative random variables. Show that sequence X_n is almost surely bounded if and only if $\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) < \infty$ for some c. (Hint: X_n a.s. bounded simply means $\mathbb{P}(\sup_n X_n = \infty) = 0$.)

Solution: Suppose that there exists such a c. Then, with probability 1, the set $S = \{n \mid X_n > c\}$ is finite. Then, $\sup_n X_n \leq \max\{c, \max_{n\in S} X_n\} < \infty$, a.s.

Conversely, if if no such c exists, then $X_n > c$, i.o. By letting k range over the integers, we see that except for a countable union of zero measure sets, then for all k, there exists n_k such that $X_{n_k} > k$, so that $\sup_n X_n = \infty$, a.s.

(Alternate Solution). The event $\{\sup_n X_n = \infty\}$ can be expressed as

$$\{\sup_n X_n = \infty\} = \bigcap_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{X_k > m\}.$$

Suppose $\sup_n X_n(\omega) = \infty$. Suppose there exists natural numbers m_0 and n_0 so that $\sup_{k\geq n_0} \leq m_0$, then

$$\begin{aligned}\sup_n X_n(\omega) &= \max\left\{\max_{j=1,\ldots,n_0-1} X_j(\omega), \sup_{k\geq n_0} X_k(\omega)\right\} \\ &\leq \max\left\{\max_{j=1,\ldots,n_0-1} X_j(\omega), m_0\right\} < \infty,\end{aligned}$$

a contradiction. Conversely suppose for all $M \in \mathbb{N}$ and for all $N \in \mathbb{N}$ there exists $k_N \geq N$ such that $X_{k_N}(\omega) > M$, then

$$\sup_n X_n(\omega) \geq X_{k_N} > M.$$

Therefore the sup is larger than any natural number and must be infinite. Hence the desired relation holds.

Rewriting the above expression

$$\left\{\sup_n X_n = \infty\right\} = \bigcap_{m=1}^{\infty} \{X_n > m \text{ i.o. }\}.$$

Hence, for all $m \in \mathbb{N}$

$$\mathbb{P}\left(\sup_n X_n = \infty\right) \leq \mathbb{P}\left(X_n > m \text{ i.o. }\right). \tag{1}$$

Moreover, as the events $\{X_n > m\}$ are nested, $\{X_n > m+1\} \subset \{X_n > m\}$, by continuity of probability

$$\mathbb{P}\left(\sup_n X_n = \infty\right) = \lim_{m\to\infty} \mathbb{P}\left(X_n > m \text{ i.o }\right). \tag{2}$$

Suppose X_n is almost surely bounded. Suppose for all $c \in \mathbb{R}$

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) = \infty.$$

then as the $\{X_n\}$ are independent by the Borel Cantelli Lemma, for all c,

$$\mathbb{P}(X_n > c \text{ i.o }) = 1.$$

In particular this holds for all $m \in \mathbb{N}$. Applying (2)

$$\mathbb{P}\left(\sup_n X_n = \infty\right) = \lim_{m\to\infty} \mathbb{P}\left(X_n > m \text{ i.o }\right) = \lim_{m\to\infty} 1 = 1,$$

a contradiction.

Conversely suppose there exists $c \in \mathbb{R}$ such that

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) < \infty \quad \Longrightarrow \quad \mathbb{P}\left(X_n > c \text{ i.o}\right) = 0,$$

by Borel Cantelli. Therefore, there exists an $m_0 \in \mathbb{N}$, $m_0 \geq c$ and by monotonicity of $\mathbb{P}$ and (1)

$$\mathbb{P}(\sup_n X_n = \infty) \leq \mathbb{P}(X_n \geq m_0) \leq \mathbb{P}(X_n \geq c) = 0.$$

Hence X_n is almost surely bounded.

Book Exercises:

Williams [13]: Appendix E4: 2, 5, 6, 7.

Grimmett [7]: Chapter 1: 2.3, 3.7, 5.3, 8.3, 8.8, 8.16.

Chapter 2: 2.2, 7.2, 7.11, 7.13.

Durrett [5]: Section 2.3 (Solution Chap 6 - same problem # unless otherwise specified):

8 (6.4), 9 (6.5), 10, 14, 15, 16, 17, 18, 19, 20.

Folland [6]: Chapter 10.2: 14, 15.

2.3 Day 3: Kolmogorov's 0-1 Law

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Tail σ-algebra $X_n \perp\!\!\!\perp$, define $\tau_n := \sigma(X_n, X_{n+1}, \ldots)$. Then $\tau := \bigcap_{n=1}^{\infty} \tau_n$.

Tail event Any $F \in \tau$ (above) is a tail event.

Example:

$\star$ $F_1 := \{\omega \in \Omega : \lim_n X_n(\omega) \text{ exists}\} \in \tau$.

$(\because) F_1 = \{\omega \in \Omega : \lim_k X_{n+k} \text{ exists}\} \in \tau_n \ (\forall n) \Rightarrow F_1 \in \bigcap_n \tau_n = \tau$.

$\star$ $F_2 := \{\omega \in \Omega : \sum_{n \geq 1} X_n(\omega) \text{ converges}\} \in \tau$.

$(\because)$Note $F_2 = \{\omega \in \Omega : \sum_{k \geq n} X_k(\omega) \text{ converges}\} \in \tau_n \ (\forall n) \Rightarrow F_2 \in \bigcap_n \tau_n = \tau$.

$\star$ $F_3 := \{\omega \in \Omega : \lim_k \dfrac{X_1 + \cdots + X_k}{k} \text{ exists}\} \in \tau$.

$(\because) F_3 = \{\omega \in \Omega : \lim_k \dfrac{X_{n+1} + \cdots + X_{n+k}}{k} \text{ exists}\} \in \tau_n \ (\forall n) \Rightarrow F_3 \in \bigcap_n \tau_n = \tau$.

Careful! $\mathscr{F}$, $\mathscr{G}_n$ a σ-algebra, $\mathscr{G}_n$ decreasing: $\bigcap_n \sigma(\mathscr{F}, \mathscr{G}_n) \stackrel{?}{=} \sigma\left(\mathscr{F}, \bigcap_n \mathscr{G}_n\right)$... Not in general!

Deciding when it is true is "a tantalizing problem"!

Example: $X_n := Y_0 Y_1 ... Y_n$, where $\{Y_n \perp\!\!\!\perp\}$ & $Y_n = \pm 1$, w.p. $\frac{1}{2}$. Let $\mathscr{Y} = \sigma(Y_1, Y_2, \ldots)$ and $\tau_n = \sigma(X_k : k > n)$. Then $\{X_n \perp\!\!\!\perp\}$ and $\bigcap_n \sigma(\mathscr{Y}, \tau_n) \neq \sigma\left(\mathscr{Y}, \bigcap_n \tau_n\right)$.

$(\because) Y_0 \in \bigcap_n \sigma(\mathscr{Y}, \tau_n)$ but $Y_0 \perp\!\!\!\perp \sigma\left(\mathscr{Y}, \bigcap_n \tau_n\right)$

MAIN THEOREMS:

Kolmogorov's 0-1 Law Let τ be the tail σ-algebra of $\{X_n \perp\!\!\!\perp\}$. Then:

(a) $\forall F \in \tau$ (above) $\implies \mathbb{P}(F) = 0$ or 1.

(b) $\forall$ r.v. X τ-measurable (above) $\implies \mathbb{P}(X = c) = 1$ ($X = c$ a.s.) for some $c \in [-\infty, +\infty]$.

Note We say that τ is "$\mathbb{P}$-trivial".

Corollary: $\{X_n \perp\!\!\!\perp\} \implies \mathbb{P}\left(\sum_n X_n \text{ converges}\right) = 0$ or 1. Which one ? See "3-Series Thm"

Proof: *(click)*

Proof outline of (a):

1. $\mathcal{X}_n = \sigma(X_1, \ldots, X_n) \perp\!\!\!\perp \tau_n = \sigma(X_{n+1}, \ldots)$. $(\because)$Use π-systems: $\mathcal{K} = \{\omega : X_i(\omega) \le x_i, 1 \le i \le n, x_i \in \overline{\mathbb{R}})\}$ and $\mathcal{J} = \{\omega : X_j(\omega) \le x_j, n+1 \le j \le n+r, x_j \in \overline{\mathbb{R}}, r \in \mathbb{N})\}$

2. $\mathcal{X}_n \perp\!\!\!\perp \bigcap_n \tau_n$ $(\because)\tau \subset \tau_n$

3. $\mathcal{X}_\infty := \sigma(X_n : n \in \mathbb{N})$: $\mathcal{X}_\infty \perp\!\!\!\perp \tau$ $(\because)\mathcal{K}_\infty := \cup_{n=1}^\infty \mathcal{X}_n$ is a π-system generating $\mathcal{X}_\infty$, $\mathcal{K}_\infty \perp\!\!\!\perp \tau$, so $\mathcal{X}_\infty \perp\!\!\!\perp \mathcal{X}_\infty$.

4. $\tau \subset \mathcal{X}_\infty \implies \tau \perp\!\!\!\perp \tau \Rightarrow \forall F \in \tau : \mathbb{P}(F) = \mathbb{P}(F \cap F) = \mathbb{P}(F)\,\mathbb{P}(F) = 0$ or 1

Proof outline of (b):

$c := \sup\{x : \mathbb{P}(X \le x) = 0\}$. Apply (a): $\mathbb{P}(X \le x) = 0, 1$, $\forall x$.

Take $c_n = c - \frac{1}{n}$, get $\mathbb{P}(X < c) = 0$. Take $c_n = c + \frac{1}{n}$, get $\mathbb{P}(X \le c) = 1$. $\square$

Exercise:

Prove that $F_1 := \{\omega \in \Omega : \lim_n X_n(\omega) \text{ exists}\} \in \tau$.

Solution: *(click)*

Recall that $\{X_n(\omega)\}_{n=1}^\infty$ converges $\iff$ $\{X_n(\omega)\}_{n=1}^\infty$ Cauchy. Hence, $\forall \varepsilon > 0, \exists N \in \mathbb{N}_{\ge k}$ s.t. $n, m \ge N$ $\Rightarrow |X_n(\omega) - X_m(\omega)| < \varepsilon$ $(\forall k \in \mathbb{N})$. Thus, $F_1 = \bigcap_{\varepsilon > 0} \bigcup_{N \ge k} \bigcap_{n,m \ge N} \{\omega \in \Omega : |X_n(\omega) - X_m(\omega)| < \varepsilon\}$. Since $\{\omega \in \Omega : |X_n(\omega) - X_m(\omega)| < \varepsilon\} = (X_n - X_m)^{-1}\big((-\varepsilon, \varepsilon)\big) \in \tau_k$, and we take countable intersections/unions etc., $F_1 \in \tau_k, \forall k \Rightarrow F_1 \in \cap_k \tau_k = \tau$. To approximate ε's with a countable collection, consider $\frac{1}{l}, l \in \mathbb{N}$.

2.4 Day 4: Integration

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS: $(\Omega, \mathscr{F}, \mu)$ measure space. Assume $0 \times \infty := 0$.

Notation $\mu(f) := \int_\Omega f(\omega)\, \mu(d\omega) = \int_\Omega f\, d\mu$.

$\mu(f; A) := \mu(f\mathbb{1}_A) := \int_A f(\omega)\, \mu(d\omega) = \int_A f\, d\mu$ for $A \in \mathscr{F}$.

Note Sum is a special kind of integral: the integral's form is tailored to the measure used.

Positive Simple Functions (SF^+) $\exists n \in \mathbb{N}$ S.T. $f = \sum_{k=1}^n a_k \mathbb{1}_{E_k}$ $(E_k \in \mathscr{F},\, a_k \geq 0)$

Canonical rep: a_k's distinct. All reps give same results.

Integral of an SF $\mu(\mathbb{1}_A) := \mu(A) \leq \infty$ $(A \in \mathscr{F})$.

Integral of a Positive Function $\mu(f) = \displaystyle\int_\Omega f\, d\mu :=^1 \sup_{\substack{0 \leq g \leq f \\ g \in SF^+}} \int_\Omega g\, d\mu \leq \infty$

Note $\mu(f) = \sup_{g \in SF^+} \displaystyle\sum_{a \in \text{Range}(g)} a\mu(\{\omega : g(\omega) = a\})$

Positive/Negative Part f meas: $f = f^+ - f^-$ with $f^+ = \max(f, 0)$ & $f^- = \max(-f, 0)$

Note $f^+, f^- \in SF^+$ and $|f| = f^+ + f^-$

Integrable Function $f \in L^1(S, \Sigma, \mu)$ if $\mu(|f|) = \mu(f^+) + \mu(f^-) < \infty$

Integral of Integrable Functions $f \in L^1$: $\int f\, d\mu := \int f^+\, d\mu - \int f^-\, d\mu$

Note $f \in L^1$: $|\mu(f)| \leq \mu(|f|)$

MAIN THEOREMS:

SF$^+$ Properties Let $f, g \in SF^+$ and $a, b \in [0, \infty]$

(a) $\mu(f) = \sum_{k=1}^n a_k \mu(A_k) \leq \infty$

(b) $f = g$ μ-a.e[2] $\Rightarrow \mu(f) = \mu(g)$.

(c) (Linearity) $af + bg \in SF^+$ and $\mu(af + bg) = a\mu(f) + b\mu(g)$.

(d) (Monotonicity) $0 \leq f \leq g$ a.e. $\Rightarrow \mu(f) \leq \mu(g)$.

(e) (Max-Min) $\max\{f, g\}$ & $\min\{f, g\} \in SF^+$.

[1] Reasonable def as SF^+ approx works.

[2] Specification of μ: $\mu\big(\{\omega \in \Omega : f(\omega) \neq g(\omega)\}\big) = 0$.

Null $f \geq 0$ measurable: $\mu(f) = 0 \Rightarrow f = 0$ a.e.

$(\because) A_n = \{\omega : f(\omega) \geq \frac{1}{n}\}$: $\mu(f) \geq \mu(f; A_n) \geq n\mu(A_n) \Rightarrow n\mu(f) \geq \mu(A_n) \Rightarrow \mu(A) = 0$.

$\{\omega : f(\omega) > 0\} = \bigcup_n A_n$

SF$^+$ Approximation $f \geq 0$ measurable $\implies \exists$ a seq. $\varphi_n \in SF^+$ S.T. $0 \leq \varphi_n \nearrow f$.

$(\because)$Take $\varphi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{\{\omega : \frac{k-1}{2^n} \leq f(\omega) < \frac{k}{2^n}\}} + n\mathbb{1}_{\{\omega : f(\omega) \geq n\}}$[3]

Monotone Convergence Thm (MCT) f_n, f meas: $0 \leq f_n \nearrow f \implies \mu(f_n) \nearrow \mu(f) \leq \infty$.

Proof: *(click)*

[$\mathbf{f} \geq \lim \mu(\mathbf{f_n})$] $\mu(f_1) \leq \mu(f_2) \leq \dots$ (monotonicity) $\Rightarrow \lim_{n\to\infty} \mu(f_n)$ exists[a]$\Rightarrow f \geq \lim_n \mu(f_n)$.

[$\mathbf{f} \leq \lim \mu(\mathbf{f_n})$] $g \in SF^+$, $g \leq f$. Fix $c \in (0,1)$, let $A_n := \{\omega : f_n(\omega) \geq cg(\omega)\}$.

So $A_n \nearrow \Omega$ and $\mu(f_n) \geq \mu(f_n; A_n) \geq c\mu(g; A_n) \Rightarrow \lim_n \mu(f_n) \geq c\mu(g) \Rightarrow \lim_n \mu(f_n) \geq \mu(g)$

$\Rightarrow \lim_n \mu(f_n) \geq \sup_{g \in SF^+ : 0 \leq g \leq f} \mu(g) = \mu(f)$. □

[a] Monotonic sequences of real numbers have limits.

Exercise:

Breaking Monotone Convergence:

Consider the sequence of Borel measurable functions $g_n : [0,1] \to \mathbb{R}$, defined by $g_n(x) = -\frac{1}{nx} I_{\{x>0\}}$.

1. $\lim_{n\to\infty} \int_0^1 g_n(x)dx = \lim_{n\to\infty} (-\infty)$.
2. $\lim_{n\to\infty} g_n(x) = 0$ for all $x \in [0,1]$ and thus $\int_0^1 \lim_{n\to\infty} g_n(x)dx = 0$.

We have a monotone seq. of functions, BUT the condition $\int_0^1 |g_1(x)| I_{\{\omega : g_1(\omega) < 0\}}(x)dx < \infty$ fails.

Furthermore, fails hypotheses of the DCT: $\sup_n |g_n| = \frac{1}{x} I_{\{x>0\}}$ is not integrable.

Corollary: Also holds with a.e. : $f_n \nearrow f$ a.e. then $\implies \mu(f_n) \nearrow \mu(f) \leq \infty$.

$(\because)$Let $N = \{\omega : f_n \not\nearrow f\}$ $(\mu(N) = 0)$: $g_n = f_n \mathbb{1}_{N^c}$ & $g = f\mathbb{1}_{N^c}$. Apply MCT (use $\int_\Omega = \int_{N^c}$)

Corollary: If $\mu(A) = 0$ then $\int_A f \, d\mu = 0 \; \forall f$ meas.

$(\because)$True if SF^+ $(\int_A \mathbb{1}_E \, d\mu = \int_\Omega \mathbb{1}_{E\cap A} \, d\mu = \mu(E \cap A) \leq 0)$

Corollary: On $([0,1], \mathscr{B}([0,1]), Leb)$: f Riemann-integ $\Rightarrow f$ Leb. meas. and Leb. integ.

$(\because)$See Stein p.57 Thm 1.5

SF$^+$ Integr Approx $f \geq 0$ meas. $\Rightarrow \exists$ a seq $\varphi_n \in SF^+$ S.T. $0 \leq \varphi_n \nearrow f$ & $\mu(\varphi_n) \nearrow \mu(f)$.

$(\because)$MCT

a.e. Equality $f = g \geq 0$ μ-a.e $\implies \mu(f) = \mu(g)$. $(\because)$MCT

[3] Basically, threshold y-axis at n, then discretize with step size $\frac{1}{2^n}$.

Fatou's Lemmata [4] $f:(S,\Sigma)\to(\mathbb{R},\mathscr{B})$, $g\in L^1$ (i.e., $\mu(g)<\infty$)

1. $f_n\geq g$ a.e. meas $\implies \mu(\liminf f_n)\leq \liminf \mu(f_n)$
2. $f_n\leq g$ a.e. meas: $\implies \mu(\limsup f_n)\geq \limsup \mu(f_n)$

Note Fatou $\implies$ MCT: $(\because) 0\leq f_n \nearrow f$ a.e. (so $\limsup=\liminf=\lim$), $g=0$:

Fatou $\Rightarrow \int f\,d\mu=\int \lim f_n\,d\mu\leq \lim\int f_n\,d\mu$

But $f_n\leq f$ a.e. $\implies \int f\,d\mu=\int\lim f_n\,d\mu\geq\lim\int f_n\,d\mu$

Proof: *(click)*

1. $g_k=\inf_{n\geq k} f_n\nearrow \liminf_n f_n \Rightarrow$ MCT: $\mu(g_k)\nearrow\mu(\liminf_n f_n)$.

Now: $\mu(g_k)\leq\mu(f_n)\ \forall n\geq k\Rightarrow \mu(g_k)\leq\inf_{n\geq k}\mu(f_n)\Rightarrow \lim_k\mu(g_k)\leq\lim_k\inf_{n\geq k}\mu(f_n)=\liminf\mu(f_n)$

2. Apply (1.) for $(g-f_n)$ $\square$

Exercise:

Breaking Fatou's Lemma:(strict inequalities can happen)

Consider the seq. of Borel meas. functions $f_n:[0,1]\to\mathbb{R}$, defined by $f_n=\begin{cases}\mathbb{1}_{[0,1/2]}, & \text{if } n \text{ odd,}\\ \mathbb{1}_{[1/2,1]}, & \text{if } n \text{ even.}\end{cases}$

Then $\limsup_n f_n=\mathbb{1}_{[0,1]}$ and $\liminf_n f_n=0$

$\implies \int\liminf f_n=0<\frac{1}{2}=\liminf\int f_n$ and $\int\limsup f_n=1>\frac{1}{2}=\limsup\int f_n$

Linearity $f,g\in L^1(S,\Sigma,\mu)$: $(af+bg)\in L^1$ and $\mu(af+bg)=a\mu(f)+b\mu(g)\ \forall a,b,\in\mathbb{R}$

Dominated Conv. Thm (DCT) f_n, f meas, $f_n\to f$ a.e. and $\exists g\in L^1$ S.T. $|f_n|\leq g$ a.s.

$\implies f_n\to f\in L^1$ i.e., $\mu(|f_n-f|)\to 0 \implies \mu(f_n)\to\mu(f)$

Corollary: Bounded Convergence Thm (BCT)

f_n, f meas, $f_n\to f$ a.s. and $\exists c\geq 0$ S.T. $|f_n|\leq c$ a.e. $\implies \mu(f_n)\to\mu(f)$

Proof: *(click)*

$|f_n-f|\leq 2g\in L^1 \implies$ Fatou: $0\leq\liminf\mu(|f_n-f|)\leq\limsup\mu(|f_n-f|)\leq\mu(\limsup|f_n-f|)=\mu(0)=0$

$\implies \lim\mu(|f_n-f|)=0 \implies |\mu(f_n)-\mu(f)|=|\mu(f_n-f)|\leq\mu(|f_n-f|)\to 0.$ $\square$

[4] Greek Plural ☺

Exercise:

Breaking Dominated Convergence:

Consider the sequence of Borel measurable functions $g_n : [0,1] \to \mathbb{R}$, defined by $g_n(x) = n.I_{(0,\frac{1}{n})}$.

1. $\lim_{n\to\infty} \int_{\mathbb{R}} g_n(x)dx = 1$.
2. $\lim_{n\to\infty} g_n(x) = 0$ for all $x \in \mathbb{R}$ and thus $\int_{\mathbb{R}} \lim_{n\to\infty} g_n(x)dx = 0$.

In this case, the smallest upper bound (the supremum) for $|g_n|$ is not integrable because
$$\int_0^1 \sup_n\{|g_n(x)|\}dx = \sum_{n=1}^{\infty} n\left(\frac{1}{n} - \frac{1}{n+1}\right) = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty,$$
so the condition about the existence of an integrable function that dominates the seq. fails to hold.

Exercise:

Show that: $\sum_{n=1}^{\infty} \mathbb{E}[|Z_n|] < \infty \implies \sum_{n=1}^{\infty} \mathbb{E}[Z_n] = \mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right]$,

i.e., $\sum_{n=1}^{\infty} \int |f_n| \, d\mu < \infty \implies \sum_{n=1}^{\infty} \int f_n \, d\mu = \int \left[\sum_{n=1}^{\infty} f_n\right] \, d\mu$

Solution: *(click)*

($\star$) Let $Y := \sum_{k=1}^{\infty} \mathbb{E}[|Z_k|]$ and $Y_n := \sum_{k=1}^{n} \mathbb{E}[|Z_k|] \Rightarrow$ MCT: $\mathbb{E}\left[\sum_{i=1}^{\infty} |Z_n|\right] = \sum_{i=1}^{\infty} \mathbb{E}[|Z_n|] = \mathbb{E}[Y] < \infty$
$\implies \mathbb{E}[Y] < \infty$.

($\star$) Now let $X := \sum_{k=1}^{\infty} \mathbb{E}[Z_k]$ and $X_n := \sum_{k=1}^{n} \mathbb{E}[Z_k]$: $|X_n| \leq Y_n \leq Y$, $\forall n \in \mathbb{N}$
$\Rightarrow$ DCT: $\mathbb{E}[X_n] \to \mathbb{E}[\lim_n X_n] = \mathbb{E}[X]$ (since by def, $\lim_n X_n = X$)

Scheffe's Lemma $f_n, f \in L^1$: $f_n \to f$ a.s. $\implies$ "$\mu(|f_n - f|) \to 0 \iff \mu(|f_n|) \to \mu(|f|)$"

Proof: *(click)*

($\Rightarrow$) $|\,|f_n| - |f|\,| \leq |f_n - f| \to 0$ so $\mu(||f_n| - |f||) \leq |\mu(|f_n - f|)| \to 0$

($\Leftarrow$) For simplicity, assume $f_n, f \geq 0$.

Use $\mu(f_n) \to \mu(f)$ to show $\mu(|f_n - f|) = \mu((f_n - f)^+) + \mu((f_n - f)^-) \to 0$

1. $(f_n - f)^- = \max(f - f_n, 0) \underbrace{\leq}_{f_n \geq 0} \max(f, 0) \underbrace{\leq}_{f \geq 0} f \in L^1$, $f_n \to f$ a.s. $\implies$ DCT: $\mu((f_n - f)^-) \to 0$
2. $\mu((f_n - f)^+) = \mu(\max(f_n - f, 0)) = \mu(f_n - f) - \mu(f_n - f; f_n < f) = \mu(f_n) - \mu(f) + \mu((f_n - f)^-) \to 0$

Note General case: Fatou $\Rightarrow \mu(f_n^{\pm}) \to \mu(f^{\pm})$ and apply the special case above. $\square$

2.5 Day 5: Radon-Nikodym

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Standard Machine To prove a "linear result" $\forall f \in L^1$, take following steps:

- ($\star$) Prove result for $f = \mathbb{1}_E$, $E \in \Sigma$.
- ($\star$) Use linearity to prove for $f \in SF^+$
- ($\star$) Use MCT to prove for $f \in (m\Sigma)^+$
- ($\star$) Decompose $f \in L^1$ as $f = f^+ - f^-$ and use linearity.

Note Equivalent to using Monotone Class Theorem

Absolute Continuity $\nu \ll \mu$ if $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

MAIN THEOREMS:

Lemma $f : \Omega \to \mathbb{R}_{\geq 0}$ meas., $\nu(A) := \int_A f \, d\mu$. Then, $\nu(\cdot)$ measure on (Ω, Σ) and $\nu \ll \mu$. $f := \frac{d\nu}{d\mu}$.

Proof: *(click)*

$A_i \in \Sigma$ disj: $\nu\left(\cup_i A_i\right) = \displaystyle\int_{\cup_i A_i} f \, d\mu = \lim_n \int \mathbb{1}_{\cup_{i=1}^n A_i} f \, d\mu = \lim_n \sum_{i=1}^n \int_{A_i} f \, d\mu = \lim_n \sum_{i=1}^n \nu(A_i) = \sum_i \nu(A_i)$ $\square$

Radon-Nikodym ν, μ on (Ω, Σ) S.T. $\nu \ll \mu$. Then, $\exists f : \Omega \to \mathbb{R}_{\geq 0}$ S.T. $\nu(A) = \int_A f \, d\mu$, $\forall A \in \Sigma$.

Chain Rule $\forall g : \Omega \to \mathbb{R}_{\geq 0}$ meas., $\int g \, d\nu = \int gf \, d\mu$ (f def. above). Why chain[5] ?

Proof: *(click)*

($\star$) For $g \in SF^+$: $g = \sum_{i=1}^n a_i \mathbb{1}_{E_i}$ (canonical rep.) $\Rightarrow \int g \, d\nu = \int \sum_{i=1}^n a_i \mathbb{1}_{E_i} d\nu = \sum_{i=1}^n \int a_i \mathbb{1}_{E_i} d\nu$

So: $\int g \, d\nu = \sum_{i=1}^n a_i \nu(E_i) = \sum_{i=1}^n a_i \int \mathbb{1}_{E_i} f \, d\mu = \int \left(\sum_{i=1}^n a_i \mathbb{1}_{E_i}\right) f \, d\mu = \int gf \, d\mu$.

($\star$) For $g \in (m\Sigma)^+$, take $g_n \in SF^+$ S.T. $0 \leq g_n \nearrow g$. Therefore:

$\int g \, d\nu \underbrace{=}_{\text{MCT}} \lim_n \int g_n \, d\nu \underbrace{=}_{\text{above}} \lim_n \int g_n f \, d\mu \underbrace{=}_{\text{MCT}} \lim_n gf \, d\mu,$ as $0 \leq (g_n f) \nearrow (gf)$ $\square$

Lebesgue's Thm $g : \mathbb{R} \to \mathbb{R}$ is Riemann integrable $\Leftrightarrow$ g is continuous λ- a.e.

[5] $\int g \, d\nu = \int g \frac{d\nu}{d\mu} \, d\mu = \int gf \, d\mu$.

Summary:

Measure Theory:	**Probability:**
$\int \mathbb{1}_A \, d\mu = \mu(A)$	$\mathbb{E}[1_A] = \mathbb{P}(A)$
$g \geq 0 \Longrightarrow \int g \, d\mu \geq 0$	$X \geq 0 \Longrightarrow \mathbb{E}[X] \geq 0$
$g = 0$ a.e. $\Longrightarrow \int g \, d\mu \geq 0$	$X = 0$ a.s. $\Longrightarrow \mathbb{E}[X] = 0$
$g \leq h$ a.e. $\Longrightarrow \int g \, d\mu \leq \int h \, d\mu$	$X \leq Y$ a.s. $\Longrightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
$g = h$ a.e. $\Longrightarrow \int g \, d\mu = \int h \, d\mu$	$X = Y$ a.s. $\Longrightarrow \mathbb{E}[X] = \mathbb{E}[Y]$
$g \geq 0$ a.e. $: \int g \, d\mu = 0 \Longrightarrow g = 0$ a.e.	$X \geq 0$ a.s. $: \mathbb{E}[X] = 0 \Longrightarrow Z = 0$ a.s.
$\int (g + h) \, d\mu = \int g \, d\mu + \int h \, d\mu$	$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
$\int \alpha g \, d\mu = \alpha \int g \, d\mu$	$\mathbb{E}[\alpha X] = \alpha \mathbb{E}[X]$
MCT: $0 \leq g_n \nearrow g$ a.e. $\Longrightarrow \int g_n \, d\mu \nearrow \int g \, d\mu$	**MCT:** $0 \leq X_n \nearrow X$ a.s. $\Longrightarrow \mathbb{E}[X_n] \nearrow \mathbb{E}[X]$
DCT: $g_n \to g$ a.e. , $\lvert g_n \rvert \leq h \in L^1$ a.e. $\Rightarrow \int g_n \, d\mu \to \int g \, d\mu$	**DCT:** $X_n \to X$ a.s. , $\lvert X_n \rvert \leq Y$ a.s. , $\mathbb{E}[\lvert Y \rvert] < \infty \Longrightarrow \mathbb{E}[X_n] \to \mathbb{E}[X]$
BCT: $\lvert g_n \rvert \leq c$ a.e. , $g_n \to g$ a.e. $\Rightarrow \int g_n \, d\mu \to \int g \, d\mu$	**BCT:** $\lvert X_n \rvert \leq c$ a.s. $X_n \to X$ a.s. $\Longrightarrow \mathbb{E}[\lvert X_n - X \rvert] \to 0, \mathbb{E}[X_n] \to \mathbb{E}[X$
Scheffe: $f_n \to f$ a.s. $\in L^1 : \mu(\lvert f_n - f \rvert) \to 0 \Leftrightarrow \mu(\lvert f_n \rvert) \to \mu(\lvert f \rvert)$	**Scheffe:** $X_n \to X$ a.s. , $\mathbb{E}[\lvert X_n \rvert] \to \mathbb{E}[\lvert X \rvert] \Longrightarrow \mathbb{E}[\lvert X_n - X \rvert] \to 0$
$g \geq 0 \Longrightarrow \nu(A) = \int_A g \, d\mu$ is a measure	$f \geq 0$: $\int f \, d\mathbb{P} = 1 \Longrightarrow \nu(A) = \int_A f \, d\mathbb{P}$ is a prob. measure
RN: $\nu \ll \mu \Rightarrow \exists f : \Omega \to \mathbb{R}_{\geq 0}$ S.T. $\nu(A) = \int_A f \, d\mu, \forall A \in \Sigma$	**RN:** $\mathbb{P}_2 \ll \mathbb{P}_1 \Rightarrow \exists Y : \Omega \to \mathbb{R}_{\geq 0}$ S.T. $\mathbb{P}_2(A) = \mathbb{E}_{\mathbb{P}_1}[Y \mathbb{1}_A] = \int_A Y \, d\mathbb{P}_1$

Chapter 3

Week 3: Expectation, Law of Large Numbers & Applications

3.1 Day 1: Expectation

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS: $(\Omega, \mathscr{F}, \mathbb{P})$ a Probability Space, $L^p = L^p(\Omega, \mathscr{F}, \mathbb{P})$.

Notation $\mathbb{P}(X) = \int X \; d\mathbb{P}$; $\mathbb{E}[X; F] = \mathbb{E}[X \mathbb{1}_F] = \int_F X \; d\mathbb{P}, \quad \forall F \in \mathscr{F}$.

Expectation X s.t. X^+ or $X^- \in L^1$: $\mathbb{E}[X] = \int X \; d\mathbb{P} = \int X^+ \; d\mathbb{P} - \int X^- \; d\mathbb{P}$ (avoid $\infty - \infty$)

Note $\mathbb{E}[X] = \int_0^\infty (1 - F_X(x)) \; dx - \int_{-\infty}^0 F_X(x) \; dx = \int_0^\infty \mathbb{P}(X > x) \; dx - \int_{-\infty}^0 \mathbb{P}(X < x) \; dx$.

Integral $\exists$ as $\mathbb{P}(X > x)$ Riemann integ. (monotonic $\Rightarrow \mu(\text{discont.}) = 0$, bdd.)

Trick: $\sum_{k=1}^\infty \mathbb{P}\left(|X| > k\right) \leq \int_0^\infty \mathbb{P}\left(|X| > x\right) dx \leq \sum_{k=0}^\infty \mathbb{P}\left(|X| > k\right)$, so $X \in L^1 \Leftrightarrow \sum_{k=1}^\infty \mathbb{P}\left(|X| > k\right) < \infty$

Convex Function $g : I \to \mathbb{R}$ s.t. $g\left(\lambda x + (1-\lambda)y\right) \leq \lambda g(x) + (1-\lambda) g(y) \; \forall \lambda \in [0,1]$

Note g convex $\implies g$ continuous on I. Also if g twice differentiable: g convex $\iff g'' \geq 0$.

L^p Spaces $p \in [1, \infty)$: $X \in L^p$ if $\mathbb{E}[|X|^p] < \infty$; $\|X\|_p := \left(\mathbb{E}[|X|^p]\right)^{\frac{1}{p}}$; L^p is a V-Space[1].

Truncation $|X| \in m\mathscr{F}$, $X_n := \min\{|X|, n\} = X \wedge n \Longrightarrow X_n \in L^p, \forall p \geq 1$, & $0 \leq X_n \nearrow |X|$.

MAIN THEOREMS:

Moments Calculation $\mathbb{E}[|X|^p] = \int_0^\infty p x^{p-1} \mathbb{P}\left(|X| > x\right) \; dx$

Proof: *(click)*

$$\int_0^\infty p x^{p-1} \mathbb{P}\left(|X| > x\right) \; dx = \int_0^\infty \int_\Omega p x^{p-1} \mathbb{1}_{|X(\omega)| > x} \; d\mathbb{P}(\omega) \; dx = \int_\Omega \int_0^{|X(\omega)|} p x^{p-1} \; dx d\mathbb{P}(\omega) = \int_\Omega |X|^p \; d\mathbb{P}(\omega)$$

□

Markov's Inequality $Z \in m\mathscr{F}$, $g : (\mathbb{R}, \mathscr{B}) \to ([0, \infty], \mathscr{B})$ increasing: $g(c)\mathbb{P}\left(Z \geq c\right) \leq \mathbb{E}[g(Z)]$

Proof: *(click)*

$\mathbb{E}[g(Z)] \geq \mathbb{E}[g(c) \mathbb{1}_{Z \geq c}] = g(c) \cdot \mathbb{P}\left(Z \geq c\right)$ □

Chernoff Inequality Y a r.v.: $\mathbb{P}\left(Y > c\right) \leq e^{-\theta c} \mathbb{E}[e^{\theta Y}] \;\; \forall \theta > 0, c \in \mathbb{R}$ (optimize over $\theta > 0$)[2].

---

[1] $X, Y \in L^p$, $|X + Y|^p \leq (|X| + |Y|)^p \leq 2^p \max\{|X|^p, |Y|^p\} \leq 2^p(|X|^p + |Y|^p)$. Take integral and conclude.

[2] Proven by taking $g : t \mapsto e^{\theta t}$. One can consider the 'best' $g(\cdot)$ too, namely $\mathbb{P}\left(Z \geq c\right) \leq \inf_{g \in \mathcal{S}} \frac{\mathbb{E}[g(Z)]}{g(c)}$, $\mathcal{S}$ is some function space. Even more curiously, consider the gap $\left|\mathbb{E}[g(Z)] - g(c)\mathbb{P}\left(Z \geq c\right)\right|$. How small can we make it, namely $\inf_{c \in \mathbb{R}} \inf_{g \in \mathcal{S}} \left|\mathbb{E}[g(Z)] - g(c)\mathbb{P}\left(Z \geq c\right)\right|$.

Chebyshev's Inequality $X \in L^2$: $\mathbb{P}\left(|X - \mathbb{E}[X]| > c\right) \leq \frac{\mathrm{Var}(X)}{c^2}$, $\forall c \geq 0$

Sum of r.v.s

- $\star$ $\mathbb{E}[X] < \infty \implies X < \infty$ a.s.
- $\star$ $Z_k \geq 0$: $\mathbb{E}[\sum Z_k] = \sum(\mathbb{E}[Z_k]) \leq \infty$
- $\star$ $\sum \mathbb{E}[|Z_k|] < \infty \implies Z_k \to 0,$ a.s. .
- $\star$ $\sum_{n=1}^{\infty} \mathbb{E}[|Z_n|] < \infty \implies \sum_{n=1}^{\infty} \mathbb{E}[Z_n] = \mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right]$.

Exercise:

(Alternative Pf.) If $\sum_k \mathbb{P}(E_k) < \infty$: $Z_k = \mathbb{1}_{E_k} \Rightarrow \sum_k \mathbb{E}[Z_k] < \infty \Rightarrow Z_k \to 0,$ a.s. $\Rightarrow Z_k = 0$ eventually w.p. 1 (as $Z_k \in \{0, 1\}$)

Jensen's Inequality $g : I \to \mathbb{R}$ cvx., I open; $X, g(X) \in L^1(\Omega, \mathscr{F}, \mathbb{P}) \Rightarrow \mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$.

Proof: *(click)*

(Kallenberg) g cvx $\Rightarrow g(x) = \sup_{a,b}(ax + b) \Rightarrow \mathbb{E}[g(X)] = \mathbb{E}[\sup_{a,b}(aX + b)] \geq \mathbb{E}[aX + b] = a\mathbb{E}[X] + b \Rightarrow$ $\mathbb{E}[g(X)] \geq \sup_{a,b}(a\mathbb{E}[X] + b) = g(\mathbb{E}[X])$. $\square$

Monotonicity of Norms $1 \leq p \leq r < \infty \Rightarrow \|X\|_p \leq \|X\|_r \Rightarrow L^r \subset L^p$.

Proof: *(click)*

$X_n = (|X| \wedge n)^p$, $g : x \mapsto x^{\frac{r}{p}}$ cvx. $\implies \left(\mathbb{E}\left[X_n\right]\right)^{\frac{r}{p}} \leq \mathbb{E}\left[X_n^{r/p}\right] = \mathbb{E}[(|X| \wedge n)^r] \leq \mathbb{E}[|X|^r]$, MCT it. $\square$

Cauchy-Schwarz $X, Y \in L^2 \implies |\mathbb{E}[XY]| \leq \mathbb{E}[|XY|] \leq \|X\|_2\|Y\|_2 \implies XY \in L^1$.

Corollary: $X, Y \in L^2 \Rightarrow \|X + Y\|_2 \leq \|X\|_2 + \|Y\|_2$.

Proof: *(click)*

W.l.o.g. $X, Y \geq 0$. $X_n = X \wedge n$, $Y_n = Y \wedge n$. $0 \leq \mathbb{E}[(aX_n + bY_n)^2] = a^2\mathbb{E}[X_n^2] + 2ab\mathbb{E}[X_nY_n] + b^2\mathbb{E}[Y_n^2]$ $(\forall a, b)$. Convert to quadratic (in $\frac{a}{b}$): $b^2\left(\left(\frac{a}{b}\right)^2\mathbb{E}[X_n^2] + 2\frac{a}{b}\mathbb{E}[X_nY_n] + \mathbb{E}[Y_n^2]\right) \geq 0 \implies \Delta \geq 0$ $\implies \mathbb{E}[X_nY_n]^2 \leq \mathbb{E}[X_n^2]\mathbb{E}[Y_n^2] \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$. MCT: $0 \leq X_nY_n \nearrow XY$. $\square$

Inequalities $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. $f, g \in L^p(S, \Sigma, \mu)$[3], and $h \in L^q(S, \Sigma, \mu)$ (generic)

- $\star$ **(Holder)** $|\mu(fh)| \leq \mu(|fh|) \leq \|f\|_p\|h\|_q$.
- $\star$ **(Minkovski)**[4] $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

[3] $\|f\|_p := \left(\mu(|f|^p)\right)^{\frac{1}{p}} = \left(\int |f|^p \, d\mu\right)^{\frac{1}{p}}$

[4] Directly implies vector space structure of L^p.

3.2 Day 2: Expectation and L^p Spaces

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Covariance $\mathrm{Cov}(X,Y) := \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Variance $\mathrm{Var}(X) = \mathbb{E}[(X-\mu_X)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathrm{Cov}(X,X)$

Inner Product $\langle X, Y\rangle = \mathbb{E}[XY]$.

Correlation $\rho_{X,Y} = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \dfrac{\langle X-\mu_X, Y-\mu_Y\rangle}{\|X-\mu_X\|_2\|Y-\mu_Y\|_2}$, $|\rho_{X,Y}| \le 1$ (Cauchy-Schwarz)

Note Interpret $\rho = \cos\theta$.

Orthogonality $U \perp V$ if $\langle U, V\rangle = \mathbb{E}[UV] = 0$.

Note If $\mu_U = \mu_V = 0$, orthogonality = uncorrelated

Cauchy Sequence $\{x_n\}_{n=1}^{\infty}$ Cauchy if $\forall\varepsilon > 0, \exists N$, s.t. $n,m \ge N \Rightarrow d(x_n, x_m) < \varepsilon$.

Note Alternatively, $\{x_n\}$ Cauchy if $\lim\limits_{N\to\infty} \sup\limits_{n,m\ge N} d(x_n,x_m) = 0$. $d(\cdot,\cdot) = \|\cdot\|_p$ for L^p spaces.

MAIN THEOREMS:

Pythagoras $U \perp V \implies \|U+V\|^2 = \|U\|^2 + \|V\|^2$ (L^2-norm).

Parallelogram Law $\|U+V\|^2 + \|U-V\|^2 = 2\|U\|^2 + 2\|V\|^2$.

Variance of Sum $X_1,\dots,X_n \in L^2$: [5] $\mathrm{Var}(X_1+\cdots+X_n) = \sum\limits_n \mathrm{Var}(X_n) + \sum\limits_{i\ne j}\mathrm{Cov}\big(X_i, X_j\big)$

Completeness of $\mathbf{L^p}$ X_n Cauchy in $L^p \implies \exists X \in L^p$ S.T. $\|X_n - X\|_p \to 0$ (i.e., $X_n \to X \in L^p$)

Proof: *(click)*

Pass to subseq. s.t. $\|X_{k_{n+1}} - X_{k_n}\|_p < 2^{-n}$, use monotonicity of norms: $\sum \mathbb{E}[|X_{k_{n+1}} - X_{k_n}|] < \infty$ $\Rightarrow \mathbb{E}\left[\sum |X_{k_{n+1}} - X_{k_n}|\right] < \infty \Rightarrow \sum |X_{k_{n+1}} - X_{k_n}| < \infty$ a.s. $\Rightarrow X := \lim_n X_{k_n}(\omega)$ exists w.p. 1. X measurable, deduce (FATOU + V-Space) $X \in L^p$ & $X_n \to X$ in L^p. $\square$

[5] To make sure $\mathrm{Var}(X) < \infty$.

Orthogonal Proj. $\mathcal{K} \subset L^2$ complete, V-Subspace.

$$\forall X \in L^2, \exists Y \in \mathcal{K} \text{ S.T.} \begin{cases} (X-Y) \perp Z, \forall Z \in \mathcal{K} \\ \|X-Y\|_2 = \inf_{Z\in\mathcal{K}} \|X-Z\|_2 \end{cases}$$

Note If $\exists \tilde{Y} \in \mathcal{K}$ s.t. one of above holds, then $\tilde{Y} = Y$, a.s.

Note Y is called a version of the orthogonal projection of X onto $\mathcal{K}$.

Note Application: project $L^2(\Omega, \mathscr{F}, \mathbb{P})$ onto $\mathcal{K} = L^2(\Omega, \mathscr{G}, \mathbb{P})$ ($\mathscr{G} \subset \mathscr{F}$ σ-algebra)

Expec. of Func. of r.v. $h : (\mathbb{R}, \mathscr{B}) \to (\mathbb{R}, \mathscr{B})$, then $\mathbb{E}[h(X)] = \int_\Omega h(X(\omega))\, d\mathbb{P}(\omega) = \int_\mathbb{R} h(x)\, d\mathbb{P}_X(x)$.

Also, $h(X) \in L^1(\Omega, \mathscr{F}, \mathbb{P}) \iff h \in L^1(\mathbb{R}, \mathscr{B}, \mathbb{P}_X)$.

Proof: *(click)*

Verify for $h = \mathbb{1}_B$ ($B \in \mathscr{B}$), extend to simple fnc., generalize via MCT (standard machinery). $\square$

PDF and Cont. r.v. X cont. r.v. if $\exists f_X : \mathbb{R} \to [0, \infty]$ Borel s.t. $F(x) = \int_{(-\infty,x]} f_X(x)\, d\lambda(x)$.

Therefore $\mathbb{P}_X(B) = \int_B f_X(x)\, d\lambda(x)$ and $f_X = \dfrac{d\mathbb{P}_X}{d\lambda}$, Radon-Nikodym.

Proof: *(click)*

$F(x) = \mathbb{P}((-\infty, x]) = \int_{(-\infty,x]} f_X(x)\, d\lambda(x)$. Define $\mu(\cdot)$ on $(\mathbb{R}, \mathscr{B})$ $\mu(B) = \int_B f_X(x) d\lambda(x)$.
$\mu(\cdot)$ valid measure, agrees with $\mathbb{P}_X$ on $\pi(\mathbb{R}) = \{(-\infty, x] : x \in \mathbb{R}\}$: uniqueness thm $\implies \mu = \mathbb{P}_X$. $\square$

Exp. for Cont. r.v. $h : (\mathbb{R}, \mathscr{B}) \to (\mathbb{R}, \mathscr{B})$, then $\mathbb{E}[h(X)] = \int_\mathbb{R} h(x) f_X(x)\, dx$

Also, $\mathbb{E}[|h(X)|] < \infty \iff \int_\mathbb{R} |h(x)| f_X(x)\, dx < \infty$.

3.3 Day 3: Strong Laws

Main Reference(s):	– David Williams, *Probability With Martingales*, 1991 [13] Chap.7
	– Rick Durrett, *Probability: Theory & Examples*, 2010 [5] Sec. 2.4

MAIN THEOREMS:

Independence and Expectation $(X \perp\!\!\!\perp Y) \in L^1 \implies \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \Rightarrow XY \in L^1$.

$(X \perp\!\!\!\perp Y) \in L^2 \implies \mathrm{Cov}\,(X,Y) = 0 \;\&\; \mathrm{Var}\,(X+Y) = \mathrm{Var}\,(X) + \mathrm{Var}\,(Y)$

Proof: *(click)*

If $X = X^+ - X^-$, $Y = Y^+ - Y^-$, then holds if proven for $X, Y \geq 0$.

So, WLOG $X, Y \geq 0$. $0 \leq X_n \nearrow X$, $0 \leq Y_n \nearrow Y$, simple r.v.'s. Prove $\mathbb{E}[X_nY_n] = \mathbb{E}[X_n]\mathbb{E}[Y_n]$ using $\sigma(X) \perp\!\!\!\perp \sigma(Y)$. Conclude, by MCT as $0 \leq X_nY_n \nearrow XY \Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. $\square$

Careful! If X and Y are not $\perp\!\!\!\perp$, $(X,Y) \in L^1 \not\Rightarrow XY \in L^1$

Example: $X : (\Omega, \mathscr{F}) \to (\mathbb{N}, 2^{\mathbb{N}})$: $\mathbb{P}\,(X = k) = \dfrac{c}{k^{2+\varepsilon}}$ (with $\varepsilon \in (0,1]$) and $Y = X$.
Then $\mathbb{E}[X] = \mathbb{E}[Y] = \sum_k \dfrac{c}{k^{1+\varepsilon}} < \infty$ and $\mathbb{E}[XY] = \mathbb{E}[X^2] = \sum_k \dfrac{c}{k^{\varepsilon}} = \infty$.

Proving $\mathbf{X_n \xrightarrow{a.s.} 0}$ Let $\{X_n\}_{n=1}^{\infty}$ r.v.

(a) $\exists s > 0$ s.t. $\sum_n \mathbb{E}[|X_n|^s] < \infty \implies X_n \xrightarrow{\text{a.s.}} 0$.

(b) $\forall \varepsilon > 0$ s.t. $\sum_n \mathbb{P}\left(|X_n| > \varepsilon\right) < \infty \implies X_n \xrightarrow{\text{a.s.}} 0$.

(c) $X_n \xrightarrow{\text{a.s.}} 0 \iff \lim_{n\to\infty} \mathbb{P}\left(\sup_{m \geq n} |X_m| > \varepsilon\right) = 0, \forall \varepsilon > 0$

Proof: *(click)*

(a) By MCT, $\sum_n \mathbb{E}[|X_n|^s] = \mathbb{E}\left[\sum_n |X_n|^s\right] < \infty \Rightarrow \sum_n |X_n|^s < \infty$, a.s. $\Rightarrow |X_n|^s \xrightarrow{\text{a.s.}} 0 \Rightarrow X_n \xrightarrow{\text{a.s.}} 0$.

(b) Fix $\varepsilon = \frac{1}{k}$. By BC1, $\{|X_n| > \frac{1}{k}\}$ occurs finitely often w.p. $1 \Rightarrow \mathbb{P}\left(\limsup_{n\to\infty} X_n > \dfrac{1}{k}\right) = 0$

$\Rightarrow \mathbb{P}\left(\bigcup_k \left\{\limsup_{n\to\infty} X_n > \dfrac{1}{k}\right\}\right) = 0$ (union of null) $\implies \limsup_{n\to\infty} X_n \leq 0$, a.s. $\Rightarrow X_n \xrightarrow{\text{a.s.}} 0$.

(c) Notice RHS equiv. to $\mathbb{P}\left(|X_m| \leq \varepsilon, \forall m \geq n\right) \to 1$ as $n \to \infty$: $A_n := \{\omega \in \Omega : |X_m(\omega)| \leq \varepsilon, \forall m \geq n\}$.
$A_n \nearrow \bigcup_n A_n = \Omega$ (since $|X_n| \xrightarrow{\text{a.s.}} 0$), hence $\mathbb{P}(A_n) \to 1$. $\square$

Est. for the Max. of Avg. (EMA) $X_n \sim X$ i.i.d. $X \in L^1$: $\mathbb{P}\left(\sup_{n\geq 1} \frac{|X_1+\cdots+X_n|}{n} > a\right) \leq \frac{\mathbb{E}[|X|]}{a}$.

($\because$)See Lecture 18 LLN from Yury 6.436.

Kolmogorov's Maximal Ineq. $\{X_i \perp\!\!\!\perp\} \in L^2$, $\mathbb{E}[X_i]=0$, $\mathrm{Var}(X_i)<\infty$, $S_n = X_1 + \ldots X_n$.

Then: $\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq x\right) \leq \frac{\mathrm{Var}(S_n)}{x^2}$, $\forall x \geq 0$

Note Chebyshev: $\mathbb{P}\left(|S_n| \geq x\right) \leq \frac{\mathrm{Var}(S_n)}{x^2}$, $\forall x \geq 0$

Note Recall Kolmogorov 0-1 law: this ineq. helps proving $\mathbb{P}(A)=1$.

Proof: *(click)*

Let $A_k := \{\omega \in \Omega : |S_k(\omega)| \geq x, |S_j(\omega)| < x,\ \forall j < k\}$.

Notice $\{\omega \in \Omega : \max_{1\leq k \leq n} |S_k(\omega)| \geq x\} = \bigcup_k A_k$ & $A_i \cap A_j = \varnothing$, $i \neq j$: then

$$\mathbb{1}_\Omega \geq \mathbb{1}_{\cup_k A_k} = \sum_k \mathbb{1}_{A_k} \implies \mathbb{E}\left[S_n^2\right] \geq \sum_k \mathbb{E}\left[\mathbb{1}_{A_k} S_n^2\right] = \sum_k \mathbb{E}\left[\mathbb{1}_{A_k}(S_k + S_n - S_k)^2\right]$$

$$\Rightarrow \mathbb{E}\left[S_n^2\right] \geq \sum_k \mathbb{E}\left[\mathbb{1}_{A_k}S_k^2 + \mathbb{1}_{A_k}2S_k(S_n-S_k) + \mathbb{1}_{A_k}(S_n-S_k)^2\right] \geq \sum_k \left(\mathbb{E}\left[\mathbb{1}_{A_k}S_k^2\right] + \mathbb{E}\left[\mathbb{1}_{A_k}2S_k(S_n-S_k)\right]\right)$$

As $S_k\mathbb{1}_{A_k} \in \sigma(X_1,\ldots,X_k)$ & $(S_n - S_k) \in \sigma(X_{k+1},\ldots,X_n)$, so $S_k\mathbb{1}_{A_k} \perp\!\!\!\perp (S_n-S_k) \Rightarrow \mathbb{E}\left[\mathbb{1}_{A_k}S_k(S_n-S_k)\right]=0$.

Hence $\mathbb{E}[S_n^2] \geq \sum_{k=1}^n \mathbb{E}[\mathbb{1}_{A_k}S_k^2] \geq \sum_{k=1}^n \mathbb{E}[x^2\mathbb{1}_{A_k}] = x^2\sum_{k=1}^n \mathbb{P}(A_k) = x^2\mathbb{P}\left(\max_{1\leq k\leq n}|S_k| \geq x\right)$. $\square$

Strong Laws of Large Number

(1) Kolmogorov (Not ID but L^2) $\{X_n\} \in L^2 \perp\!\!\!\perp$, mean $\{\mu_n\}$, variance $\{\sigma_n^2\}$ s.t. $\sum_n \frac{\sigma_n^2}{n^2} < \infty$

(2) 6.436 (IID but L^2) $\{X_n\} \in L^2$ i.i.d., mean μ, variance σ^2

(3) Khinchine (IID but L^1) $\{X_n\} \in L^1$ i.i.d., mean μ

(4) Durrett (IID but $\mathbb{E}$ exists) $\{X_n\} \in L^1$ IID with $\mathbb{E}[X_n]$ exists[6] (i.e. $\mathbb{E}[X^+]$ or $\mathbb{E}[X^+]<\infty$)

$\implies \frac{1}{n}\sum_{k=1}^n X_k \xrightarrow{\text{a.s.}} \mu$. If not I.D. $\frac{1}{n}\sum_{k=1}^n (X_k - \mu_k) \xrightarrow{\text{a.s.}} 0$.

Note SLLN needs $\mathbb{E}[|X|]<\infty$ (cf Week 2 Day 2): if $\mathbb{E}[|X|]=\infty$, then $\limsup \frac{S_n}{n} = \infty$ a.s.

Proof: *(click)*

Assume $X_i \geq 0$, generalize via $X_i = X_i^+ - X_i^-$.

(4) Durrett: WLOG $\mathbb{E}[X_i^+]=\infty, \mathbb{E}[X_i^-]<\infty$ ($\mathbb{E}[X_i]$ exists).

Let $X_i^M := X_i \wedge M$ (truncate at fixed M) & $S_n^M = \sum_{i=1}^n X_i^M \xrightarrow[\text{in } L^2]{\text{SLLN}} \lim_n \frac{S_n^M}{n} \to \mathbb{E}[X_i^M]$.

Now, $X_i \geq X_i^M \Rightarrow \liminf_n \frac{S_n}{n} \geq \lim_n \frac{S_n^M}{n} = \mathbb{E}[X_i^M] \xrightarrow{\text{MCT}} \mathbb{E}[X_i^M] \nearrow \mathbb{E}[X_i] = \infty \Rightarrow \liminf_n \frac{S_n}{n} = \infty$ a.s. $\square$

[6]Note that the mean can be infinite!

Proof: *(click)*

(1) Kolmogorov: $S_n = \sum_{k=1}^{n}(X_k - \mu_k)$ $A_k := \left\{\omega \in \Omega : \max_{2^{k-1} \le n < 2^k} \frac{|S_n|}{n} > \varepsilon\right\}$.

$$\left\{\max_{2^{k-1}\le n<2^k} \frac{|S_n|}{n} \ge \varepsilon\right\} = \left\{\exists n \in \left[2^{k-1}, 2^k\right] \text{ s.t.} |S_n| \ge n\varepsilon\right\} \subset \left\{\max_{2^{k-1}\le n<2^k} |S_n| \ge 2^{k-1}\varepsilon\right\} \subset \left\{\max_{1\le n<2^k} |S_n| \ge 2^{k-1}\varepsilon\right\}$$

Apply Kolmogorov: $\mathbb{P}(A_k) \le \mathbb{P}\left(\max_{1\le n<2^k} |S_n| \ge 2^{k-1}\varepsilon\right) \le \frac{1}{\left(2^{k-1}\varepsilon\right)^2}\operatorname{Var}(S_{2^k}) = \frac{4}{\varepsilon^2}\frac{1}{2^{2k}}\sum_{i=1}^{2^k}\sigma_i^2$

Now $\sum_{k=1}^{\infty}\mathbb{P}(A_k) \le \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty}\sum_{i=1}^{2^k}\frac{1}{2^k}\sigma_i^2 = \frac{4}{\varepsilon^2}\sum_{i=1}^{\infty}\sigma_i^2\sum_{k=\log_2 i}^{\infty}\frac{1}{2^{2k}} \le \frac{8}{\varepsilon^2}\sum_{i=1}^{\infty}\frac{\sigma_i^2}{i^2} < \infty \xRightarrow{\text{BC 1}} \mathbb{P}(A_k \text{ i.o. }) = 0.$

Since $\{A_k \text{ i.o. }\} = \left\{\frac{|S_n|}{n} \ge \varepsilon \text{ i.o.}\right\}$, so $\mathbb{P}\left(\limsup_n \frac{|S_n|}{n} \le \varepsilon\right) = 1$ Take $\varepsilon = \frac{1}{k} \to 0$: $\frac{S_n}{n} \xrightarrow{\text{a.s.}} 0$.

(2) 6.436: Notice $\mathbb{E}\left[\left(\frac{S_n}{n} - \mu\right)^2\right] = \operatorname{Var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$

Claim: $\frac{S_{k^2}}{k^2} \xrightarrow{\text{a.s.}} \mu$

$$(\because)\sum_{k=1}^{\infty}\mathbb{E}\left[\left(\frac{S_{k^2}}{k^2} - \mu\right)^2\right] = \sum_{k=1}^{\infty}\frac{\sigma^2}{k^2} < \infty \xRightarrow{\text{Lemma (a)}} \left(\frac{S_{k^2}}{k^2} - \mu\right)^2 \xrightarrow{\text{a.s.}} 0$$

OR $(\because)$Fix $\varepsilon > 0$: $\sum_{k=1}^{\infty}\mathbb{P}\left(\left|\frac{S_{k^2}}{k^2} - \mu\right| > \varepsilon\right) \le \sum_{k=1}^{\infty}\frac{1}{k^2}\frac{\sigma^2}{\varepsilon^2} < \infty \xRightarrow{\text{Lemma (b)}} \frac{S_{k^2}}{k^2} \xrightarrow{\text{a.s.}} \mu$

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

$(\because)$Take $k^2 \le n \le (k+1)^2$: $\underbrace{\frac{k^2}{(k+1)^2}}_{\to 1}\underbrace{\frac{S_{k^2}}{k^2}}_{\xrightarrow{\text{a.s.}}\mu} \le \frac{S_n}{n} \le \underbrace{\frac{(k+1)^2}{k^2}}_{\to 1}\underbrace{\frac{S_{(k+1)^2}}{(k+1)^2}}_{\xrightarrow{\text{a.s.}}\mu} \implies \frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu$

Note (Durrett Thm. 2.3.8) Suppose $0 \le X_n \nearrow$ r.v., $0 \le c_n \nearrow$. **Goal:** show $\frac{X_n}{c_n} \xrightarrow{\text{a.s.}} c$.

Idea: pick subseq. $\{c_{n_k}\}_{k=1}^{\infty}$ s.t. $\frac{c_{n_{k+1}}}{c_{n_k}} \to 1$. Show that $\frac{X_{n_k}}{c_{n_k}} \xrightarrow{\text{a.s.}} c$, and conclude.

(3) Khinchine: Idea: Truncate, so $\in L^2$, bound difference via estimate for max. of avg.

Let $Y_n = X_n \mathbb{1}_{X_n \le k}$, $Z_n = X_n \mathbb{1}_{X_n > k}$, $T_n = \frac{Y_1 + \cdots + Y_n}{n}$, and $Z^* = \sup_{n\ge 1}\frac{|Z_1 + \cdots + Z_n|}{n}$.

Claim: $\mathbb{P}\left(\sup_{m\ge n}\frac{|S_m|}{m} > \varepsilon\right) \to 0$.

$(\because)$Triangle inequality $\Rightarrow \frac{|S_m|}{m} \le \frac{|T_m|}{m} + Z^*$

$$\implies \mathbb{P}\left(\sup_{m\ge n}\frac{|S_m|}{m} > \varepsilon\right) \le \mathbb{P}\left(\sup_{m\ge n}\frac{|T_m|}{m} + Z^* > \varepsilon\right) \le \underbrace{\mathbb{P}\left(\sup_{m\ge n}\frac{|T_m|}{m} > \frac{\varepsilon}{2}\right)}_{:=A} + \underbrace{\mathbb{P}\left(Z^* > \frac{\varepsilon}{2}\right)}_{:=B}.$$

Notice $|Y_n| \le |X_n| \in L^1 \xRightarrow{\text{DCT}} |\mathbb{E}[Y_n]| \to \mathbb{E}[X_n] = 0$ as $k \to \infty$ & $Z_n \to 0$ a.s. $\Rightarrow \mathbb{E}[|Z_n|] \to_{k\to\infty} 0$.

Choose k large : $|\mathbb{E}[Y_n]| < \frac{\varepsilon}{4}, \mathbb{E}[|Z_n|] < \delta\frac{\varepsilon}{2}$. By EMA, $B \le \frac{\mathbb{E}[|Z|]}{\varepsilon/2} < \delta$.

Next, $|T_m| \le |T_m - \mathbb{E}[T_m]| + |\mathbb{E}[T_m]|$: $\sup_{m\ge n}\frac{|T_m|}{m} > \frac{\varepsilon}{2} \Rightarrow \sup_{m\ge n}\frac{|T_m - \mathbb{E}[T_m]|}{m} > \frac{\varepsilon}{4}$.

Thus $A \le \mathbb{P}\left(\sup_{m\ge n}\frac{|T_m - \mathbb{E}[T_m]|}{m} > \frac{\varepsilon}{4}\right) \to_{n\to\infty} 0$, as SLLN holds for L^2.

Hence, $\limsup_{n\to\infty}\mathbb{P}\left(\sup_{m\ge n}\frac{|S_m|}{m} > \varepsilon\right) \le \delta \implies \mathbb{P}\left(\sup_{m\ge n}\frac{|S_m|}{m}\right) \to 0$, as $n \to \infty$.

$\square$

EXERCISES:

Exercise:

(Durrett 2.5.1) Let $X_1, X_2, \ldots$ iid, $\mathbb{E}[X] = 0$, $\operatorname{Var}(X) = C < \infty$, and $S_n = X_1 + \ldots + X_n$.

Use Kolmogorov's Maximal's Inequality to show that $\frac{S_n}{n^p} \xrightarrow{\text{a.s.}} 0$ for $p > 1/2$.

Solution: *(click)*

Fix a subsequence $n = m^\alpha$ where $\alpha(2p-1) > 1$.

Let $A_m = \left\{\omega \in \Omega : \max_{(m-1)^\alpha \le k \le m^\alpha} |S_k| \ge (m-1)^{\alpha p}\varepsilon\right\}$.

By KMI: $\mathbb{P}(A_m) \le \frac{1}{\varepsilon^2 (m-1)^{2\alpha p}} \operatorname{Var}(S_{m^\alpha}) \le \frac{Cm^\alpha}{\varepsilon^2(m-1)^{2\alpha p}} \le \frac{Cm^\alpha}{\varepsilon^2 (m/2)^{2\alpha p}} \le \frac{1}{m^{\alpha(2p-1)}} \cdot \frac{C \cdot 2^{2\alpha p}}{\varepsilon^2}$

So $\sum_{m=1}^{\infty} \mathbb{P}(A_m) \le \frac{C \cdot 2^{2\alpha p}}{\varepsilon^2} \sum_{m=1}^{\infty} \frac{1}{m^{\alpha(2p-1)}} < \infty \xRightarrow{\text{BC-1}} \mathbb{P}\left(\{A_m \text{ i.o. }\}\right) = 0$

$\Longrightarrow \max_{(m-1)^\alpha \le k \le m^\alpha} |S_k| \le \varepsilon (m-1)^{\alpha p}$ for a.a. ω for sufficiently large m.

In particular, $\forall k \in [(m-1)^\alpha, m^\alpha]$: $\frac{|S_k|}{k^p} \le \frac{|S_k|}{(m-1)^{\alpha p}} \le \frac{\varepsilon \cancel{(m-1)^{\alpha p}}}{\cancel{(m-1)^{\alpha p}}} = \varepsilon$ a.e. .

Hence, $\limsup_{k \to \infty} \frac{|S_k|}{k^p} \le \varepsilon$ almost surely. Taking $\varepsilon_k = \frac{1}{k} \searrow 0$, countable, we conclude.

Alternative Solution Notice, it suffices to prove (due to Kronecker's lemma) that $\sum_{n=1}^{\infty} \frac{X_n}{n^p} < \infty$ a.s. , which holds (Kolmogorov 2-series) if $\sum_{n=1}^{\infty} \operatorname{Var}\left(\frac{X_n}{n^p}\right) = \sum_{n=1}^{\infty} \frac{C}{n^{2p}} < \infty$, which is true since $p > \frac{1}{2}$.

3.4 Day 4: Application of SLLN

Main Reference(s):	– David Williams, *Probability With Martingales*, 1991 [13] Chap.7
	– Rick Durrett, *Probability: Theory&Examples*, 2010 [5] Sec. 2.4/5

KEYWORDS:

Empirical CDF (E-CDF) $\{X_i \text{ iid}\} \sim F$, then: $F_n(x) := \frac{1}{n}\sum_{m=1}^{n} \mathbb{1}\{X_m \le x\}$

Note $F_n(x)$ counts the frequency of the observed values that are $\le x$.

MAIN THEOREMS:

Glivenko-Cantelli / CDF Approx $F_n =$ E-CDF: $\lim_{n\to\infty} \sup_x |F_n(x) - F(x)| = 0$ a.s.

Proof: *(click)*

$F_n(x-) := \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbb{1}_{X_k < x} \xrightarrow{\text{a.s.}} F(x-) = \mathbb{P}(X < x)$. Take k, $x_{j,k} := \inf\left\{y : F(y) \ge \dfrac{j}{k}\right\}$.

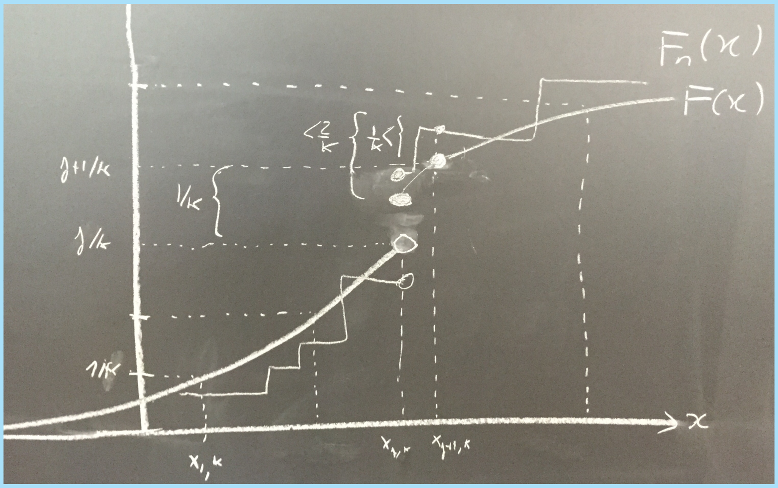
$\exists N(\omega)$ S.T. $|F_n(x_{j,k}-) - F(x_{j,k}-)| < \frac{1}{k}$, & $|F_n(x_{j,k}) - F(x_{j,k})| < \frac{1}{k}$, $1 \le j \le k-1$, $n \ge N(\omega)$.

Claim: $F(x_{j,k}-) = \lim_{x \nearrow x_{j,k}} F(x) \le \frac{j}{k}$, as $x < x_{j,k} \Rightarrow F(x) \le \frac{j}{k}$.

Claim: $F(x_{j,k}) \ge \frac{j}{k}$, as $x_{j,k} = \inf\{\dots\}$, $\exists y_n \in \{\dots\}$ S.T. $y_n \searrow x_{j,k} \Rightarrow F(x_{j,k}) = \lim_{y_n \searrow x_{j,k}} F(y_n) \ge \frac{j}{k}$ [a]

Now $x \in (x_{j-1,k}, x_{j,k}) \Rightarrow F_n(x) \le F_n(x_{j,k}-) \le F(x_{j,k}-) + \frac{1}{k} \le F(x_{j-1,k}) + \frac{2}{k} \le F(x) + \frac{2}{k}$.

Similar argument $\Rightarrow F_n(x) - F(x) \ge -\frac{2}{k} \Rightarrow |F_n(x) - F(x)| \le \frac{2}{k}$, $\forall x$, $\forall n \ge N(\omega)$.



$\square$

[a] Argument uses right continuity of CDF.

Renewal Process $\{X_i \text{ iid}\}$ with $X_i \in (0,\infty)$ a.s. , and $T_n := X_1 + \ldots + X_n$ (n^{th} arrival time).

Let $N_t := \sup\{n : T_n \le t\}$ (# arrivals before t) $\implies \dfrac{N_t}{t} \xrightarrow{\text{a.s.}} \dfrac{1}{\mathbb{E}[X]}$ (works for $\dfrac{1}{\infty} = 0$).

Proof: *(click)*

Observe that $T_{N_t} \le t < T_{N_t+1}$ and that $t < T_{N_t+1} \implies T_{N_t+1} \to \infty \implies N_t \to \infty$ (can't be bdd)[a].

$$T_{N_t} \le t < T_{N_t+1} \implies \frac{T_{N_t}}{N_t} \le \underbrace{\frac{t}{N_t}}_{\xrightarrow{\text{a.s.}}\mathbb{E}[X]} \le \underbrace{\frac{T_{N_t+1}}{N_t+1}}_{\xrightarrow{\text{a.s.}}\mathbb{E}[X]}\underbrace{\frac{N_t+1}{N_t}}_{\xrightarrow{\text{a.s.}}1} \implies \frac{t}{N_t} \xrightarrow{\text{a.s.}} \mathbb{E}[X].$$

$(\because)$Since $N_t \nearrow \infty$ a.s. $\overset{\text{SLLN}}{\Longrightarrow} \dfrac{T_n(\omega)}{n} \xrightarrow{\text{a.s.}} \mathbb{E}[X]$ so $\dfrac{T_{N_t(\omega)}(\omega)}{N_t(\omega)} \xrightarrow{\text{a.s.}} \mathbb{E}[X]$ and $\dfrac{N_t(\omega)+1}{N_t(\omega)} \xrightarrow{\text{a.s.}} 1$. $\square$

[a] To be more precise, $N_t \nearrow \infty$ a.s. , as $t \to \infty$.

Chapter 4

Week 4: Convergence of Random Series & Large Deviation

4.1 Day 1: Convergence of Random Series

Main Reference(s):	– David Williams, *Probability With Martingales*, 1991 [13] Chap.12
	– Rick Durrett, *Probability: Theory&Examples*, 2010 [5] Sec. 2.4/5

KEYWORDS:

Convergence of Series $\sum_{n=1}^{\infty} a_n$ converges iff $\lim_{N\to\infty} \sum_{n=1}^{N} a_n$ exists[1].

MAIN THEOREMS:

SLLN (Etemadi Ideas) (IID, L^1)$\{X_k\} \perp\!\!\!\perp$, $\mathbb{E}\left[|X_k|\right] < \infty$, $\mathbb{E}X_k = \mu \Rightarrow \frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu$.

Lemma 4.1.1.

(a) *Let* $Y_k := X_k \mathbb{1}\{|X_k| \le k\}^2$ *, &* $T_n = \sum_{k=1}^{n} Y_k$. *Then* $\frac{T_n}{n} \xrightarrow{a.s.} \mu \implies \frac{S_n}{n} \xrightarrow{a.s.} \mu$.

(b) $\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} \le 4\mathbb{E}\left[|X|\right] < \infty$.

Proof: *(click)*

(a) Claim: $\mathbb{P}(X_k \ne Y_k, \text{ i.o. }) = 0$. This shows $\mathbb{P}\left(\frac{|S_n - T_n|}{n} \to 0\right) = 1$.

$(\because)\sum_{k=1}^{\infty} \mathbb{P}(X_k \ne Y_k) \le \sum_{k=1}^{\infty} \mathbb{P}\left(|X_k| > k\right) = \sum_{k=1}^{\infty} \mathbb{P}\left(|X_1| > k\right) \le \int_0^{\infty} \mathbb{P}\left(|X_1| > x\right)\, dx = \mathbb{E}\left[|X_1|\right] < \infty$, BC-1.

(b) $\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} \le \sum_{k=1}^{\infty} \frac{\mathbb{E}\left[Y_k^2\right]}{k^2} = \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^{\infty} 2y\mathbb{P}\left(|Y_k| > y\right)\, dy = \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^{k} 2y\mathbb{P}\left(|Y_k| > y\right)\, dy$

So, $\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} \le \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^{k} 2y\mathbb{P}\left(|X| > y\right)\, dy \le \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^{\infty} \mathbb{1}_{\{y<k\}} 2y\mathbb{P}\left(|X| > y\right)\, dy$

So, $\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} \le \int_0^{\infty} \sum_{k=1}^{\infty} \mathbb{1}_{\{y<k\}} \frac{1}{k^2} 2y\mathbb{P}\left(|X| > y\right)\, dy = \int_0^{\infty} \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \mathbb{1}_{\{y<k\}}\right) 2y\mathbb{P}\left(|X| > y\right)\, dy$.

Claim: $\sum_{k=1}^{\infty} \frac{1}{k^2} \mathbb{1}_{\{y<k\}} \le \frac{2}{y}$. This shows that $\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} \le \int_0^{\infty} \left(\frac{2}{y}\right) 2y\mathbb{P}\left(|X| > y\right)\, dy = 4\mathbb{E}\left[|X_1|\right]$

$(\because)\sum_{k\ge y} \frac{1}{k^2} \le^{a} \int_{y-1}^{\infty} \frac{1}{x^2}\, dx = \frac{1}{y-1} \le \frac{2}{y}$. □

[a] Here, we assume $y \in \mathbb{N}$, and $y \ge 2$. For slightly more rigor, use floor functions (see Durrett p.74 [5])

[1] If $a_n \ge 0$, then either $\sum_{n=1}^{\infty} a_n < \infty$ or $= \infty$, but it necessarily exists.

[2] Different from $Y_k = X_k \mathbb{1}\{|X_k| \le n\}$ for fixed n (see SLLN proof).

Kolmogorov 2-Series ($\perp\!\!\!\perp$) $\{X_n \perp\!\!\!\perp\}$, $\sum_{n=1}^{\infty} \mathbb{E}X_n < \infty$ & $\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty \implies \sum_{n=1}^{\infty} X_n < \infty$ a.s.

Proof: *(click)*

WLOG, prove for $\tilde{X}_n \perp\!\!\!\perp \mathbb{E}[Y_n] = 0$ ($\tilde{X}_n = X_n - \mu_n$, hence $\sum_n \tilde{X}_n < \infty \iff \sum_n X_n < \infty$).

To prove convergence, need to prove $\tilde{S}_n := \sum_{k=1}^{n} \tilde{X}_k$ is Cauchy, via Kolmogorov's maximal ineq (KMI).

KMI $\Rightarrow \mathbb{P}\left(\max_{\{M \le m \le N\}} |S_m - S_M| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\mathrm{Var}(S_N - S_M) = \frac{1}{\varepsilon^2}\sum_{n=M+1}^{N} \sigma_n^2 \xrightarrow[N\to\infty]{} \frac{1}{\varepsilon^2}\sum_{n=M+1}^{\infty} \sigma_n^2$.

$\mathbb{P}\left(\sup_{m \ge M} |S_m - S_M| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\sum_{n=M+1}^{\infty} \sigma_n^2 \xrightarrow[M\to\infty]{} 0$.

$\underbrace{\sup_{n,m\ge M} |S_n - S_m|}_{:=\Delta_M} \le \underbrace{\sup_{n\ge M} |S_n - S_M|}_{:=A_M} + \underbrace{\sup_{m\ge M} |S_m - S_M|}_{:=B_M}$. Claim: $\Delta := \lim_{M\to\infty} \Delta_M = 0$ a.s.

$\Delta_M \ge \Delta_{M+1} \ge \cdots \ge 0 \Rightarrow \lim_M \Delta_M$ exists a.s.

Trick: $\lim_M \mathbb{P}(\Delta_M > \varepsilon) \le \lim_M \mathbb{P}\left(\{A_M > \varepsilon/2\} \cup \{B_M > \varepsilon/2\}\right) \le \lim_M \mathbb{P}\left(A_M > \varepsilon/2\right) + \lim_M \mathbb{P}\left(B_M > \varepsilon/2\right) = 0$

Now, $0 = \lim_M \mathbb{P}(\Delta_M > \varepsilon) = \mathbb{P}\left(\bigcap_M \{\Delta_M > \varepsilon\}\right) \ge \mathbb{P}(\Delta > \varepsilon)$ so $\mathbb{P}(\Delta > \varepsilon) = 0$ $\quad (\because) \cap_M\{\Delta_M > \varepsilon\} \supset \{\Delta > \varepsilon\}$

So $\lim_{k\to\infty} \mathbb{P}\left(\Delta > \frac{1}{k}\right) \overset{\varepsilon=1/k}{=\!=} \mathbb{P}\left(\bigcup_k \left\{\Delta > \frac{1}{k}\right\}\right) = \mathbb{P}(\Delta > 0) = 0 \implies \Delta = 0$ a.s. $\square$

Trick: $C \le A + B$ r.v.s: $\mathbb{P}(C > \varepsilon) \le \mathbb{P}\left(\{A > \varepsilon/2\} \cup \{B > \varepsilon/2\}\right) \le \mathbb{P}\left(A > \varepsilon/2\right) + \mathbb{P}\left(B > \varepsilon/2\right)$.

Kolmogorov's 3-Series ($\perp$) $\{X_n \perp\!\!\!\perp\}$, $A > 0$ & $Y_i = X_i \mathbb{1}_{|X_i| \le A}$: Then $\sum_{n=1}^{\infty} X_n$ converges a.s. $\Leftrightarrow$

- **(i)** $\sum_{n=1}^{\infty} \mathbb{P}\left(|X_n| > A\right) < \infty$
- **(ii)** $\sum_{n=1}^{\infty} \mathbb{E}[Y_n] < \infty$
- **(iii)** $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$.

Proof: *(click)*

($\Rightarrow$) CLT (Durrett [5], Example 3.4.7)

($\Leftarrow$) $\sum_{n=1}^{\infty} \mathbb{P}(X_n \ne Y_n) \le \sum_{n=1}^{\infty} \mathbb{P}\left(|X_n| > A\right) < \infty \overset{\text{BC-1}}{\Longrightarrow} \mathbb{P}(X_n \ne Y_n,\ \text{i.o.}) = 0$

This shows that $\sum_n X_n < \infty \iff \sum_n Y_n < \infty$. End: apply Kolmogorov's 2-series on $\{Y_n\}$. $\square$

Kronecker's Lemma $a_n \nearrow \infty$ & $\sum_{n=1}^{\infty} \frac{x_n}{a_n} < \infty \implies \frac{1}{a_n}\sum_{k=1}^{n} x_k \to 0$.

($\because$)See Durrett [5] p.81; Williams [13] p.117.

SLLN Alternative Proof (iid, L^1) $\{X_i iid\}$, $\mathbb{E}\left[|X_i|\right] < \infty$, then $\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu$.

Proof: *(click)*

As before, $Y_n := X_n \mathbb{1}_{|X_n| \le n}$, $T_n = \frac{Y_1 + \cdots + Y_n}{n}$.

Claim: $\left(\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu\right) \iff \left(\frac{T_n}{n} \xrightarrow{\text{a.s.}} \mu\right)$. $(\because)$see Etamadi's idea (a)

Claim: $Z_n := Y_n - \mathbb{E}\left[Y_n\right]$: Then $\frac{1}{n}\sum_{k=1}^{n} Z_k \xrightarrow{\text{a.s.}} 0$.

$(\because)\sum_{n=1}^{\infty} \text{Var}\left(\frac{Y_n}{n}\right) < \infty \Rightarrow \sum_{n=1}^{\infty} \text{Var}\left(\frac{Z_n}{n}\right) < \infty \xRightarrow[\text{2-Series}]{\text{Kolmogorov}} \sum_{n=1}^{\infty} \frac{Z_n}{n} < \infty$ a.s. $\xRightarrow[\text{Lemma}]{\text{Kronecker's}} \frac{1}{n}\sum_{k=1}^{n} Z_k \xrightarrow{\text{a.s.}} 0$.

Claim: (Kolmogorov Truncation Lemma) $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[Y_k\right] \to \mu$. Use $\tilde{Y}_n := \begin{cases} X, & \text{if } |X| \le n, \\ 0, & \text{if } |X| > n. \end{cases}$.

$(\because)$We have $\tilde{Y}_n \cong Y_n \implies \mathbb{E}[\tilde{Y}_n] = \mathbb{E}[Y_n]$.

$\tilde{Y}_n \xrightarrow{\text{a.s.}} X$ and $|\tilde{Y}_n| \le |X| \xRightarrow{\text{DCT}} \mathbb{E}[\tilde{Y}_n] \to \mathbb{E}[X] = \mu \xRightarrow{\text{Cesaro}} \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[Y_k\right] \to \mu$ $\square$

Kolmogorov Truncation Lemma $\{X_n\}_{n=1}^{\infty} \overset{\text{iid}}{\sim} X$, $X \in L^1$.

$$Y_n = \begin{cases} X_n, & \text{if } |X_n| \le n, \\ 0, & \text{if } |X_n| > n. \end{cases} \implies \begin{cases} \mathbb{E}\left[Y_n\right] \to \mathbb{E}\left[X\right] \\ \mathbb{P}\left(Y_n \ne X_n \text{ i.o. }\right) = 0 \\ \sum_{n=1}^{\infty} \frac{\text{Var}\left(Y_n\right)}{n^2} < \infty \end{cases}$$

$(\because)$see above

L^2 **Rate of Conv.** (iid, L^2)$\{X_n\ iid\}$, $\mathbb{E}[X_n] = 0$ & $\mathrm{Var}(X_n) = \sigma^2 < \infty \implies \dfrac{S_n}{n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\varepsilon}} \xrightarrow{\text{a.s.}} 0$

Note Kolmogorov's test (Durrett [5], Thm 8.8.2): $\limsup_{n\to\infty} \dfrac{S_n}{n^{\frac{1}{2}}(\log\log n)^{\frac{1}{2}}} = \sigma\sqrt{2}$ a.s.

Proof: *(click)*

$$\sum_{n=1}^{\infty} \mathrm{Var}\left(\frac{X_n}{n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\varepsilon}}\right) = \frac{\sigma^2}{a_1^2} + \sigma^2 \sum_{n=2}^{\infty} \frac{1}{n(\log n)^{1+2\varepsilon}} < \infty \xRightarrow[\text{2-Series}]{\text{Kolmogorov}} \sum_{n=1}^{\infty} \frac{X_n}{n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\varepsilon}} < \infty \text{ a.s.}$$

$$\xRightarrow[\text{Lemma}]{\text{Kronecker's}} \frac{1}{n}\sum_{k=1}^{n} \frac{X_k}{k^{\frac{1}{2}}(\log k)^{\frac{1}{2}+\varepsilon}} \xrightarrow{\text{a.s.}} 0.$$

□

L^p **Rate of Conv.** X_i iid, $\mathbb{E}[X_i] = 0$ & $\exists p \in (1,2)$ S.T. $\mathbb{E}\left[|X_i|^p\right] < \infty$. Then, $\dfrac{S_n}{n^{1/p}} \xrightarrow{\text{a.s.}} 0$.

Note This applies to case where second moment might possibly be infinity.

Converse: $p > 0$: $\dfrac{S_n}{n^{1/p}} \xrightarrow{\text{a.s.}} 0 \implies \mathbb{E}\left[|X_1|^p\right] < \infty$

Proof: *(click)*

Similar techniques, $Y_k := X_k \mathbb{1}_{|X_k| \le k^{1/p}}$. Start with $\sum_{k=1}^{\infty} \mathbb{P}(X_k \neq Y_k) \le \sum_{k=1}^{\infty} \mathbb{P}\left(|X_i|^p > k\right) \le \mathbb{E}\left[|X_i|^p\right] < \infty$.

Hence, suff. to show $\dfrac{Y_1 + \cdots + Y_n}{n^{1/p}} \xrightarrow{\text{a.s.}} 0$.

□

Infinite Mean $\{X_i\ iid\}$, $\mathbb{E}[|X_i|] = \infty$, $S_n = X_1 + \ldots + X_n$ & $\{a_n \in \mathbb{R}_+\}$ S.T. $\dfrac{a_n}{n} \nearrow$:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|X_i| \ge a_n\right) < \infty \implies \limsup_{n\to\infty} \frac{|S_n|}{a_n} = 0; \sum_{n=1}^{\infty} \mathbb{P}\left(|X_i| \ge a_n\right) = \infty \implies \limsup_{n\to\infty} \frac{|S_n|}{a_n} = \infty$$

4.2 Day 2: Large Deviations

Main Reference(s): Rick Durrett, *Probability: Theory&Examples*, 2010 [5] Sec. 2.6

KEYWORDS:

Moment Generating Function $M_X(s) := \mathbb{E}\left[e^{sX}\right]$, whenever finite.

Rate Function For any $a > 0$, $\varphi(a) := \sup_{s>0}\left(sa - \log \mathbb{E}\left[e^{sX}\right]\right)$

MAIN THEOREMS: Assume $\{X_n\}$ iid and $\mathbb{E}[X_i] = 0$.

Chernoff Bound $\{X_n\}$ iid, $\mathbb{E}[X_i] = 0 \implies \mathbb{P}(S_n \geq na) \leq e^{-n(sa - \log M_X(s))}$, assuming $M_X(s) < \infty$ over $s \in [0, c]$, for some $c > 0$.

Proof: *(click)*

$\mathbb{P}(S_n \geq na) = \mathbb{P}\left(e^{s \cdot S_n} \geq e^{s \cdot na}\right) \leq \dfrac{\mathbb{E}\left[e^{s \cdot S_n}\right]}{e^{sna}} = e^{-n(sa - \log M_X(s))}$ using Markov Ineq, and $M_{S_n}(s) = M_X(s)^n$

Note If $M(s) = \infty$ everywhere, then ineq. above trivial. $\square$

Existence of the MGF $X \geq 0$ and $\exists c > 0$ S.T. $M_X(c) < \infty$: Then $M_X(s) < \infty$ $\forall s \in (-\infty, c]$.

Note Hence, once we have MGF to be finite at a point, we realize that it is finite on an interval, therefore we can take e.g. derivatives etc.

Proof: *(click)*

$$M_X(s) = \mathbb{E}\left[e^{sX}\right] = \int_{\{\omega : X(\omega) \geq 0\}} e^{sX}\, d\mathbb{P} + \int_{\{\omega : X(\omega) < 0\}} e^{sX}\, d\mathbb{P} \leq \int_\Omega e^{cX}\, d\mathbb{P} + 1 = 1 + M_X(c) < \infty. \quad \square$$

Differentiability of the MGF $X \geq 0$ & $M_X(s) < \infty$ $\forall s \in (-\infty, c]$ for some $c > 0$. Then
$\mathbb{E}[X] \overset{(1)}{=\!=} \mathbb{E}\left[\lim_{h \searrow 0} \dfrac{e^{hX} - 1}{h}\right] \overset{(2)}{=\!=} \lim_{h \searrow 0} \dfrac{\mathbb{E}\left[e^{hX}\right] - 1}{h}$,and
$\dfrac{d^k}{(ds)^k} M_X(s) = \mathbb{E}\left[X^k e^{sX}\right]$.

Proof: *(click)*

(a) Claim: $\mathbb{E}[X^k] < \infty$, $\forall k$

($\because$)Recall $x^k \leq e^{sx}$ for x large enough ($i.e.$ $x \geq a$ for some a).

$E[X^k] = \int_\Omega X^k(\omega)\, d\mathbb{P}(\omega) = \int_{\{\omega : X(\omega) \geq a\}} X^k \, d\mathbb{P} + \int_{\{\omega : X(\omega) < a\}} X^k \, d\mathbb{P} \leq \int_{\{\omega : X(\omega) \geq a\}} e^{sX} \, d\mathbb{P} + a^k \int_\Omega 1 \, d\mathbb{P}$

So $E[X^k] \leq M_X(s) + a^k < \infty$

Alternative Proof: Taylor: $e^{sX} = \sum_{k=0}^{\infty} \frac{(sX)^k}{k!} \geq \frac{s^k X^k}{k!} \Rightarrow \mathbb{E}\left[e^{sX}\right] \geq \frac{s^k}{k!}\mathbb{E}\left[X^k\right]$.

(b) Claim: $\mathbb{E}[X^k e^{sX}] < \infty$, for every $s < c$ and $\forall k$

($\because$)Recall $x^k e^{sx} \leq e^{s'x}$ $\forall k, s, s'$ with $s' \in (s, c)$ for x large enough ($i.e.$ $x \geq a$ for some a).

$E[X^k e^{sX}] = \int_\Omega X^k(\omega) e^{sX(\omega)} \, d\mathbb{P}(\omega) = \int_{\{\omega : X(\omega) \geq a\}} X^k e^{sX} \, d\mathbb{P} + \int_{\{\omega : X(\omega) < a\}} X^k e^{sX} \, d\mathbb{P}$

So $E[X^k e^{sX}] \leq \int_{\{\omega : X(\omega) \geq a\}} e^{s'X} \, d\mathbb{P} + a^k \int_\Omega e^{sa} \, d\mathbb{P} \leq M_X(s') + a^k e^{sa} < \infty$

(c) Claim: $(e^{hX} - 1)/h \leq X e^{hX}$

($\because$)Let $f(h) := e^{hX} \xRightarrow[\text{Theorem}]{\text{Mean Value}} \exists t \in (0, h)$ S.T. $f(h) - f(0) = hf'(t) \Rightarrow e^{hX} - 1 = hXe^{tX} \leq hXe^{hX}$

(d) Claim: $\mathbb{E}[X] \overset{(1)}{=\!=} \mathbb{E}\left[\lim_{h \searrow 0} \frac{e^{hX} - 1}{h}\right]$

($\because$)$X = \lim_{h \searrow 0} \frac{e^{hX} - 1}{h}$ by definition of the derivative of e^{hX} with respect to h is Xe^{hX} evaluated at 0.

Similarly, $\frac{d}{ds} M_X(s) = \lim_{h \searrow 0} \frac{M_X(s+h) - M_X(s)}{h} = \lim_{h \searrow 0} \frac{\mathbb{E}\left[e^{(s+h)X}\right] - \mathbb{E}\left[e^{sX}\right]}{h} = \frac{\mathbb{E}\left[e^{sX}(e^{hX} - 1)\right]}{h}$.

(e) Claim: $\mathbb{E}\left[\lim_{h \searrow 0} e^{sX} \frac{e^{hX} - 1}{h}\right] \overset{(2)}{=\!=} \lim_{h \searrow 0} \frac{\mathbb{E}\left[e^{(h+s)X}\right] - \mathbb{E}\left[e^{sX}\right]}{h}$

($\because$)We have $g_h(X) := \frac{e^{hX} - 1}{h} e^{sX} \Longrightarrow g_h(X) \xrightarrow{h \searrow 0} Xe^{sX}$, and by **(b)** & **(c)**: $g_h(X) \leq Xe^{(h+s)X} \in L^1(\Omega, \mathbb{P})$

$\Longrightarrow \lim_{h \searrow 0} \frac{\mathbb{E}\left[e^{(h+s)X}\right] - \mathbb{E}\left[e^{sX}\right]}{h} = \mathbb{E}\left[g_h(X)\right] \xrightarrow[DCT]{h \searrow 0} \mathbb{E}\left[Xe^{sX}\right] = \mathbb{E}\left[\lim_{h \searrow 0} e^{sX} \frac{e^{hX} - 1}{h}\right]$ ☐

Str. Convexity of Log MGF $M(s) < \infty$ for $s \in [0, c] \Rightarrow h(s) = sa - \log M_X(s)$ is strictly concave

Proof: *(click)*

We now know $M_X(s)$ differentiable, so $\frac{d^2}{ds^2} \log M_X(s) = \frac{d}{ds} \frac{M_X'(s)}{M_X(s)} = \frac{M_X''(s) M_X(s) - M_X'(s)^2}{M_X(s)^2}$

So $\frac{d^2}{ds^2} \log M_X(s) = \frac{\mathbb{E}\left[X^2 e^{sX}\right] \mathbb{E}\left[e^{sX}\right] - \left(\mathbb{E}\left[Xe^{sX}\right]^2\right)}{M_X(s)^2} \underset{\text{Cauchy-Schwarz}}{\geq} \frac{\left(\mathbb{E}\left[Xe^{sX}\right]^2\right) - \left(\mathbb{E}\left[Xe^{sX}\right]^2\right)}{M_X(s)^2} = 0$

$\Longrightarrow \frac{d^2}{ds^2} h(s) \leq 0$.

To get strict concavity, recall that Cauchy-Schwarz is strict unless $X^2 e^{sX} = \alpha e^{sX}$ a.s. , which happens if $X = \sqrt{\alpha}$ a.s. . Unless X is a degenerate r.v., the log MGF is strictly concave. ☐

Tighter Upper Bound If $M(s) < \infty$ for $s \in [0, c]$ & $a > 0$, then:

$$\mathbb{P}(S_n \geq na) \leq e^{-n\varphi(a)},\ \forall n \geq 1;\ \text{where } \varphi(a) := \sup_{s>0}\left[sa - \log M_X(s)\right] > 0,$$

$$\text{where } \sup_{s>0}\left[sa - \log M_X(s)\right] = s^*a - \log M_X(s^*) \text{ with } s^* \text{ is S.T. } \frac{M_X'(s^*)}{M_X(s^*)} \overset{!}{=} a$$

Proof: *(click)*

Claim: : $s \leq 0 \Longrightarrow sa - \log M_X(s) \leq 0$. Hence, can restrict to $s \geq 0$.

$(\because) sa - \log M_X(s) = sa - \log \mathbb{E}\left[e^{sX}\right] \underset{\text{Jensen's}}{\leq} sa - \mathbb{E}\left[\log e^{sX}\right] = sa - s\mathbb{E}[X] = sa \leq 0.$

Claim: $\exists s \in (0, c)$ s.t. $(sa - \log M_X(s)) > 0$.

$(\because) \left.\frac{d}{ds}\right|_{s=0} (sa - \log \mathbb{E}\left[e^{sX}\right]) = a - \frac{\mathbb{E}[X]}{\mathbb{E}\left[e^{0X}\right]} = a > 0.$ □

Convexity of Rate Function $\varphi(a) := \sup_{s>0}(sa - \log M_X(s))$ is convex $\forall a > 0$.

Corollary: $\varphi(a)$ is continuous on $(0, \infty)$

Proof: *(click)*

Notice, $\varphi(a) = \sup_{s>0} f_s(a)$, where $f_s(a) = sa - \log M_X(s)$. Each $f_s(a)$ is affine in a, hence, $\varphi(\cdot)$, being pointwise sup of family of affine functions is convex.

Analysis (e.g. Rudin) $\Longrightarrow$ "convex $\Rightarrow$ continuous" □

Log MGF Asymptotics X S.T. $F_X(x) < 1\ \forall x \in \mathbb{R}$ (X does not admit a finite upper bound)

$$\Longrightarrow \lim_{s\to\infty} \frac{\log M(s)}{s} = \infty.$$

Proof: *(click)*

Fix $b > 0$: $M_X(s) = \mathbb{E}\left[e^{sX}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{sX}|\mathbb{1}_{X>b}\right]\right] \geq e^{bs}\mathbb{P}(X > b)$.

$\mathbb{P}(X > b) = 1 - F_X(b) > 0 \implies \frac{\log M(s)}{s} \geq b + \frac{1}{s}\log\left(\mathbb{P}(X > b)\right)$ $\forall s$ large enough.

$\Longrightarrow \lim_{s\to\infty} \frac{\log M(s)}{s} \geq b + \lim_{s\to\infty} \frac{1}{s} \log \mathbb{P}(X > b) = b$ $\forall b$ large enough (as $\mathbb{P}(X > b) > 0 \Rightarrow \log \mathbb{P}(X > b) > -\infty$)

As $b \to \infty$: $\lim_{s\to\infty} \frac{\log M(s)}{s} \geq \lim_{b\to\infty} b = +\infty.$ □

Corollary: $\forall a > 0,\ \lim_{s\to\infty}(sa - \log M_X(s)) = -\infty.$

$(\because) sa - \log M_X(s) = s\Big(a - \underbrace{\frac{\log M_X(s)}{s}}_{\to+\infty}\Big)$

Tight Lower Bound Assume $X \sim f_X$ is continuous, $\mathbb{E}[X] = 0$, $M_X(s) < \infty\ \forall s \in \mathbb{R}$, and $0 < \mathbb{P}(X \leq x) < 1\ \forall x \in \mathbb{R}$ [3].

Then: $\lim\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) = -\varphi(a)$, $\forall a > 0$.

Note So Chernoff bound is tight: $e^{-n(\varphi(a)+\varepsilon)} \leq \mathbb{P}(S_n \geq na) \leq e^{-n\varphi(a)}\ \forall \varepsilon > 0, \forall n$ large.

Proof: *(click)*

From upper bound, $\limsup\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) \leq -\varphi(a)$. Hence, we need to prove the other direction.

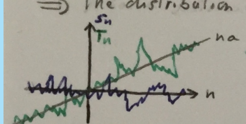
Proof = By Change of Measure!

- Fix s^* and define : $f_Y(x) = \dfrac{e^{s^* x} f_X(x)}{M_X(s^*)} \geq 0\ \forall x \in \mathbb{R}$ ($M_X(s^*)$ $\longrightarrow$ constant)
- f_Y is a PDF (⇔) $\int_{-\infty}^{\infty} f_Y(x)\,dx = \frac{1}{M_X(s^*)} \int_{-\infty}^{\infty} e^{s^* x} f_X(x)\,dx = \frac{1}{M_X(s^*)} \mathbb{E}[e^{s^* X}] = 1$
- So Y is a r.v. with PDF f_Y : $\begin{cases} M_Y(s) = \dfrac{M_X(s+s^*)}{M_X(s^*)} & \forall s \in \mathbb{R} \\ \mathbb{E}[Y] = a \end{cases}$

(⇔) $M_Y(s) = \int_{-\infty}^{\infty} e^{sy} f_Y(y)\,dy = \int_{-\infty}^{\infty} e^{sy} \cdot \dfrac{e^{s^* y} f_X(y)}{M_X(s^*)}\,dy = \dfrac{\mathbb{E}[e^{(s+s^*)X}]}{M_X(s^*)}$

(⇔) $\mathbb{E}[Y] = M_Y'(0) = \dfrac{M_X'(0+s^*)}{M_X(s^*)} = \dfrac{M_X'(s^*)}{M_X(s^*)} = a$

$\Rightarrow$ The distribution of Y is a tilted version of the distribution of X.



$\Rightarrow Y$: LLN $\Rightarrow \frac{T_n}{n} \approx \mathbb{E}[Y] = a \Rightarrow T_n \approx na$

$\Rightarrow X$: LLN $\Rightarrow \frac{S_n}{n} \approx \mathbb{E}[X] = 0 \Rightarrow S_n \approx 0$

Note : we want to find $\mathbb{P}(S_n \geq na) \leadsto$ use Y instead!
$\longrightarrow$ use $Y_1, \ldots, Y_n \sim f_Y$ iid r.v.

- Fix $a>0$, $\delta>0$: $B := \{(x_1, \ldots, x_n) : a-\delta \leq \frac{1}{n}\sum_{i=1}^{n} x_i \leq a+\delta\} \subset \mathbb{R}^n$ and let $S_n = \sum_{i=1}^{n} X_i$, $T_n = \sum_{i=1}^{n} Y_i$

$$\mathbb{P}(S_n \geq n(a-\delta)) \geq \mathbb{P}(n(a-\delta) \leq S_n \leq n(a+\delta)) = \int_{\vec{x}\in B} f_X(x_1)\cdots f_X(x_n)\,dx_1\cdots dx_n = M_X^n(s^*) \int_{\vec{x}\in B} \left(e^{-s^*(x_1+\cdots+x_n)}\right) \cdot \frac{e^{s^* x_1} f_X(x_1)}{M_X(s^*)} \cdots \frac{e^{s^* x_n} f_X(x_n)}{M_X(s^*)}\,dx_1\cdots dx_n$$

$$\geq M_X^n(s^*) \cdot e^{-s^* n(a+\delta)} \int_{\vec{x}\in B} f_Y(x_1)\cdots f_Y(x_n)\,dx_1\cdots dx_n$$

$$\Rightarrow \mathbb{P}(S_n \geq n(a-\delta)) \geq M_X^n(s^*) \cdot e^{-ns^*(a+\delta)} \cdot \mathbb{P}(T_n \in B)$$

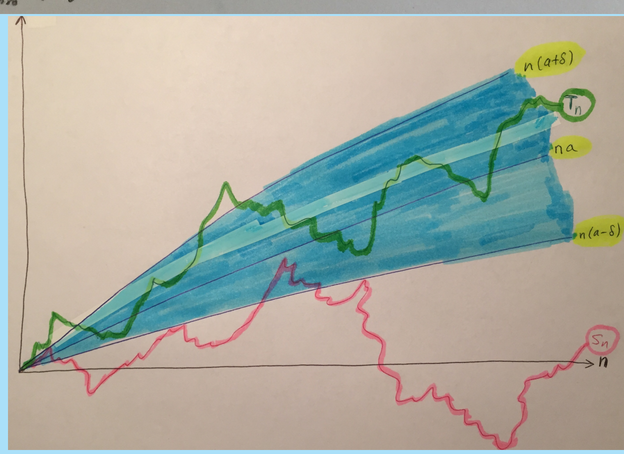
- WLLN $\Rightarrow \mathbb{P}(T_n \in B) = \mathbb{P}\left(\frac{Y_1+\cdots+Y_n}{n} \in [a-\delta, a+\delta]\right) \xrightarrow[n\to\infty]{} 1$ as $\mathbb{E}[Y] = a$

So : $\liminf\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na - n\delta) \geq \log M_X(s^*) - s^* a - s^*\delta = -\phi(a) - s^*\delta\quad \forall a>0, \forall \delta>0.$

replace $a \to a+\delta$: $\liminf\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) \geq -\phi(a+\delta) - s^*\delta \xrightarrow[\delta\to 0]{} -\phi(a) - 0$ as ϕ is continuous

$\Rightarrow \liminf\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) \geq -\phi(a) \underset{\text{upper bound}}{\geq} \frac{1}{n} \log \mathbb{P}(S_n \geq na)$

$\Rightarrow \liminf\limits_{n\to\infty} = \lim\limits_{n\to\infty} \Rightarrow \lim\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) \geq -\phi(a)$



□

[3] Namely X has full support on $\mathbb{R}$, so r.v. does not admit a finite lower/upper bound.

EXERCISES: From Durrett Chap. 2.6

Exercise:

Suppose that

$$-\nu := \limsup_{x\to\infty} \frac{\log \mathbb{P}(X > x)}{x} < 0. \tag{4.1}$$

Establish that $M_X(s) < \infty$ for all $s \in [0, \nu)$.

Solution: *(click)*

From Eq. (4.1), we know that given any $\varepsilon > 0$, we can find some x_ε such that

$$\frac{\log \mathbb{P}(X > x)}{x} \leq -\nu + \varepsilon,$$

for all $x \geq x_\varepsilon$. We will rewrite this as

$$\mathbb{P}(X > x) \leq e^{(-\nu+\varepsilon)x}. \tag{4.2}$$

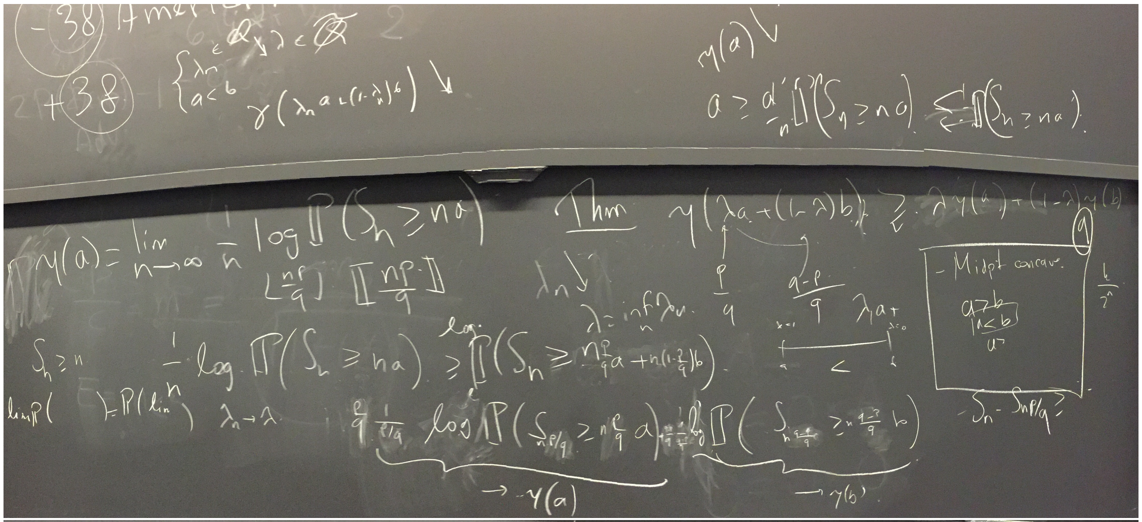
Now we want to turn the actual task of the problem, which is to show that $E[e^{sX}]$ exists. For each s, we will pick a particular ε such that Eq. (4.2) holds; the precise method of picking ε will be given a few lines below. We use a helpful formula for expectation:

$$\begin{aligned} E[e^{sX}] &= \int_{-\infty}^{+\infty} \mathbb{P}(e^{sX} > t)dt \\ &= \int_0^{x_\varepsilon} \mathbb{P}(e^{sX} > t)dt + \int_{x_\varepsilon}^{\infty} \mathbb{P}(e^{sX} > t)dt. \end{aligned}$$

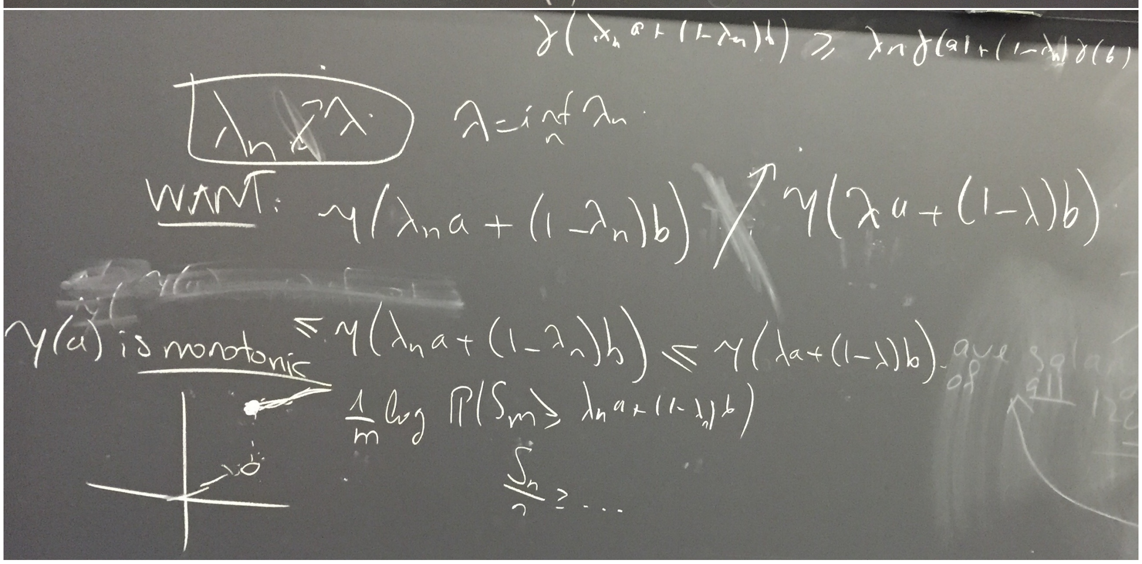
Now the first part above is the integral of a function which is at most 1 over a finite region, so it is finite; the second we bound as

$$\int_{x_\varepsilon}^{\infty} \mathbb{P}(e^{sX} > t)dt = \int_{x_\varepsilon}^{\infty} \mathbb{P}(X > \frac{\log t}{s})dt \leq \int_{x_\varepsilon}^{\infty} e^{(-\nu+\varepsilon)(\log t/s)}dt \leq \int_{x_\varepsilon}^{\infty} t^{(-\nu+\varepsilon)/s}$$

Now for any particular $s \in [0, \nu)$, we can pick ε so that $s + \varepsilon < \nu$, guaranteeing that $(-\nu + \varepsilon)/s < -1$. And now we are done, because $\int_0^\infty t^\alpha dt$ is finite when $\alpha < -1$.



Thm
Midpt concave



$\lambda = \inf_n \lambda_n$
WANT:
$\gamma(\lambda_n a + (1-\lambda_n)b) \nearrow \gamma(\lambda a + (1-\lambda)b)$
$\gamma(a)$ is monotonic

Chapter 5

Week 5: Product Measures & Conditional Expectation

5.1 Day 1: Product Measures

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

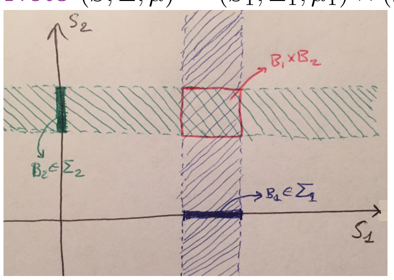
Coordinate Maps $\pi_1 : S_1 \times S_2 \to S_1$, $\pi_2 : S_1 \times S_2 \to S_2$, $\pi_1(x_1, x_2) = x_1$ & $\pi_2(x_1, x_2) = x_2$.

Product $\sigma-$Algebra (S_1, Σ_1) and (S_2, Σ_2) measurable spaces. Then:

$$\Sigma_1 \times \Sigma_2 := \sigma(\pi_1, \pi_2) = \sigma\left(\left\{\pi_1^{-1}(B_1), \pi_2^{-1}(B_2) : B_1 \in \Sigma_1, B_2 \in \Sigma_2\right\}\right) = \sigma\left(\{B_1 \times S_2, S_1 \times B_2 : S_1, \in \Sigma_1, S_2 \in \Sigma_2\}\right).$$

Note $B_1 \times S_1$ and $S_2 \times B_2$=∞-strips; $B_1 \times B_2$=rectangle

Note $(S, \Sigma, \mu) := (S_1, \Sigma_1, \mu_1) \times (S_2, \Sigma_2, \mu_2)$



Product Measure Let $E \in \Sigma := \Sigma_1 \times \Sigma_2$:

$$\mu(E) := \int_{S_1} \int_{S_2} \mathbb{1}_E \, d\mu_1 \, d\mu_2 = \int_{S_2} \int_{S_1} \mathbb{1}_E \, d\mu_2 \, d\mu_1 \text{ (using the b}\Sigma\text{ integral order swap).}$$

In particular, $E = E_1 \times E_2$ $(E_i \in \Sigma_i) \implies \mu(E) = \mu_1(E_1)\mu_2(E_2)$.

Joint Laws $\mathbb{P}_{X,Y} : \mathscr{B} \times \mathscr{B} \to [0, 1]$ S.T. $\mathbb{P}_{X,Y}(E) = \mathbb{P}\big((X, Y) \in E\big)$, $\forall E \in \Sigma_1 \times \Sigma_2$

Joint CDF $F_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ S.T. $F_{X,Y}(x, y) = \mathbb{P}(X \le x, Y \le y)$.

MAIN THEOREMS:

Product σ-Algebra Characterization $\overbrace{\Sigma_1 \times \Sigma_2}^{\infty\text{-strips}} = \sigma\Big(\Big\{ \overbrace{B_1 \times B_2}^{\text{bounded rectangles}} : B_1 \in \Sigma_1, B_2 \in \Sigma_2\Big\}\Big).$

Note The rectangle generator is easily seen to be a π-system.

Proof: *(click)*

$A := \{S_1 \times B_2, B_1 \times S_2 : B_1 \in \Sigma_1, B_2 \in \Sigma_2\}$ & $B := \{B_1 \times B_2 : B_1 \in \Sigma_1, B_2 \in \Sigma_2\}$.

$(S_1 \times B_2) \cap (B_1 \times S_2) = B_1 \times B_2 \Rightarrow B_1 \times B_2 \in \sigma(A)$, for all $B_1 \in \Sigma_1, B_2 \in \Sigma_2 \Rightarrow B \subseteq \sigma(A) \Rightarrow \sigma(B) \subseteq \sigma(A)$.

$B_1 \times \Sigma_2, S_1 \times B_2 \in B \Rightarrow A \subseteq B \Rightarrow \sigma(A) \subseteq \sigma(B)$.

Hence, $\boxed{\Sigma_1 \times \Sigma_2 = \sigma(A) = \sigma(B)}$. $\square$

Meas. Check $f : (S_1 \times S_2, \Sigma_1 \times \Sigma_2) \to (\mathbb{R}, \mathscr{B})$: $\int_{S_1} \int_{S_2} f(x_1, x_2)\, d\mu_1(x_1)\, d\mu_2(x_2)$ makes sense $\Leftrightarrow$

(i) $\forall x_1 \in S_1, x_2 \mapsto f(x_1, x_2)$ Σ_2-meas.

(iii) $x_1 \mapsto \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2)$ Σ_1-meas.

(ii)$\forall x_1,\ f(x_1, x_2)$ μ_2-integrable

(iv) $x_1 \mapsto \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2)$ μ_1-integr.

Proof: *(click)*

We will use Monotone-Class argument.

(i) Let $\mathcal{H} = \{f : (S_1 \times S_2, \Sigma_1 \times \Sigma_2) \to (\mathbb{R}, \mathscr{B})$ bdd S.T. **(i)** holds.$\}$.

- $f, \tilde{f} \in \mathcal{H}$. $x_2 \mapsto f(x_1, x_2) + \tilde{f}(x_1, x_2)$ is Σ_2-meas., sum of meas.
- $f(x_1, x_2) = 1 \implies x_2 \mapsto f(x_1, x_2) = 1$, $\forall x_1$, & $1 \in m\Sigma_2$.
- $0 \leq f_n \nearrow f \implies \forall x_1,\ 0 \leq \underbrace{f_n(x_1, x_2)}_{:= g_n^{x_1}(x_2)} \nearrow \underbrace{f(x_1, x_2)}_{:= g^{x_1}(x_2)}$. $g^{x_1}(x_2) = \lim_{n\to\infty} g_n^{x_1}(x_2)$, meas.
- $\{B_1 \times B_2 : B_1 \in \Sigma_1, B_2, \in \Sigma_2\}$ π-system. $f := \mathbb{1}_{B_1 \times B_2}(x_1, x_2)$.

  $x_1 \notin B_1 \implies x_2 \mapsto f(x_1, x_2) = 0 \in m\Sigma_2$.

  $x_1 \in B_1 \implies x_2 \mapsto f(x_1, x_2) = \mathbb{1}_{B_2} \in m\Sigma_2$.

(iii) Let $\mathcal{H} = \{f : (S_1 \times S_2, \Sigma_1 \times \Sigma_2) \to (\mathbb{R}, \mathscr{B})$ bdd S.T. **(iii)** holds.$\}$.

- $f, \tilde{f} \in \mathcal{H}$. $x_1 \mapsto \int_{S_2} (f(x_1, x_2) + \tilde{f}(x_1, x_2))\, d\mu_2(x_2) = \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2) + \int_{S_2} \tilde{f}(x_1, x_2)\, d\mu_2(x_2)$.
- $f(x_1, x_2) = 1 \implies x_1 \mapsto \int_{S_2} 1\, d\mu_2 = \mu_2(S_2)$[a].
- $0 \leq f_n \nearrow f \implies \forall x_1,\ \lim_{n\to\infty} \underbrace{\int_{S_2} f_n(x_1, x_2)\, d\mu_2(x_2)}_{\mu_1\text{-meas}} \stackrel{MCT}{=} \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2)$, so μ_1-meas.
- $\{B_1 \times B_2 : B_1 \in \Sigma_1, B_2, \in \Sigma_2\}$ π-system & $f := \mathbb{1}_{B_1 \times B_2}(x_1, x_2)$:

  $x_1 \notin B_1 \implies x_1 \mapsto \int_{S_2} 0\, d\mu_2(x_2) = 0 \in m\Sigma_1$.

  $x_1 \in B_1 \implies \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2) = \int_{S_2} \mathbb{1}_{B_2}\, d\mu_2(x_2) = \mu_2(B_2)$.

  Hence, $x_1 \mapsto \int_{S_2} \mathbb{1}_{B_1 \times B_2}(x_1, x_2)\, d\mu_2(x_2) = \mu_2(S_2)\mathbb{1}_{B_1}(x_1) \in m\Sigma_1$.

(ii) and (iv) $\exists B > 0$ S.T. $|f(x_1, x_2)| \leq B$ & $\mu_{1,2}$ finite $\Rightarrow$ integrability. $\square$

[a] This requires finiteness of μ_2. Will also comment that results hold for σ-finite spaces too. However, for non-σ-finite measures, it won't hold; as will be exhibited via an example later.

Integral Order $(b\Sigma)$ $\forall$ <u>bdd</u> & <u>meas.</u> $f : (S_1 \times S_2, \Sigma_1 \times \Sigma_2) \to (\mathbb{R}, \mathscr{B})$:

$$\int_{S_1} \int_{S_2} f(x_1, x_2)\, d\mu_2(x_2)\, d\mu_1(x_1) = \int_{S_2} \int_{S_1} f(x_1, x_2)\, d\mu_1(x_1)\, d\mu_2(x_2). \qquad (\star)$$

Proof: *(click)*

$\mathcal{H} = \{f : (S_1 \times S_2, \Sigma_1 \times \Sigma_2) \to (\mathbb{R}, \mathscr{B})$ bdd, S.T. $(\star)$ holds$\}$. <u>GOAL:</u> $\mathcal{H}$ mon. class.

- $\mathcal{H}$ vec. space. ✓
- $f(\cdot, \cdot) = 1 \Rightarrow$ RHS $=$ LHS $= \mu_1(S_1)\mu_2(S_2)$. ✓
- $0 \le f_n \nearrow f$ & $f_n \in \mathcal{H}$: $\displaystyle\int_{S_1} \int_{S_2} f\, d\mu_2\, d\mu_1 = \int_{S_1} \int_{S_2} \lim_{n\to\infty} f_n\, d\mu_2\, d\mu_1 \overset{MCT}{=\!=} \int_{S_1} \lim_{n\to\infty} \int_{S_2} f_n\, d\mu_2\, d\mu_1$

$$\overset{MCT}{=\!=} \lim_{n\to\infty} \int_{S_1} \int_{S_2} f_n\, d\mu_2\, d\mu_1 \overset{f_n \in \mathcal{H}}{=\!=} \lim_{n\to\infty} \int_{S_2} \int_{S_1} f_n\, d\mu_1\, d\mu_2 \overset{MCT}{=\!=} \int_{S_2} \lim_{n\to\infty} \int_{S_1} f_n\, d\mu_1\, d\mu_2$$

$$\overset{MCT}{=\!=} \int_{S_2} \int_{S_1} \lim_{n\to\infty} f_n\, d\mu_1\, d\mu_2 = \int_{S_2} \int_{S_1} f\, d\mu_1\, d\mu_2$$

- $f = \mathbb{1}_{E_1 \times E_2}$, then RHS & LHS eval. to $\mu_1(E_1)\mu_2(E_2)$. □

Uniqueness of Product Meas. $\mu := \mu_1 \times \mu_2 =$ unique meas. S.T. $\mu(E_1 \times E_2) = \mu_1(E_1) \times \mu_2(E_2)$

Proof: *(click)*

$\mathcal{I} = \{E_1 \times E_2 : E_1 \in \Sigma_1, E_2 \in \Sigma_2\}$, π-system generating $\Sigma_1 \times \Sigma_2$

Apply uniqueness theorem ($\mu(S_1 \times S_2) < \infty$, as finite measures). □

Fubini's Theorem (S_i, Σ_i, μ_i) $(i = 1, 2)$ σ-finite meas. space[1], $(S, \Sigma, \mu) = (S_1, \Sigma_1, \mu_1) \times (S_2, \Sigma_2, \mu_2)$.

IF either: $\boxed{f \ge 0}$ OR $\boxed{\displaystyle\int_{S_2} \int_{S_1} |f| d\mu_1\, d\mu_2 < \infty}$

THEN: $\displaystyle\mu(f) = \int_{S_1 \times S_2} f\, d(\mu_1 \times \mu_2) = \int_{S_1} \int_{S_2} f\, d\mu_2\, d\mu_1 = \int_{S_2} \int_{S_1} f\, d\mu_1\, d\mu_2$

Proof: *(click)*

Already proven for $b\Sigma$, in particular for SF^+.

For any $f \ge 0$, take $0 \le f_n \nearrow f$, $f_n \in SF^+$, and apply MCT.

If $\mu(|f|) < \infty$, then do $|f| = f^+ - f^-$ & use linearity. □

Measurability of Diag. $(S_1, \Sigma_1) = (S_2, \Sigma_2) = (\mathbb{R}, \mathscr{B})$, $D = \{(x, y) \in \mathbb{R}^2 : x = y\} \Rightarrow D \in \mathscr{B} \times \mathscr{B}$.

[1] We have proven until this point the results for finite meas. spaces. For σ-finite meas. spaces, it is all of a matter of breaking into countably many pieces of finite measure, and then arguing via MCT.

Proof: *(click)*

(1^{st} **Way**) Let $f:\mathbb{R}^2 \to \mathbb{R}$ be $f(x,y)=x-y$. Proj. maps meas. $\Rightarrow$ f $\mathscr{B}\times\mathscr{B}$ meas, $D=f^{-1}(\{0\})$.

(2^{nd} **Way**) $D^c = \big\{(x,y): x\neq y\big\} \overset{\text{Trick:}}{=\!=} \bigcup_{q\in\mathbb{Q}} \{(x,y): x>q>y\} = \bigcup_{q\in\mathbb{Q}} (q,\infty)\times(-\infty,q)$. $(q,\infty),(-\infty,q)\in\mathscr{B}$.

(3^{rd} **Way**) Vertical lines $\{x\}\times\mathbb{R}$ are meas. $\xRightarrow[\text{Invariance}]{\text{Rotation}}$ diagonal is meas. $\square$

Exercise:

Breaking Fubini: <u>NON</u> σ-finite (counting) measure on $([0,1],\mathscr{B})$.

Let $(S_i,\Sigma_i)=([0,1],\mathscr{B})$ $(i=1,2)$ & $\mu_1 =$ Lebesgue & $\mu_2 =$ counting. Then D diagonal:

$$\int_{S_1}\int_{S_2} \mathbb{1}_D(x_1,x_2)\,d\mu_2\,d\mu_1 = \int_{S_1} \mu_2(\{x_1\})\,d\mu_1(x_1) = \mu_1([0,1]) = 1.$$
$$\int_{S_2}\int_{S_1} \mathbb{1}_D(x_1,x_2)\,d\mu_1\,d\mu_2 = \int_{S_2} \mu_1(\{x_2\})\,d\mu_2(x_2) = 0.$$

Exercise:

Breaking Fubini: Discrete nonnegativity and integrability.

Let $(S_i,\Sigma_i)=(\mathbb{N},2^{\mathbb{N}})$ $(i=1,2)$ & $\mathbb{P}_1=\mathbb{P}_2=$ counting (i.e., $\mathbb{P}(A)=|A|$).

Then: $\displaystyle\int_A g\,d\mathbb{P}_1 = \sum_{a\in A} f(a)$ & $\quad \displaystyle\int_B h\,d\mathbb{P}_2 = \sum_{b\in B} h(b)$ & $\quad \displaystyle\int_C f\,d(\mathbb{P}_1\times\mathbb{P}_2) = \sum_{c\in C} f(c)$.

$$\text{Let } f \text{ be S.T.} \begin{cases} f(m,m)=1 \\ f(m,m+1)=-1 \\ f(x,y)=0 \text{ elsewhere.} \end{cases} \implies \text{visually: } \begin{matrix} 1 & -1 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{matrix}$$

So: $\displaystyle\int_{\Omega_1}\int_{\Omega_2} f\,d\mathbb{P}_1\,d\mathbb{P}_2 = \sum_n\sum_m f(n,m) = 0 \neq 1 = \sum_m\sum_n f(n,m) = \int_{\Omega_2}\int_{\Omega_1} f\,d\mathbb{P}_2\,d\mathbb{P}_1$.

<u>Fail:</u> here, f is neither nonnegative nor integrable $\left(\sum |\pm 1| = \infty\right)$.

Exercise:

Breaking Fubini: Continuous nonnegativity and integrability.

Let $(S_1,\Sigma_1)=\big((0,1),\mathscr{B}\big)$ and $(S_2,\Sigma_2)=\big((1,\infty),\mathscr{B}\big)$ & $\mathbb{P}_1=\mathbb{P}_2=$ Lebesgue.

Let $f(x,y)=e^{-xy}-2e^{-2xy}$ (can take positive and negative values)

So: $\displaystyle\int_{\Omega_1}\int_{\Omega_2} f\,d\mathbb{P}_1\,d\mathbb{P}_2 = \int_0^1\int_1^\infty f(x,y)\,dy\,dx > 0 > \int_0^1\int_1^\infty f(x,y)\,dx\,dy$.

<u>Fail:</u> here, f is neither nonnegative nor integrable.

Exp. Value : Area under Graph $X : (\Omega, \Sigma, \mathbb{P}) \to (\mathbb{R}, \mathscr{B}, \lambda)$, $X \geq 0$, p-space $(\Omega \times \mathbb{R}, \mathbb{P} \times \lambda)$.

Graph of X $= \{(\omega, x) : 0 \leq x \leq X(\omega)\} \subset \Omega \times \mathbb{R}$ (check it is $\Sigma \times \mathscr{B}$-meas.).

Area of graph : $(\mathbb{P} \times \lambda)($ graph of $X) = \int_{\Omega \times \mathbb{R}} \mathbb{1}_{\text{graph of } X} \, d(\mathbb{P} \times \lambda)$.

Do in two ways to show $\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X \geq x) \, dx$.

Joint Law: Uniqueness $\mathscr{B} \times \mathscr{B}$ gen. by. $\mathcal{I} = \{(-\infty, x] \times (-\infty, y] : x, y \in \mathbb{R}\}$, $\mathcal{I}$ π-system.

$\mathbb{P}_{X,Y}$ uniquely specified by $F_{X,Y}$.

Joint Continuity X, Y jointly cont. if $\exists f : (\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2)) \to (\mathbb{R}_{\geq 0}, \mathscr{B})$, S.T.

$\mathbb{P}_{X,Y}((-\infty, x] \times (-\infty, y]) = \int \mathbb{1}_{(-\infty,x] \times (-\infty,y]} f(x, y) \, dx \, dy$.

$(\because)$Argue that $\mathbb{P}_{X,Y}(F) = \int \mathbb{1}_F f(x, y) \, dx \, dy$, $\forall F \in \mathscr{B} \times \mathscr{B}$.

Marginal Density Recall, X cont. $\Leftrightarrow \exists f$ non-neg.$\in \sigma\mathscr{B}$ S.T. $\mathbb{P}_X(S) = \int_S f \, dx$.

Then: $f_X(x) = \int_{\mathbb{R}} f(x, y) \, dy$.

Proof: *(click)*

$$\mathbb{P}_X(S) = \mathbb{P}_{X,Y}(S \times \mathbb{R}) = \mathbb{P}\left((X, Y) \in S \times R\right) = \int\int \mathbb{1}_{S \times \mathbb{R}}(x, y) f(x, y) \, dy \, dx = \int_S \underbrace{\int_{\mathbb{R}} f(x, y) \, dy;}_{:= f_X(x)} dx \qquad \square$$

Independence and Product Measure $X \sim (\mathbb{P}_X, F_X)$&$Y \sim (\mathbb{P}_Y, F_Y)$. TFAE:

- $\star$ $X \perp\!\!\!\perp Y$
- $\star$ $\mathbb{P}_{X,Y} = \mathbb{P}_X \times \mathbb{P}_Y$
- $\star$ $F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$
- $\star$ If joint PDF exists: $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for $\lambda \times \lambda -$ a.e. (x, y)

5.2 Day 2: Conditional Expectation

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Conditional Expectation Given $(\Omega, \mathscr{F}, \mathbb{P})$, $\mathscr{G} \subset \mathscr{F}$ σ-algebra & X S.T. $\mathbb{E}\left[|X|\right] < \infty$.

$$\text{Let } (\star) = \begin{cases} (i) & \mathbb{E}\left[|Y|\right] < \infty \\ (ii) & Y \text{ is } \mathscr{G}-\text{measurable} \\ (iii) & \displaystyle\int_G Y \, d\mathbb{P} = \int_G X \, d\mathbb{P}, \ \forall G \in \mathscr{G} \end{cases}$$

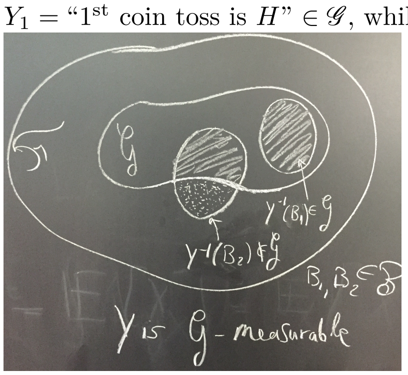
A r.v. Y satisfying $(\star)$ is called a version of the conditional expectation $\mathbb{E}\left[X|\mathscr{G}\right] =: Y$ a.s.

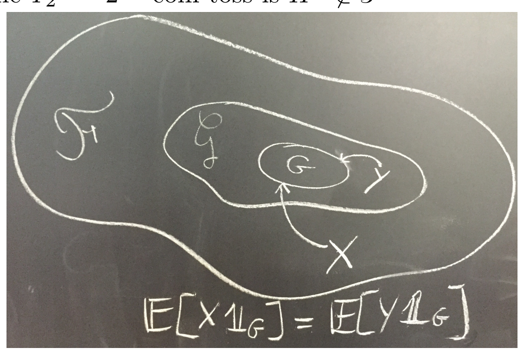
Note $\mathbb{E}[X|Z] := \mathbb{E}[X|\sigma(Z)]$ and $\mathbb{E}[X|Z_1, \ldots, Z_n] := \mathbb{E}[X|\sigma(Z_1, \ldots, Z_n)]$

Idea I can decide whether or not events of Y happen only based on the info I have in $\mathscr{G} \subset \mathscr{F}$.

Example: Coin toss: $\Omega = \{H, T, HH, HT, TH, TT\}$ $\mathscr{F} = \{H, T\}^2$ $\mathscr{G} = \{H, T\}$

$Y_1 =$ "1$^{\text{st}}$ coin toss is H" $\in \mathscr{G}$, while $Y_2 =$ "2$^{\text{nd}}$ coin toss is H" $\notin \mathscr{G}$





MAIN THEOREMS:

Practical π-system Check To check that a r.v. Y verifies $(\star)$ part (iii), enough to check on a π-system $\mathcal{I}$ S.T. $\Omega \in \mathcal{I}$ and $\sigma(\mathcal{I}) = \mathscr{G}$.

Proof: *(click)*

$\mathcal{I}$ π-system, $\mathscr{G} = \sigma(\mathcal{I})$. $A := \left\{G \in \mathscr{G} : \displaystyle\int_G Y \, d\mathbb{P} = \int_G X \, d\mathbb{P}\right\}$.

- $\Omega \in A$.
- $G, H \in A$, $G \subseteq H \Rightarrow H \setminus G \in A$ $\left(\text{since } \displaystyle\int_{H \setminus G} Y \, d\mathbb{P} = \int_H Y \, d\mathbb{P} - \int_G Y \, d\mathbb{P}\right)$.
- $A_n \in A$, $A_n \nearrow A$. We know $\mathbb{1}_{A_n} \nearrow \mathbb{1}_A$. $Y_n = Y\mathbb{1}_{A_n}$. $|Y_n| \leq |Y| \in L^1$, apply DCT.

Hence, A is a d-system containing $\mathcal{I} \Rightarrow A = d(A) \supseteq d(\mathcal{I}) = \sigma(\mathcal{I}) = \mathscr{G}$. $\square$

Existence & Uniqueness Given $(\Omega, \mathscr{F}, \mathbb{P})$, r.v. X S.T. $\mathbb{E}\left[|X|\right] < \infty$ & $\mathscr{G} \subseteq \mathscr{F}$ (sub σ-algebra):

$\implies$ There exists a r.v. Y satisfying $(\star)$ above.

If $\exists \tilde{Y}$ satisfying $(\star)$, then $\tilde{Y} = Y$ a.s. (i.e., $\mathbb{P}\left(\tilde{Y} = Y\right) = 1$).

Note $\varphi(Y)$ is a Cond. Exp of $X \iff \varphi$ is meas. & $\mathbb{E}\left[(\varphi(Y) - X)g(Y)\right] = 0 \; \forall g$ meas.[2]

Proof: *(click)*

<u>**Uniqueness:**</u> Assume $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ & $Y, \tilde{Y}$ ver. of $\mathbb{E}\left[X|\mathscr{G}\right]$. <u>GOAL:</u> $\mathbb{P}(Y = \tilde{Y}) = 1$.

By definition of $\mathbb{E}\left[X|\mathscr{G}\right]$: $Y, \tilde{Y} \in L^1$ and $\mathbb{E}\left[(Y - \tilde{Y})\mathbb{1}_G\right] = 0 \; \forall G \in \mathscr{G}$ by $(\star)$ part (iii).

$A_n := \left\{\omega \in \Omega : Y(\omega) - \tilde{Y}(\omega) \geq \frac{1}{n}\right\}$. $Y - \tilde{Y} \in m\mathscr{G} \Rightarrow A_n \in \mathscr{G}$.

$0 = \displaystyle\int_{A_n} (Y - \tilde{Y})\, d\mathbb{P} \geq \frac{1}{n}\mathbb{P}(A_n) \Rightarrow \mathbb{P}(A_n) = 0 \Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0 \Rightarrow \mathbb{P}\left(Y > \tilde{Y}\right) = 0.$

Similarly, $\mathbb{P}\left(\tilde{Y} > Y\right) = 0 \Rightarrow \mathbb{P}\left(Y = \tilde{Y}\right) = 1$.

<u>**Existence: Radon-Nikodym Proof**</u>

Recall Radon-Nikodym (RN) Theorem: $\mathbb{P}_2 \ll \mathbb{P}_1 \Rightarrow \exists Y : \Omega \to \mathbb{R}_{\geq 0}$ S.T. $\mathbb{P}_2(A) = \mathbb{E}_{\mathbb{P}_1}[Y\mathbb{1}_A] = \int_A Y\, d\mathbb{P}_1$.

WLOG assume $X \geq 0$ (otherwise use $X^+ - X^-$) & X not constant: then $\mathbb{E}[X] < \infty$ & $\mathbb{E}[X] > 0$. **(i)** ✓

Define $\mathbb{P}_2(A) := \dfrac{\mathbb{E}[X\mathbb{1}_A]}{\mathbb{E}[X]} = \dfrac{\int_A X\, d\mathbb{P}}{\int_\Omega X\, d\mathbb{P}}$ $(\forall A \in \mathscr{G})$. $\mathbb{P}_2(\cdot)$ valid prob. meas. (on $\mathscr{G}$) & $\mathbb{P}_2 \ll \mathbb{P}$.

$\xrightarrow[\text{Nikodym}]{\text{Radon}}$[a] $\exists Z \geq 0$ S.T. $\mathbb{P}_2(A) = \mathbb{E}[Z\mathbb{1}_A] = \displaystyle\int_A Z\, d\mathbb{P}$ and Z is $\mathscr{G}$-measurable.

$\implies \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Z\mathbb{1}_A]\,\mathbb{E}[X] = \mathbb{E}\Big[\underbrace{(Z\,\mathbb{E}[X])}_{:=Y}\mathbb{1}_A\Big]$.

$\implies$ If $Y := (Z\,\mathbb{E}[X])$, then $\mathbb{E}[Y\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] \; \forall A \in \mathscr{G}$ **(iii)** ✓ and Y is $\mathscr{G}$-measurable **(ii)** ✓

<u>**Existence: Orthogonal Projection Proof**</u> (see Williams [13] p. 86)

<u>Sketch:</u> use the completeness of $L^2(\Omega, \mathscr{G}, \mathbb{P})$ & $L^2(\Omega, \mathscr{G}, \mathbb{P}) \subset L^2(\Omega, \mathscr{F}, \mathbb{P})$ to construct orthogonal projection Y of X onto $L^2(\Omega, \mathscr{G}, \mathbb{P})$. **Y will be the conditional expectation of X.**

Recall that we get $\langle X - Y, Z\rangle = 0$ together with $Z = \mathbb{1}_G$ for $G \in \mathscr{G}$.

Once done with L^2: truncate $X_n = X \wedge n$ & $\xRightarrow{\text{MCT}}$ get L^1. □

[a] We apply RN to the sub-σ-algebra $\mathscr{G}$: $\mathbb{P}_2$ & $\mathbb{P}$ are measures on $(\Omega, \mathscr{G})$, where $\mathbb{P}$ can safely be restricted to sub-σ-algebra, then the RN derivative is automatically measurable w.r.t. the common σ-algebra on which measures are defined.

Positivity $0 \leq X \in L^1 \implies \mathbb{E}\left[X|\mathscr{G}\right] \geq 0$ a.s. .

Proof: *(click)*

$Z := \mathbb{E}\left[X|\mathscr{G}\right]$. $A_n = \left\{\omega : Z(\omega) < -\frac{1}{n}\right\}$

$\implies 0 \geq -\dfrac{1}{n}\mathbb{P}(A_n) \geq \displaystyle\int_{A_n} Z\, d\mathbb{P} = \int_{A_n} X\, d\mathbb{P} \geq 0 \implies \mathbb{P}(A_n) = 0 \; \forall n.$ □

[2] Try Monotone Class Argument

Properties of Conditional Expectation Let $X \in L^1$ & $\mathscr{G}, \mathcal{H}$ sub σ-algebras of $\mathscr{F}$.

(a) **Iterated Expectation:** $\mathbb{E}\left[\mathbb{E}\left[X|\mathscr{G}\right]\right] = \mathbb{E}\left[X\right]$

(b) **Full Knowledge:** X is $\mathscr{G}$-meas $\implies \mathbb{E}\left[X|\mathscr{G}\right] = X$ a.s.

(c) **Linearity:** $\mathbb{E}\left[a_1X_1 + a_2X_2|\mathscr{G}\right] = a_1\mathbb{E}\left[X_1|\mathscr{G}\right] + a_2\mathbb{E}\left[X_2|\mathscr{G}\right]$ a.s. , $\forall a_1, a_2 \in \mathbb{R}$

(d) **Positivity:** $X \geq 0 \implies \mathbb{E}\left[X|\mathscr{G}\right] \geq 0$ a.s.

Corollary: $X \leq \tilde{X} \implies \mathbb{E}\left[X|\mathscr{G}\right] \leq \mathbb{E}\left[\tilde{X}|\mathscr{G}\right]$ a.s.

(e) **MCT:** $0 \leq X_n \nearrow X \implies \mathbb{E}\left[X_n|\mathscr{G}\right] \nearrow \mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

(f) **Fatou:** $X_n \geq 0 \implies \mathbb{E}\left[\liminf_{n\to\infty} X_n|\mathscr{G}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X_n|\mathscr{G}\right]$ a.s.

$Y \in L^1$: $X_n \leq Y\ \forall n \implies \mathbb{E}\left[\limsup_{n\to\infty} X_n|\mathscr{G}\right] \geq \limsup_{n\to\infty} \mathbb{E}\left[X_n|\mathscr{G}\right]$ a.s.

(g) **DCT:** $X_n \xrightarrow{\text{a.s.}} X,\ |X_n| \leq V$ a.s. $\forall n$ & $\mathbb{E}[V] < \infty \implies \mathbb{E}\left[X_n|\mathscr{G}\right] \xrightarrow{\text{a.s.}} \mathbb{E}\left[X|\mathscr{G}\right]$

(h) **Jensen:** $c : \mathbb{R} \to \mathbb{R}$ convex: $\mathbb{E}\left[|c(X)|\right] < \infty \implies \mathbb{E}\left[c(X)|\mathscr{G}\right] \geq c\Big(\mathbb{E}\left[X|\mathscr{G}\right]\Big)$ a.s.

Corollary: $\|\mathbb{E}\left[X|\mathscr{G}\right]\|_p \leq \|X\|_p$, $\forall p \geq 1$.

(i) **Tower Property:** $\mathcal{H} \subset \mathscr{G} \subset \mathscr{F} \implies \mathbb{E}\left[\mathbb{E}\left[X|\mathscr{G}\right]\Big|\mathcal{H}\right] = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{H}\right]\Big|\mathscr{G}\right] = \mathbb{E}\left[X|\mathcal{H}\right]$ a.s.

(j) **Factorization:** Z $\mathscr{G}$-meas. & bounded $\implies \mathbb{E}\left[ZX|\mathscr{G}\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

Corollary: $X \in L^p(\Omega, \mathscr{F}, \mathbb{P})$, $Z \in L^q(\Omega, \mathscr{F}, \mathbb{P})$, $p, q > 1$:

$p^{-1} + q^{-1} = 1 \implies \mathbb{E}\left[ZX|\mathscr{G}\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

Corollary: $X \in (\mathrm{m}\mathscr{F})^+$, $Z \in (\mathrm{m}\mathscr{G})^+$: $\mathbb{E}[X], \mathbb{E}[ZX] < \infty \implies \mathbb{E}\left[ZX|G\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

(k) **Independence:** $\mathcal{H} \perp\!\!\!\perp \sigma\left(\sigma(X), \mathscr{G}\right) \implies \mathbb{E}\left[X|\sigma(\mathscr{G}, \mathcal{H})\right] = \mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

Corollary: $X \perp\!\!\!\perp \mathcal{H} \implies \mathbb{E}\left[X|\mathcal{H}\right] = \mathbb{E}\left[X\right]$ a.s.

Proof: *(click)*

(a) **Iterated Expectation:** $\mathbb{E}\left[\mathbb{E}\left[X|\mathscr{G}\right]\right] = \mathbb{E}\left[X\right]$

Taking $G = \Omega$ in $(\star, iii)$: $\mathbb{E}\left[X|\mathscr{G}\right] = \mathbb{E}\left[Y\right] = \mathbb{E}\left[Y\mathbb{1}_\Omega\right] = \mathbb{E}\left[X\mathbb{1}_\Omega\right] = \mathbb{E}\left[X\right]$.

(b) **Full Knowledge:** X is $\mathscr{G}$-meas $\implies \mathbb{E}\left[X|\mathscr{G}\right] = X$ a.s.

Notice $Y = X$ satisfies all assumptions in $(\star)$.

(c) **Linearity:** $\mathbb{E}\left[a_1X_1 + a_2X_2|\mathscr{G}\right] = a_1\mathbb{E}\left[X_1|\mathscr{G}\right] + a_2\mathbb{E}\left[X_2|\mathscr{G}\right]$ a.s. , $\forall a_1, a_2 \in \mathbb{R}$

Linear combo of $\mathscr{G}$-meas. fnc. is $\mathscr{G}$-meas & linearity of integral.

(d) **Positivity:** $X \geq 0 \implies \mathbb{E}\left[X|\mathscr{G}\right] \geq 0$ a.s. Corollary: $X \leq \tilde{X} \implies \mathbb{E}\left[X|\mathscr{G}\right] \leq \mathbb{E}\left[\tilde{X}|\mathscr{G}\right]$ a.s.

Proved above. For corollary, use linearity and positivity.

(e) **MCT:** $0 \leq X_n \nearrow X \implies \mathbb{E}\left[X_n|\mathscr{G}\right] \nearrow \mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

$Y_n := \mathbb{E}\left[X_n|\mathscr{G}\right] \Rightarrow 0 \leq Y_n \nearrow$ (from monotonicity and linearity) $\Rightarrow Y := \lim_{n\to\infty} Y_n \exists \xRightarrow{Y_n \in m\mathscr{G}} Y \in m\mathscr{G}$.

$\int_G Y\, d\mathbb{P} \overset{MCT}{=\!=} \lim_{n\to\infty} \int_G Y_n\, d\mathbb{P} \overset{def}{=\!=} \lim_{n\to\infty} \int_G X_n\, d\mathbb{P} \overset{MCT}{=\!=} \int_G X\, d\mathbb{P}\ (\forall G \in \mathscr{G})$. $\square$

Proof: *(click)*

(f) **Fatou:** $X_n \geq 0 \implies \mathbb{E}\left[\liminf_{n\to\infty} X_n|\mathscr{G}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X_n|\mathscr{G}\right]$ a.s.

$$Y \in L^1\colon X_n \leq Y\ \forall n \implies \mathbb{E}\left[\limsup_{n\to\infty} X_n|\mathscr{G}\right] \geq \limsup_{n\to\infty} \mathbb{E}\left[X_n|\mathscr{G}\right] \text{ a.s.}$$

$$Y_n := \inf_{k\geq n} X_k \nearrow \liminf_{n\to\infty} X_n \overset{\text{MCT}}{\Longrightarrow} \mathbb{E}\left[\liminf_{n\to\infty} X_n|\mathscr{G}\right] = \lim_{n\to\infty} \mathbb{E}\left[\inf_{k\geq n} X_k|\mathscr{G}\right] \leq \lim_{n\to\infty} \inf_{k\geq n} \mathbb{E}\left[X_k|\mathscr{G}\right]$$

$$Y - X_n \geq 0 \overset{\text{Fatou}}{\Longrightarrow} \mathbb{E}\left[\liminf_{n\to\infty} Y - X_n|\mathscr{G}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[Y - X_n|\mathscr{G}\right] \overset{Y\in L^1}{\Longrightarrow} \mathbb{E}\left[\liminf_{n\to\infty} -X_n|\mathscr{G}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[-X_n|\mathscr{G}\right].$$

(g) **DCT:** $X_n \xrightarrow{\text{a.s.}} X$, $|X_n| \leq V$ a.s. $\forall n$ & $\mathbb{E}[V] < \infty \implies \mathbb{E}\left[X_n|\mathscr{G}\right] \xrightarrow{\text{a.s.}} \mathbb{E}\left[X|\mathscr{G}\right]$

Consequence of two Fatou lemmas with lim sups and lim infs.

(h) **Jensen:** $c : \mathbb{R} \to \mathbb{R}$ convex: $\mathbb{E}\left[|c(X)|\right] < \infty \implies \mathbb{E}\left[c(X)|\mathscr{G}\right] \geq c\left(\mathbb{E}\left[X|\mathscr{G}\right]\right)$ a.s.

Let $S = \left\{(a,b) \in \mathbb{Q}^2 : ax + b \leq c(x), \forall x\right\} \Rightarrow s(x) = \sup_{(a,b)\in S} (ax+b)$.

$$\mathbb{E}\left[c(X)|\mathscr{G}\right] = \mathbb{E}\left[\sup_{(a,b)\in S} aX + b|\mathscr{G}\right] \geq a\mathbb{E}\left[X|\mathscr{G}\right] + b \xrightarrow[\text{RHS}]{\sup_{(a,b)}} \mathbb{E}\left[X|\mathscr{G}\right] \geq \sup_{(a,b)\in S} a\mathbb{E}\left[X|\mathscr{G}\right] + b = s(\mathbb{E}\left[X|\mathscr{G}\right]).$$

Corollary: $\|\mathbb{E}\left[X|\mathscr{G}\right]\|_p \leq \|X\|_p$, $\forall p \geq 1$.

Take $c(x) = |x|^p$, apply Jensen's & take $\mathbb{E}[\cdot]$: $\left(\mathbb{E}\left[|X|^p\right]\right)^{1/p} = \left(\mathbb{E}\left[\mathbb{E}\left[|X|^p|\mathscr{G}\right]\right]\right)^{1/p} \geq \left(\mathbb{E}\left[\left|\mathbb{E}\left[X|\mathscr{G}\right]\right|^p\right]\right)^{1/p}$

(i) **Tower Property:** $\mathcal{H} \subset \mathscr{G} \subset \mathscr{F} \implies \mathbb{E}\left[\mathbb{E}\left[X|\mathscr{G}\right]\Big|\mathcal{H}\right] = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{H}\right]\Big|\mathscr{G}\right] = \mathbb{E}\left[X|\mathcal{H}\right]$ a.s.

$\mathbb{E}\Big[\underbrace{\mathbb{E}\left[X|\mathscr{G}\right]}_{:=Z}\Big|\mathcal{H}\Big] = \underbrace{\mathbb{E}\left[X|\mathcal{H}\right]}_{:=Y}$ GOAL: $Y = \mathbb{E}\left[Z|\mathcal{H}\right]$.

✓ Y is $\mathcal{H}$-measurable.

✓ $\int_G Y\, d\mathbb{P} \overset{\forall G\in\mathcal{H}}{=\!=} \int_G X\, d\mathbb{P} \overset{\forall G\in\mathscr{G}}{=\!=} \int_G Z\, d\mathbb{P} \Rightarrow \int_G Y\, d\mathbb{P} = \int_G Z\, d\mathbb{P}$ $(\forall G \in \mathcal{H})$.

(j) **Factorization:** Z $\mathscr{G}$-meas & bounded $\implies \mathbb{E}\left[ZX|\mathscr{G}\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

$Y := \mathbb{E}\left[X|\mathscr{G}\right]$ GOAL: $YZ = \mathbb{E}\left[ZX|\mathscr{G}\right]$.

✓ YZ is $\mathscr{G}$-measurable (prod of two meas. fnc.)

✓ $X_n{:=}X\wedge n$ & $Y_n{:=}\mathbb{E}\left[X_n|\mathscr{G}\right]$ & $Z_k{=}Z\wedge k \xrightarrow[\text{OrthProj}]{Z\text{ bdd}} \mathbb{E}\left[(Y_n - X_n)Z\mathbb{1}_G\right] \overset{\text{DCT}:\,Y_n\text{ bdd}}{=\!=} \lim_{k\to\infty} \mathbb{E}\left[(Y_n - X_n)Z_k\mathbb{1}_G\right]{=}0$

$X_n Z\mathbb{1}_G \xrightarrow{\text{a.s.}} XZ\mathbb{1}_G$, and $|X_n Z\mathbb{1}_G| \leq |XZ| \in L^1 \overset{\text{DCT}}{\Longrightarrow} \mathbb{E}\left[XZ\mathbb{1}_G\right] = \mathbb{E}\left[YZ\mathbb{1}_G\right]$.

Monotone Class Proof:

$\mathcal{H} := \left\{Z : (\Omega, \mathscr{G}) \to (\mathbb{R}, \mathscr{B}) : Z \text{ is bounded and conclusion holds}\right\}$

✓ $Z = 1 \in \mathcal{H}$, as $Y = \mathbb{E}\left[X|\mathscr{G}\right] \Rightarrow \mathbb{E}\left[Y\mathbb{1}_G\right] = \mathbb{E}\left[X\mathbb{1}_G\right], \forall G \in \mathscr{G}$.

✓ $\mathcal{H}$ clearly is a vector space.

✓ $Z = \mathbb{1}_E$ $(E \in \mathscr{G}) \xrightarrow[\in\,\mathscr{G}]{G\cap E} \int_G YZ\, d\mathbb{P} = \int_{G\cap E} Y\, d\mathbb{P} = \int_{G\cap E} X\, d\mathbb{P} = \int_G XZ\, d\mathbb{P}$.

✓ $0 \leq Z_n \nearrow Z$, Z bdd, $Z_n \in \mathcal{H} \overset{\text{DCT}}{\Longrightarrow} Z \in \mathcal{H}$.

So **(j)** satisfied $\forall$ Z bdd & $\mathscr{G}$-meas. For any Z: truncate $Z_n = Z \wedge n$, and use DCT & $\mathbb{E}\left[|XZ|\right] < \infty$.

Corollary: $X \in L^p(\Omega, \mathscr{F}, \mathbb{P})$, $Z \in L^q(\Omega, \mathscr{F}, \mathbb{P})$, $p, q > 1$: $p^{-1} + q^{-1} = 1 \Rightarrow \mathbb{E}\left[ZX|\mathscr{G}\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

Corollary: $X \in (\mathrm{m}\mathscr{F})^+$, $Z \in (\mathrm{m}\mathscr{G})^+$: $\mathbb{E}[X], \mathbb{E}[ZX] < \infty \implies \mathbb{E}\left[ZX|\mathscr{G}\right] = Z\mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

For both, notice that above proof hold whenever $|XZ| \in L^1$. □

Proof: *(click)*

(k) Independence: $\mathcal{H} \perp\!\!\!\perp \sigma\left(\sigma(X), \mathscr{G}\right) \implies \mathbb{E}\left[X|\sigma(\mathscr{G}, \mathcal{H})\right] = \mathbb{E}\left[X|\mathscr{G}\right]$ a.s.

Let $Y := \mathbb{E}\left[X|\mathscr{G}\right]$. GOAL: $Y = \mathbb{E}\left[X|\sigma(\mathscr{G}, \mathcal{H})\right]$.

✓ Y is $\mathscr{G}$-meas. $\Rightarrow$ Y is $\sigma(\mathscr{G}, \mathcal{H})$ meas.

✓ Goal: $\mathbb{E}\left[X\mathbb{1}_E\right] = \mathbb{E}\left[Y\mathbb{1}_E\right]\ \forall\ E \in \sigma(\mathscr{G}, \mathcal{H})$. Idea: show on π-system $\mathcal{I} = \{G \cap H : G \in \mathscr{G}, H \in \mathcal{H}\}$

$\mathbb{E}\left[X\mathbb{1}_G\mathbb{1}_H\right] \stackrel{\perp\!\!\!\perp}{=} \mathbb{E}\left[X\mathbb{1}_G\right]\mathbb{E}\left[\mathbb{1}_H\right] \stackrel{Def}{=} \mathbb{E}\left[Y\mathbb{1}_G\right]\mathbb{E}\left[\mathbb{1}_H\right] \stackrel{\perp\!\!\!\perp}{=} \mathbb{E}\left[Y\mathbb{1}_G\mathbb{1}_H\right]$

Corollary: $X \perp\!\!\!\perp \mathcal{H} \implies \mathbb{E}\left[X|\mathcal{H}\right] = \mathbb{E}\left[X\right]$ a.s.

Take $\mathscr{G} = \{\varnothing, \Omega\}$, and recall $\mathbb{E}\left[X|\mathscr{G}\right] = \mathbb{E}\left[X\right]$. □

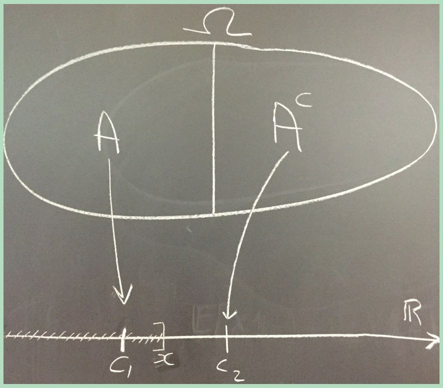
Exercise:

Let $\mathscr{G} = \{\varnothing, A, A^c, \Omega\}$. Compute $\mathbb{E}\left[X|\mathscr{G}\right]$, $\forall G \in \mathscr{G}$.

Generalization: $\{A_1, \dots, A_n\} \in \mathscr{G}$ a partition of Ω: $Y = c_1\mathbb{1}_{A_1} + \dots + c_n\mathbb{1}_{A_n}$ with $c_i = \dfrac{\mathbb{E}\left[X\mathbb{1}_{A_i}\right]}{\mathbb{P}\left(A_i\right)}$.

Solution: *(click)*

Notice Y can take at most two values. Indeed, if $\exists c_1, c_2, c_3 \in \mathbb{R}$, distinct s.t. $Y(\omega_i) = c_i$ $(i = 1, 2, 3)$ for some ω_i's, then, $Y^{-1}(\{c_1\}) \neq \varnothing, \Omega$. Only possibility is that $Y^{-1}(\{c_1\}) = A$. In this case, $Y^{-1}(\{c_2\}) = A^c$, but $\omega_3 \notin A$ & $\notin A^c$ $\boxed{\Rightarrow\Leftarrow}$



(i) Let $Y = \begin{cases} c_1, & \text{if } \omega \in A, \\ c_2, & \text{if } \omega \in A^c. \end{cases}$. WLOG assume $c_1 < c_2$.

Then: $Y^{-1}(-\infty, x] = \left\{\omega : Y(\omega) \leq x\right\} = \begin{cases} \varnothing & \text{if } x < c_1 \\ A & \text{if } c_1 \leq x < c_2 \\ \Omega & c_2 \leq x \end{cases} \implies Y^{-1}(-\infty, x] \in \mathscr{G},\ \forall x \in \mathbb{R}$

Since it's true over $\pi(Y) = \Big\{\left\{\omega : Y(\omega) \leq x\right\} : x \in \mathbb{R}\Big\}$, then, Y is $\mathscr{G}$-measurable. ✓

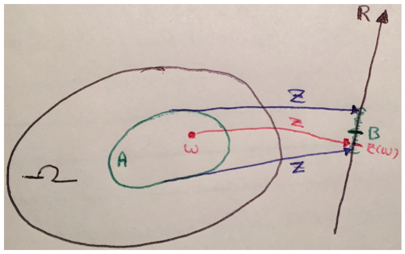
(ii) $\mathbb{E}\left[|Y|\right] = |c_1|\mathbb{P}\left(A\right) + |c_2|\mathbb{P}\left(A^c\right) \leq |c_1| + |c_2| < \infty$. ✓

(ii) $\displaystyle\int_A Y\, d\mathbb{P} = c_1\mathbb{P}(A) \stackrel{!}{=} \mathbb{E}\left[X\mathbb{1}_A\right] \Rightarrow \boxed{c_1 = \frac{\mathbb{E}\left[X\mathbb{1}_A\right]}{\mathbb{P}(A)}}$ is indeed a constant ✓,

& $\displaystyle\int_{A^c} Y\, d\mathbb{P} = c_2\mathbb{P}(A^c) \stackrel{!}{=} \mathbb{E}\left[X\mathbb{1}_{A^c}\right] \Rightarrow \boxed{c_2 = \frac{\mathbb{E}\left[X\mathbb{1}_{A^c}\right]}{\mathbb{P}(A^c)}}$ is indeed a constant ✓

5.3 Day 3: More on Conditional Expectation

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS: 

Intuition Behind $A \in \sigma(Z)$ $Z : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$: $A \in \sigma(Z) = \{Z^{-1}(B) : B \in \mathscr{B}\}$

MAIN THEOREMS:

Technical Point 1 $X \in L^1$, $\mathscr{G} \subset \mathscr{F}$, assume $\exists Y$ s.t. $Y \in m\mathscr{G}$ & $\int_G Y ; d\mathbb{P} = \int_G X \, d\mathbb{P}, \forall G \in \mathscr{G}$.

$\implies \mathbb{E}\left[|Y|\right] \leq \mathbb{E}\left[|X|\right] < \infty$.

Note This follows as conditional expectation reduces norms in L^p spaces ($p \geq 1$).

Proof: *(click)*

$$0 \leq \int_{Y>0} Y \, d\mathbb{P} = \int_{Y>0} X \, d\mathbb{P} = \left| \int_{Y>0} X \, d\mathbb{P} \right| \leq \int_{Y>0} |X| \, d\mathbb{P}$$

$$\& \ 0 \leq -\int_{Y\leq 0} Y \, d\mathbb{P} = -\int_{Y \leq 0} X \, d\mathbb{P} = \left| -\int_{Y\leq 0} X \, d\mathbb{P} \right| \leq \int_{Y \leq 0} |X| \, d\mathbb{P}.$$

$$\implies \mathbb{E}\left[|Y|\right] = \int_{Y>0} Y \, d\mathbb{P} + \int_{Y\leq 0} (-Y) \, d\mathbb{P} \leq \mathbb{E}\left[|X|\right] < \infty.$$ □

Technical Point 2 $X \geq 0$. Then $\exists Y$ S.T. $\int_G Y \, d\mathbb{P} = \int_G X \, d\mathbb{P}, \forall G \in \mathscr{G}$.

Proof: *(click)*

$X_n := X \wedge n$, $Y_n := \mathbb{E}\left[X_n | \mathscr{G}\right] \Rightarrow 0 \leq X_n \nearrow X \Rightarrow 0 \leq Y_n \nearrow$. Hence $Y := \lim_{n\to\infty} Y_n(\omega)$ exists (possibly ∞).

$$\int_G Y \, d\mathbb{P} \overset{MCT}{=\!=} \lim_{n\to\infty} \int_G Y_n \, d\mathbb{P} \overset{def}{=\!=} \lim_{n \to \infty} \int_G X_n \, d\mathbb{P} \overset{MCT}{=\!=} \int_G X \, d\mathbb{P}.$$ □

Agreement with Traditional Usage X,Z jointly continuous[3], h-Borel S.T. $\mathbb{E}\left[|h(X)|\right] < \infty$:

$$\implies g(Z) = \mathbb{E}\left[h(X)|\sigma(Z)\right], \text{ where } g(z) = \int_{\mathbb{R}} h(x) f_{x|z}(x|z)\, dx \ \&\ f_{x|z}(x|z) = \frac{f_{x,z}(x,z)}{f_Z(z)}$$

Proof: *(click)*

(i) $g(Z) \in m\sigma(Z)$. Will prove $g(\cdot)$ is Borel.

Claim: $\tilde{h}(x,z) = h(x)$ is $\mathscr{B} \times \mathscr{B}$ measurable.

$(\because)\tilde{h}^{-1}(B) = h^{-1}(B) \times \mathbb{R} \in \mathscr{B} \times \mathscr{B}$, $\forall B \in \mathscr{B}$. Hence, $h(x) f_{x|z}(x|z)$ is measurable with respect to $\mathscr{B} \times \mathscr{B}$.

$\Rightarrow \displaystyle\int h(x) f_{x|z}(x|z)\, dx$ is $\mathscr{B}$-measurable, hence **(i)** is proven.

(ii) Prove: $\displaystyle\int_A g(Z)\, d\mathbb{P} = \int_A h(X)\, d\mathbb{P}$, $\forall A \in \sigma(Z)$.

$A \in \sigma(Z) \iff \exists B \in \mathscr{B}$ S.T. $A = Z^{-1}B$. Hence, $\mathbb{1}_A(\omega) = \mathbb{1}_B(Z(\omega))$.

$$\int_A g(Z)\, d\mathbb{P} = \int_\Omega \mathbb{1}_B(Z(\omega)) g(Z(\omega))\, d\mathbb{P} = \int_R \mathbb{1}_B(z) g(z) f_z(z)\, dz.$$

$$\int_A h(X)\, d\mathbb{P} = \int \mathbb{1}_B(Z(\omega)) h(X(\omega))\, d\mathbb{P} = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_B(z) h(x)\, d\mathbb{P}_{X,Z} = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_B(z) h(x) f_{x|z}(x|z) f_z(z)\, dxdz$$

$$\xRightarrow{\text{Fubini}} = \int_{\mathbb{R}} \mathbb{1}_B(z) \underbrace{\left(\int_{\mathbb{R}} h(x) f_{x|z}(x|z) dx\right)}_{=g(z)} f_z(z) dz = \int_{\mathbb{R}} \mathbb{1}_B(z) g(z) f_z(z) dz. \qquad \square$$

Independence $\{X_i\}_{i=1}^n \perp\!\!\!\perp$, $h : (\mathbb{R}^n, \mathscr{B}^n) \to (\mathbb{R}, \mathscr{B})$ S.T. $\mathbb{E}\left[|h(X_1, \ldots, X_n)|\right] < \infty$.

Then $\mathbb{E}\left[h(X_1, \ldots, X_n)|X_1\right] = \varphi(X_1)$, for $\varphi(x) := \mathbb{E}\left[h(x, X_2, \ldots, X_n)\right]$.

Proof: *(click)*

Claim: $\varphi(\cdot)$ is $\mathscr{B}$-measurable, hence $\varphi(X_1)$ is $\sigma(X_1)$-measurable.

$$\varphi(x) = \int_\Omega h(x, X_2(\omega), \ldots, X_n(\omega))\, d\mathbb{P}(\omega) \underbrace{=\!=}_{\text{Change of Measure}} \int_{\mathbb{R}^{n-1}} h(x, x_2, \ldots, x_n)\, d\mathbb{P}_{X_2,\ldots,X_n}(x_2, \ldots, x_n).$$

$S_1 = (\mathbb{R}, \mathscr{B}, \mathbb{P}_{X_1})$, $S_2 = (\mathbb{R}^{n-1}, \mathscr{B}^{n-1}, \mathbb{P}_{X_2,\ldots,X_n})$. $h : (S_1 \times S_2) \to (\mathbb{R}, \mathscr{B}) \xRightarrow[\text{part } (iii)]{\text{“Meas. Checks”}} \varphi(x)$ is $\mathscr{B}$-function.

Claim: $\forall G \in \sigma(X_1)$, $\mathbb{E}\left[\varphi(X_1)\mathbb{1}_G\right] = \mathbb{E}\left[h(X_1, \ldots, X_n)\mathbb{1}_G\right]$.

$G \in \sigma(X_1) \iff G = X_1^{-1}(B)$. Hence, $\mathbb{1}_G(\omega) = \mathbb{1}_B(X_1(\omega))$.

$$\mathbb{E}\left[h(X_1, \ldots, X_n)\mathbb{1}_G\right] = \int_\Omega \mathbb{1}_G(\omega) h(X_1(\omega), \ldots, X_n(\omega))\, d\mathbb{P} = \int_\Omega \mathbb{1}_B(X_1(\omega)) h(X_1(\omega), \ldots, X_n(\omega))\, d\mathbb{P}.$$

$$= \int_{\mathbb{R}^n} \mathbb{1}_B(x_1) h(x_1, \ldots, x_n)\, d\mathbb{P}_{X_1,\ldots,X_n} = \int_{\mathbb{R}^n} \mathbb{1}_B(x_1) h(x_1, \ldots, x_n)\, d\mathbb{P}_{X_1} d\mathbb{P}_{X_2,\ldots,X_n}$$

$$\overset{Fubini}{=\!=} \int_{\mathbb{R}} \mathbb{1}_B(x_1) \left(\int_{\mathbb{R}^{n-1}} h(x_1, \ldots, x_n)\, d\mathbb{P}_{X_2,\ldots,X_n}\right) d\mathbb{P}_{X_1} = \int_{\mathbb{R}} \mathbb{1}_B(x_1) \varphi(x_1)\, d\mathbb{P}_{X_1}$$

$$\implies \mathbb{E}\left[h(X_1, \ldots, X_n)\mathbb{1}_G\right] = \mathbb{E}\left[\mathbb{1}_B(X_1)\varphi(X_1)\right] = \mathbb{E}\left[\mathbb{1}_G \varphi(X_1)\right]. \qquad \square$$

---

[3] i.e., $\exists\ f(x,z) : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ $\mathscr{B} \times \mathscr{B}$-measurable S.T. $\dfrac{d\mathbb{P}_{X,Z}}{d\lambda \times d\lambda}(x,z) = f(x,z)$.
So $\forall h$ Borel, we can compute $\mathbb{P}(h(X,Z) = \cdot)$ or $\mathbb{E}[h(X,Z)]$ using 2 changes of measures: $\mathbb{P} \to \mathbb{P}_{X,Z} \to f(x,z)dx\, dz$.

Existence of Cond. Probs Want to define a conditional probability $\mathbb{P}\left(X \in A|\mathscr{G}\right) := \mathbb{E}\left[\mathbb{1}_B(X)|\mathscr{G}\right]$.

Problem: $\mathbb{P}(\cdot|\mathscr{G})$ may not verify countable additivity $\implies$ $\mathbb{P}(\cdot|\mathscr{G})$ may not be a prob. meas.

$\star$Claim: Fix $\{F_n\}_{n=1}^{\infty}$ disjoint: $\exists\mathcal{N}$ null[4], S.T. $\mathbb{P}\left(\bigcup_{n=1}^{\infty} F_n|\mathscr{G}\right)(\omega) = \sum_{n=1}^{\infty}\mathbb{P}\left(F_n|\mathscr{G}\right)(\omega)\ \forall\ \omega\in\mathcal{N}^c$

$(\because) 0 \le \mathbb{1}_{\bigcup_{k=1}^{n} F_k}(X) \nearrow \mathbb{1}_{\bigcup_{k=1}^{\infty} F_k}(X)$ (used in MCT below)

$$\sum_{n=1}^{\infty}\mathbb{P}\left(F_n|\mathscr{G}\right)(\omega) \overset{Def}{=\!=} \sum_{k=1}^{\infty}\mathbb{E}\left[\mathbb{1}_{F_k}(X)|\mathscr{G}\right] \overset{Def}{=\!=} \lim_{n\to\infty}\sum_{k=1}^{n}\mathbb{E}\left[\mathbb{1}_{F_k}(X)|\mathscr{G}\right] \overset{Disj.}{=\!=} \lim_{n\to\infty}\mathbb{E}\left[\mathbb{1}_{\bigcup_{k=1}^{n} F_k}(X)|\mathscr{G}\right]$$

$$\overset{MCT}{=\!=} \mathbb{E}\left[\mathbb{1}_{\bigcup_{k=1}^{\infty} F_k}(X)\Big|\mathscr{G}\right] \overset{Def}{=\!=} \mathbb{P}\left(\bigcup_{n=1}^{\infty} F_n\Big|\mathscr{G}\right)(\omega) \text{ for } \omega \in \mathcal{N}^c. \qquad \square$$

$\star$ Let $\mathcal{A}$ be set of all countable collections of disjoint sets in $\mathscr{F}$: $\mathcal{A} := \left\{\alpha = \{F_n\}_{n=1}^{\infty} : F_n \in \mathscr{F} \text{ disj.}\right\}$.

Claim: $\mathbb{P}(\cdot|\mathscr{G})$ might not be count. additive for a.e. ω.

$(\because)$NEED: Verify count. add. for **every** countable collection of disjoint sets.

We have only proved that countable additivity holds over $\Omega \setminus \bigcup_{\alpha\in\mathcal{A}} \mathcal{N}_\alpha$.

Problem: may have $\mathbb{P}\Big(\bigcup_{\alpha\in\mathcal{A}} \mathcal{N}_\alpha\Big) \neq 0$ (an uncountable union of null sets may not be null) $\square$

Exercise:

Use of Symmetry Let $\{X_i\}_{i=1}^{n}$ iid. Then:

(i) $\mathbb{E}\left[\frac{X_i}{S_n}\right] = \frac{1}{n}\ \forall i = 1\dots n\ \ (\because) 1 = \mathbb{E}\left[\frac{S_n}{S_n}\right] = \mathbb{E}\left[\frac{X_1+\cdots+X_n}{S_n}\right] = \mathbb{E}\left[\frac{X_1}{S_n}\right] + \cdots + \mathbb{E}\left[\frac{X_n}{S_n}\right] = n\mathbb{E}\left[\frac{X_1}{S_n}\right]$

(ii) $\mathbb{E}\left[\frac{S_m}{S_n}\right] = \frac{m}{n}$, with $m \le n$ $(\because)\mathbb{E}\left[\frac{S_m}{S_n}\right] = \sum_{i=1}^{m}\mathbb{E}\left[\frac{X_i}{S_n}\right] = \frac{m}{n}$

(iii) $\mathbb{E}\left[X_i|S_n\right] = \frac{1}{n}S_n\ \forall i = 1\dots n\ \ (\because) S_n = \mathbb{E}\left[S_n|S_n\right] = \mathbb{E}\left[X_1+\cdots+X_n|S_n\right] = \sum_{i=1}^{n}\mathbb{E}\left[X_i|S_n\right] = n\mathbb{E}\left[X_1|S_n\right]$

(iv) $\mathbb{E}\left[S_m|S_n\right] = \frac{m}{n}S_n$, with $m \le n$ $(\because)\mathbb{E}\left[S_m|S_n\right] = \sum_{i=1}^{m}\mathbb{E}\left[X_i|S_n\right] = \frac{m}{n}S_n$

[4] i.e., $\mathbb{P}(\mathcal{N}) = 0$

Chapter 6

Week 6: Martingales

6.1 Day 1: Martingales

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Filtration Given $(\Omega, \mathscr{F})$, $\{\mathscr{F}_n\}_{n=1}^{\infty}$: $\mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \mathscr{F}_3 \subseteq \cdots \subseteq \mathscr{F}$ is a filtration. $F_\infty = \sigma\left(\bigcup_n \mathscr{F}_n\right)$.[1]

Natural Filtration Given $\{W_n\}_{n=1}^{\infty}$ stochastic process[2], natural filt. $\mathscr{F}_n = \sigma(W_1, \dots, W_n)$.

Adapted Process Process $X = \{X_n\}_{n=1}^{\infty}$ adapted to filtration $\{\mathscr{F}_n\}_{n=1}^{\infty}$ if X_n is $\mathscr{F}_n$-measurable.

> ***Note*** Usually, $X = f(W_1, \dots, W_n)$, $f : (\mathbb{R}^n, \mathscr{B}^n) \to (\mathbb{R}, \mathscr{B})$, and $\mathscr{F}_n$ nat. filt. of W.

Martingale Given a filtered space $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}, \mathbb{P})$: a proc. X is a martingale w.r.t. $(\mathscr{F}_n, \mathbb{P}) \Leftrightarrow$

- (**i**) X adapted: $X_n \in \mathrm{m}\mathscr{F}_n$
- (**ii**) $X_n \in L^1$: $\mathbb{E}\left[|X_n|\right] < \infty \ \forall n$
- (**iii**) $\mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] \begin{cases} \leq X_{n-1} \text{ a.s.} & \textbf{Supermartingale (Sup}\mathbb{MG}\textbf{)} \\ = X_{n-1} \text{ a.s.} & \textbf{Martingale (}\mathbb{MG}\textbf{)} \\ \geq X_{n-1} \text{ a.s.} & \textbf{Submartingale (Sub}\mathbb{MG}\textbf{)} \end{cases}$

> ***Note*** A process X is a martingale $\iff$ X is both a sub/supermartingale.
>
> ***Note*** X supermartingale $\iff$ $-X$ submartingale.
>
> ***Note*** $X_0 \in L^1$: X_n is a martingale $\Leftrightarrow \tilde{X}_n = X_n - X_0$ is martingale. So WLOG, set $X_0 = 0$.

Previsible Process Process $C = \{C_n\}_{n=1}^{\infty}$ is previsible $\Leftrightarrow C_n \in m\mathscr{F}_{n-1}$.

> ***Idea*** **Stake on game n:** C_n = bet you make at round n-often based on your history

[1] Alternate notation: $\mathscr{F}_\infty = \bigvee_{n=1}^{\infty} \mathscr{F}_n$.

[2] A stochastic process $W = \{W_n\}_{n=1}^{\infty}$ on $(\Omega, \mathscr{F})$ is a sequence of $\mathscr{F}$-measurable r.v.s indexed by $n \in \mathbb{N}$.

Gambling Strategy

⋆ **Bank Account:** X_n if bet $1^{\$}$ per game;

Y_n if bet $C_n^{\$}$ per game.

⋆ **Net Winnings For Unit Bets:** round n:

$X_n - X_{n-1}$ (bet $1^{\$}$ per round for each round).

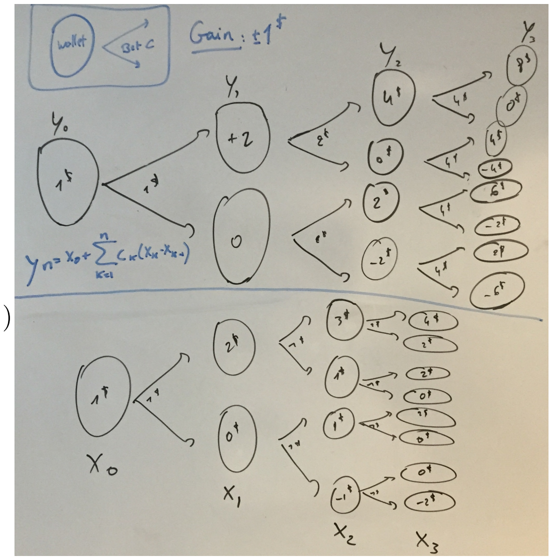
⋆ **Total Winnings up to Time n:**

$Y_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1}) =: (C \bullet X)_n$

Note $(C \bullet X)_0 = 0$ & $Y_n - Y_{n-1} = C_n(X_n - X_{n-1})$

⋆ **Martingale Transform of X by C:** $C \bullet X$

Continuous analog: stoch. integr. $\int C\, dX$



MAIN THEOREMS:

Martingale Property X a super/sub/martingale: then $\forall\ m < n$

$$\mathbb{E}\left[X_n|\mathscr{F}_m\right] \begin{cases} \leq X_m \text{ a.s.} & \textbf{Supermartingale (SupMG)} \\ = X_m \text{ a.s.} & \textbf{Martingale (MG)} \\ \geq X_m \text{ a.s.} & \textbf{Submartingale (SubMG)} \end{cases}$$

$$(\because)\mathbb{E}\left[X_n|\mathscr{F}_m\right] \overset{Tower}{=\!=} \mathbb{E}\left[\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]|\mathscr{F}_m\right] \overset{prop}{=\!=} \mathbb{E}\left[X_{n-1}|\mathscr{F}_m\right] = \cdots = \mathbb{E}\left[X_{m+1}|\mathscr{F}_m\right] = X_m$$

Martingale Transform Let C be a previsible process ($C_n \in \mathscr{F}_{n-1}$).

(i) $C \geq 0$, C bdd[3] X SupMG $\implies Y := C \bullet X$ SupMG (null @ 0).

(ii) C bdd, X MG $\implies Y := C \bullet X$ MG (null @ 0).

(iii) $X \in L^2$ & $C \in L^2 \implies (i)$ & (ii) still hold.

Proof: *(click)*

$$\mathbb{E}\left[Y_n - Y_{n-1}|\mathscr{F}_{n-1}\right] = \mathbb{E}\left[C_n(X_n - X_{n-1})|\mathscr{F}_{n-1}\right] \overset{C_n \in \mathscr{F}_{n-1}}{=\!=} C_n \mathbb{E}\left[X_n - X_{n-1}|\mathscr{F}_{n-1}\right] \begin{cases} = 0, & X \text{ MG}, \\ \leq 0, & X \text{ SupMG}, C \geq 0. \end{cases}.$$

Integrability $|Y_n| = \left|\sum_{k=1}^{n} C_k(X_k - X_{k-1})\right| \leq \sum_{k=1}^{n} |C_k||X_k - X_{k-1}| \in L^1$ if $\mathbb{E}\left[|C_k||X_k - X_{k-1}|\right] < \infty, \forall k$.

C_k bdd $(X_k \in L^1) \implies \mathbb{E}\left[|C_k||X_k - X_{k-1}|\right] < \infty$;

$C_k, X_k \in L^2 \xRightarrow[\text{Schwarz}]{\text{Cauchy}} \mathbb{E}\left[|C_k||X_k - X_{k-1}|\right] \leq \|C_k\|_2 \|X_k - X_{k-1}\|_2 < \infty.$ □

[3] $\exists K$ S.T. $C_n(\omega) \leq K, \forall n \in \mathbb{N},\ \forall \omega \in \Omega$

Exercise:

(i) $X_n \perp\!\!\!\perp \mathbb{E}[X_i] = 0$, $S_n = X_1 + \cdots + X_n$, $\mathscr{F}_n = \sigma(X_1, \ldots, X_n) \Longrightarrow$ S_n martingale w.r.t. $\mathscr{F}_n$.

Note Interesting Q : When does $\lim_{n\to\infty} S_n$ exist? [Kolmogorov 3-Series Theorem]

(ii) $X_n \geq 0 \perp\!\!\!\perp$, $\mathbb{E}[X_i] = 1$, $M_n = X_1 X_2 \ldots X_n$, $\mathscr{F}_n = \sigma(X_1, \ldots, X_n) \Longrightarrow$ M_n martingale

Note Interesting Q : $M_\infty := \lim_n M_n$ exists [Martingale Conv. Thm], when $\mathbb{E}[M_\infty] = 1$? [Kakutani's Thm]

(iii) Given $\xi \in L^1$, filt. $\mathscr{F}_n$, $M_n(\omega) = \mathbb{E}\left[\xi | \mathscr{F}_n\right](\omega) \implies M_n$ martingale w.r.t. $\mathscr{F}_n$.

$\implies$ This scenario is accumulating info about a r.v.

Note $M_n \to M_\infty = \mathbb{E}\left[\xi | \mathscr{F}_\infty\right]$ [Levy's Upward Thm]. When is $\xi = \mathbb{E}\left[\xi | \mathscr{F}_\infty\right]$? [Noisy Obs. of a r.v.]

Solution: *(click)*

(i) $\mathbb{E}\left[S_n | \mathscr{F}_{n-1}\right] = \mathbb{E}\left[S_{n-1} + X_n | \mathscr{F}_{n-1}\right] \overset{S_n \in m\mathscr{F}_{n-1}}{=\!=} S_{n-1} + \mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] \overset{X_n \perp\!\!\!\perp \mathscr{F}_{n-1}}{=\!=} S_{n-1} + \mathbb{E}[X_n] = S_{n-1}$.

(ii) $\mathbb{E}\left[M_n | \mathscr{F}_{n-1}\right] = \mathbb{E}\left[M_{n-1} X_n | \mathscr{F}_{n-1}\right] \overset{M_{n-1} \in m\mathscr{F}_{n-1}}{=\!=} M_{n-1} \mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] \overset{X_n \perp\!\!\!\perp \mathscr{F}_{n-1}}{=\!=} M_{n-1}$.

(iii) $\mathbb{E}\left[M_n | \mathscr{F}_{n-1}\right] = \mathbb{E}\left[\mathbb{E}\left[\xi | \mathscr{F}_n\right] \Big| \mathscr{F}_{n-1}\right] \overset{Tower}{=\!=} \mathbb{E}\left[\xi | \mathscr{F}_{n-1}\right] = M_{n-1}$.

6.2 Day 2: Doob's Optional Stopping Theorem

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Stopping Time $T : \Omega \to \mathbb{N} \cup \{\infty\}$ is a stopping time if $\{\omega : T(\omega) \leq n\} \in \mathscr{F}_n$, $\forall n \leq \infty$.

Note Equiv. to check $\{T = n\} \in \mathscr{F}_n$, $\forall n \leq \infty$.

$(\because)\{\omega : T(\omega) = n\}=\{\omega : T(\omega) \leq n\}\setminus\{\omega : T(\omega) \leq n-1\}\in\mathscr{F}_n$ & $\{\omega : T(\omega) \leq n\}=\bigcup_{k=0}^{n}\{\omega : T(\omega) = k\}\in\mathscr{F}_n$

Idea We stop game right after n^{th} round based on our info at the end of n^{th} round.

Example: $\{X_n\}_{n=1}^{\infty}$ adapted process, $B \in \mathscr{B}$.

$$T := \inf\{n \geq 0 : X_n \in B\} \implies T = \text{stopping time.}$$

Note T = first time X_n hits set B. $T = \infty \iff X_n$ never hits B.

$(\because)\{T \leq n\} = \bigcup_{k=0}^{n}\{\omega : X_k(\omega) \in B\} \in \mathscr{F}_n$, since $\{\omega : X_k(\omega) \in B\}$[4] $\in \mathscr{F}_k \subset \mathscr{F}_n$.

$$T = \infty \Rightarrow X_n \neq B,\ \forall n \implies \{T = \infty\} = \bigcap_{k=1}^{\infty} X_k^{-1}(B^c) \in \mathscr{F}_\infty.$$

Stopped Process $X^T := \{X_{T\wedge n}\}_{n=0}^{\infty}$, where T a stopping time, X a process.

MAIN THEOREMS:

Stopped $\mathbb{MG}$'s are $\mathbb{MG}$'s T stop. time, then $\forall n$:

$$\implies \text{stopped process } X_{T\wedge n} = \begin{cases} \text{Super } \mathbb{MG}\text{ , and } \mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0] & \text{if } X \text{ Super } \mathbb{MG} \\ \mathbb{MG}\text{ , and } \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] & \text{if } X\ \mathbb{MG} \\ \text{Sub } \mathbb{MG}\text{ , and } \mathbb{E}[X_{T\wedge n}] \geq \mathbb{E}[X_0] & \text{if } X \text{ Sub } \mathbb{MG} \end{cases}$$

Proof: *(click)*

Claim: $C_n(\omega) = \mathbb{1}_{T(\omega)\geq n} \implies C_n$ is previsible.

$(\because)C_n : \Omega \to \{0, 1\}$: need $C_n^{-1}(\{1\}) \in \mathscr{F}_{n-1}$ and $C_n^{-1}(\{0\}) = C_n^{-1}(\{1\}^c) \in \mathscr{F}_{n-1}$.

Indeed, $\{C_n = 1\} = \{T \geq n\} = \{T \leq n-1\}^c \in \mathscr{F}_{n-1}$.

Claim: $X_{T\wedge n} = (C \bullet X)_n + X_0 \xRightarrow[\text{Theorem}]{\text{Martingale Transform}} X_{T\wedge n}$ is super/sub/$\mathbb{MG}$

$(\because)(C \bullet X)_n(\omega) = \sum_{k=1}^{n} C_k(X_k - X_{k-1}) = \sum_{k=1}^{n} \mathbb{1}\{T \geq n\}(X_k - X_{k-1}) = \sum_{k=1}^{T\wedge n}(X_k - X_{k-1}) = X_{T\wedge n} - X_0.$ □

[4] $X_k : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B})$, $\mathscr{F}_k$ meas. (adapted proc.) $\implies X_k^{-1}(B) \in F_k$, $\forall B \in \mathscr{B}$, $\forall k \in \mathbb{N}$.

Careful! $\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0]\ \forall n \ \not\Rightarrow\ \mathbb{E}[X_T] = \mathbb{E}[X_0]$ in general

$(\because)$Random Walk over $\mathbb{Z}$: $T = \inf\{n \geq 0 : X_n = 1\}$

$\implies \mathbb{E}[X_{T\wedge n}] = 0 = \mathbb{E}[X_0]\ \forall n$ but $\mathbb{E}[X_T] = 1 \neq 0 = \mathbb{E}[X_0]$.

Note $\mathbb{P}(T < \infty) = 1$ but $\mathbb{E}[T] = \infty$ (later)

Idea Know $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$. Want to understand when $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$. Under some 'nice' conditions, $X_{T\wedge n} \xrightarrow{\text{a.s.}} X_T$ and $\mathbb{E}[\cdot]$'s converge, too.

Doob's Optional Stopping Thm (OPT) $T =$ stop. time, $X =$ SupMG /MG . If either holds:

(**i**) T bounded: $\exists B > 0$ s.t. $T(\omega) \leq B$ a.s. .

(**ii**) $T < \infty$ a.s. & $\exists Y \in L^1$ S.T. $|X_n(\omega)| \leq Y\ \forall(n,\omega)$

(**iii**) $\mathbb{E}[T] < \infty$ & $\exists B > 0$ S.T. $|X_{n+1}(\omega) - X_n(\omega)| \leq B\ \forall(n,\omega)$

(**iv**) $X \geq 0$ SupMG & $T < \infty$ a.s.

$\implies X_T$ integrable & $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$ or $\mathbb{E}[X_T] = \mathbb{E}[X_0]$

Proof: *(click)*

We know: $\mathbb{E}[|X_{T\wedge n}|] < \infty$ & $\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0]$ $(*)$ always holds for any X SupMG & T stop. time.

(**i**) Take $n \geq B \implies \mathbb{E}[X_T] = \mathbb{E}\Big[X_{\underbrace{T\wedge n}_{=T}}\Big] \leq \mathbb{E}[X_0]$.

(**ii**) $T < \infty$ a.s. $\implies X_{T\wedge n} \xrightarrow{T(\omega)\wedge n \xrightarrow{\text{a.s.}} T(\omega)} X_{T(\omega)}(\omega)$ a.s. , $|X_n| \leq Y \overset{\text{DCT}}{\implies} \mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.

(**iii**) $X_{T\wedge n} - X_0 = \sum_{k=1}^{T\wedge n} X_k - X_{k-1} \implies |X_{T\wedge n} - X_0| \leq \sum_{k=1}^{T\wedge n} |X_k - X_{k-1}| \leq B(T\wedge n) \leq BT \overset{\text{DCT}}{\implies} \mathbb{E}[X_T] \leq \mathbb{E}[X_0]$

(**iv**) $Y_n \overset{X_n\geq 0}{:=} X_{T\wedge n} \geq 0$, $Y_n \xrightarrow{\text{a.s.}} X_T \overset{\text{Fatou}}{\implies} \mathbb{E}[X_T] = \mathbb{E}\left[\lim_{n\to\infty} Y_n\right] \leq \liminf_{n\to\infty} \mathbb{E}[X_{T\wedge n}] \leq^{(*)} \mathbb{E}[X_0]$. $\square$

Stopping Time Calculus T_1, T_2 stop. times $\implies T_1 \wedge T_2$, $T_1 \vee T_2$, $T_1 + T_2$ are stop. times.

$(\because)\{T_1 \wedge T_2 \leq n\} = \{T_1 \leq n\} \cup \{T_1 \leq n\} \in \mathscr{F}_n$

$(\because)\{T_1 + T_2 = n\} = \bigcup_{k=0}^{n} \{T_1 = k\} \cap \{T_2 = n-k\} \in \mathscr{F}_n$.

Corollary: T stop. time $\implies T_M = T \wedge M$ stop time, for any $M \in \mathbb{N}$.

Note If $T < \infty$ a.s. , $T_M \xrightarrow{\text{a.s.}} T$ & T_M bdd.

Stopped σ-algebra $\tau =$ stopping time: $\mathscr{F}_\tau := \Big\{A \in \mathscr{F}_\infty : A \cap \{\omega : \tau(\omega) \leq n\} \in \mathscr{F}_n, \forall n\Big\}$

Note $\mathscr{F}_\tau = \sigma$-field & $X_\tau \in \mathrm{m}\mathscr{F}_\tau$

Truncated Martingale Prop X_n SupMG , $k \in \mathbb{R} \implies Y_n := X_n \wedge k$ is SupMG .

Proof: *(click)*

$$\mathbb{E}\left[Y_n|\mathscr{F}_{n-1}\right]=\mathbb{E}\left[X_n \wedge k|\mathscr{F}_{n-1}\right]\leq\min\left\{\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right],\mathbb{E}\left[k|\mathscr{F}_{n-1}\right]\right\}=\min\left\{\underbrace{\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]}_{\leq X_{n-1}},k\right\}\leq Y_{n-1}$$ □

Alt.Proof of Doob OPT (iv) indexMartingales!Doob!Optional Stopping Thm $X{\geq}0$ SupMG &

$T{<}\infty$ a.s. $\implies X_T{\in}L^1$ & $\mathbb{E}\left[X_T\right]{\leq}\mathbb{E}\left[X_0\right]$

Proof: *(click)*

X_n SupMG , $k \in \mathbb{R} \xrightarrow[\mathbb{MG}]{\text{Truncated}} Y_n := X_n \wedge k$ is SupMG .

Now, $T_M = T \wedge M$ stop. time $\implies Y_n = X_{T_M \wedge n}$ SupMG $\xrightarrow{\text{Lemma}} Z_n = Y_n \wedge k = X_{T_M \wedge n} \wedge k$ SupMG .

$\mathbb{E}\left[Z_n\right] \leq \mathbb{E}\left[Z_0\right] = \mathbb{E}\left[X_0 \wedge k\right] \implies \mathbb{E}\left[X_{T_M \wedge n} \wedge k\right] \leq \mathbb{E}\left[X_0 \wedge k\right] \xrightarrow{n \geq M} \mathbb{E}\left[X_{T_M} \wedge k\right] \leq \mathbb{E}\left[X_0\right]$.

$X_{T_M} \wedge k \xrightarrow{\text{a.s.}} X_T \wedge k$ (as $M \to \infty$) & $|X_{T_M} \wedge k| \leq k \xrightarrow{\text{DCT}} \mathbb{E}\left[X_T \wedge k\right] \leq \mathbb{E}\left[X_0 \wedge k\right]$.

$0 \leq X_T \wedge k \nearrow X_T$ (as $k \to \infty$) $\xrightarrow{\text{MCT}} \mathbb{E}\left[X_T\right] \leq \mathbb{E}\left[X_0\right]$. □

Two Stopping Times T, S two stop. times S.T. $S \leq T \implies \mathbb{E}\left[X_{T\wedge n}\right] \leq \mathbb{E}\left[X_{S \wedge n}\right]$.

Corollary: T & S **bounded**, $S \leq T$ a.s. $\implies \mathbb{E}\left[X_T\right] \leq \mathbb{E}\left[X_S\right]$.

Proof: *(click)*

Claim: $C_n := \mathbb{1}_{(S,T]}(n,\omega) = \mathbb{1}\left\{S(\omega) < n \leq T(\omega)\right\} \stackrel{\text{a.s.}}{=} \mathbb{1}\left\{T \geq n\right\} - \mathbb{1}\left\{S \geq n\right\}$ is previsible.

$(\because) C_n : \Omega \to \{0,1\} \implies$ check $C_n^{-1}(\{1\}) = \{S < n\} \cap \{n \leq T\} = \{S \leq n-1\} \cap \{T \leq n-1\}^c \in \mathscr{F}_{n-1}$.

Claim: $X_{T\wedge n} - X_{S\wedge n}$ is a SupMG $\implies \mathbb{E}\left[X_{T\wedge n} - X_{S\wedge n}\right] \leq 0$.

$(\because) X_{T\wedge n} - X_{S\wedge n} = (C \bullet X)_n$ is SupMG as C_n previsible.

Note Corollary follows from Doob's OPT (i). □

OPT Corollary M a MG : C previsible, T a stopping time, S.T.

$\mathbb{E}[T] < \infty$, C **bounded**, and $|M_n - M_{n-1}|$ **bounded** $\implies \mathbb{E}\left[(C \bullet M)_T\right] = 0$

Proof: *(click)*

Let $Y_n := (C \bullet M)_n = \displaystyle\sum_{k=1}^{n} C_k(X_k - X_{k-1})$. **WANT:** $\mathbb{E}\left[Y_T\right] = 0$. $\quad \mathbb{E}\left[Y_{T\wedge n}\right] = \mathbb{E}\left[Y_0\right] = 0$

$$Y_{T\wedge n} = \sum_{k=1}^{T\wedge n} C_k(X_k - X_{k-1}) \Longrightarrow |Y_{T\wedge n}| \leq \sum_{k=1}^{T\wedge n} \underbrace{|C_k|}_{bdd} \cdot \underbrace{|(X_k - X_{k-1})|}_{bdd} \leq B \cdot T \ (\exists B > 0 \text{ bound})$$

$\xrightarrow[\text{OPT}]{\text{Doob}} \mathbb{E}\left[Y_T\right] = \mathbb{E}\left[Y_0\right] = 0$. □

Finite $\mathbb{E}[\mathbf{T}]$: T =stop. time: $\exists N, \varepsilon > 0$ S.T. $\mathbb{P}\left(T \leq n + N | \mathscr{F}_n\right) > \varepsilon$ a.s. $\forall n \implies \mathbb{E}[T] < \infty$

Idea Use it to prove $\mathbb{E}[T] < \infty$ for Doob's OPT.

Proof: *(click)*

$$\mathbb{P}(T > kN) = \mathbb{P}\left(T > kN \cap T > (k-1)N\right) = \underbrace{\mathbb{P}\left(T > kN | T > (k-1)N\right)}_{=\mathbb{P}\left(T > kN | \mathscr{F}_{(k-1)N}\right)} \mathbb{P}\left(T > (k-1)N\right)$$

$$\mathbb{P}(T > kN) \leq (1-\varepsilon)\mathbb{P}\left(T > (k-1)N\right) \underbrace{\leq \cdots \leq}_{induction} (1-\varepsilon)^{k-1}\mathbb{P}(T > N) = (1-\varepsilon)^{k-1} \underbrace{\mathbb{P}\left(T > N | \mathscr{F}_0 = \{\varnothing, \Omega\}\right)}_{\leq (1-\varepsilon)}.$$

$$\mathbb{P}\left(\frac{T}{N} > k\right) \leq (1-\varepsilon)^k \xrightarrow[T \geq 0]{\text{Trick}} \mathbb{E}\left[\frac{T}{N}\right] = \sum_{k=0}^{\infty} \mathbb{P}\left(\frac{T}{N} > k\right) \leq \sum_{k=0}^{\infty} (1-\varepsilon)^k < \infty \xrightarrow[\text{constant}]{\text{N finite}} \mathbb{E}[T] < \infty \qquad \square$$

Exercise:

$\{X_n\}_{n=1}^{\infty}$iid$\sim X=\begin{cases} +1, & w.p.1/2, \\ -1, & w.p.1/2. \end{cases}$ $S_n = \sum_{k=1}^{\infty} X_k$ (with $S_0 = 0$), $\mathscr{F}_n = \sigma(X_1, \ldots, X_n) = \sigma(S_0, S_1, \ldots, S_n)$.

$T = \inf\{n \geq 0 : S_n = 1\} \implies \mathbb{P}(T = \infty) = 0$ while $\mathbb{E}[T] = \infty$.

Solution: *(click)*

Let $M_n^\theta = (\text{sech } \theta)^n e^{\theta S_n}$, where $\text{sech } \theta = \dfrac{1}{\cosh\theta} = \dfrac{2}{e^\theta + e^{-\theta}} = \dfrac{1}{\mathbb{E}\left[e^{\theta X_n}\right]}$.

Claim: M_n^θ is a $\mathbb{MG}$.

$(\because) M_n^\theta = \prod_{i=1}^{n} \tilde{X}_i$, where $\tilde{X}_i = \dfrac{e^{\theta X_i}}{\mathbb{E}\left[e^{\theta X}\right]}$ unit mean, indep & use Product Martingale Exercise.

Claim: $\mathbb{E}\left[M_{T\wedge n}^\theta\right] = 1$

$(\because) M_{T\wedge n}$ is $\mathbb{MG} \implies \mathbb{E}\left[M_{T\wedge n}^\theta\right] = \mathbb{E}\left[M_{T\wedge 0}^\theta\right] = 1$

Claim: $\mathbb{E}\left[M_T^\theta\right] = 1 = \mathbb{E}\left[(\text{sech } \theta)^T e^\theta\right]$.

$(\because) \theta > 0 \Rightarrow e^{\theta S_{T\wedge n}} \leq e^\theta \xRightarrow{\text{sech } \theta < 1} M_{T\wedge n}^\theta \leq e^\theta$

Claim: $\mathbb{E}\left[(\text{sech } \theta)^T\right] = e^{-\theta}$ (for $\theta > 0$)

$(\because) 1 = \mathbb{E}\left[(\text{sech } \theta)^T e^\theta\right] = e^\theta \mathbb{E}\left[(\text{sech } \theta)^T\right]$

Claim: $\mathbb{P}(T < \infty) = 1$.

$(\because) (\text{sech } \theta)^T \xrightarrow[\theta\searrow 0]{} \mathbb{1}\{T < \infty\}$ and $\xRightarrow{\text{BCT}} 1 = \mathbb{E}\left[\mathbb{1}\{T < \infty\}\right] = \mathbb{P}(T < \infty) = 1$

Claim: $\mathbb{E}[T] = \infty$.

$(\because) 1 = \mathbb{E}[S_T] = \mathbb{E}\Big[\mathbb{E}\left[S_T|\sigma(T)\right]\Big] = \mathbb{E}\left[T\mathbb{E}[X_i]\right] = \underbrace{\mathbb{E}[X_i]}_{=0}\mathbb{E}[T]$. $\mathbb{E}[T] < \infty \Rightarrow$ contradiction!

$(\because)$**[Wald's Identity]** $\mathbb{E}\left[\sum_{i=1}^T X_i|\sigma(T)\right] = T\mathbb{E}[X_i]$

Claim: $\alpha = \text{sech } \theta \implies \mathbb{E}\left[\alpha^T\right] = \sum_{n=1}^{\infty} \alpha^n \mathbb{P}(T = n) \overset{\mathbb{E}\left[(\text{sech } \theta)^T\right]=e^{-\theta}}{=\!=} e^{-\theta} \overset{\text{Claim!}}{=\!=} \dfrac{1 - \sqrt{1-\alpha^2}}{\alpha}$

Corollary: $\mathbb{P}(T = 2m-1) = (-1)^{m+1}\dbinom{1/2}{m}$

$(\because)$Recursion: from $X_0 = 0$ go to $X_1 = \pm 1$ w.p. $\dfrac{1}{2}$

$\implies f(\alpha) := \mathbb{E}\left[\alpha^T\right] = \dfrac{1}{2}\underbrace{\mathbb{E}\left[\alpha^T|X_1 = 1\right]}_{\text{so } T=1} + \dfrac{1}{2}\mathbb{E}\left[\alpha^T|X_1 = -1\right] = \dfrac{1}{2}\alpha + \dfrac{1}{2}\mathbb{E}\left[\alpha^T|X_1 = -1\right]$

Now let $T_1 = \inf\left\{n \geq 0 : S_n = 0|X_0 = -1\right\}$ & $T_2 = \inf\left\{n \geq 0 : S_n = 1|X_0 = 0\right\}$:

Then $\mathbb{E}\left[\alpha^T|X_1 = -1\right] = \mathbb{E}\left[\alpha^{1+T_1+T_2}|X_1 = -1\right] = \alpha\mathbb{E}\left[\alpha^{T_1}|X_1 = -1\right]\mathbb{E}\left[\alpha^{T_2}|X_1 = 0\right]$

$\implies f(\alpha) = \frac{1}{2}\alpha + \frac{1}{2}\alpha f(\alpha)^2 \implies f(\alpha) = \dfrac{1 \pm \sqrt{1-\alpha^2}}{\alpha}$ but the "+" cannot hold:

$(\because) 1 + \sqrt{1-\alpha^2} \geq 1$ & $\frac{1}{\alpha} \geq 1 \implies \dfrac{1 + \sqrt{1-\alpha^2}}{\alpha} \geq 1$, but $f(\alpha) = e^{-\theta} \xRightarrow{\theta > 0} f(\alpha) \leq 1$.

6.3 Day 3: Martingale Convergence Theorem

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

$\{X_n\}$ **Bounded in $\mathbf{L^1}$** $\sup_{n\geq 1} \mathbb{E}\left[|X_n|\right] < \infty \iff \exists B > 0$ S.T. $\mathbb{E}\left[|X_n|\right] < B$

Number of Upcrossing $U_N[a,b](\omega) = k$ if $\exists s_1, \ldots, s_k$ & $t_1, \ldots, t_k$ S.T. :

$0 \leq s_1 < t_1 < \cdots < s_k < t_k \leq N$ & $X_{s_i} < a$, $X_{t_i} > b$, $\forall i \in \{1, \ldots, k\}$.

MAIN THEOREMS:

Sup$\mathbb{MG}$ Boundedness $\{X_n\}$ Sup$\mathbb{MG}$ is bounded in $L^1 \iff X_n^-$ is bounded in L^1.

Note For Sub$\mathbb{MG}$: $\iff X_n^+$ is bounded in L^1.

Corollary: $\{X_n\}$ Sup$\mathbb{MG}$: $X_n \geq 0 \implies$ bdd in L^1 $\quad (\because)\mathbb{E}\left[|X_n|\right] = \mathbb{E}\left[X_n\right] \leq \mathbb{E}\left[X_0\right]$.

Proof: *(click)*

$\mathbb{E}\left[X_n\right] + \mathbb{E}\left[X_n^-\right] = \mathbb{E}\left[X_n^+\right] \implies \mathbb{E}\left[|X_n|\right] = \mathbb{E}\left[X_n^+\right] + \mathbb{E}\left[X_n^-\right] = 2\mathbb{E}\left[X_n^-\right] + \mathbb{E}\left[X_n\right] \leq 2\mathbb{E}\left[X_n^-\right] + \mathbb{E}\left[X_0\right]$

since Sup$\mathbb{MG}$ & $\mathbb{E}\left[|X_n|\right] \geq \mathbb{E}\left[X_n^-\right]$. $\square$

Useful Ineq. Let X Sup$\mathbb{MG}$ & C previsible: $\begin{cases} C_1 = \mathbb{1}\{X_0 < a\} \\ C_n = \mathbb{1}\{X_{n-1} < a\}\mathbb{1}\{C_{n-1} = 0\} + \mathbb{1}\{X_0 \leq b\}\mathbb{1}\{C_{n-1} = 1\} \end{cases}$

$\implies Y_n := (C \bullet X)_n$ Sup$\mathbb{MG}$ & $Y_N(\omega) \geq (b-a)U_N[a,b](\omega) - (X_N - a)^-$.

Note C corresponds to betting once hit below a, until hit b

Proof: *(click)*

Let $U_N[a,b](\omega) = k$. Define $\begin{cases} s_1 = \inf\{n \geq 1 : X_n < a\} \\ t_1 = \inf\{n \geq s_1 : X_n > b\} \end{cases}$ & $\begin{cases} s_i = \inf\{n \geq t_{i-1} : X_n < a\} \\ t_i = \inf\{n \geq s_i : X_n > b\} \end{cases}$

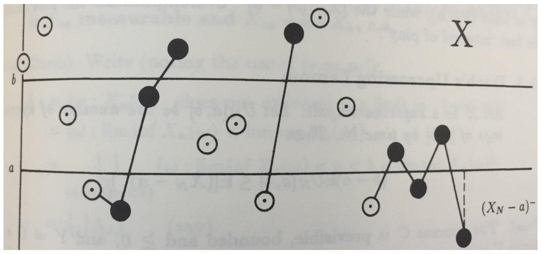
$\implies 1 \leq s_1 < t_1 < \cdots < s_k < t_k \leq N\ \forall i \in \{1, \ldots, k\}$ & $C_j = 1\ \forall s_i + 1 \leq j \leq t_i$.

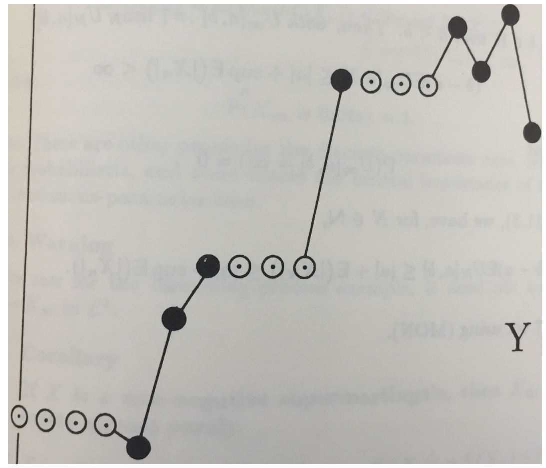
Also, let $s_{k+1} = \inf\{n \geq t_k : X_n \leq a\} \wedge N$

$$Y_N = \sum_{t=1}^{N} C_t(X_t - X_{t-1}) = \sum_{i=1}^{k} \underbrace{\sum_{j=s_i+1}^{t_i} (X_j - X_{j-1})}_{=X_{t_i} - X_{s_i} \geq (b-a)} + \sum_{j=s_{k+1}+1}^{N} (X_j - X_{j-1})$$

$$\implies Y_N \geq (b-a)k + \sum_{j=s_{k+1}+1}^{N} (X_j - X_{j-1}) = (b-a)k + X_N - X_{s_{k+1}} \geq (b-a)k - (X_N - a)^-. \quad \square$$

$C = 1$ for $\bullet$ & $C = 0$ for $\circ$.





Doob's Upcrossing Lemma X_n SupMG $\implies (b-a)\mathbb{E}\left[U_N[a,b]\right] \le \mathbb{E}\left[(X_N - a)^-\right]$.

Corollary: X_n SupMG bdd in L^1, $U_\infty[a,b](\omega) = \Big\{$number of upcrossings in $\left\{X_n(\omega)\right\}_{n=1}^{\infty}\Big\}$:

$\implies (b-a)\mathbb{E}\left[U_\infty[a,b]\right] \le \mathbb{E}\left[(X_N - a)^-\right] \le |a| + \sup\limits_{n\ge 1} \mathbb{E}\left[|X_n|\right]$

$\implies \mathbb{P}\left(\omega : U_\infty[a,b](\omega) = \infty\right) = 0$

Proof: *(click)*

Take expectation in Y_n (defined above) and use $\mathbb{E}\left[Y_n\right] \le \mathbb{E}\left[Y_0\right] = 0$.

For corollary, observe: $0 \le U_N[a,b](\omega) \nearrow U_\infty[a,b](\omega)$ & use MCT. $\square$

Doob's Forward Conv. Thm X_n SupMG bounded in L^1 $(\sup\limits_n \mathbb{E}\left[X_n^-\right] < \infty)$

$\implies X_\infty := \lim\limits_{n\to\infty} X_n$ exists, $X_n \xrightarrow{\text{a.s.}} X_\infty$ & $\mathbb{P}\left(X_\infty < \infty\right) = 1$.

Note $X_n \in m\mathscr{F}_n \subseteq \mathscr{F}_\infty \implies X_n \in m\mathscr{F}_\infty \implies X_\infty \in m\mathscr{F}_\infty$.

Corollary: $X_n \ge 0$ SupMG $\implies \lim_{n\to\infty} X_n = X_\infty$ exists a.s.

Proof: *(click)*

Let $A := \left\{\omega \in \Omega : \lim\limits_{n\to\infty} X_n(\omega) \text{ doesn't exist in } \bar{\mathbb{R}}\right\} = \left\{\omega \in \Omega : \liminf\limits_{n\to\infty} X_n(\omega) < \limsup\limits_{n\to\infty} X_n(\omega)\right\}$

$$\overset{\text{Trick:}}{=} \bigcup_{(a,b)\in\mathbb{Q}^2} \underbrace{\left\{\liminf_{n\to\infty} X_n(\omega) < a < b < \limsup_{n\to\infty} X_n(\omega)\right\}}_{:=A_{a,b}}.$$

$$\left.\begin{array}{l} \liminf < a \Rightarrow X_n < a \text{ i.o.} \\ \limsup > b \Rightarrow X_n > b \text{ i.o.} \end{array}\right\} \implies A_{a,b} \subseteq \left\{\omega \in \Omega : U_\infty[a,b](\omega) = \infty\right\}$$

$$\xrightarrow[\text{Corollary}]{\text{Upcrossing}} \mathbb{P}\left(A_{a,b}\right) = 0 \implies \mathbb{P}\left(A\right) = \mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \text{ doesn't exist in } \bar{\mathbb{R}}\right\}\right) = 0$$

$\implies X_\infty = \lim\limits_{n\to\infty} X_n(\omega)$ exists.

$$\mathbb{E}\left[|X_\infty|\right] = \mathbb{E}\left[\lim_{n\to\infty} |X_n|\right] = \mathbb{E}\left[\liminf_{n\to\infty} |X_n|\right] \underbrace{\le}_{\text{FATOU}} \liminf_{n\to\infty} \mathbb{E}\left[|X_n|\right] \le \sup_n \mathbb{E}\left[|X_n|\right] < \infty. \quad \square$$

Careful! $X_n \xrightarrow{\text{a.s.}} X_\infty \not\Longrightarrow X_n \xrightarrow{L^1} X_\infty$

Exercise:

S_n =Simple Symmetric Random Walk: $S_0 = 1$, $S_n = S_{n-1} + X_n$ (X_n iid).

Show $T := \inf\{n \geq 1 : S_n = 0\} \implies S_{T\wedge n} \xrightarrow{\text{a.s.}} S_\infty$ but not in L^1.

Solution: *(click)*

$Y_n := S_{T\wedge n} \geq 0 \xRightarrow[\text{Forward}]{\text{Doob}} Y_n \xrightarrow{\text{a.s.}} Y_\infty$ & $\mathbb{E}[Y_n] = \mathbb{E}[S_{T\wedge n}] \overset{\mathbb{MG}}{=\!=} \mathbb{E}[S_0] = 1$.

Fix $\omega \in \Omega$. $Y_n(\omega)$ conv. & integer $\implies \exists N \in \mathbb{N}$: $Y_n(\omega)$ fixed for $n \geq N$.

If fixed $\neq 0 \implies Y_{n+1} = Y_n \pm 1$ not fixed!

Exercise:

$X_n \geq 0$ iid, $\mathbb{E}[X_n] = 1$, $\mathbb{P}(X_n = 1) < 1$ & $M_n = X_1 \cdots X_n$.

$\implies M_n \xrightarrow{\text{a.s.}} 0$ & $\dfrac{1}{n}\log M_n \xrightarrow{\text{a.s.}} c < 0$.

Solution: *(click)*

M_n product $\mathbb{MG}$, $M_n \geq 0 \xRightarrow[\text{CT}]{\mathbb{MG}} M_n \xrightarrow{\text{a.s.}} M_\infty$. **<u>WANT:</u>** $M_\infty = 0$, a.s.

Suppose $\exists F \subset \Omega$ S.T. $\mathbb{P}(F) > 0$ & $M_\infty \neq 0$ on $F \Rightarrow \forall \omega \in F$: $X_n(\omega) = \dfrac{M_n(\omega)}{M_{n-1}(\omega)} \xrightarrow{\text{a.s.}} 1$.

$\xRightarrow[\text{0-1 Law}]{\text{Kolmogorov}} \mathbb{P}\left(\omega : \lim_{n\to\infty} X_n \text{ exists}\right) \in \{0,1\}$ & $\mathbb{P}\left(\omega : \lim_{n\to\infty} X_n \text{ exists}\right) \geq \mathbb{P}(F) > 0$

$\implies \mathbb{P}\left(\omega : \lim_{n\to\infty} X_n \text{ exists}\right) = 1$. Also, $\lim_{n\to\infty} X_n \in m\tau$ (tail) $\implies$ constant a.s. $\implies \lim_{n\to\infty} X_n = 1$.

$\mathbb{P}\left(|X_n - 1| \leq \dfrac{1}{k}\right) \xrightarrow{k\to\infty} \mathbb{P}\left(|X_n - 1| = 0\right) = 0$, hence $\exists k$: $\mathbb{P}\left(|X_n - 1| \leq \dfrac{1}{k}\right) < 1$.

Let $\delta = \dfrac{1}{k}$, and $\mathbb{P}\left(|X_n - 1| \leq \delta\right) < 1$ can't conv to 1, contradiction!

$$\frac{1}{n}\log M_n = \frac{1}{n}\sum_{k=1}^{n} \log X_k \xrightarrow[\text{a.s.}]{SLLN} \mathbb{E}[\log X] \underbrace{<}_{\text{Jensen}} \log \mathbb{E}[X] = 0.$$

Chapter 7

Week 7: L^2 Martingales & Uniform Integrability

7.1 Day 1: Martingales Bounded in L^2

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Bounded in L^2 X_n Sup/Sub/$\mathbb{MG}$ bdd in L^2 $\iff \sup_n \mathbb{E}[|X_n|^2] < \infty$

Note bdd in $L^2 \xRightarrow{\|\cdot\|_1 \le \|\cdot\|_2}$ bdd in L^1.

Oscillation ω: Sequence a_n: $\omega_{a_n} = \limsup_n a_n - \liminf_n a_n$.

Function f (at pt x_0): $\omega_f(x_0) = \lim_{\varepsilon \to 0} \left(\sup_{y \in B(x_0,\varepsilon)} f(y) - \inf_{y \in B(x_0,\varepsilon)} f(y) \right) = \inf_{\varepsilon > 0} \left(\sup_{y \in B(x_0,\varepsilon)} f(y) - \inf_{y \in B(x_0,\varepsilon)} f(y) \right)$

Note $\omega < \infty$: **finite** oscillation ; $\omega < \infty$: **infinite** oscillation ($\infty - \infty$ case undefined)

MAIN THEOREMS:

Orth. of Increments $M_n \in L^2$ $\mathbb{MG}$, $a \le b \le c \le d \in \mathbb{N} \implies \langle M_d - M_c, M_b - M_a \rangle = 0.$

Corollary: $M_n = M_0 + \sum_{k=1}^{n} \underbrace{(M_k - M_{k-1})}_{\text{orth. terms}} \xRightarrow{\text{Pythagoras}} \mathbb{E}\left[M_n^2\right] = \mathbb{E}\left[M_0^2\right] + \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2\right]$

Proof: *(click)*

$\mathbb{E}\left[M_d - M_c | \mathscr{F}_b\right] = \mathbb{E}\left[M_d | \mathscr{F}_b\right] - \mathbb{E}\left[M_c | \mathscr{F}_b\right] = M_b - M_b = 0$ & $M_b - M_a \in L^2(\Omega, \mathscr{F}_b, \mathbb{P})$

$\xRightarrow{\text{Orth. Proj.}} M_d - M_c - \mathbb{E}\left[M_d - M_c | \mathscr{F}_b\right] \perp (M_b - M_a) \implies \langle M_d - M_c, M_b - M_a \rangle = 0.$ $\square$

L^2 **Boundedness** $M_n \in L^2$ $\mathbb{MG}$: $\quad M$ bdd in $L^2 \iff \sum_{k=1}^{\infty} \mathbb{E}\left[(M_k - M_{k-1})^2\right] < \infty.$

L^2 **Convergence** $M_n \in L^2$ $\mathbb{MG}$: $\quad M$ bdd in $L^2 \implies M_n \to M_\infty$ both a.s. & in L^2.

Proof: *(click)*

M_n bdd $L^2 \implies M_n$ bdd $L^1 \xRightarrow[\text{CT}]{\mathbb{MG}} M_n \xrightarrow{\text{a.s.}} M_\infty$ & $\mathbb{P}(M_\infty < \infty) = 1.$

$\mathbb{E}\left[(M_m - M_n)^2\right] = \sum_{k=m+1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2\right] \xrightarrow[n\to\infty]{m\to\infty} 0$ (L^2 $\mathbb{MG}$ boundedness)

Way 1: $\xRightarrow{\text{Fatou}} \mathbb{E}\left[(M_m - M_\infty)^2\right] \le \sum_{k=m+1}^{\infty} \mathbb{E}\left[(M_k - M_{k-1})^2\right] \xrightarrow{m\to\infty} 0 \implies M_n \to M_\infty$ a.s. & in L^2

Way 2: $\xRightarrow[\text{boundedness}]{L^2\ \mathbb{MG}} \{M_n\}_{n=1}^{\infty}$ is Cauchy in $L^2 \xRightarrow[\text{complete}]{L^2 \text{ is}} \exists \tilde{M}$ S.T. $M_n \xrightarrow[n\to\infty]{L^2} \tilde{M}$

Now: $\mathbb{P}\left(|\tilde{M} - M_\infty| > \varepsilon\right) \le \underbrace{\mathbb{P}\left(|\tilde{M} - M_n| > \frac{\varepsilon}{2}\right)}_{\le \frac{\mathbb{E}[|M_n - \tilde{M}|^2]}{(\varepsilon/2)^2} \to 0} + \underbrace{\mathbb{P}\left(|M_n - M_\infty| > \frac{\varepsilon}{2}\right)}_{\to 0} \to 0$ $\square$

Sum of 0-Mean Indep r.v.s in L^2 $X_i \perp\!\!\!\perp$ S.T. $\mathbb{E}[X_i]=0$ & $\mathbb{E}\left[X_i^2\right]=\sigma_i^2$.

(i) $\displaystyle\sum_k \sigma_k^2<\infty \implies \sum_k X_k(\omega)<\infty$ a.s.

(ii) $\exists K>0$ S.T. $|X_n(\omega)|\le K$, $\forall n$, $\forall\omega$:

$\displaystyle\sum_k X_k<\infty$ a.s. $\implies \displaystyle\sum_k \sigma_k^2<\infty$.

Note By Kolmogorov 0-1 Law, $\mathbb{P}\left(\omega:\displaystyle\sum_k X_k(\omega)\text{ converges}\right)\in\{0,1\}$.

Proof: *(click)*

$M_n:=X_1+\dots X_n$ & M_n $\mathbb{MG}$ and $\in L^2$. $\mathscr{F}_n=\sigma(X_1,\dots,X_n)$[a] & $\mathscr{F}_0=\{\varnothing,\Omega\}$.

(i) $\displaystyle\sum_n\mathbb{E}\left[(M_n-M_{n-1})^2\right]=\sum_n\mathbb{E}\left[X_n^2\right]=\sum_n\sigma_n^2<\infty \xRightarrow[\text{Convergence}]{L^2} S_n\xrightarrow{\text{a.s.}} S$ & $S<\infty$ a.s.

(ii) $G_n=M_n^2-\displaystyle\sum_{k=1}^n\sigma_k^2$.

Claim: G_n is $\mathbb{MG}$.

$$(\because)\mathbb{E}\left[G_n|\mathscr{F}_{n-1}\right]=\mathbb{E}\left[(M_{n-1}+X_n)^2-\sum_{k=1}^n\sigma_k^2\Big|\mathscr{F}_{n-1}\right]=\mathbb{E}\left[M_{n-1}^2+2M_{n-1}X_n+X_n^2-\sum_{k=1}^n\sigma_k^2\Big|\mathscr{F}_{n-1}\right]$$

$$=M_{n-1}^2+2M_{n-1}\underbrace{\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]}_{=\mathbb{E}[X_n]=0}+\underbrace{\mathbb{E}\left[X_n^2|\mathscr{F}_{n-1}\right]}_{=\mathbb{E}[X_n^2]=\sigma_n^2}-\sum_{k=1}^n\sigma_k^2=M_{n-1}^2-\sum_{k=1}^{n-1}\sigma_k^2=G_{n-1}.$$

Claim: $\exists c>0$ S.T. $T_c:=\inf\left\{n\ge1:|M_n|>c\right\}$ & $\mathbb{P}(T_c=\infty)>0$.

$(\because)$Suppose, $\forall c>0$, $\mathbb{P}(T_c=\infty)=0 \implies \forall c,\forall\omega,\exists n(c,\omega)\ge1: |M_n(\omega)|>c \implies \limsup_n|M_n|=\infty$.

However, $M_n(\omega)$ conv $\implies |M_n(\omega)|$ is bdd $\implies \limsup_n|M_n(\omega)|<\infty$.

Claim: $|M_{T_c\wedge n}|\le K+c$ for every n.

$$(\because)\begin{cases}T_c\ge n\implies|M_{T_c\wedge n}|=|M_n|\le c\\ T_c\le n\implies|M_{T_c\wedge n}|=|M_{T_c}|\le|M_{T_c}-M_{T_c-1}|+|M_{T_c-1}|\le K+c\end{cases}$$

Claim: $\displaystyle\sum_{k=1}^\infty\sigma_k^2<\infty$

$$(\because)G_{T_c\wedge n}\text{ is }\mathbb{MG}\implies 0=\mathbb{E}\left[G_{T_c\wedge n}\right]=\mathbb{E}\left[M_{T_c\wedge n}^2\right]-\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\right]\xRightarrow[\le K+c]{|M_{T_c\wedge n}|}\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\right]\le(K+c)^2.$$

Hence, $$(K+c)^2\ge\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\right]=\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\Big|T_c=\infty\right]\mathbb{P}(T_c=\infty)+\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\Big|T_c<\infty\right]\mathbb{P}(T_c<\infty)$$

$$\implies (K+c)^2\ge\mathbb{E}\left[\sum_{k=1}^{T_c\wedge n}\sigma_k^2\Big|T_c=\infty\right]\mathbb{P}(T_c=\infty)=\left(\sum_{k=1}^\infty\sigma_k^2\right)\mathbb{P}(T_c=\infty)\implies\sum_k\sigma_k^2<\infty.\qquad\square$$

[a] Natural filtration of the process.

Random Signs $\varepsilon_n \sim \pm 1$, w.p. $1/2$ & $\{a_n\}_{n=1}^{\infty} \in \mathbb{R}$ non-random.

(i) $\sum_n a_n \varepsilon_n < \infty$ a.s. $\iff \sum_n \sigma_n^2 < \infty$

(ii) $\sum_n \sigma_n^2 = \infty \implies \sum_n a_n \varepsilon_n$ oscillates infinitely.

Careful! We don't know (TLTC[1]) if we can allow the random sum can converge to ∞.

Proof: *(click)*

(i) $\sigma^2 = \mathrm{Var}(a_n \varepsilon_n) = a_n^2 \mathrm{Var}(\varepsilon) = a_n^2 \implies \sum_n \sigma_n^2 < \infty \xRightarrow[\text{Result}]{\text{Prev.}} \sum_n a_n \varepsilon_n$.

& $\sum a_n \varepsilon_n < \infty \implies |a_n \varepsilon_n| = |a_n| \to 0 \implies a_n$ bdd. Prev. result.

(ii) **GOAL:** Prove that $\limsup_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k - \liminf_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k = \infty$, a.s. .

Kolmogorov's 0-1 Law $\Longrightarrow \mathbb{P}\left(\limsup_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k < \infty\right) \in \{0,1\}$ & $\mathbb{P}\left(\liminf_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k > -\infty\right) \in \{0,1\}$

Only case to eliminate:[a] $\mathbb{P}\left(\limsup_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k < \infty\right) = 1$ & $\mathbb{P}\left(\liminf_{n\to\infty} \sum_{k=1}^{n} a_k \varepsilon_k > -\infty\right) = 1$.

Claim: Under assumptions above, a_n bounded.

($\because$)Let $M_n = \sum_{k=1}^{n} a_k \varepsilon_k$. Fix ω S.T. $\limsup_{n\to\infty} M_n(\omega) = L < \infty$ & $\liminf_{n\to\infty} M_n(\omega) = N > -\infty$.

Take $\varepsilon > 0$: $\exists N$ S.T. $n \geq N \implies M_n \leq L + \varepsilon$ & $M_n \geq N - \varepsilon$

$\implies L + \varepsilon \geq M_n \geq N - \varepsilon, \forall n \geq N \implies |M_n - M_{n+1}| = |a_{n+1}| \leq L - N + 2\varepsilon, \forall n \geq N$. Hence, a_n bdd.

Claim: $\lim_{n\to\infty} M_n(\omega)$ exists, a.s.

($\because$)Take an ω, $\limsup_{n\to\infty} M_n(\omega) < \infty \implies \exists B > 0$ S.T. $M_n(\omega) \leq B, \forall n$.

Define stop. time $T := \inf\{n \geq 1 : M_n > B\}$:

$M_{T\wedge n} \leq M_T - M_{T-1} + M_{T-1} \leq |a_T| + B \leq \sup_{n\geq 1} |a_n| + B. \xRightarrow[\text{CT}]{\mathbb{MG}} M_{T\wedge n} \xrightarrow{\text{a.s.}} M_\infty$ with $\mathbb{P}(M_\infty < \infty) = 1$.

For ω above, $T(\omega) = \infty \implies M_n(\omega) \xrightarrow{\text{a.s.}} M_\infty(\omega)$ $\boxed{\Rightarrow\Leftarrow}$($\sum_n a_n^2 = \infty \implies$ can't converge).

Note Above, we 'proved' that if M_n $\mathbb{MG}$ S.T. $|M_n - M_{n-1}| \leq K$

Then $\limsup_{n\to\infty} M_n < \infty$ a.s. or $\liminf_{n\to\infty} M_n > -\infty$ a.s. $\implies M_n \xrightarrow{\text{a.s.}} M_\infty$. $\square$

[a] The $\infty - \infty$ case is not allowed by the definition.

[1] Too Lazy To Check

Symmetrization Technique: Expanding the Sample Space

$(\Omega^*, \mathscr{F}^*, \mathbb{P}^*) = (\Omega, \mathscr{F}, \mathbb{P}) \times (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$, where $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ is an exact copy of $(\Omega, \mathscr{F}, \mathbb{P})$.

Let $\omega^* = (\omega, \tilde{\omega}) \in \Omega^*$: $\begin{cases} X_n^*(\omega^*) = X_n(\omega) \\ \tilde{X}_n^*(\omega^*) = \tilde{X}_n(\tilde{\omega}) \end{cases}$ & $Z_n^*(\omega^*) := X_n^*(\omega^*) - \tilde{X}_n^*(\omega^*)$

$\implies$ Then: $\{Z_n^*\}_{n=1}^{\infty}$ are indep, with $\begin{cases} \mathbb{E}\left[Z_n^*\right] = 0 \\ \mathrm{Var}\left(Z_n^*\right) = 2\mathrm{Var}\left(X_n\right) \end{cases}$

Proof: *(click)*

Claim: $X_n^* \perp\!\!\!\perp \tilde{X}_m^*$.

$(\because)\mathbb{P}_{X_n^*, \tilde{X}_m^*}(\cdot,\cdot)$: $\mathscr{B}\times\mathscr{B}\to[0,1]$ $\mathbb{P}$-meas. (check on π-syst $\{B_1 \times B_2 : B_1, B_2 \in \mathscr{B}\}$ gen. $\mathscr{B}\times\mathscr{B}$)

$\mathbb{P}_{X_n^*, \tilde{X}_m^*}(B_1, B_2) = \mathbb{P}^*\left(\omega^* : X_n^*(\omega^*) \in B_1, \tilde{X}_m^*(\omega^*) \in B_2\right) = \mathbb{P}^*\left((\omega, \tilde{\omega}) : X_n(\omega) \in B_1, \tilde{X}_m(\tilde{\omega}) \in B_2\right)$.

$A := \left\{\omega : X_n(\omega) \in B_1\right\} \in \mathscr{F}$ & $B := \left\{\omega : \tilde{X}_m(\omega) \in B_2\right\} \in \tilde{\mathscr{F}}$.

$\xRightarrow[\text{Prod. Meas.}]{\text{Uniq. of}} \mathbb{P}^*\left(A \times B\right) = \mathbb{P}(A)\tilde{\mathbb{P}}(B) = \mathbb{P}\left(X_n \in B_1\right)\tilde{\mathbb{P}}(\tilde{X}_m \in B_2) = \mathbb{P}^*_{X_n^*}(B_1)\mathbb{P}^*_{\tilde{X}_m^*}(B_2)$.

Claim: $\mathbb{E}\left[Z_n^*\right] = 0$ & $\mathrm{Var}\left(Z_n^*\right) = 2\mathrm{Var}\left(X_n\right)$

$(\because)\begin{cases} \mathbb{E}\left[Z_n^*(\omega^*)\right] = \mathbb{E}\left[X_n^*(\omega^*)\right] - \mathbb{E}\left[\tilde{X}_n^*(\omega^*)\right] \overset{\text{copy}}{=\!=} 0 \\ \mathrm{Var}\left(Z_n^*(\omega^*)\right) \overset{\perp\!\!\!\perp}{=\!=} \mathrm{Var}\left(X_n^*(\omega^*)\right) + \mathrm{Var}\left(\tilde{X}_n^*(\omega^*)\right) = 2\mathrm{Var}\left(X_n\right) \end{cases}$ $\square$

Conv implies $\mathbb{E}[\cdot]$ & $\mathrm{Var}(\cdot)$ Conv $X_n \perp\!\!\!\perp$ & $\exists K > 0$ S.T. $|X_n(\omega)| \le K, \forall n, \omega$.

$\sum_n X_n$ converges a.s. $\implies \sum_n \mathbb{E}\left[X_n\right]$ converges & $\sum_n \mathrm{Var}\left(X_n\right) < \infty$.

Note This is a partial converse to Kolmogorov's 2-Series Theorem.

Proof: *(click)*

Take Z_n^* defined in symmet. technique above. $\mathbb{E}\left[Z_n^*\right] = 0$ & $\mathrm{Var}\left(Z_n^*\right) = 2\mathrm{Var}\left(X_n\right)$ & $|Z_n^*| \le 2K$ $(\forall n, \forall\omega)$.

$G := \left\{\omega \in \Omega : \sum_n X_n(\omega) \text{ conv}\right\}$ & $\tilde{G} := \left\{\tilde{\omega} \in \tilde{\Omega} : \sum_n \tilde{X}_n(\tilde{\omega}) \text{ conv}\right\} \implies \mathbb{P}\left(G\right) = \tilde{\mathbb{P}}(\tilde{G}) = 1 \implies \mathbb{P}^*(G \times \tilde{G}) = 1$

On $G \times \tilde{G}$, $\sum_n Z_n^*(\omega^*) = \sum_n X_n(\omega) - \tilde{X}_n(\tilde{\omega})$ conv $\implies \mathbb{P}^*(\omega^* : \sum_n Z_n^* \text{ conv}) = 1$.

$\xRightarrow[L^2 \perp\!\!\!\perp\text{r.v.s } \mathbf{(b)}]{\text{Sum of 0-Mean}} \sum_n \mathrm{Var}\left(X_n\right) = \sum_n \sigma_n^2 < \infty$.

$\xRightarrow[L^2 \perp\!\!\!\perp\text{r.v.s } \mathbf{(a)}]{\text{Sum of 0-Mean}} \sum_n (X_n - \mathbb{E}\left[X_n\right]) < \infty \implies \sum_n \mathbb{E}\left[X_n\right] < \infty$. $\square$

Kolmogorov's 3-Series (⊥) $\{X_n \perp\!\!\!\perp\}$, $A > 0$ & $Y_i = X_i \mathbb{1}_{|X_i| \le A}$: Then $\sum_{n=1}^{\infty} X_n$ converges a.s. $\Leftrightarrow$

- **(i)** $\sum_{n=1}^{\infty} \mathbb{P}\big(|X_n| > A\big) < \infty$
- **(ii)** $\sum_{n=1}^{\infty} \mathbb{E}\,[Y_n] < \infty$
- **(iii)** $\sum_{n=1}^{\infty} \mathrm{Var}\,(Y_n) < \infty$.

Note (Technical) Above holds if there is an A. In this case, it holds for any A [2].

Proof: *(click)*

($\Leftarrow$) Done: Kolmogorov's 3-Series Theorem.

($\Rightarrow$) $\sum_n X_n(\omega) < \infty$ a.s. $\implies$ $X_n \xrightarrow{\text{a.s.}} 0$ $\implies$ $\forall K\ \mathbb{P}\big(|X_n| > K \text{ i.o. }\big) = 0 \xRightarrow[1\&2]{\text{BC}} \sum_{n=1}^{\infty} \mathbb{P}\big(|X_n| > K\big) < \infty$. □

Cesaro's Lemma $a_n \nearrow \infty$ & $x_n \to x_\infty \implies \frac{1}{a_n} \sum_{k=1}^{n} (a_k - a_{k-1}) x_k \to x_\infty$.

Note Taking $a_n = n$, get standard Cesaro mean.

(∵)See Williams [13] p.117.

Kronecker's Lemma $a_n \nearrow \infty$ & $\sum_{n=1}^{\infty} \frac{x_n}{a_n} < \infty \implies \frac{1}{a_n} \sum_{k=1}^{n} x_k \to 0$.

(∵)See Durrett [5] p.81; Williams [13] p.117. Use Cesaro.

[2] *Aut Caesar, aut nullus*

7.2 Day 2: Doob Decomposition, Angle Brackets Process

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Angle-Bracket Process $M \in L^2$ $\mathbb{MG}$ & $M_0 = 0$. M_n^2 Sub$\mathbb{MG}$ & $M_n^2 = N_n + A_n$ (Doob decomp.).

$\implies A = \langle M \rangle$. Will study A for convergence of M_n.

Note $\xrightarrow[\text{Decomp}]{\text{Doob}}$ $A = \langle M \rangle$ is $\nearrow$ and has limit A_∞

MAIN THEOREMS:

Doob Decomposition Given filtration $\{\mathscr{F}_n\}$ & adapted process $\{X_n\} \in L^1$:

$$\exists \text{ decomp. } X_n = X_0 + M_n + A_n \text{ with } \begin{cases} M_n \ \mathbb{MG} \ , \ M_n = \displaystyle\sum_{k=1}^{n} X_k - \mathbb{E}\left[X_k | \mathscr{F}_{k-1}\right] \\ A_n \text{ previsible, } A_n = \displaystyle\sum_{k=1}^{n} \mathbb{E}\left[X_k | \mathscr{F}_{k-1}\right] - X_{k-1} \\ M_0 = A_0 = 0 \end{cases}$$

If $X_n = X_0 + M'_n + A'_n \implies M_n = M'_n$ & $A_n = A'_n$ a.s. , $\forall n$.

Corollary: X_n is a Sub$\mathbb{MG}$ $\iff A_n \nearrow$

Note A_n accumulates expected increase & M_n accumulates surprises.

Note Doob-Meyer Decomposition generalizes this for continuous time.

Proof: *(click)*

WLOG $X_0 = 0$ (otherwise, take $\tilde{X}_n = X_n - X_0$)

$\mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] = \mathbb{E}\left[M_n + A_n | \mathscr{F}_{n-1}\right] = M_{n-1} + A_n = X_{n-1} - A_{n-1} + A_n \Rightarrow A_n - A_{n-1} = \mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] - X_{n-1}$.

$\implies A_n = \displaystyle\sum_{k=1}^{n} \mathbb{E}\left[X_k | \mathscr{F}_{k-1}\right] - X_{k-1} \implies M_n = \displaystyle\sum_{k=1}^{n} X_k - \mathbb{E}\left[X_k | \mathscr{F}_{k-1}\right]$.

Corollary: X_n Sub$\mathbb{MG}$ $\iff \mathbb{E}\left[X_n | \mathscr{F}_{n-1}\right] - X_{n-1} \geq 0 \iff A_n \nearrow$. $\square$

Function of an $\mathbb{MG}$ X_n process, $\varphi(\cdot)$ function S.T. $\mathbb{E}\left[|\varphi(X_n)|\right] < \infty, \forall n$

$$\implies \begin{cases} X_n\ \mathbb{MG}\ ,\ \varphi(\cdot) \text{ convex } \implies \varphi(X_n) = \text{Sub}\mathbb{MG} \\ X_n\ \mathbb{MG}\ ,\ \varphi(\cdot) \text{ concave } \implies \varphi(X_n) = \text{Sup}\mathbb{MG} \\ X_n\ \text{Sub}\mathbb{MG}\ ,\ \varphi(\cdot) \text{ cvx + increasing } \implies \varphi(X_n) = \text{Sub}\mathbb{MG} \end{cases}$$

Proof: *(click)*

$\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right] = X_{n-1} \implies \mathbb{E}\left[\varphi(X_n)|\mathscr{F}_{n-1}\right] \underbrace{\geq}_{\text{Jensen}} \varphi(\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]) \geq \varphi(X_{n-1})$

$\sigma(\varphi(X_n)) \subset \sigma(X_n) \subset \mathscr{F}_n \implies \varphi(X_n) \in m\mathscr{F}_n$ & $\mathbb{E}\left[|\varphi(X_n)|\right] < \infty$

$\implies \varphi(X_n)$ Sub$\mathbb{MG}$. Other cases are analogous. $\square$

Angle-Brackets Process $A = \langle M\rangle$: $A_n \nearrow A_\infty$ exists (∞ allowed).

M bdd in $L^2 \iff \mathbb{E}\left[A_\infty\right] < \infty$.

$A_n - A_{n-1} = \mathbb{E}\left[M_n^2 - M_{n-1}^2|\mathscr{F}_{n-1}\right] = \mathbb{E}\left[(M_n - M_{n-1})^2|\mathscr{F}_{n-1}\right]$.

Note A is useful in studying the martingale. For examples, see Williams [13].

Proof: *(click)*

M_n^2 Sub$\mathbb{MG}$ $\implies 0 \leq A_n \nearrow \xRightarrow{\text{monotonic}} A_\infty$ exists.

$\mathbb{E}\left[M_n^2\right] = \mathbb{E}\left[A_n\right] + \mathbb{E}\left[N_n\right] = \mathbb{E}\left[A_n\right] \implies \sup_n \mathbb{E}\left[M_n^2\right] = \sup_n \mathbb{E}\left[A_n\right] = \mathbb{E}\left[A_\infty\right]$.

Last item: plug & check. $\square$

Interpretation of A-B Process A-B process called quadratic variation: $A_n - A_{n-1} = \text{Var}\left(X_n|\mathscr{F}_{n-1}\right)$,

hence A_n accumulates sum of variation in process.

$(\because) A_n = \sum_{n=1}^{\infty} \mathbb{E}\left[X_n^2|\mathscr{F}_{n-1}\right] - X_{n-1}^2$ and $X_{n-1} = \mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]$

$\implies A_n - A_{n-1} = \mathbb{E}\left[X_n^2|\mathscr{F}_{n-1}\right] - \left(\mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]\right)^2 = \mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n|\mathscr{F}_{n-1}\right]\right)^2 |\mathscr{F}_{n-1}\right] = \text{Var}\left(X_n|\mathscr{F}_{n-1}\right)$

Convergence of $M \leftrightarrow$ finiteness of A_∞ $A := \langle M \rangle$

(i) For a.a ω: "$A_\infty(\omega) < \infty \implies \lim_n M_n(\omega)$ exists and is finite" [3].

(ii) $\exists K > 0$ S.T. $|M_n(\omega) - M_{n-1}(\omega)| \le K$, $\forall n, \omega$[4]:

$\implies$ For a.a ω: "$\lim_n M_n(\omega) \implies A_\infty(\omega) < \infty$ exists".

Note This is an extension of Sum of 0-Mean L^2 $\perp\!\!\!\perp$r.v.s.

Proof: *(click)*

Can't use Doob's Conv. Thm.: Look at pointwise stuff, not L^2/L^1 bounded stuff. $A_\infty < \infty \not\Leftrightarrow \mathbb{E}[A_\infty] < \infty$.

(i) Claim: $T_k = \inf\{n \ge 1 : A_{n+1} > k\} \implies \{\omega \in \Omega : A_\infty < \infty\} = \bigcup_{k=1}^{\infty} \{T_k = \infty\}$.

$(\because)\{A_\infty < \infty\} = \bigcup_{k=1}^{\infty} \{A_\infty < k\} = \bigcup_{k=1}^{\infty} \underbrace{\bigcap_{n=1}^{\infty} \{A_n < k\}}_{=\{T_k=\infty\}} = \bigcup_{k=1}^{\infty} \{T_k = \infty\}$, as $A_\infty < k \implies A_n < k, \forall n$.

Claim: T_k is a stopping time.

$(\because)\{T_k = n\} = \left(\bigcap_{i=1}^{n} \underbrace{\{A_i \le k\}}_{\mathscr{F}_{i-1}}\right) \cap \underbrace{\{A_{n+1} > k\}}_{\in \mathscr{F}_n} \implies \{T_k = n\} \in \mathscr{F}_n$.

Claim: $A_{T_k \wedge n}$ is previsible.

$(\because)$Fix $B \in \mathscr{B}$: $\left\{\omega \in \Omega : A_{T_k \wedge n} \in B\right\} = \left\{\omega : T_k \le n-1, A_{T_k \wedge n} \in B\right\} \cap \left\{\omega : T_k \ge n, A_{T_k \wedge n} \in B\right\}$

So: $\left\{\omega \in \Omega : A_{T_k \wedge n} \in B\right\} = \left(\bigcup_{i=1}^{n-1} \underbrace{\overbrace{\{T_k = i\}}^{\in \mathscr{F}_i} \cap \overbrace{\{A_i \in B\}}^{\in \mathscr{F}_{i-1}}}_{\in \mathscr{F}_{n-1}}\right) \cap \left(\underbrace{\{T_k \le n-1\}^c}_{\in \mathscr{F}_{n-1}} \cap \underbrace{\{A_n \in B\}}_{\in \mathscr{F}_{n-1}}\right)$

Claim: $M^2_{T_k \wedge n} = N_{T_k \wedge n} + A_{T_k \wedge n} \xRightarrow[\text{of } M^2]{\text{Doob dec.}} A_{T_k \wedge n} = \left\langle M_{T_k \wedge n} \right\rangle$.

$(\because)N_{T_k \wedge n}$ is $\mathbb{MG}$ & $A_{T_k \wedge n} \in \mathscr{F}_{n-1}$.

Claim: $M_{T_k \wedge n}(\omega)$ bdd in $L^2 \implies \lim_{n\to\infty} M_{T_k(\omega) \wedge n}$ exists and is finite.

$(\because)\mathbb{E}\left[M^2_{T_k \wedge n}\right] = \underbrace{\mathbb{E}\left[N_{T_k \wedge n}\right]}_{=0} + \underbrace{\mathbb{E}\left[A_{T_k \wedge n}\right]}_{\le k} \le k$.

Claim: $A_\infty < \infty \implies \lim_{n\to\infty} M_{T_k \wedge n}$ exists and is finite.

$(\because)$On $\left\{\omega : T_k(\omega) = \infty\right\}$, $\lim_{n\to\infty} M_{T_k(\omega) \wedge n}(\omega) = \lim_{n\to\infty} M_n(\omega)$ exists, except a set N_k of meas. 0.

(ii)

TBD $\square$

Easy SLLN for MG $M \in L^2$ $\mathbb{MG}$, $M_0 = 0$ & $A = \langle M \rangle$.

(i) $\dfrac{1}{1+A}$ bdd + previs.

(ii) $W_n = \left(\dfrac{1}{1+A} \bullet M\right)_n = \sum_{k=1}^{n} \dfrac{M_k - M_{k-1}}{1+A_k}$ $\mathbb{MG}$ with $\langle W \rangle_\infty \le 1$ & $\lim_n W_n \exists$ a.s.

(iii) $\dfrac{M_n}{A_n} \xrightarrow{\text{a.s.}} 0$ on the set $\left\{\omega \in \Omega : A_\infty(\omega) = \infty\right\}$.

---

[3]More precisely, $B = \left\{\omega \in \Omega : A_\infty(\omega) < \infty\right\}$. Then, $\exists N$ S.T. $\mathbb{P}(N) = 0$ & $\lim_n M_n(\omega)$ exists, $\forall \omega \in B \setminus N$.

[4]*i.e.,* M has uniformly bounded increments.

Proof: *(click)*

(**i**)

(**ii**)

(**iii**)

TBD ☐

Levy's Extension of BC Lemma Suppose $E_n \in \mathscr{F}_n \forall n$. Let $\begin{cases} Z_n(\omega) := \sum_{k=1}^{n} \mathbb{1}_{E_k}(\omega) \\ \xi_k := \mathbb{P}\left(E_k | \mathscr{F}_{k-1}\right) \\ Y_n := \sum_{k=1}^{n} \xi_k \end{cases}$. Then,

(**i**) $Y_\infty < \infty \implies Z_\infty < \infty$.

(**ii**) $Y_\infty = \infty \implies \dfrac{Z_n}{Y_n} \xrightarrow{\text{a.s.}} 1$.

Note Z_n is number of occurrences of E_k $(k \leq n)$.

Note $\mathbb{E}[\xi_k] = \mathbb{P}(E_k) \implies \mathbb{E}[Y_n] = \sum_{k=1}^{n} \mathbb{P}(\xi_k) \underbrace{\nearrow}_{\text{MCT}} \mathbb{E}[Y_\infty]$.

$\sum_{k=1}^{\infty} \mathbb{P}(E_k) < \infty \implies \mathbb{E}[Y_\infty] < \infty \implies Y_\infty < \infty \overset{\text{(i)}}{\implies} Z_\infty < \infty \implies \text{BC1}$.

Note $E_n \perp\!\!\!\perp$, $\mathscr{F}_n = \sigma(\mathbb{1}_{E_1}, \ldots, \mathbb{1}_{E_n})$: $\xi_k = \mathbb{P}\left(E_k | \mathscr{F}_{k-1}\right) \overset{\perp\!\!\!\perp}{=\!=} \mathbb{P}(E_k) \implies Y_\infty = \sum_{k=1}^{\infty} \mathbb{P}(E_k)$.

$\overset{\text{(ii)}}{\implies} \dfrac{Z_n}{Y_n} \xrightarrow{\text{a.s.}} 1 \implies Z_\infty = \infty \implies \text{BC2}$

Proof: *(click)*

(**i**)

(**ii**)

TBD ☐

7.3 Day 3: Uniform Integrability

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Keyword 1

MAIN THEOREMS:

Theorem 1

Chapter 8

Week 8:

8.1 Day 1:

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Keyword 1

MAIN THEOREMS:

Theorem 1

8.2 Day 2:

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Keyword 1

MAIN THEOREMS:

Theorem 1

8.3 Day 3:

Main Reference(s): David Williams, *Probability With Martingales*, 1991 [13]

KEYWORDS:

Keyword 1

MAIN THEOREMS:

Theorem 1

Index

Bibliography

[1] R. Ash and C. Doléans-Dade. *Probability & Measure Theory*. Harcourt/Academic Press, 2000.

[2] Martin Bazant. 18.366: Random walks and diffusions. Available at `https://ocw.mit.edu/courses/mathematics/18-366-random-walks-and-diffusion-fall-2006/`, Fall 2006.

[3] P. Billingsley. *Probability & Measure*. Wiley Series in Probability and Statistics. Wiley, 1995.

[4] P. Bremaud. *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer New York, 1981.

[5] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[6] G. Folland. *Real Analysis:Modern Techniques&Application*. Pure&Applied Math. Wiley, 1999.

[7] G. Grimmett and D. Stirzaker. *1000 Exercises in Probability*. OUP Oxford, 2001.

[8] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Probability and Random Processes. OUP Oxford, 2001.

[9] Alice Guionnet. 18.176: Introduction to stochastic analysis. Available at `http://perso.ens-lyon.fr/aguionne/CoursStochProc18176.pdf`, Fall 2013.

[10] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. American Math. Soc.

[11] P. Morters and Y. Peres. *Brownian Motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[12] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer Berlin Heidelberg, 2013.

[13] D. Williams. *Probability with Martingales*. Cambridge mathematical textbooks. Cambridge University Press, 1991.