FSF3940 Probability Theory

Oral Exam

Ruo-Chun Tzeng

January 11, 2021

Outline

Outline

Kolmogorov Axioms

A probability space $(\underbrace{\Omega}_{\text{sample space}}, \underbrace{\mathcal{F}}_{\sigma\text{-field}}, \underbrace{P}_{\text{probability measure}})$ with $P : \mathcal{F} \to \mathbb{R}$ satisfying:

1. $\forall A \in \mathcal{F}$, $P(A) \geq 0$.
2. $P(\Omega) = 1$ and $P(\emptyset) = 0$.
3. Any countable disjoint $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ implies $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

The *σ-field* $\mathcal{F}$ is a collection of subsets of Ω satisfying

1. $\Omega \in \mathcal{F}$.
2. Any $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$.
3. Any countable $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ implies $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The *σ-field generated by class $\mathcal{C}$* is the smallest unique σ-field containing $\mathcal{C}$.

Random Variable & Convergence Types

A *random variable* $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if $\forall A \in \mathcal{B}(\mathbb{R})$, $X^{-1}(A) \in \mathcal{F}$.

- $\mathcal{B}(\mathbb{R})$ is called *Borel set*, generated by $\{(a, b] : -\infty \leq a < b < \infty\}$.
- X is *simple* if it has finite range.
- X is *integrable* if $\int_\Omega |X| dP < \infty$, denoted by $X \in L^1$.
- **(Measurability Theorem)** Any measurable X can be approximated *uniformly* by simple functions $\{X_n\}_{n=1}^{\infty}$.

Given a sequence $\{X_n\}_{n=1}^{\infty}$ of r.v.'s and a r.v. X,

- $X_n \to X$ *uniformly* if $\lim_{n\to\infty} \sup_{\omega\in\Omega} |X_n(\omega) - X(\omega)| = 0$.
- $X_n \to X$ *pointwisely* if $\forall \omega \in \Omega$, $\lim_{n\to\infty} |X_n(\omega) - X(\omega)| = 0$.
- $X_n \stackrel{a.s.}{\to} X$ if $\exists N \in \mathcal{F}$ with $P(N) = 0$ s.t. $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in N^c$.
- $X_n \stackrel{P}{\to} X$ if $\forall \epsilon > 0$, $\lim_{n\to\infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = 0$.

Convergence Theorem

(BCT) If $\{X_n\}_{n=1}^{\infty}$ is uniformly bounded and $X_n \xrightarrow{P} X$, then $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

(Fatou's Lemma) If $\{X_n\}_{n=1}^{\infty} \geq 0$ and $X_n \xrightarrow{P} X$, then $\mathbb{E}[X] \leq \liminf_{n\to\infty} \mathbb{E}[X_n]$.
More generally, if $\{X_n\}_{n=1}^{\infty} \geq 0$, then $\mathbb{E}[\liminf_{n\to\infty} X_n] \leq \liminf_{n\to\infty} \mathbb{E}[X_n]$.

(MCT) If $\{X_n\}_{n=1}^{\infty} \geq 0$ and $X_n \uparrow X$, then $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

(DCT) If $\{X_n\}_{n=1}^{\infty}$ with $X_n \xrightarrow{a.s./\mathrm{P}} X$, $|X_n| \leq Y$ and $Y \in L^1$, then $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

(Jensen's Inequality) Let ϕ be convex and $X, \phi(X) \in L^1$. Then, $\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X])$.

(Transformation Theorem) Let $X : (\Omega_1, \mathcal{F}_1) \mapsto (\Omega_2, \mathcal{F}_2)$, $Y : (\Omega_2, \mathcal{F}_2) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and $P : \mathcal{F}_1 \to \mathbb{R}$ be a prob. measure. Then,

$$\int_{\Omega_1} Y(X(\omega_1)) dP(\omega_1) = \int_{\Omega_2} Y(\omega_2) d \underbrace{P(X^{-1}(\omega_2))}_{\text{induced measure}}.$$

(Fubini's Theorem) Let $X \in L^1$ w.r.t. $P = P_1 \times P_2$ on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$. Then,

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} X(\omega_1, \omega_2) dP(\omega_1, \omega_2) &= \int_{\Omega_1} \int_{\Omega_2} X(\omega_1, \omega_2) dP_2(\omega_2) dP_1(\omega_1) \\ &= \int_{\Omega_2} \int_{\Omega_1} X(\omega_1, \omega_2) dP_1(\omega_1) dP_2(\omega_2).\end{aligned}$$

Outline

Distribution Function and Characteristic Function

Given a probability space $(\Omega, \mathcal{F}, P)$ and r.v. $X : \Omega \to \mathbb{R}$ measurable on $(\Omega, \mathcal{F})$.

- $\mu_X(\cdot) = P(X^{-1}(\cdot))$ is the *distribution* of X.
- $F_X(\cdot) = \mu_X((-\infty, \cdot])$ is the *distribution function* of X.
- If $X \in L^1$, $\mathbb{E}[X] = \int_{-\infty}^{\infty} x\mu_X(dx)$.
- Denote $\mu_n \xrightarrow{w} \mu$ if $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ for all $A \in \mathcal{B}(\mathbb{R})$ satisfying $\mu(\partial A) = 0$.

The *characteristic function* of X is

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\mu_X(dx), \forall t \in \mathbb{R}.$$

- φ_X is 1-to-1 to μ_X for any r.v. X.
- φ_X is uniformly continuous and $|\varphi_X(t)| \leq 1$, $\forall t \in \mathbb{R}$.
- If $\mathbb{E}[X^n] < \infty$, then φ_X is n-times continuously differentiable.

Portmanteau Theorem

(Portmanteau Theorem) The following are equivalent:

1. $X_n \Rightarrow X$.
2. $\forall$ bounded continuous function f on $\mathbb{R}$, $\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$.
3. $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$, $\forall x$ that is a continuous point of F_X.
4. $\forall$ closed set $C \subset \mathbb{R}$, $\limsup_{n\to\infty} P(X_n \in C) \leq P(X \in C)$.
5. $\forall$ open set $O \subset \mathbb{R}$, $\liminf_{n\to\infty} P(X_n \in O) \geq P(X \in O)$.
6. $\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t)$, $\forall t \in \mathbb{R}$.

Relation between convergence types:

- $X_n \stackrel{a.s.}{\to} X$ implies $X_n \stackrel{P}{\to} X$ implies $X_n \Rightarrow X$.
- If $X = C$ for some $C \in \mathbb{R}$, $X_n \Rightarrow X$ implies $X_n \stackrel{P}{\to} X$.

Independence

- Events A, B are *independent* if $P(A \cap B) = P(A)P(B)$.
- Random variables X, Y are *independent* if $P(X \in A, Y \in B) = P(X \in A)P(Y \in B), \forall A, B \in \mathcal{B}(\mathbb{R})$.

Lemma

Let X, Y be r.v.'s on $(\Omega, \mathcal{F}, P)$.
Then, X, Y are independent iff the induced probability measure

$$\mu_{(X,Y)}(A_1 \times A_2) = \mu_X(A_1)\mu_Y(A_2), \forall A_1, A_2 \in \mathcal{B}(\mathbb{R}).$$

This implies for any $f : \mathbb{R}^2 \to \mathbb{R}$,

$$\int_{\mathbb{R}^2} f(x,y)\mu_{(X,Y)}(dx \times dy) = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mu_X(dx)\mu_Y(dy).$$

Outline

Weak Law of Large Numbers

Let X, Y be independent r.v.'s. on $(\Omega, \mathcal{F}, P)$.

- $\forall A \in \mathcal{B}(\mathbb{R})$, $P((X+Y) \in A) = \int_{x\in\mathbb{R}} \int_{\{y\in\mathbb{R}: x+y\in A\}} \mu_Y(dy)\mu_X(dx)$.
- Characteristic function $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$, $\forall t \in \mathbb{R}$.
- $Var[X+Y] = Var[X] + Var[Y]$.

(Chebyshev's Inequality) If $Var[X] < \infty$, then $\forall t \in (0, \infty)$, $P(X > t) \leq \frac{\mathbb{E}[X^2]}{t^2}$.

(WLLN with finite variance) Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. r.v.'s with $Var[X_1] = \sigma^2 < \infty$.

$$\text{Then, } \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{P} \mathbb{E}[X_1].$$

(WLLN) Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. r.v.'s with $\mathbb{E}[|X_1|] < \infty$.

$$\text{Then, } \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{P} \mathbb{E}[X_1].$$

Strong Law of Large Numbers (1/2)

(Borel-Cantelli Lemma) Let $\{A_n\} \subset \mathcal{F}$. If $\sum_n P(A_n) < \infty$, then $P(\limsup_{n\to\infty} A_n) = 0$.

- By Chebyshev's inequality and Borel-Cantelli Lemma, we have the following:

(SLLN with finite 4-th moment) If i.i.d. $\{X_n\}$ with $\mathbb{E}[|X_1|^4] < \infty$, then

$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s.} \mathbb{E}[X_1].$$

Strong Law of Large Numbers (2/2)

(Kolmogorov's 1-Series Theorem) Let $\{X_n\}$ be independent with 0 mean and $\sum_{n=1}^{\infty} Var[X_n] < \infty$. Then, $\sum_{i=1}^{n} X_i$ converges a.s.

(SLLN) Let $\{X_n\}$ be i.i.d. with 0 mean and $\mathbb{E}[|X_1|] < \infty$.

$$\text{Then, } \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s.} 0.$$

- Let $\mathcal{F}^n = \sigma(X_n, X_{n+1}, \cdots)$ and *tail σ-field* $\mathcal{F}^\infty = \bigcap_n \mathcal{F}^n$.

(Kolmogorov's 0-1 Law) If $A \in \mathcal{F}^\infty$, then $P(A) = 0$ or 1.

Outline

Central Limit Theorem

(CLT with i.i.d. r.v.'s) Let $\{X_n\}$ be i.i.d. with $\overbrace{Var[X_1]}^{\sigma^2} < \infty$.

$$\text{Then, } \frac{\sum_{i\leq n} \overbrace{\frac{X_i - \mathbb{E}[X_1]}{\sigma}}^{\text{standardized } X_i}}{\sqrt{n}} \Rightarrow Z \sim \mathcal{N}(0,1).$$

(Lindeberg's CLT) Let $\{X_n\}$ be independent with $\mathbb{E}[X_1] = 0$ and $Var[X_n] = \sigma_n^2 < \infty$ such that

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i\leq n} \mathbb{E}[X_i^2 \mathbf{1}_{[|X_i|\geq \epsilon s_n]}] = 0, \forall \epsilon > 0, \text{ where } s_n^2 = \sum_{i\leq n} \sigma_i^2.$$

$$\text{Then, } \frac{\sum_{i=1}^n X_i}{s_n} \Rightarrow Z \sim \mathcal{N}(0,1).$$

Convergence Rate of CLT

(Berry-Esseen Theorem) Let $\{X_n\}$ be i.i.d. with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = \sigma^2 < \infty$, and $\mathbb{E}[|X_1|^3] = \rho < \infty$. Then, $\exists C > 0$ such that

$$\left| \underbrace{P(\frac{\sum_{i\leq n} X_i}{\sqrt{n\sigma^2}} \leq x)}_{\text{empirical distribution function}} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy \right| \leq \frac{C\rho}{\sigma^3 \sqrt{n}}, \forall x \in \mathbb{R}, n \in \mathbb{N}.$$

- Can CLT be stronger (i.e., i.p. or a.s. to $Z \sim \mathcal{N}(0,1)$)?

(Lemma) Let $\{X_n\}$ be i.i.d. with $\mathbb{E}[X_1] = 0$ and $Var[X_1] < \infty$.

$$\forall \{n_j\} \subset \{n\}_{n\in\mathbb{N}}, P(\limsup_{j\to\infty} \frac{\sum_{i\leq n_j} X_i}{\sqrt{n_j}} = \infty) = 1.$$

Law of Iterated Logarithm & Large Deviation Principles

- Any result of the form $\sum_{i\leq n} X_i/f(n) \to C$ other than $f(n) = \sqrt{n}$ and n?

(Law of Iterated Logarithm) Let $\{X_n\}$ be i.i.d. with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = 1$.

$$\text{Then, } P(\limsup_{n\to\infty} \frac{\sum_{i\leq n} X_i}{\sqrt{n \ln \ln n}} = \sqrt{2}) = P(\liminf_{n\to\infty} \frac{\sum_{i\leq n} X_i}{\sqrt{n \ln \ln n}} = -\sqrt{2}) = 1$$

- How fast does $\sum_{i\leq n} X_i/n \to 0$?

(Cramér's Theorem) Let $\{X_n\}$ be i.i.d. such that $H(\alpha) = \ln \mathbb{E}[e^{\alpha X_1}] < \infty$, $\forall \alpha \in \mathbb{R}$.

$$\text{Then, } \lim_{n\to\infty} \frac{\ln P(\sum_{i\leq n} X_i/n > \beta)}{n} = -\underbrace{\sup_{\alpha\in\mathbb{R}} (\alpha\beta - H(\alpha))}_{L(\beta)}, \forall \beta > \mathbb{E}[X_1].$$

Outline

Motivation

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G} \subset \mathcal{F}$ be a sub σ-field and a r.v. $X \in L^1$

- Naive: if $P(A) > 0$ for some $A \in \mathcal{F}$, then

$$\underbrace{\mathbb{E}[X|A]}_{\text{a value}} = \int_{\Omega \cap A} X(\omega) dP(\omega|A) = \int_A X dP / P(A), \tag{1}$$

  where $P(\cdot|A) = P(\cdot \cap A)/P(A)$.

- General: $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and

$$\int_A \underbrace{\mathbb{E}[X|\mathcal{G}]}_{\text{a r.v.}} dP = \int_A X dP, \forall A \in \mathcal{G}. \tag{2}$$

Radon-Nikodym Theorem

- A *signed measure* λ is countably additive but not necessarily nonnegative.

  Example: $\lambda(\cdot) = \int_{\cdot} X d\mu$ for nonnegative measure μ.
  Then, $\mu(A) = 0$ implies $\lambda(A) = 0$ since $\min_\omega X(\omega)\mu(A) \leq \lambda(A) \leq \max_\omega X(\omega)\mu(A)$.

- $\lambda \ll \mu$ denotes *absolutely continuity* of the signed measure λ w.r.t. the nonnegative measure μ. That is, $\mu(A) = 0$ for some $A \in \mathcal{F}$ implies $\lambda(A) = 0$.

(Radon-Nikodym Theorem) If $\lambda \ll \mu$, then $\exists$ a $\mathcal{F}$-measurable function $f \in L^1$ such that $\lambda(A) = \int_A f d\mu$, $\forall A \in \mathcal{F}$. The function $f = \frac{d\lambda}{d\mu}$ is uniquely determined μ-a.s.

(Existence and Uniqueness) Let X be $\mathcal{F}$-measurable and in L^1 w.r.t. P and $\mathcal{G} \subset \mathcal{F}$. Then, $\exists$ a $\mathcal{G}$-measurable and P-a.s. unique r.v. $\mathbb{E}[X|\mathcal{G}]$ such that

$$\int_A \mathbb{E}[X|\mathcal{G}] dP = \int_A X dP, \forall A \in \mathcal{G}.$$

Properties

(Properties of $\mathbb{E}[X|\mathcal{G}]$) Let $\mathcal{G} \subset \mathcal{F}$ and X be $\mathcal{F}$-measurable and Y be $\mathcal{G}$-measurable.

- $\mathbb{E}[Y|\mathcal{G}] \overset{a.s.}{=} Y$.
- $\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|] \overset{a.s.}{\leq} \mathbb{E}[|X|]$.
- If $X \geq 0$, $\mathbb{E}[X|\mathcal{G}] \overset{a.s.}{\geq} 0$.
- $\forall a, b \in \mathbb{R}$, $\mathbb{E}[aX + bY|\mathcal{G}] \overset{a.s.}{=} a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$.
- If $X, XY \in L^1$, $\mathbb{E}[XY|\mathcal{G}] \overset{a.s.}{=} Y\mathbb{E}[X|\mathcal{G}]$.
- If $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2] \overset{a.s.}{=} \mathbb{E}[X|\mathcal{G}_1] \overset{a.s.}{=} \mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1]$.
- $\mathbb{E}[XY|\mathcal{G}]^2 \overset{a.s.}{\leq} \mathbb{E}[X^2|\mathcal{G}]\mathbb{E}[Y^2|\mathcal{G}]$.
- BCT, MCT, DCT, Fatou's Lemma, Jensen's Inequality hold a.s.

Outline

Stationary Markov Chain

Let $(\Omega, \mathcal{B}, P)$ be a probability space and $(\mathcal{X}, \mathcal{F})$ be the state space.

- $\{X_n\}$ with $X_n : \Omega \to \mathcal{X}$ is a *Markov Chain* if

$$P(X_{n+1} \in A | \sigma(X_1, \cdots, X_n)) = P(X_{n+1} \in A | X_n), \forall A \in \mathcal{F}, \forall n \in \mathbb{N} \cup \{0\}.$$

- Markov chain $\{X_n\}$ is *stationary* with transition probability $\pi : \mathcal{X} \times \mathcal{F} \to [0, 1]$ if

$$P(X_{n+1} \in A | X_n) = \pi(X_n, A), \forall A \in \mathcal{F}, \forall n \in \mathbb{N}.$$

- **(Chapman-Kolmogorov Equation)**

$$\forall k, \ell \in \mathbb{N}, \pi^{(k+\ell)}(x, A) = \int_{y \in \mathcal{X}} \pi^{(\ell)}(y, A) \pi^{(k)}(x, dy), \forall x \in \mathcal{X}, A \in \mathcal{F}.$$

- Every stationary Markov chain can be expressed as

$$X_n = f(X_{n-1}, Y_n) \text{ for some function } f \text{ and i.i.d. } \{Y_n\}, \forall n \in \mathbb{N}.$$

Invariant Measure, Stopping Time

Let $\{X_n\}$ be a stationary Markov Chain with transition probability π.

- $\mu : \mathcal{F} \to [0, 1]$ is the *invariant measure* for $\{X_n\}$ if

$$\mu(A) = \int_{y \in \mathcal{X}} \pi(y, A)\mu(dy), \forall A \in \mathcal{F}.$$

- $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ is *stopping time* if

$$\{\tau \leq n\} \in \mathcal{F}_n, \forall n \in \mathbb{N} \text{ and } \mathcal{F}_n = \sigma(X_0, X_1, \cdots, X_n).$$

- Define

$$\mathcal{F}_\tau = \{A \in \mathcal{F}^\infty : A \cap \{\tau \leq n\} \in \mathcal{F}_n, \forall n \in \mathbb{N} \cup \{\infty\}\}.$$

  Then, (i) τ is $\mathcal{F}_\tau$-measurable and (ii) X_τ is $\mathcal{F}_\tau$-measurable on $\{\tau < \infty\}$.

- **(Strong Markov property)**

$$P(X_{\tau+1} \in A | \mathcal{F}_\tau) = \pi(X_\tau, A), \forall A \in \mathcal{F} \text{ for all } \{\tau < \infty\}.$$

Aperiodic Markov Chain

Let $\{X_n\}$ be a stationary MC with transition probability π and $\mathcal{X}$ be countable.

- $\{X_n\}$ is *irreducible* if $\forall x, y \in \mathcal{X}, \exists n \in \mathbb{N}$ such that $\pi^{(n)}(x, y) > 0$.

Let $\tau_x = \inf\{n \geq 1 : X_n = x\}$.

- A state $x \in \mathcal{X}$ is called *transient* if $P(\tau_x < \infty | X_0 = x) < 1$.
- A state $x \in \mathcal{X}$ is called *recurrent* if $P(\tau_x < \infty | X_0 = x) = 1$.
  More precisely, x is $\begin{cases} \textit{positive recurrent}, & \textit{if } \mathbb{E}[\tau_x | X_0 = x] < \infty \\ \textit{null recurrent}, & \textit{if } \mathbb{E}[\tau_x | X_0 = x] = \infty \end{cases}$.

For any $x \in \mathcal{X}$, let d_x be the gcd of $D_x = \{n \in \mathbb{N} : \pi^{(n)}(x, x) > 0\}$.

(Theorem) For any $x, y \in \mathcal{X}$, $\pi^{(n)}(x, y), \pi^{(m)}(x, y) > 0$ for some m, n implies $d_x = d_y$.

(Theorem) For any irreducible chain, all states have the same type and period d.

- An irreducible MC is *aperiodic* if $d = 1$.

Ergodic Theorem

(Theorem) Let $\{X_n\}$ be irreducible, recurrent, aperiodic on $(\mathcal{X}, \mathcal{F})$ and π be the transition probability.
If $\{X_n\}$ is null recurrent, then

$$\lim_{n\to\infty} \pi^{(n)}(x,y) = 0, \forall x, y \in \mathcal{X}.$$

If $\{X_n\}$ is positive recurrent, then

$$\lim_{n\to\infty} \pi^{(n)}(x,y) = \mu(y), \forall x, y \in \mathcal{X},$$

where $\mu(y) = \frac{1}{\mathbb{E}[\tau_y | X_0 = y]}$ is the limiting distribution.

(Ergodic Theorem) Let $\{X_n\}$ be irreducible, positive recurrent, aperiodic on $(\mathcal{X}, \mathcal{F})$ and $f \in L^1(\mu)$ for $\mu(x) = \frac{1}{\mathbb{E}[\tau_x | X_0 = x]}$. Then,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j \le n} f(X_j) \overset{a.s.}{=} \sum_{x \in \mathcal{X}} f(x)\mu(x).$$

Outline

Martingale

Given a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$ *filtered* probability space.

(Definition) $\{X_n\}$ is a *martingale* if

(i) $\underbrace{X_n \text{ is } \mathcal{F}_n\text{-measurable}, \forall n \geq 0.}_{\text{adapted}}$

(ii) $X_n \in L^1$, $\forall n \geq 0$.

(iii) $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \overset{a.s.}{=} X_n$, $\forall n \geq 0$.

Similarly, $\{X_n\}$ is a *sub-/super- martingale* if replacing $\overset{a.s.}{=}$ with $\overset{a.s.}{\geq}$ / $\overset{a.s.}{\leq}$ in (iii).

(Doob's Martingale) Given $X \in L^1$ and $\{\mathcal{F}_n\}$, $\{X_n\}$ with $X_n = \mathbb{E}[X|\mathcal{F}_n]$ is a m.g.

Relationship between martingale and submartingale:

- **(From Jensen's inequality)** Let $\{(X_n, \mathcal{F}_n)\}$ be a martingale and ϕ be a convex function. If $\phi(X_n) \in L^1$ for all $n \geq 0$, then $\{(\phi(X_n), \mathcal{F}_n)\}$ is a submartingale.
- **(Doob's Decomposition)** Every submartingale $\{(X_n, \mathcal{F}_n)\}$ can be uniquely written as $X_n = M_n + A_n + X_0$, where $\{(M_n, \mathcal{F}_n)\}$ is a martingale and $\{A_n\}$ with $A_0 = 0$ and $\underbrace{A_n \geq A_{n-1}, \forall n \geq 1}_{\text{increasing}}$ and $\underbrace{A_n \text{ is } \mathcal{F}_{n-1}\text{-measurable}, \forall n \geq 1}_{\text{predictable}}$.

Optional Sampling Theorem

Martingale Transformation: let $M = \{(M_n, \mathcal{F}_n)\}$ be a (sub-/super-) martingale and $A = \{(A_n, \mathcal{F}_n)\}$ be a predictable process.

- *(sub-/super-) martingale transform* of M by A is

$$X_n = \sum_{k \leq n} A_k (M_k - M_{k-1}), \forall n \geq 1, X_0 = 0.$$

- The above $\{X_n\}$ is a (sub-/super-) martingale if $A_n \geq 0, X_n \in L^1, \forall n$.

(Optional Sampling Theorem) Let $\{(X_n, \mathcal{F}_n)\}$ be a martingale and τ be a stopping time. Then, $\{(X_{n\wedge\tau}, \mathcal{F}_n)\}$ is a martingale and $\mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}[X_0]$, $\forall n \geq 0$.

- The above holds for sub-/super- martingale by replacing with $\mathbb{E}[X_{n\wedge\tau}] \geq \mathbb{E}[X_0]$ and $\mathbb{E}[X_{n\wedge\tau}] \leq \mathbb{E}[X_0]$, respectively.

Martingale Convergence Theorem

For any $a, b \in \mathbb{R}$ with $a < b$, let $U_n(a,b) = \sup\{m \geq 1 : \tau_{2m} \leq n\}$,where $\tau_0 = 0$, $\tau_{2k+1} = \inf\{n \geq \tau_{2k} : X_n \leq a\}$ and $\tau_{2k+2} = \inf\{n \geq \tau_{2k+1} : X_n \geq b\}$.

(Upcrossing Inequality) For a supermartingale $\{(X_n, \mathcal{F}_n)\}$,

$$\mathbb{E}[U_n(a,b)] \leq \frac{\mathbb{E}[(X_n - a)^-]}{b-a}, \forall n \geq 1, a > b.$$

(Theorem) Let $\{(X_n, \mathcal{F}_n)\}$ be a supermartingale with $\sup_n \mathbb{E}[X_n^-] < \infty$. Then, $X_\infty = \lim_{n\to\infty} X_n \in L^1$.

$\{X_n\}$ is *uniformly integrable* if $\lim_{M\to\infty} \sup_n \mathbb{E}[|X_n| \mathbf{1}_{\{|X_n|>M\}}] = 0$.

(Martingale Convergence Theorem) Let $X = \{(X_n, \mathcal{F}_n)\}$ be a martingale. Then, X contains the last element $X_\infty \in L^1$ and $\mathbb{E}[X_\infty|\mathcal{F}_n] = X_n$, $\forall n \geq 1$ iff $\{X_n\}$ is uniformly integrable.

- The above hold for sub-/super- martingale by replacing with $\{X_n^+\}$ or $\{X_n^-\}$ is uniformly integrable, respectively.