

Probability and Measure

1 Boolean Algebras and Finitely Additive Measures

- Boolean algebra \mathcal{B}
 - $\emptyset \in \mathcal{B}$
 - stable under finite union
 - stable under complementation

Example.

- (i) *trivial Boolean algebra*
- (ii) *discrete Boolean algebra*
- (iii) *family of constructable sets*

- constructable sets — finite union of locally closed sets from topological space
- locally closed sets — $O \cap C$ where O open, C closed
- finitely additive measure, m
 - $m(\emptyset) = 0$
 - $m(E \sqcup F) = m(E) + m(F)$
- sub-additive — $m(E \cup F) \leq m(E) + m(F)$
- monotone — $E \subset F \Rightarrow m(E) \leq m(F)$

Fact. *finitely additive measure is sub-additive and monotone*

- counting measure

2 Jordan Measure on \mathbb{R}^d

- box $B = I_1 \times \cdots \times I_d$
- elementary subset — finite union of boxes
- volume of box, $|B|$
- $\mathcal{E}(B)$ — family of elementary subsets of box B

Proposition 2.1. *Fixed B , then*

(i) $\mathcal{E}(B)$ Boolean algebra

(ii) every $E \in \mathcal{E}(B)$ finite union of disjoint boxes

(iii) volume well defined

$$- m(E) = \sum |B_i| \text{ for } E = \bigsqcup B_i$$

Fact. m finitely additive measure on $(B, \mathcal{E}(B))$

$$- \text{Jordan measurable} \text{ --- For all } \epsilon > 0, \exists \text{ elementary } E \subset A \subset F \text{ st } m(F \setminus E) < \epsilon$$

Fact. Jordan measurable subsets bounded

$$- m(A) \text{ for Jordan measurable } A \text{ ---}$$

$$m(A) = \inf\{m(F) : A \subset F, F \text{ elementary}\} = \sup\{m(F) : A \supset E, E \text{ elementary}\}$$

$$- \mathcal{J}(B) \text{ --- family of Jordan measurable subsets of box } B$$

Proposition 2.2. Fixed B , then

(i) $\mathcal{J}(B)$ Boolean algebra

(ii) m finitely additive measure on $(B, \mathcal{J}(B))$

Fact. $E \subset$ finite interval $[a, b] \subset \mathbb{R}$, then E Jordan measurable iff $\mathbb{1}_E(x)$ Riemann integrable

3 Lebesgue measurable sets

$$- \text{Lebesgue outer-measure} \text{ --- } E \subset \mathbb{R}^d,$$

$$m^*(E) = \inf\{\sum |B_n| : E \subset \bigcup B_n \text{ boxes}\}$$

Fact. m^* translation invariant

$$- \text{Lebesgue measurable} \text{ --- For } \epsilon > 0, \exists C = \bigcup B_n, E \subset C \text{ st}$$

$$m^*(C \setminus E) < \epsilon$$

$$- \text{Lebesgue measurable} \text{ --- } \exists B \text{ Borel st } m^*(B \triangle E) < \epsilon$$

$$- \mathcal{L} \text{ --- family of Lebesgue measurable sets}$$

Fact. \mathcal{L} translation invariant, scales naturally

Fact. Jordan measurable \Rightarrow Lebesgue measurable

Proposition 3.1.

(i) m^* extends m

(ii) \mathcal{L} Boolean algebra, stable under countable unions

(iii) m^* countably additive on $(\mathbb{R}^d, \mathcal{L})$

Lemma 3.2. m^*

(i) monotone — $A \subset B \Rightarrow m^*(A) \leq m^*(B)$

(ii) countably sub-additive — $m^*(\bigcup A_n) \leq \sum m^*(A_n)$

Fact. Jordan measure countably additive on Jordan measurable set

– continuity property — E_n non-increasing, empty intersection $\Rightarrow \lim m(E_n) = 0$

Lemma 3.3. Jordan measure has continuity property on elementary sets

Lemma 3.4. Elementary sets E_n decreasing, $A = \bigcap E_n$, then

(i) A Lebesgue measurable

(ii) $m(E_n) \rightarrow m^*(A)$

Fact. countable intersection of elementary sets Lebesgue measurable

Corollary 3.5. open and closed subsets Lebesgue measurable

– null set — $m^*(E) = 0$

Lemma 3.6. null set Lebesgue measurable

Proposition 3.7. E Lebesgue measurable, then \exists closed C , open O st

(i) $C \subset E \subset O$

(ii) $m^*(O \setminus C) < \epsilon$

Fact. E can be written as $(\bigcup C_n) \sqcup N$ or $(\bigcap O_n) \setminus N$

Example. Vitali's counter example — E set of representatives $E = \{x + \mathbb{Q}\} \subset [0, 1]$

(i) m^* not additive on all subsets of \mathbb{R}^d

(ii) E not Lebesgue measurable

4 Abstract Measure Theory

– σ -algebra — Boolean algebra, stable under countable unions

– measurable space, (X, \mathcal{A})

– measure μ —

(i) $\mu(\emptyset) = 0$

(ii) countably additive

– measure space, (X, \mathcal{A}, μ)

Example.

(i) $(\mathbb{R}^d, \mathcal{L}, m)$

(ii) $m_0(E) = m(A_0 \cap E)$ for fixed $A_0 \in \mathcal{L}$

(iii) $(X, 2^X, \#)$, $\#$ counting measure

(iv) $(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ where $\mu(I) = \sum_{i \in I} a_i$ for fixed $(a_n)_{n \geq 1}$

Proposition 4.1. (X, \mathcal{A}, μ) measure space

(i) μ monotone

(ii) μ countably sub-additive

(iii) upward monotone convergence — E_n increasing, then $\mu(\bigcup E_n) = \lim \mu(E_n) = \sup \mu(E_n)$

(iv) downward monotone convergence — $\mu(E_1) < \infty$, E_n decreasing, then $\mu(\bigcap E_n) = \lim \mu(E_n) = \inf \mu(E_n)$

– finite — $\mu(X) < \infty$

– σ -finite — $X = \bigcup E_n$, $\mu(E_n) < \infty$

– probability space

– probability measure

– σ -algebra generated by \mathcal{F} , $\sigma(\mathcal{F})$ — \mathcal{F} family of subsets

Example.

(i) $X = \sqcup X_i$

(ii) X countable, \mathcal{F} singletons

– Borel σ -algebra, $\mathcal{B}(X)$ — X topological space, generated by all open subsets

– Borel sets

Fact. $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{L}$

Fact. $\mathcal{B}(\mathbb{R}^d)$ strictly smaller than \mathcal{L} — every subset of null sets is null

Fact. $\mathcal{B}(X)$ (σ -algebra) usually larger than family of constructable sets (Boolean algebra)

– Boolean algebra generated by \mathcal{F} , $\beta(\mathcal{F})$

– explicitly described — elements of $\beta(\mathcal{F})$ are finite unions of $F_1 \cap \dots \cap F_n$, F_i or \bar{F}_i in \mathcal{F}

Myth. Borel hierarchy

– Borel measure — measure on $\mathcal{B}(X)$

Setting 1. X set, \mathcal{B} Boolean algebra, μ finitely additive measure

- continuity property — under setting 1, non-increasing (E_n) , $\mu(E_1) < \infty$, empty intersection

$$\lim \mu(E_n) = 0$$

Theorem 4.2 (Caratheodory extension theorem). *Under setting 1, \mathcal{B} continuity property, μ σ -finite, then μ uniquely extends to μ^* on $\sigma(\mathcal{B})$*

- outer-measure μ^* — $\mu^*(E) = \inf \{ \sum \mu(B_i) : E \subset \bigcup B_i, B_i \in \mathcal{B} \}$
- μ^* measurable — $\exists \bigcup B_n := C$ st $\mu^*(C \setminus E) < \epsilon$
- completion of \mathcal{B} , \mathcal{B}^* — family of μ^* measurable subsets

Fact. *completion contains all null sets*

Proposition 4.3. *Under setting 1,*

- (i) \mathcal{B}^* σ -algebra containing \mathcal{B}
- (ii) μ^* countably additive on \mathcal{B}^*
- (iii) μ^* extends μ

Myth. *X compact metric space, μ probability measure on Borel σ -algebra \mathcal{B} , no atom, then \exists measure preseving measurable isomorphism between (X, \mathcal{B}^*, μ) and $([0, 1], \mathcal{L}, m)$*

5 Uniqueness of Measures

- π -system — family \mathcal{F}
 - (i) contains \emptyset
 - (ii) stable under finite intersection

Proposition 5.1 (measure uniqueness). *(X, \mathcal{A}) measurable space, μ_1, μ_2 finite measures st*

- (i) $\mu_1 = \mu_2$ on $\mathcal{F} \cup \{X\}$
- (ii) \mathcal{F} π -system st $\sigma(\mathcal{F}) = \mathcal{A}$

then $\mu_1 = \mu_2$ on \mathcal{A}

Fact. *For general measures, if $\exists F_n \subset \mathcal{F}$ st μ_1, μ_2 finite on F_n , $X = \bigcup F_n$, then uniqueness also holds*

Lemma 5.2 (Dynkin's lemma).

- (i) \mathcal{F} π -system
 - (ii) $\mathcal{F} \subset \mathcal{C}$
 - (iii) \mathcal{C} stable under complementation, disjoint countable union
- then $\sigma(\mathcal{F}) \subset \mathcal{C}$*

- translation invariant — $m(A + x) = m(A)$ for all A, x

Proposition 5.3. *Lebesgue measure unique measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ st*

- (i) translation invariant
- (ii) $m([0, 1]^d) = 1$

6 Measurable Functions

Setting 2. $(X, \mathcal{A}), (Y, \mathcal{B})$ measurable space

- $f : X \rightarrow \mathbb{R}$ measurable function
- $f : X \rightarrow Y$ measurable map

Fact. can extend to $\{\infty\}$ or $\{-\infty\}$

Fact. continuous function measurable

Fact. $E \in \mathcal{A}$ iff $\mathbb{1}_E$ measurable

- \mathbb{R} -algebra

Proposition 6.1. $(f_n)_{n \geq 1}$ measurable functions

- (i) f, g measurable $\Rightarrow g \circ f$ measurable
- (ii) Family of measurable functions form \mathbb{R} -algebra
- (iii) $\limsup f_n, \liminf f_n, \sup f_n, \inf f_n$ measurable functions

Proposition 6.2. $f = (f_1, f_2, \dots, f_d)^T$, then f measurable iff f_i measurable

- Borel measurable (or simply Borel)

Fact. f measurable

- (i) $f^{-1}(L)$ need not measurable for $L \in \mathcal{L}$
- (ii) $f(X)$ need not measurable even for f continuous

Example. (i) f sends to trivial σ -algebra

7 Integration

- simple function — $\sum^N a_i \mathbb{1}_{A_i}$ with $a_i \geq 0$

Lemma 7.1. f simple, $f = \sum a_i \mathbb{1}_{A_i} = \sum b_j \mathbb{1}_{B_j}$, then $\sum a_i \mu(A_i) = \sum b_j \mu(B_j)$

- integral $\mu(f)$ for simple f — $\mu(f) = \sum a_i \mu(A_i) = \int f d\mu$
- integral $\mu(f)$ for non-negative f — $\mu(f) = \sup\{\mu(g) : g \leq f, g \text{ simple}\}$

Proposition 7.2 (positivity). f, g non-negative measurable, then

- $f \geq g \Rightarrow \mu(f) \geq \mu(g)$
- $f \geq g, \mu(f) = \mu(g) \Rightarrow f = g$ a.e.
- $f = g$ almost everywhere

Lemma 7.3. $f \geq 0$, then \exists increasing simple functions g_n st $g_n \rightarrow f$ pointwise

Proof. $g_n(x) = 2^{-n} \lfloor 2^n(f(x) \wedge n) \rfloor$

□

Theorem 7.4 (Monotone Convergence Theorem).

(i) (f_n) non-negative, non-decreasing

(ii) let $f(x) = \lim f_n(x)$, the pointwise limit

Then, $\mu(f) = \lim \mu(f_n)$

Lemma 7.5. Fixed g simple, then $m_g(E) := \mu(\mathbb{1}_E g)$ is a measure

Lemma 7.6 (Fatou). $f_n \geq 0$, then $\mu(\liminf f_n) \leq \liminf \mu(f_n)$

– f^+, f^-

– μ -integrable $\iff \mu(|f|) < \infty$

– integral $\mu(f)$ for integrable $f \iff \mu(f) = \mu(f^+) - \mu(f^-)$

Proposition 7.7 (Linearity of integral). f, g integrable

(i) $\alpha f + \beta g$ integrable

(ii) $\mu(\alpha f + \beta g) = \alpha \mu(f) + \beta \mu(g)$

Fact. Also holds for nonnegative f, g, α, β

Theorem 7.8 (Dominated Convergence Theorem). f, f_n measurable, g integrable

(i) $|f_n(x)| \leq g(x)$

(ii) $\lim f_n(x) = f(x)$

Then,

(i) $\lim \mu(f_n) = \mu(f)$

(ii) f integrable

Fact. condition for MCT, Fatou, DCT only need to hold μ -almost everywhere

Corollary 7.9 (Exchange of \int and \sum).

(i) $f_n \geq 0$, then $\mu(\sum^\infty f_n) = \sum^\infty \mu(f_n)$

(ii) $\sum |f_n|$ μ -integrable, then

– $\sum f_n$ integrable

– $\mu(\sum f_n) = \sum \mu(f_n)$

Corollary 7.10 (Differentiation under \int sign). U open set, $f : U \times X \rightarrow \mathbb{R}$ st

(i) $f(t, \cdot)$ μ -integrable

(ii) $f(\cdot, x)$ differentiable

(iii) (domination) \exists integrable g st $\sup_t |\frac{\partial f}{\partial t}(t, x)| \leq g(x)$

Then,

(i) $\frac{\partial f}{\partial t}(t, \cdot)$ μ -integrable

(ii) let $F(t) = \int_X f(t, x) d\mu(x)$, then

(a) F differentiable

(b) $F' = \int_X \frac{\partial f}{\partial t}(t, x) d\mu(x)$

Fact. f bounded, then f Riemann integrable iff $\{x : f(x) \text{ not continuous}\}$ has Lebesgue measure 0

Fact (invariance under affine map). $g \in GL_d(\mathbb{R})$, f integrable, then $m(f \circ g) = \frac{1}{|\det g|} m(f)$

Fact. $\phi \in C^1$, then $\int f(\phi(x)) J_\phi(x) dx = \int f(x) dx$

– Radon measure — Borel measure, finite on every compact subset

Fact (Riesz Representation for locally compact spaces).

(i) μ Radon measure, let $\Lambda(f) = \mu(f)$, then $\Lambda \in C_c(X)'$

(ii) let $\Lambda \in C_c(X)'$, Λ non-negative, then \exists Radon measure μ st $\Lambda(f) = \mu(f)$

8 Product Measure

– product σ -algebra $\mathcal{A} \otimes \mathcal{B}$ — σ -algebra generated by $A \times B$ where $A \in \mathcal{A}$, $B \in \mathcal{B}$

Fact.

(i) $\{A \times B\}$ π -system

(ii) smallest σ -algebra st projection map measurable

(iii) $\mathcal{B}(\mathbb{R}^{d_1}) \otimes \mathcal{B}(\mathbb{R}^{d_2}) = \mathcal{B}(\mathbb{R}^{d_1+d_2})$ (generally not true)

– slice, E_x — $E_x = \{y : (x, y) \in E\}$

Lemma 8.1. E $\mathcal{A} \otimes \mathcal{B}$ -measurable, then E_x \mathcal{B} -measurable

Proof. start with $A \times B$, then Dynkin's lemma □

Setting 3. $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ σ -finite

Lemma 8.2. Under setting 3, f $\mathcal{A} \otimes \mathcal{B}$ -measurable, non-negative, then

(i) $f(x, \cdot)$ \mathcal{B} -measurable for every x

(ii) $g(x) := \int f(x, y) d\nu(y)$ \mathcal{A} -measurable

– product measure $\mu \otimes \nu$

Proposition 8.3. Under setting 3, then \exists unique measure σ on $\mathcal{A} \otimes \mathcal{B}$ st $\sigma(A \times B) = \mu(A)\nu(B)$

Theorem 8.4 (Fubini-Tonelli). Under setting 3,

(i) f $\mathcal{A} \otimes \mathcal{B}$ measurable, non-negative, then $\int_{X \times Y} f d\mu \otimes \nu = \int_X \left(\int_Y f d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f d\mu(x) \right) d\nu(y)$

(ii) f $\mu \otimes \nu$ -integrable, then

(a) $f(x, \cdot)$ ν -integrable for μ -almost every x

(b) $f(\cdot, y)$ μ -integrable for ν -almost every y

(c) above also holds

Fact. justify Fubini, just need $f(x, \cdot), f(\cdot, y)$ integrable (???)

9 Probability Theory

- universe Ω
- outcome ω
- events \mathcal{F}
- probability measure \mathbb{P}
- random variable X
- expectation \mathbb{E}
- probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- law of X / distribution of X — Borel measure $\mu_X(A) = \mathbb{P}(X \in A)$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
- image measure $f_*\mu$ — $f_*\mu(A) = \mu(f^{-1}(A))$
- distribution function of X , F_X — $F_X(t) = \mathbb{P}(X \leq t)$

Proposition 9.1. F_X

(i) non-decreasing

(ii) right continuous

(iii) F_X determines μ_X uniquely

- Lebesgue-Stieltjes measure μ_F

Proposition 9.2. Given F non-decreasing, right continuous, $\lim_{-\infty} F(t) = 0$, $\lim_{\infty} F(t) = 1$, then \exists unique Borel measure μ_F st $F(t) = \mu_F((-\infty, t])$

Proof. One approach using Caratheodory. Another shown as follow. □

- "inverse function" g — $g(y) = \inf\{t : F(t) \geq y\}$

Lemma 9.3. g

(i) non-decreasing

(ii) left continuous

(iii) $g(y) \leq t$ iff $y \leq F(t)$

Fact. let m Lebesgue measure on $(0, 1)$, set $\mu(A) = g_*m(A) = m(g^{-1}(A))$, then $\mu = \mu_F$

Proposition 9.4. μ Borel probability measure, then $\exists(\Omega, \mathcal{F}, \mathbb{R})$, r.v. X st $\mu = \mu_X$

Fact. Can take $\Omega, \mathcal{F}, \mu = (0, 1)$, Borel σ -algebra, \mathbb{P}

- density

Example.

(i) uniform distribution

(ii) exponential distribution

(iii) gaussian distribution

(iv) Dirac mass

- mean
- moment of order k
- variance

10 Independence

- events (A_i) mutually independent — every finite $F \subset \mathbb{N}$, $\mathbb{P}(\bigcap_F A_i) = \prod_F \mathbb{P}(A_i)$

Fact. (A_i) independent $\Rightarrow (B_i)$ independent where $B_i = A_i$ or A_i^c

- σ -subalgebras (\mathcal{A}_i) mutually independent — $\mathcal{A}_i \subset \mathcal{F}$, every $A_i \in \mathcal{A}_i$, (A_i) mutually independent

Fact. $\Pi_i \subset \mathcal{A}_i$ π -system, $\sigma(\Pi_i) = \mathcal{A}_i$, then suffices just check $A_i \in \Pi_i$

- $\sigma(X)$
- random variables (X_i) mutually independent — $(\sigma(X_i))$ independent

Fact. Equivalence to every finite $F \subset \mathbb{N}$

- $\mathbb{P}(\bigcap_F X_i \leq t_i) = \prod_F \mathbb{P}(X_i \leq t_i)$
- $\mu_{(X_{i_1}, \dots, X_{i_m})} = \mu_{X_{i_1}} \otimes \dots \otimes \mu_{X_{i_m}}$

Fact. (X_i) independent $\Rightarrow (f_i(X_i))$ independent

Proposition 10.1. X, Y independent, non-negative (or integrable), then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

Setting 4. $\{(\Omega_i, \mathcal{F}_i, \nu_i)\}$

- cylinder set — $A \times \prod_{i>n} \Omega_i$ where $A \subset \prod_n \Omega_i$, $A \in \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n$

- Boolean algebra of cylinder set \mathcal{C}
- infinite product measure

Proposition 10.2. $\Omega = \prod \Omega_i$, $\mathcal{F} = \sigma(\mathcal{C})$, then \exists unique probability measure ν on (Ω, \mathcal{F}) st

$$\nu(B) = \nu_1 \otimes \cdots \otimes \nu_n(A) \quad \text{for every cylinder set } B$$

Fact. more general theorem Kolmogorov extension theorem

- $\limsup A_n$ — $\bigcap \bigcup A_n$ (infinitely often)

Lemma 10.3 (1st Borel Cantelli). $\sum \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\limsup A_n) = 0$

Lemma 10.4 (2nd Borel Cantelli). $\sum \mathbb{P}(A_n) = \infty$, (A_n) mutually independent, then $\mathbb{P}(\limsup A_n) = 1$

Fact. independence condition in 2nd Cantelli can be relaxed

- pairwise independence
- small correlation between events
- $(\Omega, \mathcal{F}, \mathbb{P})$ probability space
- random process / stochastic process (X_n)
- n-th term of the associated filtration $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$
- tail σ -algebra $\mathcal{T} = \bigcap \sigma(X_n, X_{n+1}, \dots)$

Theorem 10.5 (Kolmogorov 0-1 law). (X_n) mutually independent, then $\mathbb{P}(A) \in \{0, 1\}$ for all $A \in \mathcal{T}$

- Cauchy-Schwarz — $\mathbb{E}|XY| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$
- Markov's inequality — $\lambda \mathbb{P}(X \geq \lambda) \leq \mathbb{E}(X)$ for $X \geq 0, \lambda \geq 0$
- Chebychev's inequality — $\lambda^2 \mathbb{P}(|Y - \mathbb{E}(Y)| \geq \lambda) \leq \text{Var}(Y)$ for $\lambda \geq 0$

Theorem 10.6 (Strong law of large number). (X_n) i.i.d., $\mathbb{E}|X_1| < \infty$, then $\bar{X}_n \xrightarrow[a.s.]{} \mathbb{E}X_1$

Fact. $\mathbb{E}(X^4) < \infty \Rightarrow \mathbb{E}((X - \mathbb{E}X)^4) < \infty$ (Jensen on X^4)

Fact. $\mathbb{E}(|X|^n) < \infty \Rightarrow \mathbb{E}(|X|^k) < \infty$ for $k \leq n$

11 Convergence of Random Variables

- probability measures μ_n converge weakly — \forall bounded, continuous f , $\mu_n(f) \rightarrow \mu(f)$
- ◊ sequences (X_n)
- almost surely (a.s.) — $X_n(\omega) \rightarrow X(\omega)$ for \mathbb{P} almost every ω
- in probability (in measure) — $\mathbb{P}(\|X_n - X\| > \epsilon) \rightarrow 0$
- in law (in distribution) — μ_{X_n} converge weakly to μ_X

Proposition 11.1. *almost surely \Rightarrow in probability \Rightarrow in distribution*

Fact. $X_n \rightarrow X$ in law iff $F_{X_n}(x) \rightarrow F_X(x)$

Fact. To prove μ_n converge weakly to μ , suffice to check $f \in C_c^\infty$

Counter Example.

- weakly \nRightarrow in prob — i.i.d. X_n with same distribution
- in prob \nRightarrow a.s. — moving bump $\mathbb{1}_{[k/n, (k+1)/n]}$

Proposition 11.2. $X_n \rightarrow X$ in prob, then \exists subsequence $X_{n_j} \rightarrow X$ a.s.

- converge in L^1 — integrable X_n , $\mathbb{E}\|X_n - X\| \rightarrow 0$

Proposition 11.3. $L^1 \Rightarrow$ in probability

Counter Example.

- in prob \nRightarrow in L^1 — $X_n = n\mathbb{1}_{[0, 1/n]}$
- bounded — $X_n \leq C$ for constant C independent of n

Fact. If (X_n) bounded, in prob \Rightarrow in L^1

Proof. Passing to subsequence, a.s. convergence. Then DCT □

- uniformly integrable (U.I.) — integrable (X_n) , $\lim_M \limsup_n \mathbb{E}(\|X_n\| \mathbb{1}_{\|X_n\| > M}) = 0$
- dominated — $X_n \leq Y$ for integrable Y , all n

Fact. dominated \Rightarrow U.I.

- bounded in L^p — $\sup_n \mathbb{E}\|X_n\|^p < \infty$

Fact. bounded in L^p for $p > 1 \Rightarrow$ U.I.

Theorem 11.4. (X_n) integrable, then following equivalent:

- (i) $X_n \rightarrow X$ in L^1 , X integrable
- (ii) $X_n \rightarrow X$ in prob, X_n U.I.

Lemma 11.5. Y integrable, (X_n) U.I. , then $(X_n + Y)$ U.I.

12 L^p spaces

Setting 5. $(\Omega, \mathcal{A}, \mathbb{P})$ probability space, I open interval, $X : \Omega \rightarrow I$, $\phi : I \rightarrow \mathbb{R}$

Proposition 12.1 (Jensen's inequality). *Under setting 5, X integrable, ϕ convex, then $\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}(X))$*

– convex

Lemma 12.2. *convex iff $\phi = \sup_{\mathcal{F}} l$ where \mathcal{F} family of affine linear forms*

Fact. $\phi(X)^-$ always integrable

– L^p norm $\|f\|_p$

– L^∞ norm $\text{essup } |f|$

Fact. let $g = f \mathbb{1}_{f \leq \|f\|_\infty}$, then $\sup g = \text{essup } |g|$

Proposition 12.3 (Minkowski inequality). $p \in [1, \infty]$, then $\|f + g\|_p \leq \|f\|_p + \|g\|_p$

Setting 6. $\frac{1}{p} + \frac{1}{q} = 1$

Proposition 12.4 (Holder's inequality). $\int |fg| d\mu \leq \|f\|_p \|g\|_q$
Equality holds for finite p, q when $\alpha|f|^p = \beta|g|^q$ for μ -a.e.

Lemma 12.5 (Young's inequality for product). $a, b \geq 0$, then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. Equality iff $a^p = b^q$

– $L^p(X, \mathcal{A}, \mu) \longrightarrow \left\{ \|f\|_p < \infty \right\}$

– $f \equiv g \longrightarrow f = g \text{ } \mu\text{-a.e.}$

Lemma 12.6. \equiv equivalence relation, stable under addition and multiplication

– L^p space $\longrightarrow L^p(X, \mathcal{A}, \mu) / \equiv$

Proposition 12.7 (completeness of L^p spaces).

(i) L^p space with $\|\cdot\|_p$ normed vector space

(ii) complete (a.k.a. Banach space)

Proposition 12.8 (Approximation by simple functions). $p \in [1, \infty)$, V linear span of simple functions, then $V \cap L^p$ dense in L^p

Fact. linear span as we need $g^+ - g^-$

Fact. For $(\mathbb{R}^d, \mathcal{L}, m)$, $C_c^\infty(\mathbb{R}^d)$ dense in L^p

Fact. $\mu(X) < \infty$, then $L^{p'} \subset L^p$ for $p' \geq p$

Fact. X discrete, countable, then $L^{p'} \subset L^p$ for $p' \leq p$

13 Hilbert Spaces and L^2 methods

- Hermitian inner product — \mathbb{C}
- sesquilinear form — we pick linear in first argument
- Euclidean inner product — \mathbb{R}
- bilinear symmetric form

Lemma 13.1. (i) $\|\alpha x\| = |\alpha|\|x\|$

(ii) *Cauchy-Schwarz inequality* $|\langle x, y \rangle| \leq \|x\|\|y\|$

(iii) *triangle inequality* $\|x + y\| \leq \|x\| + \|y\|$

(iv) *Parallelogram identity* $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$

Corollary 13.2. $(V, \|\cdot\|)$ normed vector space

- Hilbert space — complete Hermitian/Euclidean vector space
- orthogonal projection — unique $\pi_{\mathcal{C}}(x)$ st $\|x - \pi_{\mathcal{C}}(x)\| = \inf_{\mathcal{C}} \|x - c\|$

Proposition 13.3 (orthogonal projection on closed convex sets). \mathcal{H} Hilbert, \mathcal{C} closed convex, then \exists orthogonal projection

Corollary 13.4. V closed vector subspace, then $\mathcal{H} = V \oplus V^\perp$

Fact. V^\perp closed

- bounded linear form

Fact. bounded iff continuous

Theorem 13.5 (Riesz representation theorem for Hilbert spaces). \mathcal{H} Hilbert space, l bounded linear form, then \exists unique v_0 st $l(\cdot) = \langle \cdot, v_0 \rangle$

14 Conditional Expectation

Setting 7. $(\Omega, \mathcal{F}, \mathbb{P})$ probability space, $\mathcal{G} \subset \mathcal{F}$ σ -subalgebra, X integrable

- conditional expectation $\mathbb{E}(X|\mathcal{G})$

Proposition 14.1. \exists (a.s.) unique conditional expectation Y st

- (i) \mathcal{G} -measurable
- (ii) integrable
- (iii) $\mathbb{E}(\mathbb{1}_A X) = \mathbb{E}(\mathbb{1}_A Y)$

Proposition 14.2.

- (i) *linearity* — $\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha \mathbb{E}(X|\mathcal{G}) + \beta \mathbb{E}(Y|\mathcal{G})$ a.s.

- (ii) positivity — $X \geq 0$ a.s. , then $\mathbb{E}(X|\mathcal{G}) \geq 0$ a.s.
- (iii) tower property — $\mathcal{H} \subset \mathcal{G}$ σ -subalgebra, then $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H})$ a.s.
- (iv) independence — X independent of \mathcal{G} , then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ a.s.
- (v) X \mathcal{G} -measurable, then $\mathbb{E}(X|\mathcal{G}) \leq 0$ a.s.
- (vi) Z \mathcal{G} -measurable, bounded, then $\mathbb{E}(XZ|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G})$
- (vii) MCT, Fatou, DCT holds for $\mathbb{E}(\cdot|\mathcal{G})$

15 Fourier Transform on \mathbb{R}^n

- Fourier transform $\widehat{f}(u)$ — $f \in L^1$, $\widehat{f}(u) = \int f(x)e^{i\langle u, x \rangle} dx$

Proposition 15.1.

- (i) $|\widehat{f}(u)| \leq \|f\|_1$
- (ii) $\widehat{f} \in C^0$

- characteristic function of $\widehat{\mu}$ — μ finite Borel measure, $\widehat{\mu}(u) = \int e^{i\langle u, x \rangle} d\mu(x)$

Proposition 15.2.

- (i) $|\widehat{\mu}(u)| \leq \mu(\mathbb{R}^d)$
- (ii) $\widehat{\mu} \in C^0$

Example. For Gaussian measure, let $g(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$, then $\widehat{g}(u) = \sqrt{2\pi}g(u)$

- self-dual
- character $\chi_u(x) = e^{-i\langle u, x \rangle}$

Theorem 15.3 (Fourier Inversion Formula).

- (i) μ finite Borel measure, $\widehat{\mu} \in L^1$, then
 - \exists density $\phi \in C^0$ st $d\mu = \phi(x)dx$
 - $\phi(x) = \frac{1}{(2\pi)^d} \widehat{\widehat{\mu}}(-x)$
- (ii) $f, \widehat{f} \in L^1$, then $f(x) = \frac{1}{(2\pi)^d} \widehat{\widehat{f}}(-x)$ a.e.

Fact. simple sufficient condition for $\widehat{f} \in L^1$: $f \in C^2$ and f, f', f'' integrable

- convolution $\mu * \nu$ — $\Phi_*(\mu \otimes \nu)$ where $\Phi(x, y) = x + y$
- convolution $f * g$ — $f, g \in L^1$, then $f * g = \int f(x - t)g(t)dt$

Fact. $\mu_X * \nu_Y$ equivalent to law of $X + Y$

Fact. $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$

– G_σ — density of $\mathcal{N}(0, \sigma^2 I_d)$

– $\tau_t(f)(x)$ — $f(x + t)$

Proposition 15.4 (Gaussian approximation). $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty)$, then $\lim \|f * G_\sigma - f\|_p = 0$

Lemma 15.5 (continuity of translation in L^p). $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty)$, then $\lim \|\tau_t(f) - f\|_p = 0$

Proposition 15.6.

(i) μ, ν Borel prob measures, then $\widehat{\mu * \nu} = \widehat{\mu} \widehat{\nu}$

(ii) $f, g \in L^1$, then $\widehat{f * g} = \widehat{f} \widehat{g}$

Theorem 15.7 (Levy's criterion). X_n r.v. , then following are equivalent:

(i) $X_n \rightarrow X$ in law

(ii) $\lim_n \widehat{\mu_{X_n}}(u) = \widehat{\mu_X}(u)$ for all u

Fact. $\widehat{\mu_X} = \widehat{\mu_Y}$ iff $\mu_X = \mu_Y$

Example. $\mathcal{N}(m, \sigma^2) \rightarrow \delta_m$ weakly

Fact. $\widehat{\mu_X}(0) = 1$

– positive definite — $\forall u_1, \dots, u_N \in \mathbb{R}^d, \forall t_1, \dots, t_N \in \mathbb{C}, \sum t_i \bar{t}_j \widehat{\mu_X}(u_i - u_j)$ real and ≥ 0

– normalize — $f(0) = 1$

Fact (Boucher's theorem). f normalized continuous positive-definite, then \exists unique probability measure μ st $f = \widehat{\mu}$

– linear isometry

Theorem 15.8 (Plancherel formula). $f, g \in L^1 \cap L^2$, then

(i) $\widehat{f} \in L^2$

(ii) $\|\widehat{f}\|_2 = (2\pi)^{d/2} \|f\|_2$

(iii) $\langle \widehat{f}, \widehat{g} \rangle_{L^2} = (2\pi)^d \langle f, g \rangle_{L^2}$

(iv) $\mathcal{F} : L^1 \cap L^2 \rightarrow L^2$ where $\mathcal{F}(f) = \frac{1}{(2\pi)^{d/2}} \widehat{f}$

• extends uniquely to linear isometry of L^2

• $\mathcal{F} \circ \mathcal{F}(f)(x) = f(-x)$

Fact. smoothness/decay barter

Fact. uncertainty principle

Fact. Schwarz space

16 Gaussian random variables

- gaussian (\mathbb{R}^d) — can be degenerated
- mean
- covariance matrix
- correlation coefficients

Fact. $\mathcal{N}(m, 0) := \delta_m$

Proposition 16.1. *law of Gaussian determined by mean and cov*

Proof. can assume 1-d Gaussian determined by mean and var □

Fact. *cov matrix positive semi-definite symmetric*

Proposition 16.2. N_i i.i.d $\mathcal{N}(0, 1)$, $A \in M_d(\mathbb{R})$, $b \in \mathbb{R}^d$, then

- (i) $AN + b \sim \mathcal{N}_d(b, AA^*)$
- (ii) every Gaussian $X = AN + b$ for some A, b

Fact. $\mathcal{N}(0, \lambda I_d)$ only Borel probability law with

- invariant under rotation
- independent coordinates

Proposition 16.3. $X = (X_1, \dots, X_n)$ Gaussian vector, then following equivalent:

- (i) X_i independent r.v.
- (ii) X_i pairwise independent
- (iii) Cov matrix diagonal

Theorem 16.4 (Central Limit Theorem). (X_n) i.i.d with common law μ , finite moment of order 2, then $\sqrt{n}(\bar{X}_n - \mathbb{E}(X_1)) \rightarrow \mathcal{N}(0, \text{Cov}(X_1))$ in law

Fact. $\widehat{\mu_Y}(tu) = \widehat{\mu_{\langle Y, u \rangle}}(t)$

17 Introduction to Ergodic Theory

Setting 8. measurable map $T : X \rightarrow X$

- measure preserving map — $T_*\mu = \mu$
- measure preserving system — measure space (X, \mathcal{A}, μ) with measure preserving map T
- T -invariant function — measurable $f = f \circ T$
- T -invariant subset — $T^{-1}A = A$
- invariant σ -algebra \mathcal{I} — $\{A : T^{-1}A = A\}$

Lemma 17.1. f measurable function, then following equivalent:

- (i) f T -invariant
- (ii) f measurable wrt \mathcal{I}
 - ergodic wrt T — T measure preserving, $\forall A \in \mathcal{I}, \mu(A) = 0$ or $\mu(A^c) = 0$

Fact. ergodic kind of irreducibility condition

Lemma 17.2. (X, \mathcal{A}, μ, T) measure preserving system, then following equivalent:

- (i) T ergodic
- (ii) every \mathcal{I} -measurable f , $f(x) \equiv a$ μ -a.e. for some a

Setting 9 (circle rotation). $X = \mathbb{R}/\mathbb{Z}$, $T(x) = x + a$

Proposition 17.3. T ergodic iff a irrational

Fact (Parseval's formula). $f(x) = \sum \hat{f}(n)e^{-i2\pi nx}$

Setting 10 (times 2 map on the circle). $X = \mathbb{R}/\mathbb{Z}$, $T_2(x) = 2x \mod \mathbb{Z}$

Proposition 17.4. T_2 ergodic

18 Canonical Model for Stochastic Process

Setting 11. (X_n) \mathbb{R}^d r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ probability space

- space of sequences $(\mathbb{R}^d)^{\mathbb{N}}$
- sample path map Φ — $\Phi(\omega) = (X_n(\omega))$
- shift map T — $T((x_n)_n) = (x_{n+1})_n$
- shift space
- coordinate functions x_k — $x_k((x_n)_n) = x_k$

Setting 12. $X = ((\mathbb{R}^d)^{\mathbb{N}})$ endowed σ -algebra $\mathcal{A} = \sigma(x_k)$

- cylinder set
- π_F — F finite set of indices
- law of the stochastic process μ — $\mu = \Phi_*\mathbb{P}$ prob measure on (X, \mathcal{A})
- canonical model — (X, \mathcal{A}, μ, T)

Proposition 18.1. (X_n) stochastic process, (X, \mathcal{A}, μ, T) canonical model, then following equivalent:

- (i) (X, \mathcal{A}, μ, T) measure preserving
- (ii) joint law of $(X_n, X_{n+1}, \dots, X_{n+k})$ independent of n
 - stationary

Proposition 18.2. (X_n) i.i.d. process, then

- (i) stationary
- (ii) canonical model ergodic
 - Bernoulli shift $\mu = \nu^{\otimes \mathbb{N}}$ — ν law of X_1

19 Mean Ergodic Theorem

Setting 13. (X, \mathcal{A}, μ, T) prob measure preserving system

$$- S_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$$

Theorem 19.1 (Mean ergodic theorem in L^2). $f \in L^2(X, \mathcal{A}, \mu)$, then
 $\exists T$ -invariant $\bar{f} = \mathbb{E}(f|\mathcal{I})$ st $S_n(f) \rightarrow \bar{f}$ in L^2

- adjoint A^* — \mathcal{H} Hilbert space, A bounded linear map, then by Riesz $\exists A^*$ st $\langle Ax, y \rangle = \langle x, A^*y \rangle$
- involutive — $A^{**} = A$

Fact.

$$(i) \|A^*\| = \|A\|$$

$$(ii) \|AA^*\| = \|A\|^2$$

$$- U(f) = f \circ T$$

$$- \text{co-boundaries } W = \{\phi - U\phi\}$$

Fact. \bar{f} orthogonal projection onto $W^\perp = \{g = Ug\}$

Corollary 19.2 (Mean ergodic theorem in L^p). $p \in [1, \infty)$, $f \in L^p$, then
 $\exists T$ -invariant $\bar{f} = \mathbb{E}(f|\mathcal{I})$ st $S_n(f) \rightarrow \bar{f}$ in L^p

$$- E_t = \{x : \sup_n S_n f(x) > t\}$$

Theorem 19.3 (Maximal ergodic theorem). $f \in L^1$, then $\mu(E_t) \leq \frac{1}{t} \|f\|_1$

Lemma 19.4 (the maximal inequality). $f \in L^1$, $f_n = nS_n f$, $f_0 = 0$, $P_N = \{x : \max_{0 \leq x \leq N} f_n(x) > 0\}$,
then $\int_{P_N} f d\mu \geq 0$

Theorem 19.5 (Pointwise ergodic theorem). $f \in L^1$, then $S_n(f) \rightarrow \bar{f}$ μ -a.e.

Fact. ergodic $\Rightarrow \mathbb{E}(f|\mathcal{I}) = \mathbb{E}(f)$

Fact. $f = \mathbb{1}_A$, orbit $\{T^n x\}$, then time spent in almost every orbit equidistributed

Corollary 19.6 (Strong law of large number). (X_n) i.i.d., $\mathbb{E}(\|X_1\|) < \infty$, then $\frac{1}{n} \sum S_i \rightarrow \mathbb{E}(X_1)$ a.s.

- $\mathcal{I}(X)$ — family of all T -invariant probability measures
- extremal — $\mu \in \mathcal{I}$, $\nexists \mu_1, \mu_2$ st $\mu = t\mu_1 + (1-t)\mu_2$

Proposition 19.7. $\mu \in \mathcal{I}(X)$, then ergodic iff extremal

20 Fun fact

- G_δ -set — countable intersection of open sets
- locus of continuity — set of continuous point

Fact. locus of continuity is G_δ -set

- Cantor function (a.k.a Devil staircase)