

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

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

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

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

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

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

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

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

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

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

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 — $\mathcal{A}_i \subset F$, 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

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

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*