LTCC: Measure Theoretic Probability

Alexander Gnedin

Queen Mary, University of London

website: google for LTCC - London Taught Course Centre (Queen Mary Courses)

1. Basics of Measure Theory
2. Random Variables, Independence, Integration
3. Conditioning, Martingales and Convergence
4. The Brownian Motion, Functional Convergence
5. The Invariance Principle

MTP, Lecture 1: Basics of Measure Theory

Introduction

- For ground set Ω how can we *measure* subsets $A \subset \Omega$?

If Ω is finite or countable (e.g. $\{1, \ldots, n\}, \mathbb{N}, \mathbb{Z}$), assign $\mu(\omega) \geq 0$ to each $\omega \in \Omega$, then let

$$\mu(A) := \sum_{\omega \in A} \mu(\omega).$$

for *every* $A \subset \Omega$. This defines on the power set $\mathcal{P}(\Omega)$ a function μ, which is *σ-additive*,

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k), \quad \text{for (pairwise) disjoint } A_k$$

nonnegative, $0 \leq \mu(A) \leq \infty$, and satisfies $\mu(\varnothing) = 0$.

- For uncountable spaces $\mathbb{R}, \mathbb{R}^n, \{0,1\}^\infty, \mathbb{R}^\infty, C([0,1])$ we want to consider also more involved measures that may assign positive measure to sets whose individual points receive measure zero.

Fundamental example: the Lebesgue measure on $\mathbb{R}$

$\lambda(I) := b - a$ for interval $I = (a, b], -\infty < a < b < \infty$.
For (pairwise) disjoint intervals $I_1, I_2, \ldots$ let

$$\lambda\left(\bigcup_{k=1}^{\infty} I_k\right) := \sum_{k=1}^{\infty} \lambda(I_k).$$

For single point $\lambda(x) = \lambda(\{x\}) = 0$, thus since $\mathbb{Q}$ is countable also

$$\lambda(\mathbb{Q}) = \sum_{x \in \mathbb{Q}} \lambda(x) = 0,$$

while $\lambda((a, b] \setminus \mathbb{Q}) = b - a$.

However, the Cantor set also has Lebesgue measure 0 but cardinality continuum.

- But there is no good way to define λ for *all* subsets of $\mathbb{R}$.

Example: the 'coin-tossing' space

$\Omega = \{0,1\}^\infty$ models an infinite series of independent Bernoulli trials with probability p for outcome 1 and $1-p$ for outcome 0. Then $\mathbb{P}(A) = (1-p)p$ for $A = \{\omega = (\omega_1, \omega_2, \dots) : \omega_1 = 0, \omega_0 = 1\}$

To which $A \subset \Omega$ can we assign probability? For instance, what is the probability of

$$A = \{\omega \in \Omega : \lim_{n\to\infty} \frac{\omega_1 + \cdots + \omega_n}{n} = p\}.$$

σ-algebras

Definition A family of sets $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called *σ-algebra* if

(i) $\Omega \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,

(iii) $A_k \in \mathcal{F} \quad (k \in \mathbb{N})$ disjoint sets $\Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Then $\varnothing \in \mathcal{F}$, and $\mathcal{F}$ is closed under countable intersections.

Definition $(\Omega, \mathcal{F})$ is called *measurable space*, and $A \in \mathcal{F}$ are called ($\mathcal{F}$-)*measurable sets* or *events* in the probability context.

Examples

(i) $\{\varnothing, \Omega\}$ (trivial σ-algebra),

(ii) $\mathcal{P}(\Omega)$ (the power set),

(iii) for partition $\Omega = \cup_{k=1}^{\infty} A_k$ in disjoint nonempty subsets $A_1, A_2, \ldots$, the following system of union-sets is a σ-algebra:

$$\bigcup_{k \in J} A_k, \quad J \subset \mathbb{N}$$

(iv) for any family $\{\mathcal{F}_t,\ t \in T\}$ of σ-algebras (T arbitrary set)

$$\bigcap_{t \in T} \mathcal{F}_t$$

is a σ-algebra,

(v) but the union of σ-algebras $\mathcal{F}_1 \cup \mathcal{F}_2$ typically is not a σ-algebra.

Generators

Definition For $\mathcal{G} \subset \mathcal{P}(\Omega)$, the intersection of all σ-algebras (in Ω) containing $\mathcal{G}$ is called the σ-algebra *generated* by $\mathcal{G}$, denoted $\sigma(\mathcal{G})$. This is the smallest σ-algebra containing $\mathcal{G}$.

If $\mathcal{F} = \sigma(\mathcal{G})$ we call $A \in \mathcal{G}$ *generators* of $\mathcal{F}$. A σ-algebra admitting a countable family of generators is called *separable*.

Example $\Omega = \{0, 1\}^\infty$, the 2^k *cylinder* sets

$$A(\epsilon_1, \ldots, \epsilon_n) = \{\omega \in \Omega : \omega_1 = \epsilon_1, \ldots, \omega_n = \epsilon_n\}.$$

generate a finite σ-algebra $\mathcal{F}_n$. We have *filtration* $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ The set $A = \{\omega : (\omega_1 + \cdots + \omega_n)/n \text{ converges as } n \to \infty\}$ is not in $\cup_n \mathcal{F}_n$ but $A \in \mathcal{F}$ for $\mathcal{F} = \sigma(\cup_n \mathcal{F}_n)$.

Borel σ-algebra

The σ-algebra of *Borel* sets in $\mathbb{R}$, denoted $\mathcal{B}(\mathbb{R})$ is generated by any of the families of sets:

(i) open sets,

(ii) closed sets,

(iii) intervals $(a, b]$, with $a, b \in \mathbb{R}$

(iv) intervals $(a, b]$ with $a, b \in \mathbb{Q}$ ($\mathcal{B}(\mathbb{R})$ is separable!),

(v) halflines $(-\infty, b]$, where $b \in \mathbb{R}$ (alternativley, $b \in \mathbb{Q}$)

For topological space $\mathcal{X}$, the Borel σ-algebra $\mathcal{B}(\mathcal{X})$ is the σ-algebra generated by the family of open sets. Examples are $\mathbb{R}^n, \mathbb{R}^\infty, C([a, b])$, etc.

Criteria for σ-algebra: monotone class

• Let $\mathcal{A}$ be an algebra (closed under *finite* unions), and *monotone class*: that is for $A_k \in \mathcal{A}$

$$A_1 \subset A_2 \subset \cdots \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A},$$

$$A_1 \supset A_2 \supset \cdots \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A},$$

then $\mathcal{A}$ is a σ-algebra.

Criteria for σ-algebra: $\pi - \lambda$ system

A family $\mathcal{D}$ of subsets in Ω is called a π-*system*, if closed under finite intersections, that is

$$A_1, A_2 \in \mathcal{D} \Rightarrow A_1 \cap A_2 \in \mathcal{D}.$$

A family $\mathcal{D}$ is called a λ-system if

(i) $\Omega \in \mathcal{D}$,

(ii) $A, B \in \mathcal{D}, A \subset B \Rightarrow B \setminus A \in \mathcal{D}$,

(iii) $A_1 \subset A_2 \subset \cdots \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{D}$.

Dynkin's Theorem: a π-λ-system is a σ-algebra.

Definition of a measure

Definition A *measure* on a mesurable space $(\Omega, \mathcal{F})$ is a nonnegative function $\mu : \mathcal{F} \to [0, \infty]$ such that

(i) $\mu(\varnothing) = 0$,

(iii) for disjoint $A_1, A_2, \ldots,$ with $A_n \in \mathcal{F}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

If $\mu(\Omega) < \infty$ the measure is called *finite*. If $\Omega = \cup_{k=1}^{\infty} \Omega_k$ where $\mu(\Omega_k) < \infty$, the measure is σ*-finite*.
If $\mu(\Omega) = 1$ we speak of a *probability measure* and may use notation $\mathbb{P}$.

Criteria for σ-additivity.

(i) Subadditivity: for $A_1, A_2, \ldots,$ with $A_n \in \mathcal{F}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

(iii) Monotonicity (increasing tower): $A_1 \subset A_2 \subset \cdots \Rightarrow$ $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$.

(iv) Monotonicity (decreasing tower): $A_1 \supset A_2 \supset \cdots \Rightarrow$ $\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$.

(v) Continuity at 'zero':
$A_1 \supset A_2 \supset \cdots, \bigcap_{n=1}^{\infty} A_n = \varnothing \Rightarrow \lim_{n\to\infty} \mu(A_n) = 0$.

Uniqueness of measures

Theorem Let $(\Omega, \mathcal{F})$ be a measurable space, where $\mathcal{F} = \sigma(\mathcal{D})$ for a π-system $\mathcal{D}$. Suppose $\Omega = \cup_k \Omega_k$, where $\Omega_k \in \mathcal{D}$. If two measures μ and ν coincide on $\mathcal{D}$ and $\nu(\Omega_k) = \mu(\Omega_k) < \infty$ then $\mu(A) = \nu(A)$ for all $A \in \mathcal{F}$.

Idea of proof: the collection of A's with $\mu(A) = \nu(A)$ is a λ-system, so Dynkin's theorem applies.

Construction by extension

A σ-additive μ_0 on algebra (or other set family) is called *pre-measure*. For instance, the intervals $I = (a, b] \subset \mathbb{R}$ comprise an algebra, and if $I = \cup_{n=1}^{\infty} I_n$ then $\lambda(I) = \sum_{n=1}^{\infty} \lambda(I_n)$.

Caratheodory Theorem: Suppose μ_0 is a pre-measure on $(\Omega, \mathcal{A})$, where $\mathcal{A}$ algebra. Then there exists a measure μ on $(\Omega, \sigma(\mathcal{A})$ such that

$$\mu(A) = \mu_0(A), \quad A \in \mathcal{A}.$$

Such measure μ_0 is unique if

$$\Omega = \bigcup_{k=1}^{\infty} \Omega_k, \Omega_k \in \mathcal{A}, \Omega_1 \subset \Omega_2 \subset \cdots$$

and $\mu_0(\Omega_k) < \infty, k \in \mathbb{N}$.

Examples of extension

• Lebesgue measure on $\mathcal{B}(\mathbb{R})$ is the extension from the algebra of intervals of the function 'interval length'.

• For c.d.f. F on $\mathbb{R}$ there is a unique probability measure with $\mu((-\infty, x]) = F(x), \quad x \in \mathbb{R}$. The measure may have no atoms, but have no density function (example: Cantor ladder).

• Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$ is the extension from the algebra of n-dimensional intervals $(a_1, b_1] \times \cdots \times (a_n, b_n]$. The intervals $(-\infty, b_1] \times \cdots \times (-\infty, b_n]$ comprise a π-system.

• Bernoulli(p) measure on $\{0,1\}$ is the extension from the algebra 'finite-dimensional' cylinders $A(\epsilon_1, \ldots, \epsilon_k)$. The event 'frequency of '1's is p becomes probability 1 (strong Law of Large Numbers).

In the case $p = 1/2$ the function

$$\omega \mapsto \sum_{k=1}^{\infty} \frac{\omega_k}{2^k}$$

establishes a measure-theoretic isomorphism between $\{0,1\}^{\infty}$ with Bernoulli($1/2$) and $[0,1]$ with λ. In the case $p \neq 1/2$ the pushforward of Bernoulli(p) measure is *singular* relative to λ, that is supported by a Borel set of zero Lebesgue measure.

Space $\mathbb{R}^\infty$

The Borel σ-algebra $\mathcal{B}(\mathbb{R}^\infty)$ is generated by finite-dimensional cylinders

$$C(B^n) := \{x \in \mathbb{R}^\infty : (x_1, \dots, x_n) \in B^n\}, \quad B^n \in \mathcal{B}(\mathbb{R}^n), \; n \in \mathbb{N},$$

where for B^n's we can also take n-dim intervals or smaller family $B^n = (-\infty, b_1] \times \cdots \times (-\infty, b_n]$, $b_i \in \mathbb{R}$ (or $b_i \in \mathbb{Q}$).

Probability measures P_n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $n \in \mathbb{N}$ are called *consistent* if

$$P_{n+1}(B^n \times \mathbb{R}) = P_n(B^n), \quad B^n \in \mathcal{B}(\mathbb{R}^n).$$

Kolmogorov's Extension Theorem: Suppose P_n are consistent probability measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Then there exists a unique probability measure $\mathbb{P}$ on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ such that

$$\mathbb{P}(C(B^n)) = P_n(B^n), \quad B^n \in \mathcal{B}(\mathbb{R}^n), \quad n \in \mathbb{N}.$$

Lebesgue measurable sets

Definition A family of sets $\mathcal{S} \subset \mathcal{P}(\Omega)$ is a *semiring* is $\mathcal{S}$ is closed under finite intersections and

$$A, B \in \mathcal{S} \Rightarrow A \setminus B = C_1 \cup \cdots \cup C_n,$$

for some n and disjoint $C_k \in \mathcal{S}$.
For pre-measure μ on $\mathcal{S}$ define *exterior measure*

$$\mu^*(A) = \inf \sum_{n=1}^{\infty} \mu(A_n)$$

where $A \subset \cup_n A_n$, and $A_n \in \mathcal{S}$ disjoint.
Clearly, $\mu^*(A) = \mu(A)$ for $A \in \mathcal{S}$.

Set $A \subset \Omega$ is called *Lebesgue-measurable* if for every $\epsilon > 0$ there exists $B \in \mathcal{S}$ such that

$$\mu^*(A \Delta B) < \epsilon.$$

Denote this family $L(\mathcal{S}, \mu)$.

Lebesgue's Theorem: The family $L(\mathcal{S}, \mu)$ is a σ-algebra and μ^* is a measure on $L(\mathcal{S}, \mu)$.

Example For $\mathcal{S} = \mathcal{B}(\mathbb{R})$ the σ-algebra of Lebesgue-measurable sets is

$$L(\mathcal{B}(\mathbb{R}), \lambda) = \sigma(\mathcal{B}(\mathbb{R}) \cup \mathcal{N}),$$

where $\mathcal{N}$ is the family of sets $A \subset B$, where B Borel set with $\lambda(B)$. This sort of measure extension is called *measure completion*.