Lecture 2 · 16.4.2021 · 10:15–12:00 via Zoom

Review of measure spaces, measures and integration.

We start by reviewing the basic setup of prob. theory, σ-algebras and prob. measure, the construction of probability measures, product of prob spaces, the notion of integral, and the various properties of the integral.

A probability space is triple $(\Omega, \mathscr{F}, \mathbb{P})$ where

- Ω is a set;
- $\mathscr{F} \subseteq \mathscr{P}(\Omega)$ is a σ-*algebra* of events, that is a family of subsets of Ω which is stable under complement, under finite intersections and under countable unions, and moreover it contains the empty set $\emptyset$. This represents the possible events we want to consider in our probabilistic setting.
- $\mathbb{P}$ is a probability measure, that is a function $\mathbb{P}: \mathscr{F} \to [0,1]$ such $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ and is σ-additive, that is for any disjoint family $(A_k)_k \subseteq \mathscr{F}$ it holds

$$\mathbb{P}(\cup_n A_n) = \sum_n \mathbb{P}(A_n),$$

(σ-additivity is equivalent to continuity at $\emptyset$, that is $\mathbb{P}(B_k) \to 0$ if $\cap_k B_k = \emptyset$ for an arbitrary family $(B_k)_k \subseteq \mathscr{F}$).

These axioms are due to Kolmogorov in the '40. There are fomalizations of probability which do not require σ-additivity (see de Finetti). But you can prove less in them.

More generally we call measure a positive function $\mu: \mathscr{F} \to [0, \infty]$ which satisfy all the properties of a prob. measure apart from the property on complements, i.e. $\mu(\emptyset) = 0$ and it is σ-additive.

σ-algebras are complicated to describe, so we would like to work with more manageable objects. So for any family $\mathscr{U} \subseteq \mathscr{P}(\Omega)$ we call $\sigma(\mathscr{U})$ the smallest σ-algebra which contains $\mathscr{U}$, this is the σ-algebra *generated by* $\mathscr{U}$.

Examples of σ-algebras

- $\mathscr{P}(\Omega)$ is a σ-algebra and we have always the trivial σ-algebra $\{\emptyset, \Omega\}$.
- If Ω is a topological space then we can consider the σ-algebra generated by all the open sets of Ω we call it the Borel σ-algebra and denote it with $\mathscr{B}(\Omega) = \sigma(\{\text{open sets in } \Omega\})$. The elements in $\mathscr{B}(\Omega)$ are called Borel sets.

How can we work with prob. measures?

We need first a way to work with σ-algebras easily. One important tool is Dynkin's $\pi - \lambda$ theorem:

- We say that a family of sets Λ is a λ-system if it contains $\emptyset$ and is closed under complements and countable disjoint unions;

- We say that a family of sets Π is a π-system if it is closed under finite intersections.

Note that a σ-algebra is both a λ-system and a π-system.

Examples (show the claims as exercise):

- the family of open intervals of $\mathbb{R}$ together with $\emptyset$ is a π-system (one can allow also finite unions and it is still a π-system);
- the family of rectangles $A \times B \subseteq \Omega \times \Omega$ with $A, B \in \mathscr{F}$ is a π-system;
- the family $\{B \in \mathscr{F}, \mathbb{P}(B) = \mathbb{Q}(B)\}$ for two probability measure $\mathbb{P}, \mathbb{Q}$ on $\mathscr{F}$, is a λ-system (exercise);
- consider the family $\Lambda \subseteq \mathscr{F}$ such that there exists a vector space $\mathscr{H}$ of bounded measurable real-valued functions on $\mathscr{F}$ such that $1 \in \mathscr{H}$ and $\mathscr{H}$ contains all the indicator functions $\mathbb{1}_B$ of elements $B \in \Lambda$. Then Λ is a λ-system;

> **Theorem. (Dynkin's $\pi - \lambda$ theorem)** *If Π is a π-system and Λ a λ-system then $\Pi \subseteq \Lambda$ implies that $\sigma(\Pi) \subseteq \Lambda$.*

A function $f: (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ between two measure spaces (i.e. a pair of a space and σ-algebra on it) is measurable iff $f^{-1}(F) \in \mathscr{F}$ for all $F \in \mathscr{E}$.

Note that $f^{-1}(\mathscr{E}) := \{f^{-1}(F): F \in \mathscr{E}\}$ is always a σ-algebra, for a measurable function we have moreover that $f^{-1}(\mathscr{E}) \subseteq \mathscr{F}$, that is for an $\mathscr{F}$-measurable function $f^{-1}(\mathscr{E})$ is a sub-σ-algebra of $\mathscr{F}$.

Functions measurable on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ are called <u>random variables</u>. We can speak about probabilities associated to this particular function.

For a random variable $f: \Omega \to (E, \mathscr{E})$ we can do

$$\mathscr{E} \xrightarrow{f^{-1}} \mathscr{F} \xrightarrow{\mathbb{P}} [0, 1]$$

that is we can construct a new probability measure $\mathbb{P}_f: \mathscr{E} \to [0, 1]$ by pullback of $\mathbb{P}$ with f. This is called the law (or the distribution) of f.

> Real valued random variable are like coordinates on the probability space, i.e. reduce the problem to compute probabilities in the unstructured setting of Ω to a concrete problem about algebra and analysis of real-valued functions.

If $X: (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ is a r.v. then $\sigma(X)$ is the smallest σ-algebra for which X is measurable and

$$\sigma(X) = \{X^{-1}(F): F \in \mathscr{E}\} \subseteq \mathscr{F}.$$

> We could say that $\sigma(X)$ represents the information on the measure space $(\Omega, \mathscr{F})$ obtained by looking at it via X (you can think about it as a coodinate system). Like any coordinate system it could not be complete and discard some of the underlying information. $\sigma(X)$ can be tought as the set of all possible questions you can ask about Ω using only the language given by X.

For real valued functions measurability is usually understood wrt. to the Borel σ-algebra. More generally this applies for functions with values in top. spaces.

Measurability strongly depends on the inital σ-algebra $\mathscr{F}$, when this is important we say explicitly $\mathscr{F}$-measurable.

Measurable functions, like σ-algebras are difficult to describe explicitly, but we have nice way to relate this two concepts in the case the σ-algebra is generated by some family $\mathscr{U}$.

Theorem. (Monotone class theorem) *Let $\mathscr{H}$ be a vector space of bounded real-values functions on Ω such that*

i. $1 \in \mathscr{H}$,

ii. *if $f_n \geqslant 0$ and $f_n \uparrow f$ pointwise with f bounded, then $f \in \mathscr{H}$ (so $\mathscr{H}$ is stable under monotone limits).*

Then if $\mathscr{H}$ contains the indicator functions of every element of a π-system $\mathscr{U}$ then $\mathscr{H}$ contains every bounded $\sigma(\mathscr{U})$-measurable function.

The proof is not difficult and uses Dynkin's theorem. The property (ii) tells us that we can do pointwise approximations in $\mathscr{H}$ and the proof proceed by appoximating measurable functions with simple functions which then are shown to be in $\mathscr{H}$. Recall that simple functions are those measurable functions which take only finitely many values, i.e. $f: \Omega \to \mathbb{R}$ is simple if

$$f(\omega) = \sum_{x \in B} x \mathbb{1}_{f^{-1}(\{x\})}(\omega)$$

with B a finite subset of $\mathbb{R}$. Note that $\{x\}$ is measurable wrt. $\mathscr{B}(\mathbb{R})$, this ensure that $f^{-1}(\{x\}) \in \mathscr{F}$ since f is measurable.

Usually in proofs one consider a class of functions $\mathscr{H}$ which satisfy all the above properies and then this shows that it actually consists of all the measurable functions. This allows to prove a statement about all measurable functions by proving:

1. first that it holds for indicator functions of measurable sets (is enough a generating subset),
2. you prove it for linear combinations,
3. and then you prove that it is stable under monotone pointwise limits.

Indeed one applies the above theorem using as $\mathscr{H}$ the set of all measurable functions satisfying the statement we are interested in.

Still, how we construct probability measures?

The main tool here is the Carathéodory extension theorem (stated for positive measures):

Theorem. (Carathéodory extension theorem) *Let Ω be a set, $\mathscr{U}$ an algebra of subsets of Ω and $\mu_0: \mathscr{U} \to \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$ a positive σ-additive set-function on $\mathscr{U}$. Then there exists a measure $\mu: \sigma(\mathscr{U}) \to \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$ such that*

$$\mu|_{\mathscr{U}} = \mu_0.$$

If μ_0 is σ-finite then μ is unique.

$\mathscr{U}$ being an algebra only finite unions are allowed, so σ-additivity for a set-function μ_0 is understood in the sense that for any disjoint family $(A_k)_k \subseteq \mathscr{U}$ such that $\cup_k A_k \in \mathscr{U}$ then $\mu_0(\cup_k A_k) = \sum_k \mu_0(A_k)$, i.e. the requirement is only imposed when the uncountable union is in $\mathscr{U}$, otherwise we do not impose any requirements.

A measure is σ-finite if it exist a countable measurable cover $(A_k)_k$ of Ω such that $\mu(A_k) < \infty$ for any k.

Example: Lebesgue measure on $\mathbb{R}^n$ is not finite but σ-finite.

Usually in probability theory we just work with probability measures so the uniqueness part of Carathéodory extension theorem is for free.
