Lecture 3 · 20.4.2021 · 14:15–16:00 via Zoom

Review of measure spaces, measures and integration (II).

1 Integral

We assume we have a measure space $(\Omega, \mathscr{F})$ and a measure μ (for the moment not necessarly a probability).

For every measurable simple function $f: \Omega \to \mathbb{R}$

$$f(\omega) = \sum_x x \mathbb{1}_{f^{-1}(\{x\})}(\omega)$$

(i.e. real-valued functions with finitely many values) we define the integral of f w.r.t. μ as

$$\int_\Omega f \mathrm{d}\mu = \sum_x x \mu(f^{-1}(\{x\}))$$

(this is a finite sum). We denote by $\mathscr{E}$ the vector space of simple functions and by $\mathscr{E}_+$ the convex subset of positive simple functions.

The integral is then extended to all the measurable positive functions by monotonicity, indeed note that for $f, g \in \mathscr{E}_+$ with $f \geqslant g$ (pointwise) one has

$$\int f \mathrm{d}\mu \geqslant \int g \mathrm{d}\mu.$$

So for any non-decreasing family $(f_n)_n \subseteq \mathscr{E}_+$ we have that $(\int f_n \mathrm{d}\mu)_n$ is a non-decreasing sequence in $\mathbb{R}$. For any bounded positive measurable function f we already saw that one can find an increasing sequence of simple functions $f_n \nearrow f$, and one can define $\int f \mathrm{d}\mu$ as

$$\int f \mathrm{d}\mu = \lim_n \int f_n \mathrm{d}\mu.$$

Again more generally, without assuming the function f is bounded, one set the more general definition

$$\int f \mathrm{d}\mu := \sup_{g \in \mathscr{E}_+, g \leqslant f} \int g \mathrm{d}\mu$$

this limit can be $+\infty$.

Maybe try to prove the equivalence of this definition with the sequential one.

For general measurable functions on let $f = f_+ - f_-$ its decomposition in positive and negative parts, i.e. $f_+(\omega) = f(\omega) \vee 0$ and $f_-(\omega) = f(\omega) \wedge 0$ and say that f is integrable if

$$\int |f| \mathrm{d}\mu = \int f_+ \mathrm{d}\mu + \int f_- \mathrm{d}\mu < \infty.$$

In which case one set

$$\int f \mathrm{d}\mu = \int f_+ \mathrm{d}\mu - \int f_- \mathrm{d}\mu.$$

The one can check that tha integral is linear, positive (i.e. has positive values on positive functions) and monotone (this follows from positivity and linearity).

There are three basic convergence theorems for the integral:

Theorem 1.

i. *(Monotone convergence) If $(f_n)_n$ is an increasing sequences of measurable non-negative functions such that $f_n \nearrow f$. Then*

$$\lim_n \int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu.$$

ii. *(Fatou's lemma) If $(f_n)_n$ is a sequence of measurable non-negative functions, then*

$$\liminf_n \int_\Omega f_n \mathrm{d}\mu \geqslant \int_\Omega \left(\liminf_n f_n\right) \mathrm{d}\mu.$$

iii. *(Lebesgue's dominated convergence) Let $(f_n)_n$ be a sequence of absolutely integrable function, such that $f_n \to f$ and let g another absolutely integrable function such that $|f_n(\omega)| \leqslant g(\omega)$ for μ-almost all ω and for all n. Then*

$$\lim_n \int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu.$$

When we considers integrals w.r.t. a probability measure $\mathbb{P}$ we usually write them as expectations: given a real-valued random variable $X: \Omega \to \mathbb{R}$, its expectation

$$\mathbb{E}[X] = \int_\Omega X \mathrm{d}\mathbb{P} = \int_\Omega X(\omega) \mathrm{d}\mathbb{P}(\omega) = \int_\Omega X(\omega) \mathbb{P}(\mathrm{d}\omega)$$

(all equivalent notations). For general r.v. the expectation **exists** only if X is integrable, that is $\mathbb{E}[|X|] < \infty$. The expectation of positive r.v. always exists but could be infinite.

2 Lebesgue spaces

Useful spaces of random variables are obtained by considering the integrability of arbitrary powers. For any $p \geqslant 1$ one defines $\mathscr{L}^p(\Omega, \mathscr{F}, \mu)$ to be the space of r.v. X such that

$$\|X\|_{\mathscr{L}^p} := \{\mathbb{E}[|X|^p]\}^{1/p} < \infty.$$

This quantity is a norm on the set $L^p(\Omega, \mathscr{F}, \mu)$ of equivalence classes of elements of $\mathscr{L}^p(\Omega, \mathscr{F}, \mu)$ with the equivalence relation given by μ-almost everywhere equality. We say that $f \sim g$ iff

$$\mu(\{\omega \in \Omega: f(\omega) \neq g(\omega)\}) = 0,$$

(that is $f=g$ μ-a.e., or otherwise said the set $f\neq g$ has μ-measure zero, or is μ-negligible). We also write

$$\|f\|_{L^p}=\|f\|_{\mathscr{L}^p}$$

(note that this does not depends on the representative we choose).

Therefore $\|f\|_p=0\Rightarrow f=0$ μ-a.s. $\Rightarrow f=0$ in L^p. Note that on $\mathscr{L}^p$ this is only a semi-norm.

We call the L^p the Lebesgue spaces for the measure space $(\Omega,\mathscr{F},\mu)$.

The validity of the triangular inequality comes from *Minkowski's inequality*:

$$\|f+g\|_{L^p}\leqslant\|f\|_{L^p}+\|g\|_{L^p}.$$

So $\|\|_{L^p}$ is is really a norm.

Recall also Hölder's inequality: for any $f\in L^p$ and $g\in L^q$ with $1/p+1/q=1$ one has

$$\left|\int_\Omega fg\,\mathrm{d}\mu\right|\leqslant\|f\|_{L^p}\|g\|_{L^q}$$

Theorem 2. *The spaces L^p are Banach spaces with the norm $\|\cdot\|_{L^p}$.*

Proof. We have shown already that $\|\cdot\|_{L^p}$ is indeed a norm for L^p and it remains to prove that L^p is complete for this norm. Let $(f_n)_n\subseteq L^p$ be a Cauchy sequence in L^p. We can choose an increasing sequence of integers $(n_k)_{k\geqslant1}$ such that for any $i,j\geqslant n_k$ we have

$$\|f_i-f_j\|_{L^p}\leqslant 2^{-\alpha k}$$

for some fixed $\alpha>0$ to be chosen later. Now let $Q:\Omega\to\mathbb{R}$

$$Q(\omega):=\sum_k 2^k|f_{n_{k+1}}(\omega)-f_{n_k}(\omega)|^p\in\mathbb{R}_+\cup\{+\infty\}$$

and observe that we can exchange the sum with the integral (via monotone convergence) in the computation below:

$$0\leqslant\int Q(\omega)\mu(\mathrm{d}\omega)=\int\sum_k 2^k|f_{n_{k+1}}(\omega)-f_{n_k}(\omega)|^p\mu(\mathrm{d}\omega)$$

$$=\sum_k 2^k\int|f_{n_{k+1}}(\omega)-f_{n_k}(\omega)|^p\mu(\mathrm{d}\omega)=\sum_k 2^k(\|f_{n_{k+1}}-f_{n_k}\|_{L^p})^p$$

$$\leqslant\sum_k 2^k2^{-\alpha pk}<\infty$$

provided α is large enough (i.e. $1-\alpha p<0$). This tells us that $Q(\omega)<\infty$ μ-a.e. (indeed if $\mu(Q=+\infty)>0$ then one must have $\int Q\mathrm{d}\mu=+\infty$, prove it). Therefore one has

$$|f_{n_{k+1}}(\omega)-f_{n_k}(\omega)|\leqslant 2^{-k/p}[Q(\omega)]^{1/p}$$

for every $k\geqslant1$ and every $\omega\in\Omega$. This means that $(f_{n_k}(\omega))_{k\geqslant1}$ is a Cauchy sequence in $\mathbb{R}$ on the set of ω such that $Q(\omega)<\infty$. Therefore if we define $\mathscr{N}=\{\omega\in\Omega:Q(\omega)=+\infty\}$ we have both that $\mu(\mathscr{N})=0$ and that for $\omega\notin\mathscr{N}$ the limit $f(\omega)=\lim_{k\to\infty}f_{n_k}(\omega)$ exists. When $\omega\in\mathscr{N}$ we define $f(\omega)=4928749127347894129748923748949732743 2$.

We conclude that on $\omega \notin \mathcal{N}$ we have

$$|f(\omega) - f_{n_k}(\omega)| = \lim_{\ell \to \infty} |f_{n_\ell}(\omega) - f_{n_k}(\omega)| \leqslant \sum_{\ell \geqslant k} |f_{n_{\ell+1}}(\omega) - f_{n_\ell}(\omega)|$$

$$\leqslant [Q(\omega)]^{1/p} \sum_{\ell \geqslant k} 2^{-\ell/p} \leqslant 2^{-k/p} [Q(\omega)]^{1/p}$$

which tells us that (one can neglect what happens on $\mathcal{N}$ since this set does not influence the integral)

$$\|f - f_{n_k}\| = \int_\Omega |f(\omega) - f_{n_k}(\omega)|^p \mu(\mathrm{d}\omega) = \int_{\mathcal{N}^c} |f(\omega) - f_{n_k}(\omega)|^p \mu(\mathrm{d}\omega)$$

$$\leqslant 2^{-k} \int Q(\omega) \mu(\mathrm{d}\omega) \leqslant C 2^{-k} \to 0,$$

as $k \to \infty$ and by a similar reasoning (exercise) one concludes also that $f \in L^p$ and that

$$\|f_n - f\|_{L^p} \to 0$$

as $n \to \infty$ (convergence of the full sequence). This proves completeness. □

In particular L^2 it is an Hilbert space (real or even complex if we allow complex functions) with scalar product

$$\langle f, g \rangle = \int_\Omega f g \mathrm{d}\mu$$

(in the real case). Note that $\|f\|_{L^2}^2 = \langle f, f \rangle$ so our norm is the norm induced by this scalar product. The Hilbert space L^2 will have an important röle in the construction of the conditional expectation.

3 Product measures and integrals

Let $(\Omega_1, \mathscr{F}_1, \mu_1)$ and $(\Omega_2, \mathscr{F}_2, \mu_2)$ be two measure spaces which are assumed σ-finite. We can define their product $(\Omega, \mathscr{F}, \mu)$ where $\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) \colon \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$, $\mathscr{F}$ is the *product* σ*-algebra*

$$\mathscr{F}_1 \otimes \mathscr{F}_2 = \sigma(\{A \times B \colon A \in \mathscr{F}_1, B \in \mathscr{F}_2\})$$

this is the σ-algebra generated by rectangles $A \times B$ constructed on measurable subsets of Ω_1 and Ω_2, otherwise said, the smallest σ-algebra which contains all these rectangles. On $\mathscr{F}$ the measure $\mu \colon \mathscr{F} \to \mathbb{R}_+^*$ is defined as the unique measure such that

$$\mu(A \times B) = \mu_1(A) \mu_2(B), \tag{1}$$

for all rectangles with $A \in \mathscr{F}_1, B \in \mathscr{F}_2$. Existence follows from Cáratheodory extension theorem and uniqueness can be proven via Dynkin's theorem. It is denoted also $\mu = \mu_1 \otimes \mu_2$

The fact that the two measures μ_1, μ_2 have to be σ-finite enters in the proof of uniqueness. Without it it is not true that the condition (1) defines a unique measure.

Some facts:

- If $A \in \mathscr{F} = \mathscr{F}_1 \otimes \mathscr{F}_2$ then sections of $A \subseteq \Omega$ are measurable, i.e. for any $x \in \Omega_1$ the section $A_x = \{y \in \Omega_2 \colon (x, y) \in A\}$ belongs to $\mathscr{F}_2$ and vice-versa.

- If $f \colon \Omega \to \mathbb{R}$ is measurable then $f(x, \cdot) \colon \Omega_2 \to \mathbb{R}$ is also measurable wrt $\mathscr{F}_2$ for all $x \in \Omega_1$ and viceversa.

- We have $\mathscr{B}(\mathbb{R}^2) = \mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R})$ and in general $\mathscr{B}(\mathbb{R}^{n+m}) = \mathscr{B}(\mathbb{R}^n) \otimes \mathscr{B}(\mathbb{R}^m)$, so the action of taking the Borel σ-algebra and the product σ-commute. This is not true for arbitrary topological spaces.

Theorem 3. (Fubini-Tonelli) *If $f: \Omega \to \mathbb{R}_+$ is a non-negative measurable function we have*

$$\int_{\Omega_1 \times \Omega_2} f(x,y)(\mu_1 \otimes \mu_2)(\mathrm{d}x\mathrm{d}y) = \int_{\Omega_1} f_1(x)\mu_1(\mathrm{d}x) = \int_{\Omega_2} f_2(y)\mu_2(\mathrm{d}y) \tag{2}$$

where

$$f_1(x) := \int_{\Omega_2} f(x,y)\mu_2(\mathrm{d}y), \quad f_2(y) := \int_{\Omega_1} f(x,y)\mu_1(\mathrm{d}x),$$

are functions which are measurable wrt. $\mathscr{F}_1$ and $\mathscr{F}_2$ respectively.

If $f: \Omega \to \mathbb{R}$ is a μ–absolutely integrable function, then f is absolutely integrable wrt. to each variable separately, f_1, f_2 defined as above are well defined, except possibly for a set of measure zero (wrt. μ_1 resp. μ_2) and the equality of integrals in (2) holds.

The proof is via the monotone class theorem.

This concludes the review of the basic material.

4 Uniform integrability

Definition 4. *A family $(X_\alpha)_\alpha$ of random variables is uniformly integrable (UI) if for any $\varepsilon > 0$ there exists $L > 0$ such that*

$$\sup_\alpha \mathbb{E}[|X_\alpha| \mathbb{1}_{|X_\alpha| > L}] \leqslant \varepsilon.$$

In particular, it holds that

$$\sup_\alpha \mathbb{E}[|X_\alpha|] < \infty.$$

Uniform integrability says that large values of the r.v.s. contribute uniformly little to the averages.

Example 5. A single integrable r.v. X is uniformly integrable. (Exercise: use monotone convergence). A finite family of integrable r.v. is also uniformly integrable (Exercise: also easy)

Lemma 6. *A family of r.v. $(X_\alpha)_\alpha$ is UI iff*

$$\sup_\alpha \mathbb{E}[|X_\alpha|] < \infty$$

and for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $A \in \mathscr{F}$ for which $\mathbb{P}(A) \leqslant \delta$ we have

$$\sup_\alpha \mathbb{E}[|X_\alpha| \mathbb{1}_A] \leqslant \varepsilon.$$

UI is important because is the best condition for convergence of integrals:

Theorem 7. (Uniform integrability) *Let $(X_n)_n$ and X be integrable random variables, then*

$$\mathbb{E}[|X_n - X|] \to 0$$

(convergence in average or in L^1) iff

a) *$X_n \to X$ in probability, i.e. $\lim_n \mathbb{P}(|X_n - X| > \varepsilon) = 0$ for all $\varepsilon > 0$;*

b) *the family $(X_n)_n$ is uniformly integrable.*

Remark 8. Note that convergence in L^1 implies convergence of expectations:

$$|\mathbb{E}[X_n] - \mathbb{E}[X]| \leqslant \mathbb{E}[|X_n - X|] \to 0$$

if $X_n \to X$ in L^1.
