36-752 Advanced Probability Overview Spring 2018

5. L^p Spaces and Weak Law of Large Numbers

Instructor: Alessandro Rinaldo

Associated reading: Sec 2.4, 2.5, and 4.11 of Ash and Doléans-Dade; Sec 1.5 and 2.2 of Durrett.

1 L^p Spaces

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. For $p \geq 1$, consider $\mathcal{L}^p$, the set of all functions f from Ω to $\mathbb{R}$ such that $|f|^p$ is integrable. This set has many properties of a metric space, but it has one problem that we shall see shortly. Let $\|f\|_p = \left(\int |f|^p d\mu\right)^{1/p}$. It is easy to see that $\|f\|_p \geq 0$ for all $f \in \mathcal{L}^p$. For $p \geq 1$, we will also show that $\|f+g\|_p \leq \|f\|_p + \|g\|_p$, whenever $f, g \in \mathcal{L}^p$. Also, if $f \in \mathcal{L}^p$ and $a \in \mathbb{R}$, then $af \in \mathcal{L}^p$ and $\|af\|_p = |a|\|f\|_p$, so $\mathcal{L}^p$ is a real vector space. The only problem is that $\|\cdot\|_p$ is not a true norm because $\|f\|_p = 0$ does not imply that $f = 0$ in the vector space. Every f that equals 0 a.e. $[\mu]$ will have $\|f\|_p = 0$. Hence, we modify $\mathcal{L}^p$ to create a new set L^p which consists of the equivalence classes of elements of $\mathcal{L}^p$ under the equivalence relation $f \sim g$ if and only if $f = g$ a.e. $[\mu]$. We define $\|[f]\|_p = \|f\|_p$, where $[f]$ stands for the equivalence class to which f belongs. Because $g \in [f]$ implies $\|g\|_p = \|f\|_p$, the norm is now well-defined on L^p. So, L^p is a normed linear space, hence a metric space. We will also show (later in this class) that it is complete, and hence a Banach space. People tend to ignore the fact that L^p is a set of equivalence classes rather than a set of functions, and we will do the same.

To be precise, we should include more information in the name of L^p. Indeed, we should refer to the space as $L^p(\Omega, \mathcal{F}, \mu)$, since each measure space has its own L^p space.

The special case of $p = \infty$ is handled separately. A function (equivalence class) f is in L^∞ if the function is essentially bounded.

Definition 1 (Essential Supremum). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $f : \Omega \to \overline{\mathbb{R}}$ be a measurable function. Let $\mathcal{A}$ be the collection of all elements A of $\mathcal{F}$ such that $\mu(A^C) = 0$. The* essential supremum *of f is*

$$\operatorname{ess\,sup} f = \inf_{A \in \mathcal{A}} \sup_{\omega \in A} |f(\omega)|$$

or equivalently

$$\operatorname{ess\,sup} f = \sup\{t : \mu(\omega : |f(\omega)| \geq t) > 0\}.$$

If $\operatorname{ess\,sup} f < \infty$, we say that f is essentially bounded. *The L^∞ norm of f is $\|f\|_\infty = \operatorname{ess\,sup} f$.*

The essential supremum of f is the least upper bound of $|f|$ on sets whose complements have measure 0. It is the smallest number c such that $|f| \leq c$ a.e. $[\mu]$.

The special L^p space in which $\Omega = \mathbb{Z}^+$, $\mathcal{F} = 2^\Omega$, and μ is counting measure is called ℓ^p.

Example 2. *For each $1 \leq p < \infty$, $f \in \ell^p$ if and only if its pth power is an absolutely summable sequence. A sequence f is in ℓ^∞ if and only if it is bounded.*

For probability spaces, $L^{p_1} \subseteq L^{p_2}$ whenever $p_1 \geq p_2$. Indeed, $\|X\|_{p_1} \geq \|X\|_{p_2}$ if $p_1 \geq p_2$. This follows easily from Jensen's inequality.

Proposition 3 (Jensen's inequality). *If f is a convex function, then $\mathrm{E}[f(X)] \geq f(\mathrm{E}(X))$.*

For infinite measure spaces, odd behavior is possible.

Example 4. *Consider $L^p(\mathbb{R}, \mathcal{B}^1, \lambda)$, where λ is Lebesgue measure. Let*

$$f(x) = \begin{cases} x^{-3/8} & \text{for } 0 < x < 1, \\ x^{-1} & \text{for } 1 \leq x < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f \notin L^1$, $f \in L^2$, and $f \notin L^3$.

We will prove most results for $L^p(\Omega, \mathcal{F}, \mu)$ where μ is a σ-finite measure. We will use the results mainly for probability spaces.

Lemma 5 (Linearity of L^p). *If $X, Y \in L^p$, then $X + Y \in L^p$.*

Proof: For all $p > 0$, we have

$$|a+b|^p \leq (|a| + |b|)^p \leq (2\max(|a|, |b|))^p = 2^p \max(|a|^p, |b|^p) \leq 2^p(|a|^p + |b|^p)\,.$$

It follows that

$$\int |X+Y|^p d\mu \leq 2^p \left[\int |X|^p d\mu + \int |Y|^p d\mu\right].$$

■

Some of the results on L^p spaces involve the concept of conjugate indices.

Definition 6 (Conjugate Indices). *For each $p \in (1, \infty)$, let q be the unique value in $(1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. For $p = 1$, let $q = \infty$. Then p and q are* conjugate indices.

Theorem 7 (Hölder's inequality). *Let $X \in L^p$ and $Y \in L^q$, where p and q are conjugate indices. Then $XY \in L^1$ and $\int |XY| d\mu \leq \|X\|_p \|Y\|_q$.*

Proof: If either p or q is ∞, the result is obvious. So assume that neither p nor q is ∞. Define $U = |X|^p$ and $V = |Y|^q$, so that $U, V \in L^1$. Since the weighted geometric mean of two numbers is never more than the weighted arithmetic mean (with the same weights), we have

$$\left[\frac{U}{\int U d\mu}\right]^{1/p} \left[\frac{V}{\int V d\mu}\right]^{1/q} \leq \frac{1}{p}\frac{U}{\int U d\mu} + \frac{1}{q}\frac{V}{\int V d\mu}.$$

It follows that

$$\int U^{1/p} V^{1/q} d\mu \leq \left(\int U d\mu\right)^{1/p} \left(\int V d\mu\right)^{1/q} < \infty.$$

The first of these inequalities can be rewritten $\int |XY| d\mu \leq \|X\|_p \|Y\|_q$. The second one implies that $XY \in L^1$. ∎

Example 8 (Cauchy-Schwarz inequality). *Let $p = q = 2$ in Theorem 7 to get that $X, Y \in L^2$ implies*

$$\int |XY| d\mu \leq \sqrt{\int X^2 d\mu \int Y^2 d\mu}.$$

If μ is a probability, this is the familiar Cauchy-Schwarz inequality.

Theorem 9 is the triangle inequality for L^p norms.

Theorem 9 (Minkowski's inequality). *If $X, Y \in L^p$, then*

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p.$$

Proof: The proofs are simple for $p = 1$ and $p = \infty$. So assume $p \in (1, \infty)$. First, let q be the conjugate index and notice that $(p-1)q = p$. Hence $|X + Y|^{p-1} \in L^q$ and $\||X + Y|^{p-1}\|_q = \|X + Y\|_p^{p/q}$. Write

$$|X + Y|^p = |X + Y||X + Y|^{p-1} \leq |X||X + Y|^{p-1} + |Y||X + Y|^{p-1}.$$

Theorem 7 says that

$$\int |X + Y|^p d\mu \leq (\|X\|_p + \|Y\|_p)\||X + Y|^{p-1}\|_q.$$

Rewrite this as

$$\|X + Y\|_p^p \leq (\|X\|_p + \|Y\|_p)\|X + Y\|_p^{p/q}.$$

Divide both sides by $\|X + Y\|_p^{p/q}$ to get the desired result, because $p - p/q = 1$. ∎

We give two more simple but useful inequalities for integrals and expectations.

Proposition 10 (Markov inequality). *If f is a nonnegative measurable function, then $\mu(\{\omega : f(\omega) \geq c\}) \leq \int f d\mu / c$. In particular, let X be a nonnegative random variable. Then $\Pr(X \geq c) \leq \mathrm{E}(X)/c$.*

There is also a famous corollary.

Corollary 11 (Tchebychev inequality). *Let X be a random variable and have finite mean μ. Then $\Pr(|X - \mu| \geq c) \leq \mathrm{Var}(X)/c^2$.*

Exercise 12. *Show that there is some random variable X for which $\Pr(|X - \mu| \geq c) = \mathrm{Var}(X)/c^2$. Thus, without additional assumptions, Tchebychev's inequality cannot be improved.*