

## 1 Lecture 8 – $L^p$ Spaces and Classical Inequalities

Fix a measure space  $(\Omega, \mathcal{F}, \mu)$ . This lecture introduces the spaces  $L^p(\Omega, \mathcal{F}, \mu)$ , the four workhorse tail-bound inequalities (Markov, Chebyshev, Chernoff, Jensen) and the two inequalities (Hölder and Minkowski) that give  $L^p$  the structure of a normed vector space. We close with a density theorem that lets one approximate any  $L^p$  function by simple functions supported on sets of finite measure.

### 1.1 The spaces $L^p$

#### Definition 1.1: $L^p$ space, $1 \leq p < \infty$

For  $p \in [1, \infty)$  and a measurable function  $f: \Omega \rightarrow \mathbb{R}$ , define

$$\|f\|_p = \left( \int_{\Omega} |f|^p d\mu \right)^{1/p}.$$

The space  $L^p(\Omega, \mathcal{F}, \mu)$  consists of all measurable  $f$  with  $\|f\|_p < \infty$ , modulo equality  $\mu$ -almost everywhere.

#### Definition 1.2: Essential supremum and $L^\infty$

The *essential supremum* of a measurable function  $f$  is

$$\|f\|_\infty = \text{ess sup } |f| = \inf \{ t \in [-\infty, \infty] : \mu(\{|f| > t\}) = 0 \}.$$

The space  $L^\infty(\Omega, \mathcal{F}, \mu)$  consists of all measurable  $f$  with  $\|f\|_\infty < \infty$ , again modulo  $\mu$ -a.e. equality.

**Remark 1.1.** The endpoint cases  $p = 1$  and  $p = \infty$  are honest analogues of the finite- $p$  definition: as  $p \rightarrow \infty$  one has  $\|f\|_p \rightarrow \|f\|_\infty$  when  $\mu$  is finite and  $f$  is bounded.

### 1.2 Markov, Chebyshev, Chernoff

The next three results all spring from the same one-line trick: bound the integrand below on the set  $\{f \geq t\}$ .

#### Theorem 1.3: Markov's inequality

Let  $f \geq 0$  be measurable and  $t > 0$ . Then

$$\mu(\{f \geq t\}) \leq \frac{1}{t} \int_{\Omega} f d\mu.$$

**Corollary 1.4: Chebyshev's inequality**

For any measurable  $f$  and  $m \in \mathbb{R}$ ,  $t > 0$ ,

$$\mu(\{|f - m| \geq t\}) \leq t^{-2} \int_{\Omega} (f - m)^2 d\mu.$$

**Corollary 1.5: Chernoff's inequality**

For any measurable  $f$ ,  $t \in \mathbb{R}$ , and  $\eta \geq 0$ ,

$$\mu(\{f \geq t\}) \leq e^{-\eta t} \int_{\Omega} e^{\eta f} d\mu.$$

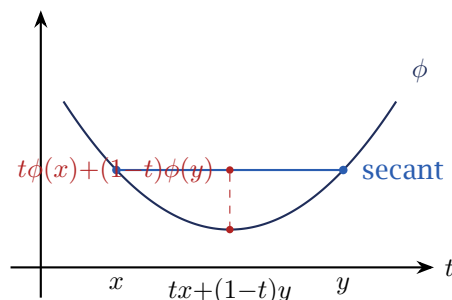
**Remark 1.2.** Both Chebyshev and Chernoff follow from Result 1.3 applied to a non-negative transform:  $(f - m)^2$  for Chebyshev,  $e^{\eta f}$  for Chernoff. Chernoff is sharpest when one optimises over  $\eta \geq 0$ .

**1.3 Convexity and Jensen's inequality****Definition 1.6: Convex function**

Let  $I \subseteq \mathbb{R}$  be an interval. A function  $\phi: I \rightarrow \mathbb{R}$  is *convex* if for all  $x, y \in I$  and  $t \in [0, 1]$ ,

$$\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y).$$

Geometrically: the secant line lies above the graph.



**Figure 1.** A convex function: the secant joining  $(x, \phi(x))$  and  $(y, \phi(y))$  lies above the graph.

**Theorem 1.7: Jensen's inequality**

Let  $(\Omega, \mathcal{F}, \mu)$  be a probability space (so  $\mu(\Omega) = 1$ ), let  $X: \Omega \rightarrow I$  be integrable with  $\mathbb{E}X = \int X d\mu$  lying in the interior of  $I$ , and let  $\phi: I \rightarrow \mathbb{R}$  be convex with  $\mathbb{E}[\phi(X)]$  well defined. Then

$$\phi(\mathbb{E}X) \leq \mathbb{E}[\phi(X)].$$

Equivalently, for any measurable  $X$  and convex  $\phi$ ,  $\phi(\int X d\mu) \leq \int \phi(X) d\mu$ .

**Remark 1.3.** Convexity guarantees a supporting line: at  $m = \mathbb{E}X$  one can choose  $a, b \in \mathbb{R}$  with  $\phi(x) \geq ax + b$  for all  $x \in I$  and  $\phi(m) = am + b$ . Taking expectations of the inequality

gives the result; integrability of  $\phi(X)$  follows because  $\phi^-(x) \leq |a||x| + |b|$ .

### 1.4 Hölder and Minkowski

We now turn to the two inequalities that pin down the geometry of  $L^p$ . Throughout,  $p, q \in [1, \infty]$  are called *conjugate exponents* when

$$\frac{1}{p} + \frac{1}{q} = 1,$$

with the conventions  $1/\infty = 0$  and  $(p, q) \in \{(1, \infty), (\infty, 1)\}$  included.

#### Theorem 1.8: Hölder's inequality

Let  $p, q \in [1, \infty]$  be conjugate exponents and let  $f, g$  be measurable. Then

$$\|fg\|_1 = \int_{\Omega} |fg| d\mu \leq \|f\|_p \|g\|_q.$$

In particular, if  $f \in L^p$  and  $g \in L^q$ , then  $fg \in L^1$ .

#### Corollary 1.9: Cauchy–Schwarz

The choice  $p = q = 2$  in Result 1.8 gives

$$\int_{\Omega} |fg| d\mu \leq \left( \int_{\Omega} f^2 d\mu \right)^{1/2} \left( \int_{\Omega} g^2 d\mu \right)^{1/2} = \|f\|_2 \|g\|_2.$$

#### Theorem 1.10: Minkowski's inequality

Let  $p \in [1, \infty]$  and let  $f, g$  be measurable. Then

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

In particular,  $L^p(\Omega, \mathcal{F}, \mu)$  is closed under addition and  $\|\cdot\|_p$  is a seminorm; on the quotient by  $\mu$ -a.e. equality it is a norm.

**Remark 1.4.** For  $p > 1$ , Minkowski's bound is obtained by writing  $|f + g|^p \leq 2^{p-1}(|f|^p + |g|^p)$  (so  $f + g \in L^p$ ) and then applying Hölder to the splitting  $\int |f + g|^p = \int |f| |f + g|^{p-1} + \int |g| |f + g|^{p-1}$  with conjugate exponents  $p$  and  $q = p/(p-1)$ .

### 1.5 Approximation in $L^p$

The next theorem says that “simple functions supported on sets of finite measure” are dense in  $L^p$  for  $p \in [1, \infty)$ . It is the standard tool for reducing analytic statements to a check on indicator functions.

**Theorem 1.11: Density of simple functions in  $L^p$** 

Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and assume there exist  $A_n \in \mathcal{F}$  with  $A_n \uparrow \Omega$  and  $\mu(A_n) < \infty$  for all  $n$  (i.e.  $\mu$  is  $\sigma$ -finite). Let

$$V_0 = \text{span}\{\mathbf{1}_A : A \in \mathcal{F}, \mu(A) < \infty\}$$

denote the simple functions supported on sets of finite measure. Then for every  $p \in [1, \infty)$ ,  $V_0 \subseteq L^p$  and for every  $f \in L^p$  and every  $\varepsilon > 0$  there exists  $v \in V_0$  with

$$\|f - v\|_p < \varepsilon.$$

**Remark 1.5.** The proof has the usual three-step shape: (i) a Dynkin  $\pi$ - $\lambda$  argument shows the class  $\mathcal{L} = \{A \in \mathcal{F} : \mathbf{1}_A \text{ is approximable}\}$  is a  $\lambda$ -system containing the generating  $\pi$ -system, hence all of  $\mathcal{F}$ ; (ii) for non-negative  $f \in L^p$  the truncations  $f_n = \min(n, 2^{-n} \lfloor 2^n f \rfloor)$  satisfy  $|f - f_n|^p \rightarrow 0$  pointwise with  $|f - f_n|^p \leq |f|^p$ , so dominated convergence gives  $\|f - f_n\|_p \rightarrow 0$ ; (iii) general  $f$  is handled by splitting  $f = f^+ - f^-$  and restricting to the exhausting sets  $A_n$ .