10-752: Advanced Probability **Spring 2018**

Lecture 9: February 27

Lecturer: Alessandro Rinaldo *Scribe: Matteo Bonvini*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Last time

Consider the usual triplet $(\Omega, \mathcal{F}, \mu)$, and, for $p \geq 1$, define:

$$\|f\|_p = \left[\int |f|^p d\mu \right]^{1/p}$$

The set L_p is the set of all equivalent classes of functions such that $\|f\|_p < \infty$, where two functions, f and g, are in the same class if $f = g$ *a.e.* $[\mu]$.

We can also define the "infinity" norm:

$$\|f\|_\infty = \inf \{\alpha \geq 0 : \mu(\{\omega : |f(\omega)| \geq \alpha\}) = 0\}$$

$\|f\|_\infty$ is generally called "the essential supremum" of a function and is not necessarily equal to $\sup_{\omega\in\Omega} |f(\omega)|$. We can think of $\|f\|_\infty$ as the smallest number α such that $f \leq \alpha$ a.e. $[\mu]$

Proposition 9.1 *If $\mu(\Omega) < \infty$, then $p \leq q \leftrightarrow \|f\|_p \leq C\|f\|_q$, where C is a constant that depends on $\mu(\Omega)$, p, and q. In particular, if $\mu(\Omega) = 1$, then $C = 1$.*

Corollary 9.2 *If $\mu(\Omega) = 1$, then the following chain of inequalities holds for $p \leq q$:*

$$\|f\|_1 \leq |f\|_p \leq |f\|_q \leq |f\|_\infty \tag{9.1}$$

and $L_q \subset L_p$.

If $\mu(\Omega) = \infty$, then there is no set relationship among L_p spaces. In particular, example 4 from lecture notes shows that if $\Omega = \mathbb{R}^+$ and μ is the Lebesgue measure, then there exists a function f such that $f \notin L_1$, $f \in L_2$ and $f \notin L_3$.

If μ is a counting measure on Ω, even if $\mu(\Omega) = \infty$, some set relationships can be established. In particular, we define the ℓ_p space to be the space of all sequences $\{x_n\}_{n=1}^\infty$ such that $(\sum_n |x_n|^p)^{1/p}$ is finite.

Proposition 9.3 *Let $1 \leq p \leq q$ and let ℓ_p and ℓ_q be spaces defined as above. Then, we have $\|f\|_\infty = \max_{\omega\in\Omega} |f(\omega)|$ and*

$$\|f\|_\infty \leq \|f\|_q \leq \|f\|_p \leq \|f\|_1$$

Proof:
If $q = \infty$, claim follows. If $q < \infty$, define $\tilde{f}(\omega) = \frac{f(\omega)}{\|f\|_p}$. Then, we have that $\sum_\omega |\tilde{f}(\omega)|^p = 1$ and $0 \leq |\tilde{f}(\omega)|^p \leq 1$. In particular, we have $|\tilde{f}(\omega)|^p \geq |\tilde{f}(\omega)|^q$ since $p \leq q$. Thus:

$$\frac{\|f\|_q}{\|f\|_p} = \left[\sum_\omega |\tilde{f}(\omega)|^q\right]^{1/q} \leq \left[\sum_\omega |\tilde{f}(\omega)|^p\right]^{1/q} = 1$$

■

Before proving Holder Inequality, we need the following lemma.

Lemma 9.4 *If $a, b > 0$ and $\lambda \in (0, 1)$, the following inequality holds:*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda) b$$

Proof: By Jensen's inequality (below), we have

$$\log(\lambda a + (1-\lambda)b) \geq \lambda \log a + (1-\lambda)\log b = \log a^\lambda + \log b^{1-\lambda} = \log(a^\lambda b^{1-\lambda})$$

Exponentiating both sides yields the desire inequality. ■

Proposition 9.5 (Holder Inequality) *Let $p, q \geq 1$ such that $1/p + 1/q = 1$. Let $X \in L_p \leftrightarrow [E|X|^p]^{1/p} < \infty$ and $Y \in L_q \leftrightarrow [E|Y|^q]^{1/q} < \infty$. Then, $XY \in L_1$ and $\|XY\|_1 \leq \|X\|_p \|X\|_q$.*

Proof:
If $p = 1$, then $q = \infty$ and claim follows. Suppose $p, q < \infty$. Let $U = |X|^p$ and $V = |Y|^q$. Notice that $U, V \in L_1$. By Lemma 9.4, we have that:

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

Then, we take the integral on both sides of the inequality:

$$\frac{\int U^{1/p} V^{1/q} \; d\mu}{\int U \; d\mu \int V \; d\mu} \leq \frac{1}{p} + \frac{1}{q} = 1$$

Substituting X and Y in the expression leads to the claim. ■

Proposition 9.6 (Jensen's Inequality) *Let f be a convex function on $I = (a, b)$, where $-\infty \leq a < b \leq \infty$ and let X be a random variable such that $P(X \in I) = 1$. Suppose further that EX exists. Then, $Ef(X) \geq f(EX)$.*

Proof:
Let ℓ be the supporting hyperplane to f at a point z. Consider $z = EX$. By convexity, we have that $f(x) \geq \ell(x)$ and $f(z) = \ell(z)$. Because ℓ is a linear function and E is a linear operator, we have

$$E[f(X)] \geq E[\ell(X)] = \ell(E[X]) = \ell(z) = f(z) = f(E[X])$$

In particular, equality holds when f is linear or $X = EX$ a.e. ■

Proposition 9.7 (Minkovsky) *If $X, Y \in L_p$ then $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$.*

Proposition 9.8 (**C_r Inequality**) *$E|X+Y|^r \le C_r\left(|X|^r + |Y|^r\right)$ and $C_r = 1$ if $r \in (0,1]$ or $C_r = 2^{r-1}$ is $r > 1$.*

Proposition 9.9 (Markov Inequality) *Let $X \ge 0$ a.e.. Then, for any $c > 0$, $Pr(X \ge c) \le \frac{EX}{c}$.*

Proof: Notice that $X \ge c \cdot I_{X \ge c}$. Therefore:

$$EX \ge c\Pr(X \ge c)$$

and the claim follows. ■

Proposition 9.10 (Chebyshev) *The following inequality holds*

$$Pr(|X - EX| \ge c) \le \frac{Var(X)}{c^2}$$

Proof: Notice that the event $|X - EX| \ge c$ is the same as $(X - EX)^2 \ge c^2$. Then the claim follows by Markov Inequality. ■

Conditional Probability & Conditional Expectation

Let us start with a triplet $(\Omega, \mathcal{F}, P)$. The intuition behind conditional probability is that we would like to measure an event leveraging on some extra information. Extra information is usually thought as the information contained in a sub-σ-field of $\mathcal{F}$.

For example, let $B_1, \ldots, B_n$ be a partition of Ω and A any a subset in $\mathcal{F}$. Let $\mathcal{C} = \sigma(B_1, \ldots, B_n) \subseteq \mathcal{F}$ and $f(\omega) = \frac{\Pr(\omega \in A \cap B_i)}{\Pr(\omega \in B_i)} = P(A|B_i)$ if $\omega \in B_i$. We notice that f is measurable with respect to $\mathcal{C}$ and is such that

$$P(A \cap B_i) = P(A|B_i)P(B_i) = \int_{B_i} f(\omega) dP(\omega)$$

In general, for any fixed set A, we can use the R-N derivative to define a measure on $(\Omega, \mathcal{C})$, given by $B \in \mathcal{C}$, defined as:

$$\nu(B) = P(A \cap B) = \int_B f(\omega) dP(\omega)$$

A valid conditional probability measure is a function $f(\omega) = P(A|C)(\omega)$ that satisfies the following properties:

1. $\Pr(A|C)$ is measurable wrt $\mathcal{C}$
2. $\forall B \in \mathcal{C}$, $\nu(B) = \Pr(\omega \in A \cap B) = \int_B \Pr(A|C)(\omega) dP(\omega)$
3. $\Pr(A|C)$ is unique a.e. $[P]$

We can define conditional expectation in a similar way. Let X be a random variable such that $E|X| < \infty$ and $\mathcal{C} \subseteq \mathcal{F}$. Then, we define the conditional expectation of X given $\mathcal{C}$, denoted $E[X|\mathcal{C}]$ as any function $h : \Omega \to \mathbb{R}$ that is $\mathcal{C}/\mathcal{B}_1$ measurable and such that

$$\int_B h(\omega) dP(\omega) = \int_B X(\omega) dP(\omega) \ \forall B \in \mathcal{C}$$

Any other function $f : \Omega \to \mathbb{R}$ satisfying these conditions is called a version of $E[X|\mathcal{C}]$.

For example, let $\mathcal{C} = \{\emptyset, \Omega\}$, notice that $EX(\omega) : \Omega \to \mathcal{R}$ is $\mathcal{C}/\mathcal{B}_1$ measurable.[1] Moreover, $\int_\Omega EXdP(\omega) = EX = \int_\Omega XdP(\omega)$ and $\int_\emptyset EXd\mu = 0 = \int_\emptyset XdP(\omega)$. Hence, we conclude that a version of $E[X|\mathcal{C}]$ when $\mathcal{C}$ is the trivial σ-field is $E[X]$.

Let (X, Y) a vector of random variables with density $f_{X,Y}$ with respect to some σ-finite dominating measure on $\mathbb{R}^2$. Let $f_Y(y) = \int_\mathbb{R} f_{X,Y}(x, y)dx$, then, assuming $f_Y(y) \neq 0$, we define $f_{X|Y}(x) = f_{X,Y}(x, y)/f_Y(y)$.

Proposition 9.11 *Let $\mathcal{C} = \sigma(Y)$, then a version of $E[X|\mathcal{C}]$ is given by $g(y) = \int_\mathbb{R} xf_{X|Y}(x, y)dx$.*

Proof: Let $\mathcal{C} = \sigma(Y)$, then $g(Y)$ is $\mathcal{C}$-measurable. Therefore, we only need to show that $\int_B g(Y)dP(\omega) = \int_B XdP(\omega)$ for all $B \in \mathcal{C}$.

Let $B \in \mathcal{C}$, then there exist $A \in \mathbb{R}$ such that $B = Y^{-1}(A)$. Hence, we have $I_B(\omega) = I_A(Y(\omega))$ and

$$\begin{aligned}
\int_B g(Y(\omega))dP(\omega) &= \int_\mathbb{R} I_A(y)g(y)d\mu(y) \\
&= \int_\mathbb{R} I_A(y)g(y)f_Y(y)dy \quad \text{[R-N derivative]} \\
&= \int_\mathbb{R} I_A(y)\int_\mathbb{R} xf_{X|Y}(x, y)dxf_Y(y)dy \\
&= \int_\mathbb{R}\int_\mathbb{R} xI_A(y)f_{X,Y}(x, y)dxdy \\
&= E_{X,Y}[X \cdot I_A(Y)] \\
&= E[X \cdot I_B(X)] = \int_B XdP(\omega)
\end{aligned}$$

■

[1] $E(X)$ is a constant, and we have that the σ-field generated by Z is the trivial σ-field if and only if Z is a constant function.