

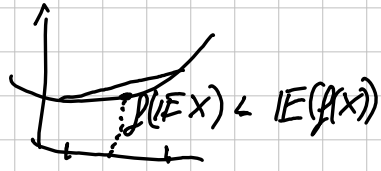
Lecture 7:

Recall: $\|X\|_p = \mathbb{E}(|X|^p)^{1/p}$; $p \geq 1$

We write $L^p(\Omega, \mathcal{F}, P)$ for all X s.t.

$$\|X\|_p < \infty.$$

• Jensen's inequality:



$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$ for convex functions f

• $\|X\|_p \geq \|X\|_q \Rightarrow L^q(\Omega, \mathcal{F}, P) \supseteq L^p(\Omega, \mathcal{F}, P)$
for $p \geq q \geq 1$

Cauchy - Schwarz inequality: If X, Y are in

$L^2(\Omega, \mathcal{F}, P)$ then XY is integrable and

$$|\mathbb{E}(XY)| \leq \mathbb{E}(|XY|) \leq \|X\|_2 \|Y\|_2 = \sqrt{\mathbb{E}(X^2) \mathbb{E}(Y^2)}.$$

Cauchy - Schwarz implies that $\langle X, Y \rangle = \mathbb{E}(XY)$

becomes a well-defined inner product.

Proof (C.S.): Truncate X, Y :

$$X_n = \min(|X|, n), \quad Y_n = \min(|Y|, n)$$

which are clearly bounded r.v.s.

For all a, b we have

$$\mathbb{E}((aX_n + Y_n)^2) \geq 0$$

$$f(a) = a^2 \mathbb{E}(X_n^2) + 2a \mathbb{E}(X_n Y_n) + \mathbb{E}(Y_n^2) \geq 0 \quad (\#)$$

Consider the above as a quadratic in a .

For $(\#)$ to hold, $f(a) = 0$ can have at most 1 solution. Hence its discriminant must be non-positive and we get:

$$(2\mathbb{E}(X_n Y_n))^2 - 4\mathbb{E}(X_n^2)\mathbb{E}(Y_n^2) \leq 0$$

$$\text{So } \mathbb{E}(X_n Y_n)^2 \leq \mathbb{E}(X_n^2)\mathbb{E}(Y_n^2)$$

Taking limits & appealing to MCT gives

$$\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2) \quad \square$$

Note: We used X_n, Y_n above to ensure $\mathbb{E}(X_n Y_n)^2$ is finite. This may, a priori, not be true.

Corollary: $\|X + Y\|_2 \leq \|X\|_2 + \|Y\|_2$

Proof: $\|X + Y\|_2^2 = \mathbb{E}((X + Y)^2)$

$$= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2)$$

$$\leq \mathbb{E}(X^2) + 2\|X\|_2\|Y\|_2 + \mathbb{E}(Y^2)$$

$$= \|X\|_2^2 + 2\|X\|_2\|Y\|_2 + \|Y\|_2^2$$

$$= (\|X\|_2 + \|Y\|_2)^2$$

□

The triangle inequality satisfies the triangle inequality

Defⁿ Let X, Y be random variables, with

$$m_x = \mathbb{E}(X), m_y = \mathbb{E}(Y). \text{ We set}$$

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}((X - m_x)(Y - m_y))$$

$$\text{Var}(X) = \text{Cov}(X, X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}((X - m_x)^2)$$

We have:

- $\text{Var}(X) \geq 0$
- If X, Y independent, $\text{Cov}(X, Y) = 0$
- $|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$ by C.-S.

⇒ the correlation coefficient

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}} \in [-1, 1]$$

It is equal to ± 1 if and only if there is a linear relation between X & Y .

The Cauchy-Schwarz inequality can be extended to the Hölder inequality:

Assume $X \in L^p$, $Y \in L^q$ for $\frac{1}{p} + \frac{1}{q} = 1$; $p, q \geq 1$.

Then, $|\mathbb{E}(XY)| \leq \mathbb{E}(|XY|) \leq \|X\|_p \|Y\|_q$.

C.S. is the case when $p=q=2$.

Remark: This holds for arbitrary measure spaces

$$|\int f \cdot g \, d\mu| \leq \int |f \cdot g| \, d\mu \leq \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} \left(\int |g|^q \, d\mu \right)^{\frac{1}{q}}$$

Corollary: $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$ for $p \geq 1$.

(Known as Minkowski's inequality)

Proof: We may assume $p > 1$. (otherwise just the triangle ineq.)

$$\begin{aligned} (\|X+Y\|_p)^p &= \mathbb{E}(|X+Y|^p) = \mathbb{E}(|X+Y| |X+Y|^{p-1}) \\ &\leq \mathbb{E}(|X| |X+Y|^{p-1}) + \mathbb{E}(|Y| |X+Y|^{p-1}) \\ &\leq \|X\|_p \| |X+Y|^{p-1} \|_q + \|Y\|_p \| |X+Y|^{p-1} \|_q \\ &\quad \text{where } \frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow q = \frac{p}{p-1} \\ &= \|X\|_p \mathbb{E}(|X+Y|^{q(p-1)})^{\frac{1}{q}} + \|Y\|_p \mathbb{E}(|X+Y|^{q(p-1)})^{\frac{1}{q}} \\ &= (\|X\|_p + \|Y\|_p) \underbrace{\mathbb{E}(|X+Y|^p)^{\frac{1}{q}}}_{(\|X+Y\|_p)^{\frac{p}{q}}} \\ &\Rightarrow (\|X+Y\|_p)^{p - \frac{p}{q}} \leq \|X\|_p + \|Y\|_p \\ &\quad \text{But } p - \frac{p}{q} = 1 \text{ and claim follows } \square \end{aligned}$$

Thus $\|\cdot\|_p$ is a norm for $p \geq 1$.

Theorem: L^p is a complete space, i.e. Cauchy sequences converge.

Densities

Suppose X is a random variable with law Λ_X : $\Lambda_X(A) = \mathbb{P}(X \in A)$; $A \in \Sigma$.

Th^m For every Borel measurable function f , we have

$$\mathbb{E}(f(X)) (= \int_{\Omega} f(X) d\mathbb{P}) = \int_{\mathbb{R}} f(x) d\Lambda_X$$

Proof: First for indicator functions:

Let $f = I_A$, then

$$\int_{\Omega} I_A(X) d\mathbb{P} = \mathbb{E}(I_{\{X \in A\}}) = \mathbb{P}(X \in A) \quad \text{and}$$

$$\int_{\mathbb{R}} I_A(x) d\Lambda_X = \int_A 1 d\Lambda_X = \Lambda_X(A) = \mathbb{P}(X \in A).$$

The property then extends by linearity to step functions and to $f \in \Sigma^+$ by the MCT.

Finally, for $f \in m\mathbb{Z}$ this is shown by splitting into f^+ & f^- . □

If X also has a density, we can express $E(f(X))$ in terms of a density.

Th^m If X has density φ (and $P(X \in A) = \int_A \varphi(x) dx$) then

$$E(f(X)) = \int_{\mathbb{R}} f(x) \varphi(x) dx$$

Proof: As before. □

Remark: φ here is the same "density" as in the Radon-Nikody sense.

Recall: that X, Y are independent if $P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A)P(Y \in B)$ for all $A, B \in \Sigma$. We get:

Th^m If X, Y are independent and integrable, then so is XY and $E(XY) = E(X)E(Y)$.

Proof: We can assume X, Y are non-negative (otherwise split and consider $X^+, X^-,$ etc.)
We can approximate X, Y by increasing

step functions $\alpha^{(r)}(X) \uparrow X, \alpha^{(r)}(Y) \uparrow Y.$

Each is a linear combination of indicator functions $I_A(X) = \begin{cases} 1 & X \in A \\ 0 & \text{otherwise} \end{cases}, I_B(Y) = \begin{cases} 1 & Y \in B \\ 0 & \text{otherwise} \end{cases}$

and for each indicator we have

$$\begin{aligned} \mathbb{E}(I_A(X) I_B(Y)) &= P(\{X \in A\} \cap \{Y \in B\}) \\ &= P(X \in A) P(Y \in B) = \mathbb{E}(I_A(X) I_B(X)). \end{aligned}$$

We extend this by linearity to

$$\mathbb{E}(\alpha^{(r)}(X) \alpha^{(r)}(Y)) = \mathbb{E}(\alpha^{(r)}(X)) \mathbb{E}(\alpha^{(r)}(Y)).$$

Taking $r \rightarrow \infty$ and using MCT

gives $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$ as required. \square

Corollary: If X, Y are independent,
then $\text{Cov}(X, Y) = 0$, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$.

Proof: First is immediate from definition.

The second follows from:

$$E((X+Y)^2) = E(X^2) + 2E(XY) + E(Y^2) \quad (1)$$

$$E(X+Y)^2 = E(X^2) + 2E(X)E(Y) + E(Y^2) \quad (2)$$

and (1) - (2) gives

$$\text{Var}(X+Y) = E(X+Y)^2 - E(X+Y)^2$$

$$= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2$$

$$= \text{Var}(X) + \text{Var}(Y). \quad \square$$

Important: $E(X)E(Y) = E(XY) \not\Rightarrow X, Y$ independent.

Counter example:

$X \backslash Y$	1	0	-1
1	$\frac{1}{4}$	0	$\frac{1}{4}$
-1	0	$\frac{1}{2}$	0

$$E(X) = E(Y) = 0$$

$$E(XY) = \frac{1}{4} \cdot (1 \cdot 1) + \frac{1}{2} \cdot (-1 \cdot 0) + \frac{1}{4} \cdot (1 \cdot (-1)) = 0$$

Not independent as $P(\{X=1\} \cap \{Y=0\}) = 0 \neq P(X=1)P(Y=0) = \frac{1}{2} \cdot \frac{1}{2}$