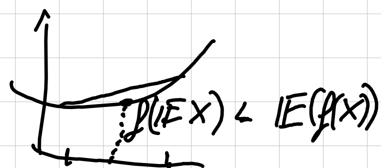
Lecture 7:

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

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

- Jensen's inequality:



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

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

<u>Cauchy - Schwarz inequality</u>: If X, Y are in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ 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:

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

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 it discriminant must be non-positive and we get:

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

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

$$\mathbb{E}(|X||Y|)^2 \leq \mathbb{E}(|X|^2)\mathbb{E}(|Y|^2) \qquad \square$$

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

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

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

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

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

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

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

The triangle inequality satisfies the triangle inequality

Def$^{\underline{n}}$ Let X, Y be random variables. with

$m_X = \mathbb{E}(X)$, $m_Y = \mathbb{E}(Y)$. We set

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}((X-m_X)(Y-m_Y))$$

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

We have:

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

$\Rightarrow$ the correlation coefficient

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

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 $p = q = 2$.

<u>Remark</u>: This holds for arbitrary measure spaces

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

<u>Corollary</u>: $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$ for $p \geq 1$.

(Known as Minkowski's inequality)

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

$$\left(\|X+Y\|_p \right)^p = \mathbb{E}(|X+Y|^p) = \mathbb{E}(|X+Y|\,|X+Y|^{p-1})$$
$$\leq \mathbb{E}(|X|\,|X+Y|^{p-1}) + \mathbb{E}(|Y|\,|X+Y|^{p-1})$$
$$\leq \|X\|_p \, \| |X+Y|^{p-1} \|_q + \|Y\|_p \, \| |X+Y|^{p-1} \|_q$$

where $\frac{1}{p} + \frac{1}{q} = 1 \iff q = \frac{p}{p-1}$

$$= \|X\|_p \, \mathbb{E}(|X+Y|^{q(p-1)})^{\frac{1}{q}} + \|Y\|_p \, \mathbb{E}(|X+Y|^{q(p-1)})^{\frac{1}{q}}$$
$$= (\|X\|_p + \|Y\|_p) \underbrace{\mathbb{E}(|X+Y|^p)^{\frac{1}{q}}}_{(\|X+Y\|_p)^{p/q}}$$

$$\Rightarrow (\|X+Y\|_p)^{p-\frac{p}{q}} \leq \|X\|_p + \|Y\|_p$$

But $p - \frac{p}{q} = 1$ and claim follows $\square$

Thus $\|.\|_p$ 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$^{\underline{m}}$ For every Borel measurable function f, we have

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

Proof: First for indicator functions:

Let $f = I_A$, then

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

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

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

Finally, for $f \in m\Sigma$ this is shown by splitting into f^+ & f^-. $\square$

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

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

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

Proof: As before. $\square$

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

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

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

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 $I_A(X) = \begin{cases} 1 & X \in A \\ 0 & \text{otherwise} \end{cases}$, $I_B(X) = \begin{cases} 1 & Y \in B \\ 0 & \text{otherwise} \end{cases}$,

and for each indicator we have

$$\mathbb{E}(I_A(X) I_B(Y)) = \mathbb{P}(\{X \in A\} \cap \{Y \cap B\})$$
$$= \mathbb{P}(X \in A)\mathbb{P}(Y \in B) = \mathbb{E}(I_A(X)\, I_B(X)).$$

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 $r \to \infty$ and using MCT gives $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ as required. $\square$

<u>Corollary:</u> If X, Y are independent, then $\mathrm{Cov}(X,Y) = 0$, $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

<u>Proof:</u> First is immediate from definition.

The secon follows from:

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

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

and (1) − (2) gives

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y).\end{aligned}$$

$\square$

<u>Important:</u> $\mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) \not\Rightarrow X, Y$ independent.

<u>Counterexample:</u>

X \ Y	1	0	−1
1	1/4	0	1/4
−1	0	1/2	0

$\mathbb{E}(X) = \mathbb{E}(Y) = 0$

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

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