APA lecture 4

Expectation (a.k.a. Lebesgue's integral)

Let $X : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$

Ex: $X \sim \text{Bernoulli}(p)$: $\mathbb{E}(X) = 1 \cdot \underbrace{\mathbb{P}(\{X=1\})}_{p} + 0 \cdot \mathbb{P}(\{X=0\}) = p$

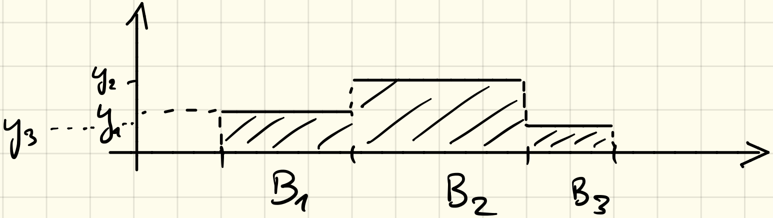
Let $g : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function

Our aim: to define $\mathbb{E}(g(X))$

3 steps:
1. g is a "simple" function
2. g is a non-negative function
3. g is a general function

<u>Step 1</u>: g is a <u>simple function</u>, i.e.

$$g(x) = \sum_{i \geq 1} y_i \cdot 1_{B_i}(x) \qquad \underset{\text{distinct}}{y_i \geq 0}, \quad \underset{\text{disjoint}}{B_i} \in \mathcal{B}(\mathbb{R})$$



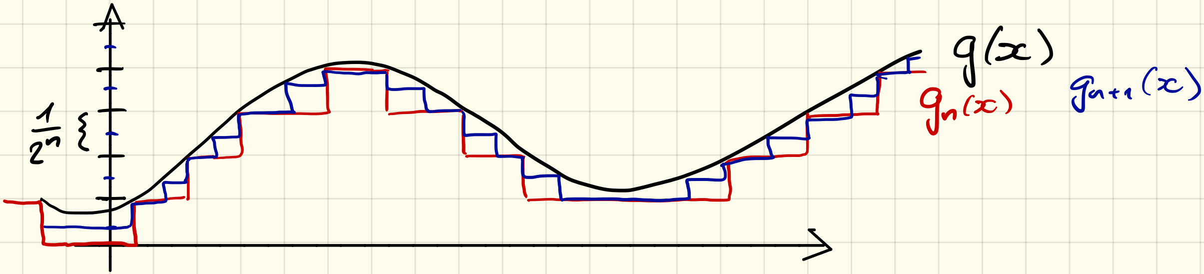
$$\mathbb{E}(g(X)) = \sum_{i \geq 1} \underbrace{y_i}_{\geq 0} \cdot \mathbb{P}(\underbrace{\{X \in B_i\}}_{= \{g(X) = y_i\}}) \in [0, +\infty]$$

$\in \mathcal{F}$

<u>Ex</u>:

- $g(x) = 1_B(x) \qquad \mathbb{E}(g(X)) = \mathbb{P}(\{X \in B\})$
- $g_t(x) = 1_{(-\infty, t]}(x) \qquad \mathbb{E}(g_t(X)) = \mathbb{P}(\{X \leq t\}) = F_X(t)$

Step 2 : g is non-negative (& Borel-measurable)



$$g_n(x) = \sum_{i \geq 1} \frac{i-1}{2^n} \cdot 1_{\left\{ \frac{i-1}{2^n} < g(x) \leq \frac{i}{2^n} \right\}} \leq g(x)$$

$$\mathbb{E}(g_n(X)) = \sum_{i \geq 1} \frac{i-1}{2^n} \cdot P\left(\left\{ \frac{i-1}{2^n} < g(X) \leq \frac{i}{2^n} \right\} \right) \in [0, +\infty]$$

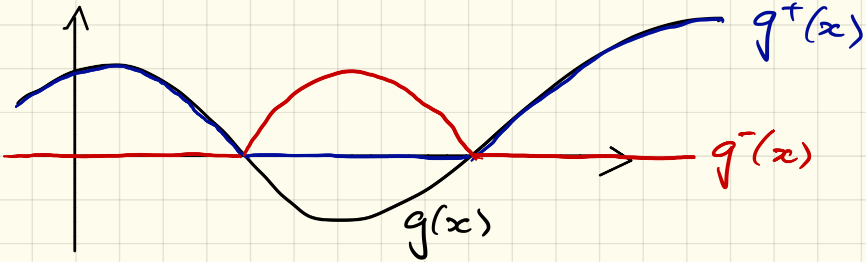
$$g_n(x) \leq g_{n+1}(x) \leq g(x) \quad \& \quad \mathbb{E}(g_n(X)) \leq \mathbb{E}(g_{n+1}(X))$$

increasing sequence

$$\mathbb{E}(g(X)) = \lim_{n \to \infty} \mathbb{E}(g_n(X)) \quad \in [0, +\infty]$$

Step 3 : g is Borel-measurable

$$\begin{cases} g^+(x) = \max(g(x), 0) \geq 0 & \text{positive part of } g \\ g^-(x) = \max(-g(x), 0) \geq 0 & \text{negative part of } g \end{cases}$$



$$g(x) = g^+(x) - g^-(x) \qquad |g(x)| = g^+(x) + g^-(x)$$

Def: If $\mathbb{E}(|g(X)|) < +\infty$, then we define

$$\mathbb{E}(g(X)) = \mathbb{E}(g^+(X)) - \mathbb{E}(g^-(X)) \in \mathbb{R}$$

Two particular cases:

a) X is a discrete r.v. with pmf $(p_x, x \in \underset{\text{countable set}}{C})$:

If $\mathbb{E}(|g(X)|) = \sum_{x \in C} |g(x)| \cdot p_x < +\infty$

Then $\mathbb{E}(g(X)) = \sum_{x \in C} g(x) \cdot p_x \in \mathbb{R}$

<u>NB</u>: $p_x = \mathbb{P}(\{X = x\})$

b) X is a continuous r.v. with pdf p_X:

If $\mathbb{E}(|g(X)|) = \int_{\mathbb{R}} |g(x)| \cdot p_X(x)\, dx < +\infty$

Then $\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) \cdot p_X(x)\, dx \in \mathbb{R}$

Terminology

- If $\mathbb{E}(|X|) < +\infty$, we say that X is <u>integrable</u>.
- If $\mathbb{E}(X^2) < +\infty$, we say that X is <u>square-integrable</u>.
- If $\exists C > 0$ s.t. $|X(\omega)| \leq C \ \forall \omega \in \Omega$, we say that X is <u>bounded</u>.
- If $\mathbb{E}(X) = 0$, we say that X is <u>centered</u>.
- If $X \sim -X$, we say that X is <u>symmetrically distributed</u>.

Implications:

- X bounded $\Longrightarrow$ X square-integrable $\Longrightarrow$ X integrable
- X symmetrically distributed $\Longrightarrow$ X centered
- $\left\{ \begin{array}{l} X, Y \text{ square-integrable} \\ \qquad\qquad \text{OR} \\ X \text{ integrable}, Y \text{ bounded} \end{array} \right\} \Longrightarrow X \cdot Y$ integrable

Proofs :

- X bounded $\Rightarrow \mathbb{E}(|X|) \leq \mathbb{E}(C) = C < +\infty$

  i.e. X integrable

- X bounded (by C) $\Rightarrow X^2$ bounded (by C^2) $\Rightarrow X$ sq-integrable
- X square-integrable : $(1 - |X|)^2 \geq 0$ so $2|X| \leq 1 + X^2$

  so $2\mathbb{E}(|X|) \leq 1 + \mathbb{E}(X^2)$

  $\Rightarrow X$ integrable

- X symmetrically distributed : $\mathbb{E}(X) = \mathbb{E}(-X)$ i.e. $2 \cdot \mathbb{E}(X) = 0$
- X, Y square-integrable : $(|X| - |Y|)^2 \geq 0$ so $2|X| \cdot |Y| \leq X^2 + Y^2$

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

- X integrable, Y bounded (by C) : $|(X \cdot Y)(\omega)| \leq |X(\omega)| \cdot C \quad \forall \omega$

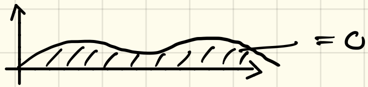
Basic properties

Linearity: if $c \in \mathbb{R}$, X, Y are integrable r.v.'s, then

$$\mathbb{E}(c \cdot X) = c \cdot \mathbb{E}(X) \quad \& \quad \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

Positivity:
- if X is integrable & non-negative, then $\mathbb{E}(X) \geq 0$
- if X is integrable & $\underbrace{\mathbb{P}(\{X \geq 0\}) = 1}_{X \geq 0 \text{ almost surely (a.s.)}}$, then $\mathbb{E}(X) \geq 0$

Strict positivity:
- if X is integrable, $X \geq 0$ a.s. and $\mathbb{E}(X) = 0$, then $X = 0$ a.s.



eq.

- if $X \geq 0$ a.s. & $\mathbb{P}(\{X > 0\}) > 0$, then $\mathbb{E}(X) > 0$

Monotonicity:
- if X, Y are integrable & $X \geq Y$ a.s., then $\mathbb{E}(X) \geq \mathbb{E}(Y)$

Variance and covariance (& independence)

Def: Let X, Y be two square-integrable r.v.'s

- The variance of X is: $Var(X) = \mathbb{E}((X-\mathbb{E}(X))^2)$

$$= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0$$

- The covariance of X, Y is: $Cov(X,Y) = \mathbb{E}((X-\mathbb{E}(X)) \cdot (Y-\mathbb{E}(Y)))$

$$= \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y) \in \mathbb{R}$$

- $Cov(X,X) = Var(X)$

Terminology: If $Cov(X,Y) = 0$, we say that X, Y are uncorrelated.

Facts:
- $Var(c \cdot X) = c^2 \cdot Var(X)$ $\leftarrow$
- $Var(X+Y) = Var(X) + Var(Y) + 2\, Cov(X,Y)$

If $X \perp\!\!\!\perp Y$, then: $Cov(X,Y) = 0$ so: $\begin{cases} \mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y) \\ Var(X+Y) = Var(X) + Var(Y) \end{cases}$

Characteristic function (or Fourier transform)

Def: The characteristic function of a random variable X is the mapping $\phi_X : \mathbb{R} \to \mathbb{C}$ defined as

$$\phi_X(t) = \mathbb{E}\left(e^{itX}\right) \qquad t \in \mathbb{R}$$

Note: $\phi_X(t)$ is always well defined, as $\mathbb{E}\left(\underbrace{|e^{itX}|}_{=1}\right) = 1 < +\infty$

Ex:

- $X = \begin{cases} +1 & \text{wp } \frac{1}{2} \\ -1 & \text{wp } \frac{1}{2} \end{cases}$ $\qquad \phi_X(t) = \frac{1}{2}\left(e^{it} + e^{-it}\right) = \cos(t)$

- $X \sim \mathcal{U}([-1,+1])$ $\qquad \phi_X(t) = \frac{1}{2}\int_{-1}^{+1} e^{itx}\,dx = \frac{e^{it} - e^{-it}}{2it} = \frac{\sin(t)}{t}$

- $X \sim N(0,1)$ $\qquad \phi_X(t) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{itx}\,dx = \dots = e^{-t^2/2}$

Properties:

(i) $\phi_X(0) = 1$

(ii) ϕ_X is continuous on $\mathbb{R}$

(iii) ϕ_X is positive semi-definite on $\mathbb{R}$, i.e.

$$\sum_{j,k=1}^{n} c_j \overline{c_k} \underbrace{\phi_X(t_j - t_k)}_{= A_{jk}} \geq 0 \quad \forall n \geq 1,\ c_1 \dots c_n \in \mathbb{C},\ t_1 \dots t_n \in \mathbb{R}$$

Proof of (iii): $\displaystyle\sum_{j,k=1}^{n} c_j \overline{c_k}\, \mathbb{E}\left(e^{i(t_j - t_k)X}\right)$

$$= \mathbb{E}\Big(\sum_{j=1}^{n} c_j e^{it_jX} \cdot \underbrace{\sum_{k=1}^{n} \overline{c_k} e^{-it_kX}}_{= \overline{\sum_{k=1}^{n} c_k e^{it_kX}}}\Big) = \mathbb{E}\Big(\Big|\sum_{j=1}^{n} c_j e^{it_jX}\Big|^2\Big) \geq 0 \quad \#$$

(iv) $\phi_X(-t) = \mathbb{E}(e^{-itX}) = \overline{\mathbb{E}(e^{itX})} = \overline{\phi_X(t)} \quad \forall t \in \mathbb{R}$

(v) $|\phi_X(t)| = |\mathbb{E}(e^{itX})| \leq \mathbb{E}(\underbrace{|e^{itX}|}_{=1}) = 1 = \phi_X(0) \quad \forall t \in \mathbb{R}$

(vi) $\phi_{cX}(t) = \mathbb{E}(e^{itcX}) = \mathbb{E}(e^{ictX}) = \phi_X(ct) \quad \forall t \in \mathbb{R} \ \forall c \in \mathbb{R}$

<u>Thm</u>: Any function $\phi : \mathbb{R} \to \mathbb{C}$ satisfying properties (i), (ii), (iii) is the characteristic function of some random variable X (so it automatically also satisfies properties (iv), (v), (vi)).

Inversion formula:

- To a given characteristic function ϕ_X corresponds a unique cdf F_X given by:

$$F_X(b) - F_X(a) = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-ita} - e^{-itb}}{it} \cdot \phi_X(t)\, dt$$

for $a < b$ which are continuity points of F_X.

- If in addition $\int_{\mathbb{R}} |\phi_X(t)|\, dt < +\infty$, then F_X admits a pdf p_X given by

$$p_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \phi_X(t)\, dt \quad \forall x \in \mathbb{R}$$

Relation between F_X & ϕ_X :

- If $\mathbb{E}(|X|) < +\infty$, then ϕ_X is continuously differentiable and $\phi_X'(0) = i \cdot \mathbb{E}(X)$:

$$\frac{d}{dt}\phi_X(t) = \frac{d}{dt}\mathbb{E}(e^{itX}) \overset{\sim}{=} \mathbb{E}\left(\frac{d}{dt} e^{itX}\right) = \mathbb{E}\left(iX e^{itX}\right)$$

so $\phi_X'(0) = \mathbb{E}(iX) = i\,\mathbb{E}(X)$

- If $\mathbb{E}(X^2) < +\infty$, then ϕ_X is twice cont. diff &

$$\phi_X''(0) = -\mathbb{E}(X^2)$$

Factorization property:

- X_1, X_2 are independent if and only if

$$\mathbb{E}\left(e^{it_1 X_1 + it_2 X_2}\right) = \phi_{X_1}(t_1) \cdot \phi_{X_2}(t_2) \quad \forall t_1, t_2 \in \mathbb{R}$$

- In particular, if X_1, X_2 are independent, then

$$\underbrace{\phi_{X_1+X_2}}_{\text{distribution} = F_{X_1} * F_{X_2}}(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t) \quad \forall t \in \mathbb{R}$$