APA lecture 2

Probability measures

(Ω = fundamental set , $\mathcal{F}$ = σ-field) $\leftarrow$ measurable space

<u>Def</u>: A <u>probability measure</u> on $(\Omega, \mathcal{F})$ is a mapping

$\mathbb{P} : \mathcal{F} \to [0,1]$ such that

(i) $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$

(ii) finite additivity : if $A_1 .. A_n \in \mathcal{F}$ are such that $A_j \cap A_k = \emptyset$ $\forall j \neq k$

then $\mathbb{P}\left(\dot{\bigcup}_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mathbb{P}(A_j)$

(ii') σ-additivity : if $(A_n, n \geq 1) \subset \mathcal{F}$ is s.t. $A_j \cap A_k = \emptyset$ $\forall j \neq k$

then $\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n)$

<u>Terminology</u>: $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space

Properties: If $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$, then:

(ii')

(iii) If $A_1 \dots A_n \in \mathcal{F}$, then $\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) \leq \sum_{j=1}^{n} \mathbb{P}(A_j)$

(iii') If $(A_n, n \geq 1) \subset \mathcal{F}$, then $\mathbb{P}\left(\bigcup_{n\geq 1} A_n\right) \leq \sum_{n \geq 1} \mathbb{P}(A_n)$

(iv) If $A, B \in \mathcal{F}$ & $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

and $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$. Also $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

(v) If $A, B \in \mathcal{F}$, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

(vi) If $(A_n, n \geq 1) \subset \mathcal{F}$ is s.t. $A_n \subset A_{n+1}$ $\forall n \geq 1$, then

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

(vii) If $(A_n, n \geq 1) \subset \mathcal{F}$ is s.t. $A_n \supset A_{n+1}$ $\forall n \geq 1$, then

$$\mathbb{P}\left(\bigcap_{n \geq 1} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

<u>Example:</u> $\Omega = \{1..6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$

- $\mathbb{P}_1(\{i\}) = \frac{1}{6}$, $\mathbb{P}_1(A) = \sum_{i \in A} \mathbb{P}_1(\{i\}) = \frac{\# A}{6}$ $\quad A \in \mathcal{F}$

  ex: $\mathbb{P}_1(\{1,3,5\}) = \sum_{i \in \{1,3,5\}} \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$

- $\mathbb{P}_2(\{i\}) = 0$ if $i \in \{1..5\}$ & $\mathbb{P}_2(\{6\}) = 1$

  $$\mathbb{P}_2(A) = \sum_{i \in A} \mathbb{P}_2(\{i\}) = \begin{cases} 1 & \text{if } 6 \in A \\ 0 & \text{otherwise} \end{cases}$$

$$\Omega = \{1..6\}, \quad \mathcal{F} = \{\underset{\substack{\uparrow \\ 0}}{\emptyset}, \underset{\substack{\uparrow \\ p}}{\{1\}}, \underset{\substack{\uparrow \\ 1-p}}{\{2..6\}}, \underset{\substack{\uparrow \\ 1}}{\Omega}\}$$

$\longleftarrow$ possible $\mathbb{P}$

<u>Example:</u> $\Omega = [0,1]$

- $\mathcal{F} = \sigma(I_1, \dots, I_n) \qquad \bigcup_{j=1}^{n} I_j = [0,1],\ I_j \cap I_k = \emptyset$
- $\mathcal{F} = \mathcal{B}([0,1]) = \sigma(\{]a,b[,\ 0 \le a < b \le 1\}, \{0\}, \{1\})$

$\mathbb{P}(]a,b[) = b-a$ length of the interval $]a,b[$

Caratheodory's extension theorem: There exists a unique extension of $\mathbb{P}$ on $\mathcal{B}([0,1])$, that is:

- $\mathbb{P}(]a,b[) = b-a \quad \forall 0 \le a < b \le 1$
- $\mathbb{P}$ is a probability measure on $\mathcal{B}([0,1])$

$\mathbb{P}$ = Lebesgue measure on $[0,1]$ (= uniform measure on $[0,1]$)

Lebesgue measure on $\mathbb{R}$ and $\mathbb{R}^2$

- on $\mathbb{R}$: $\mathcal{B}(\mathbb{R}) = \sigma\left( ]a,b[,\ a<b \right)$

  $m(]a,b[) = b-a$ Lebesgue measure of $]a,b[$

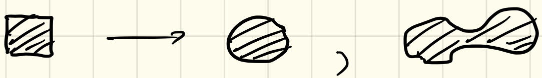
  Caratheodory's extension theorem: $\exists$ a unique extension of m on $\mathcal{B}(\mathbb{R})$.

  <u>Notation:</u> $m(B) = |B| \in [0,+\infty]$

- on $\mathbb{R}^2$: $\mathcal{B}(\mathbb{R}^2) = \sigma\left( ]a,b[\times]c,d[\ :\ a<b,\ c<d \right)$

  $m\left( ]a,b[\times]c,d[\right) = (b-a)\cdot(d-c)$

  Caratheodory $\Longrightarrow$ $\exists$ a unique extension of m on $\mathcal{B}(\mathbb{R}^2)$



<u>Notation:</u> $m(B) = |B|$

Terminology:

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space

An event $A \in \mathcal{F}$ is said to be negligible if $\mathbb{P}(A) = 0$,

resp. almost sure if $\mathbb{P}(A) = 1$.

Ex: $\cdot\ \Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$, $\mathbb{P}$ = Lebesgue measure

$\Rightarrow$ every singleton $\{x\}$ is negligible: (& $[0,1] \setminus \{x\}$ is almost sure)

$$\{x\} = \bigcap_{n \geq 1} \left]x - \frac{1}{n}, x + \frac{1}{n}\right[, \text{ so } \mathbb{P}(\{x\}) = \lim_{n \to \infty} \frac{1}{2n} = 0$$

$\cdot\ \Omega = \{1..6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}_2(A) = \begin{cases} 1 & \text{if } 6 \in A \\ 0 & \text{otherwise} \end{cases}$

$\Rightarrow$ $A = \{1..5\}$ is negligible: $\mathbb{P}_2(A) = 0$.

& $A = \{6\}$ is almost sure: $\mathbb{P}_2(A) = 1$

Proposition

- Let $(A_n, n \geq 1)$ be a collection of negligible sets. Then $\bigcup_{n\geq 1} A_n$ is also negligible.
- Let $(B_n, n \geq 1)$ be a collection of almost-sure sets. Then $\bigcap_{n \geq 1} B_n$ is also almost sure.

Proof:
- $\mathbb{P}\left(\bigcup_{n \geq 1} A_n \right) \leq \sum_{n \geq 1} \underbrace{\mathbb{P}(A_n)}_{=0} = 0$
- Take $A_n = B_n^c$

#

Consequence: $A = \mathbb{Q} \cap [0,1]$ is negligible w.r.t. the Lebesgue measure.

Distribution of a random variable

Def: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space

$X : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable random variable

The distribution of X is the mapping $\mu_X : \mathcal{B}(\mathbb{R}) \to [0,1]$

defined as:

$$\mu_X(B) = \mathbb{P}(\underbrace{\{\omega \in \Omega : X(\omega) \in B\}}_{=\{X \in B\} = X^{-1}(B) \in \mathcal{F}}) \qquad B \in \mathcal{B}(\mathbb{R})$$

Note: $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ is another probability space.

Cumulative distribution function (cdf)

Def: The cdf of a random variable X is the mapping $F_X : \mathbb{R} \to [0,1]$ defined as

$$F_X(t) = \mathbb{P}(\{X \leq t\}) \qquad t \in \mathbb{R}$$
$$= \mu_X(]-\infty, t])$$

Fact: The knowledge of μ_X is equivalent to the knowledge of F_X.

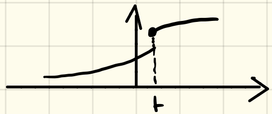
e.g: $\mu_X(]a,b]) = F_X(b) - F_X(a)$

Properties:

(i) $\lim_{t \to -\infty} F_X(t) = 0$, $\lim_{t \to +\infty} F_X(t) = 1$

(ii) F_X is non-decreasing: if $a \leq b$, then $F_X(a) \leq F_X(b)$

(iii) F_X is right-continuous: $\lim_{\varepsilon \downarrow 0} F_X(t+\varepsilon) = F_X(t)$ $\forall t \in \mathbb{R}$



Fact:

Any function $F : \mathbb{R} \to [0,1]$ satisfying properties (i), (ii) and (iii) is the cdf of a random variable X.

Rmks:

- F_X is not necessarily left-continuous
- F_X has at most a countable number of jumps

1. Discrete random variables

Def: X is a discrete random variable if there exists a finite or countable set D such that $X(\omega) \in D \quad \forall \omega \in \Omega$.

Def: for $x \in D$, $p_x = \mathbb{P}(\{X = x\})$

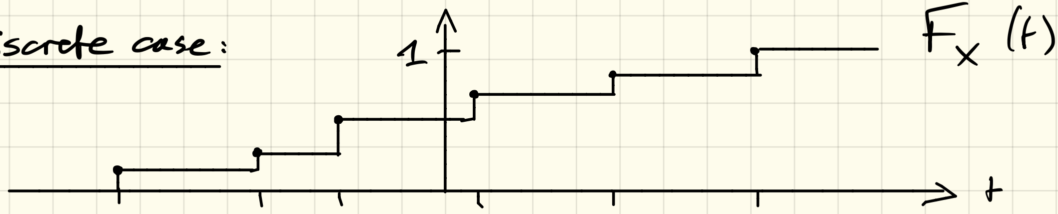
The sequence of numbers $(p_x, x \in D)$ is called the probability mass function (pmf) of X

Properties: $0 \leq p_x \leq 1 \quad \forall x \in D$, $\sum_{x \in D} p_x = 1$

$$\mu_X(B) = \sum_{x \in D \cap B} p_x \qquad \forall B \in \mathcal{B}(\mathbb{R})$$

$$F_X(t) = \sum_{x \in D,\, x \leq t} p_x \qquad \forall t \in \mathbb{R}$$

Discrete case:



2. Continuous random variables

Def: $\begin{cases} X \text{ is continuous if } \mathbb{P}(\{X \in B\}) = 0 \\ \forall B \in \mathcal{B}(\mathbb{R}) \text{ such that } \underbrace{|B|}_{=\text{Lebesgue measure of } B} = 0 \end{cases}$

In particular, $\mathbb{P}(\{X = x\}) = 0 \quad \forall x \in \mathbb{R}$

Theorem: (Radon-Nikodym)

In this case, there exists a Borel-measurable function $p_X : \mathbb{R} \to \mathbb{R}_+$ such that $\mu_X(B) = \int_B p_X(x)\, dx \quad \forall B \in \mathcal{B}(\mathbb{R})$.

Terminology: p_X is the probability density function (pdf) of X

Properties:
- $p_X(x) \geq 0 \quad \forall x \in \mathbb{R}$ (but may be > 1)
- $\int_{\mathbb{R}} p_X(x)\, dx = \mu_X(\mathbb{R}) = \mathbb{P}(\{X \in \mathbb{R}\}) = 1$

$$F_X(t) = \int_{-\infty}^{t} p_X(x)\, dx \qquad t \in \mathbb{R}$$

So F_X is continuous & "differentiable": $F_X'(t) = p_X(t)$

⚠ $p_X(x) \neq \mathbb{P}(\{X = x\})$ (which is equal to 0 $\forall x \in \mathbb{R}$)

Change of variable formula

Let X be a random variable with cdf F_X

$f: \mathbb{R} \to \mathbb{R}$ increasing (i.e. if $x_1 < x_2$, then $f(x_1) < f(x_2)$)

$Y = f(X)$. What is F_Y ?

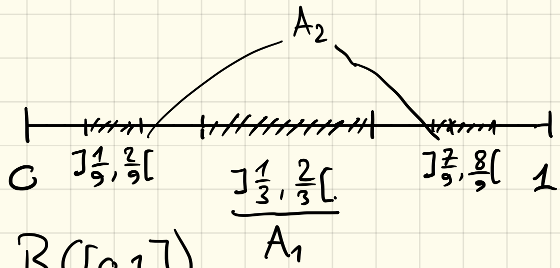
$$F_Y(t) = \mathbb{P}(\{Y \le t\}) = \mathbb{P}(\{f(X) \le t\}) = \mathbb{P}(\{X \le f^{-1}(t)\})$$
$$= F_X(f^{-1}(t)) \quad t \in \mathbb{R}$$

If in addition X is a continuous r.v. and f is differentiable:

$$p_Y(t) = \frac{d}{dt} F_Y(t) = \frac{d}{dt} F_X(f^{-1}(t)) = p_X(f^{-1}(t)) \cdot \frac{1}{\underbrace{f'(f^{-1}(t))}} \quad t \in \mathbb{R}$$

A strange cdf

Intro: the Cantor set



$\rightarrow \quad C = [0,1] \setminus \left(\bigcup_{n \geq 1} A_n \right) \in \mathcal{B}([0,1])$

What is $|C|$? $\quad |C| = 1 - \sum_{n \geq 1} |A_n|$

$|A_1| = \frac{1}{3}, \quad |A_2| = \frac{2}{9}, \quad |A_3| = \frac{4}{27} \quad \cdots \quad |A_n| = \frac{2^{n-1}}{3^n}$

$$|C| = 1 - \frac{1}{2} \sum_{n \geq 1} \left(\frac{2}{3}\right)^n = 1 - \frac{1}{2}\left(\frac{1}{1-\frac{2}{3}} - 1\right) = 1 - \frac{1}{2} \cdot 2 = 0$$

$$\begin{cases} x \in \underline{C} \text{ iff } x = \sum_{n \geq 1} \frac{a_n}{3^n} \text{ where } a_n \in \{0,2\} \\ x \in \underline{[0,1]} \text{ iff } x = \sum_{n \geq 1} \frac{b_n}{2^n} \text{ where } b_n \in \{0,1\} \end{cases} \Rightarrow C \text{ is } \underline{\text{uncountable}}$$

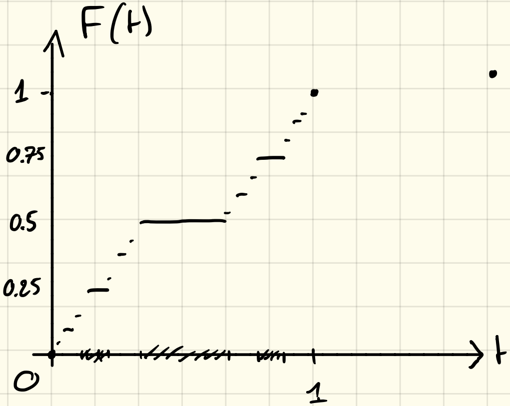
The devil's staircase

Question: Does there exist a function $F : [0,1] \to [0,1]$

such that : • $F(0) = 0$ & $F(1) = 1$ ✓

• F is continuous ✓

• $\left| \{ t \in [0,1] : F'(t) = 0 \} \right| = 1$? **Yes!** ✓



$$F'(t) = 0 \quad \forall t \in \bigcup_{n \geq 1} A_n$$

$F(t)$ is a cdf!

• X is discrete? no (F continuous)   • X is continuous? no ($p = F' = 0$ almost everywhere)