= exercise.

σ-algebra / σ-field

Let $\Omega \neq \emptyset$ be a set and let 2^{Ω} denote the power set of Ω.
Then $\mathcal{F} \subset 2^{\Omega}$ is called a σ-algebra if it satisfies the following:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A_n \in \mathcal{F}, n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ ⟸ ⚠ countable union

Trivial implications: • $A_n \in \mathcal{F}, n \in \mathbb{N} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ (use De Morgan laws)

Measurable spaces

$(\Omega, \mathcal{F})$ with $\Omega \neq \emptyset$ and $\mathcal{F}$ σ-algebra on Ω is called a measurable space.

Measurable functions (after measures & meas spaces)

Given two measurable spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ a function $f: \Omega \to \Omega'$ is said to be measurable if

$$f^{-1}(B) := \{\omega \in \Omega : f(\omega) \in B\} \in \mathcal{F}$$

for every $B \in \mathcal{F}'$.

HW: show that $1\!\!1_A : \Omega \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R})$-measurable iff $A \in \mathcal{F}$.

$$1\!\!1_A^{-1}(B) = \{\omega \in \Omega : 1\!\!1_A(\omega) \in B\} = \begin{cases} A & \text{if } A \subset B \\ & \end{cases}$$

In particular, a random variable is a measurable function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ where $\mathcal{B}(\mathbb{R})$ is the smallest sigma algebra containing every open set i.e

$$\mathcal{B}(\mathbb{R}) = \sigma(\text{open sets of } \mathbb{R})$$
$$\underset{\text{Thm}}{=} \sigma((-\infty, a], a \in \mathbb{R})$$

Measures on a measurable space

A measure μ on a measurable space $(\Omega, \mathcal{F})$ is a mapping $\mu : \mathcal{F} \mapsto [0,\infty]$ satisfying:

(i) $\mu(\emptyset) = 0$

(ii) σ-additivity ie if $\underset{n\in\mathbb{N}}{A_n \in \mathcal{F}}$, are disjoint then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Rem Why do we actually work with σ-algebras instead of just the power set of Ω?

It turns out that it is not always possible to assign a measure to all subsets of Ω that behaves nicely. E.g one cannot assign a measure μ to all subsets of $\mathbb{R}$ in such a way that $\mu(]a,b[) = b-a$ and that μ is translation invariant. However, if we restrict ourself to $\mathcal{B}(\mathbb{R})$, one
↳ see Banach Tarski paradox
can show the existence of such a measure μ ⇝ Lebesgue measure.

A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a measure st $\mathbb{P}(\Omega) = 1$.

HF: show

- $P(\emptyset) = 0$ $\quad (P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset))$
- $P(A^c) = 1 - P(A)$ $\quad (P(\Omega) = P(A \cup A^c) = P(A) + P(A^c))$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$ $\quad (P(B) = P(A \cup B \setminus A) = P(A) + \underbrace{P(B\setminus A)}_{\geq 0})$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Distribution of a random variable

The law of a random variable $X: \Omega \to \mathbb{R}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a measure μ_X on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\mu_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)) \quad \text{"push-forward measure".}$$

$$= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

The cumulative distr. funct of X is defined as $F_X(x) = \mathbb{P}(X \leq x) = \mu_X((-\infty, x])$.

Example 1 : $\Omega = \{1, 2, 3\}$, $\mathcal{F} = 2^\Omega$, $\mathbb{P}(\omega) = \frac{1}{3}$ $\forall \omega \in \Omega$

$$X: \Omega \longmapsto \mathbb{R}$$
$$\omega \longmapsto 1$$

1) compute the law of X $\quad P(X \in A) = \begin{cases} 1 & \text{if } 1 \in A \\ 0 & \text{else} \end{cases}$
2) $\sigma(X) = ?$
3) Characterise all random variables that are measurable wrt $\{\emptyset, \Omega\}$.

Example 2 $\Omega = (0,1)$ $\quad \mathbb{P} = \mathcal{L}$ $\quad \mathcal{F} = \mathcal{B}((0,1))$

$$X: \Omega \to \mathbb{R}$$
$$\omega \longmapsto \frac{1}{\lambda} \log\left(\frac{1}{1-\omega}\right) \quad \text{for some fixed } \lambda > 0.$$

Compute the cdf of X and identify its distribution.

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\left\{\omega : \frac{1}{\lambda} \log\left(\frac{1}{1-\omega}\right) \leq x\right\}\right) = \mathbb{P}\left(\left\{\omega : \omega \leq 1 - e^{-\lambda x}\right\}\right)$$

$$= \mathcal{L}(\{\omega : \omega \leq 1 - e^{-\lambda x}\})$$

$$= 1 - e^{-\lambda x} \quad \to \quad \mathcal{E}xp(1).$$

σ-algebra generated by a random variable

Let $\Omega \neq \emptyset$ and $\mathcal{A} \subset 2^{\Omega}$. The sigma algebra generated by $\mathcal{A}$ is the smallest σ-algebra $\sigma(\mathcal{A})$ containing $\mathcal{A}$.

i.e $$\sigma(A) = \bigcap_{\substack{\mathcal{A} \subseteq B \\ B \text{ sigma algebra}}} B$$

$$= \{ C \subseteq \Omega \mid C \in B \text{ for any sigma alg. } B \text{ on } \Omega \text{ st } \mathcal{A} \subseteq B \}.$$

The sigma algebra generated by a $X : \Omega \to \mathbb{R}$ is defined

as $\sigma(X) := \sigma\left(\{ X^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}) \} \right)$

It is thus the smallest σ-algebra that makes X measurable.

The sigma algebra generated by a collection of random variables $(X_i)_{i \in I}$ is that makes all mappings measurable

i.e

$$\sigma(X_i, i \in I) = \sigma\left(\bigcup_{i \in I} \sigma(X_i) \right)$$

Almost surely

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that an event $B \in \mathcal{F}$ happens a.s if $\mathbb{P}(A) = 1$.

$\boxed{\Sigma}$

E.g $X = Y$ $\mathbb{P}$-as means $\mathbb{P}(\{\omega \in \Omega : X(\omega) = Y(\omega)\}) = 1$

HF: Construct two rv X & Y st $X = Y$ $\mathbb{P}$-as but $X \neq Y$ pointwise

i.e $\mathbb{P}(X = Y) = 1$ but $\exists \omega \in \Omega$ st $X(\omega) \neq Y(\omega)$.

Expectation

The expectation of a rv $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with $\int |X(\omega)| \, d\mathbb{P}(\omega) < \infty$ is defined as

$$\mathbb{E}[X] = \int_\Omega X(\omega) \, d\mathbb{P}(\omega).$$

Properties (good exercise)

- linearity
- Jensen's inequality : $\varphi : \mathbb{R} \to \mathbb{R}$ convex, $X, \varphi(X) \in L^1 \Rightarrow \varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$
- Monotone CV thm : $X_n \geq 0$, $X_{n+1} \geq X_n$, $X_n \to X$ a.s Then

  $$\lim_{n \to \infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

- Dominated CV thm : $\left.\begin{array}{l} X_n \to X \ \ \mathbb{P}\text{.as} \\ \exists Y \in L^1 \text{ st } |X_n| \leq Y \ \forall n \end{array}\right\} \Rightarrow \lim_{n \to \infty} \mathbb{E}[X_n] = \mathbb{E}[X].$

- Fatou's Lemma

  Let $(X_n)_n$ be a sequence of non-negative random variables. Then

  $$\mathbb{E}[\liminf_{n \to \infty} X_n] \leq \liminf_{n \to \infty} \mathbb{E}[X_n]$$

  $\hookrightarrow$ good exercise prove this using MC thm applied to the sequence $(\inf_{m \geq n} X_m)_n$

<u>Thm</u>: Let $g : \mathbb{R} \to \mathbb{R}$ be a function st $g(X) \in L^1$.

- If X is a discrete rv with probability mass function p_X taking values in a countable set S then

$$\mathbb{E}[g(X)] = \sum_{t \in S} g(t)\, p_X(t)$$

- If X is an absolutely continuous r.v (with respect to Lebesgue measure) with density function f_X then

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx.$$

Modes of convergence

Let $X_1, X_2, \dots$ and X be random variables.

- $X_n \to X$ $\mathbb{P}$-as if $\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega)\}) = 1$
- $X_n \xrightarrow{P} X$ if $\forall \varepsilon > 0$ $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$
- $X_n \xrightarrow{L^p} X$ if $\mathbb{E}|X|^p < \infty$ and $\mathbb{E}|X_n - X|^p \to 0$
- $X_n \xrightarrow{\mathcal{L}} X$ if $F_{X_n}(t) \to F_X(t)$ for all $t \in \mathbb{R}$ of continuity points of $F_X(\cdot)$.

<u>Thm</u>: The following implications hold.

$$\left.\begin{array}{l} X_n \to X \quad \mathbb{P}\text{-as} \\ \quad \text{or} \\ X_n \to X \text{ in } L^p \ (p \geq 1) \end{array}\right\} \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{L}} X$$

$$X_n \xrightarrow{P} X \Downarrow \exists \text{ subsequence } (n_k)_k \text{ st } X_{n_k} \to X \text{ a.s.}$$

Conditional expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and fix a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$.
The conditional expectation of X given $\mathcal{G}$ is the $\mathbb{P}$-a.s unique integrable rv Y satisfying:

(i) Y is $\mathcal{G}$-measurable

(ii) $\mathbb{E}[X \cdot \mathbb{1}_A] = \mathbb{E}[Y \cdot \mathbb{1}_A] \quad \forall A \in \mathcal{G}$.

⚠ $\mathbb{E}[X|\mathcal{G}]$ is a r.v.

It is the "best" approximation of X given some partial information $\mathcal{G}$.
in the sense that

$$\mathbb{E}[X|\mathcal{G}] := \underset{Z \ \mathcal{G}\text{-meas}}{\operatorname{argmin}} \ \mathbb{E}[(X-Z)^2]$$

see exercise 6.

Properties (good exercise)

- linearity
- $\mathbb{E}[X|\mathcal{G}] = X$ if X is $\mathcal{G}$-meas.
- $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$ if $\mathcal{H} \subseteq \mathcal{G}$.
- Jensen's inequality
- Monotone CV thm
- Dominated CV thm.

+ see exercise sheet 1.