Albert-Ludwigs-Universität Freiburg

Lecture notes

Probability Theory I

Angelika Rohde

Summer semester 2025

Contents

1 A short repetition of Measure Theory

This chapter contains a brief summary (without proofs) of the measure theoretic foundations as already known from our Analysis III course (BSc Mathematics) or the bridging course on measure theory (MSc Data Science).

1.1 σ-Algebras and measures

Definition 1.1. *Let $\Omega \neq \varnothing$. A system $\mathscr{A}$ of subsets of Ω is called $\underline{\sigma\text{-algebra}}$ if*

(i) $\Omega \in \mathscr{A}$,

(ii) $A \in \mathscr{A} \implies A^c \in \mathscr{A}$,

(iii) $A_n \in \mathscr{A}\ \forall\, n \in \mathbb{N} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

$\mathscr{A}$ is called algebra if only (iii') instead of (iii) is granted:

(iii')

$$A, B \in \mathscr{A} \implies A \cup B \in \mathscr{A}.$$

Lemma 1.2. *Let $\Omega \neq \varnothing$ be a set and $\mathscr{A}$ a σ-algebra over Ω. Then the following statements hold true:*

(i) $A_n \in \mathscr{A}\ \forall\, n \in \mathbb{N} \implies \bigcap_{n\in\mathbb{N}} A_n \in \mathscr{A}$

(ii) $A, B \in \mathscr{A} \implies A \setminus B \in \mathscr{A}$.

Remark. *The pair $(\Omega, \mathscr{A})$ with a σ-algebra $\mathscr{A}$ over Ω is is called measurable space.*

Arbitrary intersections of σ-algebras over Ω are σ-algebras over Ω again. For any systems $\mathscr{E}$ of subsets of Ω, there exists a smallest σ-algebra $\sigma(\mathscr{E})$ with $\mathscr{E} \subset \sigma(\mathscr{E})$, namely the intersection of all those which contain $\mathscr{E}$.

Definition 1.3. *If Ω is equipped with a topology, then the σ-algebra generated by the open subsets of Ω is called Borel-σ-algebra.*

Definition 1.4. *Let $(\Omega, \mathscr{A})$ be a measurable space. A map $\mu : \mathscr{A} \to [0, \infty]$ with $\mu(\varnothing) = 0$ is called measure if it is σ-additive, meaning that*

$$A_n \in \mathscr{A}\ \forall n \in \mathbb{N},\ A_i \cap A_j = \varnothing\ \forall i \neq j \implies \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

The triple $(\Omega, \mathscr{A}, \mu)$ is called measure space. If $\mu(\Omega) = 1$, then μ is named probability measure and $(\Omega, \mathscr{A}, \mu)$ correspondingly probability space. A measure μ is called finite if $\mu(\Omega) < \infty$. It is named σ-finite if there exist sets $\Omega_i \in \mathscr{A}$ $\forall i \in \mathbb{N}$ such that $\mu(\Omega_i) < \infty$ $\forall i \in \mathbb{N}$ and $\Omega = \bigcup_{i\in\mathbb{N}} \Omega_i$.

Notation. *The following notations are occasionally used variants of "$A \cup B$", but also convey the information that the sets A and B are disjoint:*

$$A \,\dot{\cup}\, B, \ A + B.$$

Lemma 1.5. *Let $(\Omega, \mathscr{A})$ be a measurable space, μ a measure on $(\Omega, \mathscr{A})$, $A, B, A_n \in \mathscr{A}$ $\forall\, n \in N$. Then the following statements are satisfied:*

(i) $A \subset B \implies \mu(A) + \mu(B \setminus A) = \mu(B)$. *In particular,* $A \subset B \implies \mu(A) \le \mu(B)$ *(monotonicity).*

(ii) $\mu(A \cap B) + \mu(A \cup B) = \mu(A) + \mu(B)$. *In particular,* $\mu(A \cup B) \le \mu(A) + \mu(B)$ *(subadditivity).*

(iii) If $A_n \subset A_{n+1}$ $\forall\, n \in \mathbb{N}$, *then the continuity from below holds:*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

(iv) If $A_{n+1} \subset A_n$ $\forall n \in \mathbb{N}$ *and* $\mu(A_1) < \infty$, *then the continuity from above holds:*

$$\mu\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

We now pose the question under which conditions an additive functional (uniquely) extends to a measure.

Theorem 1.6 (Carathéodory's existence and uniqueness theorem)**.** *Let $\mathscr{A}$ be a system of subsets of Ω with the following properties:*

(i) $\Omega \in \mathscr{A}$,

(ii) $A, B \in \mathscr{A} \implies B \setminus A$ *is a finite disjoint union of sets of* $\mathscr{A}$,

(iii) $A, B \in \mathscr{A} \implies A \cap B \in \mathscr{A}$.

[Such a system of subsets of Ω is called half ring.] Let $\mu : \mathscr{A} \to [0, \infty]$ be a function with the following properties:

(iv) $\mu(\varnothing) = 0$,

(v) $A, B \in \mathscr{A}$ *with* $A \cup B \in \mathscr{A}$ *and* $A \cap B = \varnothing \implies \mu(A \cup B) = \mu(A) + \mu(B)$ *(additivity),*

(vi) There exist $\Omega_i \in \mathscr{A}$ *with* $\mu(\Omega_i) < \infty$ *and* $\Omega_i \subset \Omega_{i+1}$ $\forall\, i \in \mathbb{N}$ *such that* $\bigcup_{i \in \mathbb{N}} \Omega_i = \Omega$.

(vii) $A, A_n \in \mathscr{A}\ \forall n \in \mathbb{N}$ *with* $A \subset \bigcup_{n\in\mathbb{N}} A_n$

$$\implies \mu(A) \leq \sum_{n\in\mathbb{N}} \mu(A_n).$$

Then there exists a unique extension of μ to a measure on $\sigma(\mathscr{A})$.

Remark. *The system of sets*

$$\{(a, b] \cap \mathbb{R} \mid a, b \in \mathbb{R} \cup \{-\infty, +\infty\}, a \leq b\}$$

satisfies the conditions (i)-(iii).

Remark. *A probability measure μ is uniquely described by its values on a $\cap$-stable generator $\mathscr{E}$ of the σ-algebra.*

Proof. As μ is normed (i.e. $\mu(\Omega) = 1$) and additive, these values provide all the values on

$$\mathscr{E}' = \left\{ \bigcap_{i=1}^{K} A_i \,\middle|\, A_i \text{ or } A_i^c \in \mathscr{E}\ \forall\, i \leq K, K \in \mathbb{N} \right\}.$$

The set $\mathscr{E}'$ satisfies (i)-(iii):

(i) We can assume that $\varnothing \in \mathscr{E}$, as the goal is to describe μ by its values on $\mathscr{E}$, and we already know that $\mu(\varnothing) = 0$. Using $K = 1$ and $A_1 = \Omega$ (which is valid, as $\Omega^c = \varnothing \in \mathscr{E}$), $\Omega \in \mathscr{E}'$ follows from the definition of $\mathscr{E}'$.

(ii) Let $A = \bigcap_{i=1}^{K} A_i, B = \bigcap_{j=1}^{L} B_j \in \mathscr{E}'$.

$$\implies B \setminus A = B \cap A^c = \left(\bigcap_{j=1}^{L} B_j\right) \cap \left(\bigcup_{i=1}^{K} A_i^c\right) = \bigcup_{i=1}^{K} \underbrace{\overbrace{\left(A_i^c \cap \bigcap_{j=1}^{L} B_j\right)}^{\text{finite intersection of sets for which the set or its complement are in } \mathscr{E}}}_{\in \mathscr{E}'}$$

(iii) Let $A = \bigcap_{i=1}^{K} A_i, B = \bigcap_{j=1}^{L} B_j \in \mathscr{E}'$. Then,

$$A \cap B = \bigcap_{i=1}^{K} \bigcap_{j=1}^{L} (A_i \cap B_j)$$

is a finite intersection of subsets of Ω such that for every set, the set itself or its complement is in $\mathscr{E}$; hence, $A \cap B \in \mathscr{E}'$.

Therefore, the claim follows from Theorem 1.6. □

1.2 Dynkin systems and the $\pi - \lambda$ theorem

Definition 1.7. *A system of sets $\mathscr{D} \subset \mathcal{P}(\Omega)$ is called* <u>*Dynkin*</u> *system over Ω if*

(i) $\Omega \in \mathscr{D}$

(ii) $A \in \mathscr{D} \implies A^c \in \mathscr{D}$

*(iii) If $A_n \in \mathscr{D}$ are **pairwise disjoint** sets $\forall\, n \in \mathbb{N}$,*

$$\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{D}.$$

Remark. *Let $\mathscr{D}$ be a Dynkin system over Ω. (i) and (ii) imply that $\emptyset \in \mathscr{D}$ and therefore, (iii) holds true in particular for finite unions of disjoint sets. If $E \subset D \subset \Omega$ with $E, D \in \mathscr{D}$, then $D^c \in \mathscr{D}$ by (ii), D^c and E are disjoint and therefore,*

$$D \setminus E = (D^c \cup E)^c = (D^c + E)^c \in \mathscr{D}.$$

As for σ-algebras, arbitrary intersections of Dynkin systems over Ω are Dynkin systems over Ω again. For any systems $\mathscr{E}$ of subsets of Ω, there exists a smallest Dynkin system $\mathscr{D}(\mathscr{E})$ with $\mathscr{E} \subset \mathscr{D}(\mathscr{E})$, that is

$$\mathscr{D}(\mathscr{E}) = \bigcap_{\substack{\mathscr{D} \text{ Dynkin system} \\ \mathscr{E} \subset \mathscr{D}}} \mathscr{D}.$$

Proposition 1.8. *Let $\mathscr{E} \subset \mathcal{P}(\Omega)$ be stable under intersection. Then $\mathscr{D}(\mathscr{E}) = \sigma(\mathscr{E})$.*

Proof. Apparently, every σ-algebra is a Dynkin system, so $\mathscr{D}(\mathscr{E}) \subset \sigma(\mathscr{E})$. For the other direction, $\sigma(\mathscr{E}) \subset \mathscr{D}(\mathscr{E})$, it is sufficient to prove that $\mathscr{D}(\mathscr{E})$ is a σ-algebra, that is, it must be stable under intersection (because then, for $A_1, A_2 \in \mathscr{D}$, we have $A_1 \cup A_2 = A_1 + (A_2 \setminus (A_1 \cap A_2)) \in \mathscr{D}$). Let $D \in \mathscr{D}(\mathscr{E})$ be fixed and consider

$$\widetilde{D} := \{A \subset \Omega \mid A \cap D \in \mathscr{D}(\mathscr{E})\}.$$

We want to show that $\mathscr{D}(\mathscr{E})$ is contained in $\widetilde{D}$. We observe that $\widetilde{D}$ is itself a Dynkin system:

(a) $\Omega \in \widetilde{D}$ because $\Omega \cap D = D \in \mathscr{D}(\mathscr{E})$.

(b) $A \in \widetilde{D} \implies A^c \cap D = D \setminus (A \cap D) \in \mathscr{D}(\mathscr{E})$ by the previous remark, as $A \cap D \subset D$, so $A^c \in \widetilde{\mathscr{D}}$.

(c) Let $A_n \in \widetilde{D}$ be pairwise disjoint sets. $A_n \cap D \in \mathscr{D}(\mathscr{E})$ are pairwise disjoint sets for $n \in \mathbb{N}$, so

$$\left(\bigcup_{n \in \mathbb{N}} A_n \right) \cap D = \bigcup_{n \in \mathbb{N}} (A_n \cap D) \in \mathscr{D}(\mathscr{E}),$$

i.e. $\bigcup_{n\in\mathbb{N}} A_n \in \widetilde{D}$.

For all $E \in \mathscr{E}$, we have $\mathscr{E} \subset \widetilde{E}$, because by requirement, $\mathscr{E}$ is stable under intersection. Hence, $\mathscr{D}(\mathscr{E}) \subset \widetilde{E}$, as $\mathscr{D}(\mathscr{E})$ is the *smallest* Dynkin system which contains $\mathscr{E}$, and $\widetilde{E}$ is *some* Dynkin system which contains $\mathscr{E}$. In particular, $D \cap E \in \widetilde{E}$ for the fixed D from above, so $E \in \widetilde{D}$. As $E \in \mathscr{E}$ was arbitrary, this implies $\mathscr{E} \subset \widetilde{D}$, so $\mathscr{D}(\mathscr{E}) \subset \widetilde{D}$ (because $\widetilde{D}$ is a Dynkin system) for every $D \in \mathscr{D}(\mathscr{E})$. Therefore, $\mathscr{D}(\mathscr{E})$ is stable under intersection. □

Often, an $\cap$-stable system of sets is called π-system, whereas a Dynkin system is also referred to as λ-system.

Corollary 1.9 (Dynkin's $\pi - \lambda$ theorem)**.** *If $\mathscr{E}$ is a π-system and $\mathscr{L}$ is a λ-system that contains $\mathscr{E}$, then $\sigma(\mathscr{E}) \subset \mathscr{L}$.*

1.3 Measurable maps and measure integral

Definition 1.10. *Let $(\Omega, \mathscr{A}), (\Omega', \mathscr{A}')$ be measurable spaces. A map $f : \Omega \to \Omega'$ is called measurable (or $\mathscr{A}$-$\mathscr{A}'$-measurable) if*

$$f^{-1}(A') \in \mathscr{A} \ \forall A' \in \mathscr{A}'$$

where $f^{-1}(A') := \{\omega \in \Omega \mid f(\omega) \in A'\}$.

Lemma 1.11. *Let $(\Omega, \mathscr{A}), (\Omega', \mathscr{A}')$ be measurable spaces and $\mathscr{E} \subset \mathscr{A}'$ with $\sigma(\mathscr{E}) = \mathscr{A}'$ (meaning $\mathscr{E}$ is a generator of $\mathscr{A}'$). Then:*

$$f^{-1}(E') \in \mathscr{A} \ \forall E' \in \mathscr{E}' \iff f^{-1}(A') \in \mathscr{A} \ \forall A' \in \mathscr{A}'$$

This means that it is sufficient to check measurability on a generator of the σ-algebra.

Notation. *For the sake of readability, we define $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and*

$$\mathscr{B}(\overline{\mathbb{R}}) = \{B \subset \mathbb{R} \mid B \cap \mathbb{R} \in \mathscr{B}(\mathbb{R})\} = \{B \cup E \mid B \in \mathscr{B}(\mathbb{R}), E \subset \{-\infty, +\infty\}\}.$$

Proposition 1.12. *Let $(\Omega, \mathscr{A})$ be a measurable space and $f_n : \Omega \to \overline{\mathbb{R}}$ an $\mathscr{A}$-$\mathscr{B}(\overline{\mathbb{R}})$-measurable function $\forall n \in \mathbb{N}$. Then, the functions*

$$\inf_{n\in\mathbb{N}} f_n, \ \sup_{n\in\mathbb{N}} f_n, \ \liminf_{n\to\infty} f_n, \ \limsup_{n\to\infty} f_n$$

are also $\mathscr{A}$-$\mathscr{B}(\overline{\mathbb{R}})$-measurable.

Monotone limits play a crucial role for the construction of the measure integral.

Definition 1.13. *Let $A \subset \Omega$. The function*

$$\mathbb{1}_A : \Omega \to \{0,1\},\ \omega \mapsto \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

is called <u>*indicator function*</u>*. If $A_k \in \mathscr{A}$ and $c_k \in \mathbb{R}\ \forall\, k \in \{1,\dots,n\}$ for some $n \in \mathbb{N}$, then $\sum_{k=1}^{n} c_k \mathbb{1}_{A_k}$ is called* <u>*$\mathscr{A}$-elementary function*</u>*.*

Measure integral for elementary functions

Let f be an elementary function, i.e. $f = \sum_{k=1}^{n} c_k \mathbb{1}_{A_k}$ with measurable sets $A_k \subset \Omega$. Then the elementary integral is defined as

$$\int f \ \mathrm{d}\mu := \sum_{k=1}^{n} c_k \mu(A_k).$$

If there is an alternative representation $f = \sum_{k=1}^{n'} c'_k \mathbb{1}_{A'_k}$, one can easily prove that

$$\sum_{k=1}^{n} c_k \mu(A_k) = \sum_{k=1}^{n'} c'_k \mu(A'_k).$$

Hence, this integral is well-defined.

Lemma 1.14. *Let $(\Omega, \mathscr{A})$ be a measurable space and $f : \Omega \to \mathbb{R}$ an $\mathscr{A}$-$\mathscr{B}(\mathbb{R})$-measurable function with $f \geq 0$. Then the following statements hold true.*

(i) There exists a sequence $(f_n)_{n\in\mathbb{N}}$ of elementary functions with $f_n \leq f_{n+1}\ \forall n \in \mathbb{N}$ and $\lim_{n\to\infty} f_n = f$.

(ii) There exist sets $A_k \in \mathscr{A}\ \forall\, n \in \mathbb{N}$ and a sequence $(c_k)_{k\in\mathbb{N}}$ of non-negative real numbers such that $f = \sum_{k\in\mathbb{N}} c_k \mathbb{1}_{A_k}$.

That is: Any measurable function $f \geq 0$ is the monotone limit of elementary functions $f_n \nearrow f$.

The crucial idea is to define the integral for general non-negative, measurable functions f by monotone approximation: For non-negative elementary functions $f_n \nearrow f$,

$$\int \underset{=\lim_{n\to\infty} f_n}{\underset{\uparrow}{f}} \ \mathrm{d}\mu := \lim_{n\to\infty} \int f_n \ \mathrm{d}\mu = \sup_{n\in\mathbb{N}} \int f_n \ \mathrm{d}\mu$$

To ensure that this integral is well-defined, we have to verify that

$$\lim_{n\to\infty} \int f_n \ \mathrm{d}\mu = \lim_{n\to\infty} \int g_n \ \mathrm{d}\mu$$

if g_n is another sequence of elementary functions with $g_n \nearrow f$. This is the content of the next lemma.

Lemma 1.15. *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space and $f : \Omega \to \mathbb{R}$ a measurable, bounded function with $f \geq 0$. If*

$$f = \sum_{n=1}^{\infty} \alpha_n \mathbb{1}_{A_n} = \sum_{n=1}^{\infty} \beta_n \mathbb{1}_{B_n}$$

with constants $\alpha_n, \beta_n > 0 \ \forall n \in \mathbb{N}$, then

$$\sum_{n \in \mathbb{N}} \alpha_n \mu(A_n) = \sum_{n \in \mathbb{N}} \beta_n \mu(B_n).$$

We summarize the important findings and state formally the definition.

Definition 1.16. *Let $f \geq 0$ be a non-negative measurable function. By Lemma 1.14, f is the monotone limit of non-negative elementary functions $f_n = \sum_{k=1}^{n} c_k \mathbb{1}_{A_k}$, that is $f = \sum_{n \in \mathbb{N}} c_k \mathbb{1}_{A_k}$, and we define*

$$\int f \, \mathrm{d}\mu := \lim_{n \to \infty} \int f_n \, \mathrm{d}\mu = \sum_{n \in \mathbb{N}} c_n \mu(A_n).$$

By Lemma 1.15, this integral is well-defined.

A measurable (not necessarily non-negative) function $f : \Omega \to \mathbb{R}$ is called <u>(finitely) integrable</u> if $f^+ := f \cdot \mathbb{1}_{\{f \geq 0\}}$ and $f^- := -f \cdot \mathbb{1}_{\{f < 0\}}$ are integrable, i.e.

$$\int f^+ \, \mathrm{d}\mu < \infty \text{ and } \int f^- \, \mathrm{d}\mu < \infty.$$

This is the case if and only if $\int |f| \, \mathrm{d}\mu < \infty$ (which we call <u>absolute integrability</u>). In this case, we define

$$\int f \, \mathrm{d}\mu := \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu.$$

Lemma 1.17 (Properties of the integral)**.** (*i*) *The integral is linear on the finitely integrable functions, i.e.*

$$\int (\alpha f + \beta g) \, \mathrm{d}\mu = \alpha \int f \, \mathrm{d}\mu + \beta \int g \, \mathrm{d}\mu$$

for all finitely integrable functions f, g and $\alpha, \beta \in \mathbb{R}$.

(*ii*) *The integral is monotone, i.e.*

$$g \leq f \implies \int g \, \mathrm{d}\mu \leq \int f \, \mathrm{d}\mu$$

for all integrable functions f, g.

(iii) Theorem of monotone convergence: *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of finitely integrable or non-negative measurable functions with $f_n \leq f_{n+1}\ \forall\, n \in \mathbb{N}$, then*

$$\int \lim_{n\to\infty} f_n \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

(iv) Theorem of dominated convergence: *If $(f_n)_{n\in\mathbb{N}}$ is a sequence of measurable functions with $|f_n| \leq f\ \forall n \in \mathbb{N}$ for some function f with $\int f \, \mathrm{d}\mu < \infty$, such that the pointwise limit $g := \lim_{n\to\infty} f_n$ exists, then*

$$\int g \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

(v) Fatou's Lemma: *If f, f_n are finitely integrable and $f \leq f_n\ \forall n \in \mathbb{N}$, then the function $\liminf_{n\to\infty} f_n$ is finitely integrable, and*

$$\liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu \geq \int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu.$$

1.4 Image measure and transformation formula

Definition 1.18. *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space, $(\Omega', \mathscr{A}')$ a measurable space and $X : \Omega \to \Omega'$ a measurable map. The image measure μ^X on $(\Omega', \mathscr{A}')$ is defined as*

$$\mu^X(A') := \mu(X^{-1}(A'))\ \forall A' \in \mathscr{A}'.$$

We have seen already that μ^X is actually a measure on $(\Omega', \mathscr{A}')$.

Theorem 1.19 (Transformation formula)**.** *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space, $(\Omega', \mathscr{A}')$ a measurable space and $X : \Omega \to \Omega'$ a measurable map. Then:*

(i) *If $f' : \Omega' \to \mathbb{R}$ is $\mathscr{A}'$-$\mathscr{B}(\mathbb{R})$-measurable and non-negative, then*

$$\int f' \, \mathrm{d}\mu^X = \int f' \circ X \, \mathrm{d}\mu.$$

(ii) *If $f' : \Omega' \to \mathbb{R}$ is finitely absolutely integrable w.r.t. μ^X, then so is $f' \circ X$ and vice-versa. In this case, we also have:*

$$\int f' \, \mathrm{d}\mu^X = \int f' \circ X \, \mathrm{d}\mu.$$

1.5 Product spaces and Kolmogorov's consistency theorem

Definition 1.20. *Let I be an arbitrary index set and $(\Omega_i)_{i\in I}$ a family of sets. We define*

$$\Omega = \prod_{i\in I} \Omega_i := \left\{ \omega : I \to \bigcup_{i\in I} \Omega_i \,\middle|\, \omega(i) \in \Omega_i \ \forall i \in I \right\}.$$

Its elements ω are sometimes written as $\omega = (\omega_i)_{i\in I}$. Let $\pi_i : \Omega \to \Omega_i$, $\omega \mapsto \omega_i$ be the coordinate map (i.e. the projection onto the i-th coordinate).

Definition 1.21. *Let $(\Omega_i, \mathscr{A}_i)$ be a measurable space for every $i \in I$. The smallest σ-algebra over Ω with respect to which the function π_i is measurable $\forall\, i \in I$ is called the product-σ-algebra. This means that the product-σ-algebra is*

$$\bigotimes_{i\in I} \mathscr{A}_i := \sigma\left(\pi_i^{-1}(A_i) \mid A_i \in \mathscr{A}_i \ \forall\, i \in I\right).$$

Theorem 1.22. *Let $(\Omega_i, \mathscr{A}_i, \mu_i)$ be σ-finite measure spaces $\forall i \in \{1, \ldots, n\}$ for some $n \in \mathbb{N}$. Then there exists a unique σ-finite measure $\bigotimes_{i=1}^n \mu_i$ on the product space*

$$\left(\prod_{i=1}^n \Omega_i,\ \bigotimes_{i=1}^n \mathscr{A}_i \right)$$

with the property

$$\left(\bigotimes_{i=1}^n \mu_i \right)(A_1 \times \ldots \times A_n) = \prod_{i=1}^n \mu_i(A_i)$$

for all sets $A_i \in \mathscr{A}_i$ $\forall i \in \{1, \ldots, n\}$. We call this measure the product measure.

Example. *Let $\Omega_1 = \mathbb{R} = \Omega_2$, $\mathscr{A}_1 = \mathscr{B}(\mathbb{R}) = \mathscr{A}_2$, $\mu_1 = \lambda = \mu_2$. Then the product measure λ^2 satisfies $\lambda^2([a_1, b_1] \times [a_2, b_2]) = \lambda([a_1, b_1]) \cdot \lambda([a_2, b_2])$.*

One nice property of the integral with respect to product measures is that it can be iteratively boiled down to integrals over the marginals.

Theorem 1.23 (Fubini)**.** *Let $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ be σ-finite measure spaces and $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ a $(\mathscr{A}_1 \otimes \mathscr{A}_2)$-$\mathscr{B}(\mathbb{R})$-measurable function. If f is non-negative or absolutely integrable, then the maps*

$$\Omega_1 \to \overline{\mathbb{R}}, \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2)\ \mathrm{d}\mu_2(\omega_2) \quad \text{and} \quad \Omega_2 \to \overline{\mathbb{R}}, \omega_2 \mapsto \int_{\Omega_1} f(\omega_1, \omega_2)\ \mathrm{d}\mu_1(\omega_1)$$

are measurable, and

$$\int_\Omega f\ \mathrm{d}\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2)\ \mathrm{d}\mu_2(\omega_2) \right)\ \mathrm{d}\mu_1(\omega_1)$$

$$= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) \, \mathrm{d}\mu_1(\omega_1) \right) \, \mathrm{d}\mu_2(\omega_2)$$

Definition 1.24. *A topological space is called* <u>*Polish*</u>*, if there exists a complete metric inducing its topology and a countable base of the topology.*

Example. *A well-known example of a Polish space is the set of all continuous functions on the unit interval, $\mathcal{C}([0,1])$, equipped with the topology of uniform convergence:*

$$d_{\sup} : \mathcal{C}([0,1]) \times \mathcal{C}([0,1]) \to \mathbb{R}, \ (f,g) \mapsto ||f-g||_\infty = \sup_{x\in[0,1]} |f(x) - g(x)|.$$

Proposition 1.25 (Ulam's lemma)**.** *Let (Ω, τ) be a Polish space and $\mathscr{B}(\Omega)$ the Borel-σ-algebra on this topological space. Let μ be a probability measure on $(\Omega, \mathscr{B}(\Omega))$. Let $E \in \mathscr{B}(\Omega)$ and $\varepsilon > 0$. Then there exists a compact set $K \subset E$ with $\mu(E \setminus K) < \varepsilon$.*

Definition 1.26. *Let I be an index set, $(\Omega_i, \mathscr{A}_i)$ a measurable space for every $i \in I$ and*

$$(\Omega, \mathscr{A}) := \left(\prod_{i\in I} \Omega_i, \ \bigotimes_{i \in I} \mathscr{A}_i \right).$$

A family of probability measures $\{\mu_J \mid J \subset I \text{ is finite}\}$ on

$$(\Omega_J, \mathscr{A}_J) := \left(\prod_{i\in J} \Omega_i, \ \bigotimes_{i \in J} \mathscr{A}_i \right)$$

is called <u>*projective family*</u> *if for every finite set $J \subset I$ and every $K \subset J$, we have*

$$\mu_K = \mu_J^{\pi_K^J} \ \textit{(image measure of } \mu_J \textit{ under } \pi_K^J\textit{)},$$

where $\pi_K^J : \Omega_J \to \Omega_K$, $(\omega_i)_{i\in J} \mapsto (\omega_i)_{i \in K}$ is the coordinate projection from Ω_J to Ω_K.

Theorem 1.27 (Consistency theorem of Kolmogorov)**.** *Suppose that we have Polish spaces (Ω_n, τ_n) for every $n \in \mathbb{N}$ and a projective family $\{\mu_J \mid J \subset \mathbb{N} \text{ is finite}\}$ where μ_J is a probability measure on*

$$\left(\prod_{k\in J} \Omega_k, \ \bigotimes_{k \in J} \mathscr{B}(\tau_k) \right)$$

for all finite $J \subset \mathbb{N}$. Then, there exists a unique probability measure μ on

$$\left(\prod_{n\in \mathbb{N}} \Omega_n, \ \bigotimes_{n \in \mathbb{N}} \mathscr{B}(\tau_n) \right)$$

such that $\mu_J = \mu^{\pi_J}$ for the projections $\pi_J : \prod_{n\in\mathbb{N}} \Omega_n \to \prod_{n \in J} \Omega_k$, $(\omega_n)_{n\in\mathbb{N}} \mapsto (\omega_i)_{n\in J}$.

Definition 1.28. *Let* $(\Omega_1, \mathscr{A}_1)$ *and* $(\Omega_2, \mathscr{A}_2)$ *be measurable spaces. A map* $K : \Omega_1 \times \mathscr{A}_2 \to [0,1]$ *is called* <u>*probability kernel*</u> *if the following properties hold:*

(i) $\forall \omega_1 \in \Omega_1$*, the map* $K(\omega_1, \cdot)$ *is a probability measure on* $(\Omega_2, \mathscr{A}_2)$*.*

(ii) $\forall A_2 \in \mathscr{A}_2$*, the map* $K(\cdot, A_2)$ *is* $\mathscr{A}_1$*-*$\mathscr{B}([0,1])$*-measurable (where* $\mathscr{B}([0,1])$ *is the Borel-*σ*-algebra on* $[0,1]$*).*

Theorem 1.29. *Let* $(\Omega_1, \mathscr{A}_1, \mathbb{P}_1)$ *be a probability space,* $(\Omega_2, \mathscr{A}_2)$ *a measurable space and* $K : \Omega_1 \times \mathscr{A}_2 \to [0,1]$ *a probability kernel. Then, there exists a unique probability measure* $\mathbb{P}$ *on the product space* $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ *with*

$$\mathbb{P}(A_1 \times A_2) = \int_{A_1} K(\omega_1, A_2) \, \mathrm{d}\mathbb{P}_1(\omega_1) \quad \forall A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2.$$

Moreover, for every bounded, measurable function $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$*, we have:*

$$\int f \, \mathrm{d}\mathbb{P} = \int \underbrace{\int f(\omega_1, \omega_2) \, K(\omega_1, \mathrm{d}\omega_2)}_{\text{bounded, measurable function from } \Omega_1 \text{ to } \mathbb{R}} \, \mathrm{d}\mathbb{P}_1(\omega_1)$$

This is a generalization of Fubini's theorem.

1.6 Lebesgue decomposition and densities

Definition 1.30. *Let* $(\Omega, \mathscr{A})$ *be a measurable space and* μ, ν *measures on it.*

(i) μ *is called* <u>*(absolutely) continuous*</u> *with respect to* ν *(notation:* $\mu \ll \nu$*) if for every* $A \in \mathscr{A}$ *with* $\nu(A) = 0$*, we also have* $\mu(A) = 0$*.*

(ii) μ *and* ν *are called* <u>*equivalent*</u> *if* $\nu \ll \mu$ *and* $\mu \ll \nu$*.*

(iii) μ *and* ν *are called* <u>*singular*</u> *(notation:* $\mu \perp \nu$*) if there exists a set* $A \in \mathscr{A}$ *with* $\mu(A) = \nu(A^c) = 0$*.*

(iv) ν *has a density* f *with respect to* μ *if there exists a measurable function* f *with*

$$\nu(A) = \int f \cdot \mathbb{1}_A \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu.$$

In this case, we write:

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Proposition 1.31. *Let* ν *be a* σ*-finite measure. If* f_1 *and* f_2 *both are densities of* ν *w.r.t.* μ*, then*

$$\mu(\{f_1 \neq f_2\}) = 0.$$

Theorem 1.32 (Lebesgue-decomposition)**.** *Let μ and ν be σ-finite measures on some measurable space $(\Omega, \mathscr{A})$. Then, ν possesses a unique decomposition*

$$\nu = \nu_a + \nu_s$$

where $\nu_a \ll \mu$, $\nu_s \perp \mu$, ν_a has a density w.r.t. μ and

$$\mu\left(\frac{\mathrm{d}\nu_a}{\mathrm{d}\mu} = \infty\right) = 0.$$

Corollary 1.33 (Theorem of Radon-Nikodym)**.** *Let $(\Omega, \mathscr{A})$ be a measurable space and μ, ν σ-finite measures on it. Then,*

$$\nu \ll \mu \iff \nu \text{ has a density w.r.t. } \mu.$$

2 Stochastic independence

From now on, we always assume $(\Omega, \mathscr{A}, \mathbb{P})$ is a probability space.

Definition 2.1. *A familiy of set systems $(\mathscr{C}_i)_{i\in I}$ with $\mathscr{C}_i \subset \mathscr{A}$ $\forall i \in I$ is called stochastically independent if*

$$\mathbb{P}\left(\bigcap_{i\in S} C_i\right) = \prod_{i\in S} \mathbb{P}(C_i) \quad \forall \text{ finite } S \subset I \text{ and } C_i \in \mathscr{C}_i \ \forall i \in S.$$

Theorem 2.2. *Let $(\mathscr{C}_i)_{i\in I}$ be a stochastically independent family of set systems with $\mathscr{C}_i \subset \mathscr{A}$ $\forall i \in I$, which are stable under intersection. Then, the family $(\sigma(\mathscr{C}_i))_{i\in I}$ is also stochastically independent.*

Proof. Let $S = \{i_1, \dots, i_n\} \subset I$ be finite and $l \in \{1, \dots, n\}$ fixed. We define:

$$\mathscr{D}_l = \Bigg\{A \in \mathscr{A} \mid \mathbb{P}\Bigg(A \cap \bigcap_{j\neq l} C_{i_j}\Bigg) = \mathbb{P}(A) \cdot \prod_{j\neq l} \mathbb{P}(C_{i_j})$$
$$\text{for all } C_{i_j} \in \sigma(\mathscr{C}_{i_j}) \ \forall j < l \text{ and } C_{i_j} \in \mathscr{C}_{i_j} \ \forall j > l\Bigg\}$$

We now want to prove that $\mathscr{D}_l$ is a Dynkin system, which we will do by induction on l. Initialization step: We show that $\mathscr{D}_1$ is a Dynkin system. Let $C_{i_j} \in \mathscr{C}_{i_j}$ $\forall j > 1$.

(a) As $\mathscr{C}_{i_j}$ are independent, we have:

$$\mathbb{P}\Bigg(\Omega \cap \bigcap_{j\neq 1} C_{i_j}\Bigg) = \mathbb{P}\Bigg(\bigcap_{j\neq 1} C_{i_j}\Bigg) = \prod_{j\neq 1} \mathbb{P}(C_{i_j})$$

Therefore, $\Omega \in \mathscr{D}_1$.

(b) Let $D \in \mathscr{D}_1$. We have:

$$D^c \cap \bigcap_{j \neq 1} C_{i_j} = \Big(\Omega \cap \bigcap_{j>1} C_{i_j}\Big) \setminus \Big(D \cap \bigcap_{j>1} C_{i_j}\Big)$$

As

$$\mathbb{P}\Big(\Omega \cap \bigcap_{j>1} C_{i_j}\Big) = \mathbb{P}\Big(\bigcap_{j>1} C_{i_j}\Big) = \prod_{j>1} \mathbb{P}(C_{i_j}),$$

$$\mathbb{P}\Big(D \cap \bigcap_{j>1} C_{i_j}\Big) = \mathbb{P}(D) \cdot \prod_{j>1} \mathbb{P}(C_{i_j}),$$

and $D \cap \bigcap_{j>1} C_{i_j} \subset \Omega \cap \bigcap_{j>1} C_{i_j}$, we obtain:

$$\mathbb{P}\Big(D^c \cap \bigcap_{j>1} C_{i_j}\Big) = (1 - \mathbb{P}(D)) \cdot \prod_{j>1} \mathbb{P}(C_{i_j}) = \mathbb{P}(D^c) \cdot \prod_{j>1} \mathbb{P}(C_{i_j}).$$

Therefore, $D^c \in \mathscr{D}_1$.

(c) Let $A_n \in \mathscr{D}_1$ be pairwise disjoint sets $\forall n \in \mathbb{N}$. We have:

$$\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) \cap \Big(\bigcap_{j>1} C_{i_j}\Big) = \bigcup_{n \in \mathbb{N}} \underbrace{\Big(A_n \cap \bigcap_{j>1} C_{i_j}\Big)}_{\text{pairwise disjoint}}$$

$$\begin{aligned} \implies \mathbb{P}\Big(\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) \cap \Big(\bigcap_{j>1} C_{i_j}\Big)\Big) &= \sum_{n \in \mathbb{N}} \mathbb{P}\Big(A_n \cap \bigcap_{j>1} C_{i_j}\Big) && \text{by } \sigma\text{-additivity} \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(A_n) \cdot \prod_{j>1} \mathbb{P}(C_{i_j}) \\ &= \mathbb{P}\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) \cdot \prod_{j>1} \mathbb{P}(C_{i_j}) && \text{by } \sigma\text{-additivity} \end{aligned}$$

Therefore, $\bigcup_{n \in \mathbb{N}} A_n \in \mathscr{D}_1$.

As we have shown that $\mathscr{D}_1$ is a Dynkin system, and we know that it contains $\mathscr{C}_{i_1}$, we know that $\sigma(\mathscr{C}_{i_1}) = \mathscr{D}(\mathscr{C}_{i_1}) \subset \mathscr{D}_1$. The induction step follows analogously. □

Definition 2.3. *Let I be an arbitrary index set, $(\Omega, \mathscr{A}, \mathbb{P})$ a measure space and $(\Omega_i, \mathscr{A}_i)$ a measurable space for every $i \in I$. Let $X_i : \Omega \to \Omega_i$ be random variables for every $i \in I$. The family $(X_i)_{i \in I}$ is called stochastically independent if $\big(X_i^{-1}(\mathscr{A}_i)\big)_{i \in I}$ is stochastically independent.*

Theorem 2.4. *Let $\mathscr{E}_i$ be a $\cap$-stable generator of $\mathscr{A}_i$ for all $i \in I$. Then $(X_i)_{i\in I}$ is stochastically independent if $(X_i^{-1}(\mathscr{E}_i))_{i\in I}$ is stochastically independent.*

Proof. We know that $X_i^{-1}(B) \cap X_i^{-1}(C) = X_i^{-1}(B \cap C)$, so $X_i^{-1}(\mathscr{E}_i)$ is stable under intersection (for every $i \in I$). Thanks to theorem 2.2, it therefore suffices to shows that

$$\sigma(X_i^{-1}(\mathscr{E}_i)) = X_i^{-1}(\underbrace{\sigma(\mathscr{E}_i)}_{=\mathscr{A}_i}).$$

The inclusion $\sigma(X_i^{-1}(\mathscr{E}_i)) \subset X_i^{-1}(\sigma(\mathscr{E}_i))$ is easy:

$$\mathscr{E}_i \subset \sigma(\mathscr{E}_i) \implies X_i^{-1}(\mathscr{E}_i) \subset \underbrace{X_i^{-1}(\sigma(\mathscr{E}_i))}_{\sigma\text{-algebra}} \implies \sigma(X_i^{-1}(\mathscr{E}_i)) \subset X_i^{-1}(\sigma(\mathscr{E}_i)).$$

To show the other inclusion, we define

$$\mathscr{D} = \{B \in \sigma(\mathscr{E}_i) \mid X_i^{-1}(B) \in \sigma(X_i^{-1}(\mathscr{E}_i))\}$$

and aim to show that $\mathscr{D}$ is a Dynkin system over Ω_i.

(a) We have $\Omega_i \in \sigma(\mathscr{E}_i)$ and $X_i^{-1}(\Omega_i) = \Omega \in \sigma(X_i^{-1}(\mathscr{E}_i))$, so $\Omega_i \in \mathscr{D}$.

(b) Let $B \in \mathscr{D}$. We have

$$X_i^{-1}(B^c) = (X_i^{-1}(B))^c \in \sigma(X_i^{-1}(\mathscr{E}_i)),$$

so $B^c \in \mathscr{D}$.

(c) Let $B_n \in \mathscr{D}$ be pairwise disjoint sets $\forall n \in \mathbb{N}$. We have

$$X_i^{-1}\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \bigcup_{n\in\mathbb{N}} X_i^{-1}(B_n) \in \sigma(X_i^{-1}(\mathscr{E}_i)),$$

so $\bigcup_{n\in\mathbb{N}} B_n \in \mathscr{D}$.

We deduce

$$\mathscr{E}_i \subset \mathscr{D} \subset \sigma(\mathscr{E}_i) \implies \mathscr{D}(\mathscr{E}_i) \subset \mathscr{D} \subset \sigma(\mathscr{E}_i) \overset{\text{Prop. 1.8}}{\implies} \mathscr{D} = \sigma(\mathscr{E}_i)$$

and therefore, $X_i^{-1}(\sigma(\mathscr{E}_i)) \subset \sigma(X_i^{-1}(\mathscr{E}_i))$. □