Igor Kortchemski
Probality theory,
Autumn 2023
ETH

Chapter 1: σ-fields, measures

Outline: 1) σ-fields
2) measures
3) The Dynkin Lemma
4) Independence of σ fields

(see course webpage for recap on sets)

1) σ-fields

Definition Let Ω be a set. A σ-field on Ω is a collection $\mathcal{A}$ of subsets of Ω s.t.
(1) $\Omega \in \mathcal{A}$
(2) if $A \in \mathcal{A}$, then $A^c = \Omega \setminus A \in \mathcal{A}$ ("stability by complementation")
(3) if $(A_n)_{n\geq 1}$ is a sequence of elements of $\mathcal{A}$, then $\bigcup_{n\geq 1} A_n \in \mathcal{A}$ ("stability by countable union")

- $(\Omega, \mathcal{A})$ is a measurable space
- Elements of $\mathcal{A}$ are called measurable sets or events

Remarks:
- by (2), (1) is equivalent to (1') $\emptyset \in \mathcal{A}$
- by (2), if $A_n \in \mathcal{A}$ $\forall n \geq 1$, then $\bigcap_{n\geq 1} A_n \in \mathcal{A}$
- since $\emptyset \in \mathcal{A}$, $\mathcal{A}$ is stable by finite unions and intersections
- interpretation in probability: Ω represents everything that can happen in the system. Elements of $\mathcal{A}$ are the sets an observer is able to detect.
- We sometimes say "σ-algebra" instead of σ-field, it's exactly the same thing.

Examples Let Ω be a set. The following are σ-fields:
- $\mathcal{A}_1 = \{\emptyset, \Omega\}$
- $\mathcal{A}_2 = \mathcal{P}(\Omega)$ (powerset: set of all subsets of Ω)
- $\mathcal{A}_3 = \{A \subset \Omega : A \text{ or } A^c \text{ is countable}\}$

Example: $\{A \subset \mathbb{N} : A \text{ or } A^c \text{ finite}\}$ is not a σ-field.

Definition: Let $(A_n)_{n\geq 1}$ be a sequence of events of $(\Omega, \mathcal{A})$. We define
$$\limsup_{n\to\infty} A_n = \bigcap_{N\geq 0} \bigcup_{n\geq N} A_n \quad \text{and} \quad \liminf_{n\to\infty} A_n = \bigcup_{n\geq 0} \bigcap_{N\geq n} A_n$$
which are events

Remark. For $\omega \in \Omega$, we have $\omega \in \limsup_{n\to\infty} A_n \Leftrightarrow \{n : \omega \in A_n\}$ is infinite

$\Leftrightarrow$ "A_n happens infinitely often"

and $\omega \in \liminf_{n\to\infty} A_n \Leftrightarrow \exists n(\omega)$ s.t. $n \geq n(\omega) \Rightarrow \omega \in A_n$

$\Leftrightarrow$ "A_n always happens after a certain time"

For these reason, such events play a HUGE role in probability.

WARNING This should not be confused with $\liminf_{n\to\infty} x_n$, $\limsup_{n\to\infty} x_n$ for a sequence (x_n) of real numbers. Recall that
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} (\inf_{m\geq n} x_m) \in \mathbb{R} \cup \{\pm\infty\}, \quad \limsup_{n\to\infty} x_n = \lim_{n\to\infty} (\sup_{m\geq n} x_m) \in \mathbb{R} \cup \{\pm\infty\}$$
are the smallest and largest cluster points of (x_n) (= possible limits of subsequences)

Proposition: Let $(\mathcal{A}_i)_{i\in I}$ be a collection of σ-fields on a set Ω (with I being any collection of indices, not necessarily countable). Then $\bigcap_{i\in I} \mathcal{A}_i$ is a σ-field. In other words, any intersection of σ-fields is a σ-field.

Proof. First, since $\Omega \in \mathcal{A}_i$ for every $i \in I$, we have $\Omega \in \bigcap_{i\in I} \mathcal{A}_i$.

Second, if $A \in \bigcap_{i\in I} \mathcal{A}_i$, then for every $i \in I$ we have $A \in \mathcal{A}_i$, so ${}^cA \in \mathcal{A}_i$ since $\mathcal{A}_i$ is a σ-field. Hence ${}^cA \in \bigcap_{i\in I} \mathcal{A}_i$.

Third, if $(A_n)_{n\geq 1}$ is a sequence of elements of $\bigcap_{i\in I} \mathcal{A}_i$, then for every $i \in I$, $(A_n)_{n\geq 1}$ is a sequence of elements of $\mathcal{A}_i$, so $\bigcup_{n\geq 1} A_n \in \mathcal{A}_i$ since $\mathcal{A}_i$ is a σ-field.

Hence $\bigcup_{n\geq 1} A_n \in \bigcap_{i\in I} \mathcal{A}_i$.

This result allows to define the notion of generated σ-field:

Definition If $\mathcal{C} \subset \mathcal{P}(\Omega)$ is a collection of subsets of Ω, we set

$$\sigma(\mathcal{C}) := \bigcap_{\substack{\mathcal{A}\ \sigma\text{-field} \\ \text{such that } \mathcal{C} \subset \mathcal{A}}} \mathcal{A} \qquad (*),$$

called the σ-field generated by $\mathcal{C}$. It is the smallest σ-field containing all elements of $\mathcal{C}$

Examples • For $\Omega = \{1,2,3,4\}$ and $\mathcal{C} = \{\{1\},\{2\}\}$ we have $\sigma(\mathcal{C}) = \{\emptyset, \Omega, \{1\}, \{2,3,4\}, \{2\}, \{1,3,4\}, \{3,4\}, \{1,2\}\}$

• If $A \subset \Omega$, $\sigma(\{A\}) = \{\emptyset, A, A^c, \Omega\}$

Notation When $A_1, \dots, A_k \subset \Omega$, we write $\sigma(A_1, \dots, A_k)$ for $\sigma(\{A_1, \dots, A_k\})$.

The notion of generated σ-field is very useful to "construct" σ-fields. However, unlike in the previous examples, in general elements of generated σ-fields are not explicit at all.

In $(*)$ the intersection, which is taken on all σ-fields containing $\mathcal{C}$ is nonempty (since there is at least one σ-field containing $\mathcal{C}$, namely $\mathcal{P}(\Omega)$)). By the previous proposition, $\sigma(\mathcal{C})$ is a σ-field, called the σ-field generated by $\mathcal{C}$. Note that $\mathcal{C} \subset \sigma(\mathcal{C})$.

Remark If $\mathcal{C}$ is a σ-field, then $\sigma(\mathcal{C}) = \mathcal{C}$. Indeed, by taking $\mathcal{A} = \mathcal{C}$ in $(*)$, we get $\sigma(\mathcal{C}) \subset \mathcal{C}$, and the reverse inclusion always holds as noted before

Proposition: If $\mathcal{C} \subset \mathcal{C}'$, then $\sigma(\mathcal{C}) \subset \sigma(\mathcal{C}')$

Proof: We have $\mathcal{B} \subset \mathcal{B}' \subset \sigma(\mathcal{B}')$. Hence $\sigma(\mathcal{B}')$ is a σ-field containing $\mathcal{B}$. The result follows by taking $\mathcal{A} = \sigma(\mathcal{B}')$ in $(*)$

END OF LECTURE 1

Example: Take $\Omega = \{0,1\}^{\mathbb{N}}$, which can model the outcomes of throwing infinitely many times a coin. We say that a subset of Ω is a cylinder set (or, in short, a cylinder) if it is of the form
$$C_{x_1,\dots,x_n} = \{(a_i)_{i\geq 1} \in \Omega : a_1 = x_1, \dots, a_n = x_n\}$$
with $n \geq 1$ and $x_1, \dots, x_n \in \{0,1\}$. The cylinder σ-algebra $\mathcal{B}_{cyl}$ is by definition generated by all the cylinders. For example, $\{(1,1,\dots,1)\} \in \mathcal{B}_{cyl}$ because $\{(1,\dots,1)\} = \bigcap_{n\geq 1} C_{\underbrace{1,\dots,1}_{n \text{ times}}}$

Example: Take $\Omega = \mathbb{R}$ and let $\mathcal{A}$ be the σ-field generated by all the singletons $\{x\}$ with $x \in \mathbb{R}$. Then $\mathcal{A} = \{A \subset \mathbb{R} : A \text{ or } {}^cA \text{ is countable}\}$

Definition: If E is a metric space, the **Borel σ-field**, denoted by $\mathcal{B}(E)$ (or, in short, $\mathcal{B}$) is the σ-field generated by the open sets of E (or, equivalently, by the closed sets of E)

Example One can check that
$$\mathcal{B}(\mathbb{R}) = \sigma(]a,b[\,; a,b \in \mathbb{R}) = \sigma(]-\infty,a], a \in \mathbb{R}) = \sigma(]-\infty, a[, a \in \mathbb{Q})$$
(for this, the key property is that an open set in $\mathbb{R}$ can be written as a countable union of disjoint open intervals)

Definition: Let $(E,\mathcal{E})$ and $(F,\mathcal{F})$ be two measurable spaces. The product σ-field $\mathcal{E} \otimes \mathcal{F} = \sigma(A \times B : A \in \mathcal{E}, B \in \mathcal{F})$ on $E \times F$ is the smallest σ-field containing all elements of the form $A \times B$ with $A \in \mathcal{E}$, $B \in \mathcal{F}$.

2) Measures

Definition: A **measure** on a measurable space $(\Omega, \mathcal{A})$ is a function $\mu : \mathcal{A} \to \mathbb{R}_+ \cup \{+\infty\}$ such that

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

(2) if $(A_n)_{n \geq 1}$ is a sequence of elements of $\mathcal{A}$ which are pairwise disjoint (i.e. $A_i \cap A_j = \emptyset$ for $i \neq j$), then

$$\mu\Big(\bigcup_{n \geq 1} A_n\Big) = \sum_{n \geq 1} \mu(A_n)$$

where the sum belongs to $\mathbb{R}_+ \cup \{+\infty\}$

When $\mu(\Omega) < \infty$, we say that μ is finite.

When $\mu(\Omega) = 1$, we say that μ is a probability measure (usually we then use $\mathbb{P}, \mathbb{Q}$ etc. instead of μ) and $(\Omega, \mathcal{A}, \mu)$ is called a probability space. In this case, μ takes it values in $[0,1]$.

The following properties are very useful:

Proposition Let μ be a measure on $(\Omega, \mathcal{A})$

(1) For every $A, B \in \mathcal{A}$, if $A \subset B$, then $\mu(B) = \mu(A) + \mu(B \setminus A)$. In particular, $\mu(B \setminus A) = \mu(B) - \mu(A)$ if $\mu(A) < \infty$

(2) If $(A_i)_{i \geq 1}$ are measurable sets with $A_1 \subset A_2 \subset \cdots$ then $\mu\Big(\bigcup_{n \geq 1} A_n\Big) = \lim_{n \to \infty} \mu(A_n)$

(3) If $(A_i)_{i \geq 1}$ are measurable sets with $A_1 \supset A_2 \supset \cdots$ and $\mu(A_1) < \infty$ then $\mu\Big(\bigcap_{n \geq 1} A_n\Big) = \lim_{n \to \infty} \mu(A_n)$

(4) If $(A_i)_{i \geq 1}$ are measurable sets, $\mu\Big(\bigcup_{n \geq 1} A_n\Big) \leq \sum_{n=1}^{\infty} \mu(A_n)$

Proof: (1) consequence of the definition of a measure.

(2) Set $B_1 = A_1$, and for $i \geq 1$, $B_{i+1} = A_{i+1} \setminus A_i$, so $(B_i)_{i\geq 1}$ are pairwise disjoint. Introduce $A_i = \bigcup_{j=1}^{i} B_j$

and observe that $A = \bigcup_{i\geq 1} A_i = \bigcup_{i \geq 1} B_i$,

so $\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) = \lim_{n\to\infty} \mu(A_n)$

(3)

Use the complement event with (2)

(4) the first assertion gives $\mu(A \cup B) = \mu(A) + \mu(B \setminus A \cap B) \leq \mu(A) + \mu(B)$ since $B \setminus A\cap B \subset B$.

By induction, $\mu\left(\bigcup_{n=1}^{K} A_n\right) \leq \sum_{n=1}^{K} \mu(A_n)$, and then pass to the limit $K \to \infty$ by using the second assertion.

Examples: (1) The **counting measure** $\#$ on a set E is defined by $\#B = \mathrm{Card}(B)$ for a measurable set B. This measure is essentially used when E is countable.

(2) A **Dirac mass** on a set E is a measure of the form δ_a for $a \in E$, where for a measurable set A: $\delta_a(A) := \mathbb{1}_A(a) = \begin{cases} 1 & \text{if } a \in A \\ 0 & \text{if } a \notin A \end{cases}$.

(3) The Lebesgue measure λ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfies

$$\lambda((a,b)) = b - a \quad \text{for every } a < b.$$

Note that any positive linear combination of measures is a measure.

For example, on a set E with $\alpha_1, \dots, \alpha_n \geq 0$ and $a_1, \dots, a_n \in E$, $\sum_{k=1}^{n} \alpha_k \delta_{a_k}$ is a measure.

Also, for instance, $\lambda + \frac{1}{3}\delta_2$ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Remark: We keep the notation of the example with $E=\{0,1\}^{\mathbb{N}}$ and the cylinder σ-algebra $\mathcal{B}_{cyl}$. It is possible to show that there does not exist a measure μ on $(\{0,1\}^{\mathbb{N}}, \mathcal{P}(\{0,1\}^{\mathbb{N}}))$ such that

$$\mu(C_{x_1,\dots,x_n}) = \frac{1}{2^n} \qquad \text{for every cylinder } C_{x_1,\dots,x_n}.$$

However, there exists such a measure if one replaces $\mathcal{P}(\{0,1\}^{\mathbb{N}})$ with the cylinder σ-algebra. Intuitively speaking, $\mathcal{P}(\{0,1\}^{\mathbb{N}})$ is "too large" for a measure to satisfy countable additivity. This is one of the reasons one cannot always work with $\mathcal{P}(E)$ as a σ-algebra

Notation: Let μ be a measure on $(\Omega, \mathcal{A})$.

- We say that μ is **σ-finite** if there exists a sequence $(A_n)_{n\geq 1}$ of elements of $\mathcal{A}$ such that $\mu(A_n)<\infty$ for every $n\geq 1$ and $\bigcup_{n\geq 1} A_n = E$.
- We say that $x\in E$ is an **atom** of μ if $\mu(\{x\})>0$. We say that μ is **non-atomic** (or **continuous**) if μ has no atoms.

3) The Dynkin lemma

Definition: Let $\mathcal{D} \subset \mathcal{P}(\Omega)$ be a collection of subsets of Ω. We say that $\mathcal{D}$ is a Dynkin system (or λ-system)

(1) $\Omega \in \mathcal{D}$

(2) if $A \in \mathcal{D}$, $A^c \in \mathcal{D}$

(3) if $(A_n)_{n\geq 1}$ is a pairwise disjoint sequence of elements of $\mathcal{D}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$

Example It is a simple matter to see that a σ-field is a Dynkin system. However, if for instance $\Omega = \{0,1,2,3\}$ and $\mathcal{D} = \{\phi, \{0,2\}, \{0,3\}, \{1,2\}, \{1,3\}, \Omega\}$, then $\mathcal{D}$ is a Dynkin system but is not a σ-field.

Lemma Assume that $\mathcal{D} \subset \mathcal{P}(\Omega)$ is a Dynkin system and is closed by finite intersections. Then $\mathcal{D}$ is a σ-field

Proof: Take $(A_n)_{n\geq 1}$ a sequence in $\mathcal{D}$. We show that $\bigcup_{n\geq 1} A_n \in \mathcal{D}$. Set $B_1 = A_1$
For $j \geq 2$, set $B_j = A_j \setminus (A_1 \cup \cdots \cup A_{j-1})$. By construction, $B_1 \cup \cdots \cup B_j = A_1 \cup \cdots \cup A_j$ and (B_j) are pairwise disjoint
We show by strong induction that $\forall j \geq 1$, $B_j \in \mathcal{D}$

- $j=1$: $B_1 = A_1 \in \mathcal{D}$
- Assume $B_1, \ldots, B_j \in \mathcal{D}$. Then $B_{j+1} = A_{j+1} \setminus (A_1 \cup \cdots \cup A_j) = A_{j+1} \setminus (B_1 \cup \cdots \cup B_j)$

$$= A_{j+1} \cap \underbrace{(\Omega \setminus (\underset{\in \mathcal{D}}{B_1 \cup \cdots \cup B_j}))}_{\in \mathcal{D}} \in \mathcal{D} \quad \text{(finite intersection)}$$

END OF LECTURE 2.

The proof of the following result is left to the reader (it is very similar to the analog for σ-fields)

Proposition Any intersection of Dynkin systems is a Dynkin system.

As for generated σ-fields, the previous result allows us to define for $\mathcal{C} \subset \mathcal{P}(\Omega)$ the notion of generated Dynkin system

Definition If $\mathcal{C} \subset \mathcal{P}(\Omega)$ is a collection of subsets of Ω, we set

$$\lambda(\mathcal{C}) := \bigcap_{\substack{A \text{ Dynkin system} \\ \text{with } \mathcal{C} \subset A}} A$$

, called the Dynkin system **generated** by $\mathcal{C}$

Theorem (Dynkin Lemma) Let Ω be a set and let $\mathcal{C} \subset \mathcal{P}(\Omega)$ be a collection of subsets of Ω. Assume that $\mathcal{C}$ is stable by finite intersections. Then $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$. In other words, the Dynkin-system generated by $\mathcal{C}$ is equal to the σ-algebra generated by $\mathcal{C}$.

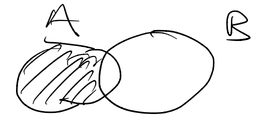
Proof: • Since $\sigma(\mathcal{C})$ is a λ-system, $\lambda(\mathcal{C}) \subset \sigma(\mathcal{C})$.

• To show $\sigma(\mathcal{C}) \subset \lambda(\mathcal{C})$, we show that $\lambda(\mathcal{C})$ is a σ-field by showing that it is stable by finite intersections.

We argue in two steps.

First, fix $A \in \mathcal{C}$ and consider $\lambda_A = \{B \in \mathcal{P}(\Omega),\ A \cap B \in \lambda(\mathcal{C})\}$.

We show that λ_A is a Dynkin-system containing $\mathcal{C}$, which will imply $\lambda(\mathcal{C}) \subset \lambda_A$



(0) by assumption, $\mathcal{C} \subset \lambda_A$.

(1) $\Omega \in \lambda_A$ since $A \in \mathcal{C}$

(2) First, observe that $A \in \lambda(\mathcal{C})$

If $B \in \lambda_A$, then $A \cap (B^c) = \Omega \setminus \big(\underbrace{(A \cap B)}_{\in \lambda(\mathcal{C})} \underbrace{\cup}_{\text{disjoint}} \underbrace{(A^c)}_{\in \lambda(\mathcal{C})} \big) \in \lambda(\mathcal{C})$ because $\lambda(\mathcal{C})$ is a Dynkin-system. Hence $B^c \in \lambda_A$.

(3) If $(B_n)_{n \geq 1}$ are pairwise disjoint elements of λ_A, then $A \cap \left(\bigcup_{n \geq 1} B_n \right) = \underbrace{\bigcup_{n \geq 1}}_{\text{disjoint union}} \underbrace{A \cap B_n}_{\in \lambda(\mathcal{C})} \in \lambda(\mathcal{C})$ since $\lambda(\mathcal{C})$ is a Dynkin system, so $\bigcup_{n \geq 1} B_n \in \lambda_A$

The first step allows to conclude that

$$\boxed{\forall A \in \mathcal{C},\ \forall B \in \lambda(\mathcal{C}),\ A \cap B \in \lambda(\mathcal{C})}$$

For the **second step**, we repeat this reasoning but we now fix $A \in \lambda(\mathcal{C})$ and similarly show that λ_A is a Dynkin system containing $\mathcal{C}$, which implies that $\lambda(\mathcal{C}) \subset \lambda_A$.

This shows that $\lambda(\mathcal{C})$ is stable by finite intersection, and completes the proof.

~

In practice, the Dynkin lemma is often used in the following form:
if $\mathcal{D}$ is a Dynkin system containing a collection $\mathcal{A}$ which is stable by finite intersections, then

$$\sigma(\mathcal{A}) \subset \mathcal{D}.$$

Indeed,
$$\sigma(\mathcal{A}) \underset{\substack{\uparrow \\ \text{Dynkin lemma}}}{=} \lambda(\mathcal{A}) \subset \mathcal{D}.$$

The following important idea appears in the previous proof:

If $P(A)$ is some property of a set A, if $\mathcal{A} \subset \mathcal{P}(\Omega)$ is a collection of subsets of Ω, to show that $\forall A \in \mathcal{A}$, $P(A)$ is true, consider the set of "good" things, i.e.

$\mathcal{G} = \{A \in \mathcal{A} : P(A) \text{ holds}\}$ (or $\mathcal{G} = \{A \subset \Omega : P(A) \text{ holds}\}$).

Indeed, if $\mathcal{B} \subset \mathcal{A}$ is such that $\forall C \in \mathcal{B}$ $P(C)$ holds

(view $\mathcal{B}$ as "nice" sets which we know how to study):

- if $\sigma(\mathcal{B}) = \mathcal{A}$ and $\mathcal{G}$ is a σ-field, then $\mathcal{A} \subset \mathcal{G}$ (so $\forall A \in \mathcal{A}$, $P(A)$ is true)
- if $\sigma(\mathcal{B}) = \mathcal{A}$, $\mathcal{B}$ is stable by finite intersections and $\mathcal{G}$ is a Dynkin system, then $\mathcal{A} \subset \mathcal{G}$.

This justifies the following definition:

Definition: Let $(\Omega, \mathcal{A})$ be a measurable space and $\mathcal{B} \subset \mathcal{A}$ a collection of measurable subsets

- We say that $\mathcal{B}$ is a **π-system** if $\mathcal{B}$ is stable by finite intersections
- We say in addition that $\mathcal{B}$ is a **generating** π-system if $\sigma(\mathcal{B}) = \mathcal{A}$.

Examples:

- The collection $\{(-\infty, a) : a \in \mathbb{R}\}$ is a generating π-system of $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
- The collection $\{(a,b) : a, b \in \mathbb{R}, a < b\}$ is a generating π-system of $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
- The collection $\{\{i\} : i \in \mathbb{N}\}$ is a generating π-system of $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.
- For $E = \{0,1\}^{\{1,2,3,\dots\}}$, the cylinders are a generating π-system of $(E, \mathcal{E}_{cyl})$.

Corollary:

Let $(\Omega, \mathcal{A})$ be a measurable space and $\mathcal{B}$ a generating π-system.

(1) Let μ, ν be finite measures such that $\mu(\Omega) = \nu(\Omega)$ and $\mu(C) = \nu(C)$ $\forall C \in \mathcal{B}$

Then $\mu = \nu$

(2) More generally, if there exists $E_n \in \mathcal{A}$ s.t. $\mu(E_n) = \nu(E_n) < \infty$, $\bigcup_{n \geq 1} E_n = E$ and $\mu(E_n \cap C) = \nu(E_n \cap C)$ for every $C \in \mathcal{B}$, then $\mu = \nu$.

Proof: We only establish (1) and leave the extension (2) to the reader.

Define $\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$.

We check that $\mathcal{D}$ is a monotone class containing $\mathcal{C}$ and the desired result will follow from the Dynkin lemma, which will imply that $\mathcal{D}$ contains $\sigma(\mathcal{C}) = \mathcal{A}$.

① $\Omega \in \mathcal{D}$ since $\mu(\Omega) = \nu(\Omega)$

② If $A \in \mathcal{D}$, then $\mu(A^c) = \mu(\Omega) - \mu(A) = \nu(\Omega) - \nu(A) = \nu(A^c)$, so $A^c \in \mathcal{D}$

③ If $(A_n)_{n \geq 1}$ is a disjoint sequence of $\mathcal{D}$:

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

so $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$. This completes the proof.

Example: this shows that there exists at most one measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mu((a,b)) = b-a$ for every $a<b$ (uniqueness of the Lebesgue measure)

4) Independence

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Two events A, B are said to be independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Interpretation When $\mathbb{P}(B) > 0$, this is equivalent to saying that $\mathbb{P}(A \mid B) \overset{\text{def}}{=} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ is equal to $\mathbb{P}(A)$.

conditional probability

Intuitively, observing that B holds does not influence the likelihood that A holds as well.

Examples (1) Throwing two dice $\Omega = \{1,..,6\}^2$, $\mathbb{P}(\{\omega\}) = \frac{1}{36}$ for $\omega \in \Omega$. Then $A = \{6\} \times \{1,..,6\}$ and $B = \{1,..,6\} \times \{6\}$ are independent

(2) Throwing one die $\Omega = \{1,..,6\}$, $\mathbb{P}(\{\omega\}) = \frac{1}{6}$ for $\omega \in \Omega$.

$A = \{1,2\}$ and $\{1,3,5\}$ are independent

Definition n events $A_1,..,A_n$ are (mutually) <u>independent</u> if for every non empty subset $\{j_1,...,j_p\}$ of $\{1,..,n\}$ we have $\mathbb{P}(A_{j_1} \cap ... \cap A_{j_p}) = \mathbb{P}(A_{j_1}) \cdots \mathbb{P}(A_{j_p})$

Equivalently, $\forall J \subset \{1,..,n\}$, $\#J \geq 1$, $\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$

Remark independence is relative to $\mathbb{P}$ (this is implicit)

Notation $(A_i)_{1 \leq i \leq n}$ are $\perp\!\!\!\perp$.

Remarks
- In general, $\mathbb{P}(A_1 \cap ... \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n) \not\Rightarrow (A_i)_{1\leq i \leq n}$ independent and $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for $i \neq j$ (<u>pairwise</u> independence) $\not\Rightarrow (A_i)_{1 \leq i \leq n}$ independent.
(By independence we <u>always</u> mean mutual independence)
- ⚠ "independent" and "disjoint" are two very different notions. For example, if $A \cap B = \emptyset$, $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$, so $\mathbb{P}(A \cap B) = 0$ and $\mathbb{P}(A)\mathbb{P}(B) > 0$, so A and B are not independent

Proposition Events $A_1,..,A_n$ are $\perp\!\!\!\perp$ if and only if $\mathbb{P}(B_1 \cap ... \cap B_n) = \mathbb{P}(B_1) \cdots \mathbb{P}(B_n)$ $(*)$ for every $B_i \in \sigma(A_i) = \{\emptyset, A_i, A_i^c, \Omega\}$ for every $1 \leq i \leq n$.

Proof $\boxed{\Leftarrow}$ If $J \subset \{1,..,n\}$ $\#J \geq 1$, take $B_i = A_i$ for $i \in J$, $B_i = \Omega$ for $i \notin J$.

$\boxed{\Rightarrow}$ Set $J = \{i : B_i \neq \Omega\}$. We have to check that $\mathbb{P}\left(\bigcap_{i \in J} B_i\right) = \prod_{i \in J} \mathbb{P}(B_i)$

Therefore it is sufficient to check that if $C_1,..,C_p$ are $\perp\!\!\!\perp$, then $C_1^c,..,C_p$ are $\perp\!\!\!\perp$.

To this end, for $\{i_1,..,i_q\} \subset \{2,..,p\}$ write

$$\mathbb{P}(C_1^c \cap C_{i_1} \cap ... \cap C_{i_q}) = \mathbb{P}(C_{i_1} \cap ... \cap C_{i_q}) - \mathbb{P}(C_1 \cap ... \cap C_{i_q})$$
$$= \mathbb{P}(C_{i_1}) \cdots \mathbb{P}(C_{i_q}) - \mathbb{P}(C_1) \cdots \mathbb{P}(C_{i_q})$$
$$= \mathbb{P}(C_1^c)\,\mathbb{P}(C_{i_1}) \cdots \mathbb{P}(C_{i_q})$$

END OF LECTURE 3

This result naturally leads to the notion of independent σ-fields, which the "good" setting for independence

Definition Let $\mathcal{B}_1, \dots, \mathcal{B}_n \subset \mathcal{A}$ be σ-fields. They are independent (notation: $\perp\!\!\!\perp$) if $\forall A_1 \in \mathcal{B}_1, \dots, \forall A_n \in \mathcal{B}_n$, $\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$.

By the previous results, events $A_1, \dots, A_n$ are $\perp\!\!\!\perp$ $\iff$ the σ-fields $\sigma(A_1), \dots, \sigma(A_n)$ are $\perp\!\!\!\perp$

The following result is useful to show independence:

Proposition Let $\mathcal{B}_1, \dots, \mathcal{B}_n \subset \mathcal{A}$ be σ-fields. For $1 \leq i \leq n$, let $\mathcal{C}_i$ be a π-system such that $\mathcal{B}_i \subset \sigma(\mathcal{C}_i)$ and $\Omega \in \mathcal{C}_i$.
Then $\forall C_1 \in \mathcal{C}_1, \dots, \forall C_n \in \mathcal{C}_n$, $\mathbb{P}(C_1 \cap \dots \cap C_n) = \mathbb{P}(C_1) \cdots \mathbb{P}(C_n) \implies \mathcal{B}_1, \dots, \mathcal{B}_n$ are independent

Proof The proof is based on the Dynkin lemma:
First fix $C_2 \in \mathcal{C}_2, \dots, C_n \in \mathcal{C}_n$ and set $\lambda_1 = \{ B_1 \in \mathcal{B}_1 : \mathbb{P}(B_1 \cap C_2 \cap \dots \cap C_n) = \mathbb{P}(B_1)\mathbb{P}(C_2) \cdots \mathbb{P}(C_n) \}$.
Then one checks that λ_1 is a λ-system containing $\mathcal{C}_1$, thus $\lambda(\mathcal{C}_1) = \sigma(\mathcal{C}_1)$ by Dynkin's lemma
thus $\forall B_1 \in \mathcal{B}_1, \forall C_2 \in \mathcal{C}_2, \dots, \forall C_n \in \mathcal{C}_n$, $\mathbb{P}(B_1 \cap C_2 \cap \dots \cap C_n) = \mathbb{P}(B_1)\mathbb{P}(C_2) \cdots \mathbb{P}(C_n)$
Then similarly fix $B_1 \in \mathcal{B}_1$, $C_3 \in \mathcal{C}_3, \dots, C_n \in \mathcal{C}_n$ and set

$$\lambda_2 = \{ B_2 \in \mathcal{B}_2 : \mathbb{P}(B_1 \cap B_2 \cap C_3 \cap \dots \cap C_n) = \mathbb{P}(B_1)\mathbb{P}(B_2)\mathbb{P}(C_3) \cdots \mathbb{P}(C_n) \},$$

which a λ-system containing $\mathcal{C}_2$ and then containing $\lambda(\mathcal{C}_2) = \sigma(\mathcal{C}_2)$.
We continue by induction to get the desired result.

Application (coalition principle): Let $\mathcal{B}_1, \dots, \mathcal{B}_n$ be $\perp\!\!\!\perp$ σ-fields. Fix $1 \leq n_1 < n_2 < \dots < n_p = n$
Then $\mathcal{D}_1 = \sigma(\mathcal{B}_1, \dots, \mathcal{B}_{n_1})$, $\dots$, $\mathcal{D}_p = \sigma(\mathcal{B}_{n_{p-1}+1}, \dots, \mathcal{B}_n)$ are $\perp\!\!\!\perp$
(notation for $\sigma(\mathcal{B}_1 \cup \dots \mathcal{B}_{n_1})$)

Proof: We check that $\mathcal{C}_j = \{ B_{n_{j-1}+1} \cap \dots \cap B_{n_j} : B_i \in \mathcal{B}_i \}$ is a generating π-system of $\mathcal{D}_j$ and the result follows by the previous proposition (we have $\Omega \in \mathcal{C}_j$)

To simplify notation, we do it for $j=1$.

* $\mathcal{C}_1$ is stable by finite intersections by definition.
* We show $\sigma(\mathcal{C}_1)=\sigma(\mathcal{B}_1,\dots,\mathcal{B}_{n_1})$ by double inclusion
  - Since $\forall A\in\mathcal{C}_1$ we have $A\in\sigma(\mathcal{B}_1,\dots,\mathcal{B}_{n_1})$, we get $\mathcal{C}_1\subset\sigma(\mathcal{B}_1,\dots,\mathcal{B}_{n_1})$ so $\sigma(\mathcal{C}_1)\subset\sigma(\mathcal{B}_1,\dots,\mathcal{B}_{n_1})$
  - We have $\mathcal{B}_1,\dots,\mathcal{B}_{n_1}\subset\mathcal{C}_1$, so $\mathcal{B}_1\cup\dots\cup\mathcal{B}_{n_1}\subset\mathcal{C}_1$, thus $\sigma(\mathcal{B}_1\cup\dots\cup\mathcal{B}_{n_1})\subset\sigma(\mathcal{C}_1)$.

Definition (independence of any family)

- Let $(\mathcal{B}_i)_{i\in I}$ be σ-fields included in $\mathcal{A}$, with I being any set. They are <u>independent</u> if any finite collection is independent, i.e:
$$\forall J\subset I,\quad Card(J)<+\infty,\quad (\mathcal{B}_j)_{j\in J} \text{ are independent.}$$
- Similarly, events $(A_i)_{i\in I}$ are <u>independent</u> if
$$\forall J\subset I,\ Card(J)<\infty,\ (\mathcal{B}_j)_{j\in J} \text{ are independent.}$$

For events $(A_n)_{n\geq 1}$, recall the notation $\limsup A_n=\bigcap_{l\geq 0}\Big(\bigcup_{n\geq l} A_n\Big)$ and $\liminf A_n=\bigcup_{l\geq 0}\Big(\bigcap_{n\geq l} A_n\Big)$

The following results are very useful to show that events have probability 0 or 1.

Borel-Cantelli lemmas

1) If $\sum_{n=1}^{\infty}\mathbb{P}(A_n)<+\infty$, then $\mathbb{P}(\limsup A_n)=0$
2) If $\sum_{n=1}^{\infty}\mathbb{P}(A_n)=+\infty$ and $(A_n)_{n\geq 1}$ are $\perp\!\!\!\perp$, then $\mathbb{P}(\limsup A_n)=1$

Remark/intuition

In 1), A_n is so unlikely that a.s. A_n happens only a finite number of times (for n sufficiently large A_n does not happen)

In 2), A_n is not that unlikely, so that a.s. A_n happens infinitely many often

Proof 1) For $n \geq 1$, $\limsup_{n\to\infty} A_n \subset \bigcup_{k\geq n} A_k$, so $\mathbb{P}(\limsup_{n\to\infty} A_n) \leq \mathbb{P}(\bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} \mathbb{P}(A_k) \xrightarrow[n\to\infty]{} 0$ as the remainder of a convergent series.

The result follows.

2) Fix $\ell \geq 1$, $n \geq \ell$ and write

$$\mathbb{P}\left(\bigcap_{k=\ell}^{n} A_k^c\right) = \prod_{k=\ell}^{n} \mathbb{P}(A_k^c) = \prod_{k=\ell}^{n} (1 - \mathbb{P}(A_k)).$$

This tends to 0 as $n \to \infty$. Indeed, using $\ln(1-x) \leq -x$ for $0 \leq x \leq 1$,

$$\prod_{k=\ell}^{n} (1 - \mathbb{P}(A_k)) = \exp\left(\sum_{k=\ell}^{n} \ln(1 - \mathbb{P}(A_k))\right) \leq \underbrace{\exp\Big(-\underbrace{\sum_{k=\ell}^{n} \mathbb{P}(A_k)}_{\xrightarrow[n\to\infty]{} -\infty}\Big)}_{\xrightarrow[n\to\infty]{} 0}$$

Hence $\mathbb{P}\left(\bigcap_{k=\ell}^{+\infty} A_k^c\right) = 0$

Hence $\mathbb{P}(\liminf A_n^c) = \mathbb{P}\left(\bigcup_{\ell=1}^{\infty} \bigcap_{k=\ell}^{+\infty} A_k^c\right) = 0$

Hence $\mathbb{P}(\limsup A_n) = 1$

⚠ $\perp\!\!\!\perp$ in 2) is important (otherwise take $A_n = A$ with $0 < \mathbb{P}(A) < 1$ for a counter-example).