

## 1 Review of Probability Spaces

The standard notation for a probability space is  $(\Omega, \mathcal{F}, \mathbb{P})$ . The components of this tuple are:

1.  $\Omega$ : this is the **sample space**, which is the collection of ALL possible outcomes.  $\omega \in \Omega$  is a **sample point**.  $\omega$  corresponds to a specific outcome.
2.  $\mathcal{F}$ : this is a  **$\sigma$ -algebra**. This is a collection of events. For  $A \in \mathcal{F}$ , we call  $A$  an **event**.  $A \subseteq \Omega$ . As we will see later, a  $\sigma$ -algebra is a collection of subsets of  $\Omega$ . This satisfies certain conditions.
3.  $\mathbb{P}$ : this is a function defined on a sigma algebra.

$$\begin{aligned}\mathbb{P} : \mathcal{F} &\rightarrow [0, 1], \\ A \in \mathcal{F} &\mapsto \mathbb{P}(A) \in [0, 1].\end{aligned}$$

We call  $\mathbb{P}(A)$  the **probability of event A**.

**Example 1.1.** Consider flipping a fair coin. Then:

$$\begin{aligned}\Omega &= \{H, T\}, \\ \mathcal{F} &= \{\{H\}, \{T\}, \{H, T\}, \emptyset\}, \\ \mathbb{P}(H) &= \frac{1}{2}, \quad \mathbb{P}(T) = \frac{1}{2}, \quad \mathbb{P}(\{H, T\}) = 1, \quad \mathbb{P}(\emptyset) = 0.\end{aligned}$$

**Example 1.2.** Will do later, It's annoying to write out.

### 1.1 Measure Theory

**Measure theory** is the foundation of modern probability theory. We will define things for a general measure space  $(S, \Sigma, \mu)$  to replace  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Definition 1.1** (Algebra). Let  $S$  be a set. A collection  $\Sigma_0$  of subsets of  $S$  is called an **algebra** if:

1.  $S \in \Sigma_0$ .
2.  $A \in \Sigma_0 \Rightarrow A^c := S \setminus A \in \Sigma_0$  (**closed under complements**).
3.  $\forall n \in \mathbb{N}, A_1, \dots, A_n \in \Sigma_0 \Rightarrow \bigcup_{j=1}^n A_j \in \Sigma_0$  (**closed under finite unions**).

Some remarks: if  $\Sigma_0$  is an algebra of  $S$ , then:

4.  $\emptyset \in \Sigma_0$ .
5. if  $A, B \in \Sigma_0$ , then  $A \cup B, A \cap B, A \setminus B, A \Delta B, B \setminus A \in \Sigma_0$ .
6.  $\forall n \in \mathbb{N}, A_1, \dots, A_n \in \Sigma_0 \Rightarrow \bigcap_{j=1}^n A_j \in \Sigma_0$ .

Note that all of these operations are finite.

**Definition 1.2** ( $\sigma$ -algebra). A collection of subsets  $\Sigma$  of  $S$  is a **sigma-algebra** if:

1.  $\Sigma$  is an algebra.
2.  $A_1, A_2, \dots \in \Sigma \Rightarrow \bigcup_{j=1}^{\infty} A_j \in \Sigma$  (closed under countable unions).

**Remark 1.1.** If  $\Sigma$  is a sigma algebra, then  $\Sigma$  satisfies (1)-(6) and:

$$A_1, A_2, \dots \in \Sigma \Rightarrow \bigcap_{j=1}^{\infty} A_j \in \Sigma. \tag{1}$$

Very often at this stage, when we want to prove something, we have to go back to the definitions.

**Definition 1.3** (Measurable Space). The pair  $(S, \Sigma)$  is called a **measurable space**. A set  $A \in \Sigma$  is a **measurable set**.

**Definition 1.4** (Sigma Algebra Generated). Let  $\mathcal{C}$  be a collection of subsets of  $S$ . The  $\sigma$ -algebra **generated** by  $\mathcal{C}$ , denoted by  $\sigma(\mathcal{C})$ , is the smallest  $\sigma$ -algebra which is a subset of  $\mathcal{C}$ :

1.  $\mathcal{C} \subseteq \sigma(\mathcal{C})$ .
2. if  $\Sigma'$  is a  $\sigma$ -algebra containing  $\mathcal{C}$ , then  $\sigma(\mathcal{C}) \subseteq \Sigma'$ .

**Remark 1.2.** 1. if  $\mathcal{C}$  is a  $\sigma$ -algebra, then  $\sigma(\mathcal{C}) = \mathcal{C}$ .  
 2.  $\sigma(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$ .  
 3. if  $\mathcal{C}_1 \subseteq \mathcal{C}_2$ , then  $\sigma(\mathcal{C}_1) \subseteq \sigma(\mathcal{C}_2)$ .

**Proposition 1.1.**

$$\sigma(\mathcal{C}) = \bigcap \{ \Sigma \mid \Sigma \text{ is a } \sigma\text{-algebra and } \mathcal{C} \in \Sigma \}. \quad (2)$$

**Fact:** if  $\{\Sigma_\alpha \mid \alpha \in I\}$  where  $I$  is some index set is any collection of  $\sigma$ -algebras of subsets of  $S$ , then  $\bigcap \Sigma_\alpha$  remains a  $\sigma$ -algebra, i.e., intersections of  $\sigma$ -algebras remain  $\sigma$ -algebras.

**Example 1.3.** Given  $S$ , let  $A, B \subseteq S$ . Then,  $\sigma(\{A\}) = \{A, A^c, \emptyset, S\}$ .

*Question: What do  $\sigma$ -algebras mean for us?*

A  $\sigma$ -algebra contains the collection of events we can study. In other words, it tells me the information available to me, from the point of view of probability. If you're not in the  $\sigma$ -algebra, then you're not measurable with respect to the  $\sigma$ -algebra.

**Example 1.4** (Borel  $\sigma$ -algebra). Take  $S = \mathbb{R}$ . Then, we define the **Borel Sigma Algebra** to be:

$$\mathcal{B}(\mathbb{R}) := \sigma(\{ \text{open subsets of } \mathbb{R} \}). \quad (3)$$

This applies to any topological space. Equivalently, only for  $\mathbb{R}$ , this reduces to:

$$\mathcal{B}(\mathbb{R}) = \sigma(\{ ]a, b[ \mid a < b, a, b \in \mathbb{R} \}) =: \Sigma_{],[}. \quad (4)$$

Note that the generating class of  $\mathcal{B}(\mathbb{R})$  is *not* unique.

**Definition 1.5** ( $\pi$ -system). Let  $S$  be a set. A collection  $\mathcal{I}$  (of subsets of  $S$ ) is called a  **$\pi$ -system** if  $\forall A, B \in \mathcal{I}$ , one has  $A \cap B \in \mathcal{I}$ .

**Definition 1.6.** Let  $S$  be a set. A collection  $\mathcal{D}$  (of subsets of  $S$ ) is called a  **$\mathcal{D}$ -system** if:

1.  $S \in \mathcal{D}$ .
2. If  $A, B \in \mathcal{D}$  and if  $A \subseteq B$ , then  $B \setminus A \in \mathcal{D}$ .
3. If  $A_n \in \mathcal{D}$  for  $n \geq 1$ , and  $A_n \uparrow A$  then  $A \in \mathcal{D}$ .

Note that these two definitions mean that we can separate the properties of  $\mathcal{B}(\mathbb{R})$  into a  $\pi$ -system and a  $\mathcal{D}$ -system. This will allow us to further decode a sigma algebra.

**Lemma 1.1.** Let  $\Sigma$  be a collection of subsets of  $S$ . Then,  $\Sigma$  is a sigma-algebra  $\iff \Sigma$  is a  $\pi$ -system and a  $\mathcal{D}$ -system.

*Proof.* " $\implies$ ": trivial.

" $\impliedby$ ": We just need to verify that  $\Sigma$  is a  $\sigma$ -algebra.

1.  $S \in \Sigma$  follows from the fact that  $\Sigma$  is a d-algebra.

2. If  $A \in \Sigma$ , then  $A^c = S \setminus A \in \Sigma$  from (2) of d-system definition.
3. If  $A_n \in \Sigma$  for  $n \geq 1$ , we need to check that  $\bigcup_{n=1}^{\infty} A_n \in \Sigma$ .

Set  $B_n := \bigcup_{j=1}^n A_j$ . Then,  $B_n$  is an increasing sequence, and  $B_n \in \Sigma$  for all  $n \in \mathbb{N}$  (we know this since we can apply deMorgan's Law, then use the fact that  $\Sigma$  is a  $\pi$ -system and is hence closed under intersections. Now, using (III) of being a d-system, we can conclude:

$$\bigcup_{n=1}^{\infty} B_n \in \Sigma \Rightarrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n \in \Sigma.$$

□

**Theorem 1.1** (Dynkin's  $\pi$ -d lemma). *Suppose that  $\mathcal{I}$  is a  $\pi$ -system of subsets of  $S$ , and  $d(\mathcal{I})$  is the d-system generated by  $\mathcal{I}$ . Then,  $d(\mathcal{I}) = \sigma(\mathcal{I})$ .*

In words, this theorem is saying that if you start with a  $\pi$ -system generating class, all you need is a d-system and you'll automatically get a  $\sigma$ -algebra.

*Proof.* We observe that it is sufficient to show that  $d(\mathcal{I})$  is a  $\pi$ -system. Why?

1. If we manage to show this, by the previous lemma, we would know that  $d(\mathcal{I})$  is a  $\sigma$ -algebra, and so by the definition of a  $\sigma$ -algebra one would get that  $\sigma(\mathcal{I}) \subseteq d(\mathcal{I})$ .
2. Since  $\sigma(\mathcal{I})$  is certainly a d-system,  $d(\mathcal{I}) \subseteq \sigma(\mathcal{I})$ .

So the GOAL: show that  $d(\mathcal{I})$  is a  $\pi$ -system. This proof requires two stages. We will use the **good set principle**, which is a common technique in set theory. Broadly speaking, you collect all items with a certain property, argue that this collection satisfies certain properties, then show that this collection is actually the whole set. So, our "good set" will be defined as follows:

$$\mathcal{D}_1 := \{ B \in d(\mathcal{I}) \mid B \cap A \in d(\mathcal{I}) \ \forall A \in \mathcal{I} \}.$$

**Claim:**  $\mathcal{D}_1$  is a d-system. We need to check against the definition of a d-system.

1.  $S \in \mathcal{D}_1$ : satisfied, since  $\forall A \in \mathcal{I}$ ,  $A \cap S = A \in d(\mathcal{I})$  since  $A \in \mathcal{I} \subseteq d(\mathcal{I})$ .
2. Let  $A_1, A_2 \in \mathcal{D}_1$ . Suppose that  $A_1 \subseteq A_2$ . We want to show that  $A_2 \setminus A_1 \in \mathcal{D}_1$ . So, we need to verify that  $(A_2 \setminus A_1) \cap A \in d(\mathcal{I})$  for any  $A \in \mathcal{I}$ :

$$A \cap (A_2 \setminus A_1) = (A_2 \cap A) \setminus (A_1 \cap A) \in d(\mathcal{I})$$

$(A_2 \cap A)$  and  $(A_1 \cap A)$  both belong to  $d(\mathcal{I})$  since by definition they are in  $\mathcal{D}_1$ . Now, since  $d(\mathcal{I})$  is a d-system, the difference is in  $d(\mathcal{I})$ .

3. Finally we need to show that when  $A_n \in \mathcal{D}_1$ , for  $n \geq 1$  and  $A_n \uparrow A_{\infty}$ , then  $A_{\infty} \in \mathcal{D}_1$ . However, since  $A \in \mathcal{I}$ :

$$\underbrace{A_n \cap A}_{\in d(\mathcal{I})} \uparrow A_{\infty} \cap A \Rightarrow A_{\infty} \cap A \in d(\mathcal{I}) \Rightarrow A_{\infty} \in \mathcal{D}_1.$$

This shows me that  $\mathcal{D}_1$  forms a d-system. Since  $\mathcal{I}$  is a  $\pi$ -system,  $\mathcal{I} \subseteq \mathcal{D}_1$ . This tells me that  $d(\mathcal{I}) \subseteq \mathcal{D}_1$ . However,  $\mathcal{D}_1$  is defined using only elements in  $d(\mathcal{I})$ , which then gives us the second inequality:  $\mathcal{D}_1 \subseteq d(\mathcal{I})$ . Therefore,  $\forall B \in d(\mathcal{I})$  and  $\forall A \in \mathcal{I}$ , one has that  $B \cap A \in d(\mathcal{I})$ . This was the intermediate step; we need to re-do this but with  $A \in d(\mathcal{I})$  now. Hence, we set:

$$\mathcal{D}_2 := \{ c \in d(\mathcal{I}) \mid B \cap C \in d(\mathcal{I}) \text{ for all } B \in d(\mathcal{I}) \}$$

From our intermediate step conclusion, we know that  $\mathcal{I} \subseteq \mathcal{D}_2$ . Next, we verify that  $\mathcal{D}_2$  is a d-system. Exercise: verify this. This shows that  $\mathcal{D}_2$  is a d-system and  $\mathcal{I} \subseteq \mathcal{D}_2$ . This shows us that  $d(\mathcal{I}) \subseteq \mathcal{D}_2$ . Hence,  $d(\mathcal{I}) = \mathcal{D}_2$ . Hence,  $\forall c \in d(\mathcal{I})$ ,  $\forall B \in d(\mathcal{I})$ ,  $B \cap C \in d(\mathcal{I})$ . This proves that  $d(\mathcal{I})$  is a  $\pi$ -system, which is what we wanted to show. □

This idea is very important in the study of measures. The reason why this theorem is important is because when constructing a measure, we only look at the  $\pi$ -system that generates the  $\sigma$ -algebra.

**Definition 1.7** (Additive). Let  $S$  be a set. Let  $\Sigma_0$  be an algebra of subsets of  $S$ . Let  $\mu_0$  be a non-negative set function defined on  $\Sigma - 0$  i.e.:

$$\mu_0 := \sigma_0 \rightarrow [0, \infty].$$

We say that  $\mu_0$  is **additive** if (1) is  $\mu_0(\emptyset) = 0$  and (2)  $\forall A, B \in \Sigma_0$  and  $A \cap B = \emptyset$ ,  $\mu_0(A \cup B) = \mu_0(A) + \mu_0(B)$ .

**Definition 1.8.** We say that  $\mu_0$  is **countably additive** if:

1.  $\mu_0(\emptyset) = 0$
2.  $\forall A_n \in \Sigma_0, n \geq 1$  such that  $A_i \cap A_j = \emptyset$  for all  $i \neq j$  and  $\cup_{n=1}^{\infty} A_n \in \Sigma_0$ , then we require:

$$\mu_0 \left( \bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

**Definition 1.9.** Let  $(S, \Sigma)$  be a measure space. If  $\mu$  is a non-negative set function defined on  $\Sigma$  and if  $\mu$  is countably additive, then  $\mu$  called a **measure** and the triple  $(S, \Sigma, \mu)$  is called a **measure space**.

1. If  $\mu(S) < \infty$ , then  $\mu$  is a **finite measure**.
2. If  $\mu(S) = 1$ , then  $\mu$  is a **probability measure**.
3. If  $\exists \{S_n | n \geq 1\} \subseteq \Sigma$  such that  $\cup_{n=1}^{\infty} S_n = S$  and  $\mu(S_n) < \infty$  for all  $n \geq 1$ , then  $\mu$  is  $\sigma$ -finite.
4. For  $N \in \Sigma$  such that  $\mu(N) = 0$ , then we say that  $N$  is a **null set**.
5. Of a statement holds everywhere except on a null set, then we say that the statement is true **almost everywhere** or **almost surely**.

**Remark:** All measures that we will discuss in this course will either be finite or  $\sigma$ -finite.

## 1.2 Properties of a Measure $\mu$

**Proposition 1.2** (Monotonicity). Let  $A, B \in \Sigma$ ,  $A \subseteq B$ . Then,  $\mu(A) \leq \mu(B)$ .

**Caution!** Do not take the difference,  $\mu(B) - \mu(A)$ , because both  $\mu(A)$  and  $\mu(B)$  might be infinite. This is undefined! If at least one of them are finite, then this is ok.

**Proposition 1.3** (Subadditivity). Let  $A_n \in \Sigma$  for all  $n \geq 1$ . Then:

$$\mu \left( \bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu(A_n). \quad (5)$$

*Proof.* We can prove this using the monotonicity of  $\mu$ . Define a new collection of sets  $B_n$ :

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 \setminus A_1 \\ B_n &= A_n \setminus \left( \bigcup_{j=1}^{n-1} A_j \right). \end{aligned}$$

The  $B_n$ 's are disjoint, and by construction  $B_n \subseteq A_n$ . This gives us that  $\forall n \geq 1$  we have:  $\mu(B_n) \leq \mu(A_n)$ . Furthermore,  $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$ . Combining all this together and using the countable additivity of  $\mu$  on the  $B_n$ 's, we obtain:

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

□

**Proposition 1.4** (Continuity from Below). Let  $A_n \in \Sigma$  for all  $n \geq 1$ . Also suppose that  $A_n \uparrow$  (“monotonic increasing sequence”). Then:

$$\mu \left( \bigcup_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (6)$$

When  $A_n \uparrow$ , we view  $\bigcup_{n=1}^{\infty} A_n$  as  $\lim_{n \rightarrow \infty} \mu(A_n)$ . Then,  $\mu(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$ . This is what continuity means: I can interchange the limit with the measure  $\mu$ .

*Proof.* Define a new sequence of disjoint sets  $B_n$  as follows:

$$\begin{aligned} B_1 &= A_1 \\ B_n &= A_n \setminus A_{n-1} \quad \forall n \geq 2. \end{aligned}$$

The  $B_n$ ’s are clearly disjoint, and  $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$ . Hence, we can use countable additivity:

$$\mu \left( \bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mu(B_j) = \lim_{n \rightarrow \infty} \mu \left( \bigcup_{j=1}^n B_j \right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

□

**Proposition 1.5** (Continuity from Above). If  $A_n \in \Sigma$  for all  $n \geq 1$ ,  $A_n \downarrow$ , AND  $\mu(A_n)$  is finite for *some*  $n \geq 1$ , then:

$$\mu \left( \bigcap_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n). \quad (7)$$

*Proof.* We’ll show this property through continuity from below. WLOG, assume that  $\mu(A_1) < \infty$ . By monotonicity, this implies that  $\mu(A_n)$  is finite  $\forall n \in \mathbb{N}$ . Using our decreasing sequence  $\{A_n\}$  of sets, construct an increasing sequence of sets as follows:

$$B_n := A_1 \setminus A_n \text{ for all } n \geq 1.$$

Then,  $B_n \uparrow$  and we can use Continuity from Below:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu(B_n) &= \mu \left( \bigcup_{n=1}^{\infty} B_n \right) \\ &= \mu \left( \bigcup_{n=1}^{\infty} (A_1 \setminus A_n) \right) \\ &= \mu \left( A_1 \setminus \bigcap_{n=1}^{\infty} A_n \right) \\ &= \mu(A_1) - \mu \left( \bigcap_{n=1}^{\infty} A_n \right) \quad (\text{this is ok since everything is finite}). \end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu \left( \bigcap_{n=1}^{\infty} A_n \right).$$

□

**Remark:** in general, the assumption “ $\mu(A_n) < \infty$  for some  $n$ ” is necessary. For example, consider the set  $S = \mathbb{R}$ ,  $A_n := ]n, \infty[$ . Consider the lebesgue measure which we will denote by  $\mu$ . Then:  $A_n$  is a decreasing sequence, but for each  $n$ :

$$\mu(]n, \infty[) = \infty \Rightarrow \lim_{n \rightarrow \infty} \mu(A_n) = \infty.$$

However, since  $A_n \downarrow \emptyset$ ,

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = 0.$$

This is a contradiction.

**Theorem 1.2.** *Given a set  $S$  and an algebra  $\Sigma_0$ . Suppose that  $\mu$  is a non-negative set function:*

$$\mu : \Sigma_0 \rightarrow [0, \infty[$$

*and  $\mu$  is finitely additive. Then,  $\mu$  is countably additive  $\iff \mu$  is continuous at  $\emptyset$ . We call this “continuity at the empty set.”*

*Proof.* To do. □

### 1.3 Existence and Uniqueness of Measure

Before proceeding, we define what it means for two measures to be equal. Given a measure space  $(S, \Sigma)$  and two measures  $\mu_1, \mu_2$ , we say that  $\mu_1 = \mu_2$  if  $\mu_1(A) = \mu_2(A)$  for all  $A \in \Sigma$ .

**Theorem 1.3.** *Given a set  $S$  and a  $\pi$ -system  $\mathcal{I}$  of subsets of  $S$ . Let  $\Sigma$  and  $\Sigma$  be the sigma algebra generated by the set  $\mathcal{I}$ . Then, if  $\mu_1(S) = \mu_2(S) < \infty$  and if  $\mu_1(A) = \mu_2(A)$  for all  $A \in \mathcal{I}$ , then  $\mu_1 = \mu_2$  on the whole sigma-algebra.*

Significance of this theorem:

1. Only  $\mathbb{R}$  valued.
2. Can extend to any  $\sigma$ -finite space.

**Theorem 1.4** (Caratheodory’s Extension Theorem). *Given a set  $S$ , suppose that  $\Sigma_0$  is an algebra and  $\mu_0 : \Sigma_0 \rightarrow [0, \infty]$  is countably additive. Then, there exists a measure defined on  $\Sigma = \sigma(\Sigma_0)$  such that  $\mu(A) = \mu_0(A)$  for all  $A \in \Sigma$ .*

We can extend a measure on  $\Sigma_0$  to  $\Sigma$ . Moreover, if  $\mu_0(S) < \infty$ , then such an extension is unique.

### 1.4 Completion of Measure / Measure Space

Sometimes it is convenient to “assume” subsets of null sets are measurable. Let  $(S, \Sigma, \mu)$  be a measurable space. Let  $N := \{A \subseteq S \mid \exists B \in \Sigma \text{ and } \mu(B) = 0 \text{ s.t. } A \subseteq B\}$ . Define:

$$\Sigma^* := \{F \subseteq S \mid \exists G, H \in \Sigma \text{ s.t. } G \subseteq F \subseteq H \text{ and } \mu(H \setminus G) = 0\} \quad (8)$$

**Proposition 1.6.**  $\Sigma^*$  is a sigma-algebra and  $\Sigma^*$  is the sigma-algebra generated by

$$\Sigma^* = \sigma(\Sigma \cup N).$$

**Proposition 1.7.** Define  $\mu^*$  to be a set function of  $\Sigma^*$  by:

$$\forall F \in \Sigma^* \text{ if } G \subseteq F \subseteq H \text{ for some } G, H \in \Sigma \text{ with } \mu(G) = \mu(H). \text{ Then, } \mu^*(F) = \mu(G) = \mu(H).$$

Then,  $\mu^*$  is a measure on  $(S, \Sigma^*)$ .

**Definition 1.10** (Complete Measure Space).  $(S, \Sigma^*, \mu^*)$  is a **complete measure space**, i.e., the completion of  $(S, \Sigma, \mu)$ .

## 2 Events and Independence

Throughout this section, assume that  $(\Sigma, \mathcal{F}, \mathbb{P})$  is a probability space.

**Definition 2.1** (Limsup and Liminf). Let  $\{A_n \mid n \geq 1\}$  be a sequence of events, i.e.,  $A_n \in \mathcal{F}$  for all  $n \geq 1$ . Then:

$$\limsup_{n \rightarrow \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m. \quad (9)$$

$$= \lim_{n \rightarrow \infty} \bigcup_{m=n}^{\infty} A_m. \quad (10)$$

$$\liminf_{n \rightarrow \infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m. \quad (11)$$

$$= \lim_{n \rightarrow \infty} \bigcap_{m=n}^{\infty} A_m. \quad (12)$$

limsup is the collection of outcomes which are in infinitely many  $A_n$ 's; the liminf is the collection of outcomes which are in every  $A_n$  after a certain cutoff.

**Remarks:**

1. Obviously,  $\liminf_n A_n \subseteq \limsup_n A_n$ . If  $\liminf_n A_n = \limsup_n A_n$ , we say that  $\lim_n A_n$  exists.
2. If  $\{A_n \mid n \in \mathbb{N}\}$  is a sequence of events and  $\{B_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$  and  $A_n \subseteq B_n$ . Then:

$$\limsup_n A_n \subseteq \limsup_n B_n$$

and

$$\limsup_n A_n \subseteq \limsup_n B_n.$$

3. **(deMorgan's Law):**  $\limsup_n A_n^c = (\liminf_n A_n)^c$  and  $\liminf_n A_n^c = (\limsup_n A_n)^c$ .
4.  $(\limsup_n A_n) \setminus (\liminf_n A_n) = \limsup_n (A_n \setminus A_{n+1})$ .
5. Let  $\{A_n \mid n \in \mathbb{N}\}$  and  $\{B_n \mid n \in \mathbb{N}\}$  be two sequences of events. Then in general:
  - (a)  $(\limsup_n A_n) \cap (\limsup_n B_n) \supseteq \limsup_n (A_n \cap B_n)$ .
  - (b)  $(\limsup_n A_n) \cup (\limsup_n B_n) = \limsup_n (A_n \cup B_n)$ .
  - (c)  $(\liminf_n A_n) \cap (\liminf_n B_n) = \liminf_n (A_n \cap B_n)$ .
  - (d)  $(\liminf_n A_n) \cup (\liminf_n B_n) \subseteq \liminf_n (A_n \cup B_n)$ .

**Theorem 2.1** (Borel-Cantelli Lemma (BC1)). If  $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$  then  $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0$ .

**Definition 2.2** (Independent). Given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , a sequence of events  $\{E_n \mid n \geq 1\} \subseteq \mathcal{F}$  is called **(mutually) independent** if  $\forall k \in \mathbb{N}, \forall 1 \leq i_1 \leq i_2 \leq \dots \leq i_k$ :

$$\mathbb{P}\left(\bigcap_{j=1}^k E_{i_j}\right) = \prod_{j=1}^k \mathbb{P}(E_{i_j})$$

**Definition 2.3.** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. Suppose that  $\{G_n \mid n \in \mathbb{N}\}$  is a sequence of  $\sigma$ -algebras (of subsets of  $\Omega$ ). Then,  $\{G_n \mid n \in \mathbb{N}\}$  is *independent* if for any choice of  $E_n \in G_n, n \geq 1$ ,  $\{E_n \mid n \geq 1\}$  is independent.

**Proposition 2.1.** Given  $\{E_n \mid n \geq 1\} \subseteq \mathcal{F}$ , we say that the collection  $\{E_n \mid n \geq 1\}$  is independent  $\iff \{\sigma(E_n) \mid n \geq 1\}$  is independent.

**Theorem 2.2.** Given  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  $\{I_n \mid n \geq 1\}$  be sequence of  $\pi$ -systems (of subsets of  $\Omega$ )  $I_n \subseteq \mathcal{F}$ ,  $\forall n \geq 1$ . Then,  $\{E_n \mid n \geq 1\}$  is independent  $\iff \{\sigma(I_n) \mid n \geq 1\}$  is independent.

**Theorem 2.3** (Borel-Cantelli Lemma (BC2)). Given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  let  $\{E_n \mid n \in \mathbb{N}\}$  be an independent sequence of events if:

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty, \text{ then } \mathbb{P}\left(\limsup_n E_n\right) = 1. \quad (13)$$

**Definition 2.4** (Tail Sigma-Algebra). Given  $(\Omega, \mathcal{F}, \mathbb{P})$  and a sequence of events  $\{E_n \mid n \geq 1\}$ . Define:

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(\{E_n, E_{n+1}, E_{n+2}, \dots\}). \quad (14)$$

Then,  $\mathcal{T}$  is defined to be the **tail sigma algebra** associated with  $\{E_n\}$ . If an event  $A \in \mathcal{T}$ , then  $A$  is a **tail event** with respect to  $\{E_n \mid n \in \mathbb{N}\}$ .

**Theorem 2.4** (Kolmogorov's 0-1 Law). If  $\{E_n \mid n \geq 1\}$  is independent, and  $\mathcal{T}$  is the tail sigma-algebra associated with  $E_n$  then for all  $A \in \mathcal{T}$ , either  $\mathbb{P}(A) = 0$  or  $\mathbb{P}(A) = 1$ .

### 3 Random Variables

**Definition 3.1** (Measurable). Let  $(S, \Sigma)$  be a measurable space and  $h : S \rightarrow \mathbb{R}$  (in certain situations, could be  $\overline{\mathbb{R}}$ ) be a function. We say that  $h$  is  $\Sigma$ -**measurable**, denoted by  $h \in m\Sigma$  if:  $\forall B \in \mathcal{B}(\mathbb{R})$ , the pre-image of  $B$  under  $h$  is measurable, i.e.,  $h^{-1}(B) \in \Sigma$ .

Remarks:

1. If  $h$  is a measurable function,  $h \in m\Sigma$ , then  $\{h = \infty\} = \{s \in S \mid h(s) = +\infty\} \in \Sigma$  and  $\{h = -\infty\} = \{s \in S \mid h(s) = -\infty\} \in \Sigma$ .
2. More generally,  $h : (S_1, \Sigma_1) \rightarrow (S_2, \Sigma_2)$ , then we say that  $h$  is  $\Sigma_1 \setminus \Sigma_2$ -measurable if  $\forall B \in \Sigma_2$ ,  $h^{-1}(B) \in \Sigma_1$ .
3. For all  $A \subseteq \mathbb{R}$ ,  $h^{-1}(A^c) = (h^{-1}(A))^c$ . Moreover, for all  $A_\alpha \in \mathbb{R}$ , where  $\alpha \in I$ ,

$$h^{-1}\left(\bigcup_{\alpha \in I} A_\alpha\right) = \bigcup_{\alpha \in I} h^{-1}(A_\alpha),$$

$$h^{-1}\left(\bigcap_{\alpha \in I} A_\alpha\right) = \bigcap_{\alpha \in I} h^{-1}(A_\alpha).$$

4. Suppose  $\mathcal{C} \subseteq \mathcal{B}(\mathbb{R})$  and  $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$ . Then, a function is measurable  $\iff \forall c \in \mathcal{C}$ ,  $h^{-1}(c) \in \Sigma$ .
5. Given  $h : S \rightarrow \mathbb{R}$  measurable,  $f : \mathbb{R} \rightarrow \mathbb{R}$  a Borel function, then  $(f \circ h) \in m\Sigma$ .
6. Given  $h_1, h_2 \in m\Sigma$ ,  $h_1 + h_2$ ,  $h_1 - h_2$ ,  $h_1 \cdot h_2$ ,  $\frac{h_1}{h_2}$  (where  $h_2 \neq 0$ ), ...,  $\in m\Sigma$ .
7. Given  $\{h_n \mid n \in \mathbb{N}\} \subseteq m\Sigma$ , we have:

$$\inf_n h_n, \sup_n h_n, \liminf_n h_n, \limsup_n h_n \in m\Sigma.$$

**Definition 3.2** (Random Variable). Consider a random probability space  $(\Sigma, \mathcal{F}, \mathbb{P})$ .  $X : \Sigma \rightarrow \mathbb{R}$  is a **Random Variable** if  $X$  is  $\mathcal{F}$ -measurable, i.e.,  $X \in m\mathcal{F}$ .



**Definition 3.3** (Sigma Algebra Generated by  $X$ ). Let  $X$  be a random variable. The  $\sigma$ -algebra generated by  $X$ , denoted by  $\sigma(X)$ , is:

$$\sigma(X) := \{X^{-1}(\mathcal{B}) \mid \mathcal{B} \in \mathcal{B}(\mathbb{R})\}. \quad (15)$$

I.e.:  $\sigma(X)$  is the smallest  $\sigma$ -algebra with respect to which the random variable  $X$  is measurable.

**Remarks:**

1. Given  $(\Sigma, \mathcal{F}, \mathbb{P})$  and  $X : \Omega \rightarrow \mathbb{R}$ . Then,  $X$  is a random variable  $\iff \forall a \in \mathbb{R} \{X \leq a\} \in \mathcal{F}$ .
2. Let  $\{X_\alpha \mid \alpha \in I\}$  be a family of random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Then, the  $\sigma$ -algebra generated by  $\{X_\alpha \mid \alpha \in I\}$  is

$$\sigma(\{X_\alpha \mid \alpha \in I\}) = \sigma(\{X_\alpha^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}), \alpha \in I\}).$$

3. Let  $\{X_n \mid n \in \mathbb{N}\}$  be a sequence of random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Set:

$$\mathcal{P} := \left\{ \bigcap_{j=1}^k \{x_{n_j} \leq a_j\} \right\}.$$

Then,  $\mathcal{P}$  is a  $\pi$ -system and  $\sigma(\mathcal{P}) = \sigma(\{X_n \mid n \in \mathbb{N}\})$ .

**Definition 3.4** (Independent Random Variables). Given  $(\Omega, \mathcal{F}, \mathbb{P})$  and a sequence of random variables  $\{X_n \mid n \in \mathbb{N}\}$ ,  $\{X_n \mid n \in \mathbb{N}\}$  is **independent** if  $\{\sigma(X_n) \mid n \in \mathbb{N}\}$  is independent.

**Proposition 3.1.**  $\{X_n \mid n \in \mathbb{N}\}$  are independent  $\iff \forall k \geq 1, \forall 1 \leq n_1 \leq n_2 \leq \dots \leq n_k, \forall a_1, a_2, \dots, a_k \in \mathbb{R}$ :

$$\mathbb{P} \left( \bigcap_{j=1}^k \{X_{n_j} \leq a_j\} \right) = \prod_{j=1}^k \mathbb{P}(X_{n_j} \leq a_j). \quad (16)$$

**Definition 3.5** (Tail Sigma Algebra). Let  $\{X_n \mid n \in \mathbb{N}\}$  be a sequence of random variables. The **tail sigma-algebra** associated with  $\{X_n \mid n \in \mathbb{N}\}$  is defined as:

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(\{X_n, X_{n+1}, \dots\}) \quad (17)$$

**Theorem 3.1** (Kolmogorov's 0-1 Law). Let  $\{X_n \mid n \in \mathbb{N}\}$  be a sequence of independent random variables and let  $\mathcal{T}$  be the tail sigma-algebra associated with  $\{X_n \mid n \in \mathbb{N}\}$ . Then, for all  $A \in \mathcal{T}$ :

$$\mathbb{P}(A) \in \{0, 1\} \quad (18)$$

If  $X \in m\mathcal{T}$ , then  $X$  is constant a.s., i.e.,  $\exists$  an  $a \in \bar{\mathbb{R}}$  such that  $\mathbb{P}(X = a) = 1$ .

## 4 Distribution Functions

**Definition 4.1** (Law/Distribution). Given  $(\Omega, \mathcal{F}, \mathbb{P})$  and a random variable  $X : \Omega \rightarrow \mathbb{R}$ . The **law/distribution** of  $X$ , denoted by  $\mathcal{L}_X$ , is the probability distribution on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  such that for all  $B \in \mathcal{B}(\mathbb{R})$ :

$$\mathcal{L}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B). \quad (19)$$

The **distribution function** of  $X$  (of  $\mathcal{L}_X$ ) is:

$$\begin{aligned} F_x : \mathbb{R} &\rightarrow [0, 1] \\ x \in \mathbb{R} &\mapsto F_X(x) := \mathbb{P}(X \leq x) = \mathcal{L}_X([-\infty, x]) \end{aligned}$$

### Remarks

1.  $F_X$  is increasing.
2.  $\lim_{x \rightarrow +\infty} F_X(x) = 1$ .
3.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$ .
4.  $F_X$  is **right continuous**:  $\forall a \in \mathbb{R}, F_X(a^+) = \lim_{x \rightarrow a^+} F_X(x) = F_X(a)$ .
5. For all  $a > b$ ,  $F_X(a) - F_X(b) = \mathbb{P}(b < X < a) = \mathcal{L}_X(]b, a])$ .

**Definition 4.2** (Independent and Identically Distributed). Let  $\{X_n \mid n \in \mathbb{N}\}$  be a sequence of random variables. We say they are **independent and identically distributed (iid)** if  $\{X_n \mid n \in \mathbb{N}\}$  is independent and for some probability measure  $\mu$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ,  $\mathcal{L}_{X_n} = \mu$  for all  $n \in \mathbb{N}$ .