Notes on Probability Theory

Kai-Jyun Wang*

Fall 2025

The notes are based on the lecture of Prof. David Anderson at University of Wisconsin-Madison in fall 2025. The course structure mainly follows Durrett.

*National Taiwan University, Department of Economics.

Contents

1. Probability Space and Random Variable

1.1. Probability Space

Definition 1.1
*Let Ω be a set. A collection of subsets $\mathcal{F}$ forms a σ-**algebra** if*

(a) $\varnothing \in \mathcal{F}$.

(b) $A \in \mathcal{F}$ *implies* $A^c \in \mathcal{F}$.

(c) If $A_i \in \mathcal{F}$ *are countably many sets,* $\cup_i A_i \in \mathcal{F}$.

The dual $(\Omega, \mathcal{F})$ *is called a **measurable space** and the sets falling in* $\mathcal{F}$ *are said to be **measurable**.*

Definition 1.2
Let $(\Omega, \mathcal{F})$ *be a measurable space. A set function* $\mu : \mathcal{F} \to [0, \infty]$ *is a **measure** if*

(a) $\mu(\varnothing) = 0$.

(b) For countably many disjoint $A_i \in \mathcal{F}$, $\mu(\cup_i A_i) = \sum_i \mu(A_i)$.

The triple $(\Omega, \mathcal{F}, \mu)$ *is called a **measure space**.*

Definition 1.3
*A **probability space** is a measure space* $(\Omega, \mathcal{F}, \mathrm{P})$ *such that* $\mathrm{P}(\Omega) = 1$.

Lemma 1.4
Let $\mathcal{S}$ *be a collection of sets. Then there exists the smallest* σ*-algebra containing* $\mathcal{S}$.

Proof. Let $\mathcal{F}$ be the intersection of all σ-algebra containing $\mathcal{S}$. $\mathcal{F}$ is non-empty since the power set is a σ-algebra containing $\mathcal{S}$. Now it is clear that $\varnothing \in \mathcal{F}$ since $\varnothing \in \mathcal{A}$ for every σ-algebra $\mathcal{A}$ contiaining $\mathcal{S}$. If $A \in \mathcal{F}$, $A \in \mathcal{A}$ for all $\mathcal{A}$ containing $\mathcal{S}$ and $A^c \in \mathcal{A}$ for all $\mathcal{A}$. Thus $A^c \in \mathcal{F}$. Finally, if $A_i \in \mathcal{F}$ are countably many sets, then each A_i lies in every $\mathcal{A}$ containing $\mathcal{S}$; so does $\cup_i A_i$ and thus $\cup_i A_i \in \mathcal{F}$. The minimality follows by the construction of $\mathcal{F}$. ∎

Definition 1.5
For any collection of sets $\mathcal{S}$*, the smallest* σ*-algebra is denoted as* $\sigma(\mathcal{S})$.

Theorem 1.6
Let $(\Omega, \mathcal{F}, \mathrm{P})$ *be a probability space. Then*

(a) If $A, B \in \mathcal{F}$ *and* $A \subset B$*, then* $\mathrm{P}(A) \le \mathrm{P}(B)$.

(b) For countably many $A_i \in \mathcal{F}$, $\mathrm{P}(\cup_i A_i) \le \sum_i \mathrm{P}(A_i)$.

(c) If $A_i \nearrow A$, $\mathrm{P}(A_i) \to \mathrm{P}(A)$.

(d) If $A_i \searrow A$, $\mathrm{P}(A_i) \to \mathrm{P}(A)$.

Proof. (a) and (b) are clear. For (c), write $E_i = A_i - A_{i-1}$ and $A_0 = \varnothing$. Then since E_i are disjoint and $A_n = \cup_{i=1}^n E_i$,

$$\mathrm{P}(A_n) = \mathrm{P}\big(\cup_{i=1}^n E_i\big) = \sum_{i=1}^n \mathrm{P}(E_i) \to \sum_i \mathrm{P}(E_i) = \mathrm{P}(\cup_i E_i) = \mathrm{P}(A)$$

as $n \to \infty$.

For (d), note that $A_i^c \nearrow A^c$. Thus $1 - \mathrm{P}(A_i) = \mathrm{P}(A_i^c) \to \mathrm{P}(A^c) = 1 - \mathrm{P}(A)$. Thus $\mathrm{P}(A_i) \to \mathrm{P}(A)$. ■

Definition 1.7

*The **Borel σ-algebra** is the σ-algebra generated by all open sets.*

Definition 1.8

*Let P be a probability measure on $(\mathbb{R}, \mathcal{B})$. The **distribution function** F is defined as*

$$F(x) = \mathrm{P}((-\infty, x])$$

for $x \in \mathbb{R}$.

Proposition 1.9

The distribution function in $(\mathbb{R}, \mathcal{B})$ satisfies that

(a) $F(x) \le F(y)$ *for all* $x \le y$.

(b) $F(x) \to F(y)$ *as* $x \to y^+$.

(c) $F(-\infty) = 0$ *and* $F(\infty) = 1$.

Proof. For (a), note that $(-\infty, x] \subset (-\infty, y]$ and

$$F(x) = \mathrm{P}((-\infty, x]) \le \mathrm{P}((-\infty, y]) = F(y).$$

For (b), notice that for $x_n \to y^+$, $(-\infty, x_n] \searrow (-\infty, y]$. Hence

$$F(x_n) = \mathrm{P}((-\infty, x_n]) \to \mathrm{P}((-\infty, y]) = F(y).$$

Similarly, taking $x_n \to \pm\infty$ gives (c). ■

Definition 1.10

*A collection $\mathcal{S}$ of sets is called an **algebra** if*

(a) $\varnothing \in \mathcal{S}$.

(b) If $A \in \mathcal{S}$, *then* $A^c \in \mathcal{S}$.

(c) If $A, B \in \mathcal{S}$, *then* $A \cup B \in \mathcal{S}$.

Remark

An algebra is closed under finite unions. It is also clear that a σ-algebra is an algebra, while the converse is not true. An example is the collection of all finite unions of intervals in $\mathbb{R}$.

Definition 1.11
*A collection $\mathcal{S}$ of sets is called a **semi-algebra** if*

(a) *If $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$.*

(b) *If $A \in \mathcal{S}$, then A^c can be written as a finite disjoint union of sets in $\mathcal{S}$.*

Remark
A semi-algebra must contain $\varnothing$ since for any $A \in \mathcal{S}$, $A^c = \cup_i A_i$, where $A_i \in \mathcal{S}$ are disjoint. Then $A \cap A_1 = \varnothing \in \mathcal{S}$.

Remark
An example of being a semi-algebra but not an algebra is the collection of all intervals of the form $(a_i, b_i]$ for $-\infty \leq a_i < b_i \leq \infty$ with the empty set.

Lemma 1.12
If $\mathcal{S}$ is a semi-algebra, then $\overline{\mathcal{S}} = \{$finite disjoint unions of sets in $\mathcal{S}\}$ forms an algebra.

Proof. It has been shown that $\varnothing \in \mathcal{S}$. For $A, B \in \overline{\mathcal{S}}$, write $A = \cup_{i=1}^n A_i$ and $B = \cup_{j=1}^m B_j$ for disjoint $A_i, B_j \in \mathcal{S}$, respectively. Then $A \cap B = \cup_{i,j}(A_i \cap B_j) \in \overline{\mathcal{S}}$. Thus $\overline{\mathcal{S}}$ is closed under intersection. Now if $A \in \overline{\mathcal{S}}$, $A = \cup_{i=1}^n A_i$ for disjoint $A_i \in \mathcal{S}$. Then $A^c = \cap_{i=1}^n A_i^c$. By the definition of semi-algebra, A_i^c can be written as a finite disjoint union of sets in $\mathcal{S}$ and thus $A_i^c \in \overline{\mathcal{S}}$. Since $\overline{\mathcal{S}}$ is closed under finite intersection, $A^c = \cap_{i=1}^n A_i^c \in \overline{\mathcal{S}}$. Finally, for $A, B \in \overline{\mathcal{S}}$, $A \cup B = (A^c \cap B^c)^c \in \overline{\mathcal{S}}$. We conclude that $\overline{\mathcal{S}}$ is indeed an algebra. ∎

Definition 1.13
*Suppose $\mathcal{S}$ is a semi-algebra. $\overline{\mathcal{S}} = \{$finite disjoint unions of sets in $\mathcal{S}\}$ is called the **algebra generated by** $\mathcal{S}$.*

Definition 1.14
*Let $\mathcal{S}$ be an algebra. A set function $\mu_0 : \mathcal{S} \to [0, \infty]$ is called a **premeasure** if*

(a) *$\mu_0(\varnothing) = 0$.*

(b) *For countable disjoint $A_i \in \mathcal{S}$ such that $\cup_i A_i \in \mathcal{S}$,*

$$\mu_0(\cup_i A_i) = \sum_i \mu_0(A_i).$$

Theorem 1.15
Let ν be a set function on a semi-algebra $\mathcal{S}$ such that $\nu(\varnothing) = 0$. Suppose that

(a) *if $A \in \mathcal{S}$ and $A = \cup_{i=1}^n A_i$ for disjoint $A_i \in \mathcal{S}$, then $\nu(A) = \sum_{i=1}^n \nu(A_i)$;*

(b) *if $A_i \in \mathcal{S}$ are countably many sets and $A = \cup_i A_i \in \mathcal{S}$, then $\nu(A) \leq \sum_i \nu(A_i)$.*

Then ν can be extended to a unique premeasure μ_0 on the algebra generated by $\mathcal{S}$.

Proof. We first show the existence. From lemma 1.12 we know that $\mathcal{S}$ generates an algebra $\mathcal{A} = \{$finite disjoint union of sets in $\mathcal{S}\}$. Define our candidate μ_0 by $\mu_0(A) = \sum_i \nu(A_i)$ for

$A = \cup_i A_i$ where A_i are disjoint. To see that μ_0 is well-defined, suppose $A = \cup_i B_i$ for disjoint $B_i \in \mathcal{S}$. Observe that

$$A_i = \cup_j (A_i \cap B_j) \quad \text{and} \quad B_j = \cup_i (A_i \cap B_j)$$

are finite disjoint unions. Then

$$\sum_i \nu(A_i) = \sum_i \sum_j \nu(A_i \cap B_j) = \sum_j \sum_i \nu(A_i \cap B_j) = \sum_j \nu(B_j)$$

by (a). Thus μ_0 is well-defined.

Now we check that μ_0 is a premeasure. Clearly $\mu_0(\varnothing) = 0$. For countably many disjoint $A_i \in \mathcal{A}$, if $A = \cup_i A_i \in \mathcal{A}$, ■

Theorem 1.16

If F is non-decreasing, right-continuous and satisfies that $F(-\infty) = 0$, $F(\infty) = 1$, then there is a probability measure such that

$$\mathrm{P}((-\infty, x]) = F(x).$$

1.2. Random Variable