STAT 547 LECTURE NOTES

PROBABILITY THEORY

UNIVERSITY OF BRITISH COLUMBIA

Jiajun ZHANG
Department of Statistics
SEPTEMBER 2026

Contents

1 Basics of Measure Theory

1.1 Preliminaries

We assume E is a state space, let $\{A_i\}_{i\in\mathcal{I}} \subset E$ be a sequence of sets in E, then the **intersection** and **union** of those sets are defined as

$$\bigcup_{i\in\mathcal{I}} A_i := \{x \in E, \exists i : x \in A_i\}; \quad \bigcap_{i\in\mathcal{I}} A_i := \{x \in E, \forall A_i, x \in A_i\}.$$

The union can be understood as "*there exists*", and the intersection can be understood as "*for all*". The **compliment** of A, denote as A^C, is defined by $A^C := E \setminus A$. The DeMorgan's law on sets states that if we have a collection of sets $A_i \in \mathcal{C}$, then

$$\left(\bigcup_{A\in\mathcal{C}} A\right)^C = \bigcap_{A\in\mathcal{C}} A^C; \quad \left(\bigcap_{A\in\mathcal{C}} A\right)^C = \bigcup_{A\in\mathcal{C}} A^C.$$

Definition 1. Given a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, the **limsup** and **liminf** of that sequence are defiend by

$$\limsup_n x_n := \lim_{n\to\infty} \sup_{k\geq n} x_k; \quad \liminf_n x_n = \lim_{n\to\infty} \inf_{k\geq n} x_k.$$

Note that $\sup_{k\geq n} x_k$ is decreasing as a function of n, hence limsup can also be written as $\limsup_n x_n = \inf_n \sup_{k\geq n} x_k$; Similarly for liminf, since $\inf_{k\geq n} x_k$ is increasing as a function of n, we then have $\liminf_n x_n = \sup_n \inf_{k\geq n} x_k$. The limit of $\lim_n x_n$ exists if and only if

$$\limsup_n x_n = \liminf_x x_n = \lim_n x_n.$$

The definition of limsup and liminf can be extended to sets:

Definition 2. Let $\{A_n\}_{n\in\mathbb{N}} \subset E$ be a sequence of sets, then the limsup and liminf of that sequence are defined by

$$\limsup_n A_n = \bigcap_n \bigcup_{k\geq n} A_k; \quad \liminf_n A_n = \bigcup_n \bigcap_{k\geq n} A_k.$$

We can interpret them as follow:

- $x \in \limsup_n A_n$ if and only if $\forall n \in \mathbb{N}$, $\exists k \geq n$, such that $x \in A_k$. We say $x \in A_k$ *infinitely often*;
- $x \in \liminf_n A_n$ if and only if there $\exists n \in \mathbb{N}$, such that $x \in A_k$ for all $k \geq n$. We say $x \in A_k$ *eventually*.

Similarly, $\lim_n A_n$ exists if and only if

$$\limsup_n A_n = \liminf_n A_n = \lim_n A_n.$$

By using DeMorgan's law, we have

$$\left(\limsup_n A_n\right)^C = \left(\bigcap_n \bigcup_{k\geq n} A_k\right)^C = \bigcup_n \left(\bigcup_{k\geq n} A_k\right)^C = \bigcup_n \bigcap_{k\geq n} A_k^C := \liminf_n A_n^C,$$

and, we can play around with other notations.

Proposition 1. $x \in \limsup_n A_n$ if and only if $\limsup_n \mathbb{1}\{x \in A_n\} = 1$; and $x \in \liminf_n A_n$ if and only if $\limsup_n \mathbb{1}\{x \in A_n\} = 1$.

Proof. We first assume $x \in \limsup_n A_n$, which means $\forall n$, $x \in \bigcup_{k \geq n} A_k$. Hence we may construct a subsequence $\{A_{n_k}\}_{k=1}^{\infty}$ where $n_k \in \mathbb{N}$ such that $x \in A_{n_k}, \forall k$, which implies $\limsup_n \mathbb{1}\{x \in A_n\} = 1$. The other direction follows by reversing the steps. The proof for liminf follows the same procedure. □

Proposition 2. We say a sequence $\{A_n\}_{n\in\mathbb{N}}$ is **monotonically increasing** if $A_1 \subset A_2 \subset \cdots$, in this case $\lim_n A_n$ exists and

$$\lim_n A_n = \bigcup_n A_n.$$

Proof. We have that

$$\limsup_n A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k = \bigcap_{n\geq 1} \bigcup_{k\geq 1} A_k = \bigcup_n A_n,$$

since $\{A_n\}$ is increasing. Also,

$$\liminf_n A_n = \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k = \bigcup_{n\geq 1} A_n,$$

since $\{A_n\}$ is decreasing. □

In fact by using DeMorgan's law $\bigcup A_n^C = \bigcap A_n$, where $\{A_n\}$ is monotonically increasing, we may further conclude that if $\{A_n\}_{n=1}^{\infty}$ is **monotonically decreasing** (meaning $A_1 \supset A_2 \supset \cdots$), then $\lim_n A_n$ exists and $\lim_n A_n = \bigcap_n A_n$.

1.2 σ-algebras

Definition 3. Let E be a non-empty set, $\mathcal{E}$ be the collection of subsets of E, $\mathcal{E}$ is called a **σ-algebra** on E if the following holds:

- $E \in \mathcal{E}$;
- If $A \in \mathcal{E}$, then $A^C = E \setminus A \in \mathcal{E}$ (it is closed under compliments);
- If $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{E}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{E}$ (it is closed under countable unions).

We denote $(E, \mathcal{E})$ to be a **measure space**.

From the definition above, some properties of σ-algebra follow immediately:

- $\varnothing \in \mathcal{E}$: This is because $E^C \in \mathcal{E}$;
- $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{E}$: This is because $\left(\bigcup_{n\in\mathbb{N}} A_n^C\right)^C \in \mathcal{E}$;
- It is also closed under finite unions and intersections.

Below are some examples of σ-algebras:

- **The trivial σ-algebra**: For any non-empty set E, $\mathcal{E} = \{E, \varnothing\}$ is a σ-algebra (the smallest σ-algebra on E);

- For any non-empty set E, its power set 2^E is also a σ-algebra (the biggest σ-algebra on E);
- Given a collection of sets $\mathcal{C}$, denote $\sigma(\mathcal{C})$ as the *smallest* σ-algebra that contains $\mathcal{C}$. We denote $\sigma(\mathcal{C})$ as the σ-algebra **generated by $\mathcal{C}$**;
- Let E be the state space, $A, B \subset E$ be sets. Then

$$\sigma(\{A, B\}) = \{\varnothing, E, A, B, A^C, B^C, A \cap B, A \cup B\};$$

- Given two measure space $(E, \mathcal{E}), (F, \mathcal{F})$, the product space $E \times F := (e, f), e \in E, f \in F$, and

$$\mathcal{E} \otimes \mathcal{F} := \sigma\{A \times B : A \in \mathcal{E}, B \in \mathcal{F}\}.$$

Definition 4. A collection $\mathcal{C}$ on E is called a **p-system** if $A, B \in \mathcal{C}$ implies $A \cap B \in \mathcal{C}$.

Sometimes p-system is also referred as π-system.

Definition 5. A collection $\mathcal{D}$ on E is called a **d-system** if it satisfies the following:

- $E \in \mathcal{D}$;
- If $A, B \in \mathcal{D}$ and $A \subset B$, then $B \setminus A \in \mathcal{D}$;
- If $A_n \nearrow A$ ($\{A_n\}$ is non-decreasing with limit A), then $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Proposition 3. A collection $\mathcal{D}$ on E is a σ-algebra on E if and only if it is both a p-system and a d-system on E.

Proof. ($\Longrightarrow$): Suppose $\mathcal{D}$ is a σ-algebra on E, then trivially it is a p-system since $(A \cap B = A^C \cup B^C)^C \in \mathcal{D}$. Since $E \in \mathcal{D}$, take any $A \in E$, $A \subset E$, so $E \setminus A \in \mathcal{D}$. Finally consider an arbitrary sequence $\{A_n\}_{n=1}^{\infty} \subset \mathcal{D}$, define $B_n = \bigcup_{i=1}^{n} A_i$, then $B_n \nearrow B$ where $B = \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

($\Longleftarrow$): Suppose $\mathcal{D}$ is both a p-system and a d-system, then we already have $E \in \mathcal{D}$. If $A \in \mathcal{D}$, $A \subset E$, then $E \setminus A = A^C \in \mathcal{D}$. Finally let $\{A_n\}_{n=1}^{\infty}$ be an arbitrary sequence, let $B_1 = A_1, B_2 = B_1 \cup (A_2 \setminus B_1), \cdots, B_n = B_{n-1} \cup (A_n \setminus B_{n-1})$. Then $\{B_n\}_{n=1}^{\infty}$ is increasing, and

$$\lim_{n \to \infty} B_n = \lim_{n \to \infty} B_{n-1} \cup (A_n \setminus B_{n-1}) \in \mathcal{D},$$

and $\bigcup B_n = \bigcup A_n \in \mathcal{D}$. □

Theorem 1 (Monotone class theorem). If $\mathcal{C}$ is a p-system, $\mathcal{D}$ is a d-system and $\mathcal{C} \subset \mathcal{D}$, then $\sigma(\mathcal{C}) \subset \mathcal{D}$.

We need two lemmas to establish this proof.

Lemma 1. Let $\mathcal{D}$ be a d-system on E, and fix $D \in \mathcal{D}$. Deifne

$$\widehat{\mathcal{D}} = \{A \in \mathcal{D} : A \cap D \in \mathcal{D}\},$$

then $\widehat{\mathcal{D}}$ is a d-system on E.

Proof. It is easy to see that $E \in \mathcal{D}, E \cap D = D \in \mathcal{D}$, so $E \in \widehat{D}$; Let $A_1, A_2 \in \widehat{D}$, $A_1 \subset A_2$, then we have $A_1 \cap D \in \mathcal{D}, A_2 \cap D \in \mathcal{D}, A_1 \cap D \subset A_2 \cap D$, and $(A_2 \setminus A_1) \cap D = (A_2 \cap D) \setminus (A_1 \cap D) \in \widehat{D}$. Let $A_n \nearrow A$, where $A_n \in \widehat{D}$, then $A_n \cap D \nearrow A \cap D$, and

$$A \cap D = \bigcup_{n=1}^{\infty} A_n \cap D \in \widehat{D}.$$

□

Lemma 2. If $\mathcal{D}_1, \mathcal{D}_2$ are two d-systems on E, then so is $\mathcal{D}_1 \cap \mathcal{D}_2$.

We shall omit this proof since it is trivial.

We now prove the monotone class theorem:

Proof. Let $\mathcal{D}'$ be the smallest d-system containing $\mathcal{C}$, i.e, $\mathcal{C} \subset \mathcal{D}'$, then we have $\sigma(\mathcal{C}) \subset \sigma(\mathcal{D}')$. Hence it suffices to show $\sigma(\mathcal{D}') = \mathcal{D}'$. If $\mathcal{D}'$ is a σ-algebra, then clearly $\sigma(\mathcal{D}') = \mathcal{D}'$. Hence by proposition 3, it remains to prove $\mathcal{D}'$ is a p-system. Fix $B \in \mathcal{C}$, and define

$$\mathcal{D}'_1 = \{A \in \mathcal{D}' : A \cap B \in \mathcal{D}'\},$$

then by construction and the previous lemma, $\mathcal{D}'_1$ is a d-system, hence $\subset \mathcal{D}' \subset D'_1$, which implies $\forall A \in \mathcal{D}', B \in \mathcal{C}, A \cap B \in \mathcal{D}'$. Now fix $A \in \mathcal{D}'$, define

$$\mathcal{D}'_2 := \{B \in \mathcal{D}' A \cap B \in \mathcal{D}'\},$$

which is again a d-system, and $\mathcal{C} \subset D'_2$, which means $\mathcal{D}' \subset \mathcal{D}'_2$. Now we have that $\forall A, B \in \mathcal{D}'$, $A \cap B \in \mathcal{D}'$, so $\mathcal{D}'$ is a p-system. □

Next we introduce one important σ-algebra on real numbers $\mathbb{R}$:

Definition 6. The **Borel σ-algebra**, denote as $\mathfrak{B}(\mathbb{R})$, is defined as

$$\mathfrak{B}(\mathbb{R}) := \sigma\{\text{all open sets in } \mathbb{R}\}.$$

In fact, it can be shown that $\mathfrak{B}(\mathbb{R})$ is also equal to $\sigma((a,b)); \sigma([a,b]), \sigma([a,b)), \sigma((a,b])$, etc. Open sets are the elements inside a **topology** $\mathcal{T}$. The definition of a topology follows.

Definition 7. We say $\mathcal{T}$ is a topology over S, if

- $\varnothing, S \in \mathcal{T}$;
- If $A, B \in \mathcal{T}$, then $A \cap B \in \mathcal{T}$;
- Given an arbitrary collection $A_i \in \mathcal{T}, i \in \mathcal{I}$, then $\bigcup_{i \in \mathcal{I}} A_i \in \mathcal{T}$.

If $S = \mathbb{R}$, then

$$\mathcal{T} := \left\{ \bigcup_{i \in \mathcal{I}} (a_i, b_i); a_i < b_i; a_i, b_i \in \mathbb{R} \right\} \cup \{\varnothing\}.$$

1.3 Measurable Functions

A function $f : X \to Y$ is a mapping such that for any $x \in X$, there is a unique $y \in Y$ such that $f(x) = y$. We now give the definition of a measurable function:

Definition 8. Given two measurable spaces $(E, \mathcal{E}), (F, \mathcal{F})$, the function $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ is **measurable** if for any $B \in \mathcal{F}$, its pre-image is in $\mathcal{E}$, i.e,

$$f^{-1}(B) = \{x \in E : f(x) \in B\} \in \mathcal{E}.$$

We often work with real valued functions $f : \mathbb{R} \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ is the extended real line ($\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$). In this setting, we consider $[-\infty, a), (b, \infty]$ as open sets. The σ-algebra we consider is simply the Borel σ-algebra $\mathfrak{B}(\mathbb{R})$.

Proposition 4. The function $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ is measurable if and only if there exists a collection $\mathcal{F}_0$ on F with $\sigma(\mathcal{F}_0) = \mathcal{F}$, and $f^{-1}(B) \in \mathcal{E}$ for all $B \in \mathcal{F}_0$.

Proof. $\Longrightarrow$ is trivial. We consider a collection $\mathcal{F}_0$ of subsets of F with $\sigma(\mathcal{F}_0) = \mathcal{F}$. Denote

$$\mathcal{F}_1 := \{B \in \mathcal{F}, f^{-1}(B) \in \mathcal{E}\},$$

then it is straightforward to show that $\mathcal{F}_1$ is a σ-algebra on F_1, hence $\mathcal{F}_1 = \sigma(\mathcal{F}_1) \supset \sigma(\mathcal{F}_0) = \mathcal{F}$. □

Proposition 5. The function $f : \mathbb{R} \to \overline{\mathbb{R}}$ is measurable on $\mathfrak{B}(\mathbb{R})$ if and only if for all $r \in \mathbb{R}$,

$$f^{-1}([-\infty, r]) := \{x \in \mathbb{R} : f(x) \leq r\} \in \mathfrak{B}(\mathbb{R}).$$

Proof. $\Longrightarrow$ is trivial. Using Proposition 4, it suffices to show the Borel σ-algebra can also be generated by sets of the form $(a, b]$. For any $r \in \mathbb{R}$, we have that

$$[-\infty, r] = \bigcup_{n=1}^{\infty} \left[-\infty, r + \frac{1}{n}\right) \in \mathfrak{B}(\mathbb{R}).$$

Recall in our setting $[-\infty, r)$ is open. □

Measurable functions have the following property as well:

Proposition 6. Let f, g be measurable functions on $(E, \mathcal{E})$, $\{f_n\}$ be a sequence of measurable functions on $(E, \mathcal{E})$. Then

- For any $a, b \in \mathbb{R}$, $af + bg$; fg; $f \circ g$; $f/g, g \neq 0$ are measurable;
- $\inf_n f_n$; $\sup_n f_n$; $\limsup_n f_n$; $\liminf_n f_n$ are measurable;

We only give the proof for a few cases.

Proof. For $af + bg$, consider any $r \in \mathbb{R}$,

$$\begin{aligned}
(af + bg)^{-1}([-\infty, r]) &= \{x \in E : af(x) + bg(x) \leq r\} \\
&= \{x \in E : af(x) \leq r - bg(x)\} \\
&= \bigcup_{q \in \mathbb{Q}} \{x \in \mathbb{R} : af(x) \leq q \leq r - bg(x)\} \\
&= \bigcup_{q \in \mathbb{Q}} \left(\{x \in E : af(x) \leq q\} \bigcap \{x \in E : r - bg(x) \geq q\} \right) \in \mathcal{E}.
\end{aligned}$$

□

Proof. For $\sup_n f_n$, take any $r \in \mathbb{R}$,

$$\begin{aligned}
\sup_n f_n^{-1}([-\infty, r]) &:= \{x \in E : \sup_n f_n(x) \leq r\} \\
&= \left\{ x \in E : \bigcap_{i=1}^{n} f_i(x) \leq r \right\} \\
&= \bigcap_{i=1}^{n} \{x \in E : f_i(x) \leq r\} \in \mathcal{E}.
\end{aligned}$$

Using the fact that $\inf_n f_n(x) = -\sup_n(-f_n(x))$, one can easily show that $\sup_n f_n$ is also measurable. Then $\liminf_n f_n, \limsup_n f_n$ is naturally the limit of $\inf_n f_n, \sup_n f_n$. $\square$

We now introduce a class of measurable functions called **simple functions**:

Definition 9 (Simple Functions). A function $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ is said to be simple, if it takes the form

$$f(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i},$$

where $a_1, \cdots, a_n \in \mathbb{R}$, $A_i \subset \mathcal{E}$.

The $\mathbb{1}_A(\cdot)$ is the **indicator function** on set A, it is defined by

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}.$$

Clearly, for a simple function f, there exists $b_1, \cdots, b_m \in \mathbb{R}$ and $B_1, \cdots, B_m \in \mathcal{E}$ a partition of $\mathcal{E}$, such that $f(x) = \sum_{j=1}^{m} b_j \mathbb{1}_{B_j}(x)$. It is called the **canonical form** of the simple function f.

The reason for introducing simple functions is because it can be used to approximate measurable functions.

Theorem 2. A positive function $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ is measurable if and only if it is the limit of an increasing sequence of positive simple functions.

Proof. ($\Longleftarrow$) is trivial. We start by introducing a sequence of function $\{d_n\}$: For any $x \in \overline{\mathbb{R}}^+$, define

$$d_n(x) := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \cdot \mathbb{1}_{\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)}(x) + n \cdot \mathbb{1}_{[n,\infty]}(x).$$

The graph of $d_n(x)$ is given as follow:

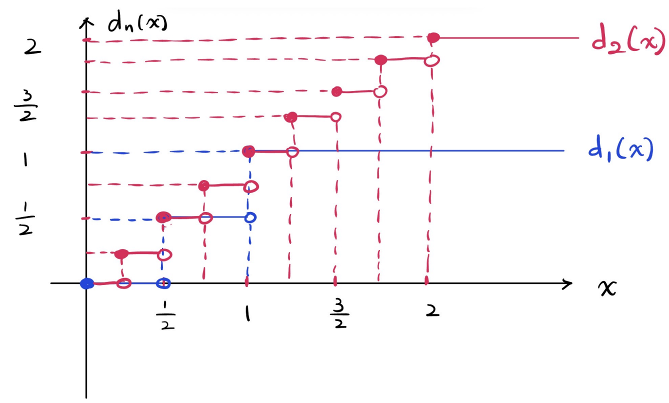

Figure 1: The graphical illustration of $d_n(x)$, with n specifically chosen to be 1 and 2.

We claim that for each n, $d_n(x)$ is a non-decreasing, right continuous simple function, and $\forall x \in \overline{\mathbb{R}}^+$, $d_n(x) \nearrow x$. Now we compose $d_n(\cdot)$ with f, and let

$$f_n(x) = d_n(f(x)).$$

Then by the properties of $d_n(x)$ and f, f_n is measurable, positive, and simple. Furthermore

$$f(x) = \lim_{n\to\infty} f_n(x),$$

and f is measurable, which is the limit of a sequence of positive simple functions. $\square$

In general, given a function f, we may decompose f as

$$f = f^+ - f^-,$$

where $f^+ := \max\{f, 0\}$, and $f^- : \max\{-f, 0\}$.

Theorem 3. A (numerical) function f on E is $\mathcal{E}$-measurable if and only if it is the limit of a sequence of simple functions.

Definition 10. Let $\mathcal{M}$ be a collection of (numerical) functions on E. Denote $\mathcal{M}^+$ as the subcollection containing positive functions in $\mathcal{M}$, and $\mathcal{M}_b$ as the subcollection of bounded functions in $\mathcal{M}$. Then $\mathcal{M}$ is called a **monotone class** if

- $1 \in \mathcal{M}$;
- If $f, g \in \mathcal{M}_b$, $a, b \in \mathbb{R}$ then $af + bg \in \mathcal{M}$;
- If $\{f_n\}_{n=1}^\infty \in \mathcal{M}^+$ and $f_n \nearrow f$, then $f \in \mathcal{M}$.

We have a monotone class theorem for functions. The theorem next can be used to show that a certain property holds for all $\mathcal{E}$-measurable functions.

Theorem 4 (**Monotone class theorem for functions**)**.** Let $\mathcal{M}$ be a monotone class of functions on E. Suppose that for some p-system $\mathcal{C}$ generating $\mathcal{E}$, such that $\mathbb{1}_A \in \mathcal{M}$ for every $A \in \mathcal{C}$, then $\mathcal{M}$ includes all positive $\mathcal{E}$-measurable functions and all bounded $\mathcal{E}$-measurable functions.

Proof. Let $\mathcal{D} := \{A \in \mathcal{E} : 1_A \in \mathcal{M}\}$. Then it is easy to prove that $\mathcal{D}$ is a d-system. By assumption, $\mathcal{D} \supset \mathcal{C}$, and $\mathcal{C}$ is a p-system that generates $\mathcal{E}$, then by monotone class theorem we have $\mathcal{D} \supset \mathcal{E}$, hence $\mathbb{1}_A \in \mathcal{M}$ for every $A \in \mathcal{E}$. In this case using the definition of a monotone class of functions, all simple functions are contained in $\mathcal{M}$. Let f be a positive $\mathcal{E}$-measurable functions, then there exists a sequence of positive simple functions $\{f_n\}_{n=1}^{\infty} \subset \mathcal{M}^+$ such that $f_n \nearrow f$, hence $f \in \mathcal{M}$. Finally let f be bounded $\mathcal{E}$-measurable functions, use $f = f^+ - f^-$, we know both f^+, f^- are positive and bounded, hence $f \in \mathcal{M}$. □

1.4 Measures

Let $(E, \mathcal{E})$ be a measurable space, a **measure** on $(E, \mathcal{E})$ is a function $\mu : \mathcal{E} \to [0, \infty]$ with certain properties. We now formally define a measure:

Definition 11. A measure μ on $(E, \mathcal{E})$ is a non-negative mapping $\mu : \mathcal{E} \mapsto [0, \infty]$ such that

- $\mu(E') \geq 0$, for all $E' \in \mathcal{E}$;
- $\mu(\varnothing) = 0$;
- (*Countable additivity*) If $\{A_n\}_{n=1}^{\infty} \subset \mathcal{E}$ are disjoint, then

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

Note that countable additivity automatically implies finite additivity.

With measure μ on the measurable space $(E, \mathcal{E})$ defined, we denote the triple $(E, \mathcal{E}, \mu)$ a **measure space**.

We will see some examples of different measures. Below, denote $(E, \mathcal{E})$ the measurable space.

- **Dirac measure**: Fix $x \in \mathcal{E}$, and for any $A \in \mathcal{E}$, the Dirac measure on x is defined by

$$\delta_x(A) := \begin{cases} 1 & x \in A; \\ 0 & x \notin A \end{cases}.$$

- **Counting measure**: Fix $D \subset E$, and for any $A \in \mathcal{E}$, the counting measure on D is defined by

$$\nu_D(A) = \begin{cases} |D \cap A| & \text{if } |D| \text{ is countable;} \\ \infty & \text{otherwise} \end{cases}$$

- **Discrete measure**: Fix $D \subset E$ be countable, and $\forall x \in D$, define $\omega_x \geq 0$, then for any $A \in \mathcal{E}$, the counting measure on D is defined by

$$\mu(A) = \sum_{x \in D} \omega_x \delta_x(A).$$

- **Lebesgue measure**: The Lebesgue measure, $\text{Leb}(\cdot)$ on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$, maps each interval to its length. That is, for any $a, b \in \mathbb{R}, a < b$,
$$\text{Leb}((a,b)) = \text{Leb}([a,b]) = \text{Leb}((a,b]) = \text{Leb}([a,b)) = b - a,$$
and singletons have Lebesgue measure 0.

- **Probability measure**: If μ is a probability measure on $(E, \mathcal{E})$, then $\mu(E) = 1$, and E is referred as the sample space.

Based on Definition 11, we are able to derive more properties about a measure μ on $(E, \mathcal{E})$:

Proposition 7. Let $(E, \mathcal{E}, \mu)$ be a measure space, then the followings hold:

- **Monotonicity**: If $A \subset B \in \mathcal{E}$, then $\mu(A) \leq \mu(B)$;
- **Countable subadditivity**: Let $\{A_n\}_{n=1}^{\infty} \in \mathcal{E}$ be an arbitrary sequence of sets, then
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Note that countable subadditivity automatically implies finite subadditivity.

Proof. Assume $A \subset B \in \mathcal{E}$. If one of them has measure ∞ then the result follows trivially. Hence suppose $\mu(A), \mu(B) < \infty$, moreover
$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

Let $B_1 := A_1$, and $B_n := A_n \backslash \left(\bigcup_{i=1}^{n-1} A_i\right)$ for every $n \geq 2$, then we know that $\{B_n, n \geq 1\} \subset \mathcal{E}$ and B_n's are disjoint. More importantly, by our construction we have $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Thus by countable additivity, we have
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$
□

Consider a monotonic sequence $\{A_n\}_{n=1}^{\infty} \in \mathcal{E}$, then we also have the following:

Proposition 8 (**Continuity From Below**)**.** Given $\{A_n\}_{n=1}^{\infty} \in \mathcal{E}$, and $A_n \subset A_{n+1}$ (A_n is increasing) for every n. Denote $\lim_n A_n = A$, then
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \mu(A).$$

Proof. Set $B_1 = A_1$, $B_n = A_n \backslash A_{n-1}$ for all $n \geq 2$. By definition we have $\{B_n, n \geq 1\} \subset \mathcal{E}$ is disjoint, and we have $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$; also for each $N \geq 1$, $\bigcup_{n=1}^{N} B_n = A_n$, which means
$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) &= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) \\ &= \lim_{N\to\infty} \mu\left(\bigcup_{n=1}^{N} B_n\right) \\ &= \lim_{n\to\infty} (A_n).\end{aligned}$$
□

Proposition 9 (**Continuity From Above**). Given $\{A_n\}_{n=1}^{\infty} \in \mathcal{E}$ such that $A_{n+1} \subset A_n$ for every $n \geq 1$ (A_n is decreasing). If $\mu(A_1) < \infty$, denote $\lim_n A_n = A$, then

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

Proof. Set $B_n = A_1 \backslash A_n$, then we know that the sequence $\{B_n, n \geq 1\}$ is increasing, then we have

$$\bigcup_{n=1}^{\infty} B_n = A_1 \backslash \bigcap_{n=1}^{\infty} A_n,$$

thus,

$$\mu\left(A_1 \backslash \bigcap_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \mu(A \backslash A_n).$$

Given that the measure of A_n is always finite, so $\mu(A_1 \backslash A_n) = \mu(A_1) - \mu(A_n)$ and

$$\mu\left(A_1 \backslash \bigcap_{n=1}^{\infty} A_n\right) = \mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right),$$

we then have

$$\mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \left(\mu(A_1) - \mu(A_n)\right),$$

which is

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

$\square$

Definition 12 (**Finite, σ-finite, Σ-finite measures**). Let μ be a measure on $(E, \mathcal{E})$. μ is said to be:

- **Finite**: If $\mu(E) < \infty$;
- **σ-finite**: If there exists a measurable partition $\{E_n\}_I$ of E such that $\mu(E_n) < \infty$ for all $n \in I$;
- **Σ-finite**: If there exists a sequence of measures $\{\mu_n\}$ such that μ_n is finite for all n and $\mu = \sum_n \mu_n$.

Given a measurable space $(E, \mathcal{E})$, we are able to define many different measures. How can we tell them apart?

Theorem 5 (**Monotone class theorem for measures**). Let $(E, \mathcal{E})$ be a measurable space, where 2 finite measures μ, ν are defined. If μ and ν agree on a p-system generating $\mathcal{E}$, then $\mu \equiv \nu$.

Proof. Let $\mathcal{C}$ be a p-system generating $\mathcal{E}$, (i.e, $\mathcal{E} = \sigma(\mathcal{C})$). We assume $\mu(A) = \nu(A), \forall A \in \mathcal{C}$ and both μ, ν are finite. Then we will show $\mu(A) = \nu(A)$ extend to all $A \in \mathcal{E}$. This is equivalent as showing

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

Since $\mathcal{D} \supset \mathcal{E} \supset \mathcal{C}$, it is enough to show $\mathcal{D}$ is a d-system (by Monotone class theorem 1). First it is obvious that $E \in \mathcal{D}, \mu(E) = \nu(E) < \infty$. Let $A, B \in \mathcal{D}, A \supset B$, then

$$\mu(A \setminus B) = \mu(A) - \mu(B) = \nu(A) - \nu(B) = \nu(A \setminus B),$$

hence $A \setminus B \subset \mathcal{D}$. Finally let $\{A_n\}_{n=1}^{\infty} \subset \mathcal{D}$, $A_n \nearrow A$, then using the continuity of μ, ν, we can easily show that $\mu(A) = \nu(A)$. Hence $\mathcal{D}$ is a d-system. □

Corollary 1. Let μ, ν be probability measures on $(\overline{\mathbb{R}}, \mathfrak{B}(\overline{\mathbb{R}}))$, then $\nu = \mu$ if and only if $\mu([-\infty, r]) = \nu([-\infty, r])$ for every $r \in \mathbb{R}$.

This later on illustrates that the cumulative distribution function (CDF) for a random variable is unique.

Definition 13. Let μ be a measure on $(E, \mathcal{E})$, a singleton x is said to be an **atom** of μ if $\mu\{x\} > 0$. The measure μ is said to be **diffuse** if it has no atoms. Furthermore, μ is said to be **purely atomic** if the set of its atoms D is countable and $\mu(E \setminus D) = 0$.

Proposition 10. Let μ be a Σ-finite measure on $(E, \mathcal{E})$, then μ admits a unique decomposition $\mu = \lambda + \nu$ where λ is diffuse and ν is purely atomic.

Definition 14. Let $(E, \mathcal{E}, \mu)$ be a measure space. A measurable set $B \in E$ is said to be **negligible** if $\mu(B) = 0$. Any subset $F \subset B$ (not necessarily measurable) is also said to be negligible. The measurable space is said to be **complete** if every negligible set is measurable.

If a property holds for all but a negligible set in E, we would say it holds for almost every x (or almost everywhere). We will very commonly use the notation "almost everywhere". If a measurable space is not complete, we may extend μ onto an enlarged $\mathcal{E}$, then the new measure space $(E, \bar{\mathcal{E}}, \bar{\mu}$ is called the **completion** of $(E, \mathcal{E}, \mu)$.

Theorem 6 (**Carathéodory's extension theorem**)**.** Let $(E, \mathcal{E}, \mu)$ be a measurable space, $\mathcal{F}$ be the collection of all negligible sets of E. Let $\bar{\mathcal{E}}$ be the σ-algebra generated by $\mathcal{E} \cup \mathcal{F}$, then $\forall B \in \bar{\mathcal{E}}$, $B = A \cup F$ where $A \in \mathcal{E}, F \in \mathcal{F}$; Furthermore $\bar{\mu}(A \cup F) = \mu(A)$ defines a unique measure $\bar{\mu}$ on $\bar{\mathcal{E}}$, $\bar{\mu}(A) = \mu(A), \forall A \in \mathcal{E}$, and $(E, \bar{\mathcal{E}}, \bar{\mu})$ is complete.

In this sense, the Lebesgue measure on $\mathbb{R}$ can be extended to $\mathfrak{B}(\bar{\mathbb{R}})$.

1.5 Integration

Given a measure space $(E, \mathcal{E}, \mu)$, we like to define the integral of a function f over E with respect to the measure μ. We shall denote this as

$$\mu(f) = \int_E f d\mu = \int_E \mu(dx) f(x).$$

Recall that any measurable function can be approximated by a sequence of simple functions. We hence first define the integral for simple functions:

Definition 15. Let $(E, \mathcal{E}, \mu)$ be a measure space. f be a simple function with its canonical form $f = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$ $(A_i \cap A_j = \varnothing, i \neq j))$, then the integral of f over E is

$$\mu(f) = \int_E \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

2 Bibliography

References

[1] Sheldon Axler. *Measure, Integration and Real Analysis*. Vol. 282. Graduate Texts in Mathematics.

[2] Erhan Çinlar. *Probability and Stochastics*. Vol. 261. Graduate Texts in Mathematics.

[3] Jiajun Zhang. *MATH 454 Lecture Notes: Honors Real Analysis 3*. McGill University, 2024.