

---

# Contents

---

## Part I 测度论

---

<b>1</b>	<b>可测空间</b>	<b>3</b>
1.1	可测集	3
1.2	正测度	4
1.3	可测函数	7
1.4	单调类	8
<b>2</b>	<b>可测函数的积分</b>	<b>9</b>
2.1	非负函数的积分	9
2.2	可积函数	11

---

## Part II 概率论

---

<b>3</b>	<b>概率论基础</b>	<b>15</b>
3.1	一般定义	15
3.1.1	概率空间	15
3.1.2	随机变量	16
3.1.3	数学期望	17



# Part I

---

## 测度论

---



# 可测空间

## 1.1 可测集

定义 1.1. 集合  $E$  上的  $\sigma$ -域  $\mathcal{A}$  指的是  $E$  的一个子集族, 其满足下面的性质:

1.  $E \in \mathcal{A}$ ;
2.  $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ ;
3. 如果一列子集  $A_n \in \mathcal{A}$ , 那么  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ .

$\mathcal{A}$  的元素被称为可测集,  $(E, \mathcal{A})$  被称为可测空间. 根据定义, 我们很容易得出下面的结果:

- $\emptyset = E^c \in \mathcal{A}$ .
- 如果一列子集  $A_n \in \mathcal{A}$ , 那么

$$\bigcap_{n \in \mathbb{N}} A_n = \left( \bigcup_{n \in \mathbb{N}} A_n^c \right)^c \in \mathcal{A}.$$

- $\mathcal{A}$  对有限并和有限交也是封闭的, 只需要从某一项  $A_n$  开始全部取空集即可.

例 1.2. 根据可测集的定义, 很容易构造出一些最简单的例子:

1.  $\mathcal{A} = \mathcal{P}(E)$ , 当  $E$  是有限集或者可数集的时候我们通常会使用这样的  $\sigma$ -域, 其他情况则很少使用.
2.  $\mathcal{A} = \{\emptyset, E\}$ , 平凡  $\sigma$ -域.
3.  $E$  的所有至多可数的子集以及所有补集至多可数的子集构成  $E$  上的一个  $\sigma$ -域.

为了产生更多的例子, 我们注意到  $E$  上任意  $\sigma$ -域的交集仍然是  $\sigma$ -域, 这导出了下面的定义.

定义 1.3. 令  $\mathcal{C}$  是  $\mathcal{P}(E)$  的子集,  $E$  上包含  $\mathcal{C}$  的最小的  $\sigma$ -域被记为  $\sigma(\mathcal{C})$ , 不难看出其是所有包含  $\mathcal{C}$  的  $\sigma$ -域的交集. 我们称  $\sigma(\mathcal{C})$  是由  $\mathcal{C}$  生成的  $\sigma$ -域.

定义 1.4. 设  $(E, \mathcal{O})$  是拓扑空间, 所有开集  $\mathcal{O}$  生成的  $\sigma$ -域  $\sigma(\mathcal{O})$  被称为  $E$  上的 Borel  $\sigma$ -域, 记为  $\mathcal{B}(E)$ .

$E$  上的 Borel  $\sigma$ -域是包含所有开集的最小的  $\sigma$ -域.  $\mathcal{B}(E)$  的元素被称为  $E$  的 Borel 子集. 显然,  $E$  中的闭集也都是 Borel 子集.

**例 1.5** ( $\mathbb{R}$  上的 Borel  $\sigma$ -域). 记  $\mathcal{C}_1$  为  $\mathbb{R}$  中开区间的集合:

$$\mathcal{C}_1 = \{(a, b) \mid a, b \in \mathbb{R}, a < b\},$$

显然有  $\mathcal{C}_1 \subseteq \mathcal{B}(\mathbb{R})$ , 于是  $\sigma(\mathcal{C}_1) \subseteq \mathcal{B}(\mathbb{R})$ . 下面我们说明  $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C}_1)$ . 我们不加证明地使用一个结论 (Lindelöf 定理):  $\mathbb{R}$  的任意开子集  $U$  都是开区间的可数并. 那么根据  $\sigma$ -域的定义, 任意开区间都在  $\sigma(\mathcal{C}_1)$  中, 故  $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C}_1)$ . 这表明  $\mathcal{B}(\mathbb{R})$  可以由所有开区间生成.

此外, 如果注意到

$$(a, b) = (-\infty, b) \cap (-\infty, a)^c,$$

还可以证明  $\mathcal{B}(\mathbb{R})$  由  $\mathcal{C}_2$  生成, 其中

$$\mathcal{C}_2 = \{(-\infty, a) \mid a \in \mathbb{R}\}.$$

在后文中, 每当我们考虑拓扑空间 (例如  $\mathbb{R}$  或者  $\mathbb{R}^d$ ) 时, 除非有特别说明, 否则我们总是假设它们配备 Borel  $\sigma$ -域.

下一个非常重要的  $\sigma$ -域是乘积  $\sigma$ -域.

**定义 1.6.** 令  $(E_1, \mathcal{A}_1)$  和  $(E_2, \mathcal{A}_2)$  是可测空间, 定义  $E_1 \times E_2$  上的  $\sigma$ -域  $\mathcal{A}_1 \otimes \mathcal{A}_2$  为

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}).$$

**引理 1.7.** 设  $E$  和  $F$  是可分 (有可数的稠密子集) 的拓扑空间,  $E \times F$  配备积拓扑, 那么  $\mathcal{B}(E \times F) = \mathcal{B}(E) \otimes \mathcal{B}(F)$ .

## 1.2 正测度

令  $(E, \mathcal{A})$  是可测空间.

**定义 1.8.**  $(E, \mathcal{A})$  上的正测度指的是一个映射  $\mu: \mathcal{A} \rightarrow [0, \infty]$ , 其满足下面的性质:

1.  $\mu(\emptyset) = 0$ ;
2. ( $\sigma$ -可加性) 对于任意可数个不相交的可测集序列  $(A_n)_{n \in \mathbb{N}}$ , 有

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

此时, 三元组  $(E, \mathcal{A}, \mu)$  被称为**测度空间**. 值  $\mu(E)$  被称为测度  $\mu$  的总质量.

需要注意的是, 我们允许  $\mu$  的值为  $+\infty$ , 此时级数  $\sum_{n \in \mathbb{N}} \mu(A_n)$  作为正向级数在  $[0, \infty]$  中总是有意义的. 根据  $\sigma$ -可加性, 如果我们令  $n > n_0$  开始  $A_n = \emptyset$ , 便可以得到有限可加性.

**命题 1.9 (测度的性质).** 根据定义, 测度  $\mu$  满足下面的性质:

1. 如果  $A \subseteq B$ , 那么  $\mu(A) \leq \mu(B)$ . 此外, 如果还满足  $\mu(A) < \infty$ , 那么

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

2. 如果  $A, B \in \mathcal{A}$ , 那么

$$\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B).$$

3. 如果  $A_n \in \mathcal{A}$  且  $A_n \subseteq A_{n+1}$ , 那么

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

4. 如果  $B_n \in \mathcal{A}$  且  $B_{n+1} \subseteq B_n$ ,  $\mu(B_1) < \infty$ , 那么

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

5. 如果  $A_n \in \mathcal{A}$ , 那么

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n).$$

*Proof.* (1) 若  $A \subseteq B$ , 那么  $B = A \cup (B \setminus A)$  是无交并, 所以

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(2) 若  $\mu(A), \mu(B)$  中有至少一个为无穷, 那么根据 (1),  $\mu(A \cup B)$  为无穷, 所以结论成立. 下面假设  $\mu(A), \mu(B)$  均有限, 记  $C = A \cap B$ , 那么  $A \cup B = (A \setminus C) \cup C \cup (B \setminus C)$  是无交并, 所以

$$\mu(A \cup B) = \mu(A \setminus C) + \mu(C) + \mu(B \setminus C) = \mu(A) + \mu(B) - \mu(C),$$

结论 (2) 成立.

(3) 令  $C_1 = A_1$ , 对于  $n \geq 2$  的时候, 令

$$C_n = A_n \setminus A_{n-1},$$

那么  $A_n = \bigcup_{k \leq n} C_k$  是无交并, 所以

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}} C_n\right) = \sum_{n \in \mathbb{N}} \mu(C_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(C_k) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(4) 令  $A_n = B_1 \setminus B_n$ , 那么  $A_n \subseteq A_{n+1}$ , 此时

$$\mu\left(\bigcap_{n \in \mathbb{N}} B_n\right) = \mu(B_1) - \mu\left(B_1 \setminus \bigcap_{n \in \mathbb{N}} B_n\right) = \mu(B_1) - \mu\left(\bigcup_{n \in \mathbb{N}} A_n\right),$$

再根据 (3), 就有

$$\mu\left(\bigcap_{n \in \mathbb{N}} B_n\right) = \mu(B_1) - \lim_{n \rightarrow \infty} \mu(A_n) = \lim_{n \rightarrow \infty} \mu(B_1 \setminus A_n) = \lim_{n \rightarrow \infty} \mu(B_n).$$

(5) 令  $C_1 = A_1$ , 对于  $n \geq 2$  的时候, 令

$$C_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k,$$

那么  $C_n$  之间互不相交, 所以

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}} C_n\right) = \sum_{n \in \mathbb{N}} \mu(C_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n). \quad \square$$

**例 1.10** (常见的测度).

1. 令  $E = \mathbb{N}$ ,  $\mathcal{A} = \mathcal{P}(\mathbb{N})$ , 定义**计数测度**为

$$\mu(A) = \text{card}(A).$$

2. 如果  $A$  是  $E$  的子集, 定义  $A$  的指示函数  $\mathbf{1}_A : E \rightarrow \{0, 1\}$  为

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

令  $(E, \mathcal{A})$  是可测空间, 固定  $x \in E$ . 对于每个  $A \in \mathcal{A}$ , 令  $\delta_x(A) = \mathbf{1}_A(x)$ , 这给出了  $(E, \mathcal{A})$  上的一个测度, 被称为  **$x$  处的 Dirac 测度**. 更一般的, 如果  $(x_n)_{n \in \mathbb{N}}$  是  $E$  中的点列,  $(\alpha_n)_{n \in \mathbb{N}}$  是  $[0, \infty]$  中的点列, 我们可以考虑测度  $\sum_{n \in \mathbb{N}} \alpha_n \delta_{x_n}$  为

$$\left(\sum_{n \in \mathbb{N}} \alpha_n \delta_{x_n}\right)(A) = \sum_{n \in \mathbb{N}} \alpha_n \delta_{x_n}(A) = \sum_{n \in \mathbb{N}} \alpha_n \mathbf{1}_A(x_n),$$

这个测度被称为  $E$  上的**点测度**.

3. 可以证明, 在  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  上存在唯一的正测度  $\lambda$  使得: 对于每个开区间  $[a, b]$ , 有  $\lambda([a, b]) = b - a$ . 这个测度  $\lambda$  被称为 **Lebesgue 测度**.

如果  $\mu$  是  $(E, \mathcal{A})$  上的正测度,  $C \in \mathcal{A}$ , 那么可以定义  $\mu$  在  $C$  上的**限制**  $\nu$  为:

$$\nu(A) = \mu(A \cap C), \quad \forall A \in \mathcal{A}.$$

不难验证  $\nu$  还是  $(E, \mathcal{A})$  上的正测度.

**定义 1.11.**

- 如果  $\mu(E) < \infty$ , 那么我们说测度  $\mu$  是**有限的**.
- 如果  $\mu(E) = 1$ , 那么我们说测度  $\mu$  是**概率测度**,  $(E, \mathcal{A}, \mu)$  是**概率空间**.
- 如果存在一列可测集  $(E_n)_{n \in \mathbb{N}}$  使得  $E = \bigcup_n E_n$  以及每个  $\mu(E_n) < \infty$ , 那么我们说测度  $\mu$  是 **$\sigma$ -有限的**.
- 如果  $x \in E$  使得单点集  $\{x\} \in \mathcal{A}$  并且  $\mu(\{x\}) > 0$ , 那么我们说  $x$  是测度  $\mu$  的一个**原子**.
- 如果测度  $\mu$  没有原子, 那么我们说  $\mu$  是**扩散测度**.



如果  $(A_n)_{n \in \mathbb{N}}$  是一列可测集, 类比数列的上下极限, 我们可以定义集合列的上下极限分别为:

$$\limsup A_n = \bigcap_{n=1}^{\infty} \left( \bigcup_{k=n}^{\infty} A_k \right), \quad \liminf A_n = \bigcup_{n=1}^{\infty} \left( \bigcap_{k=n}^{\infty} A_k \right).$$

注意到对于任意  $m$ , 都有

$$\bigcup_{n=1}^m \left( \bigcap_{k=n}^{\infty} A_k \right) = \bigcap_{k=m}^{\infty} A_k, \quad \bigcap_{n=1}^m \left( \bigcup_{k=n}^{\infty} A_k \right) = \bigcup_{k=m}^{\infty} A_k,$$

所以显然有  $\liminf A_n \subseteq \limsup A_n$ .

**引理 1.12.** 令  $\mu$  是  $(E, \mathcal{A})$  上的测度, 那么

$$\mu(\liminf A_n) \leq \liminf \mu(A_n).$$

如果  $\mu$  是有限测度, 或者更一般地,  $\mu(\bigcup_{n=1}^{\infty} A_n) < \infty$ , 那么

$$\mu(\limsup A_n) \geq \limsup \mu(A_n).$$

*Proof.* 对于任意的  $n$ , 有

$$\mu\left(\bigcap_{k=n}^{\infty} A_k\right) \leq \inf_{k \geq n} \mu(A_k),$$

所以

$$\mu(\liminf A_n) = \lim_{n \rightarrow \infty} \mu\left(\bigcap_{k=n}^{\infty} A_k\right) \leq \lim_{n \rightarrow \infty} \inf_{k \geq n} \mu(A_k) = \liminf \mu(A_n).$$

第二个结论同理. □

## 1.3 可测函数

**定义 1.13.** 令  $(E, \mathcal{A})$  和  $(F, \mathcal{B})$  是两个可测空间, 如果映射  $f: E \rightarrow F$  满足:

$$\forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{A},$$

那么我们说  $f$  是 **可测映射**. 当  $E, F$  是两个配备了 Borel  $\sigma$ -域的拓扑空间时, 我们说  $f$  是 **Borel 可测的**.

显然, 可测映射的复合是可测映射.

**命题 1.14.** 令  $(E, \mathcal{A})$  和  $(F, \mathcal{B})$  是两个可测空间, 映射  $f: E \rightarrow F$ .  $f$  可测当且仅当对于某个生成  $\mathcal{B}$  的子集族  $\mathcal{C}$  (即  $\mathcal{B} = \sigma(\mathcal{C})$ ), 有  $f^{-1}(B) \in \mathcal{A} (\forall B \in \mathcal{C})$ .

*Proof.* 只需证明充分性. 记

$$\mathcal{G} = \{B \in \mathcal{B} \mid f^{-1}(B) \in \mathcal{A}\},$$

直接验证可知  $\mathcal{G}$  是一个  $\sigma$ -域, 又因为  $\mathcal{C} \subseteq \mathcal{G}$ , 所以  $\mathcal{B} = \sigma(\mathcal{C}) \subseteq \mathcal{G} \subseteq \mathcal{B}$ , 所以  $\mathcal{G} = \mathcal{B}$ , 这就表明  $f$  是可测的. □

**例 1.15.** 若  $(F, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , 要证明  $f$  是可测的, 只需说明集合  $f^{-1}((a, b))$  是可测的, 或者  $f^{-1}((-\infty, a))$  是可测的.

**推论 1.16.** 设  $E, F$  是两个配备 Borel  $\sigma$ -域的拓扑空间, 那么连续映射  $f : E \rightarrow F$  都是可测的.

**引理 1.17.** 令  $(E, \mathcal{A}), (F_1, \mathcal{B}_1)$  和  $(F_2, \mathcal{B}_2)$  是可测空间, 乘积  $F_1 \times F_2$  配备乘积  $\sigma$ -域  $\mathcal{B}_1 \otimes \mathcal{B}_2$ , 令映射  $f_1 : E \rightarrow F_1$  和  $f_2 : E \rightarrow F_2$ , 定义  $f : E \rightarrow F_1 \times F_2$  为  $f(x) = (f_1(x), f_2(x))$ , 那么  $f$  可测当且仅当  $f_1, f_2$  都可测.

**推论 1.18.** 令  $(E, \mathcal{A})$  是可测空间,  $f, g$  是从  $E$  到  $\mathbb{R}$  的可测函数, 那么函数

$$f + g, fg, \min(f, g), \max(f, g)$$

都是可测的.

记扩充实数  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ , 其拓扑为序拓扑. 与  $\mathbb{R}$  类似,  $\bar{\mathbb{R}}$  的 Borel  $\sigma$ -域由区间  $[-\infty, a)$  生成.

**命题 1.19.** 令  $(f_n)_{n \in \mathbb{N}}$  是  $E \rightarrow \bar{\mathbb{R}}$  的可测函数列, 那么

$$\sup f_n, \inf f_n, \limsup_{n \rightarrow \infty} f_n, \liminf_{n \rightarrow \infty} f_n$$

都是可测函数. 特别地, 如果  $(f_n)$  逐点收敛, 那么极限  $\lim f_n$  是可测函数.

**定义 1.20.** 令  $(E, \mathcal{A})$  和  $(F, \mathcal{B})$  是可测空间,  $\varphi : E \rightarrow F$  是可测映射,  $\mu$  是  $(E, \mathcal{A})$  上的测度, 定义  $(F, \mathcal{B})$  上的测度  $\nu$  为

$$\nu(B) = \mu(\varphi^{-1}(B)), \quad \forall B \in \mathcal{B}.$$

$\nu$  被称为  $\mu$  在  $\varphi$  下的推前, 记为  $\varphi(\mu)$ , 有时也记为  $\varphi_*\mu$ .

## 1.4 单调类

# 可测函数的积分

## 2.1 非负函数的积分

在本章中, 我们考虑配备正测度  $\mu$  的可测空间  $(E, \mathcal{A})$ .

**简单函数** 如果可测函数  $f : E \rightarrow \mathbb{R}$  的值域是有限集, 那么我们说  $f$  的**简单函数**. 假设  $f$  的所有可能的取值为  $\alpha_1, \dots, \alpha_n$ , 不妨假设  $\alpha_1 < \alpha_2 < \dots < \alpha_n$ . 那么  $f$  可以表示为

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

其中  $A_i = f^{-1}(\{\alpha_i\}) \in \mathcal{A}$ . 注意到  $E$  是  $A_1, \dots, A_n$  的无交并. 上述公式  $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$  被称为  $f$  的标准表示.

**定义 2.1.** 令  $f$  是取值在  $\mathbb{R}_+$  中的简单函数, 标准表示为  $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ . 定义  $f$  相对于  $\mu$  的积分为

$$\int f \, d\mu = \sum_{i=1}^n \alpha_i \mu(A_i).$$

在  $\alpha_i = 0$  和  $\mu(A_i) = \infty$  的情况下, 约定  $0 \times \infty = 0$ .

注意上述定义中  $\sum_{i=1}^n \alpha_i \mu(A_i)$  的取值为  $[0, \infty]$ . 所以在上述定义中我们只考虑非负的简单函数, 这是为了避免出现  $\infty - \infty$  之类的表达式.

值得注意的是, 如果简单函数  $f$  有表达

$$f = \sum_{j=1}^m \beta_j \mathbf{1}_{B_j},$$

其中  $B_j$  仍然构成  $E$  的一个划分, 但是  $\beta_j$  不再是两两不同的. 此时  $f$  的积分仍然为

$$\int f \, d\mu = \sum_{j=1}^m \beta_j \mu(B_j).$$

这是因为对于每个  $A_i$ , 某些  $B_j$  构成了  $A_i$  的划分, 即

$$A_i = \bigcup_{\{j \mid \beta_j = \alpha_i\}} B_j,$$

那么

$$\alpha_i \mu(A_i) = \alpha_i \sum_{\{j \mid \beta_j = \alpha_i\}} \mu(B_j) = \sum_{\{j \mid \beta_j = \alpha_i\}} \beta_j \mu(B_j).$$

非负简单函数的积分满足下面的一些基本的性质.

**命题 2.2.** 令  $f, g$  是  $E$  上的非负简单函数.

1. 对于每个  $a, b \in \mathbb{R}_+$ , 有

$$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu.$$

2. 如果  $f \leq g$ , 那么

$$\int f d\mu \leq \int g d\mu.$$

我们用  $\mathcal{E}_+$  来表示  $E$  上的非负简单函数的集合.

**定义 2.3.** 令  $f : E \rightarrow [0, \infty]$  是可测函数, 定义  $f$  相对于  $\mu$  的积分为

$$\int f d\mu = \sup_{h \in \mathcal{E}_+, h \leq f} \int h d\mu.$$

$f$  相对于  $\mu$  的积分通常有很多写法, 下面的表达

$$\int f d\mu, \int f(x) d\mu(x), \int f(x) \mu(dx), \int \mu(dx) f(x)$$

表示的含义是完全相同的. 此外, 如果  $A$  是  $E$  的可测子集, 我们定义

$$\int_A f d\mu = \int f \mathbf{1}_A d\mu.$$

从现在开始, 我们用非负可测函数表示  $E \rightarrow [0, \infty]$  的可测函数 (值可以为无穷). 需要注意的是, 我们前面定义的非负简单函数值必须有限.

**命题 2.4.** 令  $f, g$  是  $E$  上的非负可测函数.

1. 如果  $f \leq g$ , 那么  $\int f d\mu \leq \int g d\mu$ .

2. 如果  $\mu(\{x \in E \mid f(x) > 0\}) = 0$ , 那么  $\int f d\mu = 0$ .

**定理 2.5 (单调收敛定理).** 令  $(f_n)_{n \in \mathbb{N}}$  是  $E$  上的一列递增的非负可测函数, 即  $f_n \leq f_{n+1}$ , 记  $f = \lim \uparrow f_n$ , 那么

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

**命题 2.6.**

1. 设  $f$  是  $E$  上的非负可测函数, 那么存在一列递增的非负简单函数  $(f_n)_{n \in \mathbb{N}}$  使得  $f = \lim \uparrow f_n$ . 如果  $f$  有界, 那么  $f_n \rightarrow f$  一致收敛.
2. 令  $f, g$  是两个  $E$  上的非负可测函数,  $a, b \in \mathbb{R}_+$ , 那么

$$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu.$$

3. 令  $(f_n)_{n \in \mathbb{N}}$  是一列  $E$  上的非负可测函数, 那么

$$\int \left( \sum_{n \in \mathbb{N}} f_n \right) d\mu = \sum_{n \in \mathbb{N}} \int f_n d\mu.$$

**推论 2.7.** 令  $g$  是非负可测函数, 对于  $A \in \mathcal{A}$ , 令

$$\nu(A) = \int_A g d\mu = \int g \mathbf{1}_A d\mu,$$

那么  $\nu$  是  $E$  上的正测度, 被称为密度  $g$  相对于  $\mu$  的测度, 记为  $\nu = g \cdot \mu$ . 此外, 对于非负可测函数  $f$ , 有

$$\int f d\nu = \int fg d\mu.$$

**命题 2.8.** 令  $f$  是非负可测函数.

1. 对于每个  $a \in (0, \infty)$ , 有

$$\mu(\{x \in E \mid f(x) \geq a\}) \leq \frac{1}{a} \int f d\mu.$$

2. 我们有

$$\int f d\mu < \infty \Rightarrow f < \infty, \mu \text{ a.e.}$$

3. 我们有

$$\int f d\mu = 0 \Leftrightarrow f = 0, \mu \text{ a.e.}$$

4. 如果  $g$  是非负可测函数,

$$f = g, \mu \text{ a.e.} \Rightarrow \int f d\mu = \int g d\mu.$$

**定理 2.9 (Fatou 引理).** 令  $(f_n)_{n \in \mathbb{N}}$  是一列非负可测函数, 那么

$$\int \liminf f_n d\mu \leq \liminf \int f_n d\mu.$$

**命题 2.10.** 令  $(F, \mathcal{B})$  是可测空间,  $\varphi: E \rightarrow F$  是可测映射. 令  $\nu$  是  $\mu$  在  $\varphi$  下的推前. 那么, 对于任意  $F$  上的非负可测函数  $h$ , 我们有

$$\int_E h(\varphi(x)) \mu(dx) = \int_F h(y) \nu(dy).$$

## 2.2 可积函数

**定义 2.11.** 令  $f: E \rightarrow \mathbb{R}$  是可测函数, 如果

$$\int |f| d\mu < \infty,$$

那么我们说  $f$  相对于  $\mu$  **可积**. 在这种情况下, 我们定义

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

如果  $A \in \mathcal{A}$ , 记

$$\int_A f \, d\mu = \int f \mathbf{1}_A \, d\mu.$$

## Part II

---

### 概率论

---





# 概率论基础

## 3.1 一般定义

### 3.1.1 概率空间

令  $(\Omega, \mathcal{A})$  是可测空间,  $\mathbb{P}$  是  $(\Omega, \mathcal{A})$  上的概率测度, 我们说  $(\Omega, \mathcal{A}, \mathbb{P})$  是**概率空间**. 因此, 概率空间是测度空间的一个特例. 然而, 概率论的观点与测度论有很大不同. 在概率论中, 我们的目标是一个“随机实验”的数学模型:

- $\Omega$  表示实验的所有可能的结果的集合.
- $\mathcal{A}$  是所有“事件”的集合. 这里的事件指的是  $\Omega$  的一个子集, 其概率可以被计算 (也就是可测集). 我们应当把事件  $A$  视为满足某一属性的所有  $\omega \in \Omega$  构成的子集.
- 对于每个  $A \in \mathcal{A}$ ,  $\mathbb{P}(A)$  表示事件  $A$  发生的概率.

当然, 一个自然的疑问是, 为什么需要考虑事件域  $\mathcal{A}$ ? 换句话说, 为什么不能对  $\Omega$  的任意子集都计算一个概率? 原因在于, 一般不可能在  $\Omega$  的幂集  $\mathcal{P}(\Omega)$  上定义我们感兴趣的概率测度 (除开  $\Omega$  是可数集这一简单情况). 例如, 取  $\Omega = [0, 1]$ , 配备 Borel  $\sigma$ -域和 Lebesgue 测度, 但是, 可以证明不可能将 Lebesgue 测度扩展到  $[0, 1]$  的任意子集上使得其仍然满足测度的定义.

**例 3.1.** 一些常见的概率模型.

1. 考虑扔两次骰子这一实验, 那么

$$\Omega = \{1, 2, \dots, 6\}^2, \quad \mathcal{A} = \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \frac{\text{card}(A)}{36}.$$

这里概率  $\mathbb{P}$  的选取意味着让所有结果都有相同的概率. 更一般地, 如果  $\Omega$  是有限集,  $\mathcal{A} = \mathcal{P}(\Omega)$ , 概率测度  $\mathbb{P}(\{\omega\}) = 1/\text{card}(\Omega)$  被称为  $\Omega$  上的**均匀概率测度**.

2. 现在我们考虑实验: 扔骰子, 直到出现 6 为止. 由于得到 6 所需的投掷次数是无界的 (即使你扔了 1000 次骰子, 仍有可能没有得到 6), 所以  $\Omega$  的正确选择是想象我们扔了无限次骰子:

$$\Omega = \{1, 2, \dots, 6\}^{\mathbb{N}}.$$

$\Omega$  上的  $\sigma$ -域  $\mathcal{A}$  被定义为包含形如

$$\{\omega \in \Omega \mid \omega_1 = i_1, \dots, \omega_n = i_n\}$$

### 3.1.2 随机变量

在本章的剩余部分, 我们都考虑一个概率空间  $(\Omega, \mathcal{A}, \mathbb{P})$ , 并且所有随机变量都将在这个概率空间上定义.

**定义 3.2.** 令  $(E, \mathcal{E})$  是可测空间, 值在  $E$  中的**随机变量**指的是一个可测映射  $X: \Omega \rightarrow E$ .

**例 3.3.** 回顾 (3.1) 中的模型.

1.  $X((i, j)) = i + j$  定义了值在  $\{2, 3, \dots, 12\}$  中的随机变量.
2.  $X(\omega) = \inf\{j \mid \omega_j = 6\}$ , 约定  $\inf \emptyset = \infty$ , 定义了值在  $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$  中的随机变量. 为了验证  $X$  的可测性, 只需要注意到

$$X^{-1}(\{k\}) = \{\omega \in \Omega \mid \omega_1 \neq 6, \dots, \omega_{k-1} \neq 6, \omega_k = 6\}.$$

**定义 3.4.** 令  $X$  是值在  $(E, \mathcal{E})$  中的随机变量, 定义随机变量  $X$  的**分布律**  $\mathbb{P}_X$  是概率测度  $\mathbb{P}$  在  $X$  下的推前. 也就是说,  $\mathbb{P}_X$  是  $(E, \mathcal{E})$  上的概率测度, 满足

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)), \quad \forall B \in \mathcal{E}.$$

两个值在  $(E, \mathcal{E})$  中的随机变量  $Y, Y'$  如果有相同的分布  $\mathbb{P}_Y = \mathbb{P}_{Y'}$ , 那么我们说  $Y$  和  $Y'$  是**同分布的**.

在概率论中, 我们通常将  $\mathbb{P}_X(B)$  写为  $\mathbb{P}(X \in B)$  而不是  $\mathbb{P}(X^{-1}(B))$ . 这里  $X \in B$  是集合  $\{\omega \in \Omega \mid X(\omega) \in B\}$  的简写, 这是一个一般性的简写规则, 在概率论中参数  $\omega$  通常被隐藏.

**离散型随机变量** 当  $E$  是有限或者可数 ( $\mathcal{E} = \mathcal{P}(E)$ ) 的时候,  $X$  的分布是点测度, 这是因为

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}\left(\bigcup_{x \in B} \{X = x\}\right) = \sum_{x \in B} \mathbb{P}(X = x) = \sum_{x \in E} p_x \delta_x(B),$$

其中  $p_x = \mathbb{P}(X = x)$ . 这就表明

$$\mathbb{P}_X = \sum_{x \in E} p_x \delta_x$$

是  $E$  上的点测度.

**例 3.5.** 我们考虑 (3.1) 中的第二个例子, 随机变量为  $X(\omega) = \inf\{j \mid \omega_j = 6\}$ . 那么

$$\begin{aligned} \mathbb{P}(X = k) &= \mathbb{P}\left(\bigcup_{1 \leq i_1, \dots, i_{k-1} \leq 5} \{\omega \mid \omega_1 = i_1, \dots, \omega_{k-1} = i_{k-1}, \omega_k = 6\}\right) \\ &= 5^{k-1} \left(\frac{1}{6}\right)^k = \frac{1}{6} \left(\frac{5}{6}\right)^{k-1}. \end{aligned}$$

注意到

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = \frac{1}{6} \frac{1}{1 - \frac{5}{6}} = 1$$

并且  $\{X = \infty\} \cup \bigcup_{k=1}^{\infty} \{X = k\} = \Omega$ , 所以

$$\mathbb{P}(X = \infty) = 1 - \sum_{k=1}^{\infty} \mathbb{P}(X = k) = 0,$$

但是  $\{X = \infty\} \neq \emptyset$ .

**具有密度的随机变量**  $\mathbb{R}^d$  上的密度函数是一个非负的 Borel 函数  $p : \mathbb{R}^d \rightarrow \mathbb{R}_+$ , 其满足

$$\int_{\mathbb{R}^d} p(x) dx = 1.$$

对于一个值在  $\mathbb{R}^d$  中的随机变量  $X$ , 如果存在密度  $p$  使得

$$\mathbb{P}_X(B) = \int_B p(x) dx$$

对于任意 Borel 子集  $B$  都成立, 那么我们说  $X$  有密度函数  $p$ . 换句话说,  $p$  是  $\mathbb{P}_X$  相对于 Lebesgue 测度的密度.

注意到密度  $p$  实际上是在相差一个 Lebesgue 零测集的意义下由  $\mathbb{P}_X$  确定的. 在我们遇到的大多数例子中,  $p$  在  $\mathbb{R}^d$  上连续, 在这种情况下,  $p$  由  $\mathbb{P}_X$  唯一确定.

在  $d = 1$  的时候, 我们有

$$\mathbb{P}(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} p(x) dx.$$

### 3.1.3 数学期望

**定义 3.6.** 令  $X$  是定义在  $(\Omega, \mathcal{A}, \mathbb{P})$  上的实随机变量, 我们定义

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) \mathbb{P}(d\omega) = \int X d\mathbb{P},$$

只要上述积分有意义, 我们就说  $\mathbb{E}[X]$  是  $X$  的**期望**.

根据前面的内容, 上述积分有意义的条件为下列二者之一:

- $X \geq 0$ , 此时  $\mathbb{E}[X] \in [0, \infty]$ .
- $X$  符号任意, 但是  $\mathbb{E}[|X|] = \int |X| d\mathbb{P} < \infty$ .

上面的定义可以拓展到多元随机变量  $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ , 此时我们定义  $\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$ . 类似的, 如果  $M$  是随机矩阵 (值在实矩阵空间中的随机变量), 我们可以定义矩阵  $\mathbb{E}[M]$  为对  $M$  的每个分量求期望构成的矩阵.

注意到若  $X = \mathbf{1}_B$ , 那么

$$\mathbb{E}[X] = \int \mathbf{1}_B d\mathbb{P} = \mathbb{P}(B).$$

对于一些特殊的随机变量, 下面的命题被频繁地使用.

**命题 3.7.** 令  $X$  是值在  $[0, \infty]$  中的随机变量, 那么

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X \geq x) dx.$$

令  $Y$  是值在  $\mathbb{Z}_+$  中的随机变量, 那么

$$\mathbb{E}[Y] = \sum_{k=0}^{\infty} k \mathbb{P}(Y = k) = \sum_{k=1}^{\infty} \mathbb{P}(Y \geq k).$$

*Proof.* 根据 Fubini 定理, 我们有

$$\mathbb{E}[X] = \mathbb{E} \left[ \int_0^\infty \mathbf{1}_{\{x \leq X\}} dx \right] = \int_0^\infty \mathbb{E}[\mathbf{1}_{\{x \leq X\}}] dx = \int_0^\infty \mathbb{P}(X \geq x) dx.$$

对于随机变量  $Y$ , 我们有

$$\mathbb{E}[Y] = \mathbb{E} \left[ \sum_{k=0}^{\infty} k \mathbf{1}_{\{Y=k\}} \right] = \int \left( \sum_{k=0}^{\infty} k \mathbf{1}_{\{Y=k\}} \right) d\mathbb{P} = \sum_{k=0}^{\infty} k \mathbb{P}(Y = k).$$

对于第二个等式, 只需注意到

$$Y = \sum_{k=1}^{\infty} \mathbf{1}_{\{Y \geq k\}}.$$

□

下面的命题是 **命题 2.10** 的特例, 其十分重要, 所以我们再次叙述一遍.

**命题 3.8.** 令  $X$  是值在  $(E, \mathcal{E})$  中的随机变量, 对于任意可测函数  $f : E \rightarrow [0, \infty]$ , 我们有

$$\mathbb{E}[f(X)] = \int_E f(x) \mathbb{P}_X(dx).$$