概 率 论 基 础

王凤雨　毛永华

(北京师范大学数学科学学院)

二〇〇九年九月十四日

目录

第一章　集类与测度

什么是测度? 简单地讲, 测度是用来测量集合大小的工具. 例如, 使用通常的测度 (Lebesgue 测度) 测量 $\mathbb{R}$ 中一个区间 $[a,b)$ 时, 得到的测量结果是 $b-a$. 那么对于一个抽象的全集 Ω, 我们如何选定其子集类并对其中的集合进行测量(即定义该子集类上的测度)? 通过学习 Lebesgue 测度, 我们知道, 通常并非 Ω 的所有子集都可以测量 (即可测), 因此我们首先需要研究如何定义可测集类. 为此, 先回顾 Lebesgue 可测集类所具备的基本特征: (1) 包含全集和空集; (2) 对于集合的可数次运算 (即集合的交、并、差) 是封闭的. 我们将具有这两个特征的由 Ω 的一些子集所组成的类, 称作 σ 代数 (域), 它将是我们理想中的"可测集类".

那么如何在一个 σ 代数上定义测度呢? 让我们回到已学过的 $\mathbb{R}$ 上的 Lebesgue 测度.

前面提过, 我们首先很容易定义区间的测度值, 即 $[a,b)$ 的测度为 $b-a\,(\forall b\geqslant a)$, 因此我们首先在由区间组成的集类上定义了该测度, 然后再通过一些合理的手段把它定义到可测集类上. 如何将这一手法推广到一般情形呢? 如同 σ 代数的定义, 我们先看看由区间所组成的集类所具备的特征: (1) 包含全集和空集; (2) 对于交封闭, 且这样两个集合的差可表成有限个同类集合的不交并. 我们将 Ω 中具有这两个特征的子集类称为半集代数. 而假设已经在一个半集代数上定义了测度, 我们再设法将其扩张到相应的 σ 代数上, 这就是本章的核心定理—测度扩张定理. 为此, 如何由半集代数生成 σ 代数, 便是我们首先要研究的内容, 其核心结果就是单调类定理.

§1.1 集类与单调类定理

§1.1.1 半集代数

如同前面的解释, 以 $\mathbb{R}$ 上的区间所组成的集类的特征为基础, 引入半集代数的概念. 在集合的运算中, 我们将用 $\sum$ 代表集合的不交并.

定义 1.1. 如果 Ω 的子集类 $\mathscr{S}$ 满足

1) $\Omega, \emptyset \in \mathscr{S}$,

2) $A, B \in \mathscr{S} \Rightarrow A \cap B \in \mathscr{S}$,

3) $A_1, A \in \mathscr{S}, A_1 \subset A \Rightarrow \exists n \geqslant 1$ 及 $A_1, A_2, \cdots, A_n \in \mathscr{S}$ 两两不交, 使 $A = \sum\limits_{i=1}^{n} A_i$,

则称之为 Ω 中的一个半集代数.

性质 1.2. 在定义 1.1 之 1) 和 2) 成立的条件下, 3) 等价于:

3′ 若 $A \in \mathscr{S}$, 则 $\exists n \geqslant 1$ 及 $A_1, A_2, \cdots, A_n \in \mathscr{S}$ 两两不交, 使得 $A^c = \sum\limits_{i=1}^{n} A_i$.

证明 3) $\Rightarrow$ 3′): 由于 $A \subset \Omega$, 由 3) 知 $\exists n \geqslant 1$ 及 $A_1, A_2, \cdots, A_n \in \mathscr{S}$ 两两不交, 且与 A 不交, 使得 $\Omega = A + \sum\limits_{i=1}^{n} A_i$, 从而 $A^c = \sum\limits_{i=1}^{n} A_i$.

3′) $\Rightarrow$ 3): 由 3′) 知 $\exists n \geqslant 1$ 及 $A_2, \cdots, A_n \in \mathscr{S}$ 两两不交, 使得 $A_1^c = \sum\limits_{i=2}^{n} A_i$, 则 $A = A_1 + \sum\limits_{i=2}^{n} A_i \cap A$. □

例 1.3. $\Omega = [0, +\infty)$, $\mathscr{S} = \{[a, b) : 0 \leqslant a \leqslant b \leqslant +\infty\}$, 则 $\mathscr{S}$ 是半集代数.

作为由半集代数到 σ 代数的过渡, 我们引入集代数, 它对于集合的有限次运算均封闭, 因而更接近 σ 代数. 后面我们将看到, 由半集代数生成集代数是非常直接的.

§1.1.2 集代数

定义 1.4. 如果 Ω 的子集类 $\mathscr{F}$ 满足

1) $\Omega \in \mathscr{F}$,

2) $A, B \in \mathscr{F} \Rightarrow A - B \in \mathscr{F}$,

则称之为 Ω 中的一个集代数 (或 Boole 代数).

性质 1.5. 在定义 1.4 之 1) 成立的条件下, 2) 与下列任意一个条件等价:

2′) $A, B \in \mathscr{F} \Rightarrow A \cup B, A^c, B^c \in \mathscr{F}$;

2″) $A, B \in \mathscr{F} \Rightarrow A \cap B, A^c, B^c \in \mathscr{F}$.

证明 我们将证明 2″) ⇒ 2′) ⇒ 2) ⇒ 2″).

2″) ⇒ 2′): 由 2″), $\mathscr{F}$ 对于余与交封闭, 从而 $A, B \in \mathscr{F}$ 蕴含 $A \cup B = (A^c \cap B^c)^c \in \mathscr{F}$.

2′) ⇒ 2): 设 $A, B \in \mathscr{F}$. 由 2′) 知 $\mathscr{F}$ 对于余与并封闭, 从而 $A - B = A \cap B^c \in \mathscr{F}$.

2) ⇒ 2″): 设 $A, B \in \mathscr{F}$. 由 2) 知 $A^c = \Omega - A, B^c = \Omega - B \in \mathscr{F}$, 进而 $A \cap B = A - B^c \in \mathscr{F}$. □

命题 1.6. 设 $\mathscr{F}$ 是 Ω 中的集代数, 则 $\forall A, B \in \mathscr{F}$, 有 $A^c, B^c, A \cap B, A \cup B, A - B \in \mathscr{F}$.

显然, 集代数一定是半集代数. 下面的定理告诉我们如何由半集代数生成集代数.

定理 1.7. 若 $\mathscr{S}$ 是半集代数, 则

$$\mathscr{F} = \left\{ \sum_{k=1}^{n} A_k : n \geqslant 1, A_k \in \mathscr{S} (1 \leqslant k \leqslant n) \text{ 两两不交} \right\}$$

是包含 $\mathscr{S}$ 的最小集代数, 记作 $\mathscr{F}(\mathscr{S})$.

证明 先证 $\mathscr{F}$ 是集代数. 显然定义 1.4 之 1) 成立. 此外, $\forall A, B \in \mathscr{F}$, $\exists A_1, A_2, \cdots, A_n \in \mathscr{S}$ 及 $B_1, B_2, \cdots, B_m \in \mathscr{S}$, 分别两两不交且 $A = \sum\limits_{i=1}^{n} A_i, B = \sum\limits_{i=1}^{m} B_i$. 则 $A \cap B = \sum\limits_{i,j} A_i \cap B_j$. 由定义 1.1 之 2) 知 $A \cap B \in \mathscr{F}$. 从而 $\mathscr{F}$ 对有限交封闭.

由性质 1.5, 为证 $\mathscr{F}$ 为集代数, 只需证明若 $A \in \mathscr{F}$, 则 $A^c \in \mathscr{F}$. 设 $A = \sum\limits_{i=1}^{n} A_i \in \mathscr{F}, A_i \in \mathscr{S}$. 则 $A^c = \bigcap\limits_{i=1}^{n} A_i^c$. 由性质 1.2 知 A_i^c 可表为 $\mathscr{S}$ 中两两不交集合之和, 故 $A_i^c \in \mathscr{F}$. 由于 $\mathscr{F}$ 对于有限交封闭, 从而 $A^c \in \mathscr{F}$.

最后如果 $\mathscr{F}' \supset \mathscr{S}$ 是集代数, 由集代数的有限并封闭性质知 $\mathscr{F}' \supset \mathscr{F}$. □

例 1.8. 例 1.3 中的 $\mathscr{S}$ 不是集代数, 对并运算不封闭. 由定理 1.7,

$$\mathscr{F}(\mathscr{S}) = \left\{ \sum_{i=1}^{n} [a_i, b_i) : n \geqslant 1, 0 \leqslant a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \cdots \leqslant a_n \leqslant b_n \right\}.$$

§ 1.1.3 σ 代数

按照 Lebesgue 可测集类的特征, 我们要求 σ 代数对于可数运算封闭. 由于并与交可以通过余运算而相互表示, 我们可以简单地要求其对于余运算及可数交 (或并) 运算封闭.

定义 1.9. 如果 Ω 的子集类 $\mathscr{A}$ 满足

1) $\Omega \in \mathscr{A}$,

2) 若 $A \in \mathscr{A}$, 则 $A^c \in \mathscr{A}$,

3) 若 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$,

则称之为 Ω 中的一个 σ 代数.

性质 1.10. σ 代数是集代数.

性质 1.11. 在定义 1.9 中, 当 1) 与 2) 成立的条件时, 3) 等价于

3′) 若 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{A}$.

证明 仅留意 $\bigcap\limits_{n=1}^{\infty} A_n = \left(\bigcup\limits_{n=1}^{\infty} A_n^c \right)^c$. □

性质 1.12. Ω 中任意多个 σ 代数的交仍是 Ω 中的 σ 代数.

证明 设 $\{\mathscr{A}_r : r \in \Gamma\}$ 是一族 σ 代数, $\mathscr{A} = \bigcap\limits_{r \in \Gamma} \mathscr{A}_r$.

1) 由于任给 $r \in \Gamma$, $\emptyset, \Omega \in \mathscr{A}_r$, 则 $\emptyset, \Omega \in \mathscr{A}$.

2) 若 $A \in \mathscr{A}$, 则 $\forall r \in \Gamma$ 有 $A \in \mathscr{A}_r$. 从而 $A^c \in \mathscr{A}_r$ $(r \in \Gamma)$, 故 $A^c \in \mathscr{A}$.

3) 若 $A_1, A_2, \cdots \in \mathscr{A}$, 则 $\forall r \in \Gamma$ 有 $A_1, A_2, \cdots \in \mathscr{A}_r$, 从而 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}_r$. 故 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$. □

例 1.13. $\mathscr{A} = \{\emptyset, \Omega\}$ 是 Ω 中的最小 σ 代数, $\mathscr{A} = \{A : A \subset \Omega\}$ 是 Ω 中的最大 σ 代数. 最大 σ 代数常被记为 2^{Ω}, 这是由于 Ω 中的每个子集唯一对应于 $\{0,1\}^{\Omega}$ 中的一个组态: $\Omega \ni \omega \mapsto \mathbf{1}_A(\omega)$, 其中 $\mathbf{1}_A$ 为 A 的示性函数.

定理 1.14. 设 $\mathscr{C}$ 是 Ω 的一个子集类. 则存在 Ω 中的 σ 代数 $\mathscr{A}_0$, 使得

1) $\mathscr{C} \subset \mathscr{A}_0$,

2) 如 $\mathscr{A}$ 为 Ω 中的 σ 代数且 $\mathscr{A} \supset \mathscr{C}$, 则 $\mathscr{A} \supset \mathscr{A}_0$.

证明 由于最大 σ 代数包含 $\mathscr{C}$, 故存在包含 $\mathscr{C}$ 的 σ 代数. 令 $\mathscr{A}_0$ 是包含 $\mathscr{C}$ 的所有 σ 代数之交, 则由性质 1.12 知 $\mathscr{A}_0$ 是 σ 代数, 且包含 $\mathscr{C}$. 此外 $\forall\sigma$ 代数 $\mathscr{A}\supset\mathscr{C}$, 有 $\mathscr{A}\supset\mathscr{A}_0$. □

我们称定理 1.14 中的 σ 代数是由 $\mathscr{C}$ 生成的 σ 代数, 记作 $\sigma(\mathscr{C})$.

下面的定理表明, 当半集代数生成 σ 代数时, 可先生成集代数, 再由集代数生成 σ 代数.

定理 1.15. 若 $\mathscr{S}$ 是 Ω 中的半集代数, 则 $\sigma(\mathscr{S})=\sigma(\mathscr{F}(\mathscr{S}))$.

证明 由于 $\sigma(\mathscr{F}(\mathscr{S}))\supset\mathscr{S}$, 则 $\sigma(\mathscr{F}(\mathscr{S}))\supset\sigma(\mathscr{S})$. 反之, 由于 $\sigma(\mathscr{S})$ 是包含 $\mathscr{S}$ 的集代数, 故 $\sigma(\mathscr{S})\supset\mathscr{F}(\mathscr{S})$, 而 $\sigma(\mathscr{S})$ 是 σ 代数, 所以 $\sigma(\mathscr{S})\supset\sigma(\mathscr{F}(\mathscr{S}))$. □

例 1.16. 在 $\mathbb{R}^d$ 中, 由所有开集生成的 σ 代数称为 Borel 域 (或 Borel σ 代数), Borel 域中的元素称为 Borel 集, 它包含所有的开集和闭集, 且与由开集类或闭集类所生成的 σ 代数相同. 在一般的拓扑空间中, 由开集或闭集类生成的 σ 代数也称为 Borel σ 代数, 或 Borel 域.

§ 1.1.4　单调类定理

相对于数列极限而言, 集合序列的极限仅对单调增或单调降两种情形有定义, 相应的极限分别是单调增序列的并与单调降序列的交. 这是两种特殊的易于验证的集合的可数运算. 我们因此把验证单调集合序列极限的封闭性作为集类是否为 σ 代数的关键步骤. 为此, 引入单调类的概念.

定义 1.17. Ω 中子集类 $\mathscr{M}$ 如果对单调序列的极限封闭, 即:

1) 如 $A_n\in\mathscr{M}, n=1,2,\cdots$, 且 $A_1\subset A_2\subset\cdots$, 则 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{M}$;

2) 如 $A_n\in\mathscr{M}, n=1,2,\cdots$, 且 $A_1\supset A_2\supset\cdots$, 则 $\bigcap\limits_{n=1}^{\infty}A_n\in\mathscr{M}$.

则称之为 Ω 中的一个单调类.

前面我们已经将由半集代数生成的集代数的结构弄清楚了, 那么如何由集代数生成 σ 代数呢? 按照定义, 我们需要把集代数中任意可数个集合拿来运算. 而下面的定理将告诉我们, 只需把单调序列的极限纳入即可产生 σ 代数.

定理 1.18. Ω 中的集类为 σ 代数当且仅当它既是集代数又是单调类.

证明 仅需证明充分性. 若 $\mathscr{F}$ 既是集代数又是单调类, $\forall A_1, A_2, \cdots \in \mathscr{F}$, 令 $B_n = \bigcup\limits_{i=1}^{n} A_i$, 则 $\{B_n\} \subset \mathscr{F}$ 且单增. 由于 $\mathscr{F}$ 是单调类, 则 $\bigcup\limits_{n=1}^{\infty} B_n = \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{F}$, 从而 $\mathscr{F}$ 为 σ 代数. □

定理 1.19. 设 $\mathscr{C}$ 是 Ω 中任一集代数, 则存在 Ω 中的单调类 $\mathscr{M}_0$ 满足:

1) $\mathscr{M}_0 \supset \mathscr{C}$,

2) $\forall$ 包含 $\mathscr{C}$ 的单调类 $\mathscr{M}$, 有 $\mathscr{M} \supset \mathscr{M}_0$.

称这样的单调类为由 $\mathscr{C}$ 生成的单调类, 记作 $\mathscr{M}(\mathscr{C})$.

定理 1.20. 设 $\mathscr{F}$ 是集代数, 则 $\mathscr{M}(\mathscr{F}) = \sigma(\mathscr{F})$.

证明 只需证明 $\mathscr{M}(\mathscr{F})$ 是 σ 代数, 而由定理 1.18 只需证明其为集代数.

a) $A \in \mathscr{M}(\mathscr{F}) \Rightarrow A^c \in \mathscr{M}(\mathscr{F})$: 令 $\mathscr{M}_1 = \{A : A^c \in \mathscr{M}(\mathscr{F})\}$. $\forall$ 单降序列 $\{A_n\} \subset \mathscr{M}_1$, 有 $\{A_n^c\} \subset \mathscr{M}(\mathscr{F})$ 且单增. 由于 $\mathscr{M}(\mathscr{F})$ 为单调类, 有 $\bigcup\limits_{n=1}^{\infty} A_n^c \in \mathscr{M}(\mathscr{F})$, 从而 $\bigcap\limits_{n=1}^{\infty} A_n = \left(\bigcup\limits_{n=1}^{\infty} A_n^c\right)^c \in \mathscr{M}_1$. 类似可证 $\mathscr{M}_1$ 对于单调增序列的极限也封闭. 因此 $\mathscr{M}_1$ 为包含 $\mathscr{F}$ 的单调类, 从而也包含 $\mathscr{M}(\mathscr{F})$. 故 $\forall A \in \mathscr{M}(\mathscr{F})$ 有 $A^c \in \mathscr{M}(\mathscr{F})$.

b) 往证如 $A \in \mathscr{F}$, 则 $\forall B \in \mathscr{M}(\mathscr{F})$ 有 $A \bigcap B \in \mathscr{M}(\mathscr{F})$. 为此令 $\mathscr{M}_A = \{B \in \mathscr{M}(\mathscr{F}) : A \cap B \in \mathscr{M}(\mathscr{F})\}$, 则 $\mathscr{M}_A \supset \mathscr{F}$. 如 $\{B_n\} \subset \mathscr{M}_A$ 单增, 则 $\{A \bigcap B_n\} \subset \mathscr{M}(\mathscr{F})$ 单增, 故 $A \cap \left(\bigcup\limits_{n=1}^{\infty} B_n\right) \in \mathscr{M}(\mathscr{F})$, 从而 $\bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{M}_A(\mathscr{F})$. 同样处理单降情形, 知 $\mathscr{M}_A$ 是单调类. 故 $\mathscr{M}_A \supset \mathscr{M}(\mathscr{F})$. 即 $\forall B \in \mathscr{M}(\mathscr{F})$, 有 $A \cap B \in \mathscr{M}(\mathscr{F})$.

c) 往证 $\forall A, B \in \mathscr{M}(\mathscr{F})$, 有 $A \cap B \in \mathscr{M}(\mathscr{F})$. 由 b) 知 $\mathscr{M}_A \supset \mathscr{F}$ 且为单调类, 故 $\mathscr{M}_A \supset \mathscr{M}(\mathscr{F})$, 从而 $A \cap B \in \mathscr{M}(\mathscr{F})$. □

上面的证明技巧很容易掌握. 一般地, 要证明某集类 $\mathscr{C}_1$ 具有某种性质, 我们将具有这种性质的所有集合组成一个新的集类 $\mathscr{C}_2$, 只需证明 $\mathscr{C}_1 \subset \mathscr{C}_2$ 即可. 有时为论证方便, 需分成数步完成, 就好比上面证明 $\mathscr{M}(\mathscr{F})$ 对交的封闭性, 分成 b)、c) 两步完成更容易些. 这样的技巧后面会反复使用.

验证单调类易于检验对可数交 (并) 封闭. 单调类与集代数一起形成 σ 代数. 下面我们引入另一对常用的可以形成 σ 代数的集类.

定义 1.21. (1) 如果 Ω 中子集类 $\mathscr{C}$ 对交封闭, 则称之为 π 系.

(2) 如果 Ω 中子集类 $\mathscr{C}$ 满足

1) $\Omega \in \mathscr{C}$;

2) $A, B \in \mathscr{C}, A \subset B \Longrightarrow B - A \in \mathscr{C}$;

3) $\{A_n\} \subset \mathscr{C}$ 单增 $\Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathscr{C}$.

则称 $\mathscr{C}$ 为 Ω 中的一个 λ 系.

性质 1.22. 若 $\mathscr{C}$ 是 λ 系, 则它是单调类.

证明 若$\{A_n\} \subset \mathscr{C}$ 单减, 则 $\{A_n^c\} \subset \mathscr{C}$ 单增. 从而 $\bigcup_{n=1}^{\infty} A_n^c \in \mathscr{C}$, 故 $\bigcap_{n=1}^{\infty} A_n = \Omega - \bigcup_{n=1}^{\infty} A_n^c \in \mathscr{C}$. □

性质 1.23. 如 $\mathscr{C}$ 同时是 π 系 和 λ 系, 则 $\mathscr{C}$ 是 σ 代数.

证明 由定理 1.18 与性质 1.22, 只需证明 $\mathscr{C}$ 是集代数, 而由 π 系定义及 λ 系当定义 2), 这是显然的. □

如同集类 $\mathscr{C}$ 产生最小 σ 代数与单调类, 它也产生包含它的最小 λ 系, 记作 $\lambda(\mathscr{C})$.

定理 1.24. 设 $\mathscr{C}$ 是 π 系, 则 $\lambda(\mathscr{C}) = \sigma(\mathscr{C})$.

证明 由于 $\lambda(\mathscr{C}) \subset \sigma(\mathscr{C})$ 且 $\lambda(\mathscr{C})$ 为单调类, 由定理 1.18 只需证 $\lambda(\mathscr{C})$ 为集代数. 由 λ 系的定义知, $\lambda(\mathscr{C})$ 对于集合的余封闭, 从而只需证明它对于交也封闭. 为此, 分两步证明.

1) 设 $A \in \lambda(\mathscr{C}), B \in \mathscr{C}$, 往证 $A \cap B \in \lambda(\mathscr{C})$. 如前解释, 令 $\mathscr{C}_B = \{A : A \cap B \in \lambda(\mathscr{C})\}$. 由于 $\mathscr{C}$ 为 π 系, 有 $\mathscr{C}_B \supset \mathscr{C}$. 由 $\mathscr{C}_B$ 的定义及 $\lambda(\mathscr{C})$ 为 λ 系的事实, 易验证 $\mathscr{C}_B$ 也是 λ 系, 从而 $\mathscr{C}_B \supset \lambda(\mathscr{C})$, 即 $\forall A \in \lambda(\mathscr{C})$ 有 $A \cap B \in \lambda(\mathscr{C})$.

2) 设 $B \in \lambda(\mathscr{C})$, 由 1) 知 $\mathscr{C}_B \supset \mathscr{C}$ 且为 λ 系, 从而 $\mathscr{C}_B \supset \lambda(\mathscr{C})$. 故 $\forall A, B \in \lambda(\mathscr{C})$ 有 $A \cap B \in \lambda(\mathscr{C})$. □

有了上面的准备, 我们得到了下面的重要定理.

定理 1.25 (单调类定理)**.** 设 $\mathscr{C}$ 和 $\mathscr{A}$ 是 Ω 的两个子集类且 $\mathscr{C} \subset \mathscr{A}$.

1) 若 $\mathscr{A}$ 是 λ 系, $\mathscr{C}$ 是 π 系, 则 $\sigma(\mathscr{C}) \subset \mathscr{A}$,

2) 若 $\mathscr{A}$ 是单调类, $\mathscr{C}$ 是集代数, 则 $\sigma(\mathscr{C}) \subset \mathscr{A}$.

这个定理是集类部分的主要定理, 非常有用, 其使用方法如下:

设已知 $\mathscr{C}$ 中元具有性质 S, 证明 $\sigma(\mathscr{C})$ 中元仍具有性质 S. 为此, 令 $\mathscr{A}=\{B: B$ 有性质 $S\}$, 则 $\mathscr{A}\supset\mathscr{C}$. 再证 $\mathscr{C}$ 为 π 系或集代数, 相应的 $\mathscr{A}$ 为 λ 系或单调类, 则 $\sigma(\mathscr{C})\subset\mathscr{A}$.

作为本节主要内容的总结, 下面给出我们所学的各种集类之间的关系图.

注 1.26. 1) σ 代数 $\Rightarrow$ 集代数 $\Rightarrow$ 半集代数 $\Rightarrow$ π 系

2) π 系 $+$ λ 系 $\Rightarrow$ σ 代数 $\Rightarrow$ λ 系

3) λ 系 $\Rightarrow$ 单调类, 单调类 $+$ 集代数 $\Rightarrow$ σ 代数

§ 1.1.5 乘积空间与乘积 σ 代数

设 $\Omega_i, i=1,\cdots,n$ 是 n 个集合(空间), $\mathscr{A}_i$ 是 Ω_i 中的 σ 代数. 令 $\mathscr{C}=\{A_1\times\cdots\times A_n: A_i\in\mathscr{A}_i\}$, $\mathscr{C}$ 中的集合称为矩形. 易见 $\mathscr{C}$ 为半集代数, 称 $\sigma(\mathscr{C})$ 为 $\mathscr{A}_1,\cdots,\mathscr{A}_n$ 的乘积 σ 代数, 记作 $\mathscr{A}_1\times\cdots\times\mathscr{A}_n$.

定理 1.27 (结合率). $\forall n\geqslant 3, 1\leqslant k\leqslant n$, 有 $\mathscr{A}_1\times\mathscr{A}_2\times\cdots\times\mathscr{A}_n=(\mathscr{A}_1\times\cdots\times\mathscr{A}_k)\times(\mathscr{A}_{k+1}\times\cdots\times\mathscr{A}_n)$.

定理 1.27 可由乘积 σ 代数的定义与单调类定理得到. 我们把它的证明留作练习题.

§1.2 集函数与测度

我们既已研究了集类, 接下来就要对集类中的集合进行测量, 进而引入测度的概念. 在许多时候, 我们允许集合的测量值为负 (例如测量某地区平均气温等). 因此, 测量的结果就是某集类中的每个集合对应着实数空间中的一个值, 这就是集函数.

§ 1.2.1 集函数

定义 1.28. 设 $\mathscr{C}$ 是 Ω 的一个子集类, 则定义在 $\mathscr{C}$ 上的取值于 $(-\infty,+\infty]$ 的函数 Φ 称为 $\mathscr{C}$ 上的一个集函数. 约定集函数不恒为 ∞.

这里, 我们允许函数值取 $+\infty$, 以便覆盖 Lebesgue 测度. 如此, 就不宜允许测度值取 $-\infty$, 否则无法讨论下面的可加性. 为使集函数可以运算, 还约定集函数至少取一个有限值.

通常研究具有以下性质的集函数.

1) 可加性 如果 $\forall A, B \in \mathscr{C}$ 不交且 $A \cup B \in \mathscr{C}$ 有 $\Phi(A+B) = \Phi(A) + \Phi(B)$, 则称 Φ 为可加集函数.

2) 有限可加性 如果 $\forall A_1, \cdots, A_n \in \mathscr{C}$ 两两不交且 $\sum\limits_{i=1}^{n} A_i \in \mathscr{C}$, 有

$$\Phi\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Phi(A_i),$$

则称 Φ 为有限可加集函数.

3) σ 可加性 如果 $\forall A_1, A_2, \cdots \in \mathscr{C}$ 两两不交且 $\sum\limits_{i=1}^{\infty} A_i \in \mathscr{C}$, 都有

$$\Phi\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Phi(A_i),$$

则称 Φ 为 σ 可加集函数.

4) 有限性 如果 $\forall A \in \mathscr{C}$, 有 $\Phi(A) \in \mathbb{R}$, 则称 Φ 是有限集函数.

5) σ 有限性 如果 $\forall A \in \mathscr{C}, \exists \{A_n\}_{n \geqslant 1} \subset \mathscr{C}$ 使得 $\Phi(A_n) \in \mathbb{R}$ $(\forall n \geqslant 1)$, 且 $A = \bigcup\limits_{n=1}^{\infty} A_n$, 则称 Φ 为 σ 有限集函数.

定义 1.29. 如果集函数 Φ 具 σ 可加性, 则称其为一个符号测度; 如果它只取非负值, 则称其为一个测度; 如果进一步有 $\Phi(\Omega) = 1$, 则称之为正规测度或概率测度. 如果 Φ 取非负值且具有有限可加性, 则称之为有限可加测度.

留意, 符号测度与有限可加测度均未必是测度, 这有点“白马非马”的味道. 下面关于集函数性质的命题是显然的.

命题 1.30. 设 Φ 是 $\mathscr{C}$ 上的集函数, 则

1) 有限可加性 $\Rightarrow$ 可加性;

2) 如果 $\emptyset \in \mathscr{C}$, 则 σ 可加性 $\Rightarrow$ 有限可加性;

3) 如果 $\mathscr{C}$ 是集代数, 则有限可加性 $\Leftrightarrow$ 可加性;

4) 如 Φ 可加且 $\emptyset \in \mathscr{C}$, 则 $\Phi(\emptyset) = 0$.

当集类的性质越好时, 其上的集函数的性质也越丰富. 下面我们讨论不同集类上集函数的性质.

性质 1.31. **1) (可减性)** 设 Φ 是集代数 $\mathscr{F}$ 上的可加集函数, $A, B \in \mathscr{F}$, $A \subset B$, 则 $\Phi(B) = \Phi(A) + \Phi(B - A)$. 如 $\Phi(A) < \infty$, 则 $\Phi(B - A) = \Phi(B) - \Phi(A)$.

2) (单调性) 设 μ 是半集代数 $\mathscr{S}$ 上的有限可加测度, 如果 $A \subset B, A, B \in \mathscr{S}$, 则 $\mu(A) \leqslant \mu(B)$.

3) (有限性) 设 Φ 是半集代数 $\mathscr{S}$ 上的有限可加集函数, 若 $\Phi(B) < \infty$, $A \subset B$, 则 $\Phi(A) < \infty$. 特别地, 如果 $\Phi(\Omega) < \infty$, 则 Φ 是有限集函数.

4) (σ 有限性) 设 Φ 是半集代数 $\mathscr{S}$ 上的有限可加集函数, 如果 $\Omega = \bigcup\limits_{n=1}^{\infty} A_n$, $A_n \in \mathscr{S}, \Phi(A_n) < \infty \ (\forall n \geqslant 1)$, 则对 $\forall A \in \mathscr{S}, \exists \{A'_n\} \subset \mathscr{S}$ 两两不交, 使得 $A = \sum\limits_{i=1}^{\infty} A'_n$ 且 $\Phi(A'_n) < \infty \ (\forall n \geqslant 1)$.

证明 只证明 2) 和 4), 余者显然.

2) 由半集代数性质知, 存在 $A_1, \cdots, A_n \in \mathscr{S}$ 两两不交且与 A 不交, 使得 $B = A + A_1 + \cdots + A_n$. 由 μ 的有限可加性与非负性知 $\mu(B) = \mu(A) + \sum\limits_{i=1}^{n} \mu(A_i) \geqslant \mu(A)$.

4) 先证 Ω 可表示成可数多个 Φ 值有限的两两不交集的并. $\forall n \geqslant 1$, 令 $B_1 = A_1, B_n = A_n - \bigcup\limits_{k=1}^{n-1} A_k$, 由半集代数定义知 $\exists B_{n1}, \cdots, B_{nk_n} \in \mathscr{S}$ 两两不交, $B_n = \sum\limits_{i=1}^{k_n} B_{n_i}$, 因而 $\Omega = \sum\limits_{n=1}^{\infty} \sum\limits_{i=1}^{k_n} B_{ni} \overset{\text{重新编号}}{=} \sum\limits_{k=1}^{\infty} B'_k$, $\{B'_k\} \subset \mathscr{S}$ 两两不交. 由 3) 知 $\Phi(A_n) < \infty \ (\forall n \geqslant 1)$ 蕴含 $\Phi(B_k) < \infty \ (\forall k \geqslant 1)$. 因而令 $A'_n = A_n \cap B'_n$ 即可.

命题 1.32. **1) (次有限可加性)** 设 μ 是集代数 $\mathscr{F}$ 上的有限可加测度. 若 $A \in \mathscr{F}, A_1, \cdots, A_n \in \mathscr{F}, A \subset \bigcup\limits_{k=1}^{n} A_k$, 则 $\mu(A) \leqslant \sum\limits_{k=1}^{n} \mu(A_k)$;

2) (次 σ 可加性) 设 μ 是集代数 $\mathscr{F}$ 上的测度. 若 $A \in \mathscr{F}, A_1, \cdots, A_n, \cdots \in \mathscr{F}, A \subset \bigcup\limits_{n=1}^{\infty} A_n$, 则 $\mu(A) \leqslant \sum\limits_{n=1}^{\infty} \mu(A_n)$.

证明 1) 由归纳法知, 只需证明 $n = 2$ 的情形. 由单调性与可加性知

$$
\begin{aligned}
\mu(A) &\leqslant \mu(A_1 \cup A_2) = \mu(A_1 + (A_2 - A_1)) \\
&= \mu(A_1) + \mu(A_2 - A_1) \leqslant \mu(A_1) + \mu(A_2).
\end{aligned}
$$

2) 令 $A_0=\emptyset$. 由单调性与 σ 可加性知

$$\begin{aligned}\mu(A)&=\mu\left(\bigcup_{n=1}^{\infty}A_n\cap A\right)=\mu\left(\sum_{n=1}^{\infty}A\cap\left(A_n-\bigcup_{i\leqslant n-1}A_i\right)\right)\\&=\sum_{n=1}^{\infty}\mu\left(A_n-\bigcup_{i\leqslant n-1}A_i\right)\leqslant\sum_{n=1}^{\infty}\mu(A_n).\quad\square\end{aligned}$$

定义 1.33. 设 Φ 是集类 $\mathscr{C}$ 上的集函数. 若任给 $A\in\mathscr{C}$ 及序列 $\mathscr{C}\ni A_n\uparrow A$ 有 $\lim\limits_{n\to\infty}\Phi(A_n)=\Phi(A)$, 则称 Φ 在 A 处下连续. 若任给 $A\in\mathscr{C}$ 及序列 $\mathscr{C}\ni A_n\downarrow A$ 使得 $\Phi(A_n)<\infty$ 对某 n 成立, 都有 $\lim\limits_{n\to\infty}\Phi(A_n)=\Phi(A)$, 则称 Φ 在 A 处上连续. 在 A 处既上连续又下连续, 则称 Φ 在 A 处连续. 如果 Φ 在 $\mathscr{C}$ 处处连续, 则称 Φ 为连续集函数.

留意, 对于上连续要求 $\exists n$ 使得 $\Phi(A_n)<\infty$. 因为如不加此限制, 许多简单的情形都被排除了. 如 $\Omega=\mathbb{R}$, Φ 为 Lebesgue 测度, $A_n=(n,\infty)$, 则 A_n 单减趋于 $\emptyset$, 但 $\Phi(\emptyset)=0\neq\infty=\lim\limits_{n\to\infty}\Phi(A_n)$.

定理 1.34. 设 Φ 是集代数 $\mathscr{F}$ 上的 σ 可加集函数, 则 Φ 连续.

证明 由于 $\emptyset\in\mathscr{F}$, 从而 Φ 有限可加. 如 $\mathscr{F}\ni A_n\uparrow A\in\mathscr{F}$, 则 $A=\bigcup\limits_{n=1}^{\infty}A_n=A_1+\sum\limits_{n=2}^{\infty}(A_n-A_{n-1})$. 如 $\exists n$ 使 $\Phi(A_n)=\infty$, 则 $\Phi(A)=\infty=\lim\limits_{n\to\infty}\Phi(A_n)$. 设 $\forall n$ 有 $\Phi(A_n)<\infty$, 则由 σ 可加性与可减性知

$$\begin{aligned}\Phi(A)&=\Phi(A_1)+\sum_{n=2}^{\infty}\Phi(A_n-A_{n-1})=\Phi(A_1)+\sum_{n=2}^{\infty}[\Phi(A_n)-\Phi(A_{n-1})]\\&=\Phi(A_1)+\lim_{n\to\infty}\sum_{k=2}^{n}[\Phi(A_k)-\Phi(A_{k-1})]=\lim_{n\to\infty}\Phi(A_n),\end{aligned}$$

从而 Φ 下连续. 另一方面, 设 A_n 单减趋于 A, 且 $\exists n_0$ 使得 $\Phi(A_{n_0})<\infty$. 则 $A_{n_0}-A_n$ 单增趋于 $A_{n_0}-A$, 故由下连续性知 $\Phi(A_{n_0}-A_n)\to\Phi(A_{n_0}-A)$. 由此及可减性得 $\Phi(A_n)\to\Phi(A)$. $\square$

推论 1.35. 集代数上的测度必连续.

下面的定理表明, 当 Φ 具有可加性时, 连续性也可导出 σ 可加性. 由此, 结合定理 1.34, 连续性通常可作为 σ 可加性的等价性质来使用.

定理 1.36. 设 Φ 是集代数 $\mathscr{F}$ 上的有限可加集函数, 若 Φ 满足下列条件之一, 则 Φ 是 σ 可加的.

a) Φ 下连续,

b) Φ 是有限集函数, 在空集 $\emptyset$ 处连续.

证明 设 a) 成立. 若 $\{A_n\}_{n\geqslant 1}\subset\mathscr{F}$ 两两不交且 $A=\sum\limits_{n=1}^{\infty}A_n\in\mathscr{F}$, 令 $B_n=\sum\limits_{k=1}^{n}A_k$, 则 B_n 单增趋于 A. 由下连续性与有限可加性知,

$$\Phi(A)=\lim_{n\to\infty}\Phi(B_n)=\lim_{n\to\infty}\Phi\left(\sum_{k=1}^{n}A_k\right)=\lim_{n\to\infty}\sum_{k=1}^{n}\Phi(A_k)=\sum_{k=1}^{\infty}\Phi(A_k).$$

设 b) 成立, $\{A_n\}_{n\geqslant 1}$ 与 $\{B_n\}_{n\geqslant 1}$ 如上. 由于 $A-B_n\in\mathscr{F}$ 且 $A-B_n\downarrow\emptyset$, 则由在 $\emptyset$ 处的连续性与可减性知 $0=\lim\limits_{n\to\infty}\Phi(A-B_n)=\Phi(A)-\lim\limits_{n\to\infty}\Phi(B_n)$. 故 $\Phi(A)=\sum\limits_{k=1}^{\infty}\Phi(A_k)$. $\square$

§ 1.2.2 测度空间

定义 1.37. 设 $\mathscr{A}$ 是 Ω 中的 σ 代数, μ 是 $\mathscr{A}$ 上的测度, 则称 $(\Omega,\mathscr{A})$ 为可测空间, $(\Omega,\mathscr{A},\mu)$ 为测度空间. 特别地, 如果 μ 是概率测度, 则称 $(\Omega,\mathscr{A},\mu)$ 为概率空间. $\mathscr{A}$ 中集合称为 $\mathscr{A}$ 可测集, 或简称可测集. 如 μ 为概率测度, 也称 $\mathscr{A}$ 中集合为事件. 通常用 $\mathbb{P}$ 表示概率测度.

由前面介绍的测度的性质容易得到下面的概率性质. 这些性质在学习初等概率论时已未加证明地介绍过. 设 $(\Omega,\mathscr{A},\mathbb{P})$ 为概率空间, 则 $\mathbb{P}$ 满足:

1) 非负性 $\mathbb{P}(A)\geqslant 0,\ \forall A\in\mathscr{A}$.

2) 正规性 $\mathbb{P}(\Omega)=1$.

3) σ 可加性 (从而有限可加性) $\mathbb{P}\left(\sum\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)$.

4) 可减性 (从而单调性) $A\subset B, A,B\in\mathscr{A}$, 则 $\mathbb{P}(B-A)=\mathbb{P}(B)-\mathbb{P}(A)\Rightarrow \mathbb{P}(B)\geqslant\mathbb{P}(A)$.

5) 加法公式$\mathbb{P}(A\bigcup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\bigcap B)$. 一般地, $\forall\,\{A_n\}_{n=1}^{\infty}\subset\mathscr{A}$, 有

$$\mathbb{P}\left(\bigcup_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}\mathbb{P}(A_k)-\sum_{1\leqslant<j\leqslant n}\mathbb{P}(A_i\cap A_j)+\cdots+(-1)^{n-1}\mathbb{P}(A_1\cap\cdots\cap A_n).$$

6) 连续性 设 $A, A_n \in \mathscr{A}, n \geqslant 1$, 则 $A_n \uparrow A \Rightarrow \mathbb{P}(A_n) \uparrow \mathbb{P}(A)$; $A_n \downarrow A \Rightarrow \mathbb{P}(A_n) \downarrow \mathbb{P}(A)$.

例 1.38 (几何概型). $\Omega \subset \mathbb{R}$ 是 Lebesgue 可测集且 $0 < |\Omega| < \infty$, 其中 $|\cdot|$ 表示 Lebesgue 测度. 设 $\mathscr{A}$ 是 Ω 中所有的 Lebesgue 可测子集组成的集类, $\mathbb{P}(A) = \frac{|A|}{|\Omega|}, A \in \mathscr{A}$. 则 $(\Omega, \mathscr{A}, \mathbb{P})$ 是一个概率空间.

§1.3　测度扩张定理及测度的完全化

正如本章开始部分所解释的, 测度通常仅方便在一个小的集类 (如半集代数) 上定义, 然后再将其扩张到 σ 代数上. 本节主要将实变函数论中定义的 Lebesgue 测度的思想抽象出来, 以得到重要的测度扩张定理. 我们先将半集代数上的测度扩张到最小集代数上, 再进一步扩张到最小 σ 代数上.

§ 1.3.1　半集代数上的测度扩张为最小集代数上的测度

定义 1.39. 设 $\mathscr{C}_1 \subset \mathscr{C}_2$ 是 Ω 的两个子集类, μ_i 是定义在 $\mathscr{C}_i\,(i=1,2)$ 上的测度 (或有限可加测度). 如果对 $\forall A \in \mathscr{C}_1$, 有 $\mu_1(A) = \mu_2(A)$, 则称 μ_2 是 μ_1 在 $\mathscr{C}_2$ 上的扩张, 而称 μ_1 是 μ_2 在 $\mathscr{C}_1$ 上的限制, 记作 $\mu_1 = \mu_2|_{\mathscr{C}_1}$.

定理 1.40. 设 μ 是半集代数 $\mathscr{S}$ 上的测度 (或有限可加测度), 则 μ 在 $\mathscr{F}(\mathscr{S})$ 上存在唯一的扩张 $\widetilde{\mu}$.

证明 由定理 1.7 知, $\forall A \in \mathscr{F}(\mathscr{S})$, $\exists B_1, \cdots, B_n \in \mathscr{S}$ 两两不交, 满足 $A = \sum\limits_{i=1}^{n} B_i$. 定义 $\widetilde{\mu}(A) = \sum\limits_{i=1}^{n} \mu(B_i)$. 先证明 $\widetilde{\mu}(A)$ 与 $\{B_i\}$ 的选取无关. 如还有 $B'_1, \cdots, B'_{n'} \in \mathscr{S}$ 两两不交, 使得 $A = \sum\limits_{i=1}^{n'} B'_i$, 则 $B'_i = \sum\limits_{j=1}^{n} B'_i \cap B_j$. 由于 $B'_i \cap B_j \in \mathscr{S}$, 由有限可加性知, $\mu(B'_i) = \sum\limits_{j=1}^{n} \mu(B'_i \cap B_j)$. 从而

$$\begin{aligned}\sum_{i=1}^{n'} \mu(B'_i) &= \sum_{i=1}^{n'} \sum_{j=1}^{n} \mu(B'_i \cap B_j) = \sum_{j=1}^{n} \sum_{i=1}^{n'} \mu(B'_i \cap B_j) \\ &= \sum_{j=1}^{n} \mu(B_j) = \widetilde{\mu}(A).\end{aligned}$$

故 $\widetilde{\mu}(A)$ 与 $\{B_i\}$ 的选取无关.

再证 $\widetilde{\mu}$ 是测度 (有限可加测度). 非负性与唯一性显然, 有限可加性也显然, 今证明 σ 可加性. 设 $A_1, A_2, \cdots \in \mathscr{F}(\mathscr{S})$ 两两不交, 且 $A = \sum\limits_{n=1}^{\infty} A_n \in \mathscr{F}(\mathscr{S})$, 取 $B_1, \cdots, B_k \in \mathscr{S}$ 两两不交, 满足 $A = \sum\limits_{i=1}^{k} B_i$. 另外, $\forall n \geqslant 1$, 取 $C_{n1}, \cdots, C_{nk_n} \in \mathscr{S}$ 两两不交, 满足 $A_n = \sum\limits_{i=1}^{k_n} C_{ni}$. 则 $\forall i \leqslant k, B_i = \sum\limits_{n=1}^{\infty} A_n \cap B_i = \sum\limits_{n=1}^{\infty} \sum\limits_{l=1}^{k_n} B_i \cap C_{nl}$ 是 $\mathscr{S}$ 中两两不交集合之并. 由 μ 的 σ 可加性知 $\mu(B_i) = \sum\limits_{j=1}^{\infty} \sum\limits_{l=1}^{k_j} \mu(B_i \cap C_{jl})$. 由此及有限可加性知

$$\begin{aligned}
\widetilde{\mu}(A) = \widetilde{\mu}\left(\sum_{i=1}^{k} B_i\right) &= \sum_{i=1}^{k} \sum_{n=1}^{\infty} \sum_{l=1}^{k_n} \mu(B_i \cap C_{nl}) \\
&= \sum_{n=1}^{\infty} \sum_{l=1}^{k_n} \sum_{i=1}^{k} \mu(B_i \cap C_{nl}) = \sum_{n=1}^{\infty} \widetilde{\mu}(A_n). \quad \square
\end{aligned}$$

推论 1.41. 设 μ 是半集代数 $\mathscr{S}$ 上的有限可加测度, $A, A_1, \cdots, A_n \in \mathscr{S}$.

a) 如 $A_1, \cdots, A_n$ 两两不交且 $\sum\limits_{i=1}^{n} A_i \subset A$, 则 $\sum\limits_{i=1}^{n} \mu(A_i) \leqslant \mu(A)$.

b) 如 $\bigcup\limits_{i=1}^{n} A_i \supset A$, 则 $\sum\limits_{i=1}^{n} \mu(A_i) \geqslant \mu(A)$.

当 μ 为 σ 可加时, 上述结论对 $n = \infty$ 也成立.

§ 1.3.2 半集代数、集代数上的测度扩张为最小 σ 代数上的测度

定理 1.42 (测度扩张定理)**.** 设 μ 是 Ω 中半集代数 $\mathscr{S}$ 上的测度, 则 μ 在 $\sigma(\mathscr{S})$ 上存在一个扩张. 若 μ 是 σ 有限的, 则该扩张是唯一的.

如同实变函数论中对于 Lebesgue 测度的处理, 我们先用覆盖的手法定义 Ω 每个子集的外测度, 再证明此外测度限于最小 σ 代数时具有 σ 可加性, 从而是一个测度. 下面外测度的定义手法与实变函数论中的完全一致.

定义 1.43. 设 μ 是半集代数 $\mathscr{S}$ 上的测度. 任意 $A \subset \Omega$, 称

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) : A \subset \bigcup_{n=1}^{\infty} A_n, A_n \in \mathscr{S}\right\}$$

为 A 的外测度, 而定义在最大 σ 代数上的集函数 μ^* 称为由 μ 生成的外测度.

性质 1.44. **1)** $\mu^*|_{\mathscr{S}} = \mu$;

2) $\mu^*(A) \leqslant \mu^*(B), \forall A \subset B$;

3) $\mu^*\left(\bigcup\limits_{n=1}^{\infty} A_n\right) \leqslant \sum\limits_{n=1}^{\infty} \mu^*(A_n)$, $\forall A_n \subset \Omega, n \geqslant 1$.

证明 1) 由于 $A \subset A$, 令 $A_1 = A, A_n = \emptyset, n \geqslant 2$, 则 $\mu^*(A) \leqslant \mu(A)$. 另一方面, 由 μ 的 σ 次可加性, 任给序列 $\{A_n\} \subset \mathscr{S}$, $\bigcup\limits_{n=1}^{\infty} A_n \supset A$, 有 $\mu(A) \leqslant \sum\limits_{n=1}^{\infty} \mu(A_n)$. 从而 $\mu^*(A) \geqslant \mu(A)$.

2) 显然.

3) 任给 $\varepsilon > 0$ 及 $n \geqslant 1$, 取 $A_{n1}, A_{n2}, \cdots \in \mathscr{S}$, 使 $\bigcup\limits_{i=1}^{\infty} A_{ni} \supset A_n$ 且 $\mu^*(A_n) \geqslant \sum\limits_{i=1}^{\infty} \mu(A_{ni}) - \frac{\varepsilon}{2^n}$. 从而 $\bigcup\limits_{n=1}^{\infty}\bigcup\limits_{i=1}^{\infty} A_{ni} \supset \bigcup\limits_{n=1}^{\infty} A_n$, 且由 μ^* 的定义知

$$\begin{aligned}\mu^*(\bigcup_{n=1}^{\infty} A_n) &\leqslant \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \mu(A_{ni}) \leqslant \sum_{n=1}^{\infty}(\mu^*(A_n) + \frac{\varepsilon}{2^n}) \\ &= \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon.\end{aligned}$$

令 $\varepsilon \downarrow 0$ 即可. □

如果 μ^* 是最大 σ 代数上的测度, 那么存在性问题就解决了, 只需令 $\widetilde{\mu}$ 为 $\mu^*|_{\sigma(\mathscr{S})}$ 即可. 然而这是不可能的, 在实变函数论中已经有反例. 为保证 μ^* 的 σ 可加性, 我们需要将 σ 代数缩小, 即寻找 σ 代数 $\mathscr{A}^* \supset \sigma(\mathscr{S})$ 使 μ^* 在 $\mathscr{A}^*$ 上为测度.

参照实变函数论中选择可测集的方法, 我们选择那些在外测度意义下可以分割其它集合而保持可加性的集合作为可测的集合.

定义 1.45. 设 $A \subset \Omega$, 如果 $\forall D \subset \Omega$ 有

$$\mu^*(D) = \mu^*(A \cap D) + \mu^*(A^c \cap D).$$

则称 A 是 μ^* 可测的.

令 $\mathscr{A}^* = \big\{A \subset \Omega : A \text{ 是 } \mu^* \text{ 可测集}\big\}$. 我们需要证明: $\mathscr{A}^*$ 是包含 $\mathscr{S}$ 的 σ 代数, μ^* 是 $\mathscr{A}^*$ 上的测度. 为此, 我们先讨论 μ^* 和 $\mathscr{A}^*$ 的性质. 由性质 1.44 之 3), 可将 μ^* 可测集的定义作如下简化.

性质 1.46. A 是 μ^* 可测集当且仅当 $\forall\ D \subset \Omega$, 有 $\mu^*(D) \geqslant \mu^*(A \cap D) + \mu^*(A^c \cap D)$.

性质 1.47. $\mathscr{A}^* \supset \mathscr{S}$.

证明 设 $A \in \mathscr{S}, D \subset \Omega$. $\forall \varepsilon > 0$, 取 $\{A_n\} \subset \mathscr{S}$ 使 $\sum\limits_{n=1}^{\infty} A_n \supset D, \mu^*(D) \geqslant \sum\limits_{n=1}^{\infty} \mu(A_n) - \varepsilon$. 则

$$\begin{aligned}\mu^*(A^c \cap D) + \mu^*(A \cap D) &\leqslant \sum_{n=1}^{\infty} [\mu(A_n \cap A) + \mu(A^c \cap A_n)] \\ &= \sum_{n=1}^{\infty} \mu(A_n) \leqslant \mu^*(D) + \varepsilon.\end{aligned}$$

令 $\varepsilon \downarrow 0$ 得 $\mu^*(D) \geqslant \mu^*(A \cap D) + \mu^*(A^c \cap D)$. 故由性质 1.46 知 $A \in \mathscr{A}^*$. □

定理 1.48. **(1)** $\mathscr{A}^*$ 是 σ 代数且 $\mathscr{A}^* \supset \sigma(\mathscr{S})$;

(2) 若 $\{A_n\} \subset \mathscr{A}^*$ 两两不交, $A = \sum\limits_{n=1}^{\infty} A_n$, 则 $\forall D \subset \Omega, \mu^*(D \cap A) = \sum\limits_{n=1}^{\infty} \mu^*(D \cap A_n)$;

(3) μ^* 在 $\mathscr{A}^*$ 上的限制是 $\mathscr{A}^*$ 上的测度;

证明 (1) 先证 $\mathscr{A}^*$ 是集代数. 因为 $\mathscr{A}^* \supset \mathscr{S}$, 故 $\emptyset, \Omega \in \mathscr{A}^*$. 易见 $A \in \mathscr{A}^*$ 蕴含 $A^c \in \mathscr{A}^*$. 现只需证 $A, B \in \mathscr{A}^* \Rightarrow A \cap B \in \mathscr{A}^*$. 事实上, 由次可加性知

$$\begin{aligned}\mu^*(D) &= \mu^*(A \cap D) + \mu^*(A^c \cap D) \\ &= \mu^*(A \cap B \cap D) + \mu^*(A \cap B^c \cap D) + \mu^*(A^c \cap D) \\ &\geqslant \mu^*(A \cap B \cap D) + \mu^*((A^c \cup B^c) \cap D).\end{aligned}$$

故由性质 1.46 知 $A \cap B \in \mathscr{A}$.

再证 $\mathscr{A}^*$ 是单调类. $\forall A_n \uparrow A, A_n \in \mathscr{A}^*$. 由次可加性有 (令 $A_0 = \emptyset$)

$$\begin{aligned}\mu^*(D) &= \mu^*(A_1 \cap D) + \mu^*(A_1^c \cap D) \\ &= \mu^*(A_1 \cap D) + \mu^*(A_2 \cap A_1^c \cap D) + \mu^*(D \cap A_2^c) \\ &= \cdots = \sum_{n=1}^{n} \mu^*((A_i - A_{i-1}) \cap D) + \mu^*(D \cap A_n^c) \qquad (1.3.1) \\ &\geqslant \sum_{i=1}^{n} \mu^*((A_i - A_{i-1}) \cap D) + \mu^*(D \cap A^c).\end{aligned}$$

令 $n\to\infty$ 得

$$\mu^*(D)\geqslant\mu^*(D\cap A^c)+\sum_{i=1}^{\infty}\mu^*(D\cap(A_{i-1}-A_i))\geqslant\mu^*(D\cap A^c)+\mu^*(D\cap A).$$

从而 $A\in\mathscr{A}^*$. 由单调类定理知 $\mathscr{A}^*$ 是 σ 代数.

(2) 设 $A=\sum\limits_{n=1}^{\infty}A_n, A_n\in\mathscr{A}^*$ 两两不交, 则 $A\in\mathscr{A}^*$. 由 σ 次可加性, 只需证明 $\mu^*(D\cap A)\geqslant\sum\limits_{n=1}^{\infty}\mu^*(D\cap A_n)$. 在 (1.3.1) 中令 $A=D$ 并以 $A\cap D$ 代替 D, 以 $\sum\limits_{i=1}^{n}A_i$ 代替 A_n, 得 $\mu^*(D\cap A)\geqslant\sum\limits_{i=1}^{n}\mu^*(D\cap A_n)$. 令 $n\uparrow\infty$ 即得所求.

(3) 在 (2) 中令 $D=\Omega$ 即得 μ^* 在 $\mathscr{A}^*$ 上的 σ 可加性. □

定理 1.42 之证. 因为 $\mathscr{A}^*\supset\sigma(\mathscr{S})$, 将 μ^* 限于 $\sigma(\mathscr{S})$ 上显然仍是测度, 且 $\mu^*(A)=\mu(A), A\in\mathscr{S}$. 从而 $\exists\mu$ 在 $\sigma(\mathscr{S})$ 上扩张. 设 μ 在$\mathscr{S}$ 上是 σ 有限的, 则存在 $\{A_n\}\subset\mathscr{S}$ 两两不交使得 $\Omega=\sum\limits_{n=1}^{\infty}A_n$ 且 $\mu(A_n)<\infty, n\geqslant 1$. 如 μ_1,μ_2 都是 μ 在 $\sigma(\mathscr{S})$ 上的扩张, 只需证 $\mu_1(A\cap A_n)=\mu_2(A\cap A_n), A\in\sigma(\mathscr{S}), n\geqslant 1$. 为此令 $\mathscr{M}_n=\{A:A\in\sigma(\mathscr{S}),\ \mu_1(A\cap A_n)=\mu_2(A\cap A_n)\}$, 则 $\mathscr{M}_n\supset\mathscr{S}$. 由 μ 在 $\mathscr{F}(\mathscr{S})$ 上扩张唯一性知 $\mathscr{M}_n\supset\mathscr{F}(\mathscr{S})$, 故由单调类定理, 只需证明 $\mathscr{M}$ 是单调类. 这由测度的连续性立得. □

推论 1.49. 若 $\mathscr{S}$ 是 Ω 中的半集代数, $\mathbb{P}$ 是 $\mathscr{S}$ 上的概率, 则 $\mathbb{P}$ 在 $\sigma(\mathscr{S})$ 上有唯一的扩张.

§ 1.3.3　测度的完全化

定义 1.50. 设 $(\Omega,\mathscr{A},\mu)$ 是一个测度空间. 如果存在 $A\in\mathscr{A}$ 使 $B\subset A$ 且 $\mu(A)=0$, 则称 B 为 μ 零测集. 如果所有 μ 零测集都属于 $\mathscr{A}$, 则称 $(\Omega,\mathscr{A},\mu)$ 是完全测度空间.

定理 1.51. 设 $(\Omega,\mathscr{A},\mu)$ 是一个测度空间, 令

$$\bar{\mathscr{A}}=\{A\cup N: A\in\mathscr{A}, N\ \text{为}\ \mu\ \text{零测集}\},$$

且 $\bar{\mu}(A\cup N)=\mu(A), A\in\mathscr{A}$, N 为 μ 零测集. 则 $(\Omega,\bar{\mathscr{A}},\bar{\mu})$ 是完全测度空间, 称之为 $(\Omega,\mathscr{A},\mu)$ 的完全化.

证明 由次可加性知可数个零测集的并仍为零测集, 故 $\bar{\mathscr{A}}$ 对可数并封闭. 只需证 $\bar{\mathscr{A}}$ 对余封闭. 设 $A\cup N\in\bar{\mathscr{A}}, A\in\mathscr{A}, N$ 为零测集. 设$B\in\mathscr{A}$ 使 $B\supset N$ 且 $\mu(B)=0$, 则 $(A\cup N)^c=A^c\cap N^c=A^c\cap B^c+A^c\cap(N^c-B^c)$. 由于 $A^c\cap(N^c-B^c)\subset\Omega-B^c=B$, 而 $\mu(B)=0$, 则 $A^c\cap(N^c-B^c)$ 是零测集. 又 $A^c\cap B^c\in\mathscr{A}$, 故$(A\cup N)^c\in\mathscr{A}$.

再证 $(\Omega,\bar{\mathscr{A}},\bar{\mu})$ 的完全性. 设 $\bar{N}$ 是 $\bar{\mu}$ 零测集, 则 $\exists\bar{B}\in\bar{\mathscr{A}}$ 使 $\bar{\mu}(\bar{B})=0$ 且 $\bar{B}\supset\bar{N}$. 设 $\bar{B}=A\cup N$, $A\in\mathscr{A}$, N 为 μ 零测集. 则 $0=\bar{\mu}(\bar{B})=\mu(A)$. 取 $B\in\mathscr{A}, B\supset N$ 使 $\mu(B)=0$. 我们有 $\bar{N}\subset\bar{B}\subset A\cup B$, 且 $\mu(A\cup B)=0$. 则 $\bar{N}$ 是 μ 零测集, 从而 $\bar{N}\in\bar{\mathscr{A}}$. □

定理 1.52. 若 μ 是半集代数 $\mathscr{S}$ 上的测度, μ^* 是由 μ 生成的外测度. 若 $A\subset\Omega$ 且 $\mu^*(A)<\infty$, 则 $\exists B\in\sigma(\mathscr{S})$, 使: (i) $A\subset B$, (ii) $\mu^*(A)=\mu(B)$, (iii) $\forall C\subset B-A$ 且 $C\in\sigma(\mathscr{S})$, 有 $\mu^*(C)=0$.

称这样的 B 为 A 的可测覆盖.

证明 $\forall n\geqslant 1$, 取 $\{F_{n_k}\}_{k\geqslant 1}\subset\mathscr{S}$ 满足 $A\subset\bigcup\limits_{k=1}^{\infty}F_{n_k}$, $\sum\limits_{k=1}^{\infty}\mu(F_{n_k})\leqslant\mu^*(A)+\frac{1}{n}$. 令 $B_n=\bigcup\limits_{k=1}^{\infty}F_{n_k}$, 则 $\mu^*(A)\leqslant\mu^*(B_n)$. 令 $B=\bigcap\limits_{n=1}^{\infty}B_n$, 则 $B\in\sigma(\mathscr{S}), B\supset A$ 且 $\mu^*(B)=\mu^*(A)$. 若 $C\in\sigma(\mathscr{S})$, $C\subset B-A$, 则 $A\subset B-C$. 故 $\mu^*(A)\leqslant\mu^*(B-C)=\mu(B)-\mu(C)$. 由此及 $\mu^*(B)=\mu^*(A)<\infty$ 知 $\mu^*(C)=0$. □

定理 1.53. 设 μ 是半集代数 $\mathscr{S}$ 上的 σ 有限测度, μ^* 是由 μ 生成的外测度, 则 $(\Omega,\mathscr{A}^*,\mu^*)$ 是 $(\Omega,\sigma(\mathscr{S}),\mu)$ 的完全化.

证明 只需证 $\mathscr{A}^*=\bar{\mathscr{A}}$. 设 $\bar{A}\in\bar{\mathscr{A}}$, 则 $\exists A\in\sigma(\mathscr{S})$, N 为 μ 零测度使 $\bar{A}=A\cup N$, 易见 $\mathscr{A}^*$ 包含所有 μ 零测集, 从而 $\bar{A}\in\mathscr{A}^*$.

反之, 如 $A\in\mathscr{A}^*, \mu^*(A)<\infty$, 令 B 是 A 的可测覆盖且 C 是 $B-A$ 的可测覆盖. 则 $A=(B-C)\cup(C-(B-A))$, 其中 $B-C\in\sigma(\mathscr{S}), C-(B-A)$ 为 μ 零测集. 故 $A\in\bar{\mathscr{A}}$. 当 $\mu^*(A)=\infty$ 时, 由 μ 的 σ 有限性知 $\exists\{A_n\}_{n\geqslant 1}\subset\mathscr{S}$ 使 $\sum\limits_{n=1}^{\infty}A_n=\Omega$ 且 $\mu(A_n)<\infty, n\geqslant 1$. 由前面论证知 $A\cap A_n\in\bar{\mathscr{A}}, n\geqslant 1$. 故 $A=\sum\limits_{n=1}^{\infty}A_n\cap A\in\bar{\mathscr{A}}$. □

定理 1.54. 设 μ 是半集代数 $\mathscr{S}$ 上的 σ 有限测度, 则 μ 在 $\mathscr{A}^*$ 上的扩张是唯一的.

证明 由于 μ 是 σ 有限的, μ 在 $\sigma(\mathscr{S})$ 上的扩张是唯一的, 即为 μ^* 在 $\sigma(\mathscr{S})$ 上之限制. 如存在 μ 的另一个 $\mathscr{A}^*$ 上的扩张 μ_1, $\forall A \cup N \in \mathscr{A}^* = \bar{\mathscr{A}}, A \in \sigma(\mathscr{S})$ 且 N 为 μ 零测集, 有

$$\begin{aligned}\mu^*(A \cup N) = \mu^*(A) = \mu_1(A) \leqslant \mu_1(A \cup N) \leqslant \mu_1(A) + \mu_1(N) \\ = \mu^*(A) + \mu_1(N) = \mu^*(A \cup N) + \mu_1(N).\end{aligned}$$

设 $B \in \sigma(\mathscr{S})$ 使 $B \supset N, \mu(B) = 0$, 有 $\mu_1(N) \leqslant \mu_1(B) = \mu(B) = 0$. 从而 $\mu^*(A \cup N) = \mu_1(A \cup N)$. $\square$

定理 1.55. 设 μ 是半集代数 $\mathscr{S}$ 上的测度, 则 $\forall A \in \mathscr{A}^*$ 使 $\mu^*(A) < \infty$, $\forall \varepsilon > 0$, $\exists A_\varepsilon \in \mathscr{F}(\mathscr{S})$ 使 $\mu^*(A \Delta A_\varepsilon) < \varepsilon$.

证明 $\forall \varepsilon > 0, \exists \{B_n\}_{n \geqslant 1} \subset \mathscr{S}$ 使 $\bigcup\limits_{n=1}^{\infty} B_n \supset A$ 且 $\mu^*(A) \leqslant \sum\limits_{n=1}^{\infty} \mu^*(B_n) \leqslant \mu^*(A) + \frac{\varepsilon}{2}$. 因为 $\mu^*(A) < \infty$, 故 $\sum\limits_{n=1}^{\infty} \mu^*(B_n) < \infty$. 取 $n_0 \geqslant 1$ 使 $\sum\limits_{n > n_0} \mu^*(B_n) < \frac{\varepsilon}{2}$. 令 $A_\varepsilon = \sum\limits_{n=1}^{n_0} B_n$, $B_\varepsilon = \sum\limits_{n > n_0} B_n$. 则 $A_\varepsilon \in \mathscr{F}(\mathscr{S})$. 由 μ^* 的 σ 次可加性知 $\mu^*(B_\varepsilon) < \frac{\varepsilon}{2}$. 从而由单调性知 $\mu^*((A_\varepsilon \cup B_\varepsilon) - A) < \frac{\varepsilon}{2}$. 由于 $A_\varepsilon - A \subset B_\varepsilon$ 且 $A - A_\varepsilon \subset (A_\varepsilon \cup B_\varepsilon) - A$, 则 $\mu^*(A \Delta A_\varepsilon) = \mu^*((A - A_\varepsilon) + (A_\varepsilon - A)) < \varepsilon$. $\square$

§1.4　补充与习题

1. 证明命题 1.6.

2. 证明性质 1.10.

3. 设 $\mathscr{C}$ 是集类, 则 $\forall A \in \sigma(\mathscr{C})$, 存在 $\mathscr{C}$ 的可数子集类 $\mathscr{C}_1$ 使 $A \in \sigma(\mathscr{C}_1)$.

4. (可数生成) σ 代数 $\mathscr{A}$ 称为可数生成的, 如果存在可数的子集类 $\mathscr{C}$ 使得 $\sigma(\mathscr{C}) = \mathscr{A}$. 证明 $\mathscr{B}^d$ 是可数生成的.

5. 设 $\mathscr{C}_n$ 是单调上升的子集类.

   (a) 若 $\mathscr{C}_n$ 是代数, 则 $\bigcup\limits_{n=1}^{\infty} \mathscr{C}_n$ 是代数;

   (b) 若 $\mathscr{C}_n$ 是 σ 代数, 举例说明 $\bigcup\limits_{n=1}^{\infty} \mathscr{C}_n$ 可以不是 σ 代数;

6. 证明定理 1.19.

7. 设 $\mathscr{C}$ 为子集类. 证明若 $A \in \sigma(\mathscr{C})$, 则存在可数子类 $\mathscr{C}_A \subset \mathscr{C}$ 使得 $A \in \sigma(\mathscr{C}_A)$.

8. 证明 σ 代数不可能是可数无穷的, 即要么是有限的, 要么是不可数的.

9. 设 $\Omega_i, i = 1, \cdots, n$ 是 n 个集合(空间), $\mathscr{A}_i$ 是 Ω_i 中的 σ 代数. 证明 $\mathscr{C} = \{A_1 \times \cdots \times A_n : A_i \in \mathscr{A}_i\}$, $\mathscr{C}$ 为半集代数.

10. 证明定理 1.27.

11. 举例说明一个集类上的可加测度未必是有限可加的.

12. 举例说明半集代数 $\mathscr{S}$ 生成的 σ 代数不能一般性的表述为

$$\sigma(\mathscr{S}) = \left\{\sum_{n=1}^{\infty} A_n : \forall n \geqslant 1, A_n \in \mathscr{S}\right\}.$$

    但如果 Ω 至多可数时, 如上的表述是正确的.

13. 设 $(\Omega_n, \mathscr{A}_n, \mu_n), n \geqslant 1$ 是一列测度空间, Ω_n 互不相交. 令 $\Omega = \sum\limits_{n=1}^{\infty} \Omega_n, \mathscr{A} = \{A \subset \Omega : \forall n \geqslant 1, A \cap \Omega_n \in \mathscr{A}_n\}, \mu(A) = \sum\limits_{n=1}^{\infty} \mu_n(A \cap \Omega_n), A \in \mathscr{A}$. 证明 $(\Omega, \mathscr{A}, \mu)$ 是一个测度空间.

14. 设 Ω 为一无穷集, 令 $\mathscr{F}$ 为由 Ω 中的有限集或余集为有限集的集合组成的集类, 而对此两类集合, $\mathbb{P}$ 分别取值 0 或 1.

    (a) 证明 $\mathscr{F}$ 是代数, $\mathbb{P}$ 是有限可加的;

    (b) 若 Ω 为可数无穷集, 则 $\mathbb{P}$ 不可能可数可加;

    (c) 若 Ω 为不可数集, 则 $\mathbb{P}$ 是可数可加.

15. 证明命题 1.30.

16. 设 Ω 为一不可数集, 令 $\mathscr{F}$ 为由 Ω 中的可数集或余集为可数集的集合组成的集类, 而对此两类集合, $\mathbb{P}$ 分别取值 0 或 1. 证明 $\mathbb{P}$ 是可数可加的.

17. 概率空间 $(\Omega, \mathscr{A}, \mathbb{P})$ 称为非原子的, 如果 $\mathbb{P}(A) > 0, A \in \mathscr{A}$, 那么存在 $B \in \mathscr{A}$ 使得 $B \subset A, 0 < \mathbb{P}(B) < \mathbb{P}(A)$.

    (a) 证明若 $\mathbb{P}(A) > 0, \varepsilon > 0$, 那么存在 $B \in \mathscr{A}$ 使得 $B \subset A, 0 < \mathbb{P}(B) < \varepsilon$;

    (b) 证明若 $\mathbb{P}(A) \geqslant x \geqslant 0$, 那么存在 $B \in \mathscr{A}$ 使得 $B \subset A, 0 < \mathbb{P}(B) = x$;

    (c) 证明存在 $\{B_1, B_2, \cdots\}$ 使得 $\mathbb{P}(A)$ 是 $\{\mathbb{P}(B_1), \mathbb{P}(B_2), \cdots\}$ 的凸组合;

    (d) 证明 $(\mathbb{R}, \mathscr{B})$ 上的 Lebesgue 测度 λ 是非原子的.

18. 证明推论 1.35.

19. 设 $([0,1], \mathscr{B}([0,1]), \mu)$ 是有限测度空间, 满足 $\mu(\{x\}) = 0, \forall x \in [0,1]$. $\forall \varepsilon >$, 证明

    (a) $\forall x \in [0,1]$, 存在区间 $I \ni x$ 使得 $\mu(I) \leqslant \varepsilon$;

(b) 存在 $[0,1]$ 的稠子集 A 使得 $\mu(A) \leqslant \varepsilon$.

20. 证明推论 1.41.

21. 举例说明 μ^* 未必是最大 σ 代数上的测度.

22. 证明性质 1.46.

23. 试构造反例说明当 μ 不是 σ 有限时, 它从半集代数扩张到最小 σ 代数上的扩张可能不唯一.

24. 设 μ^* 是 μ 生成的外测度. 则测度空间 $(\Omega, \mathscr{A}, \mu)$ 是完全的当且仅当 $\mathscr{A} \supset \{A \in \Omega : \mu^*(A) = 0\}$.

25. $\mathscr{S}$ 是半集代数, μ 是 $\mathscr{S}$ 上有限测度. 记 $(\Omega, \mathscr{A}^*, \mu^*)$ 是 μ 扩张至 $\sigma(\mathscr{S})$ 的完全化, 令

$$\mu_*(A) = \sup\left\{\sum_n \mu(A_n) : A_n \in \mathscr{S} \text{ 两两不交}, \sum_n A_n \subset A\right\},$$
$$\mathscr{A}_* = \{A \subset \Omega : \mu^*(A) = \mu_*(A)\}.$$

试证: $\mathscr{A}^* \supset \mathscr{A}_*$

26. 设 $(\Omega, \mathscr{A}, \mu)$ 为测度空间, μ^* 为由 μ 生成的外测度. 证明 $N \subset \Omega$ 为 μ 零测集当且仅当 $\mu^*(N) = 0$.

27. (a) 在 $(0,1]$ 上存在非 Borel 可测集.
    (b) 在 $2^{(0,1]}$ 上不存在平移不变的概率测度. 从而 Lebesgue 测度不可能扩张到 $(0,1]$ 上.

28. 在测度空间 $(\Omega, \mathscr{A}, \mu)$ 中, $A_i, B_i \subset \Omega$ 满足 $\mu^*(A_i \Delta B_i) = 0$, 则 $\mu^*\left(\sum\limits_{i=1}^{\infty} A_i\right) = \mu^*\left(\sum\limits_{i=1}^{\infty} B_i\right)$.

29. 令 $\mathscr{C} = \{C_{a,b} = [-b, -a) \cup (a, b] : 0 < a < b\}$, 定义 $\mu(C_{a,b}) = b - a$. 证明 μ 在 $\sigma(\mathscr{C})$ 上可以扩张为一个测度. 问 $[1,2]$ 是 μ^* 可测的吗?

30. 设函数 f 在 $[0,\infty)$ 严格单增,严格凹且 $f(0)=0$. $\forall A\subset(0,1]$, 定义 $\mu^*(A)=f(\lambda^*(A))$ (λ^* 为 Lebesgue 外测度). 证明 μ^* 为外测度, 即它满足 $\mu^*(\emptyset)=0$ 及非负性, 单调性与次可数可加性.

31. 设 $(\Omega,\mathscr{A},\mathbb{P}$ 为概率空间, $A\notin\mathscr{A}$. 令 $\mathscr{A}_1=\sigma(\mathscr{A}\cup\{A\})$, 证明 $\mathbb{P}$ 可以扩张为 $\mathscr{A}_1$上的概率测度.

32. 令 $f:\mathbb{R}\ni x\longmapsto\frac{x}{3}\in\mathbb{R}$, $A_0=[0,1]$. 则 $A_{n+1}=f(A_n)+\frac{2}{3}f(A_n)$ $(n\geqslant 0)$ 单调下降. A_n 的极限记为 C, 称之为 Cantor 集. 证明 C 的 Lebesgue 测度为 0.

第二章　随机变量与可测函数

前章我们已经建立了测度空间的概念, 由此可给出随机变量及其分布函数的公理化定义.

定义 2.1. **1)** 设 $(\Omega, \mathscr{A}, \mathbb{P})$ 是概率空间, $\xi : \Omega \to \mathbb{R}$ 是实函数. 如果 $\forall x \in \mathbb{R}, \{\omega : \xi(\omega) < x\} \in \mathscr{A}$, 则称 ξ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的实随机变量. 如果 η , ζ 是实随机变量, 则称 $\xi = \eta + \mathrm{i}\,\zeta$ 为 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的复随机变量.

2) 如果 $\xi_1, \cdots, \xi_n$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的实 (复) 随机变量, 则称向量值函数 $(\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的 n 维实 (复) 随机向量 (或变量).

3) 设 $(\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的一个实随机向量, 则称

$$F : \mathbb{R}^n \ni (x_1, x_2, \cdots, x_n) \mapsto \mathbb{P}(\xi_i < x_i : 1 \leqslant i \leqslant n)$$

为这个随机向量的分布函数.

4) 设 $(\xi_1, \cdots, \xi_n), (\eta_1, \cdots, \eta_n)$ 是概率空间 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的两个随机向量, 若

$$\mathbb{P}(\xi_i \neq \eta_i) = 0, 1 \leqslant i \leqslant n,$$

则称这两个随机向量是几乎肯定 (记作 a.s.) 相等的. 如果它们具有相同的分布函数, 则称它们是同分布的.

本章先将上面的随机变量概念延拓为一般可测空间上的可测函数, 然后着重讨论可测函数的构造及收敛定理, 它们将成为定义和研究可测函数积分与随机变量期望的理论基础.

§2.1　可测函数

§2.1.1　基本概念及性质

设 $\mathscr{B}$ 为 $\mathbb{R}$ 上的 Borel σ 代数, $\bar{\mathbb{R}}=[-\infty,\infty]$, $\bar{\mathscr{B}}=\sigma(\mathscr{B}\cup\{\infty\}\cup\{-\infty\})$, 设 $\bar{\mathbb{R}}^n$ 为 $\bar{\mathbb{R}}$ 的 n 维乘积空间, $\bar{\mathscr{B}}^n$ 为其上的乘积 σ 代数. 同样地，可以定义广义复平面 $\bar{\mathbb{C}}$ 的 n 维乘积空间 $\bar{\mathbb{C}}^n$, 及其上的乘积 σ 代数 $\bar{\mathscr{B}}_c^n$.

定义 2.2. **1)** 设 $(\Omega,\mathscr{A})$ 与 $(E,\mathscr{E})$ 是两个可测空间, f 是从 Ω 到 E 的映射. 如果 $\forall B\in\mathscr{E}, f^{-1}(B)\triangleq\{\omega\in\Omega: f(\omega)\in B\}\in\mathscr{A}$, 则称 f 为从 $(\Omega,\mathscr{A})$ 到 $(E,\mathscr{E})$ 的可测映射.

2) 特别地, 如果 f 是从 $(\Omega,\mathscr{A})$ 到 $(\bar{\mathbb{R}},\bar{\mathscr{B}})$ 上的可测映射, 则称 f 为实可测函数. 如 f 是从 $(\Omega,\mathscr{A})$ 到 $(\bar{R}^n,\bar{\mathscr{B}}^n)$ 上的可测映射, 则称 f 为 n 维实可测函数. 若 $f=f_1+\mathrm{i}\,f_2, f_1, f_2$ 是 n 维实可测函数, 则称 f 为 n 维复可测函数.

如无特别说明, 下面所设的可测函数均指实可测函数.

3) 称 $f^{-1}(B)$ 为 B 在 f 下的逆象. 设 $\mathscr{C}$ 是 E 的子集类, 则称 $\{f^{-1}(B): B\in\mathscr{C}\}$ 为 $\mathscr{C}$ 在 f 下的逆象, 记作 $f^{-1}(\mathscr{C})$.

易见, $f:(\Omega,\mathscr{A})\to(E,\mathscr{E})$ 可测当且仅当 $f^{-1}(\mathscr{E})\subset\mathscr{A}$.

性质 2.3. 设 f 是从 Ω 到 E 的映射, 则

1) $f^{-1}(E)=\Omega, f^{-1}(\varnothing)=\varnothing$;

2) $f^{-1}(B^c)=[f^{-1}(B)]^c, B\subset E$;

3) $f^{-1}(B_1-B_2)=f^{-1}(B_1)-f^{-1}(B_2), B_1, B_1\subset E$;

4) $f^{-1}\left(\bigcup\limits_{\gamma\in\Gamma}B_\gamma\right)=\bigcup\limits_{\gamma\in\Gamma}f^{-1}(B_\gamma), B_\gamma\subset E, \gamma\in\Gamma$;

5) $f^{-1}\left(\bigcap\limits_{\gamma\in\Gamma}B_\gamma\right)=\bigcap_{\gamma\in\Gamma}f^{-1}(B_\gamma), B_\gamma\subset E, \gamma\in\Gamma$.

简言之, 逆像与集合的任意运算可交换.

性质 2.4. 设 $\mathscr{E}$ 是 E 上的 σ 代数, 则 $f^{-1}(\mathscr{E})$ 是 Ω 中使得 f 可测的最小 σ 代数.

性质 2.5. 设 $\mathscr{C}$ 是 E 的子集类, $f:\Omega\to E$. 则 $f^{-1}(\sigma(\mathscr{C}))=\sigma(f^{-1}(\mathscr{C}))$.

证明 $f^{-1}(\sigma(\mathscr{C}))$ 是 σ 代数且包含 $f^{-1}(\mathscr{C})$, 故 $f^{-1}(\sigma(\mathscr{C}))\supset\sigma(f^{-1}(\mathscr{C}))$. 从

而只需证

$$\mathscr{A} \triangleq \{C \subset E : f^{-1}(C) \in \sigma(f^{-1}(\mathscr{C}))\} \supset \sigma(\mathscr{C}),$$

事实上, 我们有 1) $\mathscr{A} \supset \mathscr{C}$; 2) $f^{-1}(E) = \Omega \in \sigma(f^{-1}(\mathscr{C})) \Rightarrow E \in \mathscr{A}$; 3) $C \in \mathscr{A} \Rightarrow f^{-1}(C^c) = (f^{-1}(C))^c \in \sigma(f^{-1}(\mathscr{C})) \Rightarrow C^c \in \mathscr{A}$; 4) $\{C_n\}_{n\geqslant 1} \subset \mathscr{A} \Rightarrow f^{-1}\left(\bigcup\limits_{n=1}^{\infty} C_n\right) = \bigcup\limits_{n=1}^{\infty} f^{-1}(C_n) \in \sigma(f^{-1}(\mathscr{C})) \Rightarrow \bigcup\limits_{n=1}^{\infty} C_n \in \mathscr{A}$. 所以 $\mathscr{A} \supset \sigma(\mathscr{C})$. □

定理 2.6. **1)** f 是 $(\Omega, \mathscr{A})$ 上可测实函数当且仅当是 $\forall x \in \mathbb{R}, \{f < x\} \in \mathscr{A}$.

2) $f = (f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上的 n 维可测实函数的充分必要条件是 $\forall 1 \leqslant k \leqslant n, f_k$ 是 $(\Omega, \mathscr{A})$ 上的可测实函数.

证明 1) 必要性显然. 令 $\mathscr{S} = \{[-\infty, x) : x \in \mathbb{R}\}$, 则 $\sigma(\mathscr{S}) = \bar{\mathscr{B}}$, 故由性质 2.5 知 $f^{-1}(\bar{\mathscr{B}}) = f^{-1}(\sigma(\mathscr{S})) = \sigma(f^{-1}(\mathscr{S})) \subset \sigma(\mathscr{A}) = \mathscr{A}$. 而由性质 2.4 知 $f^{-1}(\bar{\mathscr{B}})$ 是使 f 可测的最小 σ 代数, 从而 f 是 $(\Omega, \mathscr{A})$ 上的可测函数.

2) 必要性: $\forall A_k \in \bar{\mathscr{B}}, \{f_k \in A_k\} = \{f \in \bar{R} \times \cdots \times A_k \times \cdots \times \bar{R}\} \in \mathscr{A}$. 故 f_k 可测.

充分性: 令 $\mathscr{S} = \{\{f_k < r\} : 1 \leqslant k \leqslant n, r \in \mathbb{R}\}$. 由于 $\bar{\mathscr{B}}^n = \sigma(\{\{x : x_k < r\} : 1 \leqslant k \leqslant n, r \in \mathbb{R}\})$, 我们有 $f^{-1}(\bar{\mathscr{B}}^n) = \sigma(\mathscr{S})$, 而由 $f_k\,(1 \leqslant k \leqslant n)$ 的可测性知 $\mathscr{S} \subset \mathscr{A}$. 从而 $f^{-1}(\bar{\mathscr{B}}^n) \subset \mathscr{A}$, 即 f 可测. □

定理 2.7. 若 $(\Omega_i, \mathscr{A}_i), i = 1, 2, 3$ 是三个可测空间, $(\Omega_1, \mathscr{A}_1) \xrightarrow{f} (\Omega_2, \mathscr{A}_2) \xrightarrow{g} (\Omega_3, \mathscr{A}_3)$ 都可测, 则 $g \circ f$ 是从 $(\Omega_1, \mathscr{A}_1)$ 到 $(\Omega_3, \mathscr{A}_3)$ 的可测映射.

证明 由 $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$ 立得. □

定理 2.6 之 1) 表明, 随机变量就是概率空间上有限可测函数; 定理 2.6 之 2) 表明, 一个向量值函数可测当且仅当其每个分量均可测; 而定理 2.7 是说可测映射的复合仍可测.

推论 2.8. **1)** 设 g 是 $(\bar{\mathbb{R}}^n, \bar{\mathscr{B}}^n)$ 上的实 (复) 可测函数, $f_1, \cdots, f_n$ 是 $(\Omega, \mathscr{A})$ 上的实可测函数, 则 $g(f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上的实 (复) 可测函数.

2) 设 g 是 $(\bar{\mathbb{C}}^n, \bar{\mathscr{B}}_c^n)$ 上的实 (复) 可测函数, $f_1, \cdots, f_n$ 是 $(\Omega, \mathscr{A})$ 上的复可测函数, 则 $g(f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上的实 (复) 可测函数.

推论 2.9. **1)** 设 g 是 $(\bar{\mathbb{R}}^n, \bar{\mathscr{B}}^n)$ 上的实 (复) 可测函数, $f_1, \cdots, f_n$是$(\Omega, \mathscr{A}, \mathbb{P})$ 上的随机变量. 若 $\mathbb{P}(|g(f_1, \cdots, f_n)| = \infty) = 0$, 则 $g(f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的一个实 (复) 随机变量.

2) 设 g 是 $(\bar{\mathbb{C}}^n, \bar{\mathscr{B}}_c^n)$ 上的实 (复) 可测函数, $f_1, \cdots, f_n$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的复随机变量. 如 $|g(f_1, \cdots, f_n)| < \infty$,则 $g(f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的一个实 (复) 随机变量.

§ 2.1.2 可测函数的构造

我们将从可测集的示性函数出发, 使用线性组合与极限构造出所有实可测函数.

定义 2.10. **1)** $\forall A \subset \Omega$, 定义 A 的示性函数为

$$\mathbf{1}_A(\omega) = \begin{cases} 1, & \text{若 } \omega \in A, \\ 0, & \text{其它}. \end{cases}$$

2) 若 $A_k \in \mathscr{A}, k = 1, \cdots, n$ 两两不交且 $\Omega = \sum\limits_{k=1}^{n} A_k, a_1, \cdots a_n \in \bar{\mathbb{R}}^n$, 则称函数 $f = \sum\limits_{k=1}^{n} a_k \mathbf{1}_{A_k}$ 为简单函数.

3) 若在 2) 中取 $n = \infty$, 则称 f 为初等函数.

性质 2.11. **1)** $\mathbf{1}_A$ 是 $(\Omega, \mathscr{A})$ 上的可测函数当且仅当 $A \in \mathscr{A}$;

2) 初等函数与简单函数是 $(\Omega, \mathscr{A})$ 上的可测函数.

证明 设 $f = \sum\limits_{k=1}^{n} a_k \mathbf{1}_{A_k}$, 则 $\forall B \in \bar{\mathscr{B}}, f^{-1}(B) = \bigcup\limits_{k: a_k \in B} A_k \in \mathscr{A}$. □

定理 2.12. **1)** 可测函数是简单函数列的逐点收敛极限;

2) 可测函数是初等函数列的一致极限;

3) 有界可测函数是简单函数列的一致极限;

4) 非负可测函数是非负简单函数列 (初等函数列) 的不降极限 (一致极限).

证明 1) $\forall n \geqslant 1$ 及 $\omega \in \Omega$, 令

$$f_n = \sum_{k=-n2^n}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{\{\frac{k}{2^n} \leqslant f < \frac{k+1}{2^n}\}} + n\mathbf{1}_{\{f \geqslant n\}} - n\mathbf{1}_{\{f < -n\}}$$

则 f_n 是简单函数且

$$|f_n - f| \mathbf{1}_{\{-n \leqslant f < n\}} < \frac{1}{2^n}; \text{当 } f = \infty \text{ 时}, f_n = n; \text{当 } f = -\infty \text{ 时}, f_n = -n.$$

故 f_n 逐点收敛于 f.

2) 令

$$f_n = \sum_{k=-\infty}^{\infty} \frac{k}{2^n}\mathbf{1}_{\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\}} + \infty\mathbf{1}_{\{f=\infty\}} - \infty\mathbf{1}_{\{f=-\infty\}}$$

则 f_n 为初等函数, 满足

$$|f_n - f|\mathbf{1}_{\{|f|<\infty\}} < \frac{1}{2^n};\quad \text{且当 } |f| = \infty \text{ 时}, f_n = f.$$

从而 f_n 一致收敛于 f.

3) 如 f 有界, 则由 1) 知简单函数 f_n 一致收敛于 f.

4) 如 f 非负, 则 1) 和 2) 中所构造的函数序列 f_n 单调下降. □

设 f 是 Ω 上实函数, 定义 $f^+ = \max\{f, 0\}, f^- = \max\{-f, 0\}$. 并分别称为 f 的正部与负部. 我们有 $f = f^+ - f^-, |f| = f^+ + f^-, f^+ = \frac{|f|+f}{2}, f^- = \frac{|f|-f}{2}$.

定理 2.13. 可测函数的正部与负部仍为可测函数. 因而任何可测函数可表成非负可测函数之差.

§ 2.1.3 可测函数的运算

命题 2.14. 设 $\{f_n\}_{n\geqslant 1}$ 是 Ω 上一列实函数.

1) $\{f_n\}_{n\geqslant 1}$ 的上、下极限与上下确界都存在, 且

$$\begin{aligned}\underline{\lim}_{n\to\infty} f_n &= \lim_{n\to\infty}\inf_{k\geqslant n} f_k = \sup_n \inf_{k\geqslant n} f_k,\\ \overline{\lim}_{n\to\infty} f_n &= \lim_{n\to\infty}\sup_{k\geqslant n} f_k = \inf_n \sup_{k\geqslant n} f_k.\end{aligned}$$

2) $\{f_n\}_{n\geqslant 1}$ 的极限 $\lim\limits_{n\to\infty} f_n$ 存在当且仅当 $\forall\omega\in\Omega$, $\underline{\lim}\limits_{n\to\infty} f_n(\omega) = \overline{\lim}\limits_{n\to\infty} f_n(\omega)$.

定理 2.15. **1)** 设 $\{f_n\}_{n\geqslant 1}$ 是可测空间 $(\Omega, \mathscr{A})$ 上的一个实可测函数列, 则 $\sup\limits_{n\geqslant 1} f_n$, $\inf\limits_{n\geqslant 1} f_n$, $\underline{\lim}\limits_{n\to\infty} f_n$, $\overline{\lim}\limits_{n\to\infty} f_n$ 都是实可测函数.

2) 设 $\{f_n\}_{n\geqslant 1}$ 是可测空间 $(\Omega, \mathscr{A})$ 上的一个复可测函数列. 如果 $\lim\limits_{n\to\infty} f_n$ 存在, 则它也是可测函数.

证明 注意到 $\forall x\in\mathbb{R}$,

$$\left\{\inf_{n\geqslant 1} f_n < x\right\} = \bigcup_{n\geqslant 1}\{f_n < x\} \in \mathscr{A},$$

从而 $\inf\limits_{n\geqslant 1} f_n$ 可测. 由于 $\sup\limits_{n\geqslant 1} f_n = -\inf\limits_{n\geqslant 1}(-f_n)$, 故 $\sup\limits_{n\geqslant 1} f_n$ 可测.

最后, $\forall x \in \mathbb{R}$ 有

$$\left\{\overline{\lim_{n\to\infty}} f_n < x\right\} = \bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcup_{k\geqslant n}\left\{f_k < x - \frac{1}{m}\right\} \in \mathscr{A}.$$

故 $\overline{\lim\limits_{n\to\infty}} f_n$ 可测. 而 $\underline{\lim\limits_{n\to\infty}} f_n = -\overline{\lim\limits_{n\to\infty}}(-f_n)$, 所以 $\underline{\lim\limits_{n\to\infty}} f_n$ 也可测. □

定理 2.16. 设 g 是 $D \subset \bar{\mathbb{R}}^n$ 上的连续函数, 则 g 是 $(D, D \cap \bar{\mathscr{B}}^n)$ 上可测函数. 使用 $\bar{\mathbb{C}}^n$ 代替 $\bar{\mathbb{R}}^n$ 时, 结论仍成立.

证明 不妨设 g 是实函数. $\forall m \geqslant 1$, 把 $\bar{R}^n$ 分割成可数个边长为 $1/2^m$ 的互不相交的立方块, 即

$$A_{j_1,\cdots,j_n} = \left[\frac{j_1}{2^m}, \frac{j_1+1}{2^m}\right) \times \cdots \times \left[\frac{j_n}{2^m}, \frac{j_n+1}{2^m}\right), j_1, \cdots, j_n \in \mathbb{Z} \cup \{+\infty, -\infty\},$$

当 $j = -\infty$ 时约定 $[\frac{j}{2^m}, \frac{j+1}{2^m}) = \{-\infty\}$, 当 $j = +\infty$ 时约定 $[\frac{j}{2^m}, \frac{j+1}{2^m}) = \{+\infty\}$. 把这些小方块重新排列, 记成 $\{A_i^m : i, m \in \mathbb{N}\}$.

取定 $x_{im} \in A_i^m$, 定义

$$g_m(x) = \sum_{i=1}^{\infty} \mathbf{1}_{A_i^m \cap D}(x) g(x_{im}),$$

则 g_m 可测, 且由 g 的连续性知 $g_m(x) \xrightarrow{m\to\infty} g(x)$, 从而 g 可测. □

定理 2.17. 设 $D \subset \bar{\mathbb{C}}^n$, $f_1, \cdots, f_k$ 是 $(\Omega, \mathscr{A})$ 上可测函数, $(f_1, \cdots, f_k)(\Omega) \subset D$. 设 g 是 D 上可测函数, 则 $g(f_1, \cdots, f_k)$ 是可测函数.

证明 由于可测函数的复合是可测函数. □

推论 2.18. 可测函数的有意义的和、差、积、商仍是可测的.

推论 2.19. 设 $\xi_1, \cdots, \xi_n$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的 n 个随机变量, g 是 $\mathbb{C}^n$ 上有限连续函数, 则 $g(\xi_1, \cdots, \xi_n)$ 是随机变量. 特别地, 随机变量的有意义的和、差、积、商仍是随机变量.

§ 2.1.4 函数形式的单调类定理

定义 2.20. 设 $\mathscr{L}$ 是 Ω 上一个函数族, 满足条件: $f \in \mathscr{L} \Rightarrow f^+, f^- \in \mathscr{L}$. 如果函数族 L 满足:

1) $1 \in L$;

2) L 中有限个函数的有意义的线性组合属于 L;

3) 如果 $f_n \in L, 0 \leqslant f_n \uparrow f$, f 有界或 $f \in \mathscr{L}$, 则 $f \in L$;

则称 L 为 $\mathscr{L}$ 系,

定理 2.21 (**函数形式的单调类定理**). 若 $\mathscr{L}$ 系 L 包含某一 π 系 $\mathscr{C}$ 中所有集合的示性函数, 则 L 包含一切属于 $\mathscr{L}$ 的关于 $\sigma(\mathscr{C})$ 可测的实函数.

证明 令 $\Lambda = \{A : \mathbf{1}_A \in L\}$, 则 $\Omega \in \Lambda, \Lambda$ 对真差封闭, 对单调增集的并封闭, 从而 Λ 是 λ 系. 又因 $\Lambda \supset \mathscr{C}$, 且 $\mathscr{C}$ 为 π 系, 故由单调类定理知 $\Lambda \supset \sigma(\mathscr{C})$. 由此及定义 reft2.20 之 2) 知 L 包含所有 $\sigma(\mathscr{C})$ 可测的简单函数. 设 $f \in \mathscr{L}$ 且关于 $\sigma(\mathscr{C})$ 可测, 则 $f^+, f^- \in \mathscr{L}$ 且关于 $\sigma(\mathscr{C})$ 可测, 从而存在非负简单函数列 $f_n \uparrow f^+$, 从而由 3) 知 $f^+ \in L$. 同理 $f^- \in L$, 故 $f = f^+ - f^- \in L$. □

函数形式单调类定理的应用思路是: 欲证明函数族 F 中函数具有某种性质 A_0, 引入 $\mathscr{L} \supset F$, 使 $L \triangleq \big\{f : f \text{ 具性质 } A_0\big\}$ 为 $\mathscr{L}$ 系, 再引进 π 系 $\mathscr{C}$ 使 $\mathscr{C}$ 中集合的示性函数均属于 L, 且所有关于 $\sigma(\mathscr{C})$ 可测的函数类包含 F. 于是由定理 2.21 知 F 中函数具有性质 A_0.

下面的定理是定理 2.21 的一个具体应用.

定理 2.22. 设 Ω 是一集合, $(E, \mathscr{E})$ 是可测空间, $f : \Omega \to E$ 为映射. 令 $\sigma(f) = f^{-1}(\mathscr{E})$, 则 $\varphi : \Omega \to \bar{\mathbb{R}}$ 是 $\sigma(f)$ 可测函数的充要条件是存在 $(E, \mathscr{E})$ 上可测函数 g 使 $\varphi = g \circ f$. 若 φ 有限 (有界), 则 g 可取有限 (有界) 值.

证明 充分性由复合函数的可测性即得.

必要性: 令 $L = \{g \circ f : g \in \mathscr{E}\}$, 则

1) $\mathbf{1}_\Omega = \mathbf{1}_E \circ f \in L$;

2) $\forall g_1 \circ f, g_2 \circ f \in L$ 及 $a_1, a_2 \in \mathbb{R}$ 使 $a_1(g_1 \circ f) + a_2(g_2 \circ f)$ 有意义, 我们有

$$a_1 g_1 \circ f + a_2 g_2 \circ f = [(a_1 g_1 + a_2 g_2)\mathbf{1}_A] \circ f,$$

其中 $A = \big\{x \in E : a_1 g_1(x) + a_2 g_2(x) \text{ 存在}\big\}$. 从而 $a_1 g_1 \circ f + a_2 g_2 \circ f \in L$.

3) 若 $\varphi_n \in L, \varphi_n \uparrow \varphi$, 则 $\exists g_n \in \mathscr{E}$,使 $\varphi_n = g_n \circ f$. 令 $g = \sup\limits_{n \geqslant 1} g_n$, 则 $g \in \mathscr{E}$ 且 $\varphi = g \circ f$, 故 $\varphi \in L$. 再取 $\mathscr{L}$ 为 $\sigma(f)$ 的可测函数类, 则 L 为 $\mathscr{L}$ 系. 若 $C \in \sigma(f)$, 则存在 $B \in \mathscr{E}$ 使 $C = f^{-1}(B)$, 故 $\mathbf{1}_C = \mathbf{1}_B \circ f$, 从而 L 包含所有 $\sigma(f)$ 可测的示性函数. 由定理 2.21 知, L 包含 $\mathscr{L}$. 从而第一论断得证. 最

后, 如 φ 有界 (有限), 且 $\varphi = g\circ f$, 则可以 $g\mathbf{1}_{\{|g|\leqslant||\varphi||_\infty\}}(g\mathbf{1}_{\{|g|<\infty\}})$ 代替 g. □

推论 2.23. 设 f 是 Ω 上 n 维实函数. 则 φ 关于 $f^{-1}(\bar{\mathscr{B}}^n)$ 可测的充要条件是存在 $(\bar{\mathbb{R}}^n, \bar{\mathscr{B}}^n)$ 上可测函数 g 使 $\varphi = g\circ f$.

定理 2.24. 设 $\mathscr{L}$ 是 $\bar{\mathbb{R}}^n$ 上实函数全体, L 是 $\bar{\mathbb{R}}^n$ 上包含有界连续函数的 $\mathscr{L}$ 系, 则 L 包含 $\mathscr{L}$ 中一切 Borel 可测函数.

证明 设 $\mathscr{S} = \{A : A$ 是 $\bar{\mathbb{R}}^n$ 中开区间 $\}$, 则 $\mathscr{S}$ 是 π 系, $\sigma(\mathscr{S}) = \bar{\mathbb{R}}^n$. $\forall A\in\mathscr{S}$, 令 $d(x, A^c) = \inf\{|x-y| : y\notin A\}$. $\forall m\geqslant 1$, 令

$$f_m(x) = \begin{cases} 0, & x\notin A \\ 1, & x\in A, d(x,A^c) > \frac{1}{m}, \\ md(x,A^c), & x\in A, d(x,A^c)\leqslant \frac{1}{m}. \end{cases}$$

则 f_m 连续, $f_m\uparrow \mathbf{1}_A$, 从而 $\mathbf{1}_A\in L$. 再由定理 2.21 即得所证结论. □

§2.2 分布函数与分布律

由随机向量分布函数的定义容易看出它具有下面定理中所描述的四个特征. 其中, 对于 $\mathbb{R}^n$ 上函数 F 及 $a, b\in\mathbb{R}^n, a\leqslant b$, $\Delta_{b,a}F$ 为 F 在区间 $[a,b)$ 上的差分. 当 $n=1$ 时, $\Delta_{b,a}F$ 定义为 $F(b)-F(a)$. 而对高维情形, 当 $a=(a_1,\cdots,a_n), b=(b_1,\cdots,b_n)$ 时, 令 $\Delta_{b,a}F = \Delta_{b_1,a_1}\Delta_{b_2,a_2}\cdots\Delta_{b_n,a_n}F$, 其中 Δ_{b_i,a_i} 为函数对第 i 个分量作差分.

定理 2.25. 设 F 是一个 n 元实函数, 它是一个 n 维随机向量的分布函数当且仅当

(a) F 不降且 $\triangle_{b,a}F\geqslant 0$;

(b) F 对每个分量左连续 (简称左连续);

(c) 如 $\exists 1\leqslant i\leqslant n$ 使 $x_i\to-\infty$, 则 $F(x)\to 0$;

(d) $F(\infty,\infty,\cdots,\infty)\triangleq \lim\limits_{n\to\infty} F(n,\cdots,n) = 1$.

我们仅需证明充分性, 即: 对每个满足 (a)–(d) 的函数 F, 构造概率空间 $(\Omega,\mathscr{A},\mathbb{P})$ 及其上 n 维的随机变量 ξ, 使它的分布函数就是 F. 由于往往不限

于概率测度, 因而我们将证明一个更广的结论. 为此, 先拓广分布函数的定义.

定义 2.26. $\mathbb{R}^n$ 上的有限实函数 F 如果满足左连续性及差分非负性, 即 $\forall a \leqslant b$ 有 $\Delta_{b,a}F \geqslant 0$, 则称之为 $\mathbb{R}^n$ 上一个分布函数.

定理 2.27. 设 F 是 $\mathbb{R}^n$ 上的分布函数, 则在 $\mathscr{B}^n$ 上存在唯一的测度 μ_F 满足 $\mu_F([a,b)) = \Delta_{b,a}F$, $a \leqslant b$. 因而 μ_F 在有限区间上取值有限, 它的完备化称作由 F 生成的 Lebesgue-Stieltjes (L-S) 测度.

证明 令

$$\mathscr{C} = \left\{ [a,b) \triangleq \prod_{k=1}^{n} [a_k, b_k) : a \leqslant b, a_k \in [-\infty, +\infty), b_k \in (-\infty, +\infty] \right\}$$

其中 $a_k = -\infty$ 时, $[a_k, b_k)$ 理解成 $(-\infty, b_k)$, 则 $\mathscr{C}$ 是 $\mathbb{R}^n$ 中半集代数.

定义 $\mathscr{C}$ 上的集函数: $\mu_F([a,b)) = \Delta_{b,a}F, a \leqslant b$. 当 b 或 a 的某分量为 $\pm\infty$ 时, $\mu_F([a,b))$ 理解为当该分量趋于 $\pm\infty$ 时之极限. 易证 μ_F 是有限可加的. 由于 μ_F 在有限区间上取值有限, 故而是 σ 有限的. 往证 μ_F 是 σ 可加的. 设 $A \in \mathscr{C}, \{A_k\}_{k\geqslant 1} \subset \mathscr{C}$ 两两不交且 $\sum\limits_{k=1}^{\infty} A_k = A$. 下面证明 $\sum\limits_{k=1}^{\infty} \mu_F(A_k) = \mu_F(A)$.

a) 先证 $\sum\limits_{k=1}^{\infty} \mu_F(A_k) \leqslant \mu_F(A)$.

先将 μ_F 扩张到 $\mathscr{F}(\mathscr{C})$ 上, 由有限可加性知 $\sum\limits_{k=1}^{n} \mu_F(A_k) = \mu_F\left(\sum\limits_{k=1}^{n} A_k\right) \leqslant \mu_F(A)$. 令 $n \to \infty$ 即得.

b) 再证 $\sum\limits_{k=1}^{\infty} \mu_F(A_k) \geqslant \mu_F(A)$.

不妨设 A 是有限区间, 否则以 $A^{(N)} = A \cap I^{(N)}$ 代替 A, 以 $A_k^{(N)} = A_k \cap I^{(N)}$ 代替 A_k, 再令 $N \uparrow \infty$ 即可. 这里 $I^{(N)} = [-N, N)^n$. 令 $A = [a,b), A_k = [a^{(k)}, b^{(k)}), a \leqslant b, a^{(k)} \leqslant b^{(k)}, k \geqslant 1$.

由 F 的左连续性, $\forall \varepsilon > 0, \exists \delta > 0$, 使

$$\mu_F(A) - \varepsilon < \mu_F([a, b - \vec{\delta})).$$

其中 $\vec{\delta} = (\delta, \cdots, \delta)$. 此外, 对每一 $k \geqslant 1$, 存在 $\delta^{(k)} > 0$ 使

$$\mu_F([a^{(k)} - \vec{\delta}^{(k)}, b^{(k)})) \leqslant \mu_F(A_k) + \frac{\varepsilon}{2^k}.$$

由于

$$[a, b-\vec{\delta}] \subset [a, b) = \bigcup_{k=1}^{\infty} [a^{(k)}, b^{(k)}) \subset \bigcup_{k=1}^{\infty} (a^{(k)} - \vec{\delta}^{(k)}, b^{(k)}),$$

由有限覆盖定理, 存在 $N \geqslant 1$ 使

$$[a, b-\vec{\delta}] \subset \bigcup_{k=1}^{N} (a^{(k)} - \vec{\delta}^{(k)}, b^{(k)})$$

从而

$$\begin{aligned} \mu_F([a,b)) &\leqslant \mu_F([a, b-\vec{\delta})) + \varepsilon \leqslant \varepsilon + \sum_{k=1}^{N} \mu_F((a^{(k)} - \vec{\delta}^{(k)}, b^{(k)})) \\ &\leqslant 2\varepsilon + \sum_{k=1}^{\infty} \mu_F(A_k). \end{aligned}$$

令 $\varepsilon \downarrow 0$ 即得所求. 从而 μ_F 为 $\mathscr{C}$ 上 σ 有限测度, 由测度扩张定理 (定理 1.42) 知它在 $\mathscr{B}^n = \sigma(\mathscr{C})$ 上存在唯一的扩张. □

定理 2.25 的证明 设 μ_F 是由 F 导出的 $\mathscr{B}^n$ 上的测度, 由 (c) 和 (d) 知 μ_F 是概率测度. 在概率空间 $(\Omega, \mathscr{A}, \mathbb{P}) = (\mathbb{R}^n, \mathscr{B}^n, \mu_F)$ 上定义 $\xi(x) = x$, 则 ξ 是 n 维随机变量, 且 $\mathbb{P}(\xi < x) \triangleq \mu_F((-\infty, x)) = F(x)$. □

为区别于一般分布函数, 满足定理 2.25 之 (a)–(d) 的分布函数称为概率分布函数.

例: 令 $F(x_1, \cdots, x_n) = x_1 x_2 \cdots x_n$, 则 F 是 $\mathbb{R}^n$ 上分布函数且 μ_F 为 $\mathbb{R}^n$ 上 Lebesgue 测度.

定义 2.28. 由 n 维随机变量 ξ 的分布函数生成的 $(\mathbb{R}^n, \mathscr{B}^n)$ 上概率测度, 称为 ξ 的概率分布, 记为 $\mathbb{P}_\xi$. $\mathbb{P}_\xi$ 能唯一确定概率分布的规律, 称为分布律.

如：分布函数、特征函数、离散型随机变量的分布列、连续型随机变量的分布密度等均是分布律.

§2.3 独立随机变量

定义 2.29. 设 $\left\{\xi^{(t)} = (\xi_{t,1}, \cdots, \xi_{t,m_t}) : t \in T\right\}$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上实随机向量族.

若 $\forall\{t_1,\cdots,t_l\}\subset T$ 及 $x^{(t_i)}\in\mathbb{R}^{m_{t_i}}, i=1,\cdots,l$, 有

$$\mathbb{P}(\xi^{(t_1)}<x^{(t_1)},\cdots,\xi^{(t_l)}<x^{(t_l)})=\prod_{i=1}^{l}\mathbb{P}\left(\xi^{(t_i)}<x^{(t_i)}\right).$$

则称 $\left\{\xi^{(t)}:t\in T\right\}$ 为一族独立随机向量, 简称 $\left\{\xi^{(t)}:t\in T\right\}$ 独立. 同样可定义复随机向量的独立性.

下面的性质是显然的.

性质 2.30. **1)** $\{\xi^{(t)}:t\in T\}$ 独立当且仅当 $\forall T'\subset T, |T'|<\infty, \{\xi^{(t)}:t\in T'\}$ 独立.

2) 设 $\bigcup\limits_{r\in I}T_r=T$, T_r 两两不交, $|T_r|<\infty$. 令 $\bar{\xi}^{(r)}=(\xi^{(t)}:t\in T_r)$. 若 $\{\xi^{(t)}:t\in T\}$ 独立, 则 $\{\bar{\xi}^{(r)}:r\in I\}$ 独立.

定理 2.31. 设 $\xi^{(k)}, k=1,\cdots,n$ 是 n 个实 (复) 随机向量, 则它们独立的充要条件是: $\forall B^{(m_k)}\in\mathscr{B}^{(m_k)}$, 有

$$\mathbb{P}\left(\bigcap_{k=1}^{n}\{\xi^{(k)}\in B^{(m_k)}\}\right)=\prod_{k=1}^{n}\mathbb{P}\left(\xi^{(k)}\in B^{(m_k)}\right).$$

证明 充分性显然. 今证必要性. 只证 $n=2$ 的情形, 对 $n>2$ 情形可使用归纳法.

1) 令 $\mathscr{S}_k=\{(-\infty,b^k):b^k\in\mathbb{R}^{(m_k)}\}, k=1,2$, 则 $\mathscr{S}_k$ 是 $\mathbb{R}^{m_k}$ 上的 π 系且 $\sigma(\mathscr{S}_k)=\mathscr{B}^{(m_k)}$. 给定 $(-\infty,b)\in\mathscr{B}^{m_2}$, 令

$$\mathscr{C}_1=\left\{A_1\in\mathscr{B}^{(m_1)}:\mathbb{P}(\xi^{(1)}\in A_1,\xi^{(2)}<b)=\mathbb{P}(\xi^{(1)}\in A_1)\mathbb{P}(\xi^{(2)}<b)\right\}.$$

则 $\mathscr{C}_1\supset\mathscr{S}_1$. 今证 $\mathscr{C}_1$ 是 λ 系.

显然 $\Omega\in\mathscr{C}_1$. 如 $A^{(n)}\uparrow A_1, A^{(n)}\in\mathscr{C}_1$, 则 $\{\xi^{(1)}\in A^{(n)}\}\uparrow\{\xi^{(1)}\in A_1\}$. 从而由概率的连续性 $A_1\in\mathscr{C}_1$. 此外, 如 $A_1\supset A_1'$ 且 $A_1, A_1'\in\mathscr{C}_1$, 则

$$\begin{aligned}&\mathbb{P}(\xi^{(1)}\in A_1-A_1',\ \xi^{(2)}<b)=\mathbb{P}(\xi^{(1)}\in A_1,\xi^{(2)}<b)-\mathbb{P}(\xi^{(1)}\in A_1',\xi^{(2)}<b)\\&=\mathbb{P}(\xi^{(1)}\in A_1)\mathbb{P}(\xi^{(2)}<b)-\mathbb{P}(\xi^{(1)}\in A_1')\mathbb{P}(\xi^{(2)}<b)\\&=\mathbb{P}(\xi^{(1)}\in A_1-A_1')\,\mathbb{P}(\xi^{(2)}<b).\end{aligned}$$

故 $A_1-A_1'\in\mathscr{C}_1$, 从而 $\mathscr{C}_1$ 是 λ 系. 由单调类定理知 $\mathscr{C}_1\supset\mathscr{B}^{(m_1)}$.

2) 设 $\forall A_1 \in \mathscr{B}^{(m_1)}$. 令

$$\mathscr{C}_2 = \{A_2 \in \mathscr{B}^{(m_2)} : \mathbb{P}(\xi^{(1)} \in A_1, \xi^{(2)} \in A_2) = \mathbb{P}(\xi^{(1)} \in A_1)\,\mathbb{P}(\xi^{(2)} \in A_2)\}.$$

则由 1), $\mathscr{C}_2 \supset \mathscr{S}_2$, 且同样可证明 $\mathscr{C}_2$ 是 λ 系. 从而证得结论. □

推论 2.32. 设 $\left\{\xi^{(k)} : k = 1, \cdots, n\right\}$ 独立. $f_k : \mathbb{R}^{m_k} \to \mathbb{R}^{m'_k}$ 是有限 Borel 可测函数, 则 $\left\{f_k(\xi^{(k)}) : k = 1, \cdots, n\right\}$ 独立.

证明 $\forall A_k \in \mathscr{B}^{(m'_k)}$, 有 $f_k^{-1}(A_k) \in \mathscr{B}^{(m_k)}$, 则

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{k=1}^{n}\{f_k(\xi^{(k)}) \in A_k\}\right) &= \mathbb{P}\left(\bigcap_{k=1}^{n}\{\xi^{(k)} \in f_k^{-1}(A_k)\}\right) \\
&= \prod_{k=1}^{n}\mathbb{P}(\xi^{(k)} \in f_k^{-1}(A_k)) \\
&= \prod_{k=1}^{n}\mathbb{P}(f_k(\xi^{(k)}) \in A_k). \quad \square
\end{aligned}$$

推论 2.33. $\left\{\xi^{(k)} : k = 1, \cdots, n\right\}$ 独立当且仅当 $(\xi^{(1)}, \cdots, \xi^{(n)})$ 的分布函数可表成

$$F(x^{(1)}, \cdots, x^{(n)}) = \prod_{k=1}^{n} F_k\left(x^{(k)}\right),$$

其中 F_k 是 $\mathbb{R}^{m_k}$ 上的实函数.

证明 必要性: 显然, 取 F_k 为 $\xi^{(k)}$ 之分布函数即可.

充分性: 不妨设 F_k 非负, 否则以 $|F_k|$ 代替 F_k. 由于

$$\begin{aligned}
\mathbb{P}(\xi^{(k)} < x^{(k)}) &= \mathbb{P}(\xi^{(k)} < x^{(k)}, \xi^{(i)} < \infty, i \neq k) \\
&= F(\infty, \cdots, x^{(k)}, \infty, \cdots, \infty) \\
&= F_k(x^{(k)})\prod_{i\neq k} F_i(\infty),
\end{aligned}$$

从而由 $\prod\limits_{i=1}^{n} F_i(\infty) = 1$ 知 $\xi^{(k)}$ 的分布函数为 $F_k(x^{(k)})/F_k(\infty)$. 由此知

$$F(x^{(1)}, \cdots, x^{(n)}) = \prod_{k=1}^{n} F_k(x^{(k)}) = \prod_{k=1}^{n} \frac{F_k(x^{(k)})}{F_k(\infty)}$$

蕴含 $\xi^{(1)}, \cdots, \xi^{(n)}$ 独立.

§2.4　可测函数序列的收敛

设 $(\Omega,\mathscr{A},\mu)$ 是完备测度空间. 如果某关系在某 μ 零测集外处处成立, 则称之关于 μ 几乎处处成立, 记作 μ-a.e. 或 a.e. 成立. 称这个零测集称为例外集. 在这一节, 所讨论的可测函数都是 a.e. 有限的.

§ 2.4.1　几乎处处收敛

定义 2.34. 设 $\{f_n\}$ 是可测函数列, f 是可测函数. 若存在 $N\in\mathscr{A},\mu(N)=0$, 使得 $\forall\omega\notin N$ 有 $f_n(\omega)\to f(\omega),n\to\infty$, 则称 $\{f_n\}$ 几乎处处收敛于 f, 记作 $f_n\xrightarrow{\text{a.e.}}f$.

如果 $\forall\omega\notin N$, 当 $n,m\to\infty$ 时, $f_n(\omega)-f_m(\omega)\to 0$, 则称 $\{f_n\}$ 几乎处处相互收敛, 记作 $f_n-f_m\xrightarrow{\text{a.e.}}0$.

易见, $f_n-f_m\xrightarrow{\text{a.e.}}0$ 当且仅当 $f_{n+m}-f_n\xrightarrow{\text{a.e.}}0(n\to\infty)$ 对 m 一致成立.

性质 2.35. **1)** 如 $f_n\xrightarrow{\text{a.e.}}f$, 则任何子列 $\{f_{n_k}\}$ 满足 $f_{n_k}\xrightarrow{\text{a.e.}}f$.

2) 如 $f_n\xrightarrow{\text{a.e.}}f,f_n\xrightarrow{\text{a.e.}}f'$, 则 $f=f'$ a.e.

3) 如 $f_n\xrightarrow{\text{a.e.}}f,g_n=f_n$ a.e., $f=g$ a.e., 则 $g_n\xrightarrow{\text{a.e.}}g$.

4) 如 $f_n^{(k)}\xrightarrow{\text{a.e.}}f^{(k)},k=1,\cdots,m,g$ 为 $\mathbb{R}^m$ 上连续函数, 则

$$g(f_n^{(1)},\cdots,f_n^{(m)})\xrightarrow{\text{a.e.}}g(f^{(1)},\cdots,f^{(m)}).$$

定理 2.36. 设 $\{f_n\}$ 为可测函数列, 则存在可测函数 f 使 $f_n\xrightarrow{\text{a.e.}}f$ 当且仅当 $\{f_n\}$ 几乎处处相互收敛.

证明 如 $f_n\xrightarrow{\text{a.e.}}f$, 则存在零测集 N 使 $f_n(\omega)\to f(\omega),\omega\notin N$, 从而 $\forall\omega\notin N,\{f_n(\omega)\}_{n\geqslant 1}$ 是 Cauchy 列, 即当 $n,m\to\infty$ 时, $f_n(\omega)-f_m(\omega)\to 0$. 故 $\{f_n\}_{n\geqslant 1}$ 几乎处处相互收敛.

反之, 如 $\{f_n\}$ 几乎处处相互收敛, 则存在零测集 N 使 $\forall\omega\notin N,\{f_n(\omega)\}_{n\geqslant 1}$ 是 Cauchy 列, 从而有极限, 记作 $f(\omega)$. 当 $\omega\in N$ 时, 令 $f(\omega)=0$. 由于可测函数的极限也可测, 从而 f 可测且 $f_n\xrightarrow{\text{a.e.}}f$. □

下面定理可由几乎处处收敛与几乎处处相互收敛的定义立得.

定理 2.37. 设 $f,f_n,n\geqslant 1$ 是有限可测函数.

1) $f_n \xrightarrow{\text{a.e.}} f$ 当且仅当 $\forall\varepsilon>0, \mu\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m=n}^{\infty}\{|f_m-f|\geqslant\varepsilon\}\right)=0$. 特别地, 当 μ 有限时,

$$f_n \xrightarrow{\text{a.e.}} f \text{ 当且仅当 } \forall\varepsilon>0, \mu\left(\bigcup_{m=n}^{\infty}\{|f_m-f|\geqslant\varepsilon\}\right)\to 0(n\to\infty).$$

2) $f_n-f_m \xrightarrow{\text{a.e.}} 0$ 当且仅当 $\forall\varepsilon>0, \mu\left(\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{v=1}^{\infty}\{|f_{n+v}-f_n|\geqslant\varepsilon\}\right)=0$. 特别地, 当 μ 有限时, $f_n-f_m \xrightarrow{\text{a.e.}} 0$ 当且仅当

$$\forall\varepsilon>0, \mu\left(\bigcup_{v=1}^{\infty}\{|f_{n+v}-f_n|\geqslant\varepsilon\}\right)\to 0(n\to\infty).$$

§ 2.4.2 依测度收敛

定义 2.38. 设 $\{f_n\}_{n\geqslant1}$ 是有限可测函数序列, f 可测. 如果 $\forall\varepsilon>0, \mu(|f_n-f|\geqslant\varepsilon)\to 0(n\to\infty)$, 则称 $\{f_n\}_{n\geqslant1}$ 依测度 μ 收敛于 f, 记作 $f_n \xrightarrow{\mu} f$.

如果 $\forall\varepsilon>0$,

$$\sup_{v\geqslant1}\mu(|f_{n+v}-f_n|\geqslant\varepsilon)\to 0, n\to\infty,$$

则称 $\{f_n\}$ 依测度 μ 相互收敛, 记作 $f_{n+v}-f_n \xrightarrow{\mu} 0$.

易见, 若 $f_n \xrightarrow{\mu} f$, 则 f a.e. 有限. 下面的性质是显然的.

性质 2.39. **1)** $f_n \xrightarrow{\mu} f$, 则任何子列 $f_{n_k} \xrightarrow{\mu} f$.

2) $f_n \xrightarrow{\mu} f, f_n \xrightarrow{\mu} f'$, 则 $f=f'$ a.e.

3) $f_n \xrightarrow{\mu} f, g_n=f_n$ a.e., $g=f$ a.e., 则 $g_n \xrightarrow{\mu} g$.

定理 2.40. 设 $f, f_n:\Omega\to\mathbb{R}^m$ 可测且 $D\supset f(\Omega), D\supset\bigcup\limits_{n=1}^{\infty}f_n(\Omega)$. 如 $g:D\to\mathbb{R}$ 一致连续且 $f_n \xrightarrow{\mu} f$, 则 $g(f_n) \xrightarrow{\mu} g(f)$.

证明 $\forall\varepsilon>0, \exists\delta>0$, 当 $x, y\in D$ 且 $|x-y|<\delta$ 时, 有 $|g(x)-g(y)|<\varepsilon$. 则 $\{|g(f_n)-g(f)|\geqslant\varepsilon\}\subset\{|f_n-f|\geqslant\delta\}$, 得证. □

推论 2.41. 若 $f_n \xrightarrow{\mu} f, g_n \xrightarrow{\mu} g$, 则 $f_n+g_n \xrightarrow{\mu} f+g$.

定理 2.42. 在定理 2.40 中, 若 μ 是 $(\Omega, \mathscr{A})$ 上有限测度且 D 是开集, 则 g 可换成连续函数.

证明 令

$$D_N=\{x\in\mathbb{R}^n:|x|\leqslant N,d(x,D^c)\geqslant N\},d(x,\varnothing)=\infty.$$

则 D_N 是有界闭集 (因为 $d(\cdot,D^c)$ 连续). 由当 $N\uparrow\infty$ 时 $D_N\uparrow D$ 知 $\mu(f^{-1}(D\backslash D_N))\downarrow 0$. $\forall\varepsilon\in(0,1)$, 由于 g 是 D_{N+1} 上一致连续函数, 故存在 $\delta_N>0$ 使 $\forall x,y\in D_{N+1},|x-y|<\delta_N$ 时, $|g(x)-g(y)|<\varepsilon$. 则

$$\begin{aligned}A_n:&=\{|g(f_n)-g(f)|\geqslant\varepsilon\}\subset(A_n\cap\{f_n,f\in D_{N+1}\})\cup\{f\notin D_N\}\\&\subset\{|f_n-f|\geqslant\delta_N\}\cup\{f\notin D_N\}.\end{aligned}$$

而 $\lim\limits_{n\to\infty}\mu(A_n)\leqslant 0+\mu(f^{-1}(D\backslash D_N))$. 令 $N\uparrow\infty$ 得 $\lim\limits_{n\to\infty}\mu(|g(f_n)-g(f)|\geqslant\varepsilon)=0$.

最后说明两种收敛性的关系.

定理 2.43. 设 $\{f_n\}_{n\geqslant 1}$ 是有限可测函数序列.

1) 如果 $f_n\xrightarrow{\mu}f$, 则存在子列 $\{f_{n_k}\}$ 使 $f_{n_k}\xrightarrow{\text{a.e.}}f$.

2) 如果 $f_{n+v}-f_n\xrightarrow{\mu}0$, 则存在子列 $\{f_{n_k}\}$ 及有限可测函数 f 使 $f_{n_k}\xrightarrow{\text{a.e.}}f$ 且 $f_{n_k}\xrightarrow{\mu}f$.

3) 若 μ 是有限测度, 则 $f_n\xrightarrow{\text{a.e.}}f$ 蕴含 $f_n\xrightarrow{\mu}f$.

证明 1) $\forall k\geqslant 1,\exists n_k\uparrow\infty$ 使 $\mu(|f_n-f|\geqslant 2^{-k})<2^{-k},n\geqslant n_k$. 令 $f'_k=f_{n_k}$, 则 $\mu(|f'_k-f|\geqslant\frac{1}{2^k})<2^{-k},n\geqslant n_k$. 从而 $\forall\varepsilon>0$ 及 $k'\geqslant 1$ 使 $2^{-k'}\leqslant\varepsilon$,

$$\mu\left(\bigcap_{k=1}^{\infty}\bigcup_{v=1}^{\infty}\{|f'_{k+v}-f|\geqslant\varepsilon\}\right)\leqslant\sum_{v=1}^{\infty}\mu(|f'_{k'+v}-f|\geqslant\varepsilon)\leqslant\sum_{v=1}^{\infty}2^{-(k'+v)}=2^{-k'}.$$

令 $k'\uparrow\infty$, 由定理 2.37 之 1) 即得.

2) 如 1) 取 $n_k\uparrow\infty$ 使

$$\sup_{v\geqslant 1}\mu(|f_{n_k+v}-f_{n_k}|\geqslant 2^{-k})<2^{-k},$$

则 $\forall\varepsilon>0$ 及 $k'\geqslant 1$ 使 $2^{-k'}\leqslant\varepsilon$,

$$\begin{aligned}\mu\left(\bigcap_{k=1}^{\infty}\bigcup_{v=1}^{\infty}\{|f'_{k+v}-f'_k|\geqslant\varepsilon\}\right)&\leqslant\sum_{i=0}^{\infty}\mu(|f'_{k'+i+1}-f'_{k'+i}|\geqslant\varepsilon)\\&\leqslant\sum_{i=0}^{\infty}2^{-(k'+i)}=2^{-k'+1}.\end{aligned}$$

令 $k' \uparrow \infty$, 由定理 2.37 之 2) 得 $\{f_{n_k}\}$ 几乎处处相互收敛, 从而几乎处处收敛于某有限可测函数 f.

再证 $f'_k \xrightarrow{\mu} f$. 由 $f'_k \xrightarrow{\text{a.e.}} f$ 知 $\exists$ 零测集 $N, \forall \omega \notin N, f'_k(\omega) \to f(\omega)$. 则

$$\{|f'_k - f| \geqslant \varepsilon\} \subset N \bigcup \left(\bigcup_{i=1}^{\infty} \{|f'_{k+i} - f'_{k+i-1}| \geqslant 2^{-i}\varepsilon\} \right),$$

由此知当 $\varepsilon \geqslant 2^{1-k}$ 时有

$$\begin{aligned}\mu(|f'_k - f| \geqslant \varepsilon) &\leqslant \sum_{i=1}^{\infty} \mu\left(|f'_{k+i} - f'_{k+i-1}| \geqslant 2^{-(k+i-1)}\right) \\ &\leqslant \sum_{i=1}^{\infty} 2^{-(k+i-1)} = 2^{1-k}.\end{aligned}$$

3) 设 μ 有限且 $f_n \xrightarrow{\text{a.e.}} f$. 则

$$\mu(|f_n - f| \geqslant \varepsilon) \leqslant \mu\left(\bigcup_{m=n}^{\infty} \{|f_m - f| \geqslant \varepsilon\} \right).$$

由此结合 $f_n \xrightarrow{\text{a.e.}} f$ 及测度的上连续性得

$$\lim_{n\to\infty} \mu(|f_n - f| \geqslant \varepsilon) \leqslant \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{|f_m - f| \geqslant \varepsilon\} \right) = 0. \quad \square$$

定理 2.44. 存在有限可测函数 f 使 $f_n \xrightarrow{\mu} f$ 当且仅当 $f_{n+v} - f_n \xrightarrow{\mu} 0$.

证明 必要性显然, 使用三角不等式即可. 往证充分性.

设 $f_{n+v} - f_n \xrightarrow{\mu} 0$. 由上一定理, 存在子列 $f_{n_k} \xrightarrow{\mu}$ 某 f . 则

$$\begin{aligned}&\lim_{k\to\infty} \mu(|f_k - f| \geqslant \varepsilon) \\ &\quad\leqslant \lim_{k\to\infty} \mu\left(|f_k - f_{n_k}| \geqslant \frac{\varepsilon}{2}\right) + \lim_{k\to\infty} \mu\left(|f_{n_k} - f| \geqslant \frac{\varepsilon}{2}\right) = 0. \quad \square\end{aligned}$$

§ 2.4.3 依分布律收敛

定义 2.45. 设 $\{\xi_n\}_{n\geqslant 1}$ 为随机向量序列, ξ_n 的分布函数为 F_n, ξ 的分布函数为 F. 若对 F 的任意连续点 x_0 有 $F_n(x_0) \to F(x_0)$, 则称 ξ_n 依分布律收敛于 ξ. 记作 $F_n \xrightarrow{c} F$ 或 $\xi_n \xrightarrow{d} \xi$.

定理 2.46. 如 $\xi_n \xrightarrow{\mathbb{P}} \xi$, 则 $\xi_n \xrightarrow{d} \xi$.

证明 使用概率不等式 $|\mathbb{P}(A)-\mathbb{P}(B)| \leqslant \mathbb{P}(A\triangle B)$, 其中 $A\triangle B=(A-B)\cup(B-A)$ 为 A 与 B 对称差. 设 e 为各分量为 1 的向量, 则

$$
\begin{aligned}
&|F_n(x)-F(x)| = |\mathbb{P}(\xi_n<x)-\mathbb{P}(\xi<x)|\\
&\leqslant \mathbb{P}(\xi_n<x, \xi\in(-\infty,x)^c)+\mathbb{P}(\xi_n\in(-\infty,x)^c, \xi<x)\\
&\leqslant \mathbb{P}(\xi_n<x, \xi\geqslant x+\varepsilon e)+\mathbb{P}(\xi_n\geqslant x, \xi<x-\varepsilon e)+\mathbb{P}(x-\varepsilon e\leqslant\xi<x+\varepsilon e)\\
&\leqslant \mathbb{P}(|\xi-\xi_n|\geqslant\varepsilon)+F(x+\varepsilon e)-F(x-\varepsilon e).
\end{aligned}
$$

如 x 是 F 的连续点, 先令 $n\uparrow\infty$, 再令 $\varepsilon\downarrow 0$ 即得 $F_n(x)\to F(x)$. □

推论 2.47. 设 $a\in\mathbb{R}^n$, 则 $\xi_n \xrightarrow{\mathbb{P}} a$ 当且仅当 $\xi_n \xrightarrow{d} a$.

证明 只需证明充分性. 由于 $\forall\varepsilon>0, a-\varepsilon$ 与 $a+\varepsilon$ 均为 $\xi\equiv a$ 的分布函数 F 的连续点, 且 $F(x)=\mathbf{1}_{(a,\infty)}$. 则由 $\xi_n \xrightarrow{d} a$ 知 $\mathbb{P}(|\xi_n-a|>\varepsilon)=\mathbb{P}(\xi_n<a-\varepsilon)+\mathbb{P}(\xi_n>a+\varepsilon)\to 0(n\to\infty)$. □

类似地, 容易证明如下的两个定理.

定理 2.48. 若 $\xi_n-\xi_n' \xrightarrow{\mathbb{P}} 0$ 且 $\xi_n' \xrightarrow{d} \xi$, 则 $\xi_n \xrightarrow{d} \xi$.

定理 2.49. 如果 $\xi_n \xrightarrow{d} \xi, \eta_n \xrightarrow{d} a$ 常数, 则 $\xi_n+\eta_n \xrightarrow{d} \xi+a$.

§2.5 补充与习题

1. 证明性质 2.3.

2. 证明推论 2.8.

3. 证明推论 2.9.

4. 证明定理 2.13.

5. 分布函数是否是不降的? 试举反例或给出证明.

6. 在 $(\mathbb{R}^n, \mathscr{B}^n)$ 的有限区间上取有限值的测度, 称为 L-S 测度. 证明每个 L-S 测度都是某分布函数生成的 Lebesgue-Stieljes 测度.

7. 证明若 $F(x) = \mathbb{P}(\xi < x)$ 是连续的, 则 $\eta = F(\xi)$ 具有 $(0,1)$ 上的均匀分布.

8. 证明性质 2.30.

9. 设随机变量 ξ, η 独立, 其分布函数分别为 F, G , 求 $\xi + \eta$ 的分布函数.

10. 设 $\xi_1, \xi_2, \cdots$ 是 $i.i.d.$ 随机变量序列, 分布律为 μ. 给定 $A \in \mathscr{B}, \mu(A) > 0$, 定义 $\tau = \inf\{k : \xi_k \in A\}$. 证明 ξ_τ 的分布律为 $\mu(\cdot \cap A)/\mu(A)$.

11. 设 $\xi, \tilde{\xi}$ 独立同分布, 令 $\eta = \xi - \tilde{\xi}$ (称 η 为 ξ 的对称化). 证明 $\mathbb{P}(|\eta| > t) \leqslant 2\mathbb{P}(|\xi| > \frac{t}{2})$.

12. 如果 π 系 $\mathscr{C}_1, \cdots, \mathscr{C}_n$ 独立, 那么 $\sigma(\mathscr{C}_1), \cdots, \mathscr{C}_n$ 独立.

13. (a) 设 $\{A_n\}_{n \geqslant 1}$ 为独立事件序列, 令 $\mathscr{T} = \bigcap\limits_{n=1}^{\infty} \sigma\{A_n, A_{n+1}, \cdots\}$. 证明 $\forall A \in \mathscr{T}$, 有 $P(A) = 0$ 或 1.
    (b) 设 $\{\xi_n\}_{n \geqslant 1}$ 为独立随机变量序列, 令 $\mathscr{T} = \bigcap\limits_{n=1}^{\infty} \sigma\{\xi_n, \xi_{n+1}, \cdots\}$. 证明 $\forall A \in \mathscr{T}$, 有 $P(A) = 0$ 或 1.

14. 证明性质 2.35

15. 证明定理 2.37

16. 证明 $\xi_n \xrightarrow{\mathbb{P}} \xi$ 当且仅当

$$\mathbb{E}\left(\frac{|\xi_n - \xi|}{1 + |\xi_n - \xi|}\right) \to 0.$$

17. 设 $\xi_1, \xi_2, \cdots \in \{1, 2, \cdots, r\}$ 独立服从分布 $\mathbb{P}(\xi_i = k) = p(k) > 0, 1 \leqslant k \leqslant r$. 令 $\pi_n(\omega) = p(\xi_1(\omega)) \cdots p(\xi_n(\omega))$, 证明

$$-n^{-1} \log \pi_n \xrightarrow{\mathbb{P}} H \triangleq -\sum_{k=1}^{r} p(k) \log p(k).$$

这里 H 称为 Shannon 信息熵.

18. 令 $\xi_n = \mathbf{1}_{A_n}$, 则 $\xi \xrightarrow{\mathbb{P}} 0$ 当且仅当 $\mathbb{P}(A_n) \to 0$.

19. 如果 f 关于 $\sigma(\mathscr{C})$ 可测, 那么存在 $\mathscr{C}$ 的可数子类 $\mathscr{C}_f$ 使得 $f \in \mathscr{C}_f$.

20. 设 F 为分布函数.

    (a) 证明 $\int_{\mathbb{R}} (F(x+c) - F(x))\,\mathrm{d}x = c$;

    (b) 对连续分布函数 F, 证明 $\int_{\mathbb{R}} F(x)\,\mathrm{d}F(x) = 1/2$.

21. 若随机变量序列 $\{\xi_n\}_{n \geqslant 1}$ 单调上升且 $\xi_n \xrightarrow{\mathbb{P}} \xi$, 则 $\xi_n \xrightarrow{\text{a.e.}} \xi$.

22. (a) 若 $\xi_n \xrightarrow{\text{a.e.}} \xi$, 则

$$S_n \triangleq \frac{1}{n} \sum_{k=1}^{n} \xi_k \xrightarrow{\text{a.e.}} \xi;$$

    (b) 若 $\xi_n \xrightarrow{\mathbb{P}} \xi$, 则 $S_n \xrightarrow{\mathbb{P}} \xi$ 是否成立?

23. 证明在离散的概率空间依概率收敛等价于 a.e. 收敛.

24. (Egorov 定理) 设 $(\Omega, \mathscr{A}, \mu)$ 是有限测度空间, 可测函数 $f_n \xrightarrow{\text{a.e.}} f$, 证明 $\forall \varepsilon > 0, \exists N \in \mathscr{A}, \mu(N) \leqslant \varepsilon$ 使得 f_n 在 N^c 上一致收敛到 f.

25. 若 ξ_n 依分布收敛于 ξ, 则 $\mathbb{E}|\xi| \leqslant \varliminf\limits_{n\to\infty} E|\xi_n|$.

26. 对任意对随机变量序列 ξ_n, 存在正数序列 a_n 使得 $a_n\xi_n \xrightarrow{d} 0$.

27. 设随机变量 ξ_n, ξ 分别具有密度函数 f_n, f. 证明若 $f_n \xrightarrow{\text{a.e.}} f$ 则 $\xi_n \xrightarrow{d} \xi$.

28. 举例说明当 g 仅为连续函数时, 定理 2.40 的结论不成立.

29. 证明定理 2.48 和定理 2.49.

30. 设 ξ_n, ξ 的分布函数分别为 F_n, F. 若 $\xi_n \xrightarrow{d} \xi$, 则对 F 的任意连续点 x 有 $\mathbb{P}(\xi_n \leqslant x) \to \mathbb{P}(\xi \leqslant x), \mathbb{P}(\xi_n > x) \to \mathbb{P}(\xi > x)$.

31. 设随机变量序列 $\{\xi_n\}_{n\geqslant 1}$ 独立同分布具分布函数 F. 令

$$F_n(x)\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{(-\infty,x)}(\xi_k).$$

证明 $\sup\limits_{x\in\mathbb{R}}|F_n(x) - F(x)| \xrightarrow{\mathbb{P}} 0$.

32. 设 $\{\mathbb{P}_n\}_{n\geqslant 1}$ 是 $(\Omega, \mathscr{A})$ 上一列概率测度. 若 $\forall A \in \mathscr{A}$, 有 $\mathbb{P}_n(A) \to \mathbb{P}(A)$, 则 $\mathbb{P}$ 是概率测度.

33. 如果分布函数 $F_n \Rightarrow F_\infty$, 那么存在随机变量 $\{\xi_n\}_{n\geqslant 1}$ 具有分布 F_n 使得 $\xi_n \xrightarrow{\text{a.s.}} \xi_\infty$.

第三章　数学期望与积分

我们在初等概率论中学习了如何计算离散型随机变量和连续型随机变量的数学期望, 前者的数学期望可定义为随机变量的取值关于分布列的加权和, 而后者的数学期望则可定义成恒同函数关于分布密度的积分. 那么, 如何定义一般随机变量的数学期望呢?

另一方面, 在数学分析中我们学习了函数的黎曼积分, 它可以看成是微分的逆运算. 为定义更一般函数的积分, 在实变函数论中引入了 Lebesgue 积分. 该积分的重要意义在于摆脱了黎曼积分的定义中所使用的对自变量微分的手法, 从而可以推广到一般测度空间上.

如果说可测集的测度是对该集合的测量结果, 那么函数的积分则可视为对该函数的测量结果. 为测量一个可测函数 (或定义可测函数的积分), 上一章所介绍的可测函数的构造与极限定理发挥将至关重要的作用.

当一般测度空间上的可测函数的积分获得定义之后, 一般随机变量的数学期望的定义就变得非常简单了, 它就是恒同函数关于随机变量分布的积分.

§3.1　积分的定义和性质

§ 3.1.1　积分的定义

设 $(\Omega, \mathscr{A}, \mu)$ 为完备测度空间, f 为 Ω 上实可测函数. 我们将从非负简单函数开始定义积分. 当 $f = \mathbf{1}_A (A \in \mathscr{A})$ 时, 由于对 f 的测量等同于对 A 的测量, f 的积分自然定义为 $\mu(A)$. 由此结合积分应具备的线性性质, 我们引入非负简单函数积分的定义. 由于我们容许函数值为无穷, 约定 $0 \times \infty = 0$.

定义 3.1. 设 f 是非负简单函数, $f = \sum\limits_{k=1}^{n} a_k \mathbf{1}_{A_k}, a_k \in [0,\infty], A_1, \cdots, A_n \in \mathscr{A}$ 两两不交, $\sum\limits_{k=1}^{n} A_k = \Omega$. 则称 $\int_\Omega f\,\mathrm{d}\mu = \sum\limits_{k=1}^{n} a_k\mu(A_k)$ 为 f 在 Ω 上关于 μ 的积分. 设 $A \in \mathscr{A}$, 称 $\int_A f\,\mathrm{d}\mu = \int_\Omega f\mathbf{1}_A\,\mathrm{d}\mu$ 为 f 在 A 上关于 μ 的积分.

易见，以上的定义是合理的, 即 $\int_\Omega f\,\mathrm{d}\mu$ 的值与简单函数 f 的表示无关. 由于积分是使用测度来测量函数, 通常也记 $\mu(f) = \int_\Omega f\,\mathrm{d}\mu$.

性质 3.2. 设 f, g 是非负简单函数.

1) $f \leqslant g \Rightarrow \mu(f) \leqslant \mu(g)$.

2) $\forall c \geqslant 0, \mu(cf) = c\mu(f)$.

3) 令 $\mu_f(A) = \int_A f\,\mathrm{d}\mu$, 则 μ_f 为 $\mathscr{A}$ 上的测度, 且任给非负简单函数 g 有 $\int_\Omega g\,\mathrm{d}\mu_f = \int_\Omega fg\,\mathrm{d}\mu$.

下面我们定义非负可测函数的积分.

定义 3.3. 若 f 是非负可测函数, 则称

$$\int_\Omega f\,\mathrm{d}\mu = \sup\left\{\int_\Omega g\,\mathrm{d}\mu : 0 \leqslant g \leqslant f, g \text{ 为简单函数}\right\}$$

为 f 在 Ω 上关于 μ 的积分. $\forall A \in \mathscr{A}$, 定义 $\int_A f\,\mathrm{d}\mu = \int_\Omega f\mathbf{1}_A\,\mathrm{d}\mu$.

性质 3.4. 如 $0 \leqslant f \leqslant g$, 则 $\mu(f) \leqslant \mu(g)$.

定理 3.5 (单调收敛定理). 若 $\{f_n\}_{n\geqslant 1}$ 非负可测, 且 $f_n \uparrow f$, 则 $\lim\limits_{n\to\infty} \mu(f_n) = \mu(f)$.

证明 首先，因为 $\mu(f_n)$ 不降，因而极限存在. 又由于 $f_n \leqslant f$, 故 $\lim\limits_{n\to\infty} \mu(f_n) \leqslant \mu(f)$, 只需证明相反不等式. 任给简单函数 $0 \leqslant g \leqslant f, g = \sum\limits_{j=1}^{m} a_j \mathbf{1}_{A_j} + \infty\mathbf{1}_{A_{m+1}}$ 使 $g \leqslant f$. 由于 $\lim\limits_{n\to\infty} f_n \geqslant g$, 则 $\forall \varepsilon \in (0, \min\limits_{1\leqslant j\leqslant m} a_j)$ 及 $N \geqslant 1$, 当 n 充分大时有 $f_n \geqslant \sum\limits_{j=1}^{m} (a_j - \varepsilon)\mathbf{1}_{A_j \cap \{|f_n - f| \leqslant \varepsilon\}} + N\mathbf{1}_{A_{m+1}\cap\{|f_n|\geqslant N\}} \triangleq g_n$. 由此

及 μ 的下连续性知

$$\begin{aligned}
&\lim_{n\to\infty}\mu(f_n)\geqslant\lim_{n\to\infty}\mu(g_n)\\
&=\sum_{j=1}^{m}(a_j-\varepsilon)\lim_{n\to\infty}\mu(A_j\cap\{|f_n-f|\leqslant\varepsilon\})+N\mu(A_{m+1}\cap\{|f_n|\geqslant N\})\\
&=\sum_{j=1}^{m}(a_j-\varepsilon)\mu(A_j)+N\mu(A_{m+1}).
\end{aligned}$$

由 ε 和 N 的任意性得 $\lim\limits_{n\to\infty}\mu(f_n)\geqslant\mu(g)$. 从而由 g 的任意性得 $\lim\limits_{n\to\infty}\mu(f_n)\geqslant\mu(f)$. □

定义 3.6. **1)** 设 f 是实可测函数，f^+,f^- 分别是正部与负部，若 $\mu(f^+)$ 与 $\mu(f^-)$ 至少有一个有限，则称 f 的积分存在并定义 f 在 Ω 上对 μ 的积分为

$$\mu(f)=\int_\Omega f\,\mathrm{d}\mu=\mu(f^+)-\mu(f^-).$$

若 $\mu(f)$ 有限，则称 f 可积. 对 $A\in\mathscr{A}$ 使 $\mathbf{1}_Af$ 积分存在, 定义 $\int_A f\,\mathrm{d}\mu=\int_\Omega f\mathbf{1}_A\,\mathrm{d}\mu$.

2) 设 $f=f_1+\mathrm{i}\,f_2$ 是可测复函数, 如果 f_1,f_2 积分存在，则定义 $\int_\Omega f\,\mathrm{d}\mu=\int_\Omega f_1\,\mathrm{d}\mu+\mathrm{i}\int_\Omega f_2\,\mathrm{d}\mu$.

命题 3.7. 若 $f=g$ a.e. 且积分存在, 则 $\mu(f)=\mu(g)$.

§ 3.1.2　积分的性质

下面的定理总结了可测函数积分的常用性质. 由于证明简单且与 Lebesgue 积分的有关证明完全一致, 故从略.

定理 3.8. 设 f,g 为实可测函数.

1) 线性性质

- **a)** 如果 $\int_\Omega f\,\mathrm{d}\mu+\int_\Omega g\,\mathrm{d}\mu$ 存在, 则 $f+g$ 积分存在且 $\int_\Omega(f+g)\,\mathrm{d}\mu=\int_\Omega f\,\mathrm{d}\mu+\int_\Omega g\,\mathrm{d}\mu$.
- **b)** 若 $\int_\Omega f\,\mathrm{d}\mu$ 存在且 $A\cap B=\emptyset$, 则 $\int_{A+B}f\,\mathrm{d}\mu=\int_A f\,\mathrm{d}\mu+\int_B f\,\mathrm{d}\mu$.
- **c)** 若 $c\in\mathbb{R}$, $\int_\Omega f\,\mathrm{d}\mu$ 存在, 则 $\int_\Omega cf\,\mathrm{d}\mu$ 存在, 且 $\int_\Omega cf\,\mathrm{d}\mu=c\int_\Omega f\,\mathrm{d}\mu$.

2) 单调性

- **a)** 设 f,g 积分存在且 $f\geqslant g$, a.e. 则 $\forall A\in\mathscr{A}$, $\int_A f\,\mathrm{d}\mu\geqslant\int_A g\,\mathrm{d}\mu$.

b) 设 f 积分存在, 则 $|\int_\Omega f\,\mathrm{d}\mu| \leqslant \int_\Omega |f|\,\mathrm{d}\mu$.

c) 设 $f \geqslant 0$, 则 $\int_\Omega f\,\mathrm{d}\mu = 0$ 当且仅当 $f = 0$ a.e.

d) 设 N 为零测集, 则 $\int_N f\,\mathrm{d}\mu = 0$.

3) 可积性

a) f 可积当且仅当 $\int_\Omega |f|\,\mathrm{d}\mu < \infty$; 当 f 可积时, f a.e. 有限.

b) 设 $|f| \leqslant g$ 且 g 可积, 则 f 可积.

c) 若 f, g 可积, 则 $f + g$ 可积.

d) 若 $\int_\Omega fg\,\mathrm{d}\mu$ 存在, 则 $|\int_\Omega fg\,\mathrm{d}\mu|^2 \leqslant \int_\Omega |f|^2\,\mathrm{d}\mu \int_\Omega |g|^2\,\mathrm{d}\mu$.

推论 3.9. 若 f 是 A 上取非负值的可测函数, 则 $\forall c > 0$ 有 $\mu(\{f \geqslant c\} \cap A) \leqslant \frac{1}{c}\int_A f\,\mathrm{d}\mu$.

证明 设 $g = c\mathbf{1}_{A\cap\{f\geqslant c\}}$, 则 $g \leqslant \mathbf{1}_A f$, 故 $\int_\Omega g\,\mathrm{d}\mu \leqslant \int_A f\,\mathrm{d}\mu$, 即 $c\mu(\{f \geqslant c\} \cap A) \leqslant \int_A f\,\mathrm{d}\mu$. □

§3.2 收敛定理

基于单调收敛定理, 我们引入如下两个重要的收敛定理.

定理 3.10 (Fatou-Lebesgue 定理). 设 g, h 是实可积函数, $\{f_n\}_{n\geqslant 1}$ 是实可测的函数列.

1) 若 $\forall n \geqslant 1, g \leqslant f_n$, 则 $\int_\Omega \varliminf_{n\to\infty} f_n\,\mathrm{d}\mu \leqslant \varliminf_{n\to\infty} \int_\Omega f_n\,\mathrm{d}\mu$.

2) 若 $\forall n \geqslant 1, f_n \leqslant g$, 则 $\varlimsup_{n\to\infty} \int_\Omega f_n\,\mathrm{d}\mu \leqslant \int_\Omega \varlimsup_{n\to\infty} f_n\,\mathrm{d}\mu$.

3) 若 $g \leqslant f_n \uparrow f$ 或 $\forall n \geqslant 1, g \leqslant f_n \leqslant h$, a.e., $f_n \xrightarrow{\text{a.e.}} f$, 则 $\lim_{n\to\infty} \int_\Omega f_n\,\mathrm{d}\mu = \int_\Omega f\,\mathrm{d}\mu$.

证明 由于 $g \leqslant f_n$, 则 $g^- \geqslant f_n{}^-$, 从而 $\int_\Omega f_n{}^-\,\mathrm{d}\mu < \infty$, 故 $\int_\Omega f_n\,\mathrm{d}\mu$ 存在. 同理在 2) 与 3) 中 $\int_\Omega f_n\,\mathrm{d}\mu$ 也存在.

1) 令 $g_n = \inf_{k\geqslant n}(f_k - g)$, 则 $g_n \geqslant 0$ 且

$$g_n \uparrow \varliminf_{n\to\infty}(f_n - g) = \varliminf_{n\to\infty} f_n - g.$$

由单调收敛定理,

$$\begin{aligned}&\int_\Omega \varliminf_{n\to\infty} f_n\,\mathrm{d}\mu-\int_\Omega g\,\mathrm{d}\mu=\lim_{n\to\infty}\int_\Omega \inf_{k\geqslant n}(f_k-g)\,\mathrm{d}\mu\\&\leqslant \varliminf_{n\to\infty}\int_\Omega (f_n-g)\,\mathrm{d}\mu=\varliminf_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu-\int_\Omega g\,\mathrm{d}\mu.\end{aligned}$$

2) 只需考察 $-f_n$, 由 1) 立得.

3) 当 $g\leqslant f_n\uparrow f$ 时, $0\leqslant f_n-g\uparrow f-g$, 由单调收敛定理即得. 当 $g\leqslant f_n\leqslant h$ a.e. 且 $f_n\xrightarrow{\text{a.e.}} f$ 时, 令 N 是零测集使得在 N^c 上有 $g\leqslant f_n\leqslant h, f_n\to f$. 则 $g\mathbf{1}_{N^c}\leqslant f_n\mathbf{1}_{N^c}\leqslant h\mathbf{1}_{N^c}$. 由 1) 和 2) 知 $\lim\limits_{n\to\infty}\int_\Omega f_n\,\mathrm{d}\mu=\lim\limits_{n\to\infty}\int_\Omega f_n\mathbf{1}_{N^c}\,\mathrm{d}\mu=\int_\Omega \lim\limits_{n\to\infty} f_n\mathbf{1}_{N^c}\,\mathrm{d}\mu=\int_\Omega f\mathbf{1}_{N^c}\,\mathrm{d}\mu=\int_\Omega f\,\mathrm{d}\mu$. □

定理 3.11 (控制收敛定理). 设 g 是可积函数, $|f_n|\leqslant g$ a.e. 如果 $f_n\xrightarrow{\text{a.e.}} f$ 或 $f_n\xrightarrow{\mu} f$, 则 $\int_\Omega f_n\,\mathrm{d}\mu\to\int_\Omega f\,\mathrm{d}\mu$.

证明 由定理 3.10 之 3), 只需证 $f_n\xrightarrow{\mu} f$ 之情形.

由积分的单调性, 只需证 $\lim\limits_{n\to\infty}\int_\Omega |f_n-f|\,\mathrm{d}\mu=0$. 若不成立, 则存在 $n_k\uparrow\infty$ 及 $\varepsilon>0$, 使 $\int_\Omega |f_{n_k}-f|\,\mathrm{d}\mu\geqslant\varepsilon, \forall k\geqslant 1$. 由于 $f_{n_k}\xrightarrow{\mu} f$, 存在子列 $f_{n_k'}\xrightarrow{\text{a.e.}} f$, 从而 $\lim\limits_{n\to\infty}\int_\Omega |f_{n_k'}-f|\,\mathrm{d}\mu=0$, 矛盾. □

推论 3.12. 设 $\{f_n\}_{n\geqslant 1}$ 是可测函数列, 若 f_n 非负或 $\sum\limits_{n=1}^\infty\int_\Omega |f_n|\,\mathrm{d}\mu<\infty$, 则 $\sum\limits_{n=1}^\infty f_n$ 的积分存在, 且 $\int_\Omega\sum\limits_{n=1}^\infty f_n\,\mathrm{d}\mu=\sum\limits_{n=1}^\infty\int_\Omega f_n\,\mathrm{d}\mu$.

证明 令 $g_n=\sum\limits_{k=1}^n f_n$. 当 f_n 非负时, $g_n\uparrow\sum\limits_{n=1}^\infty f_n$, 则由单调收敛定理立得. 若 $\sum\limits_{n=1}^\infty\int_\Omega |f_n|\,\mathrm{d}\mu<\infty$, 令 $g'=\sum\limits_{n=1}^\infty |f_n|, g_n'=\sum\limits_{k=1}^n |f_k|$, 则 $0\leqslant {g'}_n\uparrow g'$. 由单调收敛定理,

$$\sum_{n=1}^\infty\int_\Omega |f_n|\,\mathrm{d}\mu=\lim_{n\to\infty}\int_\Omega g_n'\,\mathrm{d}\mu=\int_\Omega g'\,\mathrm{d}\mu=\int_\Omega\sum_{n=1}^\infty |f_n|\,\mathrm{d}\mu,$$

则 g' 可积且 $|g_n|\leqslant g'$. 由于 $\sum\limits_{n=1}^\infty\int_\Omega |f_n|\,\mathrm{d}\mu<\infty$ 且 g' a.e. 有限, 故 $g_n\xrightarrow{\text{a.e.}}\sum\limits_{n=1}^\infty f_n$. 由控制收敛定理即得. □

推论 3.13. 若 $\int_\Omega f\,\mathrm{d}\mu$ 存在, 则 $\forall A\in\mathscr{A},\{A_n\}_{n=1}^\infty\subset\mathscr{A}$ 两两不交使 $A=\sum\limits_{n=1}^\infty A_n$, 有 $\int_A f\,\mathrm{d}\mu=\sum\limits_{n=1}^\infty\int_{A_n} f\,\mathrm{d}\mu$.

证明 由 $f^\pm\mathbf{1}_A=\sum\limits_{n=1}^\infty f^\pm\mathbf{1}_{A_n}$, 有 $\int_A f^\pm\,\mathrm{d}\mu=\sum\limits_{n=1}^\infty\int_{A_n} f^\pm\,\mathrm{d}\mu$. 由于 f 积分存在, 则上述级数至少有一个有限, 因此可以逐项相减, 从而得所求结论. □

定义 3.14. 设 f 积分存在, 称符号测度 $\mu_f(A)\triangleq\int_A f\,\mathrm{d}\mu, A\in\mathscr{A}$ 为 f 的**不定积分**.

推论 3.15. 若 f 可积, 则当 $\mu(A_n)\to 0$ 时, 有 $\int_{A_n} f\,\mathrm{d}\mu\to 0$.

证明 若 $\int_{A_n} f\,\mathrm{d}\mu\nrightarrow 0$, 由于 $|\int_{A_n} f\,\mathrm{d}\mu|\leqslant\int|f|\,\mathrm{d}\mu<\infty$, 则存在 $n_k\uparrow\infty$ 使 $\int_{A_{n_k}} f\,\mathrm{d}\mu\to\epsilon\neq 0$. 取 $\{n_k\}$ 的子列 $\{n'_k\}$, 使 $\mu\left(A_{n'_k}\right)\leqslant\frac{1}{2^k}$. 令 $B_k=\bigcup\limits_{i=k}^\infty A_{n'_i}$, 则 $\mu(B_k)\leqslant\frac{1}{2^{k-1}}$, 从而 $B_k\downarrow B=\bigcap\limits_{k=1}^\infty B_k$ 是零测集. 由此知 $\mathbf{1}_{A_{n'_k}}f\leqslant|\mathbf{1}_{B_k}f|\to 0,\text{a.e.}$ 由控制收敛定理知 $\int_{A_{n'_k}} f\,\mathrm{d}\mu\to 0$. 这与 $n_k\uparrow\infty,\int_{A_{n_k}} f\,\mathrm{d}\mu\to\epsilon\neq 0$ 矛盾. □

在以上收敛定理中, 序列 $n\uparrow\infty$ 可换成任何连续参数 $t\to t_0$. 从而由控制收敛定理得到下面的推论.

推论 3.16 (导数与积分号交换). 设 $T\subset\mathbb{R}$ 是开集. $\forall t\in T, f_t$ 可积. $\forall\omega\in\Omega, f_t(\omega)$ 在 t_0 点可导, 则 $\frac{\mathrm{d}}{\mathrm{d}t}f_t(\omega)|_{t_0}$ 是可测函数. 如果存在可积函数 g 及 $\varepsilon>0$ 使当 $|t-t_0|<\varepsilon$ 时 $\left|\frac{f_t-f_{t_0}}{t-t_0}\right|\leqslant g$, 则 $\left(\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega f_t\,\mathrm{d}\mu\right)|_{t_0}=\int_\Omega\frac{\mathrm{d}f_t}{\mathrm{d}t}|_{t_0}\,\mathrm{d}\mu$.

推论 3.17. 设 $\{f_t\}_{t\in(a,b)}$ 为一族实可积函数, $\frac{\mathrm{d}f_t}{\mathrm{d}t}$ 存在. 如存在可积函数 g 使得 $\left|\frac{\mathrm{d}f_t}{\mathrm{d}t}\right|\leqslant g$, 则在 (a,b) 上有 $\frac{\mathrm{d}}{\mathrm{d}t}\int f_t\,\mathrm{d}\mu=\int\frac{\mathrm{d}}{\mathrm{d}t}f_t\,\mathrm{d}\mu$.

证明 使用微分中值定理, $\forall t_0\in(a\,,b)$ 有 $\left|\frac{f_t-f_{t_0}}{t-t_0}\right|\leqslant g, t\in(a,b)$. 由推论 3.16 立得. □

推论 3.18 (交换积分次序). (1) 设 $\{f_t\}_{t\in(a,b)}$ 为一族实可积函数满足 $\forall\omega\in\Omega, f_t(\omega)$ 对 t 连续, 存在可积函数 g 使得 $\forall t\in(a,b)$ 有 $|f_t|\leqslant g$. 则

$$\int_a^b\left(\int_\Omega f_t\,\mathrm{d}\mu\right)\,\mathrm{d}t=\int_\Omega\left(\int_a^b f_t\,\mathrm{d}t\right)\,\mathrm{d}\mu.$$

(2) 如在任何有限区间上, 上式成立, 且 $\int_{-\infty}^\infty|f_t|\,\mathrm{d}t\leqslant h$, h 可积, 则 $\int_{-\infty}^\infty\left(\int_\Omega f_t\,\mathrm{d}\mu\right)\,\mathrm{d}t=\int_\Omega\left(\int_{-\infty}^\infty f_t\,\mathrm{d}t\right)\,\mathrm{d}\mu$.

证明 (1) 设 $a=t_0<t_1<\cdots<t_n=b$ 为 $[a,b]$ 的任一分割, 则

$$\int_a^b f_t\,\mathrm{d}t=\lim_{n\to\infty}\sum_{i=1}^n (t_i-t_{i-1})f_{t_i}.$$

由于 $\sum\limits_{i=1}^n (t_i-t_{i-1})f_{t_i}\leqslant (b-a)g$, 由控制收敛定理易知 $\int_\Omega f_t\,\mathrm{d}\mu$ 关于 t 连续. 故由控制收敛定理及积分的线性性得

$$\int_\Omega\left(\int_a^b f_t\,\mathrm{d}t\right)\mathrm{d}\mu=\lim_{n\to\infty}\sum_{i=1}^n (t_i-t_{i-1})\int_\Omega f_{t_i}\,\mathrm{d}\mu=\int_a^b\left(\int_\Omega f_t\,\mathrm{d}\mu\right)\mathrm{d}t.$$

(2) 由于 $\int_{-\infty}^\infty |f_t|\,\mathrm{d}t\leqslant h$, 则 $g_n=\int_{-n}^n f_t\,\mathrm{d}t$ 满足 $g_n\to\int_{-\infty}^\infty f_t\,\mathrm{d}t$ 且 $|g_n|\leqslant h$, 由控制收敛定理立得. □

推论 3.19 (交换求和号). 设 $\{f_{nm}\}_{n,m\geqslant 1}$ 为一族实数, 满足$f_{nm}\geqslant 0$ 或存在数列 $\{g_n\}$ 使 $\sum\limits_{m=1}^\infty |f_{nm}|\leqslant g_n$ $(\forall n)$ 且 $\sum\limits_{n=1}^\infty g_n<\infty$, 则 $\sum\limits_{m=1}^\infty\sum\limits_{n=1}^\infty f_{nm}=\sum\limits_{n=1}^\infty\sum\limits_{m=1}^\infty f_{nm}$.

证明 设 $\Omega=\mathbb{N},\mu$ 是计数测度, $g(n)=g_n$, 则 g 可积. 令 $f_m(n)=f_{nm}$, 则 $\sum\limits_{m=1}^\infty |f_m|\leqslant g$. 由单调收敛或控制收敛定理知

$$\sum_{m=1}^\infty\sum_{n=1}^\infty f_{nm}=\sum_{m=1}^\infty\int_\Omega f_m\,\mathrm{d}\mu=\int_\Omega\sum_{m=1}^\infty f_m\,\mathrm{d}\mu=\sum_{n=1}^\infty\sum_{m=1}^\infty f_{nm}.\quad\square$$

推论 3.20. 设 $\{f_{nm}\}_{n,m\geqslant 1}$ 为一族实数, 使得 $0\leqslant f_{nm}\uparrow f_n$ $(m\uparrow\infty)$ 或存在数列 $\{g_n\}_{n\geqslant 1}$ 使 $|f_{nm}|\leqslant g_n,\sum\limits_{n=1}^\infty g_n<\infty$, 且 $\lim\limits_{m\to\infty} f_{nm}=f_n$. 则 $\lim\limits_{m\to\infty}\sum\limits_{n=1}^\infty f_{nm}=\sum\limits_{n=1}^\infty f_n$.

§3.3　数学期望

§3.3.1　数字特征

定义 3.21. 设 ξ 是概率空间 $(\Omega,\mathscr{A},\mathbb{P})$ 上的一个随机变量, 如果 ξ 在 Ω 上关于 $\mathbb{P}$ 的积分存在, 则将其定义为 ξ 的数学期望, 记为 $\mathbb{E}\xi=\int_\Omega\xi\,\mathrm{d}\mathbb{P}$. 如 $\mathbb{E}|\xi|<\infty$, 则称期望有限.

由定义知, 数学期望具有积分的所有性质.

定理 3.22 (乘法定理). 如果概率空间 $(\Omega,\mathscr{A},\mathbb{P})$ 上的随机变量 $\xi_1,\xi_2,\cdots,\xi_n$ 是独立的, 它们全部非负或全部具有有限数学期望, 则 $\mathbb{E}(\xi_1\cdots\xi_n)=\mathbb{E}\xi_1\cdots\mathbb{E}\xi_n$.

证明 仅证 $n=2$ 情形, 对于一般情形用归纳法即可. 设 ξ,η 独立.

(1) 设 ξ,η 为非负简单函数, $\xi=\sum\limits_{i=1}^{n}a_i\mathbf{1}_{A_i}\,(a_i\neq a_j,i\neq j)$, $\eta=\sum\limits_{i=1}^{m}b_i\mathbf{1}_{B_i}\,(b_i\neq b_j,i\neq j)$. 则 $\forall i,j$, $\mathbb{P}(A_i\cap B_j)=\mathbb{P}(A_i)\mathbb{P}(B_j)$ 且

$$\xi\eta=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ib_j\mathbf{1}_{A_i\cap B_j},\ \mathbb{E}\xi\eta=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ib_j\mathbb{P}(A_i)\mathbb{P}(B_j)=\mathbb{E}\xi\mathbb{E}\eta.$$

(2) 对于非负的 ξ,η, 由 (1) 并使用非负可测函数的构造与单调收敛定理即可.

(3) 设 ξ,η 具有有限数学期望. 由于 (ξ^+,ξ^-) 与 (η^+,η^-) 独立且 $\xi=\xi^+-\xi^-,\eta=\eta^+-\eta^-$, 由 (2) 与数学期望的线性性质即得 $\mathbb{E}(\xi\eta)=\mathbb{E}\xi\mathbb{E}\eta$. □

定义 3.23. **1)** 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是 n 维随机变量, 则称 n 元函数

$$(t_1,\cdots,t_n)\mapsto\varphi_\xi(t_1,\cdots,t_n)\triangleq\mathbb{E}\mathrm{e}^{\mathrm{i}\sum\limits_{j=1}^{n}t_j\xi_j}$$

为 ξ 的特征函数.

2) 设 ξ 为随机变量, 期望 $\mathbb{E}\xi$ 存在. 称 $D\xi\triangleq\mathbb{E}|\xi-\mathbb{E}\xi|^2$ 为 ξ 的方差.

3) 设 ξ 为随机变量, $r>0$. 称 $\mathbb{E}|\xi|^r$ 为 ξ 的 r 阶矩, $\mathbb{E}|\xi-\mathbb{E}\xi|^r$ 为 ξ 的 r 阶中心矩.

4) 设 ξ,η 为两个随机变量使得 $\mathbb{E}\xi$ 与 $\mathbb{E}\eta$ 存在. 称 $b_{\xi,\eta}=\mathbb{E}(\xi-\mathbb{E}\xi)(\overline{\eta-\mathbb{E}\eta})$ 为 ξ 与 η 的相关矩. 若 $D\xi D\eta\neq0$ 且有限, 称 $r_{\xi,\eta}=\frac{b_{\xi,\eta}}{\sqrt{D\xi D\eta}}$ 为 ξ,η 的相关系数.

5) 设 $\xi=(\xi_1,\cdots,\xi_n)$. 令 $\mathbb{E}\xi=(\mathbb{E}\xi_1,\ldots,\mathbb{E}\xi_n),b_{ij}=b_{\xi_i,\xi_j}$. 称

$$\mathbf{B}(\xi)=\begin{pmatrix}b_{11}&\ldots&b_{1n}\\ \ldots&\ldots&\ldots\\ b_{n1}&\ldots&b_{nn}\end{pmatrix}$$

为 ξ 的相关方阵. 将 $\mathbf{B}(\xi)$ 的秩记成 $r(\mathbf{B}(\xi))$ 或 $r(\xi)$.

命题 3.24. 1) 随机变量 $\xi_1,\cdots,\xi_n$ 相互独立当且仅当

$$\varphi_{\xi_1,\cdots,\xi_n}(t_1,\cdots,t_n)=\varphi_{\xi_1}(t_1)\cdots\varphi_{\xi_n}(t_n),\ t_1,\cdots,t_n\in\mathbb{R}.$$

2) 如 $\xi_1,\cdots,\xi_n$ 独立且方差存在, 则 $D(\xi_1+\cdots+\xi_n)=D\xi_1+\cdots+D\xi_n$.

3) 如 ξ,η 独立且期望存在, 则 $b_{\xi,\eta}=0$.

4) 设 ξ 为随机向量使得 $\mathbf{B}(\xi)$ 有定义, 则 $\mathbf{B}(\xi)\geqslant 0$ (非负定).

5) 如果 $\mathbb{E}|\xi|^r<\infty$, 则 $\forall 0<s<r$ 有 $\mathbb{E}|\xi|^s<\infty$.

证明 只证明 1), 4), 5), 余者显然.

1) 只需证明充分性. 构造独立的 $\widetilde{\xi_1},\cdots,\widetilde{\xi_n}$ 使得其特征函数分别为 $\varphi_{\xi_1}(t_1),\cdots,\varphi_{\xi_n}(t_n)$. 则 $(\widetilde{\xi_1},\cdots,\widetilde{\xi_n})$ 与 $(\xi_1,\cdots,\xi_n)$ 具有相同的特征函数, 从而由逆转公式 (定理 6.5) 知它们同分布. 由此及 $\widetilde{\xi_1},\cdots,\widetilde{\xi_n}$ 相互独立知 $\xi_1,\cdots,\xi_n$ 也相互独立.

4) $\forall t_1,\cdots,t_n$, 有

$$\sum_{i,j=1}^n b_{ij}t_i\overline{t_j}=\mathbb{E}\left|\sum_{i=1}^n t_i(\xi_i-\mathbb{E}\xi_i)\right|^2\geqslant 0.$$

5) 只需注意当 $\forall 0<s<r$ 时 $|\xi|^s\leqslant 1+|\xi|^r$. $\square$

§ 3.3.2 L-S 积分表示

前面介绍的期望定义依赖于概率空间 $(\Omega,\mathscr{A},\mathbb{P})$. 而 $\mathbb{E}\xi$ 是 ξ 的分布性质, 仅依赖于 $\mathbb{P}_\xi$, 它是 $\mathbb{R}$ 上的概率测度. 因此, 为了便于计算, 我们将使用关于 $\mathbb{P}_\xi$ 的积分来定义期望, 并称之为期望的 L-S 积分表示. 一般地, 设 $f:(\Omega,\mathscr{A})\to(E,\mathscr{E})$ 是可测映射, μ 是 $(\Omega,\mathscr{A})$ 上的测度, 则定义 $\mu_f(B)\triangleq\mu(f^{-1}(B))\ (B\in\mathscr{E})$, 它是 $\mathscr{E}$ 上的测度，称为 μ 在 $(E,\mathscr{E})$ 上由 f 诱导的测度. ξ 的概率分布 $\mathbb{P}_\xi$ 是 $\mathbb{P}$ 在 $(\mathbb{R},\mathscr{B})$ 上的由 ξ 诱导的概率测度.

定理 3.25 (积分变换定理). 设 $f:(\Omega,\mathscr{A})\longrightarrow(E,\mathscr{E})$ 可测, g 是 $(E,\mathscr{E})$ 上的可测函数且关于 μ_f 积分存在, μ 是 $(\Omega,\mathscr{A})$ 上的测度. 则 $g\circ f$ 关于 μ 积分存在且 $\forall B\in\mathscr{E}$, 有 $\int_{f^{-1}(B)}g\circ f\,\mathrm{d}\mu=\int_B g\,\mathrm{d}\mu_f$.

证明 (1) 设 g 是示性函数, $g=\mathbf{1}_{B'},B'\in\mathscr{E}$. 则

$$\begin{aligned}\int_B g\,\mathrm{d}\mu_f&=\mu_f(B\cap B')=\mu(f^{-1}(B\cap B'))\\&=\int_{f^{-1}(B)}\mathbf{1}_{f^{-1}(B')}\,\mathrm{d}\mu=\int_{f^{-1}(B)}\mathbf{1}_{B'}\circ f\,\mathrm{d}\mu.\end{aligned}$$

(2) 由积分的线性性质知结论对简单函数成立, 再由单调收敛定理知结论对非负函数成立. 最后由于一般函数可表为正、负部之差, 从而知结论成立. □

定义 3.26. 设 μ 是 $(\mathbb{R}^n, \mathscr{B}^n)$ 上的 L-S 测度, 分布函数为 F. 设 f 是 $(\mathbb{R}^n, \mathscr{B}^n)$ 上的可测函数. 称 f 关于 μ 的积分为一个 L-S 积分, 记成 $\int_{\mathbb{R}^n} f \,\mathrm{d}\mu = \int_{\mathbb{R}^n} f \,\mathrm{d}F$.

定理 3.27. 设 $\xi = (\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 的 n 维实随机变量, 分布函数为 F, 则 $\forall G \in \mathscr{B}^n$, 有 $\mathbb{P}(\xi \in G) = \int_G \mathrm{d}F \triangleq \int_{\mathbb{R}^n} \mathbf{1}_G \,\mathrm{d}F$.

证明 由定义 3.26 立得. □

定理 3.28. 设 ξ 与 F 如定理 3.27. 令 $g_k\,(k = 1, \cdots, m)$ 是 $\mathbb{R}^n$ 上的有限实可测函数, 令 $\eta_k = g_k(\xi_1, \cdots, \xi_n)$. 则 $(\eta_1, \cdots, \eta_m)$ 的分布函数是

$$F_{\eta_1, \cdots, \eta_m}(y_1, \cdots, y_m) = \int_{\{x:\, g_k(x) < y_k, \forall k\}} \mathrm{d}F.$$

证明 由定理 3.27 知, 左式 $= \mathbb{P}(g_k(\xi) < y_k, \forall k) = \int_{\{x: g_k(x) < y_k, \forall k\}} \mathrm{d}F$. □

定理 3.29. 设 ξ 与 η 如上定理, g 是 $\mathbb{R}^n$ 上有限实函数使得 $g(\xi_1, \cdots, \xi_n)$ 的期望存在, 则 $\mathbb{E}g(\xi_1, \cdots, \xi_n) = \int_{\mathbb{R}^n} g \,\mathrm{d}F$.

证明 由积分变换定理, $\int_{\mathbb{R}^n} g \,\mathrm{d}F = \int_{\mathbb{R}^n} g \,\mathrm{d}\mathbb{P}_\xi = \int_{\xi^{-1}(\mathbb{R}^n)} g \circ \xi \,\mathrm{d}\mathbb{P} = \mathbb{E} g \circ \xi$. □

推论 3.30. 设 ξ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上实值随机变量, 分布函数为 F, 则 $\mathbb{E}\xi = \int_{\mathbb{R}} x \,\mathrm{d}F(x)$.

命题 3.31. 设 μ 为 $(\Omega, \mathscr{A})$ 上测度, $\rho \geqslant 0$ 可测. 如果可测函数 f 关于 $\nu(A) \triangleq \int_A \rho \,\mathrm{d}\mu$ 的积分存在, 则 ρf 关于 μ 的积分存在, 且 $\int_\Omega \rho f \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\nu$. 由此, 我们可记 $\mathrm{d}\nu = \rho \,\mathrm{d}\mu$.

例 3.32. 设 ξ 为期望存在的离散型随机变量, $\mathbb{P}(\xi = a_i) = p_i, \sum\limits_{i=1}^{\infty} p_i = 1$. 它在 $\mathbb{R}$ 上的分布律为 $\mathbb{P}_\xi(\{a_i\}) = p_i, \mathbb{P}_\xi(A) = \sum\limits_{a_i \in A} p_i$. 则 $\mathbb{E}\xi = \int_{\mathbb{R}} x \,\mathrm{d}F = \int_{\mathbb{R}} x \,\mathrm{d}\mathbb{P}_\xi = \sum\limits_{i=1}^{\infty} a_i p_i$.

例 3.33. 设 ξ 为期望存在的连续型随机变量, 分布密度为 $\rho(x)$. 则 $\mathbb{P}_\xi(A) = \mathbb{P}(\xi \in A) = \int_A \rho(x) \,\mathrm{d}x$, 且 $\mathbb{E}\xi = \int_{\mathbb{R}} x \,\mathrm{d}F = \int_{\mathbb{R}} x \,\mathrm{d}\mathbb{P}_\xi = \int_{\mathbb{R}} x\rho(x) \,\mathrm{d}x$.

上面两个例子表明, 我们所定义的随机变量的数学期望是对初等概率论所介绍的离散型与连续型随机变量数学期望的推广.

§3.4　r 次平均与 L^r 空间

定义 3.34. 设 $(\Omega,\mathscr{A},\mu)$ 为测度空间, $r>0$. 则称

$$L^r(\mu)=\{\xi:\xi \text{ 为 } \Omega \text{ 上可测函数}, \mu(|\xi|^r)<\infty\}$$

为 μ 的 L^r 空间. 如 $f_n,f\in L^r(\mu)$ 且 $\mu(|f_n-f|^r)\to 0\,(n\to\infty)$, 则称 f_n 关于 μ 以 r 次平均收敛于 f, 记成 $f_n\xrightarrow{L^r(\mu)}f$.

记 $||f||_r=\mu(|f|^r)^{1/r\wedge 1},r>0$. 我们将证明 $(L^r(\mu),||\cdot||_r)$ 是 Banach 空间, $L^2(\mu)$ 在内积 $\langle f,g\rangle\triangleq\mu(fg)$ 下是 Hilbert 空间. 我们还将讨论 r 次平均收敛与可测函数列其它收敛性之间的关系. 为此, 先介绍一些经典的不等式.

§ 3.4.1　几个重要不等式

命题 3.35. 若 $a\geqslant 0,b\geqslant 0,0<\alpha<1,\alpha+\beta=1$, 则 $a^\alpha b^\beta\leqslant a\alpha+b\beta$ 且等号成立当且仅当 $a=b$.

证明 由于 log 是凹函数, 故 $\log(a\alpha+b\beta)\geqslant\alpha\log a+\beta\log b=\log(a^\alpha b^\beta)$, 且易知等号成立当且仅当 $a=b$.　□

命题 3.36 (Hölder 不等式). 设 $r>1,\frac{1}{r}+\frac{1}{s}=1$. 则

$$\mu(|f\,g|)\leqslant(\mu(|f|^r))^{\frac{1}{r}}(\mu(|g|^s))^{\frac{1}{s}}.$$

当右端有限时等号成立当且仅当 $\exists c_1,c_2$ 不全为零, 使得 $c_1|f|^r+c_2|g|^s=0,\mu$-a.e.

证明 当 $f=0$ 或 $g=0$ 或右方无穷时不等式显然成立. 故不妨设 $0<\mu(|f|^r),\mu(|g|^s)<\infty$. 令

$$a=\frac{|f|^r}{\mu(|f|^r)},\ b=\frac{|g|^s}{\mu(|g|^s)},\ \alpha=\frac{1}{r},\ \beta=\frac{1}{s},$$

由命题 3.35 可得

$$\frac{|f\,g|}{||f||_r||g||_s} \leqslant \frac{|f|^r}{r\mu(|f|^r)} + \frac{|g|^s}{s\mu(|g|^s)}.$$

两边同时对 μ 取积分即得所需不等式, 等式成立当且仅当上面不等式中的等式 μ-a.e. 成立, 由命题 3.35 即 $\frac{|f|^r}{\mu(|f|^r)} = \frac{|g|^s}{\mu(|g|^s)}$, a.e. 故取 $c_1 = \frac{1}{\mu(|f|^r)}, c_2 = -\frac{1}{\mu(|g|^s)}$ 即可. □

推论 3.37 (Jessen 不等式). $\forall r > 1, \mathbb{E}|\xi| \leqslant (\mathbb{E}|\xi|^r)^{\frac{1}{r}}$, 且等号成立当且仅当 $|\xi|^r$ a.s. 为常数.

命题 3.38 (非随机的 C_r 不等式). $\forall a_1, \cdots, a_n \in \mathbb{R}$, 有 $|a_1 + \cdots + a_n|^r \leqslant n^{(r-1)^+}(|a_1|^r + \cdots + |a_n|^r)$. 当 $r > 1$ 时等号成立当且仅当 $a_1 = \cdots = a_n$; 当 $r = 1$ 时等号成立当且仅当 a_i 同号; 当 $r < 1$ 时等号成立当且仅当 a_i 中至多有一个不是零.

证明 (1) $r > 1$ 情形. 设 ξ 满足 $\mathbb{P}(\xi = a_i) = \frac{1}{n}, 1 \leqslant i \leqslant n$. 则 $\mathbb{E}|\xi| = \frac{1}{n}\sum\limits_{i=1}^n |a_i|, \mathbb{E}|\xi|^r = \frac{1}{n}\sum\limits_{i=1}^n |a_i|^r$. 由 Jessen 不等式, $n^{-r}\left(\sum\limits_{i=1}^n |a_i|\right)^r \leqslant \frac{1}{n}\sum\limits_{i=1}^n |a_i|^r$, 故所证不等式成立, 且等式成立当且仅当 $|\xi|$ 为常数, 即 $|a_i| = |a_j|, \forall i, j$. 而 $\left|\sum\limits_{i=1}^n a_i\right| = \sum\limits_{i=1}^n |a_i|$ 当且仅当 a_i 同号, 从而 $a_i = a_j, \forall i, j$.

(2) $r \leqslant 1$ 情形. 只需证 a_i 不完全为零情形. 注意

$$\frac{|a_k|}{\sum\limits_{i=1}^n |a_i|} \leqslant \frac{|a_k|^r}{\left(\sum\limits_{i=1}^n |a_i|\right)^r}, \quad r \leqslant 1.$$

两边对 k 求和即得所需不等式. 当 $r = 1$ 时, 等号成立当且仅当 a_i 同号. 而当 $r \leqslant 1$ 时, 等号成立当且仅当 $\forall k, |a_k|/\sum\limits_{i=1}^n |a_i| = 1$ 或 0, 即 a_i 只有一个非零. □

命题 3.39 (C_r **不等式**). 设 $f_1, \cdots, f_n$ 是可测函数, 则 $\mu(|f_1 + \cdots + f_n|^r) \leqslant n^{(r-1)^+}\sum\limits_{i=1}^n \mu(|f_i|^r)$, 且等号成立当且仅当

1) $r > 1$ 时, $\forall i \neq j, f_i = f_j$, a.e.; 或

2) $r < 1$ 时, $\mu(|f_i|)$ 中至多一个非零; 或

3) $r = 1$ 时, f_i a.e. 同号.

命题 3.40 (Minkowski 不等式). 设 $r \geqslant 1, f, g \in L^r(\mu)$. 则 $(\mathbb{E}|f+g|^r)^{\frac{1}{r}} \leqslant (\mathbb{E}|f|^r)^{\frac{1}{r}} + (\mathbb{E}|g|^r)^{\frac{1}{r}}$, 且等号成立当且仅当

1) $r > 1$ 时, 存在不全为零且同号的 c_1, c_2 使 $c_1 f - c_2 g = 0$, a.e.; 或

2) $r = 1$ 时, f, g a.e. 同号.

证明 仅证 $r > 1$. 由 Hölder 不等式,

$$\begin{aligned}\mu(|f+g|^r) &\leqslant \mu(|f||f+g|^{r-1}) + \mu(|g||f+g|^{r-1})\\ &\leqslant \|f\|_r(\mu(|f+g|^r))^{\frac{r-1}{r}} + \|g\|_r(\mu(|f+g|^r))^{\frac{r-1}{r}}.\end{aligned}$$

等号成立当且仅当存在不全为零的 c_1, c_2 和不全为零的 c_3, c_4 使 $|f|^r c_1 + |f+g|^r c_2 = 0, c_3|g|^r + c_4|f+g|^r = 0$ 且 f, g 同号. 由此得所需结论. □

§ 3.4.2 L^r 空间

为使 $\|\cdot\|_r$ 是范数, 我们把 $L^r(\mu)$ 中两个 μ-a.e. 相等的函数视为相同. 即 $L^r(\mu)$ 中每个元素是一个 μ-a.e. 相等下的等价类.

定理 3.41. 设 $r > 0$. 则 $(L^r(\mu), \|\cdot\|_r)$ 是一个线性赋范空间.

证明 显然, $\|f\|_r = 0$ 当且仅当 $f = 0, \mu$-a.e. 故 $\forall f \in L^r(\mu), \|f\|_r = 0$ 当且仅当 $f = 0$. 易见 $L^r(\mu)$ 是线性空间且由 C_r 不等式 (当 $r < 1$) 及 Minkowski 不等式 (当 $r \geqslant 1$) 易知 $\|\cdot\|_r$ 满足三角不等式. 从而 $(L^r(\mu), \|\cdot\|_r)$ 是一个线性赋范空间. □

定理 3.42. 设 $\{f_n\} \subset L^r(\mu)$, 则 $\{f_n\}$ 在 $\|\cdot\|_r$ 下收敛于某函数 $f \in L^r(\mu)$ 当且仅当它在 $\|\cdot\|_r$ 下为 Cauchy 列. 因而线性赋范空间 $(L^r(\mu), \|\cdot\|_r)$ 是完备线性赋范 (Banach) 空间.

证明 由三角不等式, 只需证充分性. 设 $\{f_n\}$ 在 $L^r(\mu)$ 中为 Cauchy 列. 由辛钦不等式知当 $n, m \to \infty$ 时有

$$\mu(|f_n - f_m| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^r}\mu(|f_n - f_m|^r) \to 0.$$

从而 $\{f_n\}$ 相互依测度收敛, 故存在子列 $n_k \uparrow \infty$ 及某 f 使 $f_{n_k} \xrightarrow{\text{a.e.}} f$. 由此知 $\forall m \geqslant 1$, 有 $f_m - f_{n_k} \xrightarrow{\text{a.e.}} f_m - f$ $(n_k \to \infty)$. 则由 Fatou 引理 $\mu(|f_m - f|^r) = \mu\left(\varliminf_{n_k\to\infty} |f_m - f_{n_k}|^r\right) \leqslant \varliminf_{n_k\to\infty} \mu(|f_m - f_{n_k}|^r)$. 由于 $\{f_n\}$ 在 $L^r(\mu)$ 中为 Cauchy 列, 由此令 $m \to \infty$ 得 $\lim_{m\to\infty} \mu(|f_m - f|^r) = 0$. □

命题 3.43. (1) 设 μ 有限. 若 $f_n \xrightarrow{L^r(\mu)} f$, 则 $\forall r' \in (0,r), f_n \xrightarrow{L^{r'}(\mu)} f$.

(2) $f_n \xrightarrow{L^r(\mu)} f$, 则 $\mu(|f_n|^r) \to \mu(|f|^r)$.

证明 (1) 和 (2) 可分别由 Hölder 不等式和 $||\cdot||_r$ 的三角不等式得到. □

§ 3.4.3 与各种收敛性之间的关系

定义 3.44. 设 $(\Omega, \mathscr{A}, \mu)$ 为测度空间, $\{f_t, t \in T\}$ 是一族实可测函数.

1) 若 $\lim\limits_{\mu(A)\to 0} \sup\limits_{t\in T} \mu(|f_t|\mathbf{1}_A) = 0$, 则称 $\{f_t, t \in T\}$ 积分一致连续.

2) 若 $\lim\limits_{n\to\infty} \sup\limits_{t\in T} \mu(|f_t|\mathbf{1}_{\{|f_t|\geqslant n\}}) = 0$, 则称 $\{f_t, t \in T\}$ 一致可积.

3) 若 $\sup\limits_{t\in T} \mu(|f_t|) < \infty$, 则称 $\{f_t, t \in T\}$ 积分一致有界.

留意, 当 μ 为无穷测度时, 一致可积性未必导出可积性.

定理 3.45. 设 μ 为有限测度, $\{f_n\}_{n\geqslant 1} \subset L^r(\mu)$, 则以下几条等价:

1) $f_n \xrightarrow{L^r(\mu)} f$;

2) $f_n \xrightarrow{\mu} f$ 且 $\{|f_n - f|^r\}_{n\geqslant 1}$ 积分一致连续;

3) $f_n \xrightarrow{\mu} f$ 且 $\{|f_n|^r\}_{n\geqslant 1}$ 积分一致连续;

4) $f_n \xrightarrow{\mu} f$ 且 $\{|f_n|^r\}_{n\geqslant 1}$ 一致可积.

证明 由于 $|||\mathbf{1}_A f_n||_r - ||\mathbf{1}_A(f_n - f)||_r| \leqslant ||\mathbf{1}_A f||_r$, $\{|f_n - f|^r\}_{n\geqslant 1}$ 的积分一致连续性等价于 $\{|f_n|^r\}_{n\geqslant 1}$ 的积分一致连续性, 并进一步等价于 $\{|f_n|^r\}_{n\geqslant 1}$ 的一致可积性 (本章习题 28). 故 2), 3), 4) 等价. 仅需证 1) $\Leftrightarrow$ 2).

1) $\Rightarrow$ 2) 由于 $\mu(|f_n - f| \geqslant \varepsilon) \leqslant \varepsilon^{-r}\mu(|f_n - f|^r)$, 1) 蕴含 $\{f_n\}_{n\geqslant 1}$ 依测度收敛于 f. 为证 $\{|f_n - f|^r\}_{n\geqslant 1}$ 的积分一致连续性, 对 $\forall \varepsilon > 0$, 取 $n_\varepsilon \geqslant 1$ 使 $\forall n \geqslant n_\varepsilon$ 有 $\mu(|f_n - f|^r) < \varepsilon$. 则有

$$\sup_{n\geqslant 1} \int_A |f_n - f|^r \,\mathrm{d}\mu \leqslant \varepsilon + \sum_{n=1}^{n_\varepsilon} \mu(\mathbf{1}_A |f_n - f|^r).$$

由于给定 n, $\lim\limits_{\mu(A)\to 0} \mu(\mathbf{1}_A|f_n - f|^r) = 0$, 从而

$$\varlimsup_{\mu(A)\to 0} \sup_{n\geqslant 1} \mu(\mathbf{1}_A|f_n - f|^r) \leqslant \varepsilon.$$

由 ε 的任意性, 得 $\{|f_n - f|^r\}_{n\geqslant 1}$ 积分一致连续.

2) $\Rightarrow$ 1)　令 $A_n = \{|f_n - f| \geqslant \varepsilon\}$, 则 $\mu(A_n) \to 0$. 从而由积分一致连续性, $\mu(\mathbf{1}_{A_n}|f_n - f|^r) \leqslant \sup\limits_{m\geqslant 1} \mu(\mathbf{1}_{A_n}|f_m - f|^r) \to 0, n \to \infty$. 故

$$\varlimsup_{n\to\infty} \mu(|f_n - f|^r) \leqslant \varlimsup_{n\to\infty} \mu(|f_n - f|^r \mathbf{1}_{\{|f_n - f|\geqslant\varepsilon\}}) + \varepsilon^r = \varepsilon^r.$$

由 ε 的任意性得 $f_n \xrightarrow{L^r(\mu)} f$.　　□

§3.5　σ 可加集函数的分解

前面介绍过一个积分存在函数的不定积分是一个 σ 可加集函数. 易见一个不定积分 $\varphi = \int_\bullet f\,\mathrm{d}\mu$ 具有如下性质: 如 $\mu(A) = 0$, 则 $\varphi(A) = 0$. 我们称具有这样性质的集函数是关于 μ 绝对连续的. 那么反过来一个绝对连续 σ 可加集函数能否写成不定积分? 这是本节的核心问题. 为此, 我们先将 σ 可加集函数分解为两个测度之差, 再进一步研究分布函数的分解, 从而给出该问题一个正面回答.

§ 3.5.1　σ 可加集函数的分解定理

我们知道任何一个函数可以分解成两个非负函数之差, 那么一个 σ 可加集函数可否分解成两个测度之差呢? 为此, 先考虑不定积分 $\varphi(A) = \int_A f\,\mathrm{d}\mu$. 令 $\varphi^+(A) = \int_A f^+\,\mathrm{d}\mu, \varphi^-(A) = \int_A f^-\,\mathrm{d}\mu$. 则 φ^+ 和 φ^- 为测度, 且 $\varphi = \varphi^+ - \varphi^-$. 而对于一般 σ 可加集函数我们该怎么做呢? 留意, 令 $C = \{f > 0\}, D = \{f \leqslant 0\}$, 上面的 φ^+ 与 φ^- 可分别表成 $\varphi^+(A) = \varphi(A\cap C), \varphi^-(A) = -\varphi(A\cap D)$. 因此, 对于一般的 φ, 如能找到 $D \in \mathscr{A}$ 使 $\varphi(D) = \inf\limits_{A\in\mathscr{A}} \varphi(A)$, 则可令 $\varphi^-(A) = -\varphi(D\cap A), \varphi^+(A) = \varphi(D^c\cap A)$. 为此, 先证明下面定理.

定理 3.46. 设 φ 是 $(\Omega, \mathscr{A})$ 上的σ 可加集函数, 则 $\exists D \in \mathscr{A}$ 使 $\varphi(D) = \inf\limits_{A\in\mathscr{A}} \varphi(A)$.

证明 取 $\{A_n\}$ 使 $\varphi(A_n) \downarrow \inf\limits_{A\in\mathscr{A}} \varphi(A)$. 由于 $\inf\limits_{A\in\mathscr{A}} \varphi(A) \leqslant 0$, 可设 $\varphi(A_n)$ 有限. 令$A = \bigcup\limits_{n=1}^{\infty} A_n$. 任给 $k \geqslant 1$, 有 $A = A_k + (A - A_k) =: A_{k,1} + A_{k,2}$. $\forall n \geqslant 2$,

有

$$A = A_{n,2} + \sum_{i_1,i_2,\ldots,i_{n-1}=1}^{2} A_{1,i_1} \cap A_{2,i_2} \cap \ldots \cap A_{n-1,i_{n-1}} \cap A_{n,1}.$$

随着 n 增大, 这样对 A 的划分越来越细. 对每次划分, 我们仅取出 φ 值为负的小集合. 直观上, 当划分越来越细时, 所取出的小集合之并越来越接近于所求的集合 D. 基于此, 对每个 $n \geqslant 1$, 令

$$B_n = \sum_{\substack{1\leqslant i_1,i_2,\ldots,i_n\leqslant 2 \\ \varphi(A_{1,i_1}\cap A_{2,i_2}\cap\cdots\cap A_{n,i_n})\leqslant 0}} A_{1,i_1} \cap A_{2,i_2} \cap \cdots \cap A_{n,i_n} =: \sum_{i=1}^{k_n} A'_{n,i}.$$

由 φ 的 σ 可加性及 B_n 的定义知 $\varphi(B_n) \leqslant \varphi(A_n)$. 再令

$$D \triangleq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} B_k = \lim_{n\to\infty} \bigcup_{k=n}^{\infty} B_k.$$

由于第 $n+1$ 次划分细于第 n 次划分, 因此 B_{n+1} 中包含的小集合 $A'_{n+1,i}$ 要么包含于 B_n, 要么与 B_n^c 不交. 任给 $m > n$, 有

$$B_n \cup \cdots \cup B_m = B_n + \sum_{A'_{n+1,i}\cap B_n=\emptyset} A'_{n+1,i} + \sum_{A'_{n+2,i}\cap(B_n\cup B_{n+1})=\emptyset} A'_{n+2,i} + \cdots + \sum_{A'_{m,i}\cap(B_n\cup\cdots\cup B_{m-1})=\emptyset} A'_{m,i}.$$

因此由 φ 的 σ 可加性及 $\varphi(A'_{i,j}) \leqslant 0$ 得 $\varphi(B_n \cup \cdots \cup B_m) \leqslant \varphi(B_n) \leqslant \varphi(A_n)$. 令 $m \uparrow \infty$, 由符号测度的下连续性 (留意 $\varphi(A_n)$ 有限) 得 $-\infty < \varphi\left(\bigcup\limits_{k=n}^{\infty} B_k\right) \leqslant \varphi(A_n)$. 最后由符号测度的上连续性得

$$\varphi(D) = \lim_{n\to\infty} \varphi\left(\bigcup_{k=n}^{\infty} B_k\right) \leqslant \lim_{n\to\infty} \varphi(A_n) = \inf_{A\in\mathscr{A}} \varphi(A). \quad \square$$

推论 3.47. 如果 φ 是 $\mathscr{A}$ 上 σ 可加集函数, 则存在 $D \in \mathscr{A}$, 使 $\forall A \in \mathscr{A}$ 有 $\varphi(A\cap D) = \inf\limits_{B\in A\cap\mathscr{A}} \varphi(B), \varphi(A\cap D^c) = \sup\limits_{B\in A\cap\mathscr{A}} \varphi(B)$.

证明 设 $D \in \mathscr{A}$ 使 $\varphi(D) = \inf\limits_{A\in\mathscr{A}} \varphi(A), \varphi(D^c) = \sup\limits_{A\in\mathscr{A}} \varphi(A)$. 则 $\forall A \in \mathscr{A}$ 及 $B \in A\cap\mathscr{A}, \varphi(A\cap D)+\varphi(D-A) = \varphi(D) \leqslant \varphi(B\cup(D-A)) = \varphi(B)+\varphi(D-A)$.

由于 $\varphi(D) \leqslant 0$, 则 $\varphi(A\cap D), \varphi(D-A)$ 均有限, 从而 $\varphi(A\cap D) \leqslant \varphi(B)$. 因此 $\inf\limits_{B\in A\cap\mathscr{A}} \varphi(B) \leqslant \varphi(A\cap D) \leqslant \inf\limits_{B\in A\cap\mathscr{A}} \varphi(B)$, 即 $\varphi(A\cap D) = \inf\limits_{B\in A\cap\mathscr{A}} \varphi(B)$.

另一方面, $\forall B \in A\cap\mathscr{A}, \varphi(A\cap D^c) + \varphi(A\cap D) = \varphi(A) = \varphi(B) + \varphi(B^c\cap A)$. 由于 $\varphi(A\cap D) = \inf\limits_{B\in A\cap\mathscr{A}} \varphi(B)$ 有限, $\varphi(A\cap B^c) \geqslant \inf\limits_{B\in A\cap\mathscr{A}} \varphi(B) = \varphi(A\cap D)$, 则 $\varphi(A\cap D^c) = \varphi(B) + \varphi(B^c\cap A) - \varphi(A\cap D) \geqslant \varphi(B)$. 从而 $\sup\limits_{B\in A\cap\mathscr{A}} \varphi(B) \leqslant \varphi(A\cap D^c) \leqslant \sup\limits_{B\in A\cap\mathscr{A}} \varphi(B)$, 即 $\varphi(A\cap D^c) = \sup\limits_{B\in A\cap\mathscr{A}} \varphi(B)$. □

定理 3.48 (Hahn 分解定理). 设 φ 是 $(\Omega,\mathscr{A})$ 上的 σ 可加集函数, $D\in\mathscr{A}$ 使得 $\varphi(D) = \inf\limits_{A\in\mathscr{A}} \varphi(A)$. 令 $\varphi^+(A) = \varphi(A\cap D^c), \varphi^-(A) = -\varphi(A\cap D)$, 则 φ^+, φ^- 均为 $\mathscr{A}$ 测度, 且 $\varphi = \varphi^+ - \varphi^-$.

易见 φ^+ 与 φ^- 的定义与 D 选取无关. 我们称这两个测度分别为 φ 的上，下变差, 而称 $\overline{\varphi} \triangleq \varphi^+ + \varphi^-$ 为 φ 全变差. 在文献中全变差也记作 $|\varphi|$. 留意, 一般地 $|\varphi(A)| \neq |\varphi|(A)$.

命题 3.49. 设 f 可测且 $\mu(f^-) < \infty$. 令 φ 为 f 的不定积分, 则 $\varphi^+ = \int_\bullet f^+\,\mathrm{d}\mu, \varphi^- = \int_\bullet f^-\,\mathrm{d}\mu$.

证明 令 $D = \{f \leqslant 0\}$, 则 $\varphi(D) = \int_D f\,\mathrm{d}\mu = \inf\limits_{A\in\mathscr{A}} \varphi(A)$. 因此 $-\varphi(D\cap A) = \int_{D\cap A}(-f)\,\mathrm{d}\mu = \int_A f^-\,\mathrm{d}\mu$. □

§ 3.5.2　不定积分与 Lebesgue 分解定理

设 $(\Omega,\mathscr{A},\mu)$ 是一个测度空间, 我们将考察 σ 可加集函数 φ 与 μ 的关系.

定义 3.50. **1)** 如果 $\forall A\in\mathscr{A}$ 使 $\mu(A)=0$, 有 $\varphi(A)=0$, 则称 φ 是关于 μ 绝对连续的, 记成 $\varphi \ll \mu$.

2) 如果存在 $N\in\mathscr{A}$ 使 $\mu(N)=0$, 而 $\varphi(N^c)=0$, 则称 φ 是与 μ 奇异的.

定理 3.51. 设 φ 为 σ有限可加集函数, μ 是 σ 有限测度. 则 $\varphi \ll \mu$ 当且仅当存在可测函数 f 使 $\mu(f^-) < \infty$ 且 $\varphi = \int_\bullet f\,\mathrm{d}\mu$.

充分性显然, 而必要性由下面的更一般结论导出.

定理 3.52 (Lebesgue 分解定理). 设 μ 和 φ 如定理 3.51 . 则 φ 可划分成 $\varphi = \varphi_c + \varphi_s$, 其中 φ_c 是某有限可测函数关于 μ 的不定积分, φ_s 是与 μ 奇异的 σ 可加集函数, 且这样的分解是唯一的.

证明 1) 分解的唯一性

由 σ 有限性, 无妨设 φ 是有限的. 设有两种分解 : $\varphi=\varphi_c+\varphi_s=\varphi_c'+\varphi_s'$. 设 N_1,N_2 为 μ 零测集使 $\varphi_s(N_1^c)=\varphi_s'(N_2^c)=0$. 令 $N=N_1\cup N_2$, 则 $\mu(N)=0$ 且 $\varphi_s(N^c)=\varphi_s'(N^c)=0$. $\forall A\in\mathscr{A}$ 有 $\varphi_c(A\cap N)+\varphi_s(A\cap N)=\varphi_c'(A\cap N)+\varphi_s'(A\cap N)$ 且 $\varphi_s(A\cap N^c)=\varphi_s'(A\cap N^c)=0$. 则 $\varphi_s(A)=\varphi_s(A\cap N)=\varphi_s'(A\cap N)=\varphi_s'(A)$. 类似地, $\varphi_c(A)=\varphi_c'(A)$.

2) 分解的存在性.

i) 设 μ,φ 是有限测度. 令

$$\Phi=\left\{f: f\geqslant 0,\int_A f\,\mathrm{d}\mu\leqslant\varphi(A),\forall A\in\mathscr{A}\right\},\ \alpha=\sup_{f\in\Phi}\mu(f).$$

易见 Φ 不空且 $\alpha\in[0,\varphi(\Omega)]$. 取 $\{f_n\}_{n\geqslant1}\subset\Phi$ 使 $\alpha_n\triangleq\mu(f_n)\uparrow\alpha\leqslant\varphi(\Omega)<\infty$. 令 $g_n=\sup\limits_{k\leqslant n}f_k$, 则 $0\leqslant g_n\uparrow f=\sup\limits_{k\geqslant1}f_k$. 给定 $n\geqslant1$, 令 $A_k=\{\omega: g_n(\omega)=f_k(\omega)\}$. 则 $\bigcup\limits_{k=1}^{n}A_k=\Omega$. 再令 $B_k=A_k-\bigcup\limits_{i=1}^{k-1}A_i$, 则 $\{B_k\}$ 两两不交且 $\bigcup\limits_{k=1}^{n}B_k=\Omega$. 故 $\forall A\in\mathscr{A}$

$$\int_A g_n\,\mathrm{d}\mu=\sum_{k=1}^{n}\int_{A\cap B_k}f_k\,\mathrm{d}\mu\leqslant\sum_{k=1}^{n}\varphi(B_k\cap A)=\varphi(A).$$

从而 $\int_A f\,\mathrm{d}\mu\leqslant\varphi(A)$. 由此及 α 的定义知 $\mu(f)=\alpha$.

令

$$\varphi_c(A)=\int_A f\,\mathrm{d}\mu,\varphi_s(A)\triangleq\varphi(A)-\int_A f\,\mathrm{d}\mu.$$

$\forall n\geqslant1$，令 $\varphi_n=\varphi_s-\frac{\mu}{n}$, 由 Hahn 分解的证明知，存在 $D_n\in\mathscr{A}$ 使

$$\varphi_n(D_n\cap A)\leqslant0,\varphi_n(D_n^c\cap A)\geqslant0,\forall A\in\mathscr{A}.$$

令 $D=\bigcap\limits_{n=1}^{\infty}D_n$. 则 $\forall n$ 有

$$D\subset D_n,\varphi_s(D\cap A)\leqslant\frac{1}{n}\mu(D\cap A).$$

故 $\varphi_s(D\cap A)=0,\forall A\in\mathscr{A}$.

为证 $\mu(D^c)=0$, 只需证 $\forall n$ 有 $\mu(D_n^c)=0$. 事实上,

$$\begin{aligned}\int_A\left(f+\frac{1}{n}\mathbf{1}_{D_n^c}\right)\mathrm{d}\mu&=\varphi_c(A)+\frac{1}{n}\mu(A\cap D_n^c)\\&=\varphi(A)-\varphi_s(A)+\frac{1}{n}\mu(A\cap D_n^c)=\varphi(A)-\varphi_n(A\cap D_n^c)-\varphi_s(A\cap D_n)\\&\leqslant\varphi(A)-\varphi_s(A\cap D_n)\leqslant\varphi(A).\end{aligned}$$

由此可得 $f+\frac{1}{n}\mathbf{1}_{D_n^c}\in\Phi$. 故 $\alpha\geqslant\int_\Omega(f+\frac{1}{n}\mathbf{1}_{D_n^c})\,\mathrm{d}\mu=\int_\Omega f\,\mathrm{d}\mu+\frac{1}{n}\mu(D_n^c)=\alpha$. 从而 $\frac{1}{n}\mu(D_n^c)=0$, 即 $\mu(D_n^c)=0$.

ii) 设 μ 与 φ 为 σ 有限测度. 存在 $\{A_n\}_{n\geqslant1}$ 两两不交使 $\bigcup\limits_{n=1}^{\infty}A_n=\Omega$ 且 $\mu(A_n),\varphi(A_n)<\infty\,(\forall n)$. 由 i) 存在 $\varphi_c^{(n)},\varphi_s^{(n)}$ 使

$$\begin{aligned}\varphi(A_n\cap\bullet)&=\varphi_c^{(n)}(A_n\cap\bullet)+\varphi_s^{(n)}(A_n\cap\bullet),\\\varphi_c^{(n)}(A_n\cap\bullet)&=\int_{A_n\cap\bullet}f^{(n)}d\mu.\end{aligned}$$

设 N_n 为 μ 零测集, 使得 $\forall A\in\mathscr{A},\varphi_s^{(n)}(N_n^c\cap A\cap A_n)=0$. 令

$$f=\sum_{n=1}^{\infty}\mathbf{1}_{A_n}f^{(n)},\ \varphi_c(A)=\int_A f\,\mathrm{d}\mu,\ \varphi_s(A)=\sum_{n=1}^{\infty}\varphi_s^{(n)}(A_n\cap A).$$

再令$N=\bigcup\limits_{n=1}^{\infty}N_n$. 则 $\forall A\in\mathscr{A}$ 有 $\varphi_s^{(n)}(N^c\cap A\cap A_n)\leqslant\varphi_s^{(n)}(N_n^c\cap A\cap A_n)=0$. 由此得 $\varphi_s(N^c\cap A)=\sum\limits_n{\varphi_s}^{(n)}(N^c\cap A\cap A_n)=0$. 故 φ_s 与 μ 奇异.

iii) 一般情形. 由 Hahn 分解定理有 $\varphi=\varphi^+-\varphi^-$. 而由 ii) 知, φ^+ 与 φ^- 有分解 $\varphi^+={\varphi_c}^++{\varphi_s}^+,\varphi^-={\varphi_c}^-+{\varphi_s}^-$. 则 $\varphi=({\varphi_c}^+-{\varphi_c}^-)+({\varphi_s}^+-{\varphi_s}^-)$.
□

作为 Lebesgue 分解定理的直接推论, 下面的 Radon-Nikodym 定理回答了本节的主要问题.

定理 3.53 (Radon-Nikodym 定理). 设 μ 是 $\mathscr{A}$ 上的 σ 有限测度, 若 φ 是 σ 有限并关于 μ 绝对连续的符号测度, 则存在可测函数 f 使 $\mathrm{d}\varphi=f\,\mathrm{d}\mu$ 且 f 由 φ-a.e. 唯一确定.

定理 3.54 (Radon-Nikodym 定理的推广). μ 如上, φ 是关于 μ 绝对连续的 σ 可加集函数. 则存在可测函数 f 使得 $\mathrm{d}\varphi=f\,\mathrm{d}\mu$, 其中 f 由 φ-a.e. 唯一确定.

证明 仅证 μ 是有限测度而 φ 是测度情形. 设

$$\mathscr{B}=\{A\in\mathscr{A}:\varphi\text{ 在 }A\text{ 上 }\sigma\text{ 有限}\},$$

令 $s=\sup\limits_{B\in\mathscr{B}}\mu(B)$, 取 $\{B_n\}_{n\geqslant1}\subset\mathscr{B},\mu(B_n)\uparrow s$. 令 $B=\bigcup\limits_{n=1}^{\infty}B_n$. 则 $B\in\mathscr{B}$, $s=\mu(B)$. 由于 φ 在 $B\cap\mathscr{A}$ 上 σ 有限, 由定理 3.53, 存在 f_1 使 $\varphi(A\cap B)=\int_{A\cap B}f_1\,\mathrm{d}\mu, A\in\mathscr{A}$. 令

$$f(\omega)=\begin{cases}f_1(\omega), & \omega\in B,\\ \infty, & \omega\notin B,\end{cases}$$

则 $\forall A\in\mathscr{A}$ 使 $\mu(A\cap B^c)>0$, 有 $\int_A f\,\mathrm{d}\mu=\infty$. 另一方面, 如 $\mu(A\cap B^c)>0$, 则 $\varphi(A\cap B^c)=\infty$. 如不然, 则 φ 在 $B\cup A$ 上 σ 有限且 $\mu(B\cup A)=\mu(B^c\cap A)+\mu(B)>s$, 与 $s=\sup\limits_{B\in\mathscr{B}}\mu(B)$ 矛盾. 从而 $\forall A\in\mathscr{A}$ 有

$$\int_A f\,\mathrm{d}\mu=\int_{A\cap B}f\,\mathrm{d}\mu+\int_{A\cap B^c}f\,\mathrm{d}\mu=\varphi(A\cap B)+\infty\cdot\mu(A\cap B^c)=\varphi(A).\quad\square$$

定义 3.55. 设 $(\Omega,\mathscr{A},\mu)$ 是 σ 有限测度空间, φ 是关于 μ 绝对连续的 σ 可加集函数, 则存在 μ-a.e. 唯一的 f 使 $\mathrm{d}\varphi=f\,\mathrm{d}\mu$. 称 f 为 φ 关于 μ 的 Radon-Nikodym 导数, 记成 $\frac{\mathrm{d}\varphi}{\mathrm{d}\mu}=f$.

推论 3.56. 设 ν 与 μ 是 $\mathscr{A}$ 上的 σ 有限测度, $\nu\ll\mu$. 如 f 是可测函数, 则 f 关于 ν 的积分存在当且仅当 $f\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ 关于 μ 的积分存在, 且 $\forall A\in\mathscr{A},\int_A f\,\mathrm{d}\nu=\int_A f\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,\mathrm{d}\mu$.

§ 3.5.3 分布函数的分解定理

将上面的分解定理应用于分布函数诱导的 L-S 测度与 Lebesgue 测度 $\mathrm{d}x$, 可导出下面的分布函数的分解定理.

定理 3.57. $\mathbb{R}^n$ 上任一有界的分布函数 F 都可以唯一地分解成三个分布函数之和, 即 $F=F_c+F_d+F_s$, 其中 F_c 诱导的 L-S 测度关于 $\mathrm{d}x$ 绝对连续, F_d 诱导的 L-S 测度支撑在一个至多可数集上, 而 F_s 诱导的 L-S 测度与 $\mathrm{d}x$ 是奇异的且在任何单点上取值为 0. 这样的分解在差分意义下是唯一的, 即如 F 还可分解成 $F=F_c'+F_d'+F_s'$, 则 $F_c-F_c',F_d-F_d',F_s-F_s'$ 的差分为 0. F_c,F_d,F_s 分别称作 F 的绝对连续部分, 离散部分和奇异部分.

证明 设 μ 为 F 诱导的 L-S 测度, 则由 Lebesgue 分解定理, $\mu = \mu_c + \mu_s'$, 其中 $\mu_c \ll \mathrm{d}x, \mu_s'$ 关于 μ 奇异. 令 $A = \{x \in \mathbb{R}^n : \mu_s'(\{x\}) > 0\}$, 则 A 至多可数. 定义 $\mu_d(B) = \sum\limits_{x \in B \cap A} \mu_s'\{x\}$, 令 F_d 为 μ_d 的分布函数. 最后, 令 $\mu_s = \mu_s' - \mu_d$, 则 μ_s 是有限测度, 关于 μ 奇异且 $\forall x \in \mathbb{R}^n$, 有 $\mu_s(\{x\}) = 0$. 令 $F_s = F - F_c - F_d$ 即可.

分解的唯一性由 Lebesgue 分解的唯一性和 F_s 与 F_d 的性质立得. $\square$

§3.6　补充与习题

1. 设 f 的积分存在. 证明

$$\int_\Omega f\,\mathrm{d}\mu = \lim_{n\to\infty}\sum_{i=-\infty}^{\infty}\frac{i}{2^n}\mu\left(\left\{\frac{i}{2^n}\leqslant f<\frac{i+1}{2^n}\right\}\right).$$

2. 设 f 为非负可测函数. 令

$$\bar{\int_\Omega} f\,\mathrm{d}\mu = \inf\left\{\int_\Omega g\,\mathrm{d}\mu : g\geqslant f, g\text{ 为简单函数}\right\}.$$

举例说明 $\bar{\int}_\Omega f\,\mathrm{d}\mu$ 与 $\int_\Omega f\,\mathrm{d}\mu$ 未必相同, 并解释在定义 3.3 中为何不用 $\bar{\int}_\Omega f\,\mathrm{d}\mu$?

3. 证明定理 3.8.

4. 举例说明 f 为复可测函数且积分存在, c 为复数, 但 cf 的积分未必存在. 如 f 可积呢?

5. 设 f 为可测复函数, 则 $|\int_\Omega f\,\mathrm{d}\mu|\leqslant\int_\Omega|f|\,\mathrm{d}\mu$.

6. 举例说明在定理 3.10 之 i) 中, 控制条件 $g\leqslant f_n$ 不可去.

7. 设 $\{f_{nm}\}_{n,m\geqslant 1}$ 为一族非负实数. 证明

$$\varliminf_{m\to\infty}\sum_{n=1}^{\infty}f_{nm}\geqslant\sum_{n=1}^{\infty}\varliminf_{m\to\infty}f_{nm}$$

8. 举例说明对于随机变量列, r 次平均收敛与 a.s. 收敛互不蕴含.

9. 设 φ 为有限的 σ 可加集函数, $\varphi\ll\mu$. 则 $\forall\{A_n\}\subset\mathscr{A}$ 使 $\mu(A_n)\to 0$, 有 $\varphi(A_n)\to 0$. 举例说明 φ 为 σ 有限的 σ 可加集函数时, 命题不真. (提示: $(0,1),\mu=\mathrm{d}x,\varphi=\int\frac{1}{x}\,\mathrm{d}\mu, A_n=(0,\frac{1}{n})$.)

10. 证明推论 3.37.

11. 证明命题 3.39

12. 设 $f:\mathbb{R}\to\mathbb{R}^+$, 令 $i_A=\inf\{f(x):x\in A\},A\in\mathscr{B}$, 证明

$$i_A\mathbb{P}(\xi\in A)\leqslant\int_{[\xi\in A]}f(\xi)\,\mathrm{d}\mathbb{P}\leqslant\mathbb{E}f(\xi).$$

13. 设 $\xi\geqslant 0$ 使得 $\mathbb{E}\xi^2<\infty$. 证明 $\mathbb{P}(\xi>0)\geqslant(\mathbb{E}\xi)^2/\mathbb{E}\xi^2$.

14. 设 $A_1,\cdots,A_n$ 为事件且 $A=\bigcup\limits_{i=1}^{n}A_i$. 证明

    (a) $\mathbf{1}_A\leqslant\sum_{i=1}^n\mathbf{1}_{A_i}$;
    (b) $\mathbb{P}(A)\geqslant\sum_{i=1}^n\mathbb{P}(A_i)-\sum_{i<j}\mathbb{P}(A_i\cap A_j)$;
    (c) $\mathbb{P}(A)\leqslant\sum_{i=1}^n\mathbb{P}(A_i)-\sum_{i<j}\mathbb{P}(A_i\cap A_j)+\sum_{i<j<k}\mathbb{P}(A_i\cap A_j\cap A_k)$.

15. 利用 Jensen 不等式证明几何平均值小于代数平均值, 并进一步证明 Carleman 不等式

16. 设 $\xi\geqslant 0$. 证明

$$\lim_{t\to\infty}t\int_{[\xi>t]}\frac{1}{\xi}\,\mathrm{d}\mathbb{P}=0,\quad\lim_{t\to 0}t\int_{[\xi>t]}\frac{1}{\xi}\,\mathrm{d}\mathbb{P}=0.$$

17. 随机变量 ξ,η 独立当且仅当对任意可测函数 f,g 有 $\mathbb{E}f(\xi)g(\eta)=\mathbb{E}f(\xi)\mathbb{E}g(\eta)$.

18. 设 ξ_n 是两两不相关的随机变量序列, $\mathbb{E}\xi_n=\mu,\mathrm{Var}(\xi_n)\leqslant C<\infty$. 令 $S_n=\xi_1+\cdots+\xi_n$, 则 $S_n/n\xrightarrow{L^2}\mu.$

19. 设 ξ_n 是独立同分布的随机变量序列, 服从 $(-1,1)$ 上的均匀分布. 证明

$$\frac{\xi_1^2+\cdots+\xi_n^2}{n}\xrightarrow{\mathbb{P}}\frac{1}{3}.$$

从而, 令 $A_{n,\varepsilon}=\left\{x\in\mathbb{R}^n:(1-\varepsilon)\sqrt{n/3}<|x|<(1+\varepsilon)\sqrt{n/3}\right\}$, 其中 $|x|$ 是欧氏距离, 则 $\forall\varepsilon>0$,

$$\frac{\mathrm{Leb(A_{n,\varepsilon}\cap(-1,1)^n)}}{2^n}\xrightarrow{\mathbb{P}}1.$$

20. 设 ξ_n 是独立同分布的随机变量序列, 服从 $\{1,2,\cdots,n\}$ 上的均匀分布. 令 $\tau_k^{(n)}=\inf\{m:|\{\xi_1,\cdots,\xi_m\}|=k\}, T_n=\tau_n^{(n)}$. 证明

$$\frac{T_n-n\sum_{m=1}^n 1/m}{n\log n}\xrightarrow{\mathbb{P}}0.$$

从而 $T_n/n\log n\xrightarrow{\mathbb{P}}1$.

21. (a) 设 f 为 $[0,1]$ 上的可测函数且 $\int_0^1|f(x)|\,\mathrm{d}x<\infty$. 令 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 服从 $[0,1]$ 上的均匀分布,

$$I_n=\frac{f(\xi_1)+\cdots+f(\xi_n)}{n}.$$

证明 $I_n\xrightarrow{\mathbb{P}}I\triangleq\int_0^1 f(x)\,\mathrm{d}x$,

(b) 设$\int_0^1|f(x)|^2\,\mathrm{d}x<\infty$, 估计 $\mathbb{P}(|I_n-I|\geqslant a/n^{1/2})$.

22. (a) 若随机事件 $A_1,A_2,\cdots$ 满足 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$, 则 $\mathbb{P}(\overline{\lim\limits_{n\to\infty}}A_n)=0$.

(b) 若随机事件 $A_1,A_2,\cdots$ 独立且 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)=\infty$, 则 $\mathbb{P}(\overline{\lim\limits_{n\to\infty}}A_n)=1$.

23. 设 $p_n\in[0,1)$. 利用上题证明 $\prod\limits_{n=1}^{\infty}(1-p_n)=0$ 当且仅当 $\sum\limits_{n=1}^{\infty}p_n=\infty$.

24. 设随机变量 ξ 取非负整数值, 证明 $\mathbb{E}\xi=\sum\limits_{n\geqslant 1}\mathbb{P}(\xi\geqslant n)$. 如果 ξ 取整数值呢?

25. 设随机变量 ξ 取非负实数值, 证明 $\mathbb{E}\xi=\int_0^\infty\mathbb{P}(\xi\geqslant x)\,\mathrm{d}x$. 对一般实数值的随机变量呢?

26. 设 $\phi\geqslant 0$ 使得 $\lim\limits_{x\to\infty}\phi(x)/x\to\infty$. 如果 $\mathbb{E}\phi(|\xi_t|)\leqslant C<\infty, t\in T$, 那么 $\{\xi_t,t\in T\}$ 一致可积.

27. 设 $\{f_n\}$ 为测度空间 $(\Omega,\mathscr{A},\mu)$ 上实可测函数列, $r>0$. 如 $\sup\limits_{n\geqslant 1}\mu(|f_n|^r)<\infty$, 则 $\forall s\in(0,r), \{|f_n|^s\}$ 积分一致连续.

28. 证明: 如 $\{\xi_t\}_{t\in T}$ 一致可积, 则必积分一致连续; 且当 μ 有限时, 两者等价.

29. 证明简单函数在 L^r 空间中稠密.

30. 设 $1/p+1/q=1, p,q>1$, 证明 $||f||_p=\{\mu(fg):||g||_q\leqslant 1\}$.

31. 如果随机变量序列 $\{\xi_n\}_{n\geqslant 1}$ 一致有界, 那么 ξ_n 依概率收敛当且仅当它们依 L^r 收敛.

32. 对可测函数 f, 定义**本征上确界**为

$$||f||_\infty=\inf\left\{M:\mu(\{\omega:|f(\omega)|>M\})=0\right\}.$$

(a) 证明 $||\cdot||$ 满足三角不等式.

(b) 若 $\mu(\Omega)<\infty$, 则 $||f||_\infty=\lim\limits_{r\to\infty}||f||_r$.

33. 证明 $c=\mathbb{E}\xi$ 使 $\mathbb{E}(\xi-c)^2$ 达到最小值.

34. 设 $\xi_1,\cdots,\xi_n$ 独立并且数学期望为 0, 方差有限. 令 $S_k=\sum\limits_{i=1}^k\xi_i$, 证明

(a)
$$\mathbb{P}(\max_{1\leqslant k\leqslant n}|S_k|\geqslant t)\leqslant\frac{1}{t^2}\mathrm{Var}(S_n).$$

(b) $\mathbb{P}(\max_{1\leqslant k\leqslant n}|S_k|\geqslant 3t)\leqslant\max_{1\leqslant k\leqslant n}3\mathbb{P}(|S_k|\geqslant t)$.

35. 若 ξ,η 独立且 $\xi+\eta$ 有二阶矩, 则 ξ 与 η 也有二阶矩.

36. 设随机变量 ξ 具有数学期望 m 与方差 σ^2.

(a) 证明
$$\mathbb{P}(\xi-m\geqslant t)\leqslant\frac{\sigma^2}{\sigma^2+t^2},\quad t\geqslant 0;$$

(b) 证明
$$\mathbb{P}(|\xi-m|\geqslant t)\leqslant\frac{2\sigma^2}{\sigma^2+t^2}.$$

37. 设 f 为 $\mathbb{R}^2$ 上的凸函数, 证明 $f(\mathbb{E}(\xi,\eta)) \leqslant \mathbb{E}(f(\xi,\eta))$.

38. 设分布函数 F 为连续函数.

   (a) 证明 F 一致连续;

   (b) 若 $F_n \Rightarrow F$, 则 F_n 一致收敛于 F.

39. 如果 f 可积, 那么 $\forall \varepsilon > 0$ 存在 $\delta > 0$ 使得当 $\mu(A) \leqslant \delta$ 时, $\int_A |f| \,\mathrm{d}\mu < \varepsilon$.

40. 如果 f 可积, 那么

$$\lim_{t\to\infty} \int_{\mathbb{R}} |f(x+t) - f(x)| \,\mathrm{d}x = 0.$$

41. 设随机变量 ξ 可积. 如果 $\mathbb{R}$ 上的有严格凸的函数 ϕ 使得 $\mathbb{E}\phi(\xi) = \phi(\mathbb{E}\xi)$, 那么 $\xi \equiv C$, a.s.

42. 证明一族随机变量 $\xi_t, t \in T$ 一致可积的充要条件存在函数 $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ 满足 $\lim\limits_{x\to\infty} \phi(x)/x = \infty$ 使得 $\sup\limits_{t\in T} \mathbb{E}(\phi(\xi_t) < \infty$.

43. 如果独立随机变量 ξ, η 分别具有分布律 μ, ν, 那么

$$\mathbb{P}((\xi,\eta) \in B) = \int_{\mathbb{R}} \mathbb{P}((x,\eta) \in B)\mu(\,\mathrm{d}x) = \int_{\mathbb{R}} \mathbb{P}((\xi,y) \in B)\nu(\,\mathrm{d}y), \forall B \in \mathscr{B}$$

及

$$\mathbb{P}(\xi \in A(\xi,\eta) \in B) = \int_{\mathbb{R}} \mathbb{P}((x,\eta) \in B)\mu(\,\mathrm{d}x), \forall A, B \in \mathscr{B}.$$

44. 设 μ, ν 为直线上的概率测度, $\nu \ll \mu$. 证明

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \lim_{h\to 0} \frac{\nu(x-h, x+h]}{\mu(x-h, x+h]}, \mu_{\text{a.s.}} x.$$

45. 证明 Hahn 分解定理 (定理 3.48) 中 φ^+ 与 φ^- 的定义与 D 选取无关.

46. (Cantor 集上的均匀分布) 在 $[0,1]$ 中去掉 $(1/3.2/3)$, 再分别去掉剩余两个区间的中间的开区间, 依次下去, 极限为 Cantor 集,记为 C . 定义

$$F(x)=\begin{cases}0, & x\leqslant 0;\\ 1, & x\geqslant 1;\\ \frac{1}{2}, & x\in[\frac{1}{3},\frac{2}{3};\\ \frac{1}{4}, & x\in[\frac{1}{9},\frac{2}{9}];\\ \frac{3}{4}, & x\in[\frac{7}{9},\frac{8}{9}].\\ \cdots, & \cdots\end{cases}$$

F 称为 Cantor 集上的均匀分布函数. 证明:(1) F 是连续的, (2) F 是 Lebesgue 奇异的.

47. 设 μ_1,μ_2 是有限符号测度, 令 $\mu_1\vee\mu_2=\mu_1+(\mu_2-\mu_1)^+,\mu_1\wedge\mu_2=\mu_1-(\mu_1=\mu_2)^+$, 则 $\mu_1\vee\mu_2$ 是满足 $\nu\geqslant\mu_i(i=1,2)$ 的最小符号测度; $\mu_1\wedge\mu_2$ 是满足 $\nu\leqslant\mu_i(i=1,2)$ 的最大符号测度.

48. 设 μ 为 $(\Omega,\mathscr{A})$ 上的 σ 有限测度, $\mathscr{A}$ 包含单点集. 则集合 $\{x\in\Omega:\mu(\{x\})>0\}$ 至多可数.

49. 设 $\mathscr{A}_n$ 是单调上升的 σ 代数序列, $\mathscr{F}=\sigma\left(\bigcup\limits_n\mathscr{A}_n\right)$. 设 μ,ν 分别是 $(\Omega,\mathscr{A})$ 上的有限测度和概率测度, 令 μ_n,ν_n 分别是 μ,ν 在 $\mathscr{A}_n$ 上的限制. 假设 $\mu_n\ll\nu_n$, 令 $f_n=\mathrm{d}\mu_n/\,\mathrm{d}\nu_n, f=\varlimsup\limits_n f_n$, 证明

$$\mu(A)=\int_A f\,\mathrm{d}\nu+\mu(A\cap\{f=\infty\}).$$

第四章　乘积测度空间

为什么要研究高维乃至无穷维空间呢?

有人说, 我们生活在三维空间中, 只需弄清楚三维空间就可以了. 事实是这样的吗? 当然不是. 要确定一个人的位置的确用三维坐标就足够了, 但每个人不是独立的, 他要与其他人交往, 所以我们通常需要研究一群人, 于是就需要使用更高维的空间了. 举一个例子, 考虑一个质点在直线上做随机运动, 那么给定一个时刻它的取值是一个一维的随机变量. 而要想刻画它在给定的 n 个时刻的取值情况, 我们就需要用 n 维的随机向量. 更进一步, 为刻画它在所有时刻的取值情况, 我们需要使用无穷维的随机向量 (随机过程). 许多时候, 高维空间可看成低维空间的乘积, 测度空间也是如此.

我们首先研究乘积测度空间上的积分, 讨论如何把重积分化为累次积分 (即 Fubini 定理), 然后介绍如何使用转移测度来构造乘积空间上的非乘积测度. 这将是后面研究条件期望与正则条件概率的基础, 也是未来学习随机过程的重要理论依据.

§4.1　Fubini 定理

我们首先回忆一下数学分析与实变函数论中所学过的重积分化累次积分. 设 A 为 $\mathbb{R}^2$ 中一个可测集, f 为一个关于 Lebesgue 测度可积的函数. 为计算重积分 $\int_A f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$, 先确定 x_1 再看当 x_1 固定时 x_2 的变化范围 $A_{x_1} \triangleq \{x_2\in\mathbb{R} : (x_1,x_2)\in A\}$. 则该重积分可化成累次积分 $\int_{\mathbb{R}} \mathrm{d}x_1 \int_{A_{x_1}} f(x_1,x_2)\,\mathrm{d}x_2$. 本节的目的是将这种方法推广到一般乘积测度空间上.

设 $(\Omega_1,\mathscr{A}_1,\mu_1)$ 与 $(\Omega_2,\mathscr{A}_2,\mu_2)$ 为两个 σ 有限的测度空间. 我们研究乘

积测度空间 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2,\mu_1\times\mu_2)$ 上的积分. 类似地, 给定 $\Omega_1\times\Omega_2$ 上可测集 A 与关于 $\mu_1\times\mu_2$ 积分存在的函数 f, 我们的目标是证明

$$\int_A f\,\mathrm{d}\mu_1\times\mu_2=\int_{\Omega_1}\mu_1(\,\mathrm{d}\omega_1)\int_{A_{\omega_1}}f(\omega_1,\omega_2)\mu_2(\,\mathrm{d}\omega_2). \tag{4.1.1}$$

此公式就是 Fubini 定理. 为此, 先引入截集 $A_{\omega_1}\subset\Omega_2$.

定义 4.1. 设 $A\subset\Omega_1\times\Omega_2$. $\forall\omega_1\in\Omega_1$, 称 $A_{\omega_1}\triangleq\{\omega_2\in\Omega_2:(\omega_1,\omega_2)\in A\}\subset\Omega_2$ 为 A 在 ω_1 处的截集. 同样定义 $A_{\omega_2}\subset\Omega_1,\forall\omega_2\in\Omega_2$.

显然截集具有如下性质.

性质 4.2. (1) $A\cap B=\emptyset\Rightarrow A_{\omega_i}\cap B_{\omega_i}=\emptyset$,

(2) $A\supset B\Rightarrow A_{\omega_i}\supset B_{\omega_i}$,

(3) $\left(\bigcup\limits_n A^{(n)}\right)_{\omega_i}=\bigcup\limits_n A^{(n)}_{\omega_i}$,

(4) $\left(\bigcap\limits_n A^{(n)}\right)_{\omega_i}=\bigcap\limits_n A^{(n)}_{\omega_i}$,

(5) $(A-B)_{\omega_i}=A_{\omega_i}-B_{\omega_i}$.

为证明上面的公式 (4.1.1), 首先需明确等式的右端有意义. 所以我们需证明如下事实: $\forall\omega_1\in\Omega_1$, 截集 $A_{\omega_1}\in\mathscr{A}_2$ 且 $\int_{A_{\omega_1}}f(\omega_1,\omega_2)\mu_2(\,\mathrm{d}\omega_2)$ 是 ω_1 的可测函数, 关于 μ_1 的积分存在. 下面逐步证明这些论断.

定理 4.3. 设 $A\in\mathscr{A}_1\times\mathscr{A}_2$, 则 $\forall\omega_1\in\Omega_1$ 有 $A_{\omega_1}\in\mathscr{A}_2$, $\forall\omega_2\in\Omega_2$ 有 $A_{\omega_2}\in\mathscr{A}_1$.

证明 设 $\mathscr{M}=\{A\in\mathscr{A}_1\times\mathscr{A}_2:\forall\omega_1\in\Omega_1,A_{\omega_1}\in\mathscr{A}_2,\omega_2\in\Omega_2,A_{\omega_2}\in\mathscr{A}_1\}$. 显然 $\mathscr{M}$ 包含所有可测矩形, 即包含集类 $\{A_1\times A_2:A_1\in\mathscr{A}_1,A_2\in\mathscr{A}_2\}$. 而由以上截集的性质 (性质 4.2) 知 $\mathscr{M}$ 为 σ 代数, 从而包含 $\mathscr{A}_1\times\mathscr{A}_2$. □

定理 4.4. 如果 f 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测的, 则 $\forall\omega_1\in\Omega_1,\omega_2\in\Omega_2,f_{\omega_1}(\cdot)\triangleq f(\omega_1,\cdot)$ 是 $\mathscr{A}_2$ 可测的, $f_{\omega_2}(\cdot)\triangleq f(\cdot,\omega_2)$ 是 $\mathscr{A}_1$ 可测的.

证明 由性质 2.4 和性质 4.2, 只需证 $\forall B\in\mathscr{B}$ 有 $[f^{-1}(B)]_{\omega_1}=f^{-1}_{\omega_1}(B)$. 事实上,

$$\begin{aligned}\left[f^{-1}(B)\right]_{\omega_1}&=\left\{\omega_2\in\Omega_2:(\omega_1,\omega_2)\in f^{-1}(B)\right\}\\&=\{\omega_2\in\Omega_2:f_{\omega_1}(\omega_2)\in B\}=f^{-1}_{\omega_1}(B).\end{aligned}$$ □

称 $f_{\omega_1},f_{\omega_2}$ 分别为 f 在 ω_1,ω_2 处的截函数.

定理 4.5. 设 f 是 $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2)$ 上非负可测函数, 则

$$\int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(\mathrm{d}\omega_1) \in \mathscr{A}_2, \ \int_{\Omega_2} f(\omega_2, \omega_1)\mu_2(\mathrm{d}\omega_2) \in \mathscr{A}_1.$$

证明 由可测函数的构造及积分的性质, 仅需证 $f = \mathbf{1}_A \in \mathscr{A}_1 \times \mathscr{A}_2$ 的情形. 而由单调类定理, 仅需证 $A = A_1 \times A_2$ 的情形. 此时 $\int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(\mathrm{d}\omega_1) = \mu_1(A_1)\mathbf{1}_{A_2} \in \mathscr{A}_2, \int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(\mathrm{d}\omega_2) = \mu_2(A_2)\mathbf{1}_{A_1} \in \mathscr{A}_1$. □

定理 4.6 (Fubini 定理). 若 f 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 可测实函数且积分存在, 则

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu_1 \times \mu_2 &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2)\mu_2(\mathrm{d}\omega_2)\right)\mu_1(\mathrm{d}\omega_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2)\mu_1(\mathrm{d}\omega_1)\right)\mu_2(\mathrm{d}\omega_2).\end{aligned}$$

证明 (1) 当 $f = \mathbf{1}_{A_1\times A_2}(A_i \in \mathscr{A}_i, i = 1, 2)$ 时显然成立. 由此并使用单调类定理知 $f = \mathbf{1}_A(A \in \mathscr{A}_1 \times \mathscr{A}_2)$ 时等式成立.

(2) 由积分的线性性质及 (1) 知当 f 为简单函数时等式也成立. 由此及非负可测函数的构造 (即非负可测函数是初等函数列单调上升的极限) 以及单调收敛定理, 当 f 为非负可测函数时等式也成立.

(3) 对一般的积分存在的 f, 有 $f = f^+ - f^-$. 设 f^- 的积分有限. 由此及 (2) 知

$$\infty > \int_{\Omega_1\times\Omega_2} f^-\,\mathrm{d}\mu_1 \times \mu_2 = \int_{\Omega_1} \mu_1(\mathrm{d}\omega_1) \int_{\Omega_2} f^-(\omega_1, \cdot)\,\mathrm{d}\mu_2.$$

故对 μ_1-a.e. $\omega_1, \int_{\Omega_2} f^-(\omega_1, \cdot)\,\mathrm{d}\mu_2 < \infty$. 从而 $\int_{\Omega_2} f(\omega_1, \cdot)\,\mathrm{d}\mu_2$ 存在且等于 $\int_{\Omega_2} f^+(\omega_1, \cdot)\,\mathrm{d}\mu_2 - \int_{\Omega_2} f^-(\omega_1, \cdot)\,\mathrm{d}\mu_2$. 由此结合积分的线性性质及 (2) 知等式对这样的 f 成立. □

虽然我们仅对 $A = \Omega_1 \times \Omega_2$ 时证明了 Fubini 定理, 但对于一般可测集 A , 只需将 f 换成 $f\mathbf{1}_A$ 即导出 (4.1.1). 此外, 由归纳法, 容易将该定理推广到有限乘积空间上.

设 $(\Omega_i, \mathscr{A}_i, \mu_i), 1 \leqslant i \leqslant n$ 为 n 个 σ 有限测度空间, f 为 $\Omega \triangleq \Omega_1 \times \cdots \times \Omega_n$ 上可测实函数, 关于 $\mu_1 \times \cdots \times \mu_n$ 的积分存在, 则

$$\int_\Omega f\,\mathrm{d}\mu_1 \times \cdots \times \mu_n = \int_{\Omega_{i_1}} \mathrm{d}\mu_{i_1} \int_{\Omega_{i_2}} \mathrm{d}\mu_{i_2} \cdots \int_{\Omega_{i_n}} f\,\mathrm{d}\mu_{i_n},$$

其中 $(i_1, \cdots, i_n)$ 为 $(1, \cdots, n)$ 的任一排列. 该等式的意思是右边的每一次积分均存在且最后的累次积分等于左边的重积分.

§4.2 无穷乘积概率空间

设 T 是一个无限集 (含无穷多个元素). $\forall t \in T, (\Omega_t, \mathscr{A}_t, \mathbb{P}_t)$ 是概率空间. 令 $\Omega_T = \prod\limits_{t\in T} \Omega_t = \{\omega : \omega = (\omega_t)_{t\in T}, \omega_t \in \Omega_t, t \in T\}$.

那么, 我们需要一个什么样的无穷乘积 σ 代数呢? 人们首先想到的可能是由无穷乘积矩形类 $\left\{\prod_{t\in T} A_t : A_t \in \mathscr{A}_t, t \in T\right\}$ 生成的 σ 代数. 但在这样的集合上无法定义概率测度, 因为无穷乘积 $\prod_{t\in T} \mathbb{P}_t(A_t)$ 通常不存在. 为此, 我们将仅允许 $\mathbb{P}_t(A_t)$ 中仅有有限个不是 1, 以使得该无穷乘积可化为有限乘积, 从而可定义. 这也是我们仅限于讨论无穷乘积概率测度, 而不去讨论一般的无穷乘积测度的原因.

根据上面分析, 先引入柱集的概念.

定义 4.7. 把形如
$$B^{T_N} \times \prod_{t\notin T_N} \Omega_t$$
的集合称为可测柱集, 其中 $T_N \subset T$ 为有限集, $B^{T_N} \in \mathscr{A}_{T_N} \triangleq \prod\limits_{t\in T_N} \mathscr{A}_t$. 称 B^{T_N} 为该柱集的底 (底不唯一!). 令 $\mathscr{A}^T$ 为一切可测柱集的全体, 它是一个集代数. 定义无穷乘积 σ 代数为 $\mathscr{A}_T = \sigma(\mathscr{A}^T)$, 也记成 $\prod\limits_{t\in T} \mathscr{A}_t$.

定理 4.8. 在 $(\Omega_T, \mathscr{A}_T)$ 上存在唯一概率 $\mathbb{P}$ 满足
$$\mathbb{P}\left(A^{T_N} \times \Omega_{T_N^C}\right) = \left(\prod_{t\in T_N} \mathbb{P}_t\right)\left(A^{T_N}\right), \tag{4.2.1}$$
其中
$$T_N \subset T \text{ 有限},\ \Omega_{T_N^C} = \prod_{t\notin T_N} \Omega_t,\ A^{T_N} \in \mathscr{A}_{T_N}.$$

证明 (1) (4.2.1) 定义了 $\mathscr{A}^T$ 上的一个集函数, 为此, 需证所定义的函数值与柱集表达式的选取无关.

设 $A^{T_N} \times \Omega_{T_N^C} = A^{T'_N} \times \Omega_{(T'_N)^C}$, $T_N \subset T$ 有限, $A^{T_N} \in \mathscr{A}_{T_N}$. 令 $T''_N = T_N \cap T'_N$. 则存在 $A^{T''_N} \in \prod\limits_{t\in T''_N} \mathscr{A}_t$ 使
$$A^{T_N} = A^{T''_N} \times \prod_{t\in T'_N - T''_N} \Omega_t, A^{T'_N} = A^{T''_N} \times \prod_{t\in T'_N - T''_N} \Omega_t.$$

从而 $\left(\prod\limits_{t\in T_N}\mathbb{P}_t\right)(A^{T_N})=\left(\prod\limits_{t\in T_N''}\mathbb{P}_t\right)(A^{T_N''})=\left(\prod\limits_{t\in T_N'}\mathbb{P}_t\right)(A^{T_N'})$.

(2) $\mathbb{P}$ 具有有限可加性.

设 $\{A_k\}_{k=1}^n\subset\mathscr{A}^T$ 互不相交. $\forall 1\leqslant k\leqslant n$, 令 T_k 为 T 中有限子集且 $A^{T_k}\in\prod\limits_{t\in T_k}\mathscr{A}_t$ 使 $A_k=A^{T_k}\times\Omega_{T_k^c}$, 则 $\sum\limits_{k=1}^n A_k=:A_0\in\mathscr{A}^T$. 令 $T_0\subset T$ 为有限集, $A^{T_0}\in\prod\limits_{t\in T_0}\mathscr{A}_t$ 使 $A_0=A^{T_0}\times\Omega_{T_0^c}$.

令 $T_N=\bigcup\limits_{k=0}^n T_k, A_k^{T_N}=A^{T_k}\times\prod\limits_{t\in T_N-T_k}\Omega_t$, 则 $A_k=A_k^{T_N}\times\Omega_{(T_N)^c}, n\geqslant k\geqslant 0$. 易见 $\left\{A_k^{T_N}\right\}_{k=1}^n\subset\prod\limits_{t\in T_N}\mathscr{A}_t$ 两两不交, 且 $\sum\limits_{k=1}^n A_k^{T_N}=A$. 由于 $\prod\limits_{t\in T_N}\mathbb{P}_t$ 是测度, 故由 $\mathbb{P}$ 的定义及测度的有限可加性得 $\sum\limits_{k=1}^n\mathbb{P}(A_k)=\mathbb{P}(A_0)$.

(3) 由于 $\mathscr{A}^T$ 为集代数且 $\mathbb{P}$ 有限可加, 为证 $\mathbb{P}$ 的 σ 可加性, 仅需证明其在 $\emptyset$ 处连续. 使用反证法.

设 $\{A_n\}_{n\geqslant 1}\subset\mathscr{A}^T$ 单调下降且 $\exists\varepsilon>0$ 使 $\mathbb{P}(A_n)\geqslant\varepsilon,\forall n\geqslant 1$. 往证 $\bigcap\limits_{n=1}^\infty A_n\neq\emptyset$. $\forall n\geqslant 1$, $\exists$ 有限集 $T_n\subset T$ 以及 $A_n^{T_n}\in\prod\limits_{t\in T_n}\mathscr{A}_t$ 使 $A_n=A_n^{T_n}\times\prod\limits_{t\notin T_n}\Omega_t$.

令 $T_\infty=\bigcup\limits_{n=1}^\infty T_n$, 则 T_∞ 可数, 记成 $T_\infty=\{t_1,t_2,\cdots\}$. 为证 $\bigcap\limits_{n=1}^\infty A_n\neq\emptyset$, 只需要证 $\exists(\bar\omega_{t_1},\cdots,\bar\omega_{t_n},\cdots)\in\prod\limits_{t\in T_\infty}\Omega_t$, 使 $\bigcap\limits_{j=1}^\infty A_j(\bar\omega_{t_1},\cdots,\bar\omega_{t_n},\cdots)\neq\emptyset$, 其中 $A_j(\bar\omega_{t_1},\cdots,\bar\omega_{t_n},\cdots)$ 为 A_j 在 $(\bar\omega_{t_1},\cdots,\bar\omega_{t_n},\cdots)$ 处的截集.

首先令 $B_1^{(j)}=\left\{\omega_{t_1}\in\Omega_{t_1}:\mathbb{P}^{t_1}(A_j(\omega_{t_1}))\geqslant\frac{\varepsilon}{2}\right\}$, 其中 $\mathbb{P}^{t_1}$ 为 $\mathbb{P}$ 在 $\mathscr{A}^{T\setminus\{t_1\}}$ 上的投影. 一般地, 对 $T'\subset T$, 记 $\mathbb{P}^{T'}$ 为 $\mathbb{P}$ 在 $\mathscr{A}^{T\setminus T'}$ 上的投影. 由于

$$\mathbb{P}^{t_1}(A_j(\omega_{t_1}))=\left(\prod_{t\in T_j\setminus\{t_1\}}\mathbb{P}_t\right)\left(A_j^{T_j}(\omega_{t_1})\right)$$

是关于 $\mathscr{A}_{t_1}$ 可测的, 从而 $B_1^{(j)}\in\mathscr{A}_{t_1}$ 且

$$\begin{aligned}\varepsilon\leqslant\mathbb{P}(A_j)=\mathbb{P}^{t_1}(A_j(\omega_{t_1}))&=\left(\prod_{t\in T_j\setminus\{t_1\}}\mathbb{P}_t\right)\left(A_j^{T_j}(\omega_{t_1})\right)\\&=\int_{\Omega_{t_1}}\mathbb{P}^{t_1}\left(A_j(\omega_{t_1})\right)\mathrm{d}\mathbb{P}_{t_1}\leqslant\int_{B_1^{(j)}}d\mathbb{P}_{t_1}+\int_{\Omega_{t_1}}\frac{\varepsilon}{2}\,\mathrm{d}\mathbb{P}_{t_1}\leqslant\mathbb{P}_{t_1}\left(B_1^{(j)}\right)+\frac{\varepsilon}{2},\end{aligned}$$

由此知 $\mathbb{P}_{t_1}(B_1^{(j)}) \geqslant \frac{\varepsilon}{2}$. 由 $\left\{B_1^{(j)}\right\}_{j=1}^{\infty}$ 的单调性知 $\mathbb{P}_{t_1}\left(\bigcap\limits_{j=1}^{\infty} B_1^{(j)}\right) > 0$, 故 $\exists \bar{\omega}_{t_1} \in \bigcap\limits_{j=1}^{\infty} B_1^{(j)}$.

一般地, 设对某 $k \geqslant 1$ 已有 $(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}) \in \Omega_{t_1} \times \cdots \Omega_{t_k}$ 使

$$\mathbb{P}_{\{t_1,\cdots,t_k\}}(A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}) \geqslant \frac{\varepsilon}{2^k}, \forall j.$$

令

$$B_{k+1}^{(j)} = \left\{\omega_{t_{k+1}} \in \Omega_{t_{k+1}} : \mathbb{P}_{\{t_1,\cdots,t_k,t_{k+1}\}}(A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}, \omega_{t_{k+1}}) \geqslant \frac{\varepsilon}{2^{k+1}}\right\},$$

则

$$\begin{aligned}\frac{\varepsilon}{2^k} &\leqslant \mathbb{P}_{\{t_1,\cdots,t_k\}}(A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}) \\ &= \int_{\Omega_{t_{k+1}}} \mathbb{P}_{\{t_1,\cdots,t_k,t_{k+1}\}}(A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}, \omega_{t_{k+1}}) \mathrm{d}\mathbb{P}_{t_{k+1}}(\omega_{t_{k+1}}) \\ &\leqslant \mathbb{P}(B_{k+1}^{(j)}) + \frac{\varepsilon}{2^{k+1}}.\end{aligned}$$

于是 $\mathbb{P}(B_{k+1}^{(j)}) \geqslant \frac{\varepsilon}{2^{k+1}}, \forall j$, 即 $\exists \bar{\omega}_{t_{k+1}} \in \bigcap\limits_{j=1}^{\infty} B_{k+1}^{(j)}$ 且 $\mathbb{P}_{\{t_1,\cdots,t_k,t_{k+1}\}}(A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_k}, \bar{\omega}_{t_{k+1}}) \geqslant \frac{\varepsilon}{2^{k+1}}, \forall j$. 从而 $\exists \{\bar{\omega}_{t_i} \in \Omega_{t_i}\}_{i \geqslant 1}$ 使得 $\bigcap\limits_{j=1}^{\infty} A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_n}) \neq \emptyset, \forall n \geqslant 1$.

取定 $\tilde{\omega} \in \prod_{t \notin T_\infty} \Omega_t$, 令 $\omega \in \prod_{t \in T} \Omega_t$ 使得

$$\omega_t = \begin{cases} \bar{\omega}_t, & 若\ t \in T_\infty, \\ \tilde{\omega}_t, & 若\ t \notin T_\infty. \end{cases}$$

则 $\forall j, \exists N_j$ 使 $T_j \subset \{t_1, \cdots, t_{N_j}\}$, 而 $A_j(\bar{\omega}_{t_1}, \cdots, \bar{\omega}_{t_{N_j}}) \neq \emptyset$, 从而 $\omega \in A_j, \forall j$, 即 $\omega \in \bigcap\limits_j A_j$. □

定理 4.9. 设 $\{F_t\}_{t \in T}$ 是一族概率分布函数, 那么存在一族随机变量 $\{\xi_t\}_{t \in T}$ 使 ξ_t 的分布函数恰为 F_t.

证明 设 $\mathbb{P}_t$ 是 F_t 诱导的 $(\mathbb{R}, \mathscr{B})$ 上概率测度. 令 $(\Omega_t, \mathscr{A}_t, \mathbb{P}_t) = (\mathbb{R}, \mathscr{B}, \mathbb{P}_t), \Omega = \mathbb{R}^T, \mathscr{A} = \mathscr{B}^T, \mathbb{P} = \mathbb{P}^T$. 则 $(\xi_t(\omega) \triangleq \omega_t)_{t \in T}$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上的随机变量, 且 $\mathbb{P}(\xi_t < x_t) = \mathbb{P}_t((-\infty, x_t)) = F_t(x_t)$. 独立性显然. □

§4.3 转移测度与转移概率

回顾一下乘积测度空间 $(\Omega_1\times\Omega_2, \mathscr{A}_1\times\mathscr{A}_2, \mu)$ 及 Fubini 定理, 如果 μ 不是乘积测度, 如何化累次积分? 为此, 我们需要使用转移概率.

定义 4.10. 设 $(\Omega_1,\mathscr{A}_1)$, $(\Omega_2,\mathscr{A}_2)$ 为两个可测空间. $\lambda:\Omega_1\times\mathscr{A}_2\to[0,\infty]$ 满足:

1) $\forall\omega_1\in\Omega_1, \lambda(\omega_1,\cdot)$ 为 $(\Omega_2,\mathscr{A}_2)$ 上的测度;

2) $\forall A\in\mathscr{A}_2, \lambda(\cdot,A)$ 为 $\mathscr{A}_1$ 可测函数,

则称 λ 为 $(\Omega_1,\mathscr{A}_1)$ 到 $(\Omega_2,\mathscr{A}_2)$ 上的一个转移测度, 简称为 $\Omega_1\times\mathscr{A}_2$ 上的转移测度. 如 $\exists\Omega_2$ 的可数划分 $\{B_n\}\subset\mathscr{A}_2$ 使 $\lambda(\omega_1,B_n)<\infty\,(\forall n\geqslant 1,\omega_1\in\Omega_1)$, 则称 λ 为 σ 有限的. 更进一步, 如 $\sup\limits_{\omega_1\in\Omega_1}\lambda(\omega_1,B_n)<\infty\,(\forall n\geqslant 1)$, 则称 λ 为一致 σ 有限的. 如 $\forall\omega_1\in\Omega_1, \lambda(\omega_1,\cdot)$ 为概率测度, 则称 λ 为转移概率.

转移概率的第一个应用是构造乘积空间上的非乘积测度，进而将 Fubini 定理推广到乘积空间上关于一般测度的重积分化累次积分. 为此, 我们需要下面的定理.

定理 4.11. 设 λ 为 $\Omega_1\times\mathscr{A}_2$ 上的 σ 有限转移测度, f 是 $\mathscr{A}_1\times\mathscr{A}_2$ 上的非负可测函数, 则 $\int_{\Omega_2}f(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2)$ 是 $\mathscr{A}_1$ 可测的.

证明 由可测函数的构造及积分的性质, 仅需对 f 为示性函数证明. 而由单调类定理, 仅需考察 $f=\mathbf{1}_{A\times B}, A\in\mathscr{A}_1, B\in\mathscr{A}_2$ 的情形. 此时 $\int_{\Omega_2}f(\omega_1,\omega_2)\lambda(\omega_1,\mathrm{d}\omega_2)=\lambda(\omega_1,B)\mathbf{1}_A(\omega_1)$, 因而关于 ω_1 是 $\mathscr{A}_1$ 可测的. □

定理 4.12. 设 $(\Omega_i,\mathscr{A}_i), i=1,2,\cdots,n$ 是 n 个可测空间, 设 $\Omega^{(k)}=\prod_{i=1}^k\Omega_i$, $\mathscr{A}^{(k)}=\prod_{i=1}^k\mathscr{A}_i, k=1,\cdots,n$. 若 λ_1 为 $\mathscr{A}_1$ 测度, λ_k 是$\Omega^{(k-1)}\times\mathscr{A}_k$ 上 σ 有限转移测度, $k=2,\cdots,n$, 则 $\forall B\in\mathscr{A}^{(n)}$

$$\lambda^{(n)}(B)=\int_{\Omega_1}\cdots\int_{\Omega_n}1_B(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n)\cdots\lambda_1(\mathrm{d}\omega_1).$$

是测度. 若 $\lambda_i(i\geqslant 2)$ 是一致 σ 有限的, 则 $\lambda^{(n)}$ 是 σ 有限的.

证明 由定理 4.11 知, $\lambda^{(n)}(B)$ 的定义是合理的. 由单调收敛定理知 $\lambda^{(n)}$ 是 σ 可加的, 从而为测度.

今证 $\lambda^{(n)}$ 的 σ 有限性. 由归纳法, 只需对 $n=2$ 证明.

由于 λ_1 为 σ 有限的, λ_2 为一致 σ 有限的, 则 $\exists\Omega_1$ 与 Ω_2 的可数分割 $\{A_n\}$ 及 $\{B_n\}$ 使 $\lambda_1(A_n)<\infty,\ \sup\limits_{\omega_1\in A_m}\lambda_2(\omega_1,B_n)<\infty, m,n\geqslant 1$. 则 $\{A_i\times B_j\}_{i,j\geqslant 1}$ 为 $\Omega_1\times\Omega_2$ 的可数分割且

$$\begin{aligned}\lambda^{(2)}(A_i\times B_j)&=\int_{A_i}\lambda_1(\mathrm{d}\omega_1)\int_{B_j}\lambda_2(\omega_1,\mathrm{d}\omega_2)=\int_{A_i}\lambda_2(\omega_1,B_j)\lambda_1(\mathrm{d}\omega_1)\\&\leqslant\sup_{\omega_1\in A_i}\lambda_2(\omega_1,B_j)\lambda_1(A_i)<\infty.\quad\square\end{aligned}$$

定理 4.13 (重积分化累次积分). 设 $\lambda^{(n)}$ 如定理 4.12 所定义, f 是 $(\Omega^{(n)},\mathscr{A}^{(n)})$ 上可测函数, 则 f 关于 $\lambda^{(n)}$ 的积分存在的充分必要条件为

$$\int_{\Omega_1}...\int_{\Omega_n}f^{\pm}(\omega_1,...,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n)\cdots\lambda_1(\mathrm{d}\omega_1)\triangleq I(f^{\pm})$$

中至少有一个有限. 当积分存在时,

$$\int_{\Omega^{(n)}}f\,\mathrm{d}\lambda^{(n)}=\int_{\Omega_1}\cdots\int_{\Omega_n}f(\omega_1,...,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n)\cdots\lambda_1(\mathrm{d}\omega_1).$$

证明 当 $f=\mathbf{1}_B, B\in\mathscr{A}^{(n)}$ 时, 由定理 4.12 知等式成立. 由非负可测函数的构造及积分的性质知对非负可测函数也成立. 一般地, 有 $f=f^+-f^-$. 设 f 的积分存在, 不妨设 $\int_\Omega f^-\,\mathrm{d}\lambda^{(n)}<\infty$, 则有 $I(f^-)=\int_{\Omega^{(n)}}f^-\,\mathrm{d}\lambda^{(n)}<\infty$. 反之亦然. 从而积分存在当且仅当 $I(f^{\pm})$ 中至少有一个有限. 此时有

$$\begin{aligned}&\int f\,\mathrm{d}\lambda^{(n)}=\int f^+\,\mathrm{d}\lambda^{(n)}-\int f^-\,\mathrm{d}\lambda^{(n)}=I(f^+)-I(f^-)\\&=\int_{\Omega_1}\mathrm{d}\lambda_1\left\{\int_{\Omega_2}\cdots\int_{\Omega_n}f^+\,\mathrm{d}\lambda_n\cdots\mathrm{d}\lambda_2-\int_{\Omega_2}\cdots\int_{\Omega_n}f^-\,\mathrm{d}\lambda_n...\mathrm{d}\lambda_2\right\}\\&=\cdots=\int_{\Omega_1}\cdots\int_{\Omega_n}(f^+-f^-)\,\mathrm{d}\lambda_n\cdots\mathrm{d}\lambda_1=\int_{\Omega_1}\cdots\int_{\Omega_n}f\,\mathrm{d}\lambda_n\cdots\mathrm{d}\lambda_1.\quad\square\end{aligned}$$

最后讨论无穷维情形.

定理 4.14 (Tulcea定理). 设 $(\Omega_n,\mathscr{A}_n), n\geqslant 1$ 为可数个可测空间, $(\Omega^{(n)},\mathscr{A}^{(n)})$ 如上定义. 令 $\Omega^{(\infty)}=\prod_{i=1}^{\infty}\Omega_i, \mathscr{A}^{(\infty)}=\prod_{i=1}^{\infty}\mathscr{A}_i$. 设 $\mathbb{P}_n(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n)$ 为 $\Omega^{(n-1)}\times\mathscr{A}_n$ 上的转移概率, $\mathbb{P}_1$ 是 $\mathscr{A}_1$ 上的概率, 则在 $\mathscr{A}^{(\infty)}$ 上存在唯一的概率 $\mathbb{P}^{(\infty)}$ 使 $\mathbb{P}^{(\infty)}(B^{(n)}\times\prod_{k>n}\Omega_k)=\mathbb{P}^{(n)}(B^{(n)}), B^{(n)}\in\mathscr{A}^{(n)}, n\geqslant 1$. 其中 $\mathbb{P}^{(n)}(B^{(n)})=\int_{\Omega_1}\cdots\int_{\Omega_n}\mathbf{1}_{B^{(n)}}\,\mathrm{d}\mathbb{P}_n\cdots\mathrm{d}\mathbb{P}_1$ 为 $\mathscr{A}^{(n)}$ 的概率测度.

证明 设 $\mathscr{D}$ 为可测柱集全体. 则 $\mathscr{D}$ 为集代数且 $\mathbb{P}^{(\infty)}$ 在 $\mathscr{D}$ 上有限可加, 在 ϕ 处连续 (此法与无穷乘积概率类似). 故 $\mathbb{P}^{(\infty)}$ 为 $\mathscr{D}$ 上的概率测度, 可唯一扩张成 $\mathscr{A}^{(\infty)}$ 上的一个概率测度. $\square$

§4.4 补充与习题

1. 证明性质 4.2.

2. 设 f 是 σ 有限测度空间 $(\Omega, \mathscr{A}, \mu)$ 上的非负可测函数, λ 是 Lebesgue 测度. 证明

   (a) $\int_\Omega f \,\mathrm{d}\mu = \mu \times \lambda(\{(\omega, x) : 0 < x < f(\omega)\})$;

   (b) $\forall x \in \mathbb{R}, \mu \times \lambda(\{(\omega, x) : x = f(\omega)\})$.

3. 设 $T = (r\cos\theta, r\sin\theta)$ 是 $[0,\infty) \times [0, 2\pi) \mapsto \mathbb{R}^2$ 一一映射. 定义 $[0,\infty)$ 上的测度 $\mu(A) = \int_A r \,\mathrm{d}r$, λ 是 Lebesgue 测度, 则 $\mu \times \lambda$ 是 $\mathbb{R}^2$ 上的 Lebesgue 测度.

4. 设 f, g 是非负可测函数, 证明 $\mu(fg)^m \leqslant \mu(f)^{m-1}\mu(fg^m)$.

5. 设 F, G 为 $[a, b]$ 上的分布函数, 且没有相同的不连续点. 证明

$$\int_{[a,b)} G(x) \,\mathrm{d}F(x) = F(b)G(b) - F(a)G(a) - \int_{[a,b)} F(x) \,\mathrm{d}G(x).$$

6. 若 $\emptyset \neq A_1 \times A_2 \in \mathscr{A}_1 \times \mathscr{A}_2$, 则 $A_i \in \mathscr{A}_i, i = 1, 2$.

7. 若 $\mu_2\{\omega_2 : (\omega_1, \omega_2) \in A\} = \mu_2\{\omega_2 : (\omega_1, \omega_2) \in B\}$ 对所有 ω_1 成立, 则 $\mu_1 \times \mu_2(A) = \mu_1 \times \mu_2(B)$.

8. 设 $(\Omega_i, \mathscr{A}_i), i = 1, 2, 3$ 为三个可测空间, λ 为 $\Omega_2 \times \mathscr{A}_3$ 上的 σ 有限转移概率, f 是 $\mathscr{A}_1 \times \mathscr{A}_3$ 可测函数. 若 $\forall (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$, 积分 $g(\omega_1, \omega_2) \triangleq \int_{\Omega_3} f(\omega_1, \omega_3)\lambda(\omega_2, \mathrm{d}\omega_3)$ 存在. 证明 g 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 可测的.

9. 设 μ_k 是 $\mathscr{A}_k$ 上的 σ 有限测度, $k = 1, 2, A \in \mathscr{A}_1 \times \mathscr{A}_2$, 则下述结论等价:

   (a) $\mu_1 \times \mu_2(A) = 0$,

   (b) $\mu_1(A_{\omega_2}) = 0, \mu_2$-a.e.,

   (c) $\mu_2(A_{\omega_1}) = 0, \mu_1$-a.e..

10. 若矩阵 $P=(p_{ij})_{i,j=1}^{\infty}$ 满足 $p_{ij}\geqslant 0, \sum\limits_{j=1}^{\infty}p_{ij}=1, \forall i\geqslant 1$, 称 P 为转移概率矩阵. 令 $\lambda(i,A)=\sum\limits_{j\in A}p_{ij}$, 证明 λ 是转移概率.

11. 设 $\Omega_1=\Omega_2=\mathbb{N}, \mu_1, \mu_2$ 为计数测度,

$$f(i,j)=\begin{cases} i, & i=j; \\ -i, & j=i+1; \\ 0, & \text{其他 } i,j. \end{cases}$$

证明 $\int_{\Omega_1}\left(\int_{\Omega_2}f\,\mathrm{d}\mu_2\right)\mathrm{d}\mu_1=0, \int_{\Omega_2}\left(\int_{\Omega_1}f\,\mathrm{d}\mu_1\right)\mathrm{d}\mu_2=\infty$.

12. 试给出满足下列条件的函数 $f:[0,1]\times[0,1]\to[0,1]$:

(a) $\forall x\in[0,1], f(x,y)$ 关于 y Borel 可测, $\forall y\in[0,1], f(x,y)$ 关于 x Borel 可测,

(b) f 关于 $\mathscr{B}([0,1]\times[0,1])$ 不可测,

(c) $\int_0^1\left(\int_0^1 f(x,y)\,\mathrm{d}y\right)\mathrm{d}x$ 与 $\int_0^1\left(\int_0^1 f(x,y)\,\mathrm{d}x\right)\mathrm{d}y$ 均存在但不相等.

13. 设 μ_k, ν_k 分别为 $(\Omega_k, \mathscr{A}_k)$ 上 σ 有限测度, $\nu_k\ll\mu_k, k=1,2$, 则 $\nu_1\times\nu_2\ll\mu_1\times\mu_2$ 且

$$\frac{\mathrm{d}(\nu_1\times\nu_2)}{\mathrm{d}(\mu_1\times\mu_2)}(\omega_1,\omega_2)=\frac{\mathrm{d}\nu_1}{\mathrm{d}\mu_1}(\omega_1)\frac{\mathrm{d}\nu_2}{\mathrm{d}\mu_2}(\omega_2),\ \mu_1\times\mu_2-a.e.$$

14. 设 $(\Omega_t, \mathscr{A}_t)_{t\in T}$ 为一族可测空间, $\mathscr{C}_t\subset\mathscr{A}_t, \sigma(\mathscr{C}_t)=\mathscr{A}_t, t\in T$, 则 $\prod_{t\in T}\mathscr{A}_t=\sigma(\cup_{t\in T}J_t^{-1}(\mathscr{C}_t))$, 其中 J_t 为 $\prod_{t\in T}\Omega_t\ni\omega\to\omega_t\in\Omega_t$.

15. 证明下面的集合关于 $\mathscr{B}^{\infty}$ 可测:

(a) $\left\{x\in\mathbb{R}^{\infty}:\sup\limits_n x_n<a\right\}$,

(b) $\left\{x\in\mathbb{R}^{\infty}:\sum\limits_{n=1}^{\infty}|x_n|<\infty\right\}$,

(c) $\left\{x\in\mathbb{R}^{\infty}:\lim\limits_{n\to\infty}x_n \text{ 存在有限}\right\}$,

(d) $\left\{x\in\mathbb{R}^{\infty}:\overline{\lim\limits_n}\, x_n\leqslant a\right\}$.

第五章　条件概率与条件期望

由于世间万物存在各种各样的联系，我们往往需要在一定的限制条件下研究一些量的变化。例如，为应对 2008 年爆发的金融危机, 各国政府纷纷出台经济刺激方案, 其中包括调整存贷款利息. 那么降息的幅度为多大时对目前经济的发展才是最合适的呢? 为此,需要讨论存贷款利息对经济指标 (如 GDP) 的影响, 即刻画在给定存贷款利息情况下经济指标的分布. 一般地, 我们把问题抽象为: 在给定一些随机变量的前提下研究另一些随机变量的分布. 由于给定一些随机变量的取值范围就是给定由这些随机变量生成的 σ 代数中的一个事件, 所以该问题最终归结成研究给定 σ 代数下的条件概率与条件期望. 为此, 我们从熟知的给定事件下的条件概率讲起.

设 $(\Omega, \mathscr{A}, \mathbb{P})$ 为概率空间, $B \in \mathscr{A}$ 且 $\mathbb{P}(B) > 0$. 那么 $\mathbb{P}(\cdot|B) \triangleq \frac{\mathbb{P}(\cdot \cap B)}{\mathbb{P}(B)}$ 也是 $\mathscr{A}$ 上的概率测度, 称为给定事件 B 下的条件概率. 有了条件概率就可以将积分存在的随机变量 ξ 的条件期望定义为 ξ 对于条件概率的积分 $\mathbb{E}(\xi|B)$. 同样地, 如 $\mathbb{P}(B^c) > 0$, 则我们有给定事件 B^c 下的条件概率 $\mathbb{P}(\cdot|B^c)$ 及条件期望 $\mathbb{E}(\cdot|B^c)$. 于是可以按如下方式定义 ξ 在给定 σ 代数 $\mathscr{C} = \{B, B^c, \emptyset, \Omega\}$ 下的条件期望: 如 B 发生, 则定义为 $\mathbb{E}(\xi|B)$, 否则定义为 $\mathbb{E}(\xi|B^c)$. 从而, ξ 在给定 σ 代数 $\mathscr{C}$ 下的条件期望为

$$\mathbb{E}(\xi|\mathscr{C}) = \mathbf{1}_B \mathbb{E}(\xi|B) + \mathbf{1}_{B^c} \mathbb{E}(\xi|B^c),$$

它是关于 $\mathscr{C}$ 可测的随机变量.

本章的目的是引入给定 $\mathscr{A}$ 的子 σ 代数下的条件概率和条件期望, 并用之研究转移概率及乘积空间上的一般概率测度.

§5.1 给定 σ 代数下的条件期望

首先, 我们很容易将前面引入的条件期望推广到由可数个原子生成的 σ 代数上. 一般地, 设 $\mathscr{C}$ 为 σ 代数, $B \in \mathscr{C}$. 如果 $\forall B' \in \mathscr{C}, B' \subset B$, 则有 $B' = B$ 或 $B' = \emptyset$, 我们就称 B 为 $\mathscr{C}$ 的一个原子.

定义 5.1. 设 $(\Omega, \mathscr{A}, \mathbb{P})$ 为概率空间, $\{B_n\}_{n\geqslant 1} \subset \mathscr{A}$ 为 Ω 的可数分割, 令 $\mathscr{C} = \sigma(\{B_n : n \geqslant 1\})$. 设 ξ 为期望存在的随机变量, 则称

$$\mathbb{E}(\xi|\mathscr{C}) = \sum_{n=1}^{\infty} \mathbb{E}(\xi|B_n)\mathbf{1}_{B_n}$$

为 ξ 在给定 σ 代数 $\mathscr{C}$ 下的关于 $\mathbb{P}$ 的条件期望.

如何将此定义推广到一般子 σ 代数 $\mathscr{C}$ 上呢? 为此, 我们需提炼出条件期望不依赖于 $\mathscr{C}$ 表示的特征, 然后根据此特征来重新定义条件期望. 由 $\mathbb{E}(\xi|\mathscr{C})$ 的定义容易证明下面的命题.

命题 5.2. 设 $\mathbb{E}\xi$ 存在, $\{B_n\}_{n\geqslant 1} \subset \mathscr{A}$ 为 Ω 的可数分割, $\mathscr{C} = \sigma(\{B_n : n \geqslant 1\})$. 则 $\mathbb{E}(\xi|\mathscr{C})$ 关于 $\mathscr{C}$ 可测且满足

$$\int_B \xi \,\mathrm{d}\mathbb{P} = \int_B \mathbb{E}(\xi|\mathscr{C}) \,\mathrm{d}\mathbb{P},\ \forall B \in \mathscr{C}.$$

如果 η 为关于 $\mathscr{C}$ 可测的函数 (可能取 $\pm\infty$ 值) 使得 $\mathbb{E}\xi\mathbf{1}_B = \mathbb{E}\eta\mathbf{1}_B, \forall B \in \mathscr{C}$, 则 $\eta = \mathbb{E}(\xi|\mathscr{C})$.

根据上面命题, 我们引入给定一般 σ 代数下的条件期望.

定义 5.3. 设 $\mathscr{C} \subset \mathscr{A}$ 为 σ 代数, ξ 为期望存在的随机变量. ξ 在 $\mathscr{C}$ 之下 (关于 $\mathbb{P}$) 的条件期望 $\mathbb{E}(\xi|\mathscr{C})$ 定义为满足下列公式的关于 $\mathscr{C}$ 可测的函数

$$\int_B \mathbb{E}(\xi|\mathscr{C}) \,\mathrm{d}\mathbb{P} = \int_B \xi \,\mathrm{d}\mathbb{P}, \forall B \in \mathscr{C}.$$

为保证定义 5.3 的合理性, 需证明 $\mathbb{E}(\xi|\mathscr{C})$ 的存在唯一性. 为此, 令 $\varphi(B) = \int_B \xi \,\mathrm{d}\mathbb{P}, B \in \mathscr{C}$, 则 φ 是 $\mathscr{C}$ 上 σ 可加集函数, $\varphi \ll \mathbb{P}|_{\mathscr{C}}$. 从而存在 $\mathbb{P}_{\mathscr{C}}$-a.s. 唯一的 $f \in \mathscr{C}$ 使 $\mathrm{d}\varphi = f \,\mathrm{d}\mathbb{P}|_{\mathscr{C}}$, 即: $\forall B \in \mathscr{C}$, $\varphi(B) = \int_B f \,\mathrm{d}\mathbb{P} = \int_B \xi \,\mathrm{d}\mathbb{P}$.

定义 5.4. 设 $\mathscr{C} \subset \mathscr{A}$ 为 σ 代数, $A \in \mathscr{A}$. 称 $\mathbb{P}(A|\mathscr{C}) = \mathbb{E}(\mathbf{1}_A|\mathscr{C})$ 为 A 在 $\mathscr{C}$ 下 (关于 $\mathbb{P}$) 的条件概率.

根据条件期望的定义及积分的性质, 给定 σ 代数下的条件期望继承了期望的所有性质. 留意条件期望是 $\mathbb{P}_{\mathscr{C}}$-a.s. 定义的, 所有这些性质也是在 $\mathbb{P}_{\mathscr{C}}$-a.s. 意义下成立的.

性质 5.5. 设下面涉及的随机变量期望存在.

(1) $\mathbb{E}(\mathbb{E}(\xi|\mathscr{C}))=\mathbb{E}\xi$.

(2) 若 $\xi\in\mathscr{C}$, 则 $\mathbb{E}(\xi|\mathscr{C})=\xi$.

(3) (单调性) $\xi\leqslant\eta\Rightarrow\mathbb{E}(\xi|\mathscr{C})\leqslant\mathbb{E}(\eta|\mathscr{C})$.

(4) (线性性) $\mathbb{E}(a\xi+b\eta|\mathscr{C})=a\mathbb{E}(\xi|\mathscr{C})+b\mathbb{E}(\eta|\mathscr{C}), a,b\in\mathbb{R}$.

(5) (Fatou-Lebesgue收敛定理) 设 η,ζ 可积. 若 $\forall n\geqslant 1,\eta\leqslant\xi_n,\mathbb{P}$-a.e. 则 $\mathbb{E}\left(\varliminf\limits_{n\to\infty}\xi_n|\mathscr{C}\right)\leqslant\varliminf\limits_{n\to\infty}\mathbb{E}(\xi_n|\mathscr{C})$.

若$\forall n\geqslant 1,\xi_n\leqslant\zeta$, 则 $\varlimsup\limits_{n\to\infty}\mathbb{E}(\xi_n|\mathscr{C})\leqslant\mathbb{E}\left(\varlimsup\limits_{n\to\infty}\xi_n|\mathscr{C}\right)$

(6)(控制收敛定理) 设 ξ,η 可积, 若 $\eta\leqslant\xi_n\uparrow\xi$, 或者 $\forall n\geqslant 1,|\xi_n|\leqslant\eta$, 且 $\xi_n\to\xi$ a.s. 则 $\mathbb{E}(\xi_n|\mathscr{C})\to\mathbb{E}(\xi|\mathscr{C})$, a.s.

如同积分收敛定理的证明, 上面提到的条件期望的各种收敛定理均可由单调收敛定理导出 (见习题 3). 此外, 积分与期望的一些不等式 (如 Jensen 不等式, Hölder 不等式, Minkowski 不等式等) 也可自然推广到条件期望情形.

推论 5.6. 设 $\eta\in\mathscr{C},\mathbb{E}\xi\eta,\mathbb{E}\xi$ 存在, 则 $\mathbb{E}(\xi\eta|\mathscr{C})=\eta\mathbb{E}(\xi|\mathscr{C})$.

证明 由于 $\mathbb{E}(\xi\eta|\mathscr{C})$ 与 $\eta\mathbb{E}(\xi|\mathscr{C})$ 均关于 $\mathscr{C}$ 可测, 由条件期望的定义只要证明

$$\int_C\xi\eta\,\mathrm{d}\mathbb{P}=\int_C\eta\mathbb{E}(\xi|\mathscr{C})\,\mathrm{d}\mathbb{P}, C\in\mathscr{C}.$$

由积分 (或条件期望) 的线性性质、单调收敛定理, 只需对 ξ 与 η 为示性函数时证明该等式.

设 $\xi=\mathbf{1}_A,\eta=\mathbf{1}_B,A\in\mathscr{A},B\in\mathscr{C}$. 则

$$\int_C\eta\mathbb{E}(\xi|\mathscr{C})\,\mathrm{d}\mathbb{P}=\int_{C\cap B}\xi\,\mathrm{d}\mathbb{P}=\mathbb{P}(A\cap B\cap C)=\int_C\xi\eta\,\mathrm{d}\mathbb{P}.\quad\square$$

推论 5.7. 若 $\xi_n\xrightarrow{r}\xi, r\geqslant 1$, 则 $\mathbb{E}(\xi_n|\mathscr{C})\xrightarrow{r}\mathbb{E}(\xi|\mathscr{C})$.

证明 由 Jensen 不等式及条件期望的性质,

$$\begin{aligned}\mathbb{E}|\mathbb{E}(\xi_n|\mathscr{C})-\mathbb{E}(\xi|\mathscr{C})|^r&=\mathbb{E}|\mathbb{E}(\xi_n-\xi|\mathscr{C})|^r\leqslant\mathbb{E}(\mathbb{E}(|\xi_n-\xi|^r|\mathscr{C}))\\&=\mathbb{E}|\xi_n-\xi|^r\to 0\ (n\to\infty).\quad\square\end{aligned}$$

下面的定理表明, ξ 在 $\mathscr{C}$ 下的条件期望可看成 ξ 在 $\mathscr{C}$ 的每个原子上的平均. 这一性质称为条件期望的平滑性.

定理 5.8. $\mathbb{E}(\xi|\mathscr{C})$ 在 $\mathscr{C}$ 的每个原子上取常数值. 若 $\mathbb{P}(B)>0$, B 为原子, 则 $\forall\omega\in B$, $\mathbb{E}(\xi|\mathscr{C})(\omega)=\frac{1}{\mathbb{P}(B)}\int_B\xi\,\mathrm{d}\mathbb{P}$.

证明 设 B 为 $\mathscr{C}$ 中的原子, 如 $\exists\omega_1,\omega_2\in B$ 使 $\mathbb{E}(\xi|\mathscr{C})(\omega_1)\neq\mathbb{E}(\xi|\mathscr{C})(\omega_2)$, 则 $\emptyset\neq\{\omega\in B:\mathbb{E}(\xi|\mathscr{C})(\omega)=\mathbb{E}(\xi|\mathscr{C})(\omega_1)\}\subsetneq B$, 与 B 是一原子矛盾!

设 $\mathbb{P}(B)>0$, B 为原子. 由于 $\mathbb{E}(\xi|\mathscr{C})$ 在 B 上取常值, 有

$$\mathbb{E}(\xi|\mathscr{C})\mathbb{P}(B)=\int_B\mathbb{E}(\xi|\mathscr{C})\,\mathrm{d}\mathbb{P}=\int_B\xi\,\mathrm{d}\mathbb{P},$$

从而 $\mathbb{E}(\xi|\mathscr{C})|_B=\frac{1}{\mathbb{P}(B)}\int_B\xi\,\mathrm{d}\mathbb{P}$. □

推论 5.9. 设 $\{B_n\}_{n\geqslant1}\subset\mathscr{A}$ 为 Ω 的可数分割, $\mathscr{C}=\sigma(\{B_n:n\geqslant1\})$, 则 $\mathbb{E}(\xi|\mathscr{C})=\sum\limits_{n=1}^{\infty}\mathbb{E}(\xi|B_n)\mathbf{1}_{B_n}$. 特别地, 当 $\mathscr{C}=\{\phi,\Omega\}$ 时, $\mathbb{E}(\xi|\mathscr{C})=\mathbb{E}\xi$.

命题 5.10. 设 $\mathscr{C}$ 与 $\sigma(\xi)$ 独立, 则 $\mathbb{E}(\xi|\mathscr{C})=\mathbb{E}\xi$; 如 $\mathscr{C}\subset\mathscr{C}'$, 则 $\mathbb{E}(\xi|\mathscr{C})=\mathbb{E}(E(\xi|\mathscr{C}')|\mathscr{C})$.

证明 $\forall B\in\mathscr{C}$, 有 $\mathbf{1}_B$ 与 ξ 独立. 故

$$\int_B\mathbb{E}(\xi|\mathscr{C})\,\mathrm{d}\mathbb{P}=\int_B\xi\,\mathrm{d}\mathbb{P}=\mathbb{E}\mathbf{1}_B\xi=(\mathbb{E}\mathbf{1}_B)\mathbb{E}\xi=\mathbb{P}(B)\mathbb{E}\xi=\int_B\mathbb{E}\xi\,\mathrm{d}\mathbb{P}.$$

由 $B\in\mathscr{C}$ 的任意性知 $\mathbb{E}(\xi|\mathscr{C})=\mathbb{E}\xi$.

设 $\mathscr{C}\subset\mathscr{C}'$, 则 $\forall B\in\mathscr{C}$ 有

$$\int_B\mathbb{E}(\xi|\mathscr{C}')\,\mathrm{d}\mathbb{P}=\mathbb{E}[1_B\mathbb{E}(\xi|\mathscr{C}')]=\mathbb{E}(\mathbb{E}(\xi1_B|\mathscr{C}'))=\mathbb{E}\xi1_B=\int_B\xi\,\mathrm{d}\mathbb{P}.$$

从而 $\mathbb{E}(\xi|\mathscr{C})=\mathbb{E}(\mathbb{E}(\xi|\mathscr{C}')|\mathscr{C})$. □

最后, 我们证明 $\mathbb{E}(\xi|\mathscr{C})$ 是 ξ 在关于 $\mathscr{C}$ 可测函数类中的 L^2 最佳逼近.

命题 5.11 (最佳均方逼近). 设 $\xi\in L^2(\mathbb{P}),\mathscr{C}\subset\mathscr{A}$ 为子 σ 代数. 则 $\mathbb{E}(\xi|\mathscr{C})\in L^2(\mathbb{P}_{\mathscr{C}})$, 且 $\mathbb{E}(\xi|\mathscr{C})$ 是 ξ 在 $L^2(\mathbb{P}_{\mathscr{C}})$ 中最佳均方逼近: $\forall\eta\in L^2(\mathbb{P}_{\mathscr{C}})$, 有

$$\mathbb{E}|\xi-\mathbb{E}(\xi|\mathscr{C})|^2\leqslant\mathbb{E}|\xi-\eta|^2,\quad\mathbb{E}(|\xi-\mathbb{E}(\xi|\mathscr{C})|^2|\mathscr{C})\leqslant\mathbb{E}(|\xi-\eta|^2|\mathscr{C}).$$

证明 只需证后者. 由 Jensen 不等式, $|\mathbb{E}(\xi|\mathscr{C})|^2 \leqslant \mathbb{E}(|\xi|^2|\mathscr{C})$, 故 $\mathbb{E}(\xi|\mathscr{C}) \in L^2(\mathbb{P}_{\mathscr{C}})$. $\forall \eta \in L^2(\mathbb{P}_{\mathscr{C}})$, 有

$$\begin{aligned}\mathbb{E}(|\xi-\eta|^2|\mathscr{C}) = &\mathbb{E}(|\xi-\mathbb{E}(\xi|\mathscr{C})|^2|\mathscr{C}) + \mathbb{E}(|\eta-\mathbb{E}(\xi|\mathscr{C})|^2|\mathscr{C}) \\ &- 2\mathbb{E}((\eta-\mathbb{E}(\xi|\mathscr{C}))(\xi-\mathbb{E}(\xi|\mathscr{C}))|\mathscr{C}).\end{aligned}$$

由于 $\eta-\mathbb{E}(\xi|\mathscr{C}) \in \mathscr{C}$, 有

$$\mathbb{E}((\eta-\mathbb{E}(\xi|\mathscr{C}))(\xi-\mathbb{E}(\xi|\mathscr{C}))|\mathscr{C}) = (\eta-\mathbb{E}(\xi|\mathscr{C}))\mathbb{E}((\xi-\mathbb{E}(\xi|\mathscr{C}))|\mathscr{C}) = 0.$$

故 $\mathbb{E}(|\xi-\eta|^2|\mathscr{C}) \geqslant \mathbb{E}(|\xi-\mathbb{E}(\xi|\mathscr{C})|^2|\mathscr{C})$. □

§5.2　给定函数下的条件期望

设 ξ,η 为两个随机变量, $\mathbb{E}\xi$ 存在, 令 $\mathscr{C} = \sigma(\eta)$. 易见任给 $y \in \mathbb{R}$, $\{\eta = y\}$ 是 $\mathscr{C}$ 的原子, 因而 $\mathbb{E}(\xi|\mathscr{C})$ 在 $\{\eta = y\}$ 上取常值, 从而 $\mathbb{E}(\xi|\mathscr{C})$ 是 η 的函数, 即存在可测函数 $g : \mathbb{R} \to \mathbb{R}$, 使 $\mathbb{E}(\xi|\mathscr{C}) = g(\eta)$. 一般地, 设 $f : (\Omega, \mathscr{A}) \to (\Omega', \mathscr{A}')$ 为可测映射, 令 $\sigma(f) \triangleq f^{-1}(\mathscr{A}')$, 它是 $\mathscr{A}$ 的子 σ 代数.

定理 5.12. 设 $\mathbb{E}\xi$ 存在, f 如上, 则 $\mathbb{E}(\xi|\sigma(f)) = g \circ f$. 其中 $g : \Omega' \to \mathbb{R}$ 为可测函数使得 $\forall A' \in \mathscr{A}', \int_{A'} g \,\mathrm{d}\mathbb{P}_f = \int_{f^{-1}} \xi \,\mathrm{d}\mathbb{P}$. 这里, $\mathbb{P}_f \triangleq \mathbb{P} \circ f^{-1}(A')$ 为 $\mathbb{P}$ 在 $\mathscr{A}'$ 由 f 诱导的概率测度.

证明 由于 $\mathbb{E}(\xi|\sigma(f))$ 为 $\sigma(f)$ 可测的, 则 $\exists g : (\Omega', \mathscr{A}') \to (\mathbb{R}, \mathscr{B})$ 可测, 使 $\mathbb{E}(\xi|\sigma(f)) = g \circ f$. 由积分变换公式与条件期望的定义知, $\forall A' \in \mathscr{A}', \int_{A'} g \,\mathrm{d}\mathbb{P}_f = \int_{f^{-1}(A')} g \circ f \,\mathrm{d}\mathbb{P} = \int_{f^{-1}(A')} \xi \,\mathrm{d}\mathbb{P}$. □

在定理 5.12 中令 $\xi = \mathbf{1}_A, A \in \mathscr{A}$, 我们得到下面推论.

推论 5.13. 设 f 是 $(\Omega, \mathscr{A})$ 到 $(\Omega', \mathscr{A}')$ 的可测映射, 其中 $A \in \mathscr{A}$. 则 $\mathbb{P}(A|\sigma(f)) = g \circ f$, $g : (\Omega', \mathscr{A}') \to (\mathbb{R}, \mathscr{B})$ 可测, 满足

$$\int_{B'} g \,\mathrm{d}\mathbb{P}_f = \mathbb{P}(A \cap f^{-1}(B')), \forall B' \in \mathscr{A}'.$$

在经典概率论中, 我们先定义了在给定事件 B 下的条件概率 $\mathbb{P}(\cdot|B)$, 再将条件期望 $\mathbb{E}(\xi|B)$ 定义为 ξ 关于条件概率的积分. 本节中讨论的给定 σ 代

数下的条件期望与条件概率的定义正好反过来了, 我们先定义条件期望, 再取随机变量为示性函数得到条件概率. 那么我们可否像以前那样将条件概率 $\mathbb{P}(\cdot|\mathscr{C})$ 看成一个概率测度, 并由此给出条件期望 $\mathbb{E}(\xi|\mathscr{C}) = \int_\Omega \xi \,\mathrm{d}\mathbb{P}(\cdot|\mathscr{C})$? 如果给定 $\omega \in \Omega, \mathbb{P}(\cdot|\mathscr{C})(\omega)$ 为概率测度, 则 $\int_\Omega \xi \,\mathrm{d}\mathbb{P}(\cdot|\mathscr{C})(\omega)$ 有意义. 但是, 对于每个事件 A, $\mathbb{P}(A|\mathscr{C})$ 仅是几乎唯一确定的. 为此, 我们需要对每个 A, 在 $\mathbb{P}(A|\mathscr{C})$ 的等价类中取定一个适当的代表 $\mathbb{P}^{\mathscr{C}}(\cdot, A)$ 使得对每个 $\omega \in \Omega$, $\mathbb{P}^{\mathscr{C}}(\cdot, A)$ 为概率测度. 这就是正则条件概率.

§5.3 正则条件概率

定义 5.14 (**正则条件概率**). 设 $(\Omega, \mathscr{A}, \mathbb{P})$ 为概率空间, $\mathscr{C} \subset \mathscr{A}$ 为子 σ 代数. 设 $\mathbb{P}^{\mathscr{C}}$ 为 $\Omega \times \mathscr{A}$ 上的转移概率. 如果 $\forall A \in \mathscr{A}$ 有 $\mathbb{P}^{\mathscr{C}}(\cdot, A) = \mathbb{E}(\mathbf{1}_A|\mathscr{C}) = \mathbb{P}(A|\mathscr{C})$, 则称 $\mathbb{P}^{\mathscr{C}}$ 为 $\mathbb{P}$ 在给定 $\mathscr{C}$ 下关于的正则条件概率.

下面, 我们分别讨论正则条件概率的性质, 存在性及其应用.

§5.3.1 正则条件概率的性质

定理 5.15. 若 $\mathbb{P}^{\mathscr{C}}$ 是给定 $\mathscr{C}$ 下正则条件概率, ξ 是期望存在的随机变量, 则 $\mathbb{E}(\xi|\mathscr{C}) = \int_\Omega \xi \mathbb{P}^{\mathscr{C}}(\cdot, \mathrm{d}\omega)$.

证明 由正则条件概率的定义、可测函数的构造及积分的性质立得本定理. □

定理 5.16. 设 $\mathscr{C} \subset \mathscr{C}' \subset \mathscr{A}$ 是子 σ 代数, $\mathbb{P}^{\mathscr{C}}$ 与 $\mathbb{P}^{\mathscr{C}'}$ 是正则条件概率, $\xi \in \mathscr{C}, \xi' \in \mathscr{C}'$ 使 $\mathbb{E}\xi\xi'$ 及 $\mathbb{E}\xi$ 存在. 则

$$\int (\xi'\xi)(\omega)\mathbb{P}^{\mathscr{C}}(\cdot, \mathrm{d}\omega) = \int \xi'(\omega) \left[\int \xi(\omega)\mathbb{P}^{\mathscr{C}'}(\omega', \mathrm{d}\omega)\right] \mathbb{P}^{\mathscr{C}}(\cdot, \mathrm{d}\omega')).$$

证明 由正则条件概率的定义及条件期望的性质立得. □

§5.3.2 条件分布

设 $\xi_T = \{\xi_t : t \in T\}$ 为 $(\Omega, \mathscr{A}, \mathbb{P})$ 上一族随机变量, 为研究它的分布, 只需考虑 σ 代数 $\sigma(\xi_T) \triangleq \sigma\left(\bigcup_{t \in T} \sigma(\xi_t)\right)$. 事实上, 任给 Borel 可测集合

$B \subset \mathbb{R}^T = \prod_{t \in T} \mathbb{R}$ 有 $\mathbb{P}_{\xi_T}(B) \triangleq \mathbb{P}(\xi_T \in B) = \mathbb{P}(\xi_T^{-1}(B))$. 它确定了 ξ_T 的分布. 因此, 在考虑正则条件概率时, 可把 $\mathscr{A}$ 改成 $\sigma(\xi_T)$.

定义 5.17. 设 $\mathscr{C} \subset \mathscr{A}$ 是子 σ 代数, 如果 $\mathbb{P}^{\mathscr{C}}$ 为 $\Omega \times \sigma(\xi_T)$ 上的转移概率且 $\forall A \in \sigma(\xi_T)$ 有 $\mathbb{P}^{\mathscr{C}}(\cdot, A) = \mathbb{P}(A|\mathscr{C})$, 则称 $\mathbb{P}^{\mathscr{C}}$ 为 ξ_T 在 $\mathscr{C}$ 下的正则条件分布.

通常 $\sigma(\xi_T)$ 与 $\mathscr{C}$ 之间无包含关系. 当 $\mathscr{C} \subset \sigma(\xi_T)$ 时, 正则条件分布 是 $(\Omega, \sigma(\xi_T), \mathbb{P})$ 在 $\mathscr{C}$ 下的正则条件概率. 而当 $\xi, \xi' \in \sigma(\xi_T)$ 时, 定理 5.16 对正则条件分布成立.

我们知道, ξ_T 的分布 $\mathbb{P}_{\xi_T}$ 是乘积欧氏空间 $\mathbb{R}^T$ 上的概率测度. 类似地, 我们引入 ξ_T 在给定 $\mathscr{C}$ 下的混合条件分布, 它是 $(\Omega, \mathscr{C}) \times \mathscr{B}^T$ 上的转移概率.

定义 5.18. 设 $(\Omega, \mathscr{A}, \mathbb{P})$ 是概率空间, ξ_T 是随机变量族, $\mathscr{C} \subset \mathscr{A}$ 为子 σ 代数, $\mathbb{P}^{\mathscr{C}}_{\xi_T}$ 为 $(\Omega, \mathscr{C}) \times \mathscr{B}^T$ 上的转移概率. 如果 $\forall B \in \mathscr{B}^T$ 有 $\mathbb{P}^{\mathscr{C}}_{\xi_T}(\cdot, B) = \mathbb{P}(\xi_T^{-1}(B)|\mathscr{C})$, 则称 $\mathbb{P}^{\mathscr{C}}_{\xi_T}$ 为 ξ_T 在 $\mathscr{C}$ 之下的混合条件分布.

定理 5.19. 设 $g : \mathbb{R}^T \to \mathbb{R}$ 是 Borel 可测的, $\mathbb{E}g(\xi_T)$ 存在. 若 ξ_T 在 $\mathscr{C}$ 下的条件分布 $\mathbb{P}^{\mathscr{C}}$ 与混合条件分布 $\mathbb{P}^{\mathscr{C}}_{\xi_T}$ 存在, 则

$$\mathbb{E}(g(\xi_T)|\mathscr{C}) = \int_\Omega g(\xi_T(\omega))\mathbb{P}^{\mathscr{C}}(\cdot, \mathrm{d}\omega) = \int_{\mathbb{R}^T} g(x_T)\mathbb{P}^{\mathscr{C}}_{\xi_T}(\cdot, \mathrm{d}x_T).$$

证明 当 $g = \mathbf{1}_B, B \in \mathscr{B}^T$ 时显然. 一般情形由此及可测函数的构造与积分、条件期望的性质得到. □

定理 5.20. 如 ξ_T 在 $\mathscr{C}$ 下的条件分布存在, 则其混合条件分布也存在. 若 $\xi_T(\Omega) \in \mathscr{B}^T$, 则逆命题也成立.

证明 如 ξ_T 在 $\mathscr{C}$ 下的条件分布 $\mathbb{P}^{\mathscr{C}}$ 存在, 则有 $\mathbb{P}^{\mathscr{C}}_{\xi_T}(\omega, B) = \mathbb{P}^{\mathscr{C}}(\omega, \xi_T^{-1}(B)), B \in \mathscr{B}^T$. 反之, $\mathbb{P}^{\mathscr{C}}_{\xi_T}$ 存在且 $\xi_T(\Omega) \in \mathscr{B}^T$, 则 $\forall A \in \sigma(\xi_T)$ 存在 $B \in \mathscr{B}^T$ 使 $A = \xi_T^{-1}(B)$, 故 $\xi_T(A) = B \cap \xi_T(\Omega) \in \mathscr{B}^T$. 从而可令 $\mathbb{P}^{\mathscr{C}}(\cdot, A) = \mathbb{P}^{\mathscr{C}}_{\xi_T}(\cdot, \xi_T(A))$. □

§ 5.3.3 存在性

定理 5.21. 设 $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, \mathbb{P})$ 上 n 维随机变量, $\mathscr{C}$ 为 $\mathscr{A}$ 的子 σ 代数, 则 $\mathbb{P}^{\mathscr{C}}_\xi$ 存在. 从而当 $\xi(\Omega) \in \mathscr{B}^n$ 时, ξ 在 $\mathscr{C}$ 下的条件分布存在.

证明 为构造 $\mathbb{P}^{\mathscr{C}}_\xi(\omega, \cdot)$, 需确定相应的概率分布函数. 为此, $\forall r \in \bar{\mathbb{Q}}^n$, 选定一个 $\mathscr{C}$ 可测函数 $F(r)$ 使 $F(r) = \mathbb{P}(\xi < r|\mathscr{C})$. 则 F 满足:

1) $\forall a, b \in \bar{\mathbb{Q}}^n, a \leqslant b$ 有 $\Delta_{a,b}F = \mathbb{P}(\xi \in [a, b)|\mathscr{C})$, a.s.

2) 任给序列 $r_m \to \infty$ 有 $\lim\limits_{m\to\infty} F(r_m) = 1$, a.s.

3) 任给序列 r_m 使其中某个分量趋于 $-\infty$, 有 $\lim\limits_{m\to\infty} F(r_m) = 0$, a.s.

4) $\forall r_0 \in \mathbb{Q}^n$, $\lim\limits_{\mathbb{Q}^n \ni r \uparrow r_0} F(r) = F(r_0)$, a.s.

为确定 $\mathbb{P}_\xi^{\mathscr{C}}$, 我们需要改造 F 使上述诸条点点成立. 由于 $\bar{\mathbb{Q}}^n$ 可数, 由 1) 存在零测集 N_1 使 $\forall \omega \notin N_1$ 有 $\Delta_{a,b}F(\omega) \geqslant 0, \forall a, b \in \bar{\mathbb{Q}}^n, a \leqslant b$. 此外, 由 2)与 3) 知存在零测集 N_2 使 $\forall \omega \notin N_2$ 有 $\lim\limits_{m\to\infty} F(m, \cdots, m)(\omega) = 1$ 且 $\lim\limits_{m\to\infty} F(r_m^i)(\omega) = 0, 1 \leqslant i \leqslant n$, 其中 $r_m^i \in \bar{\mathbb{Q}}^n$ 使得第 i 个分量为 $-m$, 其余分量为 ∞. 再由 4), 存在零测集 N_3 使 $\forall \omega \notin N_3, \forall r_0 \in \mathbb{Q}^n$ 有 $\lim\limits_{m\to\infty} F(r_0 - \mathbf{1}/m)(\omega) = F_{r_0}(\omega)$. 由于 $\Delta_{a,b}F \geqslant 0 (\forall a \leqslant b)$ 蕴含 F 的单调性, 易见当 $\omega \notin N \triangleq N_1 \cup N_2 \cup N_3$ 时, $F(\cdot)(\omega)$ 作为 $\mathbb{Q}^n$ 上的函数满足差分非负性、正则性与左连续性.

令

$$F^{\mathscr{C}}(\omega, r) = \begin{cases} F(r)(\omega), & \omega \in N; \\ 1_{(0,\infty)(r)}, & \omega \in N^c, r \in \mathbb{Q}^n. \end{cases}$$

$\forall x \in \mathbb{R}^n$, 令 $F^{\mathscr{C}}(\omega, x) = \lim\limits_{r \uparrow x} F^{\mathscr{C}}(\omega, r)$. 则 $\forall \omega \in \Omega, F^{\mathscr{C}}(\omega, \cdot)$ 具有概率分布函数的特征, 因而诱导 $(\mathbb{R}, \mathscr{B})$ 上唯一概率测度 $\mathbb{P}_\xi^{\mathscr{C}}(\omega, \cdot)$ 使 $\mathbb{P}_\xi^{\mathscr{C}}(\omega, (-\infty, x)) = F^{\mathscr{C}}(\omega, x), \omega \in \Omega, x \in \mathbb{R}^n$.

为证明 $\mathbb{P}_\xi^{\mathscr{C}}$ 为混合条件分布, 令

$$\Pi = \{(-\infty, r) : r \in \mathbb{Q}^n\}, \Lambda = \left\{B \in \mathscr{B}^n : \mathbb{P}_\xi^{\mathscr{C}}(\cdot, B) = \mathbb{P}(\xi \in B|\mathscr{C})\right\}.$$

则 Π 为 π 系, $\Lambda \supset \Pi$ 且 Λ 是 λ 系. 从而 $\Lambda = \mathscr{B}^n$. □

定理 5.22. 设 $(\Omega, \mathscr{A}, \mathbb{P}) = (\mathbb{R}^n, \mathscr{B}^n, \mathbb{P})$, 则任给子 σ 代数 $\mathscr{C}$, 正则条件概率 $\mathbb{P}^{\mathscr{C}}$ 总存在.

证明 令 $\xi(x) = x, x \in \mathbb{R}^n$. 则 $\sigma(\xi) = \mathscr{B}^n$, $\mathbb{P}_\xi^{\mathscr{C}}$ 即是 $\mathbb{P}$ 在给定 $\mathscr{C}$ 下的正则条件概率. □

前面我们了解到, 可以使用转移概率构造乘积空间上的概率测度. 作为正则条件概率的应用, 我们证明 $\mathbb{R}^n$ 上的任何概率测度均可通过转移概率来构造.

定理 5.23. 设 $\mathbb{P}$ 是 $(\mathbb{R}^n, \mathscr{B}^n)$ 上的概率, 则 $\exists \mathscr{B}$ 上的概率 $\mathbb{P}_1$ 及 $\mathbb{R}^{k-1} \times \mathscr{B}$ 上的转移概率 $\mathbb{P}_k(x_1, x_2, \cdots, \mathrm{d}x_k), k = 2, \cdots, n$ 使

$$\mathbb{P}(B) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \mathbf{1}_B(x_1, \cdots, x_n) \mathbb{P}_n(x_1, \cdots, x_{n-1}, \mathrm{d}x_n) \cdots \mathbb{P}_1(\mathrm{d}x_1), B \in \mathbb{R}^n.$$

证明 当 $n = 2$ 时, 令 $\mathbb{P}_2(x_1, B_2) = \mathbb{P}^{\mathscr{C}}((x_1, 0), \mathbb{R} \times B_2), B_2 \in \mathscr{B}, \mathscr{C} = \mathscr{B} \times \mathbb{R}, \mathbb{P}_1(B_1) = \mathbb{P}(B_1 \times \mathbb{R})$. 则 $\tilde{\mathbb{P}}(B) \triangleq \int_{\mathbb{R}} \mathrm{d}\mathbb{P}_1(x_1) \int_{\mathbb{R}} \mathbf{1}_B \mathbb{P}_2(x_1, \mathrm{d}x_2)$ 是 $(\mathbb{R}^2, \mathscr{B}^2)$ 上的概率测度且满足 $\tilde{\mathbb{P}}(B_1 \times B_2) = \mathbb{P}(B_1 \times B_2), B_1, B_2 \in \mathscr{B}$. 由测度扩张定理知 $\tilde{\mathbb{P}} = \mathbb{P}$. 故定理对 $n = 2$ 成立. 对一般的 n 可使用归纳法证明.

设定理对 n 成立, 往证 $n + 1$ 的情形. 设 $\mathbb{P}^{(n-1)}(B^{(n-1)}) = \mathbb{P}(B^{(n-1)} \times \mathbb{R}), B^{(n-1)} \in \mathscr{B}^{(n-1)}$ 是前 $n - 1$ 个分量的联合分布. 令 $\mathscr{C} = \mathscr{B}^{(n-1)} \times \mathbb{R}, \mathbb{P}_n = \mathbb{P}^{\mathscr{C}}((x_1, \cdots, x_{n-1}, 0), \mathbb{R}^{(n-1)} \times B)$. 如同 $n = 2$ 情形的证明, 有 $\mathbb{P}(\mathrm{d}x_1, \cdots, \mathrm{d}x_n) = \mathbb{P}_n(x_1, \cdots, x_{n-1}, \mathrm{d}x_n)\mathbb{P}^{(n-1)}(\mathrm{d}x_1, \cdots, \mathrm{d}x_{n-1})$. 由于 $\mathbb{P}^{(n-1)}$ 为 $\mathbb{R}^{n-1}$ 上的概率, 由归纳假设, $\mathbb{P}^{(n-1)}(\mathrm{d}x_1, \cdots, \mathrm{d}x_{n-1})$ 可表为 $\mathbb{P}_{(n-1)}(x_1, \cdots, x_{n-2}, \mathrm{d}x_{n-1})\mathbb{P}_{(n-2)}(x_1, \cdots, x_{n-3}, \mathrm{d}x_{n-2}) \cdots \mathbb{P}_1(\mathrm{d}x_1)$. 从而定理得证. □

§5.4 Kolmogorov 和谐定理

我们前面已经了解到, 有限乘积空间上的概率测度可由相关边缘概率测度和若干转移概率构造出来. 那么, 我们如何构造无穷维乘积空间上的概率测度呢? 一个基本的想法是使用有限乘积空间上的概率测度来生成无穷乘积空间上的概率测度. 为此, 我们需引入和谐性的概念. 如 S 为 T 的有限子集, 则记成 $S \subset\subset T$.

定义 5.24. 设 T 为一个无穷指标集, $\forall t \in T, (\Omega_t, \mathscr{A}_t)$ 为可测空间. $\forall T' \subset T$, 记 $\Omega^{T'} = \prod\limits_{t \in T'} \Omega_t$, $\mathscr{A}^{T'} = \prod\limits_{t \in T'} \mathscr{A}_t$. 设 $\forall S \subset\subset T$, $\mathscr{A}^S$ 上给定一个概率 $\mathbb{P}^S$. 如果 $\forall S \subset S' \subset\subset T$, 有 $\mathbb{P}^S(A^S) = \mathbb{P}^{S'}\left(A^S \times \prod\limits_{t \in S' \setminus S} \Omega_t\right), A^S \in \mathscr{A}^S$, 则称概率测度族 $\left\{\mathbb{P}^S : S \subset\subset T\right\}$ 是和谐的.

定理 5.25 (Kolmogorov 和谐定理). 设 $\Omega_t = \mathbb{R}_t = \mathbb{R}, t \in T, \left\{\mathbb{P}^S : S \subset\subset T\right\}$ 是和谐概率测度族. 则存在 $(\mathbb{R}^T, \mathscr{B}^T)$ 上唯一的概率测度 $\mathbb{P}$ 使 $\forall S \subset\subset T$ 及 $B^S \in \mathscr{B}^S$ 有 $\mathbb{P}(B^S \times \mathbb{R}^{T \setminus S}) = \mathbb{P}^S(B^S)$.

证明 (1) 定义的合理性由概率测度族的和谐性保证.

(2) 设 T 可数, 无妨设 $T=\mathbb{N}$. 由定理 5.23 知, $\exists\mathbb{R}$ 上概率 $\mathbb{P}_1$ 及转移概率 $\{\mathbb{P}_n(x_1,...,x_{n-1},\mathrm{d}x_n):n\geqslant 2\}$ 使 $\mathbb{P}^{\{1,2,\cdots,n\}}=\mathbb{P}_n\cdot\mathbb{P}_{n-1}\cdots\mathbb{P}_1,\forall n\geqslant 2$. 由 Tulcea 定理, 该概率测度存在唯一.

(3) 对于不可数指标集 T, 设 $\mathscr{D}$ 为可测柱集全体, 它是集代数. 易见 $\mathbb{P}$ 在 $\mathscr{D}$ 上为有限可加概率测度. 为证其 σ 可加性, 仅需证明在 $\emptyset$ 处连续. 设 $\{A_n\}_{n\geqslant 1}\subset\mathscr{D},A_n\downarrow\emptyset$. $\forall n\geqslant 1,\exists T_n\subset\subset T$ 使 $A_n=A^{T_n}\times\mathbb{R}^{T\setminus T_n},A^{T_n}\in\mathscr{B}^{T_n}$. 令 $T_\infty=\bigcup\limits_{n=1}^{\infty}T_n$, 则 T_∞ 可数. 由 T 可数时的结论知, $\mathbb{P}$ 限于 $\mathscr{D}^{T_\infty}\times\mathbb{R}^{T\setminus T_\infty}$ 可唯一扩张成 $\mathbb{R}^T$ 上的概率测度, 其中 $\mathscr{D}^{T_\infty}$ 为 $\mathbb{R}^{T_\infty}$ 上可测柱集全体. 从而由概率的连续性知 $\mathbb{P}(A_n)\downarrow 0(n\to\infty)$. $\square$

注 5.26. 定理 5.25 的证明主要用到定理 5.23 和 Tulcea 定理. 虽然在 Tulcea 定理中没有对 Ω_t 作限制, 但定理 5.23 仅处理欧氏可加的情形. 因此, 要想将定理 5.23 中的 $\mathbb{R}$ 换成一般的 Ω_t, 需证明任何有限乘积空间 Ω^S 上的概率测度均可如定理 5.23 中那样由一个边缘概率测度和若干转移概率表示出来. 为此, 需推广正则条件概率存在性定理 (即定理 5.22 或定理 5.21) . 定理 5.21 的证明依赖于分布函数的特征与有理数集的可数性与稠性. 由于度量空间 Ω 上的概率测度均可由其在球上的值确定, 如 Ω 存在可数稠子集 Ω_0, 概率测度可由它在可数集类 $\{B(x,r):x\in\Omega_0,r\in\mathbb{Q}_+\}$ 上的值确定. 据此, 可证明完备可分度量空间 (Polish 空间) 上的概率测度必存在正则条件概率, 进而将定理 5.23 和定理 5.24 推广到 Polish 空间上.

§5.5　补充与习题

1. 证明命题 5.2.

2. 设 $\mathscr{C}$ 为 $\mathscr{A}$ 的子 σ 代数, $\varphi(B)=\int_B \xi\,\mathrm{d}\mathbb{P}, B\in\mathscr{C}$, 则 φ 是 $\mathscr{C}$ 上 σ 可加集函数, $\varphi\ll\mathbb{P}|_{\mathscr{C}}$. 从而存在 $\mathbb{E}[\xi|\mathscr{C}]=\frac{\mathrm{d}\varphi}{\mathrm{d}\mathbb{P}_{\mathscr{C}}},\mathbb{P}_{\mathscr{C}}\text{-}a.e.$

3. 设随机变量列 $0\leqslant\xi_n\uparrow\xi$. 证明 $\mathbb{E}(\xi_n|\mathscr{C})\uparrow\mathbb{E}(\xi|\mathscr{C})$.

4. 证明性质 5.5.

5. 证明 Hölder 不等式 $\mathbb{E}(\xi\eta|\mathscr{C})\leqslant\mathbb{E}(|\xi|^p|\mathscr{C})^{1/p}\mathbb{E}(|\eta|^q|\mathscr{C})^{1/q}, p>1, \frac{1}{p}+\frac{1}{q}=1$.

6. 设 ξ 复随机变量, 证明 $|\mathbb{E}[\xi|\mathscr{C}]|\leqslant\mathbb{E}[|\xi||\mathscr{C}]$.

7. 叙述并证明关于条件期望的 Jensen 不等式和 Minkowski 不等式.

8. 证明推论 5.9

9. 设 $\mathscr{C}_1,\mathscr{C}_2$ 为 $\mathscr{A}$ 的两个子 σ 代数. 举例说明 $\mathbb{E}(\xi|\mathscr{C}_1\cap\mathscr{C}_2)\notin\mathbb{E}(\mathbb{E}(\xi|\mathscr{C}_1)|\mathscr{C}_2)$.

10. 记全体有理数为 $x_1,x_2,\cdots$, 令
$$F(x)=\sum_{n=1}^{\infty}2^{-n}\mathbf{1}_{(x_n,\infty)}(x)$$
则 F 是 $\mathbb{R}$ 上的分布函数.

11. (鞅) 设 $\mathscr{A}_n$ 是一列单调上升的子 σ 代数, 若随机变量列 ξ_n 满足
$$\mathbb{E}(\xi_{n+1}|\mathscr{A}_n)=\xi_n, n\geqslant 1,$$
则称之为鞅序列. 证明 $\xi_n=\mathbb{E}(\xi|\mathscr{A}_n)$ 是鞅序列.

12. (马氏过程) 设 ξ_n 为随机变量序列, 令 $\mathscr{A}_n=\sigma(\{\xi_m:m\leqslant n\})$. 若
$$\mathbb{E}(\xi_{n+1}|\mathscr{A}_n)=\mathbb{E}(\xi_{n+1}|\xi_n), n\geqslant 1,$$

则 ξ_n 称为马氏过程. 设 $\{X_n\}$ 是独立的随机变量序列, 证明 $\xi_n = \sum_{m=1}^{n} X_m$ 是马氏过程.

13. 设 $\mathscr{A}_n = \sigma(\{\xi_m : m \leqslant n\}), \mathscr{A}^n \sigma(\{\xi_m : m \geqslant n\})$, 则 ξ_n 为马氏过程当且仅当下列条件之一成立.

    (a) $\mathbb{E}(\xi_m|\mathscr{A}_n) = \mathbb{E}(\xi_{n+1}|\xi_n), m \geqslant n \geqslant 1$,

    (b) $\mathbb{E}(\eta|\mathscr{A}_n) = \mathbb{E}(\eta|\xi_n), \eta \in \mathscr{A}^n, n \geqslant 1$,

    (c) $\forall \eta \in \mathscr{A}_n, \zeta \in \mathscr{A}^n$ 使得 $\eta, \zeta, \eta\zeta$ 均可积, 则 $\mathbb{E}[\eta\zeta|\xi_n] = \mathbb{E}[\eta|\xi_n]\mathbb{E}[\zeta|\xi_n]$, $n \geqslant 1$.

14. 设矩阵 $P = (p_{ij})_{i,j=0}^{\infty}$ 满足 $p_{ij} \geqslant 0, \sum\limits_{j=0}^{\infty} p_{ij} = 1$. 这样的矩阵称为随机矩阵. 构造概率空间 $(\Omega, \mathscr{A}, \mathbb{P})$ 及其上的随机变量序列 $\{\xi_n\}_{n\geqslant 0}$ 使得 $\mathbb{P}(\xi_{n+1} = j|\xi_n = i) = p_{ij}, n \geqslant 0, i, j \geqslant 0$.

15. 证明若 $\mathbb{E}(\xi|\mathscr{C}) = \eta$ 且 $\mathbb{E}\xi^2 = \mathbb{E}\eta^2 < \infty$, 则 $\xi = \eta, \text{a.s.}$

16. 设 $\xi \in L^1(\Omega, \mathscr{A}, \mathbb{P})$. 证明随机变量族 $\big\{\mathbb{E}(\xi|\mathscr{C}) : \mathscr{C}$ 为 $\mathscr{A}$ 的子 σ 代数$\big\}$ 一致可积.

17. 设 ξ, η 是独立同分布的可积随机变量, 证明 $\mathbb{E}(\xi|\xi+\eta) = (\xi+\eta)/2$.

18. 设 $\xi_1, \xi_2, \cdots$ 独立同分布且期望有限, $S_n = \sum\limits_{k=1}^{n} \xi_k$. τ 为取正整数值期望有限的随机变量满足 $\{\tau = n\} \in \sigma\{\xi_1, \cdots, \xi_n\}$.

    (a) 证明 $\mathbb{E}S_n = \mathbb{E}\xi_1 \mathbb{E}\tau$.

    (b) 设 ξ_1 以 p (或 $1-p$) 的概率取值 1 (或 -1), τ 为 S_n 等于 $-N$或 M 最小的 n , N, M 均为正整数. 请计算 $\mathbb{E}\tau$.

19. 设 σ 代数 $\mathscr{G}$ 由 π 系 C 生成. 证明 $f \in \mathscr{G}$ 为事件 A关于 $\mathscr{G}$ 的条件期望当且仅当

$$\int_B f \, \mathrm{d}\mathbb{P} = \mathbb{P}(A \cap B), \forall B \in \mathscr{C}.$$

20. 考虑概率空间 $(\Omega, \mathscr{A}, \mathbb{P})$ 到概率空间 $(\Omega', \mathscr{A}', \mathbb{P}')$ 的可测映射 T 使得 $\mathbb{P}' = \mathbb{P} \circ T^{-1}$. 令 $\mathscr{C}'$ 为 $\mathscr{A}'$ 的子 σ 代数, $\mathscr{C} = \left\{T^{-1}A : A \in \mathscr{C}\right\}$. $\forall A' \in \mathscr{A}'$, 证明 $\mathbb{P}(t^{-1}A'|\mathscr{C})(\omega) = \mathbb{P}'(A'|\mathscr{C}')(T\omega), \mathbb{P} - \text{a.s.}\omega$.

21. (a) 设事件 A 满足 $\mathbb{P}(A) > 0$, 定义概率 $\mathbb{Q}(B) = \mathbb{P}(B|A)$. 证明对 $B \in \mathscr{A}$ 及子 σ 代数 $\mathscr{C}$,

$$\mathbb{Q}(B|\mathscr{C}) = \frac{\mathbb{P}(A \cap B|\mathscr{C})}{\mathbb{P}(A|\mathscr{C})}, \mathbb{Q} - \text{a.s.}$$

(b) 设 $\mathscr{D}$ 由分划 $A_1, A_2, \cdots$ 生成的 σ 代数, $\mathscr{C} \vee \mathscr{D} \triangleq \sigma(\mathscr{C} \cup \mathscr{D})$. 证明对 $B \in \mathscr{A}$

$$\mathbb{P}(B|\mathscr{C} \vee \mathscr{D}) = \sum_{n=1}^{\infty} \mathbf{1}_{A_n} \frac{\mathbb{P}(B \cap A_n|\mathscr{C})}{\mathbb{P}(A_n|\mathscr{C})}.$$

22. (a) 设 $\mathscr{C}_1 \subset \mathscr{C}_2$ 为子 σ 代数, $\mathbb{E}\xi^2 < \infty$. 证明 $\mathbb{E}((\xi - \mathbb{E}(\xi|\mathscr{C}_1))^2) \leqslant \mathbb{E}((\xi - \mathbb{E}(\xi|\mathscr{C}_2))^2)$.

(b) 令 $\mathrm{Var}(\xi|\mathscr{C}) = \mathbb{E}((\xi - \mathbb{E}(\xi|\mathscr{C})^2)|\mathscr{C})$. 证明

$$\mathrm{Var}(\xi) = \mathbb{E}(\mathrm{Var}(\xi|\mathscr{C})) + \mathrm{Var}(\mathbb{E}(\xi|\mathscr{C})).$$

23. 设 $\mathscr{C}_i, i = 1, 2, 3$ 为子 σ 代数, $\mathscr{C}_{ij} = \sigma(\mathscr{C}_i \cup \mathscr{C}_j), 1 \leqslant i, j \leqslant 3$. 证明下述条件等价:

(a) $\mathbb{P}(A_3|\mathscr{C}_{12}) = \mathbb{P}(A_3|\mathscr{C}_2), \forall A_3 \in \mathscr{C}_3$;

(b) $\mathbb{P}(A_1 \cap A_3|\mathscr{C}_2) = \mathbb{P}(A_1|\mathscr{C}_2)\mathbb{P}(A_3|\mathscr{C}_2), \forall A_1 \in \mathscr{C}_1, A_3 \in \mathscr{C}_3$;

(c) $\mathbb{P}(A_1|\mathscr{C}_{23}) = \mathbb{P}(A_1|\mathscr{C}_2), \forall A_1 \in \mathscr{C}_1$.

24. 设 $\mathbb{P}$ 为 $(\Omega, \mathscr{A})$ 完备可分距离空间上的概率, 则任给子 σ 代数 $\mathscr{C}$, 正则条件概率 $\mathbb{P}^{\mathscr{C}}$ 存在.

第六章　特征函数与测度弱收敛

在学习概率论时我们知道随机变量的特征函数可唯一确定它的分布, 且与分布函数相比有更好的分析性质. 特别地, 随机变量的各阶矩可由其特征函数在 0 处的相应阶导数得到. 本章研究一般有限测度的特征函数, 测度的弱收敛等价于相应特征函数的逐点收敛, 当测度列为随机变量列的分布律时, 它们还等价于随机变量列的依分布收敛. 我们还给出判别复函数为特征函数的判别法则, 并讨论一般度量空间上有限测度的弱收敛与判别法则.

§6.1　有限测度的特征函数

§6.1.1　定义与性质

定义 6.1. 设 μ 是 $(\mathbb{R}^n, \mathscr{B}^n)$ 上的有限测度, 则称函数

$$f(t) = \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\langle t,x\rangle} \mu(\,\mathrm{d}x), \quad t \in \mathbb{R}^n$$

为 μ 的特征函数, 或 μ 的 Fourier-Stieltjes 变换. 这里 $\langle \cdot, \cdot \rangle$ 是欧氏空间中的内积.

由于 μ 有限, 且被积函数有界, 故特征函数总存在. 显然特征函数 f 具有如下性质:

(1) $f(0) = \mu(\mathbb{R}^n)$;　(2) $|f(t)| \leqslant f(0)$;　(3) $\bar{f}(t) = f(-t)$.

命题 6.2. 设 μ_k 为 $\mathbb{R}^{m_k}$ 上的有限测度, f_k 为其特征函数, $k = 1, 2, \cdots, n$. 则乘积测度 $\mu \triangleq \mu_1 \times \mu_2 \times \cdots \times \mu_n$ 的特征函数为

$$f(t) = \prod_{k=1}^{n} f_k(t^{(m_k)}), \ t = (t^{(m_1)}, \cdots, t^{(m_n)}) \in \mathbb{R}^{m_1+\cdots+m_n}.$$

命题 6.3 (**增量不等式**). 设 f 为 $\mathbb{R}^n$ 上有限测度 μ 的特征函数. 则 $|f(t)-f(t+h)|^2 \leqslant 2f(0)[f(0)-\mathrm{Re}f(h)], t,h\in\mathbb{R}^n$, 从而 f 一致连续.

证明 由于 $f(0)=\mu(\mathbb{R}^n)$, 由 Schwarz 不等式

$$
\begin{aligned}
|f(t)-f(t+h)|^2 &\leqslant f(0)\int_{\mathbb{R}^n}|\mathrm{e}^{\mathrm{i}\langle t,x\rangle}-\mathrm{e}^{\mathrm{i}\langle t+h,x\rangle}|^2\mu(\mathrm{d}x)\\
&\leqslant f(0)\int_{\mathbb{R}^n}|\mathrm{e}^{\mathrm{i}\langle h,x\rangle}-1|^2\mu(\mathrm{d}x)\\
&=2f(0)\int_{\mathbb{R}^n}(1-\cos\langle h,x\rangle)\mu(\mathrm{d}x)\\
&=2f(0)(f(0)-\mathrm{Re}f(h)). \quad \square
\end{aligned}
$$

命题 6.4. 设 μ 为 $(\mathbb{R},\mathscr{B})$ 上的有限测度, $n\geqslant 1$.

(1) 如 $f^{(2n)}(0)$ 存在且有限, 则 $\forall r\in[0,2n]$, 有

$$\beta_r \triangleq \int_{\mathbb{R}}|x|^r\mu(\mathrm{d}x)<\infty.$$

(2) 若 $\int_{\mathbb{R}}|x|^n\mu(\mathrm{d}x)<\infty$, 则 $\forall 0\leqslant k\leqslant n$ 有

$$f^{(k)}(t)=\mathrm{i}^k\int_{\mathbb{R}}x^k\mathrm{e}^{\mathrm{i}tx}\mu(\mathrm{d}x),\quad t\in\mathbb{R},$$

从而

$$\int_{\mathbb{R}}x^k\mu(\mathrm{d}x)=\mathrm{i}^{-k}f^{(k)}(0).$$

§ 6.1.2 逆转公式与唯一性定理

一个有限测度可否由其特征函数唯一确定呢, 根据下面的定理, 回答是肯定的.

定理 6.5. 设 f 为有限测度 μ 的特征函数, 若 $a,b\in\mathbb{R}^n$ 且 $\mu(\partial[a,b])=\mu([a,b]\backslash(a,b))=0$. 则

$$\mu([a,b))=\lim_{T\to\infty}\frac{1}{(2\pi)^n}\int_{-T}^{T}\cdots\int_{-T}^{T}\prod_{k=1}^{n}\frac{\mathrm{e}^{-\mathrm{i}t_ka_k}-\mathrm{e}^{-\mathrm{i}t_kb_k}}{\mathrm{i}t_k}f(t_1,\cdots,t_n)\,\mathrm{d}t_1\cdots\mathrm{d}t_n.$$

证明 右端之被积函数连续有界, 因而积分存在有限, 记成 $I(T)$. 则由 f 的定义及 Fubini 定理,

$$\begin{aligned}
I(T) &= \int_{\mathbb{R}^n} \mu(\,\mathrm{d}x) \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{k=1}^{n} \frac{\mathrm{e}^{-\mathrm{i}t_k a_k} - \mathrm{e}^{-\mathrm{i}t_k b_k}}{\mathrm{i}t_k} \mathrm{e}^{\mathrm{i}\sum\limits_{k=1}^{n} t_k x_k} \,\mathrm{d}t_1 \cdots \mathrm{d}t_n \\
&= \int_{\mathbb{R}^n} \left(\prod_{k=1}^{n} \int_{-T}^{T} \frac{\mathrm{e}^{-\mathrm{i}t_k a_k} - \mathrm{e}^{-\mathrm{i}t_k b_k}}{\mathrm{i}t_k} \mathrm{e}^{\mathrm{i}t_k x_k} \,\mathrm{d}t_k \right) \mu(\,\mathrm{d}x) \\
&= 2^n \int_{\mathbb{R}^n} \left(\prod_{k=1}^{n} \int_{0}^{T} \frac{\sin t_k(x_k - a_k) - \sin t_k(x_k - b_k)}{t_k} \,\mathrm{d}t_k \right) \mu(\,\mathrm{d}x) \\
&= 2^n \int_{\mathbb{R}^n} \left(\prod_{k=1}^{n} \int_{T(x_k - b_k)}^{T(x_k - a_k)} \frac{\sin t}{t} \,\mathrm{d}t \right) \mu(\,\mathrm{d}x).
\end{aligned}$$

由于 $\int_0^x \frac{\sin t}{t}\,\mathrm{d}t$ 关于 x 有界且 $\int_{-\infty}^{\infty} \frac{\sin t}{t}\,\mathrm{d}t = \pi$, 由控制收敛定理知

$$\lim_{T\to\infty} I(T) = (2\pi)^n \mu((a,b)) = (2\pi)^n \mu([a,b)). \quad \square$$

上面的定理告诉我们, 如果 $\forall a \leqslant b$ 且 $\mu(\partial[a,b]) = 0$, 则 $\mu([a,b))$ 由特征函数唯一确定. 直观上, 如果我们证明满足 $\mu(\partial[a,b]) = 0$ 的区间足够多, 便可以由特征函数唯一确定测度 μ.

定义 6.6. 如果 $[a,b]$ 有界且 $\mu(\partial[a,b]) = 0$, 则称 $[a,b]$ 为 μ 的一个连续区间.

为证 μ 由其特征函数唯一确定, 只需证 μ 由其在连续区间上的值唯一确定.

引理 6.7. 设 μ 为 $\mathbb{R}^n$ 上有限测度, 令

$$D(\mu) = \{a \in \mathbb{R} : \exists k \in \{1, \cdots, n\} \text{ 使 } \mu(\{x : x_k = a\}) > 0\},$$

则 $D(\mu)$ 至多可数.

证明 令

$$D_{m,k}(\mu) = \left\{ a \in \mathbb{R} : \mu(\{x : x_k = a\}) \geqslant \frac{1}{m} \right\}, \quad m \geqslant 1, 1 \leqslant k \leqslant n,$$

则 $D(\mu) = \bigcup\limits_{k,m} D_{m,k}(\mu)$. 由于 μ 有限, 每个 $D_{m,k}(\mu)$ 为有限集, 从而 $D(\mu)$ 至多可数. $\square$

引理 6.8. 令 $C(\mu)=\mathbb{R}\backslash D(\mu)$. $\forall a,b\in\mathbb{R}^n$, 如 $a_k,b_k\in C(\mu)$, 则 $[a,b]$ 是 μ 的连续区间.

证明 只需注意

$$\partial[a,b]\subset\bigcup_{k=1}^n\{x_k=a_k \text{ 或 } b_k\}$$

为 μ 零测集. □

命题 6.9. 设 μ_1,μ_2 为 $\mathbb{R}^n$ 上两个有限测度, 如果 μ_1 与 μ_2 在共同的连续区间上取值相同, 则 $\mu_1=\mu_2$. 从而 $\mathbb{R}^n$ 上有限测度与其特征函数相互唯一决定.

证明 由定理 6.5, 只需证第一个论断. 由于 $D(\mu_1)\cup D(\mu_2)$ 至多可数, $C\triangleq C(\mu_1)\cap C(\mu_2)$ 在 $\mathbb{R}$ 中稠密. $\forall[a,b)\subset\mathbb{R}^n,\exists\{b_n^{(m)}\},\{a_n^{(m)}\}\subset C$ 使 $a_n^{(m)}\uparrow a,b_n^{(m)}\uparrow b$. 由测度的连续性及 C 的定义知

$$\begin{aligned}\mu_1([a,b))&=\lim_{m\uparrow\infty}\mu_1([a^{(m)},b))=\lim_{m\uparrow\infty}\lim_{m'\uparrow\infty}\mu_1([a^{(m)},b^{(m')}))\\&=\lim_{m\uparrow\infty}\lim_{m'\uparrow\infty}\mu_2([a^{(m)},b^{(m')}))=\mu_2([a,b)).\quad\square\end{aligned}$$

最后讨论有限测度与分布函数的关系. 由于分布函数加上常数后所诱导的测度不变, 为确定有限测度的分布函数, 我们需将其标准化. 设 F 为 $\mathbb{R}^n$ 上一个非负分布函数, 使得当序列 $\{x_m\}\subset\mathbb{R}^n$ 的某分量趋于 $-\infty$ 时有 $F(x_m)\to 0$, 则称之为一个标准分布函数.

命题 6.10. $\mathbb{R}^n$ 上的有限测度与诱导它的标准分布函数相互唯一确定.

证明 给定有限测度 μ , 令 $F(x)=\mu((-\infty,x))$, 则 F 为标准分布函数, 诱导 μ. 反之, 设 F 为诱导 μ 的标准分布函数, 则 $\forall a<b$ 有 $\Delta_{a,b}F=\mu([a,b))$. 令 $a\to-\infty$ 得 $F(b)=\mu((-\infty,b))$. 这里 $-\infty$ 指每个分量均为 $-\infty$ 的向量. □

§6.2 测度的弱收敛

§ 6.2.1 定义与等价定义

设 (E,ρ) 为度量空间, $\mathscr{B}$ 为 Borel σ 代数, $\mathfrak{M}$ 为 $(E,\mathscr{B})$ 有限测度全体.

引理 6.11 (正则性). 设 $\mu \in \mathfrak{M}$, 则 $\forall A \in \mathscr{B}$, 有

$$\mu(A) = \inf_{G \supset A, G \text{ 开}} \mu(G) = \sup_{C \subset A, C \text{ 闭}} \mu(C).$$

证明 为使用单调类定理, 令 $\mathscr{C}$ 为满足要证等式的 $\mathscr{B}$ 中所有集合 A 所组成的集类. 只需证明 (1) $\mathscr{C}$ 包含全体开集, 它是 π 系; (2) $\mathscr{C}$ 为 λ 系.

先证 (1). 设 A 为开集, 则显然第一个等式成立. 如 $A = E$, 则第二个等式也成立. 设 $A^c \neq \emptyset$, 则它是非空闭集. 由 ρ 的三角不等式知, 到 A^c 的距离函数 $d(\cdot, A^c) \triangleq \inf_{y \in A^c} \rho(\cdot, y)$ 为 Lipschitz 函数. 事实上, 设 $x, y \in E$ 使 $d(x, A^c) \geqslant d(y, A^c)$, 则 $\forall \varepsilon > 0, \exists y_\varepsilon \in A^c$ 使 $d(y, A^c) \geqslant \rho(y, y_\varepsilon) - \varepsilon$. 从而由三角不等式知

$$|d(x,A^c) - d(y,A^c)| = d(x,A^c) - d(y,A^c) \leqslant \rho(x,y_\varepsilon) - \rho(y,y_\varepsilon) + \varepsilon \leqslant \rho(x,y) + \varepsilon.$$

令 $\varepsilon \downarrow 0$ 得 $|d(x, A^c) - d(y, A^c)| \leqslant \rho(x, y)$.

令 $C_n = \left\{x \in E : d(x, A^c) \geqslant \frac{1}{n}\right\}$. 则 C_n 为闭集, $C_n \subset A$. 而 $\forall x \in A$, 由于 A 为开集, $\exists n \geqslant 1$ 使 $B(x, \frac{1}{n}) \subset A$. 由此, $d(x, A^c) \geqslant \frac{1}{n}$, 即 $x \in C_n$. 从而 $C_n \uparrow A (n \to \infty)$. 由测度的连续性知 $\lim_{n\to\infty} \mu(C_n) = \mu(A)$. 因此第二个等式也成立. 故 $A \in \mathscr{C}$.

再证 (2). 只需证 $\mathscr{C}$ 为单调类, 且对真差封闭. 设 $\{A_n\}_{n \geqslant 1} \subset \mathscr{C}, A_n \uparrow A\,(n \to \infty)$. 对每个 A_n, 存在开集 $G_n \supset A_n$ 使 $|\mu(G_n) - \mu(A_n)| \leqslant 2^{-n}$; 也存在闭集 $C_n \subset A_n$ 使 $|\mu(C_n) - \mu(A_n)| \leqslant 2^{-n}$. 则 $\tilde{G}_n = \bigcup_{m=n}^{\infty} G_m$ 为开集, 包含 A, 且 C_n 为闭集, 包含于 A. 我们有

$$\begin{aligned}
\overline{\lim_{n\to\infty}} |\mu(\tilde{G}_n) - \mu(A)| &= \overline{\lim_{n\to\infty}} \left| \mu\left(\bigcup_{m=n}^{\infty} G_m\right) - \mu\left(\bigcup_{m=n}^{\infty} A_m\right) \right| \\
&\leqslant \overline{\lim_{n\to\infty}} \mu\left(\bigcup_{m=n}^{\infty} (G_m - A_m)\right) \\
&\leqslant \overline{\lim_{n\to\infty}} \sum_{m=n}^{\infty} 2^{-m} = 0, \\
\overline{\lim_{n\to\infty}} |\mu(C_n) - \mu(A)| &= \overline{\lim_{n\to\infty}} |\mu(C_n) - \mu(A_n)|.
\end{aligned}$$

故 $A \in \mathscr{C}$.

再设 $A_1, A_2 \in \mathscr{C}, A_1 \supset A_2$. 往证 $A_1 - A_2 \in \mathscr{C}$. 为此, $\forall n \geqslant 1$ 取 $G_n \supset A_1$ 与 $C_n \subset A_2$ 使

$$|\mu(G_n) - \mu(A_1)| + |\mu(C_n) - \mu(A_2)| \leqslant \frac{1}{n}.$$

则 $G_n \setminus C_n$ 为开集, 包含 $A_1 - A_2$, 且

$$|\mu(A_1 - A_2) - \mu(G_n \setminus C_n)| \leqslant |\mu(G_n) - \mu(A_1)| + |\mu(C_n) - \mu(A_2)| \leqslant \frac{1}{n}.$$

故 $A \triangleq A_1 - A_2$ 满足所求的第一个等式. 对称地, 我们可证明它也满足第二个等式. 从而 $A_1 - A_2 \in \mathscr{C}$. $\square$

在 Borel 可测空间 $(E, \mathscr{B})$ 上, 引入如下通用的记号. 以 $\mathscr{B}_b$ 表示有界可测函数全体, 以 $C_b(E)$ 表示有界连续函数全体, 以 $C_0(E)$ 表示具有紧支撑连续函数全体.

引理 6.12 (**有界连续函数为测度的决定类**). 设 $\mu_1, \mu_2 \in \mathfrak{M}$. 如果 $\forall f \in C_b(E)$, 有 $\mu_1(f) = \mu_2(f)$, 则 $\mu_1 = \mu_2$.

证明 由引理 6.11, 只需证 $\forall$ 开集 G 有 $\mu_1(G) = \mu_2(G)$. 令 $g(x) = d(x, G^c), x \in E$. 则 $\forall x \in G$ 有 $g(x) > 0$, 且 g 为 Lipschitz 函数. 令 $h_n(r) = (nr+1)^+ \wedge 1$, 则 $h_n \circ g$ 为 Lipschitz 连续函数且 $h_n \circ g \uparrow \mathbf{1}_G (n \uparrow \infty)$. 由单调收敛定理及 $\mu_1(h_n \circ g) = \mu_2(h_n \circ g)$ 得到 $\mu_1(G) = \mu_2(G)$. $\square$

定义 6.13. 设 $\{\mu_n\} \subset \mathfrak{M}, \mu \in \mathfrak{M}$.

(1) 如果

$$\sup_{A \in \mathscr{B}} |\mu_n(A) - \mu(A)| \to 0, \quad n \uparrow \infty,$$

等价地,

$$\sup_{f \in \mathscr{B}, ||f|| \leqslant 1} |\mu_n(f) - \mu(f)| \to 0, \quad n \uparrow \infty,$$

则称 μ_n 一致收敛到 μ.

(2) 如果 $\forall A \in \mathscr{B}$, 有

$$\lim_{n \to \infty} \mu_n(A) = \mu(A).$$

等价地, $\forall f \in \mathscr{B}_b$, $\mu_n(f) \to \mu(f)$, 则称 μ_n 强收敛到 μ.

(3) 如果 $\forall f \in C_b(E)$ 有 $\mu_n(f) \to \mu(f)$, 则称 μ_n 弱收敛到 μ (记为 $\mu_k \xrightarrow{w} \mu$).

(4) 如果 $\forall f \in C_0(E)$ 有 $\mu_n(f) \to \mu(f)$, 则称 μ_n 淡收敛到 μ (记为 $\mu_k \xrightarrow{v} \mu$).

定义 6.14. 设 $A \in \mathscr{B}$. 如果 $\mu(\bar{A}) = \mu(A^\circ)$, 则称之为 μ 连续集.

定理 6.15. 设 $\mu_n, \mu \in \mathfrak{M}, n \geqslant 1$. 则以下命题等价.

(1)$\forall f \in C_b(E)$, 有 $\mu_n(f) \to \mu(f)$;

(2)任给有界一致连续函数 f, 有 $\mu_n(f) \to \mu(f)$;

(3)任给有界 Lipschitz 函数 f, 有 $\mu_n(f) \to \mu(f)$;

(4)任给开集 $G \subset E$, 有 $\varliminf\limits_{n\to\infty} \mu_n(G) \geqslant \mu(G), \mu_n(E) \to \mu(E)$;

(5)任给闭集 $C \subset E$, 有 $\varlimsup\limits_{n\to\infty} \mu_n(C) \leqslant \mu(C), \mu_n(E) \to \mu(E)$;

(6)任给 μ 连续集 A, 有 $\mu_n(A) \to \mu(A)$.

证明 (1) $\Rightarrow$ (2) $\Rightarrow$ (3) 显然. (4) $\Leftrightarrow$ (5) 显然.

(3) $\Rightarrow$ (5). 设 $C \subset E$ 为闭集, 令

$$f_m(x) = \frac{1}{1 + md(x,C)}, x \in E, m \geqslant 1.$$

则 f_m 为 Lipschitz 连续函数, $f_m \uparrow \mathbf{1}_C$. 故由控制收敛定理及 (3),

$$\begin{aligned} \mu(C) &= \lim_{m\to\infty} \int_E f_m \,\mathrm{d}\mu = \lim_{m\to\infty} \lim_{n\to\infty} \int_E f_m \,\mathrm{d}\mu_n \\ &\geqslant \varlimsup_{n\to\infty} \mu_n(C). \end{aligned}$$

(4) 与 (5) $\Rightarrow$ (6). 设 A 为 μ 连续集, 则 $\mu(A) = \mu(\bar{A}) = \mu(A^\circ)$. 于是由(4) 与 (5) 知

$$\mu(A) = \mu(A^\circ) \leqslant \varliminf_{n\to\infty} \mu_n(A^\circ) \leqslant \varliminf_{n\to\infty} \mu_n(A),$$

$$\mu(A) = \mu(\bar{A}) \geqslant \varlimsup_{n\to\infty} \mu_n(\bar{A}) \geqslant \varlimsup_{n\to\infty} \mu_n(A).$$

从而 (6) 成立.

(6) $\Rightarrow$ (1). $\forall f \in C_b(E)$, 找由 μ 连续集生成的简单函数来逼近 f. 由于 μ 有限, 则

$$D \triangleq \{a \in \mathbb{R} : \mu(\{f = a\}) > 0\}$$

是至多可数集. 故存在 $c > ||f||_\infty + 1$ 使 $\pm c \in D^c$, 以及 $[-c, c]$ 的一列分割

$$I_n \triangleq \{-c = r_0 < r_1 < \cdots < r_n < r_{n+1} = c\}, n \geqslant 2$$

使

$$\{r_i\} \subset D^c,\ \delta(I_n) \triangleq \max_{1 \leqslant k \leqslant n+1} (r_k - r_{k-1}) \to 0.$$

令

$$f_n = \sum_{i=1}^{n-1} r_i \mathbf{1}_{\{r_i \leqslant f < r_{i+1}\}}.$$

则

$$||f_n - f||_\infty \leqslant \delta(I_n) \to 0, \quad n \to \infty.$$

从而

$$\begin{aligned} |\mu(f) - \mu_m(f)| &\leqslant |\mu(f) - \mu(f_n)| + |\mu(f_n) - \mu_m(f_n)| + |\mu_m(f_n) - \mu_m(f)| \\ &\leqslant \delta(I_n)\left(\mu(E) + \mu_m(E)\right) + \sum_{i=1}^{n-1} r_i \left|\mu(r_i \leqslant f < r_{i+1}) - \mu_m(r_i \leqslant f < r_{i+1})\right| \end{aligned}$$

由 (6) 先令 $m \uparrow \infty$, 再令 $n \uparrow \infty$, 即得 (1). □

§ 6.2.2 胎紧性与弱紧性

测度的弱收敛定义了 $\mathfrak{M}$ 上的一个拓扑, 称作弱拓扑. 据此, 可以讨论该拓扑空间的性质, 如集合的紧性、序列的紧性等. 我们先从紧度量空间出发. 下面定理表明, 如果 E 是紧度量空间, 则 $\mathfrak{M}$ 在弱拓扑下是紧空间.

定理 6.16. 设 (E, ρ) 为紧度量空间, 如 $\{\mu_n\} \subset \mathfrak{M}$ 有界, 则 $\exists \mu \in \mathfrak{M}$ 及子列 $\{\mu_{n_k}\}$ 使 $\mu_{n_k} \xrightarrow{w} \mu$.

证明 由泛函分析知道, $C(E)$ 在一致范数下是 Polish 空间. 设 $\{f_1, f_2, \cdots\} \subset C(E)$ 为可数稠子集. $\forall n, m \geqslant 1$, 有 $|\mu_n(f_m)| \leqslant ||f_m||_\infty \cdot C$, 其中 $C = \max\limits_n \mu_n(E)$. 则 $\{\mu_n(f_m)\}_{n\geqslant 1}$ 有收敛子列. 由对角线法则, 存在子列 $\{\mu_{n_k}\}$ 及 $\{\alpha_m\} \subset \mathbb{R}$ 使

$$\lim_{k\to\infty} \mu_{n_k}(f_m) = \alpha_m, m \geqslant 1.$$

此外, 由于 $\{f_n\}_{n\geqslant 1}$ 在 $C(E)$ 中稠, $\forall f \in C(E), \forall \varepsilon > 0, \exists m_0$ 使 $||f_{m_0} - f||_\infty \leqslant \varepsilon$. 从而

$$\begin{aligned} |\mu_{n_k}(f) - \mu_{n_l}(f)| &\leqslant |\mu_{n_k}(f - f_{m_0})| + |\mu_{n_l}(f - f_{m_0})| + |\mu_{n_k}(f_{m_0}) - \mu_{n_l}(f_{m_0})| \\ &\leqslant 2\varepsilon C + |\mu_{n_k}(f_{m_0}) - \mu_{n_l}(f_{m_0})|. \end{aligned}$$

先令 $k, l \to \infty$, 再令 $\varepsilon \to 0$, 得

$$\lim_{l,k\to\infty} |\mu_{n_k}(f) - \mu_{n_l}(f)| = 0.$$

则 $\{\mu_{n_k}(f)\}$ 为 Cauchy 列, $\exists \alpha(f) \in \mathbb{R}$ 使 $\mu_{n_k}(f) \to \alpha(f)$. 易见 $\alpha : C(E) \to \mathbb{R}$ 为有界线性泛函, 由 Riesz-Markov-Kakutani 定理 (见 [7, Theorem IV.14]), 存在唯一的 $\mu \in \mathfrak{M}$ 使 $\mu(f) = \alpha(f)$. 易见 $\mu_{n_k} \xrightarrow{w} \mu$. □

当 E 非紧时, 如果 $\mathfrak{M}$ 的子集 $\mathfrak{M}'$ 中的测度可由它们在紧集上的取值一致逼近, 则可粗略地认为它们的支撑含于某紧集, 从而可归结为紧空间情形. 据此, 我们引入胎紧的概念, 并讨论它与弱紧的关系.

定义 6.17. 设 $\mathfrak{M}' \subset \mathfrak{M}$. 如果 $\mathfrak{M}'$ 一致有界且 $\forall \varepsilon > 0, \exists$ 紧集 $K \subset E$ 使

$$\sup_{\mu\in\mathfrak{M}'} \mu(K^c) < \varepsilon,$$

则称 $\mathfrak{M}'$ 为胎紧的 (tight).

定理 6.18 (Prohorov 定理). 设 (E, ρ) 为度量空间, $\{\mu_n\}_{n\geqslant 1} \subset \mathfrak{M}$.

(1) 如存在一列紧集 $\{K_m\}_{m\geqslant 1}$ 使 $K_m \uparrow E$, 且 $\{\mu_n\}_{n\geqslant 1}$ 一致有界, 则它存在一个淡收敛的子列.

(2) 如 $\{\mu_n\}_{n\geqslant 1}$ 为胎紧的, 则它存在弱收敛子列.

证明 设 $\{K_m\}_{m\geqslant 1}$ 为 E 的一列单增紧子集. 给定 m , $\exists$ 子列 $\{\mu_{m_n}\}$ 及 $\mu^{(m)} \in \mathfrak{M}(K_m)$($K_m$ 上的有限测度全体) 使

$$\mu_{m_n}|_{K_m} \xrightarrow{w} \mu^{(m)}, n \to \infty.$$

这里 $\mu_{m_n}|_{K_m}$ 是 μ_{m_n} 在 K_m 上的限制. 由对角线法则, 这些序列存在共同的子列 $\{\mu_{n'}\}$ 使

$$\mu_{n'}|_{K_m} \xrightarrow{w} \mu^{(m)} (n \to \infty), m \geqslant 1.$$

易见

$$\mu^{(m+1)}(A \cap K_{m+1}) \geqslant \mu^{(m)}(A \cap K_m), \quad \forall A \in \mathscr{B}.$$

事实上, 任给闭集 A, 令

$$h_l = \frac{1}{1 + l d(x, A)},$$

有

$$\mu^{(m+1)}(A \cap K_{m+1}) \;=\; \lim_{l\to\infty} \mu^{(m+1)}(h_l \mathbf{1}_{K_{m+1}})$$

$$\begin{aligned} &= \lim_{l\to\infty}\lim_{n'\to\infty}\mu_{n'}(h_l\mathbf{1}_{K_{m+1}}) \\ &\geqslant \lim_{l\to\infty}\lim_{n'\to\infty}\mu_{n'}(h_l\mathbf{1}_{K_m}) \\ &= \mu^{(m)}(A\cap K_m). \end{aligned}$$

从而 $\forall A\in\mathscr{B}$, 极限 $\mu(A)=\lim_{m\to\infty}\mu^{(m)}(A\cap K_m)$ 存在. 则 $\mu\in\mathfrak{M}$, 且 $\forall f\in C_b(E), \mu^{(m)}(f\mathbf{1}_{K_m})\to\mu(f\mathbf{1}_{K_m})$.

(1) 由于 $K_m\uparrow E, \forall f\in C_0(E)$ 存在 m 使 $\mathrm{supp}f\subset K_m$. 故

$$\lim_{n'\to\infty}\mu_{n'}(f)=\lim_{m\to\infty}\mu^{(m)}(f)=\mu(f).$$

(2) 取紧子集列 $\{K_m\}_{m\geqslant 1}$ 使 $\sup\limits_{n\geqslant 1}\mu_n(K_m^c)\leqslant 1/m, m\geqslant 1$. 则 $\forall f\in C_b(E)$, 有

$$\begin{aligned} &|\mu_{n'}(f)-\mu(f)| \\ &\leqslant |\mu_{n'}(f\mathbf{1}_{K_m})-\mu^{(m)}(f\mathbf{1}_{K_m})|+|\mu_{n'}(f)-\mu_{n'}(f\mathbf{1}_{K_m})|+|\mu(f)-\mu^{(m)}(f\mathbf{1}_{K_m})| \\ &\leqslant C||f||_\infty+|\mu_{n'}(f\mathbf{1}_{K_m})-\mu^{(m)}(f\mathbf{1}_{K_m})|+|\mu(f)-\mu^{(m)}(f\mathbf{1}_{K_m})|, \end{aligned}$$

先令 $n'\uparrow\infty$, 再令 $m\uparrow\infty$, 即得 $\mu_{n'}\xrightarrow{w}\mu$. □

定理 6.19. 设 E 为 Polish 空间, 则 $\mathfrak{M}'\subset\mathfrak{M}$ 是弱相对紧的当且仅当它是胎紧的.

证明 由 Prohorov 定理, 只需证明必要性. $\forall\mu\in\mathfrak{M}$ 及开集列 $G_n\uparrow E$, 有 $\mu(G_n^c)\downarrow 0$. 今取 $\{\mu_n\}\subset\mathfrak{M}'$ 使 $\forall n\geqslant 1$ 有

$$\mu_n(G_n^c)\geqslant\sup_{\mu\in\mathfrak{M}'}\mu(G_n^c)-1/n.$$

由于 $\mathfrak{M}'$ 弱相对紧, $\exists\mu_0\in\mathfrak{M}$ 及子列 $\mu_{n'}\xrightarrow{w}\mu_0$. 于是

$$\begin{aligned} \lim_{n'\to\infty}\sup_{\mu\in\mathfrak{M}'}\mu(G_{n'}^c) &\leqslant \varlimsup_{n'\to\infty}\mu_{n'}(G_{n'}^c) \\ &\leqslant \lim_{m\to\infty}\varlimsup_{n'\to\infty}\mu_{n'}(G_m^c)\leqslant\lim_{m\to\infty}\mu_0(G_m^c)=0. \end{aligned}$$

由于 E 可分, 则 E 具有可数开球覆盖: $\forall m\geqslant 1, \exists\{x_{m,j}\}\subset E$ 使 $E=\bigcup\limits_{j=1}^{\infty}B(x_{m,j},2^{-m})$. 令 $G(n,m)=\bigcup\limits_{j=1}^{\infty}B(x_{m,j},2^{-m})$. 则 $G(n,m)\uparrow E(n\uparrow\infty)$. 将前面结论应用于 $G_m=G(n,m)$, 则 $\forall\varepsilon>0, \exists N(\varepsilon,m)\geqslant 1$ 使得

$$\sup_{\mu\in\mathfrak{M}'}\mu(G(n,m)^c)\leqslant\frac{\varepsilon}{2^m},\ n\geqslant N(\varepsilon,m).$$

令 $G_\varepsilon = \bigcap\limits_{m=1}^{\infty} G(N(\varepsilon,m),m)$, 则 $K_\varepsilon \triangleq \bar{G}_\varepsilon$ 为紧集, 而且

$$\mu(K_\varepsilon^c) \leqslant \mu(G_\varepsilon^c) \leqslant \sum_{r=1}^{\infty} \mu(G(N,r)^c) \leqslant \sum_{r=1}^{\infty} \frac{\varepsilon}{2^r} = \varepsilon,\ \mu \in \mathfrak{M}'. \quad \square$$

上面证明中关于 G_ε 的相对紧来自于点集拓扑学中的 Hausdorff 定理, 即完备度量空间的有界集合的完全有界性等价于相对紧性. 设 $A \subset E$, 如 $\forall \varepsilon > 0, \exists$ 有限个半径为 ε 的开球覆盖 A, 则称 A 为完全有界的.

当 $E = \mathbb{R}^d$ 时, 我们有如下关于弱收敛的等价命题, 即可将定理 6.15 (6) 中的 μ 连续集改为 μ 连续区间.

命题 6.20. 设 $E = \mathbb{R}^d$, 则 $\mu_n \xrightarrow{w} \mu$ 当且仅当任意有限 μ 连续区间 $[a,b)$, 有 $\mu_n([a,b)) \to \mu([a,b))$.

证明 只需证明充分性. 任给 μ_n 不弱收敛于 μ, 则存在 $\delta > 0, f \in C_b(\mathbb{R}^d)$ 及子列 $n_k \to \infty$ 使

$$|\mu_{n_k}(f) - \mu(f)| \geqslant \delta, k \geqslant 1. \tag{6.2.1}$$

由引理 6.7 与引理 6.8 知, 存在一列 μ 连续区间 $I_m \uparrow \mathbb{R}^d$. 则 $\forall \varepsilon > 0$, 存在 $m \geqslant 1$ 使 $\mu(I_m^c) \leqslant \varepsilon/2$. 由于当 $n \to \infty$ 时 $\mu_n(I_m) \to \mu(I_m)$ 且 $\mu_n(\mathbb{R}^d) \to \mu(\mathbb{R}^d)$, 我们有 $\overline{\lim\limits_{n\to\infty}}\, \mu_n(I_m^c) \leqslant \varepsilon/2$. 从而存在 $n_0 \geqslant 1$ 使 $\forall n \geqslant n_0$ 有 $\mu_n(I_m^c) < \varepsilon$. 此外, 取紧集 K_1 使 $\forall n \leqslant n_0$ 有 $\mu_n(K_1^c) < \varepsilon$. 则紧集 $K = K_1 \cup \bar{I}_m$ 满足 $\mu_n(K^c) < \varepsilon, \forall n \geqslant 1$. 故 $\{\mu_{n_k}\}$ 胎紧, 存在子列 n_k' 及有限测度 μ' 使 $\mu_{n_k'} \xrightarrow{w} \mu'$. 特别地, μ' 与 μ 在共同的连续区间上取值相同, 由命题 6.9 知 $\mu' = \mu$. 这与 (6.2.1) 矛盾. $\square$

§6.3 特征函数与弱收敛

本节先讨论 $\mathbb{R}^n$ 上有限测度序列的弱收敛等价于其特征函数的收敛性, 并给出特征函数的特征.

定理 6.21. 设 $\{\mu_k\}$ 为 $\mathbb{R}^n$ 上有限测度序列. 则 $\mu_k \xrightarrow{w} \mu$ 当且仅当 μ_k 的特征函数收敛到 μ 的特征函数.

由控制收敛定理, 必要性是显然的. 而充分性是下面的定理 6.25 的推论. 为证明该定理, 我们需要使用特征函数的不定积分, 即积分特征函数.

定义 6.22. 设 f_μ 是有限测度 μ 的特征函数, 称 f_μ 的不定积分

$$\tilde{f}_\mu(u_1,\cdots,u_n)=\int_0^{u_1}\cdots\int_0^{u_n} f_\mu(t_1,\cdots,t_n)\,\mathrm{d}t_1\cdots\mathrm{d}t_n$$

为 μ 的积分特征函数.

由于 f_μ 连续, 则 f_μ 与 $\tilde{f}_\mu$ 相互唯一确定. 我们先将 $\tilde{f}_\mu$ 表示成带参数的连续函数关于 μ 的积分.

引理 6.23. μ 的积分特征函数满足

$$\tilde{f}_\mu(u_1,\cdots,u_n)=\int_{\mathbb{R}^n}\prod_{k=1}^n\frac{\mathrm{e}^{\mathrm{i}u_kx_k}-1}{\mathrm{i}x_k}\mu(\,\mathrm{d}x_1,\cdots,\,\mathrm{d}x_n),u_1,\cdots,u_n\in\mathbb{R}.$$

证明 由定义与 Fubini 定理, 对 $u=(u_1,\cdots,u_n)\in\mathbb{R}^n$,

$$\begin{aligned}\tilde{f}_\mu(u) &= \int_0^{u_1}\cdots\int_0^{u_n}\int_{\mathbb{R}^n}\mathrm{e}^{\mathrm{i}\langle t,x\rangle}\mu(\,\mathrm{d}x)\,\mathrm{d}t\\ &= \int_{\mathbb{R}^n}\mu(\,\mathrm{d}x)\int_{[0,u]}\mathrm{e}^{\mathrm{i}\langle t,x\rangle}\,\mathrm{d}t\\ &= \int_{\mathbb{R}^n}\prod_{k=1}^n\frac{\mathrm{e}^{\mathrm{i}u_kx_k}-1}{\mathrm{i}x_k}\mu(\,\mathrm{d}x_1,\cdots,\,\mathrm{d}x_n).\quad\square\end{aligned}$$

令

$$F(x,u)=\prod_{k=1}^n\frac{e^{\mathrm{i}\,u_kx_k}-1}{\mathrm{i}\,x_k},x,u\in\mathbb{R}^n.$$

则给定 u, 有 $\lim\limits_{|x|\to\infty}F(x,u)=0$. 从而 $F(\cdot,u)$ 可由有限区间上的连续函数逼近.

定理 6.24. 设 $\{\mu_k\}$ 一致有界, $\left\{\tilde{f}_k\right\}_{k\geqslant1}$ 为相应积分特征函数. 如果 $\tilde{f}_k\to\tilde{g}$, 则存在有限测度 μ 使 $\mu_k\xrightarrow{v}\mu$ 且 $\tilde{g}=\tilde{f}_\mu$.

证明 由定理 6.18, μ_k 存在子列 $\mu_{k'}$ 淡收敛于某有限测度 μ. 由于积分特征函数决定测度的分布, 只需证明 $\tilde{f}_\mu=\tilde{g}$. 由于 $\tilde{f}_\mu(u)=\mu(F(u,\cdot))$, 而 $F(u,\cdot)$ 可由紧支撑连续函数一致逼近, 易见 $\mu_{k'}\xrightarrow{v}\mu$ 与 $\tilde{f}_k\to\tilde{g}$ 蕴含 $\tilde{f}_\mu=\tilde{g}$. $\square$

定理 6.25. 若 $\{\mu_k\}$ 一致有界且 $f_k\to g$, 其中 g 在 0 处连续, 则 $\exists\mu$ 使 $\mu_k\xrightarrow{w}\mu$ 且 $f_\mu=g$.

证明 由控制收敛定理, $f_k \to g$ 蕴含 $\tilde{f}_k \to \tilde{g}$. 由定理 6.24、命题 6.20 及习题 13, 只需证 $\mu_k(\mathbb{R}^n) \to \mu(\mathbb{R}^n)$. 由于 $\tilde{g} = \tilde{f}_\mu$, 从而 $g = f_\mu$, a.e. $\mathrm{d}x$. 又由于 g 与 f_μ 在 0 处连续, 故

$$\mu(\mathbb{R}^n) = f_\mu(0) = g(0) = \lim_{k\to\infty} f_k(0) = \lim_{k\to\infty} \mu_k(\mathbb{R}^n). \quad \square$$

下面介绍定理 6.21 的两个重要应用.

定理 6.26 (**大数定律**). 设 $\{\xi_n\}$ 为 i.i.d, $\mathbb{E}\xi_n = a \in \mathbb{R}$. 则

$$\frac{1}{n}\sum_{k=1}^n \xi_k \xrightarrow{\mathbb{P}} a.$$

证明 (1) 只需证 $\eta_n := \frac{1}{n}\sum\limits_{k=1}^n (\xi_k - a)$ 的特征函数 $f_n(t) \to 1$. 事实上, 如果$f_n(t) \to 1$, 则由定理 6.21 知 $\mathbb{P}_{\eta_n} \xrightarrow{w} \delta_0$ (质量集中于 0 的概率). 由于 $\forall \varepsilon > 0, (-\varepsilon, \varepsilon)$ 为 δ_0 连续集, 从而

$$\underline{\lim}_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^n \xi_k - a\right| < \varepsilon\right) = \underline{\lim}_{n\to\infty} \mathbb{P}_{\eta_n}((-\varepsilon,\varepsilon)) = \delta_0((-\varepsilon,\varepsilon)) = 1.$$

由此知

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^n \xi_k - a\right| \geqslant \varepsilon\right) = 0.$$

(2) 令 $\xi_n' = \xi_n - a$, 则 $\eta_n = \frac{1}{n}\sum\limits_{k=1}^n \xi_k'$, 故

$$f_n(t) = \prod_{k=1}^n f_{\xi_k'}(t/n) = [f(t/n)]^n .$$

其中 $f = f_{\xi_k'}$. 由于 $\mathbb{E}\xi_k' = 0$, 由 Taylor 公式

$$f_n(t) = \left[\mathbb{E}\mathrm{e}^{\mathrm{i}t\xi_k'/n}\right]^n = (1 + o(t/n))^n.$$

于是给定 t, 有

$$\lim_{n\to\infty} \log f_n(t) = \lim_{n\to\infty} \log\left(1 + o(t/n)\right)^n = \lim_{n\to\infty} n\log\left(1 + o(t/n)\right) = 0.$$

从而 $\lim_{n\to\infty} f_n(t) = 1$. $\square$

定理 6.27 (中心极限定理). 设 $\{\xi^{(k)}\}_{k\geqslant 1}$ 为 n 维 i.i.d 随机变量列, 期望与方差有限, $\mathbb{E}\xi^{(k)}=m\in\mathbb{R}^n$, 相关矩阵 D 满秩. 则 $\forall x\in\mathbb{R}^n$, 有

$$\lim_{N\to\infty}\mathbb{P}\left(\frac{1}{\sqrt{N}}\sum_{k=1}^{N}(\xi^{(k)}-m)<x\right)=\frac{1}{(2\pi)^{n/2}|D|^{1/2}}\int_{(-\infty,x)}\mathrm{e}^{-\frac{1}{2}\langle t,D^{-1}t\rangle}\,\mathrm{d}t.$$

证明 令 $\eta^{(k)}=\xi^{(k)}-m$, 则 $\{\eta^{(k)}\}$ i.i.d, 期望为零. 令 $\eta^{(k)}$ 的特征函数为 f, 则 $\frac{1}{\sqrt{N}}\sum\limits_{k=1}^{N}\eta^{(k)}$ 的特征函数为

$$f_N(t)=\left[f(t/\sqrt{N})\right]^N, t\in\mathbb{R}^n.$$

由于 $\mathbb{E}\eta^{(k)}=0$, 由 Taylor 展开

$$f(t/\sqrt{N})=1-\frac{1}{2N}\langle t,Dt\rangle+o(t^2/N).$$

给定 $t\in\mathbb{R}^n$, 有

$$\log f(t/\sqrt{N})=-\frac{1}{2N}\langle t,Dt\rangle+o(N^{-1}).$$

从而

$$\log f_N(t)=-\frac{1}{2}\langle t,Dt\rangle+o(1)\to-\frac{1}{2}\langle t,Dt\rangle, N\to\infty.$$

即

$$\lim_{N\to\infty}f_N(t)=\mathrm{e}^{-\frac{1}{2}\langle t,Dt\rangle},$$

它是 $N(0,D)$ 的特征函数. 由定理 6.21 知 $\frac{1}{\sqrt{N}}\sum\limits_{k=1}^{N}\eta^{(k)}$ 依分布收敛到 $N(0,D)$. □

§6.4　特征函数与非负定性

设 μ 为 $\mathbb{R}^n$ 上的有限测度, f_μ 为其特征函数. 易见 $\forall m\geqslant 1$ 及 $\alpha_1,\cdots,\alpha_m\in\mathbb{C}, t^{(1)},\cdots,t^{(m)}\in\mathbb{R}^n$ 有

$$\sum_{j,k=1}^{m}f\left(t^{(j)}-t^{(k)}\right)\alpha_j\bar{\alpha}_k=\int_{\mathbb{R}^n}\left|\sum_{k=1}^{m}\alpha_k\mathrm{e}^{\mathrm{i}\langle t^{(k)},x\rangle}\right|^2\mu(\,\mathrm{d}x)\geqslant 0.$$

一般地, 称具有如上性质的复函数 f 为非负定函数, 称该性质为函数的非负定性. 本节将证明, 非负定性是特征函数的特征.

性质 6.28. 如果 f 为非负定函数, 则 $f(0)\geqslant 0, f(-t)=\bar{f}(t), |f(t)|\leqslant f(0)$.

证明 令 $m=2, t^{(1)}=0, t^{(2)}=t, \alpha_1=1, \alpha_2\in\mathbb{C}$, 则由非负定性知

$$f(0)\left[1+|\alpha_2|^2\right]+f(t)\alpha_2+f(-t)\bar{\alpha}_2\geqslant 0.$$

(1) 令 $\alpha_2=0$, 则 $f(0)\geqslant 0$;

(2) 令 $\alpha_2=1$, 则 $2f(0)+f(-t)+f(t)\geqslant 0$. 从而 $\mathrm{Im}f(t)=-\mathrm{Im}f(-t)$. 再令 $\alpha_2=\mathrm{i}$, 则 $2f(0)+\mathrm{i}(f(t)-f(-t))\geqslant 0$. 从而 $\mathrm{Re}f(t)=-\mathrm{Re}f(-t)$. 因此 $f(-t)=\bar{f}(t)$.

(3) 如 $f(t)\neq 0$, 令 $\alpha_2=-\bar{f}(t)/|f(t)|$, 则 $2f(0)\geqslant 2|f(t)|$. 从而 $f(0)\geqslant |f(t)|$. □

引理 6.29. 令 $T_c=\{kc: k\in\mathbb{Z}^n\}, c>0$. 如果 f 为非负定函数, 则存在有限测度 μ 使

$$\mu(\mathbb{R}^n)=\mu([-\pi/c,\pi/c]^n)=f(0),$$

且 μ 的特征函数 f_μ 满足

$$f_\mu(t)=f(t),\quad \forall t\in T_c.$$

定理 6.30. 如 $f(t)$ 为在 0 处连续的非负定函数, 则它为某有限测度的特征函数.

证明 由引理 6.29, 存在一列有限测度 $\{\mu_m\}$ 使 $\mu_m(\mathbb{R}^n)=\mu_m([-m\pi,m\pi]^n)=f(0)$, 且相应特征函数 f_m 满足 $f_m(t)=f(t), t\in\frac{1}{m}\mathbb{Z}^n$. $\forall t\in\mathbb{R}^n$, 取 $\left\{t^{(m)}\right\}_{m\geqslant 1}\subset T_{\frac{1}{m}}$ 使 $|t_k-t_k^{(m)}|\leqslant 1/m, 1\leqslant k\leqslant n, m\geqslant 1$. 故由 f 的连续性及 $f(t^{(m)})=f_m(t^{(m)})$ 知

$$f(t)=\lim_{m\to\infty}f(t^{(m)})=\lim_{m\to\infty}f_m(t^{(m)}).$$

由此及定理 6.21, 只需证明

$$\lim_{m\to\infty}|f_m(t)-f_m(t^{(m)})|=0. \tag{6.4.1}$$

为此, 使用增量不等式得

$$
\begin{aligned}
&|f_m(t)-f_m(t^{(m)})|\\
&\leqslant \sum_{i=0}^{n-1}|f_m(t_1,\cdots,t_i,t_{i+1}^{(m)},\cdots,t_n^{(m)})-f_m(t_1,\cdots,t_{i+1},t_{i+2}^{(m)},\cdots,t_n^{(m)})|\\
&\leqslant \sum_{i=0}^{n-1}\sqrt{2f(0)(f(0)-\mathrm{Re}f_m(e_i(t_i-t_i^{(m)}))},
\end{aligned}
\tag{6.4.2}
$$

取值 $e_i\in\mathbb{R}^n$ 为第 i 个为 1 的单位向量. 由于对 $x_i\in[-m\pi,m\pi]$ 有 $|(t_i-t_i^{(m)})x_i|\leqslant\pi$, 而 $\cos\theta$ 在 $\theta\in[-\pi,\pi]$ 上关于 $|\theta|$ 为降函数, 我们有

$$
\begin{aligned}
0&\leqslant f(0)-\mathrm{Re}f_m(e_i(t_i-t_i^{(m)}))\\
&=\int_{[-m\pi,m\pi]^n}\left(1-\cos[(t_i-t_i^{(m)})x_i]\right)\mu_m(\,\mathrm{d}x)\\
&\leqslant\int_{[-m\pi,m\pi]^n}\left(1-\cos\frac{x_i}{m}\right)\mu_m(\,\mathrm{d}x)\\
&=f(0)-\mathrm{Re}f_m\left(\frac{e_i}{m}\right).
\end{aligned}
$$

由此及 (6.4.2) 并使用 f 在 0 处的连续性立即得到 (6.4.1). □

§6.5 补充与习题

1. 证明特征函数 f 具有如下性质: $f(0) = \mu(\mathbb{R}^n)$; (2) $|f(t)| \leqslant f(0)$; (3) $\bar{f}(t) = f(-t)$.

2. 证明命题 6.2.

3. 若 $(\mathbb{R}, \mathscr{B})$ 上的有限测度 μ 满足 $\mu(-\infty, x) = \mu(x, \infty)$, 则称之为对称测度.

   (a) μ 对称当且仅当 $\mu(A) = \mu(-A), A \in \mathscr{B}$, 其中 $-A = \{x : -x \in A\}$;

   (b) μ 对称当且仅当其特征函数为实值函数.

4. 设随机变量 ξ 的特征函数 ϕ 是实的, 则 $-\xi$ 的特征函数也是 ϕ.

5. 如果有限测度 μ 的特征函数 ϕ 满足 $\int |\phi(t)|\,\mathrm{d}t < \infty$, 那么 μ 具有有界连续的密度函数
$$f(x) = \frac{1}{2\pi}\int e^{-\mathrm{i}\,tx}\phi(t)\,\mathrm{d}t.$$

6. 证明命题 6.4.

7. 称 $\mathbb{E}(e^{s\xi})$ 为随机变量 ξ 的母函数 (Laplace 变换). 试比较它与特征函数的性质.

8. 设 ξ_n 服从正态分布 $N(0, \sigma_n^2)$ 且 $\xi_n \xrightarrow{d} \xi$. 证明 $\sigma_n^2 \to \sigma \in [0, \infty)$.

9. 设 $\xi_1, \xi_2, \cdots$ 独立服从 $\{0, 1\}$ 上的均匀分布, 令
$$\xi = 2\sum_{j=1}^{\infty} \xi_j / 3^j,$$
   确定 ξ 的分布并求其特征函数.

10. 举例说明淡收敛不等价于弱收敛.

11. 证明 $\mu_k \xrightarrow{w} \mu$ 当且仅当 $\forall \mu$ 连续紧集 A 有 $\mu_k(A) \to \mu(A)$.

12. 证明 $\mu_k \xrightarrow{w} \mu$ 当且仅当 $\forall \mu$ 连续开集 A 有 $\mu_k(A) \to \mu(A)$.

13. 在 $\mathbb{R}^n$ 上, $\mu_k \xrightarrow{v} \mu$ 当且仅当 $\forall \mu$ 连续有限区间 I 有 $\mu_k(I) \to \mu(I)$.

14. 设 $g \geqslant 0$ 连续. 若 $\xi_n \xrightarrow{w} \xi$, 则

$$\varliminf_{n\to\infty} \mathbb{E}g(\xi_n) \geqslant \mathbb{E}g(\xi).$$

15. 如果分布函数 $F_n \Rightarrow F$ 且 F 连续, 则 $\sup\limits_x |F_n(x) - F(x)| \to 0$.

16. 证明一族随机变量 $\xi_t, t \in T$ 胎紧的充要条件存在函数 $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ 满足 $\lim\limits_{x\to\infty} \phi(x) = \infty$ 使得 $\sup\limits_{t\in T} \mathbb{E}(\phi(\xi_t) < \infty$.

17. 设 ξ_n 胎紧. 若 $\eta_n \xrightarrow{\mathbb{P}} 0$, 则 $\xi_n\eta_n \xrightarrow{\mathbb{P}} 0$.

18. 设 $h : \mathbb{R} \to \mathbb{R}$ 可测且其不连续点集 D_h 可测. 若 $\mu_n \xrightarrow{w} \mu$ 且 $\mu(D_h) = 0$, 则 $\mu_n \circ h^{-1} \xrightarrow{w} \mu \circ h^{-1}$.

19. 设 $(\mathbb{R}, \mathscr{B})$ 上的有限测度 μ 具有密度 p, 其特征函数为 f.

    (a) 证明 $\lim\limits_{|t|\to\infty} f(t) = 0$.

    (b) 若 p 可积的导函数 p', 则 $f(t) = o(t^{-1}), |t| \to \infty$.

    (c) 讨论 p 具有更高阶可积导函数的情况.

20. 证明

$$\mu(\{x\}) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-\mathrm{i}\,tx} f(t)\,\mathrm{d}t.$$

第七章　概率距离

设 (E,ρ) 为度量空间, $\mathscr{P}(E)$ 为 Borel σ 代数 $\mathscr{E}$ 上所有概率测度全体. 前面已经介绍过概率测度的弱收敛, 那么这种收敛可否使用概率测度空间中的距离来描述呢? 本章的主要目的是介绍 $\mathscr{P}(E)$ 上几类常用的度量, 称为概率距离. 这些度量在概率论与随机过程理论的研究中是非常重要的.

§7.1　弱拓扑的度量化

前面研究的概率测度弱收敛定义了 $\mathscr{P}(E)$ 上一个拓扑, 称为弱拓扑, 它是概率论与随机过程研究中非常常用的拓扑. 本节表明, 在许多情况下弱拓扑是可以度量化的. 设 (E,ρ) 为 Polish 空间, 则在一致范数 $||f||_\infty = \sup\limits_E |f|$ 之下有界函数空间 $C_b(E)$ 也是 Polish 空间 (参考[13, 14, 15]), 从而存在一列 $\{f_n\}_{n\geqslant 1} \subset C_b(E)$ 在 $C_b(E)$ 中稠密.

令

$$d_w(\mu,\nu) = \sum_{n=1}^{\infty} 2^{-n} \left\{ |\mu(f_n) - \nu(f_n)| \wedge 1 \right\},\ \mu,\nu \in \mathscr{P}(E).$$

定理 7.1. 设 (E,ρ) 为 Polish 空间, 则 $(\mathscr{P}(E), d_w)$ 可分度量空间, 且任给 $\{\mu_n\}_{n\geqslant 1} \subset \mathscr{P}(E)$ 与 $\mu \in \mathscr{P}(E)$, $\mu_n \xrightarrow{w} \mu$ 当且仅当 $d_w(\mu_n,\mu) \to 0$. 如 E 为局部紧的, 则 $(\mathscr{P}(E), d_w)$ 是完备的.

证明 (a) d_w 为距离.

易见, $d_w(\mu,\mu) = 0$. 如 $d_w(\mu,\nu) = 0$, 则 $\mu(f_n) - \nu(f_n) = 0 (\forall n)$. 由于 $\{f_n\}_{n\geqslant 1}$ 在 $C_b(E)$ 中稠, 由此知 $\mu(f) = \nu(f), \forall f \in C_b(E)$, 从而 $\mu = \nu$. 最后, d_w 显然满足三角不等式.

(b) 弱收敛等价于依 d_w 收敛.

显然, 如 $\mu_n \xrightarrow{w} \mu$, 则 $d_w(\mu_n,\mu) \to 0$. 反之, 设 $d_w(\mu_n,\mu) \to 0$, 往证 $\forall f \in C_b(E)$ 有 $\mu_n(f) - \mu(f) \to 0$. 给定 $f \in C_b(E)$, 由于 $\{f_n\}$ 在 $C_b(E)$ 中稠, $\forall \varepsilon > 0, \exists n_0 \geqslant 1$ 使 $||f_{n_0} - f||_\infty < \varepsilon$. 从而

$$\begin{aligned}\overline{\lim_{n\to\infty}} |\mu_n(f) - \mu(f)| &\leqslant 2\varepsilon + \overline{\lim_{n\to\infty}} |\mu_n(f_{n_0}) - \mu(f_{n_0})| \\ &\leqslant 2\varepsilon + 2^{n_0+1} \overline{\lim_{n\to\infty}} d_w(\mu_n,\mu) \\ &= 2\varepsilon.\end{aligned}$$

由 ε 的任意性知 $\mu_n(f) \to \mu(f)$.

(c) 可分性.

$\forall m \geqslant 1, U_m \triangleq \{(\mu(f_1), \cdots, \mu(f_m)) : \mu \in \mathscr{P}(E)\} \subset \mathbb{R}^m$. 由于 $\mathbb{R}^m$ 可分, 故 U_m 也可分. 从而存在可数集 $\mathscr{P}_m \subset \mathscr{P}(E)$ 使得

$$\tilde{U}_m \triangleq \{(\mu(f_1), \cdots, \mu(f_m)) : \mu \in \mathscr{P}_m\}$$

在 U_m 中稠. 则 $\mathscr{P}_\infty \triangleq \bigcup\limits_{m=1}^{\infty} \mathscr{P}_m$ 为 $\mathscr{P}(E)$ 的可数子集, 仅需证明它在 $\mathscr{P}(E)$ 中关于距离 d_w 稠密.

事实上, $\forall \mu \in \mathscr{P}(E)$, 存在 $\mu_m \in \mathscr{P}_m$ 使

$$|\mu_m(f_i) - \mu(f_i)| \leqslant \frac{1}{m}, \quad \forall 1 \leqslant i \leqslant m.$$

从而

$$d_w(\mu_m, \mu) \leqslant 2^{-m} + \frac{1}{m} \to 0 \ (m \to \infty).$$

(d) d_w 的完备性.

设 E 局部紧, $\{\mu_n\}_{n\geqslant 1} \subset \mathscr{P}(E)$ 关于 d_w 为 Cauchy 列, 则 $\forall m \geqslant 1, \{\mu_n(f_m)\}_{n\geqslant 1}$ 为 Cauchy 列, 从而收敛于某常数, 记成 $\phi(f_m)$. 此外, 给定 $f \in C_b(E), \forall \varepsilon > 0, \exists m_0 \geqslant 1$ 使得 $||f_{m_0} - f||_\infty < \varepsilon$. 从而

$$\begin{aligned}\overline{\lim_{m,n\to\infty}} |\mu_m(f) - \mu_n(f)| &\leqslant 2\varepsilon + \overline{\lim_{m,n\to\infty}} |\mu_m(f_{m_0}) - \mu_n(f_{m_0})| \\ &= 2\varepsilon.\end{aligned}$$

由 ε 的任意性知 $\{\mu_n(f)\}_{n\geqslant 1}$ 也是 Cauchy 数列, 从而收敛于某常数, 记成 $\phi(f)$. 由积分的性质易见

$$\phi : C_b(E) \to \mathbb{R}$$

是线性映射, $\phi(1)=1$, 且如 $f \geqslant 0$ 则 $\phi(f) \geqslant 0$. 由 Riesz 表示定理存在唯一的 $\mu \in \mathscr{P}(E)$ 使 $\mu(f)=\phi(f), \forall f \in C_b(E)$. 见[7, Theorem IV.18]. 由 ϕ 的构造知 $\mu_n \xrightarrow{w} \mu$, 从而由 (b) 知 $d_w(\mu_n, \mu) \to 0$. $\square$

§7.2 全变差距离与 Wasserstein 耦合

设 $(E, \mathscr{E})$ 为可测空间, $\mathscr{P}(E)$ 为其上所有概率测度全体. 我们在第六章已经引入了强收敛的概念, 它在 $\mathscr{P}(E)$ 上定义了强拓扑. 为将该拓扑度量化, 引入全变差距离:

$$||\mu-\nu||_{\mathrm{Var}} = \sup_{A \in \mathscr{E}} (\mu(A)-\nu(A)) + \sup_{A \in \mathscr{E}} (\nu(A)-\mu(A)).$$

易见

$$||\mu-\nu||_{\mathrm{Var}} = 2(\mu-\nu)^+(E) = 2(\nu-\mu)^+(E) = |\mu-\nu|(E), \tag{7.2.1}$$

其中 $|\mu-\nu|$ 为符号测度 $\mu-\nu$ 的全变差.

本节的主要目的是使用耦合来刻画全变差距离.

定义 7.2. 设 $\mu, \nu \in \mathscr{P}(E), \pi \in \mathscr{P}(E \times E)$. 如果

$$\pi(A \times E) = \mu(A), \pi(E \times A) = \nu(A), A \in \mathscr{E},$$

则称 π 为 μ 与 ν 的一个耦合. 以 $\mathscr{C}(\mu,\nu)$ 记 μ 与 ν 的耦合全体.

假设 $D_0 = \{(x,x) : x \in E\} \in \mathscr{E} \times \mathscr{E}$, 我们将证明

$$||\mu-\nu||_{\mathrm{Var}} = 2 \inf_{\pi \in \mathscr{C}(\mu,\nu)} \pi(D_0^c), \tag{7.2.2}$$

且下确界被 Wasserstein 耦合达到. 为构造 Wasserstein 耦合, 先引入两个概率测度的下端 $\mu \wedge \nu$.

命题 7.3. 任给 $\mu, \nu \in \mathscr{P}(E)$, 有

$$\mu \wedge \nu \triangleq \mu - (\mu-\nu)^+ = \nu - (\nu-\mu)^+.$$

证明 由 Hahn 分解定理, 存在 $D \in \mathscr{E}$ 使 $(\mu-\nu)(D) = \inf\limits_{A\in\mathscr{E}}(\mu-\nu)(A)$, 且

$$(\mu-\nu)^+(A) \triangleq (\mu-\nu)(D^c\cap A), (\nu-\mu)^+(A) = (\nu-\mu)(A\cap D), A\in\mathscr{E}.$$

故

$$\begin{aligned}(\mu-(\mu-\nu)^+)(A) &= \mu(A)-\mu(D^c\cap A)+\nu(D^c\cap A)\\ &= \mu(A\cap D)+\nu(A)-\nu(D\cap A)\\ &= (\nu-(\nu-\mu)^+)(A). \quad \square\end{aligned}$$

定理 7.4. 设 $D_0 \triangleq \{(x,x): x\in E\}\in\mathscr{E}\times\mathscr{E}$. 则 (7.2.2) 成立, 且右边的下确界被 Wasserstein 耦合

$$\pi_0(\,\mathrm{d}x,\,\mathrm{d}y) \triangleq (\mu\wedge\nu)(\,\mathrm{d}x)\delta_x(\,\mathrm{d}y) + \frac{(\mu-\nu)^+(\,\mathrm{d}x)(\mu-\nu)^-(\,\mathrm{d}y)}{(\mu-\nu)^-(E)},$$

达到, 其中当 $\mu=\nu$ 时, 右边第二项约定为零.

证明 (a) $\pi_0\in\mathscr{C}(\mu,\nu)$.

显然 π_0 为乘积空间 $(E\times E,\mathscr{E}\times\mathscr{E})$ 上的测度. 当 $\mu=\nu$ 时, $\pi_0(\,\mathrm{d}x,\,\mathrm{d}y) = \mu(\,\mathrm{d}x)\delta_x(\,\mathrm{d}y)$, 从而

$$\pi_0(A\times E) = \pi_0(E\times A) = \mu(A), A\in\mathscr{E},$$

即 $\pi_0\in\mathscr{C}(\mu,\nu)$. 当 $\mu\neq\nu$ 时, 有 $(\mu-\nu)^+(E)>0$. 由于 $(\mu-\nu)^- = (\nu-\mu)^+$ 且 $\mu(E)=\nu(E)=1$, 我们有 $(\mu-\nu)^-(E) = (\mu-\nu)^+(E)$. 所以

$$\begin{aligned}\pi_0(A\times E) &= (\mu\wedge\nu)(A) + \frac{(\mu-\nu)^+(A)(\mu-\nu)^-(E)}{(\mu-\nu)^-(E)}\\ &= \mu(A)-(\mu-\nu)^+(A)+(\mu-\nu)^+(A) = \mu(A),\\ \pi_0(E\times A) &= \int_E \mathbf{1}_A(x)(\mu\wedge\nu)(\,\mathrm{d}x) + (\mu-\nu)^-(A)\\ &= \nu(A)-(\mu-\nu)^-(A)+(\mu-\nu)^-(A)\\ &= \nu(A).\end{aligned}$$

对所有 $A\in\mathscr{E}$ 成立. 故 $\pi_0\in\mathscr{C}(\mu,\nu)$.

(b) $\forall\pi\in\mathscr{C}(\mu,\nu)$, 有

$$\begin{aligned}\mu(A)-\nu(A) &= \pi(A\times E)-\pi(E\times A)\\ &\leqslant \pi(\{(x,y): x\in A, y\notin A\})\\ &\leqslant \pi(D_0^c).\end{aligned}$$

从而 $\|\mu-\nu\|_{\mathrm{Var}} \leqslant 2\pi(D_0^c)$. 故为证 (7.2.2), 只需证 $\|\mu-\nu\|_{\mathrm{Var}} \geqslant 2\pi_0(D_0^c)$. 只证 $\mu \neq \nu$ 情形. 由 (7.2.1) 及 π_0 的定义知

$$\begin{aligned}\pi_0(D_0^c) &= \frac{1}{(\mu-\nu)^+(E)}\int_{D_0^c}(\mu-\nu)^+(\,\mathrm{d}x)(\mu-\nu)^-(\,\mathrm{d}y)\\ &\leqslant \frac{1}{(\mu-\nu)^+(E)}\int_{E\times E}(\mu-\nu)^+(\,\mathrm{d}x)(\mu-\nu)^-(\,\mathrm{d}y)\\ &= (\mu-\nu)^-(E) = \frac{1}{2}\|\mu-\nu\|_{\mathrm{Var}}. \quad \square\end{aligned}$$

§7.3 Wasserstein 距离

本节从一个简单的例子出发, 引入输运问题的最优运费与 L^p Wasserstein 距离, 在一定的条件下得到最优耦合的存在性与 Wasserstein 距离的完备性.

§ 7.3.1 最优输运与 Wasserstein 距离

设有 n 个地方 $x_1,\cdots,x_n$, 它们各自生产和并消费某种产品. 设某年该产品的产出与需求分布分别为 $\mu=\{\mu_1,\cdots,\mu_n\}$ 与 $\nu=\{\nu_1,\cdots,\nu_n\}$, 即 $\forall 1\leqslant i\leqslant n, \mu_i$ 与 ν_i 分别为第 i 个地方该产品的产出份额与需求份额. 我们有 $\mu_i,\nu_i\geqslant 0$ 且 $\sum\limits_{i=1}^n \mu_i = \sum\limits_{i=1}^n \nu_i = 1$. 即 μ 与 ν 可视为空间 $E=\{x_1,\cdots,x_n\}$ 上的概率测度, $\mu(\{x_i\})=\mu_i, \nu(\{x_i\})=\nu_i, 1\leqslant i\leqslant n$.

根据市场需求, 需要将产品从分布 μ 运送为 ν. 设 $\pi=\{\pi_{ij}:1\leqslant i,j\leqslant n\}$ 为一个运输方案, $\pi_{ij}\geqslant 0$ 表示从 x_i 处运往 x_j 处的产品份额. 则

$$\mu_i=\sum_{j=1}^n \pi_{ij}, \nu_i=\sum_{j=1}^n \pi_{ji}, 1\leqslant i\leqslant n.$$

从而 π 为 μ 与 ν 的一个耦合. 反之, $\forall \pi\in\mathscr{C}(\mu,\nu)$ 也对应一个运输方案.

设 $\rho_{ij}\geqslant 0$ 为将单位产品从 x_i 处输运到 x_j 处所需费用, 则任给运输方案 $\pi\in\mathscr{C}(\mu,\nu)$, 所需运费为

$$\sum_{i,j=1}^n \rho_{ij}\pi_{ij} = \int_{E\times E}\rho\,\mathrm{d}\pi.$$

从而将产品从分布 μ 运输成 ν 所需最低运费为

$$W_1^\rho(\mu,\nu) \triangleq \inf_{\pi\in\mathscr{C}(\mu,\nu)} \int_{E\times E} \rho\,\mathrm{d}\pi,$$

称为由 ρ 诱导的 μ 与 ν 之间的 L^1 Wasserstein 距离. 类似地, 可定义 L^p Wasserstein 距离.

定义 7.5. 设 (E,ρ) 为一个度量空间, $\forall p \in [1,\infty)$, 定义由 ρ 诱导的 L^p Wasserstein 距离为

$$W_p^\rho(\mu,\nu) \triangleq \inf_{\pi\in\mathscr{C}(\mu,\nu)} \left\{\int_{E\times E} \rho^p\,\mathrm{d}\pi\right\}^{1/p},\ \mu,\nu\in\mathscr{P}(E).$$

由于通常 ρ 未必有界, 因此为使 W_p^ρ 有限, 仅考虑 $\mathscr{P}(E)$ 的如下子空间:

$$\mathscr{P}_p(E) = \{\mu\in\mathscr{P}(E) : \mu(\rho(o,\cdot)) < \infty\}, p \geqslant 1.$$

这里 $o\in E$ 为某固定点. 由三角不等式易见, $\mathscr{P}_p(E)$ 的定义与 $o\in E$ 的选取无关.

§7.3.2 最优耦合与对偶公式

先证明 Polish 空间上的 L^p Wasserstein 距离的最优耦合的存在性.

定理 7.6. 设 (E,ρ) 为 Polish 空间, 则 $\forall \mu,\nu\in\mathscr{P}_p(E)$ 存在 $\pi\in\mathscr{C}(\mu,\nu)$ 使 $W_p^\rho(\mu,\nu) = \pi(\rho^p)$.

证明 由于 $\mu,\nu\in\mathscr{P}_p(E)$ 且 $\mu\times\nu\in\mathscr{C}(\mu,\nu)$, 有

$$\begin{aligned} W_p^\rho(\mu,\nu)^p &\leqslant \int_{E\times E} \rho^p(x,y)\mu(\mathrm{d}x)\nu(\mathrm{d}y) \\ &\leqslant 2^{p-1}\int_{E\times E} (\rho^p(x,o)+\rho^p(y,o))\mu(\mathrm{d}x)\nu(\mathrm{d}y) \\ &< \infty. \end{aligned}$$

故 $\forall n\geqslant 1$, 存在 $\pi_n\in\mathscr{C}(\mu,\nu)$ 使得

$$W_p^\rho(\mu,\nu)^p \geqslant \pi_n(\rho^p) - \frac{1}{n}. \tag{7.3.1}$$

因此, 如 π_n 弱收敛到某 π_0, 则 π_0 即为所求. 为此, 先证明 $\{\pi_n\}_{n\geqslant 1}$ 是胎紧的. 首先, 由定理 6.19 知有限集 $\{\mu,\nu\}$ 是胎紧的, 故 $\forall\varepsilon > 0$, 存在

紧集 $K \subset E$ 使 $\mu(K^c) + \nu(K^c) < \varepsilon$. 从而 $\forall \pi \in \mathscr{C}(\mu,\nu), \pi((K \times K)^c) \leqslant \pi(K^c \times E) + \pi(E \times K^c) = \mu(K^c) + \nu(K^c) < \varepsilon$. 故 $\mathscr{C}(\mu,\nu)$ 是胎紧的. 因而存在子列 $\{\pi_{n_k}\}_{k\geqslant 1}$ 及 $\pi_0 \in \mathscr{P}(E)$ 使得 $\pi_{n_k} \xrightarrow{w} \pi_0$ $(k \to \infty)$. 易见 $\pi_0 \in \mathscr{C}(\mu,\nu)$. 由此结合 (7.3.1) 得

$$\pi_0(\rho^p \wedge N) = \lim_{k\to\infty} \pi_{n_k}(\rho^p \wedge N) \leqslant W_p^\rho(\mu,\nu)^p.$$

再令 $N \uparrow \infty$ 即得 $\pi_0(\rho^p) \leqslant W_p^\rho(\mu,\nu)^p$. □

由 Wasserstein 距离的定义容易得到该距离的上界估计, 而下面的对偶公式则有利于获得下界估计. 为此, 先引入两个函数类. $\forall \mu,\nu \in \mathscr{P}(E)$, 令

$$\mathscr{F}_{\mu,\nu} = \left\{(f,g) : f \in L^1(\mu), g \in L^1(\nu), f(x) \leqslant g(y) + \rho(x,y)^p, \forall x,y \in E\right\}.$$

此外, 令

$$\mathscr{F}_{\mathrm{Lip}} = \left\{(f,g) : f,g \text{ Lipschitz 连续且 } \forall x,y \in E, f(x) \leqslant g(y) + \rho(x,y)^p\right\}.$$

定理 7.7 (Kontorovich 定理). 设 (E,ρ) 为 Polish 空间, 则 $\forall \mu,\nu \in \mathscr{P}_p(E)$,

$$W_p^\rho(\mu,\nu)^p = \sup_{(f,g)\in\mathscr{F}_{\mu,\nu}} \{\mu(f) - \nu(g)\} = \sup_{(f,g)\in\mathscr{F}_{\mathrm{Lip}}} \{\mu(f) - \nu(g)\}. \tag{7.3.2}$$

证明 由于 $\mathscr{F}_{\mathrm{Lip}} \subset \mathscr{F}_{\mu,\nu}$, 只需证

$$\sup_{(f,g)\in\mathscr{F}_{\mu,\nu}} \{\mu(f) - \nu(g)\} \leqslant W_p^\rho(\mu,\nu)^p \leqslant \sup_{(f,g)\in\mathscr{F}_{\mathrm{Lip}}} \{\mu(f) - \nu(g)\}.$$

$\forall (f,g) \in \mathscr{F}_{\mu,\nu}, \forall \pi \in \mathscr{C}(\mu,\nu)$, 有

$$\mu(f) - \nu(g) = \int_{E\times E} (f(x) - g(y))\pi(\,\mathrm{d}x,\,\mathrm{d}y) \leqslant \int_{E\times E} \rho(x,y)^p \pi(\,\mathrm{d}x,\,\mathrm{d}y).$$

故由 W_p^ρ 的定义知第一个不等式成立.

由于第二个不等式的证明非常繁琐, 我们略去. 感兴趣的读者可在 [6] 的第三节中找到更一般结果的详细证明, 在 [4, Chapter 5] 中也可找到 $p = 1$ 时的对偶公式. □

§ 7.3.3 $(\mathscr{P}_p(E), W_p^\rho)$ 空间

我们将讨论度量空间 $(\mathscr{P}_p(E), W_p^\rho)$ 的完备性与紧性.

定理 7.8. 设 (E,ρ) 为局部紧 Polish 空间, 则 $(\mathscr{P}_p(E), W_p^\rho)$ 也是 Polish 空间.

证明 (a) 首先证明 W_p^ρ 为度量.

显然, $W_p^\rho(\mu,\nu)=0$ 当且仅当 $\mu=\nu$, 故只需证明三角不等式. $\forall \mu_1,\mu_2,\mu_3 \in \mathscr{P}_p(E)$, 令 π_{12} 与 π_{23} 分别为 (μ_1,μ_2) 及 (μ_2,μ_3) 的最优耦合, 我们有

$$W_p^\rho(\mu_1,\mu_2)=\pi_{12}(\rho^p)^{1/p},\ W_p^\rho(\mu_2,\mu_3)=\pi_{23}(\rho^p)^{1/p}.$$

为构造 μ_1 与 μ_3 的耦合, 令 $\pi_{12}(x_1,\mathrm{d}x_2)$ 为 π_{12} 在给定 x_1 下的正则条件概率, $\pi_{23}(x_2,\mathrm{d}x_3)$ 为 π_{23} 在给定 x_2 下的正则条件概率. 则令

$$\pi_{13}(A\times B)=\mu_1(\mathrm{d}A)\int_E \pi_{23}(x_2,B)\pi_{12}(x_1,\mathrm{d}x_2).$$

易见 $\pi_{13}\in\mathscr{C}(\mu_1,\mu_3)$. 则

$$\pi(\mathrm{d}x_1,\mathrm{d}x_2,\mathrm{d}x_3)\triangleq\mu_1(\mathrm{d}x_1)\pi_{12}(x_1,\mathrm{d}x_2)\pi_{23}(x_2,\mathrm{d}x_3)$$

为 $E\times E\times E$ 上的概率测度, 且对

$$\rho_{ij}(x_1,x_2,x_3)\triangleq\rho(x_i,x_j),\ 1\leqslant i,j\leqslant 3$$

有

$$\pi(\rho_{ij}^p)=\pi_{ij}(\rho^p),\ 1\leqslant i,j\leqslant 3.$$

从而由 $L^p(\pi)$ 中的三角不等式知

$$\begin{aligned}W_p^\rho(\mu_1,\mu_3)&\leqslant\pi(\rho_{13}^p)^{1/p}\leqslant\pi((\rho_{12}+\rho_{23})^p)^{1/p}\\&\leqslant\pi(\rho_{12}^p)^{1/p}+\pi(\rho_{23}^p)^{1/p}\\&=W_p^\rho(\mu_1,\mu_2)+W_p^\rho(\mu_2,\mu_3).\end{aligned}$$

(b) 再证 W_p^ρ 的完备性.

设 $\{\mu_n\}_{n\geqslant1}\subset\mathscr{P}_p(E)$ 在 W_p^ρ 之下的 Cauchy 列, 则 $\forall$ Lipschitz 连续函数 f 及 $\pi\in\mathscr{C}(\mu_n,\mu_m)$,

$$\begin{aligned}|\mu_n(f)-\mu_m(f)|&\leqslant\int_{E\times E}|f(x)-f(y)|\pi(\mathrm{d}x,\mathrm{d}y)\\&\leqslant||f||_{\mathrm{Lip}}\int_{E\times E}\rho(x,y)\pi(\mathrm{d}x,\mathrm{d}y).\end{aligned}$$

故

$$\begin{aligned}\lim_{m,n\to\infty}|\mu_n(f)-\mu_m(f)|&\leqslant \lim_{m,n\to\infty}||f||_{\mathrm{Lip}}W_1^\rho(\mu_m,\mu_n)\\&\leqslant \lim_{m,n\to\infty}||f||_{\mathrm{Lip}}W_p^\rho(\mu_m,\mu_n)=0.\end{aligned}\tag{7.3.3}$$

由于有界 Lipschitz 连续函数在 $C_b(E)$ 中稠, 在弱拓扑度量 d_w 的定义中可令 f_n 为 Lipschitz 连续函数, 从而由 (7.3.3) 知 $\{\mu_n\}_{n\geqslant 1}$ 在 d_w 之下也是 Cauchy 列, 故存在 $\mu\in\mathscr{P}(E)$ 使 $\mu_n\xrightarrow{w}\mu$. 另一方面, 给定 $o\in E$, 有

$$\mu_n(\rho(o,\cdot)^p)\leqslant 2^{p-1}\mu_1(\rho(o,\cdot)^p)+2^{p-1}W_p^\rho(\mu_1,\mu_n)^p$$

对 $n\geqslant 1$ 有界, 故 $\exists C>0$ 使 $\forall N\geqslant 1$

$$\mu(\rho(o,\cdot)^p\wedge N)=\lim_{n\to\infty}\mu_n(\rho(o,\cdot)^p\wedge N)\leqslant C.$$

从而 $\mu\in\mathscr{P}_p(E)$ 且

$$\varliminf_{n\to\infty}\mu_n(\rho(o,\cdot)^p)\geqslant\mu(\rho(o,\cdot)^p).\tag{7.3.4}$$

此外, $\forall\varepsilon>0,\exists n_0\geqslant 1$ 使 $W_p(\mu_{n_0},\mu_n)^p\leqslant\varepsilon,\forall n\geqslant n_0$. 则

$$\begin{aligned}&\mu_n((N-\rho(o,\cdot)^p)^+)\\&\leqslant\mu_{n_0}((N-\rho(o,\cdot)^p)^+)+|\mu_n((N-\rho(o,\cdot)^p)^+)-\mu_{n_0}((N-\rho(o,\cdot)^p)^+)|\\&\leqslant\mu_{n_0}((N-\rho(o,\cdot)^p)^+)+2^{p-1}W_p^\rho(\mu_n,\mu_{n_0})^p\\&\leqslant\mu_{n_0}((N-\rho(o,\cdot)^p)^+)+2^{p-1}\varepsilon.\end{aligned}$$

故

$$\varlimsup_{n\to\infty}\mu_n(\rho(o,\cdot)^p)\leqslant\varlimsup_{n\to\infty}\mu_n(\rho(o,\cdot)^p\wedge N)+2^{p-1}\varepsilon.$$

由 ε 对任意性及 $\mu_n\xrightarrow{w}\mu$ 知

$$\varlimsup_{n\to\infty}\mu_n(\rho(o,\cdot)^p)\leqslant\varlimsup_{n\to\infty}\mu_n(\rho(o,\cdot)^p\wedge N)=\mu(\rho(o,\cdot)^p\wedge N)\leqslant\mu(\rho(o,\cdot)^p).$$

由此结合 (7.3.4) 得 $\mu(\rho(o,\cdot)^p)=\lim_{n\to\infty}\mu_n(\rho(o,\cdot)^p)$. 从而 (见习题 6) $\lim_{n\to\infty}W_p^\rho(\mu_n,\mu)=0$.

(c) 最后证明 W_p^ρ 的可分性.

$\forall N \geqslant 1$, 令

$$\mathscr{P}_p^{(N)}(E) = \left\{\mu \in \mathscr{P}_p(E) : \mathrm{supp}\mu \subset \bar{B}(o, N)\right\},$$

其中 $\bar{B}(o,N)$ 为以 o 为中心以 N 为半径的闭球. 由于 $\forall \mu \in \mathscr{P}_p(E)$, 易证当 $N \to \infty$,

$$\mu_N \triangleq \frac{\mu(\cdot \cap B(o,N)}{\mu(B(o,N))} \xrightarrow{W_p^\rho} \mu,$$

则 $\bigcup\limits_{N=1}^{\infty} \mathscr{P}_p^{(N)}(E)$ 在 $(\mathscr{P}_p(E), W_p)$ 中稠密. 因此只需证明每个 $\mathscr{P}_p^{(N)}(E)$ 是可分的. 由于 $\rho(o,\cdot)$ 在 $\bar{B}(o,N)$ 上有界, 由习题 6 知在 $\mathscr{P}_p^{(N)}(E)$ 上弱拓扑与 W_p^ρ 拓扑等价. 而由定理 7.1 知 $\mathscr{P}_p^{(N)}(E)$ 在弱拓扑之下上可分的, 因此也在 W_p^ρ 之下可分. $\square$

定理 7.9. 设集合 $\mathfrak{M} \subset \mathscr{P}_p(E)$ 在 W_p^ρ 之下是紧的当且仅当它是弱紧的且

$$\lim_{N\to\infty} \sup_{\mu\in\mathfrak{M}} \mu\left(\rho(o,\cdot)^p \mathbf{1}_{\{\rho(o,\cdot)\geqslant N\}}\right) = 0. \tag{7.3.5}$$

证明 (a) 必要性. 易见$W_p^\rho(\mu_n,\mu) \to \infty$ 蕴含 $\forall$ Lipschitz 连续函数 f 有 $\mu_n(f) \to \mu(f)$, 从而 W_p^ρ 所诱导的拓扑强于弱拓扑, 故 $\mathfrak{M}$ 如在 W_p^ρ 之下紧则也必然弱紧. 因此只需证明 $\mathfrak{M}$ 在 W_p^ρ 之下的紧性蕴含 (7.3.5).

由于 $\mathfrak{M}$ 在 W_p^ρ 之下紧, 则 $\forall \varepsilon > 0$, 存在 $\mu_1, \cdots, \mu_n \in \mathfrak{M}$ 使得

$$\min_{1\leqslant i\leqslant n} W_p^\rho(\mu_i,\mu)^p < \varepsilon, \quad \forall \mu \in \mathfrak{M}.$$

从而 $\forall \mu \in \mathfrak{M}$,

$$\begin{aligned}\mu\left((\rho(o,\cdot)^p - N)^+\right) &\leqslant \max_{1\leqslant i\leqslant n} \mu_i\left((\rho(o,\cdot)^p - N)^+\right) + 2^{p-1} W_p^\rho(\mu_i,\mu)^p \\ &\leqslant \sum_{i=1}^n \mu_i\left((\rho(o,\cdot)^p - N)^+\right) + 2^{p-1}\varepsilon.\end{aligned}$$

故

$$\overline{\lim_{N\to\infty}} \sup_{\mu\in\mathfrak{M}} \mu\left(\rho(o,\cdot)^p \mathbf{1}_{\{\rho(o,\cdot)\geqslant N\}}\right) \leqslant 2 \overline{\lim_{N\to\infty}} \sup_{\mu\in\mathfrak{M}} \mu\left(\left(\rho(o,\cdot)^p - \frac{N}{2}\right)^+\right) \leqslant 2^p\varepsilon.$$

由 ε 的任意性立得 (7.3.5).

(b) 充分性. 设 $\mathfrak{M}$ 弱紧且 (7.3.5) 成立, 往证 $\mathfrak{M}$ 在 W_p^ρ 之下紧. 为此, 只需证明 $\forall$ 序列 $\{\mu_n\}_{n\geqslant 1} \subset \mathfrak{M}$, 必存在 W_p^ρ 之下的收敛子列. 由 $\mathfrak{M}$ 的弱紧性, 不妨设 $\mu_n \xrightarrow{w} \mu$. 设 $\{x_1, x_2, \cdots\}$ 为 E 的稠子集, 则 $\forall \varepsilon > 0$, 有 $\bigcup\limits_{i=1}^\infty B(x_i, \varepsilon) \supset E$, 其中 $B(x_i,\varepsilon)$ 为以 x_i 为中心以 ε 为半径的开球. 由于集合 $\big\{\varepsilon > 0 : \exists i \geqslant 1 \text{ 使 } \mu(\partial B(x_i, \varepsilon)) > 0\big\}$ 至多可数, $\forall m \geqslant 1$, 可取 $\varepsilon_m \in (0, 1/m)$ 使 $B(x_i, \varepsilon_m)$ 均为 μ 连续集. 令

$$U_i = B(x_i, \varepsilon_m), U_{i+1} = B(x_{i+1}, \varepsilon_m) \setminus \left(\bigcup_{j=1}^{i} B(x_j, \varepsilon_m)\right),$$

则 $\{U_i\}_{i\geqslant 1}$ 为一列两两不交的 μ 连续集, $\sum\limits_{i=1}^\infty U_i = E$ 且 U_i 的半径小于 $1/m$. 令$r_n = \sum\limits_{i=1}^\infty \mu_n(U_i) \wedge \mu(U_i)$, 则 $r_n \in [0,1]$ 且 $\lim_{n\to\infty} r_n = 1$. 令

$$\begin{aligned} Q_n(\,\mathrm{d}x) &= \mu_n(\,\mathrm{d}x) - \sum_{i=1}^\infty \frac{\mu_n(U_i) \wedge \mu(U_i)}{\mu_n(U_i)} \mathbf{1}_{U_i}(x)\mu_n(\,\mathrm{d}x), \\ Q(\,\mathrm{d}x) &= \mu(\,\mathrm{d}x) - \sum_{i=1}^\infty \frac{\mu_n(U_i) \wedge \mu(U_i)}{\mu(U_i)} \mathbf{1}_{U_i}(x)\mu(\,\mathrm{d}x). \end{aligned}$$

则

$$\begin{aligned} \pi_n(\,\mathrm{d}x,\,\mathrm{d}y) \triangleq & \sum_{i=1}^\infty \mathbf{1}_{U_i}(x)\mathbf{1}_{U_i}(y) \frac{\mu_n(U_i)\wedge\mu(U_i)}{\mu_n(U_i)\mu(U_i)} \mu_n(\,\mathrm{d}x)\mu(\,\mathrm{d}y) \\ & + \frac{1}{1-r_n} Q_n(\,\mathrm{d}x)Q(\,\mathrm{d}y) \end{aligned}$$

为 μ_n 与 μ 的耦合 (如 $r_n = 1$ 则令上式最后一项为 0), 从而

$$\begin{aligned} W_p^\rho(\mu_n, \mu)^p \leqslant \pi_n(\rho^p) &\leqslant m^{-p} + \frac{2^{p-1}}{1-r_n}\left(Q_n(\rho(o,\cdot)^p) + Q(\rho(o,\cdot)^p)\right) \\ &\leqslant m^{-p} + 2^p N^p (1-r_n) + 2^{p-1} \sup_{k\geqslant 1} \mu_k\left(\rho(o,\cdot)^p \mathbf{1}_{\{\rho(o,\cdot)\geqslant N\}}\right) \\ &\quad + 2^{p-1}\mu\left(\rho(o,\cdot)^p \mathbf{1}_{\{\rho(o,\cdot)\geqslant N\}}\right). \end{aligned}$$

先令 $n \to \infty$, 再令 $N \to \infty$, 最后令 $m \to \infty$ 得 $W_p^\rho(\mu_n, \mu) \to 0 (n \to \infty)$. □

§7.4 补充与习题

1. 设 (E,ρ) 为 Polish 空间. 试写出与淡收敛拓扑等价的度量, 并对等价性加以证明. 该度量是否完备?

2. 设 $(E,\mathscr{E})$ 为可测空间, 证明 $\forall \mu,\nu\in\mathscr{P}(E)$ 有

$$||\mu-\nu||_{\mathrm{Var}}=2\sup_{A\in\mathscr{E}}|\mu(A)-\nu(A)|=|\mu-\nu|(E).$$

3. 设 $(E,\mathscr{E})$ 上的可测函数 $V\geqslant 1$. $\forall \mu\in\mathscr{P}(E)$, 定义加权变差

$$||\mu||_V\triangleq\int_E V(x)\mu(\,\mathrm{d}x).$$

证明

$$||f||_V\triangleq\sup_{\mu\in\mathscr{P}(E)}\int_E f\,\mathrm{d}\mu=\sup_{x\in E}\frac{|f(x)|}{V(x)}.$$

4. 设 (E,ρ) 为 Polish 空间, $\mathscr{P}(E)$ 为其 Borel σ 代数上所有概率测度全体. 令 $d_{\mathrm{Var}}(\mu,\nu)=||\mu-\nu||_{\mathrm{Var}}$. 证明 $(\mathscr{P}(E),d_{\mathrm{Var}})$ 是完备度量空间, 并举例说明它未必可分.

5. 设 $(E,\mathscr{E})$ 为可测空间, $(\Omega,\mathscr{A},\mathbb{P})$ 是概率空间. 如果 ξ 为 Ω 到 E 上可测映射, 则称之为 E 上一个随机变量, 其分布 $\mathbb{P}_\xi=\mathbb{P}\circ\xi^{-1}$ 为 E 上一个概率测度. 今给定 $\mu,\nu\in\mathscr{P}(E)$ 以及 $\pi\in\mathscr{C}(\mu,\nu)$, 试构造 E 上的两个随机变量 ξ,η, 使得 $\mathbb{P}_\xi=\mu,\mathbb{P}_\eta=\nu$ 且 $\mathbb{P}_{(\xi,\eta)}=\pi$.

6. 设 (E,ρ) 为 Polish 空间, $\{\mu_n\}_{n\geqslant 1}\subset\mathscr{P}_p(E)$. 证明 $W_p^\rho(\mu_n,\mu)\to 0$ 当且仅当 $\mu_n\xrightarrow{w}\mu$ 且 $\lim_{n\to\infty}\mu_n(\rho(o,\cdot)^p)=\mu_n(\rho(o,\cdot)^p)$, 其中 $o\in E$ 为一固定点.

7. (Lévy 距离) 证明在分布函数空间上

$$\rho(F,G)=\inf\{\varepsilon\geqslant 0:F(x-\varepsilon)-\varepsilon\leqslant G(x)\leqslant F(x+\varepsilon)+\varepsilon,\forall x\}$$

定义了一个距离, 并且$\rho(F_n,F)\to 0$ 当且仅当 $F_n\Rightarrow F$.

8. (Ky Fan 距离) 对随机变量 ξ, η 定义

$$\alpha(\xi,\eta) = \inf\{\varepsilon \geqslant 0 : \mathbb{P}(|\xi-\eta| > \varepsilon) \leqslant \varepsilon\}.$$

证明若 $\alpha(\xi,\eta)=\alpha$, 则对相应的分布函数有 $\rho(F_\xi, F_\eta) \leqslant \alpha$.

9. 令

$$\beta(\xi,\eta) = \mathbb{E}\left(\frac{|\xi-\eta|}{1+|\xi-\eta|}\right).$$

证明若 $\alpha(\xi,\eta)=\alpha$, 则 $\alpha^2/(1+\alpha) \leqslant \beta(\xi,\eta) \leqslant \alpha + (1-\alpha)\alpha/(1+\alpha)$.

参考文献

[1] Billingsley, P. *Probability and Measure.* Third Edition. Wiley, 1995.

[2] 陈木法. 跳过程与粒子系统. 北京师范大学出版社, 1986.

[3] Chen, Mu-Fa. *From Markov Chains to Non–equilibrium Particle Systems.* Second edition. World Scientific, 2004.

[4] Chen, Mu-Fa. *Eigenvalues, Inequalities, and Ergodic Theory.* Springer, 2005.

[5] Neveau, J. *Mathematical Foundations of the Calculus of Probability.* Holden-Day, 1965.

[6] Rachev, S. *The Monge-Kantonovich mass transference problem and its stochastic applications.* Theory of Probability and Applications, Vol. XXIX, 1985, 647-676.

[7] Reed, M. Simon, B. *Mehtod of modern amthematical physics (I).* Academic Press, 1972.

[8] Shiryayev, A N. *Probability.* Springer-Verlag, 1984.

[9] Sinclair, A. *Algorithms for Random Generation and Counting: a Markov chain approach.* Birkhäuser, 1993.

[10] 严加安. 测度论讲义(第二版). 科学出版社, 2004.

[11] 严士健、刘秀芳. 测度与概率. 北京师范大学出版社, 2003.

[12] 严士健、王隽骧、刘秀芳. 概率论基础. 科学出版社, 1982.

[13] Yosida, K. *Functional Analysis.* Springer-Verlag, 1980.

[14] 张恭庆、林源渠. *泛函分析讲义*. 上册, 北京大学出版社, 1990.

[15] 张恭庆、郭懋正. *泛函分析讲义*. 下册, 北京大学出版社, 1990.

索引