《测度论与概率论基础》（程士宏）

概要和笔记

章节目录

Chapter1 可测空间和可测映射

对于一个集合 X，我们希望衡量和比较它的子集的大小，因此测度是一个集合函数。为使测度满足一些性质，测度的定义域不能达到 X 的所有子集组成的集合这么大。在集合类中，鉴于 σ 域的优良性质，我们选择一个 X 上的 σ 域 $\mathscr{F}$ 作为测度的“定义域”，将 $(X, \mathscr{F})$ 称为可测空间。有了可测空间后，一方面我们可以定义可测函数（$\S 5$），另一方面可以建立测度（Chapter2）。

§1 集合及其运算

元素为集合的集合称为**集合系（类）**，由于测度定义在集合类上，因此首先需要对集合类做简要的讨论。

设 $\{A_n, n = 1, 2, ...\}$ 为集合序列，

定义

- $\forall n, A_n \subset A_{n+1}$，称 $\{A_n\}$ 非降，记为 $A_n \uparrow$
- $\forall n, A_n \supset A_{n+1}$，称 $\{A_n\}$ 非增，记为 $A_n \downarrow$
- $\limsup\limits_{n\to\infty} A_n := \bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty} A_k$; $\liminf\limits_{n\to\infty} A_n := \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{k=n}^{\infty} A_k$

例

- $\bigcup\limits_{k=n}^{\infty} A_k \downarrow$; $\bigcap\limits_{k=n}^{\infty} A_k \uparrow$

§2 集合系

特殊集合类

- 代数（有限交、补封闭）$\Rightarrow$ 环（有限并、差封闭）$\Rightarrow$ 半环（有限交封闭，有包含关系的两个集合之差可以写成有限个半环中不交集合的并）$\Rightarrow$ π 类（有限交封闭）
- σ 代数（含全集，可列并、补封闭）$\Rightarrow$ λ 类（含全集，有限差、单调增封闭）$\Rightarrow$ 单调系（可列单调的极限封闭）

例

$$\mathscr{D}_R := \{(a,b] : a,b \in R\} \text{ 为 } R \text{ 上的半环}$$

关系

- 代数 $+$ 单调系 $=$ σ 代数（既是代数又是单调系的集合类是 σ 代数）
- λ 类 $+$ π 类 $=$ σ 代数

对于一个集合 X，我们希望度量它的子集的大小，但不是任意的子集都可以定义大小，今后将 X 和其上的某个 σ 代数 $\mathscr{F}$ 放一起 $(X,\mathscr{F})$ 称为**可测空间**，（$\mathscr{F}$ 中的元素就是“可测集”）。

§3 σ 域的生成

包含集合类 $\mathscr{G}$ 的最小 σ 域 $\mathscr{F}$ 称为 $\mathscr{G}$ **的生成 σ 域**，记为 $\mathscr{F} = \sigma(\mathscr{G})$

生成的环、单调系和 λ 系可同理定义，分别记为 $r(\mathscr{G}), m(\mathscr{G}), l(\mathscr{G})$

任意集合类的生成环/单调系/λ 系/σ 域都是存在的

定理

- $\mathscr{A}$ 为代数，则 $\sigma(\mathscr{A}) = m(\mathscr{A})$
- $\mathscr{A}$ 为 π 类，则 $\sigma(\mathscr{A}) = l(\mathscr{A})$

R 上的 Borel 集合系

$$\mathscr{B}_R := \sigma(\mathscr{D}_R)$$

§4 可测映射和可测函数

本节不涉及可测函数深入的性质（参见 Chapter2 §5 和 Chapter3），只是在可测空间上引入了这一重要概念而已。

对于两个集合 X, Y，$f: X \to Y$ 为映射，

定义（原像）

- 集合的原像：$\forall B \subset Y, f^{-1}B := \{f \in B\} = \{x : f(x) \in B\}$ （使像落在 B 中的 x）
- 集合类的原像：$\forall Y$ 上的集合系 $\mathscr{G}, f^{-1}\mathscr{G} := \{f^{-1}B : B \in \mathscr{G}\}$ （将 $\mathscr{G}$ 中的每个元素的原像放在一起形成的集合类）

回忆实变函数中关于 Lebesgue 可测函数的定义：将 $f(x)$ 的值域做分割，每个小区间的原像作为一个集合都是 Lebesgue 可测的。

定义（可测函数）

$(X, \mathscr{F}), (Y, \mathscr{G})$ 为可测空间，

- 可测映射：$f: X \to Y$ 满足：$f^{-1}\mathscr{G} \subset \mathscr{F}$
- 可测映射的生成 σ 域：$\sigma(f) := f^{-1}\mathscr{G}$，它是使 f 可测的最小 σ 域
- 可测函数：设 $\bar{R} := R \cup \{\infty, -\infty\}$，从 $(X, \mathscr{F})$ 到 $(\bar{R}, \mathscr{B}_{\bar{R}})$ 的可测映射为可测函数
- 有限值可测函数/随机变量：从 $(X, \mathscr{F})$ 到, $(R, \mathscr{B}_R)$ 的可测映射为随机变量

§4 可测映射和可测函数

若 $\mathscr{G}$ 是一个生成 σ 域，即 $\mathscr{G} = \sigma(\mathscr{A})$，则只需要检查 $\mathscr{A}$ 中的元素的原像是否 $\in \mathscr{F}$ 即可，于是我们有：

可测函数的判定

f 为 $(X, \mathscr{F})$ 上的可测函数当且仅当 $\{f < a\} \in \mathscr{F}, \forall a \in R$

§5 可测函数的运算

可测函数的四则运算还是可测函数

可测函数列 $\{f_n, n = 1, 2, ...\}$ 的上下确界和上下极限依然是可测函数（注：在一个给定的点 $x \in X$ 上，$f_n(x)$ 构成一个数列，该数列的下确界为 $\inf_{n\to\infty} f_n$ 在 x 上的取值. 其他同理）。

简单函数

$\{A_k, k = 1, 2, ...\}$ 为 X 的有限分割（需要 $A_i \in \mathscr{F}$），a_k 为一些实数，则 $f = \sum_{k=1}^{n} a_k I_{A_k}$ 为简单函数

简单函数虽然简单，但可测函数可以用简单函数来逼近：

定理

- 对于非负可测函数 f，存在非负简单函数 $\{f_n\}$ $s.t.$ $f_n \uparrow f$
- 对于可测函数 f，存在简单函数 $\{f_n\}$ $s.t.$ $f_n \to f$
- 若 f 有界，则上述两个收敛可以加强为一致收敛

由此引出证明中的一种典型方法：先证明命题对非负简单函数成立，再推广到非负可测函数，最后推广到一般的可测函数。

Chapter2 测度空间

将集合系上可列可加的非负集函数定义为测度。但作为性质较优的集合类，σ 域的结构也较为复杂，有时很难直接在 σ 域上定义测度。我们先考虑在 X 的半环上定义测度，并用它导出 X 上的外测度，测度扩张定理保证了将这个外测度限制在半环生成的 σ 域上后将“升级”为测度。Lebesgue-Stieljes 测度的构造即为一个例子。

在测度空间 $(X, \mathscr{F}, \mu)$ 上，可测函数具有多种收敛性。另外，引入测度后，测度为零的集合常被我们忽略掉。

§1 测度的定义及性质

定义（测度：可列可加的非负集函数）

$\mathscr{E}$ 为 X 上的集合系（不一定是 σ 域），$\varnothing \in \mathscr{E}, \mathscr{E}$ 上的非负集函数 μ 满足：$\mu(\varnothing) = 0, \forall$ 可列个两两不交的集合 $A_1, A_2 \ldots \in \mathscr{E}$，只要 $\bigcup_{n=1}^{\infty} A_n \in \mathscr{E}$ 就有

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

则称 μ 为 $\mathscr{E}$ 上的测度

定义（有限和 σ 有限）

- 有限：$\forall A \in \mathscr{E}, \mu(A) < \infty$
- σ 有限：$\forall A \in \mathscr{E}$，$\exists$ 一列 $A_k \in \mathscr{E}$ 使得 $\mu(A_k) < \infty$ 且 $A \subset \bigcup_{k=1}^{\infty} A_k$（测度无穷的集合可以用可数个测度有限的集合覆盖起来）

§1 测度的定义及性质

例

- $(X, \mathscr{F})$ 为可测空间，给定 $x \in X$，集函数 $\delta_x(A) = I_A(x), A \in \mathscr{F}$ 为测度
- $\mathscr{D}_R$ 为半环，给定非降右连续函数（准分布函数）F，$\forall a, b \in R$，定义集函数

$$\mu((a,b]) = \begin{cases} F(b) - F(a) & a < b \\ 0 & a \geq b \end{cases}$$

则 μ 为 $\mathscr{D}_R$ 上的测度

将集合（空间）X，X 的部分子集生成的 σ 域 $\mathscr{F}$，以及 $\mathscr{F}$ 上的测度 μ 三位一体 $(X, \mathscr{F}, \mu)$ 称为**测度空间**

若 $\mu(X) = 1$，则称为概率空间

§1 测度的定义及性质

半环上的测度有好的性质：可列可加性，半可列可加性，有限可加性，半有限可加性，可减性，单调性，上连续性，下连续性

σ 域作为最“完整”的集合类，我们希望建立 σ 域上的测度，但在一般的 σ 域上建立测度并不容易，因此可以考虑现在半环上建立测度，再将它扩张到半环生成的 σ 域上去

§2 外测度

定义（外侧度）

$\mathscr{T}$ 为 X 的所有子集组成的集合类，$\tau : \mathscr{T} \to \bar{R}$ 为外测度，如果：$\tau(\varnothing) = 0$，单调性，半可列可加性

外侧度很容易生成！

由 μ 生成的外测度

$\mathscr{E}$ 为集合系，$\varnothing \in \mathscr{E}$，$\mu$ 为 $\mathscr{E}$ 上的非负集函数且 $\mu(\varnothing) = 0$，令

$$\tau(A) = inf\left\{\sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathscr{E}, A \subset \bigcup_{n=1}^{\infty} B_n\right\}, A \in \mathscr{T}$$

则 τ 为外测度（用最省的方式覆盖 A，若盖不住则取 ∞）

得到 X 上的外侧度后，我们尝试把 $\mathscr{T}$ 缩小一点，使得 τ 成为一个真正的测度

§2 外测度

定理（Caratheodory）

- A 为 X 的一个子集，若 $\forall D \in \mathscr{T}, \tau(D) = \tau(D \cap A) + \tau(D \cap A^c)$，则 A 为 τ-可测的。全体 τ-可测的集合放在一起称为 $\mathscr{F}_\tau$
- $\mathscr{F}_\tau$ 是一个 σ 域，$(X, \mathscr{F}_\tau, \tau)$ 是**完全测度空间**（零测集的子集还属于 $\mathscr{F}_\tau$）

总结起来，先随随便便用一个集合系和非负集函数定义一个 X 上的外侧度，再把 $\mathscr{T}$ 以一定的方式缩小一点，就可以得到一个测度空间

然而，$\mathscr{F}_\tau$ 似乎是不受我们控制的，我们既不知道 $\mathscr{F}_\tau$ 和 $\mathscr{E}$ 有无大小关系，又不知道 $\mathscr{F}_\tau$ 是不是足够大了。这是因为我们用于建立外测度的 $\mathscr{E}$ 和 μ 太随便了

§3 测度的扩张

测度扩张定理

- 对于半环 $\mathscr{D}$ 上的测度 μ，用 μ 去生成 X 上的外测度 τ，则 $(X, \sigma(\mathscr{D}), \tau)$ 为测度空间且 $\forall A \in \mathscr{D}, \mu(A) = \tau(A)$
- 若 X 可以表示成 $\mathscr{D}$ 中一列两两不交的测度有限的集合的并，则上述 τ 唯一

Lebesgue-Stieljes 测度

- 前面已经指出，给定准分布函数 F 后，μ 是半环 $\mathscr{D}_R$ 上的一个测度，将它进行扩张，得到 $\mathscr{B}_R$ 上的测度 λ_F，称为 Lebesgue-Stieljes 测度
- $(R, \mathscr{B}_R, \lambda_F), (R, \mathscr{F}_{\lambda_F}, \lambda_F)$ 都是测度空间
- 若 $F(x) = x$，则为 Lebesgue 测度

§4 测度空间的完全化

显然 $\sigma(\mathscr{D}) \subset \mathscr{F}_\tau$，一个自然的问题是：$\mathscr{F}_\tau$ 会比 $\sigma(\mathscr{D})$ 小多少（L-S 测度中即是比较 $\mathscr{B}_R$ 和 $\mathscr{F}_{\lambda_F}$）? 这涉及到测度空间的完全化

定理（测度空间完全化）

对于任意测度空间 $(X, \mathscr{F}, \mu)$，定义

$$\tilde{\mathscr{F}} := \{A \cup N : A \in \mathscr{F}, N \subset B \in \mathscr{F}, \mu(B) = 0\}$$

$$\tilde{\mu} := \tilde{\mu}(A \cup N) = \mu(A), A \in \mathscr{F}$$

则 $(X, \tilde{\mathscr{F}}, \tilde{\mu})$ 为完全测度空间

上述操作比较直观，即 $\mathscr{F}$ 中的集合并上一些零测集，形成一个新的 σ 域；而 $\tilde{\mu}$ 和 μ 也基本相等

可以证明：若 τ 为半环 $\mathscr{D}$ 上 σ 有限的测度 μ 生成的外测度，则 $(X, \mathscr{F}_\tau, \tau)$ 是 $(X, \sigma(\mathscr{D}), \tau)$ 的完全化，亦即 $\mathscr{F}_\tau$ 和 $\sigma(\mathscr{D})$ 差不多一样大

§5 可测函数的收敛性

在测度空间 $(X, \mathscr{F}, \mu)$ 中，可测函数的性质变得更为丰富

定义（几种收敛性）

- 几乎处处收敛 (a.s.)：去掉零测集后逐点收敛，即 $\mu\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|f_n - f| \geq \varepsilon\}\right) = 0$
- 几乎一致收敛 (a.u.)：去掉零测集后一致收敛，即
$$\lim_{n\to\infty} \mu\left(\bigcup_{n=m}^{\infty}\{|f_n - f| \geq \varepsilon\}\right) = 0$$
- 依测度收敛 (μ)：不收敛的测度趋于 0，即 $\lim_{n\to\infty} \mu(\{|f_n - f| \geq \varepsilon\}) = 0$

由强至弱：

1. $a.u \Rightarrow a.e, a.u. \Rightarrow \mu$，而 a.e. 和 μ 没有包含关系
2. 若测度 μ 有限（$\mu(X) < \infty$，例如概率测度），则 $a.u. \Leftrightarrow a.e. \Rightarrow \mu$

由弱至强：

1. 依测度收敛的子列可以达到 a.u.

§5 可测函数的收敛性

在概率空间中，a.u. 和 a.e. 等价，它们即是熟知的几乎必然收敛 (a.s.)；依测度收敛即为熟知的依概率收敛

熟知：几乎必然收敛强于依概率收敛

定义随机变量的**分布函数**：

$$F(x) := P(f \leq x)$$

记为 $f \sim F$. 若 $\{f_n \sim F_n\}$，且在每个连续点上 $F_n(x) \to F(x)$，则称为 $\{F_n\}$ 弱收敛 (w) 于 F，若随机变量 $f \sim F$，则称 $\{f_n\}$ 依分布收敛 (D) 于 f

熟知：依概率收敛强于依分布收敛

Chapter3 积分

我们希望在测度 μ 这个度量方式上定义出可测函数 f 的积分，可以采用逐步定义的方法。积分的性质与高等数学的相关结论吻合，另外，三大收敛定理说明了函数列的极限与其积分的关系。在 L_p 空间上，我们又引入了两种收敛模式。最后，介绍了概率论中期望，一致可积等概念的积分定义。

§1 积分的定义

测度空间 $(X, \mathscr{F}, \mu)$ 上的积分与实变函数中的几乎一样，采用逐步的方法完成

定义（积分）

- 非负简单函数的积分：设 $f = \sum_{i=1}^{n} a_i I_{A_i}$，定义

$$\int_X f d\mu := \sum_{i=1}^{n} a_i \mu(A_i)$$

- 非负可测函数的积分：定义

$$\int_X f d\mu := sup \left\{ \int_X g d\mu : g \ is \ a \ non-negative \ simple \ function \ , \ g \leq f \right\}$$

§1 积分的定义

定义（积分）

- 可测函数的积分：定义函数的正部 $f^+ := max(f,0)$，负部 $f^- := max(-f,0)$，若 $min\left(\int_X f^+ d\mu, \int_X f^- d\mu\right) < \infty$，则称 f 的积分存在；若 $max\left(\int_X f^+ d\mu, \int_X f^- d\mu\right) < \infty$，则称 f 的可积，定义

$$\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu$$

- 一般积分区域的积分：$A \in \mathscr{F}$，定义

$$\int_A f d\mu := \int_X f I_A d\mu$$

测度空间上的积分与黎曼积分的基本性质几乎没有差异，因此直接快进到一些重要的性质

§2 积分的性质

定理（绝对连续性）

f 可积，则 $\forall\varepsilon>0,\exists\delta>0,\ s.t.\ \mu(A)<\delta,\ A\in\mathscr{F}\Rightarrow\int_A|f|d\mu<\varepsilon$

直接快进到三个收敛定理：

定理

- Levi：$\{f_n\},f\ a.e.$ 非负，$f_n\uparrow f\ a.e.$，则

$$\lim_{n\to\infty}\int_X f_n d\mu=\int_X f d\mu$$

- Fatou：$\{f_n\}\ a.e.$ 非负，则

$$\int_X\left(\liminf_{n\to\infty}f_n\right)d\mu\le\liminf_{n\to\infty}\int_X f_n d\mu$$

（菜鸡联手的效率还不如正常人分开工作的下限）

§2 积分的性质

定理

- Lebesgue 控制收敛：$\{f_n\}, f$ 可测，g 非负可积，$|f_n| \le g \ a.e$，还有 $f_n \xrightarrow{a.e.} f$ 或 $f_n \xrightarrow{\mu} f$，则
$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$$

最后，设 $g : (X, \mathscr{F}, \mu) \to (Y, \mathscr{P})$ 为可测映射，可以用 g 诱导出一个 $(Y, \mathscr{P})$ 上的测度：$\forall B \in \mathscr{P}, \nu(B) := \mu(g^{-1}B)$，则 $(Y, \mathscr{P}, \nu)$ 为测度空间且积分“换元法”仍成立

§3 空间 $L_p(X, \mathscr{F}, \mu)$

定义（L_p 空间）

测度空间 $(X, \mathscr{F}, \mu), 1 \le p < \infty$，定义

$$L_p(X, \mathscr{F}, \mu) := \left\{ f : f \ measurable, \int_X |f|^p d\mu < \infty \right\}$$

注：将 a.e. 相等的 f 都看成 L_p 中的一个元素

$1 \le p < \infty$ 时，定义范数

$$||f||_p := \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}}$$

为了将 p 的范围扩展到 $(0, \infty]$，下一页简要梳理上述为什么是范数

§3 空间 $L_p(X,\mathscr{F},\mu)$

显然，$||f||_p$ 满足正定性和齐次性，故只需证明其满足三角不等式

定理

- Holder 不等式：$1<p,q<\infty,\dfrac{1}{p}+\dfrac{1}{q}=1,f,g\in L_p$，则
$$||fg||_1\le ||f||_p||g||_q$$
- Minkowski 不等式：$1\le p<\infty,f,g\in L_p$，则
$$||f+g||_p\le ||f||_p+||g||_p$$

下面对 p 扩展：
若 $p=\infty$，定义 L_p 为 a.e. 有界的可测函数的集合，定义
$$||f||_\infty := inf\{a\in R^+ : \mu(|f|>a)=0\}$$
(最紧的界)。不难发现这一定义是对 p 自然的推广，依然满足上述两个定理

§3 空间 $L_p(X, \mathscr{F}, \mu)$

若 $0 < p < 1$，定义：

$$||f||_p := \int_X |f|^p d\mu$$

它不是范数（只差齐次性）。

由于 Minkowski 不等式始终成立（$0 < p \leq \infty$），可以定义距离，进一步，有：$0 < p < 1$ 时 L_p 为完备的距离空间；$1 \leq p \leq \infty$ 时 L_p 为 Banach 空间。

§3 空间 $L_p(X, \mathscr{F}, \mu)$

在 L_p 空间上，可以引出另外两种收敛模式

定义（L_p 收敛）

- p 阶平均收敛（L_p）：$0 < p < \infty$，若 $\lim\limits_{n\to\infty} ||f_n - f||_p = 0$，则 $f_n \xrightarrow{L_p} f$
- L_p 弱收敛（(w)L_p）：$1 < p < \infty$ 或 $p = 1$ 且测度空间 σ 有限，若 $\exists f \in L_p$，使

$$\lim_{n\to\infty} \int_X f_n g d\mu = \int_X f g d\mu, \ \forall g \in L_q$$

则 $f_n \xrightarrow{(w)L_p} f$

关系：
1. p 阶平均收敛略强于依测度收敛
2. p 阶平均收敛强于 L_p 弱收敛

补注：两个记号

我们在这页补注两个记号，主要是为了方便起见，一部分是关于 L-S 积分的，另一部分是关于后续累次积分怎么才能看得更清楚

对准分布函数 F，在半环 $\mathscr{D}_R$ 上建立测度，再唯一扩张到 $\mathscr{B}_R$ 上成为 L-S 测度 λ_F，即 $(R, \mathscr{B}_R, \lambda_F)$ 为测度空间。对于可测函数 g，定义

$$\int_R g dF := \int_R g d\lambda_F$$

另外，对于测度空间 $(X, \mathscr{F}, \mu)$，定义

$$\int_X f(x)\mu(dx) := \int_X f d\mu$$

§4 概率空间的积分

本节用积分定义了随机变量的期望等重要概念，最后定义了一些与“一致”有关的概念

定义（随机变量函数的期望）

f 为 $(X,\mathscr{F},P)$ 上的随机变量，其分布为 F，g 为 $(R,\mathscr{B}_R)$ 上的可测函数，则

$$E(g\circ f):=\int_X g\circ f\ dP=\int_R gdF$$

（第一个等号是定义，第二个是定理，表明可以转成 L-S 积分计算）

定义

$\{f_t:t\in T\}$ 为随机变量集

- 一致可积：$\lim\limits_{\lambda\to\infty}\sup\limits_{t\in T}E|f_t|I_{\{|f_t|>\lambda\}}=0$
- 绝对连续：$\lim\limits_{P(A)\to 0}\sup\limits_{t\in T}E|f_t|I_A=0$，即一致地绝对连续（前面提到过可积的随机变量有绝对连续性）

§4 概率空间的积分

定理

一致可积 $\iff$ 绝对连续 + L_1 有界 $\left(\sup\limits_{t\in T} E|f_t| < \infty\right)$

定理（一致可积与收敛的关系）

$0 < p < \infty, f_n \xrightarrow{p} f$，则以下等价：

- $\{|f_n|^p\}$ 一致可积
- $f_n \xrightarrow{L_p} f$
- $||f_n||_p \longrightarrow ||f||_p$

Chapter4 符号测度

由积分可以定义不定积分，我们希望逆向地定义出集函数关于测度的导数。我们发现不定积分满足可列可加性，但是没有非负性，由此形象地定义出符号测度。Hahn 分解表明 X 可以被分割为两部分，一部分中符号测度非负，另一部分非正，这种极好的性质使得 Radon-Nikodym 定理以异常简洁的形式给出了导数存在性条件。

作为一个例子，我们定义了条件期望。条件期望作为一个可测函数，与期望有着本质的不同，但它们的积分表达式的相似性解释了其性质的相似性。

§1 符号测度

考虑不定积分

$$\varphi(A) := \int_A f d\mu$$

容易发现 φ 满足：$\varphi(\varnothing) = 0$，可列可加性。但 φ 未必是非负的。

定义（符号测度）

若除了非负性外，φ 满足测度的条件，则 φ 为符号测度

同样地，将“符号测度 σ **有限**”定义为：存在一列符号测度的绝对值有限的不交集合能够恰好成为 X 的可测分割

注：$\varphi(A)$ 不能同时取到 ∞ 和 $-\infty$

§2 Hahn 分解和 Jordan 分解

仍然先考虑不定积分 $\varphi(A) := \int_A f d\mu$，它的值是可正可负的，然而我们可以将 X 分割为两部分：$X^+ = \{f \geq 0\}$ 与 $X^- = \{f < 0\}$，不难发现 X^+ 上可测集的不定积分一定非负，而 X^- 上可测集的不定积分一定非正。再令 $\varphi^{\pm}(A) := \int_A f^{\pm} d\mu$，则有 $\varphi^{\pm}$ 均为测度且 $\varphi = \varphi^+ - \varphi^-$

对于一般的符号测度，是否也如此呢？

定理（Hahn 分解）

φ 为 $(X\mathscr{F}, \mu)$ 上的符号测度，存在 $X^{\pm} \in \mathscr{F}$ 使得

$$X^+ \cup X^- = X,\ X^+ \cap X^- = \varnothing$$

（$X^{\pm}$ 为 X 的可测分割），且 $\forall A \in \mathscr{F}$，

$$\varphi(A \cap X^+) \geq 0 \geq \varphi(A \cap X^-)$$

符号测度为零的这部分集合放在哪里都可以，无视它们后分解唯一

§2 Hahn 分解和 Jordan 分解

定理（Jordan 分解）

令 $\varphi^+ = \varphi(A \cap X^+)$，$\varphi^- = -\varphi(A \cap X^-)$，则 $\varphi^\pm$ 都是测度（而且至少有一个有限）且有

$$\varphi = \varphi^+ - \varphi^-$$

此为 Jordan 分解

§3 Radon-Nikodym 定理

由不定积分启发，我们可以定义符号测度的导数！

定义（符号测度的导数）

φ 为 $(X, \mathscr{F}, \mu)$ 上的符号测度，若存在 a.e. 意义下唯一的可测函数 f 使得 $\varphi(A) := \int_A f d\mu$，则将 f 定义为 φ 对 μ 的 Radon-Nikodym 导数，$\dfrac{d\varphi}{d\mu} := f$

由于 $\mu(A) = 0 \Rightarrow \varphi(A) = 0$，因此这是导数存在的一个必要条件，我们将这一性质称为 “**φ 对 μ 绝对连续（$\varphi \ll \mu$）**”

Radon-Nikodym 定理（导数存在条件）

- $\varphi \ll \mu, \mu$ σ 有限，则 φ 对 μ 的 R-N 导数存在
- 若还有 φ σ 有限，则导数 a.e. 有限

可以直观地感受到，定理的条件非常宽松！证明很长，需要使用 §2 中的分解

R-N 导数与一元函数对变量的导数有十分类似的性质（可加，数乘，链式等）

§5 条件期望和条件概率

考虑概率空间 $(X, \mathscr{F}, P)$，设 $\mathscr{G}$ 是 $\mathscr{F}$ 的一个子 σ 域，令 $\varphi(A) = \int_A f dP, A \in \mathscr{F}$，由于 φ 为 $\mathscr{F}$ 上的符号测度知它也是 $\mathscr{G}$ 上的符号测度，又 $\varphi \ll P$，P 有限，故由 R-N 定理知（考虑测度空间 $(X, \mathscr{G}, P)$）存在 a.s.（相差 $\mathscr{G}$ 中的零测集）唯一的（对 $\mathscr{G}$）可测函数 $E(f|\mathscr{G})$，使得 $\varphi(A) = \int_A E(f|\mathscr{G}) dP, A \in \mathscr{G}$。将 $E(f|\mathscr{G})$ 定义为条件期望，$P(A|\mathscr{G}) := E(I_A|\mathscr{G})$ 为事件 $A \in \mathscr{F}$ 关于 $\mathscr{G}$ 的条件概率

可见，条件期望是通过将 f 的不定积分限制在子 σ 域上求导得到的一个 a.s. 唯一的可测函数

§5 条件期望和条件概率

我们先给出独立性的定义，然后讨论一些条件期望的性质

定义（独立性）

- 事件独立：$\{A_t : t \in T\}$ 为事件系，若任意取事件，都有：交集的概率 = 概率的乘积，则 $\{A_t : t \in T\}$ 相互独立
- 族独立：$\{\mathscr{E}_t \subset \mathscr{F}\}$ 为事件系组成的族，若从每个事件系中取一个事件，放在一起形成的新的事件系相互独立，则 $\{\mathscr{E}_t \subset \mathscr{F}\}$ 相互独立
- 随机变量独立：随机变量生成的 σ 域相互独立

关于条件期望的性质，仅举几个不那么显然的

条件期望的性质

- $\mathscr{G} \subset \mathscr{G}_0$ 为 $\mathscr{F}$ 的子 σ 域，$E(E(f|\mathscr{G})|\mathscr{G}_0) = E(f|\mathscr{G})$
- $f \leq g$ $a.s.$，则 $E(f|\mathscr{G}) \leq E(g|\mathscr{G})$ $a.s.$
- g 关于 $\mathscr{G}$ 可测，则 $E(fg|\mathscr{G}) = gE(f|\mathscr{G})$

§5 条件期望和条件概率

就极限性质而言，形式上看与积分的性质并无差异，但积分的极限性质是数列的收敛性，而此处是函数列的收敛

定理

$\{f_n\}, f$ 积分存在，$\mathscr{G}$ 为 $\mathscr{F}$ 的子 σ 域

- 单调收敛：$0 \le f_n \uparrow f \ a.s. \Rightarrow \lim\limits_{n\to\infty} E(f_n|\mathscr{G}) = E(f|\mathscr{G}) \ a.s.$
- Fatou：$f_n \ge 0 \ a.s. \Rightarrow E\left(\liminf\limits_{n\to\infty} f_n|\mathscr{G}\right) \le \liminf\limits_{n\to\infty} E(f_n|\mathscr{G})$
- Lebesgue：$|f_n| \le g \in L_1, \ \lim\limits_{n\to\infty} f_n = f \ a.s. \Rightarrow \lim\limits_{n\to\infty} E(f_n|\mathscr{G}) = E(f|\mathscr{G})$

下面我们将对条件期望与期望高度相似的性质背后的潜在原因略做讨论，提示：需要用到 Chapter3 的补注记号

§5 条件期望和条件概率

考虑定义在 $X \times R$ 上的二元函数 $F_{f|\mathscr{G}}(\cdot,\cdot)$，我们希望 $\forall x \in X, F_{f|\mathscr{G}}(x,\cdot)$ 是一个分布函数，$\forall a \in R, F_{f|\mathscr{G}}(\cdot,a)$ 是事件 $\{f \leq a\}$ 关于 $\mathscr{G}$ 的条件期望。可以证明，上述二元函数一定存在，称它为正则条件分布函数！

自然地，对于固定的 $x \in X$，可以由分布函数 $F_{f|\mathscr{G}}(x,\cdot)$ 诱导出一个 $\mathscr{B}_R$ 上的 L-S 测度 $\mu_{f|\mathscr{G}}(x,\cdot)$，这就表明：存在定义于 $X \times \mathscr{B}_R$ 的二元函数 $\mu_{f|\mathscr{G}}(\cdot,\cdot)$，使得固定第一维后是测度，固定第二维后 $\mu_{f|\mathscr{G}}(\cdot,B)$ 是 $P(f \in B|\mathscr{G})$（a.s. 意义下）。对后者更清晰的说明是：$\mu_{f|\mathscr{G}}(\cdot,B)$ 关于 $\mathscr{G}$ 可测，且有

$$P(f^{-1}B \cap A) = \int_A \mu_{f|\mathscr{G}}(\cdot,B)dP,\ \ A \in \mathscr{F}$$

回忆 Chapter3 §4 中关于期望的定义和计算方法：$E(g(f)) = \int_R g d\lambda_F$，$\lambda_F$ 为 f 的分布 F 诱导的 L-S 测度。对于条件期望，有：

$$E(g(f)|\mathscr{G})(\cdot) = \int_R g(y)\mu_{f|\mathscr{G}}(\cdot,dy),\ a.s.$$

从两个积分表达式中可以理解到，期望和条件期望虽然本质不同，但是性质相似！

§5 条件期望和条件概率

最后，我们来讨论初等概率论中常见的 $E(f|g=y)$ 是怎么回事

设 f 为 $(X,\mathscr{F},P)$ 上的可测函数，$g:(X,\mathscr{F},P)\to(Y,\mathscr{E})$ 为可测映射，可以定义 $E(f|g):=E(f|\sigma(g))$，即满足对 $\sigma(g)=g^{-1}\mathscr{E}$ 可测，以及

$$\forall A\in\sigma(g),\ \int_A fdP=\int_A E(f|g)dP \tag{1}$$

由复合可测函数的性质，存在 $(Y,\mathscr{E})$ 上的可测函数 h，使得 $h\circ g=E(f|g)$，于是 (1) 变形为:

$$\forall B\in\mathscr{E},\ \int_{g^{-1}B} fdP=\int_{g^{-1}B} E(f|g)dP=\int_{g^{-1}B} h\circ gdP=\int_B h(y)(Pg^{-1})(dy)$$

由 $E(f|g)$ a.s. 唯一知 h 也 a.s. 唯一，定义

$$E(f|g=y):=h(y),\ P(A|g=y):=E(I_A|g=y)$$

若 g 为随机变量，则 Pg^{-1} 为 g 的分布函数导出的 L-S 测度；若 g 还是离散型的，则可以验证上述给定值时条件概率的定义与初等概率论中的吻合

Chapter5 乘积空间

可以适当地定义两个可测空间的乘积空间，而 Fubini 定理表明两个可测空间上的测度可以“转移”到乘积空间上去，并用累次积分计算。

在乘积空间上，我们定义了多维 L-S 测度。

§1 有限维乘积空间

定义（可测空间的乘积）

- 集合乘积：$\{X_k, k = 1, 2, ..., n\}$ 为有限个集合，定义

$$\prod_{k=1}^{n} X_k := \{(x_1, ..., x_n) : x_k \in X_k\}$$

- σ 域乘积：$\{(X_k, \mathscr{F}_k), k = 1, 2, ...n\}$ 为有限个可测空间，定义

$$\mathscr{D} := \left\{ \prod_{k=1}^{n} A_k : A_k \in \mathscr{F}_k, k = 1, 2, ...n \right\}$$

  $\mathscr{D}$ 为半环，定义

$$\prod_{k=1}^{n} \mathscr{F}_k := \sigma(\mathscr{D})$$

- 可测空间的乘积：

$$\left(\prod_{k=1}^{n} X_k, \prod_{k=1}^{n} \mathscr{F}_k \right)$$

§1 有限维乘积空间

乘积空间 $\prod_{k=1}^{n} X_k$ 到某空间 Ω 的映射是一个取值于 Ω 的 n 元函数；而 Ω 到 $\prod_{k=1}^{n} X_k$ 的映射是由 n 个映射排成的向量

定义（投影）

从 $\prod_{k=1}^{n} X_k$ 到 X_k 的映射 $\pi_k(x_1, ..., x_n) = x_k, k = 1, ..., n$ 为投影

性质（投影）

- 投影是 $\left(\prod_{k=1}^{n} X_k, \prod_{k=1}^{n} \mathscr{F}_k\right)$ 到 $(X_k, \mathscr{F}_k)$ 的可测映射
- 为使每个投影都可测，最小的 σ 域就是 $\prod_{k=1}^{n} \mathscr{F}_k$，即

$$\prod_{k=1}^{n} \mathscr{F}_k = \sigma\left(\bigcup_{k=1}^{n} \pi_k^{-1} \mathscr{F}_k\right)$$

§1 有限维乘积空间

从 $(X_k, \mathscr{F}_k)$ 到 $(R^n, \mathscr{B}_{R^n})$ 的可测函数 $f = (f_1, ..., f_n)$ 为随机向量，它等价于每一维都可测，即 f_k 是 $(X_k, \mathscr{F}_k)$ 到 $(R, \mathscr{B}_R)$ 的可测函数；也等价于 f_k 都是随机变量

有了乘积空间，自然有一个问题：如何定义乘积空间上的测度？

跳过一些内容，我们直接快进到**两个测度空间乘起来成为乘积测度空间**的结论

§1 有限维乘积空间

定理（Fubini）

$(X_1, \mathscr{F}_1, \mu_1), (X_2, \mathscr{F}_2, \mu_2)$ 是 σ 有限的测度空间，则在乘积空间 $(X_1 \times X_2, \mathscr{F}_1 \times \mathscr{F}_2)$ 上存在唯一测度 μ，满足：$\forall A_1 \in \mathscr{F}_1, A_2 \in \mathscr{F}_2$

$$\mu(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2)$$

且 μ σ 有限，称为乘积测度（$\mu_1 \times \mu_2$）
对 $(X_1 \times X_2, \mathscr{F}_1 \times \mathscr{F}_2, \mu_1 \times \mu_2)$ 上任意积分存在的可测函数 f，有

$$\int_{X_1 \times X_2} f d(\mu_1 \times \mu_2) = \int_{X_1} \mu_1(dx_1) \int_{X_2} f(x_1, x_2) \mu_2(dx_2)$$

（只要重积分存在，则等于累次积分且两个累次积分相等）

§1 有限维乘积空间

作为 Fubini 定理的一个例子

L-S 积分的分部公式

F,G 为准分布函数，$a \le b$，则

$$\int_{(a,b]} F(x)G(dx) = FG|_a^b - \int_{(a,b]} G(x-0)F(dx)$$

§2 多维 Lebesgue-Stieljes 测度

设 $a=(a_1,...,a_n)\in R^n$，$b=(b_1,...,b_n)\in R^n$，令

$$(a,b]:=\{x\in R^n : a_i<x_i\le b_i\}$$

为 R^n 上的左开右闭区间

定理

$\mathscr{D}_{R^n}:=\{(a,b]:a,b\in R^n\}$ 为 R^n 上的半环

定义（准分布函数）

F 为 R^n 上的实值函数，且为准分布函数如果：

- 固定其他维后，每维右连续
- $\forall a=(a_1,...,a_n)\in R^n$，$b=(b_1,...,b_n)\in R^n$，记 $C=\{c=(c_1,...,c_n): c_i=a_i \ or\ b_i, i=1,...n\}$ $\forall c\in C, n(c):=\#\{c_i=a_i\}$，则 $a\le b$ 时有

$$\sum_{c\in C}(-1)^{n(c)}F(c)\ge 0$$

（不好琢磨，其实它是非降性的多维推广）

§2 多维 Lebesgue-Stieljes 测度

若 F 是准分布函数，定义：

$$\mu_F((a,b]) := \begin{cases} \sum_{c\in C}(-1)^{n(c)}F(c) & a < b \\ 0 & a \geq b \end{cases}$$

则 μ_F 为 $\mathscr{D}_{R^n}$ 上的测度，用它产生 R^n 上的外侧度 λ_F，再把 σ 域限制为 $\mathscr{F}_{\lambda_F}$（或者 $\mathscr{B}_{R^n}$），则 $(R^n, \mathscr{B}_{R^n}, \lambda_F)$ 为测度空间

若 $F(x_1, ..., x_n) = x_1...x_n$，则为 n 维 L 测度

Chapter6 独立随机变量序列

尾事件直观上看是与前有限项都无关的事件，独立随机变量序列的尾事件发生概率非 0 即 1. 级数和的收敛性作为一个尾事件，我们自然关心它 a.s 收敛的条件。

零均值独立随机变量序列的方差和有限时，级数 a.s. 收敛；当方差不存在时，Kolmogorov 三级数定理给出了充要条件。最后，介绍了两种强大数定律。

§1 幂零律和三级数定理

$\{f_n\}$ 是概率空间 $(X,\mathscr{F},P)$ 上的一列随机变量，对 $n=1,2,\ldots$，记

$$\mathscr{G}_n := \sigma(\{f_k, k=n,n+1,\ldots\}) = \sigma\left(\bigcup_{k=n}^{\infty} f_k^{-1}\mathscr{B}_R\right)$$

它是 $\mathscr{F}$ 的子 σ 域

定义（尾 *）

记 $\mathscr{G} := \bigcap_{n=1}^{\infty} \mathscr{G}_n$，它是 $\mathscr{F}$ 的子 σ 域，称为**尾 σ 域**。尾 σ 域中的事件为**尾事件**，对 $\mathscr{G}$ 可测的随机变量为**尾随机变量**

可以直观地认为，尾事件是“发生与否与前有限项都无关的事件”

§1 幂零律和三级数定理

例（尾事件）

对每个 $k = 1, 2, \dots,$

$$\left\{\sum_{n=k}^{\infty} f_n \ convergence\right\} \in \sigma(\{f_k, f_{k+1}, \dots\})$$

（直观地看，知道了 $f_k, f_{k+1}, \dots$ 的信息，就知道 $\sum_{n=k}^{\infty} f_n$ 是否收敛）。因此

$$\left\{\sum_{n=1}^{\infty} f_n \ convergence\right\} \in \mathscr{G}$$

类似可以验证，$\{\lim_{n\to\infty} f_n \ \exists\}$，$\left\{\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_k \ \exists\right\}$ 都是尾事件，而

$\liminf_{n\to\infty} f_n$，$\limsup_{n\to\infty} f_n$，$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_k$，$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_k$ 都是尾随机变量

§1 幂零律和三级数定理

由于尾 σ 域的特殊性，有

定理（0-1 Law）

独立随机变量序列的尾事件的概率非 0 即 1；任一尾随机变量 a.s. 为常数

这就说明，独立的随机变量序列 $\{f_n, n = 1, 2, ...\}$，$\sum_{n=1}^{\infty} f_n$ 要么 a.s. 收敛，要么 a.s. 发散！（也就是使得 $\sum_{n=1}^{\infty} f_n = \infty$ 的那些 $x \in X$ 的点组成的集合的测度为 1 或 0）

这自然引出一个问题：$\sum_{n=1}^{\infty} f_n$ a.s. 收敛的条件是什么？这部分证明有些复杂，我们直接给出结论：先给出一个类似于柯西收敛准则的收敛充要条件，再给出 Kolmogorov 不等式，它是 Chebyshev 不等式的推广，然后给出方差有限时 $\sum_{n=1}^{\infty} f_n$ a.s. 收敛的条件，最后给出 Kolmogorov 三级数定理

§1 幂零律和三级数定理

记 $S_n := \sum_{k=1}^{n} f_k$ 为部分和

定理

任意随机变量序列 $\{f_n, n = 1, 2, \ldots\}$，$\sum_{n=1}^{\infty} f_n$ a.s. 收敛的充要条件为

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} \lim_{N\to\infty} P\left(\max_{n\le l\le N} |S_l - S_n| \ge \varepsilon \right) = 0$$

定理（Kolmogorov 不等式）

独立随机变量序列 $\{f_n, n = 1, 2, \ldots\}, E(f_k) = 0,\ \sigma_k^2 := Var(f_k) < \infty,\ k = 1, 2, \ldots,$，则

$$\forall \varepsilon > 0,\ P\left(\max_{1\le k\le n} |S_k| \ge \varepsilon \right) \le \frac{1}{\varepsilon^2} \sum_{k=1}^{n} \sigma_k^2$$

§1 幂零律和三级数定理

定理（方差有限）

独立随机变量序列 $\{f_n, n = 1, 2, ...\}, E(f_k) = 0,\ \sigma_k^2 < \infty,\ k = 1, 2, ...,$

- $$\sum_{n=1}^{\infty} \sigma_n^2 < \infty \Longrightarrow \sum_{n=1}^{\infty} f_n\ convergence\ a.s.$$
- $\sum_{n=1}^{\infty} f_n$ $a.s.$ 收敛，且 $\{f_n\}$ 一致有界，则 $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$

§1 幂零律和三级数定理

当方差不一定存在时

Kolmogorov 三级数定理

独立随机变量序列 $\{f_n, n=1,2,...\}$, $\sum_{n=1}^{\infty} f_n$ $a.s.$ 收敛收敛的充要条件为：存在 $C>0$，使下列三式成立

$$\sum_{n=1}^{\infty} P(|f_n|>C)<\infty$$

$$\sum_{n=1}^{\infty} E(f_n I_{\{|f_n|\leq C\}})<\infty$$

$$\sum_{n=1}^{\infty} Var(f_n I_{\{|f_n|\leq C\}})<\infty$$

§2 强大数定律

强大数定律

- Kolmogorov 强大数定律 1：$\{f_n\} \subset L_2$，独立，$\sum_{n=1}^{\infty} \frac{Var(f_n)}{n^2} < \infty$，则
$$\frac{S_n - ES_n}{n} \xrightarrow{a.s.} 0$$
- Kolmogorov 强大数定律 2：$\{f_n\}\ iid \subset L_1$，则 $\frac{S_n}{n} \xrightarrow{a.s.} Ef_1$