≪ 测度论与概率论基础 ≫ 笔记-测度空间

QingShan

2021 年 4 月 16 日

测度其实就是测量"大小"的一种函数。像线段的长度，平面上某些曲线围成的面积和容器的容积等都是测度。但是仅仅讨论这些由直接经验建立的测度是远远不够的。例如，概率从抽象角度看是对形形色色的事件发生的可能性进行测量。所以只有在抽象空间的集合上建立了测度，才有可能真正解决概率论的问题。在抽象空间中建立测度是没有什么直接经验可循的，只能采用公理化的方法。当然，归根结底，公理化方法中的那些公理也是从实际中提炼出来的。

1 测度的定义及性质

给定空间 X 上的集合系 $\mathscr{E}$。定义在 $\mathscr{E}$ 上，取值于 $[0,\infty]$ 的函数称为**非负集函数**，用希腊字母 $\mu,\nu,\tau,\cdots$ 等记之。设 μ 是 $\mathscr{E}$ 上的非负集函数。如果对任意可列个两两不交的集合 $A_1,A_2,\cdots\in\mathscr{E}$，只要 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{E}$，就一定有

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

则称 μ 具有**可列可加性**。测度的公理化定义如下。

定义 1.1 设 $\mathscr{E}$ 是 X 上的集合系且 $\varnothing\in\mathscr{E}$。如果 $\mathscr{E}$ 上的非负集函数 μ 有可列可加性并且满足 $\mu(\varnothing)=0$，则称之为 $\mathscr{E}$ 上的**测度**。如果对每个 $A\in\mathscr{E}$ 还有 $\mu(A)<\infty$，则称**测度 μ 是有限的**；如果对每个 $A\in\mathscr{E}$，存在满足 $\mu(A_n)<\infty$ 的 $\{A_n\in\mathscr{E},n=1,2,\cdots\}$，使得 $\bigcup\limits_{n=1}^{\infty}\supset A$，则称测度 μ 是 **σ 有限的**。

如果对 $\mathscr{E}$ 中任意有限个两两不交而且满足 $\bigcup\limits_{i=1}^{n}A_i\in\mathscr{E}$ 的集合 $A_1,\cdots,A_n$，

均有
$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i),$$
则称非负集函数 μ 有**有限可加性**; 如果对任何 $A, B \in \mathscr{E}, A \subset B, B - A \in \mathscr{E}$，只要 $\mu(A) < \infty$，就有
$$\mu(B - A) = \mu(B) - \mu(A),$$
则称非负集函数 μ 具有**可减性**。由测度的定义不难证明：

命题 1.1 测度具有有限可加性和可减性。

下面给出一个重要命题。

命题 1.2 设 $X = \mathbb{R}, \mathscr{E} = \mathscr{Q}_{\mathbb{R}}$，而 F 是 $\mathbb{R}$ 上非降右连续的实值函数。对任意的 $a, b \in \mathbb{R}$，令
$$\mu((a, b]) = \begin{cases} F(b) - F(a), & a < b, \\ 0, & a \geq b, \end{cases}$$
则 μ 是 $\mathscr{E}$ 上的测度。

虽然前面在很一般的集合系上定义了测度。但我们的主要目标还是讨论由 X 的子集形成的某个 σ 域 $\mathscr{F}$ 上的测度。今后，空间 X, 加上由它的子集形成的一个 σ 域 $\mathscr{F}$，再加上 $\mathscr{F}$ 上的一个测度 μ，三位一体形成的 $(X, \mathscr{F}, \mu)$ 称为**测度空间**。如果 $N \in \mathscr{F}$ 而且 $\mu(N) = 0$，则称 N 为 μ 的**零测集**。

如果测度空间 $(X, \mathscr{F}, P)$ 满足 $P(X) = 1$，则称它为**概率空间**，对应的 P 叫做**概率测度**。在概率空间 $(X, \mathscr{F}, P)$ 中，$\mathscr{F}$ 中的集合 A 又称为**事件**，而 $P(A)$ 称为**事件 A 发生的概率**。

在一般的 σ 域上建立测度是非常复杂的，通常使用的办法是把半环上的测度扩张到由它生成的 σ 域上去。为给半环上测度的扩张作必要的准备，需要先讨论半环上非负集函数的性质。而为了使这些性质的描述更简洁，则要用到下列关于集合系 $\mathscr{E}$ 上的非负集函数 μ 的四个术语。

单调性：如果对任何 $A, B \in \mathscr{E}$ 且 $A \subset B$，均有 $\mu(A) \leq \mu(B)$，则称 μ 具有单调性。

半可列可加性：如果对任意可列个集合 $A_1, A_2, \cdots \in \mathscr{E}$，只要 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{E}$，就一定有
$$\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n),$$

则称 μ 是半可列可加的；

下连续性：如果对任意 $A_1, A_2, \cdots \in \mathscr{E}, A_n \uparrow A \in \mathscr{E}$，均有

$$\mu(A) = \lim_{n\to\infty} \mu(A_n),$$

则称 μ 是下连续的。

上连续性：如果对任意 $A_1, A_2, \cdots \in \mathscr{E}, A_n \downarrow A \in \mathscr{E}$ 且 $\mu(A_1) < \infty$，均有

$$\mu(A) = \lim_{n\to\infty} \mu(A_n),$$

则称 μ 是上连续的。

命题 1.3 半环 $\mathscr{Q}$ 上有有限可加性的非负集函数 μ 具有单调性和可减性。

命题 1.4 半环 $\mathscr{Q}$ 上有可列可加性的非负集函数 μ 具有半可列可加性、下连续性和上连续性。

定理 1.5 半环 $\mathscr{Q}$ 上的测度 μ 具有单调性，可减性，半可列可加性、下连续性和上连续性。

定理 1.6 对于环 $\mathscr{R}$ 上的有限可加性非负集函数 μ 有

(1) μ 可列可加 $\Leftrightarrow$

(2) μ 半可列可加 $\Leftrightarrow$

(3) μ 下连续 $\Rightarrow$

(4) μ 上连续 $\Rightarrow$

(5) μ 在 $\varnothing$ 上连续，即对任何满足 $A_n \downarrow \varnothing$ 和 $\mu(A_1) < \infty$ 的 $\{A_n \in \mathscr{R}, n = 1, 2, \cdots\}$，有

$$\lim_{n\to\infty} \mu(A_n) = 0.$$

如果 μ 是有限的，则还有 (5) $\Rightarrow$ (1)，即以上五个性质都等价。

2 外侧度

为了把半环上的测度扩张到 σ 域上去，需要引进外侧度的概念并讨论它的性质。

定义 2.1 由 X 的所有子集组成的集合系 $\mathscr{F}$ 到 $\overline{\mathbb{R}}$ 的函数 τ 称为 X 上的**外侧度**，如果它满足：

(1) $\tau(\varnothing) = 0$;

(2) 对任何 $A \subset B \subset X$ 有 $\tau(A) \le \tau(B)$;

(3) 对任何 $\{A_n \in \mathscr{F}, n=1,2,\cdots\}$ 有 $\tau(\bigcup\limits_{n=1}^{\infty} A_n) \leq \sum\limits_{n=1}^{\infty} \tau(A_n)$.

从定义可见**外侧度是 $\mathscr{F}$ 上具有半可列可加性的非负集函数**。从这个结论加上定义 (1) 又可见**外侧度也是半有限可加的**。

下面的定理表明了外侧度可由非负集函数生成。

定理 2.1 设 $\mathscr{E}$ 是一个集合系且 $\varnothing \in \mathscr{E}$。如果 $\mathscr{E}$ 上的非负集函数 μ 满足 $\mu(\varnothing)=0$，对每个 $A \in \mathscr{F}$，令

$$\tau(A) = \inf\{\sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathscr{E}, n \geq 1; \bigcup_{n=1}^{\infty} B_n \supset A\}.$$

则 τ 是一个外侧度，称为**由 μ 生成的外侧度**。

由测度和外侧度的定义可见，每一个 $\mathscr{F}$ 上的测度一定是 X 上的外侧度。反正则未必。那么进一步可以问：把外侧度限制在比 $\mathscr{F}$ 小一些的某个集合系上，它是否会成为测度呢？下面的定理将给出肯定的回答。

设 τ 是 X 上的一个外侧度。我们把满足

$$\tau(D) = \tau(D \cap A) + \tau(D \cap A^c),\ \forall D \in \mathscr{F}$$

的 X 的子集 A 称为 τ 可测集；把由全体 τ 可测集组成的集合系记为 $\mathscr{F}_\tau$；而**完全测度空间**的概念则由下列定义给出：

定义 2.2 如果 μ 的任一零测集的子集还属于 $\mathscr{F}$，即

$$A \in \mathscr{F}, \mu(A) = 0 \Rightarrow B \in \mathscr{F}, \forall B \subset A,$$

则称测度空间 $(X, \mathscr{F}, \mu)$ 是**完全的**。

定理 2.2(Caratheodory 定理) 如果 τ 是外侧度，则 $\mathscr{F}_\tau$ 是一个 σ 域，$(X, \mathscr{F}_\tau, \tau)$ 是一个完全测度空间。

3 测度的扩张

设 μ 和 τ 分别是集合系 $\mathscr{E}$ 和集合系 $\overline{\mathscr{E}}$ 上的测度，而且 $\mathscr{E} \subset \overline{\mathscr{E}}$。如果对每个 $A \in \mathscr{E}$ 均有

$$\tau(A) = \mu(A),$$

则称 τ 为 μ 在 $\overline{\mathscr{E}}$ 上的**扩张**。当然，我们希望测度的扩张是**唯一的**，也就是说，如果 $\overline{\mathscr{E}}$ 上还有一个测度 τ' 使对每个 $A \in \mathscr{E}$，$\tau'(A) = \mu(A)$ 也成立，就必须有 $\tau' = \tau$，即 $\tau'(A) = \tau(A), \forall A \in \overline{\mathscr{E}}$。

我们的任务是把一个集合系 $\mathscr{E}$ 上的测度扩张到比它更大的集合系上去。从表面上看，上节中建立的外侧度似乎可以很好的解决这个问题：只要在集合系 $\mathscr{E}$ 上有测度 μ，就可以用定理 2.1 在 X 上生成一个外侧度 τ，而根据定理 2.2，把这个外侧度限制在 σ 域 $\mathscr{F}_\tau$ 上就得到了一个测度。但是其中有个问题是，经过上面的程序得到的 σ 域 $\mathscr{F}_\tau$ 不一定包含集合系 $\mathscr{E}$，其甚至可能会更小。产生这种现象的原因是，定理 2.2 根本就没有提到 $\mathscr{E}$ 上的测度 μ 和 $\mathscr{F}_\tau$ 上的测度 τ 究竟是什么关系！因此，为了得到测度的扩张，必须对集合系 $\mathscr{E}$ 作某种限制。我们将证明：**半环上的 σ 有限测度按上面的程序将可以得到符合要求的扩张**！

命题 3.1 设 $\mathscr{P}$ 是一个 π 系。如果 $\sigma(\mathscr{P})$ 上的测度 μ,ν 满足

(1) 对每个 $A\in\mathscr{P}$ 有 $\mu(A)=\nu(A)$;

(2) 存在两两不交的 $\{A_n\in\mathscr{P},n=1,2,\cdots\}$ 使 $\bigcup\limits_{n=1}^{\infty}A_n=X$ 且 $\mu(A_n)<\infty$ 对每个 $n=1,2,\cdots$ 成立，则对任何 $A\in\sigma(\mathscr{P})$ 有

$$\mu(A)=\nu(A).$$

定理 3.2（测度扩张定理） 对于半环 $\mathscr{Q}$ 上的测度 μ，存在 $\sigma(\mathscr{Q})$ 上的测度 τ 使对每个 $A\in\mathscr{Q}$ 有

$$\tau(A)=\mu(A);\tag{1}$$

如果命题 3.1 之 (2) 中把 $\mathscr{P}$ 换成 $\mathscr{Q}$ 后成立，则使 (1) 式成立的 τ 唯一。

此定理中的存在性可以用此测度生成的外侧度来证明。即半环上的测度 μ 生成的外侧度 τ 满足 (1) 式，这也说明了 $\mathscr{F}_\tau$ 包含了 $\mathscr{E}$，而 $\mathscr{F}_\tau$ 又为 σ 域，所以其包含了 $\sigma(\mathscr{F})$，而 τ 为 $\mathscr{F}_\tau$ 上的测度，所以把 τ 限制在 $\sigma(\mathscr{E})$ 上即可满足要求！因为半环也是 π 系，所以只要测度 μ 是 σ 有限的，那么 τ 也是 σ 有限的，这样就说明了 τ 是唯一的扩张。所以有以下推论。

推论 3.3 设 $\mathscr{Q}$ 是一个半环且 $X\in\mathscr{Q}$。对于 $\mathscr{Q}$ 上的 σ 有限测度 μ，存在 $\sigma(\mathscr{Q})$ 上的唯一测度 τ 使 (1) 式成立。

推论中要求 $X\in\mathscr{Q}$ 的条件的原因是 $\varnothing$ 也属于半环，所以就存在两两不交的 $\{A_n\in\mathscr{P},n=1,2,\cdots\}$ 使 $X-\varnothing=X=\bigcup\limits_{n=1}^{\infty}A_n$。而要求 σ 有限的原因是为了保证 $\sigma(A_n)<\infty,\forall n=1,2,\cdots$。

需要注意定理 3.2 中“命题 3.1 之 (2)”那个条件和推论 3.3 中对应的“σ 有限”条件是不能随便去掉的，否者，在 $\sigma(\mathscr{Q})$ 上扩张出来的测度可能不唯一。

定理 3.2 表明，半环 $\mathscr{Q}$ 上的测度 μ 可以扩张到 $\sigma(\mathscr{Q})$ 上去。它实际上是通过 μ 产生的外侧度 τ 扩张到了比 $\sigma(\mathscr{Q})$ 范围更大的一个 σ 域 $\mathscr{F}_\tau$ 上。作为 $\sigma(\mathscr{Q})$ 上的测度 τ 和作为 $\mathscr{F}_\tau$ 上的测度 τ 差别究竟多大? 下面的定理给出了答案。

定理 3.4 设 τ 是半环 $\mathscr{Q}$ 上测度 μ 生成的外侧度。

(1) 对每个 $A \in \mathscr{F}_\tau$，存在 $B \in \sigma(\mathscr{Q})$ 使 $B \supset A$ 且 $\tau(A) = \tau(B)$;

(2) 如果命题 3.1 之 (2) 中把 $\mathscr{P}$ 换成 $\mathscr{Q}$ 后成立，则对每个 $A \in \mathscr{F}_\tau$ 存在 $B \in \sigma(\mathscr{Q})$ 使 $B \supset A$ 且 $\tau(B - A) = 0$。

其实定理 3.4 中唯一的差别在于 $\tau(B)$ 是否是有限的。如果 $\tau(B)$ 和 $\tau(A)$ 都是无限的，那么就只能得到 (1)，无法推出 (2)。

把讲过的定理运用于命题 1.2 的例子:

$\mathbb{R}$ 上非降右连续实值函数 F 在半环 $\mathscr{Q}_\mathbb{R}$ 上定义了一个有限测度 μ。表示 $\mathbb{R} = \bigcup\limits_{n=-\infty}^{+\infty} (n, n+1]$ 即可见该测度满足定理 3.2 的附加条件。因此，它在 $\sigma(\mathscr{Q}_\mathbb{R}) = \mathscr{R}_\mathbb{R}$ 上有唯一的扩张。根据定理 2.2，μ 的外侧度 λ_F 在由它的全体可测集组成的 σ 域 $\mathscr{F}_{\lambda_F}$ 上也是一个测度。$\mathscr{R}_\mathbb{R}$ 上的测度 λ_F 和 $\mathscr{F}_{\lambda_F}$ 上的测度 λ_F 之间的关系符合定理 3.4 之 (2) 结论。

习惯上，人们把 $\mathbb{R}$ 上非降右连续实值函数 F 叫做**准分布函数**；把 $\mathscr{F}_{\lambda_F}$ 中的集合叫做 $\mathbb{R}$ 中的 **L-S(Lebesgue-Stieljes) 可测集**；把 $(\mathbb{R}, \mathscr{F}_{\lambda_F})$ 上的可测函数叫做 **L-S 可测函数**；把 $\mathscr{F}_{\lambda_F}$ 上的测度 λ_F 叫做 $\mathbb{R}$ 上的 **L-S 测度**。需要注意的是，Borel 集合系 $\mathscr{R}_\mathbb{R}$ 虽然不依赖于 F，但是对任何 F 均有 $\mathscr{R}_\mathbb{R} \subset \mathscr{F}_{\lambda_F}$。因此，$(\mathbb{R}, \mathscr{F}_{\lambda_F}, \lambda_F)$ 和 $(\mathbb{R}, \mathscr{R}_\mathbb{R}, \lambda_F)$ 都是测度空间。特别的，当 $F(x) = x, \forall x \in \mathbb{R}$ 时，$\mathscr{F}_{\lambda_F}$ 中的集合叫做 $\mathbb{R}$ 中的 **L(Lebesgue) 可测集**；$(\mathbb{R}, \mathscr{F}_{\lambda_F})$ 上的可测函数叫做 **L 可测函数**；$\mathscr{F}_{\lambda_F}$ 上的测度 λ_F 叫做 $\mathbb{R}$ 上的 **L 测度**。今后，我们把 L 可测集记做 $\mathscr{F}_\lambda$；它上面的 L 测度记为 λ。

域也是半环，因此域上的测度也可以扩张到由他生成的 σ 域上去。下面的定理将讨论域上的测度与它的扩张之间的关系。从相反的角度看，这个定理也可以理解为域上的测度对它生成的 σ 域上**测度的逼近**。

命题 3.5 设 μ 是域 $\mathscr{A}$ 上的测度；τ 为 μ 产生的外侧度。如果 $A \in \sigma(\mathscr{A})$ 且 $\tau(A) < \infty$，则对任给 $\varepsilon > 0$，存在 $B \in \mathscr{A}$ 使 $\tau(A \triangle B) < \varepsilon$。

命题 3.6 设 $\mathscr{A}$ 是一个域，μ 是 $\sigma(\mathscr{A})$ 上的测度且在 $\mathscr{A}$ 上 σ 有限。如果 $A \in \sigma(\mathscr{A})$ 且 $\mu(A) < \infty$，则对任给 $\varepsilon > 0$，存在 $B \in \mathscr{A}$ 使 $\tau(A \triangle B) < \varepsilon$。

4 测度空间的完全化

首先说明，任何一个测度空间都可以被完全化。

定理 4.1 对任何测度空间 $(X, \mathscr{F}, \mu)$,

$\tilde{\mathscr{F}} \stackrel{\text{def}}{=} \{A \cup N : A \in \mathscr{F}; \exists N \subset B \in \mathscr{F}$ 使 $\mu(B) = 0\}$

是一个 σ 域；如对每个 $A \cup N \in \tilde{\mathscr{F}}$，令

$$\tilde{\mu}(A \cup B) = \mu(A),$$

则 $(X, \tilde{\mathscr{F}}, \tilde{\mu})$ 是一个完全测度空间且对每个 $A \in \mathscr{F}$ 有

$$\tilde{\mu}(A) = \mu(A).$$

上述定理说明，在 σ 域 $\mathscr{F}$ 上补充一些 μ 的零测集的子集以后，就可以在测度空间 $(X, \mathscr{F}, \mu)$ 的基础上得到一个完全的测度空间 $(X, \tilde{\mathscr{F}}, \tilde{\mu})$。我们把这个完全的测度空间叫做原来那个**测度空间的完全化**。关于半环 $\mathscr{Q}$ 上测度 μ 在由它生成的 σ 域 $\sigma(\mathscr{Q})$ 上扩张出来的测度空间，有如下重要结论。

定理 4.2 设 τ 是半环 $\mathscr{Q}$ 上 σ 有限测度 μ 生成的外侧度，则 $(X, \mathscr{F}_\tau, \tau)$ 是 $(X, \sigma(\mathscr{Q}), \tau)$ 的完全化。

把定理 4.2 用于 L-S 测度，我们得到结论：**测度空间** $(\mathbb{R}, \mathscr{F}_{\lambda_F}, \lambda_F)$ **是测度空间** $(\mathbb{R}, \mathscr{R}_{\mathbb{R}}, \lambda_F)$ **的完全化**。

5 可测函数的收敛性

这一小节主要讨论可测函数的各种收敛性以及这些收敛性之间的关系。

设 f 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果 $\mu(|f| = \infty) = 0$，则称它是**几乎处处有限的**；如果存在 $M > 0$ 使得 $\mu(|f| > M) = 0$，则称它是**几乎处处有界的**；当 $\mu(f \neq 0) = 0$ 时，则说 f **几乎处处为 0**；

把这种说法一般化：对于测度空间 $(X, \mathscr{F}, \mu)$ 上关于 X 的元素 x 的一个命题，如果存在 $(X, \mathscr{F}, \mu)$ 中的零测集 N 使该命题对所有的 $x \in N^c$ 成立，就说这个命题**几乎处处成立**。几乎处处常常被简写为 a.e.(almost everywhere)。例如，f 几乎处处有限记为 $|f| < \infty$ a.e.。

首先，我们讨论可测函数列几乎处处收敛的概念。

定义 5.1 设 $\{f_n, n = 1, 2, \cdots\}$ 和 f 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果

$$\mu(\lim_{n \to \infty} f_n \neq f) = 0 \tag{2}$$

则说可测函数列 $\{f_n\}$ **几乎处处以 f 为极限**，记为 $f_n \xrightarrow{a.e.} f$；如果 f a.e. 有限且 $f_n \xrightarrow{a.e.} f$，则说 $\{f_n\}$ **几乎处处收敛到** f。

(2) 式也等价于

$$\mu(\lim_{n\to\infty} f_n = f) = \mu(X).$$

几乎处处收敛的判别条件如下：

命题 5.1 $f_n \xrightarrow{a.e.} f$ 当且仅当对任给的 $\varepsilon > 0$ 有

$$\mu(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|f_n - f| \geq \varepsilon\}) = 0. \tag{3}$$

根据单调不减数列的极限的定义，(3) 式等价于

$$\mu(\lim_{m\to\infty}\bigcup_{n=m}^{\infty}\{|f_n - f| \geq \varepsilon\}) = 0. \tag{4}$$

第二，我们给出几乎一致收敛的定义及其判别条件。

定义 5.2 设 $\{f_n, n = 1, 2, \cdots\}$ 和 f 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果对任何 $\varepsilon > 0$，存在 $A \in \mathscr{F}$ 使 $\mu(A) < \varepsilon$ 且

$$\lim_{n\to\infty}\sup_{x\notin A}|f_n(x) - f(x)| = 0, \tag{5}$$

则说可测函数列 $\{f_n\}$ **几乎一致收敛到** f，记为 $f_n \xrightarrow{a.u.} f$。

几乎一致收敛的判别条件如下：

命题 5.2 $f_n \xrightarrow{a.u.} f$ 当且仅当对任给的 $\varepsilon > 0$，有

$$\lim_{m\to\infty}\mu(\bigcup_{n=m}^{\infty}\{|f_n(x) - f(x)| \geq \varepsilon\}) = 0, \tag{6}$$

可以看出，(6) 式与 (4) 式中的差别为极限符号在测度内还是测度外。容易看出，一般情况下，几乎一致收敛是更强的收敛，即如果一个测度几乎一致收敛可以推出其几乎处处收敛。而如果测度 μ 满足上连续性，则 (6) 式与 (4) 式等价，即为几乎一致收敛与几乎处处收敛等价。

第三，我们定义可测函数列的依测度收敛。

定义 5.3 设 $\{f_n, n = 1, 2, \cdots\}$ 和 f 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果对任何 $\varepsilon > 0$ 均有

$$\lim_{n\to\infty}\mu(|f_n - f| \geq \varepsilon) = 0, \tag{7}$$

则说可测函数列 $\{f_n\}$ **依测度收敛到** f，记为 $f_n \xrightarrow{\mu} f$。

易见，(6) 式蕴含了 (3) 式和 (7) 式；如果 $\mu(X)<\infty$，则由测度的上连续性又知 (3) 式蕴含了 (7) 式。因此有如下定理。

定理 5.3 下列结论成立：

(1) $f_n \xrightarrow{a.u.} f \Rightarrow f_n \xrightarrow{a.e.} f$ 和 $f_n \xrightarrow{\mu} f$;

(2) 如果 $\mu(X)<\infty$，则

$$f_n \xrightarrow{a.u.} f \Leftrightarrow f_n \xrightarrow{a.e.} f \Rightarrow f_n \xrightarrow{\mu} f.$$

除定理 5.3 外，三种收敛性的关系还有如下进一步的结果。

定理 5.4 $f_n \xrightarrow{\mu} f$ 当且仅当对 $\{f_n\}$ 的任一子列，存在该子列的子列 $\{f_{n'}\}$ 使

$$f_{n'} \xrightarrow{a.u.} f.$$

考虑 $(X,\mathscr{F},P)$ 是概率空间的情形。它上面定义的随机变量将缩写为 r.v.(random variable)。易见：对于 r.v. $\{f_n\}$ 和 f，有

$$f_n \xrightarrow{a.e.} f \Leftrightarrow P(\lim_{n\to\infty} f_n = f) = 1;$$

$$f_n \xrightarrow{P} f \Leftrightarrow \lim_{n\to\infty} P(|f_n - f| < \varepsilon) = 1, \forall \varepsilon > 0.$$

这时候，$f_n \xrightarrow{a.e.} f$ 也称为**几乎必然收敛**，记作 $f_n \xrightarrow{a.s.} f$(a.s. 是 almost sure 的缩写)；$f_n \xrightarrow{P} f$ 也不泛泛地用依测度收敛这个词，而称之为 $\{f_n\}$ **依测度收敛到** f。

除了上面对一般测度空间定义过的那些收敛性以外，概率论中还有一种很重要的收敛性，叫做**依分布收敛**。讨论依分布收敛，必然要牵涉到与分布函数有关的概念和事实。今后，凡满足

$$F(-\infty) \stackrel{\text{def}}{=} \lim_{x\to-\infty} F(x) = 0 \text{ 和 } F(\infty) \stackrel{\text{def}}{=} \lim_{x\to\infty} F(x) = 1$$

的准分布函数 F 称为**分布函数**，简记为 d.f.(distribution function)。不难验证：如果 f 是概率空间 $(X,\mathscr{F},P)$ 上的随机变量，对每个 $x\in\mathbb{R}$，令

$$F(x) = P(f \le x),$$

则 F 是一个分布函数，称为 r.v. f 的分布函数。r.v. f 的分布函数是 F，也常常说成 f **服从** F，记为 $f \sim F$。设 f 是从概率空间 $(X,\mathscr{F},P)$ 到可测空间 $(Y,\mathscr{S})$ 的随机元。它在 $(Y,\mathscr{S})$ 上自然导出的概率测度

$$(Pf^{-1})(B) \stackrel{\text{def}}{=} P(f^{-1}B), \quad \forall B \in \mathscr{S},$$

称为随机元 f 的**概率分布**或简称**分布**。容易看出，r.v. f 的 d.f. F 只不过是它的概率分布在 $\mathscr{R}_{\mathbb{R}}$ 中 π 系 $\mathscr{P}_{\mathbb{R}}$ 上的值。

设 F 是一个准分布函数。对每个 $t \in (F(-\infty), F(\infty))$，令

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}.$$

我们称 $F^{\leftarrow}$ 为 F 的**左连续逆**。左连续逆有如下性质。

引理 5.5 对任何准分布函数 F，有

(1) $F^{\leftarrow}(t) \in \mathbb{R}, \forall t \in (F(-\infty), F(\infty))$;

(2) $F^{\leftarrow}$ 左连续;

(3) 对任何 $t \in (F(-\infty), F(\infty))$ 和 $x \in \mathbb{R}$，有

$$F^{\leftarrow}(t) \leq x \Leftrightarrow F(x) \geq t.$$

令 X 为 $\mathbb{R}$ 中的区间 $(0,1), \mathscr{F} = (0,1) \cup \mathscr{R}_{\mathbb{R}}$ 而 P 为 $\mathscr{F}$ 上的 L 测度，则 $(X, \mathscr{F}, P)$ 是一个概率空间。定义此概率空间上的 r.v. U:

$$U(t) = t, \quad \forall t \in (0,1).$$

易见，U 是初等概率论中讲过的在区间 (0,1) 上均匀分布的随机变量。因此，对任给的 d.f. F，利用引理 5.5 之 (3) 就能推出 r.v. $F^{\leftarrow}$ 的分布函数为

$$\begin{aligned} P(F^{\leftarrow} \circ U \leq x) &= P(\{t \in (0,1) : F^{\leftarrow}(U(t)) \leq x\}) \\ &= P(\{t \in (0,1) : F^{\leftarrow}(t) \leq x\}) \\ &= P(t \in (0,1) : t \leq F(x)) = F(x). \end{aligned}$$

这样我们就证明了一个很有用的命题：**对任何 d.f.** F**，必存在一个概率空间** $(X, \mathscr{F}, P)$ **以及它上面的 r.v.** f**，使得** $f \sim F$。

设 $\{F_n, n = 1, 2, \cdots\}$ 和 F 都是非降实值函数。如果 $F_n(x) \to F(x)$ 对每一个连续点 x 均成立，则称 $\{F_n\}$ **弱收敛到** F，记为 $F_n \xrightarrow{w} F$。随机变量列依分布收敛本质上是其对应的分布函数列的弱收敛。

定义 5.4 设 $\{f_n \sim F_n\}$ 是概率空间 $(X, \mathscr{F}, P)$ 上的随机变量列而 F 是一个分布函数。如果 $F_n \xrightarrow{w} F$，则称随机变量列 $\{f_n\}$ **依分布收敛到**分布函数 F，记为 $f_n \xrightarrow{d} F$；如果随机变量 $f \sim F$ 而且 $f_n \xrightarrow{d} F$，则称随机变量列 f_n **依分布收敛到**随机变量 f，记为 $f_n \xrightarrow{d} f$。

初等概率论中的**中心极限定理**就是依分布收敛的一个例子。以

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy, \quad \forall x \in \mathbb{R}$$

记为**标准正态分布函数**，如果随机变量列 $\{f_n\}$ 满足 $f_n \xrightarrow{d} \Phi$，那么在概率论中就称 $\{f_n\}$ 服从中心极限定理。

定理 5.6 设 $\{f_n, n=1,2,\cdots\}$ 和 f 是概率空间 $(X, \mathscr{F}, P)$ 上的随机变量，则

$$f_n \xrightarrow{P} f \Rightarrow f_n \xrightarrow{d} f.$$

设 $\{f_n\}$ 和 f 是概率空间 $(X, \mathscr{F}, P)$ 上的随机变量，把定理 5.3 之 (2) 和定理 5.6 联系在一起，就得到

$$f_n \xrightarrow{a.s.} f \Rightarrow f_n \xrightarrow{P} f \Rightarrow f_n \xrightarrow{d} f.$$

引理 5.7 如果 $\{F_n, n=1,2,\cdots\}$ 和 F 都是分布函数，则

$$F_n \xrightarrow{w} F \Rightarrow F_n^{\leftarrow} \xrightarrow{w} F^{\leftarrow}.$$

在以下定理的叙述中，我们用符号 $f \stackrel{d}{=} g$ 来表示随机变量 f 和 g 具有相同的分布函数。

定理 5.8 (Skorokhod 定理) 设 $\{f_n, n=1,2,\cdots\}$ 和 f 是概率空间 $(X, \mathscr{F}, P)$ 上的随机变量。如果 $f_n \xrightarrow{d} f$，则存在概率空间 $(\tilde{X}, \tilde{\mathscr{F}}, \tilde{P})$，在它上面定义着随机变量 $\{\tilde{f}_n\}$ 和 $\tilde{f}$ 使

$$\tilde{f}_n \stackrel{d}{=} f_n,\ \forall n=1,2,\cdots, \quad \tilde{f} \stackrel{d}{=} f$$

而且

$$\tilde{f}_n \xrightarrow{a.s.} \tilde{f}.$$

Skorokhod 定理告诉我们从依分布收敛变换到几乎必然收敛所要付出的代价更加大，我们需要把整个概率空间都换掉！而定理 5.4 中则不需要换测度空间。

通过以上的讨论，我们可以总结出这样一条规律：在测度空间 $(X, \mathscr{F}, \mu)$ 上讨论可测函数时，零测集是不必计较的。今后，这样的 f 将称为 $(X, \mathscr{F}, \mu)$ 上 a.e. **定义的可测函数**: 第一，存在一个零测集 N 使 f 在 N^c 上有定义；第二，存在一个 $(X, \mathscr{F})$ 上的可测函数 $\tilde{f}$ 使 $\mu(f \neq \tilde{f}) = 0$。

设 $\{f_n, n=1,2,\cdots\}$ 和 f 都是测度空间 $(X, \mathscr{F}, \mu)$ 上 a.e. 定义的可测函数。如果对满足 $\mu(f_n \neq \tilde{f}_n) = 0, \forall n=1,2,\cdots$ 和 $\mu(f \neq \tilde{f}) = 0$ 的可测函数 $\{\tilde{f}_n, n=1,2,\cdots\}$ 和 $\tilde{f}$ 有

$$\tilde{f}_n \xrightarrow{a.e.} \tilde{f}, \quad \tilde{f}_n \xrightarrow{\mu} \tilde{f} \quad 或 \quad \tilde{f}_n \xrightarrow{a.u.} \tilde{f},$$

则分别称 $\{f_n, n=1,2,\cdots\}$ 几乎处处、依测度或几乎一致收敛到 f 并沿用记号

$$f_n \xrightarrow{a.e.} f, \quad f_n \xrightarrow{\mu} f \quad 或 \quad f_n \xrightarrow{a.u.} f.$$

关于可测函数的命题和定理对于 a.e. 定义的可测函数也还是对的。

同样，我们把概率空间 $(X, \mathscr{F}, P)$ 上 a.e. 定义的 a.e. 有限的可测函数也叫作**随机变量**，而不像定义 4.1 那样非得要求随机变量处处有定义和处处有限。不难看出，概率空间 $(X, \mathscr{F}, P)$ 上两个随机变量 f 和 g 如果满足 $f = g$ a.s.，则 $f \stackrel{d}{=} g$。基于这一点，我们也可以讨论这种扩充了的意义下的随机变量的依分布收敛问题并得到对狭义随机变量得到的那些结论。