高等概率论讨论班 (第 8、9 次)
随机变量的收敛

统计 91 董晟渤

西安交通大学数学与统计学院

日期：2022 年 5 月

目录

1 几乎必然收敛与依概率收敛

1.1 几乎必然收敛与依概率收敛的定义

考虑随机变量序列 $\{X_n, n \geq 1\}$, 并设随机变量 $X < \infty$, a.s..

定义 1.1 (几乎必然收敛)**.** 如果

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1, \quad \text{或} \quad \mathbb{P}\left(\lim_{n\to\infty} X_n \neq X\right) = 0,$$

则称 X_n **几乎必然收敛**于 X, 记作 $X_n \xrightarrow{\text{a.s.}} X$.

命题 1.1 (几乎必然收敛的等价命题)**.** $X_n \xrightarrow{\text{a.s.}} X$ 当且仅当对任意的 $\varepsilon > 0$, 有

$$\mathbb{P}\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n - X| \geq \varepsilon\}\right\} = 0, \quad \text{或} \quad \mathbb{P}\left\{\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{|X_n - X| < \varepsilon\}\right\} = 1.$$

定义 1.2 (依概率收敛)**.** 如果对任意的 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \leq \varepsilon) = 1, \quad \text{或} \quad \mathbb{P}(|X_n - X| \leq \varepsilon) = 0,$$

则称 X_n **依概率收敛**于 X, 记作 $X_n \xrightarrow{p} X$.

定理 1.2 (Riesz)**.** 设 $X_n \xrightarrow{p} X$, 则存在子列 $\{X_{n_k}\}$, 使得 $X_{n_k} \xrightarrow{\text{a.s.}} X$.

命题 1.3 (蕴含关系)**.** 若 $X_n \xrightarrow{\text{a.s.}} X$, 则 $X_n \xrightarrow{p} X$.

定义 1.3 (依 r 阶矩收敛)**.** 设 $X_n \in L^r(\Omega)$, 如果

$$\lim_{n\to\infty} \mathbb{E}|X_n - X|^r = 0,$$

则称 X_n **依 r 阶矩收敛**于 X, 记作 $X_n \xrightarrow{L^r} X$.

命题 1.4 (蕴含关系)**.** 若 $X_n \xrightarrow{L^r}$, 则 $X_n \xrightarrow{p} X$.

命题 1.5. $X_n \xrightarrow{p} X$ 当且仅当 $\dfrac{|X_n|}{1+|X_n|} \xrightarrow{L^r} 0$.

除了蕴含关系以外, 也请留意各种经典的反例.

1.2 Borel-Cantelli 引理

将事件列 $\{E_n\}$ 的上极限, 也即

$$\limsup_{n\to\infty} E_n = \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} E_n$$

记作 E_n, i.o., 表示 $\{E_n\}$ 发生无穷多次.

定理 1.6 (Borel-Cantelli 引理)**.** 对于事件列 $\{E_n\}$, 有

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \implies \mathbb{P}(E_n, \text{i.o.}) = 0.$$

证明. 由概率的次可加性得

$$\mathbb{P}\left(\bigcup_{n=m}^{\infty} E_n\right) \le \sum_{n=m}^{\infty} \mathbb{P}(E_n),$$

因此

$$\mathbb{P}(E_n, \text{i.o.}) = \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n=m}^{\infty} E_n\right) \le \lim_{m\to\infty} \sum_{n=m}^{\infty} \mathbb{P}(E_n) = 0,$$

这便证明了该结论. □

以下设 $\{X_n\}$ 是随机变量序列, X 几乎处处有限.

推论. $X_n \xrightarrow{\text{a.s.}} X$ 当且仅当对任意的 $\varepsilon > 0$, 有

$$\mathbb{P}(|X_n - X| > \varepsilon, \text{i.o.}) = 0.$$

证明. 应用前述结论. □

推论. 若 $X_n \xrightarrow{p} X$, 则存在子列 $\{X_{n_k}\} \subset \{X_n\}$, 使得 $X_{n_k} \xrightarrow{\text{a.s.}} X$.

证明. 由 $X_n \xrightarrow{p} X$, 知对任意的 $k \in \mathbb{N}$, 都有

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| > \frac{1}{2^k}\right) = 0.$$

从而对任意的 $k \in \mathbb{N}$, 都存在 n_k, 使得

$$\mathbb{P}\left(|X_{n_k} - X| > \frac{1}{2^k}\right) \le \frac{1}{2^k} \implies \sum_{k=1}^{\infty} \mathbb{P}\left(|X_{n_k} - X| > \frac{1}{2^k}\right) \le 1 < \infty,$$

根据 Borel-Cantelli 引理, 知

$$\mathbb{P}\left(|X_{n_k}| > \frac{1}{2^k}, \text{i.o.}\right) = 0,$$

因此 $X_{n_k} \xrightarrow{\text{a.s.}} X$. □

定理 1.7 (Borel-Cantelli 引理)**.** 对于独立事件列 $\{E_n\}$, 有

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \implies \mathbb{P}(E_n, \text{i.o.}) = 1.$$

证明. 由 $\{E_n\}$ 独立知 $\{E_n^C\}$ 独立, 因此

$$\mathbb{P}\left(\bigcap_{n=m}^{\infty} E_n^C\right) = \prod_{n=m}^{\infty} \mathbb{P}\left(E_n^C\right) = \prod_{n=m}^{\infty} (1 - \mathbb{P}(E_n)) \le \exp\left(-\sum_{n=m}^{\infty} \mathbb{P}(E_n)\right) = 0,$$

因此

$$\mathbb{P}(E_n, \text{i.o.}) = 1 - \lim_{m\to\infty} \mathbb{P}\left(\bigcap_{n=m}^{\infty} E_n^C\right) = 1 - 0 = 1,$$

这便证明了该结论. □

定理(1.6)和(1.7)分别称为 Borel-Cantelli 引理的收敛部分和发散部分, 前者对 $\{E_n\}$ 无任何要求, 而后者要求 $\{E_n\}$ 是独立的. 事实上, 后者的条件可以退为 $\{E_n\}$ 是两两独立的.

定理 1.8. 对于两两独立的事件列 $\{E_n\}$, 有

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \implies \mathbb{P}(E_n, \text{i.o.}) = 1.$$

证明. 记 $I_n := I_{E_n}$, 则 $\{E_n\}$ 两两独立等价于对任意的 $m \neq n$, 都有

$$\mathbb{E}(I_m I_n) = \mathbb{E}(I_m) \cdot \mathbb{E}(I_n).$$

考虑级数 $\sum_{n=1}^{\infty} I_n(\omega)$, 它发散到 ∞ 等价于有无限多项 $I_n(\omega) = 1$, 等价于 $\omega \in E_n, \text{i.o.}$, 因此只需证明

$$\sum_{n=1}^{\infty} \mathbb{E}(I_n) = \infty \implies \sum_{n=1}^{\infty} I_n = \infty, \quad \text{a.s.}.$$

记部分和 $J_k = \sum_{n=1}^{k} I_n$, 应用 Chebyshev 不等式, 对任意的 $A > 0$, 都有

$$\mathbb{P}\left(|J_k - \mathbb{E}(J_k)| \leq A \cdot \sqrt{\text{Var}(J_k)}\right) \geq 1 - \frac{\text{Var}(J_k)}{A^2 \cdot \text{Var}(J_k)} = 1 - \frac{1}{A^2}.$$

其中, 由 $I_1, I_2, \cdots, I_k$ 不相关, 且任意阶矩都相等, 得

$$\text{Var}(J_k) = \sum_{n=1}^{k} \text{Var}(I_n) = \sum_{n=1}^{k} \mathbb{E}\left(I_n^2\right) - \sum_{n=1}^{k} (\mathbb{E}(I_n))^2 \leq \sum_{n=1}^{k} \mathbb{E}(I_n) = \mathbb{E}(J_k),$$

从而 $\sqrt{\text{Var}(J_k)} = o\left(\mathbb{E}(J_k)\right)$, 因此当 k 充分大时, 有

$$\mathbb{P}\left(J_k > \frac{1}{2} \cdot \mathbb{E}(J_k)\right) \geq 1 - \frac{1}{A^2}.$$

令 $k \to \infty$, 可得

$$\mathbb{P}\left(\lim_{k\to\infty} J_k = \infty\right) \geq 1 - \frac{1}{A^2},$$

由 A 的任意性得知

$$\sum_{n=1}^{\infty} I_n = \lim_{k\to\infty} J_k = \infty, \quad \text{a.s.},$$

这便证明了该结论. □

推论 (0-1 律的一个例子)**.** 对于两两独立的事件列 $\{E_n\}$, 有

$$\mathbb{P}\left(E_n, \text{i.o.}\right) \in \{0, 1\}.$$

(1) 若 $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$, 则 $\mathbb{P}(E_n, \text{i.o.}) = 0$;

(2) 若 $\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty$, 则 $\mathbb{P}(E_n, \text{i.o.}) = 1$.

2 淡收敛

2.1 淡收敛及其等价命题

淡收敛是对概率测度而言的一种性质.

定义 2.1 (次概率测度)**.** 设 μ 是 $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$ 上的测度, 如果 $\mu(\mathbb{R}^1) \leq 1$, 则称 μ 为**次概率测度**.

以下为了方便, 对次概率测度 μ 及 $a, b \in \mathbb{R}$, 记 $\mu(a,b] := \mu((a,b])$, 类似的记号还有 $\mu[a,b)$, $\mu(a,b)$ 和 $\mu[a,b]$, 并约定当 $a > b$ 时, 上述的值均为 0.

定义 2.2 (淡收敛)**.** 设 $\{\mu_n\}$, μ 是次概率测度, 如果存在 $\mathbb{R}$ 的稠密子集 D, 使得对任意的 $a, b \in D$, $a < b$, 都有

$$\mu_n(a,b] \to \mu(a,b],$$

则称 μ_n **淡收敛**到 μ, 称 μ_n 为 μ 的**淡极限**, 记作 $\mu_n \xrightarrow{v} \mu$.

定义 2.3 (连续性区间)**.** 设 μ 是次概率测度, $a, b \in \mathbb{R}$, 若 $\mu(a,b) = \mu[a,b]$, 或者 a, b 均不是 μ 的原子, 则称 (a,b) 是 μ 的**连续性区间**.

定理 2.1. *设 $\{\mu_n\}$, μ 是次概率测度, 下列命题等价:*

(1) 对任意的有限区间 (a,b) 和 $\varepsilon > 0$, 存在 n_0, 使得对任意的 $n > n_0$, 都有

$$\mu(a+\varepsilon, b-\varepsilon) - \varepsilon \leq \mu_n(a,b) \leq \mu(a-\varepsilon, b+\varepsilon) + \varepsilon;$$

(2) 对任意的 μ 的连续性区间 $(a,b]$, 都有

$$\mu_n(a,b] \to \mu(a,b].$$

在这里, $(a,b]$ 可以用 (a,b), $[a,b]$ 或 $[a,b)$ 代替;

(3) $\mu_n \xrightarrow{v} \mu$.

证明. (1) $\Longrightarrow$ (2): 设 (a,b) 是 μ 的连续性区间, 由测度的单调性知

$$\lim_{\varepsilon \downarrow 0} \mu(a+\varepsilon, b-\varepsilon) = \mu(a,b) = \mu[a,b] = \lim_{\varepsilon \downarrow 0} \mu(a-\varepsilon, b+\varepsilon),$$

再令 $n \to \infty$, 则有

$$\mu(a,b) \leq \liminf_{n\to\infty} \mu_n(a,b] \leq \limsup_{n\to\infty} \mu_n(a,b] \leq \mu[a,b] = \mu(a,b),$$

这便说明了 $\mu_n(a,b]\to\mu(a,b]$. 对于 (a,b), $[a,b]$ 和 $[a,b)$ 的情形, 也可以类似证明.

(2) $\Longrightarrow$ (3): 记 $C\subset\mathbb{R}$ 为 μ 的原子所构成的集合, 也即对任意的 $c\in C$, 都有 $\mu(\{c\})>0$. 假设 C 是不可数集, 则有

$$\mu(C)=\sum_{c\in C}\mu(\{c\})=\infty,$$

此与 μ 是次概率测度矛盾, 因此 C 是至多可数集. 记 $D=C^C$, 则 D 是稠密集, 并且对任意的 $a,b\in D$, $a<b$, 都有 $\mu_n(a,b]\to\mu(a,b]$, 这便说明了 $\mu_n\xrightarrow{v}\mu$.

(3) $\Longrightarrow$ (1): 设 $D\subset\mathbb{R}$ 为满足条件的稠密集, 对任意的有限区间 (a,b) 和 $\varepsilon>0$, 存在 $a_1,a_2,b_1,b_2\in D$, 使得

$$a-\varepsilon<a_1<a<a_2<a+\varepsilon,\quad b-\varepsilon<b_1<b<b_2<b+\varepsilon.$$

由 $\mu_n\xrightarrow{v}\mu$ 知, 存在 n_0, 使得对任意的 $n>n_0$, 都有

$$|\mu_n(a_i,b_j]-\mu(a_i,b_j]|<\varepsilon,\quad \forall i=1,2,\quad \forall j=1,2,$$

因此

$$\begin{aligned}\mu(a+\varepsilon,b-\varepsilon)-\varepsilon\le\mu(a_2,b_1]-\varepsilon\le\mu_n(a_2,b_1]\le\mu_n(a,b)\\ \le\mu_n(a_1,b_2]\le\mu(a_1,b_2]+\varepsilon\le\mu(a-\varepsilon,b+\varepsilon)+\varepsilon,\end{aligned}$$

这便证明了原不等式. □

推论. *设 $\{\mu_n\}$ 是次概率测度, 则 $\{\mu_n\}$ 的淡极限是唯一的.*

证明. 设 $\mu_n\xrightarrow{v}\mu$, 且 $\mu_n\xrightarrow{v}\mu'$, 记 A 为 μ 和 μ' 的原子所构成的集合, 则对任意的 $a,b\in A^C$, 都有

$$\mu(a,b]=\mu'(a,b].$$

由 μ 和 μ' 在一个 $\mathbb{R}$ 上稠密的集合 A^C 上相等, 知 $\mu\equiv\mu'$. □

将次概率测度推广到概率测度, 可以得到如下定理. 考虑到本节研究的主要是次概率测度, 在这里暂且不给出证明.

定理 2.2. *设 $\{\mu_n\}$, μ 是概率测度, 下列命题等价:*

(1) 对任意的 $\delta>0$ 和 $\varepsilon>0$, 存在 n_0, 使得对任意的 $n>n_0$, 及对任意的区间 (a,b), 都有

$$\mu(a+\delta,b-\delta)-\varepsilon\le\mu_n(a,b)\le\mu(a-\delta,b+\delta)+\varepsilon;$$

(2) 对任意的 μ 的连续性区间 $(a,b]$, 都有

$$\mu_n(a,b]\to\mu(a,b].$$

在这里, $(a,b]$ 可以用 (a,b), $[a,b]$ 或 $[a,b)$ 代替;

(3) $\mu_n\xrightarrow{v}\mu$.

证明. 略. □

2.2 次概率测度的列紧性

进一步, 我们研究所有次概率测度所构成的集合的结构. 考虑到所有的次概率测度和 $[0,1]$ 是类似的, 并且考虑到 $[0,1]$ 是列紧集, 我们也可以证明次概率测度所构成的集合是列紧的.

定理 2.3. 设 $\{\mu_n\}$ 是次概率测度, 则存在子列 $\{\mu_{n_k}\}$, 使得 $\mu_{n_k} \xrightarrow{v} \mu$.

证明. 定义函数

$$F_n(x) = \mu_n(-\infty, x], \quad \forall x \in \mathbb{R}.$$

则 F_n 是 $\mathbb{R}$ 上单调递增的右连续函数, 且 $F_n(-\infty) = 0, F_n(\infty) = \mu_n(\mathbb{R}) \leq 1$. 设 D 是 $\mathbb{R}$ 的可数稠密子集, $\{r_k, k \geq 1\}$ 是它的排列, 按照如下方式选择 $\{F_n\}$ 的一个子列:

- 数列 $\{F_n(r_1), n \geq 1\}$ 有界, 选取其的一个收敛子列 $\left\{F_n^{(1)}(r_1), n \geq 1\right\}$;
- 数列 $\left\{F_n^{(1)}(r_2), n \geq 1\right\}$ 有界, 选取其的一个收敛子列 $\left\{F_n^{(2)}(r_2), n \geq 1\right\}$;
- $\ldots$;
- 数列 $\left\{F_n^{(k-1)}(r_k), n \geq 1\right\}$ 有界, 选取其的一个收敛子列 $\left\{F_n^{(k)}(r_k), n \geq 1\right\}$;
- $\ldots$.

由此, 我们得到了若干函数列:

$$\begin{array}{llllll}
F_1^{(1)}, & F_2^{(1)}, & \cdots, & F_n^{(1)}, & \cdots, & \text{在 } r_1 \text{ 处收敛;} \\
F_1^{(2)}, & F_2^{(2)}, & \cdots, & F_n^{(2)}, & \cdots, & \text{在 } r_1, r_2 \text{ 处收敛;} \\
\cdots, & \cdots, & \cdots, & \cdots, & \cdots; & \\
F_1^{(k)}, & F_2^{(k)}, & \cdots, & F_n^{(k)}, & \cdots, & \text{在 } r_1, r_2, \cdots, r_k \text{ 处收敛;} \\
\cdots, & \cdots, & \cdots, & \cdots, & \cdots. &
\end{array}$$

选取上述函数列的对角线 $F_1^{(1)}, F_2^{(2)}, \cdots, F_k^{(k)}, \cdots$, 则 $\lim\limits_{k\to\infty} F_k^{(k)}$ 在所有的 $\{r_k, k \geq 1\}$ 处收敛, 也即在 D 上收敛. 记

$$G(r) := \lim_{k\to\infty} F_k^{(k)}(r), \quad \forall r \in D,$$

$$F(x) := \sup_{x<r\in D} G(r), \quad \forall x \in \mathbb{R},$$

则 $F(x)$ 是 $\mathbb{R}$ 上单调递增的右连续函数. 设 C 是 $F(x)$ 的连续点, 则 C 在 $\mathbb{R}$ 中稠密. 设 $x \in C$, 则对任意的 $\varepsilon > 0$, 存在 $r, r', r'' \in D$, 使得 $r < r' < x < r''$, 且 $F(r'') - F(r) < \varepsilon$, 于是

$$F(r) \leq G(r') \leq F(x) \leq G(r'') \leq F(r'') \leq F(r) + \varepsilon,$$

且

$$F_k^{(k)}(r') < F_k^{(k)}(x) < F_k^{(k)}(r''),$$

由 ε 的任意性知

$$\lim_{k\to\infty} F_k^{(k)}(x) = F(x), \quad \forall x \in C.$$

我们知道, 存在唯一的概率测度 μ, 使得 $F(x) = \mu(-\infty, x]$. 另外, 设 $F_k^{(k)}$ 所对应的次概率测度为 μ_{n_k}. 由上面的结果, 知对任意的 $a, b \in C$, 都有

$$\lim_{k\to\infty} \mu_{n_k}(a, b] = \mu(a, b],$$

从而 $\mu_{n_k} \xrightarrow{v} \mu$. □

推论. 设 $\{\mu_n\}$ 是次概率测度, 如果对任何淡收敛的子列 $\{\mu_{n_k}\}$, 都有 $\mu_{n_k} \xrightarrow{v} \mu$, 则 $\mu_n \xrightarrow{v} \mu$.

证明. 假设 μ_n 不淡收敛到 μ, 则存在连续性区间 (a, b), 使得 $\mu_n(a, b)$ 不以 $\mu(a, b)$ 为极限. 由 $[0, 1]$ 的列紧性, 存在子列 $\{\mu_{n_k}(a, b)\}$, 使得

$$\mu_{n_k}(a, b) \to a \neq \mu(a, b).$$

而由次概率密度的列紧性, $\{\mu_{n_k}\}$ 存在淡收敛的子列 $\{\mu_{n'_k}\}$, 使得 $\mu_{n'_k} \xrightarrow{v} \mu$, 因此

$$\mu_{n'_k}(a, b) \to \mu(a, b),$$

此与以上极限矛盾, 从而假设不成立. □

2.3 随机变量的依分布收敛

最后, 我们来指出这种收敛在分布函数和随机变量上的体现.

定义 2.4 (淡收敛)**.** 设 $\{F_n\}$ 和 F 是分布函数, 对应的概率测度为 $\{\mu_n\}$ 和 μ, 若 $\mu_n \xrightarrow{v} \mu$, 则称 F_n **淡收敛**于 F, 记作 $F_n \xrightarrow{v} F$.

推论. 设分布函数 F 的连续点所构成的集合为 C, 则 $F_n \xrightarrow{v} F$ 当且仅当

$$F_n(x) \to F(x), \quad \forall x \in C.$$

证明. 一方面, 设 $F_n \xrightarrow{v} F$, 则对任意的 μ 的连续性区间 $(a, b]$, 都有

$$\mu_n(a, b] = F_n(b) - F_n(a) \to \mu(a, b] = F(b) - F(a).$$

令 $a = x, b = -\infty$, 则对任意的 $x \in C$, 都有 $F_n(x) \to F(x)$.

另外一方面, 由 $F(x)$ 是分布函数知 C 在 $\mathbb{R}$ 中稠密, 若对任意的 $x \in C$, 都有 $F_n(x) \to F(x)$, 则对任意的 $a, b \in C$, 都有

$$\lim_{n\to\infty} \mu_n(a, b] = \lim_{n\to\infty} (F_n(b) - F_n(a)) = F(b) - F(a) = \mu(a, b],$$

这便说明了 $\mu_n \xrightarrow{v} \mu$, 从而 $F_n \xrightarrow{v} F$. □

定义 2.5 (依分布收敛)**.** 设 $\{X_n\}$ 和 X 是随机变量, 对应的分布函数为 $\{F_n\}$ 和 F, 若 $F_n \xrightarrow{v} F$, 则称 X_n **依分布收敛**于 X, 记作 $X_n \xrightarrow{d} X$.

推论. 设 F_n, F 是随机变量 X_n, X 的分布函数, F 的连续点所构成的集合为 C, 则 $X_n \xrightarrow{d} X$ 当且仅当

$$F_n(x) \to F(x), \quad \forall x \in C.$$

证明. 应用上述结论即可. $\square$

命题 2.4 (蕴含关系)**.** 若 $X_n \xrightarrow{p} X$, 则 $X_n \xrightarrow{d} X$.

证明. 设 F_n, F 是随机变量 X_n, X 的分布函数. 一方面, 对任意的 $x \in \mathbb{R}$ 及 $\varepsilon > 0$, 注意到

$$\begin{aligned}\{X \le x\} &= \{X \le x, X_n \le x+\varepsilon\} \cup \{X \le x, X_n > x+\varepsilon\} \\ &\subset \{X_n \le x+\varepsilon\} \cup \{|X_n - X| > \varepsilon\},\end{aligned}$$

因此

$$F(x) \le F_n(x+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon).$$

由 $X_n \xrightarrow{p} X$ 知 $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$, 令 $n \to \infty$ 得

$$F(x) \le \liminf_{n\to\infty} F_n(x+\varepsilon).$$

另外一方面, 有

$$\begin{aligned}\{X > x\} &= \{X > x, X_n > x-\varepsilon\} \cup \{X > x, X_n \le x-\varepsilon\} \\ &\subset \{X_n > x-\varepsilon\} \cup \{|X_n - X| > \varepsilon\},\end{aligned}$$

因此

$$\begin{aligned} 1 - F(x) &\le 1 - F_n(x-\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon) \\ \Longrightarrow \ F(x) &\ge F_n(x-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon).\end{aligned}$$

令 $n \to \infty$ 得

$$F(x) \ge \limsup_{n\to\infty} F_n(x-\varepsilon).$$

综上有

$$\limsup_{n\to\infty} F_n(x-\varepsilon) \le F(x) \le \liminf_{n\to\infty} F_n(x+\varepsilon),$$

若 x 是 F 的连续点, 令 $\varepsilon \to 0$, 则

$$F(x) = \lim_{n\to\infty} F_n(x) = \limsup_{n\to\infty} F_n(x) = \liminf_{n\to\infty} F_n(x),$$

这便说明了 $X_n \xrightarrow{d} X$. $\square$

命题 2.5 (蕴含关系)**.** 设 $c \in \mathbb{R}$, 则 $X_n \xrightarrow{p} c$ 当且仅当 $X_n \xrightarrow{d} c$.

证明. 只需证明当 $X_n \xrightarrow{d} c$ 时有 $X_n \xrightarrow{p} c$. 记 c 的分布函数

$$F(x) = \begin{cases} 0, & x < c, \\ 1, & x \ge c, \end{cases}$$

连续点所构成的集合为 $\mathbb{R}\setminus\{c\}$. 设 F_n 是 X_n 的分布函数, 则

$$F_n(x)=\mathbb{P}(X_n\le x)\to F(x),\quad \forall x\in\mathbb{R}\setminus\{c\}.$$

对任意的 $\varepsilon>0$, 计算得

$$\begin{aligned}\mathbb{P}(|X_n-c|>\varepsilon)&=\mathbb{P}(X_n<c-\varepsilon)+\mathbb{P}(X_n>c+\varepsilon)\\&\le F_n(c-\varepsilon)+1-F_n(c+\varepsilon),\end{aligned}$$

令 $n\to\infty$ 可得 $\mathbb{P}(|X_n-c|>\varepsilon)\to 0$, 这便说明了 $X_n\xrightarrow{p}c$. □

定理 2.6 (Slutsky 定理)**.** 设 $X_n\xrightarrow{d}X$, $Y_n\xrightarrow{p}c$.

(1) $X_n+Y_n\xrightarrow{d}X+c$;

(2) $Y_nX_n\xrightarrow{d}cX$.

证明. (1) 设 F_n,G_n,F 是 $X_n+Y_n,X_n+c,X+c$ 的分布函数, 并且根据 $X_n\xrightarrow{d}X$, 知 $X_n+c\xrightarrow{d}X+c$, 因此 $G_n\xrightarrow{v}F$. 设 x 是 F 的连续点, 对任意的 $\varepsilon>0$, 一方面, 注意到

$$\begin{aligned}\{X_n+Y_n>x+\varepsilon\}&=\{X_n+Y_n>x+\varepsilon,|Y_n-c|\le\varepsilon\}\cup\{X_n+Y_n>x+\varepsilon,|Y_n-c|>\varepsilon\}\\&\subset\{X_n+c>x\}\cup\{|Y_n-c|>\varepsilon\},\end{aligned}$$

因此

$$\begin{aligned}&1-F_n(x+\varepsilon)\le 1-G_n(x)+\mathbb{P}(|Y_n-c|>\varepsilon)\\\Longrightarrow\ &G_n(x)\le F_n(x+\varepsilon)+\mathbb{P}(|Y_n-c|>\varepsilon),\end{aligned}$$

令 $n\to\infty$ 可得

$$F(x)=\lim_{n\to\infty}G_n(x)\le\liminf_{n\to\infty}F_n(x+\varepsilon);$$

另外一方面, 注意到

$$\begin{aligned}\{X_n+Y_n\le x-\varepsilon\}&=\{X_n+Y_n\le x-\varepsilon,|Y_n-c|\le\varepsilon\}\cup\{X_n+Y_n\le x-\varepsilon,|Y_n-c|>\varepsilon\}\\&\subset\{X_n+c\le x\}\cup\{|Y_n-c|>\varepsilon\},\end{aligned}$$

因此

$$G_n(x)\ge F_n(x-\varepsilon)-\mathbb{P}(|Y_n-c|>\varepsilon),$$

令 $n\to\infty$ 可得

$$F(x)=\lim_{n\to\infty}G_n(x)\ge\limsup_{n\to\infty}F_n(x-\varepsilon).$$

考虑到 x 是 F 的连续点, 令 $\varepsilon\to 0$, 则

$$F(x)=\lim_{n\to\infty}F_n(x)=\liminf_{n\to\infty}F_n(x)=\limsup_{n\to\infty}F_n(x),$$

这便说明了 $X_n+Y_n\xrightarrow{d}X+c$.

(2) 设 F_n, G_n, F 是 Y_nX_n, cX_n, cX 的分布函数, 并且根据 $X_n \xrightarrow{d} X$, 知 $cX_n \xrightarrow{d} cX$, 因此 $G_n \xrightarrow{v} F$. 设 x 是 F 的连续点, 对任意的 $\varepsilon > 0$, 一方面, 注意到

$$\begin{aligned}\left\{Y_nX_n > x + \frac{x}{c}\cdot\varepsilon\right\} &= \left\{Y_nX_n > x + \frac{x}{c}\cdot\varepsilon, |Y_n - c| \le \varepsilon\right\} \\ &\quad \cup \left\{Y_nX_n > x + \frac{x}{c}\cdot\varepsilon, |Y_n - c| > \varepsilon\right\} \\ &\quad \subset \{cX_n > x\} \cup \{|Y_n - c| > \varepsilon\},\end{aligned}$$

因此

$$\begin{aligned}& 1 - F_n\left(x + \frac{x}{c}\cdot\varepsilon\right) \le 1 - G_n(x) + \mathbb{P}(|Y_n - c| > \varepsilon) \\ \Longrightarrow\ & G_n(x) \le F_n\left(x + \frac{x}{c}\cdot\varepsilon\right) + \mathbb{P}(|Y_n - c| > \varepsilon)\end{aligned}$$

令 $n \to \infty$ 可得

$$F(x) = \lim_{n\to\infty} G_n(x) \le \liminf_{n\to\infty} F_n\left(x + \frac{x}{c}\cdot\varepsilon\right);$$

另外一方面, 注意到

$$\begin{aligned}\left\{Y_nX_n \le x - \frac{x}{c}\cdot\varepsilon\right\} &= \left\{Y_nX_n \le x - \frac{x}{c}\cdot\varepsilon, |Y_n - c| \le \varepsilon\right\} \\ &\quad \cup \left\{Y_nX_n \le x - \frac{x}{c}\cdot\varepsilon, |Y_n - c| > \varepsilon\right\} \\ &\quad \subset \{cX_n \le x\} \cup \{|Y_n - c| > \varepsilon\},\end{aligned}$$

因此

$$G_n(x) \ge F_n\left(x - \frac{x}{c}\cdot\varepsilon\right) - \mathbb{P}(|Y_n - c| > \varepsilon),$$

令 $n \to \infty$ 可得

$$F(x) = \lim_{n\to\infty} G_n(x) \ge \limsup_{n\to\infty} F_n\left(x - \frac{x}{c}\cdot\varepsilon\right).$$

考虑到 x 是 F 的连续点, 令 $\varepsilon \to 0$, 则

$$F(x) = \lim_{n\to\infty} F_n(x) = \liminf_{n\to\infty} F_n(x) = \limsup_{n\to\infty} F_n(x),$$

这便说明了 $Y_nX_n \xrightarrow{d} cX$. □

依分布收敛的一个重要刻画需要应用到后面介绍的特征函数, 在此简单叙述结论.

定理 2.7 (连续性定理)**.** 设 f_n 和 f 是随机变量 X_n 和 X 对应的特征函数, 则 $X_n \xrightarrow{d} X$ 当且仅当

$$f_n(t) \to f(t), \quad \forall t \in \mathbb{R}.$$

证明. 略. □

3 矩收敛

3.1 矩收敛及其基本性质

考虑随机变量序列 $\{X_n, n \geq 1\}$, 并设随机变量 $X < \infty, \text{a.s.}$. 首先回忆矩收敛的定义.

定义 3.1 (依 r 阶矩收敛)**.** 设 $X_n \in L^r(\Omega)$, 如果

$$\lim_{n\to\infty} \mathbb{E}|X_n - X|^r = 0,$$

则称 X_n **依 r 阶矩收敛**于 X, 记作 $X_n \xrightarrow{L^r} X$.

以下均设 $X_n, X \in L^r(\Omega)$. 为便于研究矩收敛的性质, 在这里引入一个常用的不等式 (来自苏淳的书上).

命题 3.1 (C_r 不等式, 二元情形)**.** 设 $r > 0$, $x, y \in \mathbb{R}$, 则

$$|x+y|^r \leq C_r \cdot (|x|^r + |y|^r),$$

其中

$$C_r = \begin{cases} 1, & 0 < r \leq 1, \\ 2^{r-1}, & r > 1. \end{cases}$$

证明. 首先设 $0 < r \leq 1$, 若 $x = y = 0$, 则不等式取等; 否则

$$|x+y|^r \leq (|x|+|y|)^r = \frac{|x|}{(|x|+|y|)^{1-r}} + \frac{|y|}{(|x|+|y|)^{1-r}} \leq |x|^r + |y|^r;$$

其次设 $r > 1$, 由 Jensen 不等式得

$$\left|\frac{x+y}{2}\right|^r \leq \left(\frac{|x|+|y|}{2}\right)^r \leq \frac{|x|^r+|y|^r}{2} \implies |x+y|^r \leq 2^{r-1}\cdot(|x|^r+|y|^r).$$

这便证明了原不等式. □

推论 (C_r 不等式, n 元情形)**.** 设 $r > 0$, $x_1, x_2, \cdots, x_n \in \mathbb{R}$, 则

$$\left|\sum_{i=1}^n x_i\right|^r \leq C_r \cdot \sum_{i=1}^n |x_i|^r,$$

其中

$$C_r = \begin{cases} 1, & 0 < r \leq 1, \\ n^{r-1}, & r > 1. \end{cases}$$

证明. 分别应用数学归纳法或 n 元 Jensen 不等式即可. □

命题 3.2. 设 $r > 0$, $X_n \xrightarrow{L^r} X$, 则 $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$.

证明. 由 C_r 不等式得

$$|X_n|^r = |X_n - X + X|^r \le C_r \cdot (|X_n - X|^r + |X|^r),$$

对上式取期望, 并令 $n \to \infty$, 可得

$$\mathbb{E}(|X_n|^r) - \mathbb{E}(|X|^r) = C_r \cdot \mathbb{E}|X_n - X|^r \to 0,$$

因此 $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$. □

当 $X_n \xrightarrow{\text{a.s.}} X$ 或 $X_n \xrightarrow{d} X$ 时, 保证矩收敛的条件是有用的, 以下将分别叙述.

命题 3.3. 设 $r > 0$, $X_n \xrightarrow{\text{a.s.}} X$, 则

$$\mathbb{E}|X|^r \le \liminf_{n\to\infty} \mathbb{E}|X_n|^r.$$

证明. 由 $X_n \xrightarrow{\text{a.s.}} X$ 知 $|X_n|^r \xrightarrow{\text{a.s.}} |X|^r$, 应用 Fatou 引理得

$$\mathbb{E}|X|^r = \mathbb{E}\left(\lim_{n\to\infty} |X_n|^r\right) \le \liminf_{n\to\infty} \mathbb{E}|X_n|^r.$$

这便证明了该定理. □

引理 3.4 (Helly 第二定理)**.** 设 F_n, F 是分布函数, 且 $F_n \xrightarrow{v} F$, 则对任意的有界连续函数 $g : \mathbb{R} \to \mathbb{R}$, 都有

$$\lim_{n\to\infty} \int_{\mathbb{R}} g(x)\mathrm{d}F_n(x) = \int_{\mathbb{R}} g(x)\mathrm{d}F(x).$$

证明. 设 $\Omega = (0,1), \omega \in \Omega$, 定义

$$X_n(\omega) = \inf\{x \in \mathbb{R} : F_n(x) \le \omega\}, \quad X(\omega) = \inf\{x \in \mathbb{R} : F(x) \le \omega\},$$

则 X_n, X 的分布函数是 F_n, F, 且根据 F 的连续点在 $\mathbb{R}$ 中稠密, 知 $F_n \xrightarrow{\text{a.s.}} F$, 从而 $X_n \xrightarrow{\text{a.s.}} X$, 进而对有界连续函数 g, 有 $g(X_n) \xrightarrow{\text{a.s.}} g(X)$. 由 Lebesgue 控制收敛定理 (或者称为有界收敛定理), 得

$$\lim_{n\to\infty} \int_{\mathbb{R}} g(x)\mathrm{d}F_n(x) = \lim_{n\to\infty} \mathbb{E}g(X_n) = \mathbb{E}g(X) = \int_{\mathbb{R}} g(x)\mathrm{d}F(x),$$

这便证明了该结论. □

命题 3.5. 设 $r > 0$, $X_n \xrightarrow{d} X$, 如果存在 $p > 0$, 使得

$$\sup_{n\ge1} \mathbb{E}|X_n|^p < \infty,$$

则对任意的 $r < p$, 都有 $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$.

证明. 设 F_n, F 分别是 X_n, X 的分布函数, 则 $F_n \xrightarrow{v} F$. 对于 $A > 0$, 定义函数

$$f_A(x) = \begin{cases} |x|^r, & |x| \le A, \\ A^r, & |x| > A, \end{cases}$$

则 f_A 是有界连续函数, 根据 Helly 第二定理得

$$\lim_{n\to\infty}\int_{\mathbb{R}} f_A(x)\mathrm{d}F_n(x) = \int_{\mathbb{R}} f_A(x)\mathrm{d}F(x),$$

并且

$$\begin{aligned}\int_{\mathbb{R}} \left|f_A(x) - |x|^r\right| \mathrm{d}F_n(x) &\le \int_{|x|>A} |x|^r \mathrm{d}F_n(x) \\ &= \mathbb{E}\left(|X_n|^r \cdot I_{\{|X_n|>A\}}\right) \\ &\le \frac{1}{A^{p-r}} \cdot \mathbb{E}\left(|X_n|^p \cdot I_{\{|X_n|>A\}}\right) \\ &\le \frac{M}{A^{p-r}},\end{aligned}$$

因此当 $A\to\infty$ 时, $\int_{\mathbb{R}} f_A(x)\mathrm{d}F_n(x)$ 对 n 一致收敛于 $\int_{\mathbb{R}} |x|^r\mathrm{d}F(x) = \mathbb{E}(|X|^r)$, 进而有

$$\begin{aligned}\lim_{n\to\infty}\int_{-\infty}^{\infty} |x|^r\mathrm{d}F_n(x) &= \lim_{n\to\infty}\int_{-\infty}^{\infty} |x|^r\mathrm{d}F_n(x) \\ &= \lim_{n\to\infty}\lim_{A\to\infty}\int_{-\infty}^{\infty} f_A(x)\mathrm{d}F_n(x) \\ &= \lim_{A\to\infty}\lim_{n\to\infty}\int_{-\infty}^{\infty} f_A(x)\mathrm{d}F_n(x) \\ &= \lim_{A\to\infty}\int_{-\infty}^{\infty} f_A(x)\mathrm{d}F(x) \\ &= \int_{\mathbb{R}} |x|^r\mathrm{d}F(x),\end{aligned}$$

也即 $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$. □

3.2 随机变量序列的一致可积

现在开始探讨 $X_n \xrightarrow{p} X$ 与 $X_n \xrightarrow{L^r} X$ 之间的关系. 我们需要对 $\{X_n\}$ 加一些条件. 对随机变量序列 $\{X_n\}$, 在此引入一个新的定义.

定义 3.2 (一致可积)**.** 设 $\{X_n\}$ 是随机变量序列, 如果

$$\lim_{A\to\infty}\sup_{n\ge 1}\mathbb{E}\left(|X_n|\cdot I_{\{|X_n|>A\}}\right) = 0,$$

则称 $\{X_n\}$ **一致可积**.

一致可积也可以写成

$$\lim_{A\to\infty}\mathbb{E}\left(|X_n|\cdot I_{\{|X_n|>A\}}\right) = 0, \quad \forall n\ge 1.$$

接下来给出一致可积的等价形式.

定理 3.6. 设 $\{X_n\}$ 是随机变量序列, 则 $\{X_n\}$ 一致可积当且仅当以下两条性质同时成立:

(1) 一致有界, 也即存在 $M > 0$, 使得

$$\sup_{n\geq 1} \mathbb{E}|X_n| < M;$$

(2) 一致绝对连续, 也即对任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对任意的满足 $\mathbb{P}(E) < \delta$ 的 $E \in \mathscr{F}$, 都有

$$\sup_{n\geq 1} \mathbb{E}\left(|X_n| \cdot I_E\right) < \varepsilon.$$

证明. 一方面, 设 $\{X_n\}$ 一致可积, 则对任意的 $n \geq 1$, 都有

$$\lim_{A\to\infty} \mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) = 0,$$

从而存在 A, 使得 $\mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) < 1$, 进一步有

$$\begin{aligned}
\mathbb{E}|X_n| &= \mathbb{E}\left(|X_n| \cdot I_{\{|X_n|\leq A\}}\right) + \mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) \\
&\leq A \cdot \mathbb{P}(|X_n| \leq A) + 1 \\
&\leq A + 1,
\end{aligned}$$

上式与 n 无关, 这便说明了 $\{X_n\}$ 一致有界; 同时, 对任意的 $\varepsilon > 0$, 存在 $A > 0$, 使得

$$\mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) < \frac{\varepsilon}{2}.$$

取 $\delta = \dfrac{\varepsilon}{2A}$, 设 $E \in \mathscr{F}$, 且 $\mathbb{P}(E) < \delta$, 则

$$\begin{aligned}
\mathbb{E}\left(|X_n| \cdot I_E\right) &= \mathbb{E}\left(|X_n| \cdot I_E \cdot I_{\{|X_n|\leq A\}}\right) + \mathbb{E}\left(|X_n| \cdot I_E \cdot I_{\{|X_n|>A\}}\right) \\
&\leq A \cdot \mathbb{P}(E) + \mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) \\
&\leq A \cdot \frac{\varepsilon}{2A} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}$$

上式与 n 无关, 说明了 $\{X_n\}$ 一致绝对连续.

反之, 设 $\{X_n\}$ 一致有界且一致绝对连续, 则对任意的 $n \geq 1$, 应用 Chebyshev 不等式得

$$\mathbb{P}(|X_n| > A) \leq \frac{\mathbb{E}|X_n|}{A} < \frac{M}{A},$$

对任意的 $\delta > 0$, 只要 $A > \dfrac{M}{\delta}$, 就有 $\mathbb{P}(|X_n| > A) < \delta$. 取 $E = \{|X_n| > A\}$, 对任意的 $\varepsilon > 0$, 都存在 δ 以及 $A > \dfrac{M}{\delta}$, 使得

$$\mathbb{E}\left(|X_n| \cdot I_E\right) = \mathbb{E}\left(|X_n| \cdot I_{\{|X_n|>A\}}\right) < \varepsilon,$$

这便说明了 $\{X_n\}$ 绝对可积. □

介绍一致可积性是为了在 $X_n \xrightarrow{p} X$ 的情况下, 探究加入 $X_n \xrightarrow{L^r} X$ 的条件所能得到的结果. 以下设 $X_n, X \in L^r(\Omega)$.

定理 3.7. 设 $r > 0$, $X_n \xrightarrow{p} X$, 下列命题等价*:*

(1) $\{|X_n|^r\}$ 一致可积*;*

(2) $X_n \xrightarrow{L^r} X$*;*

(3) $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$.

证明. (1) $\Longrightarrow$ (2): 设 $\{|X_n|^r\}$ 一致可积, 对任意的 $n \geq 1$, 由 C_r 不等式得

$$|X_n - X|^r \leq 2^{r-1} \cdot (|X_n|^r + |X|^r),$$

因此 $\{|X_n - X|^r\}$ 也一致可积. 设 $\varepsilon > 0$, 由 $X_n \xrightarrow{p} X$ 得 $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$, 并且存在 $M > 0$, 使得 $|X_n - X| < M$, a.s.. 计算得

$$\begin{aligned}\mathbb{E}|X_n - X|^r &= \mathbb{E}\left(|X_n - X|^r \cdot I_{\{|X_n - X| \leq \varepsilon\}}\right) + \mathbb{E}\left(|X_n - X|^r \cdot I_{\{|X_n - X| > \varepsilon\}}\right) \\ &\leq \varepsilon^r + \mathbb{E}\left(|X_n - X|^r \cdot I_{\{|X_n - X| > \varepsilon\}}\right) \\ &\leq \varepsilon^r + M^r \cdot \mathbb{P}(|X_n - X| > \varepsilon) \\ &\to \varepsilon^r,\end{aligned}$$

再令 $\varepsilon \to 0$ 即可得 $X_n \xrightarrow{L^r} X$.

(2) $\Longrightarrow$ (3): 这是已经证明的结论.

(3) $\Longrightarrow$ (1): 设 $\mathbb{E}|X_n|^r \to \mathbb{E}|X|^r$, 对于 $A > 0$, 由 Fatou 引理得

$$\mathbb{E}\left(|X|^r \cdot I_{\{|X|^r \leq A\}}\right) \leq \liminf_{n\to\infty} \mathbb{E}\left(|X_n|^r \cdot I_{\{|X_n|^r \leq A\}}\right),$$

因此

$$\limsup_{n\to\infty} \mathbb{E}\left(|X_n|^r \cdot I_{\{|X_n|^r > A\}}\right) \leq \mathbb{E}\left(|X|^r \cdot I_{\{|X|^r > A\}}\right).$$

由 $X \in L^r(\Omega)$ 知, 当 $A \to \infty$ 时有 $\mathbb{E}\left(|X|^r \cdot I_{\{|X|^r > A\}}\right) \to 0$, 因此对任意的 $\varepsilon > 0$, 存在 $A_0 > 0$ 及 n_0, 使得当 $A > A_0$ 时, 有

$$\sup_{n > n_0} \mathbb{E}\left(|X_n|^r \cdot I_{\{|X_n|^r > A\}}\right) < \varepsilon.$$

又当 $n \leq n_0$ 时, 根据 $X_n \in L^r(\Omega)$ 知, 当 $A \to \infty$ 时有 $\mathbb{E}\left(|X_n|^r \cdot I_{\{|X|^r > A\}}\right) \to 0$, 因此

$$\lim_{A\to\infty} \sup_{n\geq 1} \mathbb{E}\left(|X_n|^r \cdot I_{\{|X_n|^r > A\}}\right) = 0,$$

也即 $\{|X_n|^r\}$ 一致可积. $\square$

推论. 设 $\{X_n\}$ 一致可积, 则 $X_n \xrightarrow{p} X$ 当且仅当 $X_n \xrightarrow{L^r} X$.

证明. 直接应用上述结论即可. $\square$