概率论中的截尾术举例

宗语轩

2023 秋, USTC

对于随机变量 X, 主要有以下三类截尾方法:

$$X_1 = XI_{\{|X|\leqslant M\}} = \begin{cases} X, & |X| \leqslant M, \\ 0, & |X| > M. \end{cases}, \quad X_2 = \begin{cases} X, & |X| \leqslant M, \\ M, & X > M, \\ -M, & X < -M. \end{cases}, \quad X_3 = \begin{cases} X, & X \leqslant M, \\ M, & X > M. \end{cases}.$$

注. 截尾法常常用于证明有关随机变量列的收敛中, 目的是常常借助题干条件来转化问题. 对于随机变量列 $\{X_n\}$, 一种方法是一致地令 $M \uparrow +\infty$, 另一种方法是对每个 X_n 取 $M(n) = k_n$, 使得 $\{k_n\}$ 递增 (如 $M = n$) 再令 $n \to +\infty$ 来处理. 这样截尾的好处是截尾后的随机变量具有有界性 (可能还利用了随机变量列的单调性). 一方面, 有界性的保证大大可以施展手脚 (比如很多定理、命题的适用条件都有有界性的要求, 以及随机变量的有界性能保证其任意阶矩的存在性, 从而便于矩不等式的施展); 另一方面, 可能存在的单调性, 尾巴 $X - X_i$ 部分保证了目标的转化条件.

我们下举一例有关截尾的例子, 这也是强大数律的证明中一个重要的转化步骤.

例 0.1. 设非负随机变量 X_i 独立同分布且 $\mathbb{E}[X_1] < \infty$, 令 $Y_n = X_n I_{\{|X_n|\leqslant n\}}$, 则有

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - Y_k) \xrightarrow{\text{a.s.}} 0.$$

从而

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\text{a.s.}} a \Longleftrightarrow \frac{1}{n}\sum_{k=1}^{n} Y_k \xrightarrow{\text{a.s.}} a$$

证明. 对 $X \geqslant 0$, 利用

$$\sum_{m=1}^{\infty} \mathbb{P}(X \geq m) \leq \mathbb{E}[X] \leq 1 + \sum_{m=1}^{\infty} \mathbb{P}(X \geq m),$$

有

$$\sum_{n=1}^{+\infty} \mathbb{P}(Y_n \neq X_n) = \sum_{n=1}^{+\infty} \mathbb{P}(X_n \geq n) = \sum_{n=1}^{+\infty} \mathbb{P}(X_1 \geq n) \leq \mathbb{E}[X_1] < \infty$$

利用 Borel-Cantelli 引理, $\mathbb{P}(X_n \neq Y_n \text{ i.o.}) = 0$, 即几乎处处 $\{X_n \neq Y_n\}$ 只发生有限次, 从而

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) \xrightarrow{\text{a.s.}} 0.$$

□

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.

发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

注. 通过截尾, 证明原随机变量列的收敛转化成证明截尾后的随机变量列的收敛.

我们再来看几个例子:

例 0.2. 随机变量列 $\{X_n\}$ 相互独立且满足

$$\mathbb{P}(X_n = \pm n) = \frac{1}{2n^3}, \quad \mathbb{P}(X_n = \pm 1) = \frac{1}{2} - \frac{1}{2n^3}, \quad n \in \mathbb{N}^*.$$

证明:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow{\mathrm{D}} N(0,1).$$

证明. 引入截尾处理后的随机变量: $Y_i = \mathrm{sgn} X_i$, 注意到

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - Y_k)$$

一方面, 利用 CLT 知,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{\mathrm{D}} N(0,1).$$

另一方面, 对 $\forall \varepsilon > 0$, 有

$$\begin{aligned}\mathbb{P}\left(\left| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - Y_k) \right| > \varepsilon \right) &\leqslant \frac{1}{\sqrt{n}\varepsilon} \mathbb{E}[| \sum_{k=1}^{n} (X_k - Y_k)|] \leqslant \frac{1}{\sqrt{n}\varepsilon} \sum_{k=1}^{n} \mathbb{E}[|X_k - Y_k|] \\ &\leqslant \frac{1}{\sqrt{n}\varepsilon} \sum_{k=1}^{+\infty} \frac{1}{k^2} \xrightarrow{n \to +\infty} 0.\end{aligned}$$

因此

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - Y_k) \xrightarrow{\mathrm{P}} 0.$$

结合 **Slutsky 引理**知,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \xrightarrow{\mathrm{D}} N(0,1).$$

□

例 0.3. 设 $X_n \xrightarrow{\mathrm{D}} X$, 且存在 $r, C > 0$, 使得 $\mathbb{E}[|X_n|^r] \leqslant C$, 证明: 对 $\forall 0 < s < r$, 均有

$$\lim_{n \to +\infty} \mathbb{E}[|X_n|^s] = \mathbb{E}[X^s].$$

证明. 对 $\forall g \in \mathrm{C_b}(\mathbb{R})$(有界连续函数), 有 $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)], n \to \infty$. 设 $M > 0$, 定义

$$g_M(x) = \begin{cases} |x|^s, & |x| < M, \\ M^s, & |x| \geqslant M. \end{cases}$$

则 $g_M(x)$ 是有界连续函数, 以及当 $M \to +\infty$ 时 $0 \leqslant g_M(x) \uparrow |x|^s$, 因此有

$$\lim_{n \to +\infty} \mathbb{E}[g_M(X_n)] = \mathbb{E}[g_M(X)].$$

同时我们有

$$0 \leqslant \mathbb{E}[|X_n|^s] - \mathbb{E}[g_M(X_n)] \leqslant \mathbb{E}[|X_n|^s I_{\{|X_n| \geqslant M\}}] \leqslant \frac{1}{M^{r-s}} \mathbb{E}[|X_n|^r I_{\{|X_n| \geqslant M\}}] \leqslant \frac{C}{M^{r-s}},$$

因此

$$\mathbb{E}[g_M(X_n)] \leqslant \mathbb{E}[|X_n|^s] \leqslant \mathbb{E}[g_M(X_n)] + \frac{C}{M^{r-s}}.$$

先令 $n \to +\infty$, 再令 $M \uparrow +\infty$, 结合 $\lim\limits_{M \to +\infty} \mathbb{E}[g_M(X)] = \mathbb{E}[|X|^s]$(利用单调收敛定理), 得

$$\lim_{n \to +\infty} \mathbb{E}[|X_n|^s] = \mathbb{E}[X^s].$$

□

例 0.4 (截尾术证明弱大数律). 设 $X_1, \cdots, X_n$ 独立同分布, $\mu = \mathbb{E}[X_1]$ 且 $\mathbb{E}[|X_1|] < +\infty$, 记 $S_n = \sum\limits_{k=1}^{n} X_k$, 则

$$\frac{S_n}{n} \xrightarrow{P} \mu.$$

证明. 引入经截断处理的随机变量:

$$X_n^{(1)} = X_n I_{\{|X_n| \leqslant M\}}, \quad X_n^{(2)} = X_n I_{\{|X_n| > M\}}.$$

其中 $X_n = X_n^{(1)} + X_n^{(2)}$. 并设 $S_n^{(1)} = \sum\limits_{i=1}^{n} X_i^{(1)}, S_n^{(2)} = \sum\limits_{i=1}^{n} X_i^{(2)}$. 则对 $\forall \varepsilon > 0$, 有

$$\begin{aligned} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\frac{(S_n^{(1)} - \mathbb{E}[S_n^{(1)}]) + (S_n^{(2)} - \mathbb{E}[S_n^{(2)}])}{n}\right| > \varepsilon\right) \\ &\leqslant \mathbb{P}\left(\left|\frac{S_n^{(1)} - \mathbb{E}[S_n^{(1)}]}{n}\right| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(\left|\frac{S_n^{(2)} - \mathbb{E}[S_n^{(2)}]}{n}\right| > \frac{\varepsilon}{2}\right) \end{aligned}$$

现在我们分别处理这两部分:

$$\mathbb{P}\left(\left|\frac{S_n^{(1)} - \mathbb{E}[S_n^{(1)}]}{n}\right| > \frac{\varepsilon}{2}\right) \leqslant \frac{\mathrm{Var}(S_n^{(1)})}{(\frac{1}{2}\varepsilon n)^2} = \frac{4\mathrm{Var}(X_1^{(1)})}{\varepsilon^2 n} \leqslant \frac{4\mathbb{E}[(X_1^{(1)})^2]}{\varepsilon^2 n} \leqslant \frac{4M^2}{\varepsilon^2 n} \xrightarrow{n \to +\infty} 0.$$

同时有

$$\mathbb{P}\left(\left|\frac{S_n^{(2)} - \mathbb{E}[S_n^{(2)}]}{n}\right| > \frac{\varepsilon}{2}\right) \leqslant \frac{\mathbb{E}|S_n^{(2)} - \mathbb{E}[S_n^{(2)}]|}{\frac{\varepsilon}{2} n} \leqslant \frac{4\mathbb{E}[|X_1| I_{\{|X_1| > M\}}]}{\varepsilon} \xrightarrow{M \to +\infty} 0,$$

其中

$$\mathbb{E}|S_n^{(2)} - \mathbb{E}[S_n^{(2)}]| \leqslant \mathbb{E}[|S_n^{(2)}|] + |\mathbb{E}[S_n^{(2)}]| \leqslant 2\mathbb{E}[|S_n^{(2)}|] \leqslant 2\sum_{k=1}^{n} \mathbb{E}[|X_k^{(2)}|] = 2n\mathbb{E}[|X_1^{(2)}|].$$

因此令 $n \to +\infty$, 再令 $M \to +\infty$, 有

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \to 0.$$

故 $\frac{S_n}{n} \xrightarrow{P} \mu$. □

注. 分别处理并运用 Markov 不等式等矩不等式. 因为 $X_n^{(1)}$ 有界, 因此其常用于高阶矩方法估计, 而尾巴部分 $X_n^{(2)}$ 通过 $\mathbb{E}[|X|] < \infty$ 用一阶矩来控制.