概率论与数理统计 (Fall 2024)
习题课讲义

2024 年 11 月 6 日

聂嵘琢, 祝国隽

B01GB006Y-02/Note8

---

1. 把一正方体的各个面依次标记 $1,2,3,4,5,6$. 在桌上随机抛掷该正方体, 考察正方体朝上的一面出现的数字. 记连续抛郑 X 次后第一次连续出现两次 6. 求 EX.

> **Ans:** 记
> $$Y=\begin{cases}1, & \text{第1 次抛掷点数为6}\\ 0, & \text{第1 次抛掷点数不为6,}\end{cases}$$
> 则 $P(Y=1)=\dfrac{1}{6}, P(Y=0)=\dfrac{5}{6}$.
> 当 $Y=0$ 时, 相当于第 1 次白投了, 期望抛郑次数为 $EX+1$.
> 当 $Y=1$ 时, 若第 2 次抛掷点数为 6 , 则期望抛掷次数为 2 ; 若第 2 次抛郑点数不为 6 , 相当于这 2 次白投了, 则期望抛掷次数为 $EX+2$. 则
> $$\begin{aligned} EX &= E[E(X \mid Y)] \\ &= E(X \mid Y=0)P(Y=0)+E(X \mid Y=1)P(Y=1) \\ &= (EX+1)\times\frac{5}{6}+\frac{1}{6}\times\left[2\times\frac{1}{6}+(EX+2)\times\frac{5}{6}\right] \\ &= \frac{35}{36}EX+\frac{7}{6} \\ &= 42. \end{aligned}$$

2. 令 $\left(X_1, X_2\right)^{\mathrm{T}} \sim N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right), \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$ 已知.

   (1) 求 $\left(X_1-\mu_1\right)^2 / \sigma_1^2$ 与 $\left(X_2-\mu_2\right)^2 / \sigma_2^2$ 的相关系数;

(2) 记 a_1, a_2 为已知常数, 求 $a_1X_1 + a_2X_2$ 与 $a_1X_1 - a_2X_2$ 的相关系数;

(3) 记 $Y = aX_1 + bX_2$ 和 $Z = aX_1 - bX_2$. 若 $EY = \theta$ 且 Y, Z 相互独立, 求 a 和 b .

Ans: (1) 由题意有 $Z_1 = \dfrac{X_1 - \mu_1}{\sigma_1} \sim N(0,1), Z_2 = \dfrac{X_2 - \mu_2}{\sigma_2} \sim N(0,1)$.

$$\begin{aligned}
&\operatorname{cov}(Z_1, Z_2) = \frac{1}{\sigma_1\sigma_2}\operatorname{cov}(X_1, X_2) = \rho \\
&\operatorname{var}\left(\left(\frac{X_1-\mu_1}{\sigma_1}\right)^2\right) = \operatorname{var}\left(Z_1^2\right) = EZ_1^4 - \left(EZ_1^2\right)^2 = 2 \\
&\operatorname{var}\left(\left(\frac{X_2-\mu_2}{\sigma_2}\right)^2\right) = \operatorname{var}\left(Z_2^2\right) = 2 \\
&\operatorname{cov}\left(\left(\frac{X_1-\mu_1}{\sigma_1}\right)^2, \left(\frac{X_2-\mu_2}{\sigma_2}\right)^2\right) = \operatorname{cov}\left(Z_1^2, Z_2^2\right) = EZ_1^2Z_2^2 - 1
\end{aligned}$$

又 $(Z_1, Z_2)^\top \sim N(0,0,1,1,\rho)$. 因此,

$$\begin{aligned}
EZ_1^2Z_2^2 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} z_1^2z_2^2\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left(z_1^2 - 2\rho z_1z_2 + z_2^2\right)\right\}dz_2dz_1 \\
&= \int_{-\infty}^{+\infty}\frac{1}{2\pi\sqrt{1-\rho^2}}z_1^2\exp\left(-\frac{z_1^2}{2}\right)\int_{-\infty}^{+\infty} z_2^2\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left(z_2 - \rho z_1\right)^2\right\}dz_2dz_1 \\
&= \int_{-\infty}^{+\infty}\frac{1}{2\pi\sqrt{1-\rho^2}}z_1^2\exp\left(-\frac{z_1^2}{2}\right)\int_{-\infty}^{+\infty}\left(\sqrt{1-\rho^2} + \rho z_1\right)^2\exp\left\{-\frac{t^2}{2}\right\}\sqrt{1-\rho^2}dtdz_1 \\
&= \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}z_1^2\exp\left(-\frac{z_1^2}{2}\right)\left[\left(1-\rho^2\right)\times 1 + \rho^2z_1^2\right]dz_1 \\
&= \left(1-\rho^2\right)\int_{-\infty}^{+\infty}\frac{z_1^2}{\sqrt{2\pi}}\exp\left(-\frac{z_1^2}{2}\right)dz_1 + \rho^2\int_{-\infty}^{+\infty}\frac{z_1^4}{\sqrt{2\pi}}\exp\left(-\frac{z_1^2}{2}\right)dz_1 \\
&= \left(1-\rho^2\right)\times 1 + \rho^2\times 3 \\
&= 1 + 2\rho^2.
\end{aligned}$$

从而, $\operatorname{cov}\left(\left(\dfrac{X_1-\mu_1}{\sigma_1}\right)^2, \left(\dfrac{X_2-\mu_2}{\sigma_2}\right)^2\right) = 1 + 2\rho^2 - 1 = 2\rho^2$, $\left(\dfrac{X_1-\mu_1}{\sigma_1}\right)^2$ 和 $\left(\dfrac{X_2-\mu_2}{\sigma_2}\right)^2$ 的相关系数为 $\rho_1 = \dfrac{2\rho^2}{\sqrt{2\times 2}} = \rho^2$.

(2)

$$\begin{aligned}\operatorname{cov}\left(a_1X_1+a_2X_2,a_1X_1-a_2X_2\right)=&a_1^2\operatorname{Cov}\left(X_1,X_1\right)-a_1a_2\operatorname{Cov}\left(X_1,X_2\right)\\&+a_1a_2\operatorname{Cov}\left(X_2,X_1\right)-a_2^2\operatorname{Cov}\left(X_2,X_2\right)\\=&a_1^2\sigma_1^2-a_2^2\sigma_2^2\end{aligned}$$

$$\begin{aligned}\operatorname{var}\left(a_1X_1+a_2X_2\right)&=a_1^2\operatorname{Var}\left(X_1\right)+a_2^2\operatorname{Var}\left(X_2\right)+2a_1a_2\operatorname{Cov}\left(X_1,X_2\right)\\&=a_1^2\sigma_1^2+a_2^2\sigma_2^2+2a_1a_2\rho\sigma_1\sigma_2\\\operatorname{var}\left(a_1X_1-a_2X_2\right)&=a_1^2\operatorname{Var}\left(X_1\right)+a_2^2\operatorname{Var}\left(X_2\right)-2a_1a_2\operatorname{Cov}\left(X_1,X_2\right)\\&=a_1^2\sigma_1^2+a_2^2\sigma_2^2-2a_1a_2\rho\sigma_1\sigma_2\end{aligned}$$

从而 $a_1x_1+a_2x_2$ 和 $a_1x_1-a_2x_2$ 的相关系数为

$$\rho_2=\frac{a_1^2\sigma_1^2-a_2^2\sigma_2^2}{\sqrt{\left(a_1^2\sigma_1^2+a_2^2\sigma_2^2\right)^2-4a_1^2a_2^2\rho^2\sigma_1^2\sigma_2^2}}$$

(3) $EY=\theta,\operatorname{cov}(Y,Z)=0,$

$$\begin{aligned}&\theta=EY=E\left(aX_1+bX_2\right)=aE\left(X_1\right)+bE\left(X_2\right)=a\mu_1+b\mu_2\\&\theta=\operatorname{Cov}(Y,Z)=\operatorname{cov}\left(aX_1+bX_2,aX_1-bX_2\right)\\&\quad=a^2\operatorname{cov}\left(X_1,X_1\right)-ab\operatorname{cov}\left(X_1,X_2\right)+ab\operatorname{cov}\left(X_2,X_1\right)-b^2\operatorname{cov}\left(X_2,X_2\right)\\&\quad=a^2\sigma_1^2-b^2\sigma_2^2.\end{aligned}$$

联立解得

$$\begin{cases}a^2=\dfrac{\sigma_2^2}{\sigma_1^2}b^2,\\b^2=\dfrac{\theta^2}{\left(\dfrac{\sigma_2}{\sigma_1}\mu_1\pm\mu_2\right)^2}\end{cases}$$

3. 令随机变量 $X\sim N(0,1),Y$ 服从 -1 和 1 等可能的两点分布, 且 X 和 Y 相互独立. 求证: $XY\sim N(0,1)$, 且 XY 与 X 不相关也不独立.

Ans: 记 $Z = XY, Z$ 的分布函数为

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(XY \leq z) \\ &= P(XY \leq z \mid Y = 1)P(Y = 1) + P(XY \leq z \mid Y = -1)P(Y = -1) \\ &= P(X \leq z)\frac{1}{2} + P(X \geq -z)\frac{1}{2} \\ &= \frac{1}{2}[\Phi(z) + 1 - \Phi(-z)] \\ &= \Phi(z) \end{aligned}$$

故 $Z = XY \sim N(0, 1)$.

$$\operatorname{cov}(XY, X) = EX^2Y - EXYEX = EX^2EY - EX^2EY = \operatorname{Var}(X)EY = 0,$$

则 XY 与 X 不相关. 下面说明 $Z = XY$ 与 X 不独立.

$$\begin{aligned} P(X > 1) &= P(Z > 1) = 1 - \Phi(1) \\ P(X > 1, Z > 1) &= P(X > 1, XY > 1) \\ &= P(X > 1, XY > 1, Y = 1) + P(X > 1, XY > 1, Y = -1) \\ &= P(X > 1, Y = 1) + P(X > 1, X < -1, Y = -1) \\ &= P(X > 1)P(Y = 1) \\ &= \frac{1 - \Phi(1)}{2}. \end{aligned}$$

由于 $P(X > 1, Z > 1) \neq P(X > 1)P(Z > 1)$, 故 $Z = XY$ 与 X 不独立.

4. 令随机变量 X 的密度函数为偶函数. 求证: X 与 $|X|$ 不相关也不独立.

Ans: 由于 X 的密度函数 $f(x)$ 是偶函数, $f(x) = f(-x), xf(x)$ 是奇函数,

$$EX = \int_{-\infty}^{\infty} xf(x)dx = 0, E(XE|X|) = E|X|EX = 0.$$

又由于 $x|x|f(x)$ 是奇函数,

$$E(X|X|) = \int_{-\infty}^{\infty} x|x|f(x)dx = 0 = EXE|X|$$

从而 $\operatorname{cov}(X,|X|)=0$. 即 X 与 $|X|$ 不相关.

下面说明 X 与 $|X|$ 不独立.

可选 $c>0$, 使 $P(|X|<c)\in(0,1)$ 且 $P(X<c)<1$, 而 $\{|X|<c\}\subset\{X<c\}$, 则 $P(X<c,|X|<c)=P(|X|<c)\neq P(X<c)P(|X|<c)$, 故 X 与 $|X|$ 不独立.

5. 令 $X_1,X_2,\cdots,X_n$ 独立同分布, 服从 $U(a,b)$. 记

$$X_{(1)}=\min\left\{X_1,X_2,\cdots,X_n\right\},X_{(n)}=\max\left\{X_1,X_2,\cdots,X_n\right\}$$

求证: $X_{(1)}\xrightarrow{p}a, X_{(n)}\xrightarrow{p}b$.

Ans: $Y_1=X_{(1)}$ 的分布函数和密度函数为

$$\begin{aligned}F_{Y_1}\left(y_1\right)&=P\left(X_{(1)}\leq y_1\right)=1-P\left(\min\left\{X_1,\ldots,X_n\right\}>y_1\right)\\&=1-P\left(X_1>y_1\right)P\left(X_2>y_2\right)\ldots P\left(X_n>y_n\right)\\&=1-\left(1-F_X\left(y_1\right)\right)^n.\end{aligned}$$

$$f_{Y_1}\left(y_1\right)=n\left(1-F_X\left(y_1\right)\right)^{n-1}f_X\left(y_1\right)=n\left(1-\frac{y_1-a}{b-a}\right)^{n-1}\frac{1}{b-a},a\leq y_1\leq b$$

$$\begin{aligned}EX_{(1)}=EY_1&=\int_a^b y_1 f_{Y_1}\left(y_1\right)dy_1\quad\left(1-\frac{y_1-a}{b-a}=t\right)\\&=\int_0^1[(1-t)(b-a)+a]nt^{n-1}dt\\&=\int_0^1\left(bnt^{n-1}-(b-a)nt^n\right)dt\\&=b-\frac{(b-a)n}{n+1}\\&=\frac{n}{n+1}a+\frac{1}{n+1}b\to a\quad(n\to\infty)\end{aligned}$$

$$\begin{aligned}EX_{(1)}^2&=\int_a^b y_1^2 f_{Y_1}\left(y_1\right)dy_1\quad\left(1-\frac{y_1-a}{b-a}=t\right)\\&=\int_0^1[(1-t)(b-a)+a]^2nt^{n-1}dt\\&=b^2-2b(b-a)\frac{n}{n+1}+(b-a)^2\frac{n}{n+2}.\end{aligned}$$

则

$$\operatorname{var}\left(X_{(1)}\right)=b^2-2b(b-a)\frac{n}{n+1}+(b-a)^2\frac{n}{n+2}-\left(\frac{na+b}{n+1}\right)^2$$
$$\longrightarrow 0,\quad (n\to\infty).$$

由 Chebyshev 不等式,

$$P\left(\left|X_{(1)}-a\right|>\varepsilon\right)\le\frac{\operatorname{Var}\left(X_{(1)}\right)}{\varepsilon^2}\longrightarrow 0,\quad (n\to\infty),$$

即 $X_{(1)}\xrightarrow{P}a$.

同理可证

$$EX_{(n)}\to b,\quad (n\to\infty)$$
$$\operatorname{var}\left(X_{(n)}\right)\to 0,\quad (n\to\infty)$$

因此, $X_{(n)}\xrightarrow{P}b$.

6. 令随机变量序列 $X_1,X_2,\cdots$ 独立同分布, 服从 $U(0,1),c$ 是给定的常数, $0<c<1$. 求证: $(\prod_{i=1}^n X_i)^{\frac{1}{n}}\xrightarrow{p}\frac{1}{\mathrm{e}}$.

Ans: 记 $\ln Z_n=\ln\left(\prod_{i=1}^n X_i\right)^{\frac{1}{n}}=\frac{1}{n}\sum_{i=1}^n\ln X_i$.

由于 $\ln X_i$ 独立同分布, 且

$$E\ln X_i=\int_0^1\ln x dx=x\ln x|_0^1-\int_0^1 x\frac{1}{x}dx=-1.$$

故由 Khinchin 大数定律,

$$\ln Z_n\xrightarrow{P}-1.$$

再利用连续映射定理,

$$Z_n\xrightarrow{P}e^{-1}.$$

这次补充材料参考宗语庆同学 (USTC) 的讲义, 基本超过同学们的学习范围, 不作为期末考试考核要求, 仅作为课外参考资料.

概率论第 8-10 次习题课

宗语轩

2022 秋, USTC

1 随机变量列的收敛与极限定理

注. 本章节中任何引理, 定理, 命题 (不包括推论和例题等) 均可以在考试中直接使用, 除非题目已做说明及题目特别要求证明他们.

1.1 基本工具与技术

命题 1.1 (矩不等式). 我们有

(1) **Hölder 不等式**: $p,q>1, \dfrac{1}{p}+\dfrac{1}{q}=1$,

$$\mathbb{E}[|XY|] \leq \left(\mathbb{E}[|X|^p]\right)^{\frac{1}{p}} \left(\mathbb{E}[|Y|^q]\right)^{\frac{1}{q}}$$

(2) **Minkowski 不等式**:

$$\left(\mathbb{E}[|X+Y|^p]\right)^{\frac{1}{p}} \leq \left(\mathbb{E}[|X|^p]\right)^{\frac{1}{p}} + \left(\mathbb{E}[|Y|^p]\right)^{\frac{1}{p}}$$

(3) **Markov 不等式**:

$$\mathbb{P}\left(|X| \geq a\right) \leq \frac{\mathbb{E}[|X|]}{a}, a>0$$

(4) **Chebyshev 不等式**:

$$\mathbb{P}\left(|X-\mathbb{E}[X]| \geq a\right) \leq \frac{1}{a^2} Var(X)$$

(5) 设 $s>0$, 若 $\mathbb{E}[|X|^s]<+\infty$, 则对 $r\in[0,s]$, 均有 $\mathbb{E}[|X|^r]<+\infty$. 因此 $X_n \xrightarrow{s} X \implies X_n \xrightarrow{r} X$.

(6) C_r **不等式**:

$$\mathbb{E}|X_1+\cdots+X_n|^r \leqslant C_{n,r}(\mathbb{E}[|X_1|^r]+\cdots\mathbb{E}[|X_n|^r]).$$

其中

$$C_{n,r} = \begin{cases} 1, & 0<r<1, \\ n^{r-1}, & r \geqslant 1. \end{cases}$$

(7) **Lyapunov 不等式**: 对 $\forall 0<r<s$, 有

$$\left(\mathbb{E}[|X|^r]\right)^{\frac{1}{r}} \leqslant \left(\mathbb{E}[|X|^s]\right)^{\frac{1}{s}}.$$

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.
发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

证明. (5) 注意到

$$|X|^r = |X|^r I_{\{|X|\leqslant 1\}} + |X|^r I_{\{|X|>1\}} \leqslant 1 + |X|^s,$$

两边取期望即得.
(6) 当 $0 < r \leqslant 1$ 时, 利用 x^r 的凹性, 有

$$|X_1 + \cdots + X_n|^r \leqslant |X_1|^r + \cdots + |X_n|^r.$$

当 $r \geqslant 1$ 时, 利用 x^r 的凸性, 有

$$\left(\frac{|X_1| + \cdots + |X_n|}{n}\right)^r \leqslant \frac{|X_1|^r + \cdots + |X_n|^r}{n}$$

再利用 $|X_1 + \cdots + X_n| \leqslant |X_1| + \cdots + |X_n|$ 即可. 综上, 有 $|X_1 + \cdots + X_n|^r \leqslant C_{n,r}(|X_1|^r + \cdots + |X_n|^r)$. 最后对两边同时取期望即得.
(7) 利用 Hölder 不等式, 并取 $Y = 1, p = \dfrac{s}{r}$, 用 $|X|^r$ 代替 $|X|$ 即可. □

注. Marvkov 不等式 (及其推广) 是在尾概率估计中一个重要的不等式, 为随机变量偏离某些值的概率给出了上界. 更一般地, 设 g 是 $\mathbb{R}$ 上的非负 Borel 可测函数, 我们有

(1) 若 g 为偶函数, 且在 $\mathbb{R}^+$ 上单调递增 (指不严格, 下同), 则对 $\forall a \geqslant 0$, 有

$$\mathbb{P}(|X| \geqslant a) = \mathbb{P}(g(|X|) \geqslant g(a)) \leqslant \frac{\mathbb{E}[g(|X|)]}{g(a)}.$$

(2) 若 g 在 $\mathbb{R}$ 上单调递增, 则对 $\forall a \geqslant 0$, 有

$$\mathbb{P}(X \geqslant a) = \mathbb{P}(g(X) \geqslant g(a)) \leqslant \frac{\mathbb{E}[g(X)]}{g(a)}.$$

下面两个简单的特例非常值得关注, 在尾概率估计中非常常见. 对 $a > 0$, 有

$$\mathbb{P}(|X| \geqslant a) \leqslant \frac{\mathbb{E}[|X|^r]}{a^r}, r > 0, \qquad \mathbb{P}(X > a) \leqslant \mathrm{e}^{-ta}\mathbb{E}[\mathrm{e}^{tX}].$$

对于前者, 越高阶矩得到有关矩的信息越多, 也在证明随机变量的收敛时放缩更易施展手脚, 减少限制. 后者估计便于利用矩母函数的性质, 由此导出一些集中不等式 (超出本课程范围).

例 1.1 (**单边 Chebyshev 不等式**)**.** 设 $Var(X) = \sigma^2 \in (0, +\infty)$, 则对 $\forall \lambda > 0$, 有

$$\mathbb{P}(X - \mathbb{E}[X] \geqslant \lambda) \leqslant \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

证明. 记 $Y = X - \mathbb{E}[X]$, 设 $u > 0$, 则有

$$\mathbb{P}(X - \mathbb{E}[X] \geqslant \lambda) = \mathbb{P}(Y + u \geqslant \lambda + u) \leqslant \mathbb{P}((Y+u)^2 \geqslant (\lambda+u)^2) \leqslant \frac{\mathbb{E}[(Y+u)^2]}{(\lambda+u)^2} = \frac{\sigma^2 + u^2}{(\lambda+u)^2} \overset{u=\frac{\sigma^2}{\lambda}}{=\!=\!=} \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

其中最后一个不等式是一个简单的函数最值问题, 最大值当 $u = \dfrac{\sigma^2}{\lambda}$ 时取等. □

注. 对称地, 有

$$\mathbb{P}(X - \mathbb{E}[X] \leqslant -\lambda) \leqslant \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

例 1.2. 设 $X_1,\cdots X_n$ 独立同分布, $\mathbb{P}(X_1=1)=\mathbb{P}(X_1=-1)=\dfrac{1}{2}$, 记 $S_n=\sum\limits_{k=1}^{n}X_i$, 证明: 对 $\forall a>0$, 均有

$$\mathbb{P}(S_n>a)\leqslant \mathrm{e}^{-\frac{a^2}{2n}}.$$

证明. 我们有

$$\mathbb{P}(S_n>a)\leqslant \mathrm{e}^{-ta}\mathbb{E}[\mathrm{e}^{tS_n}]=\mathrm{e}^{-ta}(\mathbb{E}[\mathrm{e}^{tX_1}])^n=\mathrm{e}^{-ta}\left(\frac{\mathrm{e}^t+\mathrm{e}^{-t}}{2}\right)^n\leqslant \mathrm{e}^{-ta+\frac{n}{2}t^2}=\mathrm{e}^{\frac{n}{2}(t-\frac{a}{n})^2-\frac{a^2}{2n}}\xlongequal{t=\frac{a}{n}}\mathrm{e}^{-\frac{a^2}{2n}}.$$

其中

$$\frac{\mathrm{e}^t+\mathrm{e}^{-t}}{2}=\frac{1}{2}\left(\sum_{k=0}^{+\infty}\frac{t^k}{k!}+\sum_{k=0}^{+\infty}(-1)^k\frac{t^k}{k!}\right)=\sum_{k=0}^{+\infty}\frac{t^{2k}}{(2k)!}\leqslant\sum_{k=0}^{+\infty}\frac{t^{2k}}{2^k k!}=\mathrm{e}^{\frac{t^2}{2}}.$$

□

注. $\dfrac{\mathrm{e}^t+\mathrm{e}^{-t}}{2}\leqslant \mathrm{e}^{\frac{t^2}{2}}$ 是常用的不等式放缩, 值得一记. 类似地还有 $\dfrac{\mathrm{e}^t-\mathrm{e}^{-t}}{2t}\leqslant \mathrm{e}^{\frac{t^2}{2}}$.

下面是概率测度下证明随机变量列收敛及一些放缩中常用的拆项小技巧, 请读者自行验证:

命题 1.2. 我们有

$$\mathbb{P}(|X+Y|\leqslant a+b)\geqslant \mathbb{P}(|X|\leqslant a,|Y|\leqslant b),\quad \mathbb{P}(|X+Y|\geqslant a+b)\leqslant \mathbb{P}(|X|\geqslant a)+\mathbb{P}(|Y|\geqslant b).$$

注. 独立复制的副产品: 设 X' 是 X 的独立复制 (X,X' i.i.d), 对 $\forall x,a$, 均有

$$\mathbb{P}(|X-X'|>x)=\mathbb{P}(|(X-a)-(X'-a)|>x)\leqslant \mathbb{P}(|X-a|\geqslant\frac{x}{2})+\mathbb{P}(|X'-a|\geqslant\frac{x}{2})=2\mathbb{P}(|X-a|\geqslant\frac{x}{2}).$$

回顾一下 Borel-Cantelli 引理:

定理 1.1 (**Borel-Cantelli 引理**). 记 $\{A_n,\ \text{i.o.}\}=\limsup\limits_{n\to\infty}A_n=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m=n}^{\infty}A_m$(infinitely often)

(1) 当 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$ 时有, $\mathbb{P}(A_n,\ \text{i.o.})=0$

(2) 假设 $\{A_n\}$ 相互独立, 当 $\sum\limits_{n=1}^{\infty}\mathbb{P}(A_n)=\infty$ 时有 $\mathbb{P}(A_n,\ \text{i.o.})=1$

随机变量的截尾

对于随机变量 X, 主要有以下三类结尾方法:

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

注. 截尾法常常用于证明有关随机变量列的收敛中, 运用了转化的思想. 对于随机变量列 $\{X_n\}$, 我们常常令 $M\uparrow+\infty$ 或者取 $M=n$ (更一般取 $M=k_n$, $\{k_n\}$ 递增) 再令 $n\to+\infty$ 来处理. 这样结尾的好处是截尾后的随机变量具有有界性, 同时利用了随机变量列的单调递增性, 其中 X_2,X_3 两个随机变量本身就是单调递增 (X_2 中大于 0 递增小于 0 递减) 的函数, 递增性更好. 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, 递增性或是关于 $X-X_i$ 的性质保证了目标的转化条件 (之前做过的一道作业题: $F_n(x)\to F(x)$ 逐点收敛).

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. 后续我们还会看到一些用到截尾的例子, 可见其精妙之处.

例 1.3. 设非负随机变量 X_i 独立同分布且 $\mathbb{E}[X_1]<\infty$, 令 $Y_n=X_nI_{\{|X_n|\leqslant n\}}$, 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 (见**第 5 次习题课讲义引理 2.1**)

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

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

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

□

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.

补充: 独立复制与对称化方法

这个方法在之前证明特征函数的一道例题中 (参见**第 6 次习题课讲义例 2.2**) 也出现过, 在概率中是非常重要的技术, 考虑到内容上限就不作过多补充, 之后的一些概率类课程还会专门用到这个方法, 现在只以两个例子给大家自行体会.

1 设 $f,g:\mathbb{R}\to\mathbb{R}$ 且单调递增, 证明, 对任意随机变量 X, 均有

$$\mathbb{E}[f(X)g(X)]\geqslant\mathbb{E}[f(X)]\mathbb{E}[g(X)].$$

证明. 设 X' 是 X 的独立复制, 则有

$$(f(X)-f(X'))(g(X)-g(X'))\geqslant 0.$$

因此

$$0\leqslant\mathbb{E}[(f(X)-f(X'))(g(X)-g(X'))]=2\mathbb{E}[f(X)g(X)]-2\mathbb{E}[f(X)]\mathbb{E}[g(X)].$$

□

回忆一下, 我们称随机变量 X 是对称的, 若 $X\overset{d}{=\!=}-X$.

2 设随机变量 $e_1,\cdots e_n$ 独立同分布, 且 $\mathbb{E}[e_i]=0$, 证明: 对任意 $h_1,\cdots h_n\in\mathbb{R}$, 均有

$$\mathbb{E}\left|\sum_{i=1}^n h_ie_i\right|\leqslant 2\mathbb{E}\left|\sum_{i=1}^n |h_i|e_i\right|$$

证明. 若 X_i 对称, 则

$$h_iX_i\overset{d}{=\!=}|h_i|X_i\implies(h_1X_1,\cdots,h_nX_n)\overset{d}{=\!=}(|h_1|X_1,\cdots,|h_n|X_n)\implies\sum_{i=1}^n h_iX_i\overset{d}{=\!=}\sum_{i=1}^n|h_i|X_i.$$

设 $\{e_i'\}$ 是 $\{e_i\}$ 的一个独立复制, 则 e_i-e_i' 对称. 故有

$$\begin{aligned}\mathbb{E}\left|\sum_{i=1}^n h_ie_i\right|&\leqslant\mathbb{E}\left|\sum_{i=1}^n h_i(e_i-\mathbb{E}[e_i'|e_1,\cdots,e_n])\right|=\mathbb{E}\left|\mathbb{E}(\sum_{i=1}^n h_i[(e_i-e_i')|e_1,\cdots,e_n])\right|\leqslant\mathbb{E}\left|\sum_{i=1}^n h_i(e_i-e_i')\right|\\&=\mathbb{E}\left|\sum_{i=1}^n h_i(e_i-e_i')\right|\leqslant\mathbb{E}\left|\sum_{i=1}^n|h_i|e_i\right|+\mathbb{E}\left|\sum_{i=1}^n|h_i|e_i'\right|\\&=2\mathbb{E}\left|\sum_{i=1}^n|h_i|e_i\right|.\end{aligned}$$

□

1.2 浅谈四种收敛

回顾随机变量列四种收敛的定义:

定义 1.1. 设 $X,X_1,X_2,\cdots,X_n$ 为 $(\Omega,\mathcal{F},\mathbb{P})$ 上随机变量.

(1) **几乎处处 (以概率 1 收敛)**:

$$\mathbb{P}\left(\{\omega\in\Omega:X_n(\omega)\to X(\omega),n\to\infty\}\right)=1$$

记为 $X_n\xrightarrow{\text{a.s.}}X$.

(2) **r 阶收敛** $(r>0)$:$\mathbb{E}[|X_n|^r]<\infty,\forall n$ 且

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

记为 $X_n\xrightarrow{r}X$, 特别 $r=1$ 时为 平均收敛, $r=2$ 时为 均方收敛.

(3) **依概率收敛**:

$$\forall\epsilon>0,P(|X_n-X|>\epsilon)\to 0,n\to\infty$$

记为 $X_n\xrightarrow{P}X$.

(4) **依分布收敛**:

$$F_{X_n}\xrightarrow{w}F_X$$

记为 $X_n\xrightarrow{D}X$.

注. 只有依分布收敛会 “忘记” 样本空间 (因为分布函数会 “忘记” 样本空间), 其他类型的收敛都要考虑样本空间. 因而 $X_n\to X,Y_n\to Y\nRightarrow\mathbb{P}(X=Y)=1$ 及 $X_n+Y_n\xrightarrow{\text{D}}X+Y$.

证明依概率收敛:

(1) 估计 $\mathbb{P}(|X_n - X| > \varepsilon)$, 可能要利用**命题 1.1** 中的矩不等式或其他方法.

(2) 弱大数律 (**注意条件是 i.i.d 且一阶矩存在**).

(3) 若收敛极限是常数 c, 利用 $X_n \xrightarrow{\mathrm{D}} c \Longleftrightarrow X_n \xrightarrow{\mathrm{P}} c$, 转化成证明依分布收敛.

$\vdots$

$\vdots$

例 1.4. 有一列零均值随机变量 $X_1, \cdots X_n$, $\exists c > 0$, 使得对 $\forall n \in \mathbb{N}^*, Var(X_n) < c$. 记 $S_n = \sum_{k=1}^{n} X_k$, 若

$$\mathrm{Cov}(X_i, X_j) \to 0 \quad as\ |i-j| \to +\infty,$$

证明: $\dfrac{S_n}{n} \xrightarrow{\mathrm{P}} 0$.

证明. 对 $\forall \varepsilon > 0$, 由 Chebyshev 不等式, 有

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \leqslant \frac{\mathbb{E}[S_n^2]}{\varepsilon^2 n^2} = \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n} \mathrm{Var}(X_k) + \frac{2}{\varepsilon^2 n^2}\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j) \leqslant \frac{c}{\varepsilon^2 n} + \frac{2}{\varepsilon^2 n^2}\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j).$$

对 $\forall \delta > 0$, 由题意知, $\exists N \in \mathbb{N}^*$, 使得 $|\mathrm{Cov}(X_i, X_j)| \leqslant \delta, |i-j| > N$. 又因为 $\mathrm{Cov}(X_i, X_j) < c$, 因此

$$\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j) = \sum_{\substack{1\leqslant i<j\leqslant n \\ |i-j|\leqslant N}} \mathrm{Cov}(X_i, X_j) + \sum_{\substack{1\leqslant i<j\leqslant n \\ |i-j|> N}} \mathrm{Cov}(X_i, X_j) \leqslant nNc + n^2\delta.$$

故有

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \leqslant \frac{c}{\varepsilon^2 n} + \frac{2Nc}{\varepsilon^2 n} + 2\frac{\delta}{\varepsilon^2}.$$

令 $n \to +\infty$, 再令 $\delta \to 0$, 则有 $\mathbb{P}\left(\left|\dfrac{S_n}{n}\right| > \varepsilon\right) = 0$, 即 $\dfrac{S_n}{n} \xrightarrow{\mathrm{P}} 0$.

□

证明几乎处处收敛:

(1) 利用定义转化: 存在 $\Omega_0 \subset \Omega$, 满足 $\mathbb{P}(\Omega_0) = 1$, 且对 $\forall \omega \in \Omega_0$, 有

$$X_n(\omega) \xrightarrow{n\to+\infty} X(\omega),\ n \to +\infty.$$

固定了 ω 后 $X_n(\omega), X(\omega)$ 变为常数, 于是转化成常数列极限问题: $a_n \to a$. 例如:

- $X_n \xrightarrow{\text{a.s.}} X, Y_n \xrightarrow{\text{a.s.}} Y \implies X_n + Y_n \xrightarrow{\text{a.s.}} X + Y, X_n Y_n \xrightarrow{\text{a.s.}} XY$.
- $X_n \xrightarrow{\text{a.s.}} X \implies \dfrac{1}{n}\sum_{k=1}^{n} X_n \xrightarrow{\text{a.s.}} X$. (思考: 换成依概率收敛是否成立?)
- **Cauchy 列**: $X_n \xrightarrow{\text{a.s.}} X \Longleftrightarrow \sup_{k\geqslant n}|X_k - X| \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \sup_{k\geqslant n}|X_k - X_n| \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \sup_{k,m\geqslant n}|X_k - X_m| \xrightarrow{\text{a.s.}} 0$.
- $\dfrac{X_n}{n} \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \dfrac{\max_{1\leqslant k\leqslant n}|X_k|}{n} \xrightarrow{\text{a.s.}} 0$.

(2) 运用 Borel-Cantelli 引理解决 (可能会结合**子序列**, **截尾**, 独立复制与对称化等方法):

$$\forall \varepsilon > 0, \sum_{n=1}^{+\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty \Longrightarrow \forall \varepsilon > 0, \mathbb{P}(|X_n - X| > \varepsilon, \text{i.o.}) = 0 \Longleftrightarrow X_n \xrightarrow{\text{a.s.}} X.$$

若 $\{X_n - X\}$ 相互独立, 则为等价关系. 最后进行 $\mathbb{P}(|X_n - X| > \varepsilon)$ 的估计, 可能要借助各种矩不等式或其他方法解决, 用 Markov 不等式时对矩选取合适的阶数, 例子见作业题 **Grimmett7.11.6**.

(3) 强大数律 (注意条件是 **i.i.d** 且一阶矩存在), 例子见作业题 **Grimmett7.5.1**.

(4) * 见后, 例如转化成证明有关子列依概率收敛问题 (**命题 1.4**).

⋮

⋮

例 1.5. 设 $X_1, \cdots X_n$ 独立同分布, 证明:

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \mathbb{E}|X_1| < \infty.$$

证明. 我们有

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \forall \varepsilon > 0, \mathbb{P}\left(\frac{|X_n|}{n} > \varepsilon, \text{i.o.}\right) = 0 \Longleftrightarrow \forall \varepsilon > 0, \sum_{n=1}^{+\infty} \mathbb{P}\left(\frac{|X_n|}{\varepsilon} > n\right) < \infty.$$

利用 (见**第 5 次习题课讲义引理 2.1**)

$$\sum_{m=1}^{\infty} \mathbb{P}(|X| \geq m) \leq \mathbb{E}[|X|] \leq 1 + \sum_{m=1}^{\infty} \mathbb{P}(|X| \geq m) = \sum_{m=0}^{\infty} \mathbb{P}(|X| \geq m),$$

有

$$\frac{\mathbb{E}\left[|X_n|\right]}{\varepsilon} - 1 \leqslant \sum_{n=1}^{+\infty} \mathbb{P}\left(\frac{|X_n|}{\varepsilon} > n\right) \leqslant \frac{\mathbb{E}\left[|X_n|\right]}{\varepsilon}.$$

因此

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \forall \varepsilon > 0, \sum_{n=1}^{+\infty} \mathbb{P}\left(\frac{|X_n|}{\varepsilon} > n\right) < \infty \Longleftrightarrow \mathbb{E}|X_1| < \infty.$$

□

注意到

$$X_n \xrightarrow{\text{a.s.}} X \Longleftrightarrow \forall \varepsilon > 0, \mathbb{P}(|X_n - X| > \varepsilon, \text{i.o.}) = 0 \Longleftrightarrow \mathbb{P}\left(\lim_{n\to\infty} \bigcup_{k=n}^{+\infty} \{|X_k - X| > \varepsilon\}\right) = 0$$

$$\Longleftrightarrow \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=n}^{+\infty} \{|X_k - X| > \varepsilon\}\right) = 0 \Longleftrightarrow \lim_{n\to\infty} \mathbb{P}\left(\sup_{k\geqslant n} |X_k - X| > \varepsilon\right) = 0.$$

由此推得如下命题:

命题 1.3. 我们有

(1) $X_n \xrightarrow{\text{a.s.}} X \Longleftrightarrow \sup\limits_{k\geqslant n}|X_k - X| \xrightarrow{\text{P}} 0$. 特别地, 若 $|X_n| \downarrow 0$, 则

$$X_n \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \sup_{k\geqslant n}|X_k| \xrightarrow{\text{P}} 0 \Longleftrightarrow X_n \xrightarrow{\text{P}} 0.$$

(2) 利用 Cauchy 列等价转换,

$$\begin{aligned}X_n \xrightarrow{\text{a.s.}} X &\Longleftrightarrow \sup_{k\geqslant n}|X_k - X| \xrightarrow{\text{P}} 0\\ &\Longleftrightarrow \sup_{k\geqslant n}|X_k - X_n| \xrightarrow{\text{a.s.}} 0 \quad \Longleftrightarrow \sup_{k\geqslant n}|X_k - X_n| \xrightarrow{\text{P}} 0\\ &\Longleftrightarrow \sup_{k,m\geqslant n}|X_k - X_m| \xrightarrow{\text{a.s.}} 0 \Longleftrightarrow \sup_{k,m\geqslant n}|X_k - X_m| \xrightarrow{\text{P}} 0\end{aligned}$$

问题: 对于一般情形, 能否把几乎处处收敛转化成依概率收敛问题?

引理 1.1. 我们有

$$X_n \xrightarrow{\text{P}} X \Longleftrightarrow \forall\varepsilon > 0, \sup_{k\geqslant n}\mathbb{P}(|X_k - X_n| > \varepsilon) \xrightarrow{n\to+\infty} 0,$$

注意: 区分 $\sup\limits_{k\geqslant n}\mathbb{P}(|X_k - X_n| > \varepsilon)$&$\mathbb{P}(\sup\limits_{k\geqslant n}|X_k - X_n| > \varepsilon)$, 两者不一样!

证明. $\Longrightarrow$: 对 $\forall\varepsilon > 0$, 由 $X_n \xrightarrow{\text{P}} X$ 知, 对 $\forall\delta > 0, \exists n_0 \in \mathbb{N}^*$, 当 $n \geqslant n_0$ 时, 有 $\mathbb{P}(|X_n - X| > \frac{\varepsilon}{2}) < \frac{\delta}{2}$. 故当 $k \geqslant n \geqslant n_0$ 时,

$$\mathbb{P}(|X_k - X_n| > \varepsilon) \leqslant \mathbb{P}(|X_k - X - (X_n - X)| > \varepsilon) \leqslant \mathbb{P}\left(|X_k - X| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(|X_n - X| > \frac{\varepsilon}{2}\right) < \delta.$$

$\Longleftarrow$: $\forall k \in \mathbb{N}^*$, $\exists n_k \uparrow +\infty$, 使得

$$\mathbb{P}(|X_{n_{k+1}} - X_{n_k}| > 2^{-k}) < 2^{-k} \quad \forall k \in \mathbb{N}^*$$

所以

$$\sum_{k=1}^{+\infty}\mathbb{P}(|X_{n_{k+1}} - X_{n_k}| > 2^{-k}) < 1 \implies \mathbb{P}(|X_{n_{k+1}} - X_{n_k}| > 2^{-k}, \text{i.o.}) = 0$$

记 $\Omega_0 = \left\{|X_{n_{k+1}} - X_{n_k}| > 2^{-k}, \text{i.o.}\right\}$, 则当 $\omega \notin \Omega_0$ 时,

$$\sum_{k=1}^{+\infty}|X_{n_{k+1}}(\omega) - X_{n_k}(\omega)| < \infty$$

因此 $X_{n_k}(\omega)$ 为 Cauchy 列. 故存在极限 $X(\omega)$, 满足 $X_{n_k}(\omega) \to X(\omega)$. 所以

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

因此当 $n > n_k$ 时, 利用

$$\mathbb{P}(|X_n - X| > \varepsilon) \leqslant \mathbb{P}\left(|X_{n_k} - X| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(|X_n - X_{n_k}| > \frac{\varepsilon}{2}\right)$$

可知

$$X_n \xrightarrow{\text{P}} X.$$

$\square$

命题 1.4. 我们有

(1) $X_n \xrightarrow{\mathrm{P}} X \implies$ 存在子列 $\{n_k\}, X_{n_k} \xrightarrow{\text{a.s.}} X$.

(2) $X_n \xrightarrow{\mathrm{P}} X \iff$ 对任意子列 $\{n_k\}, \exists \{n_k'\} \subset \{n_k\}, X_{n_k'} \xrightarrow{\text{a.s.}} X$.

证明. (1) $\implies$: 参考**引理 1.1** 的证明;
(2) $\implies$: 利用 (1) 及

$$a_n \to a \iff \forall \{n_k\}, \exists \{n_k'\} \subset \{n_k\}, a_{n_k'} \to a.$$

$\impliedby$: 反证. 若 $X_n \xrightarrow{\mathrm{P}} X$ 不成立, 则 $\exists \varepsilon > 0, \delta > 0, \limsup\limits_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) > \delta$, 因此

$$\exists \{n_k\}, \mathbb{P}(|X_{n_k} - X| > \varepsilon) > \delta \implies \forall \{n_k'\} \subset \{n_k\}, \mathbb{P}(|X_{n_k'} - X| > \varepsilon) > \delta.$$

矛盾. □

注. 因此, 随机变量列的几乎处处收敛给出了另一个关于子列依概率收敛的刻画条件.

证明 r 阶收敛:

(1) 利用**命题 1.1** 中的矩不等式, 如 **Hölder 不等式**、**Minkowski 不等式**、C_r **不等式**等. 看一看笔记和作业题中的几个例子.

$$X_n \xrightarrow{\mathrm{r}} X, Y_n \xrightarrow{\mathrm{r}} Y \implies X_n + Y_n \xrightarrow{\mathrm{r}} X + Y, X_nY_n \xrightarrow{\mathrm{r}} XY.$$

(2) 借助矩不等式和截尾等技术, 运用单调收敛定理、控制收敛定理、Fatou 引理等证明.
$\vdots$
$\vdots$

已学过的定理: 若 $\exists k > 0$, 使得 $\mathbb{P}(|X_n| \leqslant k) = 1, \forall n \in \mathbb{N}^*$, 则

$$X_n \xrightarrow{\mathrm{P}} X \implies X_n \xrightarrow{\mathrm{r}} X.$$

更一般地情形留给读者自行验证: 若 $|X_n| \leqslant Y$, Y 绝对可积且 $X_n \xrightarrow{\mathrm{P}} X$, 则 $X_n \xrightarrow{1} X$.

注. r 阶收敛不是本门课重点, 加之多数人没有系统学习实变的内容, 以及没有引入一致可积等概念, 因此对这门课而言一般来说不会专门在这方面出过难的题目, 后续高等概率论课程还会深入讲解这部分内容.

例 1.6. 设随机变量 $X_1, \cdots, X_n$ 独立同分布, 且服从 $[0, a]$ 上的均匀分布, 其中 $a > 0$. 记

$$M_n = \max\{X_1, \cdots, X_n\},$$

分别在 a.s., r 阶收敛的意义下证明 $M_n \to a$.

解. 显然 $M_n < a$, 对 $\forall 0 < \varepsilon < a$ 有

$$\mathbb{P}(M_n \leqslant a - \varepsilon) = \prod_{k=1}^{n} \mathbb{P}(X_k \leqslant a - \varepsilon) = \left(\frac{a-\varepsilon}{a}\right)^n.$$

因此

$$\sum_{n=1}^{+\infty}\mathbb{P}(M_n<a-\varepsilon)=\sum_{n=1}^{+\infty}\left(\frac{a-\varepsilon}{a}\right)^n<+\infty.$$

由 Borel-Cantelli 引理知 $\mathbb{P}(M_n\leqslant a-\varepsilon,\text{i.o.})=0$. 因此 $M_n\xrightarrow{\text{a.s.}}a$.
同时有

$$\begin{aligned}\mathbb{E}[|M_n-a|^r]&=\mathbb{E}[|M_n-a|^rI_{\{|M_n-a|\leqslant\varepsilon\}}]+\mathbb{E}[|M_n-a|^rI_{\{|M_n-a|>\varepsilon\}}]\\&\leqslant\varepsilon^r+a^r\mathbb{P}(|M_n-a|>\varepsilon)\\&=\varepsilon^r+a^r\left(\frac{a-\varepsilon}{a}\right)^n.\end{aligned}$$

令 $n\to+\infty$, 再令 $\varepsilon\to 0$, 则有

$$M_n\xrightarrow{r}a.$$

□

证明/利用依分布收敛:

(1) 定义: $F_{X_n}\xrightarrow{w}F_X$. 例子见作业题 **Grimmett5.12.32，5.12.39**

(2) 利用连续性定理, 用特征函数证明或者 CLT、Lindeberg-Feller CLT 等, 内容详见**第 7 次习题课讲义 2.3 节**. 例子见作业题 **Grimmett5.12.41**

(3) 弱大数律 (**注意条件是 i.i.d 且一阶矩存在**).

(4) 利用 $X_n\xrightarrow{\text{P}}X\Longrightarrow X_n\xrightarrow{\text{D}}X$.

(5) 利用依分布收敛的等价条件, 各种性质转化, 具体见后.

⋮

⋮

例 1.7. 随机变量 $X_1,X_2,\cdots X_n$ 独立同分布且 $\mathbb{E}[X_1]=0,\mathbb{E}[X_1^2]=1$, $S_n=\sum_{k=1}^{n}X_k$, 则

$$\frac{S_n}{\sqrt{n}}\xrightarrow{\text{D}}N(0,1),$$

但不存在随机变量 Z, 使得

$$\frac{S_n}{\sqrt{n}}\xrightarrow{\text{P}}Z.$$

证明. 利用 i.i.d CLT 即可推出 $\frac{S_n}{\sqrt{n}}\xrightarrow{\text{D}}N(0,1)$, 现反证法证明后一个命题: 若存在随机变量 Z, 使得 $\frac{S_n}{\sqrt{n}}\xrightarrow{\text{P}}Z$ 成立, 则由极限分布的唯一性及前一个命题知 $Z\sim N(0,1)$. 一方面,

$$\frac{S_{2n}-S_n}{\sqrt{2n}}=\frac{S_{2n}}{\sqrt{2n}}-\frac{1}{\sqrt{2}}\frac{S_n}{\sqrt{n}}\xrightarrow{\text{P}}\left(1-\frac{1}{\sqrt{2}}\right)Z.$$

另一方面,

$$\frac{S_{2n}-S_n}{\sqrt{2n}}\xrightarrow{\text{D}}\frac{S_n}{\sqrt{2n}}\xrightarrow{\text{P}}\frac{1}{\sqrt{2}}Z,$$

矛盾. 因此不存在 Z 满足 $\frac{S_n}{\sqrt{n}}\xrightarrow{\text{P}}Z$. □

另外对于连续性随机变量, 也可以通过证明密度函数列的逐点收敛性推出依分布收敛:

命题 1.5. 设 $\{X_n\}, X$ 均为连续型随机变量, 记 X_n, X 的密度函数是 $f_n(x), f(x)$, 若对 $\forall x \in \mathbb{R}$, 均有 $f_n(x) \to f(x)$, 则有 $X_n \xrightarrow{D} X$.

证明. 对 $\forall x \in \mathbb{R}$, 利用 $|X| = X^+ + X^- = 2X^+ - X$, 有

$$
\begin{aligned}
|\mathbb{P}(X_n \leqslant x) - \mathbb{P}(X \leqslant x)| = \left|\int_{-\infty}^{x} (f_n(y) - f(y))\mathrm{d}y\right| \leqslant \int_{-\infty}^{+\infty} |f_n(y) - f(y)|\mathrm{d}y = 2\int_{-\infty}^{+\infty} (f(y) - f_n(y))^+\mathrm{d}y \\
\leqslant 2\int_{-\infty}^{+\infty} f(y)\mathrm{d}y < +\infty.
\end{aligned}
$$

利用控制收敛定理, 有

$$
\lim_{n\to+\infty} \int_{-\infty}^{+\infty} |f_n(y)-f(y)|\mathrm{d}y = 2\lim_{n\to+\infty}\int_{-\infty}^{+\infty} (f(y)-f_n(y))^+\mathrm{d}y = 2\int_{-\infty}^{+\infty} \lim_{n\to+\infty}(f(y)-f_n(y))^+\mathrm{d}y = 0.
$$

进而

$$
\lim_{n\to+\infty} |\mathbb{P}(X_n \leqslant x) - \mathbb{P}(X \leqslant x)| = 0 \implies X_n \xrightarrow{D} X.
$$

□

回顾已讲过的关于依分布收敛的内容:

引理 1.2. $X_n \xrightarrow{D} X \Longleftrightarrow$ 对任意子列 $\{n_k\}$, 存在子子列 $\{n'_k\} \subset \{n_k\}$, 使得

$$
X_{n'_k} \xrightarrow{D} X.
$$

定理 1.2 (**Skorokhod 表示定理**)**.** 设 $X_n \xrightarrow{D} X$, 则存在 $(\Omega, \mathcal{F}, \mathbb{P})$ 及其上的 Y_n, Y 满足:

(1) Y_n 与 Y 同分布, Y 与 X 同分布,

(2) $Y_n \xrightarrow{\text{a.s.}} Y$.

定理 1.3. $X_n \xrightarrow{D} X \Longleftrightarrow \forall g \in \mathrm{C_b}(\mathbb{R})$(有界连续函数), 有 $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)], n \to \infty$.

下面看一个和定理 1.3 有关的例子:

例 1.8. 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. 利用定理 1.3, 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

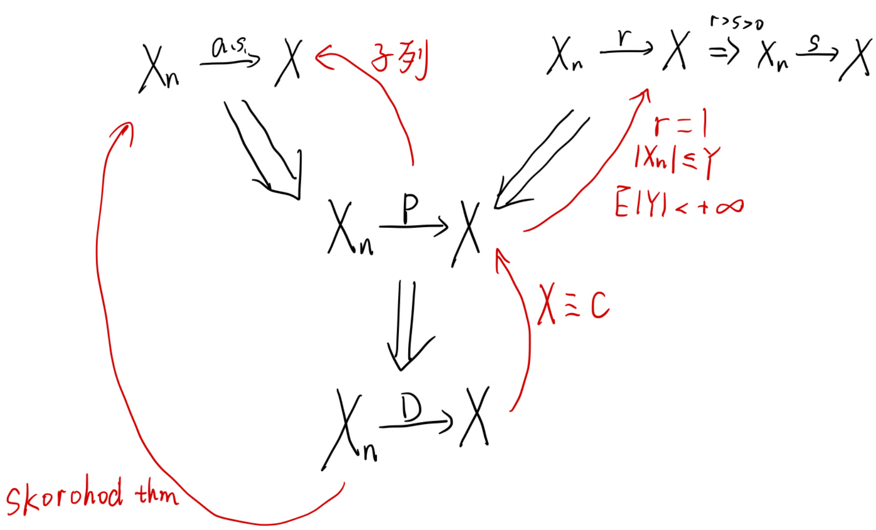
$$\lim_{n \to +\infty} \mathbb{E}[|X_n|^s] = \mathbb{E}[X^s].$$

□

注. 这里的截尾处理非常精妙, 自行体会.

1.3 结论拾零及应用

本节主要以例子来呈现. 先回顾一下各收敛间的关系:



例 1.9. 设 $f : \mathbb{R} \to \mathbb{R}$ 是连续函数, 证明:

(1) $X_n \xrightarrow{\text{a.s.}} X \implies f(X_n) \xrightarrow{\text{a.s.}} f(X)$.

(2) $X_n \xrightarrow{\text{P}} X \implies f(X_n) \xrightarrow{\text{P}} f(X)$.

(3) $X_n \xrightarrow{\text{D}} X \implies f(X_n) \xrightarrow{\text{D}} f(X)$.

证明. (1) $\exists \Omega_0 \subset \Omega$, 使得 $\mathbb{P}(\Omega_0) = 1$ 且对 $\forall \omega \in \Omega_0$, 均有 $X_n(\omega) \to X(\omega)$, 由 f 的连续性知 $f(X_n(\omega)) \to f(X(\omega))$, 因此 $f(X_n) \xrightarrow{\text{a.s.}} f(X)$.

(2) 反证: 若命题不成立, 则存在子列 $\{n_k\}$, $\varepsilon, \delta > 0$, 满足

$$\mathbb{P}(|f(X_{n_k}) - f(X)| > \varepsilon) > \delta$$