高等概率论

菠萝包萝卜

2025 年 11 月 28 日

参考书:《Probability: Theory and Examples》-Rick Durette,《概率论》-应坚刚，何萍

1 初等概率论

2 概率空间和随机变量的分布

3 独立性

4 积分论

以上部分已在另一个笔记里面有所展示

5 期望

Definition 5.1. 设概率测度 P，如果 $X \geq 0$ 是一个定义在 $(\Omega, \mathcal{F}, P)$ 上的随机变量，我们称

$$EX = \int XdP$$

是 X 的期望，但是它有可能是 ∞ 。为了方便讨论，我们考虑正负部定义

$$x^+ = \max\{x, 0\}, x^- = \max\{-x, 0\}$$

我们说 EX 存在，如果

$$EX = EX^+ - EX^-$$

是有限的；或者等价的 $EX^+ < \infty, EX^- < \infty$ 同时成立。

Theorem 5.1. *如果 $X, Y \geq 0$ 或者 $E|X|, E|Y| < \infty$, 那么*
(a) $E(X+Y) = EX + EY$
(b) $E(aX+b) = aE(X) + b$ 对任意的实数 a, b 成立
(c) 如果 $X \geq Y$ a.e. 那么 $EX \geq EY$

5.1 不等式

Theorem 5.2. *Jensen's-Inequality*：假设 ϕ 是下凸的，也就是说

$$\lambda\phi(x) + (1-\lambda)\phi(y) \geq \phi(\lambda x + (1-\lambda)y)$$

对于任意 $\lambda \in (0,1)$ 和 $x, y \in \mathbb{R}$ 那么有

$$E(\phi(X)) \geq \phi(EX)$$

当上面所有期望都存在的时候不等式总是成立的。

Proof. 定理的证明略去，因为在一般测度的结论下几乎是显然的。 □

Theorem 5.3. *Hölder-Inequality:* 如果 $p, q \in [1,\infty]^2$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$ 那么

$$E|XY| \leq ||X||_p ||Y||_q$$

这里

$$||X||_r = (E|X|^r)^{\frac{1}{r}}; ||X||_\infty = \inf\{M : P(|X| > M) = 0\}$$

Proof. 定理的证明略去，在一般测度的结论下几乎是显然的。 □

为了陈述我们下来的结论，我们需要一些记号；下面仅考虑在 $A \subset \Omega$ 上的积分，我们记

$$E(X; A) = \int_A X dP$$

Theorem 5.4. *Chebyshev's-inequality*：假设 $\phi : \mathbb{R} \to \mathbb{R}$ 有 $\phi \geq 0$ 令 $A \in \mathcal{R}$ 并且令 $i_A = \inf\{\phi(y) : y \in A\}$，那么

$$i_A P(X \in A) \leq E(\phi(X); X \in A) \leq E\phi(X)$$

Proof. $E(\phi(X); X \in A) = \int_A \phi(X)dP \geq i_A P(X \in A)$，另一边在非负函数下是显然的。 □

5.2 从积分到极限（极限理论）

Theorem 5.5. *Fatou's lemma*：如果 $X_n \geq 0$ 那么

$$\liminf_n EX_n \geq E(\liminf_n X_n)$$

Theorem 5.6. 单调收敛定理：如果 $0 \leq X_n \uparrow X$ 那么

$$EX_n \uparrow EX$$

Theorem 5.7. 控制收敛定理：如果 $X_n \to X$ *a.s.*，$|X_n| \leq Y$ 对于任意 $n \in \mathbb{N}^+$ 成立，并且 Y 是可积的，那么

$$EX_n \to EX$$

Remark. 控制收敛可以退化到有界收敛定理

Theorem 5.8. 假设 $X_n \to X$ $a.s.$ 令 g, h 为连续函数，并且满足
(i) $g \geq 0, g(x) \to \infty$ 当 $|x| \to \infty$
(ii) $\frac{|h(x)|}{g(x)} \to 0$ 当 $|x| \to \infty$
(iii) $Eg(X_n) \leq K < \infty$ 对于任意 $n \in \mathbb{N}^+$ 成立
那么 $Eh(X_n) \to Eh(X)$

我们先解释一下几个条件，首先我们要的是 X_n 以概率 1 收敛到随机变量 X ，然后 g 是个正的连续函数，并且在无穷远处发散到无穷，而 $|h(x)|$ 在无穷远处的阶数是比 g 要小的，如果 $g(X_n)$ 的期望是一致有界的，那么 $h(X_n)$ 的期望也是存在的，其实就是一个控制收敛的极限形式。这是进入极限理论的第一个新的定理，不同于前面的定理，我们会给出证明过程 (课上没给 QWQ)。

Proof. □

5.3 搞积

Theorem 5.9. 令 X 是一个来自 $(S, \mathcal{S})$ 的随机元素，并且有分布 μ,*i.e.*，$\mu(A) = P(X \in A) = P(\{\omega : X(\omega) \in A\})$ ；如果 f 是一个 $(S, \mathcal{S}) \to (\mathbb{R}, \mathcal{R})$ 的可测函数，并且 $f \geq 0$ 或者 $E|f(x)| < \infty$ 那么

$$Ef(X) = \int_S f(y)\mu(dy)$$

Remark. 积分有如下的等式

$$\int_\Omega f(h(\omega))dP = \int_S f(y)d(P \circ h^{-1})$$

Proof. 证明课程略去了, 有空补上 □

5.4 生成测度，Fubini 定理

令 $(X, \mathcal{A}, \mu_1), (Y, \mathcal{B}, \mu_2)$ 是两个 $\sigma-$ 代数空间，令

$$\Omega = X \times Y = \{(x, y) : x \in X, y \in Y\}$$
$$\mathcal{S} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$$

$\mathcal{S}$ 里面的集合称作矩形；我们可以看出 $\mathcal{S}$ 是一个半代数

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$
$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c)$$

称 $\mathcal{F} = \mathcal{A} \times \mathcal{B}$ 是由 $\mathcal{S}$ 生成的 $\sigma-$ 代数

Theorem 5.10. 在 $\mathcal{F}$ 有一个独立的测度 μ，如果

$$\mu(A \times B) = \mu_1(A)\mu_2(B)$$

Remark. μ 通常记作 $\mu_1 \times \mu_2$

Proof. 好像课上也不讲 □

Theorem 5.11. *Fubini's theorem*: 如果 $f \geq 0$ 或者 $\int |f| d\mu < \infty$ 那么

$$\int_X \int_Y f(x,y)\mu_2(dy)\mu_1(dx) = \int_{X\times Y} f d\mu = \int_Y \int_X f(x,y)\mu_1(dx)\mu_2(dy)$$

6 独立性与期望

Theorem 6.1. 设 $X_1 \cdots X_n$ 是独立随机变量，并且 X_i 有分布 μ_i 那么 $X_1 \cdots X_n$ 有分布

$$\mu_1 \times \cdots \times \mu_n$$

Proof. 课上没给证明 □

Theorem 6.2. 假设 X 和 Y 是独立的，并且有分布 μ, ν 如果 h 是 $\mathbb{R}^2 \to \mathbb{R}$ 的一个可测函数，其中 $h \geq 0$ 或者 $E|h(X,Y)| < \infty$ 那么

$$Eh(X,Y) = \iint h(x,y)\mu(dx)\nu(dx)$$

特别地，如果 $h(x,y) = f(x)g(y)$ 是 $\mathbb{R} \to \mathbb{R}$ 的可测函数，并且 $f,g \geq 0$ 或者 $E|f(X)|, E|g(Y)| < \infty$ 那么

$$Ef(X)g(Y) = Ef(X) \cdot Eg(Y)$$

Proof. 当 $f, g \geq 0$ 由 Fubini 定理, 就有

$$Ef(X)g(Y) = \iint f(x)g(y)\mu(dx)\nu(dy) = \int g(y) \int f(x)\mu(dx)\nu(dy) = Ef(X)Eg(Y)$$

如果是 $E|f(X)|, E|g(Y)| < \infty$ 直接 Fubini 就行 □

Theorem 6.3. $X_1, \ldots, X_n$ 是相互独立的，$X_i \geq 0 \forall i$ 或者 $E|X_i| < \infty$，我们有

$$E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} EX_i$$

Proof. 证明从略. □

Definition 6.1. 称两个随机变量是不相关的，如果满足 $EX^2, EY^2 < \infty$，并且有

$$EXY = EX \cdot EY$$

7 大数定律

7.1 弱大数律

Definition 7.1. 依概率收敛: 对于任意的 $\epsilon > 0$ 满足

$$\lim_{n\to\infty} P(|Y_n - Y| \geq \epsilon) = 0$$

称随机变量 Y_n 依概率收敛到 Y，记作

$$Y_n \xrightarrow{p} Y$$

Theorem 7.1. 设随机序列 $\{X_n\}_{n=1}^{\infty}$ 且 $EX_i^2 < \infty$ 是互不相关的，那么有

$$D\left(X_1 + \cdots X_n\right) = \sum_{i=1}^{n} DX_i$$

Lemma 7.2. 如果 $r > 0$ 并且 $E|Z_n|^r \to 0$ 那么 $Z_n \xrightarrow{p} 0$

Remark. 这里称 Z_n 是 r 阶平均收敛于 0。所谓 r 阶平均收敛，就是对于 $r > 0$，以及随机序列 $\{X_n\}$ 以及一个随机变量 X ，满足

$$E|X_n - X|^r \to 0$$

就是

$$X_n \xrightarrow{L_r} X$$

Proof. 只需要知道 Markov 不等式, 我们就有对任意 $\epsilon > 0$

$$0 \leq P(|Z_n - 0| \geq \epsilon) \leq \frac{E|Z_n|^r}{\epsilon^r} \to 0$$

两边夹，证毕. □

Theorem 7.3. 设 $X_1 \cdots X_n$ 互不相关，并且 $EX_i = \mu, DX_i \leq c < \infty, S_n = \sum_{i=1}^{n} X_i$ 那么

$$\frac{S_n}{n} \xrightarrow{p} \mu$$

Proof. 证明 2 阶平均收敛就行了，由于

$$0 \leq P(|\frac{S_n}{n} - \mu| \geq \epsilon) \leq \frac{D\frac{S_n}{n}}{\epsilon^2} \leq \frac{nc}{n^2\epsilon^2} = \frac{c}{n\epsilon^2} \to 0$$

证毕. □

Remark. 把互不相关换成独立同分布，并且二阶矩存在且有限，那么

$$\frac{S_n}{n} \xrightarrow{p} EX_i$$

Theorem 7.4. $\mu_n = ES_n, \sigma_n^2 = DS_n$ 如果

$$\frac{\sigma_n^2}{b_n^2} \to 0$$

那么

$$\frac{S_n - \mu_n}{b_n} \xrightarrow{p} 0$$

Proof. 还是直接去算二阶矩

$$E(\frac{S_n - \mu_n}{b_n})^2 = \frac{DS_n}{b_n^2} \to 0$$

于是 2 阶平均收敛 □

7.2 三角序列

考虑这样的随机序列 $X_{n,k}, 1 \le k \le n$ 什么意思呢，我们按下面的形式写出来

$$\begin{aligned}
&X_{1,1}\\
&X_{2,1}\ X_{2,2}\\
&X_{3,1}\ X_{3,2}\ X_{3,3}\\
&\cdots\\
&X_{n,1}\cdots\cdots\cdots X_{n,n}\\
&\cdots
\end{aligned}$$

称之为三角序列；通常我们假设行与行之间的元素是独立的，我们令

$$S_n = \sum_{j=1}^{n} X_{n,j}$$

另外再定义截断随机变量 $\bar{X} = X\mathbf{1}_{\{|X| \le M\}}$

Theorem 7.5. *截断的三级数定理: 对于每个* n*，三角序列* $\{X_{n,k}\}$ *的各个元素都是相互独立的，且* $b_n > 0, b_n \to \infty$*，在上述定义下设*

$$\bar{X}_{n,k} = X_{n,k}\mathbf{1}_{\{|X_{n,k}| \le b_n\}}$$

如果

(i)$\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0$

(ii)$\frac{\sum_{k=1}^{n} E\bar{X}_{n,k}^2}{b_n^2} \to 0$

设 $a_n = \sum_{k=1}^{n} E\bar{X}_{n,k}$ *那么*

$$\frac{S_n - a_n}{b_n} \xrightarrow{p} 0$$

Proof. 设 $\bar{S}_n = \sum_{k=1}^{n} \bar{X}_{n,k}$ 那么

$$P(|\frac{S_n - a_n}{b_n}| > \epsilon) \le P(S_n \ne \bar{S}_n) + P(|\frac{\bar{S}_n - a_n}{b_n}| > \epsilon)$$

然后分段分析

$$P(S_n \neq \bar{S}_n) \leq P(\bigcup_{k=1}^{n}\{\bar{X}_{n,k} \neq X_{n,k}\}) \leq \sum_{k=1}^{n} P(|\bar{X}_{n,k}| > b_n) \to 0$$

另一方面

$$P(|\frac{\bar{S}_n - a_n}{b_n}| > \epsilon) \leq \frac{D\bar{S}_n}{b_n^2\epsilon^2} \leq \frac{\sum_{k=1}^{n} E\bar{X}_{n,k}^2}{b_n^2\epsilon^2} \to 0$$

接下来用下面的引理 7.5 □

Lemma 7.6. 如果 $Y \geq 0, p > 0$ 那么

$$EY^p = \int_0^\infty py^{p-1}P(Y > y)dy$$

Proof. 做换序就行

$$\begin{aligned} EY^p &= \int_0^\infty y^p dF(y) = \int_0^\infty \int_0^y pt^{p-1}dtdF(y) \\ &= \int_0^\infty pt^{p-1}dt \int_t^\infty dF(y) \\ &= \int_0^\infty pt^{p-1}P(Y > t)dt \end{aligned}$$

□

Remark. 通过引理又可以得到一个结论

$$xP(|X| > x) \to 0 \Rightarrow E|X|^{1-\epsilon} < \infty$$

对任意的 $\epsilon \in (0,1)$ 成立.

事实上由于 $xP(X > x)$ 收敛，设 $Y = |X|$，对于存在一个 $B > 0$，使得 $P(Y > y) \leq \frac{1}{y}$ 对 $y \geq B$ 总是成立的，然后分段估计

$$\begin{aligned} E|X|^{1-\epsilon} &= \int_0^B (1-\epsilon)\frac{1}{y^\epsilon}P(Y > y)dy + \int_B^\infty (1-\epsilon)\frac{1}{y^\epsilon}P(Y > y)dy \\ &\leq c\int_0^B \frac{1}{y^\epsilon}dy + c\int_B^\infty y^{-1-\epsilon}dy < \infty \end{aligned}$$

Theorem 7.7. 设 $S_n = X_1 + \cdots + X_n, \mu_n = E(X_1\mathbf{1}_{\{|X_1| \leq n\}}), xP(|X_1| > x) \to 0$ 那么

$$\frac{S_n}{n} - \mu_n \xrightarrow{p} 0$$

其中 X_i 是 *i.i.d* 随机序列

Proof. 这是中国人民大学统计学学硕真题，是一个老生常谈的问题，我们就不做展开证明了，思路和 7.5 如出一辙. □

Theorem 7.8. 设 $X_1 \cdots X_n$ 是 *i.i.d.* 随机序列，且有有限期望，设 $S_n = X_1 + \cdots + X_n$ 设 $\mu = EX_1$ 那么 $\frac{S_n}{n} \xrightarrow{p} \mu$

Proof. 这个条件比 7.7 强，几乎是显然的 □

7.3 博雷尔-康特立引理

对于 Ω 的一系列子集 A_n ，我们定义上下限集

$$\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n = \limsup A_n = \lim_{m\to\infty}\bigcup_{n=m}^{\infty} A_n = \{\omega : 在无穷个A_n中出现\}$$
$$\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} A_n = \liminf A_n = \lim_{m\to\infty}\bigcap_{n=m}^{\infty} A_n = \{\omega : 仅在有限个A_n中不出现\}$$

容易证明以下事实

$$\limsup_{n\to\infty} \mathbf{1}_{A_n} = \mathbf{1}_{\limsup_{n\to\infty} A_n}, \liminf_{n\to\infty} \mathbf{1}_{A_n} = \mathbf{1}_{\liminf_{n\to\infty} A_n}$$

事实上，可以看出

$$\limsup_{n\to\infty} \mathbf{1}_{A_n} = \lim_{m\to\infty}\sup_{n\geq m} \mathbf{1}_{A_n} = \inf_m \sup_{n\geq m} \mathbf{1}_{A_n}$$

然后集列

$$\limsup_n A_n = \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n$$

于是

$$\begin{aligned}
\mathbf{1}_{\limsup A_n}(\omega) = 1 &\Leftrightarrow \omega \in \limsup A_n \\
&\Leftrightarrow \forall n \in \mathbb{N}^+, \omega \in \bigcup_{m=n}^{\infty} A_m \\
&\Leftrightarrow \forall n \in \mathbb{N}^+, \exists m \geq n, s.t. \omega \in A_m \\
&\Leftrightarrow \forall n \in \mathbb{N}^+, \exists m \geq n, s.t. \mathbf{1}_{A_m} = 1 \\
&\Leftrightarrow \forall n \in \mathbb{N}^+, \sup_{m\geq n} \mathbf{1}_{A_m} = 1 \\
&\Leftrightarrow \limsup_n \mathbf{1}_{A_n} = 1
\end{aligned}$$

另一个等式是同理的。

Theorem 7.9. *Borel-Contelli*: 如果 $\sum_{n=1}^{\infty} P(A_n) < \infty$ 那么

$$P(\limsup_n A_n) = 0$$

Proof. 设 $N = \sum_k \mathbf{1}_{A_k}$ 是一个事件发事件发生的次数。由 Fubini 定理

$$EN = \sum_k P(A_k) < \infty$$

所以 $N < \infty$ 是对于 Ω 几乎必然成立的，再由

$$P(\limsup_n A_n) = E\mathbf{1}_{\limsup A_n} = E\limsup \mathbf{1}_{A_n} = \lim_m E\sup_{n\geq m} \mathbf{1}_{A_n} \leq \lim_m E\sum_{n=m} \mathbf{1}_{A_n} = \lim_m \sum_{n=m} P(A_n) < \infty$$

即得. □

Theorem 7.10. y_n 是一个序列，如果它的任意子列 $y_{n(m)}$ 有一个子列 $y_{n(m_k)}$ 收敛到 y 那么

$$y_n \to y$$

Proof. 如果 $y_n \nrightarrow y$ 那么就存在一个含 y 的开集 G 和一个子列 $y_{n(m)}$ ，使得 $y_{n(m)} \notin G$ 对于任意的 m 成立，但是显然没有 $y_{n(m)}$ 的子列能在 G 内，矛盾！ □

Theorem 7.11. 对于随机序列 $\{\xi_n\}, \xi_n \xrightarrow{a.s.} \xi \Rightarrow \xi_n \xrightarrow{p} \xi$

Proof. 几乎必然收敛说明不收敛点集为零测集

$$\xi_n \xrightarrow{a.s.} \xi \Leftrightarrow P(\xi_n \nrightarrow \xi) = 0 \Leftrightarrow P(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} |\xi_n - \xi| \geq \epsilon) = 0$$

利用概率测度连续性又是

$$P(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} |\xi_n - \xi| \geq \epsilon) = 0 \Leftrightarrow \lim_{m \to \infty} P(\bigcup_{n=m}^{\infty} |\xi_n - \xi| \geq \epsilon) = 0$$

值得注意的是

$$0 \leq P(|\xi_n - \xi| \geq \epsilon) \leq P(\bigcup_{m=n}^{\infty} |\xi_m - \xi| \geq \epsilon) \to 0$$

夹逼，显然。 □

Theorem 7.12. $\xi_n \xrightarrow{p} \xi \Leftrightarrow \forall \xi_n$的子列$\xi_{n_k}$，存在一个 ξ_{n_k} 的子列几乎必然收敛于 ξ

Proof. 因为 $X_n \xrightarrow{p} X$ 我们可以找到一个关于 k 递降滴 $\epsilon(k)$，使得

$$P(|X_{n(m_k)} - X| > \epsilon_k) \leq 2^{-k}$$

因此

$$\sum_{k=1}^{\infty} P(|X_{n(m_k)} - X| > \epsilon_k) < \infty$$

由 B-C 引理就有 $P(|X_{n(m_k)} - X| > \epsilon_k,\ i.o.) = 0, i.e.\ X_{n(m_k)} \to X\ a.s.$

另一边，注意到如果对于每个子列 $X_{n(m)}$ 都有一个它的子列 $X_{n(m_k)}$ 几乎必然收敛到 X ，那么我们可以通过设 $y_n = P(|X_n - X| > \delta), \forall \delta > 0$ 然后利用定理 7.10 得到结论。 □

没有独立性的情况下，B-C 引理的逆命题是错的；我们设 $\Omega = (0,1), \mathcal{F} = \mathcal{B}, P$ 是一个勒贝格测度，$A_n = (0, a_n), a_n \to 0$ 那么 $\limsup A_n = \emptyset \Rightarrow P(A_n\ i.o.) = 0$，但是 $a_n \geq \frac{1}{n}$ 的时候由 p 级数的性质知道 $\sum P(A_n) = \infty$ 。

Theorem 7.13. 独立事件 A_n 满足 $\sum P(A_n) = \infty \Rightarrow P(A_n\ i.o) = 1$

Proof. 注意到 $1 - x \leq e^{-x}$ 以及余项和和部分和的关系就可以证明，证明方法在考研阶段已经熟记. □

Theorem 7.14. 如果 $X_1 \cdots X_n$ 是独立同分布序列，$E|X_i| = \infty$，那么 $P(|X_n| \geq n\ i.o.) = 1$，所以如果 $S_n = X_1 + ... + X_n$ 那么 $P(\lim S_n/n\text{存在且有限}) = 0$

Proof. 一方面

$$E|X_1| = \int_0^\infty P(|X_1| > x)dx \leq \sum_{n=0}^\infty P(|X_1| > n)$$

由于 $E|X_1|$ 发散，以及序列独立同分布，我们就可以由 B-C 引理得到 $P(|X_n| \geq n\ i.o.) = 1$。

对于后半段，注意到

$$\frac{S_n}{n} - \frac{S_{n+1}}{n+1} = \frac{S_n}{n(n+1)} - \frac{X_{n+1}}{n+1}$$

然后在集合

$$C := \{\omega : \lim S_n/n\text{存在且有限}\}$$

里面，$\frac{S_n}{n(n+1)} \to 0$ 于是在

$$C \cap \{\omega : |X_n| \geq n\ i.o.\}$$

里面就有

$$|\frac{S_n}{n} - \frac{S_{n+1}}{n+1}| > \frac{2}{3} \quad i.o.$$

这与事实 $\omega \in C$ 矛盾，因此 $\{\omega : |X_n| \geq n\ i.o.\} \cap C = \emptyset$ ，并且由于 $P(|X_n| \geq n\ i.o.) = 1$ 就有 $P(C) = 0$. □

Theorem 7.15. 如果 $A_1 \cdots A_n$ 是两两独立并且 $\sum_{n=1}^\infty P(A_n) = \infty$ 那么

$$\frac{\sum_{m=1}^n \mathbf{1}_{A_m}}{\sum_{m=1}^n P(A_m)} \to 1 \quad a.s.$$

Proof. 这是人大学硕真题。令 $X_n = \mathbf{1}_{A_n}, S_n = X_1 + ... + X_n$ 那么

$$DS_n = \sum_{k=1}^n DX_k \leq \sum_{k=1}^n EX_k^2 = ES_n$$

然后由

$$P(|S_n - ES_n| > \delta ES_n) \leq \frac{DS_n}{\delta^2 (ES_n)^2} \leq \frac{1}{\delta^2 ES_n} \to 0$$

目前说明了

$$\frac{S_n}{ES_n} \xrightarrow{p} 1$$

既然有了依概率收敛，那么是存在一个 S_n 的子列几乎必然收敛的，问题在于如何扩大到原列；令 $n_k = \inf\{n : ES_n > k^2\}, T_k = S_{n_k}$ 并且注意到 $EX_m \leq 1 \Rightarrow k^2 \leq ET_k \leq k^2 + 1$，我们用 n_k 替换之前不等式的 n，并且用 $ET_k \geq k^2$

$$P(|T_k - ET_k| > \delta ET_k) \leq \frac{1}{\delta^2 k^2}$$

所以 $\sum_{k=1}^\infty P(|T_k - ET_k| > \delta ET_k) < \infty$ 然后由 B-C 引理 $P(|T_k - ET_k| > \delta ET_k\ i.o.) = 0$，由于 δ 是任取的，这就有了 $T_k/ET_k \to 1\ a.s.$ 取一个 ω 使得 $T_k(\omega)/ET_k \to 1$ 并且观察到如果 $n_k \leq n < n_{k+1}$ 则

$$\frac{T_k(\omega)}{ET_{k+1}} \leq \frac{S_n(\omega)}{ES_n} \leq \frac{T_{k+1}(\omega)}{ET_k}$$

改写一下

$$\frac{ET_k}{ET_{k+1}}\frac{T_k(\omega)}{ET_k} \le \frac{S_n(\omega)}{ES_n} \le \frac{T_{k+1}(\omega)}{ET_{k+1}}\frac{ET_{k+1}}{ET_k}$$

夹一下就是

$$\frac{T_k}{ET_k} \le \frac{S_n}{ES_n} \le \frac{T_{k+1}}{ET_{k+1}}$$

这是在除了一个零测集上成立的，从而

$$\frac{S_n}{ES_n} \xrightarrow{a.s.} 1$$

□

8 强大数定律

Theorem 8.1. 对于独立同分布的 X_i，如果 $EX_i=\mu, EX_i^4<\infty$，$S_n=X_1+\cdots+X_n$ 那么

$$\frac{S_n}{n}\xrightarrow{a.s.}\mu$$

Proof. 不失一般性，我们假设 $\mu=0$，否则做中心化，由幂平均不等式 4 阶矩仍然有限，于是

$$ES_n^4=E(\sum_{i=1}^n X_i)^4=E\sum_{i,j,k,l}X_iX_jX_kX_l$$

既然 $E(X_i^2X_j)=E(X_i^2X_jX_k)=E(X_iX_jX_kX_l)=0$，于是

$$ES_n^4=nEX_1^4+3(n^2-n)[E(X_1^2)]^2\le Cn^2$$

这样就有

$$P(|S_n|>n\epsilon)\le\frac{ES_n^4}{n^4\epsilon^4}\le\frac{C}{n^2}$$

再代入 B-C 引理即证。 □

Theorem 8.2. 如果 X_i 两两独立同分布，且 $E|X_i|<\infty, EX_i=\mu, S_n=X_1+\cdots+X_n$ 那么

$$\frac{S}{n}\xrightarrow{a.s.}\mu$$

Proof. 为了证明这个定理，需要以下几个引理。 □

Lemma 8.3. 设 $Y_k=X_k\mathbf{1}_{(|X_k|\le k)}, T_n=Y_1+\ldots+Y_n$，这样 $\frac{T_n}{n}\xrightarrow{a.s.}\mu$。通过这条引理就有 $S_n/n\xrightarrow{a.s.}\mu$（我们需要去证明 S_n/n 与 T_n/n 不相同的部分为零测集）

Proof. 由于

$$\sum_{k=1}^n P(|X_k|>k)\le\int_0^\infty P(|X_1|>t)dt=E|X_i|<\infty\Rightarrow P(X_k\neq Y_k\ i.o.)=0$$

也就是说

$$\forall\omega,\text{如果}X_i(\omega)\neq Y_i(\omega)\text{对于}i\in I\text{成立，那么}|I|<\infty\ a.s.$$
$$\Rightarrow\forall\omega,|S_n(\omega)-T_n(\omega)|\le R(\omega)<\infty\ a.s.\text{是对任意}n\in\mathbb{N}\text{成立}$$

□

Lemma 8.4. $\sum_{k=1}^\infty D(Y_k)/k^2\le 4E|X_i|<\infty$

Proof. 我们注意到，（之后补充） □

Lemma 8.5. 如果 $y\ge 0$ 那么 $2y\sum_{k>y}\frac{1}{k^1}\le 4$

Proof. 暂时从略 □

Proof. 回到 Theorem 8.2 的证明，我们叙述思路。

因为 $X_n = X_n^+ - X_n^-$ 不失一般性设 X_n 是非负随机变量（不然就拆为正负部分别讨论）

考虑一个子列 $k(n) = [\alpha^n], \alpha > 1$ ，对于任意 $\epsilon > 0$，有

$$\begin{aligned}\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \epsilon k(n)) &\leq \sum_{n=1}^{\infty} \frac{D(T_{k(n)})}{[k(n)]^2\epsilon^2} \\ &= \frac{\sum_{n=1}^{\infty} \frac{1}{k(n)^2} \sum_{m=1}^{k(n)} D(Y_m)}{\epsilon^2} \\ &= \frac{\sum_{m=1}^{\infty} D(Y_m) \sum_{n:k(n)\geq m} k(n)^{-2}}{\epsilon^2}\end{aligned}$$

由于 $k(n) = [\alpha]^n, [\alpha^n] \geq \frac{\alpha^n}{2}$，所以

$$\sum_{n:\alpha^n \geq m} [\alpha^n]^{-2} \leq 4 \sum_{n:\alpha^n \geq m} \frac{1}{\alpha^{2n}} \leq 4(1-\alpha^{-2})^{-1} m^{-2}$$

从而

$$\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \epsilon k(n)) \leq 4(1-\alpha^{-2})^{-1}\epsilon^{-2} \sum_{m=1}^{\infty} E(Y_m^2) m^{-2} < \infty$$

这便是 $\frac{T_{k(n)} - ET_{k(n)}}{k(n)} \xrightarrow{a.s.} 0$ 由有界收敛就有

$$EY_k \to EX_1 \Rightarrow ET_{k(n)}/k(n) \to EX_1$$

于是 $T_{k(n)}/k(n) \xrightarrow{a.s.} EX_1$，当 $k(n) \leq m < k(n+1)$，有

$$\frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)}$$

(这里用到了 $Y_i \geq 0$)，再回到 $k(n) = [\alpha^n]$，我们就有 $\frac{k(n+1)}{k(n)} \to \alpha$ 并且

$$\frac{EX_1}{\alpha} \leq \liminf_{n\to\infty} \frac{T_m}{m} \leq \limsup_{m\to\infty} \frac{T_m}{m} \leq \alpha EX_1$$

因为 $\alpha > 1$ 是任取的，我们让它趋于 1，那么上个式子夹逼。 □

Theorem 8.6. 设 $X_1, X_2 \dots$ 是独立同分布的且 $EX_i^+ = \infty, EX_i^- < \infty$，如果

$$S_n = X_1 + \dots + X_n$$

那么

$$S_n/n \xrightarrow{a.s.} \infty$$

Proof. 定义 $X_i^M = \min\{X_i, M\}, S_n^M = X_1^M + \dots + X_n^M$，然后去证明

$$S_n^M/n \xrightarrow[]{n\to\infty} EX_i^M \xrightarrow[]{M\to\infty} \infty$$

□

Definition 8.1. 经验分布函数：设 X_i 是独立同分布序列，分布函数为 F ，称

$$F_n(x) = \frac{\sum_{m=1}^{n} \mathbf{1}_{\{X_m \le x\}}}{n}$$

为经验分布函数。

Theorem 8.7. *Glivenko-Cantelli:* 上述定义的经验分布 $F_n \Rightarrow F$

9 随机序列的收敛性

设 $\mathcal{F}'_n = \sigma(X_n, X_{n+1}\ldots) =$ 使得从时刻n之后的所有随机变量X_n可测的最小$\sigma-$域。设 $\mathcal{T} = \cap_n \mathcal{F}'_n$ 称为尾 $\sigma-$ 域；直观地，$A \in \mathcal{T}$ 当且仅当替换有限个 X_i 是不会影响事件 A 发生的。

Example 9.1. 如果 $B_n \in \mathcal{R}$，那么 $\{\omega : X_n \in B_n \ i.o.\} \in \mathcal{T}$，如果我们设 $X_n = \mathbf{1}_{A_n}$ 并且 $B_n = \{1\}$，这个例子便是 $\{A_n \ i.o.\}$

Proof. 设 $A_n = \{\omega : X_n \in B_n\} \in \sigma(A_n) \subset \mathcal{F}'_n \Rightarrow A_n \in \mathcal{F}'_n \subset \mathcal{F}'_m$ 其中$m \leq n$，我们希望证明 $\limsup A_n \in \mathcal{T}$，令 $C_m = \bigcup_{n=m}^{\infty} A_n \in \mathcal{F}'_m \subset \mathcal{F}'_k \ \forall k \leq m \Rightarrow \limsup A_n = \bigcap_{m=1}^{\infty} C_m \in \mathcal{F}'_k \ \forall k \Rightarrow \limsup A_n \in \cap_k \mathcal{F}'_k = \mathcal{T}$ □

Example 9.2. $S_n = X_1 + \cdots + X_n$ 那么

$$\{\lim S_n\text{存在}\} \in \mathcal{T}, \{\limsup_n S_n > 0\} \in \mathcal{T}, \{\limsup_n S_n/c_n > x\} \in \mathcal{T}\text{如果}c_n \to \infty$$

Theorem 9.1. *Kolmogorov's 0-1 law:* 如果 $X_1, X_2, \ldots$ 是独立的并且 $A \in \mathcal{T}$ 那么 $P(A) = 0$ 或者 $P(A) = 1$

Proof. 我们会去证明 A 与自身独立，那么 $P(A) = P(A \cap A) = P^2(A)$ 从而这个概率为 0 或 1

先去说明 $A \in \sigma(X_1 \ldots X_k)$ 和 $B \in \sigma(X_{k+1} \ldots,)$ 独立

再去说明 $A \in \sigma(X_1 \ldots)$ 和 $B \in \mathcal{T}$ 独立

然后由 $\mathcal{T} \subset \sigma(X_1 \ldots) \Rightarrow A \in \mathcal{T}$, 这样 A 与自身独立. □

Theorem 9.2. *Kolmogorov-inequality:* 设独立随机序列 $\{X_i\}$ 满足 $EX_i = 0, DX_i < \infty, S_n = X_1 + ... + X_n$，那么

$$P(\max_{1 \leq k \leq n} |S_k| \geq x) \leq \frac{DS_n}{x^2}$$

Proof. 设 $A_k = \{|S_n| \geq x, |S_j| < x\text{当 j<k}\}$ 这样 A_k 是不交的，我们有

$$\begin{aligned} ES_n^2 \geq \sum_{k=1}^{n} \int_{A_k} S_n^2 dP &= \sum_{k=1}^{n} \int_{A_k} S_k^2 + \sum S_k(S_n - S_k) + (S_n - S_k)^2 dP \\ &\geq \sum_{k=1}^{n} \int_{A_k} S_k^2 dP + \sum_{k=1}^{n} \int 2S_k \mathbf{1}_{A_k}(S_n - S_k) dP = \sum_{k=1}^{n} x^2 P(A_k) = x^2 P(\max_{1 \leq k \leq n} |S_k| \geq x) \end{aligned}$$

□

Theorem 9.3. X_i 独立且 $EX_i = 0, \sum_{n=1}^{\infty} Var(X_n) < \infty$ 那么 $\sum_{n=1}^{\infty} X_n$ 几乎必然收敛.

Proof. 设 $S_N = \sum_{n=1}^{N} X_n$ 然后由前面的最大值不等式就有

$$P\left(\max_{M \leq m \leq N} |S_m - S_M| > \epsilon\right) \leq \frac{D(S_N - S_M)}{\epsilon^2} = \frac{\sum_{n=M+1}^{N} D(X_n)}{\epsilon^2}$$

然后再令 $N \to \infty$ 就有

$$P\left(\sup_{m\geq M}|S_m - S_M| > \epsilon\right) \leq \frac{\sum_{n=M+1}^{\infty} D(X_n)}{\epsilon^2} \to 0$$

这是因为余项和收敛于 0，如果我们令 $w_M = \sum_{m,n\geq M}|S_m - S_n|$ 那么 w_M 随着 M 的增大而减小，并且

$$P(w_M > 2\epsilon) \leq P(\sup_{m\geq M}|S_m - S_M| > \epsilon) \to 0$$

所以 $w_M \xrightarrow{a.s.} 0$ 这意味着 $S_n(\omega)$ 以概率 1 为柯西列，因此 $\lim_{n\to\infty} S_n(\omega)$ 存在。 □

Example 9.3. X_i 独立且 $P(X_n = \frac{1}{n^\alpha}) = P(X_n = \frac{-1}{n^\alpha}) = \frac{1}{2}$，那么 $EX_n = 0, D(X_n) = \frac{1}{n^{2\alpha}}$，那么只要 $\alpha > \frac{1}{2}$ 就可以得到 $\sum_n |X_n| < \infty$ 是几乎必然的，从而 $\sum_n X_n$ 几乎必然收敛。

Theorem 9.4. *Kolmogorov 三级数定理：令 $X_1 \dots$ 是独立随机变量，设 $A > 0$ 并且设 $Y_i = \mathbf{1}_{|X_i|\leq A}$，那么 $\sum_{n=1}^{\infty} X_n$ 几乎必然收敛的充要条件是以下三个条件满足*
(i)$P(|X_n| > A) < \infty$
(ii)$\sum_{n=1}^{\infty} EY_n$收敛
(iii)$\sum_{n=1}^{\infty} D(Y_n) < \infty$

Proof. 充分性：(iii) 和定理 9.3 可以知道

$$\sum_{n=1}^{\infty}(Y_n - EY_n) < \infty$$

然后 (ii) 意味着

$$\sum_{n=1}^{\infty} Y_n < \infty$$

由 (i) 和 B-C 引理就有 $P(X_n \neq Y_n \ i.o.) = 0$ 所以

$$\sum_{n=1}^{\infty} X_n < \infty \quad a.s.$$

□

Lemma 9.5. *Kroncker 引理：$a_n \uparrow \infty, \sum_n(\frac{X_n}{a_n}) < \infty$ 那么*

$$\frac{\sum_{k=1}^{n} X_n}{a_n} \to 0$$

Proof. 证明从略 □

Theorem 9.6. *X_n 独立同分布，$EX_i = \mu < \infty, S_n = X_1 + \cdots + X_n$ 那么*

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

Proof. 设 $Y_k = X_k \mathbf{1}_{(|X_k| \le k)}$ ，设 $T_n = Y_1 + \cdots + Y_n$ 通过以前证明的办法可以去说明 $\frac{T_n}{n} \xrightarrow{a.s.} \mu$，然后令 $Z_k = Y_k - EY_k$ 所以 $EZ_k = 0$ 现在 $DZ_k = DY_k \le EY_k^2$ 并且有

$$\sum_{k=1}^{\infty} \frac{DZ_k}{k^2} \le \sum_{k=1}^{\infty} \frac{EY_k^2}{k^2} < \infty$$

再用上定理 9.3 就是

$$\sum_{k=1}^{\infty} Z_k/k 几乎必然收敛$$

再由 Kroncker 引理

$$\frac{\sum_{k=1}^{n}(Y_k - EY_k)}{n} \to 0 因此 \frac{T_n}{n} - \frac{\sum_{k=1}^{n} EY_k}{n} \to 0$$

由有界收敛定理 $EY_k \to \mu$ 这样显然有

$$\frac{\sum_{k=1}^{n} EY_k}{n} \to \mu$$

因此

$$\frac{T_n}{n} \to \mu$$

□

Theorem 9.7. X_i 是独立同分布的 $EX_i = 0, EX_i^2 = \sigma^2 < \infty, S_n = X_1 + \cdots + X_n$ 那么

$$S_n/(n^{1/2}(\log n)^{1/2+\epsilon}) \xrightarrow{a.s.} \mu$$

对任意的 $\epsilon > 0$ 成立

Proof. 设 $a_n = n^{\frac{1}{2}}(\log n)^{1/2+\epsilon}$，于是

$$\sum_{n=1}^{\infty} D(X_n/a_n) = \sigma^2(\frac{1}{a_1^2} + \sum_{n=2}^{\infty} \frac{1}{(\log n)^{1/2+\epsilon}})$$

从而由定理 9.3 就有

$$\sum_{n=1}^{\infty} \frac{X_n}{a_n} < \infty \ a.s.$$

再由 Kroncker 引理就有

$$\frac{S_n}{a_b} \to 0 \ a.s.$$

□