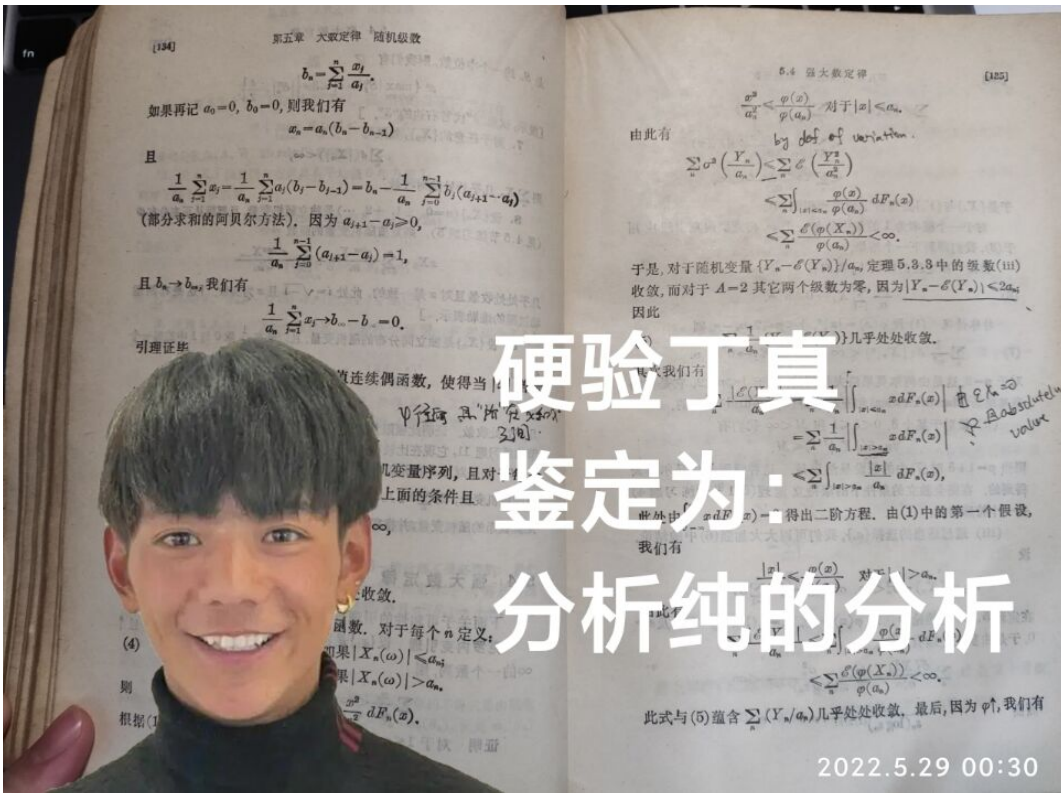
第 11 次课程在可能性理论:

董 晟 渤

Колмогорóва 强大数律



Author:
雁鸭丁真

Address:
GeodesicSeiran@gmail.com

Last TEX in 2022 年 5 月 29 日

目录

1 级数的收敛

为了让参与讨论班的同学经受一下数学中的艰难困苦, 笔者决定采用 [CKL] 中采用的 Колмогорóв 的原始证明, 以作为 "真正的精湛技巧的一个例子".

定理 1.1. *设 $\{X_n\}$ 是独立的 $r.v.$ 序列, 使得其满足 $\forall n$, 都有*

$$\mathbb{E}X_n = 0,\ \mathbb{E}X_n^2 = \sigma^2 X_n < \infty,$$

则对于 $\forall\varepsilon > 0$, 都成立不等式

$$\mathscr{P}\{\max_{1\le j\le n} |S_j| > \varepsilon\} \le \frac{\sigma^2(S_n)}{\varepsilon^2}.$$

注 1.1. 注意到如果用 $|S_n|$ 取代 $\max_j |S_n|$, 则成为 Чебышев 不等式的简单情况. 故这个命题实际上是 Чебышев 不等式的本质改进.

证明. 固定 $\varepsilon > 0$. 考虑集合

$$\Lambda = \Lambda_\varepsilon = \{\omega : \max_{1\le j\le n} |S_j| > \varepsilon\},$$

考虑其无交并, 定义函数 $\nu : \Lambda \to \mathbb{N}_+$, $\nu(\omega) = \min\{j : 1 \le j \le n, |S_j(\omega)| > \varepsilon\}$. 其定义为 "对于 $1 \le j \le n$, $|S_j(\omega)|$ 首次超过 ε 的步数 (j)". Resp., 定义

$$\Lambda_k = \{\omega : \nu(\omega) = k\},$$

即 Λ_k 是 "首次超过 ε 的步数是 k" 这一事件. 显然 $\sqcup_k \Lambda_{k=1}^n = \Lambda$. 故有

$$\begin{aligned}
\int_\Lambda S_n^2 d\mathscr{P} &= \sum_{k=1}^n \int_{\Lambda_k} S_n^2 d\mathscr{P} \\
&= \sum_{k=1}^n \int_{\Lambda_k} [S_k + (S_n - S_k)]^2 d\mathscr{P} \\
&= \sum_{k=1}^n \int_{\Lambda_k} S_k^2 + (S_n - S_k)^2 + 2S_k(S_n - S_k) d\mathscr{P}.
\end{aligned}$$

注意到根据独立变量的定义, 知道 $\chi_{\Lambda_k} S_k$ 和 $S_n - S_k$ 独立, 故

$$\begin{aligned}
\int_{\Lambda_k} S_k(S_n - S_k) &= \int_\Omega \chi_{\Lambda_k} S_k d\mathscr{P} \int_\Omega S_n - S_k d\mathscr{P} \\
&= \int_{\Lambda_k} S_k d\mathscr{P} (\sum_{j=k+1}^n \underbrace{\mathbb{E}X_j}_{=0,\,\forall j}) \\
&= 0.
\end{aligned}$$

这就得到了

$$\begin{aligned}\sigma^2(S_n) = \mathbb{E}S_n^2 &\geq \int_\Lambda S_n^2 d\mathscr{P} \\ &\geq \sum_{k=1}^n \int_{\Lambda_k} S_k^2 d\mathscr{P} \\ &\geq \underbrace{\varepsilon^2 \sum_{k=1}^n \mathscr{P}(\Lambda_k)}_{\text{by definition of } \Lambda_k} \\ &= \varepsilon^2 \mathscr{P}(\Lambda).\end{aligned}$$

这就得到了待证明的不等式. □

现在给出一个更逆天的不等式:

定理 1.2. *设 $\{X_n\}$ 是一列 r.v. , 满足 $\forall n \in \mathbb{N}_+$, 都有 $\mathbb{E}X_n < \infty$. 如果存在 A 使得 $\forall n$ 都有*

$$|X_n - \mathbb{E}X_n| \leq A < \infty,$$

则对于 $\forall \varepsilon > 0$, 都有

$$\mathscr{P}\{\max_{1\leq j\leq n} |S_j| \leq \varepsilon\} \leq \frac{(2A+4\varepsilon)^2}{\sigma^2(S_n)}.$$

证明. 记 $M_0 = \Omega$. 对于 $1 \leq k \leq n$, 记

$$\begin{aligned}M_k &= \{\omega : \max_{1\leq j\leq k} |S_j| \leq \varepsilon\}, \\ \Delta_k &= M_{k-1} - M_k,\end{aligned}$$

WLOG, 考虑 $\mathscr{P}(M_n) > 0$ 的情形, 否则原不等式显然成立.

约定记号

$$X_k' = X_k - \mathbb{E}X_k,\, S_k' = \sum_{j=1}^k X_j'.$$

且定义积分平均

$$a_k = \fint_{M_k} S_k' d\mathscr{P},$$

注意到

$$\int_{M_k} S_k' - a_k d\mathscr{P} = 0.$$

考虑积分

$$\begin{aligned}&\int_{M_{k+1}} (S_{k+1}' - a_{k+1})^2 d\mathscr{P} \\ &= \int_{M_k} (S_k' - a_k + a_k - a_{k+1} + X_{k+1}')^2 d\mathscr{P} - \int_{\Delta_{k+1}} (S_k' - a_k + a_k - a_{k+1} + X_{k+1}')^2 d\mathscr{P} \\ &=: I_1 - I_2,\end{aligned}$$

注意到

$$
\begin{aligned}
|S'_k - a_k| &= |S_k - \mathbb{E}S_k - \fint_{M_k} S_k - \mathbb{E}S_k d\mathscr{P}| \\
&= |S_k - \fint_{M_k} S_k d\mathscr{P}| \\
&\leq |S_k| + \varepsilon,
\end{aligned}
$$

$$
\begin{aligned}
|a_k - a_{k+1}| = |&\fint_{M_k} S_k d\mathscr{P} - \fint_{M_{k+1}} S_k d\mathscr{P} - \fint_{M_{k+1}} X'_{k+1} d\mathscr{P}| \\
&\leq 2\varepsilon + A.
\end{aligned}
$$

其中最后一个不等号是根据 M_k 和 X'_k 的定义. 故有

$$
\begin{aligned}
I_2 &\leq \int_{\Delta_{k+1}} (|S'_k - a_k| + |a_k - a_{k+1}| + X'_{k+1})^2 d\mathscr{P} \\
&\leq \int_{\Delta_{k+1}} (\underbrace{|S_k|}_{\leq \varepsilon} + \varepsilon + 2\varepsilon + A + A)^2 d\mathscr{P} \\
&\leq (4\varepsilon + 2A)^2 \mathscr{P}(\Delta_{k+1}).
\end{aligned}
$$

对于 I_1, 我们有

$$
\begin{aligned}
I_1 = \int_{M_k} &(S'_k - a_k)^2 + (a_k - a_{k+1})^2 + {X'_{k+1}}^2 \\
&+ 2(S'_k - a_k)(a_k - a_{k+1}) + 2(S'_k - a_k)X'_{k+1} + 2(a_k - a_{k+1})X'_{k+1} d\mathscr{P},
\end{aligned}
$$

注意到 $S'_k - a_k$ 和 X'_{k+1} 独立, 故有积分的后三项都等于 0. 从而有

$$
\begin{aligned}
I_1 &\geq \int_{M_k} (S'_k - a_k)^2 d\mathscr{P} + \int_{M_k} {X'_{k+1}}^2 d\mathscr{P} \\
&= \int_{M_k} (S'_k - a_k)^2 d\mathscr{P} + \mathscr{P}(M_k)\sigma^2(X_{k+1}).
\end{aligned}
$$

注意到等式的第二项是因为

$$
\int_{M_k} {X'_{k+1}}^2 d\mathscr{P} = \mathbb{E}\chi_{M_k} {X'_{k+1}}^2 = \mathbb{E}\chi_{M_k} \mathbb{E}{X'_{k+1}}^2 = \mathscr{P}(M_k)\sigma^2(X_{k+1}).
$$

带回原积分, 对于 $0 \leq k \leq n-1$, 得到

$$
\begin{aligned}
&\int_{M_{k+1}} (S'_{k+1} - a_{k+1})^2 d\mathscr{P} - \int_{M_k} (S'_k - a_k)^2 d\mathscr{P} \\
\geq &\mathscr{P}(M_n)\sigma^2(X_{k+1}) - (4\varepsilon + 2A)^2 \mathscr{P}(\Delta_{k+1}),
\end{aligned}
$$

此时对 k 求和, 得到

$$
\begin{aligned}
4\varepsilon^2 \mathscr{P}(M_n) &\geq \int_{M_n} (\underbrace{|S_n|}_{\leq \varepsilon, \forall \omega \in M_n} + \varepsilon)^2 d\mathscr{P} \geq \int_{M_n} (S'_n - a_n)^2 d\mathscr{P} \\
&\geq \mathscr{P}(M_n) \sum_{j=1}^{n} \sigma^2(X_j) - (4\varepsilon + 2A)^2 \mathscr{P}(\Omega \setminus M_n),
\end{aligned}
$$

移项后得到原命题成立. □

现在来证明赫赫有名的 Колмогоров 三级数定理. 这个定理同样是用标量序列去限制函数序列.

定理 1.3 (Колмогоров, 1929). *对于独立 r.v. 序列 $\{X_n\}$, 取定常数 $A>0$, 定义截尾序列*

$$Y_n(\omega)=\begin{cases}X_n(\omega), & |X_n(\omega)|\le A,\\ 0, & |X_n(\omega)|>A,\end{cases}$$

则级数 $\sum_n X_n$ a.e. 收敛, 当且仅当下述的三个数项级数都收敛:

1. $\sum_n \mathscr{P}\{|X_n|>A\}=\sum_n \mathscr{P}\{X_n\neq Y_n\}$*(即 $\{X_n\}$ 和 $\{Y_n\}$ 是等价序列),*
2. $\sum_n \mathbb{E}Y_n$,
3. $\sum_n \sigma^2(Y_n)$.

证明. "$\Longleftarrow$":

对序列 $\{Y_n-\mathbb{E}Y_n\}$ 应用第一个逆天不等式, 得到对于 $\forall m\in\mathbb{N}_+$, 都有 (注意一个简单的归一化关系)

$$\mathscr{P}\{\max_{n\le k\le n'}|\sum_{j=n}^{k}(Y_j-\mathbb{E}Y_j)|\le\frac{1}{m}\}\ge 1-m^2\sum_{j=n}^{n}\sigma^2(Y_j).$$

对 LHS 的 n,n' 分别取极限, 可以得到 $\forall m$, 有

$$\lim_{n\to\infty}\lim_{n'\to\infty}LHS=1,$$

注意到这意味着 $\sum_n Y_n-\mathbb{E}Y_n$ 的尾部 a.e. 收敛于 0, 故此级数 a.e. 收敛. 且由于 $\sum_n \mathbb{E}Y_n$ 收敛, 故 $\sum_n Y_n$ 收敛. 且由于 $\{X_n\}$ 和 $\{Y_n\}$ 等价, 故知 $\sum_n X_n$a.e. 收敛. 这就完成了充分性的证明.

"$\Longrightarrow$":

对于 $\forall A>0$, 有

$$\mathscr{P}(|X_n|>A\quad \text{i.o.})=0,$$

根据 Borel-Cantelli 引理, 知第一个级数收敛. 故 $\{X_n\}$ 和 $\{Y_n\}$ 等价, 知道 $\sum_n Y_n$a.e. 收敛. 但又因为 $|Y_n-\mathbb{E}Y_n|\le 2A$, 应用第二个逆天不等式, 得到

$$\mathscr{P}\{\max_{n\le k\le n'}|\sum_{j=n}^{k}Y_j|\le 1\}\le\frac{(4A+4)^2}{\sum_{j=n}^{n'}\sigma^2(Y_j)}.$$

这使得方差收敛, 否则 RHS 趋于 0, $\sum_n Y_n$ 的尾部不会趋于 0, 与 $\sum_n Y_n$ 收敛矛盾.

现在讨论期望. 考虑级数 $\sum_n Y_n-\mathbb{E}Y_n$, 并对常数 $2A$ 应用充分性部分. 注意到

$$\mathscr{P}\{|Y_n-\mathbb{E}Y_n|>2A\}=0,\qquad \mathbb{E}(Y_n-EY_n)=0,$$

且 $\sigma^2(Y_n)=\sigma^2(Y_n-\mathbb{E}Y_n)$(这是因为 $\mathbb{E}(Y_n-\mathbb{E}Y_n)=0$), 故知 $\sum_n Y_n-\mathbb{E}Y_n$ 是 a.e. 收敛的. 故根据等价序列, 知道 $\sum_n X_n-\mathbb{E}Y_n$ 也是 a.e. 收敛的. 而由于题设, $\sum_n X_n$ 也是 a.e. 收敛的, 这就得到了 $\sum_n \mathbb{E}Y_n$ 收敛. 必要性得到证明. □

对于独立 r.v. 序列, 有如下的 Lévy 定理改进了收敛模式之间的关系:

定理 1.4 (Lévy)**.** *对于独立 r.v. 序列 $\{X_n\}$, 则级数 $\sum_n X_n$ 依概率收敛等价于其 a.e. 收敛.*

证明. 注意到 a.e. 收敛蕴含依概率收敛, 故只需要证明反方向的箭头. 不妨假设 $\{X_n\}$ 依概率收敛, 则取定 $0 < \varepsilon < 1$, 都存在 m_0, 使得当 $n > m > m_0$ 时, 我们有

$$\mathscr{P}\{|S_{m,n}| > \varepsilon\} < \varepsilon,$$

其中 $S_{m,n} = \sum_{j=m+1}^n X_j$. 现在对于 $m < k \le n$, 考虑逆天分划

$$\sqcup_{k=m+1}^n \{\max_{m<j\le k-1} |S_{m,j}| \le 2\varepsilon, |S_{m,k}| > 2\varepsilon, |S_{k,n}| \le \varepsilon\} \subseteq \{|S_{m,n}| > \varepsilon\},$$

两边同时积分, 根据独立性得到

$$\sum_{k=m+1}^n \mathscr{P}\{\max_{m<j\le k-1} |S_{m,j}| \le 2\varepsilon, |S_{m,k}| > 2\varepsilon\} \mathscr{P}\{|S_{k,n}| \le \varepsilon\} \le \mathscr{P}\{|S_{m,n}| > \varepsilon\}.$$

注意到

$$LHS \ge \mathscr{P}\{\max_{m<j\le n} |S_{m,j}| > 2\varepsilon\} \min_{m<k\le n} \mathscr{P}\{|S_{k,n}| \le \varepsilon\}$$

这就得到了 (Ottaviani)

$$\mathscr{P}\{\max_{m<j\le n} |S_{m,j}| > 2\varepsilon\} \min_{m<k\le n} \underbrace{\mathscr{P}\{|S_{k,n}| \le \varepsilon\}}_{\ge 1-\varepsilon, \text{ by hypothesis}} \le \mathscr{P}\{|S_{m,n}| > \varepsilon\}.$$

从而对于 $m > m_0$, 则有

$$\mathscr{P}\{\max_{m<j\le n} |S_{m,j}| > 2\varepsilon\} \le \frac{1}{1-\varepsilon} \mathscr{P}\{|S_{m,n}| > \varepsilon\} < \frac{\varepsilon}{1-\varepsilon}.$$

分别令 $n \to \infty$, $m \to \infty$, $\varepsilon \to 0$, 再根据一个逆天集合论命题, 明所欲证. □

一个有趣的例子是随机选择符号的调和级数 $\sum_n \frac{X_n}{n}$, 其中 $\{X_n, n \ge 1\}$ 是以概率 $1/2$ 取 ± 1 中的每一个值的 i.i.d. r.v. 序列. 如果取 $A = 1$, 则知道三级数中的第一个为 0, 且期望为 0. 注意到方差的阶是 $1/n^2$, 故方差也收敛. 这就得到了这一随机变量和式收敛.

2 强大数定律

陈述一个技术性引理:

引理 2.1 (Kronecker)**.** 设 $\{x_k\} \subseteq \mathbb{R}$, $\{a_k\} > 0$ 且趋于正无穷. 则

$$\sum_n \frac{x_n}{a_n} < \infty \implies \frac{1}{a_n} \sum_{j=1}^n x_j \to 0.$$

现在考虑一个正值连续偶函数 φ, 其满足 $|x|$ 增加时, 有

$$\frac{\varphi(x)}{x} \nearrow, \frac{\varphi(x)}{x^2} \searrow.$$

定理 2.1. *设 $\{X_n\}$ 是一列独立 r.v. ，且对于每个 n, 有 $\mathbb{E}X_n = 0$, 又 $0 < a_n \nearrow \infty$. 如果 φ 满足上述条件, 且*

$$\sum_n \frac{\mathbb{E}\varphi(X_n))}{\varphi(a_n)} < \infty,$$

则

$$\sum_n \frac{X_n}{a_n},$$

a.e. 收敛.

证明. 记 F_n 是 X_n 的 d.f. , 对于每个 n 定义截尾变量

$$Y_n(\omega) = \begin{cases} X_n(\omega), & |X_n(\omega)| \le a_n, \\ 0, & |X_n(\omega)| > a_n, \end{cases}$$

则

$$\sum_n \mathbb{E}\frac{Y_n^2}{a_n^2} = \sum_n \int_{|x| \le a_n} \frac{x^2}{a_n^2} dF_n(x).$$

根据对 φ 的第二个假设, 对于 $|x| \le a_n$, 有

$$\frac{x^2}{a_n^2} \le \frac{\varphi(x)}{\varphi(a_n)}.$$

根据方差的定义 (第一个不等号), 有

$$\begin{aligned} \sum_n \sigma^2(\frac{Y_n}{a_n}) &\le \sum_n \mathbb{E}\frac{Y_n^2}{a_n^2} \\ &\le \sum_n \int_{|x| \le a_n} \frac{\varphi(x)}{\varphi(a_n)} dF_n(x) \\ &\le \sum_n \frac{\mathbb{E}\varphi(X_n)}{\varphi(a_n)} < \infty. \end{aligned}$$

故对于 r.v. $(Y_n - \mathbb{E}Y_n)/a_n$ 而言, 和式三级数中的方差收敛. 且如果取 $A = 2$, 则前两个级数为 0, 这是因为 $|Y_n - \mathbb{E}Y_n| \le 2a_n$. 故

$$\sum_n \frac{Y_n - \mathbb{E}Y_n}{a_n},$$

a.e. 收敛.

其次, 有

$$\begin{aligned} \sum_n \frac{|\mathbb{E}Y_n|}{a_n} &= \sum_n \frac{1}{a_n} \left| \int_{|x| \le a_n} x dF_n(x) \right| \\ &= \sum_n \frac{1}{a_n} \left| \int_{|x| > a_n} x dF_n(x) \right| \\ &\le \sum_n \int_{|x| > a_n} \frac{|x|}{a_n} dF_n(x), \end{aligned}$$

注意到第二个等号是因为 $\mathbb{E}X_n = 0$ 和绝对值符号. 根据 φ 的第一个假设, 对于 $|x| > a_n$, 有

$$\frac{|x|}{a_n} \leq \frac{\varphi(x)}{\varphi(a_n)}.$$

故有

$$\begin{aligned}\sum_n \frac{|\mathbb{E}Y_n|}{a_n} &\leq \sum_n \int_{|x|>a_n} \frac{\varphi(x)}{\varphi(a_n)} dF_n(x) \\ &\leq \sum_n \frac{\mathbb{E}\varphi(X_n)}{\varphi(a_n)} < \infty.\end{aligned}$$

这一结论和上一结论共同蕴含 $\sum_n Y_n/a_n$a.e. 收敛.

最后, 由于 φ 单调增, 有

$$\begin{aligned}\sum_n \mathscr{P}\{X_n \neq Y_n\} &= \sum_n \int_{|x|>a_n} dF_n(x) \\ &\leq \sum_n \int_{|x|>a_n} \frac{\varphi(x)}{\varphi(a_n)} dF_n(x) \\ &\leq \sum_n \frac{\mathbb{E}\varphi(X_n)}{\varphi(a_n)} < \infty.\end{aligned}$$

这就得到 $\{X_n\}$ 和 $\{Y_n\}$ 是等价序列. 和上一结论共同蕴含 $\sum_n X_n/a_n$a.e. 收敛. □

注意到根据 Kronecker 引理, 上述定理有一个推论:

$$\frac{1}{a_n}\sum_{j=1}^{n} X_j \to 0, \qquad a.e.$$

特别需要注意的是一个特殊情况: $\varphi(x) = |x|^p$, 其中 $1 \leq p \leq 2$ 的时候.

现在来讨论强大数定律: 这是弱大数律的强极限版本.

定理 2.2 (Колмогоров)**.** *设 $\{X_n\}$ 是 i.i.d. r.v. 序列, 且 $\mathbb{E}|X_1| < \infty$, 则*

$$\frac{S_n}{n} \to \mathbb{E}X_1, \quad \text{a.e.}$$

证明. 对于 $n \in \mathbb{N}_+$, 如上定义 Y_n 是每个 X_n 关于 n 的截尾变量. 注意到 “具有启发性的结论”, 有

$$\sum_n \mathscr{P}\{X_n \neq Y_n\} = \sum_n \mathscr{P}\{|X_n| > n\} = \sum_n \mathscr{P}\{|X_1| > n\} < \infty.$$

这是说 $\{X_n\}$ 和 $\{Y_n\}$ 是等价序列. 对于序列 $\{Y_n - \mathbb{E}Y_n\}$, 取 $\varphi(x) = x^2$, 此时有

$$\sum_n \frac{\sigma^2(Y_n)}{n^2} \leq \sum_n \frac{\mathbb{E}Y_n^2}{n^2} = \sum_n \frac{1}{n^2}\int_{|x|\leq n} x^2 dF(x).$$

为了使用一阶矩估计 RHS 的二阶矩, 通过分割积分区间, Fubini 定理后将一个因子 x

用端点放缩, 即

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{j=1}^{n}\int_{j-1<|x|\leq j}x^2dF(x) &= \sum_{j=1}^{\infty}\int_{j-1<|x|\leq j}x^2dF(x)\sum_{n=j}^{\infty}\frac{1}{n^2}\\ &\leq \sum_{j=1}^{\infty}j\int_{j-1<|x|\leq j}|x|dF(x)\frac{C}{j}\\ &\leq C\sum_{j=1}^{\infty}\int_{j-1<|x|\leq j}|x|dF(x)\\ &= C\mathbb{E}(|X_1|)<\infty.\end{aligned}$$

其中第一个不等号是因为 $\sum_{n=j}^{\infty}n^{-2}\leq Cj^{-1}$, $C=const.$, 故 $\sum_n\frac{\sigma^2(Y_n)}{n^2}$ 收敛, 应用如上的 $\varphi(x)=x^2$ 的特殊结论于序列 $\{Y_n-\mathbb{E}Y_n\}$, 得到结论

$$\frac{1}{n}\sum_{j=1}^{n}Y_j-EY_j\to 0,\qquad \text{a.e.}$$

又因为 $n\to\infty$ 时, $\mathbb{E}Y_n\to\mathbb{E}X_1$, 故有

$$\frac{1}{n}\sum_{j=1}^{n}Y_j\to\mathbb{E}X_1,\qquad \text{a.e.}$$

强大数律得证. □

参考文献

[CKL] A Course in Probability Theory, 钟开莱著, 刘文, 吴让泉译, 1989, 上海科学技术出版社.