Topic 2

- Convergence
  - Type 1: almost surely convergence
    - definition: $\lim_{n\to\infty} X_n(\omega) = X(\omega)$

      $P\left(\{\omega: \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1$
    - or

      $P\left(\lim_{n\to\infty} |X_n - X| < \varepsilon\right) = 1$
    - Borel-Cantelli Lemma

      Theorem 1 (Borel-Cantelli Lemma)
      We let $A_1, A_2, \ldots$ (where $A_n \in \mathcal{F}$) be a sequence of events, then we have the following statements:
        - 1) If $\sum_{n=1}^{\infty} P(A_n) < \infty$,then $P(A_n\ i.o.) = 0$.
        - 2) If all events A_n are independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(A_n\ i.o.) = 1$.
      - proof
        - 1

          ① $\sum_{n=1}^{\infty} P(A_n) < \infty$ then $P(A_n\ i.o.) = 0$. $\{A_n\ i.o.\} = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k$.

          $1_{A_n}(\omega) = \begin{cases} 1 & \text{if } \omega \in A_n \\ 0 & \text{if } \omega \notin A_n \end{cases}$ $\{A_n\ i.o.\} \Leftrightarrow \sum 1_{A_n} = \infty$

          我们用反证法. if $\{A_n\ i.o.\} > 0 \Rightarrow \mathbb{E}[\sum 1_{A_n}] = \infty$.

          $\mathbb{E}[\sum_{n=1}^{\infty} 1_{A_n}] = \mathbb{E}[\lim_{n\to\infty} \underbrace{\sum_{k=1}^{n} 1_{A_k}}_{\text{increasing}\ \&\ \geq 0}] = \lim_{n\to\infty} \mathbb{E}[\sum_{k=1}^{n} 1_{A_k}]$

          $= \lim_{n\to\infty} \sum_{k=1}^{n} P(A_k)$

          $= \sum_{k=1}^{\infty} P(A_k) < \infty$. 矛盾.
        - 2

          ② if A_n independent. $\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow P(A_n\ i.o.) = 1$ $e^x \geq x+1$

          proof: $P(\bigcap_{n=m}^{\infty} A_n^c) \overset{\text{independent}}{=} \prod_{n=m}^{\infty} P(A_n^c) = \prod_{n=m}^{\infty} (1 - P(A_n))$ $e^{-x} \geq -x+1$

          $\leq \prod_{n=m}^{\infty} e^{-P(A_n)} = e^{-\sum_{n=m}^{\infty} P(A_n)} = 0$

          $\Rightarrow P(\bigcap_{n=m}^{\infty} A_n^c) = 0 \Rightarrow P(\underbrace{\bigcup_{n=m}^{\infty} A_n}_{E_m}) = 1$ E_m is decreasing

          $P(A_n\ i.o.) = P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = P(\bigcap_{n=1}^{\infty} E_n) = P(E_\infty) = P(\lim_{k\to\infty} E_k)$

          $= \lim_{k\to\infty} P(E_k) = 1$
      - 两个之间的等价：如果a.s. 令 $A_n = \{\omega : |X_n(\omega) - X(\omega)| > \epsilon\}, P(A_n i.o.) = 0$

- $<=> \omega \in \{A_n i.o.\}^c$
- $<=> \sum_{n=1}^{+\infty} P(A_n) < +\infty$

- Type 2: converge in probability
  - definition

    $$\lim_{n\to\infty} P(\{\omega: |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0 \text{ for any } \varepsilon > 0.$$

  - 注意这里的极限提到了概率的前面，首先这可以说明并不是任何时候P和极限都可以互相交换，其次这是个更弱的条件，因为这个的意思其实是，我有无数个trajectories, 这些轨道中不处在一个区间内的概率趋于一，而as的意思是任意一个都处于这个区间之中
- Type 3: converge in distribution
  - definition

    $$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{for every } x \text{ where } F_X(x) \text{ is continuous.}$$

  - 一般利用MGF和CF来计算比较多

    Theorem 3 (**Lévy's continuity theorem**)
    We let $X_1, X_2, \cdots$ be a sequence of random variables. let X be a random variable. If either one of the following conditions holds
    - (1) The MGFs of X_n and X exists and
      $$\lim_{n\to\infty} M_{X_n}(t) = M_X(t) \text{ for all t near 0.} \quad \mathbb{E}[e^{tx}]$$
    - (2) The characteristic functions of X_n and X satisfies
      $$\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t) \quad \mathbb{E}[e^{itx}]$$

    Then $X_n \xrightarrow{d} X$

    34

  - property：已知$X_n \to^d X$
    - 如果都有限，那么期望的极限也相等

      Suppose that $|X_n| \le Y$ and $\mathbb{E}[Y] < \infty$,then
      $$\lim_{n\to\infty} \mathbb{E}[X_n] = \lim_{n\to\infty} \int_{-\infty}^{\infty} x dF_{X_n}(x) = \lim_{n\to\infty} \int_{-\infty}^{\infty} x dF_X(x) = \mathbb{E}[X]$$

    - 有限，那么加一个continuous function，那么也依分布收敛，同时根据property 1，期望的极限也相等

      (2) **(Bounded convergence theorem)** We let $g(x)$ be a continuous function such that $|g(x)| \le C < \infty$, then
      $$g(X_n) \xrightarrow{d} g(X) \text{ and } \lim_{n\to\infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$$

- Type 4: Converge in L^p norm
  - definition

    $$\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$$

- Relationship
  - a.s. --> p

$$\text{a.s.} \Rightarrow p(\lim_{n\to\infty} X_n = X) = 1$$

$$\Rightarrow p(\lim_{n\to\infty} |X_n - X| > \varepsilon) = 0$$

$$A_n = \bigcup_{m=n}^{\infty} \{w : |X_m(w) - X(w)| > \varepsilon\}. \quad A_n \text{ is decreasing.}$$

$$\lim_{n\to\infty} p(A_n) = p(A) = 0$$

$$\lim_{n\to\infty} p(|X_n - X| > \varepsilon) \le \lim_{n\to\infty} p(A_n) = p(A) = 0$$

为什么这里是小于等于、因为这里的 A_n 是取并哦.

$\Rightarrow$ P convergence

- p --> subsequence a.s.

P --> subsequence a.s.

$$\forall \varepsilon. \lim_{n\to\infty} p(|A_n - A| > \varepsilon) = 0.$$

取 $\varepsilon = \frac{1}{k}$. subsequence $p(|A_{n_k} - A| > \frac{1}{k}) \le \frac{1}{k^2}$

$$\sum p(A_{n,\frac{1}{k}}) \le \sum \frac{1}{k^2} < +\infty. \Rightarrow \text{a.s.}$$

反例来 false: 1. -1. 1. -1. . .

But: if every subsequence has a sub-subsequence. a.s.

$\Leftrightarrow$ P

  - question!!!!!

- p --> d

$p \Rightarrow d.$

$$\lim_{n\to\infty} p(|X_n(w) - X(w)| > \varepsilon) = 0 \Rightarrow \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

(a) $\lim_{n\to\infty} F_{X_n}(t) \le F_X(t)$ (b) $\lim_{n\to\infty} F_{X_n}(t) \ge F_X(t)$

(a) $p(X_n \le t) = p(X_n \le t \mid |X_n - X| < \varepsilon) \cdot p(|X_n - X| < \varepsilon)$ $\quad -\varepsilon < X_n - X < \varepsilon$

$+ p(X_n \le t \mid |X_n - X| \ge \varepsilon) \cdot p(|X_n - X| \ge \varepsilon)$ $\quad X_n - \varepsilon < X < X_n + \varepsilon \le t + \varepsilon.$

$= p(X_n \le t \ \& \ |X_n - X| < \varepsilon) + p(X_n \le t \mid |X_n - X| \ge \varepsilon) \cdot p(|X_n - X| \ge \varepsilon).$

$= p(X_n - \varepsilon < X \le t + \varepsilon) + \boxed{p(X_n \le t \mid |X_n - X| \ge \varepsilon)} p(|X_n - X| \ge \varepsilon)$

$\le p(X < t + \varepsilon) + p(|X_n - X| \ge \varepsilon)$ 把这一项放缩.

$n \to \infty,\ \varepsilon \to 0$ $\quad$ p收敛 $\Rightarrow = 0.$

$= p(X < t)$ $\quad \therefore$ (a)证毕. $t + \varepsilon < X_n - \varepsilon < X < X_n + \varepsilon$

(b) $p(X_n > t) = p(X_n > t \ \& \ |X_n - X| < \varepsilon) + p(X_n > t \mid |X_n - X| \ge \varepsilon) p(|X_n - X| \ge \varepsilon).$

$\le p(X > t - \varepsilon) + p(|X_n - X| \ge \varepsilon)$

$n \to \infty.\ \varepsilon \to 0$

$= p(X > t)$ $\quad \therefore p(X_n \le t) = 1 - p(X_n > t) \ge 1 - p(X > t) = p(X \le t)$

$$\Rightarrow \lim_{n\to\infty} F_{X_n}(t) = F_X(t).$$

- L^r --> p

$L^r \to P$

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

Markov Inequality. X is nonnegative. $a > 0 \Rightarrow p(X \ge a) \le \frac{E[X]}{a}.$

$$\lim_{n\to\infty} p(|X_n - X| \ge \varepsilon) \le \lim_{n\to\infty} p(|X_n - X|^r \ge \varepsilon^r)$$

$$\le \lim_{n\to\infty} \frac{E[|X_n - X|^r]}{\varepsilon^r}$$

$$= 0.$$

- 我们之前学过一个数列的极限一定是独特的，但是随机变量中并没有这个性质

  > (ii) No, we consider the following example: Let U be a random variable which has uniform distribution over $[0,1]$.
  > Define $X_n = U^n$, $X = 0$ and $Y = \begin{cases} 0 & if \;\; U \neq 1 \\ 1 & if \;\; U = 1 \end{cases}$.
  > One can show that $X_n \to X$ and $X_n \to Y$ almost surely (by one example in the lecture note). But $X \neq Y$.

- Law of large numbers
  - WLLN
    - iid + $E[|X_i|] < +\infty$, $\frac{X_1+X_2+...+X_n}{n} \to^{p} \mu$
    - weak version 1: iid + $E[|X_i|] < +\infty, Var(X_i) = \sigma^2 < +\infty \to$ $\frac{X_1+X_2+...+X_n}{n} \to^{P or L^2} \mu$

      $S_n = X_1 + \dots + X_n$. $\mathbb{E}[\frac{S_n}{n}] = \mathbb{E}[X_1] = \mu$. $Var(\frac{S_n}{n}) = \frac{1}{n^2} Var(S_n) = \frac{1}{n} \cdot \sigma^2$.

      L^2: $\lim_{n\to\infty} \mathbb{E}[|\frac{S_n}{n} - \mu|^2] = \lim_{n\to\infty} \mathbb{E}[(\frac{S_n}{n})^2] - 2\mathbb{E}[\frac{S_n}{n} \mu] + \mu^2$
      $= \lim_{n\to\infty} \frac{1}{n}\sigma^2 + \mu^2 - \mu^2$
      $= \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$

      P: $\lim_{n\to\infty} p(|\frac{S_n}{n} - \mu| > \varepsilon) \le \lim_{n\to\infty} p(|\frac{S_n}{n} - \mu|^2 > \varepsilon^2)$
      $\le \lim_{n\to\infty} \frac{\mathbb{E}[|\frac{S_n}{n} - \mu|^2]}{\varepsilon^2}$
      $= \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$

    - weak version 2:

      > ▪ **Theorem (Weaker version of Weak Law of Large Numbers):**
      > If $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables such that
      > 没有E和 $Var < +\infty$
      > $$tP(|X_i| > t) \xrightarrow{t\to\infty} 0$$
      > then
      > $$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{p} \mu_n \quad (converge\ in\ probability)$$
      > Where $\mu_n = \mathbb{E}[X_1 1_{\{|X_1| \le n\}}]$,

      - proof

        $Y_{1,n} = \begin{cases} X_1 & if\ |X_1| \le n \\ 0 & if\ |X_1| > n \end{cases}$

        $p(\frac{S_n}{n} = \frac{\hat{S}_n}{n}) \ge p(X_1 = Y_{1,n}\ for\ all\ n = 1, 2, \ldots) = \prod_{i=1}^{n} p(X_1 = Y_{1,n}) = \prod_{i=1}^{n} p(|X_1| \le n)$
        $= [1 - p(|X_1| > n)]^n \approx e^{-np(|X_1| > n)}$

        $n\to\infty$时. $np[|X_1| > n] \xrightarrow{n\to\infty} 0$

        $p(\frac{S_n}{n} = \frac{\hat{S}_n}{n}) \xrightarrow{n\to\infty} 1$

        $\therefore p(|\frac{S_n}{n} - \underbrace{\mathbb{E}[X_1 1_{\{|X_1| \le n\}}]}_{\mu_n}| \ge \varepsilon) = p(|\frac{S_n}{n} - \mu_n| \ge \varepsilon \mid \frac{S_n}{n} = \frac{\hat{S}_n}{n}) p(\frac{S_n}{n} = \frac{\hat{S}_n}{n})$
        $+ p(|\frac{S_n}{n} - \mu_n| \ge \varepsilon \mid \frac{S_n}{n} \neq \frac{\hat{S}_n}{n}) p(\frac{S_n}{n} \neq \frac{\hat{S}_n}{n})$
        $= \underbrace{p(|\frac{\hat{S}_n}{n} - \mu_n| \ge \varepsilon)}_{\to 0} \underbrace{p(\frac{S_n}{n} = \frac{\hat{S}_n}{n})}_{=1} + \underbrace{p(|\frac{S_n}{n} - \mu_n| \ge \varepsilon \mid \frac{S_n}{n} \neq \frac{\hat{S}_n}{n})}_{\le 1} \underbrace{p(\frac{S_n}{n} \neq \frac{\hat{S}_n}{n})}_{=0}$.

        $n\to\infty$时 由WLLN step 1.

        $\to 0$ $\therefore \frac{S_n}{n} \xrightarrow{P} \mathbb{E}[X_1 1_{\{|X_1| \le n\}}]$

    - WLLN proof

- 首先，我们不能直接用weak version 1因为方差的有限性我们不能保证，但我们可以努力往这个方向靠，创造一个新的变量$Y_{i,n}$,让他在$X_i > n$时等于0，等于人为保证了方差的有限性
- 先引入一个Lemma: $E[X^p] = \int_0^\infty px^{p-1}P(X > x)dx$

$$
\begin{aligned}
\int_0^\infty px^{p-1}P(X>x)dx &= \int_0^\infty px^{p-1}\int_{-\infty}^{+\infty}1_{\{X>x\}}dF_X(t)dx \\
&= \int_{-\infty}^{+\infty}\int_0^{+\infty} px^{p-1}1_{\{X>x\}}dx\,dF_X(t) \\
&= \int_{-\infty}^{+\infty}\boxed{\int_0^X px^{p-1}dx}\,dF_X(t) \\
&= \int_{-\infty}^{+\infty} X^p\,dF_X(t) \\
&= \mathbb{E}[X^p]
\end{aligned}
$$

- step 1: $\frac{\hat{S}_n}{n} \to^p \mu_n$

$$
P\left(\left|\frac{\hat{S}_n}{n} - \mathbb{E}[X_i 1_{\{|X_i|\le n\}}]\right| > \varepsilon\right) = P\left(\left|\frac{\hat{S}_n}{n} - \mathbb{E}\left[\frac{\hat{S}_n}{n}\right]\right| > \varepsilon\right) \le \frac{Var\left(\frac{\hat{S}_n}{n}\right)}{\varepsilon^2} = \frac{Var(Y_{i,n})}{n\varepsilon^2} \le \frac{\mathbb{E}[Y_{i,n}^2]}{n\varepsilon^2}
$$

下面我们来求 $\mathbb{E}[Y_{i,n}^2]$

Lemma $\mathbb{E}[Y_{i,n}^2] = \int_0^\infty 2y\,P(|Y_{i,n}|>y)dy = 2\int_0^n y\,P(|Y_{i,n}|>y)dy$

$$
\begin{aligned}
&= 2\int_0^n y\cdot P(y<|X_{i,n}|<n)dy \\
&\le 2\int_0^n y\,P(|X_i|>y)\,dy \quad \text{choose } n \text{ sufficiently large. such that } n>k. \\
&= 2\left[\int_0^k y\,P(|X_i|>y)dy + 2\int_k^n \underbrace{y\,P(|X_i|>y)}\right]dy \\
&\le 2\cdot\int_0^k y\,dy + 2\cdot\int_k^n \varepsilon_k\,dy \quad \underbrace{\text{how to calculate}}
\end{aligned}
$$

$\varepsilon_k = \sup\{y\,P(|X_i|>y) : y>k\}$.

$y\,P(|X_i|>y) = \mathbb{E}[y\cdot 1_{\{|X_i|>y\}}] \le \mathbb{E}[|X_i| 1_{\{|X_i|>y\}}] \xrightarrow{y\to\infty} 0$.

$\therefore \lim_{k\to\infty}\varepsilon_k = 0$

$$
= 2\frac{k^2}{2} + 2\times\varepsilon_k(n-k) = k^2 + 2\varepsilon_k(n-k).
$$

$$
\therefore P\left(\left|\frac{\hat{S}_n}{n} - \mathbb{E}[X_i 1_{\{|X_i|\le n\}}]\right| > \varepsilon\right) \le \frac{k^2+2\varepsilon_k(n-k)}{n\varepsilon^2} = \frac{k^2}{n\varepsilon^2} + \frac{2\varepsilon_k}{\varepsilon^2}\left(1-\frac{k}{n}\right) \xrightarrow[n\to\infty]{\varepsilon_k\to 0} 0 + 0\times 1 = 0
$$

- step 2: $\frac{S_n}{n} \to^p \mu_n$ (weak version 2)
- step 3: $\frac{S_n}{n} \to^p \mu$

$$
\begin{aligned}
\left|\frac{S_n}{n} - \mathbb{E}[X_i]\right| &= \left|\frac{S_n}{n} - \mathbb{E}[X_i 1_{\{|X_i|\le n\}}] + \mathbb{E}[X_i 1_{\{|X_i|\le n\}}] - \mathbb{E}[X_i]\right| \\
&\le \underbrace{\left|\frac{S_n}{n} - \mathbb{E}[X_i 1_{\{|X_i|\le n\}}]\right|}_{A} + \underbrace{\left|\mathbb{E}[X_i 1_{\{|X_i|\le n\}}] - \mathbb{E}[X_i]\right|}_{B}
\end{aligned}
$$

if $\left|\frac{S_n}{n} - \mathbb{E}[X_i]\right| > \varepsilon$. A,B至少有一个$>\frac{\varepsilon}{2}$.

因为我们有 $\lim_{n\to\infty} X_i(w) 1_{\{|X_i(w)|\le n\}} = X_i(w)$ $\therefore \lim_{n\to\infty}\mathbb{E}[X_i 1_{\{|X_i|\le n\}}] = \mathbb{E}[X_i]$

$\therefore$ B不可能大于$\frac{\varepsilon}{2}$.

考虑第一项 $\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mathbb{E}[X_i]\right| > \varepsilon\right) \le \lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mathbb{E}[X_i 1_{\{|X_i|\le n\}}]\right| > \frac{\varepsilon}{2}\right) \overset{\text{By step 2}}{=} 0$

$\therefore \frac{S_n}{n} \xrightarrow{\mu} \mathbb{E}[X_i]$

- SLLN
  - iid + $E[|X_i|] < +\infty$, $\frac{X_1+X_2+\ldots+X_n}{n} \to^{a.s.} \mu$
  - weak version 1: iid + $E[|X_i|] < \infty$, $\sum_{n=1}^\infty \frac{Var(X_n)}{n^2} < \infty \to \frac{S_n}{n} \to^{a.s.} \mu$

- 可以比较一下，这里不仅仅是p变成了a.s.，Var 的有限性也变成了Var除以n^2的和的有限性
- proof（这里我们只考虑简单的情况，$\mu=0$, 即independent + $E[X_n]=0, \sum_{n=1}^{\infty}\frac{Var(X_n)}{n^2}<\infty \to \frac{S_n}{n} \to^{a.s.} 0$）
  - 这里不再是iid了，只是独立就可以，但是期望为0
  - Lemma 1:

    Kronecker Lemma. Suppose $a_n > 0$ and $a_n \uparrow \infty$. Then $\sum_{n=1}^{\infty}\frac{x_n}{a_n} < \infty$ implies $\frac{1}{a_n}\sum_{j=1}^{n} x_j \to 0$

    b_∞

    - proof

      Toeplitz lemma: Let $c_n \geq 0$. $l_n = \sum_{j=1}^{n} c_j > 0$ for $n \geq 1$. $l_n \to \infty$.
      If $x_n \to x$ when $n\to\infty$, then $\frac{1}{l_n}\sum_{j=1}^{n} c_j x_j \to x$ when $n \to \infty$

      $b_n = \sum_{i=1}^{n} x_i/a_i \quad a_0 = b_0 = 0 \qquad b_\infty = \sum_{n=1}^{\infty} x_n/a_n < \infty. \therefore b_n \to b_\infty.$

      $x_n = (b_n - b_{n-1})a_n$

      $$\begin{aligned}\therefore \frac{1}{a_n}\sum_{j=1}^{n} x_j &= \frac{1}{a_n}\sum_{j=1}^{n}(b_j - b_{j-1})a_j = \frac{1}{a_n}\sum_{j=1}^{n} b_j a_j - b_{j-1}a_j \\ &= \frac{1}{a_n}\left(b_n a_n + \sum_{j=1}^{n-1} b_j a_j - \sum_{j=1}^{n} a_j b_{j-1}\right] \\ &= b_n + \frac{1}{a_n}\left[\sum_{j=1}^{n} a_{j-1}b_{j-1} - \sum_{j=1}^{n} a_j b_{j-1}\right] \\ &= b_n - \frac{1}{a_n}\sum_{j=1}^{n}(a_j - a_{j-1})b_{j-1}\end{aligned}$$

      $n\to\infty$, $\to b_\infty - b_\infty = 0$ $\qquad x_n = b_n \quad c_j = a_j - a_{j-1}$. $l_n = \sum_{j=1}^{n} c_j = a_n \uparrow \infty$

      $\xRightarrow{\text{lemma}} \frac{1}{a_n}\sum_{j=1}^{n}(a_j - a_{j-1})b_j \to b_\infty$

      ▨

  - Lemma 2:

    ▪ **Theorem** (**Khintchine-Kolmogorov Convergence Theorem**)
    Suppose $X_1, X_2, \cdots$ are independent with mean 0 such that $\sum_{n=1}^{\infty} var(X_n) < \infty$. Then, $\sum_{n=1}^{\infty} X_n < \infty$ a.s. (i.e., S_n converges a.s. as well as in L^2 to $\sum_{n=1}^{\infty} X_n$.)

    - proof

      Kolmogorov's inequality: $E[X_k]=0$. $Var(X_k)<\infty$.
      $\Rightarrow P(\max_{1\le k\le n}|S_k| \geq \varepsilon) \leq \frac{Var(S_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2}\sum_{k=1}^{n} Var(X_k) = \frac{1}{\varepsilon^2}\sum_{k=1}^{n} E[X_k^2]$

      define $A_{m,\varepsilon} = \{\max_{j>m}|S_j - S_m| \leq \varepsilon\}$. Then $\{\sum_{n=1}^{\infty} X_n < \infty\} = \bigcap_{\varepsilon>0}\bigcup_{m\geq 1} A_{m,\varepsilon}$.

      $$\begin{aligned}P(A_{m,\varepsilon}) = P(\max_{j>m}|S_j - S_m| \leq \varepsilon) &= 1 - P(\max_{j>m}|S_j - S_m| > \varepsilon) \\ &\geq 1 - \frac{Var(S_j - S_m)}{\varepsilon^2} \\ &= 1 - \frac{1}{\varepsilon^2}\sum_{k=m}^{j-1} Var(X_k) \\ &\xrightarrow{m\to\infty,\ j\to\infty} 1\end{aligned}$$

      $A_{m,\varepsilon}$. $m\uparrow$. $A_{m,\varepsilon}\uparrow$.

      $\therefore \lim_{m\to\infty} P(A_{m,\varepsilon}) \to 1. \Rightarrow P(\bigcup_{m\geq 1} A_{m,\varepsilon}) = 1$

      $P(\{\sum_{n=1}^{\infty} X_n < \infty\}) = P(\bigcap_{\varepsilon>0}\bigcup_{m\geq 1} A_{m,\varepsilon}) = 1$ $\Rightarrow$ a.s.

      $$\begin{aligned}E[(S_n - S_\infty)^2] &= E[(S_n - \lim_{k\to\infty} S_k)^2] = E[\lim_{k\to\infty}(S_n - S_k)^2] \leq \liminf_{k\to\infty} E[(S_n - S_k)^2] \\ &= \liminf_{k\to\infty}\sum_{j=n}^{k} Var(X_j) = \sum_{j=n}^{\infty} Var(X_j) \xrightarrow{n\to\infty} 0.\end{aligned}$$

  - proof process
    - 根据Lemma2，可以推得$\sum X_n < \infty, a_n = n$,再根据Lemma 1，可以直接得出$\frac{S_n}{n} \to^{a.s.} \mu$ (反正μ只是一个常数)
- weak version 3: iid+ $Var(X_i) < \infty, E[X^4] = C < \infty, \to \frac{S_n}{n} \to^{a.s.} \mu$

- proof

$$A_{n,\varepsilon}=\{w:\left|\frac{X_1(w)+\dots+X_n(w)}{n}-\mu\right|\geq\varepsilon\}$$

$$P(A_{n\varepsilon})=P(|S_n|\geq n\varepsilon)\leq\frac{\mathbb{E}[S_n^4]}{(n\varepsilon)^4}=\frac{1}{n^4\varepsilon^4}\mathbb{E}[(X_1+\dots+X_n)^4]$$

观察 $(X_1+\dots+X_n)^4$. 考虑 $\mu=0$.

X_i^4. 有n个. $n\mathbb{E}[X_i^4]=n\cdot C$.

$X_i^3X_j$. 有 $C_n^2\cdot C_4^1$ $\mathbb{E}[X_i^3X_j]=C_4^2C_3^1\cdot\mathbb{E}[X_i^3]\cdot\mathbb{E}[X_j]=0$

$X_i^2X_j^2$. 有 $C_n^2\cdot C_4^2$ $\mathbb{E}[X_i^2X_j^2]=C_4^2C_4^2\cdot\mathbb{E}[X_i^2]\mathbb{E}[X_j^2]$

$=C_4^2C_4^2\,Var(X_i)\cdot Var(X_j)=C_n^2C_4^2\cdot\sigma^4$

$\mathbb{E}[X_iX_jX_kX_l]=0$

$\Rightarrow \mathbb{E}[(X_1+\dots+X_n)^4]=nC+3n(n-1)\cdot\sigma^4$

$\Rightarrow\sum P(A_{n,\varepsilon})=\sum\frac{1}{n^4\varepsilon^4}(nC+3n(n-1)\sigma^4)<\infty \Rightarrow$ a.s.

- SLLN proof
  - 和WLLN相似，只是这里的Y_i需要重新修改一下，不再是之前的$Y_{i,n}$了
  - 且这里要考虑子序列，然后用子序列逼近原序列，从而证明
  - 我们在这里只证明X>0的情况，如果X不满足这个条件，可以吧X分解成$X=X^+-X^-$,也就可以得到最终的结果
  - Step 1: $\frac{\hat{S}_n}{n}\rightarrow^{a.s.}\frac{E[\hat{S}_n]}{n}\rightarrow^{a.s.}\mu$

$$Y_i=\begin{cases}X_i & \text{if } |X_i|\leq i\\ 0 & \text{if } |X_i|>i\end{cases}=X_i1_{\{|X_i|\leq i\}}$$

$$A_{n,\varepsilon}=\{w:\left|\frac{Y_1(w)+\dots+Y_n(w)}{n}-\frac{\mathbb{E}[\tilde{S}_n]}{n}\right|>\varepsilon\}$$

$$\sum P(A_{n,\varepsilon})=\sum P\left(\left|\frac{\tilde{S}_n}{n}-\frac{\mathbb{E}[\tilde{S}_n]}{n}\right|\geq\varepsilon\right)\leq\sum\frac{Var(\frac{\tilde{S}_n}{n})}{\varepsilon^2}$$

交换求和顺序

$$=\frac{1}{\varepsilon^2}\sum\left(\sum_{k=1}^{n}\frac{Var(Y_k)}{n^2}\right)\leq\frac{1}{\varepsilon^2}\sum_{n=1}^{\infty}\left(\sum_{k=1}^{n}\frac{\mathbb{E}[Y_k^2]}{n^2}\right)=\frac{1}{\varepsilon^2}\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\frac{\mathbb{E}[Y_k^2]}{n^2}.$$

很难直接考虑 $\mathbb{E}[Y_k^2]$ ⇒ 先考虑子序列. 取 $n_k=\lfloor\alpha^k\rfloor$ 先证 $\frac{\tilde{S}_{n_k}}{n_k}\xrightarrow{a.s.}\mu$

k is the smallest integer k st. $n_k=\lfloor\alpha^k\rfloor\geq m$.

$$\sum_{m=1}^{\infty}\sum_{n_k=m}^{\infty}\frac{\mathbb{E}[Y_m^2]}{n_k^2}=\sum_{m=1}^{\infty}\mathbb{E}[Y_m^2]\sum_{k:\lfloor\alpha^k\rfloor\geq m}^{\infty}\frac{1}{\lfloor\alpha^k\rfloor^2}\leq\sum_{m=1}^{\infty}\mathbb{E}[Y_m^2]\sum_{k:\lfloor\alpha^k\rfloor\geq m}^{\infty}\frac{4}{\alpha^{2k}}$$

$$=4\sum_{m=1}^{\infty}\mathbb{E}[Y_m^2]\frac{\frac{1}{\alpha^{2k}}}{1-\frac{1}{\alpha^2}}=\frac{4}{1-\alpha^{-2}}\sum_{m=1}^{\infty}\frac{\mathbb{E}[Y_m^2]}{\alpha^{2k}}\leq\frac{4}{1-\alpha^{-2}}\sum_{m=1}^{\infty}\frac{\mathbb{E}[Y_m^2]}{m^2}\leq\frac{8\mathbb{E}[|X_1|+1]}{1-\alpha^{-2}}<\infty$$

$$\therefore\sum_{k=1}^{\infty}P(A_{n_k,\varepsilon})\leq\frac{1}{\varepsilon^2}\sum\sum\frac{\mathbb{E}[Y_m^2]}{n_k^2}<\infty.\Rightarrow\frac{\tilde{S}_{n_k}}{n_k}\xrightarrow{a.s.}\frac{\mathbb{E}[\tilde{S}_{n_k}]}{n_k}.$$

$$\frac{\mathbb{E}[\tilde{S}_{n_k}]}{\tilde{S}_{n_k}}=\frac{\mathbb{E}[Y_1+\dots+Y_{n_k}]}{n_k}=\frac{\mathbb{E}[X_1]+\dots+\mathbb{E}[X_{n_k}]}{n_k}-\frac{\mathbb{E}[X_11_{\{|X_1|>1\}}]+\dots+\mathbb{E}[X_{n_k}1_{\{|X_{n_k}|>n_k\}}]}{n_k}$$

$\rightarrow 0.$

$$=\mu-\frac{\mathbb{E}[X_11_{\{|X_1|>1\}}]+\dots+\mathbb{E}[X_{n_k}1_{\{|X_{n_k}|>n_k\}}]}{n_k}.$$

$$\therefore\frac{\tilde{S}_{n_k}}{n_k}\xrightarrow{a.s.}\frac{\mathbb{E}[\tilde{S}_{n_k}]}{n_k}\xrightarrow{a.s.}\mu.$$

- 再通过子列来证明原始数列

$$\frac{\tilde{S}_{n_k}}{n_{k+1}}<\frac{\tilde{S}_n}{n}<\frac{\tilde{S}_{n_{k+1}}}{n_k}\qquad\lim_{k\to\infty}\frac{n_{k+1}}{n_k}=\frac{\lfloor\alpha^{k+1}\rfloor}{\lfloor\alpha^k\rfloor}=\alpha$$

$$\mu\alpha=\lim_{k\to\infty}\frac{\tilde{S}_{n_k}}{n_k}\left(\frac{n_k}{n_{k+1}}\right)\leq\lim_{n\to\infty}\frac{\tilde{S}_n}{n}\leq\lim_{k\to\infty}\frac{\tilde{S}_{n_{k+1}}}{n_{k+1}}\cdot\frac{n_{k+1}}{n_k}=\mu\alpha.$$

$$\xRightarrow{\alpha\to1}\lim_{n\to\infty}\frac{\tilde{S}_n}{n}=\mu.\quad a.s.$$

- Step 2: $\frac{S_n}{n} \to^{a.s.} \mu$

$$\frac{S_n}{n} = \frac{\tilde S_n}{n} + \left(\frac{S_n}{n} - \frac{\tilde S_n}{n}\right)$$

Let $A_k = \{w : X_k \neq Y_k\}$

$$\sum P(A_k) = \sum_{k=1}^{\infty} P(X_k \neq Y_k) = \sum_{k=1}^{\infty} P(|X_k| > k) = \sum_{k=1}^{\infty} P(|X_1| > k) = \int_0^{\infty} P(|X_1| > t)\,dt$$

$$= \mathbb{E}[|X_1|] < \infty$$

∴ $X_k = Y_k$. i.o.　　∴ ∃K. n>K 时 $X_k = Y_k$.

$$\therefore \frac{S_n - \tilde S_n}{n} = \sum \frac{X_k - Y_k}{n} = \sum_{k=1}^{\min(n,K)} \frac{X_k - Y_k}{n} = \sum_{k=1}^{K} \frac{X_k - Y_k}{n} \xrightarrow{n\to\infty} 0.$$

$$P\left(\lim_{n\to\infty} \frac{S_n - \tilde S_n}{n} = 0\right) \geq P(\{A_k \text{ i.o.}\}^c) = 1 \quad \therefore \frac{S_n}{n} \overset{a.s.}{\sim} \frac{\tilde S_n}{n} \overset{a.s.}{\sim} \mu$$

- Central limit theorem
  - definition

> ▪ **Central Limit Theorem for iid r.v.s: Suppose** $\{X_1, \ldots, X_n, \ldots\}$ **is a sequence of independent and identically distributed (i.i.d.) random variables with** $E(X_i) = \mu$ **and** $Var(X_i) = \sigma^2 < \infty$ **. Then as** $n \to \infty$**,**
>
> $$\frac{\bar X_n - \mu}{\sigma/\sqrt{n}} = \frac{S_n - n\mu}{\sqrt{n\,\sigma^2}} \to N(0,1) \text{ in distribution}$$

  - proof
    - weaker proof：我们先假设MGF都存在的情况下，这个使用用MGF来证明

$$\mathbb{E}\left[e^{\frac{S_n - n\mu}{\sqrt{n\sigma^2}}t}\right] = e^{-\frac{n\mu t}{\sqrt{n\sigma^2}}}\,\mathbb{E}\left[e^{\frac{S_n t}{\sqrt{n\sigma^2}}}\right]$$

$$= e^{-\frac{\sqrt{n}\mu t}{\sigma}} \prod_{i=1}^{n} \mathbb{E}\left[e^{\frac{X_i t}{\sqrt{n\sigma^2}}}\right]$$

$$= e^{-\frac{\sqrt{n}\mu t}{\sigma}} \prod_{i=1}^{n} M_{X_i}\left(\frac{t}{\sqrt{n\sigma^2}}\right)$$

$$= e^{-\frac{\sqrt{n}\mu t}{\sigma}} \cdot e^{\ln M_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right)}$$

$$= e^{-\frac{\sqrt{n}\mu t}{\sigma} + n \ln M_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right)}$$

$$M_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right) = M_{X_1}(0) + M'_{X_1}(0)\cdot\frac{t}{\sqrt{n\sigma^2}} + \frac{M''_{X_1}(0)}{2!}\left(\frac{t}{\sqrt{n\sigma^2}}\right)^2 + o(n^{-\frac{3}{2}})$$

$$= 1 + \frac{\mu t}{\sqrt{n\sigma^2}} + \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}})$$

$$\ln M_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right) = \ln\Big(1 + \underbrace{\frac{\mu t}{\sqrt{n\sigma^2}} + \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}})}_{x}\Big) \qquad x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

$$= \frac{\mu t}{\sqrt{n\sigma^2}} + \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}}) - \frac{1}{2}\Big(\frac{\mu t}{\sqrt{n\sigma^2}} + \underbrace{\frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}})}_{o(n^{-\frac{3}{2}})}\Big)^2$$

$$-\frac{n\mu t}{\sigma\sqrt{n}} + n\ln M_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right) = \cancel{-\frac{n\mu t}{\sigma\sqrt{n}}} + \cancel{\frac{n\mu t}{\sqrt{n\sigma^2}}} + \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{\sigma^2} + o(n^{-\frac{1}{2}}) - \frac{1}{2}\cdot\frac{\mu^2 t^2}{\sigma^2}$$

$$= \frac{\sigma^2 t^2}{2\sigma^2} = \frac{t^2}{2} \Rightarrow \text{normal distribution.}$$

    - true proof: 需要注意的是MGF并不一定一直存在，所以我们考虑更general 的情况，用CF来证明

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + o(t^2) = 1 + it\mathbb{E}[X] - \frac{t^2}{2}\mathbb{E}[X^2] + o(t^2).$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

$$\varphi(t) = \mathbb{E}\left[e^{it\cdot\frac{S_n - n\mu}{\sqrt{n\sigma^2}}}\right] = e^{-\frac{itn\mu}{\sqrt{n\sigma^2}}}\,\mathbb{E}\left[e^{\frac{itS_n}{\sqrt{n\sigma^2}}}\right] = e^{-\frac{itn\mu}{\sqrt{n\sigma^2}}}\cdot\prod\mathbb{E}\left[e^{\frac{itX_i}{\sqrt{n\sigma^2}}}\right]$$

$$= e^{-\frac{itn\mu}{\sqrt{n\sigma^2}}}\prod\varphi\left(\frac{t}{\sqrt{n\sigma^2}}\right) = e^{-\frac{itn\mu}{\sqrt{n\sigma^2}} + n\ln\varphi\left(\frac{t}{\sqrt{n\sigma^2}}\right)}$$

$$\varphi\left(\frac{t}{\sqrt{n\sigma^2}}\right) = 1 + \underbrace{i\mu\cdot\frac{t}{\sqrt{n\sigma^2}} - \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}})}_{x}$$

$$-\frac{itn\mu}{\sqrt{n\sigma^2}} + n\ln\varphi\left(\frac{t}{\sqrt{n\sigma^2}}\right) = \cancel{\frac{itn\mu}{\sqrt{n\sigma^2}}} + n\Big[\cancel{\frac{i\mu t}{\sqrt{n\sigma^2}}} - \frac{\mu^2+\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2} + o(n^{-\frac{3}{2}})\Big) - \frac{1}{2}\Big(\frac{i\mu t}{\sqrt{n\sigma^2}} - \frac{\mu^2+\sigma^2}{2}\frac{t^2}{n\sigma^2} + o(t^2)\Big)^2$$

$$= -\frac{\mu^2+\sigma^2}{2\sigma^2}t + o(n^{-\frac{1}{2}}) + \frac{1}{2}\cdot\frac{\mu^2 t^2}{\sigma^2} + o(n^{-\frac{1}{2}}) \qquad + o(n^{-\frac{3}{2}})$$

$$= -\frac{t^2}{2}$$

$$\therefore \to N(0,1)$$

  - Extension

- triangular arrays

  > ■ **Theorem 4 (Central Limit Theorem for triangular arrays)**
  > For each positive integer n, we let $X_{n,1}, X_{n,2}, \cdots, X_{n,n}$ be n independent random variables with $\mathbb{E}[X_{n,m}] = 0$. Suppose that
  >
  > 不要求同分布
  >
  > $$\sum_{m=1}^{n} Var(X_{n,m}) = \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2] \to \sigma^2 > 0 \quad \text{as } n \to \infty$$
  >
  > ⇒期望相加且为0
  >
  > $+\ \sigma^2 \to \sigma^2$ 注意是趋于σ^2
  >
  > For any $\varepsilon > 0$, we have
  >
  > $$\sum_{m=1}^{n} \mathbb{E}\left[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}\right] \to 0 \quad \text{as } n \to \infty$$
  >
  > $+\sum \mathbb{E}[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}] \to 0$
  >
  > Then $S_n = X_{n,1} + X_{n,2} + \cdots + X_{n,n} \to \sigma Z$ in distribution, where Z has standard normal distribution.

  - 在之前的CLT中，我们的假设是每个X_i都是iid的，但这个extension 并不要求一定是同分布的，仅仅是独立，满足其他一定的条件也可以服从标准正态分布
  - 有一个问题是这里必须是以分布收敛
  - example

    > We let $X_1, X_2, \cdots$ be a sequence of independent random variables such that $|X_i| \leq M$ and $\sum_{n=1}^{\infty} Var(X_n) = +\infty$ .We define $S_n = X_1 + X_2 + \cdots + X_n$ .Show that
    >
    > $$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var(S_n)}} \to^d Z \sim N(0,1)$$

    - proof

      $$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var(S_n)}} = \sum_{m=1}^{n} \frac{X_m - \mathbb{E}[X_m]}{\sqrt{Var(S_n)}} = \sum_{m=1}^{n} X_{n,m}$$

      $$\mathbb{E}[X_{n,m}] = 0. \quad \sum_{m=1}^{n} Var(X_{n,m}) = \sum_{m=1}^{n} Var\left(\frac{X_m - \mathbb{E}[X_m]}{\sqrt{Var(S_n)}}\right)$$

      $$= \sum_{m=1}^{n} \frac{1}{Var(S_n)} \cdot \left[Var(X_m - \mathbb{E}[X_m])\right)$$

      $$= \sum_{m=1}^{n} \frac{Var(X_m)}{Var(S_n)} = 1 = \sigma^2$$

      接下来证明

      $$\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}] \to 0$$

      $|X_i| \leq M$. $\sum_{n=1}^{\infty} Var(X_n) = +\infty$. $\therefore$ When n is large. $|X_m - \mathbb{E}[X_m]| < \varepsilon\sqrt{Var(S_n)}$

      $$X_{n,m} = \frac{X_m - \mathbb{E}[X_m]}{\sqrt{Var(S_n)}} < \varepsilon. \quad \therefore 1_{\{|X_{n,m}|>\varepsilon\}} \xrightarrow{n\to\infty} 0$$

      $$\therefore \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 1_{\{|X_{n,m}|>\varepsilon\}}] \xrightarrow{n\to\infty} 0.$$

      $\therefore$ 可以直接用 Thm 4.

- CLT for random sum