第五章、极限定理

§5.1 Bernoulli 试验场合的极限定理

- 假设$X = X_1, X_2, \cdots$ 独立同分布(i.i.d.),

$$P(X = 1) = p, \quad P(X = 0) = q = 1 - p.$$

$$S_n = X_1 + \cdots + X_n.$$

- $P(S_n = k) = C_n^k p^k q^{n-k}$, $k = 0, 1, \cdots, n$.

- 频率 $\xi_n = \frac{1}{n} S_n = \bar{X}$:

$$ES_n = np, \quad \mathrm{var}(S_n) = npq.$$

$$E\xi_n = p, \quad \mathrm{var}(\xi_n) = \frac{n \cdot \mathrm{var}(X)}{n} = \frac{1}{n}\mathrm{var}(X) = \frac{1}{n} pq.$$

- (1713年) Bernoulli 大数定律(5.1.14):

$$P\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2}\mathrm{var}\left(\frac{S_n}{n}\right) = \frac{1}{\varepsilon^2} \cdot \frac{\mathrm{var}(X)}{n} \to 0.$$

- 如果$\forall \varepsilon > 0$ 都有

$$\lim_{n\to\infty} P(|\xi_n - \xi| \geqslant \varepsilon) = 0.$$

那么, 称ξ_n 依概率收敛到ξ, 记为$\xi_n \xrightarrow{P} \xi$. (定义5.2.3)

- $S_n^* = \frac{S_n - np}{\sqrt{npq}}$: $x = \varepsilon/\sqrt{pq}$,

$$P\left(\left|\frac{S_n - np}{\sqrt{n}}\right| \geqslant \varepsilon\right) = P\left(|S_n^*| \geqslant x\right) \leqslant \frac{1}{x^2}, \quad \forall x > 0.$$

- (1733, 1778) De Moivre-Laplace 中心极限定理(定理5.1.1):

$$P(a < S_n^* \leqslant b) \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad \forall a < b.$$

- $x_k = \frac{k-np}{\sqrt{npq}}$, $\mathbb{K} = \left(np + a\sqrt{npq}, np + b\sqrt{npq}\right]$, $\Delta x = \frac{1}{\sqrt{npq}}$.

$$\frac{P(S_n = k)}{p_Z(x_k)\Delta x} = \frac{P(S_n = k)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \frac{1}{\sqrt{npq}}} \overset{\mathbb{K}}{\rightrightarrows} 1.$$

- 进一步,

$$\lim_n \frac{\sum_{k\in\mathbb{K}} P(S_n = k)}{\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx} = \lim_n \frac{\sum_{k\in\mathbb{K}} P(S_n = k)}{\sum_{k\in\mathbb{K}} p_Z(x_k)\Delta x} = 1.$$

- a 可为$-\infty$; b 可为∞.

$$P(S_n^* \leqslant x) \to \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \Phi(x), \quad \forall x \in \mathbb{R}.$$

- 假设$\xi \sim N(0,1)$. 如果$\forall x$ 都有

$$\lim_{n\to\infty} P(\xi_n \leqslant x) = P(\xi \leqslant x).$$

那么, 称ξ_n 依分布收敛于ξ, 记为$\xi_n \xrightarrow{d} \xi$.

§5.2 之依概率收敛， §5.1 & §5.3 之弱大数定律

- 如果$\forall \varepsilon > 0$ 都有

$$\lim_{n\to\infty} P(|\xi_n - \xi| \geqslant \varepsilon) = 0.$$

那么，称ξ_n 依概率收敛到ξ，记为$\xi_n \xrightarrow{P} \xi$． (定义5.2.3)

- 如果

$$\lim_{n\to\infty} E|\xi_n - \xi|^r = 0.$$

那么，称ξ_n r 阶(平均)收敛到ξ，记为$\xi_n \xrightarrow{r} \xi$． (定义5.2.4)

- $\xi_n \xrightarrow{r} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$. (定理5.2.8)

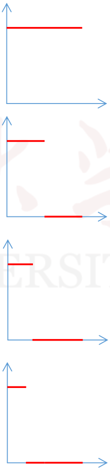
$$P(|\xi_n - \xi| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^r} E|\xi_n - \xi|^r.$$

- "$\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{r} \xi$" 不成立. (例5.2.4)

  $\Omega = (0,1]$, 几何概型.

  $\xi_n = n^{\frac{1}{r}} \times 1_{(0,\frac{1}{n}]}$,

  $\xi \equiv 0$.



- $\xi_n \xrightarrow{P} \xi$ & $E|\xi_n|^r \to E|\xi|^r \Rightarrow \xi_n \xrightarrow{r} \xi$.

  (证明不要求)

- 弱大数定律(Weak Law of Large Numbers, WLLN)形式:
  若$X_1, X_2, \cdots$ 满足 ***, 则

$$\frac{S_n}{b_n} - a_n \xrightarrow{P} 0.$$

- WLLN 的证明方法:

$$P\left(\left|\frac{S_n}{b_n} - a_n\right| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^r} E\left|\frac{S_n}{b_n} - a_n\right|^r, \quad r > 0.$$

- Chebyshev's WLLN: 两两不相关, $\mathrm{var}(X_i) \leqslant M$, $\forall i$, 则

$$\frac{S_n - ES_n}{n} \xrightarrow{P} 0. \quad (5.1.12)$$

- Markov's WLLN: $\mathrm{var}(S_n) = o(n^2)$, 则

$$\frac{S_n - ES_n}{n} \xrightarrow{P} 0. \quad (5.1.12)$$

- 又若 $\frac{ES_n}{n} \to a$, 则 $\frac{S_n}{n} \xrightarrow{P} a$.
- 习题五、26(4).

$$\xi_n \xrightarrow{P} \xi,\ \eta_n \xrightarrow{P} \eta \quad \Rightarrow \quad \xi_n + \eta_n \xrightarrow{P} \xi + \eta.$$

证明: $\forall \varepsilon > 0$,

$$\begin{aligned} & P\Big(\big|(\xi_n + \eta_n) - (\xi + \eta)\big| \geqslant \varepsilon\Big) \\ \leqslant & P\left(|\xi_n - \xi| \geqslant \frac{\varepsilon}{2}\right) + P\left(|\eta_n - \eta| \geqslant \frac{\varepsilon}{2}\right). \end{aligned}$$

- Bernoulli's WLLN (5.1.14), Poisson's WLLN (5.1.16).

习题五、44. $f:[0,1]\to\mathbb{R}$ 连续. 则存在多项式f_n 使得$f_n \overset{[0,1]}{\rightrightarrows} f$.

- 设$X_1, X_2, \cdots$ i.i.d., $P(X_1=1)=1-P(X_1=0)=x$. 则

$$\text{LLN}: \frac{S_n}{n}\to x \Rightarrow f\left(\frac{S_n}{n}\right)\approx f(x)(\text{直观}).$$

- 取f_n:

$$f_n(x)=Ef\left(\frac{S_n}{n}\right)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)C_n^k x^k(1-x)^{n-k}.$$

- $\forall\varepsilon>0$, $\exists\delta$ 使得

$$\text{若 } |y-z|<\delta, \quad \text{则 } |f(y)-f(z)|\leqslant \frac{1}{2}\varepsilon.$$

- 令$A_n = \{|\frac{S_n}{n} - x| \geqslant \delta\} = \{|S_n - nx| \geqslant n\delta\}$. 则

$$f_n(x) - f(x) = E\left(f\left(\frac{S_n}{n}\right) - f(x)\right)\left(1_{A_n} + 1_{A_n^c}\right).$$

- 取$C = 2\max\limits_{x\in[0,1]}|f(x)|$, 则

$$|f_n(x) - f(x)| \leqslant CP(A_n) + \frac{1}{2}\varepsilon.$$

- $\sigma^2 = x(1-x) \leqslant \frac{1}{4}$: 当$n \geqslant N$ 时,

$$\begin{aligned}|f_n(x) - f(x)| &\leqslant C\frac{n\sigma^2}{n^2\delta^2} + \frac{1}{2}\varepsilon \\ &\leqslant \frac{C}{4n\delta^2} + \frac{1}{2}\varepsilon \leqslant \varepsilon, \quad \forall x \in [0,1].\end{aligned}$$

有界收敛定理： 假设$\xi_n \xrightarrow{P} \xi$, 且$P(|\xi_n| \leqslant M) = 1$, $\forall n$. 那么,

$$\lim_{n\to\infty} E\xi_n = E\xi.$$

- 记$A_n = \{|\xi_n - \xi| > \varepsilon\}$.

$$\begin{aligned} E|\xi_n - \xi| &= E|\xi_n - \xi| \cdot 1_{A_n} + E|\xi_n - \xi| \cdot 1_{A_n^c} \\ &\leqslant 2M \cdot P(A_n) + \varepsilon. \end{aligned}$$

- $\xi_n \xrightarrow{1} \xi$:

$$\limsup_{n\to\infty} E|\xi_n - \xi| \leqslant \varepsilon, \quad \forall \varepsilon > 0.$$

- $|E\xi_n - E\xi| \leqslant E|\xi_n - \xi| \to 0.$

定理5.3.1. 设$X = X_1, X_2, \cdots$ i.i.d., $E|X| < \infty$, 则$\frac{S_n}{n} \xrightarrow{P} EX$.

- $E|X| < \infty \Rightarrow \lim_{x\to\infty} xP(|X| > x) = 0$.

$$\frac{x}{2}P(|X| > x) \leqslant \int_{x/2}^{x} P(|X| > y)dy \xrightarrow{x\to\infty} 0.$$

- $\frac{S_n}{n} - EX \cdot 1_{\{|X|\leqslant n\}} \xrightarrow{P} 0$. (见后面例*)
- $\mu_n = EX \cdot 1_{\{|X|\leqslant n\}} \to EX$:

$$\begin{aligned} E|X| \cdot 1_{\{|X|>n\}} &= \int_0^\infty P(|X| \cdot 1_{\{|X|>n\}} > y)dy \\ &= \int_0^\infty P(|X| > y \vee n)dy \\ &= \int_0^n P(|X| > n)dy + \int_n^\infty P(|X| > y)dy \to 0. \end{aligned}$$

例*. 假设$X_1, X_2, \cdots$ i.i.d., $\lim_{x\to\infty} xP(|X|>x)=0$, 则

$$\frac{S_n}{n} - EX\cdot 1_{\{|X|\leqslant n\}} \xrightarrow{P} 0.$$

- 截断:

$$T_n = X_1\cdot 1_{\{|X_1|\leqslant n\}} + \cdots + X_n\cdot 1_{\{|X_n|\leqslant n\}}.$$

则, $P(S_n\neq T_n) \leqslant nP\left(X\neq X\cdot 1_{\{|X|\leqslant n\}}\right) = nP(|X|>n) \to 0.$

- 故,

$$P\left(\left|\frac{S_n}{n}-\mu_n\right|\geqslant\varepsilon\right) \leqslant P(S_n\neq T_n) + P\left(\left|\frac{T_n}{n}-\mu_n\right|\geqslant\varepsilon\right).$$

- $\operatorname{var}(X\cdot 1_{\{|X|\leqslant n\}})\leqslant EX^2\cdot 1_{\{|X|\leqslant n\}}=o(n)$:

$$
\begin{aligned}
EX^2\cdot 1_{\{|X|\leqslant n\}} &= \int_0^\infty P\left(X^2 1_{\{|X|\leqslant n\}} > x\right)dx \\
&= \int_0^{n^2} P\left(\sqrt{x} < |X|\leqslant n\right)dx \\
&\leqslant \int_0^n \underset{\sim\sim\sim\sim\sim\sim\sim}{P\left(y < |X|\right)}\cdot 2y\,dy.
\end{aligned}
$$

- $P(|\frac{T_n}{n}-\mu_n|\geqslant\varepsilon)\to 0$:

$$
P\left(\left|\frac{T_n}{n}-\mu_n\right|\geqslant\varepsilon\right)\leqslant\frac{1}{\varepsilon^2}\frac{\operatorname{var}(T_n)}{n^2}=\frac{1}{\varepsilon^2}\frac{\operatorname{var}(X\cdot 1_{\{|X|\leqslant n\}})}{n}\to 0.
$$

例*(习题一38, 习题四6). 假设有n 种券, 集齐时间为S_n, 则

$$\frac{S_n}{n\ln n} \xrightarrow{P} 1.$$

- $S_n = X_1 + \cdots + X_n$:

$$X_k \sim G(p_k), \quad p_k = \frac{n-(k-1)}{n}.$$

- 若$X \sim G(p)$, 则

$$EX = \frac{1}{p}, \quad \mathrm{var}(X) = \frac{1-p}{p^2} \leqslant \frac{1}{p^2}.$$

- 于是,

$$ES_n = n\sum_{\ell=1}^{n}\frac{1}{\ell} \approx n\ln n, \quad \mathrm{var}(S_n) \leqslant n^2\sum_{\ell=1}^{n}\frac{1}{\ell^2} \leqslant Cn^2.$$

- $\frac{ES_n}{n\ln n} \to 1$, $\mathrm{var}(S_n) \leqslant Cn^2$:

$$P\left(\left|\frac{S_n - ES_n}{n\ln n}\right| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{var}(S_n)}{\varepsilon^2 n^2 (\ln n)^2} \to 0.$$

- $\frac{S_n - ES_n}{n\ln n} \xrightarrow{P} 0$, 故

$$\frac{S_n}{n\ln n} \xrightarrow{P} 1.$$

- 集齐一半的时间: $T_n = X_1 + \cdots + X_{n/2}$.

$$ET_n \approx n \sum_{\ell=n/2}^{n} \frac{1}{\ell} \approx n\ln 2,\ \mathrm{var}(T_n) \leqslant n^2 \sum_{\ell=n/2}^{n} \frac{1}{\ell^2} \leqslant \delta_n n^2.$$

故,$P\left(\left|\frac{T_n - ET_n}{n\ln 2}\right| > \varepsilon\right) \leqslant \frac{\delta_n n^2}{\varepsilon n^2 (\ln 2)^2} \to 0, \quad \frac{T_n}{n\ln 2} \xrightarrow{P} 1.$

§5.2 之几乎必然收敛、§5.4 强大数定律

- 如果

$$P\left(\underset{\sim\sim\sim\sim\sim\sim\sim\sim\sim\sim\sim\sim}{\lim_{n\to\infty}\xi_n = \xi}\right) = 1.$$

那么, 称ξ_n 几乎必然收敛到ξ, 记为$\xi_n \xrightarrow{\text{a.s.}} \xi$. (定义5.2.5)

- $\underset{\sim}{A} = \{\omega : \lim_{n\to\infty}\xi_n(\omega) = \xi(\omega)\}$.
- $\lim_{n\to\infty}\xi_n(\omega) = \xi(\omega)$ 不成立:

$$\exists \varepsilon > 0,\ \forall N \geqslant 1,\ \exists n \geqslant N \text{ 使得 } |\xi_n(\omega) - \xi(\omega)| > \varepsilon.$$

$$\exists k \geqslant 1,\ \forall N \geqslant 1,\ \exists n \geqslant N \text{ 使得 } |\xi_n(\omega) - \xi(\omega)| > \frac{1}{k}.$$

- 令 $A_{n,\varepsilon} = \{|\xi_n - \xi| > \varepsilon\}$, 则

$$\left\{\lim_{n\to\infty} \xi_n = \xi\right\}^c = \bigcup_{\varepsilon>0} \bigcap_{N\geqslant 1} \bigcup_{n\geqslant N} A_{n,\varepsilon} = \bigcup_{k\geqslant 1} \bigcap_{N\geqslant 1} \bigcup_{n\geqslant N} A_{n,\frac{1}{k}}.$$

- 当 $\varepsilon \searrow 0$ 时,

$$A_\varepsilon := \bigcap_{N\geqslant 1} \bigcup_{n\geqslant N} A_{n,\varepsilon} \quad \nearrow,$$

$$\star\star = \bigcup_{\varepsilon>0} A_\varepsilon = \bigcup_{k\geqslant 1} A_{\frac{1}{k}} = \lim_{k\to\infty} A_{\frac{1}{k}}.$$

- 结论:

$$\xi_n \xrightarrow{\text{a.s.}} \xi \text{ iff } P(A_\varepsilon) = 0, \forall \varepsilon > 0 \text{ iff } P\left(A_{\frac{1}{k}}\right) = 0, \forall k \geqslant 1.$$

- 对任意事件列$A_1, A_2, \cdots$, 令

$$\{A_n \text{ i.o.}\} = \limsup_{n\to\infty} A_n := \bigcap_{N\geqslant 1}\bigcup_{n\geqslant N} A_n.$$

- Borel-Cantelli引理(引理5.4.1): 令$s = \sum_{n=1}^{\infty} P(A_n)$.

(1) 若$s < \infty$, 则$P(A_n \text{ i.o.}) = 0$.

(2) 若$s = \infty$, 且$A_1, A_2, \cdots$ 相互独立, 则$P(A_n \text{ i.o.}) = 1$.

- $\bigcup_{n\geqslant N} A_n \searrow$: $(1 - p \leqslant e^{-p},)$

$$P(A_n \text{ i.o.}) = \lim_{N\to\infty} P\left(\bigcup_{n\geqslant N} A_n\right).$$

$$\prod_{n\geqslant N}(1 - P(A_n)) \leqslant e^{-\sum_{n\geqslant N} P(A_n)} = 0.$$

总结:

- $\xi_n \xrightarrow{\text{a.s.}} \xi$ 的定义及等价条件:

$$P\left(\left\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\}\right) = 1. \quad \text{(定义、应用、验证)}$$

$$P\left(A_{n,\varepsilon} \text{ i.o.}\right) = 0, \quad \forall \varepsilon > 0; \quad \text{(验证)}$$

$$P\left(A_{n,\frac{1}{k}} \text{ i.o.}\right) = 0, \quad \forall k \geqslant 1;$$

$$P\left(\bigcup_{n\geqslant N} A_{n,\varepsilon}\right) \searrow 0 (N \to \infty), \quad \forall \varepsilon > 0.$$

其中, $A_{n,\varepsilon} = \{|\xi_n - \xi| > \varepsilon\}$.

- $\xi_n \xrightarrow{\text{a.s.}} \xi$ 的充分条件:

$$\sum_{n=1}^{\infty} P(A_{n,\varepsilon}) < \infty, \quad \forall \varepsilon > 0.$$

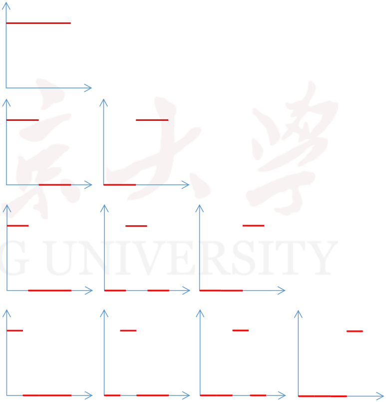
- $\xi_n \xrightarrow{\text{a.s.}} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$. (定理5.4.1)

$$P\left(\bigcup_{n\geqslant N} A_{n,\varepsilon}\right) \geqslant P(A_{N,\varepsilon}),$$

$$A_{n,\varepsilon} = \{|\xi_n - \xi| > \varepsilon\}.$$

- “$\xi_n \xrightarrow{P\text{或}r} \xi \Rightarrow \xi_n \xrightarrow{\text{a.s.}} \xi$”不成立.
  (例5.4.1)

- “$\xi_n \xrightarrow{\text{a.s.}} \xi \Rightarrow \xi_n \xrightarrow{r} \xi$” 不成立.
  (例5.2.4)



- 若$\xi_n \xrightarrow{P} \xi$, 则存在子列$\{n_k\}$ 使得$\xi_{n_k} \xrightarrow{\text{a.s.}} \xi$.

(1) $\exists n_1 \geqslant 1$ 使得

$$P(|\xi_n - \xi| > 1) < \frac{1}{2}, \quad \forall n \geqslant n_1;$$

(2) $\exists n_m > n_{m-1}$ 使得

$$P\left(|\xi_n - \xi| > \frac{1}{m}\right) < \frac{1}{2^m}, \quad \forall n \geqslant n_m.$$

(3) $\forall \varepsilon, \exists m$ 使得$\frac{1}{m} < \varepsilon$. 当$k \geqslant m$ 时, $\frac{1}{k} < \varepsilon$, $n_k \geqslant n_m$,

$$P(|\xi_{n_k} - \xi| > \varepsilon) \leqslant P\left(|\xi_{n_k} - \xi| > \frac{1}{k}\right) < \frac{1}{2^k}.$$

(4) $\sum_k P(A_{k,\varepsilon}) < \infty \Rightarrow P(A_{k,\varepsilon}, \text{ i.o.}) = 0$.

- $\xi_n \xrightarrow{P} \xi$ iff $\forall \{n_k\}, \exists \{n_{k_i}\}$ 使得$\xi_{n_{k_i}} \xrightarrow{\text{a.s.}} \xi$.
  $\Longrightarrow$: 已证
- $\Longleftarrow$: 反证法: 否则,

$$\exists \varepsilon, \delta > 0,\ n_1 < n_2 < \cdots \text{ 使得 } P\left(|\xi_{n_k} - \xi| > \varepsilon\right) > \delta.$$

- 对于$\{\xi_{n_k}, k \geqslant 1\}$, 不存在收敛子列:
  否则$\xi_{n_{k_i}} \xrightarrow{\text{a.s.}} \xi$,
  于是$\xi_{n_{k_i}} \xrightarrow{P} \xi$. 与$P\left(|\xi_{n_{k_i}} - \xi| > \varepsilon\right) > \delta$ 矛盾!
- 总结: $\xi_n \xrightarrow{\text{a.s.,r}} \eta \Rightarrow \xi_n \xrightarrow{P} \eta$, 反之不成立.

强大数定律(Strong Law of Large Numbers, SLLN)形式:

若$X_1, X_2, \cdots$ 满足 $* * *$, 则

$$\frac{S_n}{b_n} - a_n \xrightarrow{\text{a.s.}} 0. \quad (5.4.28)$$

定理 ((Borel-)Cantelli's SLLN, 定理5.4.2, 习题五、46)

若$X_1, X_2, \cdots$ 相互独立, $E(X_i - EX_i)^4 \leqslant M$, $\forall i$, 则

$$\frac{S_n - ES_n}{n} \xrightarrow{\text{a.s.}} 0.$$

- 不妨设$EX_i = 0$. $A_n = \{|\frac{S_n}{n}| \geqslant \varepsilon\}$. 往证$\sum_n P(A_n) < \infty$.
- 将证$ES_n^4 \leqslant 3n^2M$, $\forall n$. 于是,

$$P(A_n) = P\left(S_n^4 \geqslant (n\varepsilon)^4\right) \leqslant \frac{1}{n^4\varepsilon^4} ES_n^4 \leqslant \frac{3M}{\varepsilon^4} \cdot \frac{1}{n^2}.$$

(1) $ES_n^4 = \sum_{i,j,k,\ell=1}^n EX_iX_jX_kX_\ell$,

$$EX_r^4,\ EX_r^3X_s,\ EX_r^2X_s^2,\ EX_r^2X_sX_t,\ EX_rX_sX_tX_u.$$

(2) $E\star = 0$, $EX_r^4 \leqslant M$,

$$EX_r^2X_s^2 = EX_r^2 \times EX_s^2 \leqslant \sqrt{EX_r^4}\sqrt{EX_s^4} \leqslant M.$$

(3) 因此,

$$ES_n^4 \leqslant nM + C_n^2C_4^2M \leqslant 3n^2M.$$

定理

若$X = X_1, X_2, \cdots$ 独立同分布, $EX^2 < \infty$, 则

$$\frac{1}{n} S_n \xrightarrow{\text{a.s.}} EX.$$

- 不妨设$EX = 0$, $EX^2 = 1$. 令$A_n = \{|\frac{S_n}{n}| > \varepsilon\}$. 则

$$P(A_n) \leqslant \frac{1}{\varepsilon^2} E\left(\frac{S_n}{n}\right)^2 = \frac{1}{\varepsilon^2} \frac{nEX^2}{n^2} = \frac{1}{\varepsilon^2} \cdot \frac{1}{n}.$$

- $\frac{S_{m^2}}{m^2} \xrightarrow{\text{a.s.}} 0$:

$$\sum_m P(A_{m^2}) < \infty \Rightarrow P(A_{m^2} \text{ i.o.}) = 1.$$

- $|S_n| \leqslant |S_{m^2}| + T_m, \quad \forall m^2 \leqslant n < (m+1)^2 = m^2 + 2m + 1.$

$$T_m := \max_{1\leqslant k\leqslant 2m} |S_k^{(m)}|,$$

$$\text{其中，} S_k^{(m)} = X_{m^2+1} + \cdots + X_{m^2+k}.$$

- $\frac{T_m}{m^2} \xrightarrow{\text{a.s.}} 0$. 于是 $\frac{S_n}{n} \xrightarrow{\text{a.s.}} 0$.

$$P\left(\left|\frac{S_k^{(m)}}{m^2}\right| > \varepsilon\right) \leqslant \frac{1}{\varepsilon^2}\frac{k}{m^4},$$

$$\Rightarrow P\left(\frac{|T_m|}{m^2} > \varepsilon\right) \leqslant 2m \cdot \frac{1}{\varepsilon^2}\frac{2m}{m^4} = \frac{4}{\varepsilon^2} \cdot \frac{1}{m^2}.$$

定理 (Kolmogorov's SLLN, 定理5.4.4)

假设$X = X_1, X_2, \cdots$ 独立同分布, $E|X| < \infty$, 则

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} EX.$$

- 时间平均= 空间平均, (期望的含义).
- 若$X_1, X_2, \cdots$ 独立, 且$\frac{S_n}{n} \xrightarrow{\text{a.s.}} Y$, 则$Y$ 退化.
- 若i.i.d. 且$\frac{S_n}{n} \xrightarrow{\text{a.s.}} a \in \mathbb{R}$, 则$a = EX$.

(1) 在$A = \{\lim_n \frac{S_n}{n}\ \exists\}$ 上,

$$\frac{S_{n-1}}{n-1} - \frac{S_n}{n} = \frac{S_{n-1}}{(n-1)n} - \frac{X_n}{n}, \quad \text{故} \ \frac{X_n}{n} \to 0.$$

(2) 习题五、45. 假设i.i.d. 则

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{iff} \quad E|X| < \infty.$$

定理

假设$X = X_1, X_2, \cdots$ 独立同分布, $EX = \infty$, 则

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} \infty.$$

- $X = X^+ - X^-$. 不妨设$X \geqslant 0$.
- 给定$M > 0$, 令$T_n = X_1 \wedge M + \cdots + X_n \wedge M$.

$$\frac{T_n}{n} \xrightarrow{\text{a.s.}} E(X \wedge M).$$

- $S_n \geqslant T_n$: $\forall M > 0$,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geqslant E(X \wedge M), \quad \text{a.s.}$$

即, $\Omega_M = \{\star\star\}$, 则$P(\Omega_M) = 1$.

- $P(\Omega_M) = 1$. 令$\hat{\Omega} = \bigcap_{M=1}^{\infty} \Omega_M$. 则$P(\hat{\Omega}) = 1$, 且在$\hat{\Omega}$ 上,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geqslant E(X \wedge M), \quad \forall M.$$

- $E(X \wedge M) \to \infty$:

$$\begin{aligned} E(X \wedge M) &= \int_0^{\infty} P(X \wedge M > x)dx \\ &= \int_0^{M} P(X > x)dx \to \infty. \end{aligned}$$

i.i.d. 序列总结:

- EX 有意义, 则

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} EX.$$

- $EX^{\pm} = \infty$: 柯西分布(例3.3.5),

$$\frac{S_n}{n} \stackrel{d}{=} X_1.$$

例5.3.1. 样本均值& 样本方差.

- 数据: $X_1, X_2, \cdots$ 独立同分布.
- 样本均值:

$$\bar{X} = \frac{1}{n} S_n \overset{\text{a.s.}}{\to} EX.$$

- 样本方差: $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \approx$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

$$\overset{\text{a.s.}}{\to} EX^2 - (EX)^2 = \text{var}(X).$$

例5.3.2. $\int_0^1 f(x)dx = Ef(U)$.

- SLLN:

$$\frac{1}{n}\sum_{i=1}^{n} f(U_i) \xrightarrow{\text{a.s.}} \int_0^1 f(x)dx.$$

- 高维:

$$\int_0^1 \cdots \int_0^1 f(x_1, \cdots, x_n)d\vec{x} = Ef(U_1, \cdots, U_n).$$

习题五、50. 假设$f, g : [0,1] \to \mathbb{R}$, $0 \leqslant f < Cg$. 求

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \frac{f(x_1)+\cdots+f(x_n)}{g(x_1)+\cdots+g(x_n)} dx_1 \cdots dx_n.$$

- 记$X_i = f(U_i)$, $Y_i = g(U_i)$; 部分和: S_n, T_n. 令

$$W_n = \frac{S_n}{T_n}. \quad \text{再次强调: } EW_n \neq \frac{ES_n}{ET_n}.$$

- SLLN:

$$W_n = \frac{S_n/n}{T_n/n} \xrightarrow{\text{a.s.}} \frac{EX}{EY} = \frac{\int_0^1 f(x)dx}{\int_0^1 g(x)dx} = w.$$

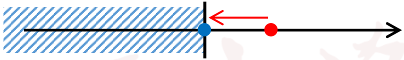
- 有界收敛定理:

$$EW_n \to Ew = w.$$

§5.2 之依分布收敛， §5.3 之中心极限定理

- 权重的收敛.
- 例5.2.1: $\xi_n = \frac{1}{n}$, $\xi = 0$.

  $F_{\xi_n}(0) = 0, \quad F_\xi(0) = 1.$



- 如果$\forall x \in C(F_\xi)$ 都有

$$\lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x).$$

  那么, 称ξ_n 依分布收敛于ξ, 记为$\xi_n \xrightarrow{d} \xi$. (定义5.2.2 & 5.2.1)
- $\mathbb{R}\backslash C(F_\xi)$ 可数.

依分布收敛与依概率收敛.

- $\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$. (定理5.2.6)

(1) $P(X \leqslant x) = P(X \leqslant x, Y \leqslant x) + P(X \leqslant x < Y)$:

$|P(X \leqslant x) - P(Y \leqslant x)| \leqslant P(X \leqslant x < Y) + P(Y \leqslant x < X)$.

(2) 若$|X - Y| < \varepsilon$ 且$X \leqslant x < Y$, 则$x < Y < x + \varepsilon$:

$|P(X \leqslant x) - P(Y \leqslant x)| \leqslant P(|X - Y| \geqslant \varepsilon) + P(|Y - x| < \varepsilon)$.

(3) $|F_{\xi_n}(x) - F_\xi(x)| \leqslant P(|\xi_n - \xi| \geqslant \varepsilon) + P(|\xi - x| < \varepsilon)$.

- “$\xi_n \xrightarrow{d} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$” 不成立. (例5.2.3)
- $\xi_n \xrightarrow{d} C \Rightarrow \xi_n \xrightarrow{P} C$. (定理5.2.7)

  例, $C = 0$:

$$P(|\xi_n| < \varepsilon) \leqslant F_{\xi_n}(\varepsilon) - F_{\xi_n}(-\varepsilon) \to F_{\xi}(\varepsilon) - F_{\xi}(-\varepsilon) = 1.$$

- $\xi_n \xrightarrow{d} \xi \Rightarrow \hat{\xi}_n \xrightarrow{\text{a.s.}} \hat{\xi}$, 其中

$$\hat{\xi}_n := F_{\xi_n}^{-1}(U) \overset{d}{=} \xi_n, \quad \hat{\xi} := F_{\xi}^{-1}(U) \overset{d}{=} \xi.$$

  (证明不要求)
- 有界收敛定理: ξ_n 有界, $\xrightarrow{d} \xi$, 则$E\xi_n \to E\xi$.

依分布收敛的等价条件.

$\xi_n \xrightarrow{d} \xi$ iff $Ef(\xi_n) \to Ef(\xi)$, $\forall f \in \mathbb{F}$.

- $\mathbb{F}_1$: $f = 1_{(-\infty,b]}$, 其中$b \in C(F_\xi)$. (定义)
- $\mathbb{F}_1'$: $f = 1_{(-\infty,b]}$, 其中$b \in C(F_\xi)$ 的某稠子集.
- $\mathbb{F}_2$: $f = 1_{(a,b]}$, 其中, $a, b \in C(F_\xi)$.
- $\mathbb{F}_2'$: 阶梯函数

$$f = \sum_{i=1}^{n} \lambda_i \cdot 1_{(a_i,b_i]},$$

其中, $a_i, b_i \in C(F_\xi)$, $1 \leqslant i \leqslant n$.

- $\mathbb{F}_3$: $f:\mathbb{R}\to\mathbb{R}$ 有界连续. $\quad\Rightarrow$: (假设$\xrightarrow{d}$, 定理5.2.2 & 5.2.3).

(1) 取a 使得$\pm a\in C(F_\xi)$,

$$P(|\xi|\geqslant a)\leqslant\varepsilon\Rightarrow P(|\xi_n|\geqslant a)\leqslant 2\varepsilon,\ \forall n\geqslant N.$$

(2) 取阶梯函数g:

$$|f(x)-g(x)|\leqslant\varepsilon,\ \forall|x|\leqslant a;\qquad g(x)=0,\ |x|>a.$$

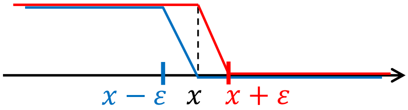
(3) $Ef(\xi_n)\cdot 1_{\{|\xi_n|\leqslant a\}}-Ef(\xi)\cdot 1_{\{|\xi|\leqslant a\}}=r_n+s_n+t_n$:

$$r_n=\star-Eg(\xi_n)\cdot 1_{\{|\xi_n|\leqslant a\}},\quad t_n=Eg(\xi)\cdot 1_{\{|\xi|\leqslant a\}}-\star,$$
$$s_n=Eg(\xi_n)\cdot 1_{\{|\xi_n|\leqslant a\}}-Eg(\xi)\cdot 1_{\{|\xi|\leqslant a\}}=Eg(\xi_n)-Eg(\xi).$$

(4) $|Ef(\xi_n)-Ef(\xi)|\leqslant 3M\varepsilon+\varepsilon+\varepsilon+|s_n|$.

- $\mathbb{F}_3$: $f:\mathbb{R}\to\mathbb{R}$ 有界连续. $\Leftarrow$:

(1) $g \leqslant 1_{(-\infty,x]} \leqslant f$.



(2) $1_{(-\infty,x]} \leqslant f$:

$$F_{\xi_n}(x) \leqslant Ef(\xi_n) \to Ef(\xi) \leqslant F_\xi(x+\varepsilon).$$

(3) $g \leqslant 1_{(-\infty,x]}$:

$$F_{\xi_n}(x) \geqslant Eg(\xi_n) \to Eg(\xi) \geqslant F_\xi(x-\varepsilon).$$

(4) $x \in C(F_\xi)$:

$$\lim_{\varepsilon\searrow 0} F_\xi(x+\varepsilon) = \lim_{\varepsilon\searrow 0} F_\xi(x-\varepsilon) = F_\xi(x).$$

- $\mathbb{F}_4$: f = 三角函数

$$\cos(tx), \quad \sin(tx), \quad t \in \mathbb{R}.$$

- 定理5.2.4 & 5.2.5 (证明不要求).

$$\xi_n \xrightarrow{d} \xi \quad \text{iff} \quad \text{特征函数收敛}: f_{\xi_n}(t) \to f_\xi(t),\ \forall t.$$

- 定理5.2.5, §4.5 连续性定理: 若

$$f_{\xi_n}(t) \to f(t),\ \forall t, \quad \text{且} f \text{ 在} t = 0 \text{ 连续}.$$

则, f 是特征函数, ξ_n 依分布收敛. (证明不要求)

习题五、35. 当$\delta_n := \max\limits_{1\leqslant k\leqslant n} p_k^{(n)} \to 0$, $\lambda_n := \sum_{k=1}^n p_k^{(n)} \to \lambda$ 时,

$$B(n,\vec{p})^{(n)} := B(1,p_1^{(n)}) * \cdots * B(1,p_n^{(n)}) \xrightarrow{d} P(\lambda).$$

- $B(n,\vec{p}^{(n)})$ 的特征函数:

$$f_n(t) = \prod_{k=1}^n \left((1-p_k^{(n)}) + p_k^{(n)} e^{it}\right)$$
$$= \prod_{k=1}^n (1+p_k^{(n)} z), \quad \text{其中} z = e^{it}-1.$$

- $P(\lambda_n) \to P(\lambda)$: 令$g_n(t)$ 表示$P(\lambda_n)$ 的特征函数, 则

$$g_n(t) = \prod_{k=1}^n e^{p_k^{(n)} z} = e^{\lambda_n z} \to e^{\lambda z}.$$

- $a_k = 1 + p_k^{(n)} z$, $b_k = e^{p_k^{(n)} z}$,

$$\|a_k\| = \|1 + p_k^{(n)} z\| = \|(1 - p_k^{(n)}) \cdot 1 + p_k^{(n)} \cdot e^{it}\| \leqslant 1.$$

$$\|b_k\| = \|e^{p_k^{(n)} z}\| = \|e^{p_k^{(n)}(\cos t - 1) + i p_k^{(n)} \sin t}\| = e^{p_k^{(n)}(\cos t - 1)} \leqslant 1.$$

- $f_n(t) = \prod_{k=1}^n a_k$, $g_n(t) = \prod_{k=1}^n b_k$, 故

$$\|f_n(t) - g_n(t)\| \leqslant \sum_{k=1}^n \|a_k - b_k\|. \quad (5.5.13)$$

- 归纳法:

$$\|\prod_{k=1}^n a_k - \prod_{k=1}^{n-1} a_k \times b_n\| \leqslant 1 \times \|a_n - b_n\|;$$

$$\|\prod_{k=1}^{n-1} a_k \times b_n - \prod_{k=1}^{n-1} b_k \times b_n\| \leqslant \|\prod_{k=1}^{n-1} a_k - \prod_{k=1}^{n-1} b_k\| \times 1.$$

- 记$s = p_k^{(n)} z$, 则$a_k = 1 + s$, $b_k = e^s$, 故$\|a_k - b_k\| \leqslant \|s\|^2$:

$$\|a_k - b_k\| = \|\sum_{\ell=2}^{\infty} \frac{s^\ell}{\ell!}\| \leqslant \sum_{\ell=2}^{\infty} \frac{\|s\|^\ell}{\ell!} \leqslant \sum_{\ell=2}^{\infty} \frac{\|s\|^2}{\ell!}$$
$$\leqslant \|s\|^2 (e-2) \leqslant (p_k^{(n)})^2 \|z\|^2.$$

- 注: 以上的估计并不精细.
- 故,

$$\|f_n(t) - g_n(t)\| \leqslant \sum_{k=1}^{n} \delta_n p_k^{(n)} \|z\|^2 \leqslant 4\delta_n \lambda_n \to 0.$$

定理5.3.1. 设$X=X_1,X_2,\cdots$ i.i.d., $E|X|<\infty$, 则$\frac{S_n}{n}\xrightarrow{P}EX$.

证明　方法II

- 不妨设$EX=0$. 只需证

$$f_n(t)=f_{\frac{S_n}{n}}(t)\to 1=e^{it\times 0}.$$

- 记$f=f_X$. 于是

$$f_n(t)=E\exp\left\{it\frac{X_1+\cdots+X_n}{n}\right\}=\left(Ee^{i\frac{t}{n}X}\right)^n=f\left(\frac{t}{n}\right)^n.$$

- 注: 根据有界收敛定理, $Ee^{i\frac{t}{n}X}\to e^{i\times 0}=1$.
往估计收敛速度.

- $e^{is} = 1 + is + \delta(s), \quad \|\delta(s)\| \leqslant \min\left\{2|s|, \frac{1}{2}s^2\right\}.$ (引理5.5.1)
- $f(\frac{t}{n})$:

$$E\left(1 + i\frac{t}{n}X + \delta\left(\frac{t}{n}X\right)\right) = 1 + 0 + E\delta\left(\frac{t}{n}X\right).$$

- $\|E\delta(\frac{t}{n}X)\| = o(\frac{1}{n})$:

$$\left\|E\delta\left(\frac{t}{n}X\right)\right\| = E\left\|\delta\left(\frac{t}{n}X\right)\right\| \cdot 1_{\{|X|>M\}} + E\left\|\delta\left(\frac{t}{n}X\right)\right\| \cdot 1_{\{|X|\leqslant M\}}$$
$$\leqslant \frac{2t}{n}E|X|1_{\{|X|>M\}} + \frac{t^2}{2n^2}M^2.$$

- $f(\frac{t}{n})^n = (1 + o(\frac{1}{n}))^n \to 1.$
- $\frac{1}{n}S_n \xrightarrow{d,P} 0.$

定理 (Lindeberg-Levy 中心极限定理, 定理5.4.4)

假设$X_1, X_2, \cdots$ 独立同分布. $0 < \mathrm{var}(X) < \infty$. 则

$$S_n^* \xrightarrow{d} Z \sim N(0,1).$$

- 不妨设$EX = 0$, $EX^2 = 1$. 记$f = f_{X_1}$, 则

$$f_{S_n^*}(t) = E\exp\left\{it\frac{S_n}{\sqrt{n}}\right\} = f\left(\frac{t}{\sqrt{n}}\right)^n.$$

- (记$s = \frac{t}{\sqrt{n}}$.) 往证

$$f(s) = Ee^{isX} = 1 - \frac{1}{2}s^2 + o(s^2).$$

$$\text{于是 } f_{S_n^*}(t) = (1 - \frac{t^2}{2n} + o(\frac{1}{n}))^n \to e^{-t^2/2}.$$

引理5.5.1. (记$y=sX$.)

$|e^{iy}-1|\leqslant|y|$, $|e^{iy}-1-iy|\leqslant\frac{|y|^2}{2}$ (5.5.10),

$|e^{iy}-1-iy+\frac{y^2}{2}|\leqslant\frac{|y|^3}{6}$ (5.5.11).

- $e^{iy}=1+iy+\frac{1}{2}(iy)^2+\varphi(y)$.

$$
\begin{aligned}
\varphi(y)=&e^{iy}-1-iy-\frac{1}{2}(iy)^2\\
=&\int_0^y i\underline{e^{iz}}dz-\int_0^y i\cdot\underline{1}dz-\frac{1}{2}(iy)^2\\
=&\int_0^y i\underline{\int_0^z ie^{iw}dwdz}-\int_0^y\int_0^z i^2dwdz\\
=&\int_0^y\int_0^z i^2(e^{iw}-1)dwdz=\int_0^y\int_0^z\int_0^w i^3e^{iu}dudwdz.
\end{aligned}
$$

- $\|\varphi(y)\|\leqslant y^2$; $\|\varphi(y)\|\leqslant\frac{|y|^3}{6}$:

$$\varphi(y)=\int_0^y\int_0^z i^2(e^{iw}-1)dwdz=\int_0^y\int_0^z\int_0^w i^3e^{iu}dudwdz.$$

- 取$y = sX$. 则

$$\begin{aligned} e^{isX} &= 1 + isX + \frac{1}{2}(isX)^2 + \varphi(sX), \\ f(s) = Ee^{isX} &= 1 + 0 - \frac{1}{2}s^2 + E\varphi(sX). \end{aligned}$$

- 由$\|\varphi(y)\| \leqslant y^2$, $\|\varphi(y)\| \leqslant \frac{|y|^3}{6}$ 可推出$\|E\varphi(sX)\| = o(s^2)$:

$$\begin{aligned} \|E\varphi(sX)\| &\leqslant E\|\varphi(sX)\| \cdot 1_{\{|X|>M\}} + E\|\varphi(sX)\| \cdot 1_{\{|X|\leqslant M\}} \\ &\leqslant E(sX)^2 \cdot 1_{\{|X|>M\}} + E\frac{|sX|^3}{6} \cdot 1_{\{|X|\leqslant M\}} \\ &\leqslant s^2 EX^2 \cdot 1_{\{|X|>M\}} + |s|^3 \cdot \frac{M^3}{6}. \end{aligned}$$

- 注:

$$f(s) = 1 - \frac{1}{2}s^2 + E\varphi(sX) = 1 - \frac{1}{2}s^2 + o(s^2).$$

- 固定t. 取$s=\frac{t}{\sqrt{n}}$, 令$n\to\infty$. 则

$$f(\frac{t}{\sqrt{n}})=1-\frac{t^2}{2n}+\varepsilon_n,$$

$$\text{其中, }\varepsilon_n:=E\varphi\left(\frac{t}{\sqrt{n}}X\right)=o(\frac{t}{\sqrt{n}})=o(\frac{1}{n}).$$

- $S_n^*\xrightarrow{d}Z\sim N(0,1)$:

$$f_{S_n^*}(t)=f\left(\frac{t}{\sqrt{n}}\right)^n=\left(1-\frac{t^2}{2n}+\varepsilon_n\right)^n\to e^{-\frac{t^2}{2}}.$$

- 中心极限定理(Central Limit Theorem, CLT)形式:

  若$X_1, X_2, \cdots$ 满足 $\underset{\sim\sim\sim}{***}$, 则

$$S_n^* \xrightarrow{d} Z \sim N(0,1). \qquad (5.1.11)$$

- Lindeberg-Levy 版本: $\underset{\sim\sim\sim}{***}$ = "i.i.d., $0 < \mathrm{var}(X) < \infty$".

- $\frac{1}{\sqrt{n}}(S_n - ES_n) \xrightarrow{d} \sigma Z \sim N(0, \sigma^2)$.

- 若$\vec{X} = \vec{X}_1, \vec{X}_2, \cdots$ i.i.d., $\vec{\mu} = E\vec{X}$, $\mathbf{\Sigma} = (\mathrm{cov}(X_i, X_j))$. 则

$$\frac{1}{\sqrt{n}}(\vec{S}_n - n\vec{\mu}) \xrightarrow{d} \sqrt{\mathbf{\Sigma}}\vec{Z} \sim N(\vec{0}, \mathbf{\Sigma}).$$

- Berry-Esseen's bound: 假设$E|X|^3 < \infty$. 那么,

$$|F_{S_n^*}(x) - \Phi(x)| \leqslant \frac{3E|X^*|^3}{\sqrt{n}}, \quad \forall x.$$

  (了解即可.)

CLT的应用.

例: 经验分布函数.

$$\text{SLLN:}\quad F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leqslant x\}} \xrightarrow{\text{a.s.}} F(x).$$

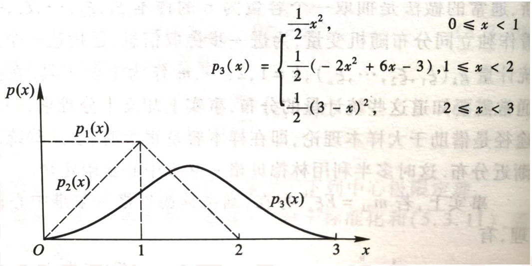
- CLT:

$$\sqrt{n}\Big(F_n(x) - F(x)\Big) \xrightarrow{d} \sqrt{F(x)\big(1 - F(x)\big)}Z$$
$$\sim N\Big(0, F(x)\big(1 - F(x)\big)\Big).$$

- (了解即可) $D_n = \max_x |F_n(x) - F(x)|$:

$$\sqrt{n}D_n \xrightarrow{d} \xi, \quad F_\xi(x) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2x^2}, \ \forall x > 0.$$

- 例5.3.4. $X_1, X_2, \cdots$ i.i.d., $\sim U(0,1)$.



- 将S_n^* "视为" $\sim N(0,1)$:

$$P(S_n \leqslant x) = P(S_n^* \leqslant x^*) \approx \Phi(x^*) = p, \quad x^* = \frac{x - ES_n}{\sqrt{\mathrm{var}(S_n)}}.$$

- n, x, p 满足关系式:

$$P(S_n \leqslant x) = p.$$

例5.3.5. 在近似计算中保留5位小数. 考虑10^4 个数, 近似值之和与真值之和的误差不超过10^{-3} 的概率.

- 近似值与真值的误差$X \sim U(-0.5\times 10^{-5}, 0.5\times 10^{-5})$.
- $n = 10^4$, $x = 10^{-3}$, 求p.
- 目标:

$$P(|S_n| \leqslant 10^{-3}) = P(|S_n^*| \leqslant x^*) = p,$$

其中, $x^* = \frac{10^{-3}}{\sqrt{n\cdot \mathrm{var}(X)}} = \frac{10^{-3}}{\sqrt{10^4 \cdot 10^{-5}\sqrt{\frac{1}{12}}}} = 2\sqrt{3} = 3.46\cdots$.

- 查表: $\Phi(3.46) = 99.97299\%$,

故$p \geqslant 2\Phi(3.46) - 1 = 99.94598\%$.

例2.4.4. 人寿保险. 死亡率$p=0.005$, 参保数$n=10^4$. 估算:

$$P(\text{死亡人数}=40),\ P(\text{死亡人数}\leqslant 70).$$

- $P(A)=C_{10000}^{40}p^{40}q^{9960}=\underline{0.0214}$.
- $x^*=\frac{x-np}{\sqrt{npq}}$:

$$P(A)\approx\varphi(40^*)\frac{1}{\sqrt{npq}}=\underline{0.0207}.$$

- 连续修正: $P(A)=P(39.5\leqslant S_n\leqslant 40.5)$,

  $x_1=39.5^*\approx -1.49$, $x_2=40.5^*\approx -1.35$:

$$P(A)=P(39.5\leqslant S_n\leqslant 40.5)=P(x_1\leqslant S_n^*\leqslant x_2)$$

$$\approx 0.93189-0.91149=\underline{0.0204}.$$

- 连续修正: $P(B)=P((-0.5)^*\leqslant S_n^*\leqslant(70.5)^*)\approx 0.998193$.
- 如果不连续修正: $P(0^*\leqslant S_n\leqslant 70^*)\approx 0.997744$.

例2.4.6. 车间有200台车床, 每台的功率为1 千瓦, 开动频率为60%. 要求正常生产的把握至少为99.9%, 问: 需多少千瓦电力?

- 已知$n = 200$, $p = 99.9\%$, 求x.
- $P(X = 1) = 1 - P(X = 0) = 0.6$. $n = 200$. $p = 99.9\%$.
- 目标: $x^* = \frac{x-200*0.6}{\sqrt{200*0.6*0.4}}$,

$$P(S_n \leqslant x) = P(S_n^* \leqslant x^*) \geqslant 99.9\% = p.$$

- 查表: $\Phi(3.09) = 99.8999\%$, $\Phi(3.1) = 99.90324\%$.
- $\star \geqslant 3.1$, $x \geqslant 120 + 3.1\sqrt{48} \in (141, 142)$. 需142 千瓦.

例5.1.2. 往常的市场占有率为15%. 为保证调查结果与真实值的误差不超过0.01 的概率至少为95%. 至少需调查多少人?

- 已知$x = 0.01$, $p = 0.95$, 求n.
- $P(X = 1) = 1 - P(X = 0) = q = 0.15$. 目标:

$$P\left(\left|\frac{S_n}{n} - q\right| \leqslant 0.01\right) \geqslant 0.95.$$

$$P\left(|S_n^*| \leqslant \frac{0.01\sqrt{n}}{\sqrt{q(1-q)}} = x^*\right) \geqslant 0.95.$$

- $\Phi(x^*) \geqslant \frac{1}{2}(1 + 0.95) = 0.975$, 查表得$x^* \geqslant 1.96$.
- $n \geqslant 196^2 q(1-q)$, 至少调查4899 人.
- 若q 未知, 则$n \geqslant 196^2 \max\limits_q q(1-q) = 196^2 \cdot \frac{1}{4} = 9604$.

§5.5 中心极限定理

- 研究对象:

$$S_n^* = \sum_{k=1}^{n} \frac{X_k - \mu}{\sqrt{\mathrm{var}(S_n)}} = \sum_{k=1}^{n} Y_k \overset{\text{iid 时}}{=} \sum_{k=1}^{n} \frac{1}{\sqrt{n}} X_k^*.$$

- 假设$X_1, X_2, \cdots$ 相互独立, $EX_k = \mu_k$, $\mathrm{var}(X_k) = \sigma_k^2$. 记

$$B_n^2 = \sum_{k=1}^{n} \sigma_k^2.$$

- 目标:

$$S_n^* = \frac{S_n - \sum_{k=1}^{n} \mu_k}{B_n} = \sum_{k=1}^{n} \frac{X_k - \mu_k}{B_n} \xrightarrow{d} Z \sim N(0,1).$$

$$S_n^* = \sum_{k=1}^{n} Y_k, \quad 其中, Y_k := \frac{X_k - \mu_k}{B_n}.$$

- Lindeberg 条件 (5.5.2):

$$\sum_{k=1}^{n} E{Y_k}^2 \cdot 1_{\{|Y_k|>\varepsilon\}} \to 0, \ \forall \varepsilon > 0.$$

- Lindeberg 条件$\Longrightarrow P(\max_{1\leqslant k\leqslant n} Y_k > \varepsilon) \to 0 \quad (n \to \infty)$.
- Feller 条件 (5.5.3): $\frac{\max_{1\leqslant k\leqslant n} \sigma_k}{B_n} \to 0.$

$$\frac{\max_{1\leqslant k\leqslant n} \sigma_k}{B_n} \to 0 \quad \text{iff} \quad B_n \to \infty \text{ 且 } \frac{\sigma_n}{B_n} \to 0. \text{ (定理5.5.1)}$$

- Lindeberg-Feller CLT (定理5.5.2):

$$(5.5.2) \quad \text{iff} \quad S_n^* \xrightarrow{d} Z \sim N(0,1) \quad \text{且 } (5.5.3).$$