期中考试:

- 时间: 11月17日(周二)上午8:00-9:50, 请7:50之前到.
- 地点: 二教105

复习

- $\xi_n \xrightarrow{P} \eta$: $P(A_n) \to 0$, 其中$A_n = \{|\xi_n - \eta| \geqslant \varepsilon\}$.
- $\xi_n \xrightarrow{\text{a.s.}} \eta$: $P(\xi_n \to \eta) = 1$ iff $P(\cup_{m\geqslant n} A_m) \to 0$.
- $S_n = X_1 + \cdots + X_n$, $\frac{1}{n} S_n \xrightarrow{P,\ \text{a.s.}} EX$.
  特别地: $EX = 0$ 时, 理解为$S_n = o(n)$.
- 切比雪夫不等式.

§4.1 随机序列的收敛性(续)

定义 (依分布收敛)

$\xi_n \xrightarrow{d} \eta$ 指$F_{\xi_n}(x) \to F_\eta(x)$ 对F_η 的任意连续点x 成立.

- $\xi_n = \frac{1}{n} \to \eta = 0$, 但$F_{\xi_n}(0) = 0$, $F_\eta(0) = 1$.
- 定理1.2. 若$\xi_n \xrightarrow{P} \eta$, 则$\xi_n \xrightarrow{d} \eta$. $A_n = \{|\xi_n - \eta| \geqslant \varepsilon\}$.

$$\begin{aligned} & |P(\xi_n \leqslant x) - P(\eta \leqslant x)| \\ \leqslant \quad & P(|\xi_n - \eta| \geqslant \varepsilon) + P(x - \varepsilon \leqslant \eta \leqslant x + \varepsilon). \end{aligned}$$

- $A_n =$ 第n 次投到正面, $B =$ 色子投到大. 则$1_{A_n} \xrightarrow{d} 1_B$.
- 假设$\xi_n, n \geqslant 1, \eta$是定义在同一个概率空间上的随机变量, $\xi_n \xrightarrow{d} \eta$ 推不出$\xi_n \xrightarrow{P} \eta$ (反例：例1.2)

§4.3 中心极限定理

假设随机变量序列$X_1, X_2, \cdots$ 满足$0 < \mathrm{var}(X_n) < \infty$, $\forall n$. 令

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

- Central Limit Theorem (CLT):
  若$S_n^* = \frac{S_n - ES_n}{\sqrt{\mathrm{var}(S_n)}} \xrightarrow{d} Z \sim N(0,1)$, 则称$X_1, X_2, \cdots$ 满足CLT.
- Linderberg-Lévy CLT (定理3.1). i.i.d. 序列满足CLT.
- Berry-Esseen bound: 假设$E|X|^3 < \infty$. 那么,
  $$|F_n(x) - \Phi(x)| \leqslant \frac{3E|X^*|^3}{\sqrt{n}}, \quad \forall x.$$
- $S_n^* = \sum_{i=1}^n \frac{X_i - \mu}{\sqrt{n}\sigma} = \sum_{i=1}^n \frac{1}{\sqrt{n}} X_i^* = \sum_{i=1}^n Y_i \overset{d}{\approx} Z.$

CLT特例证明(选读) 假设$X_1, X_2, \cdots$ i.i.d.,
$P(X_1=1)=P(X_1=-1)=1/2$. 于是$\mu=0$, $\sigma^2=1$, $S_n^*=\frac{S_n}{\sqrt{n}}$.
下面证明CLT

- $P(a \leqslant \frac{S_{2n}}{\sqrt{2n}} \leqslant b) = \sum_{a \leqslant \frac{2k}{\sqrt{2n}} \leqslant b} C_{2n}^{n+k} \frac{1}{2^{2n}}$,
  $x_k^2 = 2\frac{k^2}{n}$, $\Delta x_k = \sqrt{\frac{2}{n}}$.
- Stirling 公式: $m! \approx \sqrt{2\pi} \cdot \sqrt{m}(\frac{m}{e})^m$.
  $C_{2n}^{n+k} = \frac{(2n)!}{(n+k)!(n-k)!} \approx \frac{\sqrt{2\pi 2n}(\frac{2n}{e})^{2n}}{\sqrt{2\pi(n+k)}(\frac{n+k}{e})^{n+k}\sqrt{2\pi(n-k)}(\frac{n-k}{e})^{n-k}}$
- $C_{2n}^{n+k}\frac{1}{2^{2n}} \approx \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{2n}{(n+k)(n-k)}} \cdot \frac{n^{2n}}{(n+k)^{n+k}(n-k)^{n-k}}$
- $\sqrt{\frac{2n}{n^2-k^2}} \approx \sqrt{\frac{2}{n}} = \Delta x_k$,
- $(\frac{n^2-k^2}{n^2})^{-n}(\frac{n+k-2k}{n+k})^k \approx (1-\frac{1}{n}\frac{k^2}{n})^{-n}(1-\frac{1}{k}\frac{2k^2}{n})^k \approx$
  $e^{\frac{1}{2}x_k^2}e^{-x_k^2} = e^{-\frac{1}{2}x_k^2}$.

CLT $\Rightarrow$ WLLN.

- 不妨设$\mu = 0$, $\sigma^2 = 1$.
- WLLN: $\frac{1}{n}S_n \xrightarrow{P} 0$, CLT: $\frac{S_n}{\sqrt{n}} \xrightarrow{d} Z$.
- 粗略地, WLLN: $S_n = o(n)$, CLT: $S_n = O(\sqrt{n})$.
- 对任意$\epsilon > 0$固定，考虑$A_n = \{|\frac{1}{n}S_n| \geqslant \varepsilon\}$.

  对任意$\delta > 0$, n充分大时,

  $P(A_n) = P(|S_n^*| \geqslant \sqrt{n}\varepsilon) \leqslant P(|S_n^*| \geqslant x) \leqslant \delta(\ n \geqslant (\frac{x}{\varepsilon})^2$ 时$)$.

  - “$\leqslant$”：根据中心极限定理：n充分大时,

    $|P(|S_n^*| \geqslant x) - \Phi(x)| < \frac{1}{2}\delta$, 选取$x > 0$ s.t. $\Phi(x) \leqslant \delta$.

  因此, $P(A_n) \to 0$.

CLT: $P(S_n \leqslant x) = P(S_n^* \leqslant x^*) \approx p,$

$x^* = \frac{x - n\mu}{\sqrt{n}\sigma},\ p = \Phi(x^*),\ \Phi(-x) = 1 - \Phi(x).$

附表 1 标准正态分布数值表

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	1.40	0.9192	2.30	0.9893
0.05	0.5199	1.42	0.9222	2.33	0.9901
0.10	0.5398	1.45	0.9265	2.35	0.9906
0.15	0.5596	1.48	0.9306	2.38	0.9913
0.20	0.5793	1.50	0.9332	2.40	0.9918
0.25	0.5987	1.55	0.9394	2.42	0.9922
0.30	0.6179	1.58	0.9429	2.45	0.9929
0.35	0.6368	1.60	0.9452	2.50	0.9938
0.40	0.6554	1.65	0.9505	2.55	0.9946
0.45	0.6736	1.68	0.9535	2.58	0.9951
0.50	0.6915	1.70	0.9554	2.60	0.9953

例3.1 加法器同时收到20个i.i.d.噪声电压$V_k \sim U(0,10)$, $k=1,\cdots,20$. 记$V=\sum_{k=1}^{20} V_k$, 求$P(V>105)$.

- 此题已知$n=20$, $x=105$, 求p.
- $P(V>105)=P\left(V^*>\frac{105-20\times 5}{\sqrt{20\times\frac{100}{12}}}=:x^*\right)\approx 1-\Phi(x^*)=:p.$
- 计算得$x^*\approx 0.387$. 查表得$\Phi(x^*)=0.652$,
  从而所求之$p=1-0.652=0.348$.

例3.2 旅馆有500间客房, 每间有一台2千瓦的空调. 开房率为80%. 问: 需多少千瓦的电力能有99% 的把握保证电力足够?

- 已知n, p, 求x. 假设提供x 千瓦.
- A_i = 第i 间房开空调, $P(A_i) = 80\%$, $X_i = 2 \times 1_{A_i}$. $n = 500$.
- 要求x 满足: $P(S_n \leqslant x) \geqslant 99\% = p$.
  即$P(S_n^* \leqslant x^*) \geqslant 0.99$. 查表得$\Phi(2.33) = 0.99$.
- 故需$x^* = \frac{x-500\times2\times0.8}{\sqrt{500\times2^2\times0.8\times0.2}} \geqslant 2.33$.
  即$x \geqslant 800 + 2.33 \times \sqrt{320} = 841.68$, 从而需842 千瓦.

例(民意调查) 为保证调查结果与真实值的误差不超过$\varepsilon = 0.1$ (或0.05, 0.01) 的概率至少为95%. 至少需调查多少人?

- 已知$x = \varepsilon$, $p = 0.95$, 求n.
- $A_i =$ 第i 人支持该候选人, $P(A_i) = q =$真实值. $X_i = 1_{A_i}$.
  $\frac{1}{n}S_n =$调查结果. 目标: $P(|\frac{1}{n}S_n - q| \leqslant \varepsilon) \geqslant 0.95$.
- $P(|S_n^*| \leqslant \frac{\varepsilon\sqrt{n}}{\sqrt{q(1-q)}} = x^*) \geqslant 0.95$.
  要求$\Phi(x^*) - \Phi(-x^*) = \Phi(x^*) - (1 - \Phi(x^*)) \geqslant 0.95$,
  即$\Phi(x^*) \geqslant \frac{1}{2}(1 + 0.95) = 0.975$, 查表得$x^* \geqslant 1.96$.
- 要求$\frac{\varepsilon\sqrt{n}}{\sqrt{q(1-q)}} \geqslant 1.96$, 即$\sqrt{n} \geqslant \frac{1.96}{\varepsilon} \times \sqrt{q(1-q)}$.
- 为“保证”, 要求$\sqrt{n} \geqslant \frac{1.96}{\varepsilon} \times \max_q \sqrt{q(1-q)} = \frac{1.96}{2\varepsilon}$.
  若$\varepsilon = 0.1$, 则至少需调查$n_0 = \lceil 9.8^2 \rceil = 97$;
  若$\varepsilon = 0.05$, 则$n_0 = 385$; 若$\varepsilon = 0.01$, 则$n_0 = 9604$ 人.

习题四、13. $e^{-n}\sum_{k=0}^{n}\frac{1}{k!}n^k \to \frac{1}{2}$.

- $X_1, X_2, \cdots$ 独立同分布, $X_1 \sim \mathcal{P}(1)$, 则$S_n \sim \mathcal{P}(n)$，从而, 左边$= P(S_n \leqslant n)$.
- $P(S_n \leqslant n) = P(S_n^* \leqslant 0) \to P(Z \leqslant 0) = \frac{1}{2}$, 其中$Z \sim N(0,1)$.