概率论第 7 次习题课讲义

宗语轩

2022 秋, USTC

作业说明

(1) 见**第 6 次习题课讲义例 2.1**, 对这几类常见的离散型与连续型随机变量的特征函数建议把结果记住, 上学期期末考一道题目就考了与之相关的内容:

> 4. (15分) 记
> $$\phi(t) = \left(\frac{\sin t}{t}\right)^2,$$
> 试用概率方法证明对实数$t_1, \ldots, t_n$矩阵$H_n = (\phi(t_i - t_j))_{i,j=1}^n$非负定.

另外对这 7 类模型, 只要不是题目特别要求求解他们, 都可以直接利用其结果而无需推导. **5.12.33(a)** 答案没有 $X \sim P(\lambda)$ 特征函数的计算过程, 这里简单补一下 (一般也可以省略)

$$\phi_X(t) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX}] = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \cdot \mathrm{e}^{\mathrm{i}tk} = \sum_{k=0}^{+\infty} \frac{(\lambda \mathrm{e}^{\mathrm{i}t})^k}{k!}\mathrm{e}^{-\lambda} = \mathrm{e}^{\lambda \mathrm{e}^{\mathrm{i}t}} \cdot \mathrm{e}^{-\lambda} = \mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}t}-1)}$$

(2) **5.12.33(b)** Γ 分布参数和一般形式反了, 不是指数分布那个.

(3) **5.12.39** 可以直接通过分布函数验证, 或者同答案用特征函数 + 连续性定理做.

(4) **5.12.41** 可以直接验证 Linderberg 条件, 或者同答案.

(5) **7.1.1(c)** 答案有笔误, 不是 "$|X| \geqslant I_\varepsilon$", 而是 "$|X| \geqslant \varepsilon I_\varepsilon$".

(6) **7.2.5(a)**"$c = 0$" 的情况要单独讨论. 本题是 Slutsky's theorem(s) 的简单形式, 之后会讲.

(7) 原则性错误: $X_n \xrightarrow{D} N(0, n), X_n \to a_n$ 等.

1 专题选讲

1.1 特征函数与矩

X 的特征函数: $\phi_(t) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX}]$. 一般地, $(\mathbb{E}[\mathrm{e}^{\mathrm{i}tX}])' \neq \mathbb{E}[(\mathrm{e}^{\mathrm{i}tX})'] = \mathbb{E}[\mathrm{i}X\mathrm{e}^{\mathrm{i}tX}]$.

---

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.
发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

例 1.1. 特征函数的导数不一定存在. 例如随机变量 X 满足 $\mathbb{P}(X=5^k)=2^{-k}, k=1,2,\cdots$, 则有

$$\phi(t)=\sum_{k=1}^{+\infty}\frac{1}{2^k}\mathrm{e}^{\mathrm{i}t5^k}.$$

而 $\phi(t)$ 处处不可导.

但这也不是无可救药, 如果随机变量的 k 阶矩存在, 仍然能得到特征函数前 k 阶导的信息, 进而能得到矩的信息.

定理 1.1. 若 $\mathbb{E}[|X|^k]<\infty$, 则 $\forall j\le k$ 有

$$\phi^{(j)}(t)=\int_{-\infty}^{+\infty}(\mathrm{i}x)^j\mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x),\quad \phi^{(j)}(0)=i^j\mathbb{E}[X^j]$$

进而 $\phi(t)=\sum\limits_{j=0}^{k}\dfrac{(it)^j}{j!}\mathbb{E}[X^j]+o(t^k), t\to 0$

注. 在一定条件下, 可以利用特征函数来计算矩.

对于一个随机变量, 其 k 阶矩存在可以推出其特征函数 k 阶可导. 而反之是否成立?

定理 1.2. 设 k 为偶数, 若 $\phi(t)$ 在 0 处附近有 k 阶导数, 则 $\mathbb{E}[|X|^k]<\infty$.

证明. 先看 $k=2$. 假设 $\phi''(t)$ 在 0 附近存在, 则有

$$\phi''(0)=\lim_{h\to 0}\frac{\phi(h)+\phi(-h)-2\phi(0)}{h^2}=\lim_{h\to 0}\int_{-\infty}^{+\infty}\frac{\mathrm{e}^{\mathrm{i}xh}+\mathrm{e}^{-\mathrm{i}xh}-2}{h^2}\mathrm{d}F(x)=-2\lim_{h\to 0}\int_{-\infty}^{+\infty}\frac{1-\cos hx}{h^2}\mathrm{d}F(x).$$

利用 Fatou 引理, 我们有

$$\mathbb{E}[X^2]=\int_{-\infty}^{+\infty}x^2\mathrm{d}F(x)=2\int_{-\infty}^{+\infty}\lim_{h\to 0}\frac{1-\cos hx}{h^2}\mathrm{d}F(x)\leqslant\lim_{h\to 0}2\int_{-\infty}^{+\infty}\frac{1-\cos hx}{h^2}\mathrm{d}F(x)=-\phi''(0)<\infty.$$

所以 $k=2$ 时成立. 现在我们开始归纳: 假设当 $n=2k$ 时命题成立, 即 $\mathbb{E}[X^{2k}]<\infty$, 则有

$$\phi^{(2k)}(t)=\int_{-\infty}^{+\infty}(\mathrm{i}x)^{2k}\mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x).$$

设 $G(x):=\displaystyle\int_{-\infty}^{x}y^{2k}\mathrm{d}F(y)$, 则 $\dfrac{G(x)}{G(+\infty)}$ 为某个随机变量 Y 的分布函数, 其特征函数为

$$\phi_Y(t)=\frac{1}{G(+\infty)}\int_{-\infty}^{+\infty}x^{2k}\mathrm{e}^{\mathrm{i}tx}\mathrm{d}F(x)=\frac{(-1)^k\phi^{(2k)}(t)}{G(+\infty)}.$$

因此 $\phi_Y''(t)$ 在 0 附近存在, 故有

$$\int_{-\infty}^{+\infty}x^2\frac{\mathrm{d}G(x)}{G(+\infty)}<\infty\implies\frac{1}{G(+\infty)}\int_{-\infty}^{+\infty}x^{2k+2}\mathrm{d}F(x)<\infty\implies\mathbb{E}[X^{2k+2}]<\infty.$$

□

注. k 为奇数时上述结论不成立. 例如 X 满足 $\mathbb{P}(X=\pm j)=\dfrac{1}{2Cj^2\log j}, j=2,3,\cdots$, 则

$$\phi(t)=\frac{1}{C}\sum_{j=2}^{+\infty}\frac{\cos jt}{j^2\log j},$$

其中 $C=\displaystyle\sum_{j=2}^{+\infty}\frac{1}{j^2\log j}$ 为正则化常数. $\phi(t)$ 在 0 附近可导且连续, 但 $\mathbb{E}[|X|]=\displaystyle\sum_{j=2}^{+\infty}\frac{1}{j\log j}=+\infty$.

利用特征函数, 我们还可以通过矩来得到分布:

定理 1.3. 若 $\displaystyle\limsup_{n\to+\infty}\frac{(\mu_{2n})^{\frac{1}{2n}}}{2n}=r<\infty$, 则至多存在一个分布 F, 使得

$$\int_{-\infty}^{+\infty}x^n\mathrm{d}F(x)=\mu_n.$$

证明. 见本文最后第 4 部分, 可以暂时忽略. □

注. **Riesz** 条件 $\displaystyle\limsup_{n\to+\infty}\frac{(\mu_{2n})^{\frac{1}{2n}}}{2n}=r<\infty$ 可以换成 **Carleman** 条件 $\displaystyle\sum_{k=1}^{+\infty}\frac{1}{\mu_{2k}^{1/2k}}=+\infty$. 上述结论依然成立.

据此显然有如下推论:

推论 1.1. 对随机变量 X,Y, 若 $\displaystyle\limsup_{n\to+\infty}\frac{(\mathbb{E}[X^{2n}])^{\frac{1}{2n}}}{2n}=r<\infty$, 且

$$\mathbb{E}[X^n]=\mathbb{E}[Y^n],\quad\forall n\in\mathbb{N}^*.$$

则 X 与 Y 同分布.

同时也能推出第 6 次习题课讲义中的矩收敛定理 (**定理 2.8**), 这里不再赘述.

下面这个例子体现出随机变量在矩的信息转化成分布的过程中, 特征函数起到中转站的作用:

例 1.2. 随机变量 X 满足 $\mathbb{E}[|X|]<+\infty$. 若对任意有界光滑函数 f 均有

$$\lambda\mathbb{E}[f(X+1)]=\mathbb{E}[Xf(X)],$$

其中常数 $\lambda>0$. 证明: X 服从参数为 λ 的泊松分布.

证明. 因为 $\mathrm{e}^{\mathrm{i}tx}=\cos tx+\mathrm{i}\sin tx$ 有界光滑, 故

$$\lambda\mathbb{E}[\mathrm{e}^{\mathrm{i}t(X+1)}]=\mathbb{E}[X\mathrm{e}^{\mathrm{i}tX}].$$

记 $\phi(t)=\mathbb{E}[\mathrm{e}^{\mathrm{i}tX}]$, 因为 $\mathbb{E}[|X|]<+\infty$, 故有

$$\phi'(t)=\mathbb{E}[\mathrm{i}X\mathrm{e}^{\mathrm{i}tX}].$$

从而

$$\mathrm{i}\lambda\mathrm{e}^{\mathrm{i}t}\phi(t)=\phi'(t).$$

结合初值条件 $\phi(0)=1$, 由常微分方程的理论可知,

$$\phi(t)=\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}t}-1)},$$

由唯一性定理知, X 服从参数为 λ 的泊松分布. □

1.2 连续性定理与弱收敛

课堂上给出了连续性定理的内容, 我们这节的目标是证明连续性定理. 在此之前先引入一些与之有关的定理, 引理和概念.

上节课我们讲了特征函数里的 Parseval 等式, 在这里就要派上其用场.

定理 1.4 (**Parseval 等式**)**.** 对随机变量 X, Y, 有

$$\int_{-\infty}^{+\infty} \phi_X(t)\mathrm{d}F_Y(t) = \int_{-\infty}^{+\infty} \phi_Y(t)\mathrm{d}F_X(t).$$

我们取 Y 为 $[-u, u]$ 上的均匀分布, 则 $\phi_Y(t) = \dfrac{\sin ut}{ut}$, 利用 Parseval 等式, 则有

$$\begin{aligned}\frac{1}{2u}\int_{-u}^{u} \phi_X(t)\mathrm{d}t &= \int_{-\infty}^{+\infty} \frac{\sin ut}{ut}\mathrm{d}F_X(t) = \int_{|ut|<2} \frac{\sin ut}{ut}\mathrm{d}F_X(t) + \int_{|ut|\geqslant 2} \frac{\sin ut}{ut}\mathrm{d}F_X(t)\\ &\leqslant \int_{|ut|<2} \mathrm{d}F_X(t) + \frac{1}{2}\int_{|ut|\geqslant 2} \mathrm{d}F_X(t) = 1 - \frac{1}{2}\mathbb{P}(|Xu| \geqslant 2)\\ &= 1 - \frac{1}{2}\mathbb{P}\left(|X| \geqslant \frac{2}{u}\right).\end{aligned}$$

由此得到

命题 1.1. 设 X 是随机变量, 则对 $\forall u > 0$, 均有

$$\mathbb{P}(|X| \geqslant \frac{2}{u}) \leqslant \frac{1}{u}\int_{-u}^{u} (1 - \phi_X(t))\mathrm{d}t.$$

下面这个引理给出证明弱收敛的一个等价转化形式:

引理 1.1. $X_n \xrightarrow{D} X \iff$ 对任意子列 $\{n_k\}$, 存在子子列 $\{n'_k\} \subset \{n_k\}$, 使得

$$X_{n'_k} \xrightarrow{D} X.$$

提示. 左推右显然, 右推左反证即可.

一个重要的准备工作是要引入**胎紧**这一概念. 之前我们在数学分析中学过 “一致连续”, “一致收敛” 等概念, 其中的 “一致” 都是强调函数列 “整体” 的行为. 对于随机变量 X 的分布函数 F 而言, 由其定义知

$$\mathbb{P}(|X| > M) = (F(-M) + 1 - F(M)) \xrightarrow{M\to+\infty} 0.$$

但当研究对象为分布函数列时, 又会涉及到一致性的问题 (不胎紧的例子可以参见**例 1.4(1)**), 这里我们通过胎紧性来描述:

定义 1.1. F_n 是一列分布函数. 称 F_n 是胎紧的, 若

$$\sup_n (F_n(-M) + 1 - F_n(M)) \xrightarrow{M\to+\infty} 0.$$

注. 记 X_n 的分布函数为 F_n, 则

$$\{F_n\}\ \text{胎紧} \iff \sup_n \mathbb{P}(|X_n| > M) \xrightarrow{M\to+\infty} 0 \iff \limsup_{n\to+\infty} \mathbb{P}(|X_n| > M) \xrightarrow{M\to+\infty} 0$$

证明连续性定理的关键是下面的 Helly 定理.

定理 1.5 (**Helly 选择定理**). 设 F_n 为分布函数, 则存在子序列 $\{n_k\}$, 使得 $F_{n_k} \xrightarrow{v} F$.
记 $F_n \xrightarrow{v} F$ 为 F_n淡收敛至 F, 指对 $\forall x \in C_F, F_n(x) \to F$, 但 F 不一定有分布函数 (F 右连续且单调递增但 $F(-\infty)=0, F(+\infty)=1$ 不一定成立).

证明. 参考 **Durrett Theorem 3.2.12**(已放在群文件中). □

命题 1.2. 接上, 对任意收敛子序列 F_{n_k} 都收敛到分布函数的充要条件是 $\{F_n\}$ 胎紧.

证明. 充分性: 设 $F_{n_k} \xrightarrow{v} F$, $\{F_n\}$ 胎紧, 即对 $\forall \varepsilon > 0$, 均有 $M_\varepsilon > 0$, 使得

$$\limsup_{n\to+\infty}(1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon)) < \varepsilon.$$

记 $r < -M_\varepsilon, s > M_\varepsilon$ 为 F 的连续点. 因为 $F_{n_k}(r) \to F(r)$,$F_{n_k}(s) \to F(s)$, 故有

$$1 - F(s) + F(r) = \lim_{k\to+\infty}(1 - F_{n_k}(s) + F_{n_k}(r)) \leqslant \limsup_{n\to+\infty}(1 - F_n(M_\varepsilon) + F_n(-M_\varepsilon)) < \varepsilon.$$

再利用 F 的单调递增性知,

$$\limsup_{x\to+\infty}(1 - F(x) + F(-x)) < \varepsilon.$$

由 $\varepsilon > 0$ 的任意性知, F 是分布函数.
必要性: 反证: 假设 $\{F_n\}$ 不是胎紧的, 即 $\exists \varepsilon > 0$ 及子列 $\{n_k\}$ 使得

$$1 - F_{n_k}(k) + F_{n_k}(-k) \geqslant \varepsilon$$

对所有 $k \in \mathbb{N}^*$ 成立. 取子子列 $\{n_{k_j}\} \subset \{n_k\}$ 使得 $F_{n_{k_j}} \xrightarrow{v} F$. 记 $r < 0 < s$ 是 F 的连续点, 则有

$$1 - F(s) + F(r) = \lim_{j\to+\infty}(1 - F_{n_{k_j}}(s) + F_{n_{k_j}}(r)) \geqslant \liminf_{j\to+\infty}(1 - F_{n_{k_j}}(k_j) + F_{n_{k_j}}(-k_j)) \geqslant \varepsilon.$$

令 $s \to +\infty, r \to -\infty$, 则 F 不是分布函数. □

准备工作已就绪, 现在我们来证明连续性定理:

定理 1.6 (**Lévy-Cramér 连续性定理**). F_n 为分布函数, $\phi_n(t) = \int e^{itx}\mathrm{d}F_n$

(1) 若 F 为分布函数, 且 $F_n \xrightarrow{w} F$, 则 $\phi_n(t) \to \phi(t), \forall t \in \mathbb{R}$(且内闭一致收敛, 这里暂且不证), 这里 $\phi(t) = \int e^{itx}\mathrm{d}F$

(2) 若 $\phi(t) := \lim_{n\to\infty}\phi_n(t)$ 存在, 且 $\phi(t)$ 在 $t = 0$ 处连续, 则 ϕ 为某分布函数 F 的特征函数, 且 $F_n \xrightarrow{w} F$.

证明. 对 (1) 设 X_n, X 的分布函数分别是 F_n, F, 则有 $X_n \xrightarrow{D} X$. 注意到 $\mathrm{e}^{\mathrm{i}tx}$ 有界连续, 因此 $\mathbb{E}[\mathrm{e}^{\mathrm{i}tX_n}] \to \mathbb{E}[\mathrm{e}^{\mathrm{i}tX}]$.
(2) 先利用**引理 1.1**, 即证: 对任意子列 $\{n_k\}$, 存在子子列 $\{n_k'\} \subset \{n_k\}$, 使得

$$F_{n_k'} \xrightarrow{w} F.$$

对任意子列 $\{n_k\}$, 我们分以下四个步骤完成证明:

1. 利用 **Helly 选择定理 (定理 1.5)** 得到淡收敛子子列 $\{n_k'\}$,
2. $\phi(t)$ 在 $t=0$ 处连续保证了 $\{F_n\}$ 的胎紧, 因为

$$\limsup_{n\to+\infty}\mathbb{P}(||X_n|>r)\leqslant\frac{r}{2}\limsup_{n\to+\infty}\int_{-\frac{2}{r}}^{\frac{2}{r}}(1-\phi_n(t))\mathrm{d}t=\frac{r}{2}\int_{-\frac{2}{r}}^{\frac{2}{r}}(1-\phi(t))\mathrm{d}t\xrightarrow{r\to+\infty}2(1-\phi(0))=0,$$

其中 F_n 是随机变量 X_n 的分布函数, 第一个不等式用到了**命题 1.1**, 后面的等号用到了特征函数的一致连续性.

3. $\phi(t)=\lim\limits_{n\to\infty}\phi_n(t)$ 保证了子子列收敛的唯一性.
4. 利用等价条件**命题 1.2** 即得证.

$\square$

1.3 中心极限定理 (CLT)

课堂上我们已经用特征函数的方法证明了 i.i.d 形式的 CLT

定理 1.7 (**i.i.d CLT**)**.** 设 $\{X_k\}$ 独立同分布, 且 $\mu=\mathbb{E}[X_1],\sigma^2=Var(X_1),\sigma>0$, 则

$$\frac{S_n-n\mu}{\sqrt{n\sigma^2}}\xrightarrow{D}N(0,1).$$

我们进一步考虑更一般的情形.

问题: 如何利用特征函数证明依分布收敛?

命题 1.3. 设 $c_n\to c\in\mathbb{C}$, 则有

$$\lim_{n\to\infty}\left(1+\frac{c_n}{n}\right)^n=\mathrm{e}^c$$

证明. 先看下面两个引理:

引理 1.2. 设 $b\in\mathbb{C},|b|\leqslant 1$, 则有 $|\mathrm{e}^b-b-1|\leqslant|b|^2$.

引理 1.2 的证明: 利用 $\mathrm{e}^b=\sum\limits_{k=0}^{+\infty}\frac{b^k}{k!}$, 有

$$|\mathrm{e}^b-b-1|\leqslant\frac{|b|^2}{2}(1+\frac{|b|}{3}+\frac{|b|^2}{3\cdot 4}+\cdots)\leqslant\frac{|b|^2}{2}(1+\frac{1}{3}+\frac{1}{3\cdot 4}+\cdots)\leqslant|b|^2.$$

引理 1.3. 设 $z_1,\cdots z_n,w_1,\cdots,w_n\in\mathbb{C}$ 且模长均不超过 r, 则

$$\left|\prod_{k=1}^n z_k-\prod_{k=1}^n w_k\right|\leqslant r^{n-1}\sum_{k=1}^n|z_k-w_k|.$$

引理 1.3 的证明: 我们有

$$\begin{aligned}\left|\prod_{k=1}^n z_k-\prod_{k=1}^n w_k\right|&=\left|\prod_{k=1}^n z_k-\prod_{k=1}^{n-1}z_k\cdot w_n+\prod_{k=1}^{n-1}z_k\cdot w_n-\prod_{k=1}^{n-2}z_k\cdot w_{n-1}w_n+\cdots+z_1\prod_{k=2}^n w_k-\prod_{k=1}^n w_k\right|\\&\leqslant\left|\prod_{k=1}^n z_k-\prod_{k=1}^{n-1}z_k\cdot w_n\right|+\left|\prod_{k=1}^{n-1}z_k\cdot w_n-\prod_{k=1}^{n-2}z_k\cdot w_{n-1}w_n\right|+\cdots+\left|z_1\prod_{k=2}^n w_k-\prod_{k=1}^n w_k\right|\end{aligned}$$

$$
\begin{aligned}
&= \left|\prod_{k=1}^{n-1} z_k\right| |z_n - w_n| + \left|\prod_{k=1}^{n-2} z_k \cdot z_n\right| |z_{n-1} - w_{n-1} + \cdots + \left|\prod_{k=2}^{n} w_k\right| |z_1 - w_1| \\
&\leqslant r^{n-1} \sum_{k=1}^{n} |z_k - w_k|.
\end{aligned}
$$

回到原命题. 设 $|c_n| \leqslant r, r > 0$. 注意到

$$
\left|1 + \frac{c_n}{n}\right| \leqslant 1 + \frac{|c_n|}{n} \leqslant 1 + \frac{r}{n} \leqslant \mathrm{e}^{\frac{r}{n}}, \quad |\mathrm{e}^{\frac{c_n}{n}}| \leqslant \mathrm{e}^{\frac{|c_n|}{n}} \leqslant \mathrm{e}^{\frac{r}{n}}
$$

利用 $\lim\limits_{n\to+\infty} \mathrm{e}^{c_n} = \mathrm{e}^{c}$ 及上述引理可得

$$
\left|\left(1 + \frac{c_n}{n}\right)^n - \mathrm{e}^{c_n}\right| \leqslant (\mathrm{e}^{\frac{r}{n}})^{n-1} \cdot n \cdot \left|\mathrm{e}^{\frac{c_n}{n}} - 1 - \frac{c_n}{n}\right| \leqslant \mathrm{e}^{r} \cdot n \cdot \frac{|c_n|^2}{n^2} \leqslant \mathrm{e}^{r} \cdot \frac{r^2}{n} \xrightarrow{n\to+\infty} 0.
$$

□

注. 用特征函数证明依分布收敛, 对 i.i.d 和的情形, 想办法写成 $\left(1 + \dfrac{c_n}{n}\right)^n$ 的形式.

例 1.3. 设 $X_1, \cdots X_n$ 独立同分布, X_1 对称且满足 $\mathbb{P}(|X_1| > x) = x^{-2}$, $x \geqslant 1$. 记 $S_n = \sum\limits_{k=1}^{n} X_k$, 则有

$$
\frac{S_n}{\sqrt{n \log n}} \xrightarrow{D} N(0, 1).
$$

注. X_1 无二阶矩, 因为 $\mathbb{E}[X^2] = \displaystyle\int_1^{+\infty} x^{-1} \mathrm{d}x = +\infty$.

证明. 利用特征函数, 有

$$
\begin{aligned}
\mathbb{E}[\mathrm{e}^{\mathrm{i}t\frac{S_n}{\sqrt{n\log n}}}] &= (\mathbb{E}[\mathrm{e}^{\mathrm{i}t\frac{X_1}{\sqrt{n\log n}}}])^n = \left(\int_{|x|>1} \mathrm{e}^{\mathrm{i}t\frac{X_1}{\sqrt{n\log n}}} |x|^{-3} \mathrm{d}x\right)^n = \left(2\int_1^{+\infty} \cos\left(t\frac{X_1}{\sqrt{n\log n}}\right) x^{-3} \mathrm{d}x\right)^n \\
&= \left(1 + 2\int_1^{+\infty} \left(\cos\left(t\frac{X_1}{\sqrt{n\log n}}\right) - 1\right) x^{-3} \mathrm{d}x\right)^n
\end{aligned}
$$

为转化 $\left(1 + \dfrac{c_n}{n}\right)^n$ 的形式, 考虑 $c_n := 2n\displaystyle\int_1^{+\infty} \left(\cos\left(t\frac{X_1}{\sqrt{n\log n}}\right) - 1\right) x^{-3} \mathrm{d}x$, 令 $y = \dfrac{x}{\sqrt{n\log n}}$, 则

$$
c_n = 2n\int_{\frac{1}{\sqrt{n\log n}}}^{+\infty} (n\log n)^{-1} (\cos ty - 1) y^{-3} \mathrm{d}y = \frac{1}{\log n} 2\int_{\frac{1}{\sqrt{n\log n}}}^{+\infty} (\cos ty - 1) y^{-3} \mathrm{d}y.
$$

把 n 转化成 x, 对 x 取极限, 并利用 L'Hospital 法则, 有

$$
\begin{aligned}
\lim_{x\to+\infty} \frac{2\int_{\frac{1}{\sqrt{x\log x}}}^{+\infty} (\cos ty - 1) y^{-3} \mathrm{d}y}{\log x} &= \lim_{x\to+\infty} \frac{2(x\log x)^{\frac{3}{2}} (\frac{1}{\sqrt{x\log x}})' (\cos\frac{t}{\sqrt{x\log x}} - 1)}{\frac{1}{x}} \\
&= \lim_{x\to+\infty} x(1 + \log x)(\cos\frac{t}{\sqrt{x\log x}} - 1) \\
&= -\frac{t^2}{2}.
\end{aligned}
$$

因此 $\lim\limits_{n\to+\infty} c_n = -\dfrac{t^2}{2}$. 由**命题 1.3** 可知

$$
\lim_{n\to+\infty} \mathbb{E}[\mathrm{e}^{\mathrm{i}t\frac{S_n}{\sqrt{n\log n}}}] = \mathrm{e}^{-\frac{t^2}{2}} \implies \frac{S_n}{\sqrt{n\log n}} \xrightarrow{D} N(0, 1).
$$

□

课堂上亦给出了更一般形式的 CLT:

Linderberg 条件: 对 $X_1,\cdots,X_n$, 记 $a_k=\mathbb{E}[X_k], b_k^2=Var[X_k], B_n^2=\sum\limits_{k=1}^n b_k^2, F_k$ 为 X_k 分布函数.

$$\forall\epsilon>0,\ \lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^n\int_{|x-a_k|>\epsilon B_n}(x-a_k)^2\mathrm{d}F_k=0 \tag{L}$$

定理 1.8 (**Linderberg-Feller CLT**). 设 $X_1,\cdots,X_n$ 相互独立, 满足(L), 则

$$\frac{1}{B_n}\sum_{k=1}^n(X_k-a_k)\to N(0,1)$$

注 1. 事实上, 我们可以把上述定理推广到随机变量组列 (双下标) 中, 结论依然成立 (具体参考 **Durrett Theorem 3.4.10**):

定理 1.9. 对每个 $n\in\mathbb{N}^*$, 随机变量 $X_{n,m},1\leqslant m\leqslant n$ 相互独立, 且满足

(1) $\mathbb{E}[X_{n,m}]=0,\forall m,n.$

(2) $\sum\limits_{m=1}^n X_{n,m}^2\to\sigma^2>0.$

(3) 对 $\forall\varepsilon>0$, 均有 $\lim\limits_{n\to+\infty}\sum\limits_{m=1}^n\mathbb{E}[X_{n,m}^2 I_{\{|X_{n,m}|\geqslant\varepsilon\}}]=0.$

则当 $n\to+\infty$ 时, 有

$$\sum_{m=1}^n X_{n,m}\xrightarrow{D}N(0,1).$$

注 2. 对于 Linderberg-Feller CLT 一般有两种证明方法. 一种方法是从弱收敛的角度通过特征函数证明逐点收敛, 证明可以参考 **Durrett Theorem 3.4.10**; 另一种方法利用依分布收敛的各种等价刻画的性质, 运用 Lindeberg 替换术解决, 更加 technical, 证明见本讲义第 3 部分阅读材料.

下面看一个相关的例子:

例 1.4. 设 $\{X_n\}$ 为相互独立的随机变量列, 记 $S_n=\sum\limits_{k=1}^n X_k$, 在以下两个条件下试选择合适的数列 $\{a_n\},\{B_n\}$ 并证明

$$\frac{S_n-a_n}{B_n}\xrightarrow{\mathrm{D}}N(0,1).$$

(1) $\mathbb{P}(X_k=\sqrt{k})=\mathbb{P}(X_k=-\sqrt{k})=\frac{1}{2}.$

(2) $\mathbb{P}(X_k=1)=\frac{1}{k},\mathbb{P}(X_k=0)=1-\frac{1}{k}.$

证明. (1) $\mathbb{E}[X_k]=0,Var(X_k)=k$. 可推测

$$a_n=0,\quad B_n=\sqrt{\sum_{k=1}^n Var(X_k)}=\sqrt{\frac{n(n+1)}{2}}.$$

下面验证其满足 Linderberg 条件. 对 $\forall\varepsilon>0$, 取正整数 N, 使得 $N>\frac{2}{\varepsilon^2}$. 则当 $n>N$ 时, 均有

$$\frac{\varepsilon B_n}{\sqrt{n}}=\varepsilon\sqrt{\frac{n+1}{2}}>\varepsilon\sqrt{\frac{N}{2}}=1.$$

因此

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x|>\epsilon B_n}x^2\mathrm{d}F_k=0.$$

(2) $\mathbb{E}[X_k]=\frac{1}{k}, Var(X_k)=\frac{1}{k}-\frac{1}{k^2}$. 可推测

$$a_n=\sum_{k=1}^{n}\frac{1}{k},\quad B_n=\sqrt{\sum_{k=1}^{n}Var(X_k)}=\sqrt{\sum_{k=1}^{n}\frac{1}{k}-\sum_{k=1}^{n}\frac{1}{k^2}}.$$

下面验证其满足 Linderberg 条件. 记 $C_n=\sum_{k=1}^{n}\frac{1}{k^2}$, 则 $1\leqslant C_n\leqslant\frac{\pi^2}{6}$, 因此 B_n 递增至 $+\infty$. 因为 $|0-\mathbb{E}[X_k]|,|1-\mathbb{E}[X_k]|\leqslant 1$, 故对 $\forall\varepsilon>0,\exists N\in\mathbb{N}^*$, 使得 $B_N>\frac{1}{\varepsilon}$. 则当 $n>N$ 时, 均有

$$\varepsilon B_n>\varepsilon B_N>1.$$

因此

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x|>\epsilon B_n}x^2\mathrm{d}F_k=0.$$

□

注 1. 由 Slutsky's theorem 可知 $\{a_n\},\{B_n\}$ 不一定唯一. 如 (2) 中取 $a_n=\log n, B_n=\sqrt{\log n}$ 上述结论依然成立.

注 2. 这类题目也可以通过验证 Lyapunov 条件来解决 (并非万能方法).

2 补充习题

1 Grimmett 5.10.3, 5.12.32, 5.12.34, 5.12.42, 5.12.50, 5.12.52 (其中部分题与精选题重合)

2 设随机变量列 $\{X_k\}$ 相互独立且服从 $[0,1]$ 上的均匀分布, 令 $Y_n=(\prod_{i=1}^{n}X_i)^{-\frac{1}{n}}$, 证明:

$$\sqrt{n}(Y_n-\mathrm{e})\xrightarrow{D}N(0,\mathrm{e}^2).$$

提示. 特征函数计算, 并注意到 $-\ln X_k$ 是参数为 1 的指数分布.

3 设随机变量 X_n 服从参数为正整数 n 的泊松分布, 试选择合适的数列 $\{a_n\},\{b_n\}$ 并证明

$$\frac{X_n-a_n}{b_n}\xrightarrow{\mathrm{D}}N(0,1).$$

提示. 注意到 $X_n\sim P(n)$ 可分解为 $X_n=\sum_{k=1}^{n}Y_k$, 其中 Y_k 独立同分布于 $P(1)$.

4 设 $\{X_k\}$ 相互独立且服从指数分布, 其中 $\mathbb{E}[X_k]=\mu_k$. 证明: 若

$$\lim_{n\to+\infty}\frac{\max\{\mu_1,\cdots,\mu_n\}}{\sqrt{\sum_{k=1}^n\mu_k^2}}=0,$$

则

$$\frac{1}{\sqrt{\sum_{k=1}^n\mu_k^2}}\sum_{k=1}^n(X_k-\mu_k)\xrightarrow{D}N(0,1).$$

提示. 验证 Linderberg 条件或者 Lyapunov 条件. 以验证 Linderberg 条件为例, $B_n=\sqrt{\sum_{k=1}^n\mu_k^2}$. 对 $\forall\varepsilon>0$, 一方面,

$$\int_{\mu_k+\varepsilon B_n}^{+\infty}(x_k-\mu_k)^2\frac{1}{\mu_k}\mathrm{e}^{-\frac{x_k}{\mu_k}}\mathrm{d}x_k\xlongequal{y_k=\frac{x_k-\mu_k}{\mu_k}}\frac{\mu_k^2}{\mathrm{e}}\int_{\frac{\varepsilon B_n}{\mu_k}}^{+\infty}y_k^2\mathrm{e}^{-y_k}\mathrm{d}y_k\leqslant\frac{\mu_k^2}{\mathrm{e}}\int_{\frac{\varepsilon B_n}{\max\mu_k}}^{+\infty}y_k^2\mathrm{e}^{-y_k}\mathrm{d}y_k$$

所以

$$\frac{1}{B_n^2}\sum_{k=1}^n\int_{\mu_k+\varepsilon B_n}^{+\infty}(x_k-\mu_k)^2\frac{1}{\mu_k}\mathrm{e}^{-\frac{x_k}{\mu_k}}\mathrm{d}x_k=\frac{1}{\mathrm{e}}\int_{\frac{\varepsilon B_n}{\max\mu_k}}^{+\infty}y_k^2\mathrm{e}^{-y_k}\mathrm{d}y_k\xrightarrow{n\to+\infty}0.$$

另一方面, 存在 $N\in\mathbb{N}^*$, 当 $n>N$ 时, 有 $\frac{\max\mu_k}{B_n}<\varepsilon$, 即 $\mu_k-\varepsilon B_n<0$. 因此 Lindeberg 条件成立.

3 用 Lindeberg 替换术证明 Linderberg-Feller 中心极限定理 *

[1] $X_1,X_2,\cdots,X_n$ 相互独立, $\mathbb{E}[X_k]=0,b_k^2=Var(X_k),B_n=\sqrt{b_1^2+\cdots+b_n^2}$.

Linderberg 条件 (1922):

$$\forall\epsilon>0,\ \lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^n\int_{|x|>\epsilon B_n}x^2\mathrm{d}F_k=0\tag{L}$$

Linderberg-Feller CLT: 设 $\{X_i\}$ 相互独立, 满足(L), 则

$$\frac{1}{B_n^2}\max_{1\le k\le n}b_k^2\to0\tag{Feller}$$

且

$$\frac{1}{B_n}\sum_{k=1}^nX_k\xrightarrow{D}N(0,1)$$

1922 年 Linderberg 观察到

(1) 当 $\{Y_k\}$ 为独立正态随机变量, $\mathbb{E}[Y_k]=0,Var(Y_k)=b_k^2$ 时,

$$\mathbb{E}\left[g\left(\frac{1}{B_n}\sum_{k=1}^nY_k\right)\right]=\mathbb{E}[g(G)]+o(1),G\sim N(0,1)$$

其中 g 是一个好的测试函数, 如有界连续函数.

[1]可参考 Tao 的讲义:
https://terrytao.wordpress.com/2010/01/05/254a-notes-2-the-central-limit-theorem/
https://terrytao.wordpress.com/2015/11/02/275a-notes-4-the-central-limit-theorem/
这两篇 notes 包含了 Lindeberg Replacement Tricks 和 Moments Method in CLT

(2) 对 $\{X_k\}$ 三阶矩存在时,

$$\mathbb{E}\left[g\left(\frac{1}{B_n}\sum_{k=1}^{n}X_k\right)\right]=\mathbb{E}\left[g\left(\frac{1}{B_n}\sum_{k=1}^{n}Y_k\right)\right]+o(1)$$

(3) 使用截断术去掉三阶矩假设

回顾: $X_n\xrightarrow{D}X\Leftrightarrow\forall g\in C_{\mathrm{b}}(\mathbb{R})$(有界连续函数全体), $\mathbb{E}[g(X_n)]\to\mathbb{E}[g(X)]\Leftrightarrow\forall g,g',g'',g'''\in C_{\mathrm{b}}(\mathbb{R}),\mathbb{E}[g(X_n)]\to\mathbb{E}[g(X)]$

引入 $\{Y_k\}$ 独立正态随机变量, 且与 $\{X_k\}$ 独立, $\mathbb{E}[Y_k]=0,b_k^2=Var(Y_k)$ (思考: 两个分布函数 F,G, 试 CHECK 存在独立随机变量 X 与 Y, 使得对应分布函数为 F,G)

令 $\zeta_k=X_1+\cdots+X_{k-1}+Y_{k+1}+\cdots+Y_n$, 则

$$\zeta_n+X_n=\sum_{k=1}^{n}X_k:=S_n$$

$$\zeta_1+Y_1=\sum_{k=1}^{n}Y_k\overset{D}{=\!=}B_nY,Y\sim N(0,1)$$

更一般 $1\le k<n$ 有

$$\zeta_k+X_k=\zeta_{k+1}+Y_{k+1}$$

$$\mathbb{E}\left[g\left(\frac{S_n}{B_n}\right)\right]-\mathbb{E}[g(Y)]=\sum_{k=1}^{n}\left(\mathbb{E}\left[g\left(\frac{\zeta_k+X_k}{B_n}\right)\right]-\mathbb{E}\left[g\left(\frac{\zeta_k+Y_k}{B_n}\right)\right]\right)$$

相当于每一步做替换 $X_n\to Y_n,X_{n-1}\to Y_{n-1},\cdots,X_1\to Y_1$, 只要使总的误差足够小即可.

观察: 因为 ζ_k,X_k,Y_k 独立, 有

$$\mathbb{E}\left[g'\left(\frac{\zeta_k}{B_n}\right)(X_k-Y_k)\right]=0$$

$$\mathbb{E}\left[g''\left(\frac{\zeta_k}{B_n}\right)(X_k^2-Y_k^2)\right]=0$$

引入 $h(t)=\sup_x\{g(x+t)-g(x)-g'(x)t-\dfrac{1}{2}g''(x)t^2\}$ 则易知 $\exists K>0$, 使得 $h(t)\le Kt^2\wedge|t|^3$ 因此, 对 $g\left(\dfrac{\zeta_k+X_k}{B_n}\right),g\left(\dfrac{\zeta_k+Y_k}{B_n}\right)$ 在 $\dfrac{\zeta_k}{B_n}$ 点处做 Taylor 展开, 常数项、一次项、二次项均消去, 得

$$\left|\mathbb{E}\left[g\left(\frac{\zeta_k+X_k}{B_n}\right)-g\left(\frac{\zeta_k+Y_k}{B_n}\right)\right]\right|\le\mathbb{E}\left[h\left(\frac{X_k}{B_n}\right)\right]+\mathbb{E}\left[h\left(\frac{Y_k}{B_n}\right)\right]$$

只须证明:

$$\sum_{k=1}^{n}\mathbb{E}\left[h\left(\frac{X_k}{B_n}\right)\right]\to0\tag{1}$$

$$\sum_{k=1}^{n}\mathbb{E}\left[h\left(\frac{Y_k}{B_n}\right)\right]\to0\tag{2}$$

先证明 (2): 由 $h(t) \le K|t|^3$

$$\begin{aligned}
\sum_{k=1}^{n}\mathbb{E}\left[h\left(\frac{Y_k}{B_n}\right)\right] &\le K\sum_{k=1}^{n}\mathbb{E}\left[\left|\frac{Y_k}{B_n}\right|^3\right] \\
&= K\sum_{k=1}^{n}\frac{b_k^3}{B_n^3}\mathbb{E}[|Y|^3] \\
&\le K\max_{1\le k\le n} b_k\sum_{k=1}^{n}\frac{b_k^2}{B_n^3}\mathbb{E}[|Y|^3] \\
&= K\max_{1\le k\le n} b_k\frac{1}{B_n}\mathbb{E}[|Y|^3] \to 0 \text{ (Feller)}
\end{aligned}$$

再来证明(1): 为了使用 Linderberg 条件, 做划分 $\{|X_k| \le \epsilon B_n\}, \{|X_k| > \epsilon B_n\}$, 对前一部分用三次项控制, 后一部分用二次项控制

$$\begin{aligned}
\sum_{k=1}^{n}\mathbb{E}\left[h\left(\frac{X_k}{B_n}\right)\right] &= \sum_{k=1}^{n}\mathbb{E}\left[h\left(\frac{X_k}{B_n}\right)\left(I_{\{|X_k|\le\epsilon B_n\}} + I_{\{|X_k|>\epsilon B_n\}}\right)\right] \\
&\le K\sum_{k=1}^{n}\frac{1}{B_n^3}\mathbb{E}\left[|X_k|^3 I_{\{|X_k|\le\epsilon B_n\}}\right] \\
&\quad + K\sum_{k=1}^{n}\frac{1}{B_n^2}\mathbb{E}\left[|X_k|^2 I_{\{|X_k|>\epsilon B_n\}}\right] \\
&\le K\sum_{k=1}^{n}\frac{\epsilon}{B_n^2}\mathbb{E}\left[|X_k|^2 I_{\{|X_k|\le\epsilon B_n\}}\right] \\
&\quad K\sum_{k=1}^{n}\frac{1}{B_n^2}\mathbb{E}\left[|X_k|^2 I_{\{|X_k|>\epsilon B_n\}}\right] \\
&\le \epsilon K + K\sum_{k=1}^{n}\frac{1}{B_n^2}\int_{|x|>\epsilon B_n} x^2 \mathrm{d}F_k \to 0
\end{aligned}$$

最后一步使用了 Linderberg 条件, 证毕.

思考题: 设 $\{X_k\}$ 是独立同随机变量, 且 $\mathbb{E}[X_k]=0, Var(X_k)=1, \mathbb{E}[X_k^3]<\infty, g\in C_b^3(\mathbb{R})$. 证明:

(1) $\displaystyle\left|\mathbb{E}\left[g\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_k\right)\right] - \mathbb{E}[g(Y)]\right| \leqslant \frac{C_1}{\sqrt{n}}.$

(2) $\displaystyle\left|\mathbb{P}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_k \leqslant t\right) - \mathbb{P}(Y\leqslant t)\right| \leqslant C_2 n^{-\frac{1}{8}}.$

其中 $Y\sim N(0,1)$, C_1, C_2 为常数.

4 选读: 定理 1.3 的证明 *

定理. 设 $\limsup_{k\to\infty}\frac{1}{2k}(\gamma_{2k})^{\frac{1}{2k}} = r < \infty$, 则至多存在一个分布函数 F, 使得

$$\gamma_k = \int x^k \mathrm{d}F, k=0,1,\cdots$$

证明. 令 $\mu_k = \displaystyle\int |x|^k \,\mathrm{d}F(x)$, 则由 Cauchy-Schwarz 不等式:

$$\mu_{2k+1} \leq \sqrt{\mu_{2k}\mu_{2k+2}}$$

因此

$$\limsup_{k\to\infty} \frac{1}{k}(\mu_k)^{\frac{1}{k}} = r < \infty$$

有如下事实:

$$e^{it} - \sum_{m=0}^{n-1} \frac{(it)^m}{m!} = \frac{i^n}{(n-1)!}\int_0^t (t-s)^{n-1} e^{is}\mathrm{d}s$$

$$\left| e^{it} - \sum_{m=0}^{n-1} \frac{(it)^m}{m!} \right| \leq \frac{|t|^n}{n!}, \forall t \in \mathbb{R}$$

于是:

$$\left| e^{i\theta x}(e^{itx} - \sum_{m=0}^{n-1} \frac{(itx)^m}{m!}) \right| \leq \frac{|tx|^n}{n!}$$

取期望:

$$\left| \phi(\theta+t) - \sum_{m=0}^{n-1} \frac{t^m}{m!}\phi^{(m)}(\theta) \right| \leq \frac{|t|^n}{n!}\mu_n$$

又当 n 充分大时, 对 $\epsilon > 0$ 有

$$\mu_n \leq (r+\epsilon)^n n^n$$

利用 $\dfrac{n^n}{n!} \leq e^n$ 知

$$\phi(\theta+t) = \phi(\theta) + \sum_{m=1}^{\infty} \frac{t^m}{m!}\phi^{(m)}(\theta) \tag{3}$$

当 $|t| < \dfrac{1}{er}$ 时, $\dfrac{|t|^n}{n!}\mu_n \leq e^n(r+\epsilon)^n |t|^n \leq \left(\dfrac{r+\epsilon}{r}\right)^n$, 故上述级数可以展开.

对 $\theta = 0, \phi(0) = 1, \phi^{(m)}(0)$ 由矩 γ_m 给出, 故决定 $\phi(t), |t| < \dfrac{1}{er}$. 由 (3), 再次取 $\theta = \pm\dfrac{1}{er}$ 时, 决定 $\phi(t), |t| < \dfrac{2}{er}$. 如此下去, 矩决定 $\phi(t)$, 进而确定 F. □