

第二十一讲：随机变量的收敛性

张鑫

Email: x.zhang.seu@foxmail.com

东南大学 数学学院

- **几乎必然收敛:** 如果

$$P(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1,$$

则称随机变量列 $\{X_n\}$ 几乎必然收敛到 X , 记作 $X_n \xrightarrow{a.e.} X$, 或 $X_n \rightarrow X, a.s.$

- **依概率收敛:** 如果对任意的 $\epsilon > 0$ 有

$$\lim_{n \rightarrow \infty} P(|X_n(\omega) - X(\omega)| > \epsilon) = 0,$$

则称 X_n 依概率收敛到 X , 记作 $X_n \xrightarrow{P} X$.

- **r 阶平均收敛:** 设 $r > 0$ 为常数, 如果随机变量 X 与 $X_n (n \geq 1)$ 的 r 阶矩皆有限, 并且有

$$\lim_{n \rightarrow \infty} E|X_n - X|^r = 0,$$

则称 $\{X_n\}$ 为 r 阶平均收敛到 X , 简称 r 阶收敛, 记作 $X_n \xrightarrow{L^r} X$.



- 若记 $E := \{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$, 则

$$E = \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{m}\} \in \mathcal{F};$$

- $X_n \xrightarrow{a.s.} X$ 等价于 $P(E) = 1$ 或 $P(\bar{E}) = 0$;
- $X_n \xrightarrow{a.s.} X$ 的充分必要条件是对任意的 $\epsilon > 0$ 有

$$\lim_{k \rightarrow \infty} P(\bigcup_{n=k}^{\infty} |X_n - X| \geq \epsilon) = 0$$

- 若 $X_n \xrightarrow{a.s.} X$, 则 $X_n \xrightarrow{P} X$.

依概率收敛不能推出几乎必然收敛



例 1.1 取 $\Omega = [0, 1)$, \mathcal{F} 为区间 $[0, 1)$ 上的 Borel 集类, P 为 Lebesgue 测度。令

$$X_{11}(\omega) = 1; \quad X_{21}(\omega) = \begin{cases} 1, & \omega \in [0, \frac{1}{2}), \\ 0, & \omega \in [\frac{1}{2}, 1); \end{cases} \quad X_{22}(\omega) = \begin{cases} 1, & \omega \in [\frac{1}{2}, 1), \\ 0, & \omega \in [0, \frac{1}{2}); \end{cases}$$

一般地, 对每个自然数 k , 将 $[0, 1)$ 等分为 k 个区间, 并取

$$X_{ki}(\omega) = \begin{cases} 1, & \omega \in [\frac{i-1}{k}, \frac{i}{k}), \\ 0, & \text{其他}, \end{cases} \quad i = 1, \dots, k$$

定义

$$Y_1 = X_{11}, Y_2 = X_{21}, Y_2 = X_{22}, Y_3 = X_{31}, \dots$$

对于随机变量序列 Y_n , 对任意的 $\epsilon > 0$ 有

$$P(|X_{ki}(\omega)| \geq \epsilon) \leq \frac{1}{k} \rightarrow 0,$$

但是收敛的点集 $E = \emptyset$, 即 Y_n 不几乎必然收敛。



引理 1.1 (Markov 不等式) 设随机变量 X 的 r 阶矩有限, 则对任意 $\epsilon > 0$ 有

$$P(|X| \geq \epsilon) \leq \frac{1}{\epsilon^r} E|X|^r$$

定理 1.1 若 $X_n \xrightarrow{L^r} X$, 则 $X_n \xrightarrow{P} X$. 反之不真.

例 1.2 仍取区间 $(0, 1)$ 上的几何概率空间 (Ω, \mathcal{F}, P) . 固定 $r > 0$, 令

$$X_n(\omega) = \begin{cases} n^{1/r}, & \omega \in (0, 1/n), \\ 0, & \omega \in [1/n, 1), \end{cases} \quad X(\omega) = 0.$$

易知 $X_n \xrightarrow{a.s.} X$, 从而 $X_n \xrightarrow{P} X$, 但 $X_n \not\xrightarrow{L^r} X$.

定义 1.1 设随机变量 X, X_1, X_2, \dots , 的分布函数分别为 $F(x), F_1(x), F_2(x), \dots$, 如果对 $F(x)$ 的任一连续点 x 均有

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \quad (1)$$

则称 $\{F_n(x)\}$ 弱收敛于 $F(x)$, 记作 $F_n(x) \xrightarrow{W} F(x)$. 此时也称 $\{X_n\}$ 按分布收敛于 X , 记作 $X_n \xrightarrow{W} X$.

注 1.1 后面我们用 C_F 表示函数 $F(x)$ 的连续点集, 即

$$C_F := \{x : x \text{ 为 } F(x) \text{ 的连续点} \}$$



例 1.3 考虑随机变量序列 $\{X_n\}$ 服从如下退化分布

$$P(X_n = \frac{1}{n}) = 1, n = 1, 2, \dots,$$

它们的分布函数分别为 $F_n(x) = \begin{cases} 0, & x < \frac{1}{n} \\ 1, & x \geq \frac{1}{n} \end{cases}$

► 则在点点收敛要求下， $\{F_n(x)\}$ 的极限函数为

$$g(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

显然上述极限不满足右连续性，即 $g(x)$ 不是一个分布函数。

► 因为 $F_n(x)$ 在 $x = 1/n$ 处有跳跃，故当 $n \rightarrow \infty$ 时，跳跃点的位置趋于 0，很自然的认为 $\{F_n(x)\}$ 应该收敛于点 $x = 0$ 的退化分布，即

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}. \text{ 但是对任意的 } n, \text{ 有 } F_n(0) = 0, \text{ 而 } F(0) = 1,$$

所以 $\lim_{n \rightarrow \infty} F_n(0) = 0 \neq 1 = F(0)$ ，而这个不收敛的点恰为 $F(x)$ 的不连续点。

定理 1.2 若 $X_n \xrightarrow{P} X$, 则 $X_n \xrightarrow{W} X$.

证明: 设随机变量 X, X_1, X_2, \dots , 的分布函数分别为 $F(x), F_1(x), F_2(x), \dots$, 只需证在 $F(x)$ 的连续点上均有 $F_n(x) \rightarrow F(x)$ 即可. 而要证明这一点我们只需要说明, 对任意的 x 均有

$$F(x-0) \leq \liminf_{n \rightarrow \infty} F_n(x) \leq \limsup_{n \rightarrow \infty} F_n(x) \leq F(x+0)$$

注意到, 对任意的 $x' < x$,

$$\begin{aligned}\{X \leq x'\} &= \{X_n \leq x, X \leq x'\} \cup \{X_n > x, X \leq x'\} \\ &\subset \{X_n \leq x\} \cup \{|X_n - X| \geq x - x'\}\end{aligned}$$

从而有 $F(x') \leq F_n(x) + P(|X_n - X| \geq x - x')$. 再由 $X_n \xrightarrow{P} X$ 知

$$F(x') \leq \liminf_{n \rightarrow \infty} F_n(x) \Rightarrow F(x-0) \leq \liminf_{n \rightarrow \infty} F_n(x)$$

类似可证, 当 $x'' > x$ 时, $\limsup_{n \rightarrow \infty} F_n(x) \leq F(x+0)$.



例 1.4 设随机变量 X 的分布列为

$$P(X = -1) = P(X = 1) = 1/2$$

令 $X_n = -X$, 则 X_n 与 X 同分布, 即 X_n 与 X 有相同的分布, 故 $X_n \xrightarrow{W} X$, 但对任意的 $\epsilon \in (0, 2)$,

$$P(|X_n - X| \geq \epsilon) = P(2|X| \geq \epsilon) = 1 \not\rightarrow 0$$

若 X_n 依分布收敛至常数 c , 则 X_n 概率收敛至常数



定理 1.3 若 c 为常数, 则 $X_n \xrightarrow{P} c \Leftrightarrow X_n \xrightarrow{W} c$

只需证明充分性即可。记 X_n 的分布函数为 $F_n(x)$, $n = 1, 2, \dots$, 常数 c

的分布函数为 $F(x) = \begin{cases} 0, & x < c, \\ 1, & x \geq c. \end{cases}$ 故对任意的 $\epsilon > 0$,

$$\begin{aligned} P(|X_n - c| \geq \epsilon) &= P(X_n \geq c + \epsilon) + P(X_n \leq c - \epsilon) \\ &\leq P(X_n > c + \epsilon/2) + P(X_n \leq c - \epsilon) \\ &= 1 - F_n(c + \epsilon/2) + F_n(c - \epsilon) \\ &\rightarrow 1 - F(c + \epsilon/2) + F(c - \epsilon) = 0 \end{aligned}$$

从而 $X_n \xrightarrow{P} c$.



定理 1.4 设 $F(x), F_n(x) (n \geq 1)$ 均为分布函数, 如果对于 $F(x)$ 的连续点集 C_F 的某个稠密子集 D 上的一切 x 均有

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \forall x \in D \quad (2)$$

则 $F_n \xrightarrow{W} F$.

证明: 由 $D \subset C_F$ 立得必要性. 下设(2)式成立. 对任何 $x \in C_F$, 取

$$y < x < z \text{ 且 } y, z \in D,$$

则有

$$F_n(y) \leq F_n(x) \leq F_n(z). \quad (3)$$

上式令 $n \rightarrow \infty$ 并利用(2)式可得

$$F(y) = \lim_n F_n(y) \leq \varliminf_n F_n(x) \leq \overline{\lim}_n F_n(x) \leq \lim_n F_n(z) = F(z).$$

再令 $y \uparrow x$ 及 $z \downarrow x$ 便得证(1)式. 即 $F_n \xrightarrow{W} F$. 证毕.



定理 1.5 (Helly 第一定理) 任意分布函数列 $F_n(x)$ 均存在子列 $F_{n(k)}(x)$ 以及一个有界、非降、右连续的函数 $F(x)$ 使得对任意的 $x \in C_F$, 均有

$$\lim_{k \rightarrow \infty} F_{n(k)}(x) = F(x).$$

注 1.2 极限函数 $F(x)$ 不一定为分布函数. 比如, 若 $a + b + c = 1$, $G(x)$ 为分布函数, 令 $F_n(x) = a1_{(x \geq n)} + b1_{(x \geq -n)} + cG(x)$, 则易知

$$F_n(x) \rightarrow F(x) = b + cG(x),$$

但

$$\lim_{x \downarrow -\infty} F(x) = b \neq 0, \quad \lim_{x \uparrow \infty} F(x) = b + c = 1 - a \neq 1$$



- 令 q_1, q_2, \dots 为有理数排列, $m_0(j) \equiv j$. 由于对每个 k ,

$$F_m(q_k) \in [0, 1], \quad \forall m \geq 1.$$

因此, 存在 $m_{k-1}(j)$ 的子列 $m_k(j)$ 使得

$$\lim_{j \rightarrow \infty} F_{m_k(j)}(q_k) = G(q_k).$$

- 令 $F_{n(k)}(x) := F_{m_k(k)}(x)$, 则

$$\lim_{k \rightarrow \infty} F_{n(k)}(q) = G(q), \quad \forall q \in \mathbb{Q}.$$

- 令 $F(x) = \inf\{G(q) : q \in \mathbb{Q}, q > x\}$, 则易知 $F(x)$ 右连续. 事实上,

$$\begin{aligned} \lim_{x_n \downarrow x} F(x_n) &= \inf\{G(q) : q \in \mathbb{Q}, q > x_n \text{ 对某个 } n\} \\ &= \inf\{G(q) : q \in \mathbb{Q}, q > x\} = F(x) \end{aligned}$$



- 右连续性证明方法 2: 对任何 $\varepsilon > 0$, 总存在 $q_k > x$ 使 $G(q_k) < F(x) + \varepsilon$, 且对于一切 $y \in (x, q_k)$ 有

$$F(x) \leq F(y) < G(q_k) < F(x) + \varepsilon.$$

先令 $y \downarrow x$, 再令 $\varepsilon \downarrow 0$ 得 $F(x) \leq F(x+0) \leq F(x)$. 右连续得证.

- 令 $x \in C_F$ 并取 $r_1, r_2, r_3 \in \mathbb{Q}$ 使得 $r_1 < r_2 < x < r_3$, 则

$$F(x) - \epsilon < F(r_1) \leq F(r_2) \leq F(x) \leq F(r_3) < F(x) + \epsilon$$

- 由于

$$F_{n(k)}(r_2) \rightarrow G(r_2) \geq F(r_1), \quad F_{n(k)}(r_3) \rightarrow G(r_3) \leq F(r_3),$$

因此, 当 k 充分大时,

$$F(x) - \epsilon < F_{n(k)}(r_2) \leq F_{n(k)}(x) \leq F_{n(k)}(r_3) < F(x) + \epsilon$$

- 综上所述可知

$$\lim_{k \rightarrow \infty} F_{n(k)}(x) = F(x), \quad x \in C_F.$$

定义 1.2 称分布函数列 $\{F_n(x)\}_{n \geq 1}$ 为胎紧的 (tight), 如果对任意的 $\epsilon > 0$, 存 M_ϵ 使得

$$\limsup_{n \rightarrow \infty} (1 - F_n(M_\epsilon) + F_n(-M_\epsilon)) = \limsup_{n \rightarrow \infty} P(|X_n| \geq M_\epsilon) \leq \epsilon$$

定理 1.6 分布函数列 $\{F_n(x)\}_{n \geq 1}$ 的子序列收敛到某个分布函数 $F(x)$ 当且仅当 $\{F_n(x)\}_{n \geq 1}$ 是胎紧的.

证明:

- ▶ 假设 $\{F_n(x)\}_{n \geq 1}$ 是胎紧的, 且 $F_{n(k)}(x) \rightarrow F(x)$. 下证 $F(x)$ 为分布函数, 即证: $F(+\infty) - F(-\infty) = 1$.
- ▶ 令 $r, s \in C_F$ 且 $r < -M_\epsilon, s > M_\epsilon$, 由 $F_n(r) \rightarrow F(r), F_n(s) \rightarrow F(s)$ 得

$$\begin{aligned} 1 - F(s) + F(r) &= \lim_{k \rightarrow \infty} (1 - F_{n(k)}(s) + F_{n(k)}(r)) \\ &\leq \lim_{k \rightarrow \infty} (1 - F_{n(k)}(M_\epsilon) + F_{n(k)}(-M_\epsilon)) \\ &\leq \limsup_{n \rightarrow \infty} (1 - F_n(M_\epsilon) + F_n(-M_\epsilon)) \leq \epsilon \end{aligned}$$

- ▶ 令 $r \rightarrow -\infty, s \rightarrow \infty$ 后再令 $\epsilon \rightarrow 0$ 可得 $F(+\infty) - F(-\infty) = 1$



- ▶ 若极限函数是分布函数, 下面用反证法证明 $\{F_n(x)\}_{n \geq 1}$ 是胎紧的.
- ▶ 若 $\{F_n(x)\}_{n \geq 1}$ 不是胎紧的, 则存在 $\epsilon > 0$ 以及子列 $n(k) \rightarrow \infty$ 使得对任意的 k 均有

$$1 - F_{n(k)}(k) + F_{n(k)}(-k) \geq \epsilon,$$

- ▶ 由 Helly 第一定理可知, $\{F_{n(k)}(x)\}_{k \geq 1}$ 存在子列 $\{F_{n(k_j)}\}_{j \geq 1} \rightarrow F(x)$
- ▶ 令 $r, s \in C_F$, $r < 0 < s$, 则

$$\begin{aligned} 1 - F(s) + F(r) &= \lim_{j \rightarrow \infty} (1 - F_{n(k_j)}(s) + F_{n(k_j)}(r)) \\ &\geq \liminf_{j \rightarrow \infty} (1 - F_{n(k_j)}(k_j) + F_{n(k_j)}(-k_j)) \geq \epsilon \end{aligned}$$

- ▶ 令 $s \rightarrow \infty, r \rightarrow -\infty$ 可知 $F(x)$ 不是分布函数, 与假设矛盾.



定理 1.7 (Helly 第二定理) 设分布函数列 $\{F_n(x)\}$ 弱收敛于分布函数 $F(x)$ 当且仅当对任何有界连续函数 $g(x)$ 有

$$\int_{\mathbb{R}} g(x) dF_n(x) \rightarrow \int_{\mathbb{R}} g(x) dF(x).$$

或等价的

$$E[g(X_n)] \rightarrow E[g(X)].$$

定理 1.8 若分布函数列 $\{F_n(x)\}_{n \geq 1}$ 弱收敛于分布函数 $F(x)$, 则存在分布函数为 $F_n(x)$ 的随机变量序列 $\{Y_n\}_{n \geq 1}$ 及分布函数为 $F(x)$ 的随机变量 Y , 使得 $Y_n \rightarrow Y, a.s..$

证明:

- ▶ 令 $\Omega = (0, 1), \mathcal{F} = \mathcal{B}((0, 1)), P$ 为 $(0, 1)$ 上的 Lebesgue 测度.
- ▶ 令

$$Y_n(y) = F_n^{-1}(y) := \sup\{x : F_n(x) < y\}$$

$$Y(y) = F^{-1}(y) := \sup\{x : F(x) < y\}$$

$$a_y = \sup\{x : F(x) < y\}, \quad b_y = \inf\{x : F(x) > y\}$$

$$\Omega_0 = \{y : (a_y, b_y) = \emptyset\}$$

- ▶ 由于 (a_y, b_y) 互不相交且每个非空区间均包含不同的有理数, 因此可知 $\Omega - \Omega_0$ 中的元素个数可数, 从而 $P(\Omega - \Omega_0) = 0$.



- ▶ 下证, 对任意的 $y \in \Omega_0$, 我们有 $F_n^{-1}(y) \rightarrow F^{-1}(y)$.
- ▶ 首先, 我们有 $\liminf_{n \rightarrow \infty} F_n^{-1}(y) \geq F^{-1}(y)$,
 - ▶ 令 $x \in C_F$ 使得 $x < F^{-1}(y)$, 则由 $y \in \Omega_0$ 知 $F(x) < y$
 - ▶ 由 $x \in C_F$ 以及 $F_n(x) \rightarrow F(x)$ 可知, 当 n 充分大时, $F_n(x) < y$, 也即 $F_n^{-1}(y) \geq x$
 - ▶ 令 $n \rightarrow \infty$ 取下极限后再令 $x \rightarrow F^{-1}(y)$ 可得本结果.
- ▶ 其次, 我们可证 $\limsup_{n \rightarrow \infty} F_n^{-1}(y) \leq F^{-1}(y)$
 - ▶ 令 $x \in C_F$ 使得 $x > F^{-1}(y)$, 则由 $y \in \Omega_0$ 知 $F(x) > y$
 - ▶ 由 $x \in C_F$ 以及 $F_n(x) \rightarrow F(x)$ 可知, 当 n 充分大时, $F_n(x) > y$, 也即 $F_n^{-1}(y) \leq x$
 - ▶ 令 $n \rightarrow \infty$ 取上极限后再令 $x \rightarrow F^{-1}(y)$ 可得本结果.



► 必要性

- 由引理可知, 存在 $\{Y_n\}_{n \geq 1}$ 及 Y 使得 $Y_n \rightarrow Y, a.s.$
- 由于 $g(x)$ 是连续函数, 从而 $g(Y_n) \rightarrow g(Y), a.s.$
- 从而由控制收敛定理可知

$$E[g(X_n)] = E[g(Y_n)] \rightarrow E[g(Y)] = E[g(X)],$$

也即

$$\int_{\mathbb{R}} g(x) dF_n(x) \rightarrow \int_{\mathbb{R}} g(x) dF(x).$$



► 充分性

► 令

$$g_{x,\epsilon}(y) = \begin{cases} 1 & y \leq x \\ 0 & y \geq x + \epsilon \\ \text{linear} & x \leq y \leq x + \epsilon \end{cases}$$

► 由 $g_{x,\epsilon}$ 的定义及连续性可知

$$\limsup_{n \rightarrow \infty} P(X_n \leq x) \leq \limsup_{n \rightarrow \infty} E g_{x,\epsilon}(X_n) = E g_{x,\epsilon}(X) \leq P(X \leq x + \epsilon)$$

$$\liminf_{n \rightarrow \infty} P(X_n \leq x) \geq \liminf_{n \rightarrow \infty} E g_{x-\epsilon,\epsilon}(X_n) = E g_{x-\epsilon,\epsilon}(X) \geq P(X \leq x - \epsilon)$$

► 令 $\epsilon \rightarrow 0$ 可得

$$\limsup_{n \rightarrow \infty} P(X_n \leq x) \leq P(X \leq x),$$

$$\liminf_{n \rightarrow \infty} P(X_n \leq x) \geq P(X < x) \stackrel{x \in C_F}{=} P(X \leq x)$$

► 由上可得, 对任意的 $x \in C_F$, 我们有

$$\lim_{n \rightarrow \infty} F_n(x) = F(x).$$

定理 1.9(连续性定理) 分布函数列 $\{F_n(x)\}$ 弱收敛到分布函数 $F(x)$ 的充分必要条件是: 相应的特征函数列 $\{\varphi_n(t)\}$ 逐点收敛到相应的特征函数 $\varphi(t)$.

证明:

- 必要性显然. 事实上, 由 $\cos(tx), \sin(tx)$ 为有界连续函数以及 Helly 第二定理可知

$$\begin{aligned}\varphi_n(t) &:= \int_{\mathbb{R}} e^{itx} dF_n(x) \\&= \int_{\mathbb{R}} \cos(tx) dF_n(x) + i \int_{\mathbb{R}} \sin(tx) dF_n(x) \\&\rightarrow \int_{\mathbb{R}} \cos(tx) dF(x) + i \int_{\mathbb{R}} \sin(tx) dF(x) \\&= \int_{\mathbb{R}} e^{itx} dF(x) =: \varphi(t).\end{aligned}$$

下证充分性. 先证 $\{F_n(x)\}_{n \geq 1}$ 是胎紧的.

► 注意到

$$\begin{aligned} u^{-1} \int_{-u}^u (1 - \varphi_n(t)) dt &= u^{-1} \int_{-u}^u \int_{\mathbb{R}} (1 - e^{itx}) dF_n(x) dt \\ &= u^{-1} \int_{\mathbb{R}} \int_{-u}^u (1 - e^{itx}) dt dF_n(x) = 2 \int_{\mathbb{R}} \left(1 - \frac{\sin ux}{ux}\right) dF_n(x) \\ &\geq 2 \int_{|x| \geq \frac{2}{u}} \left(1 - \frac{1}{|ux|}\right) dF_n(x) \geq \int_{|x| \geq \frac{2}{u}} dF_n(x) = P(|X_n| \geq \frac{2}{u}) \end{aligned}$$

► 另一方面,

$$\lim_{n \rightarrow \infty} u^{-1} \int_{-u}^u (1 - \varphi_n(t)) dt = u^{-1} \int_{-u}^u (1 - \varphi(t)) dt \xrightarrow{u \rightarrow 0} 0$$

► 从而

$$P(|X_n| \geq \frac{2}{u}) \leq u^{-1} \int_{-u}^u (1 - \varphi_n(t)) dt \leq 2\epsilon$$

其次, 证明对任意的有界连续函数 $g(x)$ 均有 $E[g(X_n)] \rightarrow E[g(X)]$.

- ▶ 首先注意到, 对于 $\{F_n(x)\}_{n \geq 1}$ 的任一子列 $\{F_{n(k)}(x)\}_{k \geq 1}$, 若 $F_{n(k)}(x) \xrightarrow{W} F(x)$, 则分布函数 $F(x)$ 的特征函数必为 $\varphi(t)$.
- ▶ 由 $\{F_n(x)\}_{n \geq 1}$ 的胎紧性可知, $\{F_n(x)\}_{n \geq 1}$ 的任一子列 $\{F_{n(k)}(x)\}_{k \geq 1}$ 均存在子列 $\{F_{n(k_j)}(x)\}_{j \geq 1}$ 弱收敛到某个分布函数 $F(x)$, 且 $F(x)$ 的特征函数为 $\varphi(t)$.
- ▶ 由于 $\{F_{n(k_j)}(x)\}_{j \geq 1}$ 弱收敛到某个分布函数 $F(x)$, 故由 Helly 第二定理可知, 对任意的有界连续函数 $g(x)$ 均有

$$\int g(x) dF_{n(k_j)}(x) \rightarrow \int g(x) dF(x)$$

- ▶ 故综上所述可知, $\int g(x) dF_n(x)$ 的任一子列 $\int g(x) dF_{n(k)}(x)$ 均存在子列 $\int g(x) dF_{n(k_j)}(x) \rightarrow \int g(x) dF(x)$. 从而, $\int g(x) dF_n(x) \rightarrow \int g(x) dF(x)$, 即

$$E[g(X_n)] \rightarrow E[g(X)].$$

1. 如果 $X_n \xrightarrow{P} X$, 且 $X_n \xrightarrow{P} Y$. 试证: $P(X=Y) = 1$.
2. 如果 $X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$. 试证: (1) $X_n + Y_n \xrightarrow{P} X + Y$; (2) $X_n Y_n \xrightarrow{P} XY$.
3. 试证: $X_n \xrightarrow{P} X$ 的充要条件为: $\lim_{n \rightarrow \infty} E \left(\frac{|X_n - X|}{1 + |X_n - X|} \right) = 0$.
4. 设随机变量序列 $\{X_n\}$ 独立同分布, 其密度函数为

$$p(x) = \begin{cases} e^{-(x-\alpha)}, & x \geq \alpha, \\ 0, & x < \alpha. \end{cases}$$

令 $Y_n = \min \{X_1, X_2, \dots, X_n\}$, 试证: $Y_n \xrightarrow{P} \alpha$.

第二十二讲：大数定律

张鑫

Email: x.zhang.seu@foxmail.com

东南大学 数学学院



- ▶ 记 S_n 为 n 重伯努利试验中事件 A 出现次数, 称 $\frac{S_n}{n}$ 为事件 A 出现的频率;
- ▶ 记 p 为一次试验中 A 发生的概率;
- ▶ 概率是频率的稳定值, 何谓稳定值? 似乎与频率的极限有关;
- ▶ 考虑 $|\frac{S_n}{n} - p|$, 如果能类似于函数列极限情形即对任意的 $\epsilon > 0$, 存在 N 使得当 $n \geq N$ 时

$$|\frac{S_n}{n} - p| < \epsilon, \forall \omega \in \Omega.$$

- ▶ 但是上述情形无法做到, 事实上, 对任意的 $\epsilon < 1 - p$,

$$\begin{aligned} P(|\frac{S_n}{n} - p| < \epsilon) &\leq P(\frac{S_n}{n} - p < \epsilon) \leq P(\frac{S_n}{n} - p < 1 - p) \\ &= P(\frac{S_n}{n} < 1) = 1 - P(\frac{S_n}{n} = 1) = 1 - P(X_1 = 1, \dots, X_n = 1) \\ &= 1 - p^n < 1 \end{aligned}$$

定理 1.10 设 S_n 为 n 重伯努利试验中事件 A 出现次数, p 为一次试验中 A 发生的概率, 则对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|\frac{S_n}{n} - p| < \epsilon) = 1, \quad \text{或等价的} \quad \lim_{n \rightarrow \infty} P(|\frac{S_n}{n} - p| \geq \epsilon) = 0$$

证明: 因为 $S_n \sim B(n, p)$, 故 $S_n \sim B(n, p), E(\frac{S_n}{n}) = p, D(\frac{S_n}{n}) = \frac{p(1-p)}{n}$.
从而由切比雪夫不等式可得

$$1 \geq P(|\frac{S_n}{n} - p| < \epsilon) \geq 1 - \frac{D(\frac{S_n}{n})}{\epsilon^2} = 1 - \frac{p(1-p)}{n\epsilon^2} \rightarrow 1 \quad n \rightarrow \infty$$

故

$$\lim_{n \rightarrow \infty} P(|\frac{S_n}{n} - p| < \epsilon) = 1$$



- ▶ 随着 n 的增大, 事件 A 发生的频率 $\frac{S_n}{n}$ 与其概率 p 的偏差大于预先给定的精度 ϵ 的可能性越来越小, 要多小有多小, 这便是频率稳定于概率的含义;
- ▶ 事实上, 由切比雪夫不等式可得

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{D\left(\frac{S_n}{n}\right)}{\epsilon^2} = \frac{p(1-p)}{\epsilon^2} \frac{1}{n}$$

- ▶ 伯努利大数定律提供了用频率来确定概率的理论依据;
- ▶ 比如要估计某种产品的不合格品率 p , 则可从该种产品中随机抽取 n 件, 当 n 很大时, 这 n 件产品中的不合格品的比例可作为不合格品率 p 的估计值.



- 注意到，在伯努利大数定律中，如果记

$$X_i = \begin{cases} 1, & \text{第} i \text{试验} A \text{发生} \\ 0, & \text{第} i \text{试验} A \text{不发生} \end{cases}, \text{ 则 } S_n = \sum_{i=1}^n X_i, \text{ 并且}$$

$$\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i, p = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i)$$

- 伯努利大数定律的结论为：对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n [X_i - E(X_i)]\right| < \epsilon\right) = 1$$

也即

$$\frac{1}{n} \sum_{i=1}^n [X_i - E(X_i)] \xrightarrow{P} 0$$



定义 1.3 设有一随机变量序列 $\{X_n\}$, 如果对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n [X_i - E(X_i)]\right| < \epsilon\right) = 1 \quad \text{即} \quad \frac{1}{n} \sum_{i=1}^n [X_i - E(X_i)] \xrightarrow{P} 0$$

则称该随机变量序列 $\{X_n\}$ 满足大数定律。如果进一步有

$$\frac{1}{n} \sum_{i=1}^n [X_i - E(X_i)] \xrightarrow{a.s.} 0$$

则称 $\{X_n\}$ 满足强大数定律。

定理 1.11 设 $\{X_n\}$ 为一列两两不相关的随机变量序列, 若每个 X_i 的方差存在且有共同的上界 M , 即 $D(X_i) \leq M, \forall i \in \mathbb{N}$, 则 $\{X_n\}$ 服从大数定律, 即对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| < \epsilon\right) = 1$$

特别的, 若 $\{X_n\}$ 独立同分布, 且 $D(X_n) < \infty$, 则 $\{X_n\}$ 服从大数定律.

证明: 注意到

$$\begin{aligned} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| \geq \epsilon\right) &\leq \frac{D\left(\frac{1}{n} \sum_{i=1}^n X_i\right)}{\epsilon^2} \\ &= \frac{1}{\epsilon^2} \frac{1}{n^2} \sum_{i=1}^n D(X_i) \\ &\leq \frac{1}{\epsilon^2} \frac{M}{n} \rightarrow 0 \end{aligned}$$



定理 1.12 设随机变量序列 $\{X_n\}$ 满足如下马尔科夫条件

$$\frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) \rightarrow 0$$

则 $\{X_n\}$ 服从大数定律, 即对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| < \epsilon\right) = 1$$

定理 1.13 设 $\{X_n\}$ 为一列独立同分布随机变量序列, 若每个 X_i 的期望存在, 则 $\{X_n\}$ 服从大数定律, 即对任意的 $\epsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)| < \epsilon) = 1.$$

特别的, 如果 $E|X_i|^k, k = 1, 2, \dots$, 存在, 则 $\{X_n^k\}$ 服从大数定律.

证明: 设 $\{X_n\}$ 独立同分布, 且 $E(X_i) = a, i = 1, \dots$, 则只需要证明

$$Y_n := \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} a, \quad n \rightarrow \infty$$

由依概率收敛的性质, 仅需证 $\varphi_{Y_n}(t) \rightarrow e^{iat}$ 即可. 事实上, 由

$$\begin{aligned} \varphi_{Y_n}(t) &= [\varphi(\frac{t}{n})]^n = [1 + \varphi'(0)\frac{t}{n} + o(\frac{t}{n})]^n = [1 + ia\frac{t}{n} + o(\frac{t}{n})]^n \\ &= \exp\left\{\frac{\ln(1 + ia\frac{t}{n} + o(\frac{t}{n}))}{1/n}\right\} \rightarrow e^{iat} \end{aligned}$$



定理 1.14 对于事件列 $\{A_j\}$, 有

► 若 $\sum_{n=1}^{\infty} P(A_n) < \infty$, 则

$$P(\limsup A_n) = \lim_{k \rightarrow \infty} P(\cup_{n=k}^{\infty} A_n) = 0$$

► 若 $\{A_n\}$ 相互独立, 则 $\sum_{n=1}^{\infty} P(A_n) = \infty$ 当且仅当

$$P(\limsup A_n) = \lim_{k \rightarrow \infty} P(\cup_{n=k}^{\infty} A_n) = 1$$

证明: 1. 注意到 $\lim_{k \rightarrow \infty} P(\cup_{n=k}^{\infty} A_n) \leq \lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} P(A_n) = 0$.

2. 由于

$$\begin{aligned} P(\cup_{n=k}^{\infty} A_n) &= \lim_{m \rightarrow \infty} P(\cup_{n=k}^m A_n) = \lim_{m \rightarrow \infty} (1 - P(\cap_{n=k}^m \bar{A}_n)) \\ P(\cap_{n=k}^m \bar{A}_n) &= \prod_{n=k}^m P(\bar{A}_n) = \prod_{n=k}^m (1 - P(A_n)) \\ &\leq \prod_{n=k}^m \exp(-P(A_n)) = \exp(-\sum_{n=k}^m P(A_n)) \xrightarrow{m \rightarrow \infty} 0 \end{aligned}$$



定理 1.15 (Borel 强大数定律) 设 S_n 为 n 重伯努利试验中事件 A 出现次数, p 为一次试验中 A 发生的概率, 则

$$\frac{S_n}{n} \xrightarrow{a.s.} p$$

分析: 记 $A_n = \{|S_n/n - p| \geq \varepsilon\}$, 则定理中的几乎必然收敛的充分必要条件是

$$P\left(\limsup_n A_n\right) = 0.$$

再 Borel-Cantelli 引理可知, 只需证明

$$\sum_{n=1}^{\infty} P(A_n) < +\infty.$$

若像证明弱大数定律那样用切贝谢夫不等式, 我们只能得到

$$P(A_n) \leq \frac{1}{\varepsilon^2} D\left(\frac{S_n}{n}\right) = \frac{pq}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$



- ▶ 用 $r = 4$ 时的马尔科夫不等式可得 $P(A_n) \leq \frac{1}{\varepsilon^4} E \left| \frac{S_n}{n} - p \right|^4$.
- ▶ 将 S_n 写为 n 个独立伯努利分布随机变量之和 $S_n = \sum_{k=1}^n X_k$, 从而

$$E \left[\left| \frac{S_n}{n} - p \right|^4 \right] = \frac{1}{n^4} \sum_{1 \leq i, j, k, l \leq n} E [(X_i - p)(X_j - p)(X_k - p)(X_l - p)].$$

- ▶ 注意 X_1, \dots, X_n 相互独立, 故上式和号中仅有

$$E [(X_i - p)^4] = pq(p^3 + q^3) \neq 0,$$

$$E [(X_i - p)^2 (X_j - p)^2] = p^2 q^2 \neq 0, (i \neq j).$$

- ▶ 注意到前者共 n 项, 后者共 $C_n^2 C_4^2 = 3n(n-1)$ 项, 故我们有

$$E \left[\left| \frac{S_n}{n} - p \right|^4 \right] = \frac{pq}{n^4} [n(p^3 + q^3) + 3pq(n^2 - n)] < \frac{1}{4n^2}.$$

- ▶ 代入马尔科夫不等式右方可得 $P(A_n) \leq \frac{1}{4n^2 \varepsilon^4}$, 定理得证.



注 1.3

- ▶ 上述定理证明, 本质上只用到独立同分布序列的四阶中心矩 $E[(X_k - p)^4]$ 以及 $E[(X_k - p)^2 (X_j - p)^2]$ 有限, 并没有用到伯努利情形的其它特性. 所以, 在四阶中心矩有限的前提下, 其证明可以照搬到一般独立同分布情形.
- ▶ 在如下定理中, 我们在明显减弱的条件下, 得到一般独立同分布场合的强大数定律.
 - ▶ 证明中采用恰当选取子序列的方法, 仅用切比雪夫不等式就得到了满足 Borel-Cantelli 引理要求的估计式.
 - ▶ 而且运用著名的“截尾法”, 连切比雪夫不等式中“方差有限”的要求都可以绕过.



定理 1.16 假定 $\{X_n\}$ 是相互独立同分布的随机变量序列，如果它们有有限的期望 $E(X_n) = a$ ，则强大数定律成立，即 $n \rightarrow \infty$ 时有

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{a.s.} a.$$

- 记部分和 $S_n = \sum_{k=1}^n X_k$. 如非负情形成立, 则有

$$\frac{S_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k^+ - \frac{1}{n} \sum_{k=1}^n X_k^- \xrightarrow{a.s.} E(X_1^+) - E(X_1^-) = a.$$

故不妨假定 $\{X_k\}$ 是非负随机变量列.

- 注意到

$$P\left\{\frac{1}{n} |S_n - E(S_n)| \geq \varepsilon\right\} \leq \frac{E[\sum_{i=1}^n (X_i - E(X_i))]^k}{\varepsilon^k n^k}, k = 2, 4$$

- 高阶矩不存在, 需要截断

- 方法 1: $X_n^*(\omega) = \begin{cases} X_n(\omega), & X_n(\omega) \leq M \\ 0, & X_n(\omega) > M. \end{cases} = X_n(\omega) 1_{[0, M]}(X_n(\omega)).$
- 方法 2: $X_n^*(\omega) = \begin{cases} X_n(\omega), & X_n(\omega) \leq n, \\ 0, & X_n(\omega) > n. \end{cases} = X_n(\omega) 1_{[0, n]}(X_n(\omega)).$
- $S_n^* = \sum_{i=1}^n X_i^*$

► 对截断方法 1:

► 若 $P\{X_1 > M\} > 0$, 则

$$\sum_{n=1}^{\infty} P(X_n \neq X_n^*) = \sum_{n=1}^{\infty} P(X_n > M) = \sum_{n=1}^{\infty} P(X_1 > M) = +\infty,$$

► 无法推出 $X_n - X_n^* \rightarrow 0, a.s.$ 从而无法得到 $\frac{S_n - S_n^*}{n} \rightarrow 0, a.s.$

► 对截断方法 2:

► $\sum P(X_n \neq X_n^*) = \sum P(X_1 > n) \leq \int_0^{+\infty} P(X_1 > t) dt = E(X_1) < +\infty,$

► 故 $X_n - X_n^* \rightarrow 0, a.s.$ 更进一步有, $\frac{S_n - S_n^*}{n} \rightarrow 0, a.s.$

► 注意到

$$\frac{S_n}{n} = \frac{S_n - S_n^*}{n} + \frac{S_n^* - E(S_n^*)}{n} + \frac{E(S_n^*)}{n}$$

► $E(X_n^*) = E[X_n(\omega)1_{[0,n]}(X_n(\omega))] = E[X_1(\omega)1_{[0,n]}(X_1(\omega))] \xrightarrow{n \rightarrow \infty} E(X_1)$

► 从而 $\frac{E(S_n^*)}{n} = \frac{\sum_{i=1}^n E(X_i^*)}{n} \rightarrow E(X_1) = a$

► 只需证

$$\frac{S_n^* - E(S_n^*)}{n} \xrightarrow{a.s.} 0$$

► 注意到

$$\begin{aligned} P\left(\frac{1}{n} |S_n^* - E(S_n^*)| \geq \varepsilon\right) &\leq \frac{D(S_n^*)}{\varepsilon^2 n^2} = \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^n D(X_i^*) \\ &\leq \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^n \int_0^\infty x^2 1_{[0,i]}(x) dF(x) \leq \frac{1}{\varepsilon^2 n} \int_0^\infty x^2 1_{[0,n]}(x) dF(x) \end{aligned}$$

► 从而

$$\sum_{n=1}^\infty P\left(\frac{1}{n} |S_n^* - E(S_n^*)| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \int_0^\infty x^2 \sum_{n=1}^\infty \frac{1}{n} 1_{[0,n]}(x) dF(x)$$

► 若 $\sum_{n=1}^\infty \frac{1}{n} 1_{[0,n]}(x) < \frac{M}{x}$, 则由 X_i 期望存在可知上式小于 ∞ , 问题得证.

- ▶ 遗憾的是, 我们无法做到 $\sum_{n=1}^{\infty} \frac{1}{n} 1_{[0,n]}(x) < \frac{M}{x}$.
- ▶ 但我们如果将 n 换成特定的子列 $u_n = [\theta^n] \geq \frac{\theta^n}{2}$, 并令 N 是使 $u_n \geq x$ 的最小自然数, 我们则有可能做到

$$\sum_{n=1}^{\infty} \frac{1}{u_n} 1_{[0,u_n]}(x) = \sum_{n=N}^{\infty} \frac{1}{u_n} \leq 2 \sum_{n=N}^{\infty} \theta^{-n} = M\theta^{-N} \leq \frac{M}{x}$$

- ▶ 综上, 我们可以证明, 对任意的 $\theta > 1$ 及相应的 $u_n = [\theta^n]$, 我们均有

$$\frac{S_{u_n}^* - E(S_{u_n}^*)}{n} \xrightarrow{a.s.} 0$$

- ▶ 从而可证

$$\frac{S_{u_n}}{u_n} = \frac{S_{u_n} - S_{u_n}^*}{u_n} + \frac{S_{u_n}^* - E(S_{u_n}^*)}{u_n} + \frac{E(S_{u_n}^*)}{u_n} \xrightarrow{a.s.} a$$

- 若 $u_n \leq k \leq u_{n+1}$, 则由 X_k 的非负性知

$$\frac{u_n}{u_{n+1}} \frac{S_{u_n}}{u_n} \leq \frac{S_k}{k} \leq \frac{u_{n+1}}{u_n} \frac{S_{u_{n+1}}}{u_{n+1}}$$

- 注意 $u_{n+1}/u_n \rightarrow \theta$, 故令 n 及相应的 k 趋于无穷可得, 以概率 1 有

$$\frac{1}{\theta} \cdot a \leq \liminf_{k \rightarrow \infty} \frac{S_k}{k} \leq \limsup_{k \rightarrow \infty} \frac{S_k}{k} \leq \theta \cdot a.$$

- 上式对任意的 $\theta > 1$ 均成立, 故特别的取 $\theta = (\ell + 1)/\ell$, 则以概率 1 有

$$\frac{\ell}{\ell + 1} \cdot a \leq \liminf_{k \rightarrow \infty} \frac{S_k}{k} \leq \limsup_{k \rightarrow \infty} \frac{S_k}{k} \leq \frac{\ell + 1}{\ell} \cdot a.$$

- 对每个 $\ell \geq 1$, 记上式不成立的例外集为 E_ℓ , 且取 $E = \bigcup_{\ell=1}^{\infty} E_\ell$.
► 显然 $P(E) = 0$, 且对任意的 $\omega \notin E$, 上式对一切 $\ell \geq 1$ 成立.
► 令 $\ell \rightarrow \infty$ 便得证对任意的 $\omega \in \bar{E}$ 恒有

$$\lim_k \frac{S_k}{k} = \lim_n \frac{1}{n} \sum_{k=1}^n X_k = a$$

- ▶ 不妨假定 $\{X_k\}$ 是非负随机变量列.
- ▶ 截断: $X_n^*(\omega) = \begin{cases} X_n(\omega), & X_n(\omega) \leq n, \\ 0, & X_n(\omega) > n. \end{cases} = X_n(\omega)1_{[0,n]}(X_n(\omega)).$
- ▶ 令 $S_n^* = \sum_{i=1}^n X_i^*$ 及 $u_n = [\theta^n]$
- ▶ 证明: $\frac{1}{u_n} S_{u_n} \xrightarrow{a.s.} a.$
 - ▶ 只需证明

$$\frac{S_n - S_n^*}{n} \xrightarrow{a.s.} 0, \quad \frac{S_{u_n}^* - E(S_{u_n}^*)}{u_n} \xrightarrow{a.s.} 0, \quad \frac{E(S_n^*)}{n} \rightarrow a$$

- ▶ 由 θ 的任意性证明

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



问题 1.1 设 $0 \leq f(x) \leq 1$, 计算 $f(x)$ 在区间 $[0, 1]$ 上的积分近似值

$$J = \int_0^1 f(x) dx$$

解：注意到

$$J = \int_0^1 f(x) dx = \int_0^1 \int_0^{f(x)} dy dx = P(Y \leq f(X)) := p$$

其中 (X, Y) 为方形区域 $[0, 1] \times [0, 1]$ 上的均匀分布。此时我们有

- ▶ X, Y 分别服从 $[0, 1]$ 上的均匀分布且 X, Y 独立；
- ▶ 故计算积分的值实际上就是计算事件 $A = \{Y \leq f(X)\}$ 的概率 p ；
- ▶ (X, Y) 为方形区域 $[0, 1] \times [0, 1]$ 均匀分布，因此可以看成在方形区域内的随机投点试验；
- ▶ 由伯努利大数定律知，我们可以用重复随机投点试验中事件 A 出现的频率：即随机点 (X, Y) 落入区域 $\{(x, y) : y \leq f(x)\}$ 中的频率作为 p 的估计值。



1. 先用计算机产生 $(0, 1)$ 均匀分布的 $2n$ 个随机数, 组成 n 对随机数 $(x_i, y_i), i = 1, \dots, n$, 这里的 n 可以很大;
2. 对 n 对数据 $(x_i, y_i), i = 1, \dots, n$, 记录满足 $y_i \leq f(x_i)$ 的次数, 即事件 A 发生的频数 S_n ;
3. 根据频数 S_n , 计算频率 $\frac{S_n}{n}$, 即为积分 J 的近似值.



问题 1.2 设 $c \leq g(x) \leq d, x \in [a, b]$, 计算 $g(x)$ 在区间 $[a, b]$ 上的积分近似值

$$J_{a,b} = \int_a^b g(x) dx$$

解:

$$\begin{aligned} J_{a,b} &= \int_a^b g(x) dx \stackrel{x=a+(b-a)y}{=} \int_0^1 g(a + (b-a)y) (b-a) dy \\ &= (b-a)(d-c) \int_0^1 \left(\frac{g(a + (b-a)y) - c}{d-c} + \frac{c}{d-c} \right) dy \\ &= S_0 \int_0^1 f(y) dy + c(b-a) \end{aligned}$$

其中

$$S_0 = (b-a)(d-c), \quad f(y) = \frac{g(a + (b-a)y) - c}{d-c}$$



定理 1.17 (Kolmogorov 极大不等式) 设随机变量 X_1, \dots, X_n 相互独立, 它们都有零期望与有限方差。令 $S_n = \sum_{k=1}^n X_k$, 则对任意的 $\epsilon > 0$ 有

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} D(S_n)$$

定理 1.18 (Kolmogorov 强大数定律) 设 $\{X_n\}$ 是独立随机变量序列, 满足

$$\sum_{k=1}^{\infty} \frac{D(X_k)}{k^2} < +\infty,$$

则强大数定律成立, 即当 $n \rightarrow \infty$ 时有

$$\frac{1}{n} \sum_{k=1}^n [X_k - E(X_k)] \xrightarrow{a.s.} 0.$$

第二十三讲：中心极限定理

张鑫

Email: x.zhang.seu@foxmail.com

东南大学 数学学院



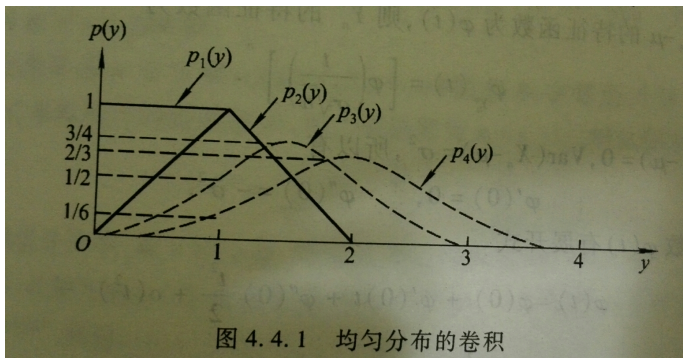
中心极限定理：问题的由来

- ▶ 实际生活中，我们会经常会进行误差分析，而误差的产生是由大量微小的相互独立的随机因素叠加而成。举个例子来说：机床加工机械轴使其直径符合规定要求，加工后的机械轴的直径会受以下各种随机因素的影响：
 - ▶ 机床振动与转速的影响；
 - ▶ 刀具装配与磨损的影响；
 - ▶ 钢材成份与产地的影响；
 - ▶ 量具的误差，测量技术的影响；
 - ▶ 车间的温度，湿度，照明，工作电压的影响等等。
- ▶ 若记每个因素使每个加工轴的直径产生的误差为 $X_i, i = 1, \dots, n$ ，则最终对加工轴直径产生的误差为

$$Y_n = \sum_{i=1} X_i$$

- ▶ 显然在实际中，我们必然要关心当 $n \rightarrow \infty$ 时 Y_n 的分布是什么？
- ▶ 为什么我们不根据 X_i 的分布用卷积公式来计算 Y_n 的分布？

例 1.5 设 $\{X_n\}$ 为独立同分布的随机变量序列，其共同的分布为区间 $(0, 1)$ 上的均匀分布。记 $Y_n = \sum_{i=1}^n X_i$ ，试计算 Y_n 的分布。





问题：寻求 Y_n 的极限分布

- 从上面均匀分布的例子可知：

$$E(Y_n) = \sum_{i=1}^n E(X_i) = \frac{n}{2} \rightarrow \infty$$

$$D(Y_n) = \sum_{i=1}^n D(X_i) = \frac{n}{12} \rightarrow \infty$$

- 考虑 Y_n 的极限分布时，需对 Y_n 进行标准化：

$$Y_n^* := \frac{Y_n - E(Y_n)}{\sqrt{D(Y_n)}}$$

- 在什么情况下

$$\lim_{n \rightarrow \infty} P(Y_n^* \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-u^2/2} du$$



定理 1.19 (林德伯格 - 莱维 (Lindeberg-Lévy) 中心极限定理) 设 $\{X_n\}$ 是独立同分布的随机变量序列, 且 $E(X_i) = \mu, D(X_i) = \sigma^2 > 0$ 存在, 若记

$$Y_n^* = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}},$$

则对任意的 $y \in R$ 有

$$\lim_{n \rightarrow \infty} P(Y_n^* \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-u^2/2} du$$



证明: 要证 Y_n^* 依分布收敛于标准正态分布, 只需要说明其特征函数收敛于标准正态分布的特征函数即可. 注意到 $Y_n^* = \sum_{i=1}^n \frac{X_i - \mu}{\sigma\sqrt{n}}$. 故若记 $Z_i := X_i - \mu$ 的特征函数为 $\varphi(t)$, 则

$$\begin{aligned}\varphi(t) &= \varphi(0) + \varphi'(0)t + \varphi''(0)\frac{t^2}{2} + o(t^2) \\ &= 1 + iE(Z_i) + i^2E(Z_i^2)\frac{t^2}{2} + o(t^2) \\ &= 1 - \frac{\sigma^2 t^2}{2} + o(t^2) \\ \varphi_{Y_n^*} &= \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) \right]^n = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n\sigma^2}\right) \right]^n \rightarrow e^{-t^2/2}\end{aligned}$$

- ▶ 在随机模拟中经常会需要产生正态随机数, 下面我们介绍一种利用中心极限定理通过 $(0, 1)$ 上的均匀分布的随机数来产生正态分布 $N(\mu, \sigma^2)$ 随机数的方法;
- ▶ 中心极限定理理论分析: 设 X_1, \dots, X_n, \dots 为一列独立同 $(0, 1)$ 上的均匀分布的随机变量序列, 则 $E(X_i) = 1/2, D(X_i) = 1/12$, 故

$$\frac{\sum_{i=1}^n X_i - n/2}{\sqrt{n/12}} \rightarrow \sim N(0, 1)$$

特别的, 若 $n = 12$, 则 $X_1 + X_2 + \dots + X_{12} - 6$ 近似服从标准正态分布.

- ▶ 产生正态随机数的方法:
 - ▶ 从计算机中产生 12 个 $(0, 1)$ 上均匀分布随机数, 记为 x_1, \dots, x_{12} ;
 - ▶ 计算 $y = x_1 + \dots + x_{12} - 6$, 由上面分析知可将 y 看成 $N(0, 1)$ 的一个随机数;
 - ▶ 计算 $z = \mu + \sigma y$, 可将 z 看成 $N(\mu, \sigma^2)$ 分布的随机数;
 - ▶ 重复上面步骤 n 次, 即可得到 n 个 $N(\mu, \sigma^2)$ 分布随机数.

定理 1.20 (棣莫弗 - 拉普拉斯中心极限定理) 设 n 重伯努利试验中, 事件 A 在每次试验中出现的概率为 $p \in (0, 1)$, 记 S_n 为 n 重伯努利试验中事件 A 出现的次数, 则对任意的 $y \in R$ 均有

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq y\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-u^2/2} du$$

在上述定理中, 若记 $\Phi(y) = \beta$, $Y_n^* := \frac{S_n - np}{\sqrt{np(1-p)}}$, 则

$$P(Y_n^* \leq y) \approx \Phi(y) = \beta$$



- ▶ 设 $\{X_n\}$ 是一个相互独立的随机变量序列，其期望与方差有限分别为

$$E(X_i) = \mu_i, \quad D(X_i) = \sigma_i^2, \quad i = 1, 2, \dots,$$

- ▶ 考虑随机变量的和 $Y_n = X_1 + \dots + X_n$ ，则

$$E(Y_n) = \mu_1 + \dots + \mu_n, \quad \sigma(Y_n) := \sqrt{D(Y_n)} = \sqrt{\sigma_1^2 + \dots + \sigma_n^2}$$

- ▶ 考虑 Y_n 的标准化: $Y_n^* := \frac{Y_n - E(Y_n)}{\sigma(Y_n)} = \sum_{i=1}^n \frac{X_i - \mu_i}{\sigma(Y_n)}$

- ▶ 要求 Y_n^* 中的各项 $\frac{X_i - \mu_i}{\sigma(Y_n)}$ 均匀的小，即对任意的 $\tau > 0$ ，要求事件

$$A_{ni} = \left\{ \frac{X_i - \mu_i}{\sigma(Y_n)} > \tau \right\} = \{|X_i - \mu_i| > \tau \sigma(Y_n)\}$$

发生的可能性小或直接要求其概率趋于 0.

- ▶ 为达目的，我们要求 $\lim_{n \rightarrow \infty} P(\max_{1 \leq i \leq n} |X_i - \mu_i| > \tau \sigma(Y_n)) = 0$

$$\lim_{n \rightarrow \infty} P(\max_{1 \leq i \leq n} |X_i - \mu_i| > \tau \sigma(Y_n)) = 0$$



注意到

$$\begin{aligned} P(\max_{1 \leq i \leq n} |X_i - \mu_i| > \tau \sigma(Y_n)) &= P(\cup_{i=1}^n |X_i - \mu_i| > \tau \sigma(Y_n)) \\ &\leq \sum_{i=1}^n P(|X_i - \mu_i| > \tau \sigma(Y_n)) = \sum_{i=1}^n \int_{|x - \mu_i| > \tau \sigma(Y_n)} dF_i(x) \\ &\leq \frac{1}{\tau^2 \sigma^2(Y_n)} \sum_{i=1}^n \int_{|x - \mu_i| > \tau \sigma(Y_n)} (x - \mu_i)^2 dF_i(x) \end{aligned}$$

故只要下面的 **林德伯格条件** 满足

$$\lim_{n \rightarrow \infty} \frac{1}{\tau^2 \sigma^2(Y_n)} \sum_{i=1}^n \int_{|x - \mu_i| > \tau \sigma(Y_n)} (x - \mu_i)^2 dF_i(x) = 0$$

就可保证 Y_n^* 中各加项充分的小

定理 1.21 设独立随机变量序列 $\{X_n\}$ 满足林德伯格条件

$$\lim_{n \rightarrow \infty} \frac{1}{\tau^2 \sigma^2(Y_n)} \sum_{i=1}^n \int_{|x - \mu_i| > \tau \sigma(Y_n)} (x - \mu_i)^2 dF_i(x) = 0$$

则对任意的 x , 有

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\sigma(Y_n)} \sum_{i=1}^n (X_i - \mu_i) \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

注 1.4 假设 $\{X_n\}$ 独立同分布且方差有限, 则容易验证上述的林德伯格条件满足, 从而我们先前在独立同分布情形下的中心极限定理是林德伯格中心极限定理的特例.



定理 1.22 设 $\{X_n\}$ 为独立随机变量序列, 如果存在常数列 $\{L_n\}$ 使得

$$\max_{1 \leq k \leq n} |X_k| \leq L_n, \text{ 且 } \lim_n \frac{L_n}{B_n} = 0$$

则 Lindeberg 条件成立, 从而 $\{X_n\}$ 满足中心极限定理.



定理 1.23 设 $\{X_n\}$ 为独立随机变量序列, 若存在 $\delta > 0$ 满足

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma(Y_n)^{2+\delta}} \sum_{i=1}^n E(|X_i - \mu_i|^{2+\delta}) = 0$$

则对任意的 x , 有

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\sigma(Y_n)} \sum_{i=1}^n (X_i - \mu_i) \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$