中国科学技术大学
University of Science and Technology of China

数理统计近似定理第一章

2025 年 3 月 9 日

对任意单变量分布函数 F，定义 p 分位数为

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}$$

引理

设 F 为一个分布函数。函数 $F^{-1}(t)$（其中 $0 < t < 1$）是单调不减且左连续的，并且满足：

(i) 对于任意 $x \in (-\infty, \infty)$，有 $F^{-1}(F(x)) \leq x$;

(ii) 对于任意 $t \in (0, 1)$ ，有 $F(F^{-1}(t)) \geq t$。

因此：

(iii) $F(x) \geq t$ 成立当且仅当 $x \geq F^{-1}(t)$，其中 $x \in (-\infty, \infty)$，$t \in (0, 1)$。

(i) 依概率收敛：$X_n \xrightarrow{P} X$ 或者 $\mathrm{p}-\lim_n X_n = X$;

(ii) 依概率 1 收敛：$X_n \xrightarrow{wp1} X$ 或者 $\mathrm{p1}-\lim_n X_n = X$，等价定义为 $\lim_{n\to\infty} P(|X_m - X| < \varepsilon,\ \text{all } m \geq n) = 1, \quad \text{each } \varepsilon > 0$;

(iii) 依 r 阶矩收敛：$X_n \xrightarrow{rth} X$ 或者 $\mathrm{L}_r-\lim_n X_n = X$，同时给定 $(\Omega, \mathcal{A}, P)$ 和 $r > 0$，记 $L_r(\Omega, \mathcal{A}, P)$ 为随机变量 Y 的空间，使得 $E|Y|^r < \infty$。在 L_r 中，通常的距离由 $d(Y, Z) = \|Y - Z\|_r$ 给出，其中

$$\|Y\|_r = \begin{cases} E|Y|^r, & 0 < r < 1, \\ [E|Y|^r]^{1/r}, & r \geq 1. \end{cases}$$

(iv) 依分布收敛：$X_n \xrightarrow{d} X$ 或者 $\mathrm{d}-\lim_n X_n = X$

一组随机变量序列 $\{X_n\}$，其相应的分布函数为 $\{F_n\}$，如果对于每个 $\varepsilon > 0$ 都存在 M_ε 和 N_ε，使得对于所有 $n > N_\varepsilon$ 都有

$$F_n(M_\varepsilon) - F_n(-M_\varepsilon) > 1 - \varepsilon$$

则称该序列在概率上有界。我们将使用记号 $X_n = O_p(1)$。很容易看出 $X_n \xrightarrow{d} X \Rightarrow X_n = O_p(1)$。

更一般地，对于两组随机变量序列 $\{U_n\}$ 和 $\{V_n\}$，记号 $U_n = O_p(V_n)$ 表示序列 $\{U_n/V_n\}$ 是 $O_p(1)$。进一步，记号 $U_n = o_p(V_n)$ 表示 $U_n/V_n \xrightarrow{p} 0$。有 $U_n = o_p(V_n) \Rightarrow U_n = O_p(V_n)$。

(i) $X_n \stackrel{wp1}{\to} X \Rightarrow X_n \stackrel{p}{\to} X$;

(ii) $X_n \stackrel{rth}{\to} X \Rightarrow X_n \stackrel{p}{\to} X$;

(iii) $X_n \stackrel{p}{\to} X \Rightarrow X_n \stackrel{d}{\to} X$;

(iv) 对于任意 $\varepsilon > 0$，如果

$$\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty$$

则 $X_n \xrightarrow{wp1} X$;

(v) 如果 $\sum_{n=1}^{\infty} E|X_n - X|^r < \infty$，则 $X_n \xrightarrow{wp1} X$。实际上，该定理的假设可以得到更强的结论，$\sum_{n=1}^{\infty} |X_n - X|^r$ wp1 收敛;

(vi) 假设 $X_n \stackrel{p}{\to} X$，对于所有 n，$|X_n| \leq |Y|$ 几乎必然成立，且 $E|Y|^r < \infty$。则 $X_n \stackrel{L^r}{\to} X$。

对于以上四种收敛，在一定条件下我们可以得到 $E[X_n^s]$ 和 $E|X_n^s|$ 的收敛性

在进行这些结果之前，我们引入三个特殊概念并研究它们的相互关系。如果

$$\lim_{c\to\infty}\sup_n E\{|Y_n|I(|Y_n|>c)\}=0$$

则称随机变量序列 $\{Y_n\}$ 是一致可积的。

定义在 $\mathscr{A}$ 上的集合函数序列 $\{Q_n\}$ 关于 $\mathscr{A}$ 上的测度 P 是一致绝对连续的，如果给定 $\varepsilon>0$，存在 $\delta>0$ 使得

$$P(A)<\delta\Rightarrow\sup_n|Q_n(A)|<\varepsilon$$

如果给定 $\varepsilon > 0$ 和 $\mathscr{A}$ 中递减到 ϕ 的序列 $\{A_n\}$，存在 M 使得

$$m > M \Rightarrow \sup_n |Q_n(A_m)| < \varepsilon$$

则称序列 $\{Q_n\}$ 在 ϕ 处是等连续的。

引理 (i) 序列 $\{Y_n\}$ 在 $(\Omega, \mathscr{A}, P)$ 上的一致可积性等价于以下两个条件:

(a) $\sup_n E|Y_n| < \infty$

(b) 由 $Q_n(A) = \int_A |Y_n| dP$ 定义的集合函数 $\{Q_n\}$ 相对于 P 是一致绝对连续的。

(ii) 使 $\{Y_n\}$ 一致可积的充分条件是

$$\sup_n E|Y_n|^{1+\epsilon} < \infty$$

对于某个 $\epsilon > 0$。

(iii) 使 $\{Y_n\}$ 一致可积的充分条件是存在一个随机变量 Y 使得 $E|Y| < \infty$ 且

$$P(|Y_n| \geq y) \leq P(|Y| \geq y), \quad \forall n \geq 1, \forall y > 0$$

(iv) 对于每个相对于测度 P 绝对连续的集合函数 Q_n，ϕ 处的等距连续性意味着相对于 P 的一致绝对连续性。

(i) 如果$X_n \xrightarrow{d} X$ 且$\{X_n^r\}$ 是一致可积的，则$E[X_n^r] \to E[X^r]$ 和$E[|X_n|^r] \to E[|X|^r]$；

(ii) 假设 $X_n \xrightarrow{rth} X$。如果 $E|X|^r < \infty$，那么 $\lim_n E\{X_n^r\} = E\{X^r\}$ 且 $\lim_n E|X_n|^r = E|X|^r$；

(iii) 假设 $X_n \xrightarrow{p} X$ 并且满足以下条件之一：

- (i) $E|X|^r < \infty$ 且 $\{X_n\}$ 是一致可积的，
- (ii) $\sup_n E|X_n|^r < \infty$ 且由 $Q_n(A) = \int_A |X_n|^r dP$ 定义的集合函数 $\{Q_n\}$ 在 Φ 处是等距连续的。

(iv) 假设 $X_n \xrightarrow{wp1} X$。如果 $\overline{\lim}_n E|X_n|^r \leq E|X|^r < \infty$，那么 $\lim_n E\{X_n^r\} = E\{X^r\}$ 且 $\lim_n E|X_n|^r = E|X|^r$。

(A) 设分布函数 $F, F_1, F_2, \ldots$ 具有相应的特征函数 $\phi, \phi_1, \phi_2, \ldots$。以下陈述是等价的：

(i) $F_n \Rightarrow F$;

(ii) $\lim_n \phi_n(t) = \phi(t)$，对于每个实数 t;

(iii) $\lim_n \int g\, dF_n = \int g\, dF$，对于每个有界连续函数 g;

(B) 设分布函数 F_n 具有有限阶矩 $\alpha_k^{(n)} = \int t^k\, dF_n(t)$ 对于 $k = 1, 2, \ldots$ 和 $n = 1, 2, \ldots$。假设极限 $\alpha_k = \lim_n \alpha_k^{(n)}$ 存在（有限），对于每个 k。则

(i) 极限 $\{\alpha_k\}$ 是分布函数 F 的矩；

(ii) 如果由 (i) 给出的 F 是唯一的，则 $F_n \Rightarrow F$。

(C) 设 $\{f_n\}$ 是绝对连续分布函数的密度序列，且 $\lim_n f_n(x) = f(x)$，对于每个实数 x。如果 f 是一个密度函数，则 $\lim_n \int |f_n(x) - f(x)|\, dx = 0$。

定理. 在 $\mathbb{R}^k$ 中，随机向量 X_n 在分布上收敛于随机向量 X 当且仅当 X_n 的各分量的每个线性组合在分布上收敛于 X 的相同线性组合的分量。

定理 (Pólya). 如果 $F_n \Rightarrow F$ 且 F 是连续的，则

$$\lim_{n\to\infty} \sup_t |F_n(t) - F(t)| = 0.$$

引理 A. 如果 X_n 是 $AN(\mu_n, \sigma_n^2)$，那么 X_n 也是 $AN(\bar{\mu}_n, \bar{\sigma}_n^2)$ 当且仅当

$$\frac{\bar{\sigma}_n}{\sigma_n} \to 1, \frac{\bar{\mu}_n - \mu_n}{\sigma_n} \to 0.$$

引理 B. 如果 X_n 是 $AN(\mu_n, \sigma_n^2)$，那么 $a_n X_n + b_n$ 也是 $AN(\mu_n, \sigma_n^2)$ 当且仅当

$$a_n \to 1, \frac{\mu_n(a_n - 1) + b_n}{\sigma_n} \to 0.$$

引理. 如果 $F_n \Rightarrow F$，那么集合

$$\{t : 0 < t < 1, F_n^{-1}(t) \not\to F^{-1}(t), n \to \infty\}$$

包含的元素至多是可数的。

对于收敛数列，我们考虑是否可以寻找到它的某种子列使得它的子列在保持原收敛性下具备其他收敛性。

定理. 如果 $X_n \xrightarrow{p} X$，那么存在一个子序列 X_{n_k}，使得 $X_{n_k} \xrightarrow{wp1} X$，当 $k \to \infty$。

考虑以下问题：已知 $X_n \xrightarrow{m} 0$，在什么情况下“平滑”序列

$$X_n^* = \frac{\sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i} \left(w_i \geq 0, \sum_{i=1}^{\infty} w_i = \infty \right)$$

几乎必然收敛？

定理. $\{X_n^*\}$ 几乎必然收敛到 0 的一个充分条件是

$$\sum_{n=1}^{\infty} \frac{E|X_n|^r}{n} < \infty$$

设 $\mathscr{B}_{[0,1]}$ 表示区间 [0,1] 中的 Borel 集，$m_{[0,1]}$ 表示限制在 [0,1] 上的 Lebesgue 测度。

定理在 $\mathbb{R}^k$ 中，假设 $X_n \xrightarrow{d} X$。那么存在定义在概率空间 $([0,1], \mathscr{B}_{[0,1]}, m_{[0,1]})$ 上的随机 k 向量 $Y, Y_1, Y_2, \ldots,$ 使得

$$\mathscr{L}(Y) = \mathscr{L}(X) \quad \text{和} \quad \mathscr{L}(Y_n) = \mathscr{L}(X_n), \quad n = 1, 2, \ldots,$$

并且

$$Y_n \xrightarrow{wp1} Y, \quad \text{即}, \quad m_{[0,1]}(Y_n \to Y) = 1.$$

定理. 设 $X_1, X_2, \ldots$ 和 X 是定义在概率空间上的随机 k-向量，g 是定义在 $\mathbb{R}^k$ 上的向量值 Borel 函数。假设 g 在 P_X-概率 1 下是连续的。那么

(i) $X_n \xrightarrow{wp1} X \Rightarrow g(X_n) \xrightarrow{wp1} g(X)$；

(ii) $X_n \xrightarrow{p} X \Rightarrow g(X_n) \xrightarrow{p} g(X)$；

(iii) $X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X)$。

定理 A. 设 $\{X_i\}$ 是独立同分布的，具有分布函数 F。存在常数 $\{a_n\}$ 使得

$$\frac{1}{n}\sum_{i=1}^{n} X_i - a_n \xrightarrow{p} 0$$

当且仅当

$$t[1-F(t)+F(-t)] \to 0, \quad t \to \infty,$$

在这种情况下，我们可以选择 $a_n = \int_{-n}^{n} x\,dF(x)$。
上述充要条件的一个充分条件是 $\int_{-\infty}^{\infty} |x|\,dF(x)$ 的有限性，但在这种情况下，以下结果断言了更强的收敛性。

定理 B (Kolmogorov). 设 $\{X_i\}$ 是独立同分布的。存在有限常数 c 使得

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{wp1} c$$

当且仅当 $E\{X_1\}$ 是有限的并且等于 c。

定理 C (Chebyshev). 设 $X_1, X_2, \ldots$ 是不相关的，具有均值 $\mu_1, \mu_2, \ldots$ 和方差 $\sigma_1^2, \sigma_2^2, \ldots$。如果 $\sum_{i=1}^{n} \sigma_i^2 = o(n^2)$，当 $n \to \infty$，则

$$\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} \mu_i \xrightarrow{p} 0.$$

定理 D (Kolmogorov). 设 $X_1, X_2, \ldots$ 是独立的，具有均值 $\mu_1, \mu_2, \ldots$ 和方差 $\sigma_1^2, \sigma_2^2, \ldots$。如果级数 $\sum_{i=1}^{\infty} \sigma_i^2/i^2$ 收敛，则

$$\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} \mu_i\right) \xrightarrow{wp1} 0.$$

定理 E. 设 $X_1, X_2, \ldots$ 具有均值 $\mu_1, \mu_2, \ldots$，方差 $\sigma_1^2, \sigma_2^2, \ldots$，并且协方差 $\mathrm{Cov}\{X_i, X_j\}$ 满足

$$\mathrm{Cov}\{X_i, X_j\} \leq \rho_{j-1}\sigma_i\sigma_j (i \leq j),$$

其中 $0 \leq \rho_k \leq 1$ 对所有 $k = 0, 1, \ldots$ 成立。如果级数 $\sum_{i=1}^{\infty} \rho_i$ 和 $\sum_{i=1}^{\infty} \sigma_i^2 (\log i)^2/i^2$ 都收敛，则定理 D 成立。

定理 A（Lindeberg-Levy）。设 $\{X_i\}$ 是独立同分布的，均值为 μ，有限方差为 σ^2。则

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \xrightarrow{d} N(0, \sigma^2),$$

即

$$\frac{1}{n}\sum_{i=1}^{n} X_i \text{ 是AN} \left(\mu, \frac{\sigma^2}{n}\right).$$

定理 B。设 $\{X_i\}$ 是独立同分布的随机向量，均值为 μ，协方差矩阵为 Σ。则

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right)\xrightarrow{d}N(0,\Sigma),$$

即，

$$\frac{1}{n}\sum_{i=1}^{n}X_i\ \text{是AN}\left(\mu,\frac{1}{n}\Sigma\right).$$

定理 C。设 $\{X_i\}$ 是独立同分布的，具有分布函数 F。则存在常数 $\{a_n\}, \{b_n\}$ 使得

$$\frac{1}{n}\sum_{i=1}^{n} X_i \text{ 是AN}(a_n, b_n)$$

当且仅当

$$\frac{t^2[1 - F(t) + F(-t)]}{U(t)} \to 0, \quad t \to \infty,$$

其中 $U(t) = \int_{-1}^{t} x^2 dF(x)$。
(条件等价于 $U(t)$ 在 ∞ 处变化缓慢的条件，即对于每个 $\alpha > 0$，$U(\alpha t)/U(t) \to 1, t \to \infty$。)

定理 A（Lindeberg-Feller）。设 $\{X_i\}$ 相互独立，具有均值 $\{\mu_i\}$，有限方差 $\{\sigma_i^2\}$，以及分布函数 $\{F_i\}$。假设 $B_n^2 = \sum_{i=1}^n \sigma_i^2$ 满足

$$\frac{\sigma_n^2}{B_n^2} \to 0, \quad B_n \to \infty, \quad \text{当} n \to \infty. \tag{1}$$

则

$$\frac{1}{n}\sum_{i=1}^n X_i \text{ 是AN} \left(\frac{1}{n}\sum_{i=1}^n \mu_i, \frac{1}{n^2}B_n^2\right)$$

当且仅当满足 Lindeberg 条件

$$\sum_{i=1}^n \int_{|t-\mu_i|>\epsilon B_n} \frac{(t-\mu_i)^2}{B_n^2}\, dF_i(t) \to 0, \quad n \to \infty, \quad \text{每个}\epsilon > 0, \tag{2}$$

推论。设 $\{X_i\}$ 相互独立，具有均值 $\{\mu_i\}$ 和有限方差 $\{\sigma_i^2\}$。假设对于某个 $\nu > 2$，

$$\sum_{i=1}^{n} E|X_i - \mu_i|^\nu = o(B_n^\nu), \quad n \to \infty.$$

则

$$\frac{1}{n}\sum_{i=1}^{n} X_i \text{ 是AN} \left(\frac{1}{n}\sum_{i=1}^{n} \mu_i, \frac{1}{n^2} B_n^2\right).$$

定理 B。设 $\{X_i\}$ 是具有均值 $\{\mu_i\}$、协方差矩阵 $\{\Sigma_i\}$ 和分布函数 $\{F_i\}$ 的独立随机向量。假设 $\frac{\Sigma_1+\cdots+\Sigma_n}{n} \to \Sigma, \quad n \to \infty$, 并且

$$\frac{1}{n}\sum_{i=1}^{n}\int_{\|x-\mu_i\|>\epsilon\sqrt{n}} \|x-\mu_i\|^2\, dF_i(x) \to 0, \quad n \to \infty, \quad \text{每个}\epsilon > 0.$$

则

$$\frac{1}{n}\sum_{i=1}^{n} X_i \text{ 是AN}\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i, \frac{1}{n}\Sigma\right).$$

在中心极限定理（CLT）中，当样本量 n 足够大时，标准化的样本和 S_n 的分布会趋近于标准正态分布 N(0,1)。然而，实际应用中，我们常常需要了解这种近似的误差大小，即 S_n 的分布与正态分布之间的差异。

定理（Berry-Esseen）。对于独立同分布（i.i.d.）的随机变量序列 $\{X_i\}$，假设它们的均值为 μ，方差为 $\sigma^2 > 0$，则对于所有 n，有

$$\sup_{t\in\mathbb{R}} |G_n(t) - \Phi(t)| \le \frac{C \cdot \mathbb{E}|X_i - \mu|^3}{\sigma^3\sqrt{n}},$$

其中，$G_n(t) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le t\right)$ 是标准化和的分布函数，$\Phi(t)$ 是标准正态分布，C 是一个常数。最初的常数 C 由 Berry 和 Esseen 给出为 $\frac{33}{4}$，后来 Zolotarev 将其降低到 0.91，van Beek 进一步优化为 0.7975。Esseen 还给出了一个“渐近最优”的常数:

$$C = \frac{3+\sqrt{10}}{6\sqrt{2\pi}} = \overline{\lim}_{n\to\infty} \sup_F \left\{ \frac{\sigma^3\sqrt{n}}{\mathbb{E}|X_1 - \mu|^3} \sup_t |G_n(t) - \Phi(t)| \right\}..$$

对于独立但不同分布的随机变量，误差上界的形式为：

$$\sup_{t\in\mathbb{R}} |G_n(t) - \Phi(t)| \le C \cdot \frac{\sum_{i=1}^{n} \mathbb{E}|X_i - \mu_i|^3}{\left(\sum_{i=1}^{n} \mathrm{Var}(X_i)\right)^{3/2}},$$

其中 C 是一个常数。

当 t 很大而 n 固定时，Berry-Esseen 定理的误差界限会变得过于粗糙。Cramér 定理研究了大偏差概率的渐近行为。假设随机变量 X_i 具有矩生成函数，并且 $t_n = o(n^{1/2})$（即 t_n 的增长速度慢于 $n^{1/2}$），则有：

$$\frac{1 - G_n(t_n)}{1 - \Phi(t_n)} \to 1 \quad (n \to \infty).$$

特别地，当 $t_n \sim c(\log n)^{1/2}$ 时，这种情形被称为“中等偏差”，Cramér 结果表明：

$$\frac{1 - G_n(t_n)}{1 - \Phi(t_n)} \to 1.$$

对于 $t_n \sim cn^{1/2}$ 的情况，Chernoff 界刻画了 $1 - G_n(t_n)$ 收敛到 0 的指数速率，这在统计应用中非常重要。

另一种解决问题的方法是细化 Berry-Esseen 界，以反映对 t 和 n 的依赖性。在这个方向上，之前的界被替换为

$$|G_n(t) - \Phi(t)| \leq C\frac{E|X_1 - \mu|^3}{\sigma^3 n^{1/2}}\frac{1}{1+t^2}, \quad \text{对所有} t,$$

其中 C 是一个常数（见 Ibragimov 和 Linnik (1971)）。同时在对涉及的分布函数有更严格的假设下，可以给出 $G_n(t) - \Phi(t)$ 的渐近展开式。例如，这种形式的一个简单结果是

$$|G_n(t) - \Phi(t)| \leq \frac{1}{n^{1/2}}\frac{E\{(X_1 - \mu)^3\}}{6\sigma^3(2\pi)^{1/2}}(1 - t^2)e^{-(1/2)t^2} + o(n^{-1/2})$$

CLT 的收敛速率有各种应用。例如，Bahadur 和 Ranga Rao (1960 年）利用这一结果建立了样本平均数的大偏差定理，该定理随后在渐近相对效率的考虑中发挥了作用。Rubin 和 Sethuraman（1965a, b）提出了“中等偏差”结果（如上所述)，并进行了类似的应用。

另一种应用涉及迭代对数定律，将在接下来讨论。

迭代对数律（Law of the Iterated Logarithm, LIL）补充了强大数定律（SLLN）和中心极限定理（CLT），描述了平均数序列或部分和序列中发生的极端波动。经典的独立同分布（i.i.d.）情形由以下定理给出：

定理 A (Hartman 和 Wintner)：设 $\{X_i\}$ 是 i.i.d.，均值为 μ，有限方差 σ^2，则

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{2\sigma^2 n \log\log n}} = 1 \quad ,$$

几乎必然成立。

换句话说：对于任意 $\epsilon > 0$，只有有限多个事件

$$\frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{2\sigma^2 n \log\log n}} > (1+\epsilon)$$

发生，然而有无限多个事件

$$\frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{2\sigma^2 n \log\log n}} > (1-\epsilon)$$

发生。

LIL 补充了 CLT，描述了随机变量序列X_i 的极端波动。CLT 表明该序列在分布上收敛于 N(0,1)，但没有提供关于这些随机变量围绕期望值 0 的波动信息。LIL 断言该序列的极端波动本质上是精确的量级 $(2\text{loglogn})^{\frac{1}{2}}$

LIL 补充了 SLLN（但假设存在第二阶矩）。就 SLLN 处理的平均数而言，LIL 断言极端波动本质上是精确的量级

$$\frac{\sigma}{\sqrt{n}}(2\log\log n)^{1/2}.$$

因此，几乎必然地，对于任意 $\epsilon > 0$，无限序列的“置信区间”

$$\left(\sum_{i=1}^{n} X_i - (1+\epsilon)\frac{\sigma}{\sqrt{n}}(2\log\log n)^{1/2}, \sum_{i=1}^{n} X_i + (1+\epsilon)\frac{\sigma}{\sqrt{n}}(2\log\log n)^{1/2}\right)$$

除了有限多个例外外，包含 μ。

对于不必然同分布的独立随机变量 $\{X_i\}$ 的 LIL 版本由 Kolmogorov 在 1929 年给出：

定理 B (Kolmogorov)：设 $\{X_i\}$ 是独立的，均值 $\{\mu_i\}$ 和有限方差 $\{\sigma_i^2\}$。假设 $B_n^2 = \sum_{i=1}^n \sigma_i^2 \to \infty$ 并且对于某个常数序列 $\{m_i\}$，几乎必然地

$$\frac{|X_n - \mu_n|}{\sigma_k} \leq m_n = o\left(\frac{B_n}{(\log\log B_n)^{1/2}}\right),$$

则

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^n (X_i - \mu_i)}{\sqrt{2B_n^2 \log\log B_n}} = 1 \ wp1.$$

定理 C（Chung）。设 $\{X_i\}$ 是独立随机变量，具有均值 $\{\mu_i\}$ 和有限方差 $\{\sigma_i^2\}$。假设 $B_n^2 = \sum_{i=1}^n \sigma_i^2 \to \infty$ 并且对于某个 $\varepsilon > 0$，有

$$\sum_{i=1}^n \frac{E|X_i - \mu_i|^3}{B_n^3} = O\left(\frac{1}{(\log B_n)^{1+\varepsilon}}\right), \quad n \to \infty.$$

则

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n (X_i - \mu_i)}{(2B_n^2 \log\log B_n)^{1/2}} = 1 \text{ wp1}.$$

定理 D（Petrov）。设 $\{X_i\}$ 是独立随机变量，$\{B_n\}$ 是满足以下条件的数列

$$B_n \to \infty, \quad \frac{B_{n+1}}{B_n} \to 1, \quad n \to \infty.$$

假设对于某个 $\varepsilon > 0$，有

$$\sup_t \left| P\left(\frac{\sum_{i=1}^n X_i}{B_n} \le t\right) - \Phi(t) \right| = O\left(\frac{1}{(\log B_n)^{1+\varepsilon}}\right), \quad n \to \infty$$

则

$$\overline{\lim}_{n\to\infty} \frac{\sum_{i=1}^n X_i}{(2B_n^2 \log\log B_n)^{1/2}} = 1 \text{ wp1}.$$

扩展 CLT 为随机过程

将经典 CLT 中的部分和序列:

$$S_n^* = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu)$$

扩展为连续函数空间中的随机过程，以研究路径依赖统计量（如极值、积分）的渐近性质。

- **离散时间点定义**：当 $t = \frac{k}{n}$ 时：

$$Y_n\left(\frac{k}{n}\right) = \frac{\sum_{i=1}^{k}(X_i - \mu)}{\sigma\sqrt{n}}$$

- **线性插值扩展**：对于 $t \in [0,1]$：

$$Y_n(t) = \frac{\sum_{i=1}^{[nt]}(X_i - \mu) + (nt - [nt])(X_{[nt]+1} - \mu)}{\sigma\sqrt{n}}$$

其中 $[nt]$ 为取整函数（最大整数部分）

空间嵌入

$Y_n(\cdot)$ 是连续函数空间 $C[0,1]$ 中的随机元素，轨迹为分段线性连续函数。

1. **与经典 CLT 的联系**：

$$Y_n(1) = S_n^* \quad (\text{直接对应 CLT 的标准化部分和})$$

2. **极值统计量表示**：

$$M_n = \sup_{0\leq t\leq 1} Y_n(t) = g(Y_n(\cdot)), \quad g(x(\cdot)) = \sup_t x(t)$$

分析框架升级

- 从单点收敛到路径空间收敛
- 允许使用泛函分析工具

理论意义

- 统一分析框架
- 推广 CLT 适用范围
- 为 Donsker 定理奠基

实际应用

- 极值分布分析
- 路径依赖统计量
- 随机波动建模

$$Y_n(t) = \begin{cases} \frac{\sum_{i=1}^{k}(X_i - \mu)}{\sigma\sqrt{n}}, & t = \frac{k}{n} \\ \text{线性插值}, & t \in \left(\frac{k}{n}, \frac{k+1}{n}\right) \end{cases}$$

$\downarrow$ 弱收敛到 Wiener 过程 $\downarrow$

基本设定

随机变量序列 $\{X_n\}$，分布函数 F_n，概率测度 P_n，极限分布 F 与 P

1. **分布函数逐点收敛**（经典 CLT 形式）:

$$\forall t \in C(F), \quad \lim_{n\to\infty} F_n(t) = F(t)$$

2. **概率测度连续性集收敛**:

$$\forall A \in \mathcal{B},\ P(\partial A) = 0,\ \lim_{n\to\infty} P_n(A) = P(A)$$

3. **连续有界函数积分收敛**:

$$\forall g \in C_b(S), \quad \lim_{n\to\infty} \int_S g\,dP_n = \int_S g\,dP$$

等价性

Billingsley (1968) 证明三者等价

核心思想

- 从 $\mathbb{R}$ 推广到任意度量空间 S
- 特别适用于函数空间 $C[0,1]$
- 定义 Borel σ-代数

随机过程应用

部分和过程收敛:

$$Y_n(\cdot) \xrightarrow{d} W(\cdot) \text{ 于 } (C[0,1], \rho)$$

其中 $\rho(x,y) = \sup_t |x(t) - y(t)|$

定理 (Donsker (1951))

设 $\{X_i\}$ 为 *IID* 序列，满足：

$$E[X_1] = \mu, \quad Var(X_1) = \sigma^2 < \infty$$

定义部分和过程：

$$Y_n(t) = \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{\lfloor nt \rfloor}(X_i - \mu) + (nt - \lfloor nt \rfloor)(X_{\lfloor nt \rfloor + 1} - \mu)\right)$$

则成立弱收敛：

$$Q_n \Rightarrow W \quad \text{于 } (C[0,1], \mathcal{B})$$

其中 W 为 *Wiener* 测度。

单点收敛 (t=1)

$$Y_n(1) = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$$

- 验证 $B_\alpha = \{x : x(1) \le \alpha\}$
- ∂B_α 的 W-测度为 0

有限维分布收敛

对任意 $0 < t_1 < \cdots < t_k \le 1$:

$$(Y_n(t_1), \ldots, Y_n(t_k)) \xrightarrow{d} (W(t_1), \ldots, W(t_k))$$

通过 CLT 多变量版本验证

泛函连续映射定理

若 $g: C[0,1] \to \mathbb{R}$ 连续，则:

$$g(Y_n) \xrightarrow{d} g(W)$$

最大值分布推导

定义泛函:

$$g(x) = \sup_{t\in[0,1]} x(t)$$

则:

$$M_n = g(Y_n) \xrightarrow{d} g(W) = \sup_t W(t)$$

通过反射原理计算:

$$P\left(\sup_t W(t) \le \alpha\right) = \sqrt{\frac{2}{\pi}} \int_0^{\alpha} e^{-u^2/2} du \quad (\alpha > 0)$$

不变性原理

- 极限分布与 $\{X_i\}$ 具体分布无关
- 只需存在有限二阶矩

其他路径泛函示例

- **波动范围**: $g(x) = \sup x(t) - \inf x(t)$
- **首达时**: $\tau_\alpha = \inf\{t : x(t) \geq \alpha\}$
- **积分泛函**: $g(x) = \int_0^1 x(t)^2 dt$

现代发展

- 非 IID 扩展: 鞅差分序列、混合序列
- 高维推广: 多参数随机过程
- 量化收敛速率: Komlós-Major-Tusnády 逼近

核心公理

Wiener 测度 W 是 $C[0,1]$ 上唯一满足以下条件的概率测度:

1. **起点固定**:

$$W(\{x(\cdot) \in C[0,1] : x(0) = 0\}) = 1$$

2. **独立增量**: 对任意 $0 \le t_0 < t_1 < \cdots < t_k \le 1$，增量

$$x(t_1) - x(t_0), \ \ldots, \ x(t_k) - x(t_{k-1})$$

   相互独立

3. **正态性**:

$$x(t) - x(s) \sim N(0, t-s) \quad (0 \le s < t \le 1)$$

存在唯一性

由 Kolmogorov 扩展定理与连续性修正保证

几乎必然性质

- 处处连续但无处可微
- 无限变差但有限二次变差

$$\lim_{n\to\infty}\sum_{i=1}^{2^n}|x(\frac{i}{2^n})-x(\frac{i-1}{2^n})|^2=1 \quad W\text{-a.s.}$$

- Hölder 连续：对任意 $\alpha<1/2$

数学奇观

- 违反直觉的路径行为
- 随机分析理论的起点

定理 (函数空间弱收敛)

设 $\{X_i\}$ 为 IID 序列，$E[X_i]=\mu$，$Var(X_i)=\sigma^2$，则标准化部分和过程：

$$Y_n(t)=\frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{\lfloor nt\rfloor}(X_i-\mu)+\text{线性项}\right)$$

满足：

$$Q_n \Rightarrow W \quad \text{在 } (C[0,1],\mathcal{B})$$

其中 Q_n 为 $Y_n(\cdot)$ 的分布。

特殊情形 → 经典 CLT

取 $t=1$ 时：

$$Y_n(1)\xrightarrow{d} N(0,1)$$

本质提升

从单点到路径空间收敛，解锁泛函分析工具

统计推断

- Kolmogorov-Smirnov 检验统计量：

$$D_n = \sup_t |F_n(t) - F(t)| \xrightarrow{d} \sup_t |W(F(t))|$$

- 极值分布计算（反射原理）

金融数学

- Black-Scholes 模型：

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

- 期权定价与风险对冲

1. **有限维分布定义**

$$(x(t_1), \ldots, x(t_k)) \sim N(0, \Sigma), \ \Sigma_{ij} = \min(t_i, t_j)$$

2. **Kolmogorov 扩展定理**将有限维分布唯一扩张到 $\mathbb{R}^{[0,1]}$
3. **连续性修正**通过 Kolmogorov 连续性定理:

$$E[|x(t) - x(s)|^{2m}] \leq C|t - s|^m \Rightarrow \text{连续修正存在}$$

- **定义**:

$$\lim_{n\to\infty} E_\theta(T_n) = g(\theta), \quad \forall\theta \in \Theta$$

- **意义**:
  - 期望随样本量增加逼近真实参数
  - 估计量合理性的基本要求

- **收敛模式**:
  - 弱: $T_n \xrightarrow{P} g(\theta)$
  - 强: $T_n \xrightarrow{\text{a.s.}} g(\theta)$
  - r 阶均方: $T_n \xrightarrow{\text{rth}} g(\theta)$
- **强相合性优势**:
  1. 长期抽样不偏离真值
  2. 存在临界点 N 保证后续估计稳定（医学案例）
  3. 区分不同估计方法的有效性

- **标准化形式**:

$$\hat{T}_n = \frac{T_n - a_n}{b_n} \xrightarrow{d} \text{非退化分布}$$

- **应用场景**:
  - 置信区间 $T_n \pm d_n$ 的近似计算
  - 假设检验临界值的确定
- **必要性**:
  - 仅相合性无法提供概率近似
  - 需结合分布收敛性

- **定义**：比较两种统计方法在样本量趋向无穷时的效率比值。若方法 A 和方法 B 在相同性能标准下所需样本量分别为 n_1, n_2，则

$$\mathrm{ARE} = \lim_{n\to\infty} \frac{n_1}{n_2}$$

- **评价标准**：
  - **方差准则**：若 T_{A_n} 和 T_{B_n} 的渐近方差分别为 $\sigma_A^2(\theta)/n_1, \sigma_B^2(\theta)/n_2$

$$\frac{\sigma_A^2(\theta)}{\sigma_B^2(\theta)}$$

  - **概率集中度**：若 $P(|T_n - g(\theta)| > \epsilon)$ 以指数速率收敛，则 ARE 由收敛速率比值决定。

高级概念

局部渐近正态性、渐近充分性、局部渐近极小极大性等未展开讨论

核心框架

上述内容已覆盖统计推断渐近分析的核心工具