概率论:Week 16

渐进理论研究随机变量及其分布的极限行为. 这是统计学分析的重要工具.

1 收敛模式

随机变量和向量存在多种收敛模式, 设$r>0$是一个常数, 对于任何$c=(c_1,\cdots,c_k)\in\mathbb{R}^k$, 定义

$$\|c\|_r=\left(\sum_{j=1}^k |c_j|^r\right)^{\frac{1}{r}}$$

如果$r\geq 1$, 那么$\|c\|_r$是0和c之间的L^r距离. 当$r=2$时, 范数的下标r可以忽略.

定义. 设$X,X_1,X_2,\cdots$是定义在概率空间上的$k-$维随机变量,

1. 我们称序列$\{X_n\}$几乎处处(a.s.)收敛到X, 记之为

$$X_n\overset{a.s.}{\to}X,$$

当且仅当 $\lim_{n\to\infty}X_n=X \quad a.s.$. 有时, 为了打印方便, 也可能记为$X_n\to_{a.s.}X$.

2. 我们称序列$\{X_n\}$依概率收敛到X, 记之为

$$X_n\overset{p}{\to}X \quad 或 \quad X_n\to_p X$$

当且仅当对于任给的固定的$\epsilon>0$, $\lim_{n\to\infty}P(\|X_n-X\|>\epsilon)=0$.

3. 我们称序列$\{X_n\}$$r$阶矩收敛(或$L^r$收敛)到$X$, 记之为

$$X_n\overset{r}{\to}X \quad 或 \quad X_n\to_{L^r}X$$

当且仅当$\lim_{n\to\infty}E\|X_n-X\|_r^r=0$, 其中$r>0$为常数.

4. 设F,F_n是$\mathbb{R}^k$上的累积分布, P,P_n是他们的概率测度, 我们称F_n弱收敛到F或者P_n弱收敛到P, 记之为$F_n\to_w F$或者$P_n\to_w P$, 当且仅当对于每个F的连续点x,

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

若$F_{X_n}\to_w F_X$, 则称X_n依分布收敛于X, 记之为$X_n\to_d X$.

注记. 前三个定义的收敛性某种程度上衡量的都是说: 当n充分大时, X_n和X作为初始概率空间上的函数十分接近. 而最后的依分布收敛仅仅依赖分布F_{X_n},F_X, 并不要求X_n和X在函数上有某种程度上的接近. 而且它仅仅要求逐点收敛性.

例 1 (P302, 例)**.** *考察随机变量$X \sim f$, 那么$I(X \le \frac{1}{n}) \sim F_n$, 一个想法是: X_n是否趋于某个随机变量, 根据以上的定义, X_n依分布收敛于$X_0 \sim I(X \le 0)$. 但是, $I(X \le n)$就不收敛了.*

注记. 关于弱收敛, 书上给出来一些分析上面的结果, 在泛函分析中, 我们也会学到弱收敛和$*-$弱收敛, 这两个课程中所谓的"弱收敛"在我看来是不同的, 但是我们在证明命题时, 往往可以用相似的证明思路.

性质. *分布函数在有理数上逐点收敛于某一函数, 那么在该函数的连续点x上, 也逐点收敛于该函数.*

特别的, 如果F是一个连续函数, 那么我们有如下的一致收敛定理:

定理 (Polya, 1920)**.** *如果 $F_n \to_w F$ 且 F 在 $\mathbb{R}^k$ 上连续, 那么*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^k} |F_n(x) - F(x)| = 0.$$

证明. 对每个固定的正整数k, 考虑以下的序列

$$-\infty = x_{0k} \le x_{1k} \le \cdots \le x_{jk} \le \cdots \le x_{kk} = \infty,$$

其中$x_{jk} = \inf\left\{x : F(x) \ge \frac{j}{k}\right\}$.

于是根据F的连续性, $F(x_{jk}) \le \frac{j}{k}$, 因此

$$0 \le F\left(x_{j+1,k}\right) - F\left(x_{jk}\right) \le \frac{1}{k}, \quad 0 \le j < k.$$

于是对于任给的x, 存在j使得$x \in \left(x_{jk}, x_{j+1,k}\right]$满足

$$F_n(x) - F(x) \le F_n\left(x_{j+1,k}\right) - F\left(x_{jk}\right) \le F_n\left(x_{j+1,k}\right) - F\left(x_{j+1,k}\right) + \frac{1}{k}$$

和

$$F_n(x) - F(x) \ge F_n\left(x_{jk}\right) - F\left(x_{j+1,k}\right) \ge F_n\left(x_{jk}\right) - F\left(x_{jk}\right) - \frac{1}{k}$$

于是

$$\sup_{x\in\left(x_{jk}x_{j+1,k}\right]} |F_n(x) - F(x)| \le \left|F_n\left(x_{j+1,k}\right) - F\left(x_{j+1,k}\right)\right| + \left|F_n\left(x_{jk}\right) - F\left(x_{jk}\right)\right| + \frac{1}{k},$$

因此

$$\sup_x |F_n(x) - F(x)| \le \max_{1\le j\le k} \left|F_n\left(x_{jk}\right) - F\left(x_{jk}\right)\right| + \max_{0\le j\le k} \left|F_n\left(x_{jk}\right) - F\left(x_{jk}\right)\right| + \frac{2}{k}.$$

先固定k对n取极限, 再对k取极限即可.

□

注记. $\sup_{x\in\mathbb{R}^k} |F_n(x) - F(x)|$往往也被称为*Kolmogorov metric*, 记之为$d_K(F_n, F)$.

定理 (Helly定理)**.** *1. 任一一致有界的非降函数列 $\{F_n(x)\}$ 中必有一子序列 $\left\{F_{n_k}(x)\right\}$ 弱收敛于某一有界的非降函数 $F(x)$.*

2. 设 $f(x)$ 是 $[a,b]$ 上的连续函数, 又 $\{F_n(x)\}$ 是在 $[a,b]$ 上弱收敛于函数 $F(x)$ 的一致有界非降函数序列, 且 a 和 b 是 $F(x)$ 的连续点, 则

$$\lim_{n\to\infty}\int_a^b f(x)\mathrm{d}F_n(x)=\int_a^b f(x)\mathrm{d}F(x)$$

定理 (Borel-Cantelli)**.** *(i)* 若随机事件序列 $\{A_n\}$ 满足

$$\sum_{n=1}^{\infty}P\left(A_n\right)<\infty$$

则

$$P\left\{\varlimsup_{n\to\infty}A_n\right\}=0,\quad P\left\{\varliminf_{n\to\infty}A_n^c\right\}=1$$

(ii) 若 $\{A_n\}$ 是相互独立的随机事件序列, 则

$$\sum_{n=1}^{\infty}P\left(A_n\right)=\infty$$

成立的充要条件为

$$P\left\{\varlimsup_{n\to\infty}A_n\right\}=1\text{ 或 }P\left\{\varliminf_{n\to\infty}A_n^c\right\}=0$$

证明. 回顾De Morgan法则, 由于

$$\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right)^c=\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}A_n^c$$
$$\left(\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}A_n\right)^c=\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n^c$$

我们有

$$\varliminf_{n\to\infty}\mathrm{A}_n^c=\left(\varlimsup_{n\to\infty}A_n\right)^c$$
$$\varlimsup_{n\to\infty}A_n^c=\left(\varliminf_{n\to\infty}A_n\right)^c$$

于是

$$\begin{aligned}P\left\{\varlimsup_{n\to\infty}A_n\right\}&=P\left\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n\right\}\\&\leqslant P\left\{\bigcup_{n=k}^{\infty}A_n\right\}\leqslant\sum_{n=k}^{\infty}P\left\{A_n\right\}\to 0\quad(k\to\infty)\end{aligned}$$

再根据$P(B^c)=1-P(B)$, 于是$P\left\{\varliminf_{n\to\infty}A_n^c\right\}=1$.

对于定理的第二部分, 先证必要性. 注意到 $\{A_n\}$ 的独立性, 有

$$\begin{aligned}P\left\{\varliminf_{n\to\infty}A_n^c\right\}&=P\left\{\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}A_n^c\right\}\leqslant\sum_{k=1}^{\infty}P\left\{\bigcap_{n=k}^{\infty}A_n^c\right\}\quad(\text{次可加性})\\&=\sum_{k=1}^{\infty}\prod_{n=k}^{\infty}P\left(A_n^c\right)=\sum_{k=1}^{\infty}\prod_{n=k}^{\infty}\left[1-P\left(A_n\right)\right]\end{aligned}$$

由于Taylor展示,

$$0 \leqslant 1 - P(A_n) \leqslant \exp\{-P(A_n)\}$$

则从

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

可得

$$\prod_{n=k}^{\infty} [1 - P(A_n)] \leqslant \lim_{N\to\infty} \exp\left\{-\sum_{n=k}^{N} P(A_n)\right\} = 0$$

所以

$$P\left\{\varliminf_{n\to\infty} A_n^c\right\} = \sum_{k=1}^{n} 0 = 0$$

再证充分性. 若 $P\left\{\varlimsup_{n\to\infty} A_n\right\} = 1$. 假定 $\sum_{n=1}^{\infty} P(A_n) < \infty$, 则由 (i)得到 $P\left\{\varlimsup_{n\to\infty} A_n\right\} = 0$, 产生矛盾. 因 $P(A_n) \geqslant 0$, 故只可能是 $\sum_{n=1}^{\infty} P(A_n) = \infty$,引理证毕. □

下面的引理给出了 a.s. 收敛一个有用的特征.

性质. 对于概率空间上的 k 维随机向量 $X, X_1, X_2, \ldots, X_n \to_{a.s.} X$,当且仅当对于每个 $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty} \{\|X_m - X\| > \epsilon\}\right) = 0.$$

证明. 不妨只证明一维的情况. 首先若 $\xi_n(\omega)(n = 1, 2, \cdots), \xi(\omega)$ 是随机变量, 则

$$\left\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} = \left\{\omega : \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \left(|\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right)\right\}$$

事实上, 按照极限的定义, 前面的集合表达的意思是: 对于任意的$\epsilon > 0$, 只要n充分大, 就有 $|\xi_n(\omega) - \xi(\omega)| < \epsilon$.因此

$$\begin{aligned}\left\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} &= \bigcap_{\epsilon\in(0,+\infty)} \varliminf_{k\to\infty} \{x \mid |f_k(x) - f(x)| < \epsilon\} \\ &= \left\{\omega : \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \left(|\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right)\right\}\end{aligned}$$

从推导的过程中(第一个等号)也可以看出, 几乎处处收敛也可以表达为: 对任意的 $\varepsilon > 0$, 成立

$$P\left\{\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} (|\xi_n(\omega) - \xi(\omega)| \geqslant \varepsilon)\right\} = 0$$

利用概率的连续性可知, 上式等价于

$$\lim_{k\to\infty} P\left\{\bigcup_{n=k}^{\infty} (|\xi_n(\omega) - \xi(\omega)| \geqslant \varepsilon)\right\} = 0$$

□

下面的结果说明了四种收敛模式之间的关系.

定理. 设$X, X_1, X_2, \cdots$是k维随机变量.

1. 如果 $X_n \to_{a.s.} X$, 那么 $X_n \to_p X$.
2. 如果对于 $r > 0$, 有 $X_n \to_{L_r} X$, 那么 $X_n \to_p X$.
3. 如果 $X_n \to_p X$, 那么 $X_n \to_d X$.
4. 如果 $X_n \to_d X$ 和 $P(X = c) = 1$, 其中 $c \in \mathcal{R}^k$ 是一个常数向量, 那么 $X_n \to_p c$.

证明. 1. 由于

$$\lim_{k\to\infty} P\left\{\bigcup_{n=k}^{\infty}(|\xi_n(\omega) - \xi(\omega)| \geqslant \varepsilon)\right\} = 0$$

以及

$$P\left\{\bigcup_{n=k}^{\infty}(|\xi_n(\omega) - \xi(\omega)| \geqslant \varepsilon)\right\} \geq P\{|\xi_k(\omega) - \xi(\omega)| \geqslant \varepsilon\} \geq 0$$

对上式两边取极限即可.

2. 根据Markov不等式, 对于任意固定的$\varepsilon > 0$,

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

对上式两边的n取极限即可.

3. 设x是$F_X(x)$的连续点, $\varepsilon > 0$, 得出

$$\begin{aligned} F_{X_n}(x) &= P(X_n \leq x) \\ &= P\big[\{X_n \leq x\} \cap \{|X_n - X| < \varepsilon\}\big] + P\big[\{X_n \leq x\} \cap \{|X_n - X| \geq \varepsilon\}\big] \\ &\leq P(X \leq x + \varepsilon) + P(|X_n - X| \geq \varepsilon) \end{aligned}$$

两边对n取极限, 根据依概率收敛, 右边的极限是存在的, 而左边未必, 因此

$$\overline{\lim_{n\to\infty}} F_{X_n}(x) \leq F_X(x + \varepsilon)$$

为了达到下界, 可以考虑

$$P(X_n > x) \leq P(X \geq x - \varepsilon) + P(|X_n - X| \geq \varepsilon)$$

同样有

$$\underline{\lim_{n\to\infty}} F_{X_n}(x) \geq F(x - \varepsilon)$$

从而

$$F_X(x - \varepsilon) \leqslant \underline{\lim_{n\to\infty}} F_n(x) \leqslant \overline{\lim_{n\to\infty}} F_n(x) \leqslant F_X(x + \varepsilon)$$

令$\varepsilon \downarrow 0$即可.

4. 设$\varepsilon > 0$给定, 于是

$$\lim_{n\to\infty} P\big[|X_n - X| \leq \varepsilon\big] = \lim_{n\to\infty} F_{X_n}(b + \varepsilon) - \lim_{n\to\infty} F_{X_n}[(b - \varepsilon) - 0] = 1 - 0 = 1$$

□

2 连续性定理

不加证明的指出Levy-Cramer连续性定理.

定理. 设$\phi_X, \phi_{X_1}, \phi_{X_2}, \cdots$分别是$X, X_1, X_2, \cdots$的特征函数(*ch.f.*)那么$X_n \to_d X$当且仅当对于所有的$t \in \mathbb{R}$, $\lim\limits_{n\to\infty} \phi_{X_n}(t) = \phi_X(t)$.

注记. 书上的结论更强, 在于逆极限定理中没有事先给出X的特征函数.

3 Landau(Big-O/Little-O)记号

比较

- O记号和确定的序列.
- O_P记号和随机的序列.

一个问题是抽样分布的"收敛速度"是多少? 一个直观的想法是: 描述抽样分布的分散情况怎样随着n的增大而减小. 我们需要严格的写出它的定义.

我们先看一些确定的序列的收敛性.

例 2 (deterministic sequences)**.**

$$A_n = \frac{1}{n}, \quad B_n = \frac{1}{n^2}$$

哪个序列是"渐进更大"的?

n	A_n	B_n
1	1	1
2	1/2	1/4
3	1/3	1/9
10	1/10	1/100
100	1/100	1/10000

显然从数据上看B_n比A_n收敛得更快, 两者都收敛到0, 而且

$$\frac{B_n}{A_n} \to 0$$

在这个意义下B_n渐进的比A_n小

定义 (Landau(Big-O/Little-O) Notation)**.** 对于两个数列, 如果$\dfrac{B_n}{A_n} \to 0$, 则称B_n渐进地比A_n小, 写作$B_n = o(A_n)$.

例 3 (数列收敛性举例)**.** 用数学分析的极限理论不难证明:

$$
\begin{aligned}
&1/n^3 = o\left(1/n^2\right)\\
&1/(n+3) = o(1/\sqrt{n})\\
&1/n^2 = o(\sin(1/n))\\
&1/\log(n) = o(1)\\
&e^{-tn} = o\left(n^{-k}\right) \textit{ for any } t, k > 0
\end{aligned}
$$

例 4 (如果不收敛到0)**.**

$$A_n = \frac{2}{n}, \quad B_n = \frac{1}{n}$$

定义. 若B_n/A_n是有界的, 则称B_n是渐进不大于A_n的, 记之为$B_n = O(A_n)$. 若$A_n = O(B_n), B_n = O(A_n)$, 则称$A_n, B_n$同阶*(of the same order)*, 记之为$B_n = \Theta(A_n)$.

"Zooming in"

定义. 给定一个随机序列Z_n和一个数列d_n, 若$Z_n/d_n \xrightarrow{P} 0$, 则称$Z_n$是渐进小于$d_n$的, 记之为$Z_n = o_p(d_n)$.

定义 (概率有界(Bounded in probability))**.** Z_n是概率有界的, 若

$$\lim_{C\to\infty} \sup_{i=1,\ldots,\infty} P(|Z_i| > C) = 0$$

也就是说, 在充分远的地方没有很大的概率.

注记. 若$Z_n = o_P(d_n)$, 则$Z_n = O_P(d_n)$.

于是根据概率有界, 若Z_n/d_n概率有界($O_P(1)$), 则称Z_n概率有界d_n.

$$
\begin{aligned}
Z_n = o_p(1) \quad &\Longleftrightarrow \quad Z_n \to_p 0\\
Z_n = O_p(1) \quad &\Longleftrightarrow \quad Z_n \text{ is bounded in probability}
\end{aligned}
$$

概率有界性往往和一分部收敛联系起来.

4　变换的收敛性

这一节中的内容在习题中出现, 这里强调其作为方法的重要性.

定理 (连续映射定理)**.** 考虑g是一个连续映射, 那么

1. $X_n \to_{a.s.} X \Rightarrow g(X_n) \to_{a.s.} g(X)$;
2. $X_n \to_p X \Rightarrow g(X_n) \to_p g(X)$
3. $X_n \to_d X \Rightarrow g(X_n) \to_d g(X)$

定理 (Slusky定理)**.** 设$X, X_1, \cdots, Y_1, \cdots$是概率空间上的随机变量, 且满足$X_n \to_d X, Y_n \to_p c$, 那么

1. $X_n + Y_n \to_d X + c$
2. $Y_n X_n \to_d cX$
3. 若$c \neq 0$, 则$\dfrac{X_n}{Y_n} \to_d \dfrac{X}{c}$

5 大数定律

- Bernoulli 大数定律
- Chebyshev 大数定律
- Markov 大数定律
- Poisson 大数定律
- Khinchin (弱)大数定律
- Kolmogorov (强)大数律

注记. 大数定律的一个直接应用是: m-估计(z-估计)的无偏性.

切比雪夫大数定律. 设 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是由两两不相关的随机变量所构成的序列, 每一随机变量都有有限的方差, 并且它们有公共上界

$$D\xi_1 \leqslant C, \quad D\xi_2 \leqslant C, \quad \cdots, \quad D\xi_n \leqslant C, \cdots$$

则对任意的 $\varepsilon > 0$, 皆有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| < \varepsilon\right\} = 1$$

马尔可夫大数定律. 对于随机变量序列 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$, 若

$$\frac{1}{n^2}D\left(\sum_{k=1}^{n}\xi_k\right) \to 0$$

成立, 则对任意 $\varepsilon > 0$, 均有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| < \varepsilon\right\} = 1$$

伯努利大数定律. 设 μ_n 是 n 次伯努利试验中事件 A 出现的次数, 而 p 是事件 A 在每次试验中出现的概率, 则对任意 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1$$

泊松大数定律. 如果在一个独立试验序列中, 事件 A 在第 k次试验中出现的概率等于 p_k, 以 μ_n 记在前 n 次试验中事件 A 出现的次数, 则对任意 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - \frac{p_1 + p_2 + \cdots + p_n}{n}\right| < \varepsilon\right\} = 1$$

目前接触到的最常用的概型是独立同分布模型, 因此以下的辛钦大数定律更为重要.

定理 (Khinchin)**.** 设 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是相互独立的随机变量序列, 它们服从相同的分布, 且具有有限的数学期望

$$a = E\xi_n$$

则对任意的 $\varepsilon > 0$, 有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right| < \varepsilon\right\} = 1$$

证明. 由于 $\xi_1, \xi_2, \cdots, \xi_n$ 具有相同分布,故有同一特征函数, 设为 $f(t)$, 因为数学期望存在, 故 $f(t)$ 可展开成

$$f(t) = f(0) + f'(0)t + o(t) = 1 + \mathrm{i}at + o(t)$$

而 $\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ 的特征函数为

$$\left[f\left(\frac{t}{n}\right)\right]^n = \left[1 + \mathrm{i}a\frac{t}{n} + o\left(\frac{t}{n}\right)\right]^n$$

对于固定的 t

$$\left[f\left(\frac{t}{n}\right)\right]^n \to \mathrm{e}^{\mathrm{i}at} \quad (n \to \infty)$$

极限函数 $\mathrm{e}^{\mathrm{i}at}$ 是连续函数, 它是退化分布 $I_a(x)$ 所对应的特征函数,由逆极限定理知 $\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ 的分布函数弱收玫于 $I_n(x)$, 再由依分布收敛到常数, 因此 $\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ 依概率收敛于常数 a, 从而证明了定理. □

独立同分布场合的强大数律. (科尔莫戈罗夫) 设 $\xi_1, \xi_2, \cdots$ 是相互独立相同分布的随机变量序列, 则

$$\frac{1}{n}\left(\xi_1 + \xi_2 + \cdots + \xi_n\right) \xrightarrow{\text{a.s.}} a$$

成立的充要条件是 $E\xi_i$ 存在且等于 a.

6 中心极限定理

不同于Poisson极限定理, 这里我们考虑的更多是一般分布的极限行为, 书上从棣莫弗-拉普拉斯中心极限定理作为背景, 但真正应用广泛的是独立同分布场合下的Lindeberg-Levy中心极限定理.

定理 (Lindeberg-Levy)**.** 若 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是一串相互独立相同分布的随机变量序列, 且

$$E\xi_k = \mu, \quad D\xi_k = \sigma^2$$

我们来讨论标准化随机变量和

$$\zeta_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (\xi_k - \mu)$$

若 $0 < \sigma^2 < \infty$, 则

$$\lim_{n\to\infty} P\{\zeta_n < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\, \mathrm{d}t$$

证明. 记 $\xi_k - \mu$ 的特征函数为 $g(t)$, 则 ζ_n 的特征函数为 $\left[g\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$. 由于 $E\xi_k = \mu, D\xi_k = \sigma^2$ 故 $g'(0) = 0, g''(0) = -\sigma^2$. 因此 - 324 .

$$g(t) = 1 - \frac{1}{2}\sigma^2 t^2 + o\left(t^2\right)$$

所以

$$\left[g\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left[1 - \frac{1}{2n}t^2 + o\left(\frac{t^2}{n}\right)\right]^n \to \mathrm{e}^{-t^{2/2}}$$

由于 $\mathrm{e}^{-t^2/2}$ 是连续函数, 它对应的分布函数为 $N(0,1)$, 因此由逆极限定理知

$$P\{\zeta_n < x\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\, \mathrm{d}t$$

定理证毕. □

对于独立但不同分布的场合, 设 $\xi_1, \xi_2 \cdots, \xi_n, \cdots$ 是一个相互独立的随机变量序列, 它们具有有限的数学期望和方差:

$$a_k = E\xi_k, \quad b_k^2 = D\xi_k \quad (k = 1, 2, \cdots, n, \cdots)$$

记

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

作标准化和数

$$\zeta_n = \sum_{k=1}^{n} \frac{\xi_k - a_k}{B_n}$$

定义 (Lindeberg条件)**.** 对于任何 $\tau > 0$, 成立

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

定义 (Feller条件)**.**

$$\lim_{n\to\infty} \max_{k\leqslant n} \frac{b_k}{B_n} = 0$$

定理 (等价性)**.** *Feller*条件的等价条件是

$$\lim_{n\to\infty} B_n = \infty$$

$$\lim_{n\to\infty} \frac{b_n}{B_n} = 0$$

定理 (Lindeberg-Feller)**.** 对和数 ζ_n, 成立

$$\lim_{n\to\infty} P\left\{\zeta_n < x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\, \mathrm{d}t$$

与费勒条件的充要条件是林德贝格条件成立.