第五节　中心极限定理

一　林德贝格条件与费勒（Feller）条件

本节将最后解决古典的中心极限定理，古典的中心极限定理讨论的是独立和的分布函数向正态分布收敛的最普遍条件.

设$X_1, X_2, \cdots, X_n, \cdots$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：

$$a_k = EX_k,\quad b_k^2 = DX_k. \quad (k = 1, 2, \cdots, n, \cdots)$$

记 $B_n^2 = \sum_{k=1}^{n} b_k^2$，　作$\sum_{k=1}^{n} X_k$的标准化

$$Y_n = \sum_{k=1}^{n} \frac{X_k - a_k}{B_n} \quad , \qquad (5.5.1)$$

我们需要寻找Y_n的分布函数趋于正态分布函数的充要条件.

现在，这个问题从某种意义上来讲已经得到了最后解决. 1922年林德贝格提出了充分条件；1935年费勒进一步指出，在某种条件下，这个条件也是必要的. 这样就搞清了向正态分布收敛的充要条件.

二　林德贝格－费勒定理

定理5.5.2　对(5.5.1)中定义的和数Y_n，成立

$$\lim_{n\to\infty} P\{Y_n < x\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt = \Phi(x)$$

与费勒条件

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

的充要条件是林德贝格条件

对于任何$\tau > 0$，有

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

成立.

费勒条件的含义：

定理5.5.1 费勒条件

$$\lim_{n\to\infty}\max_{k\le n}\frac{b_k}{B_n}=0 \qquad (5.5.3)$$

等价于

$$\lim_{n\to\infty}B_n=\infty, \qquad \lim_{n\to\infty}\frac{b_n}{B_n}=0. \qquad (5.5.4)$$

注：量$\frac{b_k}{B_n}$是Y_n的项$\frac{X_k-a_k}{B_n}$的标准差，可以看作分量X_k对总和Y_n的贡献，因此费勒条件相当于说：

总和是大量“可忽略的”分量之和.

林德贝格条件的含义

与独立同分布场合比较，这里保留了独立性的假定，但是去掉了同分布的要求. 今后我们将以$F_k(x)$记X_k的分布函数. 显然为了讨论Y_n的极限分布，要使问题的提法有意义，对各个加项必须有一定要求. 例如若允许从第二项起都等于0，则极限分布由$F_1(x)$完全确定，这时就很难有什么有意义的结果.

为了使极限分布是正态分布, 还要求各个项“均匀地小”，即任何一项在求和式中不起决定性的作用，怎样明确表达这个要求呢？下面先作一个启发性的推导.

设A_k表示下述事件：

$$\left\{\left|\frac{X_k - a_k}{B_n}\right| > \tau\right\} = \left\{\left|X_k - a_k\right| > \tau B_n\right\}, \quad k = 1, 2, \cdots, n.$$

则有

$$P\left\{\max_{1\le k\le n}\left|X_k - a_k\right| > \tau B_n\right\} = P\left\{\bigcup_{k=1}^{n}\left(\left|X_k - a_k\right| > \tau B_n\right)\right\}$$

$$= P\left\{A_1 \bigcup A_2 \bigcup \cdots \bigcup A_n\right\} \le \sum_{k=1}^{n} P(A_k)$$

$$= \sum_{k=1}^{n} \int_{\left|x - a_k\right| > \tau B_n} dF(x)$$

$$\le \frac{1}{(\tau B_n)^2} \sum_{k=1}^{n} \int_{\left|x - a_k\right| > \tau B_n} (x - a_k)^2 \, dF(x)$$

$$=\frac{1}{\tau^2 B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2\,dF(x)$$

因此，只要对于任何$\tau>0$，成立

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2\,dF(x)=0. \quad (5.5.2)$$

就可以保证总和(5.5.1)中各加项“均匀地小”. 上述条件(5.5.2)称为林德贝格条件. 林德贝格证明了条件(5.5.2)是和数(5.5.1)的分布函数趋于正态分布函数的充分条件.

但是林德贝格条件不是中心极限定理成立的必要条件(参看习题53).

三　若干推论

林德贝格条件给出了中心极限定理成立的普遍条件，由它可以推出许多特殊的结果. 容易证明独立同分布场合的林德贝格－莱维定理是定理**5.5.2**的特例. 下面我们再来给出两个有用的结果.

定理5.5.3　若$X_1, X_2, \cdots, X_n, \cdots$是独立随机变量序列，存在常数$K_n$，使

$$\max_{1\le j\le n}|X_j| \le K_n,\ (n=1,2,\cdots) \text{且} \lim_{n\to\infty}\frac{K_n}{B_n}=0,$$

则
$$P\left\{\sum_{k=1}^{n}\frac{X_k-a_k}{B_n}<x\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt.$$

定理5.5.4（李雅普诺夫）如果对相互独立的随机变量序列$X_1, X_2, \cdots, X_n, \cdots$能选择这样一个正数$\delta > 0$，使当$n \to \infty$时，

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} E\left|X_k - a_k\right|^{2+\delta} \to 0,$$

则

$$P\left\{\frac{1}{B_n} \sum_{k=1}^{n} (X_k - a_k) < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$