

第五章 极限定理

- 第一节 收敛性
- 第二节 大数定律
- 第三节 中心极限定理

§ 5.1 收敛性

- 一、分布函数弱收敛
- 二、连续性定理
- 三、随机变量的收敛性

一、分布函数弱收敛

例1 令

$$F_n(x) = \begin{cases} 0, & x \leq -1/n \\ 1, & x > -1/n \end{cases}$$

这是一个退化分布，它可以解释为一个单位质量全部集中在 $x = -1/n$ 的分布。

当 $n \rightarrow \infty$ 时，我们自然认为 $\{F_n(x)\}$ 应该收敛于一个单位质量全部集中在 $x = 0$ 这一点的分布，即

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

但是， $F_n(0) \equiv 1$ ，而 $F(0) = 0$ ，显然， $\lim_{n \rightarrow \infty} F_n(0) \neq F(0)$ 。

因此，看来要求分布函数列在所有的点都收敛到极限分布函数是太严了。

上例中不收敛的点是极限分布函数 $F(x)$ 的不连续点，于是我们提出如下定义。

定义 对于分布函数列 $\{F_n(x)\}$ ，如果存在一个非降函数 $F(x)$ 使

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

在 $F(x)$ 的每一连续点上都成立，则称 $F_n(x)$ 弱收敛于 $F(x)$ 并记为 $F_n(x) \xrightarrow{w} F(x)$ 。

这样得到的极限函数是一个有界的非降函数，但下例说明它不一定是一个分布函数。

分布函数列的弱收敛极限不一定是分布函数。

例2 取 $F_n(x) = \begin{cases} 0, & x \leq n \\ 1, & x > n \end{cases}$

显然， $\lim_{n \rightarrow \infty} F_n(x) = 0$ 对一切 x 成立。

但 $F(x) = 0$ 不是分布函数。

当然，当 $F(x)$ 是一个分布函数时，分布函数的左连续性保证了 F 在不连续点上的值完全由它在连续点集 C_F 上的值唯一确定，因此此时分布函数列的弱收敛极限是唯一的。

二、连续性定理 (Levy-Cramer)

下面我们将导出一个分布函数列弱收敛到一个极限分布的充要条件。这个结果同时说明了存在于分布函数和特征函数之间的对应是连续的，这个性质对于特征函数成为研究极限定理的主要工具有基本的重要性。

定理 5.2.4 (正极限定理)

设分布函数列 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$ ，则相应的特征函数列 $\{f_n(t)\}$ 收敛于特征函数 $f(t)$ ，且在 t 的任一有限区间内收敛是一致的。

定理2 (逆极限定理)

设特征函数列 $\{f_n(t)\}$ 收敛于某一函数 $f(t)$ ，且 $f(t)$ 在 $t=0$ 连续，则相应的分布函数列 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$ ，而 $f(t)$ 是 $F(x)$ 的特征函数。

通常把正逆极限定理合称为连续性定理。

连续性定理 (Levy-Cramer)

分布函数列 $F_n(x)$ 弱收敛到某一个分布函数 $F(x)$ ，当且仅当 $F_n(x)$ 对应的特征函数列 $f_n(t)$ 收敛到 $F(x)$ 的特征函数 $f(t)$ 。

三、随机变量的收敛性

定义1. (依分布收敛)

设随机变量 $\xi_n(\omega), \xi(\omega)$ 的分布函数分别为 $F_n(x)$ 及 $F(x)$

如果

$$F_n(x) \xrightarrow{w} F(x)$$

则称 $\{\xi_n(\omega)\}$ 依分布收敛于 $\xi(\omega)$ ，记为

$$\xi_n(\omega) \xrightarrow{L} \xi(\omega)$$

或

$$\xi_n(\omega) \xrightarrow{d} \xi(\omega)$$

定义2 (依概率收敛)

如果对任意 $\varepsilon > 0$ ，有下式成立：

$$\lim_{n \rightarrow \infty} P\{|\xi_n(\omega) - \xi(\omega)| \geq \varepsilon\} = 0$$

则称 $\{\xi_n(\omega)\}$ 依概率收敛于 $\xi(\omega)$ ，并记为

$$\xi_n(\omega) \xrightarrow{P} \xi(\omega)$$

定义3 (r阶矩收敛)

设对随机变量 ξ_n 及 ξ 有 $E|\xi_n|^r < \infty, E|\xi|^r < \infty$ ，

其中 $r > 0$ 为常数，如果

$$\lim_{n \rightarrow \infty} E|\xi_n - \xi|^r = 0$$

则称 $\{\xi_n\}$ **r阶(矩)收敛**于 ξ ，并记为

$$\xi_n \xrightarrow{r} \xi$$

在 r 阶收敛中，最重要的是 $r=2$ 的情况，称为**均方收敛**。

定义4 (几乎处处收敛)

如果

$$P\{\lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)\} = 1$$

则称 $\{\xi_n(\omega)\}$ **以概率1收敛**于 $\xi(\omega)$ ，又称 $\{\xi_n(\omega)\}$ **几乎处处收敛**于 $\xi(\omega)$ ，记为

$$\xi_n(\omega) \xrightarrow{a.s.} \xi(\omega).$$

定理1 $\xi_n(\omega) \xrightarrow{P} \xi(\omega) \Rightarrow \xi_n(\omega) \xrightarrow{L} \xi(\omega).$

证明：因为，对 $x' < x$ ，有

$$\begin{aligned} \{\xi < x'\} &= \{\xi_n < x, \xi < x'\} + \{\xi_n \geq x, \xi < x'\} \\ &\subset \{\xi_n < x\} + \{\xi_n \geq x, \xi < x'\} \end{aligned}$$

所以，我们有

$$F(x') \leq F_n(x) + P\{\xi_n \geq x, \xi < x'\}$$

因为 $\{\xi_n\}$ 依概率收敛于 ξ ，则

$$P\{\xi_n \geq x, \xi < x'\} \leq P\{|\xi_n - \xi| \geq x - x'\} \rightarrow 0$$

因而有 $F(x') \leq \lim_{n \rightarrow \infty} F_n(x)$

同理, 对 $x'' > x$,

$$\{\xi \geq x''\} = \{\xi_n < x, \xi \geq x''\} + \{\xi_n \geq x, \xi \geq x''\} \\ \subset \{\xi_n < x, \xi \geq x''\} + \{\xi_n \geq x\}$$

类似可得: $\lim_{n \rightarrow \infty} F_n(x) \leq F(x'')$

所以对 $x' < x < x''$, 有

$$F(x') \leq \lim_{x \rightarrow \infty} F_n(x) \leq \lim_{n \rightarrow \infty} F_n(x) \leq F(x'')$$

如果 x 是 $F(x)$ 的连续点, 则令 x', x'' 趋于 x 可得

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

定理证毕。

定理1逆命题不成立。

例3 若样本空间 $\Omega = \{\omega_1, \omega_2\}$, $P(\omega_1) = P(\omega_2) = \frac{1}{2}$, 定义随机变量 $\xi(\omega)$ 如下: $\xi(\omega_1) = -1, \xi(\omega_2) = 1$, 则 $\xi(\omega)$ 的分布列为

$$\begin{pmatrix} -1 & 1 \\ 1/2 & 1/2 \end{pmatrix} \quad (2)$$

若对一切 n , 令 $\xi_n(\omega) = -\xi(\omega)$, 显然 $\xi_n(\omega)$ 的分布列也是 (2), 因此 $\xi_n(\omega) \xrightarrow{L} \xi(\omega)$ 。

但是, 对任意的 $0 < \varepsilon < 2$, 因

$$P\{|\xi_n(\omega) - \xi(\omega)| > \varepsilon\} = P(\Omega) = 1$$

因此, $\{\xi_n(\omega)\}$ 不依概率收敛于 $\xi(\omega)$ 。

但是在特殊场合却有下面的结果。

定理2 设 C 是常数, 则 $\xi_n(\omega) \xrightarrow{P} C \Leftrightarrow \xi_n(\omega) \xrightarrow{L} C$ 。

证明: 由前面的定理可知只须证明由依分布收敛于常数可推出依概率收敛于常数。事实上, 对任意 $\varepsilon > 0$,

$$\begin{aligned} P\{|\xi_n - C| \geq \varepsilon\} &= P\{\xi_n \geq C + \varepsilon\} + P\{\xi_n \leq C - \varepsilon\} \\ &= 1 - F_n(C + \varepsilon) + F_n(C - \varepsilon) \\ &\xrightarrow{n \rightarrow \infty} 1 - F(C + \varepsilon) + F(C - \varepsilon) \\ &= 1 - 1 + 0 = 0. \end{aligned}$$

r 阶收敛与依概率收敛的关系。

定理3 $\xi_n \xrightarrow{r} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$ 。

证明: 先证对于任意 $\varepsilon > 0$, 成立

$$P\{|\xi_n - \xi| \geq \varepsilon\} \leq \frac{E|\xi_n - \xi|^r}{\varepsilon^r}$$

这个不等式是Chebyshev不等式的推广, 称作**Markov不等式**。

$$P\{|\xi - C| \geq \varepsilon\} \leq \frac{E|\xi - C|^r}{\varepsilon^r}$$

事实上, 若以 $F(x)$ 记 $\xi_n - \xi$ 的分布函数, 则有

$$\begin{aligned} P\{|\xi_n - \xi| \geq \varepsilon\} &= \int_{|x| \geq \varepsilon} dF(x) \leq \int_{|x| \geq \varepsilon} \frac{|x|^r}{\varepsilon^r} dF(x) \\ &\leq \frac{1}{\varepsilon^r} \int_{|x| \geq \varepsilon} |x|^r dF(x) \leq \frac{E|\xi_n - \xi|^r}{\varepsilon^r} \\ &\xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

定理3逆命题不成立。 由依概率收敛或几乎处处收敛不能推出以 r 阶收敛。

定理4 $\xi_n(\omega) \xrightarrow{a.s.} \xi(\omega) \Rightarrow \xi_n(\omega) \xrightarrow{P} \xi(\omega)$ 。

• 证明: 见 (p329~334, 第4节“强大数定律”)

由依概率收敛或矩收敛不能推出几乎处处收敛。

小结: r.v.的收敛性

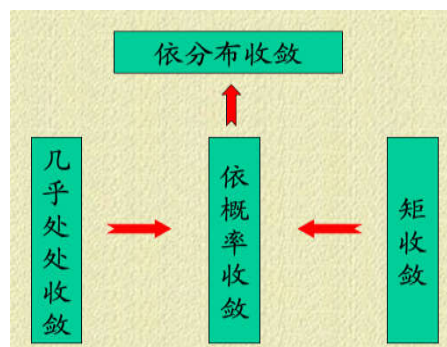
1. 四种收敛性的定义

依分布收敛: $\xi_n \xrightarrow{d} \xi$ $F_n(x) \xrightarrow{w} F(x)$

依概率收敛: $\xi_n \xrightarrow{P} \xi$ $\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| \geq \varepsilon\} = 0$

r 阶矩收敛: $\xi_n \xrightarrow{r} \xi$ $r > 0$
 $\lim_{n \rightarrow \infty} E|\xi_n - \xi|^r = 0$

几乎处处收敛: $\xi_n \xrightarrow{a.s.} \xi$ $P\{\lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)\} = 1$



§ 5.2 大数定律

(The law of large numbers)

一、大数定律的客观背景

大量随机试验中 $\begin{cases} \text{事件发生的频率稳定于某一常数} \\ \text{测量值的算术平均值具有稳定性} \end{cases}$



大量抛掷硬币
正面出现频率



生产过程
中的废品率



文章中字
母使用频率

二、Bernoulli大数定律: 概率的频率定义的理论基础

定理(伯努利大数定律) 设 μ_n 是 n 重Bernoulli试验中事件 A 发生的次数, p 是事件 A 发生的概率, 则对任给的 $\varepsilon > 0$,

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

或

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{\mu_n}{n} - p\right| \geq \varepsilon\right\} = 0$$



雅各布·伯努利
Bernoulli

证明. 根据 Chebyshev Ineq

$$P\left\{\left|\frac{\mu_n}{n} - p\right| \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2} D\left(\frac{\mu_n}{n}\right) \leq \frac{1}{4n\varepsilon^2}$$

□

三、Chebyshev大数定律:

定理(切比雪夫大数定律) 设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是两两不相关的随机变量序列, 它们都有有限的方差, 并且方差有共同的上界, 即 $D(\xi_i) \leq C, i = 1, 2, \dots$, 则对任意的 $\varepsilon > 0$, 皆有

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



切比雪夫

证明: 因为 $E\left(\frac{1}{n} \sum_{i=1}^n \xi_i\right) = \frac{1}{n} \sum_{i=1}^n E(\xi_i)$,

$$D\left(\frac{1}{n} \sum_{i=1}^n \xi_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(\xi_i) \leq \frac{C}{n}$$

由Chebyshev不等式, 对于任意的正实数 ε 有

$$0 \leq P\left\{\left|\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n E(\xi_i)\right| \geq \varepsilon\right\} \leq \frac{D\left(\frac{1}{n} \sum_{i=1}^n \xi_i\right)}{\varepsilon^2} \leq \frac{C}{n\varepsilon^2}$$

$$\text{所以 } \lim_{n \rightarrow +\infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n E(\xi_i)\right| \geq \varepsilon\right\} = 0$$

作为切比雪夫大数定律的特殊情况，有下面的推论。

推论1 (独立同分布下的大数定律)

设 ξ_1, ξ_2, \dots 是独立同分布的随机变量序列，且 $E\xi_i = \mu, D\xi_i = \sigma^2, i=1,2,\dots$ ，则对任给 $\varepsilon > 0$,

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

推论2 (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

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

证明：定义 ξ_k 为第 k 次试验中事件 A 出现的次数，则

$$E\xi_k = p_k, D\xi_k = p_k(1-p_k) \leq \frac{1}{4}, \mu_n = \sum_{k=1}^n \xi_k,$$

再利用切比雪夫大数定律立刻推出结论。

明显，当 $p_k \equiv p$ ，泊松大数定律即为伯努利大数定律。

马尔可夫注意到在切比雪夫的论证中，只要

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

则大数定律就能成立，通常称这个条件为马尔可夫条件，这样我们也就得到了下面的马尔可夫大数定律。

定理 (马尔可夫大数定律)

对于随机变量列 $\xi_1, \xi_2, \dots, \xi_n, \dots$ ，如果成立

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

则对任意的 $\varepsilon > 0$ ，皆有

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

切比雪夫大数定律显然可由马尔可夫大数定律推出；更重要的是马尔可夫大数定律已经没有任何关于独立性的假定。

例 设 $\{X_n\}$ 是同分布于 $\text{Exp}(1/\lambda)$ 的随机变量序列且序列中每一项仅与相邻两项相关，而与其它项不相关。判定该随机变量序列 $\{X_n\}$ 是否服从大数定律？

解：由于 $X_n \sim \text{Exp}(1/\lambda) \quad n=1,2,\dots$ ，从而有

$$D(X_n) = \lambda^2 \quad n=1,2,\dots$$

于是有

$$\begin{aligned} \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) &= \frac{1}{n^2} \left[\sum_{i=1}^n D(X_i) + 2 \sum_{i=1}^{n-1} \text{Cov}(X_i, X_{i+1}) \right] \\ &\leq \frac{1}{n^2} \left[\sum_{i=1}^n \lambda^2 + 2 \sum_{i=1}^{n-1} \lambda^2 \right] = \frac{1}{n^2} [n\lambda^2 + 2(n-1)\lambda^2] \\ \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) &= \frac{1}{n^2} [n\lambda^2 + 2(n-1)\lambda^2] \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

四、Khinchin大数定律

我们前面已经通过切比雪夫不等式建立起多种大数定律，在那里都假定了方差的存在性，但是在独立同分布场合，并不需要这个要求，这就是有名的辛钦大数定律告诉我们的。

定理 (辛钦大数定律) 设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 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

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

$$\text{即 } \frac{1}{n} \sum_{i=1}^n \xi_i \xrightarrow{P} a.$$

证明：由于 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 具有相同分布，故有相同的特征函数，设为 $f(t)$ ，因为数学期望存在，故 $f(t)$ 可展开成：

$$f(t) = f(0) + f'(0)t + o(t) = 1 + iat + o(t)$$

而 $\frac{1}{n} \sum_{i=1}^n \xi_i$ 的特征函数为

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

对于固定的 t ,

$$\left[f\left(\frac{t}{n}\right) \right]^n \rightarrow e^{iat} \quad (n \rightarrow \infty)$$

极限函数 e^{iat} 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，有

$$\frac{1}{n} \sum_{i=1}^n \xi_i \xrightarrow{L} a.$$

最后由依分布收敛和依概率收敛的关系定理知：

$$\frac{1}{n} \sum_{i=1}^n \xi_i \text{ 依概率收敛于常数 } a, \text{ 从而证明了定理。}$$

利用对随机变量“截尾”的技巧，Kolmogorov 在 Khintchine 大数定律的条件下，把结论加强为强收敛；即独立同分布时只要期望存在，序列 **部分和的算术平均几乎处处收敛**。

定理 (Kolmogorov 大数定律) 设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是独立同分布 (i.i.d) 的随机变量序列，则

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

当且仅当数学期望 $E\xi_i$ 存在，且 $E\xi_i = a$ 。

小结：大数定律的意义

➤ **Khintchine 大数定理**

这一定理表明：同一量 X 在相同条件下观测 n 次，当观测次数 n 充分大时，“观测值的算术平均值接近期望值”是一个大概率事件，即下式以大概率成立：

$$\frac{1}{n} \sum_{i=1}^n X_i \stackrel{n \text{ 充分大}}{\approx} E(X)$$

为寻找随机变量的期望值提供了一条实际可行的途径

➤ **Bernoulli 大数定理**

这一定理表明：在相同条件下重复同一随机试验 n 次，当试验次数 n 充分大时，“事件 A 发生的频率接近其概率”是一个大概率事件，即下式以大概率成立：

$$f_A \stackrel{n \text{ 充分大}}{\approx} P(A)$$

寻找随机事件概率提供了一条实际可行的途径

五、大数定律的应用

例1 (用蒙特卡洛方法计算定积分) 为计算积分

$$J = \int_a^b g(x) dx$$

可以通过下面的概率方法实现。

任取一系列相互独立的，都具有 $[a, b]$ 中均匀分布的随机变量 $\{\xi_i\}$ ，则 $\{g(\xi_i)\}$ 也是一列相互独立同分布的随机变量，而且

$$Eg(\xi_i) = \frac{1}{b-a} \int_a^b g(x) dx = \frac{J}{b-a}$$

既然

$$J = (b-a) \cdot Eg(\xi_i)$$

因此只要能求得 $Eg(\xi_i)$ ，便能得到 J 的数值。

为求 $Eg(\xi_i)$ ，自然想到大数定律，因为

$$\frac{g(\xi_1) + g(\xi_2) + \dots + g(\xi_n)}{n} \xrightarrow{P} Eg(\xi_i)$$

这样一来，只要能生成随机变量序列 $\{g(\xi_i)\}$ 就能对前面的积分进行数值计算。

而生成 $\{g(\xi_i)\}$ 的关键是生成相互独立同分布的 $\{\xi_i\}$ ，这里的 ξ_i 服从 $[a, b]$ 上的均匀分布。

现在已经可以把上述想法变成现实。这就是在电子计算机上产生服从均匀分布 $[a, b]$ 的随机数 $\{\xi_i\}$ 。

强大数律保证了这种算法失效的概率为0。

这种通过概率论的想法构造模型从而实现数值计算的方法，随着电子计算机的发展，已形成一种新的计算方法——概率计算方法，亦称蒙特卡洛(Monte Carlo)方法。它在原子物理、公用事业理论中发挥了不少作用，这个方法的理论根据之一就是大数定律。

至于计算积分，蒙特卡洛方法的实用场合是计算重积分

$$I = \int_K g(P) dP$$

其中 P 是 m 维空间中的点。

§ 5.3 中心极限定理 (The central limit theory)

一、中心极限定律的客观背景

在实际问题中许多随机变量是由相互独立随机因素的综合（或和）影响所形成。例如：炮弹射击的落点与目标的偏差，就受着许多随机因素（如瞄准，空气阻力，炮弹或炮身结构等）综合影响的。每个随机因素的对弹着点（随机变量和）所起的作用都是很小的。那么弹着点服从怎样分布呢？



定理 (Lindeberg-Levy中心极限定理)

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立，服从同一分布，且具有数学期望和方差： $E(X_i) = \mu$, $D(X_i) = \sigma^2 < \infty$ ($i = 1, 2, \dots$)，则对于任意的实数 x ，有

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

即 $\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{L} \xi$, 其中 $\xi \sim N(0, 1)$.

(一句话, 当 $n \rightarrow \infty$ 时, $\sum_{i=1}^n X_i$ 的标准化近似服从标准正态分布)

定理表明: $\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{a} N(0, 1)$.

由正态分布的性质, 有 $\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$.

这就是说: 当 n 充分大时, 只要 X_1, X_2, \dots, X_n 独立同分布, 无论他们服从什么分布, 一定有

$$\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$$

即: 一个由许多独立同分布随机变量作用形成的随机变量, 其概率分布一定是近似正态分布。

证明: 设 $X_n - \mu$ 的特征函数为 $f(t)$, 则由于

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

则其特征函数为 $f_n(t) = \left(f\left(\frac{t}{\sqrt{n}\sigma}\right) \right)^n$

又由于 $\frac{f'(0)}{i} = E(X_n - \mu) = 0$, 知 $f'(0) = 0$

$$-f''(0) + (f'(0))^2 = D(X_n - \mu) = \sigma^2 \text{ 知 } f''(0) = -\sigma^2$$

而 $f(t)$ Taylor展开为

$$f(t) = f(0) + f'(0)t + \frac{f''(0)}{2!}t^2 + o(t^2) = 1 - \frac{\sigma^2}{2}t^2 + o(t^2)$$

所以,

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

由于 $e^{-t^2/2}$ 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

$$\lim_{n \rightarrow \infty} P\{\eta_n < x\} = \Phi(x)$$

证毕.

定理 (De Moivre-Laplace中心极限定理)

设随机变量 $X \sim B(n, p)$, 则对于任意的实数 x , 有

$$\lim_{n \rightarrow +\infty} P\left\{\frac{X - np}{\sqrt{npq}} < x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

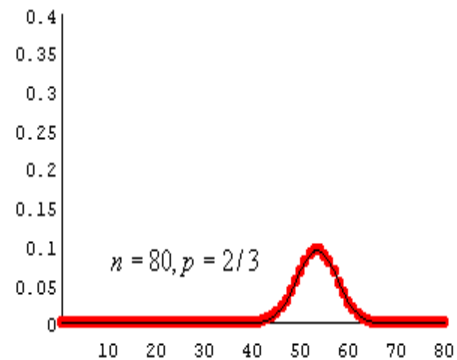
证明: 因为 $X \sim B(n, p)$, 则有 $X = \sum_{i=1}^n X_i$, 其中 $X_i \sim B(1, p)$, 且相互独立, 即 X 为独立和. 易知

$$E(X) = np, D(X) = npq$$

由Lindeberg-Levy中心极限定理知

$$\lim_{n \rightarrow +\infty} P\left\{\frac{X - np}{\sqrt{npq}} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

理解: 在定理条件下, 总有 $X \overset{a}{\sim} N(np, npq)$.



小结: 中心极限定理的意义

➤ Lindeberg-Levy中心极限定理

对于独立同分布随机变量序列 $\{X_n\}$, 不管他们服从什么分布, 只要存在有限数学期望和方差, 当 n 充分大时, 就有

$$\sum_{i=1}^n X_i \overset{\text{近似}}{\sim} N(n\mu, n\sigma^2)$$

所以, $\sum_{i=1}^n X_i$ 的有关概率问题可利用正态分布求解。

➤ De Moivre-Laplace中心极限定理

对于随机变量 $X \sim B(n, p)$, 当 n 充分大时, 有 $X \overset{\text{近似}}{\sim} N(np, npq)$, 因此, 二项分布的概率问题可利用正态分布解决。

Lindeberg-Levy中心极限定理有着广泛应用。在实际工作中, 只要 n 足够大, 便可以把独立同分布的随机变量之和当作是正态变量。此做法在数理统计中用得普遍。

例1 (正态随机数的产生)

设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是相互独立、均服从 $[0, 1]$ 均匀分布的随机变量, 这时Lindeberg-Levy中心极限定理的条件得到满足, 故 $\xi_1 + \xi_2 + \dots + \xi_n$ 渐近于正态变量。

一般 n 取不太大的值就可满足实际要求, 在蒙特卡洛方法中, 一般取 $n=12$ 。

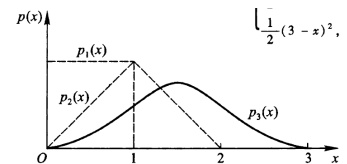


图 5.3.1 均匀分布卷积

◆ 在二项分布计算中的应用

由中心极限定理, 当 p 不太接近于0或1, 而 n 又不太小时, 对二项分布的近似计算有下面的公式:

$$P\{k_1 \leq \mu_n \leq k_2\} = P\left\{\frac{k_1 - np}{\sqrt{npq}} \leq \frac{\mu_n - np}{\sqrt{npq}} \leq \frac{k_2 - np}{\sqrt{npq}}\right\} \\ \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

实际计算中, 往往用下面的修正公式计算效果更好。

$$P\{k_1 \leq \mu_n \leq k_2\} \approx \Phi\left(\frac{k_2 - np + 0.5}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np - 0.5}{\sqrt{npq}}\right)$$

例2 (用电问题)

某车间有200台车床, 它们独立地工作着, 开工率各为0.6, 开工时耗电各为1千瓦, 问供电所至少要提供给这个车间多少电力才能以99.9%的概率保证这个车间不会因供电不足而影响生产?

解: 这是实验次数 $n = 200$ 的伯努利实验, 若把某台车床在工作看作成功, 则出现成功的概率为0.6。

记某时工作着的车床数为 ξ , 则 ξ 是随机变量, 服从 $p = 0.6$ 的二项分布, 设供给车间 r 千瓦电, 要求

$$P\{\xi \leq r\} = \sum_{k=0}^r \binom{200}{k} (0.6)^k (0.4)^{200-k} \geq 0.999$$

我们可以利用中心极限定理来计算这个概率。

$$P\{0 \leq \xi \leq r\} = \sum_{k=0}^r \binom{200}{k} (0.6)^k (0.4)^{200-k}$$

$$\approx \Phi\left(\frac{r - 200 \times 0.6 + 0.5}{\sqrt{200 \times 0.6 \times 0.4}}\right) - \Phi\left(\frac{-200 \times 0.6 - 0.5}{\sqrt{200 \times 0.6 \times 0.4}}\right)$$

$$\approx \Phi\left(\frac{r - 119.5}{\sqrt{48}}\right) - \Phi(-17.39) \approx \Phi\left(\frac{r - 119.5}{\sqrt{48}}\right) \geq 0.999$$

查表得: $\frac{r - 119.5}{\sqrt{48}} \geq 3.1$ 所以 $r \geq 141$

这个结果表明 $P\{\xi \leq 141\} \geq 0.999$, 所以我们若供电141千瓦, 那么由于供电不足而影响生产的可能性小于0.001, 相当于8小时工作中有半分钟受影响, 这在一般工厂中是允许的。

例3 将一颗骰子独立掷100次, 则点数之和不少于300的概率是多少?

解: 设 X_k 为第 k 次掷出的点数, $k=1, 2, \dots, 100$, 则

X_1, \dots, X_{100} 独立同分布.

$$E(X_i) = \frac{7}{2}, D(X_i) = \frac{1}{6} \sum_{k=1}^6 k^2 - \frac{49}{4} = \frac{35}{12}$$

由Lindeberg-Levy中心极限定理

$$P\left\{\sum_{i=1}^{100} X_i \geq 300\right\} \approx 1 - \Phi\left(\frac{300 - 100 \times \frac{7}{2} - 0.5}{10 \sqrt{\frac{35}{12}}}\right) = 1 - \Phi(-2.96) = 0.9985$$

例4 (保险问题) 在一家保险公司里有10000个人参加寿命保险, 每人每年付12元保险费。在一年内一个人死亡的概率为0.006, 死亡时其家属可向保险公司领得1000元, 问:

(1) 保险公司亏本的概率有多大?

(2) 其他条件不变, 为使保险公司一年的利润不少于60000元的概率不低于90%, 赔金至多可设为多少?

解: 设 X 表示一年内死亡的人数, 则 $X \sim B(n, p)$, 其中

$$n = 10000, p = 0.006,$$

由中心极限定理

$$\begin{aligned} (1) & P\{10000 \times 12 - 1000X < 0\} \\ &= 1 - P\{X \leq 120\} \\ &\approx 1 - \Phi(7.75) \\ &= 0; \end{aligned}$$

(2) 设赔偿金为 a 元, 则

$$P\{10000 \times 12 - aX \geq 60000\} \geq 0.9$$

$$\text{即 } P\{X \leq 60000/a\} \geq 0.9$$

由中心极限定理, 上式等价于

$$\Phi\left(\frac{\frac{60000}{a} - 10000 \times 0.006 + 0.5}{\sqrt{10000 \times 0.006 \times 0.994}}\right) - \Phi\left(\frac{0 - 10000 \times 0.006 - 0.5}{\sqrt{10000 \times 0.006 \times 0.994}}\right) \geq 0.9$$

$$\Rightarrow a \leq 865$$