第五章

极限定理

本章要点

1 收敛性

2 大数定律

3 中心极限定理

5.1 收敛性

- 分布函数弱收敛
- 连续性定理
- 随机变量的收敛性

一、分布函数弱收敛

例5.1.1

令

$$F_n(x) = \begin{cases} 0, & x \le -\frac{1}{n}; \\ 1, & x > -\frac{1}{n}. \end{cases}$$

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$x = -\frac{1}{n}$的分布。

一、分布函数弱收敛

例5.1.1

令

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

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$x = -\frac{1}{n}$的分布。

当$n \to \infty$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，$F_n(0) = 1$，而$F(0) = 0$，显然，$\lim_{n\to\infty} F_n(0) \neq F(0)$.

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。

但是，$F_n(0) = 1$，而$F(0) = 0$，显然，$\lim_{n\to\infty} F_n(0) \neq F(0)$.

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5.1.1

对于分布函数列$\{F_n(x)\}$,如果存在一个非降函数$F(x)$使

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

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$F_n(x) \xrightarrow{W} F(x)$.

定义5.1.1

对于分布函数列$\{F_n(x)\}$,如果存在一个非降函数$F(x)$使

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

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，我们也可以选得它是左连续的,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.

分布函数列的弱收敛极限不一定是分布函数.

例5.1.2

取

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

显然，$\lim\limits_{n\to\infty} F_n(x) = 0$ 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，其不连续点集最多只包含可列个点，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。

当然，当$F(x)$是一个分布函数时，其不连续点集最多只包含可列个点，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。

当一个分布函数列弱收敛到一个连续的分布函数时，这种收敛还是一致的。（习题20）

以下我们研究一个分布函数序列弱收敛到一个分布函数的充要条件，为此先建立一些重要的分析结果。

以下我们研究一个分布函数序列弱收敛到一个分布函数的充要条件，为此先建立一些重要的分析结果。

引理

设$\{F_n(x)\}$是实变量x的非降函数序列，D是R上的稠密集。若对于D中的所有点，序列$\{F_n(x)\}$收敛于$F(x)$，则对$F(x)$的一切连续点x有

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

证明： 设$x \in R$为任意点，选$x' \in D, x'' \in D$使$x' \leq x \leq x''$，

证明： 设$x \in R$为任意点，选$x' \in D, x'' \in D$使$x' \leq x \leq x''$，由非降性知，

$$F_n(x') \leq F_n(x) \leq F_n(x'').$$

证明： 设$x \in R$为任意点，选$x' \in D, x'' \in D$使$x' \le x \le x''$，由非降性知，

$$F_n(x') \le F_n(x) \le F_n(x'').$$

两边令$n \to \infty$，

$$F(x') \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x'').$$

证明： 设$x \in R$为任意点，选$x' \in D, x'' \in D$使$x' \le x \le x''$，由非降性知，

$$F_n(x') \le F_n(x) \le F_n(x'').$$

两边令$n \to \infty$,

$$F(x') \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x'').$$

因为D在R上稠密，故

$$F(x-0) \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x+0).$$

证明： 设$x \in R$为任意点，选$x' \in D, x'' \in D$使$x' \leq x \leq x''$，由非降性知，

$$F_n(x') \leq F_n(x) \leq F_n(x'').$$

两边令$n \to \infty$,

$$F(x') \leq \varliminf_{n\to\infty} F_n(x) \leq \varlimsup_{n\to\infty} F_n(x) \leq F(x'').$$

因为D在R上稠密，故

$$F(x-0) \leq \varliminf_{n\to\infty} F_n(x) \leq \varlimsup_{n\to\infty} F_n(x) \leq F(x+0).$$

所以对于$F(x)$的连续点x,引理成立。

下面是海莱(Helly)得到的两个重要定理。

下面是海莱(Helly)得到的两个重要定理。

定理5.1.1

（海莱第一定理）任一一致有界的非降函数列$\{F_n(x)\}$中必有一子序列$\{F_{n_k}(x)\}$弱收敛于某一有界的非降函数$F(x)$.

注：证明用对角线法。

下面是海莱(Helly)得到的两个重要定理。

定理5.1.1

（海莱第一定理）任一一致有界的非降函数列$\{F_n(x)\}$中必有一子序列$\{F_{n_k}(x)\}$弱收敛于某一有界的非降函数$F(x)$.

注：证明用对角线法。

定理5.1.2

（海莱第二定理）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

$$\lim_{n\to\infty}\int_a^b f(x)dF_n(x) = \int_a^b f(x)dF(x).$$

定理5.1.3

（推广的海莱第二定理）设$f(x)$在$(-\infty,+\infty)$ 上有界连续，又$\{F_n(x)\}$是$(-\infty,+\infty)$ 上弱收敛于函数$F(x)$的一致有界非降函数序列，且

$$\lim_{n\to\infty} F_n(-\infty) = F(-\infty), \quad \lim_{n\to\infty} F_n(+\infty) = F(+\infty),$$

则

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f(x)dF_n(x) = \int_{-\infty}^{+\infty} f(x)dF(x).$$

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。

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5.1.4

（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。

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5.1.4

（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。

证明在P308。

定理5.1.5

（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。

定理5.1.5

（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。

证明在P308-310。

定理5.1.5

（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。

证明在P308-310。

通常把正逆极限定理合称为连续性定理。

定理5.1.5

（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。

证明在P308-310。

通常把正逆极限定理合称为连续性定理。

定理5.1.6

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$\{f_n(t)\}$在任意有限区间内一致收敛到某个函数$f(t)$。

三、随机变量的收敛性

定义5.1.2

（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5.1.3

（依概率收敛）

如果对任意$\varepsilon > 0$，有下式成立：

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

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

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

定义5.1.4

（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

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

则称$\{\xi_n\}$ r阶（矩）收敛于ξ，并记为

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

在r阶矩收敛中，最重要的是$r = 2$的情况，称为均方收敛。

定义5.1.5

（几乎处处收敛）

如果

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

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5.1.7

设$\{X_n\}, \{Y_n\}$是两个随机变量序列，如果$X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$，则有

1. $X_n - X \xrightarrow{P} 0$;
2. $X_n \xrightarrow{P} X' \Longrightarrow P\{X = X'\} = 1$;
3. $X_n \xrightarrow{P} a, Y_n \xrightarrow{P} b$，$a, b$是常数，$b \neq 0$，$X_n/Y_n \xrightarrow{P} a/b$;
4. $X_n \pm Y_n \xrightarrow{P} X \pm Y$;
5. $X_n Y_n \xrightarrow{P} XY$;
6. $g(x)$是R上的连续函数，$g(X_n) \xrightarrow{P} g(X)$。

证明： 仅证明$X_n \pm Y_n \xrightarrow{P} X \pm Y$。

证明： 仅证明$X_n \pm Y_n \xrightarrow{P} X \pm Y$。
因为$|X_n \pm Y_n - (X \pm Y)| = |X_n - X \pm (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$，
从而对任意的ϵ,

$$\{|X_n - X \pm (Y_n - Y)| \geq \epsilon\} \subset \{|X_n - X| \geq \frac{\epsilon}{2}\} \cup \{|Y_n - Y| \geq \frac{\epsilon}{2}\},$$

$$\begin{aligned} & P\{|X_n \pm Y_n - (X \pm Y)| \geq \epsilon\} = P\{|X_n - X \pm (Y_n - Y)| \geq \epsilon\} \\ \leq \ & P(\{|X_n - X| \geq \frac{\epsilon}{2}\} \cup \{|Y_n - Y| \geq \frac{\epsilon}{2}\}) \\ \leq \ & P\{|X_n - X| \geq \frac{\epsilon}{2}\} + P\{|Y_n - Y| \geq \frac{\epsilon}{2}\}. \end{aligned}$$

由于$\lim\limits_{n\to\infty} P(|X_n - X| \geq \frac{\epsilon}{2}) = 0$, $\lim\limits_{n\to\infty} P(|Y_n - Y| \geq \frac{\epsilon}{2}) = 0$, 所以

$$\lim_{n\to\infty} P\{|X_n \pm Y_n - (X \pm Y)| \geq \epsilon\} = 0.$$

即$X_n \pm Y_n \xrightarrow{P} X \pm Y$。

定理5.1.8

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

定理5.1.8

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

证明： 因为，对$x' < x$，有

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

定理5.1.8

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

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

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

定理5.1.8

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

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

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

因而有$F(x') \leq \varliminf_{n\to\infty} F_n(x)$.

同理，对$x'' > x$，

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

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

所以对$x' < x < x''$，有

$$F(x') \leq \underline{\lim_{n\to\infty}} F_n(x) \leq \overline{\lim_{n\to\infty}} F_n(x) \leq F(x'').$$

如果x是$F(x)$的连续点，则令x', x''趋于x可得

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

定理证毕。

定理5.1.7逆命题不成立!

定理5.1.7逆命题不成立!

例5.1.3

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

$$\begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

若对一切n，令$\xi_n(\omega) = -\xi(\omega)$，显然$\xi_n(\omega)$的分布列和$\xi(\omega)$的一样，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。

但是在特殊场合却有下面的结果。

定理5.1.9

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

但是在特殊场合却有下面的结果。

定理5.1.9

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$\varepsilon > 0$，

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

r阶矩收敛与依概率收敛的关系

定理5.1.10

$$\xi_n \xrightarrow{r} \xi \Rightarrow \xi_n \xrightarrow{P} \xi.$$

r阶矩收敛与依概率收敛的关系

定理5.1.10

$$\xi_n \xrightarrow{r} \xi \Rightarrow \xi_n \xrightarrow{P} \xi.$$

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。

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

事实上,若以$F(x)$记$\xi_n-\xi$的分布函数，则有

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

事实上,若以$F(x)$记$\xi_n-\xi$的分布函数，则有

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

定理5.1.9逆命题不成立。下例说明：由依概率收敛或几乎处处收敛不能推出以r阶矩收敛。

例5.1.4

取$\Omega=(0,1]$，$\mathcal{F}$为$(0,1]$中博雷尔点集全体所构成的σ域，$\mathcal{P}$为勒贝格测度。定义$\xi(\omega)=0$及

$$\xi_n(\omega)=\begin{cases} n^{\frac{1}{r}}, & 0<\omega\le\frac{1}{n}; \\ 0, & \frac{1}{n}<\omega\le 1. \end{cases}$$

显然对一切$\omega\in\Omega, \xi_n(\omega)\to\xi(\omega)$，又对于任意的$\varepsilon>0$，

$$P\{|\xi_n(\omega)-\xi(\omega)|\ge\varepsilon\}\le P\{\omega: 0<\omega\le\frac{1}{n}\}=\frac{1}{n}.$$

因此$\xi_n(\omega)\xrightarrow{P}\xi(\omega)$，但是$E|\xi_n(\omega)-\xi(\omega)|^r=(n^{\frac{1}{r}})^r\cdot\frac{1}{n}=1$.
实际上这里ξ_n在$(0,1]$中每个点都收敛，说明几乎处处收敛也不能蕴含矩收敛。

定理5.1.11

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

事件序列

设$A_1, A_2, \ldots, A_n, \ldots$是一列事件，则

$\bigcup\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 中至少发生一个，

$\bigcap\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 同时发生。

事件序列

设$A_1, A_2, \ldots, A_n, \ldots$是一列事件，则

$\bigcup\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 中至少发生一个，

$\bigcap\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 同时发生。

记 $\overline{\lim\limits_{n\to\infty}} A_n := \bigcap\limits_{k=1}^{\infty}\bigcup\limits_{n=k}^{\infty} A_n$，称 $\overline{\lim\limits_{n\to\infty}} A_n$ 为事件序列$\{A_n\}$的上限事件；它表示A_n 发生无穷多次。

事件序列

设$A_1, A_2, \ldots, A_n, \ldots$是一列事件，则

$\bigcup\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 中至少发生一个，

$\bigcap\limits_{n=k}^{\infty} A_n$ 表示事件序列$A_k, A_{k+1}, \ldots$ 同时发生。

记 $\overline{\lim\limits_{n\to\infty}} A_n := \bigcap\limits_{k=1}^{\infty}\bigcup\limits_{n=k}^{\infty} A_n$，称 $\overline{\lim\limits_{n\to\infty}} A_n$ 为事件序列$\{A_n\}$的上限事件；它表示A_n 发生无穷多次。

因为$\omega \in \bigcap\limits_{k=1}^{\infty}\bigcup\limits_{n=k}^{\infty} A_n$ 当且仅当ω 属于无穷多个A_n。

记 $\underline{\lim}_{n\to\infty} A_n := \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$，称 $\underline{\lim}_{n\to\infty} A_n$ 为事件序列$\{A_n\}$的下限事件；它表示A_n 至多只有有限个不发生。

记 $\varliminf_{n\to\infty} A_n := \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$，称 $\varliminf_{n\to\infty} A_n$ 为事件序列$\{A_n\}$的下限事件；它表示A_n 至多只有有限个不发生。

因为$\omega \in \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$ 当且仅当存在一个N，使$\omega \in \bigcap_{n=N}^{\infty} A_n$。

记$\underline{\lim}_{n\to\infty} A_n := \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$，称$\underline{\lim}_{n\to\infty} A_n$ 为事件序列$\{A_n\}$的下限事件；它表示A_n 至多只有有限个不发生。

因为$\omega \in \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$ 当且仅当存在一个N，使$\omega \in \bigcap_{n=N}^{\infty} A_n$。

显然

$$\overline{\lim_{n\to\infty}} A_n \supset \underline{\lim}_{n\to\infty} A_n.$$

记 $\underline{\lim}_{n\to\infty} A_n := \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$，称 $\underline{\lim}_{n\to\infty} A_n$ 为事件序列$\{A_n\}$的下限事件；它表示A_n 至多只有有限个不发生。

因为$\omega \in \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n$ 当且仅当存在一个N，使$\omega \in \bigcap_{n=N}^{\infty} A_n$。

显然

$$\overline{\lim_{n\to\infty}} A_n \supset \underline{\lim}_{n\to\infty} A_n.$$

特别当

$$\overline{\lim_{n\to\infty}} A_n = \underline{\lim}_{n\to\infty} A_n$$

时，记 $\lim_{n\to\infty} A_n = \overline{\lim_{n\to\infty}} A_n = \underline{\lim}_{n\to\infty} A_n$，称它为事件$\{A_n\}$的极限事件。

利用德摩根定理

$$\overline{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n} = \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} \overline{A}_n,$$

$$\overline{\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} A_n} = \bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} \overline{A}_n,$$

利用德摩根定理

$$\overline{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}A_n}=\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}\overline{A}_n,$$

$$\overline{\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}A_n}=\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\overline{A}_n,$$

因此

$$\varliminf_{n\to\infty}\overline{A}_n=\overline{\varlimsup_{n\to\infty}A_n},\qquad \varlimsup_{n\to\infty}\overline{A}_n=\overline{\varliminf_{n\to\infty}A_n}.$$

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

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

对任一正整数m，存在一个正整数N，使当$n > N$时均有

$$|\xi_n(\omega) - \xi(\omega)| < \frac{1}{m},$$

即对任一正整数m，ω属于

$$\left\{|\xi_n(\omega) - \xi(\omega)| < \frac{1}{m}\right\}$$

的下限事件。

概率1收敛的等价表示

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

2.
$$P\left\{\bigcup_{m=1}^{\infty}\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(|\xi_n(\omega)-\xi(\omega)|\geq\frac{1}{m}\right)\right\}=0.$$

3. 对任意的$\varepsilon>0$，成立

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

[2]和[3]的等价

[2]$\Rightarrow$ [3]。

[2]和[3]的等价

[2]$\Rightarrow$ [3]。

对$\varepsilon > 0$,

$$\left\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}(|\xi_n(\omega)-\xi(\omega)|\geq\varepsilon)\right\}\subset\left\{\bigcup_{m=1}^{\infty}\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(|\xi_n(\omega)-\xi(\omega)|\geq\frac{1}{m}\right)\right\}$$

所以

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

[3]$\Rightarrow$ [2]。

[3]⇒ [2]。

$$
\begin{aligned}
&P\left\{\bigcup_{m=1}^{\infty}\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(|\xi_n(\omega)-\xi(\omega)|\geq\frac{1}{m}\right)\right\}\\
\leq\ &\sum_{m=1}^{\infty}P\left\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(|\xi_n(\omega)-\xi(\omega)|\geq\frac{1}{m}\right)\right\}=0.
\end{aligned}
$$

概率1收敛的等价表示

1. 对任意的$\varepsilon > 0$，成立

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

2. 对任意的$\varepsilon > 0$，成立

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

3. 对任意的$\varepsilon > 0$，成立

$$\lim_{k\to\infty}P\left\{\bigcap_{n=k}^{\infty}(|\xi_n(\omega)-\xi(\omega)|<\varepsilon)\right\}=1.$$

定理5.1.12

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

定理5.1.12

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

证明： 概率1收敛等价于对任意的$\varepsilon > 0$，成立

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

由于

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

所以

$$\lim_{k\to\infty} P\{|\xi_k(\omega) - \xi(\omega)| \geq \varepsilon\} = 0.$$

证明： 见（$\mathrm{P}329 - 334$，第4节“强大数定律”）

反例：依概率收敛不能导致几乎处处收敛

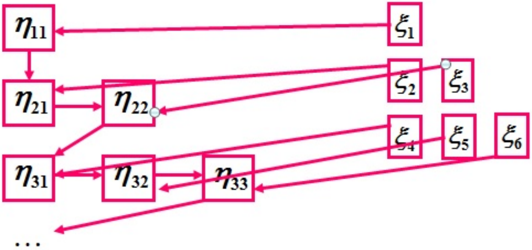
下例说明：
由依概率收敛或矩收敛不能推出几乎处处收敛。
$k \in N$，把$(0,1]k$等分，定义k个随机变量：

$$\eta_{ki}(\omega)=\begin{cases}1, & \omega\in\left(\frac{i-1}{k},\frac{i}{k}\right]; \\ 0, & \text{其他},\end{cases} \quad i=1,2,\ldots,k.$$

取P为勒贝格测度，则

$$\forall \varepsilon>0, P(|\eta_{ki}(\omega)|\ge\varepsilon)\le\frac{1}{k},\quad i=1,2,\ldots,k.\quad(*)$$

将η_{ki}依次记为ξ_n，如下图:



即

$$\xi_n(\omega) = \eta_{ki}(\omega), \quad n = i + \frac{k(k-1)}{2}.$$

由(∗)式，$\forall \varepsilon > 0$,

$$\lim_{n\to\infty} P(|\xi_n(\omega)| \geq \varepsilon) = \lim_{k\to\infty} P(|\eta_{ki}(\omega)| \geq \varepsilon) = 0.$$

从而，

$$\xi_n(\omega) \xrightarrow{P} 0.$$

同时，$\forall \omega \in (0,1]$，总有无数个$\eta_{ki}(\omega) = 1$，以及无数个$\eta_{ki}(\omega) = 0$，所以ξ_n处处不收敛。

上述ξ_n，满足

$$E|\xi_n|^r = E|\eta_{ki}|^r = \frac{1}{k} \to 0, \quad n \to \infty,$$

即$\xi_n(\omega) \xrightarrow{r} 0$. 可见，矩收敛也不能蕴涵几乎处处收敛。

小结：随机变量的收敛性

1. 四种收敛性的定义

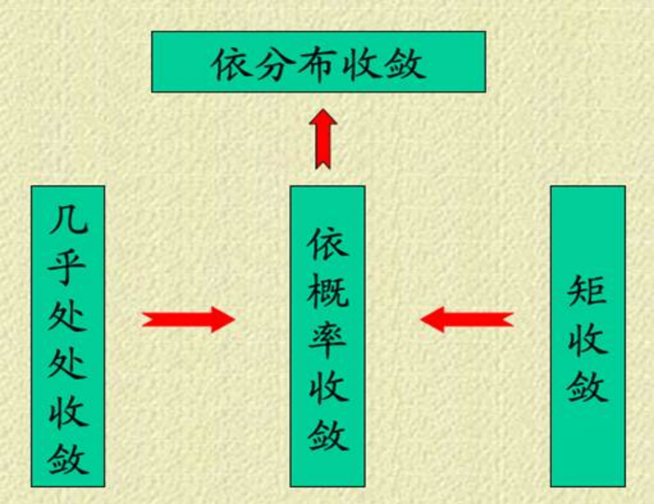
依分布收敛：$\xi_n \xrightarrow{d} \xi$ $\quad F_n(x) \xrightarrow{w} F(x)$

依概率收敛：$\xi_n \xrightarrow{P} \xi$ $\quad \lim\limits_{n\to\infty} P\{|\xi_n - \xi| \ge \varepsilon\} = 0$

r 阶矩收敛：$\xi_n \xrightarrow{r} \xi$ $\quad r > 0$, $\lim\limits_{n\to\infty} E|\xi_n - \xi|^r = 0$

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

小结：随机变量的收敛性



5.2 大数定律

一、大数定律的客观背景

5.2 大数定律

一、大数定律的客观背景

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

5.2 大数定律

一、大数定律的客观背景

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

大量抛掷硬币
正面出现频率

5.2 大数定律

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

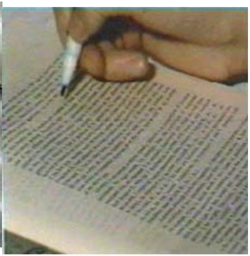
5.2 大数定律

一、大数定律的客观背景

大量随机试验中
- 事件发生的频率稳定于某一常数
- 测量值的算术平均值具有稳定性

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

大数定律(law of large numbers, ab. LLN)

大数定律(law of large numbers, ab. LLN)

定义5.2.1

若$\xi_1, \xi_2, \ldots, \xi_n, \ldots$ 是随机变量序列，令

$$\eta_n = \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n}.$$

如果存在这样的一个常数序列$a_1, a_2, \ldots, a_n, \ldots$, 对任意的$\varepsilon > 0$，恒有

$$\lim_{n\to\infty} P(|\eta_n - a_n| < \varepsilon) = 1.$$

则称序列$\{\xi_n\}$服从大数定律（或大数法则）。

强大数定律(strong law of large numbers, ab. SLLN)

强大数定律(strong law of large numbers, ab. SLLN)

定义5.2.2

若$\xi_1, \xi_2, \ldots, \xi_n, \ldots$ 是独立随机变量序列，若

$$P\left\{\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(\xi_i - E\xi_i) = 0\right\} = 1.$$

则称序列$\{\xi_n\}$满足强大数定律。

二、Bernoulli大数定律：概率的频率定义的理论基础

Bernoulli

定理5.2.1

（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$\varepsilon > 0$，

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

或

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

定理5.2.1

（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$\varepsilon > 0$,

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

或

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

证明：根据Chebyshev不等式

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

博雷尔(Borel)强大数定律

定理5.2.2

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，那么

$$P\left\{\lim_{n\to\infty}\frac{\mu_n}{n}=p\right\}=1.$$

博雷尔-康特立（Borel-Cantelli）引理

引理2.1

(I) 若随机事件序列$\{A_n\}$满足$\sum_{n=1}^{\infty} P(A_n) < \infty$,则

$$P\{\overline{\lim_{n\to\infty}} A_n\} = 0, \qquad P\{\underline{\lim_{n\to\infty}} \overline{A}_n\} = 1.$$

(II) 若$\{A_n\}$是相互独立的随机事件序列，则

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

成立的充分必要条件为

$$P\{\overline{\lim_{n\to\infty}} A_n\} = 1 \quad 或 \quad P\{\underline{\lim_{n\to\infty}} \overline{A}_n\} = 0.$$

引理证明

(i) 由于

$$P\{\overline{\lim_{n\to\infty}} A_n\} = P\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n\} \le P\{\bigcup_{n=k}^{\infty} A_n\} \le \sum_{n=k}^{\infty} P(A_n) \to 0(k \to \infty).$$

引理证明

(i) 由于

$$P\{\overline{\lim_{n\to\infty}} A_n\} = P\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n\} \le P\{\bigcup_{n=k}^{\infty} A_n\} \le \sum_{n=k}^{\infty} P(A_n) \to 0(k\to\infty).$$

(ii) 先证必要性。注意到$\{A_n\}$的独立性，有

$$P\{\underline{\lim_{n\to\infty}}\overline{A}_n\} = P\{\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty}\overline{A}_n\} \le \sum_{k=1}^{\infty} P\{\bigcap_{n=k}^{\infty}\overline{A}_n\}$$

$$= \sum_{k=1}^{\infty}\prod_{n=k}^{\infty} P(\overline{A}_n) = \sum_{k=1}^{\infty}\prod_{n=k}^{\infty}[1-P(A_n)].$$

引理证明

由于

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

则从

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

可得

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

所以,

$$P\{\varliminf_{n\to\infty} \overline{A}_n\} = 0.$$

引理证明

再证充分性。若$P\{\overline{\lim_{n\to\infty}} A_n\} = 1$。假定$\sum_{n=1}^{\infty} P(A_n) < \infty$，则由(i)得到$P\{\overline{\lim_{n\to\infty}} A_n\} = 0$，矛盾。

博雷尔(Borel)强大数定律

定理5.2.3

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，那么

$$P\left\{\lim_{n\to\infty}\frac{\mu_n}{n}=p\right\}=1.$$

博雷尔(Borel)强大数定律

证明： 只须对任意的$\varepsilon > 0$，成立

$$P\left\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right)\right\}=0.$$

博雷尔(Borel)强大数定律

证明： 只须对任意的$\varepsilon > 0$，成立

$$P\left\{\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty}\left(\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right)\right\}=0.$$

根据博雷尔-康特立（Borel-Cantelli）引理，只要能证明级数

$$\sum_{n=1}^{\infty}P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}$$

对任何$\varepsilon > 0$都收敛就可以了。

博雷尔(Borel)强大数定律

证明： 切比雪夫不等式$P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{1}{4n\varepsilon^2}$不够了。要用马尔科夫不等式。

$$P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{1}{\varepsilon^4}E\left|\frac{\mu_n}{n}-p\right|^4.$$

博雷尔(Borel)强大数定律

证明： 切比雪夫不等式$P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{1}{4n\varepsilon^2}$不够了。要用马尔科夫不等式。

$$P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{1}{\varepsilon^4}E\left|\frac{\mu_n}{n}-p\right|^4.$$

我们把μ_n表示成独立伯努利0-1变量$\xi_1,\xi_2,\ldots,\xi_n$之和，这样

$$\frac{\mu_n}{n}-p=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-p).$$

$$
\begin{aligned}
& E\left(\frac{\mu_n}{n}-p\right)^4 \\
= \quad & \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[(\xi_i-p)(\xi_j-p)(\xi_k-p)(\xi_l-p)\right].
\end{aligned}
$$

$$
\begin{aligned}
& E\left(\frac{\mu_n}{n} - p\right)^4 \\
= \quad & \frac{1}{n^4} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} E\left[(\xi_i - p)(\xi_j - p)(\xi_k - p)(\xi_l - p)\right].
\end{aligned}
$$

注意到各ξ_i的独立性以及$E(\xi_i - p) = 0$，因此上面的和式中只有$E(\xi_i - p)^4$和$E[(\xi_i - p)^2(\xi_j - p)^2], i \neq j$的项才不等于0。

$$
\begin{aligned}
& E\left(\frac{\mu_n}{n}-p\right)^4 \\
= \quad & \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[(\xi_i-p)(\xi_j-p)(\xi_k-p)(\xi_l-p)\right].
\end{aligned}
$$

注意到各ξ_i的独立性以及$E(\xi_i-p)=0$，因此上面的和式中只有$E(\xi_i-p)^4$和$E[(\xi_i-p)^2(\xi_j-p)^2], i\neq j$的项才不等于0。

$$
E(\xi_i-p)^4=(1-p)^4p+p^4(1-p)=q^4p+p^4q=pq(p^3+q^3).
$$

共有n项。

$$
\begin{aligned}
& E\left(\frac{\mu_n}{n}-p\right)^4 \\
= \quad & \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[(\xi_i-p)(\xi_j-p)(\xi_k-p)(\xi_l-p)\right].
\end{aligned}
$$

注意到各ξ_i的独立性以及$E(\xi_i-p)=0$，因此上面的和式中只有$E(\xi_i-p)^4$和$E[(\xi_i-p)^2(\xi_j-p)^2], i\neq j$的项才不等于0。

$$
E(\xi_i-p)^4=(1-p)^4p+p^4(1-p)=q^4p+p^4q=pq(p^3+q^3).
$$

共有n项。

$$
E[(\xi_i-p)^2(\xi_j-p)^2]=E(\xi_i-p)^2E(\xi_j-p)^2=pqpq=p^2q^2, \quad i\neq j.
$$

共有$C_4^2C_n^2=3n(n-1)$ 项。

因此

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

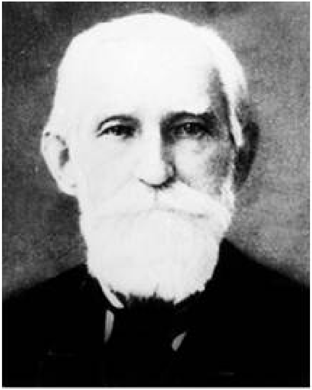
因此

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

于是，

$$P\left\{\left|\frac{\mu_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{1}{\varepsilon^4}E\left|\frac{\mu_n}{n}-p\right|^4<\frac{1}{4\varepsilon^4n^2},$$

保证收敛。



切比雪夫，П. Л.

切比雪夫

定理5.2.4

（切比雪夫大数定律）设$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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$D(\xi_i) \le C, i = 1, 2, \ldots$，则对任意的$\varepsilon > 0$，皆有

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

定理5.2.4

（切比雪夫大数定律）设$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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$D(\xi_i) \leq C, i = 1, 2, \ldots$，则对任意的$\varepsilon > 0$，皆有

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

证明：因为

$$E(\frac{1}{n}\sum_{k=1}^{n}\xi_k) = \frac{1}{n}\sum_{k=1}^{n}E(\xi_k),$$

$$D(\frac{1}{n}\sum_{k=1}^{n}\xi_k) = \frac{1}{n^2}\sum_{k=1}^{n}D(\xi_k) \leq \frac{C}{n}.$$

由Chebyshev不等式，对于任意的正实数ε有

$$0 \le P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| \ge \varepsilon\right\} \le \frac{D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right)}{\varepsilon^2} \le \frac{C}{n\varepsilon^2}.$$

由Chebyshev不等式，对于任意的正实数ε有

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

所以

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}E\xi_k\right| \geq \varepsilon\right\} = 0.$$

这个结果在1866年被俄国数学家切比雪夫所证明，它是关于大数定律的一个相当普遍的结论，许多大数定律的古典结果是它的特例。

此外，证明这个定律所用的方法后来称为矩法，也很有创造性，在这基础上发展起来的一系列不等式是研究各种极限定理的有力工具。

作为切比雪夫大数定律的特殊情况，有下面的推论。

作为切比雪夫大数定律的特殊情况，有下面的推论。

推论5.2.1

（独立同分布下的大数定律）

设$\xi_1, \xi_2, \ldots$是独立同分布的随机变量序列，

且$E\xi_i = \mu, D\xi_i = \sigma^2, i = 1, 2, \ldots$，则对任给$\varepsilon > 0$，

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

推论5.2.2

（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

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

推论5.2.2

（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

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

证明：定义ξ_k为第k次试验中事件A出现的次数，则

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

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

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

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5.2.5

（马尔可夫大数定律）

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

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

则对任意的$\varepsilon > 0$，皆有

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

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5.2.1

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

例5.2.1

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

解：由于$X_n \sim Exp(\frac{1}{\lambda}), n = 1, 2, \ldots$，从而有$D(X_n) = \lambda^2, n = 1, 2, \ldots$，于是有

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

例5.2.1

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

解：由于$X_n \sim Exp(\frac{1}{\lambda}), n = 1, 2, \ldots$，从而有$D(X_n) = \lambda^2, n = 1, 2, \ldots$，于是有

$$\begin{aligned}\frac{1}{n^2}D(\sum_{i=1}^{n} X_i) &= \frac{1}{n^2}\left[\sum_{i=1}^{n} D(X_i) + 2\sum_{i=1}^{n-1} Cov(X_i, X_{i+1})\right] \\ &\le \frac{1}{n^2}\left[\sum_{i=1}^{n} \lambda^2 + 2\sum_{i=1}^{n-1} \lambda^2\right] = \frac{1}{n^2}[n\lambda^2 + 2(n-1)\lambda^2].\end{aligned}$$

$$\frac{1}{n^2}D(\sum_{i=1}^{n} X_i) \le \frac{1}{n^2}[n\lambda^2 + 2(n-1)\lambda^2] \xrightarrow{n\to+\infty} 0.$$

四、Khintchin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，并不需要有这个要求，这就是有名的辛钦大数定律告诉我们的。

四、KHINTCHIN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，并不需要有这个要求，这就是有名的辛钦大数定律告诉我们的。

定理5.2.6

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

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

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

证明： 由于$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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

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

证明：由于$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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

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

对于固定的t，

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

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

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

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

定理5.2.7

（Kolmogrov强大数定律）设$\xi_1,\xi_2,\ldots,\xi_n,\ldots$是独立随机变量序列，且$\sum_{n=1}^{\infty}\frac{D(\xi_n)}{n^2}<\infty$，则成立

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[\xi_i-E(\xi_i)]=0\right)=1.$$

证明在P341-342。

利用对随机变量“截尾”的技巧，Kolmogrov 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。

利用对随机变量“截尾”的技巧，Kolmogrov 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5.2.8

(*Kolmogrov*) 设$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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$E\xi_i = a$.

证明在P342-344。

证明： 若ξ的分布函数为$F(x)$，我们来证明不等式：

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} \leq E|\xi| \leq 1 + \sum_{n=1}^{\infty} P\{|\xi| \geq n\}. \tag{1}$$

证明： 若ξ的分布函数为$F(x)$，我们来证明不等式：

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} \leq E|\xi| \leq 1 + \sum_{n=1}^{\infty} P\{|\xi| \geq n\}. \tag{1}$$

事实上，

$$E|\xi| = \int_{-\infty}^{\infty} |x| dF(x) = \sum_{k=0}^{\infty} \int_{k \leq |x| < k+1} |x| dF(x).$$

证明： 若ξ的分布函数为$F(x)$，我们来证明不等式：

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} \leq E|\xi| \leq 1 + \sum_{n=1}^{\infty} P\{|\xi| \geq n\}. \tag{1}$$

事实上，

$$E|\xi| = \int_{-\infty}^{\infty} |x| dF(x) = \sum_{k=0}^{\infty} \int_{k \leq |x| < k+1} |x| dF(x).$$

因此，

$$\sum_{k=0}^{\infty} kP\{k \leq |\xi| < k+1\} \leq E|\xi| \leq \sum_{k=0}^{\infty} (k+1)P\{k \leq |\xi| < k+1\}.$$

我们有，

$$\sum_{k=0}^{\infty} kP\{k \le |\xi| < k+1\} = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P\{k \le |\xi| < k+1\}$$

$$= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P\{k \le |\xi| < k+1\} = \sum_{n=1}^{\infty} P\{|\xi| \ge n\}.$$

我们有，

$$
\sum_{k=0}^{\infty} kP\{k \le |\xi| < k+1\} = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P\{k \le |\xi| < k+1\}
$$

$$
= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P\{k \le |\xi| < k+1\} = \sum_{n=1}^{\infty} P\{|\xi| \ge n\}.
$$

和

$$
\sum_{k=0}^{\infty}(k+1)P\{k \le |\xi| < k+1\} = \sum_{k=0}^{\infty} kP\{k \le |\xi| < k+1\} + 1
$$

$$
= \sum_{n=1}^{\infty} P\{|\xi| \ge n\} + 1.
$$

我们有，

$$
\begin{aligned}
& \sum_{k=0}^{\infty} kP\{k \le |\xi| < k+1\} = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P\{k \le |\xi| < k+1\} \\
= \; & \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P\{k \le |\xi| < k+1\} = \sum_{n=1}^{\infty} P\{|\xi| \ge n\}.
\end{aligned}
$$

和

$$
\begin{aligned}
& \sum_{k=0}^{\infty}(k+1)P\{k \le |\xi| < k+1\} = \sum_{k=0}^{\infty} kP\{k \le |\xi| < k+1\} + 1 \\
= \; & \sum_{n=1}^{\infty} P\{|\xi| \ge n\} + 1.
\end{aligned}
$$

因此，(1)式成立。

(1)式说明$E|\xi| < \infty$ 的充要条件是

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} < \infty.$$

(1)式说明$E|\xi| < \infty$ 的充要条件是

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} < \infty.$$

记$S_n = \sum_{i=1}^{n} \xi_i$。若$\frac{S_n}{n} \xrightarrow{a.s.} a$，$a$是一个有限数，则

$$\frac{\xi_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1} \xrightarrow{a.s.} 0.$$

(1)式说明$E|\xi| < \infty$ 的充要条件是

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} < \infty.$$

记$S_n = \sum_{i=1}^{n} \xi_i$。若$\frac{S_n}{n} \xrightarrow{a.s.} a$，$a$是一个有限数，则

$$\frac{\xi_n}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} \xrightarrow{a.s.} 0.$$

事件$\{|\xi_n| \geq n\}$ 发生无穷多次的概率为0,

$$P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left\{\frac{|\xi_n|}{n} \geq 1\right\}) = 0$$

(1)式说明$E|\xi| < \infty$ 的充要条件是

$$\sum_{n=1}^{\infty} P\{|\xi| \geq n\} < \infty.$$

记$S_n = \sum_{i=1}^{n} \xi_i$。若$\frac{S_n}{n} \xrightarrow{a.s.} a$，$a$是一个有限数，则

$$\frac{\xi_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1} \xrightarrow{a.s.} 0.$$

事件$\{|\xi_n| \geq n\}$ 发生无穷多次的概率为0，

$$P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left\{ \frac{|\xi_n|}{n} \geq 1 \right\}) = 0$$

加上已知ξ_i 相互独立，并利用博雷尔-康特立引理(ii)，可知

$$\sum_{n=1}^{\infty} P\{|\xi_n| \geq n\} < \infty.$$

(1)式可以知道$E|\xi| < \infty$.

(1)式可以知道$E|\xi| < \infty$.

因为

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

所以

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

(1)式可以知道$E|\xi| < \infty$.
因为
$$\frac{S_n}{n} \xrightarrow{a.s.} a,$$
所以
$$\frac{S_n}{n} \xrightarrow{p} a.$$
又由辛钦大数定律知道,
$$\frac{S_n}{n} \xrightarrow{p} E(\xi_i).$$

(1)式可以知道$E|\xi| < \infty$.
因为

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

所以

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

又由辛钦大数定律知道,

$$\frac{S_n}{n} \xrightarrow{p} E(\xi_i).$$

所以$E(\xi_i) = a$. 必要性已证。

下证充分性。用截尾法。令

$$\xi_n^* = \begin{cases} \xi_n, & |\xi_n| < n; \\ 0, & |\xi_n| \geq n. \end{cases}$$

下证充分性。用截尾法。令

$$\xi_n^* = \begin{cases} \xi_n, & |\xi_n| < n; \\ 0, & |\xi_n| \geq n. \end{cases}$$

先验证$\{\xi_n^*\}$满足柯尔莫哥洛夫强大数定律条件。以$F(x)$记ξ_i的分布函数，则

$$D(\xi_n^*) \leq E[(\xi_n^*)^2] \leq \sum_{k=1}^{n} k^2 P(k-1 \leq |\xi_n| \leq k).$$

$$
\begin{aligned}
\sum_{n=1}^{\infty} \frac{D(\xi_n^*)}{n^2} &\le \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{k^2}{n^2} P(k-1 \le |\xi_n| \le k) \\
&= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \frac{k^2}{n^2} P(k-1 \le |\xi_n| \le k) \\
&= \sum_{k=1}^{\infty} k^2 P(k-1 \le |\xi_n| \le k) \sum_{n=k}^{\infty} \frac{1}{n^2}.
\end{aligned}
$$

$$
\begin{aligned}
\sum_{n=1}^{\infty}\frac{D(\xi_n^*)}{n^2} &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{k^2}{n^2}P(k-1\le|\xi_n|\le k)\\
&= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\frac{k^2}{n^2}P(k-1\le|\xi_n|\le k)\\
&= \sum_{k=1}^{\infty}k^2P(k-1\le|\xi_n|\le k)\sum_{n=k}^{\infty}\frac{1}{n^2}.
\end{aligned}
$$

由于

$$
\sum_{n=k}^{\infty}\frac{1}{n^2}<\frac{1}{k^2}+\sum_{n=k+1}^{\infty}\frac{1}{n(n-1)}=\frac{1}{k^2}+\frac{1}{k}\le\frac{2}{k}.
$$

$$
\begin{aligned}
\sum_{n=1}^{\infty}\frac{D(\xi_n^*)}{n^2} &\le \sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{k^2}{n^2}P(k-1\le|\xi_n|\le k)\\
&= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\frac{k^2}{n^2}P(k-1\le|\xi_n|\le k)\\
&= \sum_{k=1}^{\infty}k^2P(k-1\le|\xi_n|\le k)\sum_{n=k}^{\infty}\frac{1}{n^2}.
\end{aligned}
$$

由于

$$\sum_{n=k}^{\infty}\frac{1}{n^2}<\frac{1}{k^2}+\sum_{n=k+1}^{\infty}\frac{1}{n(n-1)}=\frac{1}{k^2}+\frac{1}{k}\le\frac{2}{k}.$$

故由已知$E(\xi_i)$存在，

$$\sum_{n=1}^{\infty}\frac{D(\xi_n^*)}{n^2}\le 2\sum_{k=1}^{\infty}kP(k-1\le|\xi_n|\le k)<\infty.$$

因此由柯尔莫哥洛夫强大数定律，

$$P(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[\xi_i^*-E(\xi_i^*)]=0)=1.$$

因此由柯尔莫哥洛夫强大数定律,

$$P(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[\xi_i^* - E(\xi_i^*)] = 0) = 1.$$

显然 $\lim_{n\to\infty} E\xi_n^* = E\xi_1 = a$, 因此 $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} E\xi_i^* = E\xi_1 = a$.

因此由柯尔莫哥洛夫强大数定律,

$$P(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[\xi_i^*-E(\xi_i^*)]=0)=1.$$

显然 $\lim_{n\to\infty}E\xi_n^*=E\xi_1=a$, 因此 $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E\xi_i^*=E\xi_1=a$. 我们有

$$\left|\frac{1}{n}\sum_{i=1}^{n}(\xi_i-a)\right|\le\left|\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\xi_i^*)\right|$$

$$+\left|\frac{1}{n}\sum_{i=1}^{n}[\xi_i^*-E(\xi_i^*)]\right|+\left|\frac{1}{n}\sum_{i=1}^{n}[E(\xi_i^*)-a]\right|.$$

为证

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

只须证明

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\xi_i^*) \xrightarrow{a.s.} 0$$

即可。

为证

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

只须证明

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\xi_i^*) \xrightarrow{a.s.} 0$$

即可。

$$\sum_{i=1}^{\infty}P(\xi_i\neq\xi_i^*)=\sum_{i=1}^{\infty}P(|\xi_i|\geq i)\leq E|\xi_1|<\infty.$$

为证
$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{a.s.} a,$$
只须证明
$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\xi_i^*) \xrightarrow{a.s.} 0$$
即可。
$$\sum_{i=1}^{\infty}P(\xi_i\neq\xi_i^*)=\sum_{i=1}^{\infty}P(|\xi_i|\geq i)\leq E|\xi_1|<\infty.$$

由博雷尔-康特立引理(i)知，以概率1有
$$\xi_i\neq\xi_i^*, \text{只对有限个} i \text{成立}.$$

因此，

$$P(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\xi_i^*)=0)=1.$$

定理的证明完成。

小结：大数定律的意义

KHINTCHIN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：

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

小结：大数定律的意义

Khintchin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：

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

为寻找随机变量的期望值提供了一条实际可行的途径。

小结：大数定律的意义

Bernoulli大数定理

这一定理表明：在相同条件下重复同意随机试验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：

$$f_A \approx P(A) \qquad \text{当}n\text{充分大时.}$$

小结：大数定律的意义

BERNOULLI大数定理

这一定理表明：在相同条件下重复同意随机试验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：

$$f_A \approx P(A) \qquad \text{当}n\text{充分大时}.$$

为寻找随机事件概率提供了一条实际可行的途径。

小结：大数定律的意义

伯努利大数定律建立了大量重复独立试验中事件出现频率的稳定性，正因为这种稳定性，概率的概念才有客观意义。
同时，它还提供了通过试验来确定事件概率的方法。即用事件发生的频率作为它相应概率的估计。这类方法称为参数估计。它是统计中的重要研究课题

小结：大数定律的意义

观察个别现象时是连同一切个别的特性来观察的。这些个别的特性往往蒙蔽了事物的规律性。在大量观察中个别因素的影响将相互抵消而使总体稳定。

例如，虽然每个气体分子的运动带有很大的随机性，但是作为气体平均特征的压力、温度等却是稳定的，大数定律说明了这种稳定性。

古人用“定律”来称呼这类命题，可见其重视程度。

五、大数定律的应用

蒲丰投针问题中解法的理论依据就是大数定律。

五、大数定律的应用

蒲丰投针问题中解法的理论依据就是大数定律。

针长L，线距a。当投针次数n很大时，用针与线相交的频率m/n近似针与线相交的概率p，从而求得π 的近似值.

$$\pi \approx \frac{2Ln}{am}.$$

五、大数定律的应用

例5.2.2

（用蒙特卡洛方法计算定积分）计算积分

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

五、大数定律的应用

例5.2.2

（用蒙特卡洛方法计算定积分）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

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

$$\text{即} \qquad J = (b-a)\cdot Eg(\xi_i).$$

因此只要能求得$Eg(\xi_i)$，便能得到J的数值。

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

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

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。

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

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

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$\{\xi_i\}$服从$[a,b]$上的均匀分布。

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

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

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$\{\xi_i\}$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。

1 收敛性

2 大数定律

3 中心极限定理

5.3 中心极限定理

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?

伯努利大数定律只断言$\frac{\mu_n}{n}$接近于p，而棣莫弗一拉普拉斯极限定理则给出了$\frac{\mu_n}{n}$的渐进分布的更精确表达。

定理5.3.1

（棣莫弗—拉普拉斯(*De Moivre-Laplace*)中心极限定理）

若μ_n是n重伯努利试验中事件A出现的次数，$0 < p < 1$, 则对任意有限区间$[a, b]$,

1. 当$a \leq x_k := \dfrac{k - np}{\sqrt{npq}} \leq b$ 及$n \to \infty$ 时，一致地有

$$P(\mu_n = k) \div \left(\frac{1}{\sqrt{npq}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \to 1. \qquad (2)$$

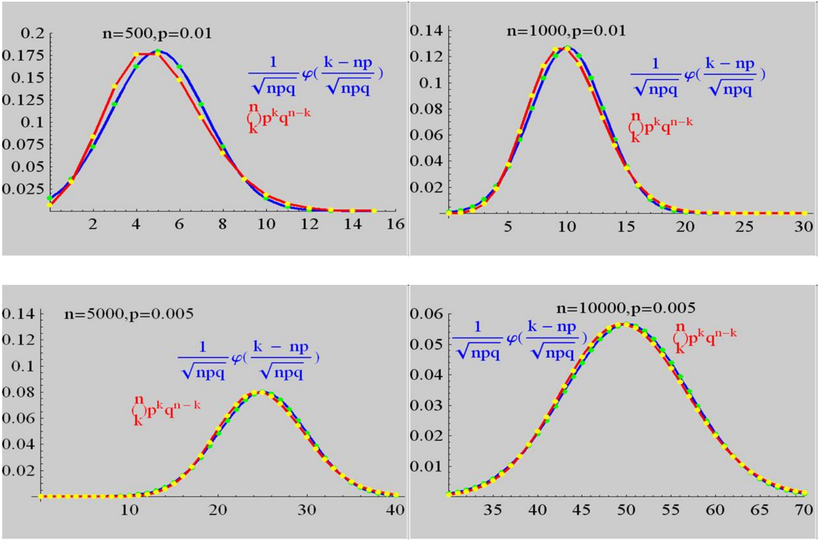
2. 当$n \to \infty$时，一致地有

$$P\left(a \leq \frac{\mu_n - np}{\sqrt{npq}} < b \right) \longrightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx. \qquad (3)$$

注5.3.1

1. 这里的第一个结果提供了$P(\mu_n = k)$的渐进表达式，这类结果称为局部极限定理。
2. 这里的第二个结果给出了标准化随机变量$\dfrac{\mu_n - np}{\sqrt{npq}}$的渐进分布，称为积分极限定理。它是一般中心极限定理的特例。
3. 这两个结果既有区别，又有联系。

局部极限定理的图形解释



证明：先证局部极限定理，我们将给出一个比(2)式更精确的渐进式。

证明：先证局部极限定理，我们将给出一个比(2)式更精确的渐进式。

因x_k只能在有限区间$[a, b]$中取值，故当$n \to \infty$ 时

$$k = np + x_k\sqrt{npq} \to \infty, \tag{4}$$

$$j := n - k = nq - x_k\sqrt{npq} \to \infty. \tag{5}$$

证明：先证局部极限定理，我们将给出一个比(2)式更精确的渐进式。

因x_k只能在有限区间$[a, b]$中取值，故当$n \to \infty$ 时

$$k = np + x_k\sqrt{npq} \to \infty, \tag{4}$$

$$j := n - k = nq - x_k\sqrt{npq} \to \infty. \tag{5}$$

由斯特林公式：

$$m! = \sqrt{2\pi m}m^m e^{-m} e^{\theta_m}, \quad (\frac{1}{12m+1} < \theta_m < \frac{1}{12m}).$$

$$
\begin{aligned}
P(\mu_n = k) &= \frac{n!}{k!j!}p^k q^j = \frac{\sqrt{2\pi n}n^n e^{-n}}{\sqrt{2\pi k}k^k e^{-k}\sqrt{2\pi j}j^j e^{-j}}p^k q^j e^{\theta_n-\theta_k-\theta_j} \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}\left(\frac{np}{k}\right)^{k+\frac{1}{2}}\left(\frac{nq}{n-k}\right)^{n-k+\frac{1}{2}} e^{\theta}, \qquad (6)
\end{aligned}
$$

其中$\theta = \theta_n - \theta_k - \theta_j$,

$$
|\theta| < \frac{1}{12}\left(\frac{1}{n} + \frac{1}{k} + \frac{1}{n-k}\right).
$$

由(4)和(5)

$$\frac{k}{np} = 1 + x_k\sqrt{\frac{q}{np}}, \quad \frac{n-k}{nq} = 1 - x_k\sqrt{\frac{p}{nq}}.$$

由(4)和(5)

$$\frac{k}{np} = 1 + x_k\sqrt{\frac{q}{np}}, \quad \frac{n-k}{nq} = 1 - x_k\sqrt{\frac{p}{nq}}.$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \quad -1 < x \le 1.$$

只有x的绝对值很小的时候才收敛得快。

当n充分大时，$x_k\sqrt{\frac{q}{np}}$和$x_k\sqrt{\frac{p}{nq}}$ 都很小。

由(4)和(5)

$$\frac{k}{np}=1+x_k\sqrt{\frac{q}{np}},\quad \frac{n-k}{nq}=1-x_k\sqrt{\frac{p}{nq}}.$$

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots,\quad -1<x\le 1.$$

只有x的绝对值很小的时候才收敛得快。

当n充分大时，$x_k\sqrt{\frac{q}{np}}$和$x_k\sqrt{\frac{p}{nq}}$ 都很小。

所以，$p=0$或者$p=1$时不能用；此外，当p或q很小时，渐进展开式引起的误差也较大，这时用泊松逼近公式。

因此,

$$
\begin{aligned}
& \ln\left(\sqrt{2\pi npq}P(\mu_n=k)\right) \\
= \quad & \theta-\left(k+\frac{1}{2}\right)\ln\frac{k}{np}-\left(n-k+\frac{1}{2}\right)\ln\frac{n-k}{nq} \\
= \quad & \theta-\left(np+x_k\sqrt{npq}+\frac{1}{2}\right)\ln\left(1+x_k\sqrt{\frac{q}{np}}\right)- \\
& \left(nq-x_k\sqrt{npq}+\frac{1}{2}\right)\ln\left(1-x_k\sqrt{\frac{p}{nq}}\right) \\
= \quad & \theta-\left(np+x_k\sqrt{npq}+\frac{1}{2}\right)\cdot\left(x_k\sqrt{\frac{q}{np}}-\frac{x_k^2q}{2np}+\frac{x_k^3q\sqrt{npq}}{3n^2p^2}-\frac{x_k^4q^2}{4n^2p^2}+\cdots\right) \\
& -\left(nq-x_k\sqrt{npq}+\frac{1}{2}\right)\cdot\left(-x_k\sqrt{\frac{p}{nq}}-\frac{x_k^2p}{2nq}-\frac{x_k^3p\sqrt{npq}}{3n^2q^2}-\frac{x_k^4p^2}{4n^2q^2}+\cdots\right) \\
= \quad & \theta-\frac{x_k^2}{2}+\frac{q-p}{6\sqrt{npq}}(x_k^3-3x_k)+\frac{1}{12npq}\left[3(p^2+q^2)x_k^2-(p^3+q^3)x_k^4\right]+o\left(\frac{1}{n}\right)
\end{aligned}
$$

因此,

$$
\begin{aligned}
& P(\mu_n = k) \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2} + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2}\right\} \exp\left\{\frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} e^{-\frac{x_k^2}{2}} \left[1 + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right].
\end{aligned}
$$

因此，

$$
\begin{aligned}
& P(\mu_n = k) \\
= & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2} + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2}\right\} \exp\left\{\frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} e^{-\frac{x_k^2}{2}} \left[1 + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right].
\end{aligned}
$$

取其第一项即得(2)式，因此已证得了局部极限定理。

因此,

$$
\begin{aligned}
& P(\mu_n = k) \\
= \ & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2} + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= \ & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \exp\left\{-\frac{x_k^2}{2}\right\} \exp\left\{\frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right\} \\
= \ & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} e^{-\frac{x_k^2}{2}} \left[1 + \frac{q-p}{6\sqrt{npq}}(x_k^3 - 3x_k) + O\left(\frac{1}{n}\right)\right].
\end{aligned}
$$

取其第一项即得(2)式,因此已证得了局部极限定理。
显然,我们得到了更精确的估计式。又知当$p = q$及$x_k^3 - 3x_k = 0$时,近似效果尤佳。

下面我们转入证明积分极限定理：

$$
\begin{aligned}
& P\left(a \le \frac{\mu_n - np}{\sqrt{npq}} < b\right) \\
= \quad & P(np + a\sqrt{npq} \le \mu_n < np + b\sqrt{npq}) = \sum_{k=k_1}^{k_2} P(\mu_n = k)
\end{aligned}
$$

下面我们转入证明积分极限定理：

$$
\begin{aligned}
& P\left(a \leq \frac{\mu_n - np}{\sqrt{npq}} < b\right) \\
= \quad & P(np + a\sqrt{npq} \leq \mu_n < np + b\sqrt{npq}) = \sum_{k=k_1}^{k_2} P(\mu_n = k)
\end{aligned}
$$

上式中k_1为不小于$np+a\sqrt{npq}$的最小整数，k_2 为小于$np+b\sqrt{npq}$的最大整数。

由局部极限定理，对任意的$\varepsilon > 0$，当n充分大时有

$$P(\mu_n = k) = \frac{1}{\sqrt{npq}}(\varphi(x_k) + \varepsilon_k), \quad |\varepsilon_k| < \varepsilon, k = k_1, k_1 + 1, \ldots, k_2.$$

由局部极限定理，对任意的$\varepsilon > 0$，当n充分大时有

$$P(\mu_n = k) = \frac{1}{\sqrt{npq}}(\varphi(x_k) + \varepsilon_k), \quad |\varepsilon_k| < \varepsilon, k = k_1, k_1 + 1, \ldots, k_2.$$

我们有

$$P\left(a \leq \frac{\mu_n - np}{\sqrt{npg}} < b\right) = \sum_{k=k_1}^{k_2} \frac{1}{\sqrt{npq}}\varphi(x_k) + \sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}}.$$

由局部极限定理，对任意的$\varepsilon > 0$，当n充分大时有

$$P(\mu_n = k) = \frac{1}{\sqrt{npq}}(\varphi(x_k) + \varepsilon_k), \quad |\varepsilon_k| < \varepsilon, k = k_1, k_1 + 1, \ldots, k_2.$$

我们有

$$P\left(a \leq \frac{\mu_n - np}{\sqrt{npg}} < b\right) = \sum_{k=k_1}^{k_2} \frac{1}{\sqrt{npq}}\varphi(x_k) + \sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}}.$$

因为

$$\left|\sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}}\right| \leq \frac{1}{\sqrt{npq}}(k_2 - k_1 + 1)\varepsilon \leq \frac{(b-a)\sqrt{npq} + 1}{\sqrt{npq}}\varepsilon,$$

故当$n \to \infty$时，注意x_k的增量为$\frac{1}{\sqrt{npq}}$，就得到

$$P\left(a \leq \frac{\mu_n - np}{\sqrt{npq}} < b\right) \longrightarrow \int_a^b \varphi(x)dx.$$

定理至此完全证毕。

故当$n \to \infty$时，注意x_k的增量为$\frac{1}{\sqrt{npq}}$，就得到

$$P\left(a \le \frac{\mu_n - np}{\sqrt{npq}} < b\right) \longrightarrow \int_a^b \varphi(x)dx.$$

定理至此完全证毕。
利用$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 1$，不难证明在积分极限定理中，当$a = -\infty, b = \infty$时仍然成立。

定理5.3.2

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

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

(续Lindeberg-Levy中心极限定理)即

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

（续Lindeberg-Levy中心极限定理）即

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

当 $n \to \infty$ 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

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \overset{a}{\sim} N(0,1), \quad n \to +\infty.$$

由正态分布的性质，有

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

由正态分布的性质，有

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

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

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

由正态分布的性质，有

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

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

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

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

$$f_n(t)=\left[f\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n.$$

证明：设$X_n-\mu$的特征函数为$f(t)$，则由于

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

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$f'(0)=0$.

证明：设$X_n-\mu$的特征函数为$f(t)$，则由于

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

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$f'(0)=0$.

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

证明：设$X_n-\mu$的特征函数为$f(t)$，则由于

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

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$f'(0)=0.$

$-f''(0)+(f'(0))^2=D(X_n-\mu)=\sigma^2$知$f''(0)=-\sigma^2.$
而$f(t)$的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，

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

由于$e^{-\frac{t^2}{2}}$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

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

证毕。

用这个定理可推出棣莫弗—拉普拉斯积分极限定理。

定理5.3.3

（*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

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

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

$$E(X) = np, \quad D(X) = npq.$$

由Lindeberg-Levy中心极限定理知

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

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

$$E(X) = np, \quad D(X) = npq.$$

由Lindeberg-Levy中心极限定理知

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

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，只有存在有限数学期望和方差，当n充分大时，就有

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

所以，$\sum\limits_{i=1}^{n} X_i$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$X \sim B(n,p)$，总有$X \overset{n\to+\infty}{\sim} N(np,npq)$，因此，当$n$充分大时，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很普遍。

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很普遍。

例5.3.1

（正态随机数的产生）

设$\xi_1, \xi_2, \ldots, \xi_n, \ldots$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$\xi_1 + \xi_2 + \cdots + \xi_n$ 渐进于正态变量。

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$n = 12$.

在二项分布计算中的应用

由积分极限定理，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：

$$\begin{aligned} P\{k_1 \le \mu_n \le k_2\} &= P\left\{\frac{k_1 - np}{\sqrt{npq}} \le \frac{\mu_n - np}{\sqrt{npq}} \le \frac{k_2 - np}{\sqrt{npq}}\right\} \\ &\approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right). \end{aligned}$$

在二项分布计算中的应用

由积分极限定理，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：

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

实际计算中，往往用下面的修正公式计算效果更好。

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

在二项分布计算中的应用

在实际计算中，为什么要求p不太接近于0或1?

在二项分布计算中的应用

在实际计算中，为什么要求p不太接近于0或1?

因为如果p接近于0或1, 则pq接近于0，误差太大。

例5.3.2

（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？

例5.3.2

（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。

例5.3.2

（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，问题是要求r，使

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

我们可以利用中心极限定理来计算这个概率。

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

$$\approx \Phi\left(\frac{r-200\times0.6+0.5}{\sqrt{200\times0.6\times0.4}}\right)-\Phi\left(\frac{-200\times0.6-0.5}{\sqrt{200\times0.6\times0.4}}\right)$$

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

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

$$
\begin{aligned}
&\sum_{k=0}^{r}\binom{200}{k}(0.6)^k(0.4)^{200-k}\\
\approx\ & \Phi\left(\frac{r-200\times 0.6+0.5}{\sqrt{200\times 0.6\times 0.4}}\right)-\Phi\left(\frac{-200\times 0.6-0.5}{\sqrt{200\times 0.6\times 0.4}}\right)\\
\approx\ & \Phi\left(\frac{r-119.5}{\sqrt{48}}\right)-\Phi(-17.39)\approx\Phi\left(\frac{r-119.5}{\sqrt{48}}\right)\geq 0.999.
\end{aligned}
$$

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

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。

近似数定点运算的误差分析

数值计算时，任何数x都只能用一定位数的有限小数y来近似，这就产生了一个误差$\xi = x - y$。

在下面的讨论中，我们假定参加运算的数都用十进制定点表示，每个数都用四舍五入的方法取到小数点后五位，这时相应的舍入误差可以看作是

$$[-0.5 \times 10^{-5}, 0.5 \times 10^{-5})$$

上的均匀分布。

现在如果要求n个数$x_i(i = 1, 2, \ldots, n)$的和S，在数值计算中就只能求出相应的有限位小数$y_i(i = 1, 2, \ldots, n)$的和T，并用T作为S的近似值。

自然要问，这样做造成的误差$\eta = S - T$是多少？

现在如果要求n个数$x_i(i=1,2,\ldots,n)$的和S，在数值计算中就只能求出相应的有限位小数$y_i(i=1,2,\ldots,n)$的和T，并用T作为S的近似值。

自然要问，这样做造成的误差$\eta=S-T$是多少？

因为我们有

$$S=\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}(y_i+\xi_i)=\sum_{i=1}^{n}y_i+\sum_{i=1}^{n}\xi_i.$$

故

$$\eta=\sum_{i=1}^{n}\xi_i.$$

一种传统的估计方法是这样的：由于

$$|\xi_i| \le 0.5 \times 10^{-5},$$

所以

$$|\eta| \le \sum_{i=1}^{n} |\xi_i| \le n \times 0.5 \times 10^{-5}.$$

一种传统的估计方法是这样的：由于

$$|\xi_i| \le 0.5 \times 10^{-5},$$

所以

$$|\eta| \le \sum_{i=1}^{n} |\xi_i| \le n \times 0.5 \times 10^{-5}.$$

以$n = 10000$为例，所得的误差估计为$|\eta| \le 0.05$。

这种估计方法显然太保守，下用概率论方法估计。

一种传统的估计方法是这样的：由于

$$|\xi_i| \le 0.5 \times 10^{-5},$$

所以

$$|\eta| \le \sum_{i=1}^{n} |\xi_i| \le n \times 0.5 \times 10^{-5}.$$

以$n = 10000$为例，所得的误差估计为$|\eta| \le 0.05$。

这种估计方法显然太保守，下用概率论方法估计。

这时直接求$\eta = \sum_{i=1}^{n} \xi_i$的分布不容易，但当$n$较大时用极限定理作为工具，则能使问题很快得到解决。因为

$$\sigma = \sqrt{D\xi_i} = \sqrt{\frac{(1 \times 10^{-5})^2}{12}} = \frac{0.5 \times 10^{-5}}{\sqrt{3}}, \quad m = E\xi_i = 0.$$

如果假定舍入误差ξ_i是相互独立的，n又较大，那么根据Lindeberg-Levy中心极限定理知：

$$P\left\{\left|\sum_{i=1}^{n}\xi_i\right| < k\sqrt{n}\sigma\right\} \approx \frac{1}{\sqrt{2\pi}}\int_{-k}^{k} e^{-\frac{t^2}{2}}\,dt.$$

取$k=3$时上式右边为0.997，因此我们能以99.7%的概率断言：

$$|\eta| < 3\times 100\times\frac{0.5\times 10^{-5}}{\sqrt{3}} = 0.866\times 10^{-3}.$$

这仅仅是传统估计法中误差上限的60分之一。