第四章　大数定律与中心极限定理

本章主要是解决概率论中的一些基本问题，如频率稳定性，正态分布的普适性问题。

§4.1　随机变量序列的两种收敛性

与数列的收敛性不同，随机变量序列的收敛性有多种，依概率收敛（Convergence in Probability）、按分布收敛（Convergence in Distribution）、几乎处处收敛（Almost Everywhere Convergence）、依均方收敛（Convergence in Mean Square）等。本节介绍依概率收敛和按分布收敛。

4.1.1 依概率收敛（Convergence in Probability）

数学分析中的极限 $\lim\limits_{n\to\infty} x_n = a$，也就是对任意给定的 $\varepsilon > 0$，存在正整数 N，当 $n > N$ 时，$|x_n - a| < \varepsilon$ 恒成立。在概率论中频率与概率的关系也是一种极限，但与数学分析中的极限不同。

不妨设事件 A 在每次试验中发生的概率为 p，在 n 重伯努利试验中的发生次数为 S_n，即发生的频率为 $\frac{S_n}{n}$。直观上，随 n 的无限增大，频率 $\frac{S_n}{n}$ 将以概率 p 为“极限”，但对很小的正数 $\varepsilon > 0$，无论 n 有多大，都不能使得 $\left|\frac{S_n}{n} - p\right| < \varepsilon$ 恒成立。如掷硬币试验，正面朝上的概率 $p = 0.5$，掷 n 次硬币，正面朝上的次数为 S_n，理论上，无论 n 有多大，都有可能次次正面朝上（当然发生的概率非常小），即频率 $\frac{S_n}{n} = 1$。对给定的正数 $\varepsilon = 0.1$，无论 n 有多大，都可能有 $\left|\frac{S_n}{n} - 0.5\right| = 0.5 > \varepsilon = 0.1$。

不过，当 n 很大时，$\left|\frac{S_n}{n} - p\right| < \varepsilon$ 虽然不是恒成立，但可以看出成立的概率将很接近于 1，可以证明，当 $n \to +\infty$ 时，$\left|\frac{S_n}{n} - p\right| < \varepsilon$ 成立的概率将趋于 1，即 $\lim\limits_{n\to+\infty} P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} = 1$。

如果随机变量序列 $\{X_n\}$ 和常数 a，对任意给定的 $\varepsilon > 0$，满足 $\lim\limits_{n\to+\infty} P\{|X_n - a| < \varepsilon\} = 1$，则称 $\{X_n\}$ 依概率收敛于 a，这就是直观上所说的 X_n 稳定于 a。

定义　（依概率收敛）设 $\{X_n\}$ 是一个随机变量序列，X 是一个随机变量。如果对任意的 $\varepsilon > 0$，有

$$\lim_{n\to+\infty} P\{|X_n - X| < \varepsilon\} = 1 。$$

则称 $\{X_n\}$ 依概率收敛于 X，记为 $X_n \xrightarrow{P} X$。

定理　收敛于常数时，四则运算保持依概率收敛性。若 $X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，则

（1）$X_n \pm Y_n \xrightarrow{P} a \pm b$；（2）$X_n Y_n \xrightarrow{P} ab$；（3）$\frac{X_n}{Y_n} \xrightarrow{P} \frac{a}{b}, \quad (b \neq 0)$。

证明：（1）因

$$|(X_n \pm Y_n)-(a \pm b)|=|(X_n-a)\pm(Y_n-b)| \le |X_n-a|+|Y_n-b|,$$

对任意的$\varepsilon>0$，有

$$\{|(X_n \pm Y_n)-(a \pm b)| \ge \varepsilon\} \subset \left\{|X_n-a| \ge \frac{\varepsilon}{2}\right\} \cup \left\{|Y_n-b| \ge \frac{\varepsilon}{2}\right\}$$

由概率的单调性和半可加性可得

$$0 \le P\{|(X_n \pm Y_n)-(a \pm b)| \ge \varepsilon\} \le P\left\{|X_n-a| \ge \frac{\varepsilon}{2}\right\}+P\left\{|Y_n-b| \ge \frac{\varepsilon}{2}\right\},$$

又因$X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，有

$$\lim_{n\to+\infty} P\left\{|X_n-a| \ge \frac{\varepsilon}{2}\right\}=0,\quad \lim_{n\to+\infty} P\left\{|Y_n-b| \ge \frac{\varepsilon}{2}\right\}=0,$$

故$\lim\limits_{n\to+\infty} P\{|(X_n \pm Y_n)-(a \pm b)| \ge \varepsilon\}=0$，即$X_n \pm Y_n \xrightarrow{P} a \pm b$；

（2）因

$$|X_nY_n-ab|=|(X_n-a)Y_n+a(Y_n-b)| \le |X_n-a|\cdot|Y_n|+|a|\cdot|Y_n-b|,$$

对任意的$\varepsilon>0$，有

$$\{|X_nY_n-ab| \ge \varepsilon\} \subset \left\{|X_n-a|\cdot|Y_n| \ge \frac{\varepsilon}{2}\right\} \cup \left\{|a|\cdot|Y_n-b| \ge \frac{\varepsilon}{2}\right\}。$$

由概率的单调性和半可加性可得

$$0 \le P\{|X_nY_n-ab| \ge \varepsilon\} \le P\left\{|X_n-a|\cdot|Y_n| \ge \frac{\varepsilon}{2}\right\}+P\left\{|a|\cdot|Y_n-b| \ge \frac{\varepsilon}{2}\right\}。$$

当$a=0$时，$\left\{|a|\cdot|Y_n-b| \ge \frac{\varepsilon}{2}\right\}=\varnothing$；当$a \ne 0$时，$\left\{|a|\cdot|Y_n-b| \ge \frac{\varepsilon}{2}\right\}=\left\{|Y_n-b| \ge \frac{\varepsilon}{2|a|}\right\}$，因$Y_n \xrightarrow{P} b$，

有$\lim\limits_{n\to+\infty} P\left\{|Y_n-b| \ge \frac{\varepsilon}{2|a|}\right\}=0$。因此，总有$\lim\limits_{n\to+\infty} P\left\{|a|\cdot|Y_n-b| \ge \frac{\varepsilon}{2}\right\}=0$。

对任意的正数$M>0$，有

$$\left\{|X_n-a|\cdot|Y_n| \ge \frac{\varepsilon}{2}\right\} \subset \{|Y_n| \ge M\} \cup \left\{|X_n-a| \ge \frac{\varepsilon}{2M}\right\},$$

可得

$$0 \le P\left\{|X_n-a|\cdot|Y_n| \ge \frac{\varepsilon}{2}\right\} \le P\{|Y_n| \ge M\}+P\left\{|X_n-a| \ge \frac{\varepsilon}{2M}\right\},$$

因$Y_n \xrightarrow{P} b$，有$\lim\limits_{n\to+\infty} P\{|Y_n-b| \ge 1\}=0$。而$\{|Y_n-b| \ge 1\}=\{Y_n \le b-1\} \cup \{Y_n \ge b+1\}$，有$\{|Y_n| \ge |b|+1\} \subset \{|Y_n-b| \ge 1\}$，即

$$0 \le P\{|Y_n| \ge |b|+1\} \le P\{|Y_n - b| \ge 1\},$$

可得 $\lim\limits_{n\to+\infty} P\{|Y_n| \ge |b|+1\} = 0$。取 $M = |b|+1$，有

$$0 \le P\left\{|X_n - a| \cdot |Y_n| \ge \frac{\varepsilon}{2}\right\} \le P\{|Y_n| \ge |b|+1\} + P\left\{|X_n - a| \ge \frac{\varepsilon}{2(|b|+1)}\right\},$$

因 $X_n \xrightarrow{P} a$，有 $\lim\limits_{n\to+\infty} P\left\{|X_n - a| \ge \frac{\varepsilon}{2(|b|+1)}\right\} = 0$，可得

$$\lim_{n\to+\infty} P\left\{|X_n - a| \cdot |Y_n| \ge \frac{\varepsilon}{2}\right\} = 0。$$

故 $\lim\limits_{n\to+\infty} P\{|X_n Y_n - ab| \ge \varepsilon\} = 0$，即 $X_n Y_n \xrightarrow{P} ab$；

（3）因

$$\left|\frac{X_n}{Y_n} - \frac{a}{b}\right| = \left|\frac{bX_n - aY_n}{bY_n}\right| = \frac{|b(X_n - a) - a(Y_n - b)|}{|bY_n|} \le \frac{|b| \cdot |X_n - a| + |a| \cdot |Y_n - b|}{|b| \cdot |Y_n|},$$

对任意的$\varepsilon > 0$，有

$$\left\{\left|\frac{X_n}{Y_n} - \frac{a}{b}\right| \ge \varepsilon\right\} \subset \left\{\frac{1}{|Y_n|} \cdot |X_n - a| \ge \frac{\varepsilon}{2}\right\} \cup \left\{\frac{|a|}{|b| \cdot |Y_n|} \cdot |Y_n - b| \ge \frac{\varepsilon}{2}\right\}。$$

由概率的单调性和半可加性可得

$$0 \le P\left\{\left|\frac{X_n}{Y_n} - \frac{a}{b}\right| \ge \varepsilon\right\} \le P\left\{\frac{1}{|Y_n|} \cdot |X_n - a| \ge \frac{\varepsilon}{2}\right\} + P\left\{\frac{|a|}{|b| \cdot |Y_n|} \cdot |Y_n - b| \ge \frac{\varepsilon}{2}\right\}$$

因 $Y_n \xrightarrow{P} b$，$b \ne 0$，有 $\lim\limits_{n\to+\infty} P\left\{|Y_n - b| \ge \frac{|b|}{2}\right\} = 0$。而 $\left\{|Y_n - b| \ge \frac{|b|}{2}\right\} = \left\{Y_n \le b - \frac{|b|}{2}\right\} \cup \left\{Y_n \ge b + \frac{|b|}{2}\right\}$，

有 $\left\{|Y_n| \le \frac{|b|}{2}\right\} \subset \left\{|Y_n - b| \ge \frac{|b|}{2}\right\}$，即

$$0 \le P\left\{|Y_n| \le \frac{|b|}{2}\right\} \le P\left\{|Y_n - b| \ge \frac{|b|}{2}\right\},$$

可得 $\lim\limits_{n\to+\infty} P\left\{|Y_n| \le \frac{|b|}{2}\right\} = 0$。

因

$$\left\{\frac{1}{|Y_n|} \cdot |X_n - a| \ge \frac{\varepsilon}{2}\right\} \subset \left\{|Y_n| \le \frac{|b|}{2}\right\} \cup \left\{|X_n - a| \ge \frac{|b|\varepsilon}{4}\right\}$$

可得

$$0 \le P\left\{\frac{1}{|Y_n|} \cdot |X_n - a| \ge \frac{\varepsilon}{2}\right\} \le P\left\{|Y_n| \le \frac{|b|}{2}\right\} + P\left\{|X_n - a| \ge \frac{|b|\varepsilon}{4}\right\},$$

因 $X_n \xrightarrow{P} a$，有 $\lim\limits_{n\to+\infty} P\left\{|X_n - a| \ge \frac{|b|\varepsilon}{4}\right\} = 0$，可得

$$\lim_{n\to+\infty} P\left\{\frac{1}{|Y_n|}\cdot|X_n - a| \ge \frac{\varepsilon}{2}\right\} = 0\text{。}$$

当 $a = 0$ 时，$\left\{\frac{|a|}{|b|\cdot|Y_n|}\cdot|Y_n - b| \ge \frac{\varepsilon}{2}\right\} = \varnothing$；当 $a \ne 0$ 时，

$$\left\{\frac{|a|}{|b|\cdot|Y_n|}\cdot|Y_n - b| \ge \frac{\varepsilon}{2}\right\} \subset \left\{|Y_n| \le \frac{|b|}{2}\right\} \cup \left\{|X_n - a| \ge \frac{b^2\varepsilon}{4|a|}\right\},$$

可得

$$0 \le P\left\{\frac{|a|}{|b|\cdot|Y_n|}\cdot|Y_n - b| \ge \frac{\varepsilon}{2}\right\} \le P\left\{|Y_n| \le \frac{|b|}{2}\right\} + P\left\{|X_n - a| \ge \frac{b^2\varepsilon}{4|a|}\right\},$$

因 $X_n \xrightarrow{P} a$，有 $\lim\limits_{n\to+\infty} P\left\{|X_n - a| \ge \frac{b^2\varepsilon}{4|a|}\right\} = 0$。因此，总有

$$\lim_{n\to+\infty} P\left\{\frac{|a|}{|b|\cdot|Y_n|}\cdot|Y_n - b| \ge \frac{\varepsilon}{2}\right\} = 0\text{。}$$

故 $\lim\limits_{n\to+\infty} P\left\{\left|\frac{X_n}{Y_n} - \frac{a}{b}\right| \ge \varepsilon\right\} = 0$，即 $\frac{X_n}{Y_n} \xrightarrow{P} \frac{a}{b}$。

4.1.2 按分布收敛（Convergence in Distribution）、弱收敛（Weak Convergence）

随机变量的特性由其分布函数完全确定，对于一个随机变量序列 $\{X_n\}$ 需要考虑其分布函数序列 $\{F_n(x)\}$ 收敛情况。但即使 $\{F_n(x)\}$ 处处收敛，其极限函数不一定是一个分布函数。

如 X_n 服从均匀分布 $U\left(-\frac{1}{n}, \frac{1}{n}\right)$，当 $n \to +\infty$ 时，直观上可看出 X_n 将“收敛于”在常数 0 处单点分布的随机变量 X。但 X_n 与 X 的分布函数 $F_n(x)$ 与 $F(x)$ 分别为

$$F_n(x) = \begin{cases} 0, & x < -\frac{1}{n}; \\ \frac{nx+1}{2}, & -\frac{1}{n} \le x < \frac{1}{n}; \\ 1, & x \ge \frac{1}{n}. \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0; \\ 1, & x \ge 0. \end{cases}$$

而

$$\lim_{n\to+\infty}F_n(x)=\begin{cases}0, & x<0;\\ \frac{1}{2}, & x=0;\\ 1, & x>0.\end{cases}$$

不满足右连续性，不是一个分布函数，$\lim_{n\to+\infty}F_n(x)\neq F(x)$。

可见，分布函数序列$\{F_n(x)\}$的极限函数不一定是一个分布函数，但该极限函数与一个分布函数在连续点处是相同的。因此，只要求$\{F_n(x)\}$在分布函数$F(x)$的连续点处收敛于$F(x)$。

定义 设随机变量序列$\{X_n\}$与随机变量X，分布函数分别为$F_n(x)$与$F(x)$，如果对于$F(x)$的任一连续点x，都有

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

则称$\{F_n(x)\}$弱收敛于$F(x)$，记为$F_n(x)\xrightarrow{W}F(x)$；也称$\{X_n\}$按分布收敛于$X$，记为$X_n\xrightarrow{L}X$。

当$\{X_n\}$按分布收敛于X时，又称$n\to+\infty$时，$\{X_n\}$的极限分布是X的分布。

由于分布函数$F(x)$单调不减，不连续点最多只有可列个。这样，若$F_n(x)\xrightarrow{W}F(x)$，则$\{F_n(x)\}$不收敛于$F(x)$的点最多只有可列个。

两种收敛的关系是依概率收敛必然按分布收敛，但按分布收敛不一定有依概率收敛。

定理 依概率收敛必然按分布收敛，即$X_n\xrightarrow{P}X\Rightarrow X_n\xrightarrow{L}X$。

证明：因$X_n\xrightarrow{P}X$，则对任意的$\varepsilon>0$，有$\lim_{n\to+\infty}P\{|X_n-X|\geq\varepsilon\}=0$。设$x$是$F(x)$的任一连续点，有

$$\{X\leq x-\varepsilon\}\subset\{X_n\leq x\}\bigcup\{X_n-X\geq\varepsilon\},$$

可得

$$F(x-\varepsilon)=P\{X\leq x-\varepsilon\}\leq P\{X_n\leq x\}+P\{X_n-X\geq\varepsilon\}=F_n(x)+P\{X_n-X\geq\varepsilon\},$$

因$0\leq P\{X_n-X\geq\varepsilon\}\leq P\{|X_n-X|\geq\varepsilon\}$且$\lim_{n\to+\infty}P\{|X_n-X|\geq\varepsilon\}=0$，则$F(x-\varepsilon)$不超过$F_n(x)$的任一收敛子序列的极限，即$F(x-\varepsilon)\leq\varliminf_{n\to+\infty}F_n(x)$，再令$\varepsilon\to0^+$，有$F(x-0)\leq\varliminf_{n\to+\infty}F_n(x)$；

又因

$$\{X_n\leq x\}\subset\{X\leq x+\varepsilon\}\bigcup\{X-X_n\geq\varepsilon\},$$

可得

$$F_n(x)=P\{X_n\leq x\}\leq P\{X\leq x+\varepsilon\}+P\{X-X_n\geq\varepsilon\}=F(x+\varepsilon)+P\{X-X_n\geq\varepsilon\},$$

因$0\leq P\{X-X_n\geq\varepsilon\}\leq P\{|X_n-X|\geq\varepsilon\}$且$\lim_{n\to+\infty}P\{|X_n-X|\geq\varepsilon\}=0$，则$F_n(x)$的任一收敛子序列的极限不超过$F(x+\varepsilon)$，即$\varlimsup_{n\to+\infty}F_n(x)\leq F(x+\varepsilon)$，再令$\varepsilon\to0^+$，有$\varlimsup_{n\to+\infty}F_n(x)\leq F(x+0)$。从而

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

因 x 是 $F(x)$ 的任一连续点，有 $F(x-0)=F(x+0)=F(x)$，则

$$\varliminf_{n\to+\infty} F_n(x) = \varlimsup_{n\to+\infty} F_n(x) = F(x)\text{，}$$

故 $\lim\limits_{n\to+\infty} F_n(x)=F(x)$，即 $X_n \xrightarrow{L} X$ 。

但按分布收敛不一定有依概率收敛，如随机变量 X 与随机变量序列 $\{X_n\}$ 独立同分布，都服从标准正态分布 $N(0,1)$，则分布函数 $F(x)$ 与 $F_n(x)$ 相同，都是 $\Phi(x)$，当然有 $F_n(x)\xrightarrow{W}F(x)$，即 $X_n\xrightarrow{L}X$ 。但因为 X_n-X 服从正态分布，且 $E(X_n-X)=0$， $\mathrm{Var}(X_n-X)=2$，即 $X_n-X\sim N(0,2)$，对任意的 $\varepsilon>0$，有

$$\lim_{n\to+\infty} P\{|X_n-X|\ge\varepsilon\} = 2\lim_{n\to+\infty} P\{X_n-X\ge\varepsilon\} = 2-2\Phi\left(\frac{\varepsilon}{\sqrt{2}}\right)>0\text{，}$$

故 $\{X_n\}$ 不是依概率收敛于 X 。

事实上，按分布收敛只是对随机变量序列的分布有要求，而对于随机变量取值之间的关系没有要求，而依概率收敛则更进一步对随机变量取值之间的关系有要求。

不过，当极限为常数时，依概率收敛的充分必要条件是按分布收敛。

定理　当极限为常数时，依概率收敛的充分必要条件是按分布收敛，即 $X_n\xrightarrow{P}c \Leftrightarrow X_n\xrightarrow{L}c$ 。

证明：必要性前面定理已证明，再证明充分性。设 $X_n\xrightarrow{L}c$，且 X_n 的分布函数为 $F_n(x)$ 。而常数 c 为单点分布，其分布函数为

$$F(x)=\begin{cases}0, & x<c;\\ 1, & x\ge c.\end{cases}$$

当 $x\ne c$ 时，x 是 $F(x)$ 的连续点，$F_n(x)$ 收敛于 $F(x)$ 。若 $x<c$， $\lim\limits_{n\to+\infty}F_n(x)=0$；若 $x>c$， $\lim\limits_{n\to+\infty}F_n(x)=1$ 。则

$$1\ge P\{|X_n-c|<\varepsilon\} = P\{c-\varepsilon<X_n<c+\varepsilon\} \ge P\left\{c-\varepsilon<X_n\le c+\frac{\varepsilon}{2}\right\}$$

$$=F_n\left(c+\frac{\varepsilon}{2}\right)-F_n(c-\varepsilon)\text{，}$$

因

$$\lim_{n\to+\infty}\left[F_n\left(c+\frac{\varepsilon}{2}\right)-F_n(c-\varepsilon)\right]=1\text{，}$$

故由夹逼准则（Squeeze Rule）可得 $\lim\limits_{n\to+\infty}P\{|X_n-c|<\varepsilon\}=1$，即 $X_n\xrightarrow{P}c$ 。

§4.2 特征函数

现实生活中，独立随机变量和广泛存在，它在概率论中具有重要地位，但其分布的计算需要用到卷积运算。而函数 $f(x)$ 的傅立叶变换（Fourier Transformation）$\varphi(t)=\int_{-\infty}^{+\infty}\mathrm{e}^{itx}f(x)dx$（其中 $i=\sqrt{-1}$ 是虚数单位）可将函数复杂的卷积运算转换为简单的乘法运算，并且可用傅立叶逆变换（Inverse Fourier Transform）将 $\varphi(t)$ 还原为 $f(x)=\dfrac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-itx}\varphi(t)dt$。

对连续随机变量 X 的密度函数 $p(x)$ 作傅立叶变换，就是数学期望 $E(\mathrm{e}^{itX})$，将该期望值称为随机变量 X 的特征函数。

4.2.1 特征函数的定义

定义 设 X 是一个随机变量，称

$$\varphi(t)=E(\mathrm{e}^{itX}),\quad -\infty<t<+\infty$$

为 X 的特征函数（Characteristic Function）；又称

$$M(u)=E(\mathrm{e}^{uX}),\quad -\infty<u<+\infty$$

为 X 的矩母函数（Moment-generating Function）。

由欧拉公式 $\mathrm{e}^{i\theta}=\cos\theta+i\sin\theta$ 可知 $|\mathrm{e}^{itX}|=1$，所以特征函数 $\varphi(t)=E(\mathrm{e}^{itX})$ 总是存在。

对于离散随机变量，设 X 的分布列为 $P\{X=x_k\}=p_k,\quad k=1,2,\cdots$，则 X 的特征函数为

$$\varphi(t)=\sum_{k=1}^{+\infty}\mathrm{e}^{itx_k}\,p_k,\quad -\infty<t<+\infty\text{；}$$

对于连续随机变量，设 X 的密度函数为 $p(x)$，则 X 的特征函数为

$$\varphi(t)=\int_{-\infty}^{+\infty}\mathrm{e}^{itx}\,p(x)dx,\quad -\infty<t<+\infty\text{。}$$

常用分布的特征函数

（1）单点分布：分布列 $P\{X=a\}=1$，特征函数为

$$\varphi(t)=\mathrm{e}^{iat}\text{；}$$

（2）二项分布 $b(n,p)$：分布列 $P\{X=k\}=C_n^k p^k(1-p)^{n-k},\quad k=0,1,2,\cdots,n$，特征函数为

$$\varphi(t)=\sum_{k=0}^{n}\mathrm{e}^{ikt}\cdot C_n^k p^k(1-p)^{n-k}=\sum_{k=0}^{n}\mathrm{e}^{ikt}\cdot C_n^k(p\,\mathrm{e}^{it})^k(1-p)^{n-k}=(1-p+p\,\mathrm{e}^{it})^n\text{；}$$

特别是，当 $n=1$ 时，两点分布，分布列 $P\{X=1\}=p,\quad P\{X=0\}=1-p$，特征函数为

$$\varphi(t)=1-p+p\,\mathrm{e}^{it}\text{；}$$

（3）泊松分布 $P(\lambda)$：分布列 $P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\quad k=0,1,2,\cdots$，特征函数为

$$\varphi(t)=\sum_{k=0}^{+\infty}\mathrm{e}^{ikt}\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\cdot\sum_{k=0}^{+\infty}\frac{(\lambda\,\mathrm{e}^{it})^k}{k!}=\mathrm{e}^{-\lambda}\cdot\mathrm{e}^{\lambda\mathrm{e}^{it}}=\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}\text{；}$$

（4）均匀分布 $U(a,b)$：密度函数 $p(x)=\dfrac{1}{b-a}$，　$a<x<b$，特征函数为

$$\varphi(t)=\int_a^b \mathrm{e}^{itx}\cdot\frac{1}{b-a}dx=\frac{1}{b-a}\cdot\frac{\mathrm{e}^{itx}}{it}\bigg|_a^b=\frac{\mathrm{e}^{ibt}-\mathrm{e}^{iat}}{it(b-a)};$$

（5）指数分布 $Exp(\lambda)$：密度函数 $p(x)=\lambda\,\mathrm{e}^{-\lambda x}$，　$x>0$，特征函数为

$$\varphi(t)=\int_0^{+\infty}\mathrm{e}^{itx}\cdot\lambda\,\mathrm{e}^{-\lambda x}\,dx=\int_0^{+\infty}\lambda\,\mathrm{e}^{-(\lambda-it)x}\,dx=\lambda\cdot\frac{\mathrm{e}^{-(\lambda-it)x}}{-(\lambda-it)}\bigg|_0^{+\infty}=\frac{\lambda}{\lambda-it};$$

（6）伽玛分布 $Ga(\alpha,\lambda)$：密度函数 $p(x)=\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\,\mathrm{e}^{-\lambda x}$，　$x>0$，特征函数为

$$\varphi(t)=\int_0^{+\infty}\mathrm{e}^{itx}\cdot\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\,\mathrm{e}^{-\lambda x}\,dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^{+\infty}x^{\alpha-1}\,\mathrm{e}^{-(\lambda-it)x}\,dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^{+\infty}\left(\frac{y}{\lambda-it}\right)^{\alpha-1}\mathrm{e}^{-y}\cdot\frac{dy}{\lambda-it}$$

$$=\frac{\lambda^\alpha}{\Gamma(\alpha)}\cdot\frac{1}{(\lambda-it)^\alpha}\int_0^{+\infty}y^{\alpha-1}\,\mathrm{e}^{-y}\,dy=\frac{\lambda^\alpha}{\Gamma(\alpha)}\cdot\frac{1}{(\lambda-it)^\alpha}\Gamma(\alpha)=\left(\frac{\lambda}{\lambda-it}\right)^\alpha;$$

（7）标准正态分布 $N(0,1)$：密度函数 $p(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$，　$-\infty<x<+\infty$，特征函数为

$$\varphi(t)=\int_{-\infty}^{+\infty}\mathrm{e}^{itx}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,dx=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2-2itx}{2}}\,dx=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-it)^2}{2}}\cdot\mathrm{e}^{-\frac{t^2}{2}}\,dx=\mathrm{e}^{-\frac{t^2}{2}}。$$

4.2.2　　特征函数的性质

特征函数具有以下性质：

（1）$|\varphi(t)|\le\varphi(0)=1$；

（2）$\varphi(-t)=\overline{\varphi(t)}$，其中 $\overline{\varphi(t)}$ 表示 $\varphi(t)$ 的共轭复数；

（3）设 a,b 为常数，则 $\varphi_{aX+b}(t)=\mathrm{e}^{ibt}\,\varphi_X(at)$；

（4）若 X 与 Y 相互独立，则 $\varphi_{X+Y}(t)=\varphi_X(t)\cdot\varphi_Y(t)$；

（5）若 X 的 l 阶矩存在，则 X 的特征函数 l 阶可导，且对 $1\le k\le l$，有 $\varphi^{(k)}(0)=i^kE(X^k)$；

（6）特征函数 $\varphi(t)$ 半正定，即对任意一组互不相同的实数 $t_1,t_2,\cdots,t_n$，矩阵 $(\varphi(t_i-t_j))_{n\times n}$ 半正定；

（7）特征函数 $\varphi(t)$ 一致连续，即对任意正数 $\varepsilon>0$，存在 $\delta>0$，当 $|h|<\delta$ 时，有 $|\varphi(t+h)-\varphi(t)|<\varepsilon$；

（8）随机变量的分布函数由其特征函数唯一确定。

证明：（1）因

$$|\varphi(t)|=|E(\mathrm{e}^{itX})|\le E(|\mathrm{e}^{itX}|)=E(1)=1,$$

且 $\varphi(0)=E(\mathrm{e}^0)=1$，故 $|\varphi(t)|\le\varphi(0)=1$；

（2）由欧拉公式可得

$$\varphi(-t)=E(\mathrm{e}^{-itX})=E(\cos tX-i\sin tX)=\overline{E(\cos tX)+iE(\sin tX)}=\overline{E(\mathrm{e}^{itX})}=\overline{\varphi(t)}\text{；}$$

（3）因 a,b 为常数，则

$$\varphi_{aX+b}(t)=E(\mathrm{e}^{it(aX+b)})=E(\mathrm{e}^{itaX}\cdot\mathrm{e}^{itb})=\mathrm{e}^{ibt}\,E(\mathrm{e}^{iatX})=\mathrm{e}^{ibt}\,\varphi_X(at)\text{；}$$

（4）因 X 与 Y 相互独立，则

$$\varphi_{X+Y}(t)=E(\mathrm{e}^{it(X+Y)})=E(\mathrm{e}^{itX}\cdot\mathrm{e}^{itY})=E(\mathrm{e}^{itX})E(\mathrm{e}^{itY})=\varphi_X(t)\varphi_Y(t)\text{；}$$

注：此性质可推广到多个随机变量，设 $X_1,X_2,\cdots,X_n$ 相互独立且 $X=X_1+X_2+\cdots+X_n$，则

$$\varphi_X(t)=\varphi_{X_1}(t)\varphi_{X_2}(t)\cdots\varphi_{X_n}(t)\text{。}$$

（5）因 X 的 l 阶矩存在，对 $1\le k\le l$，有 X 的 k 阶矩存在。又因

$$\frac{d^k}{dt^k}(\mathrm{e}^{itX})=(iX)^k\cdot\mathrm{e}^{itX}\text{，}\quad |(iX)^k\cdot\mathrm{e}^{itX}|=|i^k|\cdot|X^k|\cdot|\mathrm{e}^{itX}|=|X^k|\text{，}$$

可得 $E[(iX)^k\cdot\mathrm{e}^{itX}]$ 对 $-\infty<t<+\infty$ 都存在。将期望看作积分，即该含参变量积分一致可积，可知求导与积分可交换次序，可得

$$\varphi^{(k)}(t)=E\left[\frac{d^k}{dt^k}(\mathrm{e}^{itX})\right]=E[(iX)^k\cdot\mathrm{e}^{itX}]=i^kE(X^k\,\mathrm{e}^{itX})\text{，}$$

故 $\varphi^{(k)}(0)=i^kE(X^k)$；

（6）对任意一组互不相同的实数 $t_1,t_2,\cdots,t_n$，由性质（2）知 $\varphi(t_i-t_j)=\overline{\varphi(t_j-t_i)}$，则

$$[(\varphi(t_i-t_j))_{n\times n}]^T=\overline{(\varphi(t_i-t_j))_{n\times n}}\text{，}$$

即矩阵 $(\varphi(t_i-t_j))_{n\times n}$ 为 Hermite 矩阵（若为实矩阵，就是对称阵）。对任意一组复数 $c_1,c_2,\cdots,c_n$，记复向量 $\alpha=(c_1,c_2,\cdots,c_n)^T$，则

$$\alpha^T(\varphi(t_i-t_j))_{n\times n}\overline{\alpha}=\sum_{i=1}^n\sum_{j=1}^n c_i\overline{c_j}\varphi(t_i-t_j)=\sum_{i=1}^n\sum_{j=1}^n E(c_i\overline{c_j}\,\mathrm{e}^{it_iX}\,\mathrm{e}^{-it_jX})=E\left(\sum_{i=1}^n\sum_{j=1}^n c_i\,\mathrm{e}^{it_iX}\,\overline{c_j\,\mathrm{e}^{it_jX}}\right)$$

$$=E\left(\sum_{i=1}^n c_i\,\mathrm{e}^{it_iX}\cdot\overline{\sum_{j=1}^n c_j\,\mathrm{e}^{it_jX}}\right)=E\left(\left|\sum_{i=1}^n c_i\,\mathrm{e}^{it_iX}\right|^2\right)\ge 0\text{，}$$

故特征函数 $\varphi(t)$ 半正定；

（7）因

$$|\varphi(t+h)-\varphi(t)|=|E(\mathrm{e}^{i(t+h)X}-\mathrm{e}^{itX})|\le E(|\mathrm{e}^{itX}(\mathrm{e}^{ihX}-1)|)=E(|\mathrm{e}^{ihX}-1|)\text{，}$$

$$|\mathrm{e}^{ihX}-1|=|\cos(hX)-1+i\sin(hX)|=\sqrt{[\cos(hX)-1]^2+[\sin(hX)]^2}=\sqrt{2-2\cos(hX)}$$

$$=\left|2\sin\frac{hX}{2}\right|\text{，}$$

因 $\lim\limits_{c\to+\infty}P\{|X|>c\}=0$，对任意正数 $\varepsilon>0$，存在 $M>0$，使得 $P\{|X|>M\}<\dfrac{\varepsilon}{3}$，则

$$|\varphi(t+h)-\varphi(t)|\le E\left(\left|2\sin\frac{hX}{2}\right|\right)=\int_{-M}^{M}\left|2\sin\frac{hx}{2}\right|dF(x)+\int\limits_{|x|>M}\left|2\sin\frac{hx}{2}\right|dF(x)$$

$$\le\int_{-M}^{M}|hx|\,dF(x)+\int\limits_{|x|>M}2dF(x)\le|h|\cdot M\int_{-M}^{M}dF(x)+2P\{|x|>M\}<|h|M+\frac{2\varepsilon}{3},$$

取 $\delta=\dfrac{\varepsilon}{3M}$，当 $|h|<\delta$ 时，有 $|\varphi(t+h)-\varphi(t)|<\varepsilon$，故特征函数 $\varphi(t)$ 一致连续；

（8）首先证明逆转公式：设 $F(x)$ 是 X 的分布函数，对于 $F(x)$ 的任意两个连续点 $x_1<x_2$，满足

$$F(x_2)-F(x_1)=\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\mathrm{e}^{-itx_1}-\mathrm{e}^{-itx_2}}{it}\varphi(t)dt。$$

因

$$\int_{-T}^{T}\frac{\mathrm{e}^{-itx_1}-\mathrm{e}^{-itx_2}}{it}\varphi(t)dt=\int_{-T}^{T}\frac{\mathrm{e}^{-itx_1}-\mathrm{e}^{-itx_2}}{it}E(\mathrm{e}^{itX})dt=\int_{-T}^{T}E\left[\frac{\mathrm{e}^{it(X-x_1)}-\mathrm{e}^{it(X-x_2)}}{it}\right]dt$$

且

$$\left|\frac{\mathrm{e}^{it(X-x_1)}-\mathrm{e}^{it(X-x_2)}}{it}\right|=\left|\mathrm{e}^{it(X-x_2)}\frac{\mathrm{e}^{it(x_2-x_1)}-1}{it}\right|=\left|\frac{\mathrm{e}^{it(x_2-x_1)}-1}{it}\right|=\left|\frac{2}{t}\sin\frac{t(x_2-x_1)}{2}\right|\le|x_2-x_1|,$$

即 $\dfrac{\mathrm{e}^{it(X-x_1)}-\mathrm{e}^{it(X-x_2)}}{it}$ 有界，根据实变函数的 Fubini 定理知期望与积分可交换次序，有

$$\int_{-T}^{T}\frac{\mathrm{e}^{-itx_1}-\mathrm{e}^{-itx_2}}{it}\varphi(t)dt=E\left[\int_{-T}^{T}\frac{\mathrm{e}^{it(X-x_1)}-\mathrm{e}^{it(X-x_2)}}{it}dt\right]$$

$$=E\left[\int_{-T}^{T}\frac{\cos t(X-x_1)+i\sin t(X-x_1)-\cos t(X-x_2)-i\sin t(X-x_2)}{it}dt\right]$$

$$=E\left[\int_{-T}^{T}\frac{\sin t(X-x_1)-\sin t(X-x_2)}{t}-i\frac{\cos t(X-x_1)-\cos t(X-x_2)}{t}dt\right]$$

$$=E\left[2\int_{0}^{T}\frac{\sin t(X-x_1)-\sin t(X-x_2)}{t}dt\right],$$

根据狄利克雷积分 $\int_0^{+\infty}\dfrac{\sin x}{x}dx=\dfrac{\pi}{2}$，并且 $\dfrac{\sin x}{x}$ 是偶函数，计算 $\int_0^{+\infty}\dfrac{\sin at}{t}dt$，

当 $a>0$ 时，令 $x=at$，有 $\int_0^{+\infty}\dfrac{\sin at}{t}dt=\int_0^{+\infty}\dfrac{\sin at}{at}\cdot adt=\int_0^{+\infty}\dfrac{\sin x}{x}dx=\dfrac{\pi}{2}$；

当 $a<0$ 时，令 $x=at$，有 $\int_0^{+\infty}\dfrac{\sin at}{t}dt=\int_0^{+\infty}\dfrac{\sin at}{at}\cdot adt=\int_0^{-\infty}\dfrac{\sin x}{x}dx=-\int_{-\infty}^{0}\dfrac{\sin x}{x}dx=-\dfrac{\pi}{2}$；

当 $a=0$ 时，$\int_0^{+\infty}\dfrac{\sin at}{t}dt=0$，

则 $\int_0^{+\infty}\dfrac{\sin at}{t}dt=\dfrac{\pi}{2}\mathrm{sgn}(a)$，可得

$$\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx_1}-e^{-itx_2}}{it}\varphi(t)dt=\lim_{T\to+\infty}\frac{1}{2\pi}E\left[2\int_0^T\frac{\sin t(X-x_1)-\sin t(X-x_2)}{t}dt\right]$$

$$=E\left[\frac{1}{\pi}\int_0^{+\infty}\frac{\sin t(X-x_1)-\sin t(X-x_2)}{t}dt\right]=E\left[\frac{1}{2}\mathrm{sgn}(X-x_1)-\frac{1}{2}\mathrm{sgn}(X-x_2)\right]$$

$$=\int_{-\infty}^{+\infty}\frac{1}{2}[\mathrm{sgn}(x-x_1)-\mathrm{sgn}(x-x_2)]dF(x)=\int_{x_1}^{x_2}dF(x)=F(x_2)-F(x_1),$$

再根据逆转公式得分布函数的表示式。设 x 是 $F(x)$ 的任一连续点，由实变函数可知，单调不减函数 $F(x)$ 最多只有可列个不连续点，取 y 是 $F(x)$ 的连续点，并令 $y\to-\infty$，有

$$F(x)=\lim_{y\to-\infty}[F(x)-F(y)]=\lim_{y\to-\infty}\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ity}-e^{-itx}}{it}\varphi(t)dt,$$

故随机变量的分布函数由其特征函数唯一确定。

推论 若 X 是连续随机变量，$\varphi(t)$ 是其特征函数，如果 $\int_{-\infty}^{+\infty}|\varphi(t)|dt<+\infty$，则密度函数为

$$p(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dx。$$

证明：因

$$F(x)=\lim_{y\to-\infty}\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ity}-e^{-itx}}{it}\varphi(t)dt,$$

记 $F_T(x,y)=\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ity}-e^{-itx}}{it}\varphi(t)dt$。因 $\int_{-\infty}^{+\infty}|\varphi(t)|dt<+\infty$，且

$$\left|\frac{e^{-ity}-e^{-itx}}{it}\right|=\left|\frac{2}{t}\sin\frac{t(x-y)}{2}\right|\le|x-y|,$$

即 $\frac{e^{-ity}-e^{-itx}}{it}$ 一致有界，则含参变量积分 $F_T(x,y)$ 求导与积分可交换次序，可得

$$\frac{\partial}{\partial x}F_T(x,y)=\frac{1}{2\pi}\int_{-T}^{T}\frac{\partial}{\partial x}\left(\frac{e^{-ity}-e^{-itx}}{it}\right)\varphi(t)dt=\frac{1}{2\pi}\int_{-T}^{T}e^{-itx}\varphi(t)dt,$$

故

$$p(x)=F'(x)=\lim_{y\to-\infty}\lim_{T\to+\infty}\frac{\partial}{\partial x}F_T(x,y)=\lim_{T\to+\infty}\frac{1}{2\pi}\int_{-T}^{T}e^{-itx}\varphi(t)dt=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dx。$$

注：$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}p(x)dx$ 为 $p(x)$ 的傅立叶变换，而 $p(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$ 为 $\varphi(t)$ 的傅立叶逆变换。

一般正态分布 $N(\mu,\sigma^2)$ 可由标准正态分布 $N(0,1)$ 通过线性变换得到，因此根据性质（3）由标准正态分布 $N(0,1)$ 的特征函数得到一般正态分布 $N(\mu,\sigma^2)$ 的特征函数。

设 $X\sim N(\mu,\sigma^2)$，有 $Y=\frac{X-\mu}{\sigma}\sim N(0,1)$，$Y$ 的特征函数为 $\varphi_Y(t)=e^{-\frac{t^2}{2}}$，且 $X=\sigma Y+\mu$，故一般正态分布 $N(\mu,\sigma^2)$ 的特征函数为

$$\varphi_X(t)=\varphi_{\sigma Y+\mu}(t)=e^{i\mu t}\varphi_Y(\sigma t)=e^{i\mu t}e^{-\frac{\sigma^2t^2}{2}}=e^{i\mu t-\frac{\sigma^2t^2}{2}}。$$

例　利用特征函数性质证明正态分布的可加性：若 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且 X 与 Y 相互独立，则 $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$。

证明：因 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且 X 与 Y 相互独立，可得 X 与 Y 的特征函数分别为

$$\varphi_X(t)=\mathrm{e}^{i\mu_1 t-\frac{\sigma_1^2t^2}{2}}，\quad \varphi_Y(t)=\mathrm{e}^{i\mu_2 t-\frac{\sigma_2^2t^2}{2}}，$$

则 $X+Y$ 的特征函数为

$$\varphi_{X+Y}(t)=\varphi_X(t)\varphi_Y(t)=\mathrm{e}^{i(\mu_1+\mu_2)t-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}，$$

这正是正态分布 $N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$ 的特征函数，故根据特征函数的唯一性可知

$$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)。$$

注：此结论可推广到多个随机变量，设 $X_1,X_2,\cdots,X_n$ 相互独立，且 $X_i\sim N(\mu_i,\sigma_i^2),\quad i=1,2,\cdots,n$，则有

$$\sum_{i=1}^n X_i\sim N\left(\sum_{i=1}^n\mu_i,\sum_{i=1}^n\sigma_i^2\right)。$$

例　证明泊松分布的正态逼近：设 $X\sim P(\lambda)$，记 X 的标准化随机变量为 $Y_\lambda=\dfrac{X-\lambda}{\sqrt{\lambda}}$，则 Y_λ 按分布收敛于标准正态分布 $N(0,1)$。

证明：因 $X\sim P(\lambda)$，有 X 的特征函数为 $\varphi(t)=\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}$，而

$$Y_\lambda=\frac{X-\lambda}{\sqrt{\lambda}}=\frac{1}{\sqrt{\lambda}}X-\sqrt{\lambda}，$$

其特征函数为

$$\varphi_{Y_\lambda}(t)=\mathrm{e}^{-i\sqrt{\lambda}t}\,\varphi\left(\frac{t}{\sqrt{\lambda}}\right)=\mathrm{e}^{-i\sqrt{\lambda}t}\cdot\mathrm{e}^{\lambda(\mathrm{e}^{\frac{it}{\sqrt{\lambda}}}-1)}=\mathrm{e}^{\lambda(\mathrm{e}^{\frac{it}{\sqrt{\lambda}}}-1-\frac{it}{\sqrt{\lambda}})}。$$

因 $\mathrm{e}^x=1+x+\dfrac{x^2}{2}+o(x^2)$，即

$$\mathrm{e}^{\frac{iv}{\sqrt{\lambda}}}=1+\frac{iv}{\sqrt{\lambda}}+\frac{-v^2}{2\lambda}+o\left(\frac{1}{\lambda}\right),$$

$$\mathrm{e}^{\frac{iv}{\sqrt{\lambda}}}-1-\frac{iv}{\sqrt{\lambda}}=-\frac{v^2}{2\lambda}+o\left(\frac{1}{\lambda}\right),$$

$$\lim_{\lambda\to+\infty}\varphi_{Y_\lambda}(t)=\lim_{\lambda\to+\infty}\mathrm{e}^{\lambda(\mathrm{e}^{\frac{it}{\sqrt{\lambda}}}-1-\frac{it}{\sqrt{\lambda}})}=\lim_{\lambda\to+\infty}\mathrm{e}^{\lambda\left[-\frac{v^2}{2\lambda}+o\left(\frac{1}{\lambda}\right)\right]}=\lim_{\lambda\to+\infty}\mathrm{e}^{-\frac{v^2}{2}+o(1)}=\mathrm{e}^{-\frac{v^2}{2}},$$

这正是标准正态分布的特征函数，则根据特征函数的唯一性可知，$\lim\limits_{\lambda\to+\infty}F_{Y_\lambda}(y)=\Phi(y)$，故 Y_λ 按分布收敛于标准正态分布。

§4.3　大数定律

在大量的重复试验中，频率稳定于概率，平均值稳定于数学期望。直观上，这里的“稳定”类似于数学分析中的极限，但又有差别，本节将在数学上给予严格表述，并证明上述结论。

4.3.1　伯努利大数定律（Bernoulli Law of Large Numbers）

设事件A在每次试验中发生的概率为p，在n重伯努利试验中的发生次数为S_n，即频率为$\dfrac{S_n}{n}$。直观上，随n的无限增大，频率$\dfrac{S_n}{n}$将以概率p为“极限”。

定理　（伯努利大数定律）设S_n为n重伯努利试验中事件A的发生次数，p为每次试验中A发生的概率，则对任意的$\varepsilon > 0$，有

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

证明：因S_n服从二项分布$b(n, p)$，有$E(S_n) = np$，$\mathrm{Var}(S_n) = np(1-p)$，则

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{1}{n}\cdot np = p，$$

$$\mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathrm{Var}(S_n) = \frac{1}{n^2}\cdot np(1-p) = \frac{p(1-p)}{n}，$$

由切比雪夫不等式，可得

$$1 \ge P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} \ge 1 - \frac{\mathrm{Var}\left(\dfrac{S_n}{n}\right)}{\varepsilon^2} = 1 - \frac{p(1-p)}{n\varepsilon^2}，$$

因$\lim\limits_{n\to+\infty}\left[1 - \dfrac{p(1-p)}{n\varepsilon^2}\right] = 1$，故由夹逼准则可得

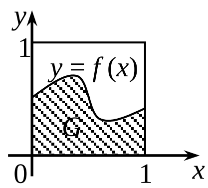
$$\lim_{n\to+\infty} P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} = 1。$$

伯努利大数定律表明频率依概率收敛于概率p，这就是在大量重复试验中频率稳定于概率p的含义。

注：此定理又称为弱大数定律，此外还有强大数定律，即频率几乎处处收敛收敛于概率p。

现在常用的统计模拟方法，蒙特卡洛方法（Monte Carlo method）中的随机投点法就是依据伯努利大数定律，以计算机为模拟工具，进行大量重复试验，用频率作为概率的估计值进行计算。

如计算定积分$\int_0^1 f(x)dx$，且$0 \le f(x) \le 1$。设随机变量(X, Y)服从区域$D = \{(x, y) \mid 0 \le x \le 1, 0 \le y \le 1\}$上的均匀分布，且区域$G = \{(x, y) \mid 0 \le x \le 1, 0 \le y \le f(x)\}$，则



$$P\{(X, Y) \in G\} = \frac{区域\,G\,的面积}{区域\,D\,的面积} = \int_0^1 f(x)dx。$$

这样，首先用计算机产生在$(0,1)$区间上均匀分布的n对随机数$(x_i, y_i),\quad i = 1, 2, \cdots, n$，记录其中满足

不等式 $y_i \le f(x_i)$ 的数据对的个数 S_n，用频率 $\frac{S_n}{n}$ 作为概率 $P\{(X,Y)\in G\}$ 的估计值，这样就得到 $\int_0^1 f(x)dx$ 的估计值。

对于一般的定积分 $\int_a^b g(x)dx$，可通过变量替换 $y=\frac{x-a}{b-a}$ 化为关于 y 的函数在 0 与 1 之间的积分，即

$$\int_a^b g(x)dx=(b-a)\int_0^1 g[a+(b-a)y]dy,$$

进一步，若 $m\le g(x)\le M$，通过函数变换 $f(y)=\frac{g[a+(b-a)y]-m}{M-m}$，使得 $0\le f(y)\le 1$，可得

$$\int_a^b g(x)dx=(b-a)(M-m)\int_0^1 f(y)dy+m(b-a),$$

这样，用蒙特卡洛方法计算出 $\int_0^1 f(y)dy$，进而就得到 $\int_a^b g(x)dx$ 的值。

附：用蒙特卡洛随机投点法计算定积分 $\int_0^1 \frac{1}{1+x^2}dy$，MATLAB 程序：

```
n=input('number of tests=');k=0;
for i=1:n
    x=rand;y=rand;
    if y<=1/(1+x^2);
        k=k+1;
    end
end
J=k/n
```

4.3.2 常用的几个大数定律

一．大数定律的一般形式

直观上，在大量重复试验中，平均值稳定于数学期望，也就是对于一个随机变量序列 $\{X_n\}$，其平均值 $\frac{1}{n}\sum_{i=1}^{n}X_i$ 将依概率收敛于其数学期望 $\frac{1}{n}\sum_{i=1}^{n}E(X_i)$。

定义 设随机变量序列 $\{X_n\}$ 满足

$$\lim_{n\to+\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1,$$

则称 $\{X_n\}$ 服从大数定律。

伯努利大数定律中，设 S_n 表示 n 重伯努利试验中事件 A 的发生次数，p 表示每次试验中 A 发生的概率，令

$$X_i=\begin{cases}1, & \text{第 } i \text{ 次试验中事件 } A \text{ 发生;}\\ 0, & \text{第 } i \text{ 次试验中事件 } A \text{ 没有发生.}\end{cases}\quad i=0,1,2,\cdots,$$

则有 $S_n=\sum_{i=1}^{n}X_i$，且频率 $\frac{S_n}{n}=\frac{1}{n}\sum_{i=1}^{n}X_i$，概率 $p=\frac{1}{n}\sum_{i=1}^{n}E(X_i)$，可见伯努利大数定律也就是说明 n 重伯努

利试验中事件 A 在每次试验中的发生情况 $\{X_n\}$ 服从大数定律。

此外，还有一些其它的大数定律，分别说明对于随机变量序列 $\{X_n\}$，在满足某些不同的条件下，将服从大数定律。

二．切比雪夫大数定律（Chebyshev Law of Large Numbers）

定理 （切比雪夫大数定律）设 $\{X_n\}$ 是两两不相关的随机变量序列，每个 X_n 的方差都存在且有共同的上界，即 $\mathrm{Var}(X_n) \le c,\quad n = 1, 2, \cdots$，则 $\{X_n\}$ 服从大数定律，即对任意的 $\varepsilon > 0$，有

$$\lim_{n\to+\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} = 1 \text{。}$$

证明：因 $\{X_n\}$ 两两不相关，有

$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)\text{，}\quad \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) \le \frac{1}{n^2}\cdot nc = \frac{c}{n}\text{，}$$

由切比雪夫不等式，可得

$$1 \ge P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} \ge 1 - \frac{1}{\varepsilon^2}\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = 1 - \frac{c}{n\varepsilon^2}\text{，}$$

因 $\lim\limits_{n\to+\infty}\left(1 - \dfrac{c}{n\varepsilon^2}\right) = 1$，故由夹逼准则可得

$$\lim_{n\to+\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} = 1 \text{。}$$

伯努利大数定律的条件要求 $\{X_n\}$ 独立，且都服从相同的两点分布，而切比雪夫大数定律的条件只要求 $\{X_n\}$ 两两不相关，不必同分布，只需要方差存在且有共同的上界，条件比伯努利大数定律放宽。

三．马尔可夫大数定律（Markov Law of Large Numbers）

在切比雪夫大数定律的证明过程中，可以看出只要随机变量序列 $\{X_n\}$ 满足 $\lim\limits_{n\to+\infty}\dfrac{1}{n^2}\sum\limits_{i=1}^{n}\mathrm{Var}(X_i) = 0$，就能证明 $\{X_n\}$ 满足大数定律。

定理 （马尔可夫大数定律）设随机变量序列 $\{X_n\}$ 满足

$$\lim_{n\to+\infty}\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = 0\text{，}$$

则 $\{X_n\}$ 服从大数定律，即对任意的 $\varepsilon > 0$，有

$$\lim_{n\to+\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon\right\} = 1。$$

证明： 由切比雪夫不等式，就可证得。

马尔可夫大数定律的条件更加放宽，去掉了$\{X_n\}$两两独立或不相关的条件，只需要它们的方差满足一定的有界条件即可。

四．辛钦大数定律（Wiener-khinchin Law of Large Numbers）

定理 （辛钦大数定律）设$\{X_n\}$是独立同分布的随机变量序列，且数学期望$E(X_n)=\mu$存在，则$\{X_n\}$服从大数定律，即对任意的$\varepsilon>0$，有

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

证明： 因$\{X_n\}$独立同分布，设X_n的特征函数为$\varphi(t)$，记$Y_n=\frac{1}{n}\sum_{k=1}^{n} X_k$，有$Y_n$的特征函数为

$$\varphi_{Y_n}(t) = \left[\varphi\left(\frac{t}{n}\right)\right]^n,$$

因$\varphi(0)=1$，$\varphi'(0)=iE(X_n)=i\mu$，有$\varphi(t)=\varphi(0)+\varphi'(0)t+o(t)=1+i\mu t+o(t)$，即

$$\varphi\left(\frac{t}{n}\right) = 1 + \frac{i\mu t}{n} + o\left(\frac{1}{n}\right),$$

则

$$\lim_{n\to+\infty} \varphi_{Y_n}(t) = \lim_{n\to+\infty}\left[1+\frac{i\mu t}{n}+o\left(\frac{1}{n}\right)\right]^n = \lim_{n\to+\infty}\left[1+\frac{i\mu t}{n}\right]^n = \mathrm{e}^{i\mu t},$$

这正是在μ点处单点分布的特征函数，故Y_n按分布收敛于μ，即$Y_n=\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{L} \mu$，可得$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$。

蒙特卡洛方法中的平均值法就是依据辛钦大数定律，以计算机为模拟工具，进行大量重复试验，用平均值作为数学期望的估计值进行计算。

如计算定积分$\int_0^1 f(x)dx$。设随机变量X服从均匀分布$U(0,1)$，有$E[f(X)]=\int_0^1 f(x)dx$。

首先用计算机产生在$(0,1)$区间上均匀分布的n个随机数x_i，$i=1,2,\cdots,n$，再用平均值$\frac{1}{n}\sum_{i=1}^{n} f(X_i)$作为数学期望$E[f(X)]$的估计值，即得到$\int_0^1 f(x)dx$的估计值。

附： 用蒙特卡洛平均值法计算定积分$\int_0^1 \frac{1}{1+x^2}dy$，MATLAB 程序：

```
n=input('number of tests=');
x=rand(n,1); y=1./(1+x.^2);
J=mean(y)
```

§4.4　中心极限定理

4.4.1　　独立随机变量和

现实生活中，许多随机现象都是众多独立随机因素共同影响的结果，中心极限定理研究独立随机变量和的分布，证明了在一定条件下，独立随机变量和的分布逼近一个正态分布。

设$\{X_n\}$是独立随机变量序列，记$Y_n=\sum_{k=1}^{n}X_k$，并标准化，可得$Y_n^*=\dfrac{Y_n-E(Y_n)}{\sqrt{\mathrm{Var}(Y_n)}}$，下面将证明当$\{X_n\}$满足某些条件时，$Y_n^*$按分布收敛于标准正态分布$N(0,1)$。

4.4.2　　独立同分布下的中心极限定理（Central Limit Theorem）

定理　（林德贝格-列维中心极限定理）设随机变量序列$\{X_n\}$独立同分布，且数学期望$E(X_i)=\mu$，方差$\mathrm{Var}(X_i)=\sigma^2>0$，记$Y_n^*=\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}$，则$Y_n^*$按分布收敛于标准正态分布$N(0,1)$，即对任意实数$y$，有

$$\lim_{n\to+\infty}P\{Y_n^*\le y\}=\Phi(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}\,dt。$$

证明：设$X_i-\mu$的特征函数为$\varphi(t)$，有$Y_n^*=\dfrac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i-\mu)$的特征函数为

$$\varphi_{Y_n^*}(t)=\left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n。$$

因$E(X_i-\mu)=0$，$\mathrm{Var}(X_i-\mu)=\sigma^2>0$，有

$$\varphi(0)=1，\ \varphi'(0)=iE(X_i-\mu)=0，\ \varphi''(0)=i^2E(X_i-\mu)^2=-\sigma^2，$$

可得

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

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)=1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)，$$

则

$$\lim_{n\to+\infty}\varphi_{Y_n^*}(t)=\lim_{n\to+\infty}\left[1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)\right]^n=\lim_{n\to+\infty}\left[1-\frac{t^2}{2n}\right]^n=\mathrm{e}^{-\frac{t^2}{2}}，$$

这正是标准正态分布的特征函数，故Y_n^*按分布收敛于标准正态分布$N(0,1)$，$\lim\limits_{n\to+\infty}F_{Y_n^*}(y)=\Phi(y)$，即

$$\lim_{n\to+\infty}P\{Y_n^*\le y\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}\,dt。$$

林德贝格-列维中心极限定理说明独立同分布随机变量序列，无论服从什么分布，只要其数学期望存在，方差大于 0，其随机变量和的分布都逼近一个正态分布。

4.4.3　　二项分布的正态近似

将林德贝格-列维中心极限定理用于二项分布（独立同分布的两点分布之和），即得

定理　（棣莫弗-拉普拉斯中心极限定理）设 n 重伯努利试验中事件 A 的发生次数为 S_n，每次试验中 A 发生的概率为 $p\quad(0<p<1)$，记 $Y_n^*=\dfrac{S_n-np}{\sqrt{np(1-p)}}$，则 Y_n^* 按分布收敛于标准正态分布 $N(0,1)$，即对任意实数 y，有

$$\lim_{n\to+\infty}P\{Y_n^*\le y\}=\Phi(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}\,dt\text{。}$$

棣莫弗-拉普拉斯定理是历史上第一个中心极限定理。

4.4.4　　独立不同分布下的中心极限定理

现实生活中，影响一个随机变量的众多随机因素并不都是同分布的。直观上，只要每一个因素的影响都很小，其标准化随机变量的极限分布将服从标准正态分布。

设 $\{X_n\}$ 是独立随机变量序列，且 $E(X_i)=\mu_i$，$\mathrm{Var}(X_i)=\sigma_i^2$，$i=1,2,\cdots$，记

$$Y_n^*=\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}}=\frac{1}{\sqrt{\sum_{i=1}^{n}\sigma_i^2}}\sum_{i=1}^{n}(X_i-\mu_i)\text{，}$$

并记 $B_n=\sqrt{\sum_{i=1}^{n}\sigma_i^2}$，即 $Y_n^*=\dfrac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)$。

为了使得 Y_n^* 中每一项 $\dfrac{X_i-\mu_i}{B_n}$ 影响都很小，即对任意的 $\tau>0$，要求概率 $P\{|X_i-\mu_i|>\tau B_n\}$ 很小，而

$$P\{|X_i-\mu_i|>\tau B_n\}=\int_{|x-\mu_i|>\tau B_n}p_i(x)dx\le\frac{1}{\tau^2B_n^2}\int_{|x-\mu_i|>\tau B_n}(x-\mu_i)^2p_i(x)dx\text{，}$$

这样，令

$$\lim_{n\to+\infty}\frac{1}{\tau^2B_n^2}\sum_{i=1}^{n}\int_{|x-\mu_i|>\tau B_n}(x-\mu_i)^2p_i(x)dx=0\text{，}$$

就可保证 Y_n^* 中每一项 $\dfrac{X_i-\mu_i}{B_n}$ 的影响都很小，这称之为林德贝格条件。不加证明地给出以下定理。

定理　（林德贝格中心极限定理）独立随机变量序列 $\{X_n\}$，$E(X_i)=\mu_i$，$\mathrm{Var}(X_i)=\sigma_i^2$，$i=1,2,\cdots$，记 $B_n=\sqrt{\sum_{i=1}^{n}\sigma_i^2}$，$Y_n^*=\dfrac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)$。如果对任意的 $\tau>0$，$\{X_n\}$ 满足林德贝格条件

$$\lim_{n\to+\infty}\frac{1}{\tau^2 B_n^2}\sum_{i=1}^{n}\int_{|x-\mu_i|>\tau B_n}(x-\mu_i)^2 p_i(x)dx=0,$$

则 Y_n^* 按分布收敛于标准正态分布 $N(0,1)$，即对任意实数 y，有

$$\lim_{n\to+\infty}P\{Y_n^*\le y\}=\lim_{n\to+\infty}P\left\{\frac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)\le y\right\}=\Phi(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}dt。$$

林德贝格条件要求比较宽，但较难验证，而下面的李雅普诺夫条件虽然比林德贝格条件要求稍严格一些，却比较容易验证。不加证明地给出以下定理。

定理 （李雅普诺夫中心极限定理）独立随机变量序列 $\{X_n\}$，$E(X_i)=\mu_i$，$\mathrm{Var}(X_i)=\sigma_i^2$，$i=1,2,\cdots$，

记 $B_n=\sqrt{\sum_{i=1}^{n}\sigma_i^2}$，$Y_n^*=\frac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)$。如果存在 $\delta>0$，$\{X_n\}$满足李雅普诺夫条件

$$\lim_{n\to+\infty}\frac{1}{B_n^{2+\delta}}\sum_{i=1}^{n}E(|X_i-\mu_i|^{2+\delta})=0,$$

则 Y_n^* 按分布收敛于标准正态分布 $N(0,1)$，即对任意实数 y，有

$$\lim_{n\to+\infty}P\{Y_n^*\le y\}=\lim_{n\to+\infty}P\left\{\frac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)\le y\right\}=\Phi(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\mathrm{e}^{-\frac{t^2}{2}}dt。$$

4.4.5　中心极限定理应用举例

中心极限定理基本思想：如果一个随机变量受众多随机因素的影响，并且每一个因素的影响都很小，则其标准化随机变量将服从标准正态分布。

一般当 n 很大时，随机变量之和、二项分布等经标准化后等都可用标准正态分布 $N(0,1)$ 近似。

对于取值整数的离散随机变量，如二项分布、泊松分布，如果用中心极限定理作正态分布近似计算，可将在每个整数点取值的概率看作分散分布在该点 ± 0.5 区间之内，求概率时最好是将取值范围的整数通过加或减 0.5 进行修正。如概率 $P\{k_1\le X\le k_2\}$ 修正为 $P\{k_1-0.5<X\le k_2+0.5\}$，而概率 $P\{k_1<X<k_2\}$ 修正为 $P\{k_1+0.5<X\le k_2-0.5\}$。

例　从一大批次品率为 10%的产品中，随机抽取 400 件，求次品数超过 50 件的概率，次品数在 36 到 44 件之间的概率以及次品数恰为 40 件的概率。

解：设 X 表示取得的次品数，有 $X\sim b(400,0.1)$。因 $n=400$ 很大，且数学期望 $E(X)=np=40$，方差 $\mathrm{Var}(X)=np(1-p)=36$，有 $\frac{X-40}{6}\sim N(0,1)$，所求概率为

$$P\{X>50\}=P\{X>50.5\}\approx 1-\Phi\left(\frac{50.5-40}{6}\right)=1-\Phi(1.75)=1-0.9599=0.0401;$$

$$P\{36\le X\le 44\}=P\{35.5<X<44.5\}\approx\Phi\left(\frac{44.5-40}{6}\right)-\Phi\left(\frac{35.5-40}{6}\right)=\Phi(0.75)-\Phi(-0.75)$$

$$=2\Phi(0.75)-1=2\times 0.7734-1=0.5468;$$

$$P\{X=40\}=P\{39.5<X<40.5\}\approx\Phi\left(\frac{40.5-40}{6}\right)-\Phi\left(\frac{39.5-40}{6}\right)=\Phi(0.0833)-\Phi(-0.0833)$$

$$=2\Phi(0.0833)-1=2\times0.5332-1=0.0664\text{。}$$

例　掷硬币 100 次，求正面朝上的频率超过 0.6 的概率。

解： 设 X 表示正面朝上的次数，有 $X\sim b(100,0.5)$。频率 $\frac{X}{100}>0.6$，即 $X>60$。因 $n=100$ 较大，且数学期望 $E(X)=np=50$，方差 $\mathrm{Var}(X)=np(1-p)=25$，有 $\frac{X-50}{5}\,\dot\sim\,N(0,1)$，所求概率为

$$P\{X>60\}=P\{X>60.5\}\doteq1-\Phi\left(\frac{60.5-50}{5}\right)=1-\Phi(2.1)=1-0.9821=0.0179\text{。}$$

例　一商店中某商品日销量服从参数 $\lambda=4$ 的泊松分布，每月月初进货，问月初进货后应库存多少件，才能以 99%以上的概率满足本月的需要？（每月以 30 天计，且每日销量相互独立）

解： 以 X_i 表示第 i 日的销量，$X_i\sim P(4)$ 且相互独立，有 $E(X_i)=\lambda=4$，$\mathrm{Var}(X_i)=\lambda=4$，月销量为 $\sum_{i=1}^{30}X_i$。因 $n=30$ 较大，且 $E\left(\sum_{i=1}^{30}X_i\right)=30\times4=120$，$\mathrm{Var}\left(\sum_{i=1}^{30}X_i\right)=30\times4=120$，有 $\frac{\sum_{i=1}^{30}X_i-120}{\sqrt{120}}\,\dot\sim\,N(0,1)$。

设月初应库存 x 件该商品，事件“满足本月的需要”，即 $\sum_{i=1}^{30}X_i\le x$，所求概率为

$$P\left\{\sum_{i=1}^{n}X_i\le x\right\}=P\left\{\frac{\sum_{i=1}^{n}X_i-120}{\sqrt{120}}\le\frac{x-120}{\sqrt{120}}\right\}\approx\Phi\left(\frac{x-120}{\sqrt{120}}\right)\ge0.99\text{，}$$

有 $\frac{x-120}{\sqrt{120}}\ge2.3263$，故 $x\ge145.4839$，取 $x=146$。

例　每个元件使用寿命（小时）服从指数分布 $Exp\left(\frac{1}{1000}\right)$，求：（1）一个元件使用寿命在 800 到 1200 小时之间的概率；（2）100 个元件平均寿命在 800 到 1200 小时之间的概率。

解： 以 X_i 表示第 i 个元件的使用寿命，则 X_i 的密度函数为

$$p(x)=\begin{cases}\frac{1}{1000}\mathrm{e}^{-\frac{x}{1000}}, & x>0;\\ 0, & x\le0.\end{cases}$$

且 $E(X_i)=\frac{1}{\lambda}=1000$，$\mathrm{Var}(X_i)=\frac{1}{\lambda^2}=10^6$。

（1）一个元件直接根据指数分布进行计算

$$P\{800<X_i<1200\}=\int_{800}^{1200}\frac{1}{1000}\mathrm{e}^{-\frac{x}{1000}}dx=-\mathrm{e}^{-\frac{x}{1000}}\Big|_{800}^{1200}=\mathrm{e}^{-0.8}-\mathrm{e}^{-1.2}=0.1481\text{；}$$

（2）100 个元件平均寿命 $\overline{X}=\dfrac{1}{100}\sum_{i=1}^{100}X_i$ ，因 $n=100$ 较大，且 X_i 相互独立，有

$$E\left(\sum_{i=1}^{100}X_i\right)=\sum_{i=1}^{100}E(X_i)=10^5，\quad \mathrm{Var}\left(\sum_{i=1}^{100}X_i\right)=\sum_{i=1}^{100}\mathrm{Var}(X_i)=10^8，$$

即 $\dfrac{\sum_{i=1}^{100}X_i-10^5}{10^4}=\dfrac{\overline{X}-1000}{100}\sim N(0,1)$，故

$$P\{800<\overline{X}<1200\}=P\left\{-2<\frac{\overline{X}-1000}{100}<2\right\}\doteq\Phi(2)-\Phi(-2)=2\times0.9772-1=0.9544\ 。$$

例　讨论各国足球联赛理论保级分。

背景：如英超、意甲、西甲、法甲，共 20 支足球队，每年双循环比赛。每轮比赛 10 场，全年共 38 轮。每场比赛胜得 3 分，平得 1 分，负得 0 分。全年比赛得分最低的 3 支球队降级。

假设：各球队实力相当，每场比赛胜、平、负的概率各占三分之一；每场比赛结果相互独立。

分析：设 X_i 表示某球队第 i 轮比赛的得分，其分布列为

X_i	0	1	3
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

则

$$E(X_i)=0\times\frac{1}{3}+1\times\frac{1}{3}+3\times\frac{1}{3}=\frac{4}{3}，\quad \mathrm{Var}(X_i)=E(X_i^2)-[E(X_i)]^2=\frac{10}{3}-\left(\frac{4}{3}\right)^2=\frac{14}{9}，$$

该队 38 轮总分 $\sum_{i=1}^{38}X_i$ ， $n=38$ 较大，

$$E\left(\sum_{i=1}^{38}X_i\right)=38\times\frac{4}{3}\approx50.67，\quad \mathrm{Var}\left(\sum_{i=1}^{38}X_i\right)=38\times\frac{14}{9}\approx59.11，$$

则 $\dfrac{\sum_{i=1}^{38}X_i-50.67}{\sqrt{59.11}}\sim N(0,1)$，设 x 为理论保级分，即 $\sum_{i=1}^{38}X_i\ge x$ 表示保级。又已知 20 支球队中有 3 支降级，17 支保级，则

$$P\left\{\sum_{i=1}^{38}X_i\ge x\right\}=1-\Phi\left(\frac{x-50.67}{\sqrt{59.11}}\right)=\frac{17}{20}=0.85，\quad -\frac{x-50.67}{\sqrt{59.11}}=1.0364，$$

故 $x=42.6982$ ，取 $x=43$，英超足球联赛理论保级分为 43 分。

一般地，设共有 n 支足球队，每年双循环比赛，全年共 $2n-2$ 轮，全年比赛得分最低的 m 支球队降级，则理论保级分为 $x=\dfrac{4}{3}(2n-2)-\sqrt{\dfrac{14}{9}(2n-2)}\cdot\Phi^{-1}\left(1-\dfrac{m}{n}\right)$，其中 $\Phi(\cdot)$ 表示标准正态分布的分布函数。