中心极限定理

- $\boldsymbol{L\acute{e}vy - Lindeberg}$ 中心极限定理
- $\boldsymbol{De\ Moivre - Laplace}$ 中心极限定理

Lévy-Lindeberg 中心极限定理

设$\{X_n: n=1,2,\ldots\}$为独立同分布随机变量序列. $E(X_n)=\mu$,

$D(X_n)=\sigma^2>0, n=1,2,\ldots$, 则随机变量

$$\frac{1}{\sqrt{n}\,\sigma}\left(\sum_{k=1}^{n} X_k - n\mu\right)$$

的分布函数$F_n(x)$收敛到标准正态分布函数$\Phi(x)$，即对任意实数x, 成立

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\,\sigma}\left(\sum_{k=1}^{n} X_k - n\mu\right) \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\,\sigma}\left(\sum_{k=1}^{n} X_k - n\mu\right) \le x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$$

- 对独立同分布随机变量序列$\{X_n : n = 1, 2, \dots\}$,

无论 X_n 服从什么分布，只要n充分大，随机变量

$$\frac{1}{\sqrt{n}\,\sigma}\left(\sum_{k=1}^{n} X_k - n\mu\right)$$

就近似服从标准正态分布$N\,(0, 1)$.

$$\frac{1}{\sqrt{n}\,\sigma}\left(\sum_{k=1}^{n} X_k - n\mu\right) \quad 近似服从 \quad N(0,1)$$

- 设$\{X_n: n = 1, 2, \ldots\}$为独立同分布随机变量序列, 满足

$$E(X_n) = \mu, \qquad D(X_n) = \sigma^2 > 0,$$

则随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N(n\mu, n\sigma^2)$.

$$E\left(\sum_{k=1}^{n} X_k\right) = n\mu \qquad D\left(\sum_{k=1}^{n} X_k\right) = n\sigma^2$$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N\,(n\mu, n\sigma^2)$

例1. 炮火轰击敌方防御工事100次， 每次命中的炮弹数服从同一分布，

其数学期望为2, 均方差为1.5，若各次轰击命中的炮弹数相互独立，

求100次轰击至少命中180发炮弹的概率.

解：设X表示100次轰击命中的炮弹数, 则要求 $P(X \geq 180)$.

设X_k表示第k次轰击命中的炮弹数, $X_1, X_2, \ldots, X_{100}$独立同分布，且

$$E(X_k) = 2, \quad D(X_k) = 1.5^2 \qquad k = 1, \ldots, 100$$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布 $N(n\mu, n\sigma^2)$

$$X = \sum_{k=1}^{100} X_k, \quad E(X) = 200, \quad D(X) = 15^2,$$

由Lévy-Lindeberg中心极限定理知X近似服从$N(200, 15^2)$.

$$P(X \geq 180) = P\left(\frac{X - 200}{15} \geq \frac{180 - 200}{15}\right)$$

$$= 1 - \Phi(-1.3) = \Phi(1.3) \approx 0.91$$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N\,(n\mu, n\sigma^2)$

例2. 计算机在进行数字计算时, 为简单起见，现对小数点后第一位进行四舍五入运算，误差X可以认为服从均匀分布$U[-0.5, 0.5]$. 若在一项计算中进行了100次数字计算，求平均误差落在区间$[-\frac{\sqrt{3}}{20}, \frac{\sqrt{3}}{20}]$上的概率.

解：设$X_k(k=1,\ldots,100)$表示第k次的误差, 则 $X_1, X_2, \ldots, X_{100}$独立同分布，且

$$E(X_k)=0,\quad D(X_k)=\frac{1}{12}$$

下面求 $P\left(\frac{1}{100}\sum_{k=1}^{100} X_k \in \left[-\frac{\sqrt{3}}{20}, \frac{\sqrt{3}}{20}\right]\right)$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N\left(n\mu, n\sigma^2\right)$

$$E\left(\frac{1}{100}\sum_{k=1}^{100} X_k\right) = 0 \qquad D\left(\frac{1}{100}\sum_{k=1}^{100} X_k\right) = \frac{1}{100}D(X_k) = \frac{1}{1200}$$

由Lévy-Lindeberg中心极限定理知 $\frac{1}{100}\sum_{k=1}^{100} X_k$近似服从$N\left(0, \frac{1}{1200}\right)$.

$$P\left(-\frac{\sqrt{3}}{20} \le \frac{1}{100}\sum_{k=1}^{100} X_k \le \frac{\sqrt{3}}{20}\right) = P\left(-3 \le \frac{\sqrt{3}}{5}\sum_{k=1}^{100} X_k \le 3\right) = \Phi(3) - \Phi(-3) \approx 0.9973$$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N(n\mu, n\sigma^2)$

在n重伯努利试验中，令

$$X_k = \begin{cases} 1, & \text{若事件}A\text{在第}k\text{次试验中发生} \\ 0, & \text{若事件}A\text{在第}k\text{次试验中不发生} \end{cases} \qquad k = 1, \ldots, n$$

则$X_1, \ldots, X_n$独立同分布，且 $E(X_k) = p$，$D(X_k) = p(1-p)$，

记 $X = \sum_{k=1}^{n} X_k$ 则 $E(X) = np$，$D(X) = np(1-p)$，

故由有Lévy-Lindeberg中心极限定理知 X近似服从$N(np, np(1-p))$.

De Moivre-Laplace中心极限定理

设 n_A 为n重伯努利试验中事件A发生的次数，

又A在每次试验中发生的概率为p $(0 < p < 1)$，则对任意实数x, 成立

$$\lim_{n\to\infty} P\left(\frac{n_A - np}{\sqrt{np(1-p)}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N(n\mu, n\sigma^2)$

例3　对于一个学生而言，来参加家长会的家长人数是一个随机变量.

设一个学生无家长、1名家长、2名家长来参加会议的概率分别

为 0.05，0.8，0.15. 若学校共有400名学生，设各学生参加会议

的家长数相互独立，且服从同一分布.

(1) 求参加会议的家长数X 超过450的概率；$P(X > 450)$.

(2) 求有1名家长来参加会议的学生数不多于340的概率.

随机变量 $\sum_{k=1}^{n} X_k$ 近似服从正态分布$N\,(n\mu, n\sigma^2)$

解： (1) 求参加会议的家长数X 超过450的概率，即 $P\,(X > 450)$.

设X_k表示第k位学生来参加会议的家长数, $k = 1, 2, \ldots, 400$, 则

X_k	0	1	2
p_k	0.05	0.8	0.15

故 $E(X_k) = 1.1$， $D(X_n) = 0.19$,

$$X = \sum_{k=1}^{400} X_k \text{ 近似服从} N\,(400 \times 1.1, 400 \times 0.19)$$

随机变量 $X=\sum_{k=1}^{400} X_k$ 近似服从$N\,(400\times 1.1,\, 400\times 0.19)$

$$P(X>450)=P\left\{\frac{X-400\times 1.1}{\sqrt{400\times 0.19}}>\frac{450-400\times 1.1}{\sqrt{400\times 0.19}}\right\}$$

$$=1-P\left\{\frac{X-400\times 1.1}{\sqrt{400\times 0.19}}\le 1.147\right\}$$

$$=1-\Phi(1.147)\approx 0.1357$$

当n充分大时，可以利用正态分布近似计算二项分布

X_k	0	1	2
p_k	0.05	0.8	0.15

(2) 求有1名家长来参加会议的学生数不多于340的概率.

设Y表示只有一名家长来参加会议的学生数则求$P(Y \leq 340)$.

$$Y \sim B(400, 0.8)$$

由De Moivre-Laplace中心极限定理知,

Y近似服从$N\,(400 \times 0.8,\ 400 \times 0.8 \times 0.2)$.

随机变量 Y近似服从$N(400\times 0.8, 400\times 0.8\times 0.2)$

$$P(Y\le 340)=P\left\{\frac{Y-400\times 0.8}{\sqrt{400\times 0.8\times 0.2}}\le\frac{340-400\times 0.8}{\sqrt{400\times 0.8\times 0.2}}\right\}$$

$$=P\left\{\frac{Y-400\times 0.8}{\sqrt{400\times 0.8\times 0.2}}\le 2.5\right\}\approx\Phi(2.5)=0.9938$$

当n充分大时，可以利用正态分布近似计算二项分布

例4. 电视台作某节目A收视率的调查，在每天节目A播出时随机地向当地居民打电话，问是否在看电视，再问是否在看节目A。设回答在看电视的居民数为n，为保证以95%的概率使调查误差不超过10%，问n至少应取多少？

解：设回答在看电视的居民中在收看节目A的人数为n_A，p为节目A的收视率

依题意要求最小的n使得 $P\left(\left|\frac{n_A}{n}-p\right| \le 0.1\right) \ge 0.95$

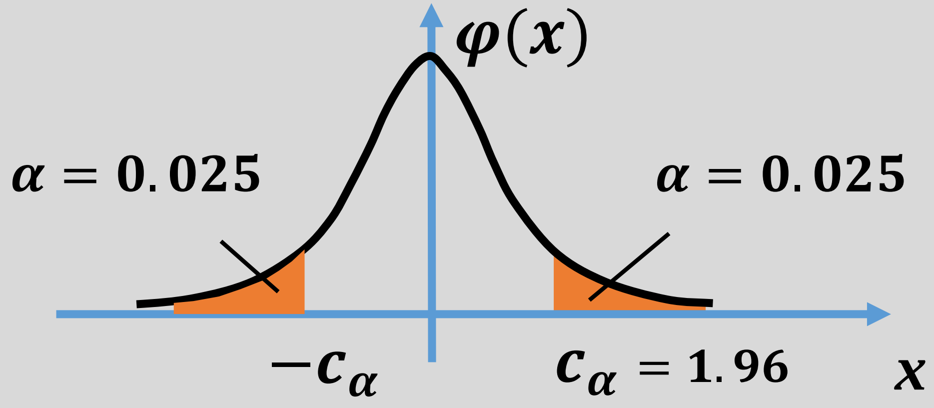
注意到 $n_A \sim B(n,p)$ 由De Moivre-Laplace中心极限定理，$\frac{n_A-np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)$

$$\frac{n_A - np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)$$

$$0.95 \le P\left(\left|\frac{n_A}{n} - p\right| \le 0.1\right) = P\left(\left|\frac{n_A - np}{\sqrt{np(1-p)}}\right| \le \frac{\sqrt{n}}{10\sqrt{p(1-p)}}\right)$$

$\Rightarrow$ 求最小的n使 $P\left(|X| \le \frac{\sqrt{n}}{10\sqrt{p(1-p)}}\right) \ge 0.95$ 其中 $X \sim N(0,1)$

$$P\left(X \le \frac{\sqrt{n}}{10\sqrt{p(1-p)}}\right) \ge 0.975$$



查表可得 $\frac{\sqrt{n}}{10\sqrt{p(1-p)}} \ge 1.96$ 求得 $n \ge 19.6^2 p(1-p)$

求最小的n使得$P\left(\left|\frac{n_A}{n}-p\right|\le 0.1\right)\ge 0.95$

$$h(p)=p(1-p)\Longrightarrow h'(p)=1-2p=0$$

$$\Longrightarrow h_{max}(p)=h\left(\frac{1}{2}\right)=\frac{1}{4}$$

$$\Longrightarrow n\ge 19.6^{2}\times\frac{1}{4}=96.04\Longrightarrow n=97$$

对任意 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\left(\left|\frac{n_A}{n}-p\right|<\varepsilon\right)=1$

$\Longrightarrow$ 估计 $P\left(\left|\frac{n_A}{n}-p\right|\le\varepsilon\right)=P\left(\left|\frac{n_A-np}{\sqrt{np(1-p)}}\right|\le\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right)$

小　结

- Lévy-Lindeberg 中心极限定理
- De Moivre-Laplace 中心极限定理