# §6.3 中心极限定理

**定理1　　独立同分布的中心极限定理**

设随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，且有期望和方差：

$$E(X_k) = \mu, \quad D(X_k) = \sigma^2 > 0, \quad k = 1, 2, \cdots$$

则对于任意实数 $x$，

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

即：

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right) = \Phi(x)$$

即 $n$ 足够大时，$\dfrac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$ 的分布函数近似于标准正态随机变量的分布函数。

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{\text{近似}}{\sim} N(0,1)$$

$\sum_{k=1}^{n} X_k$ 近似服从 $N(n\mu, n\sigma^2)$

**定理2** **德莫佛 — 拉普拉斯中心极限定理（DeMoivre-Laplace）**

设 $Y_n \sim B(n, p)$，$0 < p < 1$，$n = 1,2,\ldots$

则对任一实数 $x$，有

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

说明：若定理1中 $X_1, X_2, \cdots, X_n, \cdots$

独立同分布为0-1分布，$P(X_i=1) = p$

则：$Y_n = \sum_{i=1}^{n} X_i \sim B(n, p)$

即对任意的 $a < b$,

$$\lim_{n\to\infty} P\left( a < \frac{Y_n - np}{\sqrt{np(1-p)}} \le b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

$Y_n \sim N(np, np(1-p))$ (近似)

## 中心极限定理的意义

在实际问题中，若某随机变量可以看作是有相互独立的大量随机变量综合作用的结果，每一个因素在总的影响中的作用都很微小，则综合作用的结果服从正态分布.

# 中心极限定理的应用

**例1** 设有一大批种子，其中良种占1/6. 试估计在任选的6000粒种子中，良种所占比例与1/6比较上下不超过1%的概率.

**解** 设 $X$ 表示6000粒种子中的良种数，则

$$X \sim B(6000, 1/6)$$

$$E(X) = 1000,\ D(X) = \frac{5000}{6}$$

$$X \overset{\text{近似}}{\sim} N\left(1000, \frac{5000}{6}\right)$$

$$P\left(\left|\frac{X}{6000}-\frac{1}{6}\right| \le 0.01\right) = P\left(\left|X-1000\right| \le 60\right) = P\left(940 \le X \le 1060\right)$$

$$\approx \Phi\left(\frac{1060-1000}{\sqrt{5000/6}}\right) - \Phi\left(\frac{940-1000}{\sqrt{5000/6}}\right)$$

$$= \Phi\left(\frac{60}{\sqrt{5000/6}}\right) - \Phi\left(\frac{-60}{\sqrt{5000/6}}\right)$$

$$= 2\Phi\left(\frac{60}{\sqrt{5000/6}}\right) - 1 \quad \approx 0.9624$$

# 比较几个近似计算的结果

用二项分布(精确结果)

$$X \sim B(6000, 1/6)$$

$$P\left(\left|\frac{X}{6000}-\frac{1}{6}\right| \le 0.01\right) = P\left(940 \le X \le 1060\right)$$

$$= \sum_{k=940}^{1060} C_{6000}^{k} \left(\frac{1}{6}\right)^{k} \left(1-\frac{1}{6}\right)^{6000-k}$$

$$\approx 0.9590$$

用Poisson 分布 $\lambda = np = 6000 \cdot \frac{1}{6} = 1000$

$$P\left(\left|\frac{X}{6000} - \frac{1}{6}\right| \le 0.01\right) = P\left(940 \le X \le 1060\right)$$

$$= \sum_{k=940}^{1060} \frac{1000^k \cdot e^{-1000}}{k!}$$

$$\approx 0.9379$$

用Chebyshev 不等式 $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\geq 0.7685$

用中心极限定理 $P\left(\left|\frac{X}{6000}-\frac{1}{6}\right|<0.01\right)\approx 0.9624$

**例2** 某车间有200台车床，每台独立工作，开工率为0.6. 开工时每台耗电量为 $r$ 千瓦. 问供电所至少要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** 设至少要供给这个车间 $a$ 千瓦的电力

设 $X$ 为200 台车床的开工数.

$$X \sim B(200, 0.6)\ ,\ X \sim N(120, 48)\ (\text{近似})$$

问题转化为求 $a$ , 使

$$P(0 \le rX \le a) = 99.9\%$$

由于将 $X$ 近似地看成正态分布，故

$$P(0 \le rX \le a) = P(0 \le X \le \frac{a}{r})$$

$$= \Phi\left(\frac{\frac{a}{r} - 120}{\sqrt{48}}\right) - \Phi\left(\frac{0-120}{\sqrt{48}}\right)$$

$$\approx \Phi\left(\frac{\frac{a}{r} - 120}{\sqrt{48}}\right) - \Phi(-17.32)$$

$$\approx \Phi\left(\frac{\frac{a}{r} - 120}{\sqrt{48}}\right)$$

反查标准正态函数分布表，得

$$\Phi(3.09) = 99.9\%$$

令

$$\frac{\frac{a}{r} - 120}{\sqrt{48}} = 3.09$$

解得

$$a = (3.09\sqrt{48} + 120)r$$

$$\approx 141r \text{ (千瓦)}$$

**例3** 检查员逐个地检查某种产品，每检查一只产品需要用10秒钟．但有的产品需重复检查一次，再用去10秒钟．假设产品需要重复检查的概率为 0.5, 求检验员在 8 小时内检查的产品多于1900个的概率.

**解** 检验员在 8 小时内检查的产品多于1900个即检查1900个产品所用的时间小于 8 小时.

设 $X$ 为检查1900 个产品所用的时间(单位:秒)

设 $X_k$ 为检查第 $k$ 个产品所用的时间(单位：秒)，$k = 1,2,\ldots,1900$

| $X_k$ | 10 | 20 |
| --- | --- | --- |
| $P$ | 0.5 | 0.5 |

$$E(X_k) = 15, \quad D(X_k) = 25$$

$X_1, X_2, \cdots, X_{1900}$ 相互独立,且同分布, $X = \sum_{k=1}^{1900} X_k$

$$E(X) = 1900 \times 15 = 28500$$

$$D(X) = 1900 \times 25 = 47500$$

$$X \overset{\text{近似}}{\sim} N(28500, 47500)$$

$$P(10\times 1900 \le X \le 3600\times 8)$$

$$= p(19000 \le X \le 28800)$$

$$\approx \Phi\left(\frac{28800-28500}{\sqrt{47500}}\right)-\Phi\left(\frac{19000-28500}{\sqrt{47500}}\right)$$

$$\approx \Phi\left(1.376\right)-\Phi\left(-43.589\right)$$

$$\approx 0.9162$$

## 解法二

$$\frac{X-1900\cdot 10}{10}\text{——}1900\text{个产品中需重复检查的个数}$$

$$\frac{X-1900\cdot 10}{10}\sim B(1900,0.5)\overset{\text{近似}}{\sim} N\left(950,475\right)$$

$$P(10\times 1900\le X\le 3600\times 8)$$

$$=P\left(19000\le X\le 28800\right)$$

$$=P\left(0\le \frac{X-19000}{10}\le \frac{28800-19000}{10}\right)$$

$$= P\left(0 \le \frac{X - 19000}{10} \le 980\right)$$

$$\approx \Phi\left(\frac{980 - 950}{\sqrt{475}}\right) - \Phi\left(\frac{0 - 950}{\sqrt{475}}\right)$$

$$\approx \Phi(1.376) - \Phi(-43.589)$$

$$\approx 0.9162$$