5.2 中心极限定理

北京化工大学数学系

苏贵福

在客观实际中有许多随机变量，它们是由大量的相互独立的随机因素的综合影响所形成的. 而其中每一个别因素在总的影响中所起的所用都是微不足道的. 这种随机变量往往近似地服从正态分布. 这种现象就是中心极限定理的客观背景.

定理1 (独立同分布的中心极限定理) 设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立, 服从同一分布, 且具有数学期望和方差:

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

则随机变量之和$\sum_{k=1}^{n} X_k$的标准化变量

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

的分布函数$F_n(x)$对任意x满足

$$\begin{aligned}\lim_{n\to\infty} F_n(x) &= \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\} \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).\end{aligned}$$

♣ 定理1表明: n个相互独立且服从同一分布的随机变量 $X_1, X_2, \cdots, X_n$的和$\sum_{k=1}^n X_k$的极限分布为正态分布. 而$\sum_{k=1}^n X_k$的标准化变量近似服从标准正态分布.

♣ 一般地, 在一定条件下随机变量序列$\{X_n\}$的部分和$\sum_{k=1}^n X_k$, 经标准化后所得的随机变量序列的分布函数收敛到标准正态分布的分布函数$\Phi(x)$这一结果, 统称为中心极限定理.

例1 计算器在进行加法时, 将每个加数舍入最靠近它的整数. 设所有舍入误差$X_1, X_2, \cdots, X_n$是独立的且在$(-0.5, 0.5)$上服从均匀分布.

(1) 若将1500个数相加, 求误差总和的绝对值超过15的概率.

(2) 最多可有几个数相加才使得误差总和的绝对值小于10的概率不小于0.90?

解 (1) 已知$X_1, X_2, \cdots, X_{1500}$相互独立, 且服从$(-0.5, 0.5)$上 的均匀分布, 故$E(X_k) = 0$, $D(X_k) = \frac{1}{12}$, $k = 1, 2, \cdots, 1500$. 由定理1知

$$\frac{\sum_{K=1}^{1500} X_k - 1500 \times 0}{\sqrt{1500 \times \frac{1}{12}}}$$

近似服从标准正态分布.

因此所求概率为

$$P\left\{\left|\sum_{k=1}^{1500} X_k\right| > 15\right\} = P\left\{\left|\frac{\sum_{k=1}^{1500} X_k}{\sqrt{125}}\right| > \frac{15}{\sqrt{125}}\right\}$$

$$\approx 2 - 2\Phi\left(\frac{15}{\sqrt{125}}\right)$$

$$= 2[1 - \Phi(1.3416)] = 0.1796.$$

(2) 根据题意

$$P\left\{\left|\sum_{k=1}^{n} X_k\right| < 10\right\} = P\left\{\left|\frac{\sum_{k=1}^{n} X_k - 0}{\sqrt{n/12}}\right| < \frac{10}{\sqrt{n/12}}\right\} \geq 0.9.$$

即$2\Phi\left(\frac{10\sqrt{12}}{\sqrt{n}}\right) - 1 \geq 0.9$, 亦即$\Phi\left(\frac{10\sqrt{12}}{\sqrt{n}}\right) \geq 0.95$. 查标准正态分布表得$\frac{10\sqrt{12}}{\sqrt{n}} \geq 1.645$, 故$n \leq 443$.

定理2 (德莫佛-拉普拉斯中心极限定理) 设随机变量Y_n $(n=1,2,\cdots)$ 服从参数为n和p $(0<p<1)$的二项分布. 则对任意的x总有

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

♠ 定理2表明: 正态分布是二项分布的极限分布. 因此, 当n充分大时, 我们通常用

$$P\left\{\frac{Y_n-np}{\sqrt{np(1-p)}}\leq x\right\}\approx\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

来近似计算二项分布的概率.

例2 一船舶在某海域航行, 已知每遭受一次波浪的冲击, 纵摇角大于3°的概率为$p=\frac{1}{3}$. 若船舶遭受了90000次波浪冲击. 试问其中有 $29500-30500$ 次纵摇角大于3°的概率是多少?

解 我们将船舶每遭受一次波浪冲击看作是一次试验, 并假设各次试验是相互独立的. 在90000次波浪冲击中纵摇角大于3°的次数记作X. 则X是一随机变量, 且有$X\sim b(9000,\frac{1}{3})$. 其分布律为

$$P\{X=k\}=\binom{90000}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{90000-k},\ k=1,2,\cdots,90000.$$

因此所求概率为

$$P\{29500\le X\le 30500\}=\sum_{k=29500}^{30500}\binom{90000}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{90000-k}.$$

要直接计算上述概率比较麻烦. 下面利用德莫佛-拉普拉斯定理求它的近似值. 即有

$$
\begin{aligned}
P\{29500 \le X \le 30500\} =& P\left\{\frac{29500-np}{\sqrt{np(1-p)}} \le \frac{X-np}{\sqrt{np(1-p)}} \le \frac{30500-np}{\sqrt{np(1-p)}}\right\} \\
\approx & \int_{\frac{29500-np}{\sqrt{np(1-p)}}}^{\frac{30500-np}{\sqrt{np(1-p)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\
=& \Phi\left(\frac{30500-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{29500-np}{\sqrt{np(1-p)}}\right).
\end{aligned}
$$

将$n = 90000$和$p = \frac{1}{3}$代入上式可得

$$
P\{29500 \le X \le 30500\} \approx \Phi\left(\frac{5\sqrt{2}}{2}\right) - \Phi\left(-\frac{5\sqrt{2}}{2}\right) = 0.9995.
$$

例3 设随机变量$X_1, X_2, \cdots$相互独立, 且 X_i均服从$\mathcal{P}(\lambda)$, 则

$$\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\lambda}{\sqrt{n\lambda}} > x\right\} = 1 - \Phi(x)$$

并计算当$n = 100$, $\lambda = 2$时, $P\{\sum_{i=1}^{n} X_i > 200\}$的近似值.

解 因为随机变量序列均服从泊松分布, 所以 $E(X_i) = \lambda$, $D(X_i) = \lambda$. 由独立同分布的中心极限定理直接得

$$\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\lambda}{\sqrt{n\lambda}} > x\right\} = 1 - \Phi(x)$$

当$n = 100$, $\lambda = 2$时, 有

$$P\left\{\sum_{i=1}^{n} X_i > 200\right\} = P\left\{\frac{\sum_{i=1}^{n} X_i - 100\cdot 2}{10\sqrt{2}} > \frac{200 - 100\cdot 2}{10\sqrt{2}}\right\} = 1 - \Phi(0) = 0.5$$