第 五 章 大数定律和中心极限定理

第二节 中心极限定理

1、中心极限定理的定义

2、列维-林德伯格中心极限定理

3、棣莫佛-拉普拉斯中心极限定理

4、小结、思考

一、中心极限定理

设随机变量 $X \sim N(0,1)$，$\{X_k\}$，$k = 1,2,\ldots$相互独立，且数学期望和方差都存在，若标准化随机变量序列

$$\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k)}{\sqrt{\sum_{k=1}^{n} D(X_k)}}$$

依分布收敛于 X，则称随机变量序列$\{X_k\}$，$k = 1,2,\ldots$ 服从中心极限定理。■

若随机变量序列$\{X_k\}$，$k=1,2,\ldots$满足中心极限定理，有

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

$$= \Phi(x)$$

即$\sum_{k=1}^{n} X_k$的标准化随机变量的极 限分布为标准正态分布．

所以当n 足够大时，可以认为

$$\frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k)}{\sqrt{\sum_{k=1}^{n} D(X_k)}} \sim N(0,1)$$

近似成立，从而

$$\sum_{k=1}^{n} X_k \sim N\left(\sum_{k=1}^{n} E(X_k), \sum_{k=1}^{n} D(X_k)\right)$$

近似成立。

二. 中心极限定理

1 独立同分布中心极限定理

设$\{X_k\}$，$k=1,2\ldots$是一个相互独立、具有同分布的随机变量序列，且$E(X_k)=\mu$，$D(X_k)=\sigma^2$。

则随机变量序列$\{X_k\}$满足中心极限定理，即有

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

1’ 独立同分布中心极限定理

设$\{X_k\}$，$k=1,2\ldots$是一个相互独立、具有同分布的随机变量序列，且$E(X_k)=\mu$， $D(X_k)=\sigma^2$。

当n充分大时

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

近似服从N(0,1) ■

独立同分布中心极限定理的应用

1）求随机变量之和 $S_n = \sum_{i=1}^{n} X_i$ 取值的概率。

产 品 检 验

2）已知 $S_n = \sum_{i=1}^{n} X_i$ 取值的概率，反求**n**。

产 品 测 重

3）数理统计中大样本推断的理论基础。

例： 检验员逐个地检查某种产品，每次花10 s检查一个，但也可能有的产品需要重复检查一次再用去10s，假设每个产品需要重复检查的概率为0. 5 ，求在8h内检查员检查的产品多于1900个的概率。

分析：问题等价于求检查员检查1900个产品所花的总时间 X 不超过8h的概率 P{ X≤ 8 *3600 }

解：设检查第i个产品所花时间为X_i(i = 1,2,3,…,1900) 则检查1900个产品所花的总时间为：

$$X = \sum_{i=1}^{1900} X_i$$

显然 $X_1, X_2, \ldots, X_{1900}$ 为相互独立的随机变量,且

$$X_i = \begin{cases} 10 & 第\ i\ 个产品没有重复检查; \\ 20 & 第\ i\ 个产品需要重复检查一\quad 次; \end{cases}$$

同时 $P(X_i=10) = P(X_i=20) = 0.5\ (i = 1,2,3,\ldots,1900)$

即X_i相互独立都服从同一分布。

$$E(X_i) = 10 \times 0.5 + 20 \times 0.5 = 15$$

$$D(X_i) = E(X_i^2) - [E(X_i)]^2 = 25$$

$$i = 1,2,\ldots,1900$$

由独立同分布中心极限定理，知

$$X = \sum_{i=1}^{1900} X_i \dot{\sim} N(1900 \times 15,\ 1900 \times 25)$$

所求概率为:

$$P\left(X \le 8 \times 3600\right) \approx \Phi\left(\frac{8 * 3600 - 1900 * 15}{\sqrt{1900 * 25}}\right)$$

$$= \Phi\left(\frac{6}{\sqrt{19}}\right) = 0.9162 \qquad \blacksquare$$

例： 在天平上重复独立称一重为a的物体，各次称量的结果X_i 同服从正态分布 N(a，0.04），若以 $\overline{X}_n$ 表示n次称量结果的算术平均值，为使

$$P\left(\left|\overline{X}_n - a\right| < 0.1\right) \geq 0.95$$

问至少要称量多少次？

解:由题设知 $X_1, X_2, \ldots, X_n$ 相互独立,且$X_i \sim N(a, 0.04)$

$$E(X_i) = a \qquad D(X_i) = 0.04 \qquad i = 1, 2, \ldots, n$$

由独立同分布中心极限定理，知

$$X = \sum_{i=1}^{n} X_i \dot{\sim} N(na, 0.04n)$$

由正态分布性质，有

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \overset{\cdot}{\sim} N\left(a, \frac{0.04}{n}\right)$$

由题设，有

$$P\left(\left|\overline{X}_n - a\right| < 0.1\right) = P\left(-0.1 < \overline{X}_n - a < 0.1\right)$$

$$\approx \Phi\left(\frac{0.1}{0.2}\sqrt{n}\right) - \Phi\left(-\frac{0.1}{0.2}\sqrt{n}\right)$$

$$= 2\Phi\left(0.5\sqrt{n}\right) - 1 \geq 0.95$$

$$\Phi\left(0.5\sqrt{n}\right) \geq 0.975$$

查表有 $\Phi(1.96) = 0.975$ 因 $\Phi(x)$ 单调不减，有

$$0.5\sqrt{n} \geq 1.96 \qquad n \geq 15.36$$

至少要称16次才能满足要求。 ■

2 棣莫佛-拉普拉斯中心极限定理

设随机变量序列{ Y_n }，$Y_n \sim B(n, p)$，n =1,2… 。则对于任意的实数 x，有

$$\lim_{n\to\infty} P\left\{ \frac{Y_n - np}{\sqrt{np(1-p)}} \le x \right\} = \Phi(x)$$

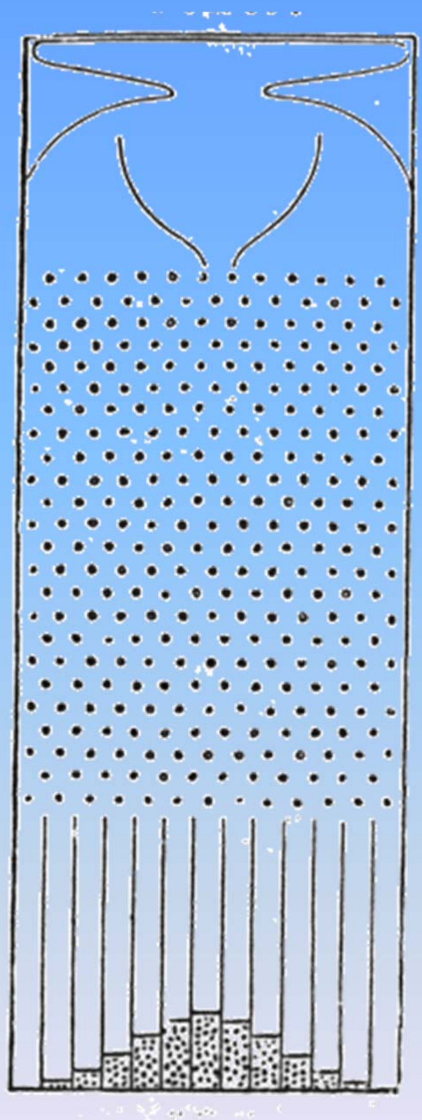
证明： 对于任意正整数n，随机变量Y_n 可表示为

$$Y_n = X_1 + X_2 + \ldots + X_n$$

$X_1, X_2, \ldots, X_n$ 相互独立，$X_i \sim B(1, p)$，且有

$$E(X_i) = p \quad , D(X_i) = p(1-p)$$

所以相互独立的随机变量序列{ X_i }, i =1,2,…满足中心极限定理。易知定理成立 ■



高尔顿板能不能演示中心极限定理？

由棣莫弗-拉普拉斯中心极限定理，若$X \sim B(n, p)$，对于足够大的n，有

$$P\{m_1 < X \le m_2\}$$

$$= P\left\{\frac{m_1 - np}{\sqrt{np(1-p)}} < \frac{X - np}{\sqrt{np(1-p)}} \le \frac{m_2 - np}{\sqrt{np(1-p)}}\right\}$$

$$\approx \Phi\left(\frac{m_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{np(1-p)}}\right)$$

棣莫佛—拉普拉斯中心极限定理的应用

1）计算在大n情况下二项分布概率的近似值。

2）已知在大n情况下二项分布在某范围内取值的概率，求该范围。

3）近似计算 与 用频率估计概率的 有关问题。

航 船 的 稳 定 性

例：一船舶在某海区航行，已知每遭受一次波浪的冲击，纵摇角大于3°的概率为 $p = 1/3$，若船舶遭受了90000次波浪冲击，问其中有 $29500 \sim 30500$ 次纵摇角大于3°的概率是多少？

解：假定船舶遭受波浪的各次冲击是独立的。记 X 为90000次冲击下纵摇角大于3°的次数，故有

$$X \sim B(90000, \frac{1}{3}), \qquad n = 90000, \ p = \frac{1}{3}$$

所求事件的概率为

$$P(29500 < X \leq 30500)$$

$$= P\left\{ \frac{29500 - np}{\sqrt{np(1-p)}} < \frac{X - np}{\sqrt{np(1-p)}} \leq \frac{30500 - np}{\sqrt{np(1-p)}} \right\}$$

航 船 的 稳 定 性

$$P(29500 < X \le 30500)$$

$$\approx \Phi\left\{\frac{30500 - np}{\sqrt{np(1-p)}}\right\} - \Phi\left\{\frac{29500 - np}{\sqrt{np(1-p)}}\right\}$$

$$= \Phi\left\{\frac{5\sqrt{2}}{2}\right\} - \Phi\left\{-\frac{5\sqrt{2}}{2}\right\}$$

$$= 2\Phi\left\{\frac{5\sqrt{2}}{2}\right\} - 1 = 0.995 \quad \blacksquare$$

因为 $\dfrac{X - np}{\sqrt{np(1-p)}}$ 近似服从 $N(0,1)$

小　结

1. 中心极限定理在近似计算与理论研究中都有很重要的作用。

列维—林德伯格中心极限定理	⇒	棣莫佛—拉普拉斯中心极限定理