第 5 章教学要求：

1. 了解切比雪夫大数定理、伯努利(Bernoulli)大数定理、辛钦(Khinchine)大数定理.

2. 了解列维−林德伯格(Levy-Lindberg)中心极限定理(独立同分布的中心极限定理)和棣莫弗−拉普拉斯(De Moivre-Laplace)中心极限定理.

3. 了解棣莫弗−拉普拉斯中心极限定理在实际问题中的应用.

第 5 章　大数定理与中心极限定理

§5.1 大数定理

第 1 章已指出：频率的稳定值是概率. 伯努利大数定理以严格的数学形式表达了频率的这种稳定性. 凡是阐述大量随机现象的平均结果具有稳定性的定理都称为大数定理.

大量的重复试验的平均结果如何用数学语言描述？

大量的重复试验即 n 重伯努利试验概型：各次试验是相互独立的，且在每次试验中某事件 A 发生的概率 p 保持不变.

例如，抛一枚均匀的硬币，设 A 表示“出现正面”，试验数据显示，在相同的条件下，A 发生频率的稳定值为 0.5.

如果设

$$X_i=\begin{cases}1, & \text{第}i\text{次试验}A\text{发生，}\\ 0, & \text{第}i\text{次试验}A\text{不发生，}\end{cases}\quad i=1,2,\cdots,$$

则显然 $X_1,X_2,\cdots$ 是相互独立、有相同分布(简称独立同分布)的随机变量序列，且

$$E(X_i)=\frac{1}{2},\ D(X_i)=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4},\quad i=1,2,\cdots.$$

令

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\ ,$$

这是 n 次试验中事件 A 发生的频率，在这里又称为 n 个随机变量 $X_1,X_2,\cdots,X_n$ 的**算数平均**，当 n 很大时，$\bar{X}$ 的取值应该稳定在 0.5.

$\bar{X}$ 是随机变量，该如何表示 $\bar{X}$ 的取值稳定在 0.5 呢？

一种提法是：当 n 足够大时，$\bar{X}$ 与 0.5 有较大偏差的概率很小. 其数学语言的表述是，对于任意正数 ε，有

$$\lim_{n\to\infty}P\{|\bar{X}-0.5|\ge\varepsilon\}=0.$$

或

$$\lim_{n\to\infty}P\{|\bar{X}-0.5|<\varepsilon\}=1.$$

定义 5.1　设 $Y_1,Y_2,\cdots,Y_n,\cdots$ 是一个随机变量序列，如果∃a，使得∀ε>0，有

$$\lim_{n\to\infty}P\{|Y_n-a|<\varepsilon\}=1,$$

则称 $Y_1,Y_2,\cdots,Y_n,\cdots$ **依概率收敛**于 **a**，并记 $Y_n\xrightarrow{P}a$.

依概率收敛有如下性质：

若 $X_n\xrightarrow{P}a$，函数 g(x)在 a 处连续，则

$$g(X_n)\xrightarrow{P}g(a).$$

若 $X_n\xrightarrow{P}a, Y_n\xrightarrow{P}b$，函数 g(x,y)在(a,b)处连续，则

$$g(X_n,Y_n)\xrightarrow{P}g(a,b).$$

更多元随机变量的序列有类似的结果.（这是矩估计的理论基础）

定理 5.1（切比雪夫大数定理的特殊情况）　设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，且具有相同的数学期望和方差：$E(X_i)=\mu, D(X_i)=\sigma^2\,(i=1,2,\cdots)$. 若令

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,$$

则对于任意正数 ε，有

$$\lim_{n\to\infty}P\{|\overline{X}-\mu|<\varepsilon\}=1.$$

切比雪夫大数定理表明：相互独立、有相同的数学期望和方差的随机变量 $X_1,X_2,\cdots,X_n$ 的算数平均依概率收敛于算数平均的期望，即

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}E(X_i).$$

定理 5.2（伯努利大数定理）　设 n_A 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意正数 ε，有

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}=1.$$

伯努利大数定理表明：事件发生的频率依概率收敛于事件发生的概率．因此，当试验次数很大时，常用事件发生的频率来代替事件发生的概率.

定理 5.3（辛钦大数定理）　设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，服从同一分布，具有数学期望

$E(X_i)=\mu$, $i=1,2,\cdots$，则对于任意正数 ε，有

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

辛钦大数定理表明：相互独立同分布、具有数学期望的随机变量 $X_1,X_2,\cdots,X_n$ 的算数平均依概率收敛于算数平均的期望.

例 设 $X_1,X_2,\cdots,X_n$ 独立同分布于参数为 2 的指数分布，则当 $n\to\infty$ 时，$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 依概率收敛于____.

§5.2 中心极限定理

如何计算关于随机变量 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 的事件的概率？切比雪夫不等式只给出了估计关于 $\bar{X}$ 的特定事件的概率：

$$P\left\{\left|\bar{X}-E(\bar{X})\right|\geq\varepsilon\right\}\leq\frac{D(\bar{X})}{\varepsilon^2}.$$

大数定理则指出：

$$\lim_{n\to\infty}P\{\left|\bar{X}-E(X)\right|\geq\varepsilon\}=0.$$

一般说来，概率计算依赖于随机变量的概率分布，而 $\bar{X}$ 的概率分布一般不好求，所以计算关于 $\bar{X}$ 的事件的概率不容易．本节讨论的中心极限定理则能较好地解决这个问题.

定理 5.4（列维–林德伯格(Levy-Lindberg)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，具有数学期望和方差 $E(X_i)=\mu,D(X_i)=\sigma^2>0(i=1,2,\cdots)$，则对任意实数 x，有

$$\lim_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leq x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x).$$

定理 5.4 又称为独立同分布中心极限定理.

定理 5.4 的含义：因为

$$\frac{\sum_{i=1}^{n}X_i-E\left(\sum_{i=1}^{n}X_i\right)}{\sqrt{D\left(\sum_{i=1}^{n}X_i\right)}}=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma},\quad n=1,2,\cdots,$$

所以 $\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$ 其实是 $\sum_{i=1}^{n}X_i$ 的标准化变量．因此定理 5.4 实际上表明，均值为 μ、方差为 σ^2(σ>0)的独立同分布的随机变量和 $\sum_{i=1}^{n}X_i$ 的标准化变量，当 n 很大时，

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \stackrel{\cdot}{\sim} N(0,1)\text{,} \tag{5.2}$$

而且 n 越大近似程度越好.

由正态变量的性质，亦即

$$\sum_{i=1}^{n} X_i \stackrel{\cdot}{\sim} N(n\mu, n\sigma^2)\text{,} \tag{5.3}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \stackrel{\cdot}{\sim} N(\mu, \frac{1}{n}\sigma^2)\text{.} \tag{5.4}$$

后一个结果是数理统计中大样本统计推断的基础.

中心极限定理的意义：无论随机变量序列 $X_i(i=1,2,\cdots)$ 服从什么分布，只要满足四个条件，那它们的和 $\sum_{i=1}^{n} X_i$，当 n 很大时，就近似地服从正态分布. 这是正态分布在概率统计中占有重要地位的一个基本原因.

例 5.1 (P_{157})

定理 5.5 (棣莫弗-拉普拉斯(De Moivre-Laplace)定理) 设随机变量 $\eta_n \sim B(n,p), n=1,2,\cdots$，则对任意实数 x，有

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

定理 5.5 表明：二项变量 $\eta_n \sim B(n,p)$ 近似地服从正态分布 $N(np, np(1-p))$.

例 5.2 (P_{158})

例 5.3 (P_{158})

作业 (P_{159})：1.–4.　6.　5.　7.

(P_{160})：1.　3.　4.　5.

补充：

1. 大数定理　设 $X_1, X_2, \cdots, X_n, \cdots$ 是随机变量序列，令 $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i (n=1,2,\cdots)$. 若存在常数序列 $a_1, a_2, \cdots, a_n, \cdots$，使得对于任意的正数 ε，总有

$$\lim_{n\to\infty}\{|Y_n - a_n| < \varepsilon\} = 1\text{,}$$

则称随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$ 服从**大数定理**(或大数法则).

2. 契比雪夫大数定理　设 $X_1, X_2, \cdots, X_n, \cdots$ 是两两不相关的随机变量序列，每一随机变量都有有限的方差，且有共同的上界，即 $D(X_i) \le C(i=1,2,\cdots)$，则对于任意的正数 ε，总有

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

3. 普阿松大数定理　在一个独立试验序列中，设事件 A 在第 k 次试验中发生的概率为 p_k，以 μ_n 表示在前 n 次试验中事件 A 发生的次数，则对于任意的正数 ε，总有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\mu_n - \frac{1}{n}\sum_{k=1}^{n} p_k\right| < \varepsilon\right\} = 1.$$

大数定理的重要意义　人类的大量经验告诉我们，概率很接近 1 的随机事件在一次试验中几乎一定要发生，即可以把该事件看作实际中的必然事件；同样，概率很小的随机事件在一次试验中几乎不发生，即可以把该事件看作实际中的不可能事件. 至于概率小到何种程度才能看作不可能事件则要视事件的重要性而定. 如，有 1％可能性染菌的药物应该弃用，但有 1％可能性为次品的纽扣问题则不大.

概率论的基本命题之一就是要建立概率接近 1 或 0 的规律，大数定理就是这种命题中最重要的一个.

观察个别现象是是连同一切个别的特性来观察的，这些个别的特性往往蒙蔽了事物的规律性：在大量观察中个别因素的影响将相互抵消而使总体稳定. 如：虽说每个气体分子的运动带有很大的随机性，但作为气体平均特征的压力、温度等确实是稳定的. 大数定理说明了这种稳定性.

伯努里大数定理建立了大量重复试验中事件出现的频率的稳定性. 正因为这种稳定性，概率的概念才有了客观意义.

4. 中心极限定理　若独立的随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$ 的前 n 项和的标准化序列

$$Y_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} E(X_i)}{\sqrt{\sum_{i=1}^{n} D(X_i)}},\quad n = 1, 2, \cdots$$

对于任意的实数 y，总有

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

则称随机变量序列 $X_1, X_2, \cdots, X_n, \cdots$ 服从**中心极限定理**.