

probability  
probability

## 第五章

# 大数定律和中心极限定理

probability  
probability

**在实际应用中，常遇到如下问题**

- 1. 现实中为什么大量随机变量服从正态分布？  
依据什么来断定一个随机变量服从正态分布？**
- 2. “频率的稳定性”到底是什么意思？在实际应用中有什么作用？**
- 3. 在计算机上如何模拟现实研究对象？根据什么来认定这种模拟是正确的？ .....**

# 中心极限定理

---

切比雪夫不等式



切比雪夫大数定理

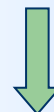


伯努里大数定律



小概率事件原理

林德伯格—列维  
中心极限定理



棣莫佛—拉普拉斯  
中心极限定理

辛钦大数定律

## §5.1 大数定律

### 一. 概率不等式

#### 1. 马尔可科夫(Markov)不等式

设随机变量  $Y$  的  $k$  阶绝对原点矩  $E\{|Y|^k\} < +\infty$ , 则对于任意的  $\varepsilon > 0$ , 有

$$P\{|Y| \geq \varepsilon\} \leq \frac{E\{|Y|^k\}}{\varepsilon^k}, \quad k = 1, 2, \dots$$

**证明**

## 2. 切比雪夫(Chebyshev)不等式

对马尔可科夫不等式特别取  $k = 2$ , 令  $Y = X - E(X)$ ,  $E\{|Y|^2\} = D(X)$  存在, 有切比雪夫不等式成立.

设随机变量  $X$  的数学期望  $E(X)$  和方差  $D(X)$  都存在, 则对于任意的  $\varepsilon > 0$ , 有

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

或者 
$$P\{|X - E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

概率估计

重复试验次数估计

## 二. 大数定律的定义

### 1. 依概率收敛

**定义5.1.1** 设 $X_n, n=1,2,\dots$ 是一个随机变量序列,  $X$  是一个随机变量或常数, 若对于任意的 $\varepsilon > 0$ , 有

$$\lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0$$

或者 
$$\lim_{n \rightarrow \infty} P\{|X_n - X| < \varepsilon\} = 1$$

称随机变量序列 $\{X_n\}$ 依概率收敛于 $X$ ，记为

$$X_n \xrightarrow{P} X \quad \text{或者} \quad \lim_{n \rightarrow \infty} X_n = X, (P)$$

**注1** 在定义中，随机变量 $X$ 也可以是常数 $a$ ，  
称随机变量序列 $\{X_n\}$ 依概率收敛于常数 $a$ 。

**注2** 随机变量序列依概率收敛不同于微积分中数列或函数列的收敛性。

依概率收敛意义的例子

## 结论

随机变量序列 $\{X_n\}$ 依概率收敛于 $X$ , 指当 $n$ 足够大时, 有足够大的概率保证 $X_n$ 任意接近于 $X$ , 但 $X_n$ 仍然有可能与 $X$ 相差很大.

## 2 大数定律的定义

设 $X_n, n=1,2,\dots$ 是随机变量序列, 其数学期望都存在, 若对于任意的 $\varepsilon > 0$ , 有

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

称随机变量序列 $\{X_n\}$ 服从大数定律.



## 思考 能否说

“ $\frac{1}{n} \sum_{i=1}^n X_i$  依概率收敛于  $\frac{1}{n} \sum_{i=1}^n E(X_i)$ ”. ?

**服从大数定律的概率意义：**  $\{X_k\}$ ,  $k=1,2,\dots$   
的前 $n$  项算术平均将紧密地聚集在其数学期望的附近.

## 三. 大数定律

### 定理5.1.1 (切比雪夫大数定律)

设 $X_k$ ,  $k=1,2,\dots$ 是相互独立的随机变量序列,其数学期望和方差都存在, 且存在常数 $C$ , 使得

$$D(X_k) < C, \quad k = 1, 2, \dots$$

则随机变量序列 $\{X_k\}$ ,  $k=1,2,\dots$ 服从大数定律.

**证明**

$$E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E(X_i)$$

$$D\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n^2} \sum_{i=1}^n D(X_i) \leq \frac{nC}{n^2} = \frac{C}{n}$$

根据切比雪夫不等式，对于任意的  $\varepsilon > 0$ ，有

$$\begin{aligned} 1 &\geq P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| < \varepsilon\right\} \\ &\geq 1 - \frac{D\left(\frac{1}{n} \sum_{i=1}^n X_i\right)}{\varepsilon^2} \\ &\geq 1 - \frac{C}{n\varepsilon^2} \rightarrow 1, \quad (\text{as } n \rightarrow \infty). \end{aligned}$$

**例5.1.4**

## 推论 独立同分布大数定律

设  $X_k$ ,  $k=1,2,\dots$  是相互独立且同分布的随机变量序列, 且  $E(X_k) = \mu$ ,  $D(X_k) = \sigma^2$ ,  $k=1,2,\dots$  则  $\{X_k\}$ ,  $k=1,2,\dots$  服从大数定律, 即对任意的  $\varepsilon > 0$ , 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| < \varepsilon\right\} = 1$$

**注 1.** 此定理是切比雪夫大数定律的一个推论.

2. 为在实际应用中用将大量重复测量值的算术平均值作为精确值的估计提供了理论依据.

## 5.1.2 (贝努里(Bernulli)大数定律)

设 $\frac{m}{n}$ 是 $n$ 次重复独立试验中事件 $A$ 发生的频率,

$p$ 是事件 $A$ 在每次试验中发生的概率, 则对任意的 $\forall \varepsilon > 0$ , 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{m}{n} - p\right| < \varepsilon\right\} = 1$$

**注 1.** 此定理是切比雪夫大数定律的推论.

2. 此定理以严格的数学形式描述了频率的稳定性.

**3.小概率事件原理** 概率很小的事件，在一次试验中几乎是不可能发生的，从而在实际中可看成不可能事件.

## 定理5.1.3 辛钦大数定律

设 $X_k$ ,  $k=1,2,\dots$ 是相互独立同分布的随机变量序列,若 $X_k$ 的数学期望存在, 则 $\{X_k\}$ 服从大数定律.



**例5.1.1** 设随机变量序列  $X_1, X_2, \dots, X_n \dots$  服从如下的分布:

$$P\{X_n = 0\} = 1 - \frac{1}{n} \qquad P\{X_n = 2n\} = \frac{1}{n}$$

则对任意的  $\varepsilon > 0$ , 有

$$\lim_{n \rightarrow \infty} P\{|X_n| \geq \varepsilon\} = \lim_{n \rightarrow \infty} P\{X_n = 2n\} = 0$$

故  $\{X_n\}$  依概率收敛于 **0**. 但无论对多大的  $n$ ,  $X_n$  都可能取**远离 0** 的值  $2n$ .



## 马尔科夫不等式证明

**证** 仅证明连续型随机变量的情形，设随机变量 $Y$ 的概率密度函数为 $f_Y(y)$ ，有

$$\begin{aligned} P\{|Y| > \varepsilon\} &= \int_{\{y: |y| \geq \varepsilon\}} f_Y(y) dy \\ &\leq \int_{\{y: |y| \geq \varepsilon\}} \frac{|y|^k}{\varepsilon^k} f_Y(y) dy \\ &\leq \frac{1}{\varepsilon^k} \int_{-\infty}^{+\infty} |y|^k f_Y(y) dy = \frac{E\{|Y|^k\}}{\varepsilon^k} \end{aligned}$$





**例5.1.2** 设随机变量 $X_1, X_2, \dots, X_{10}$ 相互独立并且服从相同的分布, 已知它们的数学期望等于0, 方差等于1,  $Y = X_1 + X_2 + \dots + X_{10}$ , 请估算概率  $P\{-10 < Y < 10\}$  之值。

**解**  $E(Y) = 0, D(Y) = 10,$

$$P\{-10 < Y < 10\} = P\{|Y| < 10\} = P\{|Y - E(Y)| < 10\}$$

由切比雪夫不等式, 有

$$P\{-10 < Y < 10\} \geq 1 - \frac{10}{10^2} = 0.9$$



**例5.1.3** 将一枚均匀硬币连续抛  $n$  次,试用切比雪夫不等式求出  $n$ ,使下式成立.

$$P\{|f_n(A) - P(A)| < 0.01\} \geq 0.99$$

其中  $A = \{ \text{出现正面} \}$

**解** 有  $P(A) = 1/2$ , 令

$$X_i = \begin{cases} 1, & \text{第} i \text{次出现正面;} \\ 0, & \text{否则,} \end{cases} \quad (i = 1, 2, \cdots, n)$$

所以出现正面的总次数为  $\sum_{i=1}^n X_i$ , 从而有

$$f_n(A) = \frac{1}{n} \sum_{i=1}^n X_i$$

而且  $E[f_n(A)] = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{2};$

$$D[f_n(A)] = \frac{1}{n^2} \sum_{i=1}^n D(X_i) = \frac{1}{n} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{4n},$$

**由切比雪夫不等式可得**

$$P\{|f_n(A) - P(A)| < 0.01\} = P\{|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{2}| < 0.01\}$$

$$\geq 1 - \frac{D[f_n(A)]}{0.01^2} = 1 - \frac{1}{0.01^2 \times 4n} \geq 0.99$$

$$\text{从而有 } 4n \geq \frac{1}{0.01^3} \Rightarrow n \geq 250,000(\text{次})$$



## 例5.1.4 泊松大数定律

设 $X_k, k=1,2,\dots$ 是相互独立的随机变量序列,

$$P\{X_n = 1\} = p_n, P\{X_n = 0\} = 1 - p_n = q_n.$$

则随机变量序列 $\{X_k\}, k=1,2,\dots$ 服从大数定律.

**分析** 根据切比雪夫大数定理仅需证明存在常数 $C$ , 使

$$D(X_n) < C, k = 1, 2, \dots$$

有

$$D(X_n) = p_n(1 - p_n) \leq \frac{1}{4}, \quad k = 1, 2, \dots$$

