西安電子科技大學
XIDIAN UNIVERSITY

第5章 大数定律及中心极限定理

概率论与数理统计课程组

CHAPTER 5

大数定律及中心极限定理

§5.1 Chebyshev（切比雪夫）不等式

§5.2 大数定律

§5.3 中心极限定理

由概率论通向数理统计的理论基础

CHAPTER 5

大数定律及中心极限定理

§5.1　Chebyshev（切比雪夫）不等式
§5.2　大数定律
§5.3　中心极限定理

5.1 Chebyshev（切比雪夫）不等式

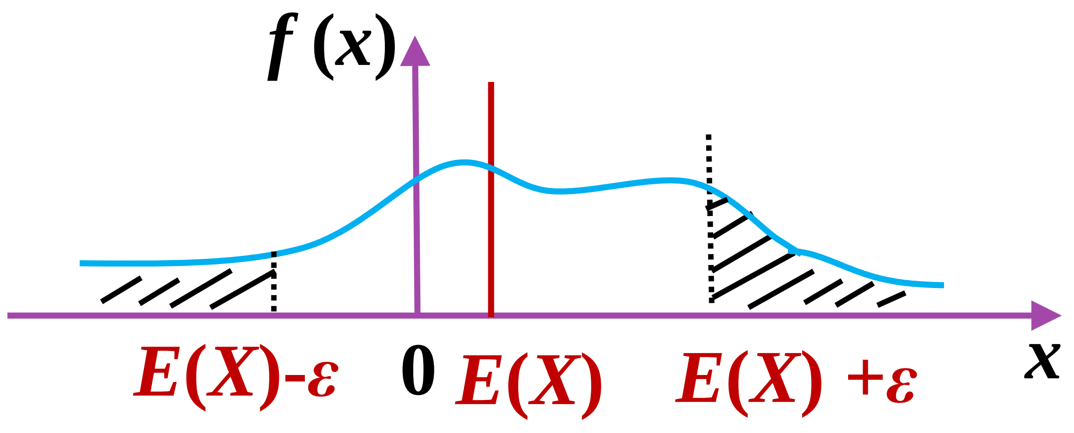
Chebyshev（切比雪夫）不等式

定理

设随机变量X的数学期望$E(X)$、方差$D(X)$存在，则对于任意$\varepsilon>0$，不等式

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

成立，称为**切比雪夫(*Chebyshev*)不等式**



证明

仅对连续型随机变量给予证明

设X的概率密度为$f(x)$，则有

$$P\left(\left|X-E(X)\right|\geq\varepsilon\right)=\int_{|x-E(X)|\geq\varepsilon}f(x)dx\leq\int_{|x-E(X)|\geq\varepsilon}\underbrace{\boxed{\frac{\left|x-E(X)\right|^2}{\varepsilon^2}}}_{\geq 1}f(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}\left[x-E(X)\right]^2f(x)dx=\frac{D(X)}{\varepsilon^2}$$

等价形式为 $$P\left(\left|X-E(X)\right|<\varepsilon\right)\geq 1-\frac{D(X)}{\varepsilon^2}$$

Chebyshev（切比雪夫）不等式

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

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

说明：

1. 注意两式中≤、≥准确写法，逆事件和不等式变形。
2. 在分布未知，而期望$E(X)$和方差$D(X)$已知的情况下，可以估计概率$P(|X-E(X)|\geq\varepsilon)$或$P(|X-E(X)|<\varepsilon)$的界限。
3. 可用于证明方差性质5。
4. 可用于证明大数定理。

例 已知正常男性成年人的血液中，平均每毫升含白细胞数是 7300，标准差是700，试估计每毫升男性成年人血液中含白细胞数在 5200 至 9400 之间的概率。

解 设每毫升男性成年人血液中所含白细胞数为X，则 $E(X)=7300$，$D(X)=700^2$，从而

$$P(5200<X<9400)=P(5200-7300<X-7300<9400-7300)$$

$$=P(-2100<X-7300<2100)=P(|X-7300|<2100)$$

由*Chebyshev*不等式得

$$P(5200<X<9400)=P(|X-7300|<2100)=P(|X-E(X)|<2100)\geq 1-\frac{D(X)}{2100^2}=1-\frac{700^2}{2100^2}=\frac{8}{9}$$

即每毫升男性成年人血液中所含白细胞数在5200至9400之间的概率不小于 $\frac{8}{9}$

例 在n重*Bernoulli*试验中，若已知每次试验事件A 出现的概率为0.75，试利用*Chebyshev*不等式估计n，使A出现的频率在0.74至0.76之间的概率不小于0.90。

解 *Bernoulli*试验中事件A 出现次数为X，则有$X\sim B(n,0.75)$

$$E(X)=np=0.75n,\quad D(X)=npq=0.1875n$$

$$f_n(A)=\frac{X}{n}$$

$$P\left(0.74<\frac{X}{n}<0.76\right)=P\left(|X-0.75n|<0.01n\right)\geq 1-\frac{0.1875n}{(0.01n)^2}=1-\frac{1875}{n}\geq 0.90$$

$$\Rightarrow n\geq 18750$$

*Chebyshev*不等式可在只知$E(X)$和$D(X)$ 前提下（分布未知）估计随机变量X落在均值左右ε界限内的概率

本节回顾

Chebyshev（切比雪夫）不等式

设随机变量X的数学期望$E(X)$、方差$D(X)$存在，则对于任意$\varepsilon>0$，以下不等式成立

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

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

1. 注意两式中≤、≥准确写法，逆事件和不等式变形。
2. 在分布未知，而期望$E(X)$和方差$D(X)$已知的情况下，可以估计概率$P(|X-E(X)|\geq\varepsilon)$或$P(|X-E(X)|<\varepsilon)$的界限。

CHAPTER 5

大数定律及中心极限定理

§5.1　Chebyshev（切比雪夫）不等式

§5.2　大数定律

§5.3　中心极限定理

5.2 大数定律

大数定律是叙述随机变量序列的前若干项的算数平均值在某种条件下收敛于这些项均值的算数平均值

定义1

设$X_1, X_2, \ldots$是随机变量序列，如果存在数列$a_1, a_2, \ldots$使得对任意的$\varepsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a_n\right| \geq \varepsilon\right) = 0$$

则称随机变量序列$\{X_i\}$服从大数定律。

定义2

设$X_1, X_2, \ldots$是随机变量序列，X是随机变量，若对任意的$\varepsilon>0$，有

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

则称随机变量序列$\{X_i\}$依概率收敛于X，记为 $X_n \xrightarrow{P} X, n \to \infty$

Chebyshev（切比雪夫）大数定律

设$X_1, X_2, \ldots$ 是相互独立的随机变量序列，具有相同的期望和方差，数学期望 $E(X_k)=\mu$，方差$D(X_k)=\sigma^2(k=1, 2, \ldots)$，作前$n$个变量的算数平均 $\frac{1}{n}\sum_{k=1}^{n} X_k$

则对于任意$\varepsilon>0$，有 $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k-\mu\right|\geq\varepsilon\right)=0$ 或 $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k-\mu\right|<\varepsilon\right)=1$

该表述形式是*Chebyshev*不等式的等价形式

也可记为依概率收敛于μ的形式，即 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu, \quad n\to\infty$

证明

在方差$D(X_k)=\sigma^2$，$k=1, 2, \ldots$存在的情况下证明

$$E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k)=\frac{1}{n}\cdot n\mu=\mu$$

由独立性 $D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}$

由Chebyshev不等式

$$0\le P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|\ge\varepsilon\right)=P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-E\left[\frac{1}{n}\sum_{k=1}^{n}X_k\right]\right|\ge\varepsilon\right)$$

$$\le\frac{D\left[\frac{1}{n}\sum_{k=1}^{n}X_k\right]}{\varepsilon^2}=\frac{\sigma^2/n}{\varepsilon^2}\to 0 \qquad n\to\infty$$

由此可证 $\lim\limits_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|\ge\varepsilon\right)=0$ ■

说明：$\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}$ 是一个随机事件

而当$n\to\infty$时事件的概率趋近于1

对于相互独立且具有相同均值μ和相同方差σ^2的随机变量$X_1, X_2, \ldots$，当n很大时它们算数平均 $\frac{1}{n}\sum_{k=1}^{n}X_k$ 很可能接近于μ

注意：不能简单理解为 $E(\frac{1}{n}\sum_{k=1}^{n}X_k)$ 接近为μ

要讨论的是一个随机变量，不是随机变量的期望值

*Chebyshev*大数定律（更一般的形式）

设$X_1, X_2, \ldots$ 是相互独立的随机变量序列，具有相同的期望，即数学期望$E(X_k)=\mu$(k =1, 2, …)，若存在常数C>0，使得$D(X_k)\le C$ (k =1, 2, …)

则对于任意ε>0，有 $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \ge \varepsilon\right) = 0$

其依概率收敛于μ的形式为 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu, \quad n\to\infty$

Markov（马尔可夫）大数定律

虽各项期望方差不同，但仍可用*Chebyshev*不等式证明

设$X_1, X_2, \ldots$ 是随机变量序列，且 $\lim\limits_{n\to\infty} \frac{1}{n^2} D\left[\sum_{k=1}^{n} X_k\right] = 0$

则对于任意ε>0，有 $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E(X_k)\right| \ge \varepsilon\right) = 0$ 即 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \frac{1}{n}\sum_{k=1}^{n} E(X_k),\ n\to\infty$

例 设$X_1, X_2, \dots$， X_n是相互独立的随机变量序列，且X_i (i=1, 2, ……)有如下分布律，

$$P(X_i=-ia)=\frac{1}{2i^2}, P(X_i=0)=1-\frac{1}{i^2}, P(X_i=ia)=\frac{1}{2i^2}$$

问$X_1, X_2, \dots$， X_n是否满足*Chebyshev*大数定律。

解 由题意知， $X_1, X_2, \dots$， X_n相互独立

$$E(X_i)=-ia\times\frac{1}{2i^2}+0\times(1-\frac{1}{i^2})+ia\times\frac{1}{2i^2}=0, i=1,2,\cdots$$

$$E(X_i^2)=(-ia)^2\times\frac{1}{2i^2}+0^2\times(1-\frac{1}{i^2})+(ia)^2\times\frac{1}{2i^2}=a^2, i=1,2,\cdots$$

$$D(X_i)=E(X_i^2)-\left[E(X_i)\right]^2=a^2, i=1,2,\cdots$$

故$X_1, X_2, \dots$， X_n具有相同的数学期望的相同的方差，且方差有界，所以$X_1, X_2, \dots$， X_n满足*Chebyshev*大数定律

以上大数定律说明在方差满足一定的条件下，随机变量序列服从大数定律。但在方差不存在时，随机变量序列是否服从大数定律呢? 由 *Khintchine* 大数定律回答这个问题。

Khintchine（辛钦）大数定律

设$X_1, X_2, \ldots$ 是**相互独立且服从同一分布**的随机变量序列，具有有限的数学期望，即数学期望 $E(X_k)=\mu(k=1, 2, \ldots)$，

则对于任意$\varepsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right) = 0$$

其依概率收敛于μ的形式为

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu, \qquad n\to\infty$$

注意：“同分布”未必要求每一项的分布参数一模一样

上述大数定律的成立条件不同，但是均表明了当n充分大时，随机变量序列的前n项的算数平均会收敛到**这些项期望的算术平均值**

Bernoulli大数定律是对第1章中提出的“频率稳定性”，给出理论上的论证

Bernoulli（伯努利）大数定律

设n_A是n次独立重复试验（n重Bernoulli试验）事件A发生的次数，p是事件A在每次试验中发生的概率，即$P(A)=p$，则对任意的$\varepsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right)=0 \quad \text{即} \quad \frac{n_A}{n}\xrightarrow{P} p,\qquad n\to\infty$$

证明 $n_A \sim B(n,p)$

$$E\left(\frac{n_A}{n}\right)=\frac{1}{n}E(n_A)=\frac{1}{n}\cdot np=p \qquad D\left(\frac{n_A}{n}\right)=\frac{1}{n^2}D(n_A)=\frac{1}{n^2}\cdot npq=\frac{p(1-p)}{n}$$

由Chebyshev不等式，对任意$\varepsilon>0$

$$P\left(\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right)\leq\frac{p(1-p)}{n\varepsilon^2} \qquad \lim_{n\to\infty} P\left(\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right)=0$$

也可直接用Chebyshev大数定律，$X_1, X_2, \ldots$ 是相互独立的随机变量，$E(X_i)=p,\ D(X_i)=p(1-p)$

$$n_A=\sum_{k=1}^{n}X_k \qquad \frac{n_A}{n}=\frac{1}{n}\sum_{k=1}^{n}X_k \qquad \lim_{n\to\infty} P\left(\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right)=\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-p\right|\geq\varepsilon\right)=0 \quad \blacksquare$$

说明：

1. *Bernoulli*大数定律建立了在大量重复独立试验中事件出现频率的稳定性，正因为这种稳定性，概率的概念才有客观意义。

2. *Bernoulli*大数定律还提供了通过试验来确定事件概率的方法，既然频率$f_n(A)$与概率p有较大偏差的可能性很小，便可通过多次试验确定某事件发生的频率并把它作为相应的概率估计。

3. 大数定律是参数估计的理论依据（采用大量的样本观测值估计分布中的未知参数是可行的）

本节回顾

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{P} \mu, \quad n\to\infty$$

- **Chebyshev大数定律**

设$X_1, X_2, \ldots$ 是**相互独立**的随机变量序列，具有相同的期望，即数学期望$E(X_k)=\mu$(k =1, 2, …)，若存在常数C>0，使得$D(X_k)\le C$ (k =1, 2, …)

- **Markov（马尔可夫）大数定律**

设$X_1, X_2, \ldots$ 是随机变量序列，且 $\lim\limits_{n\to\infty}\frac{1}{n^2}D\left[\sum_{k=1}^{n}X_k\right]=0$

- **Khintchine（辛钦）大数定律**

设$X_1, X_2, \ldots$ 是**相互独立且服从同一分布**的随机变量序列，具有有限的数学期望，即数学期望$E(X_k)=\mu$(k =1, 2, …)，

- **Bernoulli（伯努利）大数定律**

设n_A是n次独立重复试验（n重Bernoulli试验）事件A发生的次数，p是事件A在每次试验中发生的概率，即$P(A)=p$，则对任意的$\varepsilon>0$，有

$$\frac{n_A}{n} \xrightarrow{P} p, \quad n\to\infty$$