概率论与数理统计

第五章　大数定律与中心极限定理

本章要解决的问题

答复

1. 为何能以某事件发生的频率作为该事件的概率的估计？
2. 为何能以样本均值作为总体期望的估计？

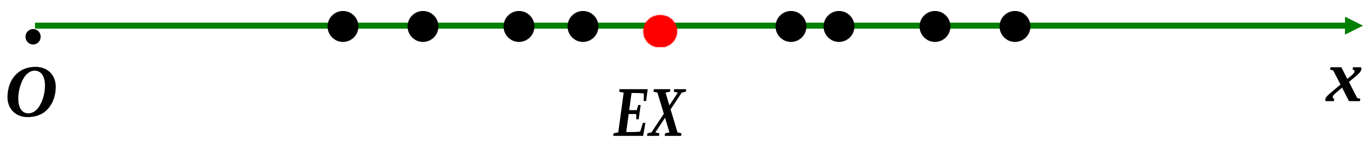
大数定律

3. 为何正态分布在概率论中占有极其重要的地位？
4. 大样本统计推断的理论基础是什么？

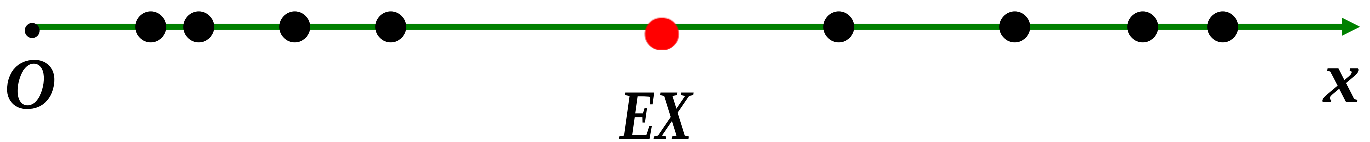
中心极限定理

第一节　大数定律

- 切比雪夫不等式
- 大数定理



DX越小，$P\{|X-EX|\geq\varepsilon\}$越小



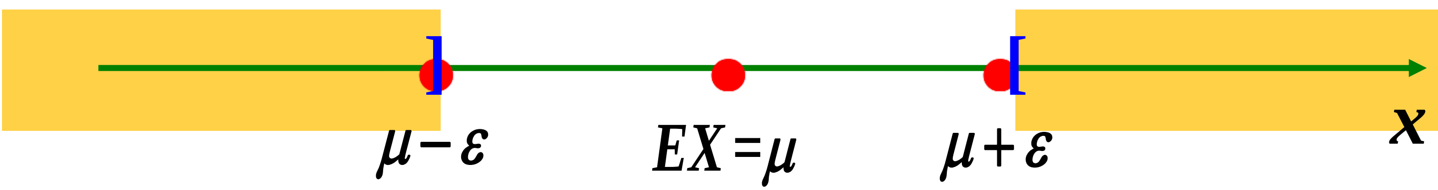
DX越大，$P\{|X-EX|\geq\varepsilon\}$越大

一、切比雪夫不等式

定理 设随机变量 X 的方差 DX 存在,则对 $\forall\ \varepsilon > 0$, 有

$$P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}$$

切比雪夫不等式



我们只就连续型随机变量的情况来证明.

证 设X的概率密度为$f(x)$,则有

$$P\{|X-\mu|\geq\varepsilon\}=\int_{|X-\mu|\geq\varepsilon} f(x)dx$$

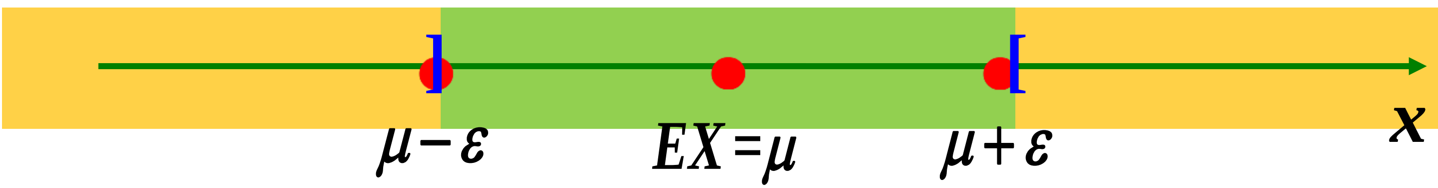
$$\leq\int_{|X-\mu|\geq\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)\mathrm{d}x$$

$$=\frac{DX}{\varepsilon^2}$$

注:

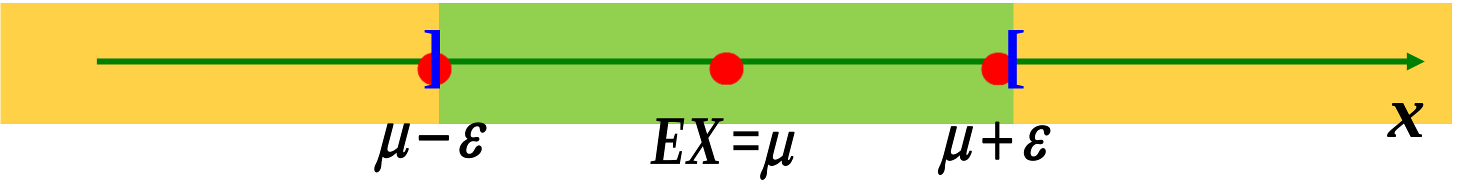
$$P\{|X-EX|\geq\varepsilon\}\leq\frac{DX}{\varepsilon^2}\qquad P\{|X-EX|<\varepsilon\}?$$



$$P\{|X-EX|\geq\varepsilon\}+P\{|X-EX|<\varepsilon\}=1$$

注:

$$P\{|X-EX|\geq\varepsilon\}\leq\frac{DX}{\varepsilon^2}\Leftrightarrow P\{|X-EX|<\varepsilon\}\geq 1-\frac{DX}{\varepsilon^2}$$



$$P\{|X-EX|\geq\varepsilon\}+P\{|X-EX|<\varepsilon\}=1$$

例题 $P\{|X-EX|\geq\varepsilon\}\leq\frac{DX}{\varepsilon^2}\Leftrightarrow P\{|X-EX|<\varepsilon\}\geq 1-\frac{DX}{\varepsilon^2}.$

1. $X\sim N(\mu,\sigma^2)$,根据切比雪夫不等式$P\{|X-\mu|<3\sigma\}$

$$P\{|X-EX|<\varepsilon\}\geq 1-\frac{DX}{\varepsilon^2}.$$

$$P\{|X-\mu|<3\sigma\}\geq 1-\frac{DX}{9\sigma^2}=1-\frac{\sigma^2}{9\sigma^2}=\frac{8}{9}\approx 0.89$$

2. 设$EX=11, DX=9$,则$P\{2<X<20\}\geq \dfrac{8}{9}$

$$P\{|X-EX|<\varepsilon\}\geq 1-\frac{DX}{\varepsilon^2}.$$

$$P\{2<X<20\}=P\{-9<X-11<9\}$$

$$=P\{|X-11|<9\}$$

$$\geq 1-\frac{DX}{9^2}=\frac{8}{9}$$

3. 设$EX=-2, DX=1, EY=2, DY=4, \rho_{XY}=-0.5$

根据切比雪夫不等式, 则$P\{|X+Y|\geq 6\}\leq$ $\frac{1}{12}$

$$P\{|X-EX|\geq \varepsilon\}\leq \frac{DX}{\varepsilon^2}$$

$$P\{|X+Y|\geq 6\}=P\{|X+Y-E(X+Y)|\geq 6\}\leq \frac{D(X+Y)}{\varepsilon^2}$$

$$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y)$$

$$=1+4+2\times(-0.5)\times 1\times 2=3$$

4. 练习 $X\sim N(1,1), Y\sim N(0,1), EXY=-0.1$, 根据切比雪夫不等式,则$P\{-4<X+Y<6\}\geq$ 0.816

二、依概率收敛的定义

定义 设$X_1, X_2, \cdots, X_n, \cdots$是随机变量序列,$a$ 是一个常数.若对$\forall \varepsilon > 0$,有

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

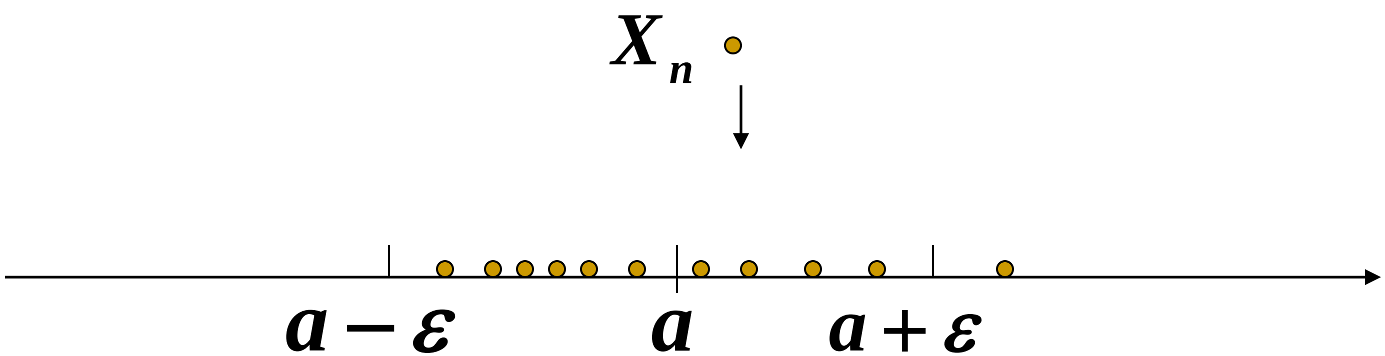
或 $$\lim_{n\to\infty} P\{|X_n - a| \geq \varepsilon\} = 0$$

则称序列$X_1, X_2, \cdots, X_n, \cdots$依概率收敛于$a$.

记为 $X_n \xrightarrow{P} a.$

$X_n \xrightarrow{P} a$ 意思是：当 $n \to \infty$ 时，X_n 落在

$(a-\varepsilon, a+\varepsilon)$ 内的概率越来越大.



而 $X_n \to a$ 意思是：$\forall \varepsilon > 0, \exists n_0$，当 $n > n_0$

$$|X_n - a| < \varepsilon$$

注意 $\{X_n\}$依概率收敛于a,意味着对任意给定的$\varepsilon > 0$,当n充分大时,事件$|X_n - X| < \varepsilon$的概率很大,接近于1;并不排除事件$|X_n - X| \ge \varepsilon$的发生,而只是说它发生的可能性很小.

依概率收敛比高等数学中的普通意义下的收敛弱些,它具有某种不确定性.

三、大数定理

定理5.2（贝努利大数定律）

在n重贝努利试验中成功的次数为Y_n，而每次试验成功的概率为p，则对$\forall \varepsilon > 0$，

$$\lim_{n\to\infty} P\{|\frac{Y_n}{n} - p| \geq \varepsilon\} = 0$$

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

$$\frac{Y_n}{n} \xrightarrow{P} p.$$

证　由于

$$E\left(\frac{Y_n}{n}\right)=\frac{1}{n}E(Y_n)=\frac{1}{n}\cdot np=p$$

$$D\left(\frac{Y_n}{n}\right)=\frac{1}{n^2}D(Y_n)=\frac{1}{n^2}\cdot npq=\frac{1}{n}pq$$

由切比雪夫不等式

$$P\left\{\left|\frac{Y_n}{n}-p\right|\geq\varepsilon\right\}\leq\frac{D\left(\frac{Y_n}{n}\right)}{\varepsilon^2}=\frac{pq}{n\varepsilon^2}$$

上式中令 $n\to\infty$ 得

$$\lim_{n\to\infty}P\{|\frac{Y_n}{n}-p|\geq\varepsilon\}=0$$

定理5.3（切比雪夫大数定理）

设随机变量$X_1, X_2, \cdots, X_n, \cdots$是相互独立的随机变量序列,若存在常数$C$,使得$DX_i < C (i = 1, 2, \ldots)$,

则对$\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i| \geq \varepsilon\} = 0$$

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

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} EX_i.$$

证 由于

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

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

由切比雪夫不等式

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}EX_i\right|\ge\varepsilon\right\}\le\frac{D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)}{\varepsilon^2}\le\frac{C}{n\varepsilon^2}$$

上式中令 $n\to\infty$ 得

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}EX_i|\ge\varepsilon\}=0$$

推论（切比雪夫大数定理的特殊情况）

设随机变量$X_1, X_2, \cdots, X_n, \cdots$是独立同分布的随机变量序列,具有有限的数学期望和方差,若$EX_i = \mu, DX_i = \sigma^2 (i = 1, 2, \ldots)$, 则对$\forall \varepsilon > 0$

$$\lim_{n \to \infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \varepsilon\} = 0$$

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

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

证 由于

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

$$D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}$$

由切比雪夫不等式

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|\geq\varepsilon\right\}\leq\frac{D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)}{\varepsilon^2}\leq\frac{\sigma^2}{n}$$

上式中令 $n\to\infty$ 得

$$\lim_{n\to\infty}P\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|\geq\varepsilon\}=0$$

说明

1、定理中$\{|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|<\varepsilon\}$是指一个随机事件)

当$n\to\infty$时,这个事件的概率趋于1.

2、定理以数学形式证明了随机变量$X_1,\cdots X_n$

的算术平均$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$接近数学期望$EX_i=\mu$

$(i=1,2\cdots,)$,这种接近说明其具有的稳定性.

这种稳定性的含义说明算术平均值是依概率收敛的意义下逼近某一常数.

下面给出的独立同分布下的大数定律，不要求随机变量的方差存在.

定理5.4（辛钦大数定律）


辛钦

设$X_1, X_2, \cdots, X_n, \cdots$是独立同分布随机变量序列，具有有限的数学期望$EX_i = \mu, (i = 1, 2, \ldots)$，则对$\forall \varepsilon > 0$

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \varepsilon\} = 0$$

$$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon\} = 1 \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

四、小结

大数定律

伯努利 $\lim_{n\to\infty} P\{|\frac{Y_n}{n}-p|<\varepsilon\}=1$ $Y_n \sim b(n,p)$

切比雪夫 $\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i-\mu|<\varepsilon\}=1$ $\begin{cases} E(X_i)=\mu \\ D(X_i)=\sigma^2 \end{cases}$

辛钦 $\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{i=1}^{n} X_i-\mu|<\varepsilon\}=1$ $E(X_i)=\mu$

大数定律以严格的数学形式表达了随机现象最根本的性质之一：**平均结果的稳定性**

大数定理

1. 设$X_1, X_2, \ldots, X_n \ldots$独立同分布，且$E(X_n) = 0$,

则$\lim\limits_{n\to\infty} P\{\sum\limits_{i=1}^{n} X_i < n\} = (\qquad)$

2. 设$X_1, X_2, \ldots, X_n \ldots$, 独立同服从参数为2的指数分布,

则当$n \to \infty$时,$Y_n = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2$ 依概率收敛于