浙江大学
ZHEJIANG UNIVERSITY

第36讲 大数定律

问题的提出：

上一讲中，提到的“频率的稳定值记为概率”，意味着

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

这个结论可以用“大数定律”来描述.

定理 1: (贝努里大数定律):

记n_A为n重贝努里试验中事件A发生的次数,

并记事件A在每次试验中发生的概率为p

$(0 < p < 1)$.则对于$\forall \varepsilon > 0$,有

$$\lim_{n \to +\infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0,$$

即 $\frac{n_A}{n} \xrightarrow{P} p$, 当$n \to +\infty$.

证明： 由于n_A为n重贝努里试验中事件A发生的次数，

故 $n_A \sim B(n,p)$， 那么

$E(n_A)=np,\ D(n_A)=np(1-p),$

根据切比雪夫不等式，则对于$\forall\varepsilon>0$,有

$$0 \le P\left\{\left|\frac{n_A}{n}-p\right| \ge \varepsilon\right\} = P\left\{\left|n_A-np\right| \ge n\varepsilon\right\}$$

$$\le \frac{np(1-p)}{n^2\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2} \to 0,\ \text{当}\ n \to +\infty.$$

从而 $\lim\limits_{n\to+\infty} P\left\{\left|\frac{n_A}{n}-p\right| \ge \varepsilon\right\}=0.$

贝努里大数定律的重要意义：

- 提供了用大量重复独立试验中事件出现频率的极限值来确定概率的理论依据，使得概率的概念才有严格的意义.
- 提供了通过试验来确定事件概率的方法 ——可以通过做试验确定某事件发生的频率并把它作为相应的概率估计．例如：想估计某产品的不合格品率p，可以随机抽取n（n较大）件，将n件产品的不合格品的比例作为p的估计.

大数定律 (Laws of Large Numbers)

内容：设$X_1, X_2, \cdots, X_n, \cdots$是一列随机变量,则在一定条件下,随机变量序列$Y_n = \dfrac{X_1 + \ldots + X_n}{n}$, 收敛到$\mu$, 当$n \to \infty$.

问题：

(1)随机变量序列Y_n收敛到μ的含义？ 依概率收敛

(2) μ是什么？ 当X_i期望相同时,$\mu = E(X_i)$

(3)一定条件是什么？

定理2(切比雪夫大数定律的推论):

$X_1, X_2, \cdots, X_n, \cdots$为相互独立的随机变量, 且具有相同的期望$\mu$,

相同的方差σ^2，那么 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$, 当 $n \to +\infty$.

证明: 记$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, 则$E(Y_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$, $D(Y_n) = \frac{1}{n^2} D(\sum_{i=1}^{n} X_i) = \frac{\sigma^2}{n}$.

对Y_n应用切比雪夫不等式, 得

$$0 \le P\{|Y_n - E(Y_n)| \ge \varepsilon\} \le \frac{D(Y_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0, \text{ 当} n \to +\infty.$$

例1: 设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立, 且它们的分布律为

$$P\{X_i=\sqrt{i}\}=P\{X_i=-\sqrt{i}\}=\frac{1}{2i},\ P\{X_i=0\}=1-\frac{1}{i},\ i=1,\ 2,\ \cdots.$$

请讨论$\frac{1}{n}\sum_{i=1}^{n}X_i$的收敛性

解: 由于对任意的$i\geq 1$, 有 $E(X_i)=0$,

$$D(X_i)=E(X_i^2)=0+(\sqrt{i})^2\cdot\frac{1}{2i}+(-\sqrt{i})^2\cdot\frac{1}{2i}=1,$$

所以$\{X_i, i\geq 1\}$相互独立, 期望、方差相同, 由定理2知

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu,\ 当\ n\to+\infty.$$

$D(X)$存在 $\Rightarrow E(X)$存在 　$D(X)$存在 $\Leftarrow E(X)$存在 

前面的定理要求随机变量的方差存在, 但当随机变量服从相同分布时，就不需要这一要求.

定理3(辛钦大数定律):

$X_1, X_2, \cdots, X_n, \cdots$为独立同分布的随机变量,

且其期望存在，记为μ，那么

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu, \text{当 } n \to +\infty.$$

辛钦大数定律的意义

- 提供了求随机变量X的数学期望$E(X)$的近似值的方法:

将随机变量X独立重复地观察n次, 记第k次观测值为 X_k,

则 $X_1, X_2, \cdots, X_n$ 相互独立, 且与X具有同样的分布.

那么, 当$E(X)$存在时, 由辛钦大数定律, 可知当n充分大时, 可将n次的平均 $\frac{1}{n}\sum_{i=1}^{n} X_i$ 作为$E(X)$的近似.

- 若目的是寻求X的期望，则这样做可以不必考虑X的分布！如可用浙大300个学生的平均身高作为整个浙大学生的平均身高的近似值!

例2: 设随机变量$X_1, X_2, \cdots, X_n, \cdots$, 相互独立同分布, $X_1 \sim U(-1,1)$. 则

(1) $\frac{1}{n}\sum_{i=1}^{n} X_i$, (2) $\frac{1}{n}\sum_{i=1}^{n} |X_i|$, (3) $\frac{1}{n}\sum_{k=1}^{n} X_i^2$ 分别依概率收敛吗?

如果依概率收敛, 分别收敛于什么? (当 $n \to +\infty$时)

大数定律的Excel模拟可以看实验9.

解: 由辛钦大数定律, $X_1, X_2, \cdots, X_n, \cdots$, 相互独立同分布, $E(X_1)$存在;

$|X_1|, |X_2|, \cdots, |X_n|, \cdots$, 相互独立同分布, $E(|X_1|)$存在;

$X_1^2, X_2^2, \cdots, X_n^2, \cdots$, 相互独立同分布, $E(X_1^2)$存在;

故 $\frac{1}{n}\sum_{i=1}^{n} X_i$, $\frac{1}{n}\sum_{i=1}^{n} |X_i|$, $\frac{1}{n}\sum_{i=1}^{n} X_i^2$均依概率收敛.

注意到 $X_1 \sim U(-1,1)$，那么 $E(X_1)=0$， 故 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} 0$,

同理， $E(|X_1|) = \int_{-1}^{1} |x| \frac{1}{2} dx = \frac{1}{2}$,

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} |X_i| \xrightarrow{P} \frac{1}{2},$$

$$E(X_1^2) = \int_{-1}^{1} x^2 \frac{1}{2} dx = \frac{1}{3},$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} \frac{1}{3}.$$

例3: 设随机变量$X_1, X_2, \cdots, X_n, \cdots$独立同分布，

$X_1 \sim U(0,1)$, 则$\sqrt[n]{X_1 X_2 \ldots X_n}$依概率收敛吗？

如果依概率收敛，收敛于什么？

分析： 不能直接使用大数定律，因为不是算术平均的形式.

回想： 关于依概率收敛，还有一个很好的性质

若$X_n \xrightarrow{P} a$, $f(x)$在点a连续，则

$f(X_n) \xrightarrow{P} f(a)$, 当$n \to \infty$时.

解： 记$Y_n = \sqrt[n]{X_1 \ldots X_n}$, 令$Z_n = \ln Y_n = \dfrac{1}{n}(\ln X_1 + \ldots + \ln X_n)$.

则$\ln X_1, \cdots, \ln X_n, \cdots$相互独立同分布，又

$$E(\ln X_1) = \int_0^1 \ln x dx = -1,$$

那么由辛钦大数定律知，故 $Z_n \xrightarrow{P} -1$, 当 $n \to +\infty$.

利用依概率收敛的性质，得

$$Y_n = e^{Z_n} \xrightarrow{P} e^{-1}, \text{当 } n \to +\infty.$$