大 数 定 律

引言

$\boxed{\text{大量试验次数下，频率具有稳定性}}$

$$\frac{n_A}{n} \to p = P(A), (n \to \infty)$$

依概率收敛

$\boxed{\text{伯努利大数定律}}$

$$\frac{n_A}{n} \to p = P(A), (n \to \infty)$$

在n重伯努利试验中，令

$$X_k = \begin{cases} 1, & 若事件A在第k次试验中发生 \\ 0, & 若事件A在第k次试验中不发生 \end{cases} \qquad k = 1, 2, \dots, n$$

$$\Longrightarrow \quad X_1, X_2, \dots, X_n : i.i.d. \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

$$\Longrightarrow \quad n_A = \sum_{k=1}^{n} X_k \qquad \frac{n_A}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k \to P(A) = p, \qquad (n \to \infty)$$

$$\frac{n_A}{n} = \frac{1}{n}\sum_{k=1}^{n} X_k \rightarrow p = P(A), (n \rightarrow \infty)$$

$$X_1, X_2, \ldots, X_n : i.i.d. \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

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

$$\Longrightarrow \frac{1}{n}\sum_{k=1}^{n} X_k \rightarrow p = \frac{1}{n}\sum_{k=1}^{n} E(X_k), \qquad (n \rightarrow \infty)$$

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to p = E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right), \qquad (n \to \infty)$$

$\{X_n: n = 1, 2, \dots\}$是一个随机变量序列

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to p = \frac{1}{n}\sum_{k=1}^{n} E(X_k)\ , \qquad (n \to \infty)$$

依概率收敛

大数定律

大 数 定 律

- 依概率收敛的定义

- 大数定律
  - 切比雪夫大数定律
  - 伯努利大数定律
  - 辛钦大数定律

依概率收敛的定义

设$\{X_n: n=1,2,\ldots\}$为随机变量序列，X为随机变量. 若对任意 $\varepsilon>0$，有

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

则称X_n依概率收敛于0，并记为$X_n \xrightarrow{P} 0$.

若$X_n - X \xrightarrow{P} 0$, 则称$X_n$依概率收敛于$X$，并记为$X_n \xrightarrow{P} X$.

依概率收敛的定义

$X_n \xrightarrow{P} X \quad \Longleftrightarrow \quad$ 若对任意 $\varepsilon > 0$，有

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

$\Longleftrightarrow$ 对任意 $\varepsilon > 0$，当n 充分大时，事件

$\{|X_n - X| > \varepsilon\}$ 发生的概率很小.

大 数 定 律

若随机变量序列$\{X_n: n = 1, 2, \ldots\}$满足

$$\lim_{n\to\infty} P\left(\left|\overset{\bar{X}}{\boxed{\frac{1}{n}\sum_{k=1}^{n} X_k}} - \overset{E(\bar{X})}{\boxed{\frac{1}{n}\sum_{k=1}^{n} E(X_k)}}\right| > \varepsilon\right) = 0$$

即 $$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \frac{1}{n}\sum_{k=1}^{n} E(X_k)$$

则称随机变量序列$\{X_n: n = 1, 2, \ldots\}$服从大数定律.

切比雪夫大数定律

设$\{X_n: n = 1, 2, \ldots\}$为独立随机变量序列. 若存在常数$C$, 使得

$D(X_n) \le C, n = 1, 2, \ldots,$ 则对任意 $\varepsilon > 0$，有

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

即 $$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \frac{1}{n}\sum_{k=1}^{n} E(X_k).$$

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

证明：令 $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$, 则 $E(\overline{X}) = E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} E(X_k)$

由于随机变量$X_1, \dots, X_n$独立，故

$$D(\overline{X}) = D\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} D(X_k)$$

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

由切比雪夫不等式，有

$$P\left(\left|\underbrace{\frac{1}{n}\sum_{k=1}^{n} X_k}_{\bar{X}} - \underbrace{\frac{1}{n}\sum_{k=1}^{n} E(X_k)}_{E(\bar{X})}\right| > \varepsilon\right) \leq \frac{\overbrace{\frac{1}{n^2}\sum_{k=1}^{n} D(X_k)}^{D(\bar{X})}}{\varepsilon^2}$$

$$\boxed{D(X_n) \leq C, n = 1, 2, \dots,} \qquad \leq \frac{C}{\varepsilon^2 n} \longrightarrow 0 \quad (n \to \infty)$$

切比雪夫大数定律

设$\{X_n: n = 1, 2, \dots\}$为独立同分布随机变量序列. $E(X_n) = \mu$,

$D(X_n) = \sigma^2, n = 1, 2, \dots,$ 则由切比雪夫大数定律有

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

即 $$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu.$$

大量试验次数下，算术平均值稳定性的解释.

若$\{X_n\}$独立同分布，$E(X_n)=\mu$，$D(X_n)=\sigma^2$，则 $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$
$n=1,2,\dots,$

在n重伯努利试验中，令

$$X_k=\begin{cases}1, & \text{若事件}A\text{在第}k\text{次试验中发生}\\ 0, & \text{若事件}A\text{在第}k\text{次试验中不发生}\end{cases}\qquad k=1,\dots,n$$

则$X_1,\dots,X_n$独立同分布，且 $E(X_k)=p$， $D(X_k)=p(1-p)$，

记 $n_A=\sum_{k=1}^{n} X_k$ 由切比雪夫大数定律有 $\frac{n_A}{n} \xrightarrow{P} p.$

伯努利大数定律

设n_A为n重伯努利试验中事件A发生的次数，

又A在每次试验中发生的概率为$p\ (0 < p < 1)$,

则对任意 $\varepsilon > 0$，有 $$\lim_{n\to\infty} P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = 1$$

大量试验次数下，频率稳定性的解释.

辛钦大数定律

设$\{X_n: n = 1, 2, \dots\}$为独立同分布随机变量序列. $E(X_n) = \mu$，则对任意 $\varepsilon > 0$，有

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

即 $$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu.$$

辛钦大数定律

例：设$\{X_n: n=1,2,\dots\}$为独立同分布随机变量序列, X_1服从幂律分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{2}{x^3}, & x\geq 1\\ 0, & x<1\end{cases}$$

试证明此随机变量序列 服从大数定律，并求$\frac{1}{n}\sum_{k=1}^{n}X_k$依概率收敛的极限.

证明： $E(X_1)=\int_1^{\infty}x\dfrac{2}{x^3}dx=2 \xrightarrow{\text{辛钦大数定律}} \dfrac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{P}E(X_1)=2$

$\int_1^{\infty}(x-2)^2\dfrac{2}{x^3}dx=\infty \longrightarrow D(X_1)$不存在, 不满足切比雪夫大数定律

小 结

$X_n \xrightarrow{P} X \iff$ 若对任意 $\varepsilon > 0$，有

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

大数定律
- 切比雪夫大数定律
- 伯努利大数定律
- 辛钦大数定律