第五章

张小玲　安徽大学物质科学与信息技术研究院

第五章

大数定律和中心极限定理

- 第一节 大数定律
- 第二节 中心极限定理

Chapter 5 Large Number Law and Central Limit Theorem

Law of Large Numbers (LLN)	Chebyshev’s LLN/weak LLN
	Bernoullis’s LLN
	Khinchin’s LLN
Central Limit Theorem (CLT)	Lindberger-Levy CLT
	De Moirve-Laplace CLT
	Liapunov CLT

Introduction

- The Law of Large Numbers shows broadly that the frequency and some average of many independent observations are **stable and predictable**.

- While the Central Limit Theorem says that the distribution of a sum or average of many small random quantities can be **approximated** by **Gaussian distribution .**

第一节　大数定律

一、相关定义

二、大数定律

5.1.1 相关定义

1. 定义

定义5.1.1（独立同分布）设随机变量$X_1,X_2,\cdots X_n$相互独立，且服从同一分布，则称$X_1, X_2,\cdots X_n$是独立同分布的随机变量序列.

定义5.1.2（依概率收敛）设$X_1,X_2,\cdots X_n$为随机变量序列，若存在一个随机变量X,对任意的$\varepsilon>0$，有

$$\lim_{n\to+\infty} P\{|X_n-X|\geq\varepsilon\}=0 \quad \text{或} \quad \lim_{n\to+\infty} P\{|X_n-X|<\varepsilon\}=1$$

则称随机变量序列 $X_1,X_2,\cdots X_n$依概率收敛于 X.

记为 $X_n \xrightarrow{P} X.$

5.1.1 相关定义

2. 依概率收敛的基本性质

1) 若$X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$,则$X_n \pm Y_n \xrightarrow{P} X \pm Y$;

2) 若$X_n \xrightarrow{P} X, k$为常数，则$kX_n \xrightarrow{P} kX$;

3) 若$X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y$,则$X_nY_n \xrightarrow{P} XY$;

4) 若$X_n \xrightarrow{P} X$, 且$g(x)$为$\mathbb{R}$上的连续函数，则$g(X_n) \xrightarrow{P} g(X)$;

Chebyshev's law of large numbers

Chebyshevs inequality

The Chebyshevs inequality provides a bound estimation on the probability of a random variable departing from its mean by a certain amount.

Theorem(Chebyshevs Inequality) If X is a random variable with finite mean μ and variance σ^2 , then for any value $\varepsilon > 0$,

$$P\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

or

$$\mathrm{P}\{|X-\mu| < \varepsilon\} \geq 1-\frac{\sigma^2}{\varepsilon^2}$$

Chebyshev's law of large numbers

Example

Suppose we flip a fair coin 1000 times. How likely it is to obtain between 40% and 60% heads?

Solution Let the r.v. $Xi(i = 1; 2; \cdots; 1000)$ be 1 if the coin lands on head and 0 otherwise, i.e.

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ outcome is head} \\ 0 & \text{if } i^{th} \text{ outcome is tail} \end{cases}$$

then, $\mu = E\,(Xi) = 0.5$,

$\sigma^2 = Var\,(Xi) = 1/4$;

Chebyshev's law of large numbers

By Chebyshevs inequality, the probability to obtain between 40% and 60% heads can be estimated by :

$$\mathrm{P}\left\{1000 \times 40\% < X_1 + X_2 + \cdots + X_{1000} < 1000 \times 60\%\right\}$$

$$= \mathrm{P}\left\{400 < X_1 + X_2 + \cdots + X_{1000} < 600\right\}$$

$$= \mathrm{P}\left\{\frac{4}{10} < \frac{X_1 + X_2 + \cdots + X_{1000}}{1000} < \frac{6}{10}\right\}$$

$$= \mathrm{P}\left\{\left|\frac{X_1 + X_2 + \cdots + X_{1000}}{1000} - \frac{1}{2}\right| < \frac{1}{10}\right\}$$

$$\geq 1 - \frac{1/4}{1000(1/10)^2} = 0.975$$

Chebyshev's law of large numbers

Notes

As Chebyshevs inequality is valid for all r.v. X with finite mean and variance, we cannot expect the estimated bound on the probability will be very close to the actual probability in most cases. That means Chebyshevs inequality usually gives **very loose bounds** on probability.

Example

Consider a normal random variable with mean μ and variance σ^2 , the following table shows the probability bound by Chebyshevs inequality and the true value.

Probability	Chebyshev's bound	$X \sim N(\mu, \sigma^2)$
$P(\lvert X-\mu\rvert \geq \sigma)$	≤ 1	0.3173
$P(\lvert X-\mu\rvert \geq 2\sigma)$	$\leq 1/2^2 = 0.25$	0.0465
$P(\lvert X-\mu\rvert \geq 3\sigma)$	$\leq 1/3^2 = 0.11$	0.00270
$P(\lvert X-\mu\rvert \geq 4\sigma)$	$\leq 1/4^2 \approx 0.06$	0.000063

5.1.1 相关定义

3. 大数定律

定义5.1.3（大数定律）设$X_1, X_2, \cdots X_n$为随机变量序列，$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ 若存在常数序列 a_n, 使对于任意给定的 $\varepsilon>0$,

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

则称随机变量序列 X_1, X_2, …, X_n, … 服从大数定律.

迄今为止,人们已发现很多大数定律. 所谓大数定律,简单地说， 就是大量数目的随机变量所呈现出的规律,这种规律一般用随机变量序列的某种收敛性来刻画.

4. 切比雪夫大数定律

定理5. 1. 1 设随机变量X_1，X_2，…，X_n，… 相互独立，且方差具有共同的上界, 即$D(X_i)\leqslant C$，$i=1$，2，… . 则对于任意给定的 $\varepsilon>0$，有

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

或

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

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

5.1.2 大数定律

切比雪夫大数定律表明：独立随机变量序列$\{X_n\}$，如果方差有共同的上界，当n充分大时，随机变量序列$\{X_n\}$的前n项和的算数平均$\overline{X_n}=\frac{1}{n}\sum_{i=1}^{n}X_i$以很大概率逼近于它的数学期望$\frac{1}{n}\sum_{i=1}^{n}E(X_i)$,或者说$\overline{X_n}=\frac{1}{n}\sum_{i=1}^{n}X_i$差不多不再是随机的了，取值接近于其数学期望的概率接近于1.

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

切比雪夫大数定律给出了平均值稳定性的科学描述

5.1.2 大数定律

6. 伯努利大数定律

定理5.1.3 设n_A是n次重复独立重复试验中事件A发生的次数，p是事件A在每次试验中发生的概率，则对任意的$\varepsilon > 0$，

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

证明 设$X_i = \begin{cases} 1, & \text{在第}i\text{次试验中事件}A\text{发生,} \\ 0, & \text{在第}i\text{次试验中事件}A\text{不发生} \end{cases}$ $i = 1, 2, \ldots, n.$

则有$n_A = X_1 + X_2 + \cdots + X_n$， $\dfrac{n_A}{n} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

显然随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，且都服从参数为$p$的0-1分布，

$$\frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_1) = p, \quad D(X_1) = p(1-p) \quad i=1, 2, \cdots, n$$

5.1.2 大数定律

由定理5.1.1（切比雪夫大数定律）得，

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

因为 $\dfrac{n_A}{n} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ $\quad \dfrac{1}{n}\sum\limits_{i=1}^{n} E(X_i) = E(X_1) = p,$

所以 $\lim\limits_{n\to+\infty} P\left\{\left|\dfrac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0$ （伯努利大数定律）

说明：伯努利大数定律从理论上证明了频率的稳定性，即事件A发生的频率$\frac{n_A}{n}$总在它的概率p的附近摆动，当试验次数n充分大时,就可以利用事件发生的频率$\frac{n_A}{n}$来近似地代替事件的概率p.

5.1.2 大数定律

7. 辛钦大数定律

定理5.1.4 设$\{X_n\}$是独立同分布的随机变量序列，且$E(X_i)=\mu$.则对任意的$\varepsilon>0$，

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

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

由辛钦大数定律可知：如果$\{X_n\}$是独立同分布的随机变量序列，且$E\left(\left|X_i\right|^k\right)$存在，其中$k$为正整数，则$\{X_n^k\}$服从大数定律，即对任意的$\varepsilon>0$，

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

5.1.2 大数定律

关于辛钦大数定律的说明：

(1) 辛钦大数定律与切比雪夫大数定律的推论一样都阐明了大量随机试验的平均结果具有稳定性, 但是辛钦大数定律去掉了随机变量序列方差存在的条件(在独立同分布场合，不需要该条件).

(2) 伯努利定律是辛钦定律的特殊情况.

小结

三个大数定理
- 契比雪夫大数定理
- 伯努利大数定理
- 辛钦定理

频率的稳定性是概率定义的客观基础， 而伯努利大数定理以严密的数学形式论证了频率的稳定性.