第五章 大数定律及中心极限定理

• 中山大学 • 计算机学院 • 郑培嘉 • Email：zhpj@mail.sysu.edu.cn

目录

1. 大数定律
2. 中心极限定理

1. 大数定律

大数定律

随机事件A的频率$f_n(A)$当重复试验的次数n增大时总呈现出稳定性，稳定在某一个常数的附近。频率的稳定性是概率定义的客观基础。本节我们将对频率的稳定性作出理论的说明。

弱大数定理（辛钦大数定理）

设$X_1, X_2, \ldots$是相互独立，服从同一分布的随机变量序列，且具有数学期望$E(X_k) = \mu \ (k = 1,2,\ldots)$。作前$n$个变量的算术平均$\frac{1}{n}\sum_{k=1}^{n} X_k$，则对于任意$\varepsilon > 0$，有

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

大数定律

证： 我们只在随机变量的方差$D(X_k)=\sigma^2(k=1,2,\cdots)$存在这一条件下证明上述结果。因为

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

又由独立性得

$$D(\frac{1}{n}\sum_{k=1}^{n}X_k)=\frac{1}{n^2}D(\sum_{k=1}^{n}X_k)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{1}{n^2}(n\sigma^2)=\frac{\sigma^2}{n}$$

由切比雪夫不等式得

$$1\geq P\{|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu|<\varepsilon\}\geq 1-\frac{\sigma^2/n}{\varepsilon^2}$$

上式中令$n\rightarrow\infty$，即得$\lim_{n\rightarrow\infty}P\{|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu|<\varepsilon\}=1$

大数定律

$\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\}$是一个随机事件。上式表明，当$n \to \infty$时这个事件的概率趋于1。即对于任意正数$\varepsilon$，当$n$充分大时，

不等式 $\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon$ 成立的概率很大。

通俗地说，辛钦大数定理是说，对于独立同分布且具有均值μ的随机变量$X_1, \ldots, X_n$，当n很大时它们的算术平均$\frac{1}{n}\sum_{k=1}^{n} X_k$很可能接近于$\mu$。

大数定律

设$Y_1, Y_2, \dots, Y_n, \dots$是一个随机变量序列，$a$是一个常数。若对于任意正数$\varepsilon$，有

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

则称序列$Y_1, Y_2, \dots, Y_n, \dots$**依概率收敛于$\boldsymbol{a}$**，记为

$$Y_n \xrightarrow{P} a$$

依概率收敛的序列有以下的性质

设$X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，又设函数$g(x,y)$在点(a,b)连续，则

$$g(X_n, Y_n) \xrightarrow{P} g(a,b)$$（证略）

大数定律

上述定理又可叙述为：

弱大数定理（辛钦大数定理）

设随机变量$X_1, X_2, \ldots, X_n, \ldots$相互独立，服从同一分布且具有数学期望$E(X_k) = \mu\ (k = 1,2, \ldots)$，则序列$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$依概率收敛于$\mu$，即$\overline{X} \xrightarrow{P} \mu.$

大数定律

伯努利大数定理（辛钦大数定理推论）

设f_A是n次独立重复实验中事件A发生的次数，p是事件A在每次试验中发生的概率，则对于任意正数$\varepsilon > 0$, 有

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

或

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

大数定律

证：因为$f_A \sim b(n,p)$，由第四章§2例6，有

$$f_A = X_1 + X_2 + \cdots + X_n$$

其中$X_1, X_2, \dots, X_n$相互独立，且都服从以p为参数的$(0-1)$分布，因而$E(X_k) = p(k = 1,2, \dots, n)$。

由辛钦大数定理即得

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

即

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

大数定律

伯努利大数定理的结果表明，对于任意$\varepsilon > 0$，只要重复独立试验次数n充分大，事件

$$\{|\frac{f_A}{n} - p| < \varepsilon\}$$

实际上几乎是必定要发生的，这就是我们所说的**频率稳定性**的真正含义。由实际推断原理，在实际应用中，当试验次数很大时，便可以用事件的频率来代替事件的概率。