大数定律及中心极限定理

October 5, 2016

1 大数定律

1.1 契比雪夫定理的特殊情况

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，且具有相同的数学期望和方差：$E(X_k) = \mu, D(X_k) = \sigma^2 (k = 1, 2, \cdots)$。作前 n 个随机变量的算术平均

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k, \tag{1}$$

则对于任意正数 ε，有

$$\lim_{n\to\infty} P\{|\bar{X} - \mu| < \varepsilon\} = \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1 \tag{2}$$

设 $Y_1, Y_2, \cdots, Y_n, \cdots$ 是一个随机变量序列，a 是一个常数。若对任意正数 ε，有

$$\lim_{n\to\infty} P\{|Y_n - a| < \varepsilon\} = 1, \tag{3}$$

称序列 $Y_1, Y_2, \cdots, Y_n, \cdots$依概率收敛于$a$。记为

$$Y_n \xrightarrow{P} a. \tag{4}$$

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

$$g(X_n, Y_n) \xrightarrow{P} g(a,b). \tag{5}$$

定理一：

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，且具有相同的数学期望和方差：$E(X_k) = \mu, D(X_k) = \sigma^2 (k = 1, 2, \cdots)$。则序列

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \tag{6}$$

依概率收敛于 μ，即 $\bar{X} \xrightarrow{P} \mu$。

1.2 伯努利大数定理

设 n_A 是 n 次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，则对于任意正数 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \tag{7}$$

或

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| \geqslant \varepsilon\right\} = 0 \tag{8}$$

伯努利大数定理表明事件发生的频率 n_A/n 依概率收敛于事件的概率 p。频率的稳定性；当试验次数很大时，可以用事件发生的频率来代替事件的概率；

1.3 辛钦定理

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，且具有数学期望 $E(X_k) = \mu (k = 1, 2, \cdots)$，则对于任意正数 ε，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1 \tag{9}$$

2 中心极限定理

2.1 独立同分布的中心极限定理

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，且具有数学期望和方差：$E(X_k)=\mu, D(X_k)=\sigma^2>0(k=1,2,\cdots)$，则随机变量之和 $\sum\limits_{k=1}^{n} X_k$ 的标准化变量：

$$Y_n=\frac{\sum\limits_{k=1}^{n} X_k-E\left(\sum\limits_{k=1}^{n} X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} X_k\right)}}=\frac{\sum\limits_{k=1}^{n} X_k-n\mu}{\sqrt{n}\sigma} \tag{10}$$

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\begin{aligned}\lim_{n\to\infty} F_n(x) &= \lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}\\ &= \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\mathrm{d}t=\Phi(x)\end{aligned} \tag{11}$$

2.2 李雅普诺夫 (Liapunov) 定理

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，它们具有数学期望和方差：

$$E(X_k) = \mu_k, D(X_k)=\sigma_k^2>0\ k=1,2,\cdots$$

记

$$B_n^2=\sum_{k=1}^{n}\sigma_k^2$$

若存在正数 δ，使得当 $n\to\infty$ 时,

$$\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n} E\{|X_k-\mu_k|^{2+\delta}\}\to 0, \tag{12}$$

则随机变量之和 $\sum\limits_{k=1}^{n} X_k$ 的标准化变量：

$$Z_n = \frac{\sum\limits_{k=1}^{n} X_k - E\left(\sum\limits_{k=1}^{n} X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n} X_k\right)}} = \frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} \mu_k}{B_n} \tag{13}$$

的分布函数 $F_n(x)$ 对于任意 x，满足

$$\begin{aligned} \lim_{n\to\infty} F_n(x) &= \lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} \mu_k}{B_n} \leqslant x\right\} \\ &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}t = \Phi(x) \end{aligned} \tag{14}$$

在定理的条件下，随机变量

$$Z_n = \frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} \mu_k}{B_n} \tag{15}$$

当 n 很大时，近似地服从正态分布 $N(0,1)$。由此，当 n 很大时，$\sum\limits_{k=1}^{n} X_k = B_n Z_n + \sum\limits_{k=1}^{n} \mu_k$ 近似地服从正态分布 $N\left(\sum\limits_{k=1}^{n} \mu_k, B_n^2\right)$。

无论各个随机变量 $X_k(k=1,2,\cdots)$ 服从什么分布，只有满足定理的条件，那么它们的和 $\sum\limits_{k=1}^{n} X_k$ 当 n 很大时，就近似地服从正态分布。

2.3 棣莫弗—拉普拉斯 (De Moivre—Laplace) 定理

设随机变量 $\eta_n(n=1,2,\cdots)$ 服从参数为 $n,p(0<p<1)$ 的二项分布，则对于任意 x，有

$$\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leqslant x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \mathrm{d}t = \Phi(x) \tag{16}$$