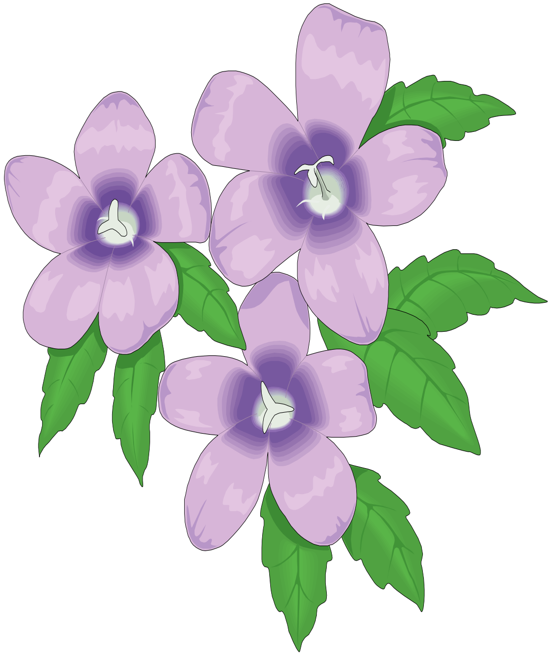
第五章 随机变量序列的极限



§5.1 大数定律

定义5.1　　设$X_1, X_2, \cdots$是一个随机变量序列。如果存在一个常数c，使得，对任意一个**ε**>0，总有 $\lim_{n\to\infty} P(|X_n - c| < \varepsilon) = 1$，那么，称随机变量序列 $X_1, X_2, \cdots$依概率收敛于c，记作$X_n \xrightarrow{P} c$ 。

$$\text{即对}\forall \varepsilon > 0, P(|X_n - c| \geq \varepsilon) \to 0, n \to \infty$$

$\varepsilon > 0$,当$n \to \infty$时，n重贝努利试验中，事件A发生的频数

$$N_A = X_1 + X_2 + \cdots + X_n,\text{频率} f_n(A) = \frac{N_A}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\frac{N_A}{n} \xrightarrow{P} p,\text{即}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} E(X_i), \quad \frac{1}{n}\sum_{i=1}^{n} X_{i*} \xrightarrow{P} 0.$$

- 依概率收敛性具有下列性质：

定理5.1 如果 $X_n \xrightarrow{P} c$，$Y_n \xrightarrow{P} b$，且函数g(x,y)在(a,b)处连续，那么

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

定理5.2（切比雪夫大数定律）

设 $X_1, X_2, ..$ 是两两不相关的随机变量序列。如果存在常数c，使得

$D(X_i) \le c, i=1,2,\cdots$ 那么，$\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_{i*} \xrightarrow{P} 0$

定理5.3（独立同分布情形下大数定律）

设 $X_1, X_2, ..$ 是一个独立同分布的随机变量序列，且 $E(X_i) \triangleq \mu$, $D(X_i) \triangleq \sigma^2$

（方差不存在也可，辛钦大数定律），i=1，2，...。那么，$\bar{X} \xrightarrow{P} \mu$

定理5.4（贝努利大数定律）

设 $X_1, X_2, \ldots$ 是一个独立同分布的随机变量序列，且每一个 X_i

都服从0–1分布B(1,p)，那么，$\bar{X} \xrightarrow{P} p$ 。

贝努利大数定律是定理5.3 的特例，其中，**μ**=p。

例5.1 设$X_1, X_2, \ldots$是独立同分布的随机变量序列，且 ，

$E(X_i)=\mu \quad D(X_i)=\sigma^2, i=1,2,\cdots$ 那么

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} \sigma^2+\mu^2$$

证

$$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)=\frac{1}{n}\sum_{i=1}^{n} E(X_i^2)$$

$$=\frac{1}{n}\sum_{i=1}^{n}(\sigma^2+\mu^2)=\sigma^2+\mu^2$$

§5.2 中心极限定理

定理5.5**(独立同分布情形下的中心极限定理)**

- 设 $X_1, X_2, \ldots$ 是一个独立同分布的随机变量序列，且

$E(X_i) \triangleq \mu, D(X_i) \triangleq \sigma^2, i = 1, 2, \cdots$ 则对任意一个x， $-\infty < x < \infty$

总有

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \le x \right) = \Phi(x)$$

其中，$\Phi(x)$是N(0,1)的分布函数。

- 定理5.5也称为列维–林德伯格（Levy–Lindberg）中心极限定理。
- 不论原来服从什么样的分布，只要条件满足，和以标准正态分布为其极限分布，即当**n**足够大时，可近似认为 $(\sum_{i=1}^{n} X_i) \sim N(n\mu, n\sigma^2)$

例5.2　某人要测量甲乙两地之间的距离，限于测量工具，他分成1200段来测量，每段测量误差（单位：cm）服从区间（-0.5,0.5）上的均匀分布，且相互独立，试求总距离误差的绝对值超过20厘米的概率。

例　为了测定一台机床的质量，把它分解成75个部件来测量。
假定每个部件的称量误差（单位：kg）
服从区间（－1,1）的均匀分布，且每个部件的称量误差相互独立
，试求机床重量的总误差的绝对值不超过10kg的概率。

定理5.6(德莫弗–拉普拉斯（De Moivre–Laplace）中心极限定理)

- 设 $X_1, X_2, ...$ 是一个独立同分布的随机变量序列，且每一个 X_i 都服从0–1分布B(1,p)，则对任意一个x，-∞<x<∞，总有

例5.3　一本20万字的长篇小说进行排版，假定每个字被排错的概率为 10^{-5}，试求这本小说出版后发现6个以上错字的概率，假定每个字是否被错排是相互独立的。

例5.4　某厂知道自己的产品不合格率p较高，因此，打算在每盒（100只）中多装几只产品，假定p＝0.2，试问每盒至少应多装几只产品才能保证顾客不吃亏的概率至少有99%?