第六章　极限定理

随机现象蕴涵着某种统计规律性。它主要有两个方面：一是随着重复次数n的增加，事件发生的频率与事件发生的概率非常接近,n趋于无穷大时，频率会“收敛”于概率，就是大数定律所要表述的问题；二是一种随机现象如果是由很多不确定因素累加导致成的，如果每种因素之间不相互影响，而且每种因素表现不是特别突出，则这种随机现象近似服从正态分布，此即中心极限定理。

§6.1 大数定律

切比雪夫不等式(Chebyshev)不等式

定理6.1（Chebyshev不等式） 设随机变量X的数学期望和方差均存在，则对$\forall \varepsilon > 0$,

$$P(|X-E(X)|\ge \varepsilon)\le \frac{D(X)}{\varepsilon^2} \quad or \quad P(|X-E(X)|< \varepsilon)\ge 1-\frac{D(X)}{\varepsilon^2}$$

证明：此处就离散型随机变量的情形加以证明。

设X的分布律为 $P(X=x_k)=p_k, k=1,2,\cdots$，

$$P(|X-E(X)|\ge \varepsilon)=\sum_{|x_k-E(X)|\ge\varepsilon} p_k \le \sum_{|x_k-E(X)|\ge\varepsilon} \frac{|x_k-E(X)|^2}{\varepsilon^2} p_k$$

$$\le \frac{1}{\varepsilon^2}\sum_k p_k|x_k-E(X)|^2=\frac{D(X)}{\varepsilon^2}。$$

以上定理说明：当X的方差D(X)越来越小时，事件发生的概率也越来越小，即X的取值偏离E(X)的程度越来越小。 这正说明，方差是描述随机变量与其期望值偏离程度的量。我们可以用这个不等式来估计这样的事件发生的概率，尽管我们连它的分布都不知道。

例：一机床制造长度为 50cm的工件，由于随机扰动，工件长度有误差。统计表明长度的均方差为2.5mm。若工件的实际长度在（49.25,50.75）cm之间算合格，试估计该机床制造工件的合格率。

解：以X表工件的长度，由于X的分布情况不明，但由题意知 $E(X)=50$, $\sqrt{D(X)}=2.5$，故用Chebyshev不等式来估计。因 $|X-50|=|X-E(X)|<0.75 \Leftrightarrow$ 机床合格，所以所求概率为

$$P(|X-50|<0.75)\geq 1-\frac{D(X)}{0.75^2}=1-\frac{0.25^2}{0.75^2}=\frac{8}{9}。$$

由以上例题及定理知，若X为随机变量，其数学期望及方差为 $E(X)=u, D(X)$ 由Chebyshev不等式有

$$P(|X-u|<k\sigma)\geq 1-\frac{1}{k^2},\ k>0.$$

由上面公式可以估计一些概率，但美中不足的是估计不太准确，因为它需要和应用的信息太少，这也是用该公式简便所付出的代价。

大数定律

在前面说过，随着重复次数的增加，频率 $f_n(A)$ 就越来越接近其概率。那么，如何用数学语言来描述这里的“接近” 呢?

定义6.1 设 $\{X_n\}$ 为一随机变量序列，a 为常数，若对任意的 $\varepsilon > 0$ ，有

$$\lim_{n\to\infty} P(|X_n - a| < \varepsilon) = 1 \quad or \quad \lim_{n\to\infty} P(|X_n - a| \geq \varepsilon) = 0$$

称 $\{X_n\}$ 依概率收敛于 a，记为 $X_n \xrightarrow{P} a.$

切比雪夫不等式的应用

定理6.2 设随机变量$\{X_n\}, n=1,2,\ldots,$ 中，$E(X_n)=\mu_n, D(X_n)=\sigma_n^2$存在，且$n\to\infty$时，有$\sigma_n^2\to 0$，则 $X_n-\mu_n \xrightarrow{P} 0$

证：由切比雪夫不等式

$$1\geq P(|X_n-\mu_n|<\varepsilon)\geq 1-\frac{\sigma_n^2}{\varepsilon^2}$$

两边取极限即可

大数定律描述的内容：

设$\{X_k\}$为随机变量序列，$\overline{X}_n$ @ $\frac{1}{n}\sum_{k=1}^{n} X_k$，描述$\overline{X}_n$ - E($\overline{X}_n$) $\xrightarrow{P} 0$，

即 $\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E(X_k) \xrightarrow{P} 0$的这一类定理称为大数定理。

定理6.3 Chebyshev大数定律：设$\{X_n\}$为一相互独立的随机变量序列，其数学期望 及方差 均存在，且存在常数 使得c $D(X_i) \le c, i = 1,$ 如果记

$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$， 则 $\overline{X}_n \xrightarrow{P} E(\overline{X}_n)$ 。即对$\forall \varepsilon>0$，

$$\lim_{n\to\infty} P(|\overline{X}_n - E(\overline{X}_n)| \ge \varepsilon) = 0 \text{ 或 } \lim_{n\to\infty} P(|\overline{X}_n - E(\overline{X}_n)| < \varepsilon) = 1。$$

证明：由Chebyshev不等式知 $0\le\lim\limits_{n\to\infty}P(|\overline{X}_n-E(\overline{X}_n)|\ge\varepsilon)$

$$\le\lim_{n\to\infty}\frac{D(\overline{X}_n)}{\varepsilon^2}=\lim_{n\to\infty}\frac{D(\frac{1}{n}\sum_{i=1}^{n}X_i)}{\varepsilon^2}$$

$$=\lim_{n\to\infty}\frac{\frac{1}{n^2}D(\sum_{i=1}^{n}X_i)}{\varepsilon^2}\le\lim_{n\to\infty}\frac{1}{n^2\varepsilon^2}\sum_{i=1}^{n}C=\lim_{n\to\infty}\frac{C}{n\varepsilon^2}=0,$$

$i.e.\quad \overline{X}_n\xrightarrow{P}E(\overline{X}_n)$。

由Chebyshev大数定律，我们有如下推论：

推论2（**独立同分布大数定律**）：设$\{X_n\}$为一相互独立的随机变量序列，且 $E(X_i)=u, D(X_i)=\sigma^2(i=1,2,\cdots)$，则 $\overline{X}_n=\frac{1}{n}\sum\limits_{i=1}^{n}X_i\xrightarrow{P}u$

推论2 Bernoulli大数定律：设 n_A 为n次Bernoulli事件中A发生的次数，p是在一次试验中事件A发生的概率，记 $f_n(A)=\dfrac{n_A}{n}$，则 $f_n(A)\xrightarrow{P} p$， i.e.

$$\lim_{n\to\infty} P(|f_n(A)-p|<\varepsilon)=1, \forall\varepsilon>0。$$

证明令 $X_i=\begin{cases}1, & \text{在第i次里事件A发生}\\ 0, & \text{否则}\end{cases}$，则 $X_i \sim B(1,p), E(X_i)=p$,

$i=1,2,\cdots$， 且 $X_1, X_2,\cdots, X_n,\cdots$ 相互独立，则 $f_n(A)=\dfrac{1}{n}\sum_{i=1}^{n} X_i=\overline{X}_n$,

于是由独立同分布大数定律知，

$$\frac{1}{n}\sum_{i=1}^{n} X_i = f_n(A)\xrightarrow{P} p = E(X_i)$$

例：设$\{X_n\}$是独立且同分布的随机变量序列，如果记

$\overline{X}_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$，且 $X_k \sim U(4,6)(k=1,2,\cdots)$，则 $\overline{X}_n$在 $n\to\infty$

时依概率收敛于何值？

解　显然 $\{X_n\}$满足独立同分布大数定律且$E(X_k)=5$

$k=1,2,\cdots$，故 $\overline{X}_n \xrightarrow{P} 5$, $i.e. \lim\limits_{n\to\infty} P(|X_n-5|<\varepsilon)=1.$

§6.2 中心极限定理

本节里，总设$\{X_k\}$为独立的随机变量序列。如果记$S_n = \sum_{k=1}^{n} X_k$，我们将看到，在一定条件下，S_n的极限分布总是正态分布，这就是中心极限定理所表达的主要内容。

在这一节里，我们主要讨论两个中心极限定理：**Levi-Lindberg**独立同分布中心极限定理 和 （二项分布以正态分布为极限的）**De. Movire-Laplace**中心极限定理。

定理6.4 Levi-Lindberg独立同分布中心极限定理 设$\{X_k\}$为独立同分布的随机变量序列$E(X_k)=u, D(X_k)=\sigma^2\neq 0, k=1,2,$

$$\cdots, S_n=\sum_{k=1}^{n}X_k, Y_n=\frac{\sum_{k=1}^{n}X_k-nu}{\sqrt{n}\sigma}, \overline{X}_n=\frac{1}{n}\sum_{k=1}^{n}X_k, \overline{Y}_n=\frac{\overline{X}_n-u}{\sigma/\sqrt{n}},$$

（1）当n充分大时，近似地有$S_n \sim N(nu, n\sigma^2)$，标准化$S_n$有

$$\lim_{n\to\infty}P(Y_n\le x)=\Phi(x), \forall x\in R \quad or \quad \lim_{n\to\infty}Y_n \sim N(0,1),$$

即当n充分大时，近似地有 $Y_n \sim N(0,1).$

（2）当n充分大时，近似地有 $\overline{X}_n \sim N(u,\frac{\sigma^2}{n})$。标准化$\overline{X}_n$

有$\lim_{n\to\infty}P(\overline{Y}_n\le x)=\Phi(x), \forall x\in R \quad or \quad \lim_{n\to\infty}\overline{Y}_n \sim N(0,1)$。

定理6.5De. Movire-Laplace中心极限定理（二项分布以正态分布为极限的）：设随机变量 $X_n \sim B(n,p), n=1,2,\cdots$，

则当n充分大时，有$X_n \overset{\cdot}{\sim} N(np, npq)$，即

$$\lim_{n\to\infty} P\left(\frac{X_n - np}{\sqrt{npq}} \le x\right) = \Phi(x), \quad \forall x \in R.$$

证明:X_n 可视为n个相互独立的服从**0-1**分布$B(1,p)$的随机变量$Y_1, Y_2, \cdots, Y_n$的和，故有 $X_n = \sum_{k=1}^{n} Y_k \sim B(n,p).$

又因为$E(Y_k) = p, D(Y_k) = pq \ne 0, k=1,2,\cdots$,由上面定理有

$$\lim_{n\to\infty} P\left(\frac{X_n - np}{\sqrt{npq}} \le x\right) = \Phi(x), \quad \forall x \in R.$$

类似与前面的讨论，我们有如下：

推论：设随机变量$X \sim B(n, p)$，则当n充分大时，近似地有$X \sim N(np, npq)$，从而有近似公式

$$P(a < X \le b) = \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right). \quad (**)$$

注:（1） Poission定理说，当 $p \le 0.1$ 时，二项分布可用Poission分布作近似计算，而上述定理则告诉我们，不论p取何值，二项分布均可用正态分布来作近似计算。实际上，在n很大而p很小时，用正态分布作近似计算不如用Poission分布作近似计算准确。

（2）n很大是一个较模糊的概念，一般而言，当 $n \ge 50$ 时，其近似程度满足一般要求，当然，n越大越好。

（3）对于概率，$P(a \le X \le b), P(a < X < b), P(a \le X < b),$ 均可认为是 $P(a < X \le b)$，因为当n很大时 $P(X = a), P(X = b)$ 可忽略不计，故它们的计算公式都可用（＊＊）或书上的公式。

例6.4

计算机在进行数值计算时，其取整误差 $X \sim U(-0.5,0.5)$ 若在一项计算中进行了100次数值计算，求平均取整误差绝对值小于0.1的概率。

解：令 $X_1, X_2, \ldots X_{100}$ 表示各次数值计算的取整误差，则 $X_1, X_2, \ldots X_{100}$ 独立同分布于 $U(-0.5,0.5)$ 且 $E(X_k)=0, D(X_k)=\dfrac{1}{12}$

平均误差为 $Y=\dfrac{1}{100}\sum_{k=1}^{100} X_k$

例6.4

$$Y=\frac{1}{100}\sum_{k=1}^{100}X_k \quad E(Y)=0, D(Y)=\frac{1/12}{100}$$

由中心极限定理，近似地有

$$Y \mathrel{\dot\sim} N(0,\frac{1}{1200})$$

于是 $P(|Y|<0.1)=P(-0.1<Y<0.1)$

$$\approx\Phi(\frac{0.1}{\sqrt{1/1200}})-\Phi(\frac{-0.1}{\sqrt{1/1200}})$$

$$=2\Phi(2\sqrt{3})-1=0.9996$$

例用一机床制造大小相同的零件，标准重为1kg，由于随机误差，每个零件的重量在（0.95，1.05）(kg)上均匀分布。设每个零件重量相互独立。

（1）制造1200个零件，问总重量大于1202kg的概率时多少?

（2）最多可以制造多少个零件，可使零件重量误差总和的绝对值小于2kg的概率不小于0.9?

解：（1）设X_i表示第i个零件的重量，X表示零件的总重量，则

$$X_i \sim U(0.95, 1.05), E(X_i) = 1, D(X_i) = \frac{(1.05-0.95)^2}{12} = \frac{1}{1200},$$

$$i = 1, 2, \cdots, 1200, \qquad X = \sum_{i=1}^{1200} X_i.$$

由中心极限定理有 $X \sim N(1200\times 1, 1200\times \frac{1}{1200}) = N(1200, 1).$

于是，所求概率为

$$P(X > 1202) = 1 - \Phi(\frac{1202-1200}{1}) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228.$$

（2）设Y_i为第i个零件的误差，则

$$Y_i: U(-0.05, 0.05), E(Y_i)=0, D(Y_i)=\frac{1}{1200}$$

又设$Y=\sum_{i=1}^{n} Y_i$表示n个零件误差的总和，由中心极限定理近似地有$Y: N(0,\frac{n}{1200})$。于是有 $P(|Y|<2)=$

$$P\left(\left|\frac{Y-0}{\sqrt{n/1200}}\right|<\frac{2}{\sqrt{n/1200}}\right)=2\Phi\left(\frac{40\sqrt{3}}{\sqrt{n}}\right)-1\ge 0.9 \Rightarrow \Phi\left(\frac{40\sqrt{3}}{\sqrt{n}}\right)\ge 0.95$$

$$\Rightarrow \frac{40\sqrt{3}}{\sqrt{n}}\ge 1.645 \Rightarrow n\le 1773.82 \Rightarrow n=1773.$$

例2：某车间有**200**台机器，各台机器正常工作与否相互独立，且每台机器的开工概率为**0.6**，正常工作时各需**3kw**的供电。问至少供给多少电力才能以**99.99%**的概率保证此车间不因供电而影响生产？

解：设$X_i=\begin{cases}1, & \text{第i台机器正常开工} \\ 0, & \text{第i台机器未正常开工}\end{cases}$，则 X_i：$B(1,0.6)$，

$E(X_i)=0.6$，$D(X_i)=0.24$，$i=1,2,\cdots,200$，$X=\sum_{i=1}^{200}X_i$

为该车间正常开动的机器台数，且 X：$B(200,0.6)$，

$E(X)=200\times0.6=120$，$D(X)=200\times0.24=48$。由

De. Movire-Laplace中心极限定理近似地有X：$N(120,48)$。

设生产时供给电力y kwh，由题意有 $P(y\geq 3X)\geq 0.9999$，

即$P(X\leq\frac{y}{3})=\Phi(\frac{\frac{y}{3}-120}{\sqrt{48}})\geq 0.9999$，查表得 $\frac{\frac{y}{3}-120}{\sqrt{48}}\geq 3.7$，解得

$y\geq 436.9$。 即至少应供给**436.9**kwh的电力才能以

99.99%的概率保证此车间不因供电不足而影响生产。