浙江大学
ZHEJIANG UNIVERSITY

第35讲　依概率收敛, 切比雪夫不等式

问题的提出：

在第4讲中，曾提到“频率的稳定值记为概率”，这个“稳定”是何含义?

记n_A为n重贝努里试验中事件A发生的次数, 则n_A/n为事件A出现的频率. 若在一次试验中A发生的概率为p. 当试验的次数充分大时, 频率的稳定值为p, 是指

$$\lim_{n\to\infty}\frac{n_A}{n}=p$$

$\Leftrightarrow$ 对于$\forall\,\varepsilon>0,\exists\,N$, 当n>$N$, 均有 $\left|\frac{n_A}{n}-p\right|<\varepsilon.$

$\Leftrightarrow$ 对于$\forall\,\varepsilon>0$, 只要n充分大, 必定有 $\left|\frac{n_A}{n}-p\right|<\varepsilon.$

$\Leftrightarrow$对于$\forall\,\varepsilon>0$, 只要n充分大, "必定"没有$\left\{\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right\}.$

频率“稳定于”概率应从可能性角度来解释, 即

对于$\forall\ \varepsilon > 0$, 只要n充分大, $\left\{ \left|\frac{n_A}{n} - p\right| \geq \varepsilon \right\}$发生的可能

性很小，而且随着n的增大，越来越小.

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

这种收敛性称为“依概率收敛”！

定义： 设$Y_1, Y_2, \cdots, Y_n, \cdots$为一个随机变量序列, c为一常数,
若对于$\forall \varepsilon > 0$, 均有：

$$\lim_{n\to+\infty} P\{|Y_n - c| \ge \varepsilon\} = 0,$$

成立, 则称随机变量序列$\{Y_n, n \ge 1\}$**依概率收敛于**c,
记为：$Y_n \xrightarrow{P} c$,　　当 $n \to +\infty$.

例1：设 $X_n \sim N(0,\frac{1}{n}),\ n=1,\ 2,\ \cdots$, 则 $X_n \xrightarrow{P} 0$, 当 $n \to +\infty$.

证明：对于任意的 $\varepsilon > 0$,

$$P(|X_n - 0| \geq \varepsilon) = P(X_n \geq \varepsilon) + P(X_n \leq -\varepsilon)$$

$$= 1 - \Phi(\frac{\varepsilon - 0}{\sqrt{1/n}}) + \Phi(\frac{-\varepsilon - 0}{\sqrt{1/n}})$$

$$= 2[1 - \Phi(\varepsilon\sqrt{n})] \to 0,\quad \text{当 } n \to +\infty \text{ 时}.$$

性质： 若 $X_n \xrightarrow{P} a,\ Y_n \xrightarrow{P} b$, 当 $n \to \infty$时, 函数$g(x,y)$

在点(a,b)处连续, 那么

$g(X_n, Y_n) \xrightarrow{P} g(a,b)$, 当$n \to \infty$时.

如：当$n \to \infty$时, $X_n + Y_n \xrightarrow{P} a + b$,

$X_n \times Y_n \xrightarrow{P} a \times b$, $X_n / Y_n \xrightarrow{P} a / b\ (b \neq 0)$.

特别地, 若$X_n \xrightarrow{P} a$, $f(x)$在点a连续, 则

$f(X_n) \xrightarrow{P} f(a)$, 当$n \to \infty$时.

定理（切比雪夫不等式）：

设随机变量X具有数学期望

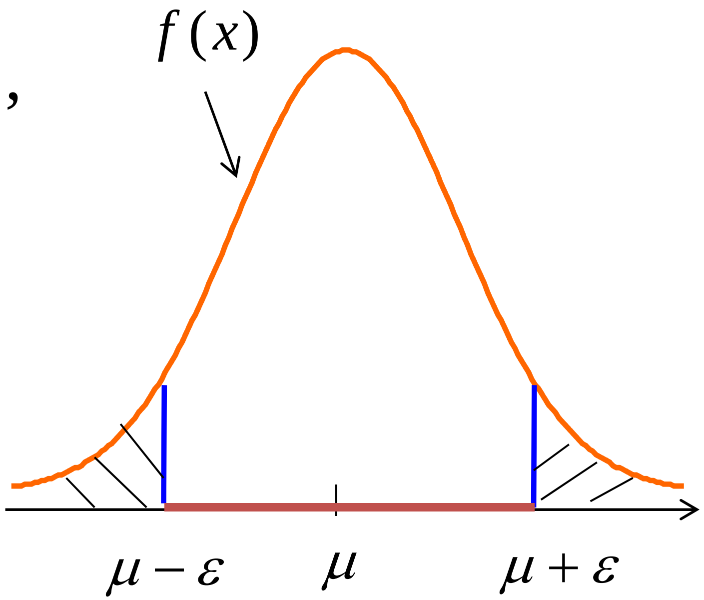
$E(X)=\mu$，方差$D(X)=\sigma^2$，

则对于任意$\varepsilon>0$，都有：

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

定理的**等价形式**为：

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



证明： 对于任意 $\varepsilon > 0$，令 $Z = \begin{cases} \varepsilon, & \text{当} |X-\mu| \geq \varepsilon \text{时}; \\ 0, & \text{当} |X-\mu| < \varepsilon \text{时}. \end{cases}$

则 $Z \leq |X-\mu|$，那么 X 的方差 $(D(X) = E[(X-\mu)^2])$ 存在时，$E(Z^2)$ 也存在，且 $E(Z^2) \leq D(X)$.

而据 Z 的定义，知 $E(Z^2) = \varepsilon^2 P\{|X-\mu| \geq \varepsilon\}$,

故而 $\varepsilon^2 P\{|X-\mu| \geq \varepsilon\} \leq D(X)$，即

$P\{|X-\mu| \geq \varepsilon\} \leq \dfrac{D(X)}{\varepsilon^2}$ 成立.

切比雪夫不等式

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

适用范围：对于期望、方差存在的随机变量.

——范围广

重要性：可以对于随机变量落在期望附近的区域内或外给出一个界的估计.

比如： 设$E(X)=\mu$, $D(X)=\sigma^2$, 取$\varepsilon=K\sqrt{D(X)}=K\sigma$,

$$\text{则有}\ P(|X-\mu|\geq K\sigma)\leq\frac{D(X)}{K^2\sigma^2}=\frac{1}{K^2}.$$

取 $K=3$, 则 $P(|X-\mu|\geq 3\sigma)=P(X\notin(\mu-3\sigma,\mu+3\sigma))\leq\frac{1}{9}$.

而当$X\sim N(\mu,\sigma^2)$时,

$P(|X-\mu|\geq 3\sigma)=1-P(|X-\mu|<3\sigma)=1-(\Phi(3)-\Phi(-3))$

$=1-0.9974=0.0026<\frac{1}{9}$.

Chebyshev不等式应用范围广，但是结果比较粗糙.

例2: 某天文机构想测量宇宙中两颗行星的距离, 进行了n次独立的观测, 测量结果分别X_i(光年), i=1, 2, ⋯, n. 若 $E(X_i)=\mu$ (为两颗行星的真实距离, 未知), $D(X_i)=5$. 现取这n次观测的平均作为实际距离μ的估计.

（1）若n=100, 那么估计与实际值之间的误差在±0.5光年之内的概率至少有多大? （2）若要以不低于95%的把握控制估计与实际值之间的误差在±0.5光年之内, 至少要观测多少次?

解: 由题意, 估计与实际值之间的误差即为$\frac{1}{n}\sum_{i=1}^{n}X_i-\mu$.

由于对$i=1, 2, \cdots, n$，都有$E(X_i)=\mu,\ D(X_i)=5$,

且X_i相互独立, 故

$$E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \mu, D(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{5}{n}.$$

(1) 当$n=100$时，由切比雪夫不等式知

$$P\{|\frac{1}{100}\sum_{i=1}^{100} X_i - \mu| < 0.5\} \quad \geq 1 - \frac{5/100}{0.5^2} = 0.8;$$

(2) 同样利用切比雪夫不等式，要使得

$$P\{|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < 0.5\} \geq 1 - \frac{5/n}{0.5^2} \geq 0.95,$$

n需满足$n \geq 400$.