第五章　大数定律及中心极限定理

1. 切比雪夫不等式
2. 大数定律
3. 中心极限定理

从随机现象中去寻求必然的法则，应该研究大量随机现象.

极限定理的内容很广泛,其中最重要的有两种

大数定律　与　中心极限定理

§1 切比雪夫不等式

切比雪夫不等式

设随机变量X具有期望 $E(X)=\mu$ 和方差 $D(X)=\sigma^2$，则对于任意给定的正数ε，有

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

证明 若X为连续型随机变量，设X的概率密度为$f(x)$,

$$P\{|X-\mu|\geq\varepsilon\}=\int_{|x-\mu|\geq\varepsilon} f(x)\mathrm{d}x$$

$$\leq\int_{|x-\mu|\geq\varepsilon}\frac{(x-\mu)^2}{\varepsilon^2}f(x)\mathrm{d}x$$

$$\leq\int_{-\infty}^{+\infty}\frac{(x-\mu)^2}{\varepsilon^2}f(x)\mathrm{d}x$$

$$=\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)\mathrm{d}x=\frac{\sigma^2}{\varepsilon^2}$$

切比雪夫不等式

设随机变量X具有期望 $E(X)=\mu$ 和方差 $D(X)=\sigma^2$，则对于任意给定的正数ε，有

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

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

例 设随机变量X的$E(X)=\mu, D(X)=\sigma^2$,则由切比雪夫

不等式, $P\{|X-\mu|<3\sigma\} \geq$ <u>8/9</u>. 估计

精确计算

特别的,若$X \sim N(\mu,\sigma^2)$, 则$P\{|X-\mu|<3\sigma\} =$ <u>0.9974</u>.

例　设$E(X)=-2, E(Y)=2, D(X)=1, D(Y)=4, \rho_{XY}=-0.5$,
由切比雪夫不等式, $P\{|X+Y|\geq 6\}\leq$ __1/12___.

例 将一颗骰子连续投掷6次, 设出现的点数和为X, 用切比雪夫不等式估计概率$P\{|X-21|<7\}$.

解 设X_i表示第i次投掷所得的点数, $i=1, 2, 3, 4, 5, 6$, 则$X=\sum_{i=1}^{6} X_i$.

$X_1, X_2, \ldots X_6$相互独立且具有相同的分布律,

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

$$E(X)=\sum_{i=1}^{6} E(X_i)=6\times\frac{7}{2}=21 \qquad D(X)=\sum_{i=1}^{6} D(X_i)=6\times\frac{35}{12}=\frac{35}{2}$$

$$P\{|X-21|<7\}\geq 1-\frac{\frac{35}{2}}{7^2}=\frac{9}{14}$$

§2 大数定律

1. 依概率收敛

2. 大数定律

依概率收敛

定义 设 $X_1, X_2, \ldots X_n, \ldots$是一个随机变量序列, a是一个常数. 如果对于任意给定的正数ε, 有

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

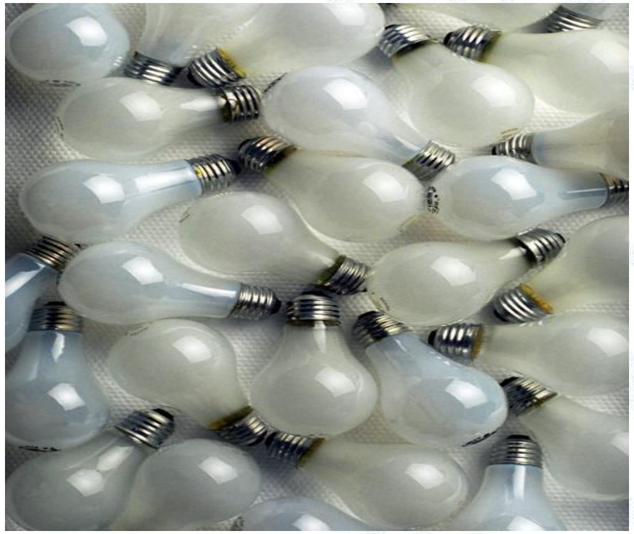
则称随机变量序列$X_1, X_2, \ldots X_n, \ldots$依概率收敛于$a$, 记作 $X_n \xrightarrow{P} a$.

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

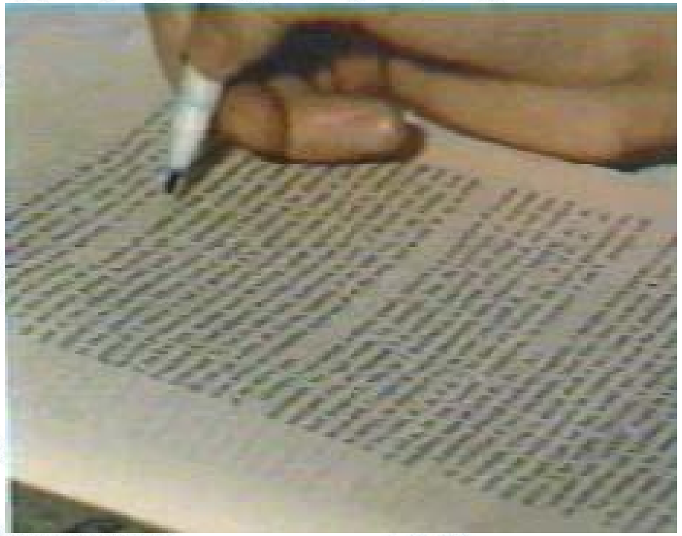
大数定律

客观背景 大量的随机现象平均结果的稳定性.

大量抛掷硬币
正面出现频率



生产过程中的
废品率



字母使用频率

大数定律

大数定律

切比雪夫定理
(Chebyshev)

伯努利定理
(Bernoulli)

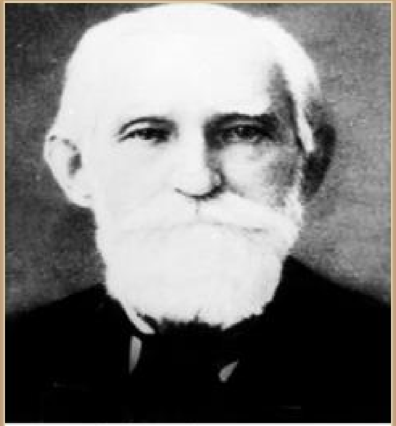
辛钦定理
(Khintchine)

切比雪夫(Chebyshev)大数定律

定理 设 $X_1, X_2, \ldots X_n, \ldots$是相互独立的随机变量序列,具有数学期望$E(X_k)$及方差$D(X_k)$, $k=1,2,\ldots$,且方差一致有上界,即存在正数M 使得 $D(X_k) \le M$, $k=1,2,\ldots$,

则对任意给定的正数ε, 恒有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E(X_k)\right| < \varepsilon\right\} = 1$$



切比雪夫, П. Л.

证明 由$E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k),D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k),$

根据切比雪夫不等式

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}\geq 1-\frac{\sum_{k=1}^{n}D(X_k)}{n^2\varepsilon^2}$$

由于方差一致有上界,$\sum_{k=1}^{n}D(X_k)\leq nM,$

$$1\geq P\leq\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}\geq 1-\frac{M}{n\varepsilon^2},$$

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}=1$$

切比雪夫定理表明，独立的随机变量序列$X_1, X_2, \ldots X_n, \ldots$，如果存在数学期望和方差，且方差一致有上界，则$X_1, X_2, \ldots X_n$的算术平均值与它们的数学期望的算术平均值之差当$n \to \infty$时依概率收敛于零.

即当n充分大时，$\dfrac{1}{n}\sum_{i=1}^{n} X_i$差不多不再是随机的了，取值接近于其数学期望的概率接近于1.

切比雪夫大数定律给出了平均值稳定性的科学描述

推论 设随机变量序列$X_1, X_2, \ldots X_n, \ldots$相互独立,并且具有相同的数学期望$E(X_k)=\mu$和方差$D(X_k)=\sigma^2$, $k=1,2, \ldots$,则其算术平均值$\frac{1}{n}\sum_{k=1}^{n} X_k$当$n \to \infty$时依概率收敛于数学期望$\mu$, 即对任意给定的正数$\varepsilon$,恒有

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

实际问题中使用

算术平均值的依据

伯努利(Bernoulli)定理

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，则对任给的正数ε，有

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

雅各布第一·伯努利

证明 $n_A \sim B(n,p)$，由

$$E\left(\frac{n_A}{n}\right)=\frac{1}{n}E(n_A)=p,\qquad D\left(\frac{n_A}{n}\right)=\frac{1}{n^2}D(n_A)=\frac{p(1-p)}{n},$$

根据切比雪夫不等式

$$1\geq P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}\geq 1-\frac{p(1-p)}{n\varepsilon^2}$$

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

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

当n充分大时,频率n_A/n与概率有较大偏差的可能性很小.

伯努利定理表明,当n充分大时, 事件A在n次独立重复试验中发生的频率n_A/n依概率收敛于其发生的概率p.

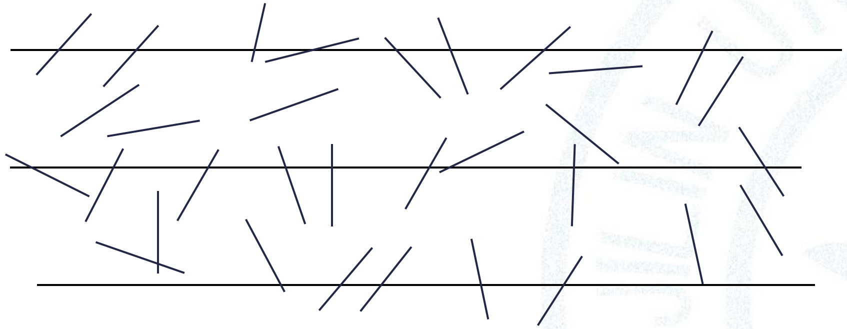
伯努利定理表达了频率的稳定性,

是实际问题中用频率代替概率的依据

大数定律是蒲丰投针问题中π解法的理论依据

线距 a

针长 l



$$\pi \approx \frac{2lN}{an}$$

当投针次数n很大时,用针与线相交的频率n/N近似针与线相交的概率p,从而求得π的近似值.

辛钦(Khintchine)大数定律

定理 设随机变量序列$X_1, X_2, \ldots X_n, \ldots$相互独立,服从同一分布, 具有数学期$E(X_k)=\mu,\ k=1,2,\ldots,$ 则对任意给定的正数ε ,有

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

注 独立同分布下的大数定律,不要求随机变量的方差存在.

辛钦

辛钦定理表明，独立同分布随机变量序列$X_1, X_2, \ldots X_n, \ldots$，具有数学期望$\mu$，则

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

如果$E(X_k^l) = \mu_l (k = 1, 2, \cdots)$ 存在，则 $\frac{1}{n}\sum_{k=1}^{n} X_k^l \xrightarrow{P} \mu_l = E(X_k^l).$

辛钦大数定律为寻找随机变量的期望值提供了一条实际可行的途径.

辛钦定理是数理统计中求参数点估计中矩估计法的理论基础

§3　中心极限定理

1.　依分布收敛

2.　中心极限定理

依分布收敛

定义 设随机变量$X, X_1, X_2, \ldots X_n, \ldots$的分布函数依次是$F(x)$, $F_1(x), F_2(x), \ldots F_n(x), \ldots$, 如果对于$F(x)$的每个连续点$x$, 都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

则称随机变量 序列$X_1, X_2, \ldots X_n, \ldots$依分布收敛于$X$,记作$X_n \xrightarrow{L} X$.

客观背景 在实际问题中,常常需要考虑许多随机因素所产生总影响.

例如炮弹射击的落点与目标的偏差是所有随机因素的总影响.

炮弹外形差异引起的偏差

空气阻力引起的偏差

火药量的微小差异引起的偏差

发射角度的微小差异引起的偏差

瞄准时的偏差 ······

自从高斯指出测量误差服从正态分布之后, 人们发现, 正态分布在自然界中极为常见.

事实上,如果一个量是由大量相互独立的随机因素的影响所造成,而每一个因素在总影响中所起的作用微小, 则这种量一般都服从或近似服从正态分布.


高斯

研究独立随机变量之和所特有的规律性问题.

考虑独立随机变量之和 $\sum_{k=1}^{n} X_k$.无穷个随机变量之和可能趋于∞,故不研究n个随机变量之和本身.

考虑随机变量之和的标准化变量

$$Z_n = \frac{\sum_{k=1}^{n} X_k - E\left(\sum_{k=1}^{n} X_k\right)}{\sqrt{D\left(\sum_{k=1}^{n} X_k\right)}}$$

的分布函数的极限.

可以证明,满足一定的条件,上述极限分布是标准正态分布.

中心极限定理

中心极限定理

独立同分布
列维-林德伯格定理
(Levy-Lindberg)

二项分布正态近似
棣莫弗-拉普拉斯定理
(De Moivre-Laplace)

一般结论
李雅普诺夫定理
(Liapunov)

独立同分布中心极限(列维-林德伯格Levy-Lindberg)定理

定理 设$X_1, X_2, \ldots X_n, \ldots$相互独立, 服从同一分布, 且具有数学期和方差 $E(X_k)=\mu, D(X_k)=\sigma^2 \neq 0(k=1,2,\ldots)$,

$$F_n(x)=P\{Y_n \le x\}=P\left\{\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\right\}, -\infty < x < +\infty$$

则

$$\lim_{n\to\infty} F_n(x)=\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$Y_n=\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$$

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

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

$$Y_n=\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \xrightarrow{L} U \sim N(0,1)$$

$$\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{\text{近似}}{\sim} N(0,1) \qquad \xrightarrow{L}$$

$$\sum_{k=1}^{n} X_k \overset{\text{近似}}{\sim} N(n\mu, n\sigma^2)$$

虽然在一般情况下,很难求出$X_1+X_2+\ldots+X_n$的分布的确切形式,但当n很大时,可以求出近似分布.

独立同分布中心极限定理表明,当n充分大时,n个具有期望和方差的独立同分布的随机变量之和近似服从正态分布.

例 已知一射击运动员在一次射击中所得环数X的概率分布，

X	6	7	8	9	10
P	0.05	0.05	0.1	0.3	0.5

(已知$\Phi(1.35)=0.9115$)

求1)在100次独立射击中所得总环数介于900与930环之间的概率.

解 设 X_k表示第 k ($k=1,2,\ldots$)次射击所得环数，则$X_1, X_2, \ldots X_{100}$相互独立，都与X同分布，100次射击所得总环数为 $\sum_{k=1}^{100} X_k$，

$$\mu = E(X_k) = E(X) = 9.15,\ \sigma^2 = D(X_k) = D(X) = 1.23$$

且 $\dfrac{\sum_{k=1}^{100} X_k - 100\times 9.15}{\sqrt{100\times 1.23}}$ 近似服从$N(0,1)$

$$P\{900<\sum_{k=1}^{100} X_k<930\}=P\{\frac{900-100\times 9.15}{\sqrt{100\times 1.23}}<\frac{\sum_{k=1}^{100} X_k-100\times 9.15}{\sqrt{100\times 1.23}}<\frac{930-100\times 9.15}{\sqrt{100\times 1.23}}\}$$

$$=P\{-1.35<\frac{\sum_{k=1}^{100} X_k-100\times 9.15}{\sqrt{100\times 1.23}}<1.35\}\approx\Phi(1.35)-\Phi(-1.35)$$

$$=2\Phi(1.35)-1=2\times 0.9115-1=0.8230.$$

2)利用切比雪夫不等式估计在100次独立射击中所得总环数介于900与930环之间的概率.

解 100次射击所得总环数为$\sum_{k=1}^{100} X_k$,

$$E(X)=100\times 9.15=915, D(X)=100\times 1.23=123,$$

由切比雪夫不等式

$$P\{900<X<930\}=P\{900-915<X-915<930-915\}$$

$$=P\{|X-915|<15\}\geq 1-\frac{123}{15^2}=0.45$$

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

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

例 生产线产品成箱包装,每箱的重量是随机的.假设每箱平均重50kg,标准差为5kg,若用最大载重为5t的汽车承运,试利用中心极限定理说明,每辆车最多可以装多少箱,才能保障不超载的概率大于0.9772. (已知$\Phi(2)=0.9772$.)

解 设每辆车最多可以装n箱,第k箱重量X_k,则载重量为$\sum\limits_{k=1}^{n} X_k$,

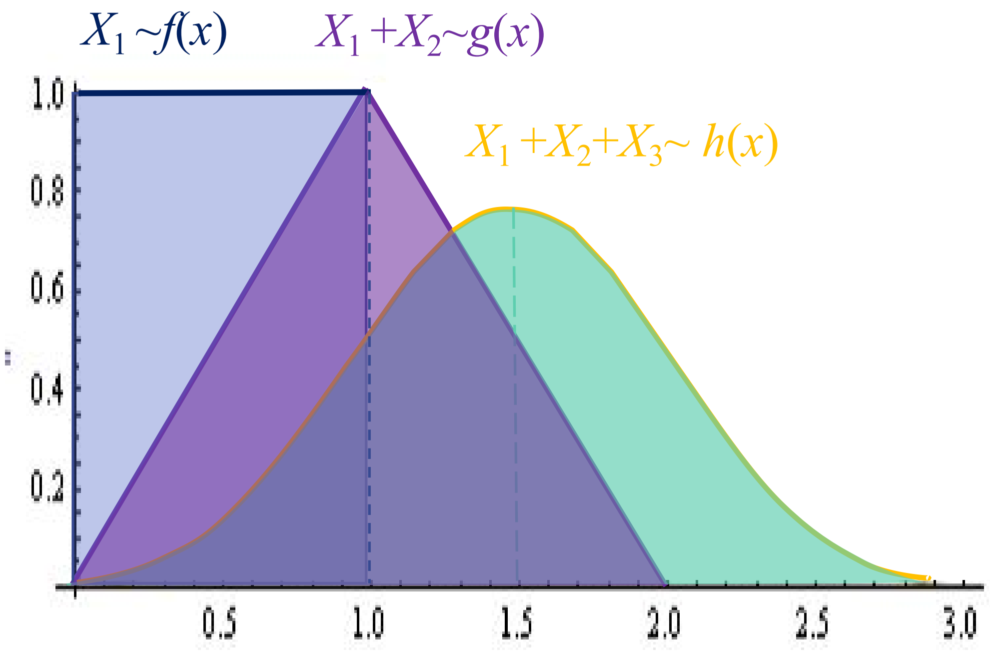
且$X_1, X_2, \ldots X_n$相互独立同分布, $E(X_k)=50$, $D(X_k)=25$.

$$P\{\sum_{k=1}^{n} X_k \le 5000\} > 0.9772,$$

$$\Phi(\frac{5000-50n}{5\sqrt{n}}) \approx P\{\frac{\sum\limits_{k=1}^{n} X_k - 50n}{5\sqrt{n}} \le \frac{5000-50n}{5\sqrt{n}}\} > 0.9772 = \Phi(2),$$

$$\frac{5000-50n}{5\sqrt{n}} > 2 \qquad \text{故} n < 98.$$

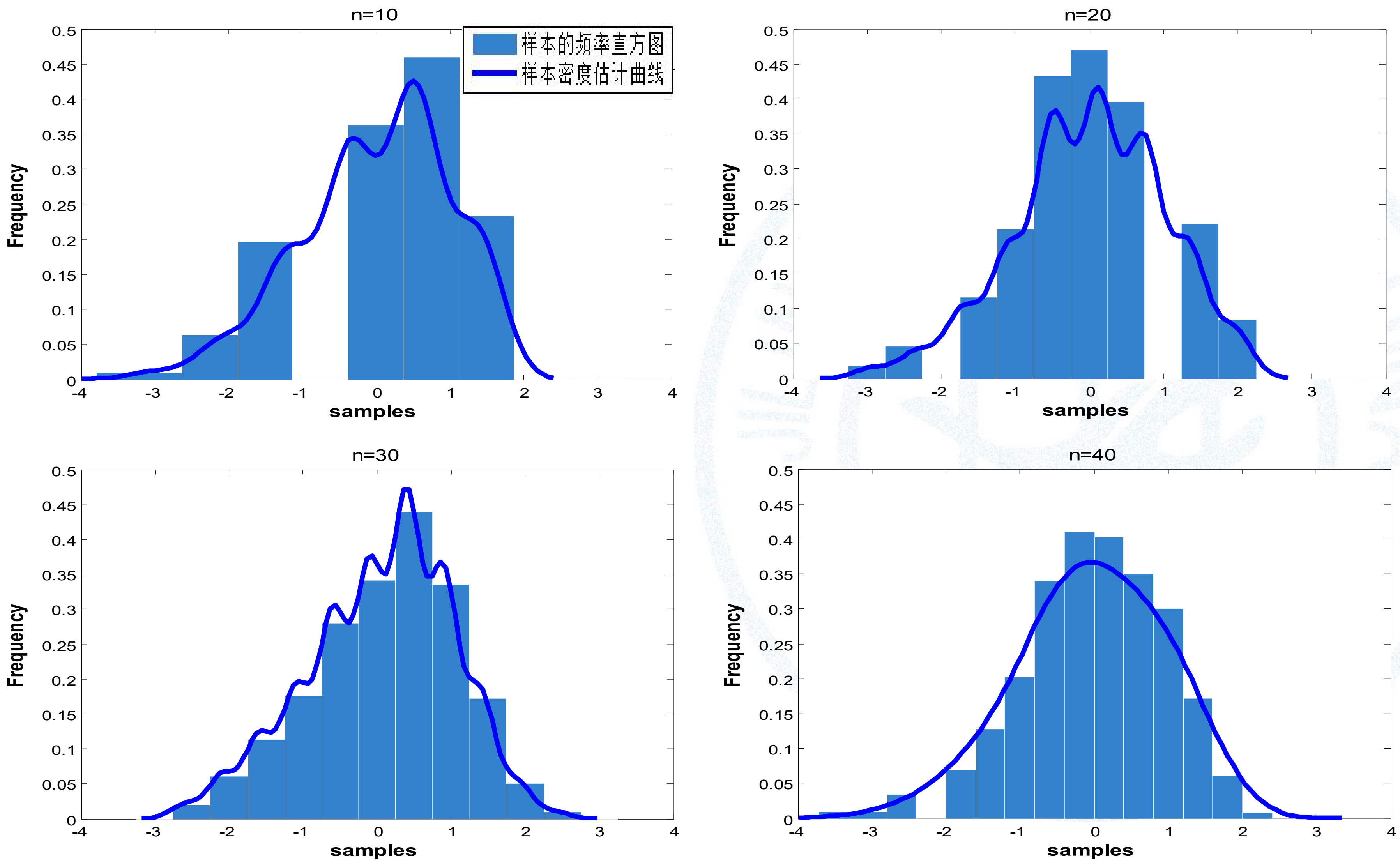
独立的随机变量之和的直观演示



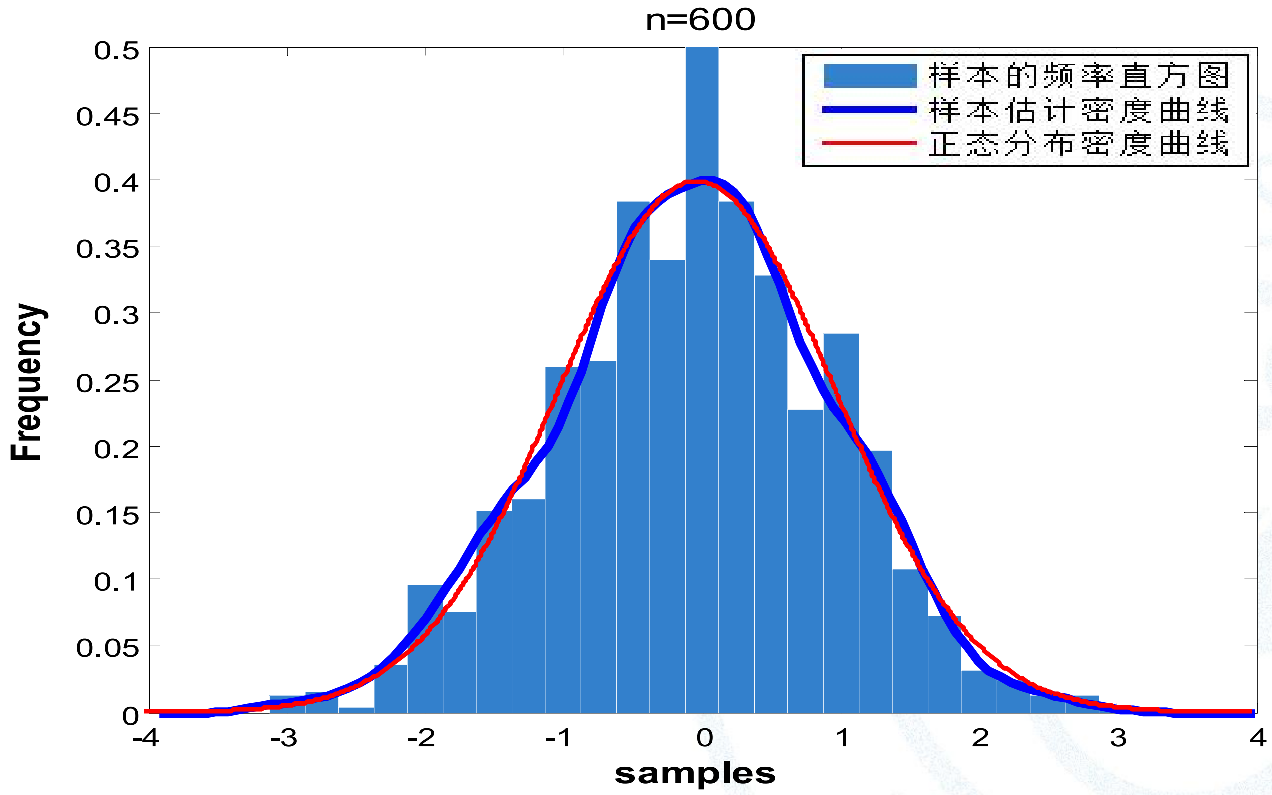
区间(0,1)上均匀分布之和的密度曲线

正态分布是大量独立同分布的随机变量之和的极限分布.

二项分布的随机模拟



二项分布的随机模拟



当n很大时,用连续的正态分布逼近离散的二项分布

棣莫佛－拉普拉斯(DeMoivre-Laplace)定理

设随机变量Y_n服从参数$n, p\,(0<p<1)$的二项分布, 则对任意x, 有

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

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \xrightarrow{L} U \sim N(0,1)$$

棣莫佛—拉普拉斯定理是独立同分布中心极限定理的一种特殊结论

正态分布是二项分布的极限分布

证明 将Y_n分解为n个相互独立，且均服从同一(0-1)分布的随机变量$X_1, X_2, \cdots, X_n$之和，即有

$$Y_n = \sum_{k=1}^{n} X_k,$$

其中$X_k (k=1,2,\cdots,n)$的分布律为

由于 $$P\{X_k = 0\} = 1 - p, \quad P\{X_k = 1\} = p, \quad 0 < p < 1,$$

$$E(X_k) = p, \; D(X_k) = p(1-p) \quad (k = 1,2,\cdots,n),$$

根据独立同分布中心极限定理, 对任意实数x, 有

$$\lim_{n\to\infty} P\left\{\frac{Y_n - np}{\sqrt{np(1-p)}} \le x\right\} = \lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - np}{\sqrt{np(1-p)}} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

二项分布的近似

注 当n充分大时,

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)\ ;$$

$$Y_n \overset{\text{近似}}{\sim} N(np\,,\,np(1-p))\ .$$

棣莫佛－拉普拉斯定理表明,当n很大,$0<p<1$是一个定值时(或者说$np(1\text{-}p)$也不太大时), 二项变量Y_n的分布近似正态分布 $N(np,np(1\text{-}p))$.

泊松定理表明, 当二项分布$B(n,p)$的参数n很大, 而p很小时, 可以用参数为$\lambda=np$的泊松分布来近似计算.

例 某工厂有200台同类型机器,每台机器实际工作时间占全部工作时间的75%.各台机器工作相互独立.求任意时刻有144至160台机器正在工作的概率. (已知$\Phi(1.63)=0.9484, \Phi(0.98)=0.8365.$)

解 设X每表示任一时刻正在工作的机器台数, $X \sim B(200, 0.75).$

由棣莫弗-拉普拉斯定理,

$$P\{144 \le X \le 160\} = P\{\frac{144-200\times 0.75}{\sqrt{200\times 0.75\times 0.25}} \le \frac{X-200\times 0.75}{\sqrt{200\times 0.75\times 0.25}} \le \frac{160-200\times 0.75}{\sqrt{200\times 0.75\times 0.25}}\}$$

$$= P\{-0.98 \le \frac{X-150}{\sqrt{37.5}} \le 1.63\} \approx \Phi(1.63)-\Phi(-0.98)$$

$$= 0.9484-(1-0.8365) = 0.7849.$$

例 某单位为了解人们对某决议的态度进行抽样调查.设该单位每人赞成该决议的概率为$p(0<p<1)$,p未知,且人们赞成与否相互独立.问要调查多少人才能使赞成该决议的人数频率作为p的近似值,其误差不超过0.01的概率达0.95以上.

解 设X_n为调查中n个人中赞成的人数, $X_n \sim B(n,p)$.

$$P\left\{\left|\frac{X_n}{n}-p\right|\le 0.01\right\}\ge 0.95, \qquad P\left\{\left|\frac{X_n-np}{\sqrt{np(1-p)}}\right|\le \frac{0.01n}{\sqrt{np(1-p)}}\right\}\ge 0.95,$$

由棣莫弗-拉普拉斯定理, $\dfrac{X_n-np}{\sqrt{np(1-p)}}$ 近似服从$N(0,1)$,

$$0.95\le P\left\{\left|\frac{X_n-np}{\sqrt{np(1-p)}}\right|\le \frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right\}\approx 2\Phi\left(\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right)-1,$$

$$0.95 \le P\left\{\left|\frac{X_n - np}{\sqrt{np(1-p)}}\right| \le \frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right\} \approx 2\Phi\left(\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1,$$

$$\Phi\left(\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right) \ge 0.975 = \Phi(1.96),$$

$$\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}} \ge 1.96,$$

$$n \ge 196^2 p(1-p), \quad 由 p(1-p) \le \frac{1}{4},$$

$$n \ge 196^2 \times \frac{1}{4} = 9604.$$

例 某单位有260部电话分机,每个分机有4%的时间使用外线.各分机是否使用外线相互独立.问总机要有多少外线能有95%保证各个分机用外线时不必等候?

解 设总机有x条外线.设X_n表示某时刻同时使用外线的分机数,$X_n \sim B(260, 0.04)$.则$P\{X_n < x\} = 0.95$,

由棣莫弗-拉普拉斯定理,$\dfrac{X_n - np}{\sqrt{np(1-p)}}$近似服从$N(0,1)$,

$$0.95 = P\{\frac{X_n - np}{\sqrt{np(1-p)}} < \frac{x - np}{\sqrt{np(1-p)}}\} \approx \Phi(\frac{x - np}{\sqrt{np(1-p)}}),$$

查表得,$\dfrac{x - 260 \times 0.04}{\sqrt{260 \times 0.04 \times 0.96}} \approx 1.65$,解得$x \approx 15.61$.

取整数$x = 16$.总机至少需要16条外线.

李雅普诺夫(Liapunov)定理*

定理　设$X_1,X_2,\ldots,X_n,\ldots$相互独立,具有$E(X_k)=\mu_k, D(X_k)=\sigma_k^2\neq 0 (k=1,2,\ldots)$,

记$B_n^2=\sum_{k=1}^{n}\sigma_k^2$. 设$Z_n=\dfrac{\sum_{k=1}^{n}X_k-E(\sum_{k=1}^{n}X_k)}{\sqrt{D(\sum_{k=1}^{n}X_k)}}=\dfrac{\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}\mu_k}{B_n}$的分布函数

$$F_n(x)=P\{Z_n\le x\}=P\{\frac{\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}\mu_k}{B_n}\le x\},-\infty<x<+\infty.$$

如果存在正数δ, 使得当$n\to\infty$时, 有$\dfrac{1}{B_n^{2+\delta}}\sum_{k=1}^{n}E(|X_k-\mu_k|^{2+\delta})\to 0$,则对任意实数$x$, 恒有

$$\lim_{n\to\infty}F_n(x)=\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$$

$$Z_n = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n} \xrightarrow{L} U \sim N(0,1)$$

在李雅普诺夫定理条件下,当n很大时,

$$Z_n = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n}$$ 近似服从标准正态分布$N(0,1)$,

$\sum_{k=1}^{n} X_k = B_n Z_n + \sum_{k=1}^{n} \mu_k$近似服从正态分布$N(\sum_{k=1}^{n} \mu_k, B_n^2)$.

实际问题中很多个独立随机变量的和变量

近似服从正态分布的理论依据

例 对敌人的防御地区进行100次轰炸,每次轰炸命中目标的炸弹数是一个随机变量,其数学期望为2,方差为1.69.求在100次轰炸中有180颗到220颗炸弹命中目标的概率.

解 设X_k表示第k次轰炸命中的炸弹数 ,则100次轰炸命中的炸弹数为 $\sum_{k=1}^{100} X_k$, $E(X)=\sum_{k=1}^{100} E(X_k)=200, D(X)=\sum_{k=1}^{100} D(X_k)=169.$

根据李雅普诺夫定理，渐进地服从正态分布，

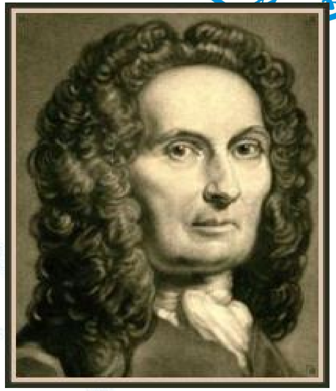
$$P\{180<X<220\}=P\{|X-200|\le 20\}=P\{\frac{|X-200|}{13}\le\frac{20}{13}\}$$

$$\approx 2\Phi(1.54)-1\approx 0.8764.$$

中心极限定理的发展历史

1733年，棣莫弗针对二项分布近似计算，最早勾画出正态分布；

$$X \sim B(n,\frac{1}{2}), \qquad P\left\{\left|\frac{X}{n}-\frac{1}{2}\le\frac{c}{\sqrt{n}}\right|\right\}\approx\int_{-2c}^{2c}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x$$



亚伯拉罕·棣莫弗
De Moivre,Abraham

1770年，拉普拉斯推广其理论，提出二项的分布的极限分布是正态分布；

十九世纪，统计学家建立中心极限定理理论，指出在一定条件下，大量随机变量之和都将近似服从正态分布.

皮埃尔-西蒙·拉普拉斯
Laplace,Pierre-Simon

中心极限定理是概率论中最著名的结果之一,它不仅提供了计算独立随机变量之和的近似概率的简单方法,而且有助于解释为什么很多自然群体的经验频率呈现出钟形曲线这一值得注意的事实.