哈爾濱工業大學

第5章随机变量的数字特征与极限定理

第28讲 中心极限定理

中心极限定理

- 本讲我们来研究独立随机变量和 $\sum_{i=1}^{n} X_i$ 当 $n \to \infty$ 时的分布问题.

- 在什么条件下极限分布会是正态的呢？
- 在概率论中，习惯于把和的分布收敛于正态分布这一类定理都叫做中心极限定理.

独立同分布下的中心极限定理

定理1 设$X_1, X_2, \cdots$是独立同分布的随机变量序列，且

$E(X_i)=\mu, D(X_i)=\sigma^2>0(i=1,2,\cdots)$存在，则对充分大的$n$，有

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

注意

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

$$D\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} D(X_i)=n\sigma^2.$$

独立同分布下的中心极限定理

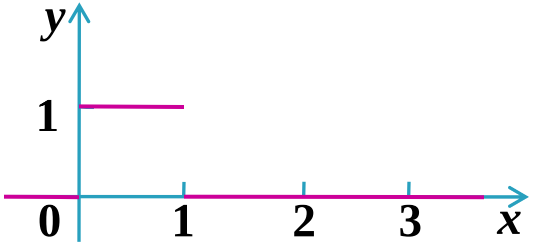
$$\sum_{i=1}^{n} X_i \overset{\text{近似}}{\sim} N(n\mu, n\sigma^2)$$

$$P(a < \sum_{i=1}^{n} X_i \le b) \approx \Phi(\frac{b - n\mu}{\sqrt{n}\sigma}) - \Phi(\frac{a - n\mu}{\sqrt{n}\sigma}).$$

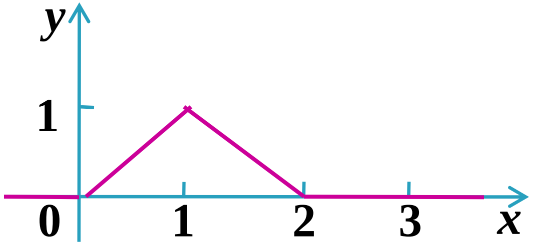
此定理也称为林德伯格-莱维(Lindeberg-Levi)中心极限定理.

中心极限定理的客观背景

$X_1 \sim U(0,1)$,概率密度为$f(x)$

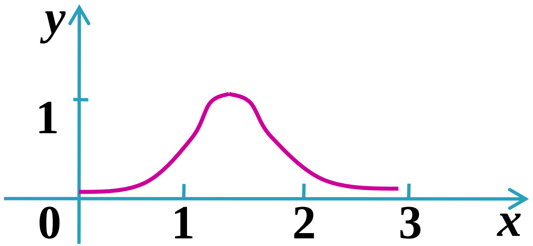


X_1+X_2的概率密度$g(x)$

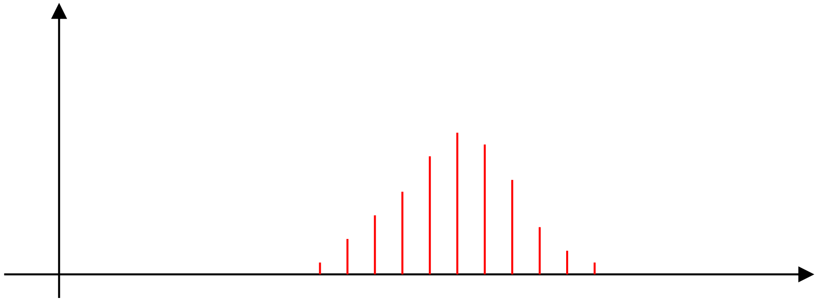


X_1, X_2, X_3独立

同分布于$U(0,1)$.



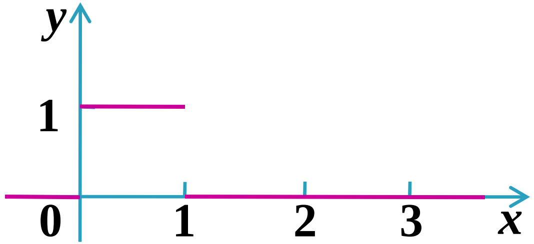
$X_1+X_2+X_3$的概率密度$h(x)$



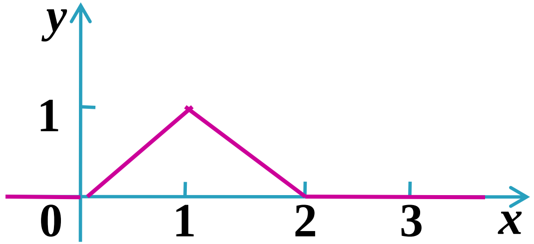
例:20个0-1分布的和的分布

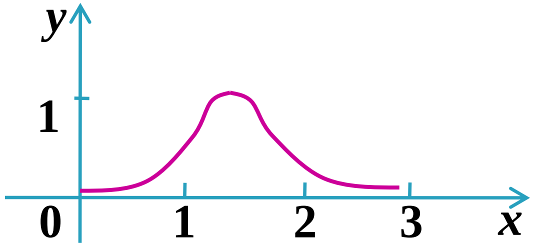
中心极限定理的客观背景

$X_1 \sim U(0,1)$,概率密度为$f(x)$



X_1+X_2的概率密度$g(x)$





$X_1+X_2+X_3$的概率密度$h(x)$

X_1, X_2, X_3独立同分布于$U(0,1)$.

应用

例1 计算机在进行加法时，对每个被加数取整(取为最接近它的整数)，设所有的取整误差是相互独立的，且它们均在(-0.5, 0.5)上服从均匀分布. 若将1500个数相加，问误差总和的绝对值不超过15的概率是多少？

解 设X_i为第i个数的误差(i=1,2,⋯,1500)，

则 $X_i \sim U(-0.5, 0.5)$ 且 $X_1, \cdots, X_{1500}$ 独立，

令$Z= X_1+\cdots+X_{1500}$ 则

$$Z \overset{\text{近似}}{\sim} N(E(Z), D(Z)).$$

应用

又 $X_i \sim U(-0.5, 0.5), E(X_i) = 0, D(X_i) = 1/12$

$$E(Z) = E\left(\sum_{i=1}^{1500} X_i\right) = 1500E(X_i) = 0,$$

$$D(Z) = D\left(\sum_{i=1}^{1500} X_i\right) = 1500D(X_i) = 1500/12 = 125,$$

所求概率

$$P(|Z| \le 15) = P(-15 \le Z \le 15) = \Phi\left(\frac{15}{\sqrt{125}}\right) - \Phi\left(\frac{-15}{\sqrt{125}}\right)$$

$$= 2\Phi(1.34) - 1 = 0.8198.$$

棣莫佛－拉普拉斯定理

定理2 设随机变量Y_n服从参数$n, p\,(0<p<1)$的二项分布，则对充分大的n，有

$$Y_n \overset{近似}{\sim} N(np, npq), \quad (q = 1 - p).$$

即

$$P(a < Y_n \le b) \approx \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right).$$

在实际中, $0.1<p<0.9$, $npq>9$时, 用正态近似;

当 $p\le 0.1$(或$p\ge 0.9$) 且 $n\ge 10$时，用泊松近似.

棣莫佛—拉普拉斯定理

例2 某保险公司多年的资料表明，在索赔户中，被盗索赔户占20%, 以X表示在随机抽查100个索赔户中因被盗而向保险公司索赔的户数，求$P(14 \le X \le 30)$.

解 $X \sim B(100, 0.2)$,

$$P(14 \le X \le 30) \approx \Phi\left(\frac{30-100\times 0.2}{\sqrt{100\times 0.2\times 0.8}}\right)-\Phi\left(\frac{14-100\times 0.2}{\sqrt{100\times 0.2\times 0.8}}\right)$$

$$=\Phi(2.5)-\Phi(-1.5)$$

$$=0.9938+\Phi(1.5)-1$$

$$=0.9938+0.9332-1=0.927.$$

谢　　谢！