概率论与数理统计

第 6 讲

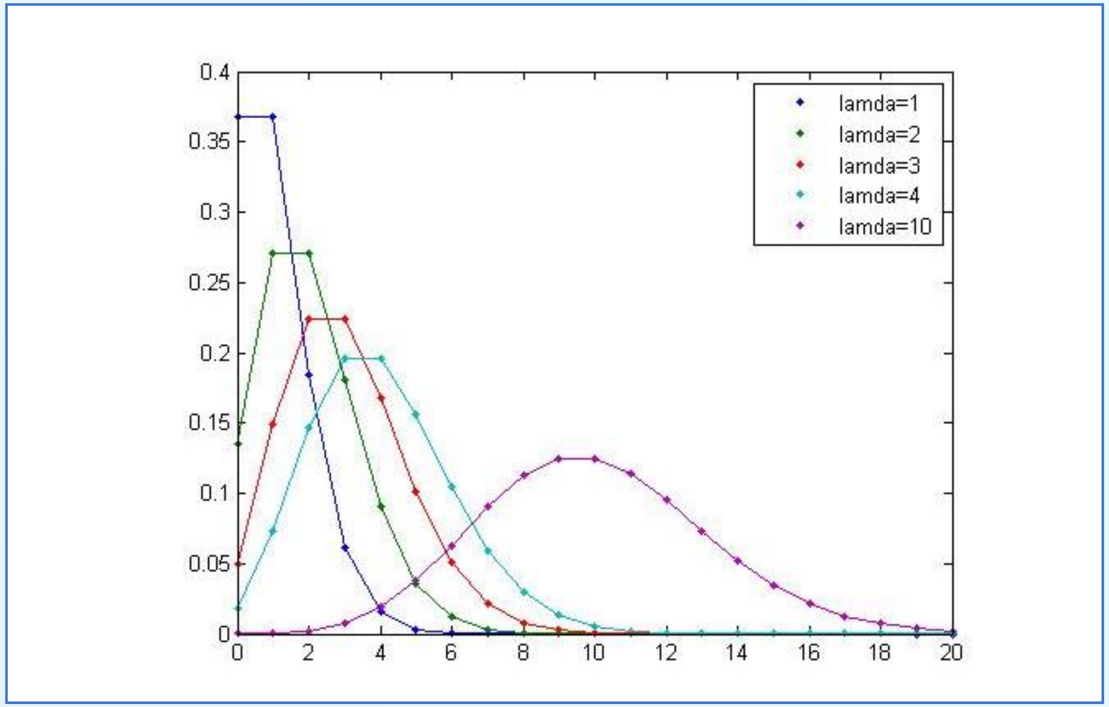
随机变量的分布函数

6.泊松分布

若离散型随机变量X的分布律为

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda} \qquad k=0, 1, 2,\ldots$$

其中常数$\lambda>0$，则称X服从参数为λ的泊松分布，记为$X\sim P(\lambda)$ 或$X\sim\pi(\lambda)$。



一. 重要的离散型分布（续）

Poisson分布是概率论中重要的分布之一，自然界及工程技术中的许多随机变量都服从Poisson分布。

电话总机在某一段时间内收到的呼叫次数；
公共汽车站在一固定时间内来到的乘客数；
放射物在某一时间间隔内发射的粒子数；
每米布的瑕疵点数；
天空中某时段的流星数等等
在一定条件下，都服从或近似服从Poisson分布。

例1.假设电话交换台每小时接到的呼叫次数X服从参数$\lambda=3$的泊松分布，求

(1)每小时恰有4次呼叫的概率；

(2)一小时内呼叫不超过5次的概率。

解： （1） $P\{X=4\}=\dfrac{3^4e^{-3}}{4!}=0.168$

（2） $P\{X\le 5\}=\sum\limits_{k=0}^{5}P\{X=k\}=\sum\limits_{k=0}^{5}\dfrac{3^k e^{-3}}{k!}=0.9161$

例2. 设随机变量$X \sim \pi(\lambda)$，且已知$P\{X=1\}=P\{X=2\}$，求$P\{X=4\}$。

解: 随机变量X的分布律为 $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,\cdots$

由$P\{X=1\}=P\{X=2\}$，得 $\dfrac{\lambda^1}{1!}e^{-\lambda}=\dfrac{\lambda^2}{2!}e^{-\lambda}$

由此得方程 $\lambda^2-2\lambda=0$

解得 $\lambda=2$，$\lambda=0$（舍去）

所以 $P\{X=4\}=\dfrac{2^4}{4!}e^{-2}\approx 0.0902$

例3.设一辆出租车一天内穿过的路口数服从泊松分布$\pi(\lambda)$，各个路口的红绿灯是独立工作的，在每个路口遇到红灯的概率是p。

(1)已知一辆出租车一天内路过了k个路口，求遇到红灯数的分布；

(2)求一辆出租车一天内遇到的红灯数的分布。

解：（1）设这辆出租一天内车路过的路口数是Y，遇到的红灯数是X。每到一个路口是否遇到红灯相当于一次伯努利试验，遇到红灯是关心的事件。则有

$$P\{X=m \mid Y=k\}=C_k^m p^m(1-p)^{k-m}, m=0,1,\cdots,k$$

（2）$\{Y=k\}$，$k=0,1,2,\ldots$构成完备事件组。利用全概率公式得到

$$P\{X=m\}=\sum_{k=0}^{\infty}P\{Y=k\}P\{X=m\mid Y=k\}=\sum_{k=m}^{\infty}P\{Y=k\}P\{X=m\mid Y=k\}$$

$$=\sum_{k=m}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}C_k^m p^m(1-p)^{k-m}=\sum_{k=m}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}\frac{k!}{m!(k-m)!}p^m(1-p)^{k-m}$$

$$=\frac{e^{-\lambda}p^m}{m!}\sum_{k=m}^{\infty}\frac{\lambda^k}{(k-m)!}(1-p)^{k-m}=\frac{e^{-\lambda}p^m}{m!}\sum_{k=m}^{\infty}\frac{\lambda^m\lambda^{k-m}}{(k-m)!}q^{k-m}$$

$$=\frac{(\lambda p)^m e^{-\lambda}}{m!}\sum_{k=m}^{\infty}\frac{(\lambda q)^{k-m}}{(k-m)!}=\frac{(\lambda p)^m e^{-\lambda}}{m!}\sum_{j=0}^{\infty}\frac{(\lambda q)^j}{j!}$$

$$=\frac{(\lambda p)^m e^{-\lambda}}{m!}e^{\lambda q}=\frac{(\lambda p)^m}{m!}e^{-\lambda(1-q)}=\frac{(\lambda p)^m}{m!}e^{-\lambda p}$$

即 $X\sim\pi(\lambda p)$

例4.设有15000件产品，其中有150件次品。现任取100件，求次品数恰好为两件的概率。

解：设X表示任取的100件中的次品数，则有 $X \sim H(100, 150, 15000)$

$$P\{X=2\}=\frac{C_{150}^{2}C_{14850}^{98}}{C_{15000}^{100}}=0.1855$$

产品总数15000很大，150相对较小，能否做如下近似?

$$P\{X=2\}=\frac{C_{150}^{2}C_{14850}^{98}}{C_{15000}^{100}}\overset{?}{\approx}C_{100}^{2}(0.01)^{2}(0.99)^{98}$$

$b(100,0.01)$

定理1 设 $\lim\limits_{N\to\infty}\dfrac{M}{N}=p$, $0<p<1$, 对固定的正整数n和$m=0,1,\ldots,n$都有

$$\lim_{N\to\infty}\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=C_n^m p^m(1-p)^{n-m}$$

证明：

$$\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=\frac{\dfrac{M!}{m!(M-m)!}\dfrac{(N-M)!}{(n-m)!(N-M-(n-m))!}}{\dfrac{N!}{n!(N-n)!}}$$

$$=\frac{n!}{m!(n-m)!}\frac{\dfrac{M!}{(M-m)!}\dfrac{(N-M)!}{(N-M-(n-m))!}}{\dfrac{N!}{(N-n)!}}$$

二. 超几何分布与二项分布，二项分布与泊松分布的关系

$$=\frac{n!}{m!(n-m)!}\frac{\dfrac{M!}{(M-m)!}\dfrac{(N-M)!}{(N-M-(n-m))!}}{\dfrac{N!}{(N-n)!}}$$

$$=\frac{n!}{m!(n-m)!}\frac{M(M-1)\cdots(M-m+1)(N-M)(N-M-1)\cdots(N-M-(n-m)+1)}{N(N-1)\cdots(N-n+1)}$$

$$=C_n^m[\overbrace{\frac{M}{N}\cdot\frac{M-1}{N}\cdots\frac{M-m+1}{N}}^{m\text{项}}][\overbrace{\frac{N-M}{N}\cdot\frac{N-M-1}{N}\cdots\frac{N-M-(n-m)+1}{N}}^{n-m\text{项}}][\overbrace{\frac{N}{N}\cdot\frac{N}{N-1}\cdots\frac{N}{N-n+1}}^{n\text{项}}]$$

二. 超几何分布与二项分布，二项分布与泊松分布的关系

$$C_n^m[\overbrace{\frac{M}{N}\cdot\frac{M-1}{N}\cdots\frac{M-m+1}{N}}^{m\text{项}}][\overbrace{\frac{N-M}{N}\cdot\frac{N-M-1}{N}\cdots\frac{N-M-(n-m)+1}{N}}^{n-m\text{项}}][\overbrace{\frac{N}{N}\cdot\frac{N}{N-1}\cdots\frac{N}{N-n+1}}^{n\text{项}}]$$

$$\because \lim_{N\to\infty}\frac{M}{N}=p \qquad \therefore \frac{M}{N}\cdot\frac{M-1}{N}\cdots\frac{M-m+1}{N}\to p^m \qquad \frac{N}{N}\cdot\frac{N}{N-1}\cdots\frac{N}{N-n+1}\to 1$$

$$\frac{N-M}{N}\cdot\frac{N-M-1}{N}\cdots\frac{N-M-(n-m)+1}{N}\to(1-p)^{n-m}$$

所以有 $$\lim_{N\to\infty}\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=C_n^m p^m(1-p)^{n-m}$$

定理1 设 $\lim\limits_{N\to\infty}\dfrac{M}{N}=p$，$0<p<1$，对固定的正整数$n$和$m=0,1,\ldots,n$都有

$$\lim_{N\to\infty}\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}=C_n^m p^m(1-p)^{n-m}$$

由定理1，当N 很大，n较小时，可以用二项分布近似计算超几何分布。若设$X\sim H(n, M, N)$，令$p=M/N$，则

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\approx C_n^m p^m(1-p)^{n-m}$$

定理2 （泊松定理）设常数$\lambda>0$，n是任意的正整数，记$np_n=\lambda$，则对于任意的非负整数k，均有

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

证明：由$np_n=\lambda$有

$$C_n^k p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

对于固定的正数k，当$n\to+\infty$时

$$\frac{n(n-1)\cdots(n-k+1)}{n^k}\to 1 \quad \left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda} \quad \left(1-\frac{\lambda}{n}\right)^{-k}\to 1$$

所以有

$$\boxed{\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}}$$

定理2 （泊松定理）设常数$\lambda>0$，n是任意的正整数，记$np_n=\lambda$，则对于任意的非负整数k，均有

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!}$$

定理的条件$np_n=\lambda$，意味着当n很大时，p_n很小，此时可以用泊松分布近似计算二项分布。若设$X\sim b(n,p)$，令$\lambda=np$，则有

$$P\{X=k\}=C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,\cdots,n$$

一般的统计书籍认为，$n\geq 50$，$np<5$时就有很好的近似效果。

例5.设有15000件产品，其中150件次品。现任取100件，求次品数恰好为两件的概率。

解：设X为任取100件产品中的次品数。 则易知X服从超几何分布，所以有

$$P\{X=2\}=\frac{C_{150}^{2}C_{14850}^{98}}{C_{15000}^{100}}$$

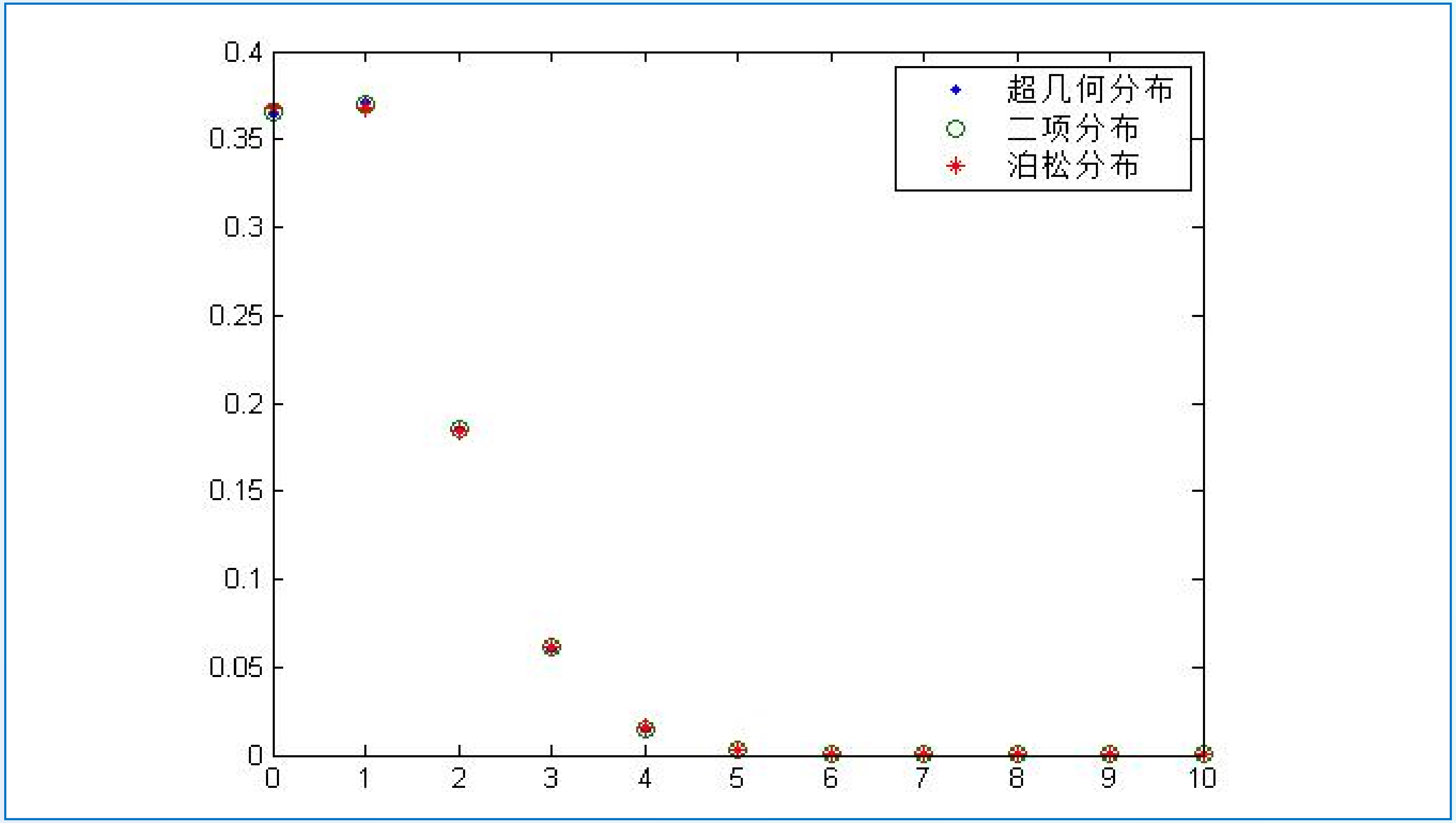
由定理1，令p=150/15000=0.01,即可以用二项分布b(100,0.01)近似X，所以有

$$P\{X=2\}=\frac{C_{150}^{2}C_{14850}^{98}}{C_{15000}^{100}}\approx C_{100}^{2}(0.01)^{2}(0.99)^{98}=0.1849$$

又由定理2知，取λ=100×0.01=1，则用泊松分布π(1)近似二项分布b(100,0.01). 所以有

$$P(X=2)=\frac{C_{150}^{2}C_{14850}^{98}}{C_{15000}^{100}}\approx C_{100}^{2}(0.01)^{2}(0.99)^{98}\approx\frac{e^{-1}}{2}=0.1839$$

二. 超几何分布与二项分布，二项分布与泊松分布的关系



例6.某人进行射击，设每次射击的命中率为0.001，他独立射击了5000次，试求他至少命中1次的概率。

解：设X为5000次射击中命中的次数，则有$X \sim b(5000, 0.001)$

$$P\{X \geq 1\} = 1 - P\{X < 1\} = 1 - P\{X = 0\}$$

$$= 1 - C_{5000}^{0}(0.001)^{0}(0.999)^{5000} \approx 1 - \frac{5^{0}e^{-5}}{0!} \approx 0.9935$$

小概率事件虽不易发生，但重复次数多了，就成了大概率事件。

由此可见日常生活中“提高警惕，防火防盗”的重要性。

三. 分布函数

对于离散型随机变量，用分布律完全描述其统计规律性。方法是枚举出每一个可能的取值，并指出取每个值的概率。

但是有些随机变量的取值不能一一列出，如：测量时的误差ε、元件的寿命T等；其取值充满一个区间，因此不能用分布律描述其统计特性；

而且对于这类随机变量，考察它取某一值的概率没有意义，如测量某一物体的长度时，要求误差必须为0.01，这在原则上不能排除，但实际中可能性极小，只能认为是0。

在实际应用中，对于这类随机变量，更关心的是它取值在某个区间之内的概率。

三. 分布函数

在测量某一物体的长度时，关心测量误差X在给定常数a，b之间的概率，即 $P\{a<X\le b\}$；

当考察元件的寿命时，若规定寿命大于8000小时为合格品，则我们感兴趣的是合格品率，即$T>8000$的概率。

即需要计算如下事件的概率 $P\{X>a\}$, $P\{X\in(a, b]\}$

事实上，如果我们定义 $\boxed{F(x)=P\{X\le x\}}$

则 $P\{X>a\}=1-P\{X\le a\}=1-F(a)$ $P\{X\in(a,b]\}=P\{a<X\le b\}=\underbrace{P\{X\le b\}}_{F(b)}-\underbrace{P\{X\le a\}}_{F(a)}$

分布函数

三. 分布函数

设X是一随机变量，对任意的实数x，令

$$F(x)=P\{X\le x\}$$

则称$F(x)$为X的分布函数(Distribution Function)。

说明:

(1)分布函数是一个普通的函数，正是通过它，可以用数学分析的工具来研究随机变量；

(2)如果将X看作数轴上随机点的坐标，那么分布函数$F(x)$的值就表示X落在区间$(-\infty, x]$的概率；

三. 分布函数

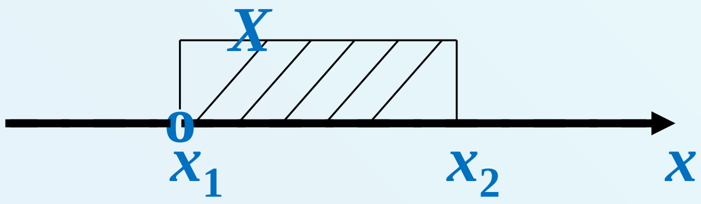
设X是一随机变量，对任意的实数x，令

$$F(x)=P\{X\leq x\}$$

则称$F(x)$为X的分布函数(Distribution Function)。

(3) 对于任意的实数$x_1, x_2(x_1<x_2)$，有

$$P\{x_1 < X \leq x_2\} = P\{X \leq x_2\} - P\{X \leq x_1\} = F(x_2) - F(x_1)$$



因此，只要知道了随机变量X的分布函数，它的统计特性就可以得到全面的描述。

例7.将一枚硬币抛掷三次，X表示三次中正面出现的次数，求X的分布律及分布函数。

解：X的分布律为

X	0	1	2	3
P	1/8	3/8	3/8	1/8

下面求X的分布函数 $F(x)=P\{X\le x\}$

当$x<0$时， $F(x)=P\{X\le x\}=0$

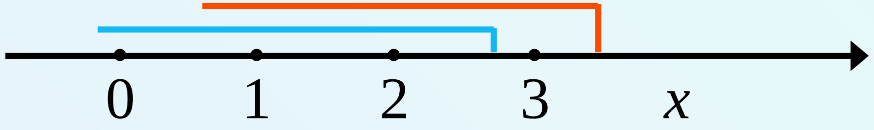
当$0\le x<1$时，$F(x)=P\{X\le x\}=P\{X=0\}=1/8$

当$1\le x<2$时， $F(x)=P\{X\le x\}=P\{X=0\}+P\{X=1\}=1/8+3/8=1/2$

三. 分布函数

当$2 \le x < 3$时，$F(x)=P\{X \le x\}=P\{X=0\}+P\{X=1\}+P\{X=2\}= 1/8+3/8+3/8=7/8$

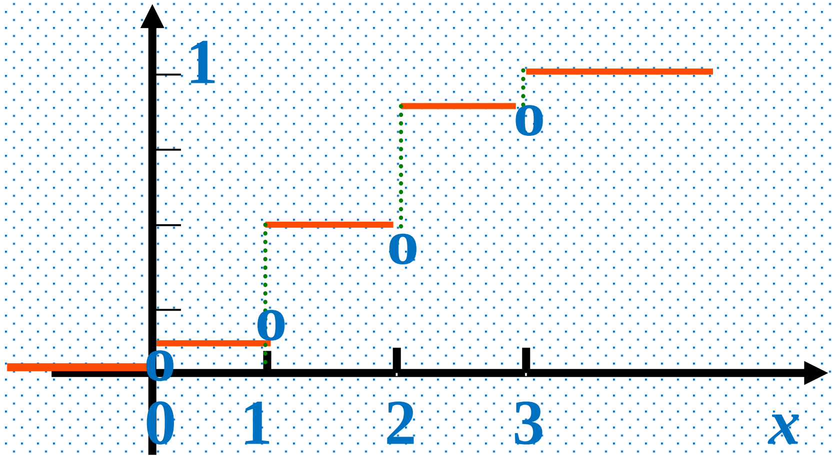
当$x \ge 3$时，$F(x)=P\{X \le x\}=P\{X=0\}+P\{X=1\}+P\{X=2\}+P\{X=3\}= 1/8+3/8+3/8+1/8=1$



所以

$$F(x)=\begin{cases} 0, & x<0 \\ 1/8, & 0 \le x < 1 \\ 4/8, & 1 \le x < 2 \\ 7/8, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

分布函数$F(x)$的图形



三. 分布函数

离散型随机变量的分布函数

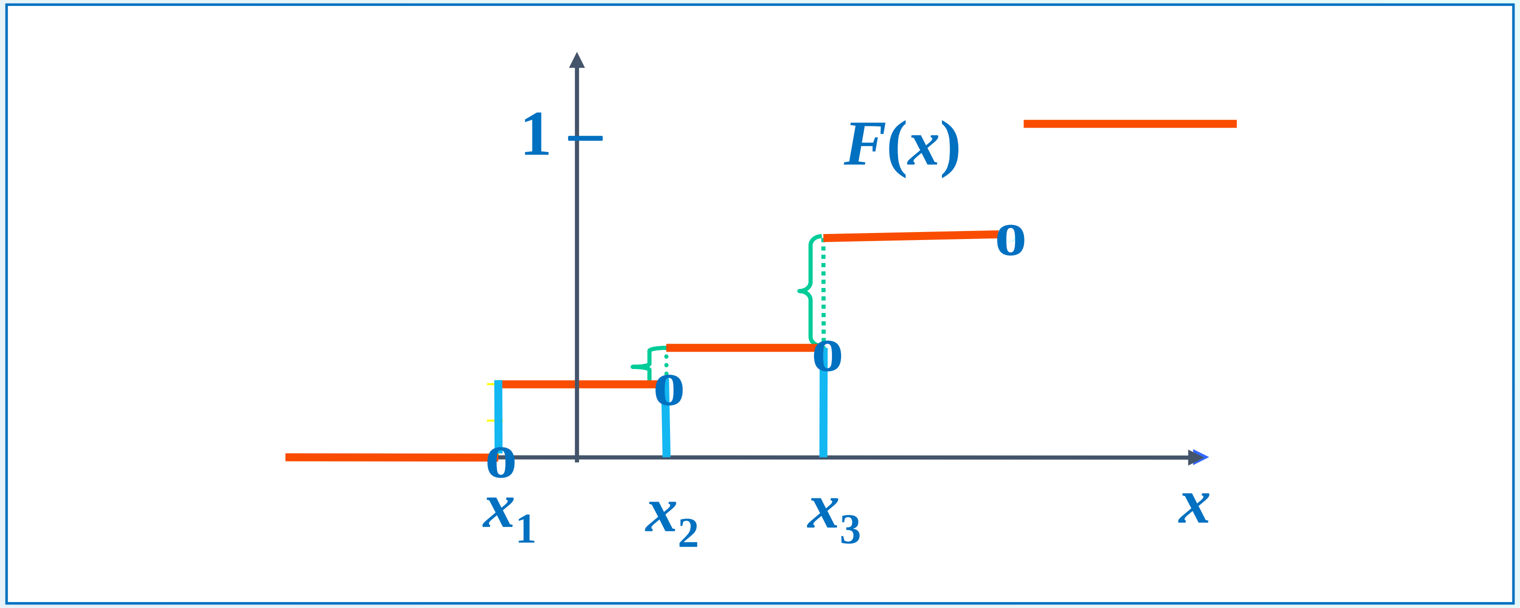
若X的分布律为 $P\{X=x_i\}=p_i,\quad i=1, 2, \ldots,$

则X的分布函数为 $F(x)=P\{X\le x\}=\sum\limits_{x_i\le x}P\{X=x_i\}=\sum\limits_{x_i\le x}p_i$

注 意

在由分布律求解分布函数$F(x)$的时候，自变量x的取值范围一般是左闭右开的区间。

三. 分布函数



离散型随机变量X的分布函数是单调增加的，右连续的，具有跳跃型间断点$\{x_i$, i=1,2,…$\}$的阶梯函数，在间断点处的跃度为

$$p_i = P\{X = x_i\} = F(x_i) - F(x_i - 0)$$

随机变量分布函数$F(x)$的性质

1. 单调性

若$x_1<x_2$，则$F(x_1)\le F(x_2)$

2. 非负性

对任意的实数x，均有$0\le F(x)\le 1$，且 $\lim\limits_{x\to-\infty}F(x)=0,\lim\limits_{x\to\infty}F(x)=1$

3. 右连续性

对任意的实数x_0，有 $\lim\limits_{x\to x_0^+}F(x)=F(x_0)$

随机变量分布函数$F(x)$的性质

注 意

如果一个函数具有上述性质，则一定是某个随机变量X的分布函数。也就是说，上述三条性质是鉴别一个函数是否是某随机变量的分布函数的充分必要条件。

三. 分布函数

给定随机变量X的分布函数的概念后，有关X的各种事件的概率都可以用分布函数表示。例如：对任意实数a和b, 有

$$P\{a < X \le b\} = F(b) - F(a)$$

$$P\{a < X < b\} = F(b-0) - F(a)$$

$$P\{a \le X \le b\} = F(b) - F(a-0)$$

$$P\{a \le X < b\} = F(b-0) - F(a-0)$$

$$P\{X > a\} = 1 - F(a)$$

$$P\{X \ge a\} = 1 - F(a-0)$$

$$P\{X = a\} = F(a) - F(a-0)$$

特别地，若$F(x)$在a处连续　则 $P\{X = a\} = 0$

例8.设 $F(x)=\begin{cases}\sin x, & 0\le x\le \pi\\ 0, & \text{其它}\end{cases}$ 试说明$F(x)$能否是某个随机变量的分布函数。

解：注意到函数$F(x)$在$[\pi/2,\pi]$上是减函数，不满足性质(1)；

或者 $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=0$ 不满足性质(2)，即$F(x)$不能是分布函数。

例9.设随机变量X的分布函数为 $F(x)=A+B\operatorname{arctg}x$, $-\infty<x<+\infty$. 试求 (1)常数A与B的值；(2)$P\{1<X\leq 1\}$.

解：（1）因为 $\lim\limits_{x\to-\infty} F(x)=A+B\cdot(-\dfrac{\pi}{2})=0$ $\qquad \lim\limits_{x\to\infty} F(x)=A+B\cdot(\dfrac{\pi}{2})=1$

所以 $A=\dfrac{1}{2}, B=\dfrac{1}{\pi}$

（2） $P\{-1<X\leq 1\}=F(1)-F(-1)=\dfrac{1}{2}$

例10.设离散型随机变量X的分布函数为 $F(x)=\begin{cases} 0, & x<-1 \\ a, & -1\le x<1 \\ \frac{2}{3}-a, & 1\le x<2 \\ a+b, & x\ge 2 \end{cases}$

且$P\{X=2\}=0.5$。试确定常数a、b的值，并求X的分布律。

解：由分布函数的性质$F(+\infty)=1$知：$a+b=1$

又由右连续性得 $P\{X=2\}=F(2)-F(2-0)=a+b-(\frac{2}{3}-a)=\frac{1}{2}$

解方程组得

$$a=\frac{1}{6},\quad b=\frac{5}{6}$$

所以分布函数为

$$F(x)=\begin{cases}0, & x<-1\\ \frac{1}{6}, & -1\le x<1\\ \frac{1}{2}, & 1\le x<2\\ 1, & x\ge 2\end{cases}$$

分布函数的间断点就是X可能的取值：-1，1，2

$$P\{X=-1\}=F(-1)-F(-1-0)=\frac{1}{6}-0=\frac{1}{6}$$

$$P\{X=1\}=F(1)-F(1-0)=\frac{1}{2}-\frac{1}{6}=\frac{1}{3}$$

$$P\{X=2\}=F(2)-F(2-0)=1-\frac{1}{2}=\frac{1}{2}$$

所以X的分布律为

X	-1	1	2
P	1/6	1/3	1/2

作业： 17,18,20,21,22

第 6 讲

谢谢观看