概率统计课件之二

第二章 随机变量及分布

主讲教师 邓小艳

为进一步研究随机试验的结果，我们将引入随机变量这一重要的概念，并对其概率分布进行研究。

随机变量及分布

§2-1 随机变量

§2-2 一维离散型R.V.及概率分布

§2-3 一维连续型随机R.V.及概率分布

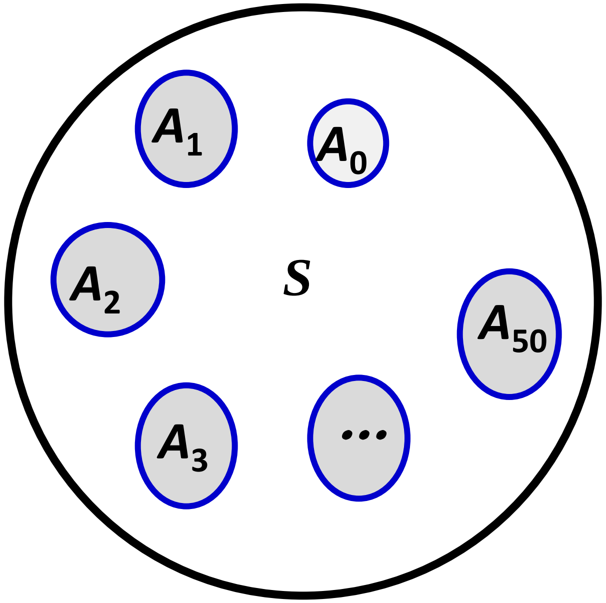
§2-5 二维离散型及R.V.概率分布

§2-6 二维连续型R.V.及概率分布

§2-8 相互独立的随机变量

§2-9 随机变量函数的分布

引例1：设有500件产品，其中60件次品，从中任取50件。



E: 从500件产品中任取50件

$$N(S) = C_{500}^{50}$$

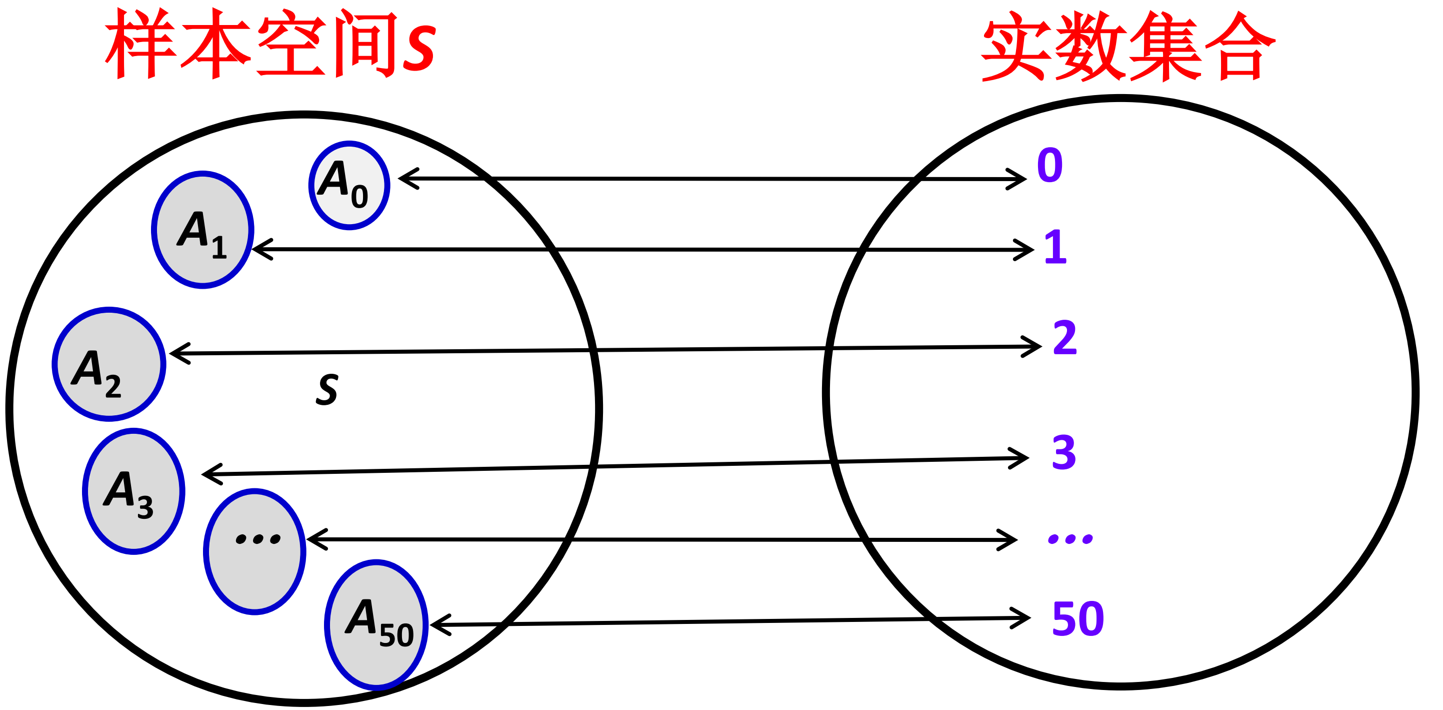
$$S = \{A_0, A_1, A_2, \cdots, A_{50}\}$$

A_k ：“从500件产品中任取50，恰好取到k件次品”

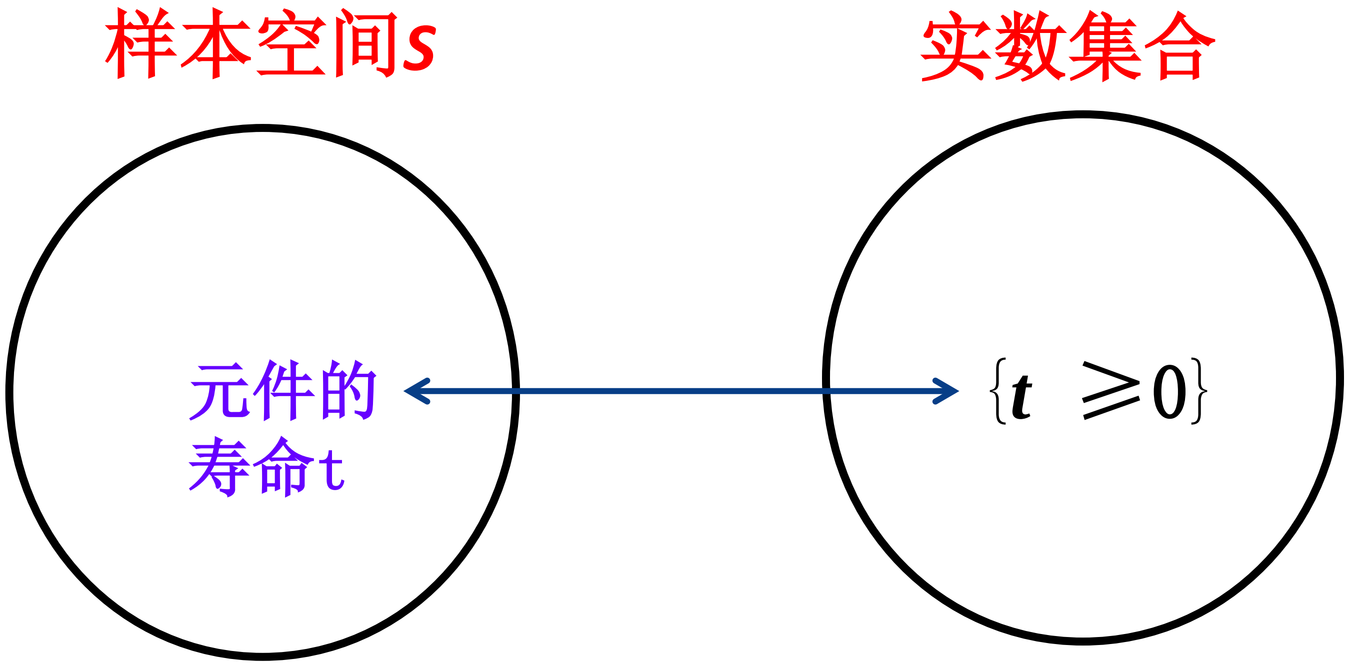
$$(k = 0,1,2,\cdots,50)$$

设X:“恰好取到的次品数”，$X = 0,1,2,\cdots,50$

设X:"恰好取到的次品数"，$X = 0,1,2,\cdots,n$



引例2：检查某种元件的寿命。



两个引例的相同点：用变量表示随机试验结果，并且其取值随结果的不同而不同.

随机变量的定义

设E为随机试验，样本空间为 S. 若对S中的每个e，均有一个实数$X(e)$与之对应，这样得到一个定义在样本空间 上的实值单值函数$X(e)$，则称$X(e)$为随机变量，简记为*R.V.X*。

〈注〉(1) 直观而言，*R.V.*是用来表示随机试验结果的变量，它的取值随试验结果的不同而不同。

(2) 随机事件通过随机变量来表示，当随机变量取到某个值，对应的某随机事件就发生了.

(3)引入随机变量后，对随机现象统计规律的研究，就由对事件及事件概率的研究转化为对随机变量及其取值规律的研究. 随机事件是从静态的观点来研究随机现象，而随机变量则是从动态的观点来研究随机现象.

2. 分类

离散型***R.V.*** ——X的取值为有限个或可列个

连续型***R.V.*** ——X在某个区间上取值(稠密性)

随机变量及分布

§ 2-1 随机变量

§ 2-2 一维离散型R.V.及概率分布

§ 2-3 一维连续型随机R.V.及概率分布

§ 2-5 二维离散型及R.V.及概率分布

§ 2-6 二维连续型R.V.及概率分布

§ 2-8 相互独立的随机变量

§ 2-9 随机变量函数的分布

例1 **(1)掷一枚硬币，可能正面向上，也可能反面向上，用变量表示这一结果；**

令**X=0** —— “正面向上”

X=1 —— “反面向上”

X的分布律为：

X	0	1
$P(X)$	p_0	p_1

(2)掷一枚均匀的骰子，用变量表示试验结果；

解：设X：“出现的点数”

X	1	2	3	4	5	6
$P(X)$	p_1	p_2	p_3	p_4	p_5	p_6

(3)将一枚硬币抛两次，用变量表示试验结果。

解：设Y“出现的正面的次数”

X	0	1	2
$P(X)$	p_0	p_1	p_2

研究目的：

(1) X可能取哪些值 —— 事件A（出现哪种结果）

(2) X取这些值的概率 —— P(A)（出现这种结果的概率）

概率分布律

一、概率分布律

1. 定义：设 $x_k(k=1,2,3,\cdots)$ 为随机变量R.V.X的所有可能取值，若 $P\{X=x_k\}=p_k$ 满足:

(1) $p_k \geqslant 0, k=1,2,\cdots$

(2) $\sum_{k=1}^{\infty} p_k = 1$

这两个性质可用于判断是否为分布律

则称 $P\{X=x_k\}=p_k, k=1,2,\cdots$ 为随机变量 X 的概率分布律，简称分布律。

分布律的表示方法

(1) 公式法

$$P\{X = x_k\} = p_k, k = 1,2,\cdots$$

(2) 表格

X	X_1	x_2	x_3	…
$P(X)$	p_1	p_2	P_3	…

(3)线条图、概率直方图

例2：设盒中装有两个红球和三个白球，现从中任取三球，求取得红球个数的概率分布。

解：设X：“取得红球个数”

X的所有可能取值为：{0，1，2}

$$P\{X=0\}=\frac{C_3^3}{C_5^3}=\frac{1}{10} \qquad P\{X=1\}=\frac{C_2^1C_3^2}{C_5^3}=\frac{6}{10}$$

$$P\{X=2\}=\frac{1}{10}$$

$\therefore$ X的分布律为：

X	0	1	2
$P(X)$	1/10	3/5	3/10

例2：一个口袋中有5只同样大小的球，编号为1，2，3，4，5，从中同时取出3只球，用*X*表示取出的3只球中的最大号码，求*X*的分布律。

解：*X*的所有可能取值为：{3，4，5}

$$P\{X=3\}=P\{\text{取出球中最大号码是}3\}=\frac{C_2^2}{C_5^3}=\frac{1}{10}$$

$$P\{X=4\}=\frac{C_3^2}{C_5^3}=\frac{3}{10} \qquad P\{X=5\}=\frac{C_4^2}{C_5^3}=\frac{6}{10}$$

∴ *X*的分布律为：

X	**3**	**4**	**5**
***P*(*X*)**	**1/10**	**3/10**	**6/10**

二、几种常见的离散型R.V.的概率分布

1. 0-1分布（两点分布，伯努利分布）

X的概率分布律为：

X	0	1
$P(X)$	$1-p$	p

或

$$P\{X=k\}=C_1^k p^k q^{1-k},\ k=0,1,\ q=1-p$$

2. 超几何分布

M 次
N-*M* 正 → 任取n(n<N)件

令X表示“取到的次品数”

则X的概率分布为：

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

3. 二项分布

定义：在n重伯努利试验中，事件A发生的次数X有分布律：

$$P\{X=k\}=C_n^k p^k q^{n-k}, k=0,1,2\cdots,n, q=1-p$$

则称此为参数是(n, p)的二项分布。记作：

$$X \sim B(n, p)$$

例3：一本100页的书上有200个错字，每个错字等可能的出现在每一页上(每个错字是否出现在指定页上相互独立)。求在指定页上：(1)出现错字次数的分布律；(2)恰有三个错字的概率；(3)错字个数在2~6之间的概率。

分析: 将观察每一个错字是否出现在指定页上看成一次试验，它是伯努利试验。令 A：“某错字在指定页上出现”，则P(A)=1/100，因此，观察200个错字在指定页上是否出现，是n=200，p=1/100的200重伯努利试验。设X：“指定页上出现错字次数”，则：$X \sim B(200,1/100)$

解：设X：“指定页上出现错字次数”，则：

$$X \sim B(200, 1/100)$$

(1) 出现错字次数的分布律为：

$$P\{X = k\} = C_{200}^{k} \times 0.01^{k} \times 0.99^{200-k}, k = 0,1,2 \cdots ,200$$

(2) 恰有三个错字的概率为：

$$P\{X = 3\} = C_{200}^{3} \times 0.01^{3} \times 0.99^{197}$$

(3) 错字个数在2~6之间的概率为：

$$P\{2 \leqslant k \leqslant 6\} = \sum_{k=2}^{6} C_{200}^{k} \times 0.01^{k} \times 0.99^{200-k}$$

4. 泊松分布

(1)定义：若R.V.X的概率分布为：

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

则称X服从参数为λ的泊松分布。记作：$X \sim \pi(\lambda)$

泊松定理

定理：设在n重伯努利试验中，事件A发生的概率为p (0<p<1)，事件A在n重伯努利试验中恰好发生k次的概率为 $P_n(k)$，当$n \to \infty$时，有：

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \lambda = np$$

(2) 二项分布的近似计算

若$X \sim B(n, p)$，则当n较大，p较小时有：

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

<注>实际计算时，当n>=10，p<=0.1时，可利用查泊松分布表求二项分布的近似值。

例3续：若X~B(200，1/100)，这里 $\lambda = np = 2$ ，则查表得：

(2) 恰有三个错字的概率为：

$$P\{X = 3\} = P\{X \leqslant 3\} - P\{X \leqslant 2\}$$

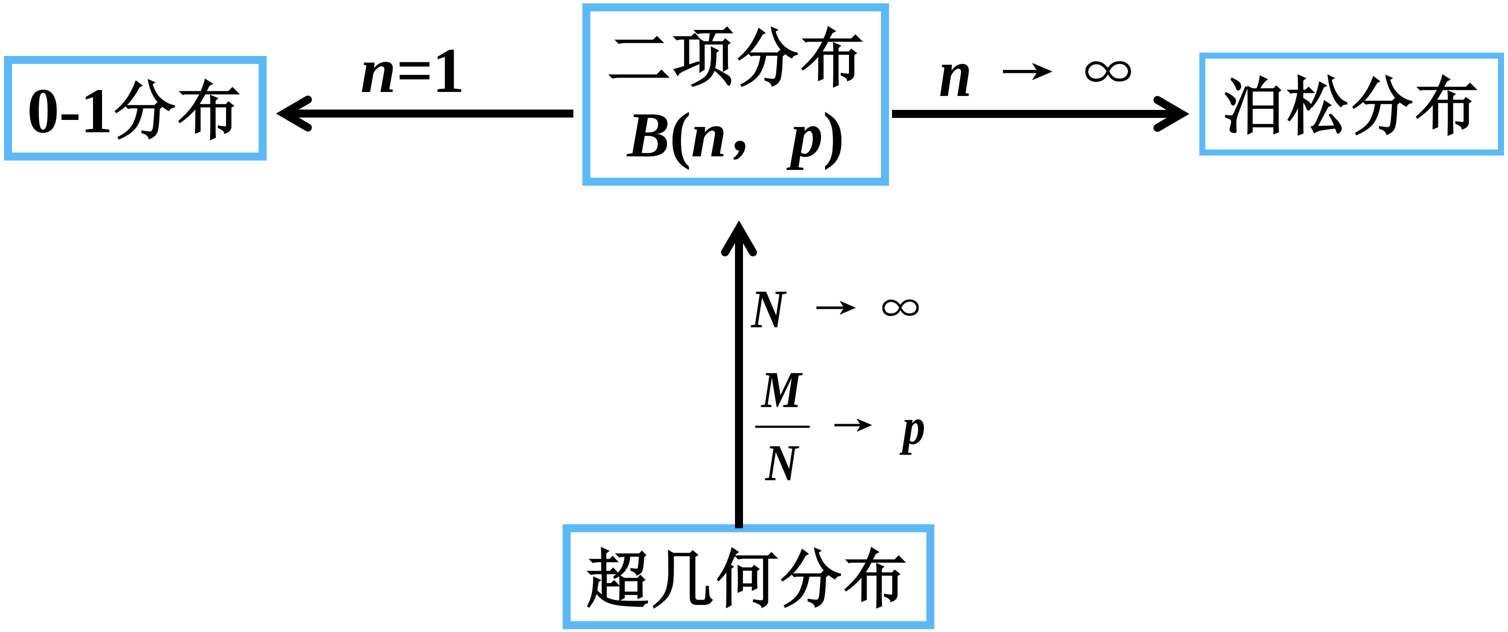
$$\approx 0.8571 - 0.6767 = 0.1804$$

(3) 错字个数在2~6之间的概率为：

$$P\{2 \leqslant k \leqslant 6\} = P\{k \leqslant 6\} - P\{k \leqslant 1\}$$

$$\approx 0.9955 - 0.4060 = 0.5895$$

5、几种分布的关系



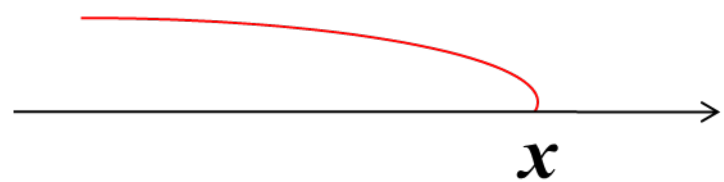
三、*R.V.*的分布函数

1. 定义及性质

定义： 设X是一个$R.V.$，x为任意实数，则称

$$F(x)=P\{\overset{R.V.}{\uparrow\atop X}\leqslant \overset{\text{实变量}}{\uparrow\atop x}\}$$

为$R.V.\ X$的分布函数。



注意： $F(x)$ 的值表示 随机变量X落入区间$(-\infty, x]$的概率。

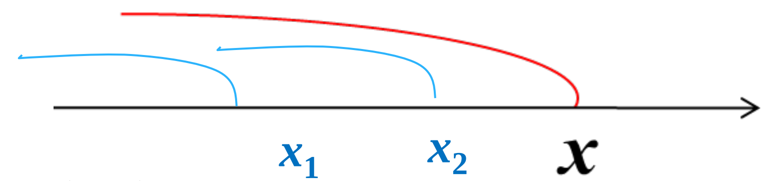
性质:

① $0 \le F(x) \le 1$ 且 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$, $F(\infty)=\lim\limits_{x\to\infty}F(x)=1$

② **$F(x)$是单调不减函数.**

$$F(x_2)-F(x_1)=P\{X\leqslant x_2\}-P\{X\leqslant x_1\}$$
$$=P\{x_1<X\leqslant x_2\}\ge 0$$

□ $F(x_2)\ge F(x_1)$



③ **X在任一区间 上的概率**

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

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

$$P\{x_1\le X\le x_2\}=F(x_2)-F(x_1)+P\{X=x_1\}$$

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

④ **$F(x)$ 是右连续函数**

$$\lim_{x\to x_0^+}F(x)=F(x_0),\qquad(-\infty<x_0<\infty).$$

离散型随机变量的分布函数

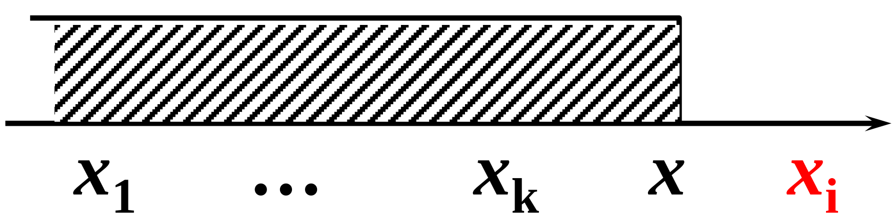
设X是离散型随机变量，其分布律为

$$P\left\{X=x_k\right\}=p_k, k=1,2,\cdots$$

则X的分布函数为：

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

即小于或等于x的那些点处的概率之和.



例4： X**的分布律为：**

X	-1	2	3
p_k	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

求： **(1)** X**的分布函数；**

(2) $P\{X \leqslant \frac{1}{2}\}$；$P\{X > 1\}$；$P\{\frac{3}{2} < X \leqslant \frac{5}{2}\}$；

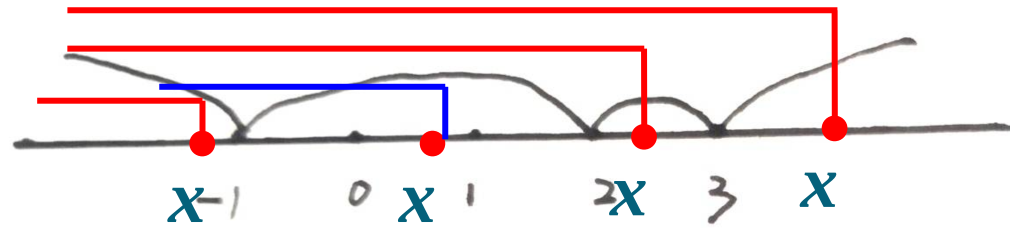
$P\{2 \leqslant X \leqslant 3\}$；$P\{1 < X < \frac{5}{2}\}$；

$P\{1 < X < 3\}$；

例4：X的分布律为：

X	-1	2	3
p_k	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

解：(1) X的分布函数为



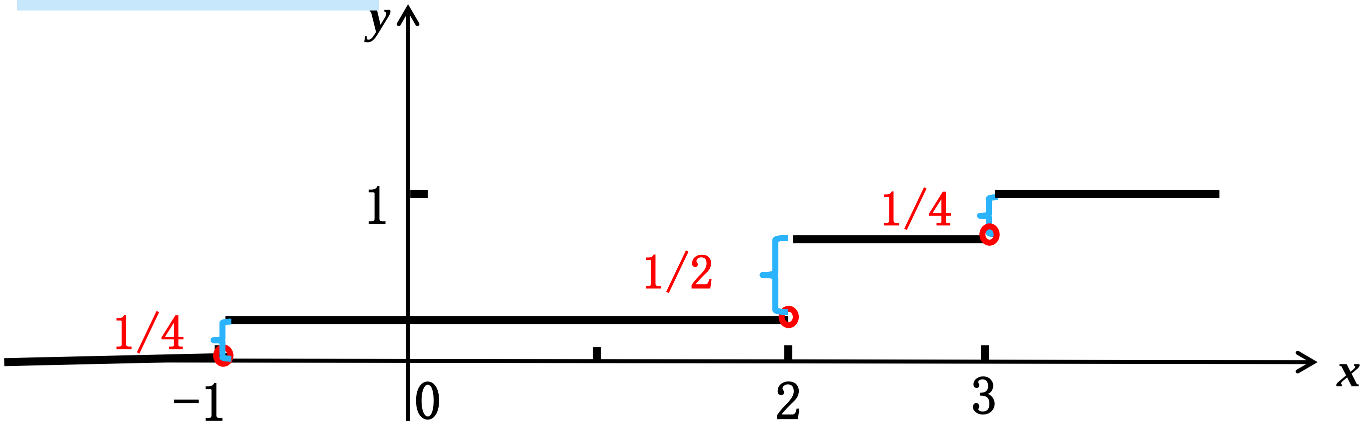
$$F(x) = P\{X \leqslant x\}$$

$$= \begin{cases} 0 & x < -1 \\ \frac{1}{4} & -1 \leqslant x < 2 \\ \frac{1}{4} + \frac{1}{2} & 2 \leqslant x < 3 \\ \frac{1}{4} + \frac{1}{2} + \frac{1}{4} & x \geqslant 3 \end{cases}$$

左闭右开
右连续

步骤：①定义②分区间（左闭右开）③求概率（和事件的概率）

分布函数图



离散型R.V.分布函数的性质：

① 分布函数是分段函数，分段区间是由X的取值点划分成的左闭右开区间；

② 函数值从0到1，逐段递增，图形呈现阶梯形跳跃递增；

③ 分布函数值跳跃高度是X取值区间中新增点的对应概率值；

④ 右连续函数.

(2) 利用分布函数的性质求解

$$P\{X \leqslant \frac{1}{2}\} = F(\frac{1}{2}) = \frac{1}{4}$$

$$P\{X > 1\} = 1 - F(1) = 1 - \frac{1}{4} = \frac{3}{4}$$

$$P\{\frac{3}{2} < X \leqslant \frac{5}{2}\} = F(\frac{5}{2}) - F(\frac{3}{2}) = \frac{1}{2}$$

$$P\{2 \leqslant X \leqslant 3\} = F(3) - F(2) + P\{X = 2\} = \frac{3}{4}$$

$$P\{1 < X < \frac{5}{2}\} = F(\frac{5}{2}) - F(1) - P\{X = \frac{5}{2}\} = \frac{1}{2}$$

$$P\{1 < X < 3\} = F(3) - F(1) - \mathrm{P}\{X = 3\} = \frac{1}{2}$$

(2) 另解：利用和事件概率求解

$$P\{X \leqslant \frac{1}{2}\} = P\{X = -1\} = \frac{1}{4}$$

$$P\{X > 1\} = P\{\{X = 2\} \cup \{X = 3\}\} = P\{X = 2\} + P\{X = 3\} = \frac{3}{4}$$

$$P\{\frac{3}{2} < X \leqslant \frac{5}{2}\} = P\{X = 2\} = \frac{1}{2}$$

$$P\{2 \leqslant X \leqslant 3\} = P\{X = 2\} + P\{X = 3\} = \frac{3}{4}$$

$$P\{1 < X < \frac{5}{2}\} = P\{X = 2\} = \frac{1}{2}$$

$$P\{1 < X < 3\} = P\{X = 2\} = \frac{1}{2}$$

例5：设10件产品中恰好有3件次品，现接连进行不放回抽样，每次抽一件，直到抽到正品为止，求：

解: (1) X的所有可能取值为：{1，2，3，4}

$$P\{X = 1\} = \frac{C_7^1}{C_{10}^1} = \frac{7}{10} \qquad P\{X = 2\} = \frac{C_3^1 \times C_7^1}{C_{10}^1 \times C_9^1} = \frac{7}{30}$$

$$P\{X = 3\} = \frac{C_3^1 \times C_2^1 \times C_7^1}{C_{10}^1 \times C_9^1 \times C_8^1} = \frac{7}{120}$$

$$P\{X = 4\} = \frac{C_3^1 \times C_2^1 \times C_1^1 \times C_7^1}{C_{10}^1 \times C_9^1 \times C_8^1 \times C_7^1} = \frac{1}{120}$$

∴ X的分布律为：

X	1	2	3	4
$P(X)$	7/10	7/30	7/120	1/120

(2) *X*的分布函数；

$$F(x)=P\{X\leqslant x\}=\begin{cases} 0 & x<-1 \\ \dfrac{7}{10} & -1\leqslant x<2 \\ \dfrac{7}{10}+\dfrac{7}{30} & 2\leqslant x<3 \\ \dfrac{7}{10}+\dfrac{7}{30}+\dfrac{7}{120} & 3\leqslant x<4 \\ 1 & x\geqslant 4 \end{cases}$$

(3) $P\{X=3.5\}=0$

$$P\{X>2\}=1-F(2)=\frac{1}{15}$$

$$P\{1<X\leqslant 3\}=F(3)-F(1)=\frac{7}{24}$$

$$P\{1<X<3\}=F(3)-F(1)-P\{X=3\}=\frac{7}{30}$$

小　　结

离散型随机变量

- 定义
- 分布律
- 几种常见的离散型随机变量的分布
  - **0-1** 分布
  - 二项分布
  - 超几何 分布
  - 泊松分布
- **4.**分布函数
  - 定义
  - 性质
  - 离散型**R.V.**的分布函数