第五课

## 2．1 随机变量

现在把事件与实数接合起来。

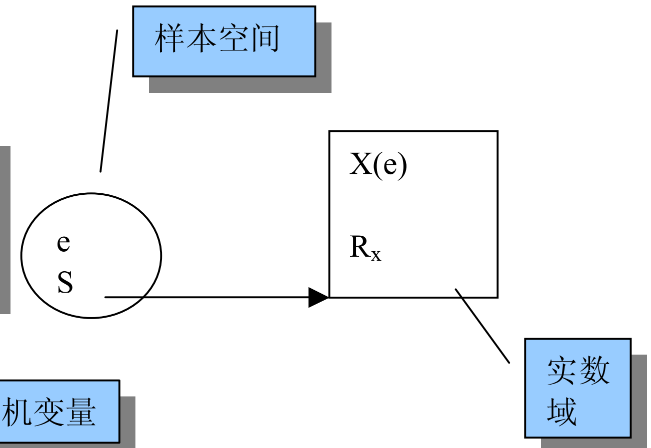
例 1：“抛硬币”，结果“出现 H”或“出现 T”，用数 1 表示“出现 H”；数 0 表示“出现 T”，引入一个变量 X，如：S={e}

使得 $X = X(e) = \begin{cases} 1 & 当 e = H \\ 0 & 当 e = T \end{cases}$ X(e)

随机变量

定义：

设 E 是随机试验，它的样本空间是 S={e}，如果对于每一个 e∈S，有实数 X(e)和它对应，这样得到一个定义在 S 上的实值单值函数 X(e)，称 X(e)为随机变量。




随机变量表示方法例题：

一般用 X，Y，Z 表示随机变量

例 2：某射手每次射击击中目标的概率是 0.8，现连续射击 30 次，则]击中目标的次数 X 就是一个随机变量。

{X=k} ⟶ “30 次恰好有 k 次击中” ； {X<5} ⟶ “30 次击中次数小于 5 次”

## 2．2 随机变量的分布函数

定义：设 X 是一随机变量，x 是任意实数，称函数 $F(x)=P\{X\le x\}$ $-\infty<x<+\infty$ 为 X 的分布函数。

意义：表示 X 落在区间（-∞，x）上的概率。

样本空间 $\xrightarrow{X(e)}$ 实数域 $\xrightarrow{F(x)}$ [0，1] 区间

性质：

1、0≤F(x)≤1 (-∞<x<+∞)

2、若 $x_1<x_2$，有 $F(x_1)\le F(x_2)$

3、$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0\,,F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$

4、F(x)右连续函数，即 F(x+0)=F(x)

例：有 5 件产品产品，其中两件是次品，从中任取两件，这两件产品中的次品数 X 是一随机变量，取值 0，1，2 且：

没取到次品 $P\{X=0\}=\dfrac{C_3^2}{C_5^2}=\dfrac{3}{10}$

取到一个次品 $P\{X=1\}=\dfrac{C_2^1C_3^1}{C_5^2}=\dfrac{6}{10}$

取到二个次品 $P\{X=2\}=\dfrac{C_2^2}{C_5^2}=\dfrac{1}{10}$

由分布函数定义：

要讨论整个实数域，由于 X 取 0，1，2 所以按这样分范围。

分布函数为：

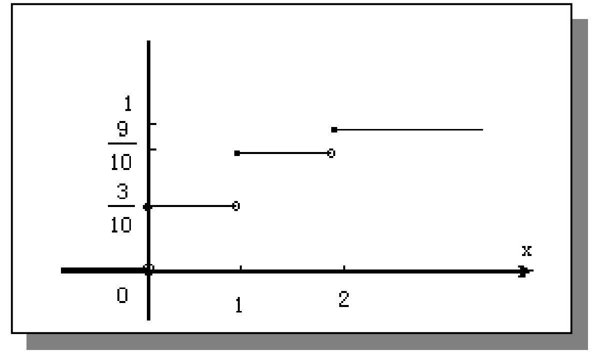
$$F(x)=\begin{cases}0 & x<0\\ \dfrac{3}{10} & 0\le x<1\\ \dfrac{9}{10} & 1\le x<2\\ 1 & x\ge 2\end{cases}$$

当 $x<0$ 时，$P\{X\leqslant x\}=P\{\phi\}=0$

当 $0\leqslant x<1$ 时，$P\{X\leqslant x\}=P\{X=0\}=3/10$

当 $1\leqslant x<2$ 时，$P\{X\leqslant x\}=P\{X=0\}+P\{X=1\}=9/10$

当 $2\leqslant x$ 时，$P\{X\leqslant x\}=P\{X=0\}+P\{X=1\}+P\{X=2\}=1$




一般，对于 $P\{a<X\leqslant b\}=P\{X\leqslant b\}-P\{X\leqslant a\}=F(b)-F(a)$

X 落入(a,b)的概率

## 2．3 离散型随机变量的概率分布

定义：

有些随机变量 X，可能取得的值是有限或可列（可数）的，叫离散型随机变量

### 一、 概率分布

例 1：一射手进行打靶练习，规定射入区域 $e_1$ 得 2 分，射入 $e_2$ 得 1 分，脱靶得 0 分（$e_3$）。射击一次所得的分数 X 是一个离散型的随机变量，X 可能取的值为 0，1，2，然而在射击前，X 取什么值不能预知，但 X 取各个可能值的概率可被确定：

| X | 0 | 1 | 2 |
|---|---|---|---|
| $p_k$ | 0 | 0.2 | 0.8 |

一般：

定义：

设离散型随机变量 X 所有可能取的值为 $x_k$ $(k=1,2,\ldots.)$，X 取各个可能值的概率，即事件$\{X=x_k\}$概率为 $P\{X=x_k\}=p_k$ $(k=1,2,\ldots.)$，如满足下面两条件：

1、$p_k\geqslant 0$ $k=1,2,\ldots$ ；2、$\sum_{k=1}^{\infty} p_k=1$

称 $P\{X=x_k\}=p_k$ 为离散型随机变量 X 的概率分布或分布律。

可以用表来表示：

| X | $X_1$ | … | $x_n$ | … |
|---|---|---|---|---|
| $p_k$ | $P_1$ | … | $p_n$ | … |

分布函数为 $F(x)=P\{X\le x\}=\sum_{x_k\le x}P\{X=x_k\}$

有许多实际问题都可以用到这些分布模型。

## 二、 个重要的离散型分布

### 1 两点分布（0-1）分布

新生婴儿的性别登记，检查产品的质量是否合格，抛硬币等问题，都可以用（0-1）分布来描述。

X 只可能取 0 和 1，概率分布为：P{X=1}=p，P{X=0}=1-p　　(0<p<1)

称 X 服从（0-1）分布

### 2 离散型均匀分布

这不是我们前面学习过的古典概型吗。

随机变量 X 的分布律为：

| X | $x_1$ | $x_2$ | ... | $x_n$ |
|---|---|---|---|---|
| p | 1/n | 1/n | … | 1/n |

### 3 二项分布

（1）努利试验：

设试验 E 只有二个可能的结果：A 及 $\overline{A}$，记 P(A)=p，$P(\overline{A})=1-p=q$ （0<p<1），
将 E 独立地重复进行 n 次，则称这一串重复的试验为 n 重贝努利试验，简称贝努利试验（Bernoulli）。

贝努利试验是一种很重要的数学模型，它有广泛的应用，是研究得最多的模型之一。

（2）二项式分布：

设 X 为 n 重贝努利试验中事件 A 出现的次数，概率分布为

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

称 X 服从二项分布，记为 X~B（n,p）

性质：

1、 $p_k\geqslant 0$

2、 $\sum_{k=0}^{n}p_k=\sum_{k=0}^{n}C_n^k p^k q^{n-k}=(p+q)^n$

例 2：某人进行射击，设每次射中的概率为 0.02，独立射击 400 次，试求击中的次数大于等于 2 的概率。

解：设 X 为击中次数，X~B(n,p)=B(400,0.02)，分布律为：$P\{X=k\}=C_{400}^{k}0.02^{k}0.98^{400-k}$ (k=0,1,2,…,400)

于是所求概率为：$P\{X\geqslant 2\}=1-P\{X<2\}=1-[P\{X=0\}+P\{X=1\}]=1-(0.98^{400}+400\times 0.02\times 0.98^{399})$

下面的计算太难，先见下面：

> 二项分布是在贝努利试验的前提下得来的，注意贝努利试验是假设把时间这一空间划分成小的间隔，假设一个事件在每一间隔内出现或不出现。如果事件可以在任何时刻或空间任一点出现，在任何给定的间隔里事件出现的可能多于一次，这种情况下用贝努利模拟就不如用下面的泊松分布来模拟了。

## 4 泊松分布

> 随机变量 X 可能取值为 0,1,2,…,n,…，取各个值的概率为
>
> $$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!}$$ λ >0, k=0,1,2,…,n,…，称 X 服从参数为 λ 的泊松分布（Poisson）

> 性质：
>
> 1. $p_k\geqslant 0$
>
> 2. $\sum_{k=0}^{\infty}p_k=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{-k}}{k!}=e^{-\lambda}e^{\lambda}=1$

泊松分布可作为二项分布的极限分布：如取 np= λ 当 n→∞时 （p 很小，n 很大时）有

$$C_n^k p^k q^{n-k} = P\{X=k\} \approx \frac{\lambda^k e^{-\lambda}}{k!}$$ （证明见：书 41 页）

对例 2 有 n=400，p=0.02 ,np=8=λ ,P{X=0}≈$e^{-8}$ ,P{X=1}≈$8e^{-8}$

所以：P{X≥2}≈1-($e^{-8}$+$8e^{-8}$)=0. 997

结果说明两点：

1 每次命中率很小（0.02），但射击 400 次，击中目标二次几乎可以肯定；

2 如果真的在 400 次射击时，击中目标竟不到两次，那么说明每次射中率为 0.02 这一结果都是不合理的（比这一结果还要低）。

例 3：为了保证设备正常工作，需配备适量的维修工人，现有同类型设备 300 台，各台工作是相互独立的，发生故障的概率都是 0.01。在通常情况下一台设备的故障可由一个人来处理，问至少需要配备多少工人，才能保证当设备发生故障但不能及维修的概率小于 0.01?

解：设需要配备 N 个维修工，记同一时该发生故障的设备台数为 X，那么 X~B(300,0.01)，

要解决的问题为：P{X>N}≤0.01，当 n 很大，p 很小可用 λ=np=3 泊松分布来近似

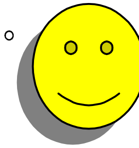
$$P\{X>N\} = 1 - P\{X \le N\}$$

$$\approx 1 - \sum_{k=0}^{N} \frac{3^k e^{-3}}{k!}$$

$$= \sum_{k=N+1}^{\infty} \frac{3^k e^{-3}}{k!} \le 0.01$$ 查附表 3（372 页）

| x | λ=2. 5 | λ=3. 0 | λ=3. 5 |
|---|---|---|---|
| 7 | 0.014187 | 0.033509 | 0.065288 |
| 8 | 0.004247 | 0.011905 | 0.026739 |
| 9 | 0.001140 | 0.003803 | 0.009874 |
| … | … | … | … |

所以：X=N+1=9，有 N=8，即要达到要求，应配备 8 人工人。



通过这个问题的学习，大家可以发现用概率的方法来处理实际问题是非常有效的。

例 4：尽管在几何教科书中已经讲过用圆规和直尺三等分一个任意角是不可能的。但每一年总有一些“发明者”撰写关于用圆规和直尺将角三等分的文章。设某地区每年撰写此类文章的篇数 X 服从参数为 6 的泊松分布。求明年没有此类文章的概率。

解：因为：X~P(6)

所以：$P\{X=0\}=\dfrac{6^0 e^{-6}}{0!}=e^{-6}\approx 0.00248$

这是一个很小的数，说明明年没有此类文章的可能性是很小，反过说，有此类文章的可能性很大。