# 2.2 离散型随机变量及其分布律

## 2.2.1 离散型随机变量及其分布律

1．离散型随机变量

**定义** 设X为一随机变量，如X的全部可能取到的值是有限个或可列无限多个，则称随机变量X为离散型随机变量。

2． 离散型随机变量的分布律

**定义** 设X为离散型随机变量，X的所有可能取的值为 $x_k(k=1, 2\ldots)$，记 X 取 $x_k$ 的概率为

$$P\{X=x_k\}=p_k \quad (k=1, 2, \ldots),$$

则称下面一组等式

$P\{X=x_k\}=p_k \quad (k=1, 2, \ldots)$为X的分布律，简写为*d. l.*。

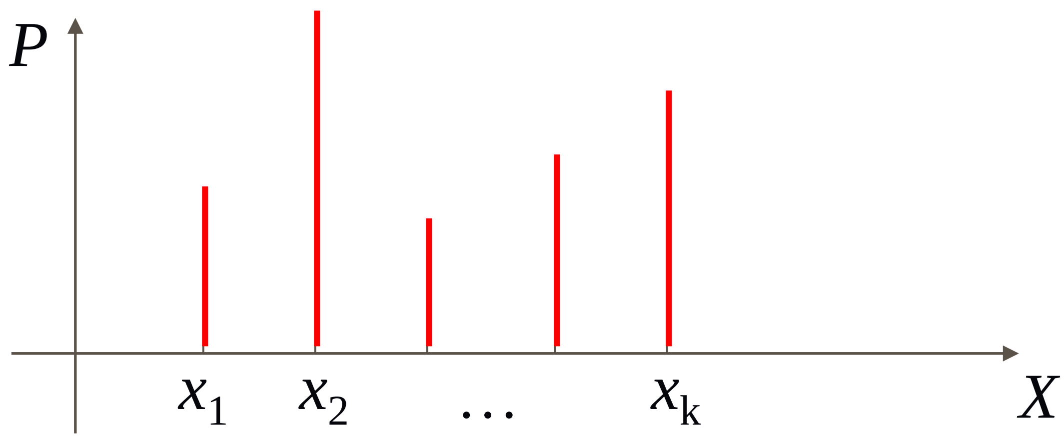
Back

分布律的表示方法：

（1）分布律可以用表格的形式表示：$x_n$一般**从小到大**排列。

| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ | $\cdots$ |
|---|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\cdots$ | $p_n$ | $\cdots$ |

（2）分布律可以用图形表示




Back

例1．袋中5个球编号1—5，从中同时取出3个，以$X$表示取出球的最大编号，求$X$的分布律.

解： $P\{X=3\}=1/C^3_5=1/10$,

$P\{X=4\}= C^2_3 /C^3_5=3/10$,

$P\{X=5\}= C^2_4 /C^3_5=6/10$

$X$的分布律为

| $X$ | 3 | 4 | 5 |
|---|---|---|---|
| P | 1/10 | 3/10 | 6/10 |

例2 重复独立地进行试验，直到事件A出现 $r\ (r\geq 1)$次为止，求试验次数 $X$ 的分布律.

解：设每次试验事件 A 出现的概率为 $p$, 若当第 $k$ 次试验时，事件A出现$r$次，则前$k-1$次试验事件A恰出现$r-1$次，于是

$$P\{X=k\}=C_{k-1}^{r-1}p^{r-1}q^{k-r}\cdot p=C_{k-1}^{r-1}p^{r}q^{k-r}$$

$$k=r,r+1,\ldots$$

称 $X$ 服从Pascal分布。

当$r$=1时，

$$P\{X=k\}=pq^{k-1}\quad k=1,2,\ldots.$$ 称$X$服从几何分布。

◄Back

## 分布律的性质：

由概率的性质可知分布律具有下述性质

（1）非负性： $p_k \geq 0$; $k=1$，2，…

（2）规范性： $$\sum_{i=1}^{\infty} p_i = 1$$

证明:设离散型 *r. v. X*的取值为 $x_1,\ldots, x_n,\ldots$

则事件组$\{X=x_1\}$，…，$\{X=x_n\}$，…构成了Ω的一个划分。

$$\therefore \sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} P(X = x_k) = P\left(\bigcup_{k=1}^{\infty}\{X = x_k\}\right) = 1.$$

## 分布律与分布函数的关系

（1）已知随机变量$X$的分布律，可求出$X$的分布函数：

①设一离散型随机变量X的分布律为

$$P\{X=x_k\}=p_k\quad(k=1，2，\ldots)$$

由概率的可列可加性可得$X$的分布函数为

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

即$F(x)=\sum_{x_k\le x}p_k$

这里的和式是所有满足$x_k\le x$的$k$求和的。分布函数$F(x)$在$x=x_k(k=1,2,\ldots)$处有跳跃，其跃跳值为$p_k=P\{X=x_k\}$。

◄Back

②已知随机变量X的分布律，亦可求任意随机事件的概率。

例如，求事件{X∈B}（B为实轴上的一个区间）的概率P{ X∈B}时，只需将属于B的X的可能取值找出来，把X取这些值的概率相加，即可得概率P{ X∈B}，即 $$P\{X \in B\}=\sum_{x_k \in B} p_k$$

因此，离散型随机变量的分布律完整地描述它的概率分布情况。

（2）已知随机变量X的分布函数，可求出X的分布律：

◀Back

设一离散型随机变量$X$的分布函数为$F(x)$，并设$F(x)$的所有间断为$x_1,x_2,\ldots$，那么，$X$的分布律为

$$P\{X=x_k\}=F(x_k)-F(x_k-0) \qquad k=1,2,3,\cdots.$$

例3：设随机变量X的分布律为

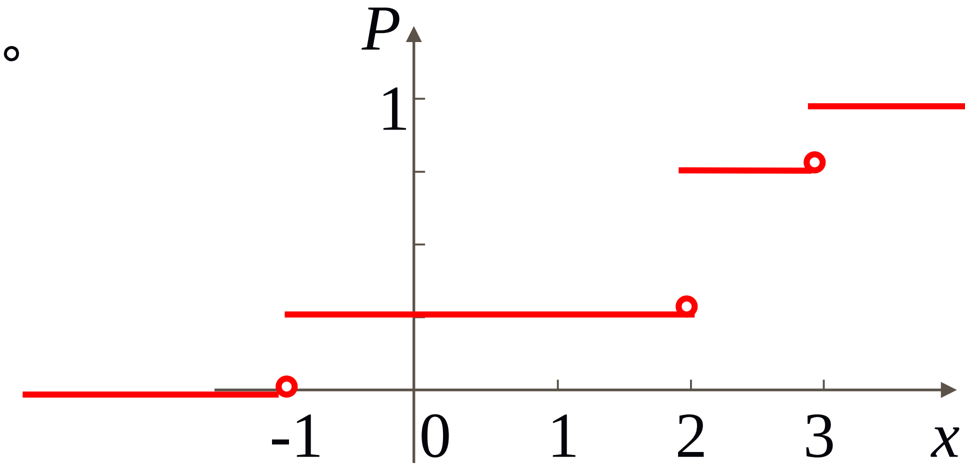
| $X$ | $-1$ | $2$ | $3$ |
|---|---|---|---|
| $P$ | $1/4$ | $1/2$ | $1/4$ |

求X的分布函数，并求 $P\{2\le X\le 3\}$, $P\left\{\frac{3}{2}\le X\le\frac{5}{2}\right\}$, $P\left\{X\le\frac{1}{2}\right\}$

解：由概率的有限可加性，得所求分布函数为

$$F(x)=\begin{cases}0 & x<-1\\ \frac{1}{4} & -1\le x<2\\ \frac{1}{4}+\frac{1}{2} & 2\le x<3\\ \frac{1}{4}+\frac{1}{2}+\frac{1}{4} & x\ge 3\end{cases}\quad \text{即}\quad F(x)=\begin{cases}0 & x<-1\\ \frac{1}{4} & -1\le x<2\\ \frac{3}{4} & 2\le x<3\\ 1 & x\ge 3\end{cases}$$

$F$（$x$）的图形如下图所示，它是一条阶梯形的曲线，在x＝－1，2，3处有**跳跃点**，**跳跃值**分别为1/4，1/2，1/4。




◄Back

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

$$P\left\{\frac{3}{2} \le X \le \frac{5}{2}\right\} = P\{X = 2\} = \frac{1}{2}$$

$$P\left\{X \le \frac{1}{2}\right\} = P\{X = -1\} = \frac{1}{4}$$

Back

例4 已知随机变量X的分布律为

| $X$ | $-2$ | $0$ | $3$ | $5$ |
| --- | --- | --- | --- | --- |
| $P$ | $1/4$ | $a$ | $1/2$ | $1/12$ |

试求（1）待定系数$a$，（2）概率$P\{X>-1/2\}$。

解：（1）由分布律的性质可知

$$\frac{1}{4}+a+\frac{1}{2}+\frac{1}{12}=1$$

即可求得$a=1/6$。

（2）
$$P\left\{X>-\frac{1}{2}\right\}=P\{X=0\}+P\{X=3\}+P\{X=5\}$$

$$=\frac{1}{6}+\frac{1}{2}+\frac{1}{12}=\frac{3}{4}$$

Back

## 2.2.2 几种常用的离散型随机变量的分布

### 1. （0—1）分布：

设随机变量$X$只可能取0与1两个值，它的分布律为

$$P\{X=k\}=p^k(1-p)^{1-k},\ k=0,\ 1.\ (0<p<1)$$

则称$X$服从（0—1）分布，记为X~（0—1）分布。

（0—1）分布的分布律用表格表示为：

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $1-p$ | $p$ |

易求得其分布函数为：

$$F(x)=\begin{cases}0 & x<0\\ 1-p & 0\le x<1\\ 1 & x\ge 1\end{cases}$$

⇦ Back ⇨

## 2. 伯努利试验、二项分布

考虑一个简单的试验，它只出现（或只考虑）两种结果，如。

一般地，试验E只有两种结果A和 $\bar{A}$，而 $P(A)=p$（$0<p<1$），**则称E为伯努利试验或伯努利概型。**

Back

设E为贝努利试验，将E独立地重复进行n次，而且每次试验中结果A出现的概率保持不变。我们把这n次独立重复伯努利试验总起来看成一个试验，称这种试验叫**n重伯努利试验**。总之，n重伯努利试验有下面四个约定：

（1）每次试验的结果只能是两个可能的结果A和 $\bar{A}$之一，

（2）A在每次试验中出现的概率p保持不变，

（3）各次试验相互独立，

（4）共进行了n次.

请思考： 古典概型与伯努里概型不同，有何区别？

伯努里概型对试验结果没有等可能的要求.

**定理**　对于n重伯努利试验，事件A在n次试验中出现k次的概率为

$$P_n(k)=C_n^k p^k q^{n-k} \qquad k=0,1,\ldots,n, \quad q=1-p$$

**证明：** 由n重伯努利试验，事件A在某指定的k次试验中出现，而在其余n-k次试中不出现的概率为

$$p^k(1-p)^{n-k} = p^k q^{n-k}$$

而在n次试验中事件A发生k次共有$C_n^k$种不同情况，对应的事件为互不相容的，由概率的可加性

$$P_n(k)=C_n^k p^k q^{n-k} \qquad k=0,1,\ldots,n, \quad q=1-p$$

**例5**：某射手的一次命中率为0.9，向目标射击5次，求恰好命中3次的概率。

解　X表示命中次数，

则$X \sim b(5, 0.9)$

$$P\{X = 3\} = C_5^3 0.9^3 \times 0.1^2$$

例6：对某种药物的疗效进行研究，假定这药对某种疾病的治愈率0.8，现有10个人患此病的病人同时服用此药，求其中至少有6个病人治愈的概率。

解：假定“病人服用此药后治愈”为事件A，按题意

P(A)=0.8，  $P(\overline{A})=0.2$

10人同时服用此药可视为10重贝努利试验，因而由公式所求事件至少有6个病人治愈的概率为

$$P=\sum_{k=6}^{10}P_{10}(k)=\sum_{k=6}^{10}C_n^k 0.8^k 0.2^{10-k}\approx 0.97$$

Back

**定义**：若离散型随机变量X的分布律为

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

其中**0<*p*<1,*q*=1-*p***, 则称X服从参数为*n,p*的二项分布，

记为**X~*b*(*n*,*p*)**.

(1) 试验模型：在**n**重伯努利试验中，若以**X**表示事件**A**出现的次数，则**X**是一随机变量，**X**可能取的值为**0**，**1**，**2**，**…**，**n**，由二项概率公式可得**X**的分布律为

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

即X服从二项分布。

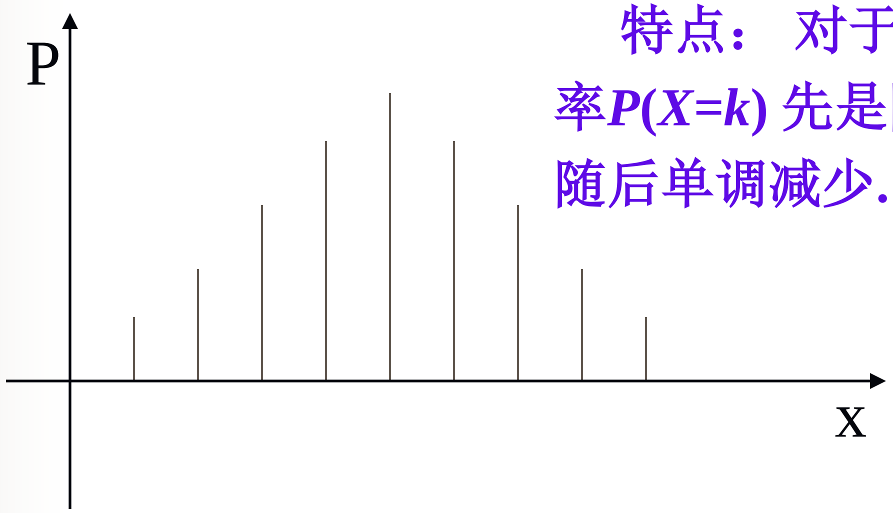
二项分布描述的是*n*重伯努里试验中**A**出现次数*X*的概率分布.

Back

（2）因为 $\sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1$ ，其中 $C_n^k p^k q^{1-k}$ 恰为二项式 $(p+q)^n$ 的一般项，故称为二项分布。

（3）当n=1时，二项分布为（0—1）分布，即 X~b( 1 ,p)。

（4）二项分布的分布律为：




特点：对于固定$n$及$p$，当$k$增加时 ,概率$P(X=k)$ 先是随之增加直至 达到最大值,随后单调减少.

Back

例7 某公司研发楼有四层，每层有20台电脑，发生故障的概率都是0.01，考虑两种方案配备电脑维护人员：

其一是由四人维护，每人一层（20台），

其二是由三人共同维护80台。

试比较这两种方法在电脑发生故障时不能及时维修的概率的大小？

解:方案一

审E E:某人维护20台电脑

选A A：电脑出故障 $p = P(A) = 0.01$

记X X：20台电脑中同时发生故障的电脑台数

得分布 $X \sim B(20, 0.01)$

欲求 $P\{X \geq 2\}$

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$
$$= 1 - C_{20}^{0} p^{0} q^{20} - C_{20}^{1} p^{1} q^{19} = 0.0169$$

Back

记B：电脑不及时维修

$A_i$ ：第i层电脑不能及时维修，i=1，2，3，4

$$\mathrm{P}(B) = P(A_1 \cup A_2 \cup A_3 \cup A_4)$$
$$= 1 - P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})P(\overline{A_4})$$
$$= 1 - (1 - 0.0169)^4 = 0.0659$$

方案二：Y：80台电脑中同时出现故障的台数

$$Y \sim B(80，0.01)$$

$$P\{Y \geq 4\} = 1 - P\{Y \leq 3\}$$

$$= 1 - \sum_{k=0}^{3} C_{80}^{k}(0.01)^{k}(0.99)^{80-k}$$

$$= 0.0087$$

说明管理模式的重要性。

Back

例8　某人射击，每次命中率为0.02，求在独立进行400次射击中，至少击中2次的概率？

解：设X表示射击400次击中的次数，由题意 X~b(400，0.02)。

$$即: P\{X=k\}=C_{400}^{k}\cdot(0.02)^{k}\cdot(0.98)^{400-k},\ k=0,1,2,\cdots,400$$

$$求: P\{X\geq 2\}=?\ 所以$$

$$P\{X\geq 2\}=1-P\{X<2\}=1-P\{X=0\}-P\{X=1\}$$

$$=1-(0.98)^{400}-400\times(0.02)\times(0.98)^{399}$$

直接计算很繁，下面介绍Possion定理。

**泊松定理**： 设λ>0是一常数，n是任意正整数，设 $np_n=\lambda$，则对于任一固定的非负整数k，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$p_n=\lambda/n$有

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

$$=\frac{\lambda^k}{k!}\left[1\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}.$$

对于任意固定的k，当n→∞时

$$\left[1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\to 1$$

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

故有

$$C_n^k p_n^k (1-p_n)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}, \qquad \text{其中 } \lambda = np_n$$

意义：定理的条件$np_n=\lambda$（常数）意味着当n很大时，$p_n$必定很小。因此，上述定理表明当n很大、p很小时有以下近似式

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}$$

由泊松近似公式计算上题：

$$P\{X \geq 2\} \approx 1-\frac{8^0}{0!}e^{-8}-\frac{8^1}{1!}e^{-8}=1-9e^{-8}\approx 0.9973$$

$$P\{X \geq 2\} \approx 1-\sum_{k=0}^{1}\frac{\lambda^k e^{-\lambda}}{k!}=1-0.003=0.997$$

分析结果：不能忽视小概率事件，迟早会发生。罕见事件原理！

实际计算时，当$n \geq 20, p \leq 0.05$时这种近似颇佳！

◀Back

## 3. 泊松分布：

(1) 设离散型随机变量X的分布律为

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

其中λ>0是常数。则称X服从参数为λ的泊松分布，记为X~π(λ)。

$$P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!} > 0 \quad \sum_{k=0}^{\infty} P\{X = k\} = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Back

历史上，*泊松分布是作为二项分布的近似*，于1837年由法国数学家泊松引入的．

泊松，S.-D.

近数十年来，泊松分布日益显示其重要性,成为概率论中最重要的几个分布之一.

在实际中，许多随机现象服从或近似服从泊松分布.

我们把在每次试验中出现概率很小的事件称作稀有事件.
如地震、火山爆发、特大洪水、意外事故等等

由泊松定理，$n$重贝努里试验中稀有事件出现的次数近似地服从泊松分布.

Back

(2) 泊松分布背景：

例如，

在一个时间间隔内某路口通过的汽车数、
一本书一页中的印刷错误数、
某地区在一天内邮递遗失的信件数、
某一医院在一天内的急诊病人数、
某一地区一个时间间隔内发生交通事故的次数、
在一个时间间隔内某种放射性物质发出的、经过计数器的$\alpha$粒子数等

都服从泊松分布，

泊松分布也是概率论中的一种重要分布。

## (3) 泊松分布产生的一般条件

在自然界和人们的现实生活中, 经常要遇到在随机时刻出现的某种事件. 我们把在随机时刻相继出现的事件所形成的序列, 叫做随机事件流.

若事件流具有平稳性、无后效性、普通性，则称该事件流为泊松事件流（泊松流）.

下面简要解释平稳性、无后效性、普通性.

平稳性: 在任意时间区间内，事件发生$k$次($k \geq 0$)的概率只依赖于区间长度而与区间端点位置无关.

无后效性: 在不相重叠的时间段内，事件的发生是相互独立的.

普通性： 如果时间区间充分小，事件出现两次或两次以上的概率可忽略不计.

因此：泊松分布在管理科学、运筹学以及自然科学的某些问题中都占有重要的地位 .

$$P\{X(\mathrm{t})=k\}=\frac{(\lambda t)^k e^{-(\lambda t)}}{k!} \qquad k=0,1,\cdots .$$

例9 有300台机器，工作相互独立。发生故障概率为0.01，通常，一台机器的故障可由一人来修理（一人修一台），问至少需要多少工人，才能保证当设备发生故障但不能及时修理的概率不大于0.01。

解：设需要配备修理工人数为N个，设备同时发生故障的台数为X台，由题知求最小的N为多少，即使 $P\{X>N\}\leq 0.01$.

因为 $X\sim b(300, 0.01)$，由于 $n$ 很大，$p$ 很小，故用泊松分布近似

Back

$$\therefore P\{X > N\} = P\{X \ge N+1\} = \sum_{k=N+1}^{300} P\{X = k\}$$

$$= \sum_{k=N+1}^{300} C_{300}^{k} (0.01)^{k} (0.99)^{300-k}$$

$$\approx \sum_{N+1}^{\infty} \frac{3^{k} e^{-3}}{k!} \le 0.01$$

查表可得:

N+1=k≥9 => N=8（最小的）。

◀Back

## 三、其它常见的分布

（ⅰ）超几何分布

设一堆同类产品共N件，其中有M个次品，现从中任取n个(为方便计算。假定n≤N-M)，则这n个中所含的次品数X是个离散型随机变量，X的分布律为

$$P\{X=m\}=\frac{C_M^m\cdot C_{N-M}^{n-m}}{C_N^n}\qquad m=0,1,\cdots,l.$$

其中l=min(M，n)，这个概率分布称为超几何分布。

(ⅱ)几何分布

在独立重复试验中，设A在每次试验中发生的概率均为p，记X为A首次发生时的试验次数。 则不难验证，X具有如下分布律

$$P\{X=k\}=q^{k-1}p \qquad k=1,2,\cdots \qquad q=1-p.$$

这个概率分布称为几何分布。

(ⅲ) 巴斯卡分布

在独立重复试验中，若记X为A在第r次发生时的试验次数, 则X的分布律为

$$P\{X=k\}=C_{k-1}^{r-1}q^{k-r}p^{r} \qquad k=r,r+1,\cdots \qquad q=1-p.$$

这个分布称为巴斯卡分布。

Back

分布函数

离散型r.v.的分布律

分布函数的性质

连续型r.v.的分布密度

分布律与分布函数的关系

概率密度与分布函数的关系

Back