第二章 随机变量

在古典概型中的几个问题

设$\{\Omega, \mathscr{A}, P\}$为随机试验$E$的概率空间

问题一　样本空间Ω中的元素与试验有关，从数学角度看，希望Ω是抽象的集合

问题二　非等可能事件的概率怎么计算？

问题三　在概率论中怎么应用微积分理论？

问题四　…………

解决问题的第一步：

将样本空间 Ω 抽象化

例 抛一枚硬币,考察正、反面出现的情况,则

$$\Omega = \{H, T\}$$

有具体含意的样本空间

令

$$X(\omega) = \begin{cases} 0, & \omega = H \\ 1, & \omega = T \end{cases}$$

称为随机变量

则在上述映射下，新的“样本空间”为

$$\tilde{\Omega} = \{\, 0,\ 1\, \}$$

直线上的抽象点集

样本点对应关系为

$$\{X = 0\} \Leftrightarrow \{H\}$$

$$\{X = 1\} \Leftrightarrow \{T\}$$

例 将一枚硬币连抛三次，观察正、反面出现的情况，则样本空间为

$$\Omega = \{TTT, TTH, THT, HTT, THH, HTH, HHT, HHH\}$$

定义随机变量

$$X = \text{正面出现的次数}$$

则

$$\Omega \xrightarrow{X} \Omega' = \{0, 1, 2, 3\}$$

考虑事件

$$\{X = 0\} = \{TTT\}$$

$$\{X = 2\} = \{THH, HTH, HHT\}$$

$$\{X \le 2\} = \{TTT, TTH, THT, HTT, THH, HTH, HHT\}$$

$$= \{X \le 3\} - \{X = 3\}$$

$$= \Omega - \{HHH\}$$

很多试验产生的结果本身就是随机变量

例 某地区的日平均气温 X，日平均降水量 Y.

例 电子产品的寿命 Y.

例 某城市的日耗电量 W是一随机变量.

连续取值

例 一人连续对目标射击 n 次，击中目标次数 X.

例 从一大批产品中随机抽取 n件进行测试，其测得的次品数N.

离散取值

常见两类随机变量 $\begin{cases}\text{离散型随机变量}\\ \text{连续型随机变量}\end{cases}$

注一 通常用大写字母 X,Y,Z,W 等表示随机变量，用小写字母 x,y,z,w 等表示实数(非随机)

注二 随机变量 random variable，简记为 r.v.

随机变量的分类 $\begin{cases} \text{离散型 r.v} \\ \text{非离散型 r.v} \end{cases}$

定义 若 r.v X 仅取有限或可列个值，则称 X 为**离散型随机变量.**

例 将一枚硬币连抛三次，观察正、反面出现的情况，定义

$$X = \text{正面出现的次数}$$

X 的取值为 $0, 1, 2, 3,$ 故 X 是离散型 r.v.

例 用同一支枪对目标进行射击，直到击中目标为止，则射击次数 X 是离散型 r.v.

例 114查号台一天接到的呼叫次数 X 是离散型 r.v.

问 电子产品的寿命 X 是否是离散型 r.v ？ ✗

设 X为离散型 r.v,设 X所有可能的取值为

$$x_1, x_2, \cdots, x_k, \cdots$$

且

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

易知 X 的统计规律完全由数列$\{x_k\}$,$\{p_k\}$确定

定义 称

$$P\{X = x_k\} = p(x_k) \quad (k = 1, 2, \cdots)$$

为离散型 r.v X的概率质量函数(PMF)或频率函数.

离散型随机变量的频率函数包括两方面:

① r.v的所有可能的取值

② r.v取各个值的概率

例 将一枚硬币连抛三次，观察正、反面出现的情况，记X为正面出现的次数,求 X 的频率函数.

解 X 的取值为 $0,1,2,3$,其样本空间为

$$\Omega=\{TTT,TTH,THT,HTT,THH,HTH,HHT,HHH\}$$

故 X 的频率函数为

$$P\{X=0\}=\frac{1}{8}$$

$$P\{X=1\}=\frac{3}{8}$$

$$P\{X=2\}=\frac{3}{8}$$

$$P\{X=3\}=\frac{1}{8}$$

所有样本点遍历一次

全部和为1

问 频率函数有什么特点？

频率函数的基本性质：

① $p(x_k) \geq 0,\ k = 1, 2, \cdots$

② $\sum_{k=1}^{\infty} p(x_k) = 1$

频率函数的本质特征

证 ② $\sum_{k=1}^{\infty} p(x_k) = \sum_{k=1}^{\infty} P\{X = x_k\}$

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

$$= P(\Omega) = 1$$

本质特征的含义：

- 离散型r.v的频率函数必满足性质①②
- 满足性质①②的数列$\{p_k\}$必是某离散型r.v的频率函数

频率函数的几种表示方法

1 解析式法

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

2 列表法

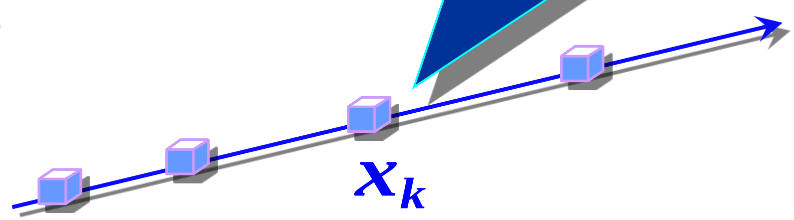
X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

3 矩阵法

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}$$

想象 离散型r.v的概率分布规律相当于向位于 x_k 处的"盒子"中扔球

扔进第 k 个"盒子"的可能性是 p_k



例 一球队要经过四轮比赛才能出线.设球队每轮被淘汰的概率为 $p = 0.5$, 记 X 表示球队结束比赛时的比赛次数, 求 X 的频率函数.

解 r.v X 可能的取值为 $1, 2, 3, 4$.记

$$A_k = \{\text{通过第 } k \text{ 轮比赛}\}, k = 1, 2, 3, 4$$

则

$$P\{X = 1\} = P(\bar{A}_1) = p$$

$$P\{X = 2\} = P(A_1\bar{A}_2) = P(\bar{A}_2 \mid A_1)P(A_1) = p(1 - p)$$

$$P\{X = 3\} = P(A_1A_2\bar{A}_3)$$

$$= P(\bar{A}_3|A_1A_2)P(A_2|A_1)P(A_1) = p(1 - p)^2$$

$$P\{X = 4\} = P(A_1A_2A_3\bar{A}_4) + P(A_1A_2A_3A_4) = (1 - p)^3$$

代入 $p = 0.5$,求得 X 的频率函数为

X	1	2	3	4
p_k	0.5	0.25	0.125	0.125

例（掷两颗骰子）掷两颗均匀的骰子，观测其点数，令X为两骰子点数之和，求

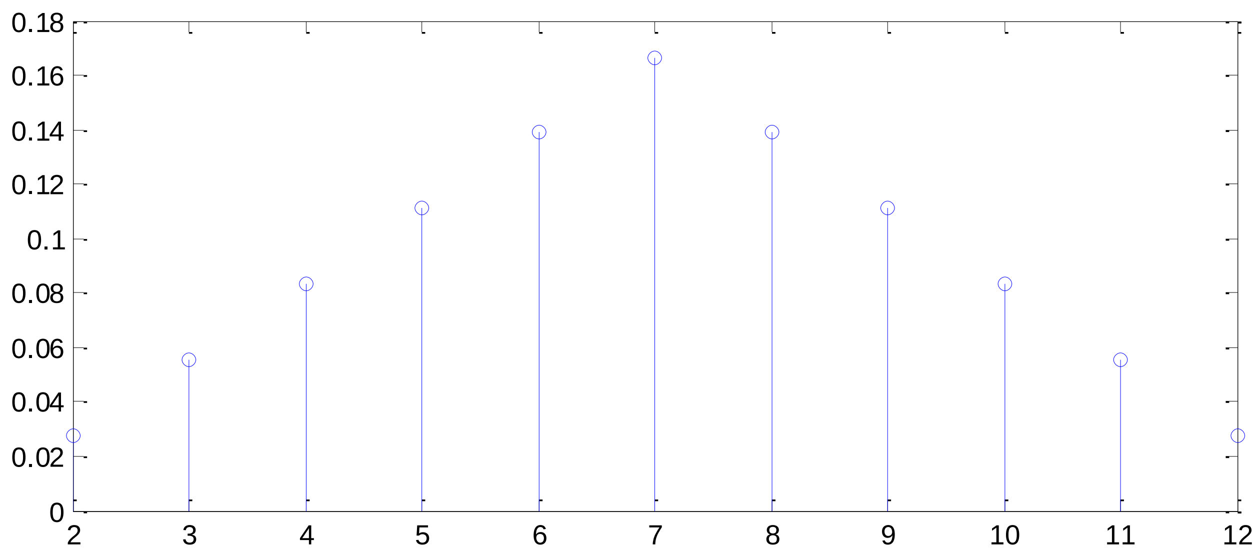
a) X的频率函数，并作图；

b) X为奇数的概率.

解 a)

X	2	3	4	5	6	7	8	9	10	11	12
p_k	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

b) $\dfrac{18}{36}=\dfrac{1}{2}$



问题 question

随机变量 $X=X(\omega)$ 是样本点 ω 的函数，是"随机函数"，不能应用微积分工具.

怎样将"随机函数"化为"普通函数"？

对于 r.v $X, \forall\ x\in(-\infty,\infty)$

$$\{X\le x\}=\{\omega\mid X(\omega)\le x\}\in\mathcal{A}$$

是事件

定义 **称函数**

$$F(x)=P\{X\le x\}\ ,\ -\infty<x<\infty$$

为 r.v X **的累积分布函数(CDF)，简称分布函数.**

注 r.v的分布函数是关于自变量 x 的普通的函数，它不再是随机的！

例 设 X 的频率函数为

X	1	2	3
p_k	0.3	0.2	0.5

求 r.v X 的分布函数.

解 由分布函数的定义有

$$\forall\ x<1, F(x)=P\{X\le x\}=P(\Phi)=0$$

$$\forall\ 1\le x<2, F(x)=P\{X\le x\}=P\{X=1\}=0.3$$

$$\forall\ 2\le x<3, F(x)=P\{X\le x\}$$

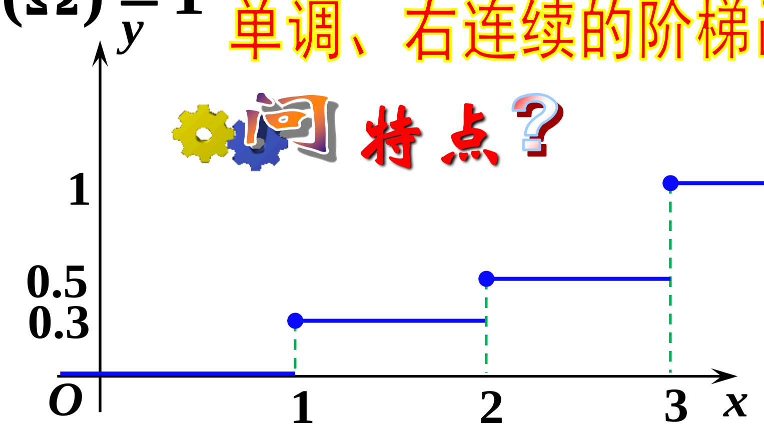
$$=P\{X=1\}+P\{X=2\}=0.5$$

$$\forall\ x\ge 3, F(x)=P\{X\le x\}=P(\Omega)=1$$

单调、右连续的阶梯函数

问特点?

$$\therefore\quad F(x)=\begin{cases}0, & x<1,\\ 0.3, & 1\le x<2,\\ 0.5, & 2\le x<3,\\ 1, & x\ge 3\end{cases}$$



r.v X 的分布函数

$$F(x)=P\{X\le x\}, \quad -\infty<x<\infty$$

分布函数的基本性质：

① $F(x)$是单调不减函数

② $0\le F(x)\le 1$且

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,\ F(+\infty)=\lim_{x\to+\infty}F(x)=1$$

③ $F(x)$右连续函数即

$$F(x+0)=\lim_{t\to x^+}F(t)=F(x)$$

注 性质①②③是分布函数的本质特征

- r.v的分布函数必满足性质①②③
- 满足性质①②③的$F(x)$必是某r.v的分布函数

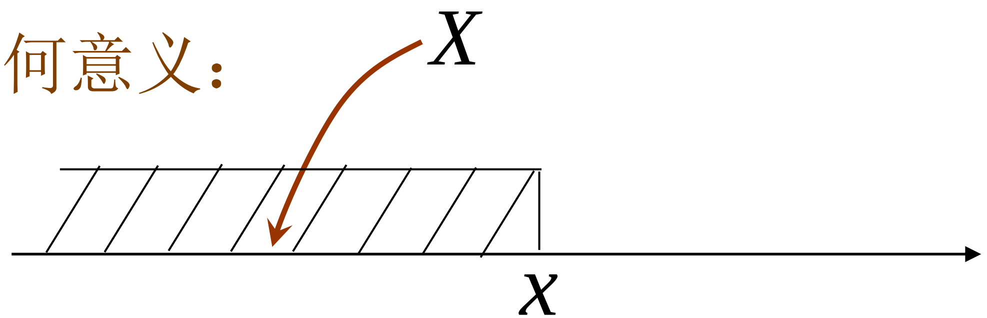
分布函数

随机变量X,对实变量x, $P(X \le x)$ 应为x的函数

定义：随机变量X,对任意实数x,称函数

$F(x) = P(X \le x)$为X的概率分布函数，简称分布函数。

$F(x)$的几何意义：



$F(x)$的性质：

1) $0 \le F(x) \le 1$

2) $F(x)$单调不减，且$F(-\infty) = 0$，$F(+\infty) = 1$

$\because \quad 0 \le P(x_1 < X \le x_2) = F(x_2) - F(x_1)$

问 怎样利用分布函数计算概率 $P\{a < X \le b\}\ (a < b)$?

分析：$\because\ \{a < X \le b\} = \{X \le b\} - \{X \le a\}$

$\{X \le a\} \subset \{X \le b\}$

$\therefore\ P\{a < X \le b\} = P\{X \le b\} - P\{X \le a\}$

$= F(b) - F(a)$

问 怎样计算概率 $P\{X = c\}$ (c为常数)?

分析：$\forall\ \Delta t > 0,\ \{X = c\} \subset \{c - \Delta t < X \le c\}$

$$\therefore\ P\{X = c\} = \lim_{\Delta t \to 0^+} P\{c - \Delta t < X \le c\}$$

$$= \lim_{\Delta t \to 0^+} [F(c) - F(c - \Delta t)]$$

$$= F(c) - F(c - 0)$$

一个有趣的现象

若 $F(x)$在 $x = c$ 处连续，则 $P\{X = c\} = 0$

若不连续呢？

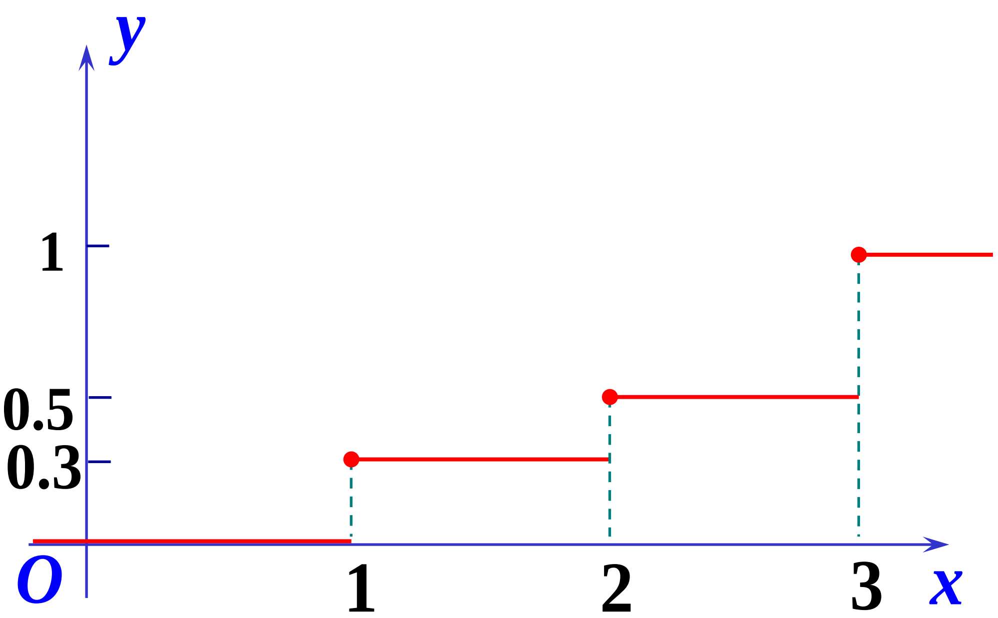
例 设 X 的频率函数为

X	1	2	3
p_k	0.3	0.2	0.5

求 r.v X 的分布函数.

解

$$F(x)=\begin{cases} 0, & x<1, \\ 0.3, & 1\le x<2, \\ 0.5, & 2\le x<3, \\ 1, & x\ge 3 \end{cases}$$



由于分布函数是一个普通的函数，

所以可以应用微积分工具来研究随机现象

几种重要的离散型随机变量

（0）单点分布

如果 r.v X 的频率函数为

$$P\{X=c\}=1$$

则称 r.v X 服从**单点分布**，其中 c 为常数.

注

1. 严格说单点分布并不具有"随机性",视为随机变量完全是理论上的需要
2. 单点分布也称为**退化分布**
3. 某事件发生的概率为1,则称该事件"几乎处处"发生

例如 $P\{X=c\}=1$ 记为 $X\overset{a.e}{=\!=}c$ 或 $X=c\ \ (a.e)$

$P\{X=Y\}=1$ 记为 $X\overset{a.e}{=\!=}Y$ 或 $X=Y\ \ (a.e)$

（一）（0-1）两点分布（伯努利随机变量）

如果 r.v X 的频率函数为

$$P\{X=1\}=p\ ,\ P\{X=0\}=1-p$$

则称 r.v X服从 **(0–1)两点分布**，其中 $0<p<1$为常数.

(0 -1) 分布的实际背景

若一个试验只产生两个结果，则可以用服从(0-1)分布的r.v来描述

例 一门课程的考试是“及格”还是“不及格”

例 刚出生的新生儿是“男”还是“女”

例 产品检验的结果是“合格”还是“不合格”

例 射击结果是“击中目标”还是“没有击中目标”

例 (收入分布) **我国2012年家庭人均收入R(千元)分布如下:**

R (千元)	1	2	4.5	9	15.9	25.8	34.3
收入低于R的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

若规定家庭人均收入2千元为贫困线. 又令

$$X=\begin{cases}1, & R\le 2\\ 0, & R>2\end{cases}$$

则*X*服从参数为0.1的两点分布,

***X*也可以看成是某家庭是否是贫困家庭的示性函数.**

（二）伯努利试验与二项分布

伯努利试验： 只产生两个结果 $A, \bar{A}$ 的试验

问 伯努利试验产生什么样的随机变量？

n重伯努利试验： 将伯努利试验独立重复进行 n次的试验

例 某战士用步枪对目标进行射击，记

$$A=\{\text{击中目标}\}, \bar{A}=\{\text{没击中目标}\}$$

每射击一次就是一个伯努利试验，如果对目标进行 n次射击，则是一个 n重伯努利试验.

例 从一批产品中随机抽取一个产品进行检验，记

$$A=\{\text{合格}\}, \bar{A}=\{\text{不合格}\}$$

每检验一个产品就是一个伯努利试验.

问 独立地抽 n件产品进行检验，是否是 n重伯努利试验？

注 **在伯努利试验中，令**

$$P(A)=p\ ,\ P(\bar{A})=1-p$$

“重 复”是指在每次试验中概率 $P(A)=p$ 保持不变

“独 立”是指各次试验的结果互不影响

记

$$A_i=\{\text{第 } i \text{次试验结果}\},\ i=1,2,\cdots,n$$

$\forall\ \ 1\le i_1<i_2<\cdots<i_k\le n$ **有**

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$

令

$$X=n\text{重伯努利试验中事件 } A \text{发生的次数}$$

则 X 是一个离散型 r.v.

问题 question　X 的频率函数是什么？

$X = n$重伯努利试验中事件 A发生的次数

① X的取值为$0, 1, 2, \cdots, n$

② $\{X = k\} \Leftrightarrow n$次独立试验中 $\begin{cases} A \text{ 发生 } k \text{次} \\ \bar{A} \text{ 发生 } n-k \text{次} \end{cases}$

$$\Leftrightarrow \bigcup_{1 \le i_1 < \cdots < i_k \le n} A_{i_1} \cdots A_{i_k} \bar{A}_{j_1} \cdots \bar{A}_{j_{n-k}}$$

$$\therefore \quad P\{X = k\} = P\{\bigcup_{1 \le i_1 < \cdots < i_k \le n} A_{i_1} \cdots A_{i_k} \bar{A}_{j_1} \cdots \bar{A}_{j_{n-k}}\}$$

$$= \sum_{1 \le i_1 < \cdots < i_k \le n} P\{A_{i_1} \cdots A_{i_k} \bar{A}_{j_1} \cdots \bar{A}_{j_{n-k}}\}$$

$$= \sum_{1 \le i_1 < \cdots < i_k \le n} \overbrace{p \cdots p}^{k} \overbrace{(1-p) \cdots (1-p)}^{n-k}$$

$$= C_n^k p^k (1-p)^{n-k} \qquad (k = 0, 1, 2, \cdots, n)$$

$X=n$重伯努利试验中事件 A发生的次数

记 $q=1-p$，从而 X 的频率函数为

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

易知

① $P\{X=k\}\geq 0 \quad (k=0,1,2,\cdots,n)$

② $\sum_{k=0}^{n} P\{X=k\}=\sum_{k=0}^{n} C_n^k p^k q^{n-k}=(p+q)^n=1$

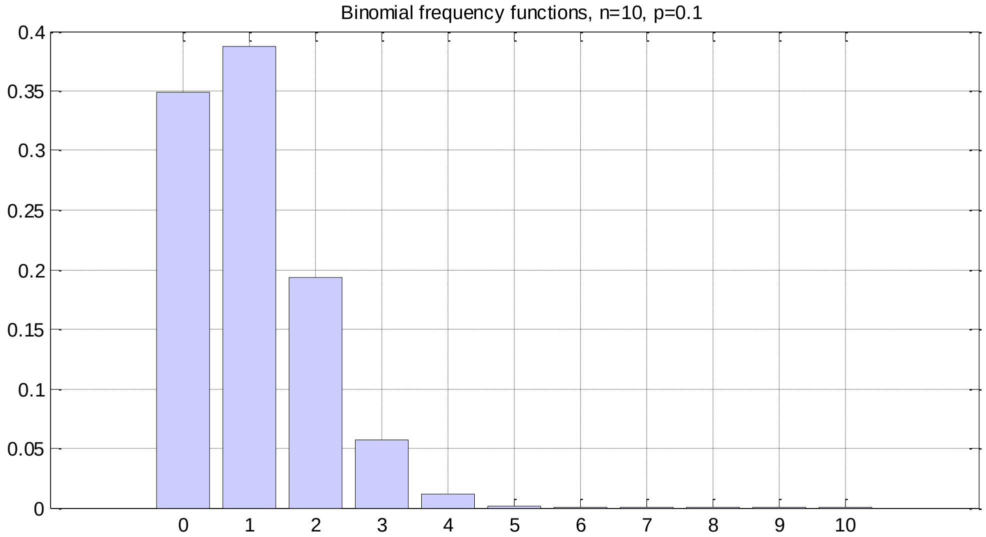
定义 若 r.v X 的频率函数为

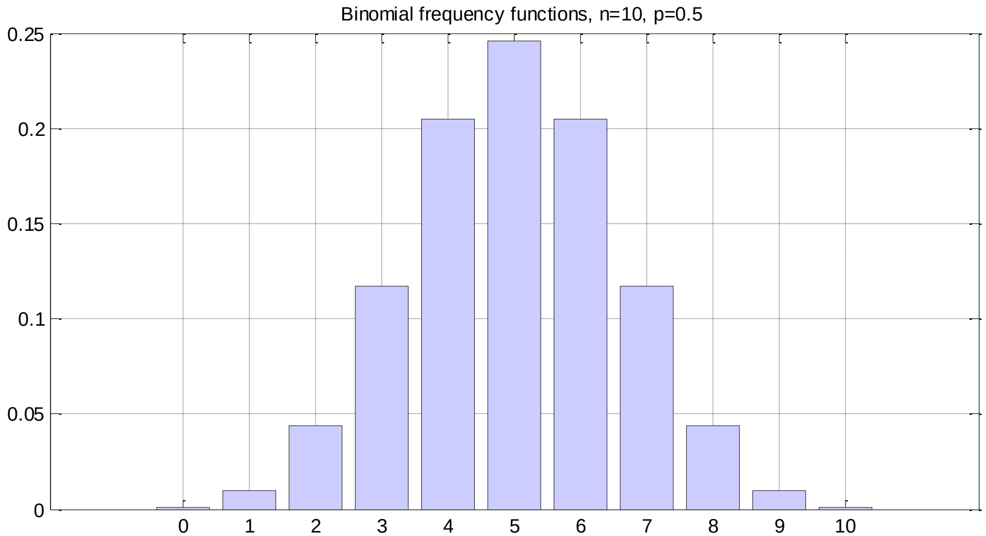
$$P\{X=k\}=C_n^k p^k q^{n-k} \quad (k=0,1,2,\cdots,n)$$

则称 X 服从参数为(n,p)的二项分布，记为 $X\sim b(n,p)$

特别当$n=1$时，$b(1,p)$ 就是(0-1)两点分布，即

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





二项分布的频率函数

n=10，p=0.1

n=10，p=0.5

最可能出现次数

二项分布 $X \sim b(n,p)$

$\{X=k\} \Leftrightarrow n$ 次独立试验中 $\begin{cases} A \text{ 发生 } k \text{ 次} \\ \bar{A} \text{ 发生 } n-k \text{ 次} \end{cases}$

$b(k;n,p) := C_n^k p^k q^{n-k}$ 当 n,p 固定时，随 k 的变化情况

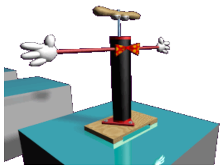
$$\frac{b(k;n,p)}{b(k-1;n,p)} = \frac{(n-k+1)p}{kq} = 1 + \frac{(n+1)p-k}{kq}$$

当 $k<(n+1)p$ 时，$b(k;n,p)$ 随 k 增加而增加

当 $k=(n+1)p$ 时，$b(k;n,p)=b(k-1;n,p)$

当 $m=(n+1)p$ 为正整数时，$b(m;n,p)=b(m-1;n,p)$

最可能出现次数 中心项



实际背景：

二项分布产生于 n 重伯努利试验

例 一大批电子元件有10%已损坏,若从这批元件中随机选取20只来组成一个线路,问这线路能正常工作的概率是多少?

解 因为元件的数量很大，所以取20只元件可看作是有放回抽样 ,记 X表示20只元件中好品的数量，则

$$X \sim b(20, 0.9)$$

$$P\{\text{线路正常}\} = P\{X = 20\}$$

$$= C_{20}^{20} \times 0.9^{20} \times 0.1^{20-20}$$

$$= 0.9^{20}$$

$$\approx 0.1216$$

例 已知100个产品中有5个次品，现从中有放回地取3次，每次任取1个，求在所取的3个中恰有2个次品的概率.

解 因为这是有放回地取3次，因此这3次试验的条件完全相同且独立，它是伯努利试验.

依题意，每次试验取到次品的概率为0.05.

设 X 为所取的3个中的次品数， 则 $X \sim b(3,0.05)$，

于是，所求概率为：

$$P\{X=2\}=C_3^2(0.05)^2(0.95)=0.007125$$

请注意：

若将本例中的“有放回”改为“无放回”，那么各次试验条件就不同了，此试验就不是伯努利试验.

此时，只能用古典概型求解.

$$P\{X=2\}=\frac{C_{95}^{1}C_{5}^{2}}{C_{100}^{3}}\approx 0.00618$$

例（加州枪劫案） 据报道,加州某地发生一起枪劫案,目击嫌疑人有两个：一个男的理平头黑人,一个女的黑发梳马尾型. 不久抓到一对具上述特征的夫妇(情侣)，能否判他们有罪?

数学家通过计算机模拟, 得出一对夫妇具有上述特征的概率为 $p = 8.3\times10^{-8}$，这是一个小概率事件.

陪审团在无其它证据的情况下，裁决他们有罪.

而加州高院推翻了该裁决.

高院认为犯罪的认定应有唯一性. 用下列算式：设X为具上述特征的夫妇数，则 $X \sim b(n, p)$. 要计算

$$P\{X>1 \mid X\ge 1\} = \frac{P\{X>1\}}{P\{X\ge 1\}} = \frac{P\{X\ge 1\}-P\{X=1\}}{P\{X\ge 1\}} = \frac{1-(1-p)^n-np(1-p)^{n-1}}{1-(1-p)^n}$$

n估计为 8.3×10^{6}， 而上式 ≈ 0.2966，不算小.

保险业是最早应用概率论的行业之一.保险公司为了估计企业的利润，需要计算各种各样的概率.

例 若一年中某类保险者里面每个人死亡的概率等于0.005，现有10000个人参加这类人寿保险，试求在未来一年中在这些保险者里面，(1) 有40个人死亡的概率; (2) 死亡人数不超过70个的概率.

解 记 X为未来一年中在这些人中死亡的人数，则

$$X \sim b(10000, 0.005)$$

(1) $$P\{X=40\}=C_{10000}^{40}\times 0.005^{40}\times 0.995^{9960}$$

$$\approx 0.0214$$

(2) $$P\{X\le 70\}=\sum_{k=0}^{70} C_{10000}^{k}\times 0.005^{k}\times 0.995^{10000-k}$$

$$\approx 0.997$$

例 设有80台同类型设备,各台工作是相互独立的,发生故障的概率都是0.01,且一台设备的故障能由一个人处理.考虑两种配备维修工人的方法,(1) 由4人维护,每人负责20台;(2) 由3人共同维护80台.试比较这两种方法在设备发生故障时不能及时维修的概率大小.

解 ① 记 X_i 表示同一时刻第 i 人维护的 20 台设备中同时发生故障的台数，则

$$X_i \sim b(20, 0.01)\,,\quad i = 1, 2, 3, 4$$

则80台设备中发生故障而不能及时维修的概率为

$$P\left(\bigcup_{i=1}^{4}\{X_i \ge 2\}\right) \ge P\{X_1 \ge 2\}$$

$$= 1 - P\{X_1 = 0\} - P\{X_1 = 1\}$$

$$= 1 - C_{20}^{0} \times 0.01^{0} \times 0.99^{20} - C_{20}^{1} \times 0.01^{1} \times 0.99^{19}$$

$$\approx 0.0169$$

例 设有80台同类型设备,各台工作是相互独立的,发生故障的概率都是0.01,且一台设备的故障能由一个人处理.考虑两种配备维修工人的方法,(1) 由4人维护,每人负责20台;(2) 由3人共同维护80台.试比较这两种方法在设备发生故障时不能及时维修的概率大小.

解 ② 记 X表示80台设备中同一时刻发生故障的台数

$$X \sim b(80, 0.01)$$

则80台设备中发生故障而不能及时维修的概率为

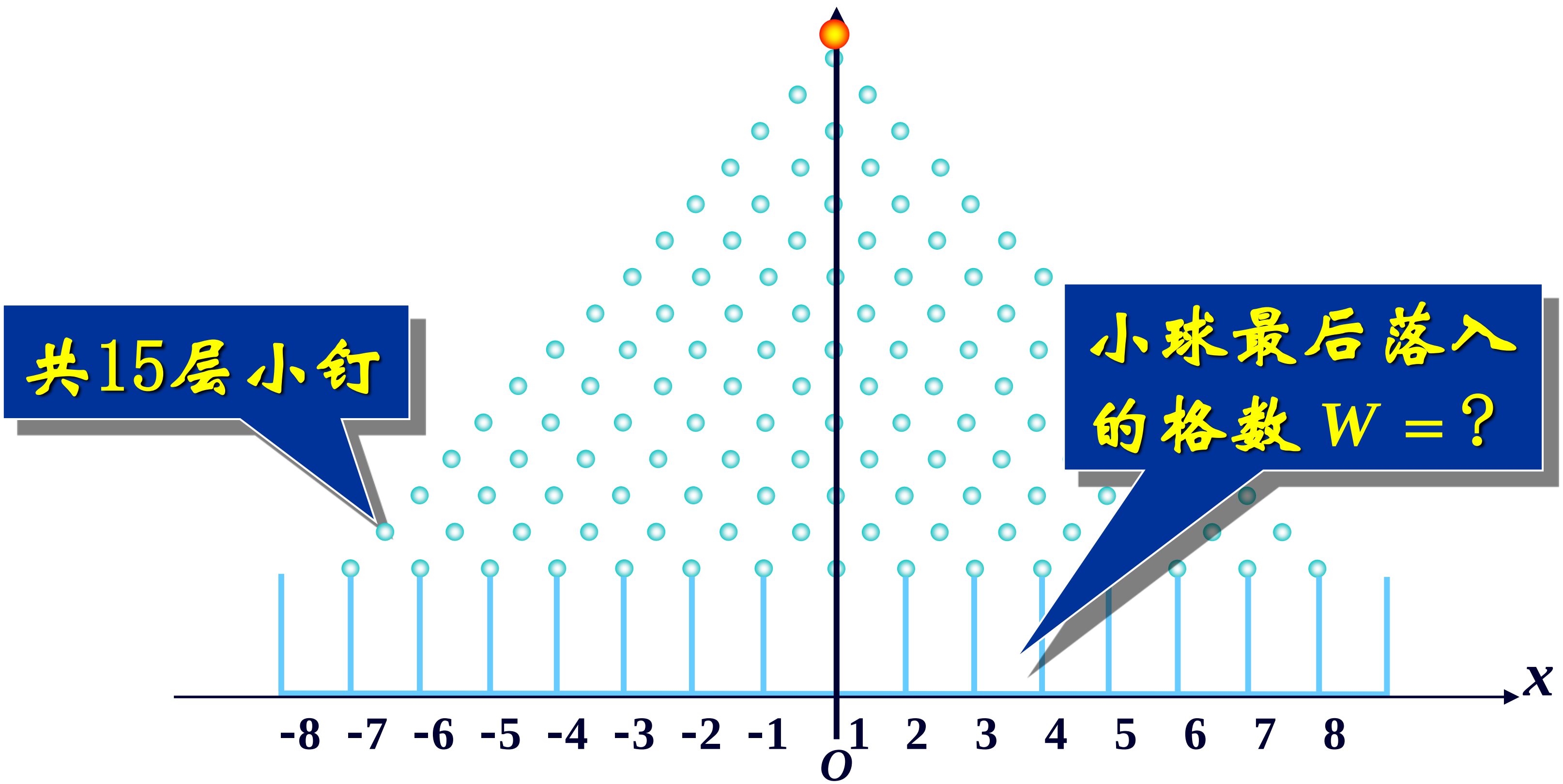
$$P\{X \geq 4\} = 1 - \sum_{i=0}^{3} P\{X = i\}$$

$$= 1 - \sum_{i=0}^{3} C_{80}^{i} \times 0.01^{i} \times 0.99^{80-i}$$

$$\approx 0.0087$$

从两种计算结果可见，方法(2)工人的劳动强度增加了(每人平均维护约27台)，但是工作效率大大提高.

高尔顿钉板试验



记小球向右落下的次数为 X,则 $X \sim b(15, 0.5)$

记小球向左落下的次数为 Y,则 $Y \sim b(15, 0.5)$

$$\therefore \quad W = [X - Y + \mathrm{sign}(X - Y)]/2$$

（三）几何分布和负二项分布

几何分布也是由独立的伯努利试验构造而成.

每次试验成功的概率是 p，X 表示直到第一次成功所做的试验次数.

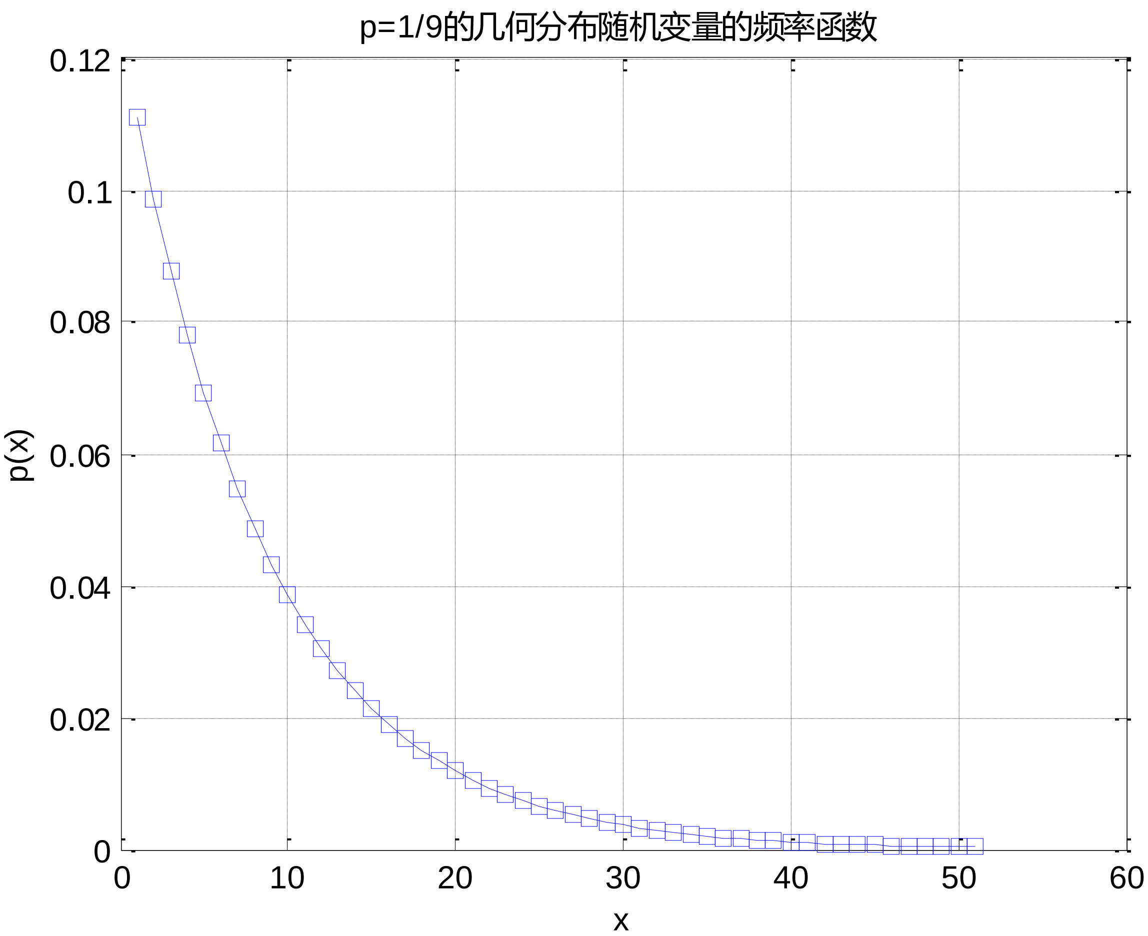
$X=k$： 前面的k-1次试验失败，第k次试验成功，故

$$p(k)=P\{X=k\}=(1-p)^{k-1}p,\quad k=1,2,3,\ldots$$

可以验证$p(k)$满足频率函数的本质特征.

例 **据说某国家彩票命中概率为1/9.**

那么一个人在猜中时所购买的彩票总数就是 $p=1/9$的几何分布随机变量.



（三）几何分布和负二项分布

负二项分布是几何分布的一般化.

每次试验成功的概率是 p，连续独立地试验直到成功 r 次为止. 令 X 表示试验次数.

$X=k$：

由独立性假设, 任一特定试验序列发生的概率是 $p^r(1-p)^{k-r}$.

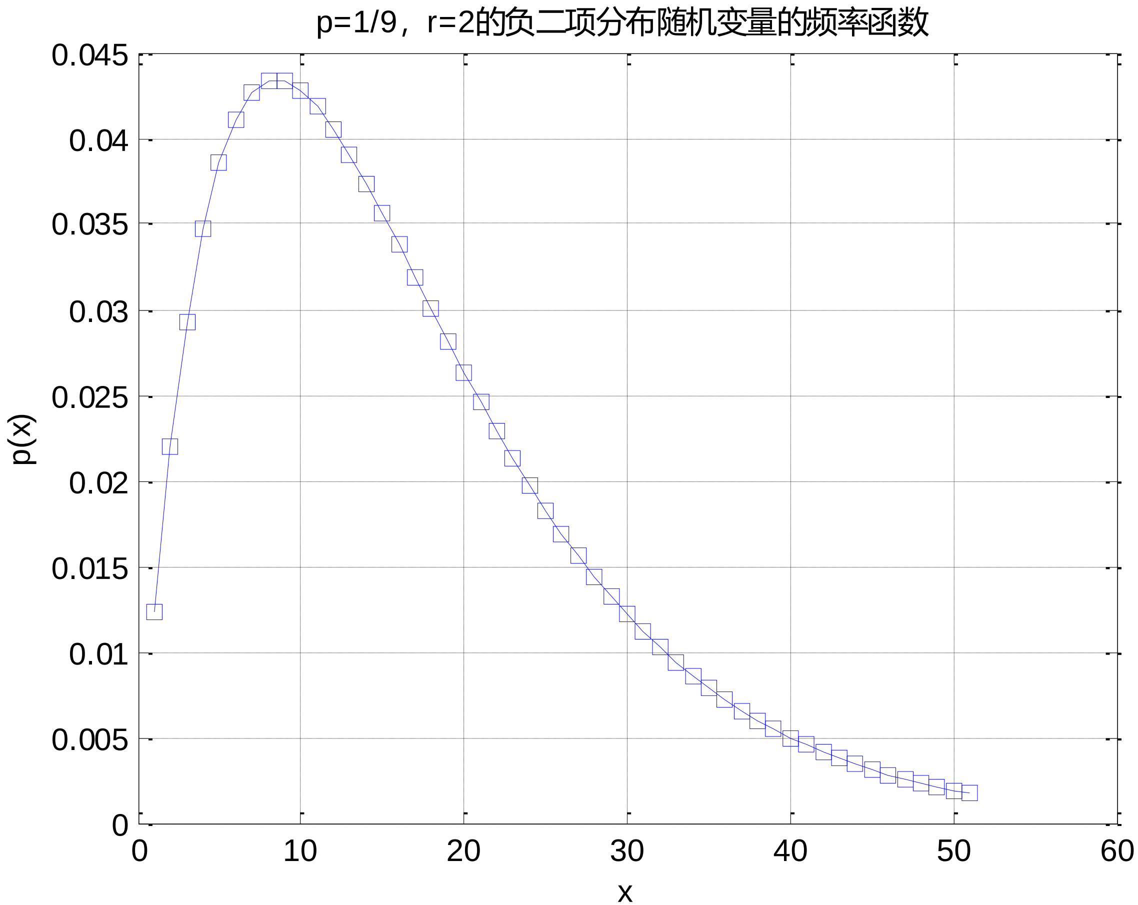
最后一次试验结果是成功的, 剩余的$r-1$次成功出现在剩余的$k-1$次试验中.

$$p(k) = P\{X = k\} = C_{k-1}^{r-1} p^r (1-p)^{k-r}, \quad k = 1, 2, 3, \ldots$$

$r=1$ 的负二项分布就是几何分布.

例 据说某国家彩票命中概率为1/9.

那么一个人直到猜中第二次彩票的试验次数的分布是负二项分布. $p(k)=(k-1)p^2(1-p)^{k-2}, \quad k=1,2,3,\ldots$



（四）超几何分布

假设盒中有n个球，其中r个黑球，n-r个白球.

从盒中无重复地抽取m个球.

令X表示抽到的黑球个数.

$$P\{X=k\}=\frac{C_r^k C_{n-r}^{m-k}}{C_n^m},\quad k=1,\ldots,m$$

X 是参数为r, n, m的超几何分布随机变量.

例 假定在美国18家主要计算机公司中，有12家位于加州的硅谷. 如果从这18家中随机抽取3家，则至少有1家位于硅谷的概率是多少?

$$P\{X\geq 1\}=1-P\{X<1\}=1-P\{X=0\}=1-\frac{C_{12}^0 C_6^3}{C_{18}^3}$$

$$=0.9755$$

（五）泊松流与泊松分布

泊松流： **随着时间的推移,在时间轴上源源不断出现的随机粒子流称为泊松流.**

例 典型的泊松流：随机服务系统

例 典型的泊松流：稀疏现象的发生

例 典型的泊松流：物理学中的现象

问题 在泊松流中，记区间$(0,t]$中出现的质点数为X，问r.v X服从什么分布？

定义 设 r.v X 的取值为$0,1,2,\cdots$，取值概率为

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

其中$\lambda>0$为参数，则称X服从参数为λ的**泊松分布**，记为 $X\sim\pi(\lambda)$ 或 $X\sim P(\lambda)$

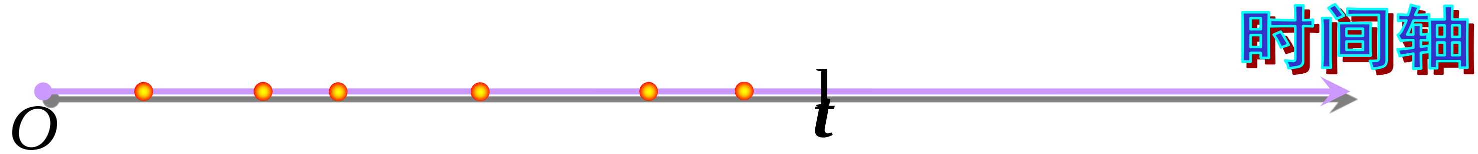
泊松分布的性质：

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

注 泊松分布是构造随机现象的“基本粒子”之一

泊松分布与泊松流的关系

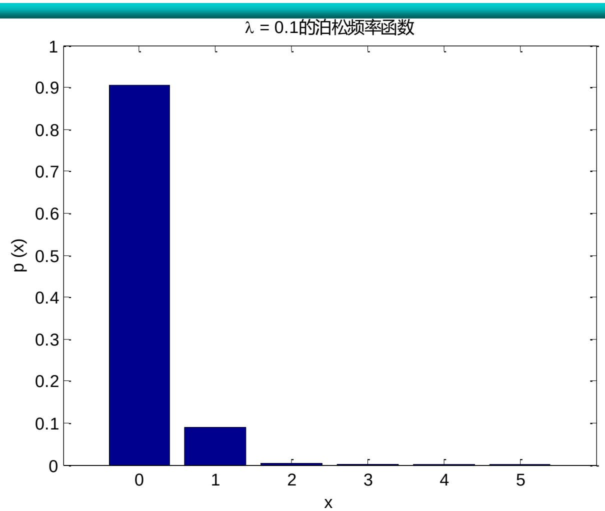
在泊松流中,记时间间隔 $(0,t]$ 中出现的质点数为 X

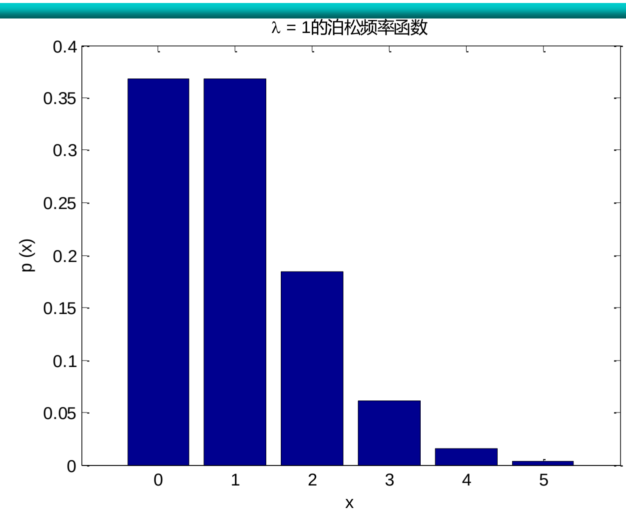


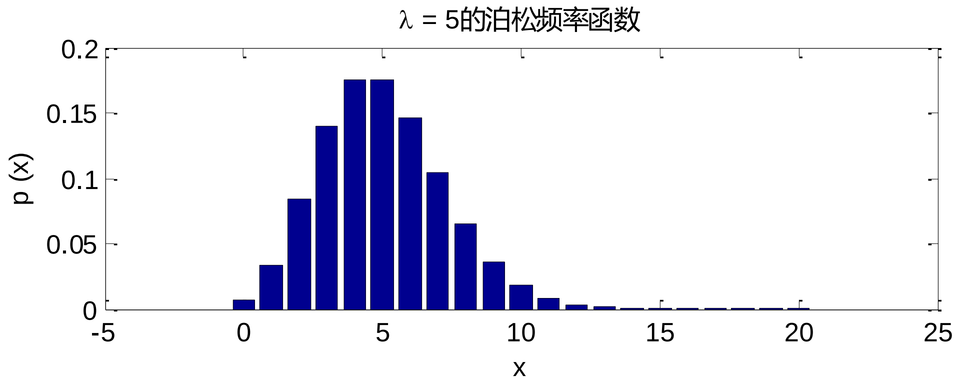
则 $X \sim P(\lambda t)$, 即有

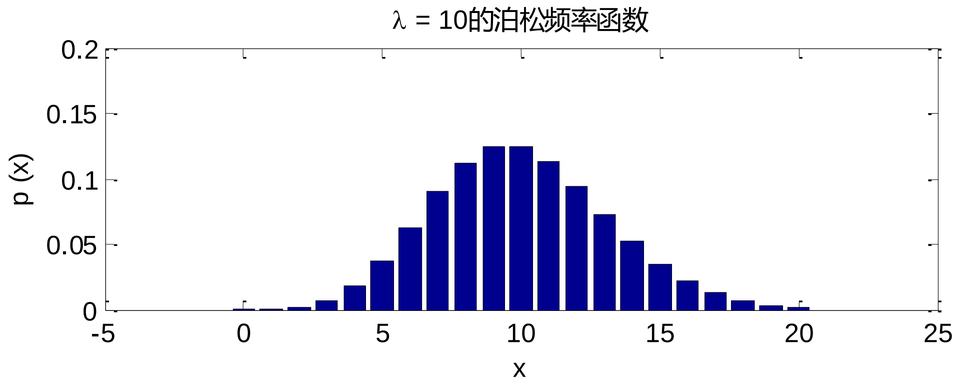
$$P\{X=k\}=\frac{(\lambda t)^k}{k!}e^{-\lambda t}, k=0,1,2,\cdots$$

其中参数 $\lambda>0$,称为 泊松强度.









泊松分布的频率函数

$\lambda=0.1$	$\lambda=1$
$\lambda=5$	
$\lambda=10$	

泊松定理

设 $\lambda > 0$ 是一常数，n 为正整数，$\lim\limits_{n\to\infty} np_n = \lambda$，

则对任一非负整数 k，有

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

说明

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

(当n很大p很小时的计算公式)

例 已知某疾病发病率为0.001，某单位共有5000人，问该单位患有这种疾病的人数不超过5人的概率.

解：设该单位患有这种疾病的人数为X，则有

$X\sim b(5000，0.001)$，则所求概率为

$$P\{X\le 5\}=\sum_{k=0}^{5}C_{5000}^{k}0.001^{k}0.999^{5000-k}$$

取$\lambda=np=5$，用泊松定理近似计算得

$$P\{X\le 5\}\approx\sum_{k=0}^{5}\frac{5^{k}}{k!}e^{-5}=0.616$$

● 泊松分布

【实验】$X \sim b(5000, 0.001)$，求$P\{X \leq 5\}$;

$np=5$, $X \underset{近似}{\sim} P(5)$，求$P\{X \leq 5\}$

➢实验：

	A	B
1	$n=$	5000
2		
3	$P\{X \leq 5\}$	0.6159
4	$P\{X \leq 5\} \approx$	0.6159

=BINOMDIST(5, 5000,0.001,TRUE)

=POISSON(5,5,TRUE)

- 在应用中，诸如服务系统中对服务的呼叫数，产品的缺陷（如布匹上的疵点、玻璃内的气泡等）数，一定时期内出现的稀有事件（如意外事故、自然灾害等）个数，放射性物质发射出的离子数等等，都以泊松分布为其概率模型.
- 这是因为上述例子本来就是n大p小的二项分布．以服务系统中的呼叫数为例，服务设施的用户n很大，每个用户在指定时间内使用这个设施的概率p很小，而且各用户使用情况又独立.
- 因此，服务系统中的呼叫数应是n大p小的二项分布，由泊松定理，可以近似认为服从 $\lambda = np$ 泊松分布．上述应用表明泊松分布广泛用于社会生活的许多方面，它在运筹学、管理科学中占有突出的地位.

练习 如果某房地产公司每天售出1.6套住宅，且住宅的销售量服从泊松分布，则一天内恰好售出4套住宅的概率是多少？一天内没有售出住宅的概率是多少？每天售出5套以上住宅的概率是多少？每天至少售出10套住宅的概率是多少？两天内恰好售出4套住宅的概率是多少？

$$P\{X=4\mid\lambda=1.6\}=0.055131$$

$$P\{X=0\mid\lambda=1.6\}=0.201897$$

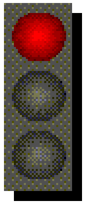
$$P\{X\geq 5\mid\lambda=1.6\}=0.006040$$

$$P\{X\geq 10\mid\lambda=1.6\}=0.000007$$

$$P\{X=4\mid\lambda=3.2\}=0.1781$$

END

- 例：某人骑了自行车从学校到火车站，一路上要经过3个独立的交通灯，设各灯工作独立，且设各灯为红灯的概率为p，0<p<1，以Y表示一路上遇到红灯的次数。



(1) 求Y的概率分布律；

(2) 求恰好遇到2次红灯的概率。

解：这是三重贝努利试验 $Y \sim b(3, p)$

$$(1)\quad P(Y=k)=C_3^k p^k (1-p)^{3-k},\ k=0,1,2,3$$

$$(2)\quad P(Y=2)=C_3^2 p^2 (1-p)$$

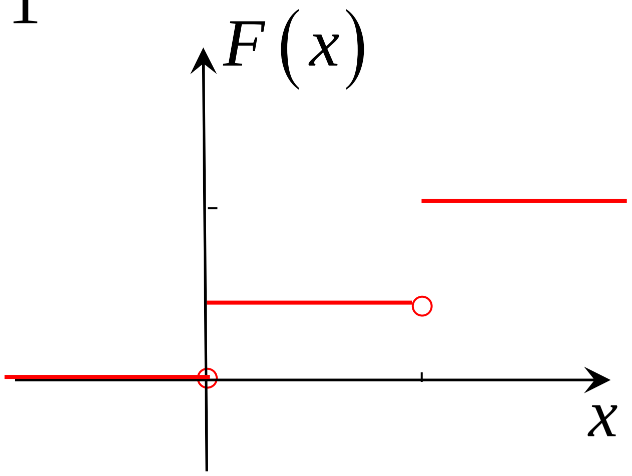
例：

X	0	1
p	q	p

求X的概率分布函数$F(x)$及$P(X \geq 1)$的值。

解：

$$F(x)=P\{X \leq x\}=\begin{cases} 0 & x<0 \\ q & 0 \leq x<1 \\ 1 & x \geq 1 \end{cases}$$



$$P(X \geq 1)=p$$

比较 $P(X \geq 1)=p$ 与 当$x \geq 1$时，$F(x)=1$

课后作业 P46：1、15、31

补充题

1. 设随机变量 X 的频率函数为：

$$P(X=x)=c\left(\frac{2}{3}\right)^x, \quad x=1,2,3$$

求 c 的值。

2. 设随机变量 X 服从泊松分布，求 k 使 $P(X=k)$ 达到最大。

3. 设在 15 只同类型零件中有 2 只为次品，在其中取 3 次，每次任取 1 只，作不放回抽样，以 X 表示取出的次品个数，求：
   (1) X 的分布律；
   (2) X 的分布函数并作图；
   (3)

$$P\{X\le\frac{1}{2}\},P\{1<X\le\frac{3}{2}\},P\{1\le X\le\frac{3}{2}\},P\{1<X<2\}.$$

4. 有 2500 名同一年龄和同社会阶层的人参加了保险公司的人寿保险.在一年中每个人死亡的概率为 0.002，每个参加保险的人在 1 月 1 日须交 12 元保险费，而在死亡时家属可从保险公司领取 2000 元赔偿金.求：
   (1) 保险公司亏本的概率；
   (2) 保险公司获利分别不少于 10000 元、20000 元的概率.