第二章 随机变量及其分布

为了更好的揭示随机现象的规律性并利用数学工具描述其规律，引入随机变量来描述随机试验的不同结果

例 电话总机某段时间内接到的电话次数，可用一个变量 X 来描述

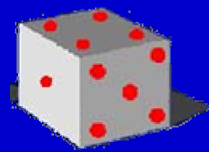
例 抛掷一枚硬币可能出现的两个结果，也可以用一个变量来描述

$$X(\omega)=\begin{cases}1, & \text{正面向上}\\ 0, & \text{反面向上}\end{cases}$$

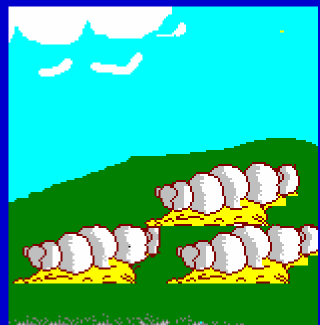
§2.1 随机变量及离散型随机变量的分布

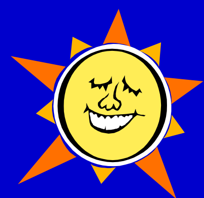
一、随机变量的概念

1、有些试验结果本身与数值有关（本身就是一个数）.

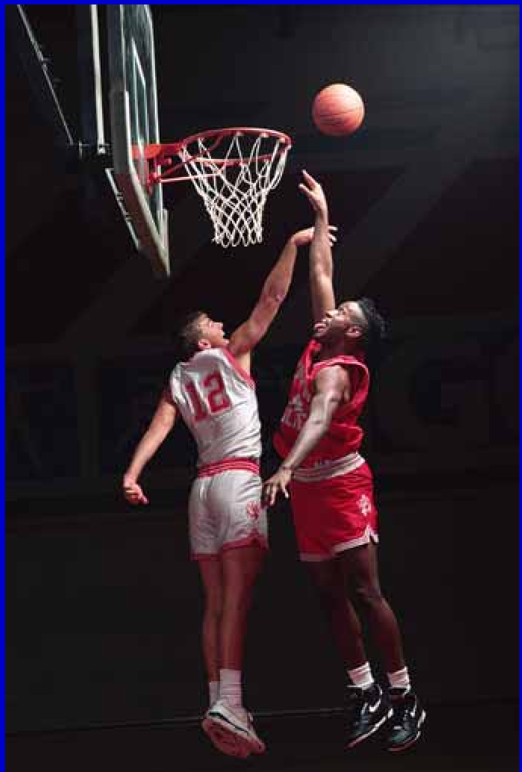
例如，掷一颗骰子面上出现的点数；

每天从赣州下火车的人数；

昆虫的产卵数；

十月份赣州的最高温度；

2、在有些试验中，试验结果看来与数值无关，但我们可以引进一个变量来表示它的各种结果.也就是说，把试验结果数值化.



正如裁判员在运动场上不叫运动员的名字而叫号码一样，二者建立了一种对应关系.

由此引入随机变量的概念

定义 设E是一随机试验，Ω 是它的样本空间， 若

$$\forall \omega \in \Omega \xrightarrow{\text{按一定法则}} \exists \text{ 实数 } X(\omega)$$

则称 Ω 上的单值实值函数 $X(\omega)$为**随机变量**

随机变量一般用 $X, Y, Z, \ldots$或小写希腊字母 ξ, η, ζ 表示

随机变量是 $\Omega \to R$ 上的映射，这个映射具有如下的特点：

- 定义域：Ω
- 随机性：随机变量 X 的可能取值不止一个，试验前只能预知它的可能的取值但不能预知取哪个值
- 概率特性：X 以一定的概率取某个值或某些值

- 引入随机变量后，用随机变量的等式或不等式表达随机事件
- 在同一个样本空间可以同时定义多个随机变量
- 随机变量的函数一般也是随机变量
- 可以根据随机事件定义随机变量
  设 A 为随机事件，则可定义

$$X_A = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \overline{A} \end{cases}$$

称 X_A 为事件A 的示性变量

如，若用X表示电话总机在9:00~10:00接到的电话次数，则

$$\{X > 100\}$$

—— 表示“某天9:00 ~ 10:00 接到的电话次数超过100次”这一事件

再如，用随机变量

$$X(\omega)=\begin{cases}1, & \text{正面向上}\\ 0, & \text{反面向上}\end{cases}$$

描述抛掷一枚硬币可能出现的结果，则

$$\{X(\omega)=1\} \quad \text{——} \quad \text{正面向上}$$

也可以用

$$Y(\omega)=\begin{cases}0, & \text{正面向上}\\ 1, & \text{反面向上}\end{cases}$$

描述这个随机试验的结果

例如，要研究某地区儿童的发育情况，往往需要多个指标，例如，身高、体重、血压等

$\Omega = \{$儿童的发育情况 $\omega\}$

$X(\omega)$ — 身高

$Y(\omega)$ — 体重

$Z(\omega)$ — 血压

各随机变量之间可能有一定的关系，也可能没有关系——即 相互独立

随机变量的分类

离散型随机变量

非离散型随机变量 — 其中一种重要的类型为

连续型随机变量

二、随机变量的分布函数

1、定义 设 X 为随机变量, 对每个实数 x , 随机

事件 $\{X \le x\}$ 的概率

$$P\{X \le x\}$$

定义的一个 x 的实值函数，称为随机变量 X 的分布函数（$d.f.$)，记为 $F(x)$,即

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

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.

2、分布函数的性质

1) $F(x)$ 单调不减，即

$$\forall\ x_1 < x_2,\ F(x_1) \le F(x_2)$$

2) $0 \le F(x) \le 1$ 且

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

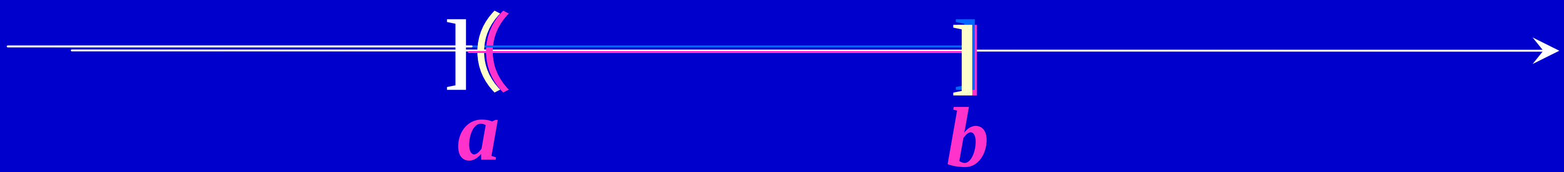
3) $F(x)$ 右连续，即

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

注：凡是满足以上3条性质的函数都是某个随机变量的分布函数

利用分布函数可以计算

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



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

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

请填空

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

三、离散型随机变量及其分布

1、离散型随机变量的概念

定义 若随机变量 X 的可能取值是有限多个或可列无限多个，则称 X 为离散型随机变量

描述离散型随机变量的概率特性常用它的分布律或概率分布，即

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

分布律的性质

1） $p_k \geq 0, \ k = 1,2,\cdots$ —————— 非负性

2） $\sum_{k=1}^{\infty} p_k = 1$ —————— 规范性

2、离散型随机变量的分布函数与分布律互求

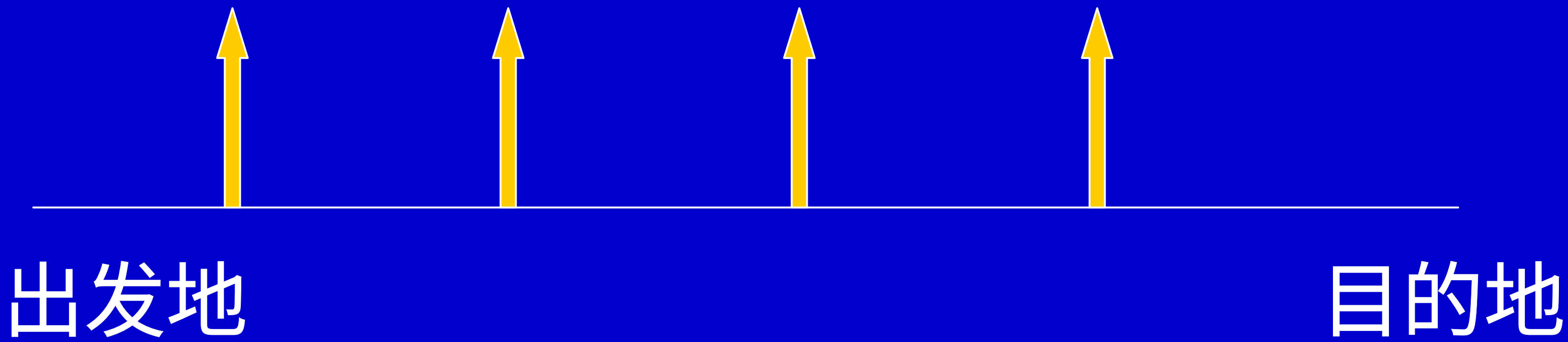
$$F(x) = P\{X \le x\} = P(\bigcup_{x_k \le x} \{X = x_k\})$$

$$= \sum_{x_k \le x} P\{X = x_k\} = \sum_{x_k \le x} p_k$$

$$p_k = P\{X = x_k\} = F(x_k) - F(x_{k-1})$$

$F(x)$ 是分段阶梯函数，在 X 的可能取值 x_k 处发生间断，间断点为第一类跳跃间断点，在间断点处有跃度 p_k

例1 设一汽车在开往目的地的途中需经过 4 盏信号灯，每盏信号灯独立地以概率 p 允许汽车通过。令 X 表示首次停下时已通过的信号灯的盏数，求 X 的分布律与 $p = 0.4$ 时的分布函数。



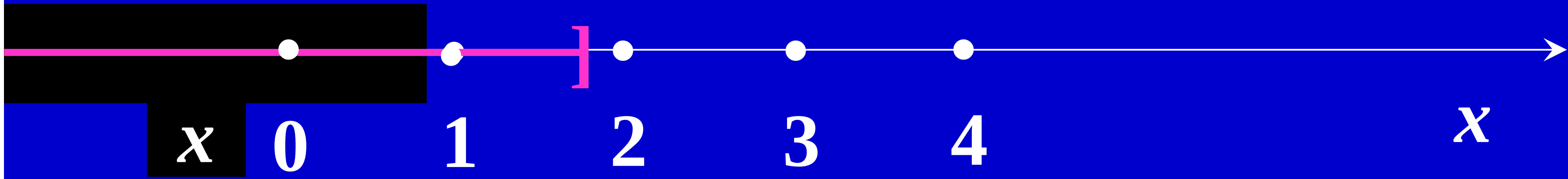
解

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

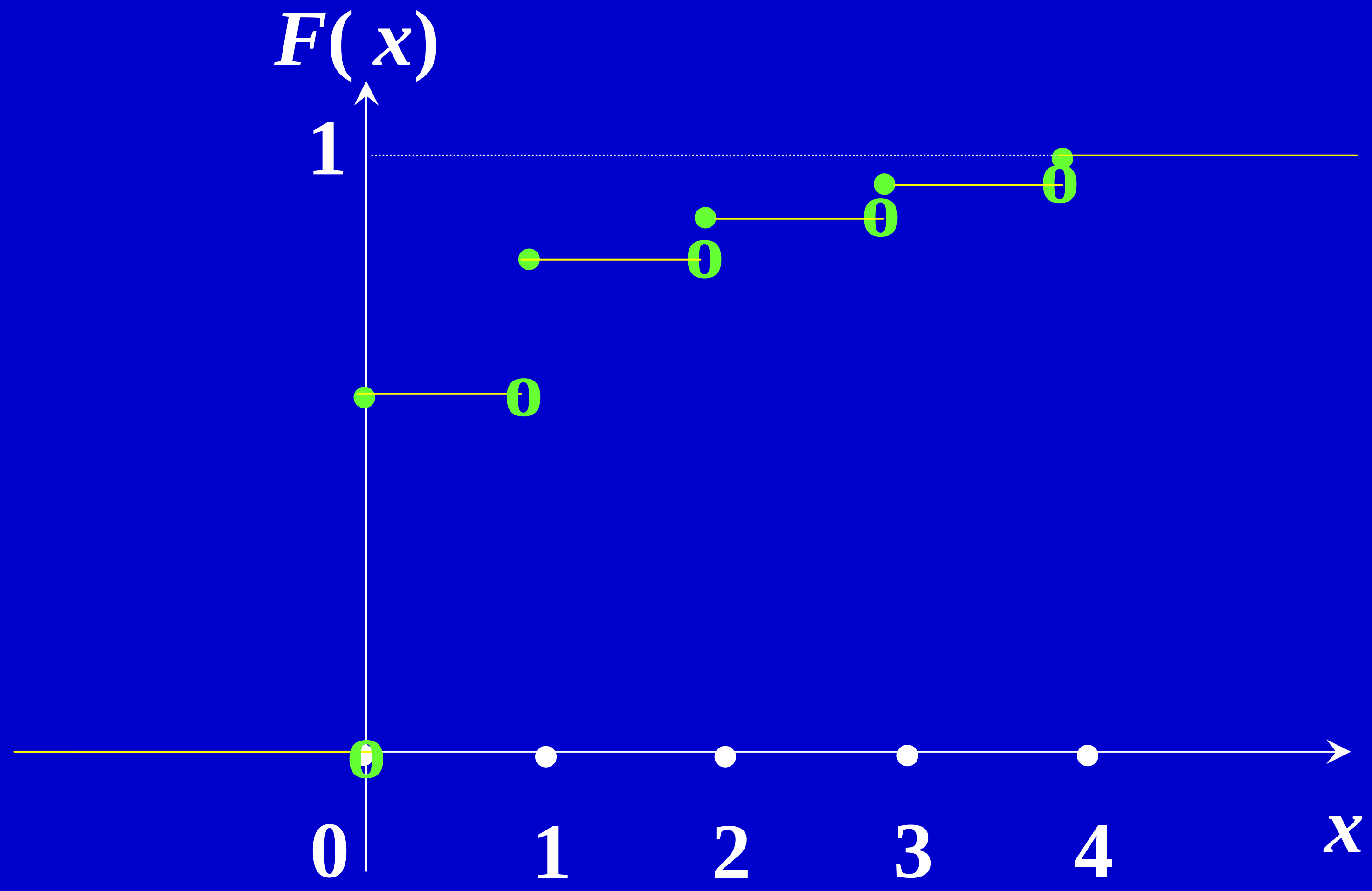
$$P\{X = 4\} = p^4, \qquad k = 4$$

当 $p=0.4$

k	0	1	2	3	4
p_k	0.6	0.4×0.6	$0.4^2\times0.6$	$0.4^3\times0.6$	0.4^4



$$F(x)=\begin{cases}0, & x<0\\ 0.6, & 0\le x<1\\ 0.6+0.6\times0.4, & 1\le x<2\\ 0.6+0.6\times0.4+0.6\times0.4^2, & 2\le x<3\\ 0.6(1+0.4+0.4^2+0.4^3), & 3\le x<4\\ 1 & x\ge 4\end{cases}$$



F(x)
1
0
1
2
3
4
x

分布律或分布函数可用来计算有关事件的概率

例2 在上例中，分别用分布律与分布函数计算下述事件的概率：

$$P\{1 < X \le 3\},\ P\{1 \le X \le 3\},$$

$$P\{X \ge 2\},\ P\{X > 2\}, P\{X = 2\}$$

解 $P\{1 < X \le 3\} = P\{X = 2\} + P\{X = 3\}$

$$= 0.4^2 \times 0.6 + 0.4^3 \times 0.6 = 0.1344$$

或 $P\{1 < X \le 3\} = F(3) - F(1)$

$$= 0.4^2 \times 0.6 + 0.4^3 \times 0.6 = 0.1344$$

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

$$= 0.6(0.4 + 0.4^2 + 0.4^3) = 0.3744$$

或

$$P\{1 \le X \le 3\} = P\{1 < X \le 3\} + P\{X=1\}$$

$$= F(3) - F(1) + P(X=1)$$

$$= F(3) - F(1) + \left[F(1) - F(1-0)\right]$$

$$= F(3) - F(1-0)$$

$$= 0.4^2 \times 0.6 + 0.4^3 \times 0.6 + \left[0.6 \times 0.4\right]$$

$$= 0.3744$$

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

$$= 1 - [P\{X = 0\} + P\{X = 1\}]$$

$$= 0.84$$

或

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

$$= 1 - \left[P\{X \leq 2\} - P\{X = 2\}\right]$$

$$= 1 - F(2 - 0)$$

$$= 0.84$$

此式应理解为极限$\lim\limits_{x \to 2^-} F(x)$

$$P\{X>2\}=1-P\{X\le 2\}$$
$$=1-P\{X=0\}-P\{X=1\}-P\{X=2\}$$
$$=0.936$$

或 $P\{X>2\}=1-P\{X\le 2\}$
$$=1-F(2)$$
$$=0.936$$

$$P\{X=2\}=F(2)-F(2-0)=0.096$$

或 $P\{X=2\}=0.096$

对离散型随机变量用分布律比用分布函数计算这些概率更方便

例3 一门大炮对目标进行轰击，假定此目标必须被击中r 次才能被摧毁。若每次击中目标的概率为p $(0 < p < 1)$, 且各次轰击相互独立，一次一次地轰击直到摧毁目标为止。求所需轰击次数 X 的概率分布。

解 $P\{X = k\} = P\{$前 $k-1$次击中 $r-1$次，第 k 次击中目标$\}$

$$= C_{k-1}^{r-1} p^{r-1} (1-p)^{k-r} \cdot p$$

$$= C_{k-1}^{r-1} p^{r} (1-p)^{k-r}$$

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

3、常见的离散型随机变量的分布

(1) 0 – 1 分布

$X = x_k$	0	1
P_k	$1\text{-}p$	p

$0 < p < 1$

应用场合 凡是随机试验只有两个可能的结果, 常用0 – 1分布描述，如产品是否格、人口性别统计、系统是否正常、电力消耗是否超负荷等等.

注 其分布律可写成

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

(2) 二项分布 $B(n, p)$

背景：n 重Bernoulli 试验中，每次试验感兴趣的事件A 在 n 次试验中发生的次数 —— X 是一离散型随机变量

若$P(A) = p$，则

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

称 X 服从参数为n, p 的二项分布，记作

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

$0-1$ 分布是 $n = 1$ 的二项分布

二项分布的取值情况　设 $X \sim B(8, \frac{1}{3})$

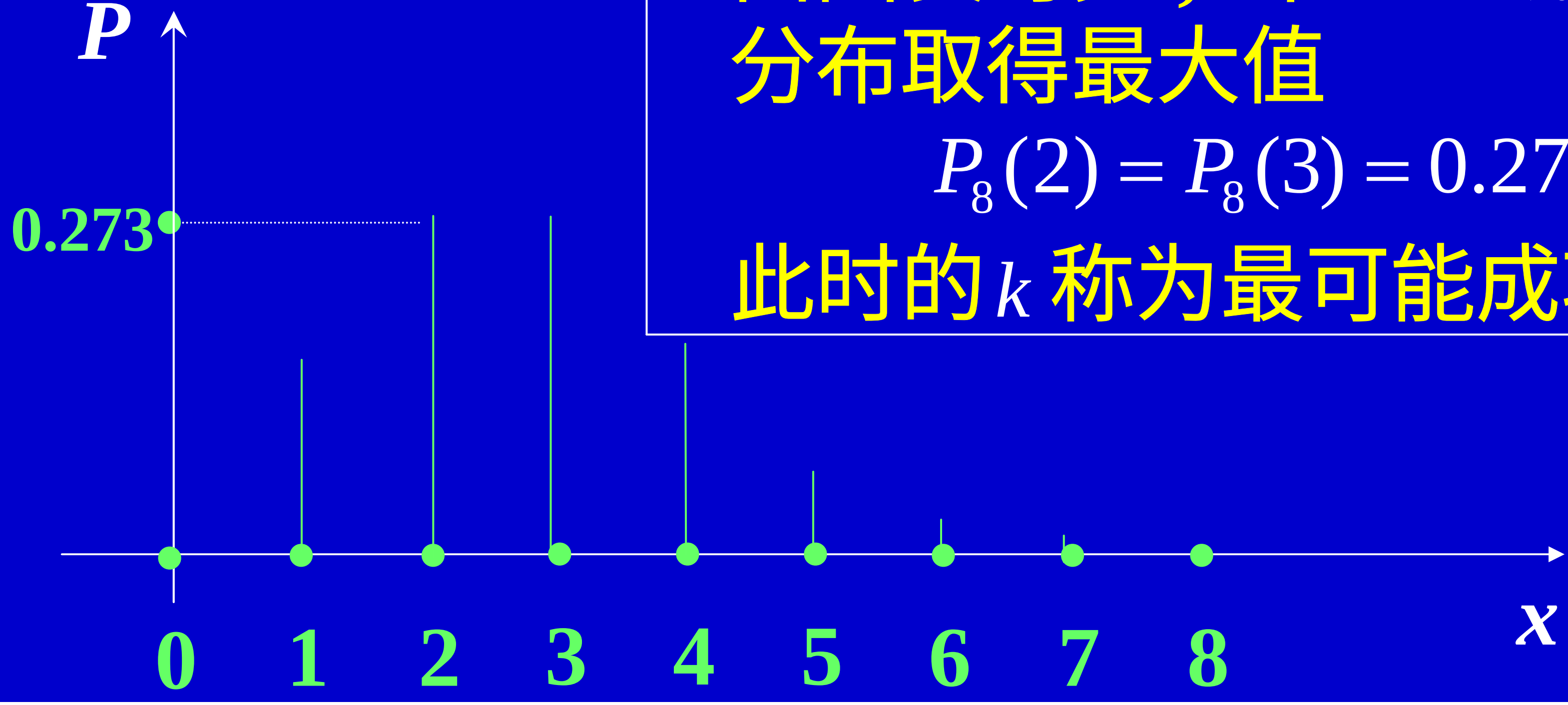
$$P_8(k) = P\{X = k\} = C_8^k \left(\tfrac{1}{3}\right)^k \left(1 - \tfrac{1}{3}\right)^{8-k},\quad k = 0, 1, \cdots, 8$$

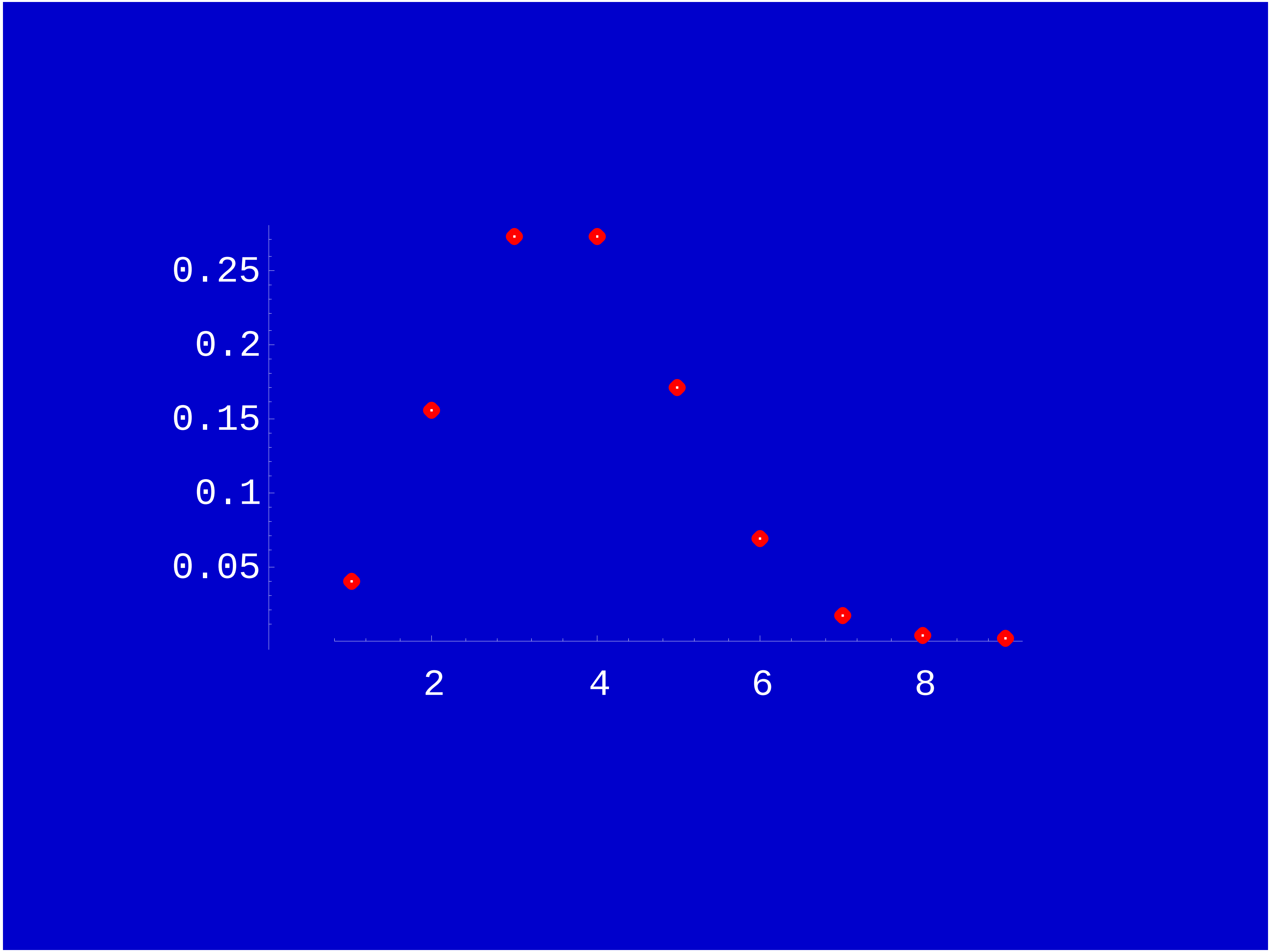
0	1	2	3	4	5	6	7	8
.039	.156	.273	.273	.179	.068	.017	.0024	.0000

由图表可见，当 $k = 2$ 或 3 时，分布取得最大值

$$P_8(2) = P_8(3) = 0.273$$

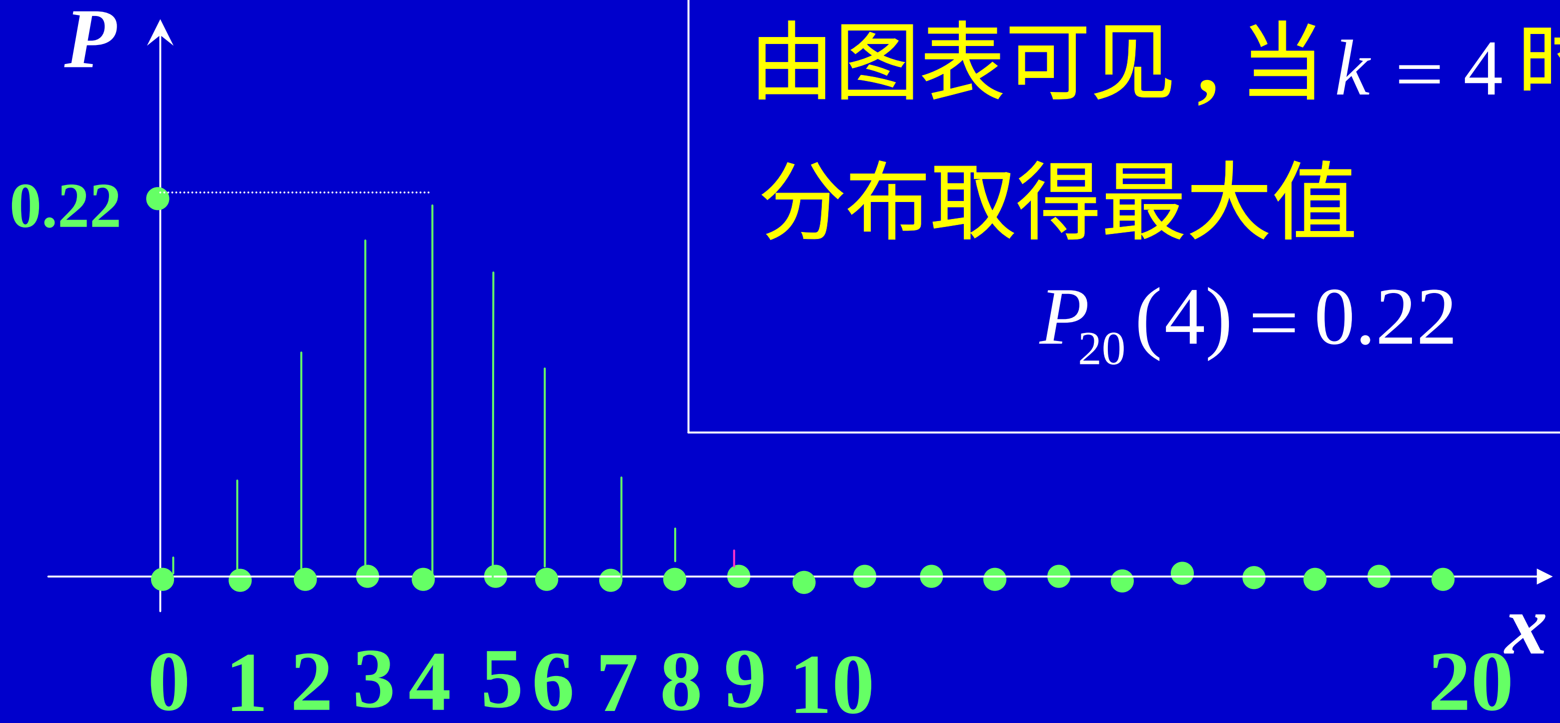
此时的 k 称为最可能成功次数




0.25
0.2
0.15
0.1
0.05
2
4
6
8

设$X \sim B(20, 0.2)$

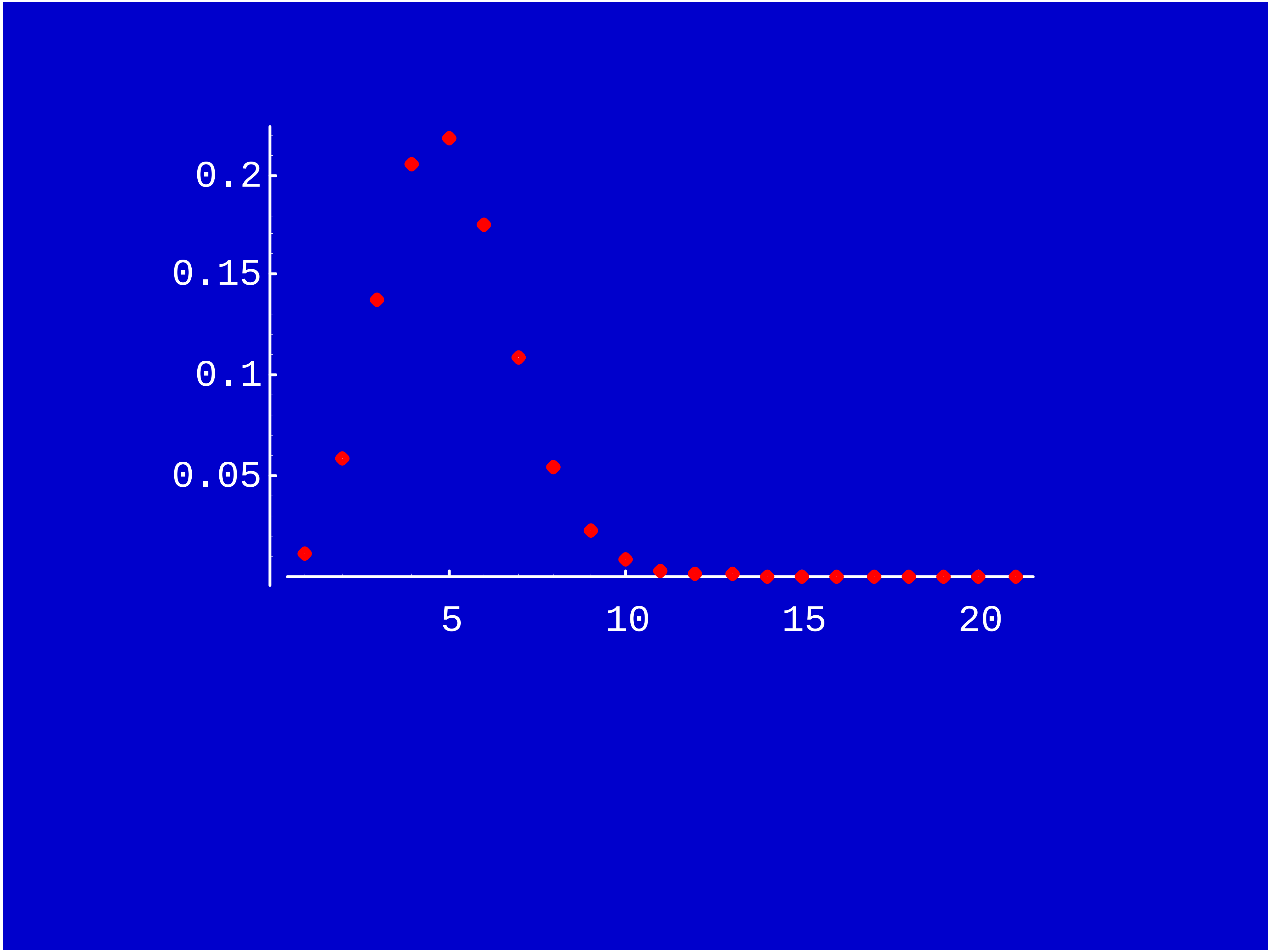
0	1	2	3	4	5	6	7	8	9	10	$11 \sim 20$
.01	.06	.14	.21	.22	.18	.11	.06	.02	.01	.002	$< .001$



由图表可见，当$k = 4$时，

分布取得最大值

$$P_{20}(4) = 0.22$$



0.2
0.15
0.1
0.05
5
10
15
20

Possion定理

设 $\lim\limits_{n\to\infty} np_n = \lambda > 0$

则对固定的 k

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

$$k = 0,1,2,\cdots$$

Poisson定理说明：若$X \sim B(n, p)$, 则当 n 较大，p 较小，而 $np = \lambda$ 适中，则可以用近似公式

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

$$k = 0,1,2,\cdots$$

证　记　$np_n = \lambda_n$

$$C_n^k p_n^k (1-p_n)^{n-k}$$

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

$$= \left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(\frac{\lambda_n{}^k}{k!}\right)\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}\cdot(-\lambda_n)\left(\frac{n-k}{n}\right)}$$

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

类似地，从装有 a 个白球，b 个红球的袋中不放回地任取 n 个球，其中恰有k 个白球的概率为

$$\frac{C_a^k C_b^{n-k}}{C_{a+b}^n}$$

当 $a+b\to\infty, \dfrac{a}{a+b}\to p$ 时，

对每个 n 有 $$\frac{C_a^k C_b^{n-k}}{C_{a+b}^n}\to C_{a+b}^k p^k(1-p)^{n-k}$$

结论

超几何分布的极限分布是二项分布

二项分布的极限分布是 Poisson 分布

例5 某厂产品不合格率为0.03，现将产品装箱，若要以不小于90%的概率保证每箱中至少有100个合格品，则每箱至少应装多少个产品？

解 设每箱至少应装100+ n 个，每箱的不合格品个数为X，则$X \sim B(100+n, 0.03)$

由题意 $$P(X \le n) = \sum_{k=0}^{n} P_{100+n}(k) \ge 0.9$$

$$(100+n)0.03=3+0.03n\approx 3 \qquad \text{取 } \lambda = 3$$

应用Poisson定理

$$\sum_{k=0}^{n} P_{100+n}(k) \approx \sum_{k=0}^{n} \frac{3^k}{k!}e^{-3} = 1 - \sum_{k=n+1}^{\infty} \frac{3^k}{k!}e^{-3} \ge 0.9$$

$$\sum_{k=n+1}^{\infty}\frac{3^k}{k!}e^{-3}\le 0.1$$ 查Poisson分布表λ =3一栏

得 $n+1=6$， $n=5$

所以每箱至少应装105个产品,才能符合要求.

在Poisson 定理中，

$$\frac{\lambda^k}{k!}e^{-\lambda} > 0$$

$$\begin{aligned}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda} &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\\ &= e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots\right)\\ &= e^{-\lambda}\cdot e^{\lambda} = 1\end{aligned}$$

由此产生了一种离散型随机变量的概率分布
— Poisson 分布

(3) Poisson 分布 $\pi(\lambda)$ 或 $P(\lambda)$

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

其中 $\lambda>0$ 是常数，则称 X 服从参数为 λ 的**Poisson 分布**，记作 $\pi(\lambda)$ 或 $P(\lambda)$

应用场合

在一定时间间隔内：

电话总机接到的电话次数；

一匹布上的疵点个数；

大卖场的顾客数；

市级医院急诊病人数；

一个容器中的细菌数；

某一地区发生的交通事故的次数

放射性物质发出的粒子数；

一本书中每页印刷错误的个数；

等等

都可以看作是源源不断出现的随机质点流，若它们满足一定的条件，则称为**Poisson**流，在长为 t 的时间内出现的质点数 $X_t \sim P(\lambda t)$