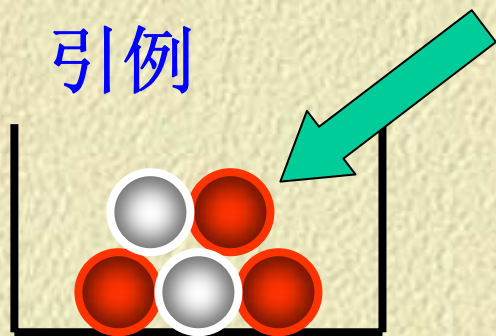


第二节 离散型随机变量的概率分布（分布律）

一. 离散型随机变量的分布律

引例



如图中所示，现从中任取 3 个球，取到的白球数 X 是一个随机变量。

X 可能取的值是 0, 1, 2

取每个值的概率为：

$$P(X = 0) = \frac{C_3^3}{C_5^3} = \frac{1}{10}$$

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

$$P(X = 2) = \frac{C_3^1 C_2^2}{C_5^3} = \frac{3}{10}$$

$$\text{且：} \sum_{i=1}^3 P(X = i) = 1$$

1. 定义：设离散型随机变量 X 所有可能取的值为 $x_k, k = 1, 2, \dots$ 其各个可能取值即事件 $\{X = x_k\}$ 的概率为：

$$P(X = x_k) = p_k \quad k = 1, 2, \dots$$

则称 $P(X = x_k) = p_k$ 为离散型随机变量 X 的概率分布或分布律。

注：分布律可以列表给出：

X	x_1	x_2	\dots	\dots	x_n	\dots	\dots
P_k	p_1	p_2	\dots	\dots	p_n	\dots	\dots

2. 性质

$$(1). p_k \geq 0,$$

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

$$(2). \sum_{k=0}^{\infty} p_k = 1$$

用这两条性质判断一个函数是否是概率函数

注 ▲ **一般**：求分布律时需验证这两条性质。若成立则称其为分布律，否则不能表明所求的是分布律。

▲ 具有离散型随机变量才具有分布律

例1. 设在15只同类型的零件中有两只次品，现从中抽取3只，以 X 表示取出3只中所含次品的个数。求： X 的分布律。

解： 由题意， X 的可能取值：0, 1, 2

X 的各种可能取值的概率如下：

$$P(X = 0) = \frac{C_{13}^3 C_2^0}{C_{15}^3} = \frac{22}{35}$$

$$P(X = 1) = \frac{C_{13}^2 C_2^1}{C_{15}^3} = \frac{12}{35}$$

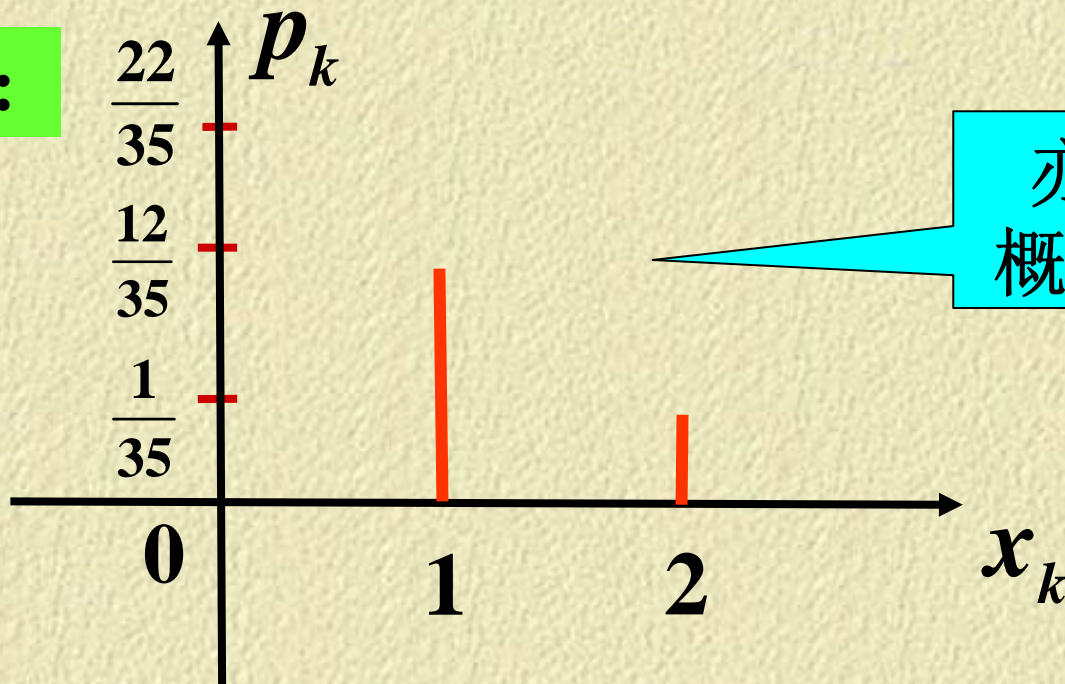
$$P(X = 2) = \frac{C_{13}^1 C_2^2}{C_{15}^3} = \frac{1}{35}$$

所以其分
布律为:

X	0	1	2
P_k	$\frac{22}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

(显然每个 $P_k > 0$, $\sum_{k=0}^3 P_k = 1$)

图形:



亦称其为
概率分布图

例2. 思考题：投篮命中率为0.9，求两次独立投篮投中次数 X 的概率分布

答案： $P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{2.2}{35} + \frac{12}{35}$

解： X 可能取值为 0、1、2 则：

$$P(X = 0) = (0.1)(0.1) = 0.01$$

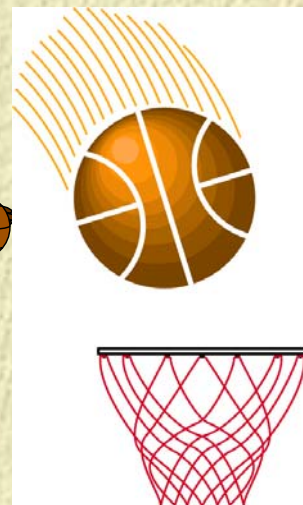
$$P(X = 1) = 2(0.9)(0.1) = 0.18$$

$$P(X = 2) = (0.9)(0.9) = 0.81$$

且 $P(X = 0) + P(X = 1) + P(X = 2) = 1$

故得其分布律为：

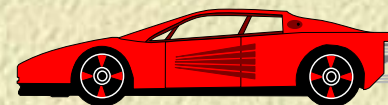
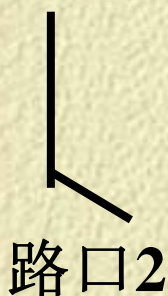
X	0	1	2
P_k	0.01	0.18	0.81



例3. 一汽车沿一街道行驶，需要通过三个均设有红绿灯信号灯的路口，每个信号灯为红或绿与其它信号灯为红或绿相互独立，且红绿两种信号灯显示的时间相等。以 X 表示该汽车首次停车时已通过的路口的个数。求： X 的概率分布。

解：依题意， X 可取值 0, 1, 2, 3

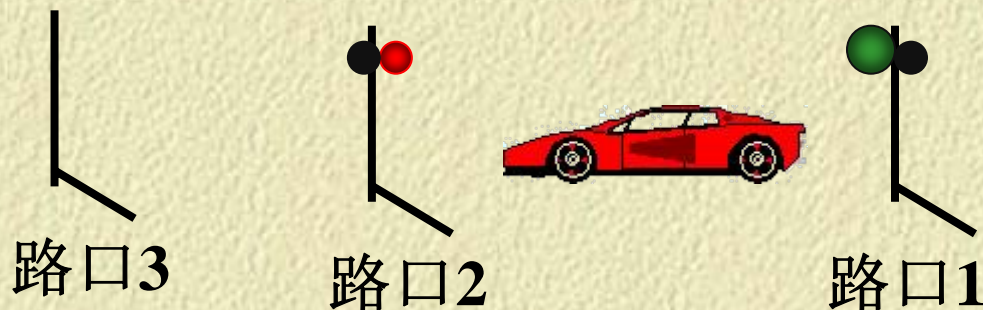
设 $A_i = \{ \text{第} i \text{个路口遇红灯} \}, i = 1, 2, 3$



则： $P(X=0) = P(A_1) = 1/2$

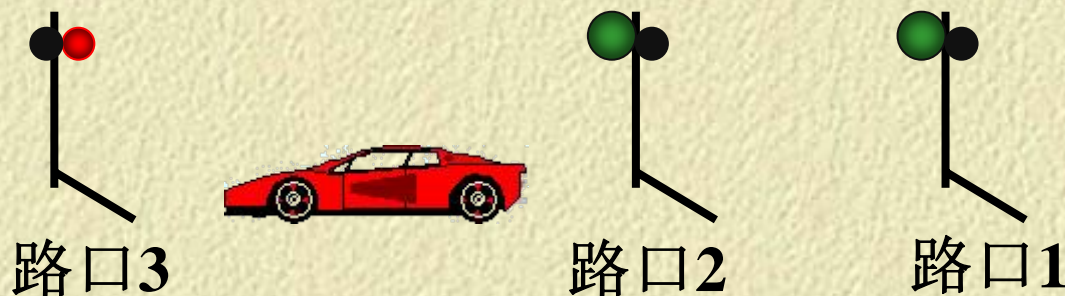
设 $A_i = \{ \text{第} i \text{个路口遇红灯} \}, i = 1, 2, 3$

X 表示该汽车首次停车时已通过的路口个数



$$P(X = 1) = P(\bar{A}_1 A_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

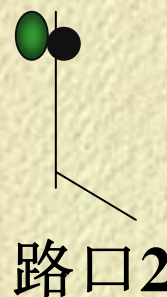
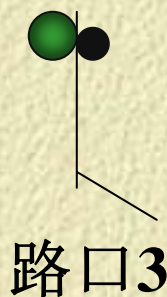
设 $A_i = \{ \text{第} i \text{个路口遇红灯} \}, i = 1, 2, 3$



$$P(X = 2) = P(\bar{A}_1 \bar{A}_2 A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

X 表示该汽车首次停车时已通过的路口的个数

设 $A_i = \{ \text{第 } i \text{ 个路口遇红灯} \}, i = 1, 2, 3$



$$P(X = 3) = P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

于是得其分布律为：

X	0	1	2	3
P_k	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

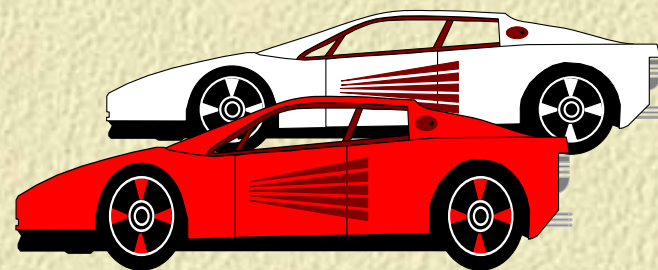
显然：

$$\sum_{i=0}^3 P(X = i) = 1$$

例4. 某加油站为某汽车公司代营出租汽车业务，每出租一辆汽车，可从出租公司得到 3 元。因代营业务，每天加油站要多付给职工服务费 60 元。设每天出租汽车数 X 是一个随机变量，它的概率分布如下：

X	10	20	30	40
P_k	0.15	0.25	0.45	0.15

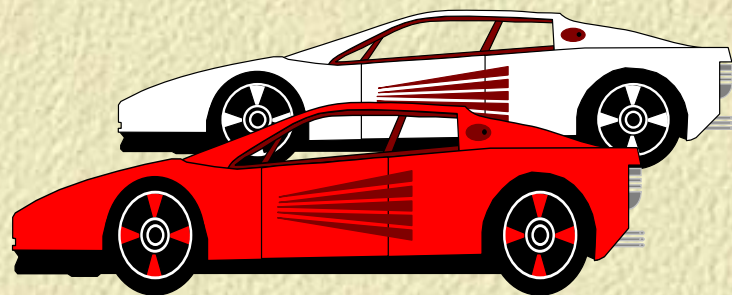
求：因代营业务得到的收入大于当天的额外支出费用的概率。



分析 因为加油站代营每出租一辆车，可得 3 元。
若设每天出租汽车数为 X
则加油站因代营业务得到的收入为 $3X$ (元)
而每天加油站要多付给职工服务费 60(元)，
即当天的额外支出费用。

所以 加油站因代营业务得到的收入大于当天的额外支出费用的概率为：

$$P(3X > 60) = P(X > 20)$$



由已知,

X	10	20	30	40
P_k	0.15	0.25	0.45	0.15

所以得:
$$P(X > 20) = P(X = 30) + P(X = 40)$$
$$= 0.45 + 0.15 = 0.6$$

结论 加油站因代营业务得到的收入大于当天的额外支出费用的概率为 **0.6**;
故加油站的经营决策者应该考虑是否继续代营此项业务或应该考虑是否**调整**当天的额外支出费用。

二. 几种常见的离散型随机变量的分布

1.(0—1)分布

若随机变量 X 只能取 0 与 1 两个值,它的分布律为:

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

则称 X 服从 (0--1)分布, 记为: $X \sim (0,1)$

列表:

X	0	1
P_k	$1-p$	p

例5. 某次射击，已知某射手的命中率为0.8.

求：射击一次命中目标次数 X 的分布律.

解： \because 它只发一弹，要么打中，要么打不中，分别记为 1 与 0

\therefore 分布律为：

X	0	1
P_k	0.2	0.8

注： (0—1)分布的应用很广，比如：

检查产品的质量(正品与次品)
有奖储蓄券是否中奖(中与不中)
对婴儿性别进行登记(男与女)
高射炮射击敌机是否击中等等.

2. 二项分布

(1). 贝努利概型

1⁰. n 次相互独立试验:

重复进行 n 次试验，若各次试验的结果互不影响，即每次试验结果出现的概率都不受其它各次试验结果的影响。则称这 n 次试验是相互独立的。

说明: 把在相同的条件下重复进行 n 次独立试验的概率模型, 称为 n 次独立试验模型.

2⁰. 贝努利概型: 设随机试验 E 只有两种可能的结果 A 与 \bar{A} 且在每次试验中 A 与 \bar{A} 出现的概率为:

$$P(A)=p, \quad P(\bar{A})=1-p=q \quad (0 < p < 1)$$

则称这样的 n 次重复独立试验概型为: **n 重贝努利概型.**

例6. 若设生男孩的概率为 p , 生女孩的概率为 $q = 1-p$, 令 X 表示随机抽查出生的 4 个婴儿中“男孩”的个数.

求: X 的概率分布.

解: 显然, 它属于贝努利概型.



X 表示随机抽查的 4 个婴儿中男孩的个数，
现设生男孩的概率为 p .



男

女

$X=0$



$X=1$



$X=2$



$X=3$



$X=4$



X 的概率函数是：

X 的取值为：0, 1, 2, 3, 4.

$$P\{X = k\} = C_4^k p^k (1-p)^{4-k}, \quad k = 0, 1, 2, 3, 4$$

例7. 现若将一枚均匀骰子抛掷 3 次,
令: X 表示 3 次中出现“4”点的次数
求: X 的概率函数



解: 显然, 它属于贝努利概型.

X 的取值为: 0, 1, 2, 3



从而, X 的概率函数为:

$$P\{X = k\} = C_3^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{3-k}, \quad k = 0, 1, 2, 3$$

3.0 定理（贝努利定理）

设一次试验中事件A发生的概率为 p , ($0 < p < 1$)

则在 n 次贝努利试验中事件A 恰发生 k 次概率为:

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \quad (k = 0, 1, 2, \dots, n)$$

证明: 按独立事件的概率计算公式可知:

n 次试验中事件A 在某 k 次 (例如前 k 次)

发生而其余 $n-k$ 次不发生的概率应为:

$$\underbrace{p \cdot p \cdots p}_k \cdot \underbrace{(1-p)(1-p) \cdots (1-p)}_{n-k} = p^k (1-p)^{n-k}$$

由于现在只考虑事件A 在n 次试验中发生 k 次而不论在哪 k 次发生，所以它应有 C_n^k 种不同的发生方式。而且它们是相互独立的，故在 n 次试验中 A 发生 k 次的概率（依概率的加法定理）为：

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \quad k = 0, 1, 2, \dots, n$$

注 ▲ 显然它满足：

$$P(X=k) \geq 0,$$

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

概率 $P_n(k)$ 就等于二项式 $[(1-p)+px]^n$ 的展开式中 x^k 的系数，这也是二项分布的名称的由来。

例8. 设某炮手射击的命中率为 **0.8**，为炸毁某个目标， 经预测只要命中两发就够炸毁。

问： 希望发射5发炮弹就能炸毁目标的可能性有多大？

解： 设 **A**：发射 5 发炮弹就炸毁了目标 则：

$$\begin{aligned} P(A) &= P(\text{至少中两发}) \\ &= P(\text{恰好中两发}) + P(\text{恰好中三发}) \\ &\quad + P(\text{恰好中四发}) + P(\text{恰好中五发}) \\ &= C_5^2(0.8)^2(1-0.8)^3 + C_5^3(0.8)^3(1-0.8)^2 \\ &\quad + C_5^4(0.8)^4(1-0.8) + C_5^5(0.8)^5(1-0.8)^0 \\ &= 0.98 \end{aligned}$$

(2). 二项分布

若用 X 表示 n 重贝努利概型中事件 A 发生的次数，它的分布律为：

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \quad k = 0, 1, 2, \dots, n$$

则称 X 服从参数为 n, p ($0 < p < 1$) 的二项分布。

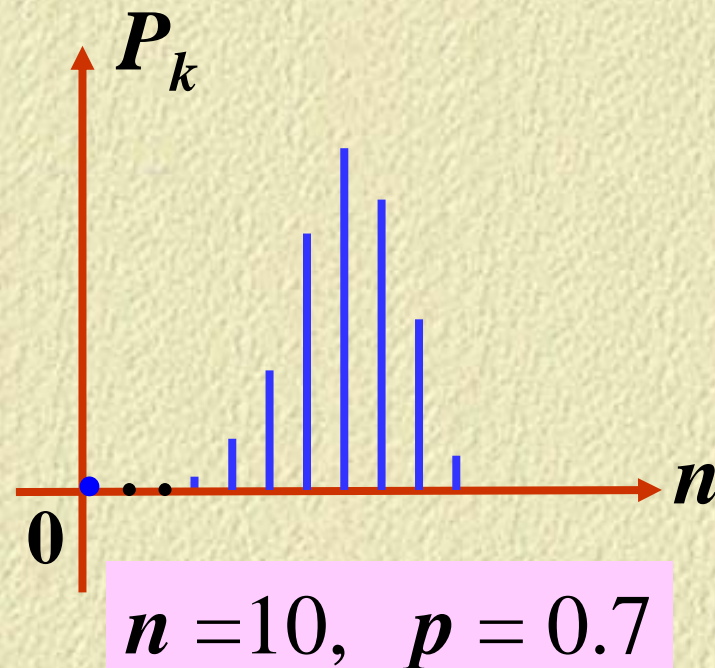
记为： $X \sim B(n, p)$

列表：	X	0	1	2.....n
	$P_n(k)$	$P_n(0)$	$P_n(1)$	$P_n(2).....P_n(n)$

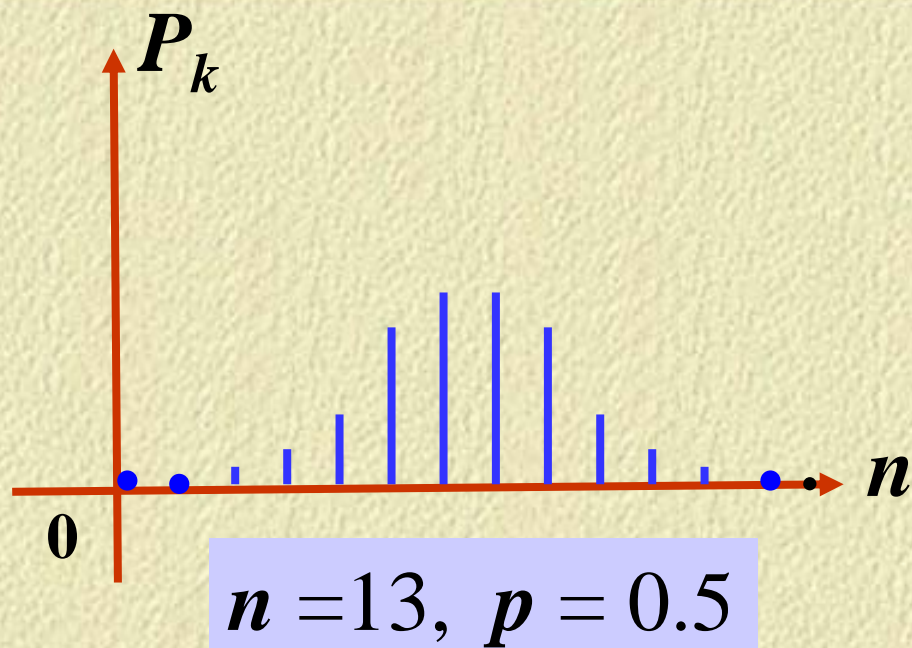
注 ▲ 特别当 $n=1$ 时，二项分布即为 (0---1) 分布

▲ 二项分布 $X \sim B(n, p)$ 的图形特点：

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。



当 $(n+1)p$ 为整数时
概率 $P(X=k)$ 在
 $k=(n+1)p$ 和
 $k=(n+1)p-1$ 处
达到最大值.



当 $(n+1)p$ 不为整数时, 概率 $P(X=k)$
在 $k=[(n+1)p]$ 达到最大值

其中: $[x]$ 表示不超过 x 的最大整数

例9. 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个产品中恰有 2个次品的概率。

解： 因为，这是有放回地取 3 次，
所以，这3 次试验的条件完全相同且独立。

因此，它是贝努利试验。

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：

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

注 若将例9中的“有放回”改为“无放回”，那么各次试验条件就不同了，就不是贝努里概型，此时，只能用古典概型求解。

贝努利概型与古典概型有何区别？

$$P(X=2) = \frac{C_{95}^1 C_5^2}{C_{100}^3}$$

贝努里概型对试验结果没有等可能的要求，但要求

(1) 每次试验条件相同，各次试验相互独立

(2) 每次试验只考虑两个互逆结果 A 或 \bar{A}

且 $P(A) = p$, $P(\bar{A}) = 1 - p$

例10. 若一年中参加人寿保险者里面每个人死亡的概率为0.005，现有10000个这类人参加人寿保险.

试求：在未来一年中在这些保险者里面：

- (1). 有10人死亡的概率
- (2). 死亡人数不超过10人的概率.

解： 显然，这是**贝努利概型**.

设 X ：在未来一年中这些保险者中的死亡人数.

则： $X \sim B(10000, 0.005)$

(1). 有10人死亡的概率为：

$$P(X = 10) = C_{10000}^{10} (0.005)^{10} (0.995)^{9990}$$

(2). 死亡人数不超过 10 人的概率是:

$$P(X < 10) = \sum_{k=0}^{10} C_{10000}^k (0.005)^k (0.995)^{10000-k}$$

这些计算是非常麻烦的，现给出一个当 n 很大， p 很小时的近似计算公式，即二项分布的 **Possion 逼近**.

泊松(Possion)
定理

设 $X \sim B(n, p)$, $\lambda > 0$ 是一常数，
且 $\lambda = nP_n$ ，则对任一固定的非负
整数 k 有：

$$\lim_{n \rightarrow \infty} P_n(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

证明: $\because P_n(k) = C_n^k p_n^k (1-p)^{n-k}$

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

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

$\rightarrow e^{-\lambda}$

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

$$\therefore \lim_{n \rightarrow \infty} P_n(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

注：一般的用 $\frac{\lambda^k e^{-\lambda}}{k!}$ 去近似二项分布的 $P_n(k)$ 当：

$n \geq 20, p \leq 0.05$ 时近似效果颇佳

$n \geq 100, np \leq 10$ 时近似效果更好

泊松定理中 $\frac{\lambda^k e^{-\lambda}}{k!}$ 的值有表可查

见本教材
第二版的
P 372的
附表3

例11. 用泊松定理中的近似公式计算例 10

解：(1) $\because \lambda = 10000 \times 0.0005 = 50$

$$\therefore P(X = 10) \approx \frac{(50)^{10} \cdot e^{-50}}{10!}$$

$$(2) P(X < 10) \approx \sum_{k=0}^{10} \frac{(50)^k e^{-50}}{k!}$$

1万人参加
保险，每人的
死亡率为
0.005.

求：10人死
亡；小于10
人死亡的概
率。

注：这里 $\lambda = 50$ 附表3 没有列入， n 确实很大时更进一步的计算将在第五章介绍中心极限定理之后再来解决比较方便.

若现将“每个人死亡的概率改为 0.0005”，则 $\lambda = 5$

$$\text{从而: } P(X = 10) = \frac{5^{10} \cdot e^{-5}}{10!} = \frac{9765625 \times 0.0067379}{3628800} \\ = 0.01813$$

$$P(X < 10) = 1 - P(X \geq 10) \approx 1 - \sum_{k=10}^{\infty} \frac{5^{10} e^{-5}}{10!} \\ = 1 - 0.031828 \\ = 0.96817$$

例12 设有80 台同类型设备，各工作是相互独立的，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。现考虑两种配备维修工人的方法：其一是由4人维护，每人负责 20台；其二是由 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）在第一种配备方法中

A_i ：第 i 人维护的20台中发生故障不能及时维修

X ：第 1 人维护的20台中同一时刻发生故障的台数

则： 在 80 台中发生故障而不能及时维修的概率为：

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \geq P(A_1) = P(X \geq 2)$$

而: $X \sim B(20, 0.01)$, $\lambda = np = 0.2$

$$\therefore P(X \geq 2) \approx \sum_{k=2}^{\infty} \frac{(0.2)^k \cdot e^{-0.2}}{k!} = 0.0175231$$

即: $P(A_1 \cup A_2 \cup A_3 \cup A_4) \geq 0.0175$

(2) 在第二种配备方法中

设 Y : 80 台中同一时刻发生故障的设备台数

而: $Y \sim B(80, 0.01)$, $\lambda = np = 0.8$

则, 在 80 台中发生故障而不能及时维修的概率为:

$$P(Y \geq 4) \approx \sum_{k=4}^{\infty} \frac{(0.8)^k e^{-0.8}}{k!} = 0.0091$$

结论 经比较, 采用第二种配备方法虽然人员减少, 每个人的任务加重(每人平均维护 27 台), 但质量不仅没降低, 反而提高了, 故应采用第二种配备方法。

3. 泊松分布

定义: 若随机变量 X 的所有可能取值为: $0, 1, 2, \dots$ 而它的分布律 (它所取值的各个概率) 为:

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

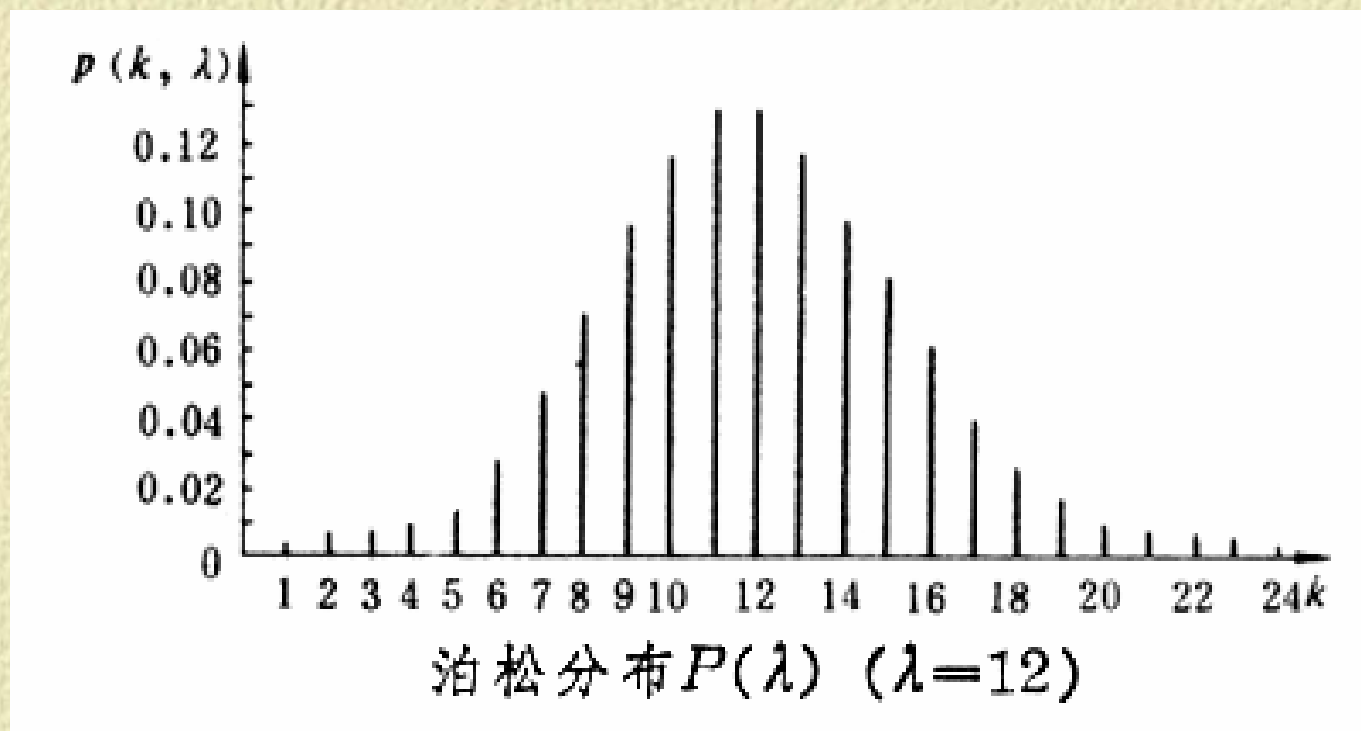
其中 $\lambda > 0$ 是常数

则称 X 服从参数为 λ 的泊松分布, 记为: $X \sim P(\lambda)$

注 ▲ 泊松分布满足分布律的两个条件:

$$P(X = k) \geq 0, \quad \sum_{k=0}^{\infty} P(X = k) = 1$$

▲ 泊松分布 $X \sim P(\lambda)$ 的图形特点:



▲ 二项分布与泊松分布的关系

由泊松分布的定义及泊松定理可知：

当 $n \rightarrow \infty$ 泊松分布是二项分布的近似。



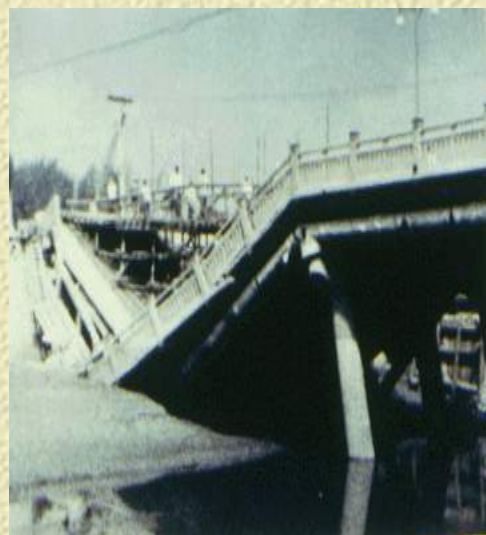
泊松, S. -D.

（这是1837年由法国数学家泊松引入的）

▲ 在实际中，许多随机现象服从或近似服从泊松分布。若把在每次试验中出现概率很小的事件称作稀有事件。比如：

比如：

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。

▲ 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。

一般把在随机时刻相继出现的事件所形成的序列称为随机事件流。

若随机事件流具有平稳性、无后效性、普通性，

则称该事件流为泊松事件流（泊松流）

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。

例如：

一放射性源放射出的
 α 粒子数



例如

某电话交换台收到的电话呼叫数



一个售货员接待的顾客数



到某机场降落的飞机数



一台纺纱机的断头数

.....

都可以看作泊松流.

例13. 一家商店拟采用科学管理，由该商店过去的销售记录知道，某种商品每月的销售数可以用参数 $\lambda = 5$ 的泊松分布来描述，现为了以 95% 以上的把握保证不脱销。

问：商店在月底至少应进该种商品多少件？

解： 设 X ：该商品每月的销售数

由已知 X 服从参数 $\lambda = 5$ 的泊松分布： $X \sim P(5)$

又设商店在月底应进该种商品 m 件

则问题为 求满足 $P(X \leq m) > 0.95$ 的最小的 m

销售数

进货数

求满足 $P(X \leq m) > 0.95$ 的最小的 m

也即求: $P(X > m) \leq 0.05$

或
$$\sum_{k=m+1}^{\infty} \frac{e^{-5} 5^k}{k!} \leq 0.05$$

查泊松分布表得:

$$\sum_{k=10}^{\infty} \frac{e^{-5} 5^k}{k!} \approx 0.032,$$

$$\sum_{k=9}^{\infty} \frac{e^{-5} 5^k}{k!} \approx 0.068$$

于是得 $m+1=10$,

即: $m=9$ (件)