2.4 常见的离散型随机变量的分布

(1) 0 – 1 分布

$X = x_k$	1	0
P_k	p	$1-p$

$0 < p < 1$

应用场合 凡是随机试验只有两个可能的结果，

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

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

服从(0-1)分布的试验叫做贝努利试验。

随机试验有两个可能的结果：$A, \overline{A}$

设 $P(A) = p, \quad 0 < p < 1$

将此试验独立的重复 n 次

称为 n 重Bernoulli 试验概型：

n 重Bernoulli 试验概型感兴趣的问题为：
在 n 次试验中事件 A 出现 k 次的概率，记为

$$P_n(k)$$

(2) **二项分布** $B(n,p)$

背景：n 重Bernoulli 试验中，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

$(X=k)$ 即事件A在n次试验中发生了k次

这k次可能是n次试验中任意的某个k次，共有C_n^k种情况

对于某个指定的k次，其余n-k次A没有发生，概率为$p^k(1-p)^{n-k}$

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

称 X 服从参数为 n, p 的二项分布，记作

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

上述定义满足概率的基本性质，即：

$$C_n^k p^k(1-p)^{n-k} \geq 0$$

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

$0-1$ 分布是 $n=1$ 的二项分布

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

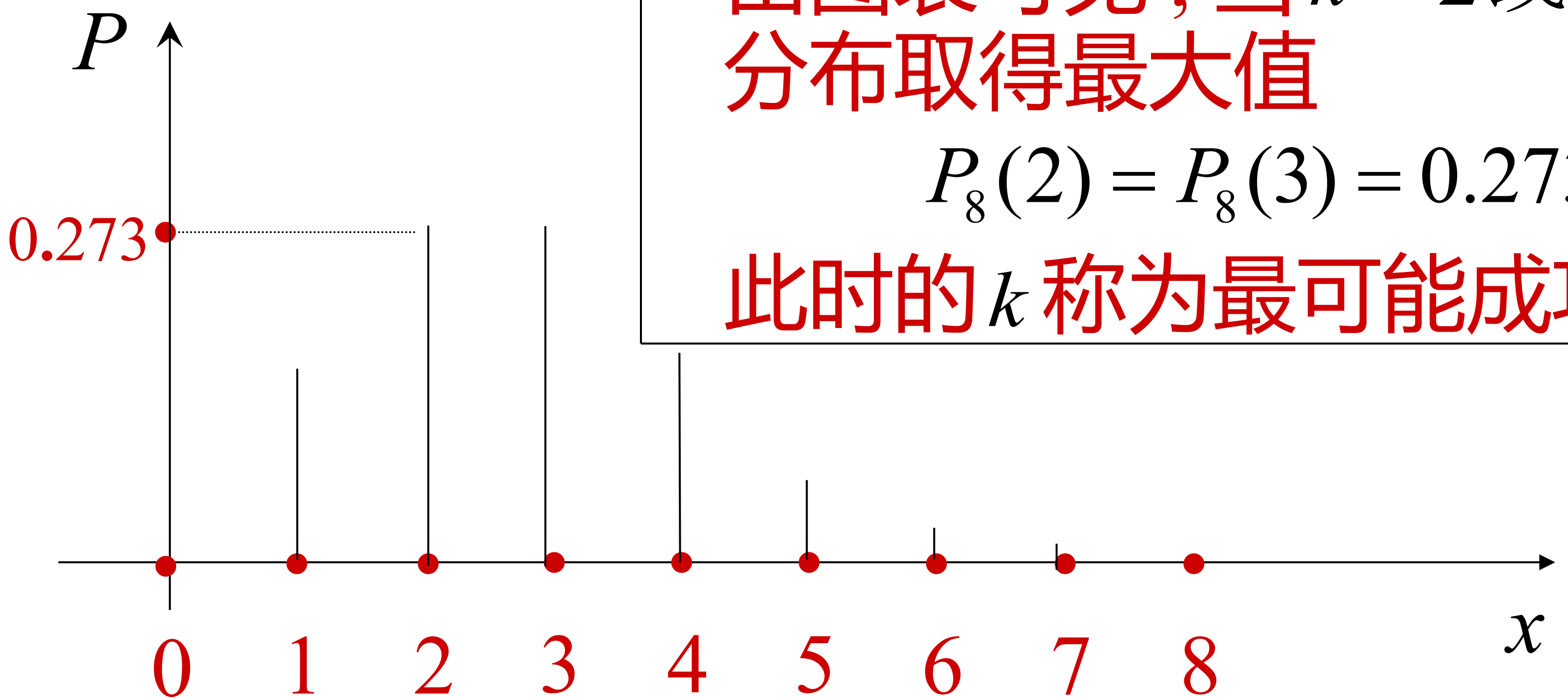
$$P_8(k) = P(X = k) = C_8^k (\tfrac{1}{3})^k (1 - \tfrac{1}{3})^{8-k}, \quad k = 0, 1, \cdots, 8$$

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

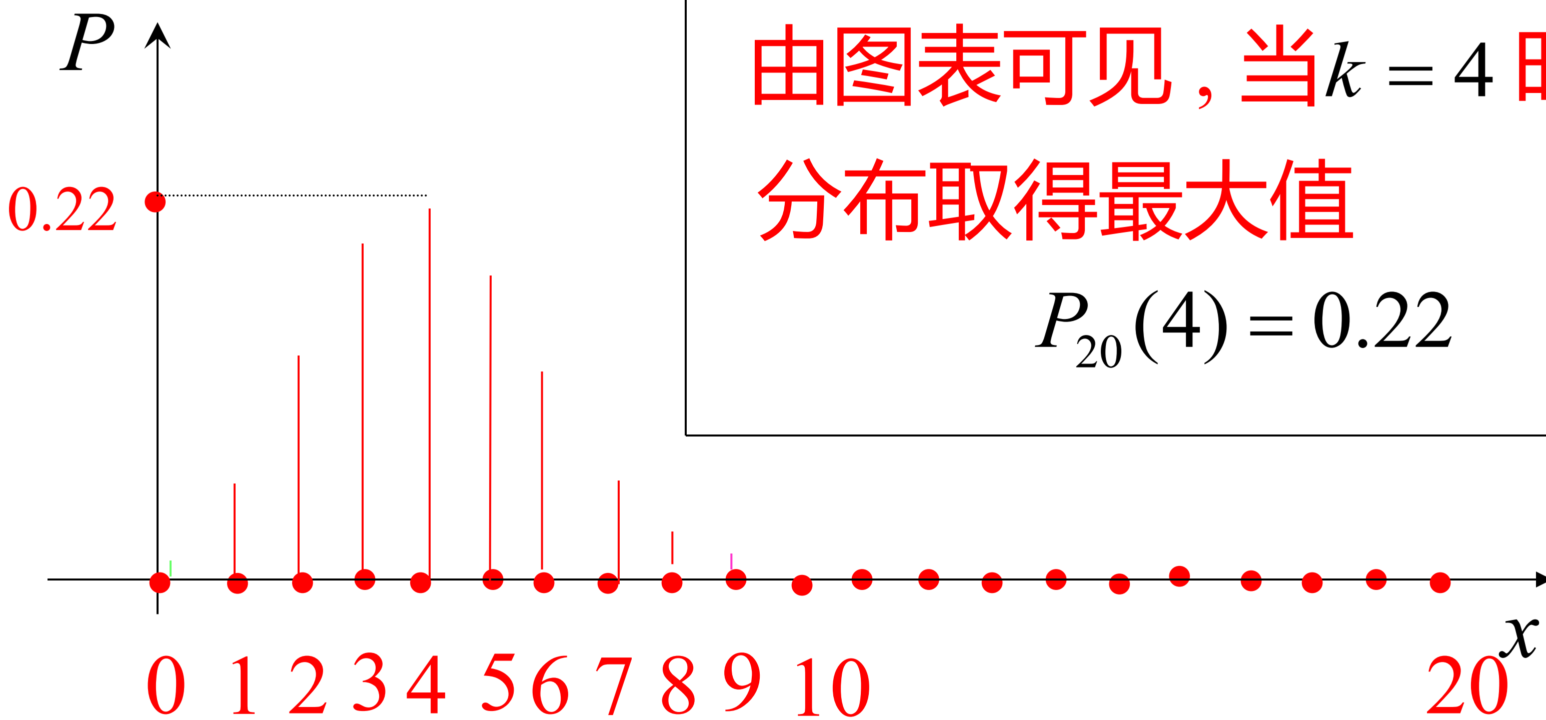
$$P_8(2) = P_8(3) = 0.273$$

此时的 k 称为最可能成功次数



设$X \sim B(20, 0.2)$

0	1	2	3	4	5	6	7	8	9	10	11 ~ 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

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

二项分布中最可能出现次数的定义与推导

若 $P(X=k)\geq P(X=j)$，$j=X$ 可取的一切值

则称k为最可能出现的次数

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

$$\left.\begin{array}{l}\dfrac{p_{k-1}}{p_k}=\dfrac{(1-p)k}{p(n-k+1)}\leq 1\\[2ex] \dfrac{p_k}{p_{k+1}}=\dfrac{(1-p)(k+1)}{p(n-k)}\geq 1\end{array}\right\}$$

$\Rightarrow$ $(n+1)p-1\leq k\leq(n+1)p$

- 当$(n+1)p=$整数时，在$k=[(n+1)p]$与$[(n+1)p]-1$处的概率取得最大值

  当$(n+1)p \neq$整数时，在$k=[(n+1)p]$处的概率取得最大值

例1 独立射击5000次，每次的命中率为0.001，求 (1) 最可能命中次数及相应的概率；

(2) 命中次数不少于2 次的概率.

解 (1) $k = [(n+1)p] = [(5000+1)0.001] = 5$

$$P_{5000}(5) = C_{5000}^{5}(0.001)^{5}(0.999)^{4995} \approx 0.1756$$

(2) 令X表示命中次数，则 $X \sim \mathrm{B}(5000, 0.001)$

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - \sum_{k=0}^{1} C_{5000}^{k}(0.001)^{k}(0.999)^{5000-k}$$

$$= 0.9574$$

问题 如何计算 $P(X \geq 2500)$ ？

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

$$C_n^k p^k (1-p)^{n-k} \approx e^{-\lambda} \frac{\lambda^k}{k!} \quad k=0,1,2,\cdots$$

证 记 $np=\lambda$ 则 $p=\frac{\lambda}{n}$ 那么有

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

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

$$=\frac{\lambda^k}{k!}\{(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^{-k}\}(1-\frac{\lambda}{n})^n$$

$$= \frac{\lambda^k}{k!}\{(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^{-k}\}(1-\frac{\lambda}{n})^n$$

当 $n \to \infty$

$$\{(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^{-k}\} \to 1$$

$$(1-\frac{\lambda}{n})^n = (1-\frac{\lambda}{n})^{-\frac{n}{\lambda}\cdot(-\lambda)} \to e^{-\lambda}$$

故 $\lim\limits_{n\to\infty} C_n^k p^k (1-p)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$

称为Poisson定理

类似地，设一批产品中有M件正品，N件次品。从中任意取n件，则恰好取到k个次品的概率为 $\dfrac{C_N^k C_M^{n-k}}{C_{M+N}^n}$

当 $M+N\to\infty, \dfrac{N}{M+N}\to p$ 时，

对每个 n 有 $\dfrac{C_N^k C_M^{n-k}}{C_{M+N}^n} \to C_{M+N}^k p^k (1-p)^{n-k}$

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

注：

$$\frac{C_N^k C_M^{n-k}}{C_{M+N}^n} = \frac{n!}{k!(n-k)!} \cdot \frac{N!M!(M+N-n)!}{(M+N)!(N-k)!(M-n+k)!}$$

$$= C_n^k \left(\frac{N}{M+N}\right)^k \left(\frac{M}{M+N}\right)^{n-k}$$

$$\cdot \frac{\dfrac{N\cdot(N-1)\cdots(N-k+1)}{N^k} \cdot \dfrac{M\cdot(M-1)\cdots(M-n+k+1)}{M^{n-k}}}{\dfrac{(N+M)\cdot(N+M-1)\cdots(N+M-n+1)}{(M+N)^n}}$$

利用Poisson定理再求**例1** (2)

解 令X表示命中次数，则 $X \sim B(5000, 0.001)$

令 $\lambda = np = 5$

$$P(X \geq 2) = 1 - \sum_{k=0}^{1} C_{5000}^{k} (0.001)^k (0.999)^{5000-k}$$

$$\approx 1 - \sum_{k=0}^{1} e^{-5} \frac{5^k}{k!}$$

$$= 1 - e^{-5}\left(\frac{5^0}{0!} + \frac{5!}{1!}\right) = 0.9596$$

此结果也可直接查 P.357 附表1 Poisson 分布表得到，它与用二项分布算得的结果0.9574仅相差千分之二点二.

例2 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$X \sim \mathrm{B}(100+n, 0.03)$

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

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

应用Poisson定理

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

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

在实际计算中，当 $n \geq 20, p \leq 0.05$时，可用上述公式近似计算；而当$n \geq 100, np \leq 10$时, 精度更好

k	按二项分布				按Possion公式
	$n=10$ $p=0.1$	$n=20$ $p=0.05$	$n=40$ $p=0.025$	$n=100$ $p=0.01$	$\lambda=np=1$
0	0.349	0.358	0.369	0.366	0.368
1	0.305	0.377	0.372	0.370	0.368
2	0.194	0.189	0.186	0.185	0.184
3	0.057	0.060	0.060	0.061	0.061
4	0.011	0.013	0.014	0.015	0.015

例3 设有同类型设备90台，每台工作相互独立，每台设备发生故障的概率都是 0.01. 在通常情况下，一台设备发生故障可由一个人独立维修，每人同时也只能维修一台设备.

(1) 问至少要配备多少维修工人，才能保证当设备发生故障时不能及时维修的概率小于0.01?

(2) 问3个人共同负责90台还是3个人各自独立负责30台设备发生故障不能及时维修的概率低？

解 (1) 设 需要配备 N 个维修工人,设 X 为90 台设备中发生故障的台数，则 $X \sim B(90, 0.01)$

$$P(X>N)=\sum_{k=N+1}^{90} C_{90}^{k}(0.01)^{k}(0.99)^{N-k}$$

令 $\lambda=90\times 0.01=0.9$

则 $P(X>N)\approx\sum_{k=N+1}^{90} e^{-0.9}\frac{0.9^{k}}{k!}$

$$=\sum_{k=N+1}^{\infty} e^{-0.9}\frac{0.9^{k}}{k!}-\sum_{k=91}^{\infty} e^{-0.9}\frac{0.9^{k}}{k!}$$

$$\approx\sum_{k=N+1}^{\infty} e^{-0.9}\frac{0.9^{k}}{k!}<0.01$$

查附表2得 $\underline{N=4}$

(2)三个人共同负责90台设备发生故障不能及时维修的概率为

$$P(X>3) \approx \sum_{k=4}^{90} e^{-0.9} \frac{0.9^k}{k!}$$

$$= \sum_{k=4}^{\infty} e^{-0.9} \frac{0.9^k}{k!} - \sum_{k=91}^{\infty} e^{-0.9} \frac{0.9^k}{k!}$$

$$\approx \sum_{k=4}^{\infty} e^{-0.9} \frac{0.9^k}{k!}$$

$$= 0.013459$$

设30台设备中发生故障的台数为 $Y \sim B(30, 0.01)$

设每个人独立负责30台设备，第 i 个人负责的30台设备发生故障不能及时维修为事件 A_i

则 $$P(A_i) = P(Y \geq 2) \approx \sum_{k=2}^{\infty} e^{-0.3} \frac{0.3^k}{k!}$$

$$= 0.0369 \quad i = 1,2,3$$

三个人各独立负责30台设备发生故障不能及时维修为事件 $A_1 \cup A_2 \cup A_3$

$$P(A_1 \cup A_2 \cup A_3) = 1 - \prod_{i=1}^{3} P(\overline{A}_i)$$

$$= 1 - (1 - 0.0369)^3 \approx 0.1067 > 0.013459$$

故 三个人共同负责90 台设备比各自负责好！

在Poisson 定理中，

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

$$\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$$

$$=e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots\right)$$

$$=e^{-\lambda}\cdot e^{\lambda}=1$$

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** $\pi(\lambda)$

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

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

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，在长为 t 的时间内出现的质点数 $X_t \sim \pi(\lambda t)$

例4 设一只昆虫所生虫卵数为随机变量 $X\sim\pi(\lambda)$, 每个虫卵发育成幼虫的概率为 p. 设各个虫卵是否能发育成幼虫是相互独立的. 求一只昆虫所生的虫卵发育成的幼虫数 Y 的概率分布.

解 昆虫 $\longrightarrow$ X 个虫卵 $\longrightarrow$ Y 个幼虫

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

$$P(Y=m|X=k)=C_k^m p^m(1-p)^{k-m},$$

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

$$(Y=m)\subset\bigcup_{k=m}^{\infty}(X=k),\quad m=0,1,2,\cdots$$

$$(X=k)\cap(X=l)=\varnothing,\quad k\neq l$$

由全概率公式

$$P(Y=m)=\sum_{k=m}^{\infty}P(X=k)P(Y=m|X=k)$$

$$=\sum_{k=m}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}C_k^m p^m(1-p)^{k-m}$$

$$=e^{-\lambda}\frac{(\lambda p)^m}{m!}\sum_{k=m}^{\infty}\frac{\lambda^{k-m}}{(k-m)!}(1-p)^{k-m}$$

$$\overset{令k-m=s}{=}\quad e^{-\lambda}\frac{(\lambda p)^m}{m!}\sum_{s=0}^{\infty}\frac{\lambda^s}{s!}(1-p)^s$$

$$=e^{-\lambda}\frac{(\lambda p)^m}{m!}e^{\lambda(1-p)}\quad=e^{-\lambda p}\frac{(\lambda p)^m}{m!}$$

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

故 $Y\sim\pi(\lambda p)$

这只昆虫有后代的概率

$$p\left\{\bigcup_{m=1}^{\infty}(Y=m)\right\}$$

$$=1-p(Y=0)$$

$$=1-e^{-\lambda p}$$