第二节 离散型随机变量及其分布律

一、离散型随机变量及其分布律的概念

定义 1 若随机变量 X 的取值为有限个或可列无限多个，

就称 X 为离散型随机变量.

定义 2 设 X 为离散型随机变量，其所有可能的取值为

$x_1, x_2, \cdots, x_i, \cdots$，且

$$P\{X = x_i\} = p_i, \quad i = 1, 2, \cdots.$$

就称上式为离散型随机变量 X 的分布律或概率分布.

离散型随机变量 X 的分布律或概率分布也记为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

或

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix},$$

其中 $x_1, x_2, \cdots, x_i, \cdots$ 互不相同，且可能为有限个.

性质 1（分布律的性质）设离散型随机变量 X 的分布律为 $X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$，则有

(1) $p_i \geq 0$，$i = 1, 2, \cdots$；　(2) $\sum_i p_i = 1$．

结论 1 设 L 为任意实数集合，则 $P\{X \in L\} = \sum_{x_i \in L} p_i$．

结论 2 X 的分布函数

$$F(x) = P\{X \leq x\} = \sum_{x_i \leq x} p_i, \quad -\infty < x < +\infty .$$

例 1 设盒子中有8个正品和2个次品，现依次不放回地将其逐个取出，记 X 为首次取到正品时的所取产品个数，试求 X 的分布律和分布函数 $F(x)$.

解 由题意知 X 的可能取值为1, 2, 3，由乘法公式和古典概型概率计算公式得 X 的分布律为

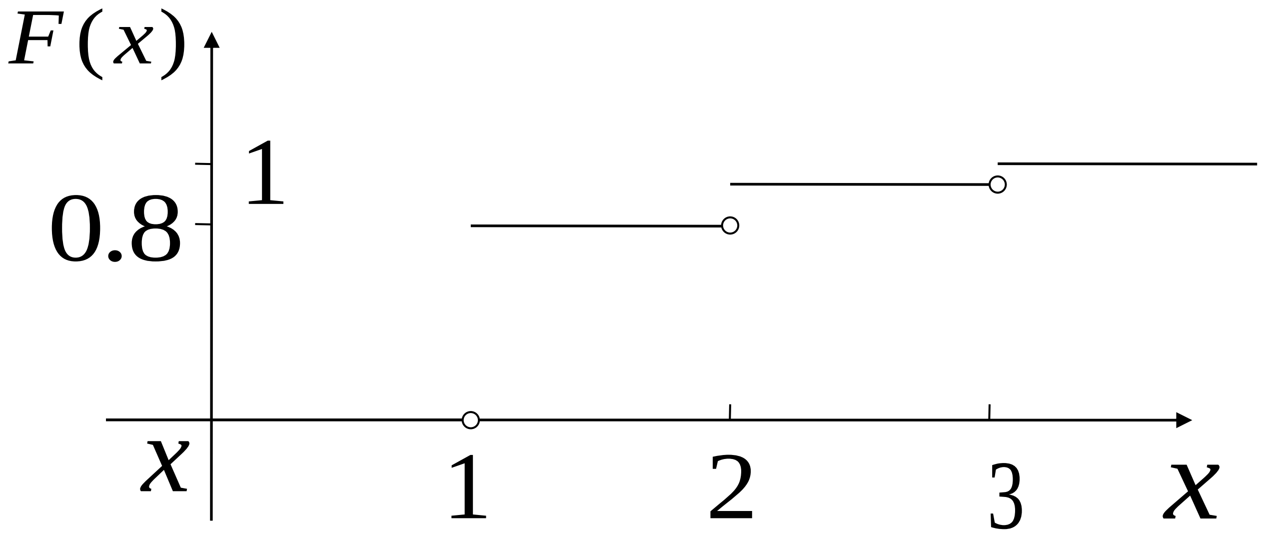
$$P\{X=1\}=\frac{8}{10}=\frac{4}{5}，\quad P\{X=2\}=\frac{2}{10}\times\frac{8}{9}=\frac{8}{45},$$

$$P\{X=3\}=\frac{2}{10}\times\frac{1}{9}\times\frac{8}{8}=\frac{1}{45}\text{ 或}=1-\frac{4}{5}-\frac{8}{45}=\frac{1}{45},$$

即
$$X\sim\begin{pmatrix}1 & 2 & 3\\ \frac{4}{5} & \frac{8}{45} & \frac{1}{45}\end{pmatrix}.$$

X 的分布函数为

$$F(x)=P\{X\le x\}=\begin{cases} 0, & x<1, \\ 0+\dfrac{4}{5}=\dfrac{4}{5}, & 1\le x<2, \\ \dfrac{4}{5}+\dfrac{8}{45}=\dfrac{44}{45}, & 2\le x<3, \\ \dfrac{44}{45}+\dfrac{1}{45}=1, & x\ge 3. \end{cases}$$



注 1 离散型随机变量分布函数的四个特征：

单调上升；右连续；阶梯形；在 $x = x_i$ 处跳跃间断.

例 2 设随机变量 X 的分布律为 $P\{X=i\}=\dfrac{k}{2^i}$，$i=1,2,\cdots$，试求常数 k，以及 X 取奇数的概率.

解 由 $\sum_{i=1}^{\infty}P\{X=i\}=\sum_{i=1}^{\infty}\dfrac{k}{2^i}=k\sum_{i=1}^{\infty}\dfrac{1}{2^i}=k$，以及分布律的性质可得 $k=1$. 由上可知，X 的分布律为

$$P\{X=i\}=\frac{1}{2^i},\quad i=1,2,\cdots,$$

所以 X 取奇数的概率为

$$\sum_{i=1}^{\infty}P\{X=2i-1\}=\sum_{i=1}^{\infty}\frac{1}{2^{2i-1}}=\frac{\frac{1}{2}}{1-(\frac{1}{2})^2}=\frac{2}{3}.$$

二、几种常见的离散型随机变量的概率分布

1. 0-1两点分布

定义 3 如果随机变量 X 的分布律为

X	0	1
p	$1-p$	p

就称 X 服从0－1两点分布，记为 $X \sim B(1,p)$.

一般地，在随机试验 E 中，如果样本空间 Ω 只包含两个样本点 $\Omega=\{\omega_1, \omega_2\}$，则 $X=\begin{cases}0, & \omega=\omega_1\\ 1, & \omega=\omega_2\end{cases}$.

在现实生活中，0–1 两点分布有着广泛的应用. 例如某产品合格与不合格; 某课程的考试及格与不及格; 某事件 A 发生与不发生等许多现象都能够刻划成 0–1 两点分布.

2. 二项分布（贝努里（Bernulli）概型）

贝努里概型

定义 4 将随机试验 E 重复进行 n 次，如果满足每次试验对应的样本空间相同；各次试验结果相互独立；就称之为 n 重独立重复试验.

在 n 重独立重复试验中，如果每次试验仅有两种结果：A 或 $\overline{A}$，就称之为 n 重贝努里试验.

定理 1 在 n 重贝努里试验中，事件 A 恰好发生 k 次的概率为 $C_n^k p^k (1-p)^{n-k}$，其中 $k=0,1,2,\cdots,n$， $p=P(A)$，$0<p<1$.

例 3 在 4 重贝努里试验中，某事件 A 发生的概率小于 $\frac{1}{2}$，且 A 恰好发生 2 次的概率为 $\frac{8}{27}$，求 A 恰好发生 1 次的概率.

解 设事件 A 发生的概率为 p，则 $p<\frac{1}{2}$. 由题意知

$$C_4^2 p^2 (1-p)^2 = \frac{8}{27},$$

解得 $p=\frac{1}{3}$，所以 A 恰好发生 1 次的概率为

$$C_4^1 \times \frac{1}{3} \times \left(1-\frac{1}{3}\right)^3 = \frac{32}{81}.$$

例 4 某人向同一目标独立重复射击，每次射击命中目标的概率为 $p\ (0<p<1)$，求此人第 4 次射击恰好第 2 次命中目标的概率.

解 由于第 4 次射击恰好第 2 次命中目标，可知前 3 次射击中恰有一次命中目标. 由贝努里概型知“前 3 次射击中恰有一次命中目标”的概率为 $C_3^1 p(1-p)^2$.

而事件“第 4 次射击恰好第 2 次命中目标”即为“前 3 次射击中恰有一次命中目标，且第 4 次射击命中目标”，故此人“第 4 次射击恰好第 2 次命中目标”的概率为

$$C_3^1 p(1-p)^2 p = 3p^2(1-p)^2 .$$

二项分布

定义 5 如果随机变量 X 的分布律为

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

就称 X 服从二项分布，记为 $X \sim B(n,p)$，其中 n 为正整数，$0<p<1$.

注 2 又 $C_n^k p^k (1-p)^{n-k}$ 为二项式 $[p+(1-p)]^n$ 的展开式中的各项，因此称 X 服从二项分布.

由贝努里概率模型，在 n 重贝努里试验中，记 X 表示事件 A 发生的次数，则 $X \sim B(n, p)$，其中 $p = P(A)$，因此二项分布也称为**贝努里分布**.

设 $X \sim B(n, p)$，当 $n=1$ 时，可得 X 的分布律为

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

故 $0-1$ 分布为二项分布中 $n=1$ 时的特例.

例 5 设某射手独立地向一目标射击 4 次，每次击中目标的概率为 0.6，求该射手在 4 次射击中，命中目标次数 X 的分布律，并问 X 取何值时的概率最大.

解 将每次射击看成一次随机试验，所需考查的试验结果只有击中目标和没有击中目标，因此整个射击过程为 4 重的贝努里试验．故由题意知，$X \sim B(4, 0.6)$，即

$$P\{X=k\}=C_4^k \times 0.6^k \times 0.4^{4-k}, \quad k=0,1,2,3,4.$$

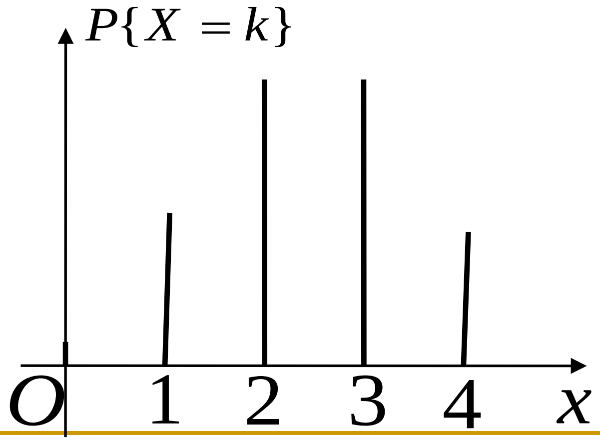
可具体计算得，$P\{X=0\}=C_4^0 \times 0.6^0 \times 0.4^4 = 0.0256$，

$P\{X=1\}=C_4^1 \times 0.6^1 \times 0.4^3 = 0.1536$,

$P\{X=2\}=C_4^2 \times 0.6^2 \times 0.4^2 = 0.3456$,

$P\{X=3\}=C_4^3 \times 0.6^3 \times 0.4^1 = 0.3456$,

$P\{X=4\}=C_4^4 \times 0.6^4 \times 0.4^0 = 0.1296$.



例 6 设某机械产品的次品率为0.005，试分别求任意1000个产品中恰有10个次品的概率和不多于5个次品的概率.

解 设 X 表示1000个产品中次品的个数，则 $X \sim B(1000, 0.005)$.

所以1000个产品中恰有10个次品的概率为

$$P\{X=10\}=C_{1000}^{10}\times 0.005^{10}\times 0.995^{990}\ .$$

1000个产品中不多于5个次品的概率为

$$P\{X\le 5\}=\sum_{k=0}^{5} C_{1000}^{k}\times 0.005^{k}\times 0.995^{1000-k}\ .$$

上面两个计算结果虽然精确，但其计算量都非常大，目前无法求出其值（包括近似值）. 在后续内容中，将陆续介绍泊松定理（第二章）和中心极限定理（第五章），利用这些定理可以近似计算出它们的值.

3. 泊松分布

定义 6 如果随机变量 X 的分布律为

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

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

一般来说，在一定的时间内，“稀有事件”发生的次数 X 服从泊松分布.

譬如，一铸件上的砂眼数，一本书中的错字个数，一平方米内玻璃上的气泡数等等服从泊松分布.

例 7 设随机变量 $X \sim P(\lambda)$，且 $P\{X=1\}=P\{X=2\}$，求 $P\{X=3\}$ 和 $P\{X\geq 1\}$.

解 由 $P\{X=1\}=P\{X=2\}$ 知 $\dfrac{\lambda}{1!}e^{-\lambda}=\dfrac{\lambda^2}{2!}e^{-\lambda}$，解得 $\lambda=2$.

所以 $X\sim P(2)$. 故

$$P\{X=3\}=\frac{2^3}{3!}e^{-2}=\frac{4}{3}e^{-2};$$

$$P\{X\geq 1\}=1-P\{X=0\}=1-\frac{2^0}{0!}e^{-2}=1-e^{-2}.$$

定理 2（泊松定理） 在 n 重贝努里试验中，事件 A 在每次试验中发生的概率为 p_n，其中 $0 < p_n < 1$，且 p_n 与试验次数 n 有关，若 $\lim\limits_{n\to\infty} np_n = \lambda\ (\lambda > 0)$，则对任意非负整数 k，有

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

注 2 定理 2 建立了二项分布与泊松分布的一种联系.

设 $X \sim B(n,p)$，则当 n 充分大，p 很小，而 $np=\lambda$ 较适中时，

有 $X \overset{\text{近似}}{\sim} P(\lambda)$，即 $C_n^k p^k (1-p)^{n-k} \approx \dfrac{\lambda^k}{k!}e^{-\lambda}$，$k = 0,1,2,\cdots,n$.

将二项分布中繁琐的概率计算在一定条件下，转化为泊松分布的概率计算，并通过查阅泊松分布表来实现.

例 8 利用泊松定理近似计算例 6 中的概率 $P\{X=10\}$ 和 $P\{X\le 5\}$.

解 在例 6 中，$X\sim B(1000,0.005)$，其中 $np=5$，由泊松定理知 $X \overset{\text{近似}}{\sim} P(5)$，查泊松分布表得

$$P\{X=10\}=C_{1000}^{10}\times 0.005^{10}\times 0.995^{990}\approx \frac{5^{10}}{10!}e^{-5}$$

$$=\sum_{k=10}^{\infty}\frac{5^k}{k!}e^{-5}-\sum_{k=11}^{\infty}\frac{5^k}{k!}e^{-5}=0.032-0.014=0.018,$$

$$P\{X\le 5\}=\sum_{k=0}^{5}C_{1000}^{k}\times 0.005^{k}\times 0.995^{1000-k}\approx\sum_{k=0}^{5}\frac{5^k}{k!}e^{-5}$$

$$=1-\sum_{k=6}^{\infty}\frac{5^k}{k!}e^{-5}=1-0.384=0.616.$$

4. 几何分布

定义 7 如果随机变量 X 的分布律为

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

其中 $0<p<1$，就称 X 服从参数为 p 的几何分布，

记为 $X\sim G(p)$.

在一系列独立重复试验中，事件 A 首次发生时所进行的试验次数 X 服从几何分布.

例 9 在射击训练中，设某选手每次击中目标的概率为 0.95，击中目标时取得十环的概率为0.4，且射击训练独立重复进行．记 X 为首次取得十环时的射击次数，求 X 的分布律.

解 设事件 A 表示该选手击中目标，B 表示该选手取得十环，则 $P(A)=0.95,\ P(B|A)=0.4$，且 $B\subset A$.

故由乘法公式

$$P(B)=P(AB)=P(A)P(B|A)=0.95\times 0.4=0.38,$$

所以由几何分布的实际应用知，X 的分布律为

$$P\{X=k\}=[P(\overline{B})]^{k-1}P(B)=(1-0.38)^{k-1}\times 0.38$$

$$=0.38\times 0.62^{k-1},\quad k=1,2,3,\cdots.$$