哈爾濱工業大學

第13讲 离散型随机变量

离散型随机变量

- 只能取有限个值或可列无穷多个值的随机变量X 称为离散型随机变量.

- 概率分布列

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

或

X	x_1	x_2	$\ldots$	x_k	$\ldots$
P	p_1	p_2	$\ldots$	p_k	$\ldots$

为离散型随机变量X的概率分布列，简称分布列或分布律.

分布列的性质

(1) $p_k \geq 0, \quad (k = 1, 2, \ldots)$,

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

可以判断数列$\{p_k\}$是否是分布列.

例1 设随机变量X的分布列为

$$P(X=k)=a\frac{\lambda^k}{k!},\quad \lambda>0, k=0,1,2,\ldots$$

求常数a.

解 由分布列的性质 $P(X=k)\geq 0$, 即 $a\geq 0$,

$$\sum_k P(X=k)=1 \Longrightarrow \sum_{k=0}^{\infty} a\frac{\lambda^k}{k!}=a\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=a\mathrm{e}^{\lambda}=1.$$

解得 $a=\mathrm{e}^{-\lambda}$.

由泰勒展示得 $\mathrm{e}^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$

例2 一汽车开往某地需通过三个交通岗, 各个交通岗出现什么信号是相互独立的, 每个交通岗出现红灯和绿灯的概率均为1/2，以X表示汽车首次遇到红灯前已通过的交通岗个数, 求X的分布列.

解 X可取值0, 1, 2, 3.

令 A_i=“第i个路口遇红灯”， $i=1,2,3$.

则A_1, A_2, A_3相互独立，且$P(A_i)=1/2$，$i=1,2,3$.

$$P(\bar{A}_i)=1/2, i=1,2,3.$$

X表示汽车首次遇到红灯前已过的交通岗数

解 令 A_i=“第i个路口遇红灯”， i=1,2,3.

$$P(X=0)=P(A_1)=1/2,$$

$$P(X=1)=P(\bar{A}_1A_2)=1/2\cdot 1/2=1/4.$$

X表示汽车首次遇到红灯前已过的交通岗数

解　令 A_i= “第i个路口遇红灯”，　i=1,2,3.

$$P(X=2)=P(\bar{A}_1\bar{A}_2A_3)=1/2\cdot1/2\cdot1/2=1/8$$

$$P(X=3)=P(\bar{A}_1\bar{A}_2\bar{A}_3)=1/2\cdot1/2\cdot1/2=1/8.$$

X表示汽车首次遇到红灯前已过的交通岗数

- 即X的分布列为

X	0	1	2	3
P	1/2	1/4	1/8	1/8


几个常用的离散型分布

■ 两点分布(伯努利分布、(0–1)分布)

定义 若随机变量X的分布列是

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

或

X	0	1
P	$1-p$	p

$(0<p<1)$.

称X服从两点分布或伯努利分布，也称为$(0-1)$分布，记为$X\sim B(1,\ p)$.

若 $P(X=a)=1$，称X服从退化分布.

二项分布(Binomial)

二项概率公式

定义 若随机变量X的分布列是

$$P(X=k)=C_n^k p^k q^{n-k}$$

$k=0, 1, \ldots, n.\quad 0<p<1,\quad q=1-p.$

称X服从参数为n, p的二项分布, 记为

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

当$n=1$时，$P(X=k)=p^k q^{1-k}$ 为两点分布.

二项分布满足分布列的两个性质．即

$$P(X=k)=C_n^k p^k q^{n-k}\geq 0,\ (k=0,1,\dots n),$$

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

二项分布描述的是n重伯努利试验中“成功”出现次数X的概率分布.

例3 设 $X \sim B(2, p)$, $Y \sim B(3, p)$,

若$P(X \geq 1) = 5/9$, 求$P(Y \geq 1)$.

解 由$P(X \geq 1) = 5/9$, 知

$$P(X < 1) = P(X = 0) = (1 - p)^2 = 4/9,$$

得 $p = 1/3$.

再由 $Y \sim B(3, p) = B(3, 1/3)$，可得

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1 - 1/3)^3$$

$$= 19/27.$$

例4 某电子管使用时数在3000小时以上的概率是0. 1，求三个同种电子管在使用3000小时以后最多只坏一个的概率.

解 设X表示三个电子管在使用3000小时已坏的个数． $X \sim B(3, 0.9)$,

$$P(X=k)=C_3^k(0.9)^k(0.1)^{3-k}, k=0,1,2,3.$$

$$P(X \leq 1)=P(X=0)+P(X=1)$$

$$=(0.1)^3+3(0.9)(0.1)^2=0.028.$$

几个常用的离散型分布

泊松分布(Poisson)

定义 若随机变量X的分布列

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

称X服从参数为λ的泊松分布, 记为$X\sim P(\lambda)$.

可以验证

$$P(X=k)\geq 0,\quad (k=0,\ 1,\ 2,\ \dots\ ),$$

$$\sum_{k=0}^{\infty}P(X=k)=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1.$$

泊松分布的应用

- 交换台在某时间段内接到呼叫的次数;
- 商场在某时间段内接待的顾客数;
- 某时间段内到达飞机场的飞机数;
- 一放射性源放射出的 α 粒子数;
- …

都可以看作泊松分布.

例5 某种铸件的砂眼数服从参数为0.5的泊松分布. 试求该铸件至多有1个砂眼和至少有2个砂眼的概率.

解 用X表示铸件的砂眼数, 由题意知$X \sim P(0.5)$

则该铸件至多有1个砂眼的概率为

$$P(X \le 1) = P(X = 0) + P(X = 1)$$

$$\boxed{P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}}$$

$$= \frac{0.5^0}{0!} e^{-0.5} + \frac{0.5^1}{1!} e^{-0.5} = 0.91.$$

该铸件至少有2个砂眼的概率为

$$P(X \ge 2) = 1 - P(X \le 1) = 0.09.$$

几何分布(Geometric)

定义 若随机变量*X*的分布列

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

称*X*服从参数为*p*的几何分布, 记为$X \sim G(p)$.

几何分布的应用

在伯努利试验中,设每次试验成功的概率均为$p(0<p<1)$, 独立重复试验直到出现首次成功为止, 所需试验次数*X*服从几何分布.

几何分布满足分布列的两个性质

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

$$\sum_{k=1}^{\infty}P(X=k)=\sum_{k=1}^{\infty}q^{k-1}p=p\sum_{k=1}^{\infty}q^{k-1}$$

$$=p\cdot\frac{q^{1-1}}{1-q}=p\cdot\frac{1}{p}=1.$$

几何分布的无记忆性

设$X\sim G(p)$, n, m为任意的两个正整数，则

$$\boldsymbol{P(X > n + m \mid X > n) = P(X > m).}$$

$$\boxed{P(X=k)=q^{k-1}p}$$

证明 $$P(X>n+m\mid X>n)=\frac{P(X>n+m,X>n)}{P(X>n)}$$

$$=\frac{P(X>n+m)}{P(X>n)}=\frac{\sum\limits_{k=n+m+1}^{\infty}(1-p)^{k-1}p}{\sum\limits_{k=n+1}^{\infty}(1-p)^{k-1}p}=\frac{\dfrac{(1-p)^{n+m}}{1-(1-p)}}{\dfrac{(1-p)^{n}}{1-(1-p)}}$$

$$=(1-p)^{m}=\sum_{k=m+1}^{\infty}(1-p)^{k-1}p=P(X>m).$$

超几何分布

定义 设有N件产品，其中有M件次品．今从中任取n件不同产品，则这n件中所含的次品数X的分布列为

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\quad (k=0,1,\dots l=\min(M,n))$$

规定当$i>m$时，$C_m^i=0$.

称X服从超几何分布.

- 二项分布用来描述有放回抽样.
- 超几何分布用来描述不放回抽样.
- 当总体*N*很大，抽样数***n***较小时，

可用二项分布来逼近超几何分布.

定理 设在超几何分布中，n是一个取定的正整数，而 $\lim\limits_{N\to\infty}\dfrac{M}{N}=p,\quad 0<p<1,$

则 $\lim\limits_{N\to\infty}\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k},(k=0,1,\cdots,n).$

对固定$n\geq 1$，当N充分大时，有

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}\approx C_N^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k},(k=0,1,\cdots,\min(M,n)).$$

在实际中，一般当$n\leq 0.1\cdot N$时，可用此近似公式.

谢　　谢！