哈爾濱工業大學

第3章　随机变量及其分布

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

或

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

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 $a\geq 0$,

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

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

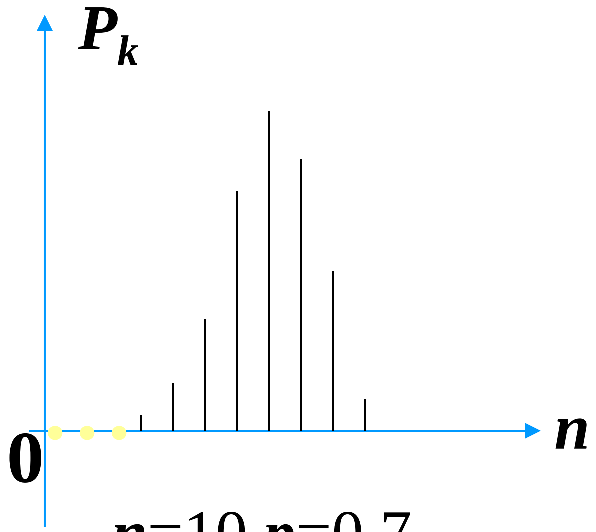
$$P(X=k)=\mathrm{C}_n^k p^k q^{n-k}\geq 0,\ (k=0,1,\ldots n),$$

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

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.

二项分布的图形特点：$X \sim B(n,p)$

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.



n=10,p=0.7

当$(n+1)p$不为整数时，二项概率$P(X=k)$在$k=[(n+1)p]$达到最大值；

($[x]$ 表示不超过 x 的最大整数)

当$(n+1)p$为整数时，二项概率$P(X=k)$在$k=(n+1)p$和$k=(n+1)p-1$处达到最大值.

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 .

则 $X \sim B(3, 0.9)$,

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

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

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

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

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

称X服从参数为λ的泊松分布，记为$X\sim P(\lambda)$.

可以验证

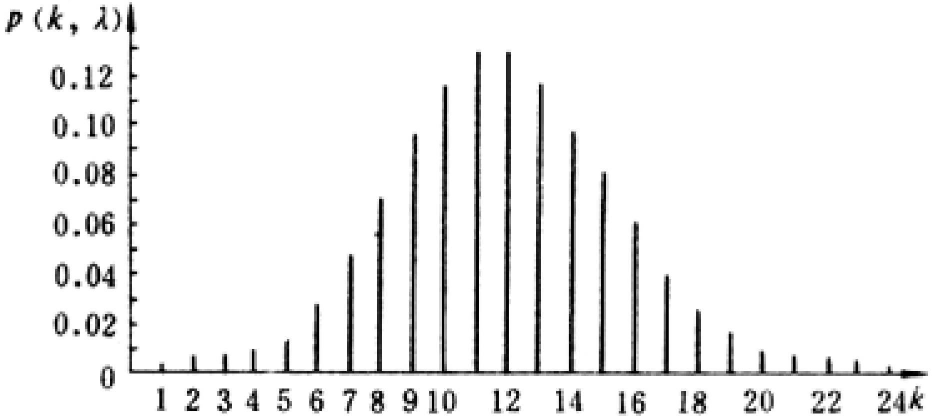
$$P(X=k)\geq 0,\quad (k=0,\ 1,\ 2,\ \ldots\ ),$$

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

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.



泊松分布 $P(\lambda)$ ($\lambda=12$)

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. 我们把在随机时刻相继出现的事件所形成的序列，叫做随机事件流.

若事件流具有平稳性、无后效性、普通性，则称该事件流为泊松事件流（泊松流）.

下面简要解释**平稳性、无后效性、普通性.**

泊松分布产生的一般条件

平稳性

在任意时间区间内，事件发生k次(k≥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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

几何分布的无记忆性

例设X为取正整数值的离散型随机变量，证明：X 是服从参数为p的几何分布的充要条件是X具有无记忆性，即$P(X=k+n|X>k)$与k 无关（k, n为任意正整数） $\boxed{P(X=k)=(1-p)^{k-1}p}$

证明 对任意正整数 k, n有： $P(X=k+n|X>k)$与k 无关，

令$P(X=k+1|X>k)=P(X=1)=p$, $q_k = P(X>k), k=0,1,2,\cdots,\quad q_0=1$

$$p_k = P(X=k) \quad P(X=k+1|X>k) = \frac{p_{k+1}}{q_k} = p$$

$$p_{k+1} = q_k - q_{k+1} \Rightarrow \frac{q_{k+1}}{q_k} = 1-p \Rightarrow \underline{q_{k+1}} = (1-p)q_k = \cdots = (1-p)^{k+1}q_0 = (1\underline{-p)}^{k+1}$$

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

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

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

规定当$i>m$时, $C_m^i=0.$

称X服从超几何分布.

定理 设在超几何分布中，n是一个取定的正整数，而

$$\lim_{N\to\infty}\frac{M}{N}=p,\quad 0<p<1,$$

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

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.

练习题

1 某射手有5发子弹，射击一次命中概率为0.5，如果命中就停止射击，否则一直到子弹用尽，求耗用子弹数的分布列

X	1	2	3	4	5
p	0.5	0.5^2	0.5^3	0.5^4	$0.5^5\times2$

练习题

2 甲、乙两人进行乒乓球比赛，在每一局比赛中，甲获胜的概率为1/3，当采用5局3胜的比赛规则时，求比赛局数的分布列

X	3	4	5
P	9/27	10/27	8/27

$$P(X=3)=\left(\frac{1}{3}\right)^3+\left(\frac{2}{3}\right)^3=\frac{9}{27}$$

$$P(X=4)=C_3^2\left(\frac{1}{3}\right)^2\frac{2}{3}\frac{1}{3}+C_3^2\left(\frac{2}{3}\right)^2\frac{1}{3}\frac{2}{3}=\frac{10}{27}$$

$$P(X=5)=C_4^2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2\frac{1}{3}+C_4^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^2\frac{2}{3}=\frac{8}{27}$$

练习题

3 一批产品共10件，其中6件正品，4件次品，每次从中任取一件，在下述三种情况下，分别求直至取到正品时所需次数X的概率分布列：

（1）每次取出的产品不放回；

（2）每次取出的产品放回；

（3）每次取出一件次品后，放回一件正品到这批产品中.

解（1）X的取值：1，2，3，4，5

X	1	2	3	4	5
p	0.6	4/15	1/10	1/35	1/210

练习题

3（2）几何分布

$$P(X=k)=q^{k-1}p=(0.4)^{k-1}(0.6),\ \ k=1,2\cdots$$

（3）$P(X=1)=\frac{6}{10}\qquad P(X=2)=\frac{4}{10}\cdot\frac{7}{10}=\frac{28}{100}$

$$P(X=3)=\frac{4}{10}\cdot\frac{3}{10}\cdot\frac{8}{10}=\frac{96}{1000}\quad P(X=4)=\frac{4}{10}\cdot\frac{3}{10}\cdot\frac{2}{10}\cdot\frac{9}{10}=\frac{216}{10000}$$

$$P(X=5)=\frac{4}{10}\cdot\frac{3}{10}\cdot\frac{2}{10}\cdot\frac{1}{10}=\frac{24}{10000}$$

第2章 小测验

在一天中学习哈工大概率论与数理统计MOOC的人数服从参数为λ的泊松分布，而学习的人中，每人提问的概率为p, 每人至多提问1次，且每人是否提问相互独立，求一天中恰有m个人提问的概率？

解

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