第四周　常见随机变量

这一周我们介绍几种常见的随机变量。我们希望能够从各种随机变量产生的机理角度进行说明，从而使它们的性质展开更加自然，同时也能更深入地理解它们之所以常见的内在原因。本周学习的分布包括：二项分布，负二项分布，泊松分布，几何分布，指数分布，正态分布。

****

4.1 二项分布与负二项分布

伯努利（Bernoulli）试验

一个随机试验只有“成功”和“失败”两种可能的结果，其中出现“成功”的概率为 $p\left(0<p<1\right)$，则称此随机试验为一个参数为 p 的伯努利试验。

由参数为 p 的伯努利试验定义一个随机变量 X，

$$X=\begin{cases}1, & \text{伯努利试验成功}\\ 0, & \text{否则}\end{cases}$$

则称 X 是参数为 p 的伯努利随机变量，或称 X 服从参数为 p 的伯努利分布。

****

例 4.1.1　抛一颗均匀色子，如果出现偶数点称为试验“成功”，出现奇数点为试验“失败”，则随机变量

$$X=\begin{cases}1, & \text{抛出的点数为偶数,}\\ 0, & \text{抛出的点数为奇数.}\end{cases}$$

是一个参数为 $p=\dfrac{1}{2}$ 的伯努利随机变量。

****

二项分布

将参数为 p 的伯努利试验独立地重复 n 次，定义随机变量 X 为试验成功的次数，则 X 的

分布律为：

$$\begin{pmatrix} 0 & 1 & 2 & \cdots & k & \cdots & n \\ p_0 & p_1 & p_2 & \cdots & p_k & \cdots & p_n \end{pmatrix},$$

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

此分布即称为二项分布，记为 $X \sim B(n,p)$，也称 X 服从参数为 (n,p) 的二项分布。

利用二项式定理可验证： $\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = \left[p+(1-p)\right]^n = 1$ ，

****

例 4.1.2 甲、乙两棋手约定进行 10 局比赛，每局棋甲获胜的概率是 0.6，乙获胜的概率为 0.4。如果各局比赛独立进行，试问甲获胜、战平和失败的概率？

X 表示甲获胜的局数，则 $X \sim b(10,0.6)$

$$P(\text{甲胜}) = P(X>5) = \sum_{k=6}^{10} C_{10}^{k} 0.6^k 0.4^{10-k} = 0.6330,$$

$$P(\text{乙胜}) = P(X<5) = \sum_{k=0}^{4} C_{10}^{k} 0.6^k 0.4^{10-k} = 0.1663,$$

$$P(\text{战平}) = P(X=5) = C_{10}^{5} 0.6^5 0.4^5 = 0.2007 \text{。}$$

****

例 4.1.3 一个通讯系统由 n 个部件组成，每个部件独立工作且能正常运行的概率均为 p，如果构成系统的部件中至少有一半以上能正常运行，则称系统是“有效”的。试问当 p 取何值时，由 5 个部件组成的系统要比由 3 个部件组成的系统更有效？

解　设 n 个部件能正常运行的数目为随机变量 X_n，则 $X_n \sim B(n,p)$

由 5 个部件组成的系统是“有效”的概率为： $P(X_5>2)$

$$P(X_5>2) = P(X_5=3)+P(X_5=4)+P(X_5=5) = C_5^3 p^3 (1-p)^2 + C_5^4 p^4 (1-p) + C_5^5 p^5$$

由 3 个部件组成的系统是“有效”的概率为： $P(X_3>1)$

$$P\left(X_3>1\right)=P\left(X_3=2\right)+P\left(X_3=3\right)=C_3^2p^2(1-p)+C_3^3p^3$$

当 $C_5^3p^3(1-p)^2+C_5^4p^4(1-p)+C_5^5p^5>C_3^2p^2(1-p)+C_3^3p^3$ 时，

由 5 个部件组成的系统要比由 3 个部件组成的系统更有效。

整理后得，p 应满足 $3(1-p)^2(2p-1)>0$，

即当 $p>\frac{1}{2}$ 时，5 个部件组成的系统要更有效些。

****

例 4.1.4 某人将一枚均匀的硬币随机抛了 10 次，已知有 6 次抛出正面，问他是在前 6 次抛出正面的概率？

解 设随机变量 X 为 10 次抛掷中正面出现的次数，则 $X\sim B\left(10,\frac{1}{2}\right)$。

记事件 A 为“前 6 次抛出正面且后 4 次抛出反面”，由题意，要求的概率为

$$P(A\mid X=6)=\frac{P(A\cap\{X=6\})}{P(X=6)}=\frac{P(A)}{P(X=6)}=\frac{(\frac{1}{2})^{10}}{C_{10}^6(\frac{1}{2})^6(\frac{1}{2})^4}=\frac{1}{C_{10}^6}\approx 0.0048$$

****

负二项分布

连续不断且独立地重复进行一个参数为 p 的伯努利试验，记 X 为第 r 次“成功”出现时所需的试验次数，则事件 $\{X=k\}$ 等价于{第 k 次试验“成功”且前 $k-1$ 次试验中恰好“成功” $r-1$ 次}，故由试验的独立性以及二项分布的性质可得

$$P(X=k)=p\cdot P\left(B\left(k-1,p\right)=r-1\right)=pC_{k-1}^{r-1}p^{r-1}q^{k-1-(r-1)}=C_{k-1}^{r-1}p^rq^{k-r},$$

其中 $q=1-p$，$k=r,r+1,\cdots$.

若随机变量 X 的分布律为

$$P(X=k)=C_{k-1}^{r-1}p^rq^{k-r},\quad q=1-p,\ k=r,r+1,\cdots.$$

则称 X 服从参数为 r,p 的负二项分布，记为 $X\sim NB(r,p)$.

****

例4.1.5　甲、乙两人进行比赛，直到某一人先赢到5局为止，假设每局比赛独立，且每局甲胜的概率为0.58，乙胜的概率为0.42。求

（1） 比赛在第7局结束的概率？ （2） 在第7局结束的条件下，获胜方为甲的概率？

解　设X为甲赢5局时所需的比赛局数，Y为乙赢5局时所需的比赛局数，则

$$X \sim NB(5, 0.58);\quad Y \sim NB(5, 0.42)$$

设事件A为{甲最终获的比赛胜利} ，事件B为{比赛在第7局结束}，则

（1） $P(B) = P(X=7) + P(Y=7)$

$$= C_6^4(0.58)^5(0.42)^2 + C_6^4(0.42)^5(0.58)^2 = 0.17 + 0.066 = 0.24,$$

（2） $P(A \mid B) = \dfrac{P(AB)}{P(B)} = \dfrac{P(X=7)}{P(X=7)+P(Y=7)} = \dfrac{0.17}{0.24} = 0.71$。

****

负二项的名称来自广义的牛顿二项式公式

Newton 二项式公式 （$\forall a \in R$，$|t| < 1$）

$$\left(1+t\right)^a = \sum_{k=0}^{\infty} C_a^k t^k = 1 + C_a^1 t + C_a^2 t^2 + \cdots,$$

C_a^k是整数组合数的推广，$C_a^k = \dfrac{a\cdot(a-1)\cdots(a-k+1)}{k!}$。

整理负二项分布随机变量$X \sim Nb(r,p)$的分布律

$$P\left(X=r+k\right) = C_{r+k-1}^{r-1}\left(1-p\right)^k p^r = C_{r+k-1}^{r-1} q^k p^r = C_{-r}^k p^r\left(-q\right)^k \quad q = 1-p,\ k = 0,1,2,\cdots$$

$$\sum_{k=0}^{\infty} P\left(X=r+k\right) = p^r \sum_{k=0}^{\infty} C_{-r}^k\left(-q\right)^k = p^r\left(1-q\right)^{-r} = 1$$

****