概率论

随机事件与概率

概率的性质

概率的单调性

若$A \supset B$，则$P(A-B)=P(A)-P(B)$，$P(A)>P(B)$

对于任意两个事件A、B，有$P(A-B)=P(A)-P(AB)$

加法公式

$$P(\bigcup_{i=1}^{n} A_i)=\sum_{i=1}^{n} P(A_i)-\sum_{1\le i\le j\le n} P(A_iA_j)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

$$P(\bigcup_{i=1}^{n} A_i)\le \sum_{i=1}^{n} P(A_i)$$

条件概率

$$P(A|B)=\frac{P(AB)}{P(B)}$$

乘法公式

$$P(\prod_{i=n}^{i} A_i)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1\cdots A_{n-1})$$

全概率公式

若$B_1,\cdots,B_n$为样本空间Ω的一个分割，则

$$P(A)=\sum_{i=1}^{n} P(B_i)P(A|B_i)$$

贝叶斯公式

若$B_1,\cdots,B_n$为样本空间Ω的一个分割，则

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n}}P(B_j)P(A|B_j)$$

独立性

设A、B、C为三个事件，如果有：

$$\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \end{cases}$$

则称**A、B、C两两独立**

如果**还有**

$$P(ABC)=P(A)P(B)P(C)$$

则称**A、B、C相互独立**

n个相互独立事件内的任意一部分事件内都是独立的，而且任意一部分事件与另一部分事件也是独立的

将相互独立事件的任意事件换为对立事件，所得事件仍是独立的

如果有A、B、C两两独立，则不能证明AB与C独立，也不能证明$A-B$与C独立，也不能证明$A\cup B$与C独立

随机变量及其分布

数学期望与方差

数学期望

$$E[g(X)]=\begin{cases}\sum g(x_i)p(x_i)\\ \int g(x)p(x)dx\end{cases}$$

若C为常数，$E(C)=C$

对于任意常数a，$E(aX)=aE(X)$

$$E[g_1(X)\pm g_2(X)]=E[g_1(X)]\pm E[g_2(X)]$$

方差与标准差

$$Var(X)=E(X-E(X))^2=\begin{cases}\sum(x_i-E(X))^2p(x_i)\\ \int(x-E(X))^2p(x)dx\end{cases}$$

$Var(X)=E(X^2)-[E(X)]^2$

如果C为常数，则$Var(C)=0$

$Var(aX+b)=a^2Var(X)+b$

切比雪夫不等式

$$P(|X-E(X)|\geq\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}$$

常用离散分布

二项分布

X为n重伯努利实验中成功的次数，p为每次实验成功的概率，则X的分布称为**二项分布**，$X\sim b(n,p)$

特别地，当n=1时，称为**贝努利分布**，或者**0-1分布**，或者**二点分布**

$$P(X=k)=C_n^kp^k(1-p)^{n-k}$$

$$E(X)=np$$

$$var(X)=np(1-p)$$

泊松分布

泊松分布，记为$X \sim P(\lambda)$

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

$$E(X)=\lambda$$

$$Var(X)=\lambda$$

二项分布的泊松近似

如果当$n \to \infty$，有$np_n \to \lambda$，则

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

当**n很大，而p很小**的时候，可以用泊松近似

超几何分布

有N件产品，其中M件不合格品，如果从中抽取n件，则不合格品的个数X服从**超几何分布**，记为$X \sim h(n,N,M)$

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

$$E(X)=n\frac{M}{N}$$

超几何分布的二项近似

当$n << N$时，可以近似地把超几何分布看做二项分布：

$$P(X=k)\approx C_n^k p^k (1-p)^{n-k}, p=\frac{M}{N}$$

几何分布

在伯努利实验中，每次实验成功的概率为p，记X为首次成功的次数，则称X服从**几何分布**，记作$X \sim Ge(p)$

$$p(X=k)=(1-p)^{k-1}p$$

$$E(X)=\frac{1}{p}$$

$$Var(X)=\frac{1-p}{p^2}$$

几何分布的无记忆性

若$X \sim Ge(p)$，则

$$P(X > m+n|X > m) = P(X > n)$$

负二项分布

如果在伯努利实验中，实验成功概率为p，如果X为实验第r次成功时的实验次数，则称X服从**负二项分布**或者**帕斯卡分布**，记作$X \sim Bb(r,p)$

$$P(X = k) = C_{k-1}^{r-1}p^r(1-p)^{k-r}$$

$$E(X) = \frac{r}{p}$$

$$Var(X) = r\frac{1-p}{p}$$

负二项分布可以看出是r个独立同分布的几何分布之和

常用连续分布

正态分布

X服从**正态分布**，可以记作$X \sim N(\mu, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

当$\mu = 0, \sigma = 1$时，称为**标准正态分布**，通常记作U

$$p(u) = \varphi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$$

$$F(u) = \Phi(u)$$

$$E(X) = \mu, Var(X) = \sigma^2$$

均匀分布

X在区间（a,b）上的**均匀分布**记作$X \sim U(a,b)$

$$F(x) = \frac{x-a}{b-a}, a \le x \le b$$

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

指数分布

X服从**指数分布**，记作$X \sim Exp(\lambda)$

$$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$$

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}$$

$$Var(X)=\frac{1}{\lambda^2}$$

$$E(X)=\frac{1}{\lambda}$$

指数分布的无记忆性

若$X \sim Exp(\lambda)$，则

$$P(x>s+t|x>s)=P(x>t)$$

随机变量函数的分布

设X是连续随机变量，$Y=g(X)$是另一个随机变量，若$Y=g(x)$严格单调，且其反函数$h(y)$有连续导数，则Y的概率密度函数为

$$p_Y(y)=\begin{cases}P_x[h(y)]|h'(y)|, & y\text{概率密度不为零时}\\ 0, & \text{其他}<0\end{cases}$$

多维变量及其分布

常用多维分布

多项分布

进行n次独立实验，每次实验有r个互不相容的结果$A_1,\cdots,A_r$，每个结果发生的概率分别为P_i，记X_i为n次独立重复实验中A_i出现的次数，则X符合r**项分布**，又称为**多项分布**，记为$M(n,p_1,\cdots,p_r)$

$$P(X_1=n_1,\ldots,X_r=n_r)=\frac{n!}{n_1!\cdots n_r!}p_1^n\cdots p_r^n$$

多维超几何分布

袋子中有N个球，共有r号球，其中i号球有N_i个，从中任一取出n个，记X_i为取出i号球的个数，则称为**多维超几何分布**

$$P(X_1=n_1,\cdots,X_r=n_r)=\frac{C_{N_1}^{n_1}\cdots C_{N_r}^{n_r}}{C_N^n}$$

多维均匀分布

若D为$\mathbb{R}^n$中的一个有界区域，如果X满足**多维均匀分布**，则:

$$p(x_1,\cdots,x_n)=\begin{cases}\dfrac{1}{S_D}, & (x_1,\cdots,x_n)\in D\\ 0, & \text{其他}\end{cases}$$

二元正态分布

如果(X,Y)服从**二元正态分布**，则记为$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，其中$\rho$为**相关系数**

在二维正态分布中，不相关与独立是等价的

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]}$$

边际分布与随机变量的独立性

边际分布

由(X,Y)的分布函数，可以求得X的分布函数，称为**X的边际分布**

$$F_X(x)=F(x,\infty)$$

边际分布列与边际密度函数

$$p(X=x_i)=\sum_{j=1}^{\infty}P(X=x_i,Y=y_j)$$

$$P_X(x)=\int_{-\infty}^{\infty}p(x,y)dy$$

具有相同边际分布的多维联合分布可以是不同的

随机变量的独立性

$X_1,...,X_n$**相互独立**等价于：

$$F(x_1,\cdots,x_n)=\prod F_i(x_i)$$
$$p(x_1,\cdots,x_n)=\prod p_i(x_i)$$

X与Y相互独立，或者说XY可分离，有两重含义：

- $p(x,y)=p_X(x)p_Y(y)$
- $p(x,y)$的非零区域可以分解为两个一维区域的乘积空间

多维随机变量的分布

可加性

同一类分布的独立随机变量的和仍属于该类分布称为**可加性**

泊松分布的可加性

如果$X\sim P(\lambda_1),Y\sim P(\lambda_2)$，且XY相互独立，则$Z=X+Y\sim P(\lambda_1+\lambda_2)$

注意：X-Y不服从泊松分布

二项分布的可加性

如果$X \sim b(n,p), Y \sim b(m,p)$，且XY相互独立，则$Z = X+Y \sim b(m+n,p)$

正态分布的可加性

如果$X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$，且XY相互独立，则$Z = X+Y \sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

最大值与最小值分布

最大值分布

如果$X_1,\cdots,X_n$是相互独立的n个随机变量，其分布分别为$F_i(x)$，$Y = \max\{X_1,\cdots,X_n\}$，则

$$\begin{aligned} F_Y(y) &= P(\max\{X_1,\cdots,X_n\} \le y) \\ &= P(X_1 \le y,\cdots,X_n \le y) \\ &= P(X_1 \le y)\cdots P(X_n \le y) \\ &= \prod F_i(y) \end{aligned}$$

最小值分布

如果$X_1,\cdots,X_n$是相互独立的n个随机变量，其分布分别为$F_i(x)$，$Y = \min\{X_1,\cdots,X_n\}$，**则**

$$\begin{aligned} F_Y(y) &= P(\min\{X_1,\cdots,X_n\} \le y) \\ &= 1-P(\min\{X_1,\cdots,X_n\} > y) \\ &= 1-P(X_1 > y,\cdots,X_n > y) \\ &= 1-P(X_1 > y)\cdots P(X_n > y) \\ &= 1-\prod[1-F_i(y)] \end{aligned}$$

如果X_i均服从参数为λ的指数分布，则$\min X_i$仍服从指数分布，参数为$n\lambda$

多维随机变量的特征数

数学期望与方差

- $E(X+Y) = E(X)+E(Y)$
- 如果X与Y相互独立，$E(XY) = E(X)E(Y)$
- 如果X与Y相互独立，$Var(X \pm Y) = Var(X)+Var(Y)$

协方差

$$Cov(X,Y) = E[(X-E(X))(Y-E(Y))]$$

特别地，Cov(X, X) = Var(X)

- 当$Cov(X,Y) > 0$，称X与Y**正相关**
- 当$Cov(X,Y) < 0$，称X与Y**负相关**
- 当$Cov(X,Y) = 0$，称X与Y**无关**

1. $Cov(X,Y) = E(XY)-E(X)E(Y)$
2. 如果随机变量X与Y相互独立，则$Cov(X,Y) = 0$，**反之不然**
3. 对于任意随机变量X，Y，$Var(X \pm Y) = Var(X)+Var(Y) \pm Cov(X,Y)$
4. $Cov(X,Y) = Cov(Y,X)$
5. 如果C为常数，则$Cov(X,C) = 0$
6. $Cov(aX,bY) = abCov(X,y)$

7. $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$

相关系数

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}$$

相关系数与协方差同号，它是**标准化后的协方差**

$$Corr(X,Y)=Cov(\frac{X-\mu_X}{\sigma_X},\frac{Y-\mu_Y}{\sigma_Y})$$

- 施瓦茨不等式：

$$[Cov(X,Y)]^2\le\sigma_X^2\sigma_Y^2$$

- $|Corr(X,Y)|\le 1$
- $Corr(X,Y)=\pm1$的充要条件为 **几乎处处** 有线性关系，即存在a,b使得$P(Y=aX+b)=1$，当$a>0$时，$Corr(X,Y)=1$，当$a<0$时，$Corr(X,Y)=-1$

在二维正态分布中， 不相关与独立是等价的

条件分布与条件期望

条件分布

$$p(y|x)=\frac{p(x,y)}{p_X(x)}$$

贝叶斯公式和全概率公式

全概率公式：

$$p_Y(y)=\int_{-\infty}^{\infty}p_X(x)p(y|x)dx$$

贝叶斯公式：

$$p(y|x)=\frac{p_Y(y)p(x|y)}{\int_{-\infty}^{\infty}p_Y(y)p(x|y)dy}$$

条件数学期望

$$E(X)=E(E(X|Y))$$

大数定理与中心极限定理