概率论与数理统计（待补充）

一、概率论基础知识

随机试验

1. 可在相同条件下重复进行
2. 试验的所有可能结果不止一个，且试验前知道一切可能结果
3. 试验前不知哪一个可能结果出现，试验后能客观确定出现的是哪一个结果

样本空间Ω表示一个试验的所有可能结果的集合，一个可能的结果称为**样本点**，记为ω

样本空间的子集称为一个随机事件，简称**事件**

称Ω为**必然事件**，空集为**不可能时间，**只含有一个样本点的事件称为**基本事件**

事件关系及运算

> 1. 事件的包含与相等
> 2. 事件的和（并）
> 3. 事件的积（交）
> 4. 事件的差 A-B
> 5. 互斥事件　　AB不可能同时发生
> 6. 互逆事件，对立事件 AB不可能同时发生，但一定会发生一个
> 7. 完备事件组 样本空间的有限划分

定义1.1在n次重复试验中若事件A发生了k次，则称k为事件A发生的频数，称$\frac{k}{n}$为事件A发生的频率记为

$$f_n(A)=\frac{k}{n} \qquad 0<=f_n(A)<=1 \qquad f_n(\Omega)=1, f_n(\Phi)=0 \qquad f_n(\bigcup_{i=1}^{r})=\sum_{i=1}^{r} f_n(A_i)$$

P为**概率**，大于零，和为1，互斥事件可加，**空集概率为0但概率为0不一定是空集**，$P(\overline{A})=1-P(A)$，$P(A-B)=P(A)-P(AB)$，$P(A\bigcup B)=P(A)+P(B)-P(AB)$

古典概型所有结果概率相等

超几何概率，m红球剩余白球$p_k=\frac{C_m^k C_{N-m}^{n-k}}{C_N^n}, k=0,1,\cdots,m$

几何概率，长度、面积、体积的比

条件概率，
$P(A|B)=\frac{P(AB)}{P(B)}$，$P(\overline{A}|B)=1-P(A|B)$，$P(A\bigcup B|C)=P(A|C)+P(B|C)-P(AB|C)$

乘法公式

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n->1})$$

全概率公式

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$$

贝叶斯公式

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{P(B)}=\frac{P(A_i)P(B|A_i)}{>}\sum_{j=1}^{n}P(A_j)P(B|A_j),\quad i=1,2,\cdots,n$$

相互独立

- 若$P(AB)=P(A)P(B)$称事件A与B相互独立
- 若事件A与B相互独立，则A与$\overline{B}$，$\overline{A}$与B，$\overline{A}$与$\overline{B}$也相互独立
- 若n个事件$P(A_iA_j)=P(A_i)P(A_j)$，则称这n个事件**两两独立**
- 若n个事件$P(A_{i1}A_{i2}\cdots A_{ik})=P(A_{i1})P(A_{i2})\cdots P(A_{ik})$，则称这n个事件相互独立
- 串联电路可靠度$P(A_1)P(A_2)\cdots$并联电路可靠度$1-P(\overline{A}_1)P(\overline{A}_2)\cdots$

伯努利概型

特别，n重独立试验中每次试验结果只有两个:“成功”与“失败”即A与$\overline{A}$且$0<P(A)<1$这样的试验称为n重**伯努利试验**，相应的数学模型叫做**伯努利概型**

二项概率

$$\sum_{k=0}^{n}P_n(k)=\sum_{k=0}^{n}C_n^kp^k(1-p)^{n-k}=1$$

多项概率公式

$$\frac{n!}{r_1!r_2!\cdots r_k!}p_1^{r_1}p_2^{r_2}\cdots p_k^{r_k},\quad >r_1+r_2+\cdots+r_k=n$$

二、随机变量及其分布

随机变量

> 设Ω为一试验的样本空间如果对每一个样本点$\omega \in \Omega$规定一个实数 $X(\omega)$，这样就定义了一个定义域为Ω的实值函数$X = X(\omega)$，称X为随机变量.

随机变量分布函数

> - 右连续性
> - 归一性
> - 负无穷是0正无穷是1
> - 单调不减
>
> $$F(x) = P(X \leqslant x), \quad x \in R$$

离散型随机变量

> 设离散型随机变量X的取值为$x_1, x_2, \cdots, x_n, \cdots$，且$X$取各值的概率为
>
> $$p_k = P(X = x_k), \quad k = 1, 2, \cdots$$
>
> 称X为概率分布，或概率函数，也可称为分布律

> $$X \backsim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

> 若随机变量两两互斥，则有
>
> $$F(x) = P(X \leqslant x) = P(\bigcup_{x_k \leqslant x} (X = x_k)) = \sum_{x_k \leqslant x} p_k$$

常见离散分布

> - 几何分布$X \backsim G(p)$，$p_k = p(X = k) = pq^{k-1}, \quad 1, 2, \cdots$
> - 超几何分布$X \backsim H(n, m, N)$，$p_k = P(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n}, \quad k = 0, 1, \cdots, n$
> - 二项分布$X \backsim B(n, p)$，$p_k = P(X = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \cdots, n$
> - 泊松分布$X \backsim P(\lambda)$，$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \cdots$

连续性随机变量及其分布

$F(x)=\int_{-\infty}^{x} f(t)dt$为$X$为**连续型随机变量**，$f(x)$称为$X$的**概率密度函数**，简称**密度函数**或**密度**

均匀分布$X \backsim U(a,b)$

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a \leqslant x \leqslant b,\\ 0, & \text{其他},\end{cases}$$

指数分布$X \backsim e(\lambda)$

Γ分布$X \backsim \Gamma(\alpha,\beta)$

$$f(x)=\begin{cases}\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}, & x>0,\\ 0, & x \leqslant 0,\end{cases}$$

$$\begin{aligned}&(1)\ \Gamma(\alpha+1)=\alpha\Gamma(\alpha),\ \text{当}\ \alpha>0;\\ &(2)\ \Gamma(1)=1,\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi};\end{aligned}$$