随机变量及其概率分布

> **随机变量：** 设E是随机试验，样本空间Ω，如果对于**每一个结果 w** 属于Ω，**都有一个实数X(w)与之对应**，这样就得到一个定义在Ω上的实值函数 X = X(w),**这个实数X称为随机变量。***随机变量由随机试验结果决定*
>
> eg：{w | X(w) = a} 表示事件，w样本点，a表示实验结果，记为：{X=a} 或P(X=a) 表示事件a
>
> P{$1 \leq X \leq 3$} 表示事件在1和3之间发生的概率
>
> ---
>
> **离散型随机变量：** 若随机变量X**只取有限个**或**可列无限多个值，则称X为离散型随机变量**
>
> **分布率：** 所有随机变量与及其对应的概率关系，**其中每个变量概率大于0，所有变量概率总和为1**

连续型随机变量及其概率密度函数

> 存在非负可积函数y=f(x)，其中x表示$a < X \leq b$之间事件，y表示(a,b]之间概率，则f(x)为事件X的概率密度函数
>
> $$P\{a < X \leq b\} = \int_a^b f(x)dx \quad X\text{连续随机变量，}f(x)\text{概率密度函数,}$$
>
> $$\int_a^b f(x)dx \text{ 为事件}P(a < X \leq b)\text{的概率，其表达式为}X - f(x)$$
>
> **密度函数性质**
>
> - f(x)≥ 0
> - $\int_{-\infty}^{+\infty} f(x) = 1$
> - **连续性概率密度函数f(x)，随机变量x表示具体值时，概率为0**
> - P{$a \leq X \leq b$} = P{$a < X \leq b$} = P{$a \leq X < b$}= P{$a < X < b$}
> - 概率为0的事件不一定是不可能事件，概率为1的事件不一定是必然事件

分布函数

> X取值不超过x的概率： **F(x) = P(X$\leq$x)** F(x) 为分布函数 F(x) $\in [0, 1]$
>
> **分布函数性质**
>
> - F(x) $\in [0, 1]$，x$\in (-\infty, +\infty)$
> - $x_1 < x_2$，$F(x_1) \leq F(x_2)$，其中x_2事件范围包含x_1
> - $lim_{x \to +\infty} F(+\infty) = 1$，$lim_{x \to -\infty} F(-\infty) = 0$
> - F(x)右连续，则离散型概率函数右连续，连续性概率函数也连续
> - P{$X \leq a$ } = F(a)
> - P{$X > a$ } =1 - P{x$\leq$ a} =1- F(a)
> - P{ a<X$\leq$b } = P{X$\leq$b} - P{X$\leq$a} = F(b) - F(a)
> - P{X=a} = F(a) - F(a-0)
> - P{a$\leq$ $X \leq b$} = F(b) - F(a-0)
> - P{X < a} = F(a-0)
> - P{x $\geq$ a} = 1 - F(a-0)
>
> **连续性分布函数性质** (连续概率函数与分布函数转换)
>
> - F(x) = P{X $\leq$ x} = $\int_{-\infty}^{x} f(t)dt \quad F'(x) = f(x)$

> - P{a<X<b} = F(b) - F(a)
>
> *分布函数转概率函数 与 概率转分布式函数*

常见随机变量概率函数与分布函数

> - **0-1分布：** $P\{X=k\}=p^k(1-p)^{1-k},k=0,1$ *离散型概率分布*
>
> > 只先进行一次事件试验，该事件发生的概率为p，不发生的概率为1-p，*只有两种实验结果*
> >
> > | X | 0 | 1 |
> > |---|---|---|
> > | p_k | $1-p$ | p |
>
> - **几何分布：** P{X=k} = $(1-p)^{k-1}p$，记为 X~G(p)，*离散型概率分布*
>
> > 进行多次实验，**第k次首次发生**概率
>
> - **二项分布：** $P\{X=k\}=C_n^kp^k(1-p)^{n-k}$ 记为 X~B(n,p) *离散型概率分布*
>
> > n次实验中，事件A发生了k次概率
> >
> > *二项分布最值*
> >
> > 1. (n+1)p 不为整数时，X=(n+1)p 达到最大值（最可能发生事件）
> > 2. (n+1)p 为整数，X=(n+1)p，与X=(n+1)p-1 同时为最大值（最可能发生事件）
>
> - **泊松分布：** P{X = k} = $\frac{\lambda^k}{k!}e^{-\lambda}$ $\lambda>0,\text{其中}k=0,1,2,3...$ 记为 X~P(λ) *离散*概率分布
>
> > *泊松分布的参数λ是单位时间(或单位面积)内**随机事件的平均发生次数。***
> >
> > **二项分布用泊松分布近似条件**
> >
> > - $n\geq 100$
> > - $np\leq 10$
> > - $C_n^kp^kq^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda},\quad \lambda=np，q=1-p$
>
> - **超几何分布：** P{X=k} = $\frac{C_{N_1}^KC_{N_2}^{m-k}}{C_N^n}$ *离散型概率分布*
>
> > N个元素，其中N_1属于一类元素，N_2属于二类元素，取n个元素其中k个是一类元素概率
>
> ---
>
> - **均匀分布：** 记为 X~U[a,b] *连续型概率分布*
>
> > $$f(x)=\begin{cases}\frac{1}{b-a} & a\leq x\leq b\\ 0 & \text{其他}\end{cases}$$
> >
> > $$\text{其分布函数：}F(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & a\leq x<b\\ 1 & b\leq x\end{cases}$$
> >
> > $$X-U[a,b]\ \text{当}[c,d]\subset[a,b]\text{时}\ P\{c\leq x\leq d\}=\int_c^d\frac{1}{b-a}dt=\frac{d-c}{b-a}$$
> >
> > *在相同长度间隔的分布概率是等可能的*
>
> - **指数分布：** $\lambda>0$ 记为 X~E(λ) *连续型概率分布*

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & 其他\end{cases}$$

$$其分布函数：F(x)=\begin{cases}1-e^{-\lambda x} & x>0\\ 0 & 其他\end{cases}$$

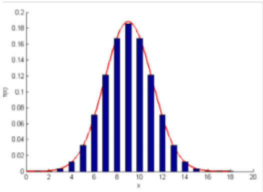
指数分布特点：P{X > s+t | X > s} = P{x > t}

通常用于描述对发生的缺陷数或系统故障数的测量结果。例如某种产品或零件经过一段时间t0的工作后,仍然如同新的产品一样,不影响以后的工作寿命值，或者说，经过一段时间t0的工作之后，该产品的寿命分布与原来还未工作时的寿命分布相同

- **正态分布：** 记为 X~N(u,σ^2)

$$\Phi x=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-u)^2}{2\sigma^2}} \quad -\infty<x<+\infty$$

$$其分布函数：F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-u)^2}{2\sigma^2}}dt$$



一般正态分布密度函数性质：

1. 以x=u为对称轴
2. x=u时，f(x)为最大值为$\frac{1}{\sqrt{2\pi}\sigma}$
3. σ固定，u变化，函数左右移动，u固定，σ变化，函数上下压缩

标准正态分布： 当u=0，$\sigma=1$时

$$密度函数：\phi_0(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

$$分布函数：F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$$

性质：

1. 关于y周对称 (偶函数)，$\phi_0(x)=\phi_0(-x)$
2. $F(-x)=1-F(x)$
3. $F(0)=\frac{1}{2}$

一般正态分布函数转标准正态分布函数

$$F(x)=F_0(\frac{x-u}{\sigma})$$

$$P\{a<X\le b\}=F(b)-F(a)=F_0(\frac{b-u}{\sigma})-F_0(\frac{a-u}{\sigma})$$

$$F(x)：一般正态分布函数$$

$$F_0(x)：标准正态分布函数$$

随机变量函数的分布

> *求分布函数关于X（随机变量）的复合函数*
>
> **离散型变量函数得到值之后概率相加**
>
> | X | -1 | 0 | 1 |
> | --- | --- | --- | --- |
> | P | 0.2 | 0.4 | 0.4 |
>
> 求变量$Y=X^2$的分布律
>
> | Y | 1 | 0 |
> | --- | --- | --- |
> | P | 0.6 | 0.4 |
>
> **连续型变量函数的密度函数**
>
> **随机变量函数的密度函数**
>
> $$f_Y(y)=\begin{cases} f_X(h(y))\,|h^{`}(y)| & a<y<b \\ 0 & \text{其他} \end{cases}$$
> $$g(x)\text{是随机变量}X\text{的函数}$$
> $$x=h(y)\text{：是变量函数}y=g(x)\text{的反函数，}h^{`}(y)\text{关于}y\text{变量的导数}$$
>
> - 当随机变量函数为$Y=kx+b(k\neq 0)$线性函数时候，其密度函数为 *同上*
>
> $$f_Y(x)=\frac{1}{|k|}f_X(\frac{x-b}{k})$$
> $$\text{原随机变量密度函数为}f_X(x)$$