哈爾濱工業大學

第3章　随机变量及其分布

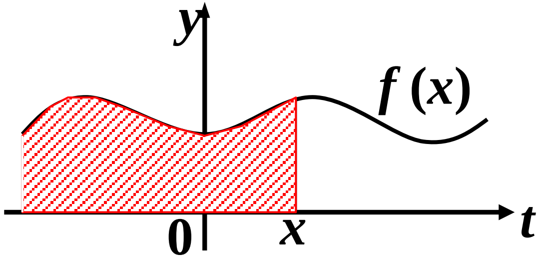
第15讲 连续型随机变量

连续型随机变量

定义　设随机变量X的分布函数为$F(x)$，若存在一个非负的函数$f(x)$，对任何实数x，有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

称X为连续型随机变量，称$f(x)$为X的概率密度函数，简称概率密度. 也可记为 $f_X(x)$.



连续型随机变量

- 由定义，可得下面两个结论

(1) 连续型随机变量的分布函数一定是连续的；

(2) 对$f(x)$的连续点，有

$$\left.\begin{aligned} F'(x) &= f(x) \\ F(x) &= \int_{-\infty}^{x} f(t)\mathrm{d}t \end{aligned}\right\} \ F(x)\text{与}f(x)\text{可以互推.}$$

概率密度的性质

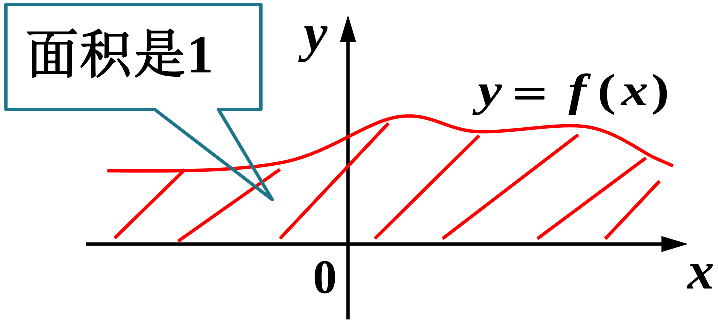
(i) $f(x) \geq 0$,

(ii) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1.$

这两条是判定函数 $f(x)$ 是否为概率密度函数的充要条件.

$$F(+\infty) = \int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1.$$

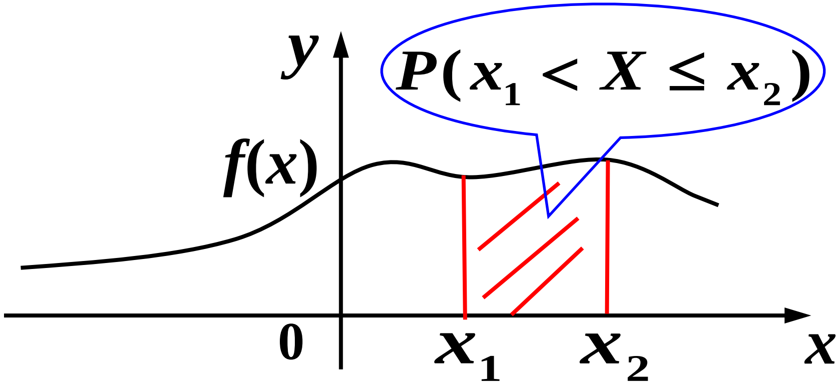
$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$



概率密度的性质

(iii) $P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\mathrm{d}x.$

$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$



连续型随机变量

- 连续型随机变量取任一指定值的概率为0，

即 $P(X=a)=0.$

这是因为

$$0 \le P(X=a) \le P(a-\Delta x < X \le a) = F(a) - F(a-\Delta x),\ \Delta x > 0.$$

由$F(x)$连续得

$$\lim_{\Delta x \to 0}(F(a) - F(a-\Delta x)) = 0 \Rightarrow P(X=a)=0.$$

连续型随机变量

$$P(X=a)=0 \nRightarrow (X=a)=\varnothing.$$

同理 $P(A)=0$不能推出$A=\varnothing$,

$P(B)=1$不能推出$B=S$.

$$\boxed{P(a<X\le b)=P(a\le X<b)=P(a\le X\le b)=P(a<X<b).}$$

连续型随机变量

- $f(x)$的值是如何反应概率呢？

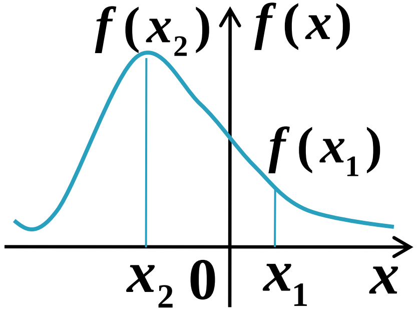
若x是$f(x)$的连续点，则

$$f(x)=F'(x)=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}$$

$$\lim_{\Delta x\to 0}\frac{P(x<X\le x+\Delta x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\int_x^{x+\Delta x}f(t)\mathrm{d}t}{\Delta x}.$$

$$\boxed{P(x<X\le x+\Delta x)\approx f(x)\Delta x.}$$

表明X落在x附近领域$(x,x+\Delta x)$的概率约等于$f(x)\Delta x$.



$f(x_2)>f(x_1)$

例1 设连续型随机变量X的概率密度

$$f(x)=\begin{cases} cx^2+x, & 0\le x\le 0.5, \\ 0, & \text{其他}. \end{cases}$$

求(1) 系数c；(2) X的分布函数$F(x)$；(3) $P(-0.5<X<0.3)$.

解

$$\underline{1}=\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\underline{\int_0^{0.5}(cx^2+x)\mathrm{d}x}=\left(\frac{c}{3}x^3+\frac{1}{2}x^2\right)\Bigg|_0^{0.5}$$

$$=\frac{c}{24}+\frac{1}{8}, \Rightarrow c=21.$$

$$\boxed{P(0\le X\le 0.5)=1.}$$

$$(2)F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

$$f(x)=\begin{cases}21x^2+x, 0\le x\le 0.5,\\ 0, \qquad 其它.\end{cases}$$

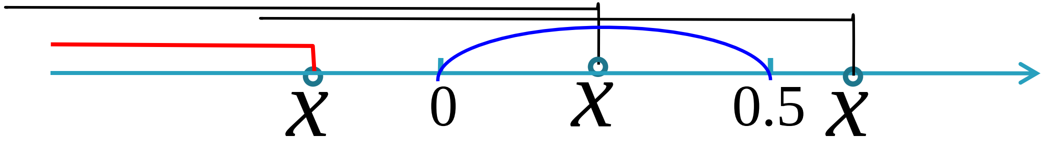
$$\boxed{由P(0\le X\le 0.5)=1.}$$

1^0当$x<0$时，$F(x)=\int_{-\infty}^{x} 0\mathrm{d}t=0;$

2^0当$x\ge 0.5$时, $F(x)=P(X\le x)=1.$

3^0当$0\le x<0.5$时，$F(x)=\int_{-\infty}^{0} 0\mathrm{d}t+\int_{0}^{x}(21t^2+t)\mathrm{d}t$

$$=7x^3+x^2/2;$$



$$f(x)=\begin{cases}21x^2+x, 0\le x\le 0.5,\\ 0, \quad \text{其它}.\end{cases}$$

即， $$F(x)=\begin{cases}0, & x<0,\\ 7x^3+x^2/2, 0\le x<0.5,\\ 1, & x\ge 0.5.\end{cases}$$

$$(3)P(-0.5<X<0.3)=\int_{-0.5}^{0.3}f(x)\mathrm{d}x=\int_0^{0.3}(21x^2+x)\mathrm{d}x$$

$$=(7x^3+x^2/2)\Big|_0^{0.3}=0.234.$$

或 $$P(-0.5<X<0.3)=F(0.3)-F(-0.5)$$

$$=7\cdot(0.3)^3+(0.3)^2/2-0=0.234.$$

练习

1. 设

$$X \sim f(x)=\begin{cases} x, & 0\le x\le 1 \\ 2-x, & 1\le x\le 2 \\ 0, & \text{其它} \end{cases}$$

求$F(x)$

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

解

$$F(x)=\begin{cases} 0, & x<0, \\ \int_0^x t\mathrm{d}t=\dfrac{x^2}{2}, & 0\le x<1, \\ \int_0^1 t\mathrm{d}t+\int_1^x (2-t)\mathrm{d}t=2x-1-\dfrac{x^2}{2}, & 1\le x<2, \\ 1, & x\ge 2. \end{cases}$$

例2 设随机变量X的分布函数为

$$F(x)=\begin{cases} 0, & x<0, \\ x^2, & 0\le x\le 1, \\ 1, & x>1. \end{cases}$$ 求X的概率密度.

解

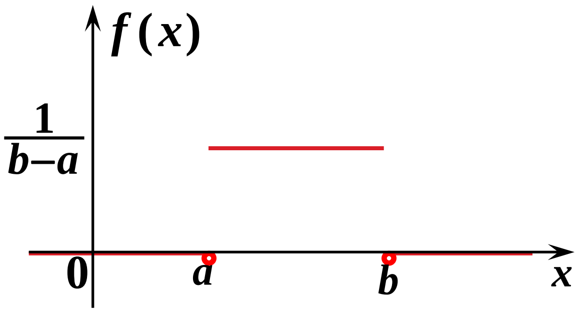
$$f(x)=F'(x)=\begin{cases} 2x, 0\le x\le 1, \\ 0, \text{ 其他}. \end{cases}$$

几种重要的连续型随机变量

均匀分布(Uniform)

定义 若随机变量X的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b,\\ 0, & 其他.\end{cases}\qquad (a<b).$$



称X在区间$[a, b]$上服从均匀分布, 记为$X \sim U[a, b]$.

$$f(x)\ge 0; \qquad \int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_a^b \frac{1}{b-a}\mathrm{d}x=1.$$

满足概率密度性质.

均匀分布

➢ 均匀的含义是等可能

若 $X \sim U[a, b]$, (x_1, x_2)为$[a, b]$中的任一子区间,

则

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{\mathrm{d}x}{b-a} = \frac{1}{b-a}(x_2 - x_1).$$

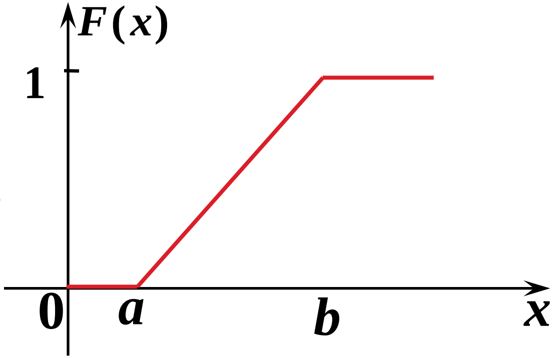
说明：X落在长度相等的各个子区间的可能性是相等的. 属于几何概率.

若$X \sim U[a,b]$,则$P(a \le X \le b) = 1.$

概率密度的性质

若$X \sim U[a, b]$，则X的分布函数为

$$F(x)=\begin{cases}0, & x<a, \\ \dfrac{x-a}{b-a}, & a \le x < b, \\ 1, & x \ge b.\end{cases}$$



$$\boxed{F(x)=\int_{-\infty}^{x} f(t)dt.}$$

均匀分布的用途

- 设通过某站的汽车10分钟一辆，则乘客候车时间X在$[0, 10]$上服从均匀分布；
- 某电台每隔20分钟发一个信号，我们随手打开收音机，等待时间X在$[0, 20]$上服从均匀分布；
- 随机投一根针于坐标纸上，它和坐标轴的夹角X在$[0, \pi]$上服从均匀分布.

例3 设随机变量$X\sim U(1,6)$, 求方程$x^2+Xx+1=0$有实根的概率.

解 由$X\sim U(1,6)$得X的概率密度为

$$f(x)=\begin{cases}\frac{1}{5}, & 1<x<6,\\ 0, & \text{其他}.\end{cases}$$

方程有实根 $\Leftrightarrow \Delta\ge 0 \Leftrightarrow X^2-4\ge 0 \Leftrightarrow X\ge 2$或$X\le -2$,

故, 方程有实根的概率为

$$P(X\ge 2)+P(X\le -2)=\int_2^6 \frac{1}{5}\mathrm{d}x+0=\frac{4}{5}.$$

练习

1. 设随机变量X在(2,5)上服从均匀分布，现对X进行三次独立观测，试求至少有两次观测值大于3的概率.

解 因为随机变量X在(2,5)上服从均匀分布，所以X的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{3}, & 2<x<5\\ 0, & \text{其它}\end{cases}$$

事件“对X的观测值大于3”的概率为 $P\{X>3\}=\int_3^5 \dfrac{1}{3}dx=\dfrac{2}{3}$

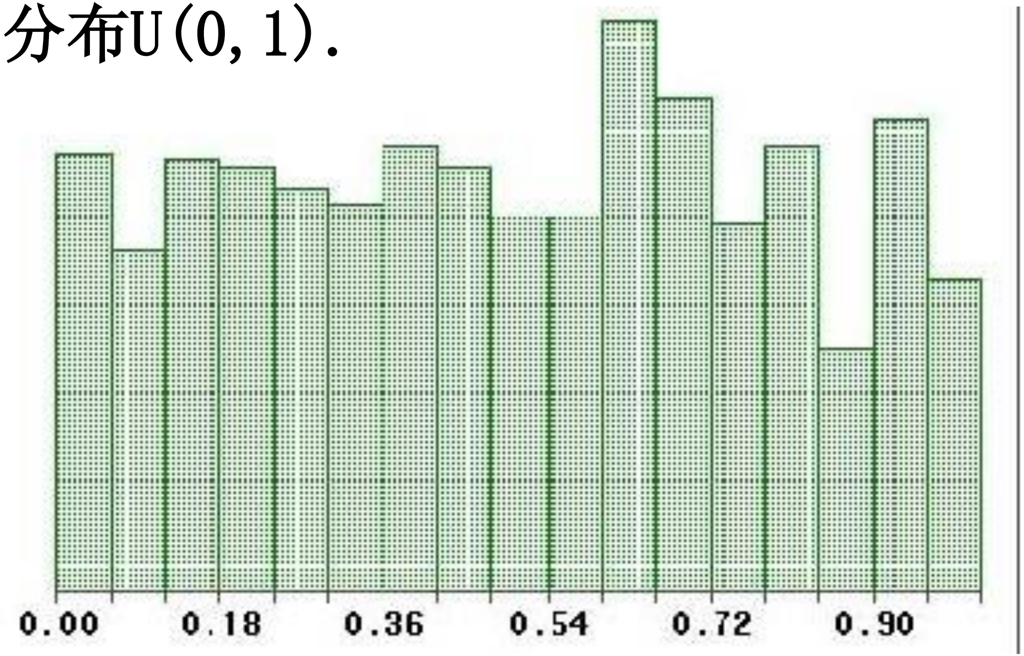
设Y表示三次独立观测中观测值大于3的次数，则 $Y\sim B\left(3,\dfrac{2}{3}\right)$

于是

$$P\{Y\geq 2\}=C_3^2\left(\frac{2}{3}\right)^2\frac{1}{3}+C_3^3\left(\frac{2}{3}\right)^3=\frac{20}{27}$$

- 区间(0, 1)上的均匀分布U(0,1)在计算机模拟中起着重要的作用.
- 实用中，用计算机程序可以在短时间内产生大量服从(0, 1)上均匀分布的随机数. 它是由一种迭代过程产生的.
- 严格地说，计算机中产生的U(0,1) 随机数并非完全随机，但很接近随机，故常称为伪随机数.

- 如取n足够大，独立产生n个U(0,1)随机数，则从用这 n 个数字画出的频率直方图就可看出，它很接近于(0, 1) 上的均匀分布U(0,1).



指数分布(Exponential)

■ **定义** 若连续型随机变量X的概率密度

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}\quad (\lambda>0).$$

称X服从参数为 λ 的指数分布. 记为$X\sim E(\lambda)$.

满足：$f(x)\ge 0,\quad \int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{0}^{+\infty}\lambda e^{-\lambda x}\mathrm{d}x=1.$

分布函数为

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}\quad \lambda>0.$$

指数分布常用来近似地表示各种寿命的分布.

指数分布的无记忆性

$$X \sim E(\lambda),\quad \forall s>0, t>0,$$

$$P(X>s+t \mid X>s) = P(X>t).$$

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

因为

$$P\{X>s+t \mid X>s\}=\frac{P\{X>s+t, X>s\}}{P\{X>s\}}$$

$$=\frac{P\{X>s+t\}}{P\{X>s\}}=\frac{1-F(s+t)}{1-F(s)}=\frac{\mathrm{e}^{-\lambda(s+t)}}{\mathrm{e}^{-\lambda s}}=\mathrm{e}^{-\lambda t}$$

$$=P(X>t).$$

例4设随机变量X服从参数为1的指数分布，a为常数且大于零，则$P(X<a+1|X>a)=$______

解 已知 $X \sim E(1)$

$$F(x)=\begin{cases}1-\mathrm{e}^{-x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

$$P(X<a+1\,|\,X>a)=1-P(X\ge a+1\,|\,X>a)=1-P(X\ge 1)$$

$$=P(X<1)=F(1)=1-\mathrm{e}^{-1}$$

例5　设机器相邻两次故障的时间间隔（小时）X服从参数为1/5的指数分布，求在机器已经无故障工作了8小时的情况下，再无故障工作10小时的概率.

解　已知

$$X \sim E(1/5). \qquad F(x)=\begin{cases}1-\mathrm{e}^{-x/5}, & x>0,\\ 0, & x\le 0.\end{cases}$$

$$P(X>18\mid X>8)=P(X>10)=1-P(X\le 10)$$

$$=1-F(10)=1-(1-\mathrm{e}^{-10/5})=\mathrm{e}^{-10/5}=\mathrm{e}^{-2}.$$

谢　　谢！