§2.5 连续型随机变量

连续型随机变量的概念

定义 设 X 是一随机变量，若存在一个非负可积函数 $f(x)$，使得

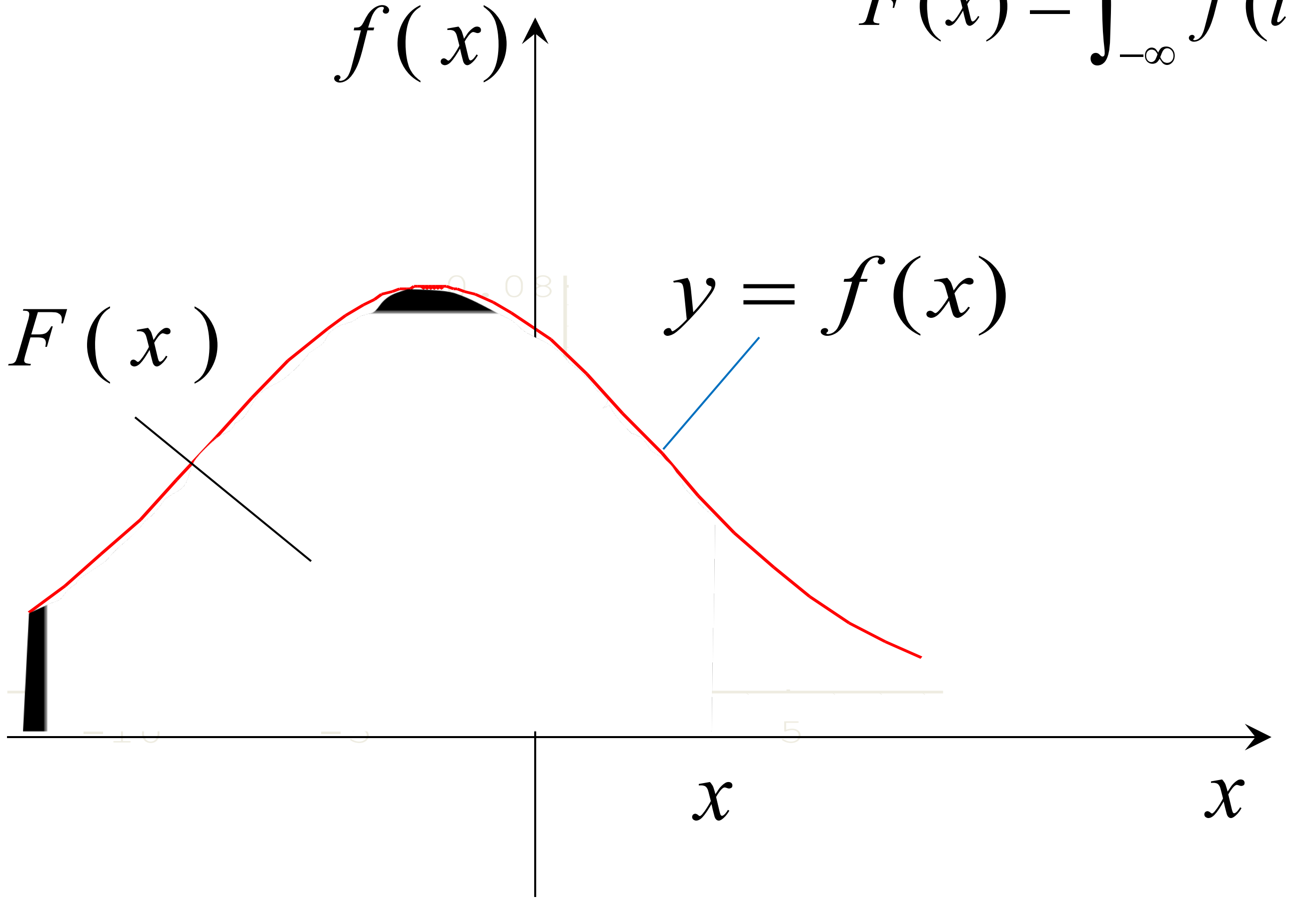
$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t \qquad -\infty < x < +\infty$$

其中 $F(x)$ 是它的分布函数

则称 X 是连续型随机变量，$f(x)$ 是它的概率密度函数(**p.d.f.**)，简称为密度函数或概率密度

分布函数$F(x)$与密度函数$f(x)$的几何意义

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



$f(x)$的基本性质

- $f(x) \geq 0$
- $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = F(+\infty) = 1$

另外，$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t \qquad -\infty < x < +\infty$

积分上限函数

(1)连续函数;

(2)在$f(x)$的连续点可导;

(3) $f(x)=F'(x)$。

$f(x)$的含义：

$$f(x_0) = F'(x_0)$$

$$= \lim_{\Delta x \to +0} \frac{F(x_0 + \Delta x) - F(x_0)}{\Delta x}$$

$$= \lim_{\Delta x \to +0} \frac{P(x_0 < X \le x_0 + \Delta x)}{\Delta x}$$

$$f(x_0) \approx \frac{P(x_0 < X \le x_0 + \Delta x)}{\Delta x}$$

$f(x_0)$描述了X在x_0附近单位长度的区间内取值的概率

注意: 对于连续型随机变量X, $P(X=a)=0$

这里 a 可以是随机变量 X 的一个可能的取值

事实上 $(X=a)\subset(a-\Delta x<X\le a)$ $\Delta x>0$

$$0\le P(X=a)\le P(a-\Delta x<X\le a)=\int_{a-\Delta x}^{a}f(x)\mathrm{d}x$$

$$0\le P(X=a)\le \lim_{\Delta x\to+0}\int_{a-\Delta x}^{a}f(x)\mathrm{d}x=0$$

$\Longrightarrow$ $P(X=a)=0$

命题 连续型随机变量取任一常数的概率为零

概率为1 (零) 的事件末必发生 (不发生)

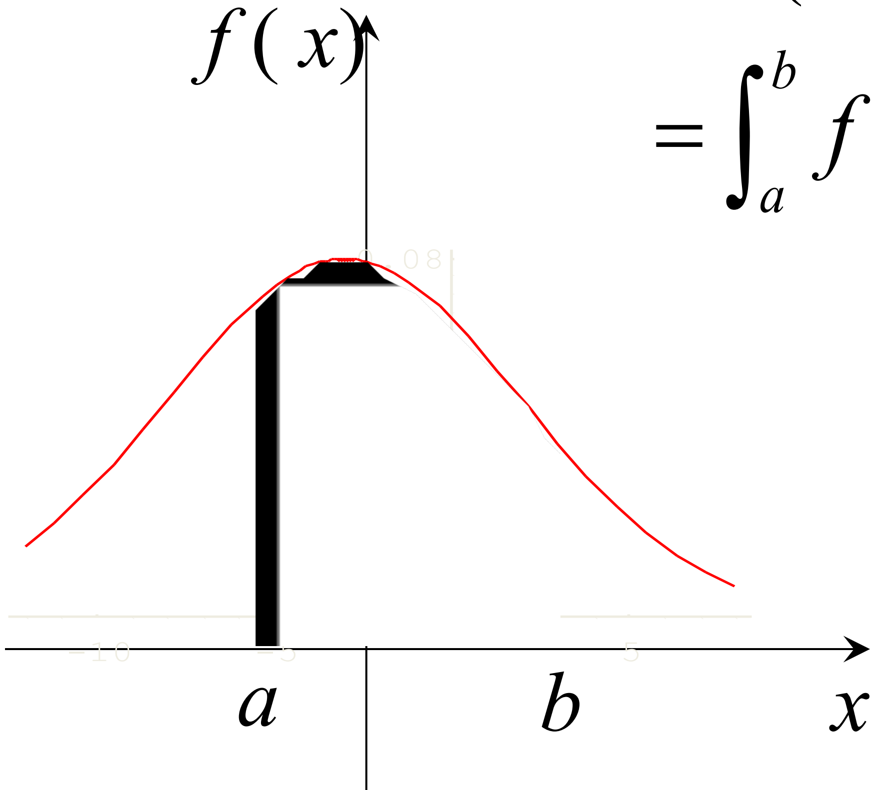
对于连续型随机变量X

$$P(a < X \le b) = P(a \le X \le b)$$

$$= P(a < X < b)$$

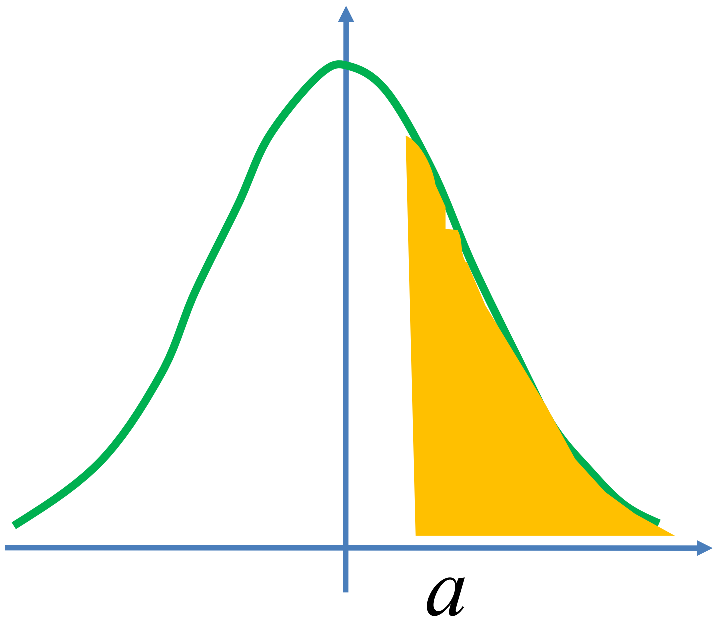
$$= P(a \le X < b)$$

$$= \int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$



$$P(X \le b) = P(X < b) = F(b)$$

$$P(X > a) = P(X \ge a) = 1 - F(a)$$



例1 有一批晶体管，已知每只的使用寿命 X 为连续型随机变量，其概率密度函数为

$$f(x)=\begin{cases}\dfrac{c}{x^2}, & x>1000\\ 0, & \text{其他}\end{cases} \qquad (c\text{ 为常数})$$

(1) 求常数 c

(2) 已知一只收音机上装有3只这样的晶体管，每只晶体管能否正常工作相互独立，求在使用的最初1500小时只有一个损坏的概率.

解 (1) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int_{1000}^{+\infty} \frac{c}{x^2}\mathrm{d}x = c\cdot(-\frac{1}{x})\Big|_{1000}^{+\infty} = \frac{c}{1000} \overset{令}{=} 1$

$\Longrightarrow c = 1000$

(2) 设事件 A 表示一只晶体管的寿命小于1500小时

$$P(A) = P(0 \le X < 1500) = \int_{1000}^{1500} \frac{1000}{x^2}\mathrm{d}x = \frac{1}{3}$$

设在使用的最初1500小时三只晶体管中损坏的只数为 $Y \sim B\left(3, \frac{1}{3}\right)$

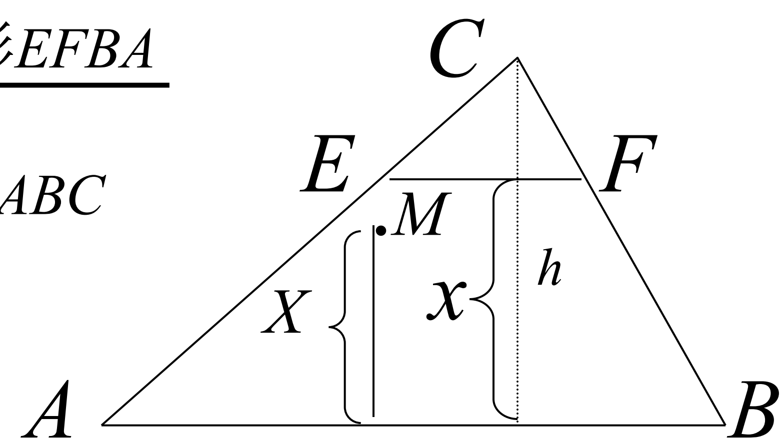
$$P(Y=1) = C_3^1\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

例2 在高为h的 $\triangle ABC$ 中任取一点M ,点M到AB的距离为X，求X的概率密度函数$f(x)$.

解 当$0 \le x \le h$ 时

作 EF//AB，使EF与AB间的距离为x

$$F(x)=P(X\le x) = \frac{S_{梯形EFBA}}{S_{\Delta ABC}}$$



$$=1-\frac{S_{\Delta CEF}}{S_{\Delta ABC}} = 1-\left(\frac{h-x}{h}\right)^2$$

于是

$$F(x)=\begin{cases} 0 & x<0 \\ 1-\left(\dfrac{h-x}{h}\right)^2 & 0\le x\le h \\ 1 & x>h \end{cases}$$

$$\therefore f(x)=F'(x)=\begin{cases} 2\dfrac{h-x}{h} & 0\le x\le h \\ 0 & \text{其他} \end{cases}$$

2.6 常见的连续性随机变量的分布

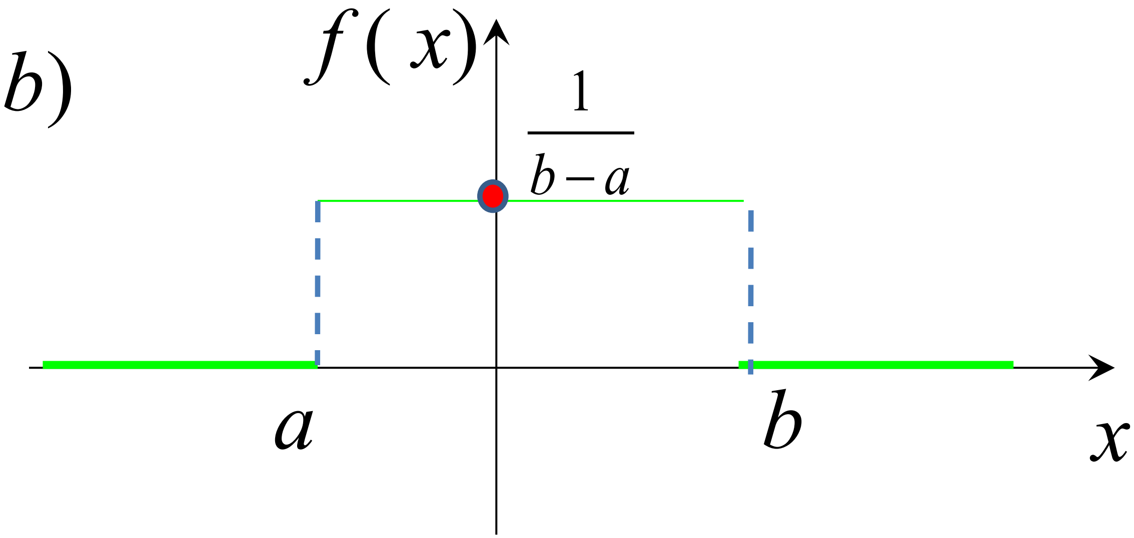
(1) 均匀分布

若 X 的密度函数为：

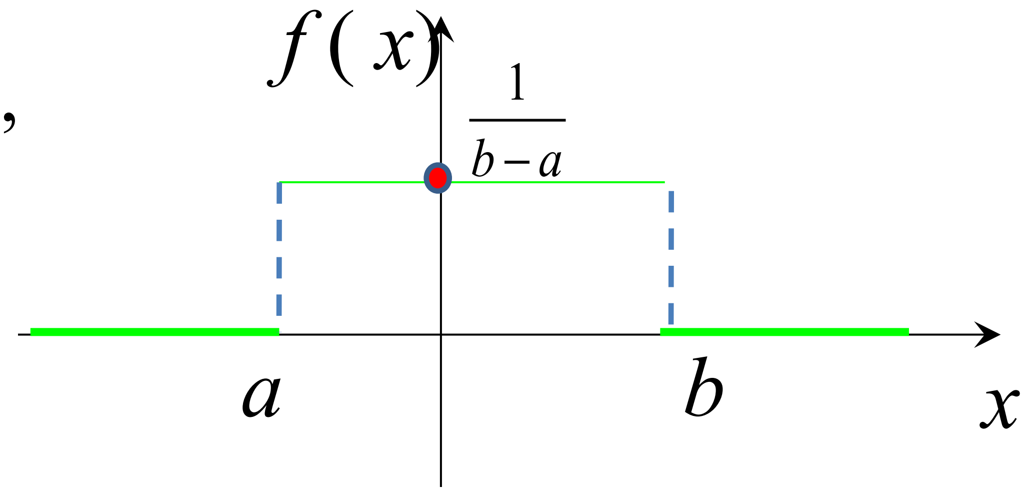
$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$$

则称 X 服从区间(a,b)上的均匀分布

记作 $X\sim U(a,b)$



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



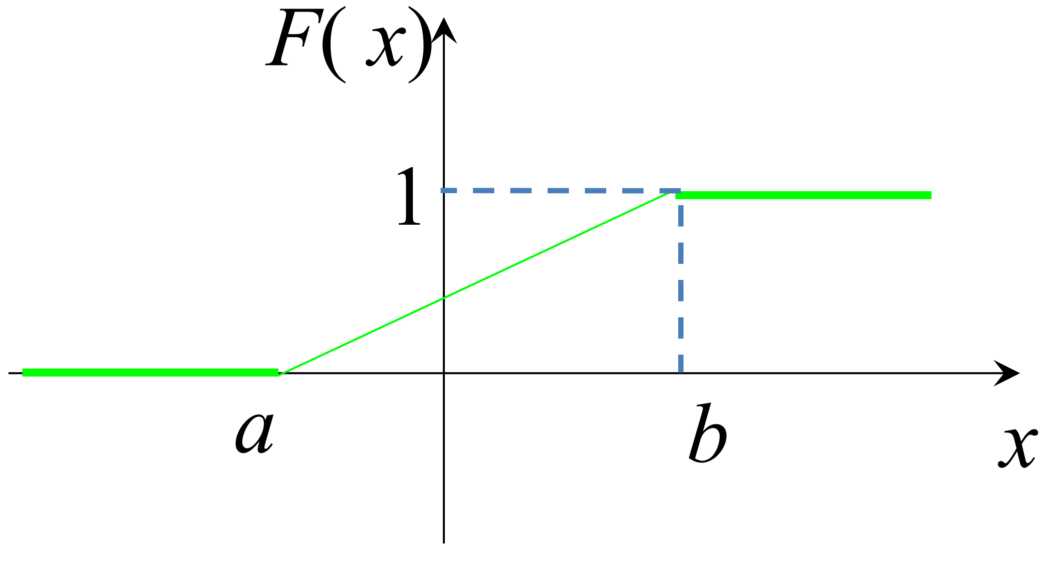
$$F(x)=\int_{-\infty}^{x} f(x)dx$$

当$x<a$时，$f(x)=0$， $F(x)=\int_{-\infty}^{x} 0dx \ =0$

当$a\le x<b$时，$F(x)=\int_{-\infty}^{a} 0dx+\int_{a}^{x}\frac{1}{b-a}dx \ =\frac{x-a}{b-a}$

当$x\ge b$时，$F(x)=\int_{-\infty}^{a} 0dx+\int_{a}^{b}\frac{1}{b-a}dx+\int_{b}^{+\infty} 0dx \ =1$

X的分布函数为 $F(x)=\begin{cases} 0, & x<a, \\ \dfrac{x-a}{b-a}, & a\le x<b, \\ 1 & x\ge b \end{cases}$



$$\forall (c,d) \subset (a,b)$$

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

即 X 的取值在(a,b)内任何长为 $d-c$ 的小区间的概率与小区间的位置无关，只与其长度成正比．这正是几何概型的情形．

应用场合

在进行大量数值计算时，如果在小数点后第 k 位进行四舍五入，则产生的误差可以看作

服从 $U\left(-5\times 10^{-k}, 5\times 10^{-k}\right)$

例3 秒表的最小刻度差为0.01秒. 若计时精度是取最近的刻度值, 求使用该秒表计时产生的随机误差X的概率密度, 并计算误差的绝对值不超过0.004秒的概率.

解 由题设知随机误差 X 等可能地取得区间 [-0.005, 0.005]上的任一值，则

$$X \sim U[-0.005 \quad 0.005]$$

$$f(x) = \begin{cases} 100, & |x| \le 0.005 \\ 0, & \text{其他} \end{cases}$$

所以 $P(|X| \le 0.004) = \int_{-0.004}^{0.004} 100 dx = 0.8$

(2) 指数分布

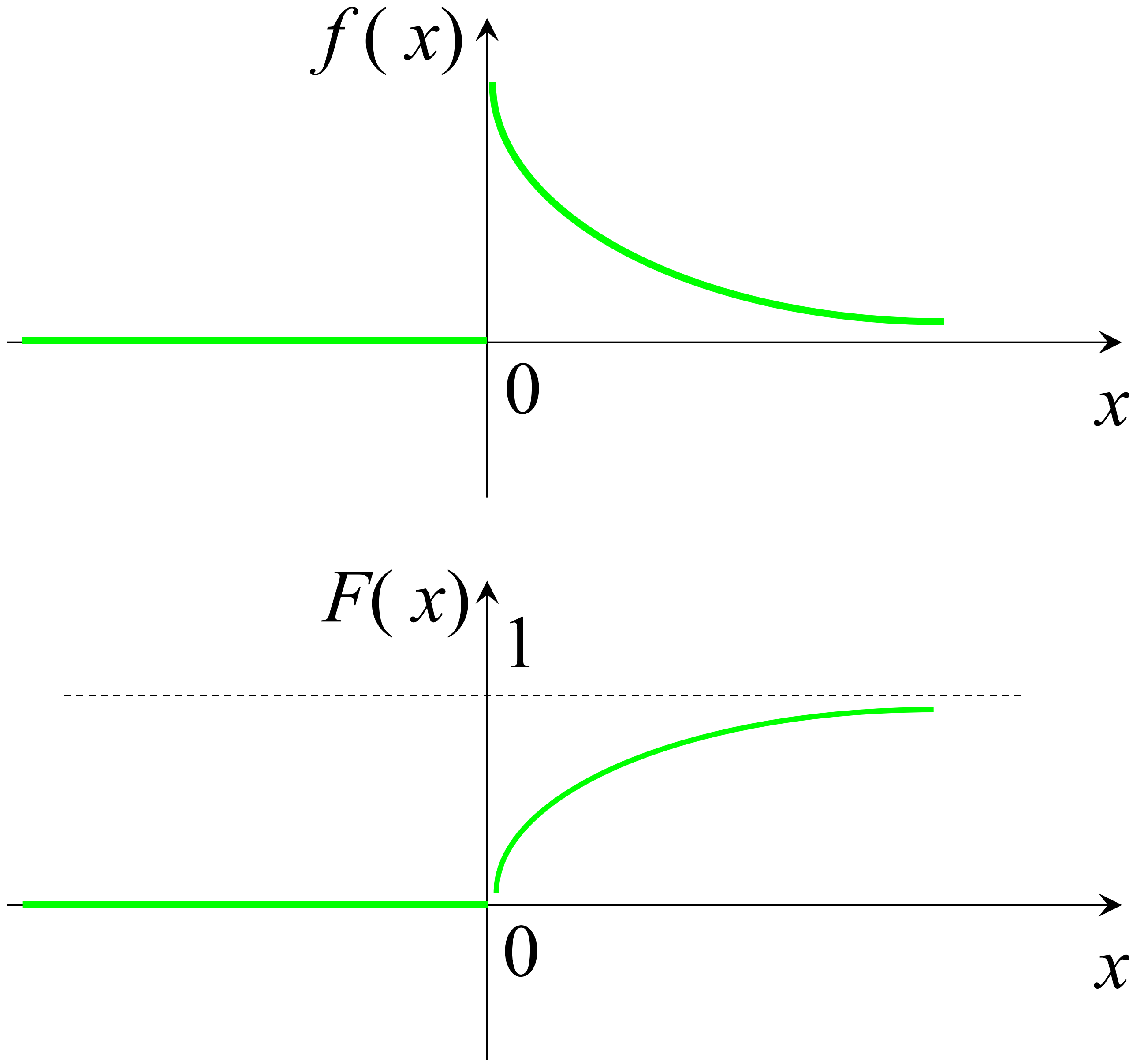
若 X 的密度函数为

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

则称 X 服从 参数为λ的指数分布

记作 $X\sim E(\lambda)$

X 的分布函数为 $F(x)=\begin{cases}0, & x<0\\ 1-e^{-\lambda x}, & x\geq 0\end{cases}$



$f(x)$
0
x
$F(x)$
1
0
x

对于任意的 $0 < a < b$,

$$P(a < X < b) = \int_a^b \lambda e^{-\lambda x} \mathrm{d}\, x$$

$$= e^{-\lambda a} - e^{-\lambda b} = F(b) - F(a)$$

应用场合 用指数分布描述的实例有：

随机服务系统中的服务时间

电话问题中的通话时间

无线电元件的寿命

动物的寿命

$\left.\begin{array}{l}\text{无线电元件的寿命}\\\text{动物的寿命}\end{array}\right\}$ 指数分布常作为各种“寿命”分布的近似

指数分布的“无记忆性”

若 $X \sim E(\lambda)$,则

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

事实上

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

$$= \frac{1 - P(X \le s + t)}{1 - P(X \le s)} = \frac{1 - F(s + t)}{1 - F(s)}$$

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

例4 假定一大型设备在任何长为 t 的时间内发生故障的次数 $N(t)$ 服从参数为λt 的Poisson分布,

(1) 求相继两次故障的时间间隔 T 的概率分布

(2) 求设备已经无故障运行 8 小时的情况下，再无故障运行 10 小时的概率.

解 (1) $F_T(t) = P(T \le t)$

$$= \begin{cases} 0, & t < 0 \\ 1 - P(T > t), & t > 0 \end{cases}$$

$$P(T > t) = P(N(t) = 0)$$

$$= \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

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

$$f(t)=\begin{cases}0, & t<0\\ \lambda e^{-\lambda t}, & t>0\end{cases}$$

即 $T\sim E(\lambda)$

(2) 由指数分布的“无记忆性”

$$\begin{aligned}P(T>18\,|\,T>8) &= P(T>8+10\,|\,T>8)\\ &= P(T>10)\\ &= e^{-10\lambda}\end{aligned}$$