§2.3 连续型随机变量

在2.2节中，我们已经学习了离散型 随机变量，**离散型随机变量的特征是：取值有限或可列多个**。除此之外，有的随机变量取值不是有限或可列集，比如如测量的误差、电子元件的寿命、人的身高等。这类随机变量称为非离散型的，本节讲学习非离散型随机变量中重要的一类:**连续型随机变量**。

定义：设 X 是随机变量，F(x)是X的分布函数，如果存在一**非负**函数 f(x)使得

$$F(x)=\int_{-\infty}^{x} f(t)dt,\ \forall\quad x\in R$$

则称X 是**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**。

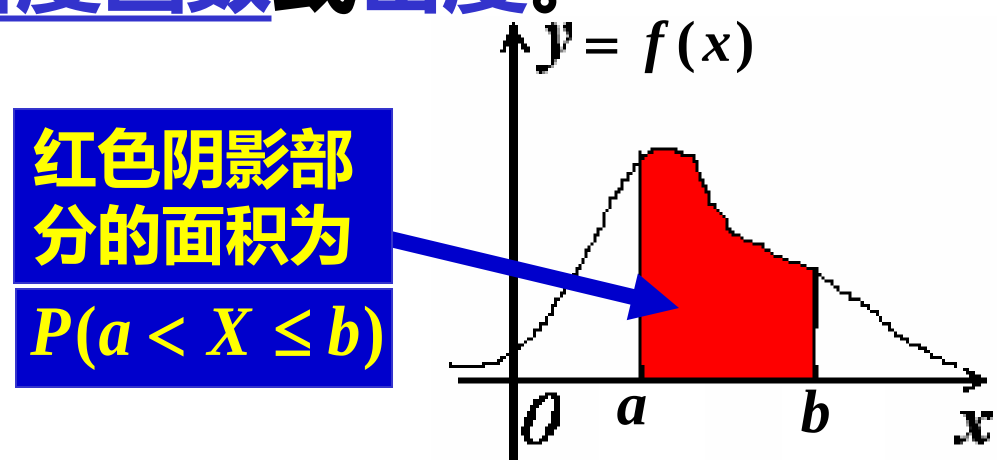
连续型随机变量的等价定义：

定义：设 X 是随机变量，如果存在一非负函数 f(x)满足

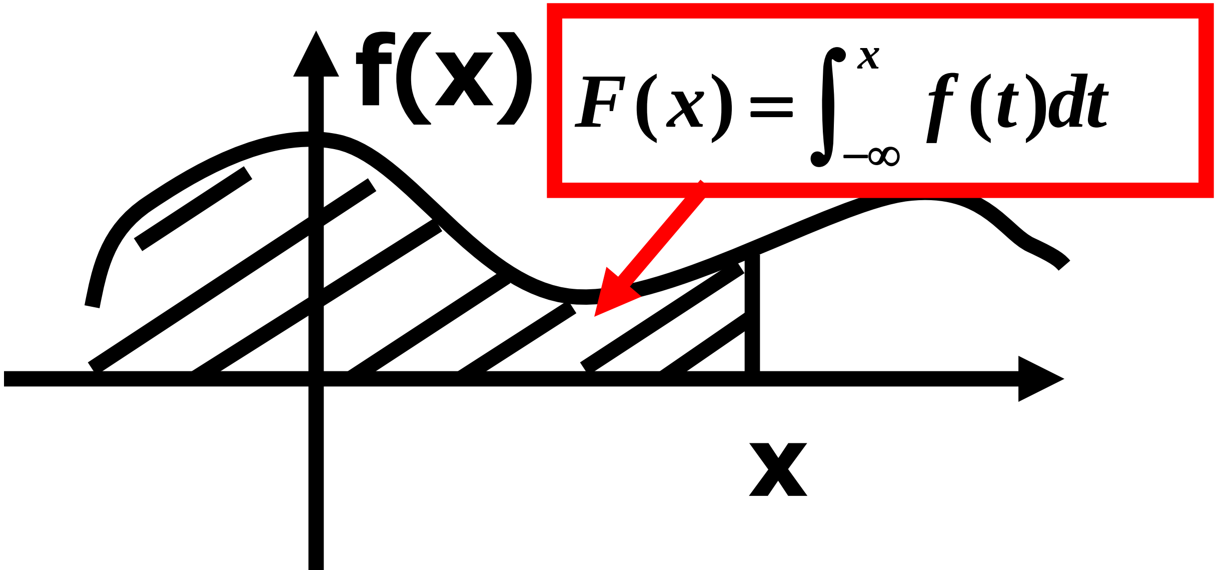
$$P(a < X \le b) = \int_a^b f(x)dx,\ \ \forall\ \ a,b \in R,\ \ a < b,$$

则称X 是连续型随机变量，f(x)称为 X 的概率密度函数，简称密度函数或密度。

由定积分的几何意义，其解释如图：



连续性随机变量分布函数的几何意义：



由连续性随机变量**的定义知，连续型随机变量的密度函数有如下性质：**

1）非负性：$f(x)\geq 0;$

2）归一性：$\int_{-\infty}^{+\infty} f(x)dx=1.$

注： 连续型随机变量的密度函数与离散型随机变量 的分布律有两条相同的性质：非负性、归一性。但密度函数不是概率！即： $f(x) \neq P(X=x)$

注： 连续型随机变量由密度函数唯一确定。非负性和归一性两条性质揭示了密度函数的本质，也就是说，如果一个函数满足以上具有这两条性质，则它一定是某连续型随机变量的密度函数。即

f(x)满足1), 2) $\Leftrightarrow$ f(x)是某连续型随机变量的密度函数。

由密度函数的定义知，连续型随机变量的密度函数还具有如下性质：

3）连续型随机变量取任意单个值的概率为0。

注　性质3）是连续型随机变量与离散型随机变量的一个重要区别。连续型随机变量取任何点的概率为0，而离散型随机变量取某点的概率要视此点是否是它的可能取值而定。因此，离散型随机变量要“点点计较”，而连续型随机变量可“点点不计较”。因此有

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

$$=P(a\le X\le b)=\int_a^b f(x)dx,\forall a,b\in R,\ \ a<b.$$

更一般地，我们有如下结论：

定理：设 X 是连续型随机变量，f(x)为其密度函数，如果对任意 $G\subset R$， **f(x)在G上可积，则**

$$P(X\in G)=\int_G f(x)dx.$$

连续型随机变量 的分布函数与密度函数的关系

定理：设 X 是连续型随机变量，f(x)为其密度函数，分布函数为F(x)，则

1）$F(x)$是连续函数，

$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)dt,\quad -\infty<x<\infty.$$

2）F(x) 在f(x)的连续点，有

$$\frac{d}{dx}F(x)=f(x).$$

注 对于连续型随机变量而言，密度函数唯一地确定了随机变量及其分布函数，且分布函数连续。但是，若知道一个连续型随机变量的分布函数，其密度函数不一定唯一，但一个连续型随机变量的任意两个密度函数之间至多有可数个点的取值不同。

例2.13 已知随机变量X的密度为

$$f(x)=\begin{cases}ax+b, 1<x<3\\0, \quad else\end{cases}$$

且P(2<X<3)=2P(1<X<2),求常数a,b及分布函数F(x).

解：由密度函数的规范性有：

$$1=\int_{-\infty}^{\infty} f(x)dx=\int_{1}^{3}(ax+b)dx=4a+2b$$

$$2P(1<X<2)=2\times\int_{1}^{2}(ax+b)dx=3a+2b$$

$$P(2<X<3)=\int_2^3 (ax+b)dx = 2.5a+b$$

$$\therefore \begin{cases} 4a+2b=1, \\ 2(3a+2b)=2.5+b, \end{cases} \Rightarrow$$

$$a=\frac{1}{3}, b=-\frac{1}{6}, \text{即} f(x)=\begin{cases} \frac{1}{3}x-\frac{1}{6}, 1<x<3 \\ 0, \qquad else \end{cases}$$

下求分布函数**F(x)**,由于密度函数是分段函数,
我们分段讨论:

$$(1) -\infty < x \le 1,$$

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0$$

$$(2) 1 < x < 3, F(x) = \int_{-\infty}^{x} f(t)dt$$

$$= \int_{-\infty}^{1} 0dt + \int_{1}^{x} (\frac{1}{3}t - \frac{1}{6})dt = \frac{1}{6}x(x-1)$$

$$(3) x \ge 3, F(x) = \int_{-\infty}^{x} f(t)dt$$

$$= \int_{-\infty}^{1} 0dt + \int_{1}^{3} (\frac{1}{3}t - \frac{1}{6})dt + \int_{3}^{x} 0dt = 1$$

$$F(x)=\begin{cases} 0, & x\le 1 \\ \frac{1}{6}x(x-1), & 1<x<3 \\ 1, & x\ge 3 \end{cases}$$

例 设X是连续型随机变量，其分布函数F(x)为：

$$F(x)=\begin{cases} a, & x<1 \\ bx\ln x+cx+d, & 1\le x\le e \\ d, & x>e \end{cases}$$

试确定参数a,b,c,d的值。

解 连续型随机变量的分布函数在实数域上处处连续，故F(x)在x=1，x=e处连续，于是

$$F(1)=\lim_{x\to 1-}F(x)=a\Rightarrow c+d=a, \qquad (1)$$

$$F(e)=\lim_{x\to e+}F(x)=d\Rightarrow be+ce+d=d, (2)$$

又

$$0=F(-\infty)=\lim_{x\to-\infty}F(x)=a \qquad (3)$$

$$1=F(+\infty)=\lim_{x\to+\infty}F(x)=d \qquad (4)$$

由等式（1）-（4）得到

$$a=0, b=1, c=-1, d=1.$$

几种常见的连续型分布

密度函数反映的是连续型随机变量取值的统计规律性。本质上不同的密度函数，对应着不同的随机变量。我们有如下一些常见的随机变量。

均匀分布 (Uniform Distribution)

定义2.9：若随机变量X 的密度函数为

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

则称X 在 $[a,b]$ 上服从均匀分布，记为 $X \sim U(a,b)$

其中a, b (a<b)为分布参数。

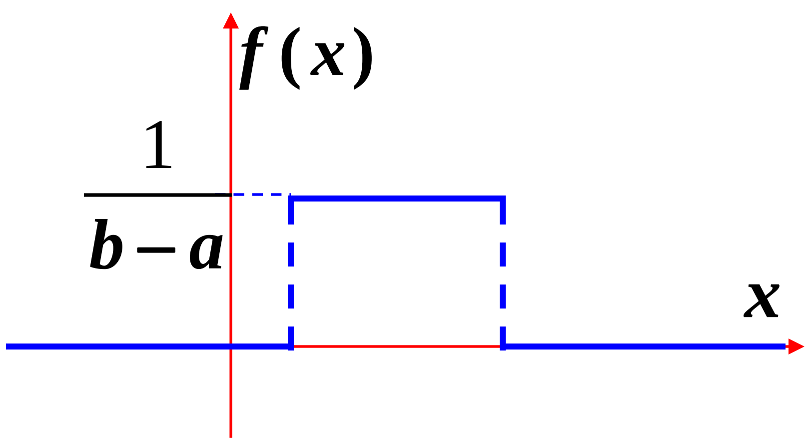
显然，若 $X \sim U(a,b)$ ，$[c,c+l]\subset[a,b]$ 是长度为l的子区间，则

$$P(c<X\leq c+l)=\int_c^{c+l}\frac{1}{b-a}dx=\frac{l}{b-a},$$

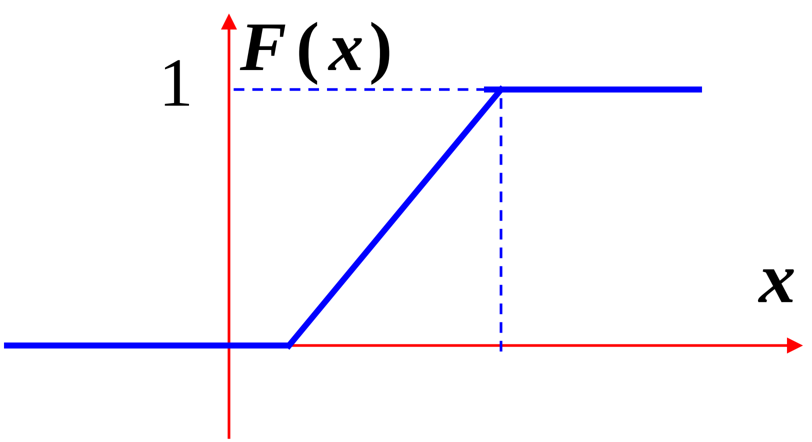
即X 落入$[a,b]$的任一子区间内的概率只与区间的长度相关且成正比，与位置无关。这正是前面所讲到的几何概型：在区间上的等可能分布，即在区间上的取值是均匀的，这就是它称为均匀分布的原因。

由分布函数的定义知，若 $X \sim U(a,b)$，则X的分布函数为

$$F(x)=\begin{cases} 0, & x<a, \\ \int_a^x \frac{1}{b-a}dx=\frac{x-a}{b-a}, & a\le x\le b, \\ 1, & x>b. \end{cases}$$



均匀分布的密度函数



均匀分布的分布函数

例：设公共汽车站从上午7时起每隔15分钟来一班车。如果乘客到达此站的时间为 7: 00—7: 30 之间的任意时刻。求该乘客候车时间不超过5分钟的概率。

解 设该乘客于7时X 分到达此站，则 $X \sim U(0,30)$,

密度函数为

$$f(x)=\begin{cases}\frac{1}{30}, & 0\le x\le 30,\\ 0, & else\end{cases}$$

则所求概率为

$$P(10\le X\le 15)+P(25\le X\le 30)$$

$$=\int_{10}^{15}\frac{1}{30}dx+\int_{25}^{30}\frac{1}{30}dx=\frac{1}{3}.$$

指数分布 X~ $e(\lambda)$

定义2.10 设随机变量X的密度函数为

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

则称X服从参数为 λ 的Exponential 分布

指数分布的分布函数为：

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

指数分布常见于元件的使用寿命，
随机服务系统的服务时间等。

例：某产品的寿命X(单位：年)服从参数为2的指数分布,试求(1)P(X≥s)=?　　(2)P(X≥s+t |X≥s)

解：X的密度函数 $f(x)=\begin{cases}2e^{-2x}, x>0\\ 0 \quad ,else\end{cases}$

（1）$P(X \geq s)=\int_{s}^{+\infty} 2e^{-2x}dx = e^{-2s}$

(2)P(X≥s+t |X≥s) $=\dfrac{P(X \geq s+t)}{P(X \geq s)}$

$=\dfrac{\int_{s+t}^{\infty} f(x)dx}{\int_{s}^{\infty} f(x)dx} = e^{-2t}$

$=\mathrm{P}(X \geq t)$ 与s无关。

若$X \sim e(\lambda)$,则

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

$= P(X > t).$

这叫指数分布的

无记忆性

Γ函数与Γ分布

定义2.11 Γ函数的定义：

$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx$$

$\Gamma(x)$称为具有以下重要性质：

1)对每一个$\alpha>0,\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$;

2)$\Gamma(1)=1,\Gamma(1/2)=\sqrt{\pi}$;

3)若n为正整数，则$\Gamma(n)=(n-1)!$

Γ分布

定义2.12 设随机变量X的密度函数为

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

则称随机变量X服从参数为 α,β 的Γ分布

记为X~Γ(α,β).

注意： α=1时，Γ分布为指数分布e(β),

即：Γ(1,β)=e(β).

例2.15 某厂生产的元件其寿命 $X \sim \Gamma(2, \frac{1}{2})$

1)随机取一个元件,求该元件寿命大于4万小时的概率;

2)随机取10个元件,求至少有1个元件寿命大于4万小时的概率;

解:由题意,X的密度函数为:

$$f(x)=\begin{cases}\frac{1}{4}x\,e^{-\frac{x}{2}}, x>0\\ 0, \qquad else.\end{cases}$$

$$1)P(X>4)$$

$$\boxed{\int_a^b u\,dv = uv\big|_a^b - \int_a^b v\,du}$$

$$=\int_4^{+\infty} f(x)dx$$

$$=\int_4^{+\infty}(-\frac{1}{2}x)d(e^{-\frac{x}{2}})$$

$$=-\frac{1}{2}xe^{-\frac{x}{2}}\Big|_4^{+\infty}-\int_4^{+\infty}e^{-\frac{x}{2}}d(-\frac{1}{2}x)=3e^{-2}\approx 0.406$$

2)每只元件寿命只有两种可能：“寿命大于4万小时”或“寿命小于等于于4万小时”。于是10只元件的寿命情况就可以看作10重贝努利实验，设Y表示10只元件中寿命大于4的只数,则Y~B(10,p)，p=P(X>4).

$$P(Y\geq 1)=1-P(Y=0)=1-0.594^{10}$$