第十一讲 分布函数

◆定义:

随机变量X ,对任意实数x, 称函数:

$$F(x) = P(X \le x)$$

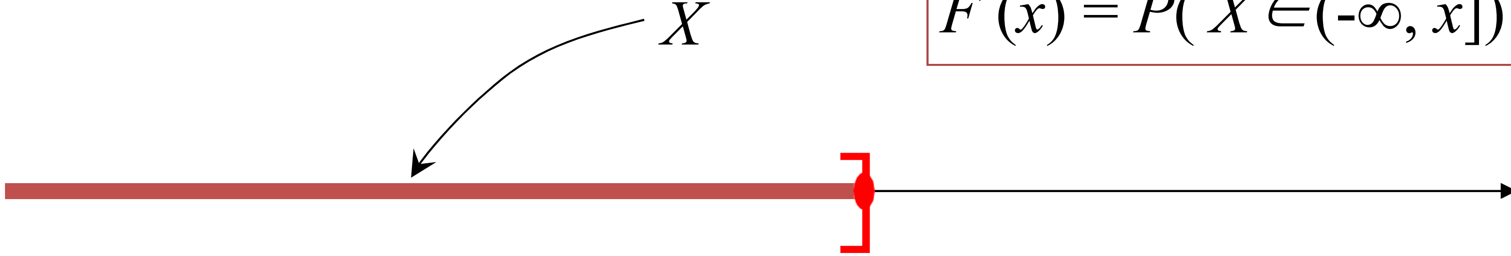
有时也写为

$$F_X(x) = P(X \le x)$$

为 X 的概率分布函数，简称**分布函数**.

- 任何随机变量都有相应的分布函数，所有的实数均有定义
- F (x)的几何意义

$$F(x) = P(X \in (-\infty, x])$$



若将X看成数轴上随机点的坐标则分布函数$F(x)$的值就表示X落在$(-\infty, x]$的概率 因而$0 \le F(x) \le 1$.

分布函数的用途:

可以给出随机变量落入任意一个范围的可能性

X的分布函数为$F(x)$,a, b为两个实数,a < b, 则

$$P(a<X\leqslant b)=P(X\leq b)-P(X\leq a)=F(b)-F(a);$$

$$P(a<X<b)=P(a<X\leq b-0)=F(b-0)-F(a);$$

$$P(X=b)=P(X\leq b)-P(X<b)=F(b)-F(b-0).$$

$$P(a<X<b)=P(a<X\leq b)-P(X=b)$$

$$P(a\leq X\leq b)$$

$$P(a\leq X<b)$$

分布函数的双重含义：首先具有概率含义，它完整描述随机变量的统计规律性； 其次是普通函数，可以用微积分的方法全面研究随机变量。

◆ 例1：

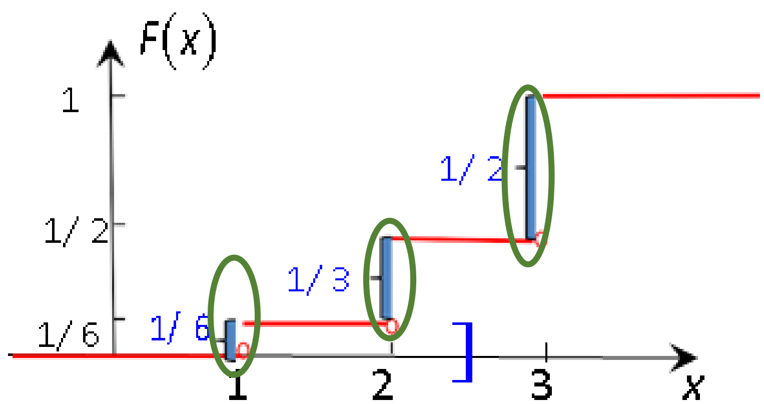
一罐子中有6个大小形状一致的球，球号分别为1，2，2，3，3，3。现随机摸一球（假设摸到各球的可能性相同），用X表示摸到球的球号。求X的分布函数。

解：由题意知，X的分布律为

X	1	2	3
P	1/6	1/3	1/2

那么X的分布函数为

$$F(x)=\begin{cases} 0\ , & x<1; \\ 1/6\ , & 1\le x<2 \\ 1/2\ , & 2\le x<3 \\ 1\ , & ,x\ge 3 \end{cases}$$



一般地，离散型随机变量的分布函数为阶梯函数.

设离散型随机变量X的分布律为P{X= x_k }= p_k , k = 1, 2,⋯

X的分布函数为F (x) =$\sum_{x_k \le x} p_k$

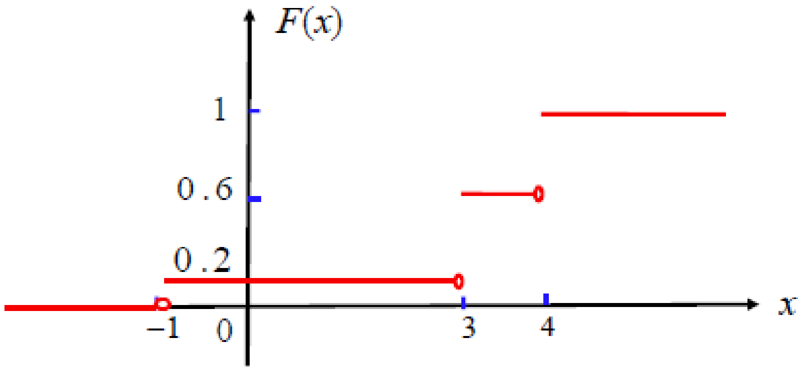
$F(x)$在x = x_k , (k = 1, 2⋯)处有跳跃，其跳跃值为p_k= $P\{X = x_k\}$.

要求：**由分布律推分布函数；反过来由分布函数推分布律。**

◆ 例2：

设随机变量X 的分布函数如下，求X的分布律。

$$F(x)=\begin{cases} 0\ , & x<-1; \\ 0.2\ , & 1\le x<3 \\ 0.6\ , & 3\le x<4 \\ 1\ , & x\ge 4 \end{cases}$$



解：$F(x)$只在-1,3,4有跳跃，跳的幅度分别是0.2,0.4,0.4.

故分布律为

X	-1	3	4
P	0.2	0.4	0.4

分布函数$F(x)$的性质：

(1) $0 \le F(x) \le 1$;

(2) $F(x)$单调不减;

$\because$ 对于任意$x_1 < x_2$,有 $0 \le P(x_1 < X \le x_2) = F(x_2) - F(x_1)$.

(3) $F(-\infty) = 0$, $F(+\infty) = 1$;

(4) $F(x)$是右连续函数, 即 $F(x+0) = F(x)$.

满足上述性质的实函数$F(x)$一定是某个随机变量X的分布函数.

◆ 例3：

设一醉汉游离于A, B两点间, A, B之间距离3个单位. 该醉汉落在A, B间任一子区间的概率与区间长度成正比. 设他在离A点距离X远处, 求X的分布函数。

解：根据题意，$P(0 \le X \le 3) = 1$,

故当$x<0$时, $F(x) = P(X \le x) = 0$; 当$x \ge 3$时, $F(x) = 1$;

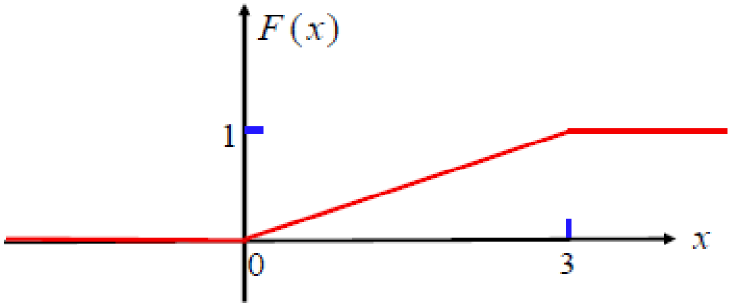
设比例系数为 k, 则 $P(0 \le X \le 3) = 3k = 1, \Rightarrow k = 1/3$.

而当$0 \le x < 3$时, $F(x) = P(X \le x) = P(X < 0) + P(0 \le X \le x) = 0 + \frac{1}{3} \cdot x$

从而X的分布函数为:

$$F(x) = P(X \le x) = \begin{cases} 0 \quad , & x < 0; \\ x/3 \quad , & 0 \le x < 3; \\ 1 \quad , & x \ge 3. \end{cases}$$

$$F(x)=P(X\le x)=\begin{cases}0 & , \quad x<0;\\ x/3 & ,0\le x<3;\\ 1 & , \quad x\ge 3.\end{cases}$$



从图可见：此分布函数是个连续函数，因此X不是离散型随机变量.

X是连续型随机变量.

THE END

第十二讲 连续型随机变量及其概率密度

◆ 定义：

对于随机变量X的分布函数F (x),若存在非负的函数f (x),使对于任意实数 x 有：

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

则称X为连续型随机变量,其中f (x)称为X的概率密度函数,简称概率密度.

有时也写为$f_X(x)$

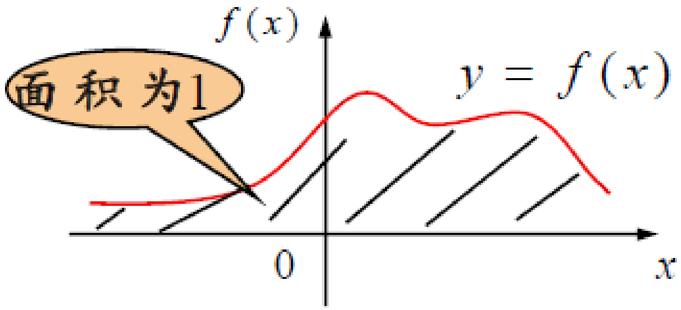
$f(x)$的性质：

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

(1) $f(x)\geq 0$;

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

$\because F(+\infty)=1$



f(x)的性质：

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

(3)对于任意的实数x_1 ,x_2 ($x_1 < x_2$)

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f(t)\,dt;$$

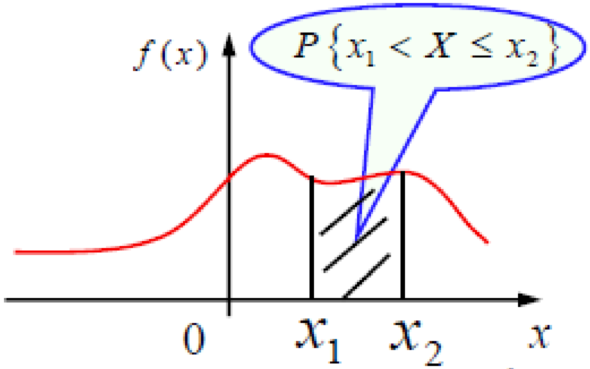
$$\because LHS = P(X \le x_2) - P(X \le x_1) = F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(t)dt - \int_{-\infty}^{x_1} f(t)dt$$

$\Rightarrow$ 对任意的实数 $a, P(X=a)=0.$

$\Rightarrow$ 且 $P(x1 < X \le x2) = P(x1 < X < x2)$

对于连续型的随机变量X ,有

$P(X \in D) = \int_D f(x)dx,$ **任意** $D \subset R.$



f(x)的性质:

(4) 在$f(x)$连续点x, $F'(x)=f(x)$.

即在f(x)的连续点,

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

$$P(x<X\le x+\Delta x)\approx f(x)\bullet\Delta x$$

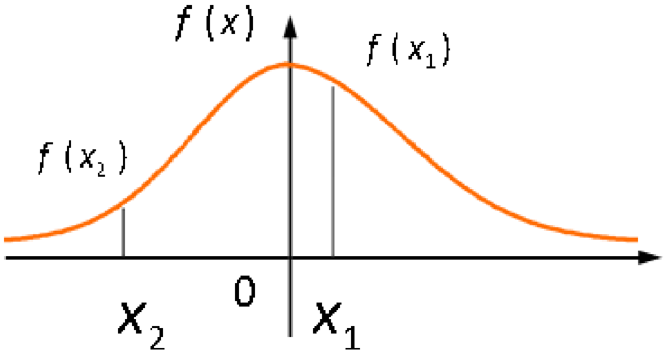
这表示X落在点 x 附近$(x, x+\Delta x]$的概率近似等于 $f(x)\Delta x$.

说明：

（1）$f(x)$值的含义；

当 Δx 充分小时，

$$P(x < X \le x + \Delta x) \approx f(x) \bullet \Delta x$$



$f(x_2) < f(x_1)$

落入的概率x_1点周围大于x_2

（2）$f(x)$的值是可以大于1的；

（3）$f(x) \underset{\frac{d}{dx}F(x)}{\overset{\int_{-\infty}^{x} f(t)dt}{\rightleftarrows}} F(x).$

密度函数与概率分布函数可互推

◆ 例1：

设X的概率密度为

$$f(x)=\begin{cases} cx+\dfrac{1}{6} & 0<x<2 \\ 0 & \text{其他} \end{cases}$$

求: (1)常数c的值; (2) X的概率分布函数$F(x)$; (3) $P(-1< X <1)$的值.

解： $(1)1=\int_{-\infty}^{+\infty}f(x)dx=\int_{-\infty}^{0}f(x)dx+\int_{0}^{2}f(x)dx+\int_{2}^{+\infty}f(x)dx$

$$=\int_{-\infty}^{0}0dx+\int_{0}^{2}(cx+\frac{1}{6})dx+\int_{2}^{+\infty}0dx=\underline{\int_{0}^{2}(cx+\frac{1}{6})dx}=(\frac{c}{2}x^2+\frac{1}{6}x)\Big|_0^2$$

$$=\frac{c}{2}\times 2^2+\frac{1}{6}\times 2 \quad \rightarrow c=\frac{1}{3}$$

$P\{X\in（0，2)\}=1$

$\{x: f(x)>0\}$
即为X的取值范围
（支撑）

$$f(x)=\begin{cases}\frac{x}{3}+\frac{1}{6} & 0<x<2;\\ 0 & 其他\end{cases}$$

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

注意到 $P\{X\in$（0，2）$\}=1$

1、当 $x<0$ 时，$F(x)=P\{X\le x\}=\int_{-\infty}^{x}0dt=0$

2、当 $x\ge 2$ 时，$(0,2)\subset$（$-\infty$，x], 故 $F(x)=P\{X\le x\}=1$;

3、当 $0\le x<2$ 时，$F(x)=P\{X\le x\}=\int_{-\infty}^{x}f(t)dt=\int_{-\infty}^{0}f(t)dt+\int_{0}^{x}f(t)dt$

$$=\int_{-\infty}^{0}0dt+\int_{0}^{x}(\frac{t}{3}+\frac{1}{6})dt=(\frac{t^2}{6}+\frac{t}{6})|_0^x=\frac{x^2}{6}+\frac{x}{6}$$

$$F(x)=\begin{cases}0\ , & x<0;\\ \frac{x^2}{6}+\frac{x}{6} & ,0\le x<2;\\ 1\ , & x\ge 2.\end{cases}$$

$$f(x)=\begin{cases} x/3+1/6, & 0<x<2; \\ 0, & \text{其他}. \end{cases}$$

$$(3)\ P(-1<X<1) = \int_{-1}^{1} f(x)dx$$
$$= \int_{-1}^{0} f(x)dx + \int_{0}^{1} f(x)dx$$
$$= \int_{-1}^{0} 0dx + \int_{0}^{1} (\frac{x}{3}+\frac{1}{6})dx = 0+(\frac{x^2}{6}+\frac{x}{6})\Big|_0^1 = \frac{1}{3}.$$

$$F(x)=\begin{cases} 0, & x<0; \\ \frac{x^2}{6}+\frac{x}{6}, & 0\le x<2; \\ 1, & x\ge 2. \end{cases}$$

或 $P(-1<X<1)=F(1)-F(-1)$

$$=\frac{1^2}{6}+\frac{1}{6}-0=\frac{1}{3}.$$

◆ **例3：**

在第11讲中，得到X的分布函数为 $F(x)=\begin{cases}0, & x<0;\\ \dfrac{x}{3}, & 0\le x<3;\\ 1, & x\ge 3.\end{cases}$

求 X 的概率密度函数 $f(x)$.

解： 由于 $f(x)=F'(x)$,

故 X 的概率密度 $f(x)=\begin{cases}\dfrac{1}{3} & 0<x<3\\ 0 & \text{其他}\end{cases}$

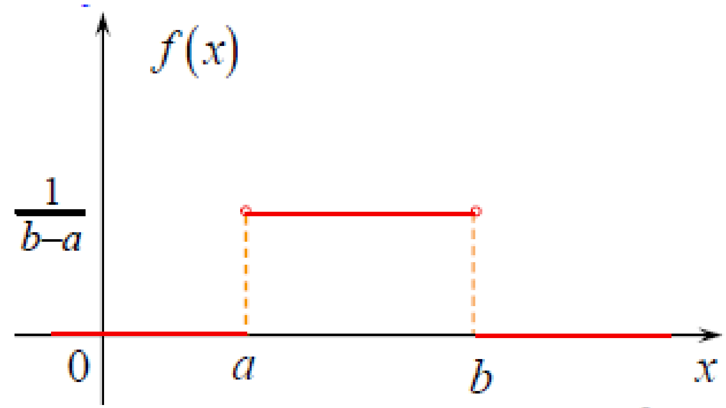
THE END

第十三讲均匀分布和指数分布

◆ 均匀分布的定义：

若X的概率密度函数为 $f(x)=\begin{cases}\dfrac{1}{b-a} & x\epsilon(a,b)\\ 0 & \text{其他}\end{cases}$

其中a<b，就称X服从（a，b）上的均匀分布（Uniform），记为X~U（a，b）或Y~Unif（a，b）.



其实 $f(x)=\begin{cases}c & x\epsilon(a,b)\\ 0 & \text{其他}\end{cases}$

$\because\int_{-\infty}^{+\infty}f(x)dx=1$，即$\int_a^b c\,dx=1 \Rightarrow c=1/(b-a)$

◆性质：均匀分布具有等可能性

即，对于任意的$a<k<k+l<b$，

均有$P（k<X<k+l）=\int_k^{k+l}\frac{1}{b-a}dt=\frac{l}{b-a}$——与$k$无关，仅与$l$有关.

即，服从U（a，b）上的均匀分布的随机变量X落入（a，b）中的任意子区间上的概率只与其区间长度有关与区间所处的位置无关.

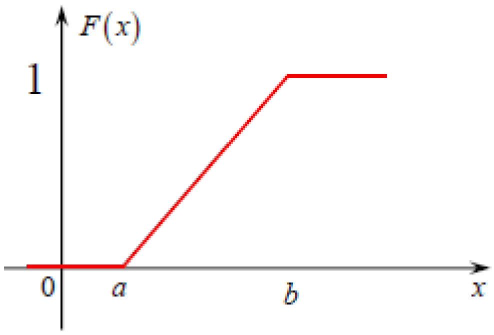
即，X落入（a，b）中的**等长度**的任意子区间上是**等可能**的.

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

且分布函数为

$\because$ 当 $a \le x < b$ 时, $F(x)=\int_{-\infty}^{x} f(t)dt = \int_{a}^{x} \frac{1}{b-a} dt = \frac{x-a}{b-a}$

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



◆ **例1：**

在区间（-1，2）上随机取一数X，求：（1）试写出X的概率密度函数;（2）该数在（-0.5，1）中的概率；（3）该数为正数的概率.

解：(1) X应在区间（-1，2）服从均匀分布 故的概率密度函数为

$$f(x)=\begin{cases}\frac{1}{3} & x\epsilon(-1,2)\\ 0 & \text{其他}\end{cases}$$

(-0.5，1)的长度

(2) $P(-0.5<X<1)=\int_{-0.5}^{1} f(\mathrm{x})dx=\int_{-0.5}^{1}\frac{1}{3}dx=\frac{1-(-0.5)}{3}=\frac{1}{2}.$

(0，+∞)∩(-1,2)=(0,2)的长度

(3) $P(X>0)=\int_{0}^{+\infty} f(\mathrm{x})dx=\int_{0}^{2}\frac{1}{3}dx=\frac{2-0}{3}=\frac{2}{3}.$

(-1,2)的长度

◆均匀分布的概率计算

若X-U（a，b），则对于$\forall I\subset R$，有

法一：$P(X\in I)=\int_I f(x)\,dx$

法二：$P(X\in I)=\dfrac{I\cap(a,b)\text{的长度}}{(a,b)\text{的长度}}$

指数分布

◆ 指数分布的概率计算

若X的概率密度函数为
$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta} & x>0\\ 0 & x\le 0\end{cases}$$

若其中$\theta>0$,为常数，就称X服从参数为λ的指数分布*(Exponential)*,

记为$X\sim E(\lambda)$ 或$X\sim Exp(\lambda)$.

分布函数 $$F(x)=\begin{cases}1-e^{-x/\theta} & x>0\\ 0 & x\le 0\end{cases}$$

◆性质：指数分布具有无记忆性

分布函数 $F(x) = \begin{cases} 1 - e^{-x/\theta} & x > 0 \\ 0 & x \le 0 \end{cases}$

对于 $t_0 > 0,\ t > 0,$

$$P(X > t_0 + t \mid X > t_0) = \frac{P(X > t_0 + t \cap X > t_0)}{P(X > t_0)}$$

$$= \frac{P(X > t_0 + t)}{P(X > t_0)} = \frac{1 - F(t_0 + t)}{1 - F(t_0)}$$

$$= \frac{e^{-(t_0+t)/\theta}}{e^{-t_0/\theta}} = e^{-t/\theta}$$

$$= P(X > t)$$

中山大學 SUN YAT-SEN UNIVERSITY

◆ **例2：**

设某人电话通话时间X（分钟）服从指数分布，概率密度为

$$f(x)=\begin{cases}\frac{1}{15}e^{\frac{-x}{15}} & x>0\\ 0 & x\le 0\end{cases}$$

求：

(1) 她的通话时间在10～20分钟之间的概率；

(2) 若她已打了10分钟，求她继续通话超过15分钟的概率.

(即，若她已打了10分钟，求她总共通话超过25分钟的概率.)

解：

(1) $P(10 < X < 20) = \int_{10}^{20} f(\mathrm{x})dx = \frac{1}{15}\int_{10}^{20} e^{\frac{-x}{15}}dx = e^{\frac{-2}{3}} - e^{\frac{-4}{3}}$

或利用分布函数，

$P(10 < X < 20) = F(20) - F(10) = 1 - e^{\frac{-20}{15}}\left(1 - e^{\frac{-10}{15}}\right) = e^{\frac{-2}{3}} - e^{\frac{-4}{3}}$

$$f(x) = \begin{cases} \frac{1}{15}e^{-\frac{x}{15}}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{15}}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

（2）若她已打了10分钟，求她继续通话超过15分钟的概率

(3)根据无记忆性

$P(\mathrm{X} > 25 \mid X > 10) = P(X > 15) = \int_{15}^{+\infty} \frac{1}{15} e^{\frac{-x}{15}}dx = e^{-1}.$

◆ 指数分布的用途：

- ◆ 指数分布可以用来表示独立随机事件发生的时间间隔，比如旅客进机场的时间间隔、中文维基百科新条目出现的时间间隔等等;
- ◆ 在排队论中，一个顾客接受服务的时间长短也可用指数分布来近似;
- ◆ 无记忆性的现象(连续时).
- ◆ 已知在已过去的13小时中没有发生交通事故, 那么在未来的2小时内不发生交通事故的概率

THE END