第二章 随机变量及其分布

§2.1　随机变量及其分布函数

2.1.1 随机变量的引入和定义

在实际问题中，随机试验的结果可以用数量来表示，由此就产生了随机变量的概念.

(1) 有些试验结果本身就是数，随机事件与实数之间存在着客观联系.

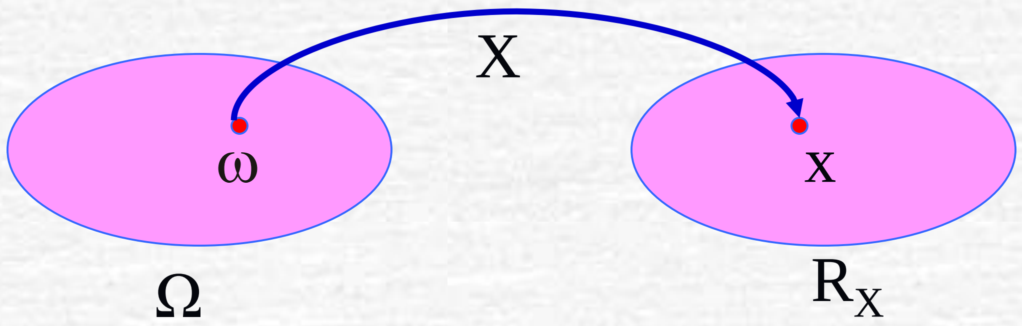
Back

例1：考虑下面的试验E：在区间[0,1]上任取一点，记录该点的坐标，试验E的样本空间$\Omega=\{\omega|0\leq\omega\leq1\}$,每个样本点ω是[0,1]上的一个数字。

若令：X表示“在[0,1]上任取一点的坐标”,则

I：它是在[0,1]上取值的一个变量，而且它的取值依赖于试验结果ω，这种依赖关系可用一个样本点ω**的函数**来表达:$X=X(\omega)$, $\omega\in\Omega$.

如：当$\omega=1/2$时，$X=X(1/2)=1/2$.



II：$\forall\, x \in (-\infty, +\infty)$，$\{X \le x\} = \{\omega \mid X(\omega) \le x\}$是一个事件，因而可求出其概率。

如 $x = -1$, $P\{X \le -1\} = P\{\omega \mid X(\omega) \le -1\} = P(\Phi) = 0$.

$x = 2/3$, $P\{X \le 2/3\} = P\{\omega \mid X(\omega) \le 2/3\} = 2/3$.

$x = 1$, $P\{X \le 1\} = P\{\omega \mid X(\omega) \le 1\} = P(\Omega) = 1$.

（2）有些试验中，试验结果看来与数值无关，随机事件与实数之间没有客观联系，但是我们可以引进一个变量来表示它的各种结果.也就是说，把试验结果数值化.

例2：系球队参加学校比赛， 实验E为：记录一场比赛结果。试验E的样本空间Ω={“胜” “平” “负”}，若设各结果对应的分数为Ω={“胜” “平” “负”}=**{2分，1分，0分}**，令

X表示“该队参加一场比赛的分数”

Ⅰ. **X是变量**

$$X = X(\omega) = \begin{cases} 2 & \omega = \omega_1 \\ 1 & \omega = \omega_2 \\ 0 & \omega = \omega_3 \end{cases} \qquad \omega \in \Omega$$

若根据以往的记录，该队参加一场比赛赢球的概率1/2，输及平的概率1/4，于是X的取值有概率规律。

P{X=2}= **P ({ω_1})** =1/2 ,

P{X=1}= **P ({ω_2})**= 1/4,

P{X=0}= **P ({ω_3})**= 1/4.

Back

II． $\forall x \in (-\infty, +\infty)$,$\{X \le x\} = \{\omega | X(\omega) \le x\}$是一个事件。

如：

当x=0.3时，$P\{X \le 0.3\} = P\{\omega | X(\omega) \le 0.3\}$

$$= P\{X=0\} = P(\{\omega_3\}) = 1/4.$$

当x=1时,$P\{X \le 1\} = P(\{\omega_2, \omega_3\}) = 1/4 + 1/4 = 1/2$

定义：设$X(\omega)$是定义在样本空间$\Omega=\{\omega\}$上的单值实函数，如果对于每个$\omega\in\Omega$,有一个实数$X(\omega)$与之对应, 且$\forall \boldsymbol{x}\in(-\infty, +\infty)$, $\{\omega|X(\omega) \le \boldsymbol{x}\}$是一个事件,则称$X(\omega)$为**随机变量**。

一般地，$X(\omega)$写为X，$\{X\le \boldsymbol{x}\}=\{\omega|X(\omega) \le \boldsymbol{x}\}$。

随机变量常用大写字母X,Y,Z,U,V,W……表示，或写成r.v.X 。

随 机 变 量

Back

例1：考察“抛硬币”这一试验，它有两种可能结果：

“出现H”或“出现T”。我们用数字“1”代表“出现H”，用数“0”代表“出现T”。

引入一个变量X，X与试验结果的关系为

$$X = X(\omega) = \begin{cases} 0 & \omega = T \\ 1 & \omega = H \end{cases}$$

因此X是一个定义在Ω=｛H，T｝上的单值实值函数，X是一个随机变量。而且｛X＝1｝表示事件“出现正面”，反之，事件“出现正面”可用｛X＝1｝来表示。X的值域$R_X=\{0,1\}$.

例2：考虑“测试灯泡寿命”这一试验。试验结果本身是用数字描述的，令X表示灯泡的寿命（以小时计），这个变量X是定义在样本空间 $\Omega=\{t|t\geqslant 0\}$上的函数，具体地说就是

$$X=X(\omega)=t,\quad \omega=t\in\Omega$$

X是随机变量。它的值域为$R_X=[0, +\infty)$，而且$\{X<500\}$表示事件“任取出的灯泡的寿命小于500小时”，事件“任取的灯泡的寿命大于500小时且不超过1000小时”可用X表示为$\{500<X\leq 1000\}$。

注释：

（1）随机变量X是一个从样本空间到实数空间的**函数**，它的定义域为样本空间Ω。它的值域Rx为全体实数集或它的一个子集。

（2）从随机变量的定义来看，它与通常的函数概念没有什么不同，把握这个概念的关键之点在于试验前后之分：**在试验前**，我们不能预知它取何值，这要凭机会，“**随机**”的意思就在这里，一旦试验完成后，它的取值就确定了。

（3）随机变量与事件的关系；

1，随机变量的取值是随机的，“它取一个值”或“它取值于某一给定的区间”为随机事件。

2，当用某个随机变量来描述某一随机现象时，随机事件能用随机变量来表示。

随机事件↔随机变量的值

试验的全部结果 → 随机变量

随机事件这个概念实际上是包容在随机变量这个更广的概念内.

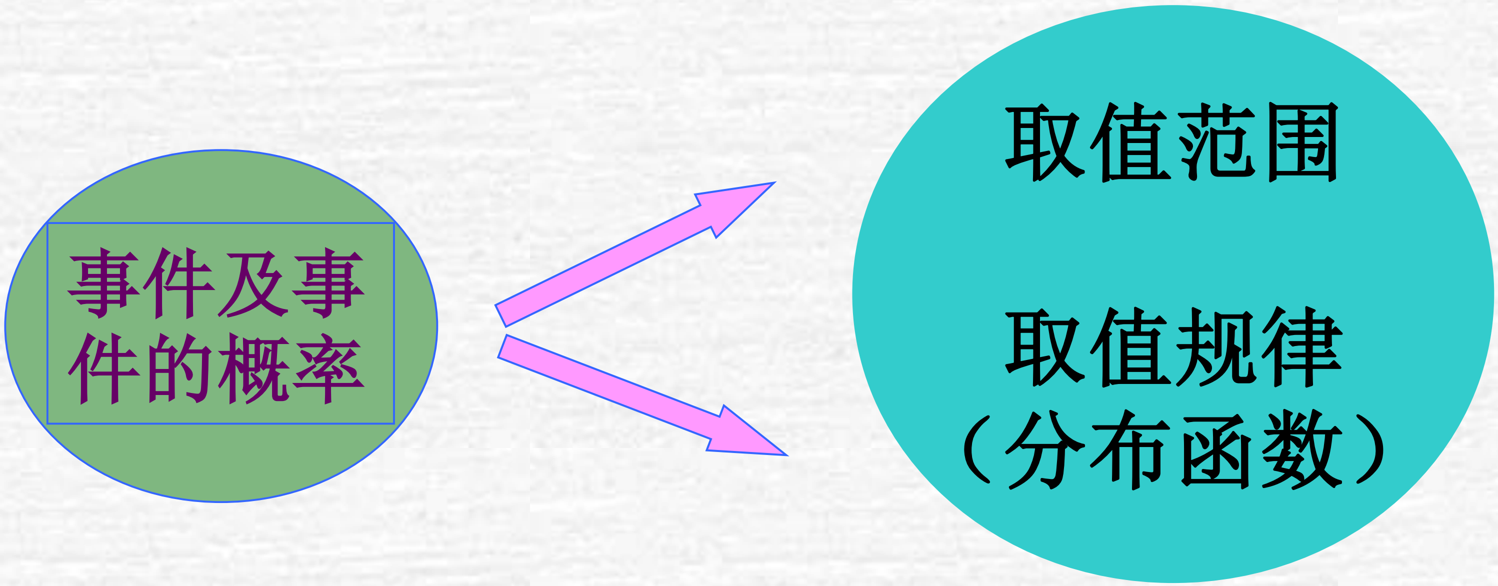
随机事件是从**静态**的观点来研究随机现象，而随机变量则是一种**动态**的观点，就象数分中常量与变量的区别.

随机变量概念的产生是概率论发展史上的重大事件. 引入随机变量后，对随机现象统计规律的研究，就由对事件及事件概率的研究*扩大为对随机变量及其取值规律*的研究.

对于r.v.X更重要的是搞清：

（I）它的取值范围；（II）取值的概率规律。

（试验结果）　　　　　　（事件的概率）



通过研究随机变量整体把握随机事件。

Back

2.1.2 r.v.的分布函数及其性质

通常我们称一个r.v.X取值的概率规律，为X的分布，下面就来介绍描述r.v.分布的一种方法 — r.v.的分布函数。

1.定义：设X为r.v.，$\forall x \in (-\infty, +\infty)$,称函数

$F(x)=P\{X \leqslant x\}$为X的分布函数，可简记为d.f.

Back

例如（1）$P\{a<X\le b\}=F(b)-F(a)$.

(因为$\{a<X\le b\}=\{X\le b\}-\{X\le a\}$,且$\{X\le b\}\supset\{X\le a\}$)

（2）$P\{X=b\}=F(b)-F(b\text{-}0)$. (由（1）连续性证明)

（3）$P\{X<b\}=F(b\text{-}0)$.

($\{X<b\}=\{X\le b\}-\{X=b\}$,$\{X\le b\}\supset\{X=b\}$,再由（2）)

（4）$P\{X>b\}=1-F(b)$. ($\{X>b\}=\overline{\{X\le b\}}$)

（5）$P\{X\ge b\}=1-F(b\text{-}0)$. (因为$\{X\ge b\}=\overline{\{X<b\}}$)

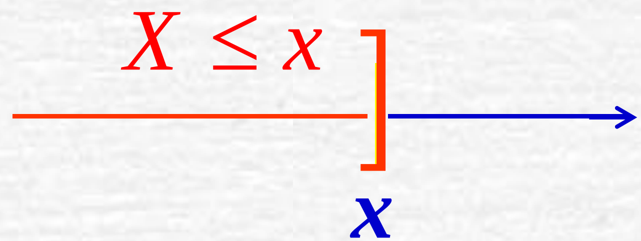
（6）$P\{a<X<b\}=F(b\text{-}0)-F(a)$.

◄Back

注释:

（1）由分布函数的定义知，分布函数$F(x)$在x处的函数值是事件$F(x)=P\{X\le x\}=P(\{\omega\,|X(\omega)\le x\}$的概率。

若把X看成数轴上的随机点的坐标，那么分布函数$F(x)$在x处的函数值，就是**表示**X落在区间$(-\infty, x]$上的概率。



（2）若已知x的分布函数$F(x)$，我们就知道了X落在任一区间的概率，从这个意义上讲，分布函数**完整地描述了随机变量的统计规律。**

（3）d.f.是在$(-\infty, +\infty)$上值域为 [0，1]的普通函数，它具有良好的分析性质，许多概率论的问题归结为函数的运算从而**利用数字分析**出许多结果，这是引入分布函数的好处之一，再加上d.f.对任意随机变量都有定义，因此d.f.在理论上有极重要的地位。

例1. 设X=0 P=1/4; X=1 P=1/4; X=2 P=1/2:

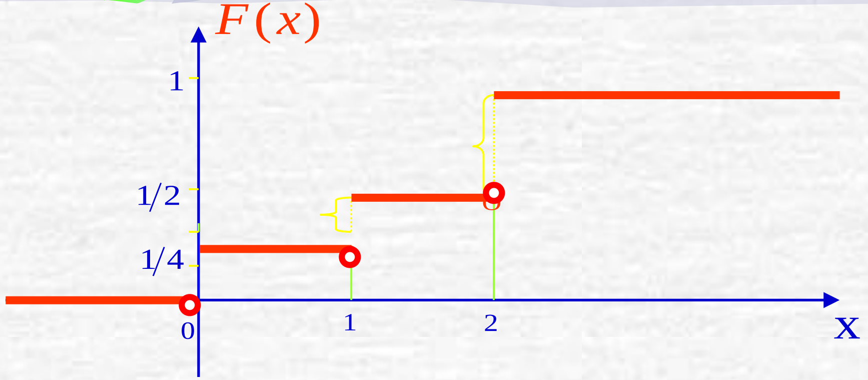
解: 当$x<0$时, $F(x)=P\{X\le x\}=0$;

当$0\le x<1$时, $F(x)= P\{X\le x\}=P\{X=0\}=1/4$;

当$1\le x<2$时, $F(x)= P\{X\le x\}=P\{X=0\}+P\{X=1\}=1/2$;

当$x\geqslant 2$时, $F(x)=P\{X\le x\}= P\{X=0\}+P\{X=1\}+P\{X=2\}=1$.

$$\therefore F(x)=\begin{cases} 0 & x<0 \\ \dfrac{1}{4} & 0\le x<1 \\ \dfrac{1}{2} & 1\le x<2 \\ 1 & x\ge 2 \end{cases}$$



*X*是随机变量, *x*是自变量.

F(*x*) 是r.v *X*取值不大于 *x* 的概率.

◀Back

例2 在区间 $[0, a]$ 上任意投掷一个质点，以 X 表示这个质点的坐标. 设这个质点落在 $[0, a]$中任意小区间内的概率与这个小区间的长度成正比，试求 X 的分布函数.

解： 设 $F(x)$ 为 X 的分布函数，

当 $x<0$ 时，$F(x) = P(X \leq x) = 0$

当 $x > a$ 时，$F(x) = 1$

当 $0 \le x \le a$ 时， $P(0 \le X \le x) = kx$ （k为常数）

由于 $P(0 \le X \le a) = 1$ $ka=1$，$k = 1/a$

$F(x) = P(X \le x) = P(X<0) + P(0 \le X \le x) = x / a$

$$\therefore \quad F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{a}, & 0 \le x \le a, \\ 1, & x > a. \end{cases}$$

Back

2. 分布函数的性质：

（1） $F(x)$为不减函数：当$x_1 < x_2$时，有$F(x_1) \leqslant F(x_2)$。

（2） $0 \le F(x) \le 1$, 且 $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to +\infty} F(x) = 1$.

（3）$F(x)$右连续，即对$\forall x$: $F(x)=F(x+0)$.

证：(1) $\because \forall x_1 < x_2, \{X \le x_1\} \subset \{X \le x_2\}$.

$\therefore P\{X \le x_1\} \le P\{X \le x_2\}$, 即$F(x_1) \le F(x_2)$

(2) 由定义可知：$0 \leqslant F(x) \leqslant 1$, 后两式与性质(3)都可用概率连续性证明。

(3) 即要证$\forall x_0$, $F(x_0+0)=F(x_0)$,

即证：$F(x_0+0) = \lim\limits_{x \to x_0+0} F(x) = F(x_0)$.

小结：

1.随机变量是概率论中最重要的概念之一，用随机变量描述随机现象是近代概率中最重要的方法。

2. 分布函数完整的描述了随机变量。

分布函数是在$(-\infty,+\infty)$上值域为 [0，1]的普通函数，它具有良好的分析性质，反之，若任意一个实值函数满足以上3条性质，则该函数一定是一个分布函数。

分布函数是研究随机变量的重要工具。