随机变量及其分布

第二章

随机变量概念的产生

上一章中，随机试验的结果用基本事件的集合表示。

局限性

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.

全面性

为更好地揭示随机现象的规律性并利用数学工具描述其规律，有必要引入随机变量来描述随机试验的不同结果.

1、有些试验结果本身与数值有关（本身就是一个数）.

例如，掷一颗骰子面上出现的点数；

七月份济南的最高温度；

电脑的使用寿命……

2、在有些试验中，试验结果看来与数值无关，但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 检测一件产品可能出现的两个结果，也可以用一个离散变量来描述

$$X(\omega)=\begin{cases}1, & \text{次品}\\ 0, & \text{正品}\end{cases}$$

这种对应关系 随 机 变 量

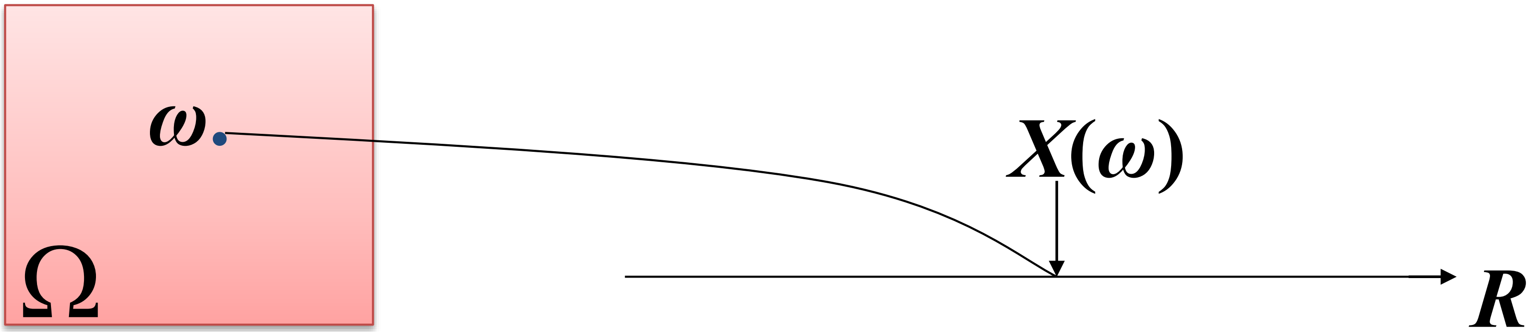
§2.1 随机变量及其分布函数

一.随机变量（random variable）

定义 设 Ω 是试验E的样本空间，若

$$\forall \omega \in \Omega \xrightarrow{\text{按一定法则}} \exists \text{ 唯一实数 } X(\omega)$$

则称 $X(\omega)$ 为 Ω上的 随机变量



随机变量 是 $\Omega \to R$ 上的映射mapping,

此映射具有如下特点

- <u>定义域</u>　事件域Ω
- <u>随机性</u>　r.v. X 的可能取值x不止一个，试验前只能预知它的可能的取值，但不能预知取哪个值
- <u>概率特性</u>　X 以一定的概率取某个值x

随机变量通常用大写字母
X,Y,Z或希腊字母ξ(ksi)，η(eta)，ζ (zeta)等表示

而表示随机变量所取的值(instantiation)时,一般采用小写字母x,y,z等.

引入随机变量的意义

(1) 任何随机现象可被 r.v.描述

有了随机变量，随机试验中的各种事件，就可以通过随机变量的关系式表达出来.

如：单位时间内某电话交换台收到的呼叫次数用X表示，它是一个随机变量.

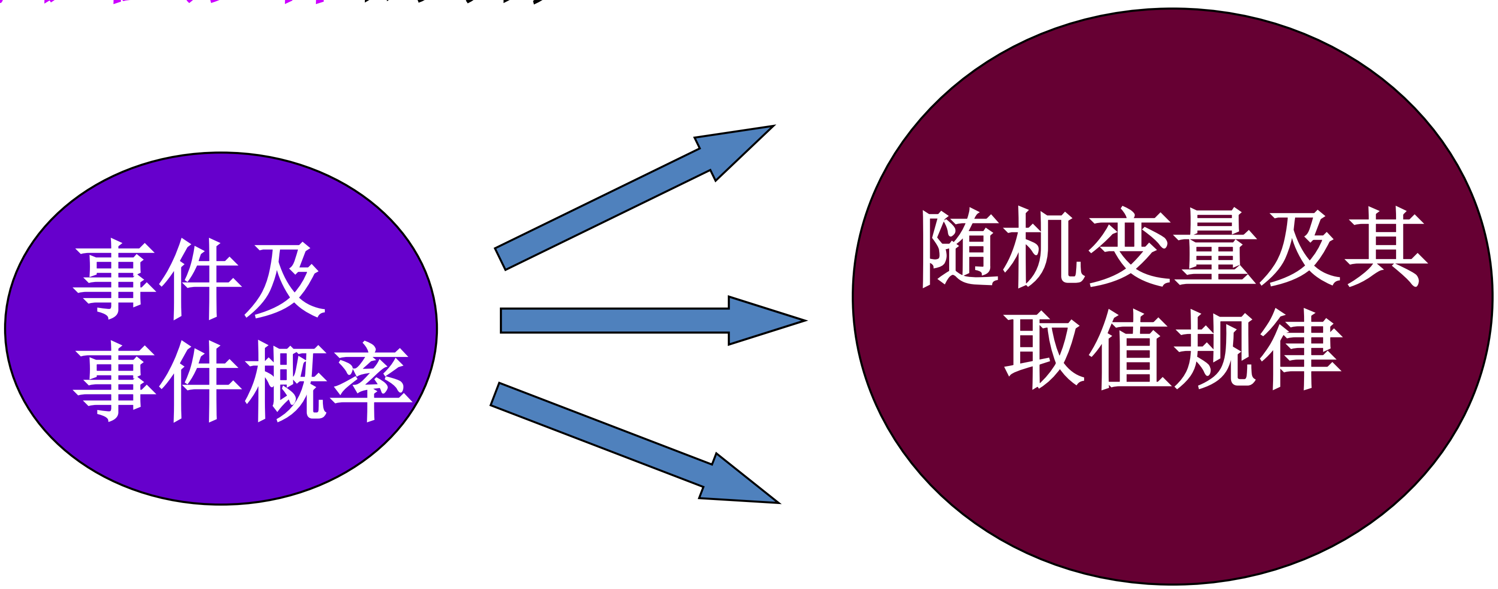
事件{收到不少于1次呼叫} $\Leftrightarrow$ $\{X \geq 1\}$

{没有收到呼叫} $\Leftrightarrow$ $\{X = 0\}$

(2) 借助微积分方法研究规律

可见，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. 也可以说，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，就象数学分析中常量与变量的区别那样.

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.



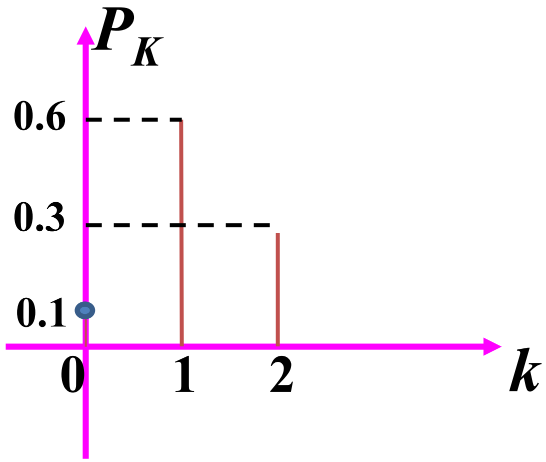
随机变量的分类

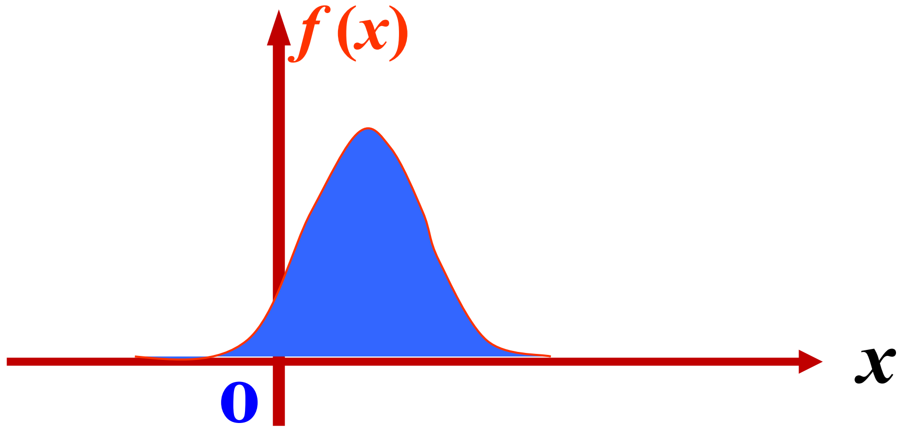
随机变量
- 离散型 Discrete r.v.
- 非离散型 -----连续型Continuous r.v.

所有取值可以逐个一一列举

有无穷多取值不能一一列举充满一个区间

为了对离散型的和连续型的 *r.v.* 以及更广泛类型的*r.v.*给出一种统一的描述方法，引进了**分布函数**的概念.





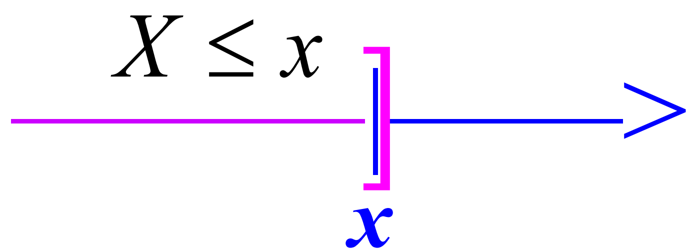
二. 随机变量的分布函数

Distribution Function/Cumulative Distribution Function

定义 设 X 为 $r.v.$, x 是任一实数,称函数

$$F(x) = P(X \le x), \quad -\infty < x < +\infty$$

为 X 的**概率分布函数**，简称**分布函数**.



如果将 X 看作数轴上随机点的坐标，那么分布函数 $F(x)$ 的值就表示 X落在区间 $(-\infty, x]$ 的概率.

用分布函数计算 X 落在 $(a, b]$ 里的概率：

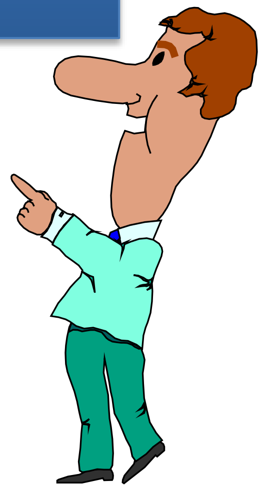
$$P(a < X \le b) = P(X \le b) - P(X \le a)$$

$$= F(b) - F(a)$$

因此，只要知道了随机变量X的分布函数，它的统计特性就可以得到全面的描述.

$$F(x) = P(X \le x), \quad -\infty < x < \infty$$

分布函数是一个普通的函数，正是通过它，我们可以用数学分析的工具来研究随机变量.



连续型的随机变量取值在任意一点的概率都是0，区间是开区间还是闭区间无关。但是对应事件并不是不可能事件。

分布函数的性质

(1) $F(x)$ 单调不减non-decreasing，即

$$\forall\ x_1 < x_2,\ \ F(x_1) \le F(x_2)$$

(2) $F(x)$ 归一化normalized, $0 \le F(x) \le 1$

$$\lim_{x\to+\infty} F(x) = 1,\ \ \lim_{x\to-\infty} F(x) = 0$$

(3) $F(x)$ 右连续right-continuous，即

$$\lim_{x\to x_0^+} F(x) = F(x_0)$$

如果一个函数具有上述性质，则一定是某个*r.v.* *X* 的分布函数. 也就是说，性质(1)--(3)是鉴别一个函数是否是某*r.v.*的分布函数的充分必要条件.

例 设随机变量 X 的分布函数为：

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

(1)$P\left(X\le\frac{1}{2}\right)$?

(2)$P\left(\frac{1}{2}<X\le\frac{3}{2}\right)$?

(3)$P\left(X>\frac{3}{2}\right)$?

$$(1)P\left(X\le\frac{1}{2}\right)=F\left(\frac{1}{2}\right)=\frac{1}{6}$$

$$(2)P\left(\frac{1}{2}<X\le\frac{3}{2}\right)=F\left(\frac{3}{2}\right)-F\left(\frac{1}{2}\right)=\frac{3}{4}-\frac{1}{6}=\frac{7}{12}$$

$$(3)P\left(X>\frac{3}{2}\right)=1-F\left(\frac{3}{2}\right)=1-\frac{3}{4}=\frac{1}{4}$$

例 设随机变量 *X* 的分布函数为

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

其中 $\lambda>0$ 为常数，求常数 *a* 与 *b* 的值。

解： 由分布函数的性质 $F(+\infty)=1$，可得：

$$F(+\infty)=\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}(a+be^{-\lambda x})=a=1$$

即 $a=1$. 由 $F(x)$的右连续性，可得

$$\lim_{x\to0^+}F(x)=\lim_{x\to0^+}(a+be^{-\lambda x})=a+b=F(0)=0$$

即 $b=-1$。