一、样本空间　样本点

定义　随机试验 E 的**所有可能结果**组成的集合称为 E 的样本空间, 记为 S .

样本空间的元素 , 即试验E 的每一个结果, 称为样本点.

实例1　抛掷一枚硬币,观察字面,花面出现的情况.

$H \rightarrow$ 字面朝上

$T \rightarrow$ 花面朝上

$S_1 = \{H, T\}.$

课堂练习

写出下列随机试验的样本空间.

生产产品直到得到10件正品,记录生产产品的总件数.

答案 $S=\{10,11,12,\cdots\}.$

一、随机变量的引入

1. 为什么引入随机变量?

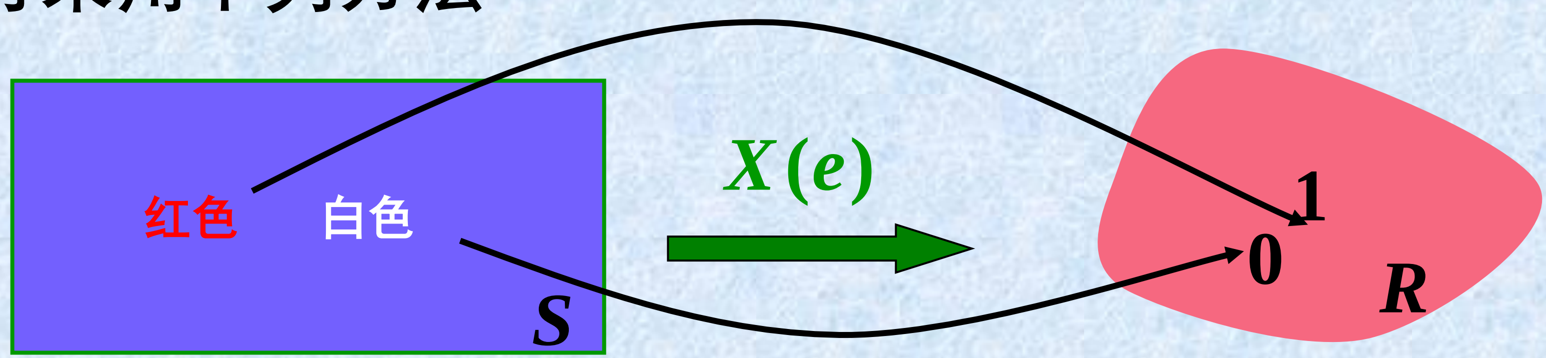
概率论是从数量上来研究随机现象内在规律性的，为了更方便有力的研究随机现象，就要用数学分析的方法来研究，因此为了便于数学上的推导和计算，就需将任意的随机事件数量化．**当把一些非数量表示的随机事件用数字来表示时，就建立起了随机变量的概念．**

2. 随机变量的引入

实例1　在一装有红球、白球的袋中任摸一个球，观察摸出球的颜色.

$S=\{\underbrace{\text{红色、白色}}_{\text{非数量}}\}\ \xrightarrow{\ ?\ }\ $ 将 S 数量化

可采用下列方法



即有　　X (红色)=1，　X (白色)=0.

$$X(e)=\begin{cases}1, & e=\text{红色},\\ 0, & e=\text{白色}.\end{cases}$$

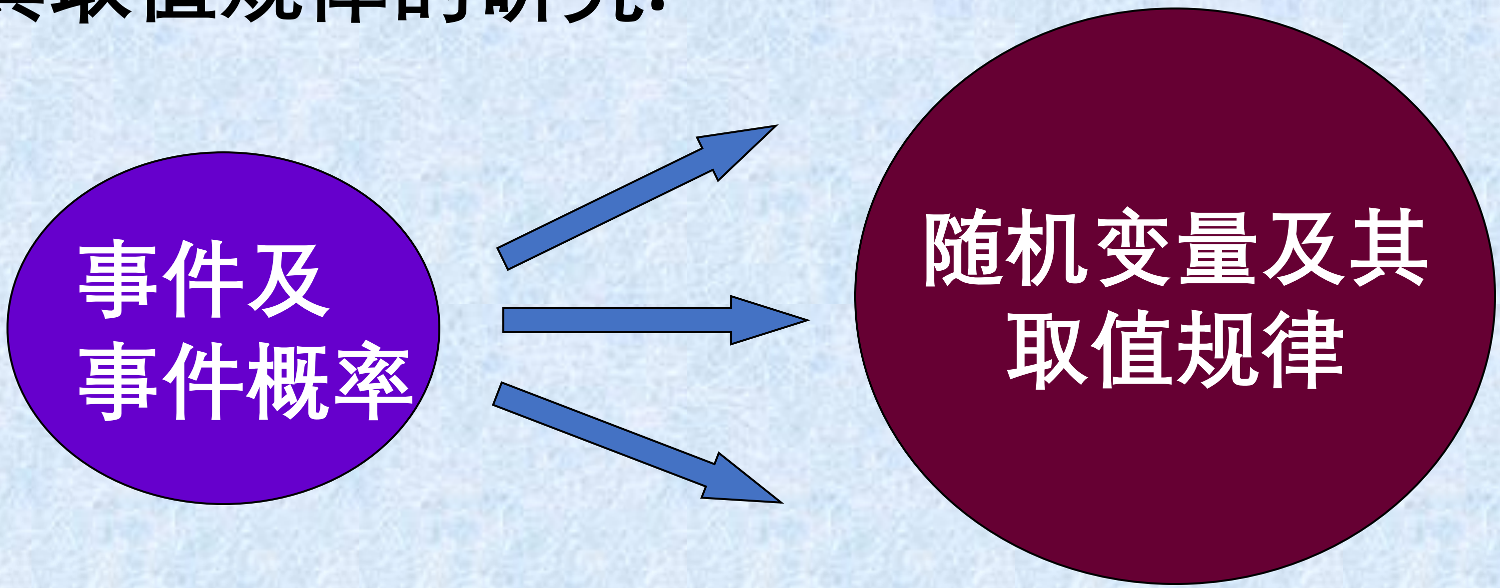
这样便将非数量的 S={红色，白色} 数量化了.

二、随机变量的概念

1.定义

设 E 是随机试验，它的样本空间是 $S=\{e\}$. 如果对于每一个 $e\in S$, 有一个实数 $X(e)$ 与之对应，这样就得到一个定义在 S 上的单值实值函数 $X(e)$, 称 $X(e)$ 为随机变量.

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.



3.随机变量的分类

- 随机变量
  - 离散型
  - 非离散型
    - 连续型
    - 其它

(1)离散型　随机变量所取的可能值是有限多个或无限可列个, 叫做离散型随机变量.

(2)连续型　随机变量所取的可能值可以连续地充满某个区间,叫做连续型随机变量.

一、离散型随机变量的分布律

定义

设离散型随机变量 X 所有可能取的值为 $x_k\ (k=1,2,\cdots)$, X 取各个可能值的概率, 即事件 $\{X=x_k\}$ 的概率, 为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots.$$

称此为离散型随机变量 X 的分布律.

说明

(1) $p_k \geq 0,\quad k=1,2,\cdots;$

(2) $\sum_{k=1}^{\infty} p_k = 1.$

用这两条性质判断一个函数是否是概率函数

离散型随机变量的分布律也可表示为

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

一、分布函数的概念

1.概念的引入

对于随机变量X, 我们不仅要知道X 取哪些值, 要知道 X 取这些值的概率 ; 而且更重要的是想知道 X 在任意有限区间(a,b)内取值的概率.

例如 求随机变量 X 落在区间$(x_1, x_2]$内的概率.

$$P\{x_1 < X \le x_2\} = \boxed{P\{X \le x_2\}} \underbrace{\;-\;}_{?} \boxed{P\{X \le x_1\}}$$

$$F(x_2) \qquad\qquad F(x_1)$$

分布函数

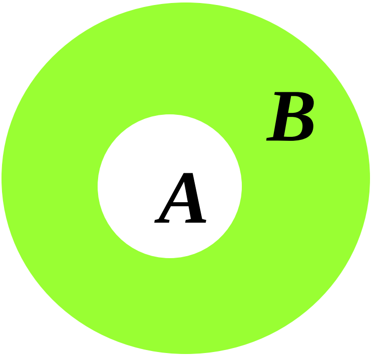
$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1).$$

(3) 设 A,B 为两个事件,且 $A \subset B$,则

$$P(A) \le P(B), \quad P(B-A) = P(B) - P(A).$$

证明　因为　$A \subset B$,

所以　$B = A \cup (B-A)$.



又　$(B-A) \cap A = \varnothing$,

得　$P(B) = P(A) + P(B-A)$.

于是　$\boxed{P(B-A) = P(B) - P(A).}$

又因　$P(B-A) \ge 0$,　故 $P(A) \le P(B)$.

2.分布函数的定义

定义 设 X 是一个随机变量, x 是任意实数,函数

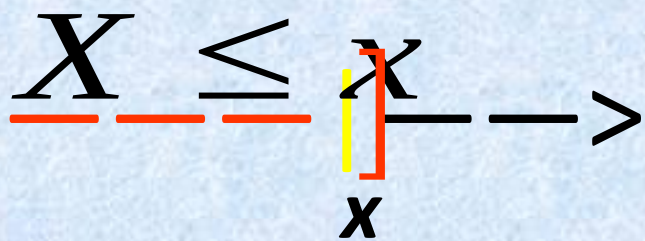
$$F(x) = P\{X \le x\}$$

称为 X 的分布函数.

说明

(1) 分布函数主要研究随机变量在某一区间内取值的概率情况.

(2)分布函数 $F(x)$ 是 x 的一个普通实函数.



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.

重要公式

(1) $P\{a < X \leq b\} = F(b) - F(a)$,

(2) $P\{X > a\} = 1 - F(a)$.

离散型随机变量的分布函数

设离散型随机变量X 的概率函数(分布律)是

$$P\{ X=x_k \} = p_k , \qquad k =1,2,3,\ldots$$

由概率的可列可加性,得X的分布函数为:

$$F(x) = P(X \le x) = \sum_{x_k \le x} p_k$$

由于$F(x)$ 是 X 取不超过x的诸值 x_k 的概率之和,故又称 $F(x)$ 为累积概率函数.

例1 将一枚硬币连掷三次, X 表示 “三次中正面出现的次数”, 求 X 的分布律及分布函数, 并求下列概率值 $P\{1 < X < 3\}, P\{X \geq 5.5\}, P\{1 < X \leq 3\}$.

解 设 $H-$正面, $T-$反面, 则

$$S=\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\},$$

因此分布律为

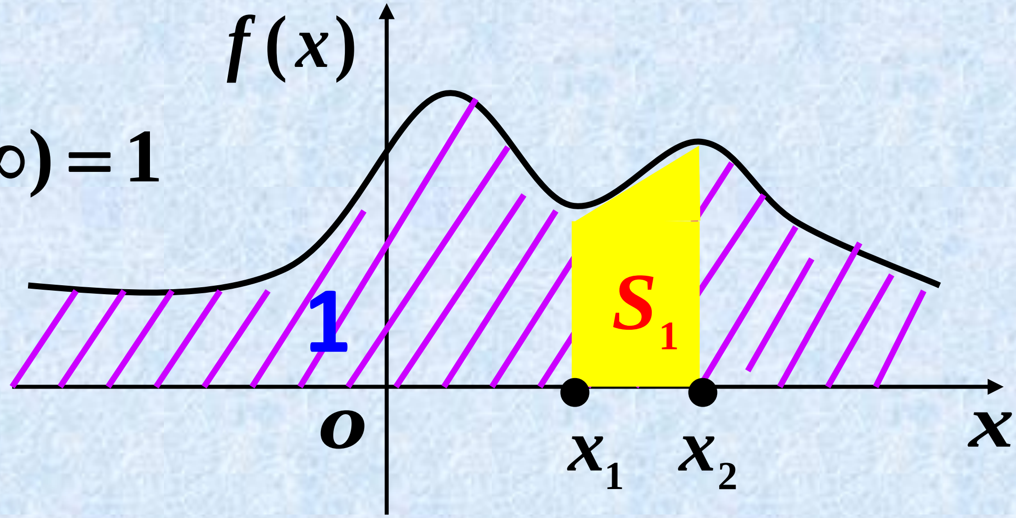
X	0	1	2	3
p	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

连续型随机变量的分布函数及其概率密度函数

1.定义 如果对于随机变量 X 的分布函数 $F(x)$, 存在非负函数, 使对于任意实数 x 有 $F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t$, 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.

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

$$S_1=\int_{x_1}^{x_2} f(x)\,\mathrm{d}x$$



性质 (1) $f(x) \geq 0;$

(2) $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1;$

这两条性质是判定一个函数 *f(x)*是否为某*r.vX*的概率密度函数的充要条件.

(3) $P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\,\mathrm{d}x;$

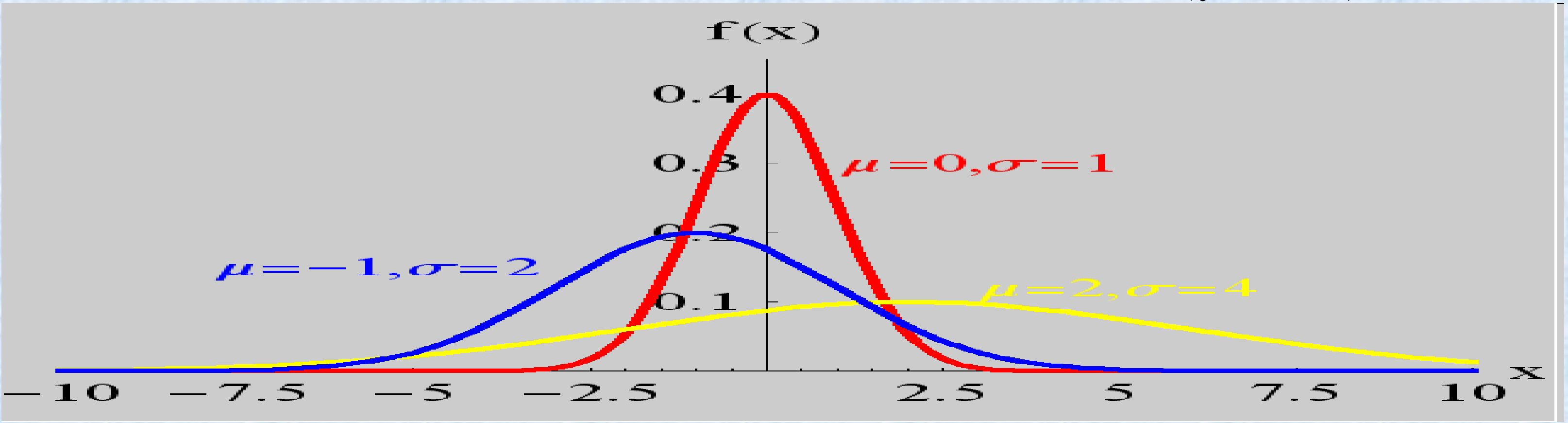
(4) 若 $f(x)$ 在点 x 处连续,则有 $F'(x) = f(x).$

3. 正态分布(或高斯分布)

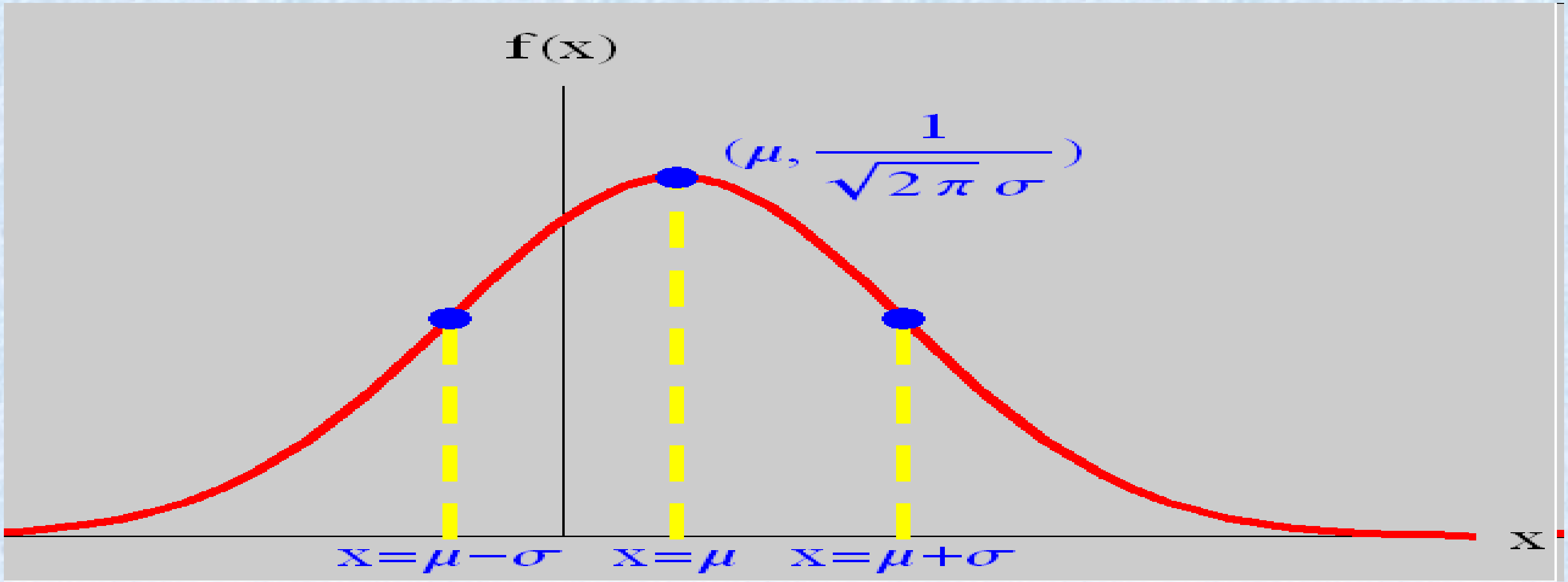
定义　设连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty<x<+\infty,$$

其中 $\mu,\sigma(\sigma>0)$ 为常数,则称 X 服从参数为 μ,σ 的正态分布或高斯分布,记为　$X\sim N(\mu,\sigma^2)$.



正态概率密度函数的几何特征



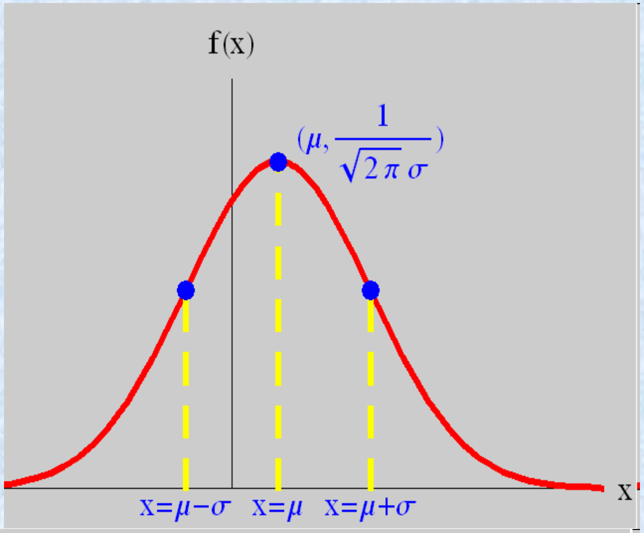
(1) 曲线关于 $x=\mu$ 对称;

(2) 当$x=\mu$时, $f(x)$取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$;

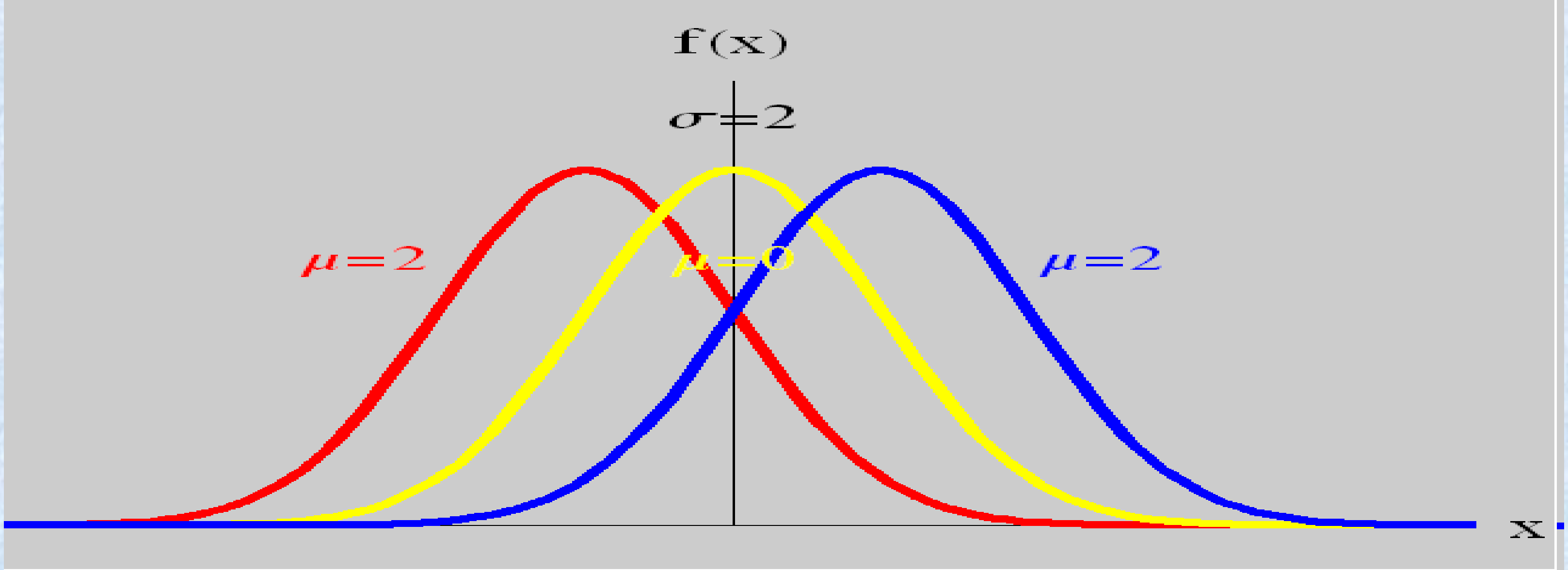
(3) 当 $x\to\pm\infty$ 时, $f(x)\to 0$;

(4) 曲线在 $x=\mu\pm\sigma$ 处有拐点;

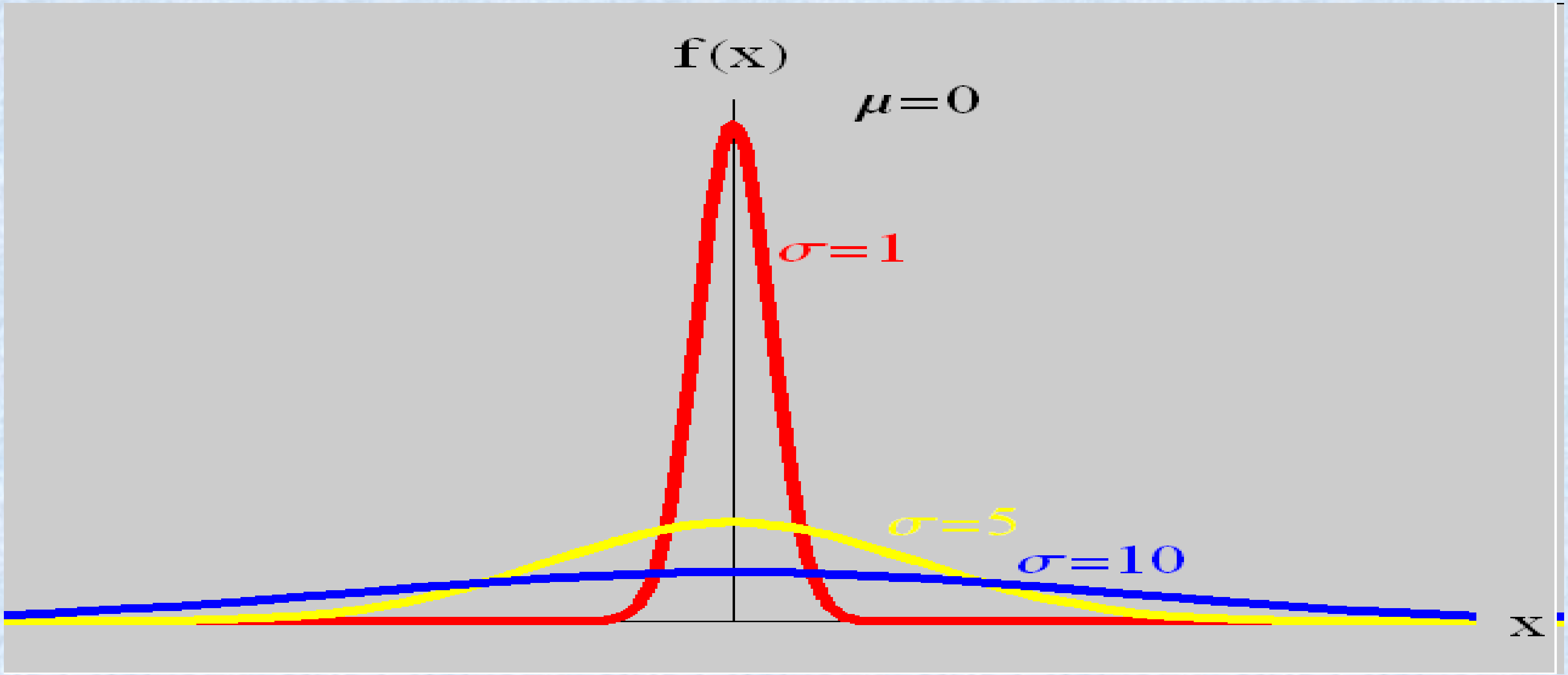
(5) 曲线以 x 轴为渐近线;



(6) 当固定 σ, 改变 μ 的大小时, $f(x)$ 图形的形状不变, 只是沿着 x 轴作平移变换;

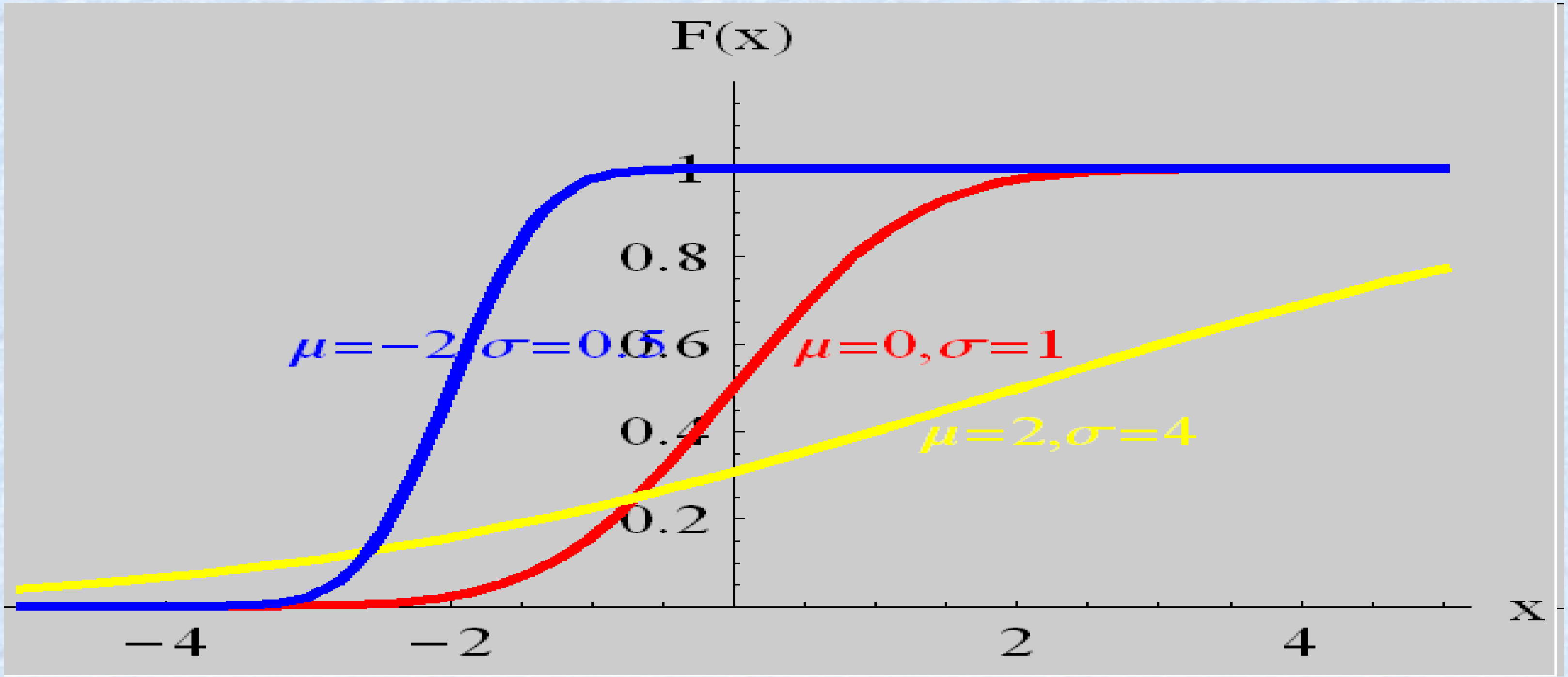


(7) 当固定 μ, 改变 σ 的大小时, $f(x)$ 图形的对称轴不变,而形状在改变, σ 越小,图形越高越瘦, σ越大,图形越矮越胖.



正态分布的分布函数

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t$$



3. 正态分布是概率论中最重要的分布

正态分布有极其广泛的实际背景, 例如测量误差, 人的生理特征尺寸如身高、体重等 ,正常情况下生产的产品尺寸:直径、长度、重量高度, 炮弹的弹落点的分布等, 都服从或近似服从正态分布.可以说,正态分布是自然界和社会现象中最为常见的一种分布, 一个变量如果受到大量微小的、独立的随机因素的影响, 那么这个变量一般是一个正态随机变量.

一、边缘分布函数

问题: 已知 (X,Y) 的分布, 如何确定 X,Y 的分布?

$$F(x,y)=P\{X\le x,Y\le y\},\ F(x)=P\{X\le x\},$$

$$P\{X\le x\}=P\{X\le x,Y<\infty\}=F(x,\infty)=F_X(x)$$

(X,Y)关于X的边缘分布函数.

定义 设 $F(x,y)$ 为随机变量 (X,Y) 的分布函数,

则 $$F(x,y) = P\{X \le x, Y \le y\}.$$

令 $y \to \infty$, 称 $P\{X \le x\} = P\{X \le x, Y < \infty\} = F(x,\infty)$

为随机变量 (X,Y) 关于X的边缘分布函数.

记为 $F_X(x) = F(x,\infty)$.

同理令 $x \to \infty$,

$$F_Y(y) = F(\infty, y) = P\{X < \infty, Y \le y\} = P\{Y \le y\}$$

为随机变量 (X,Y)关于Y的边缘分布函数.

一、随机变量的相互独立性

1.定义 设$F(x,y)$及$F_X(x),F_Y(y)$分别是二维随机变量(X,Y)的分布函数及边缘分布函数.若对于所有x,y有

$$P\{X \le x, Y \le y\} = P\{X \le x\}P\{Y \le y\},$$

即

$$F(x,y) = F_X(x)F_Y(y),$$

则称随机变量 X 和 Y 是相互独立的.

等价:设 (X,Y)是连续型*r.v*，若对任意的 *x*, *y*, 有

$$f(x,y) = f_X(x)f_Y(y)$$

几乎处处成立，则称*X*,*Y*相互独立 .

1. 离散型随机变量的数学期望

定义 设离散型随机变量 X 的分布律为

$$P\{X = x_k\} = p_k, \quad k = 1,2,\cdots.$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛, 则称级数 $\sum_{k=1}^{\infty} x_k p_k$ 为随机变量 X 的数学期望, 记为 $E(X)$. 即

$$E(X) = \sum_{k=1}^{\infty} x_k p_k.$$

2.连续型随机变量数学期望的定义

设连续型随机变量 X 的概率密度为 $f(x)$,若积分

$$\int_{-\infty}^{+\infty} x f(x) \mathrm{d} x$$

绝对收敛,则称积分 $\int_{-\infty}^{+\infty} x f(x) \mathrm{d} x$ 的值为随机变量 X 的数学期望,记为 $E(X)$.即

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d} x.$$

二、数学期望的性质

1. 设 C 是常数, 则有 $E(C)=C$.

2. 设 X 是一个随机变量,C 是常数, 则有

$$E(CX)=CE(X).$$

3. 设 X, Y 是两个随机变量, 则有

$$E(X+Y)=E(X)+E(Y).$$

4. 设 X, Y 是相互独立的随机变量, 则有

$$E(XY)=E(X)E(Y).$$

2. 方差的定义

设 X 是一个随机变量,若$E\{[X-E(X)]^2\}$存在,则称 $E\{[X-E(X)]^2\}$ 为 X 的方差, 记为 $D(X)$ 或 $\mathrm{Var}(X)$, 即$\cdots$

$$D(X)=\mathrm{Var}(X)=E\{[X-E(X)]^2\}.$$

称 $\sqrt{D(X)}$ 为标准差或均方差, 记为$\sigma(X)$.

3. 方差的意义

方差是一个常用来体现随机变量 X 取值分散程度的量. 如果 $D(X)$ 值大, 表示 X 取值分散程度大, $E(X)$ 的代表性差; 而如果 $D(X)$ 值小, 则表示X 的取值比较集中, 以 $E(X)$ 作为随机变量的代表性好.

4. 随机变量方差的计算

(1) 利用定义计算

离散型随机变量的方差

$$D(X)=\sum_{k=1}^{+\infty}[x_k-E(X)]^2 p_k,$$

其中 $P\{X=x_k\}=p_k, k=1,2,\cdots$是 X 的分布律.

连续型随机变量的方差

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)\,\mathrm{d}x,$$

其中 $f(x)$ 为X的概率密度.

(2) 利用公式计算

$$D(X)=E(X^2)-[E(X)]^2.$$

证明

$$\begin{aligned}D(X)&=E\{[X-E(X)]^2\}\\&=E\{X^2-2XE(X)+[E(X)]^2\}\\&=E(X^2)-2E(X)E(X)+[E(X)]^2\\&=E(X^2)-[E(X)]^2\\&=E(X^2)-E^2(X).\end{aligned}$$

5. 方差的性质

(1) 设 C 是常数, 则有 $D(C)=0.$

(2) 设 X 是一个随机变量, C 是常数, 则有

$$D(CX)=C^2D(X).$$

(3) 设 X, Y 相互独立, $D(X)$, $D(Y)$ 存在, 则

$$D(X\pm Y)=D(X)+D(Y).$$

若不独立,

$$\begin{aligned} D(X\pm Y) &= E\{[(X\pm Y)-E(X\pm Y)]^2\} \\ &= D(X)+D(Y)\pm 2\,\mathrm{Cov}(X,Y). \end{aligned}$$

重要分布的期望和方差方差

分　　布	参数	数学期望	方差
两点分布	$0 < p < 1$	p	$p(1-p)$
二项分布	$n \geq 1,$ $0 < p < 1$	np	$np(1-p)$
泊松分布	$\lambda > 0$	λ	λ
均匀分布	$a < b$	$(a+b)/2$	$(b-a)^2/12$
指数分布	$\theta > 0$	θ	θ^2
正态分布	$\mu,\sigma > 0$	μ	σ^2

二、基本定理

定理一（契比雪夫定理的特殊情况） 契比雪夫

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，且具有相同的数学期望和方差：$E(X_k) = \mu$, $D(X_k) = \sigma^2$ $(k = 1, 2, \cdots)$，作前 n 个随机变量的算术平均 $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$，则对于任意正数 ε 有

$$\lim_{n\to\infty} P\{|\overline{X} - \mu| < \varepsilon\} = \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1.$$

定理二（伯努利大数定理）

伯努利

设 n_A 是 n 次独立重复试验中事件 A 发生的次数, p 是事件 A 在每次试验中发生的概率, 则对于任意正数 $\varepsilon > 0$, 有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \text{ 或 } \lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0.$$

关于伯努利定理的说明:

伯努利定理表明事件发生的频率 $\frac{n_A}{n}$ 依概率收敛于事件的概率 p, 它以严格的数学形式表达了频率的稳定性.

故而当 n 很大时, 事件发生的频率与概率有较大偏差的可能性很小. 在实际应用中, 当试验次数很大时, 便可以用事件发生的频率来代替事件的概率.

定理三（辛钦定理）

辛钦资料

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，且具有数学期望 $E(X_k) = \mu$ $(k = 1,2,\cdots)$，

则对于任意正数 ε，有 $\lim\limits_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1.$

关于辛钦定理的说明:

(1) 与定理一相比, 不要求方差存在;

(2) 伯努利定理是辛钦定理的特殊情况.