第二章　随机变量及其分布

第一节　随机变量及其分布

第二节　离散型随机变量及其分布律

第三节　连续型随机变量及其密度函数

第四节　随机变量函数的分布

§2.1 随机变量及其分布函数

一、随机变量的概念

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.

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.

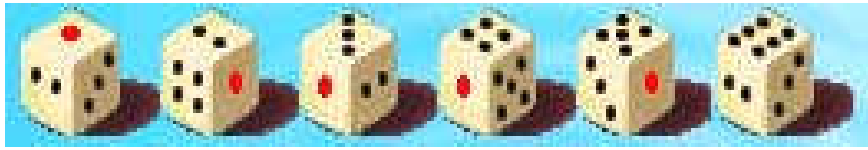
分析 $\Omega = \{$红色、白色$\}$ $\xrightarrow{?}$ 将 Ω 数量化

$\underbrace{\text{红色、白色}}_{\text{非数量}}$

令 X（红色）=1 ， X（白色）=0.

$$X(\omega) = \begin{cases} 0 & \omega = white \\ 1 & \omega = red \end{cases}$$

实例2　抛掷骰子, 观察出现的点数.



则有　$\Omega = \{1, 2, 3, 4, 5, 6\}$

样本点本身就是数量

$$X(\omega) = \omega$$

恒等变换

$$X(1)=1,\ X(2)=2,\ X(3)=3,\ X(4)=4,\ X(5)=5,\ X(6)=6.$$

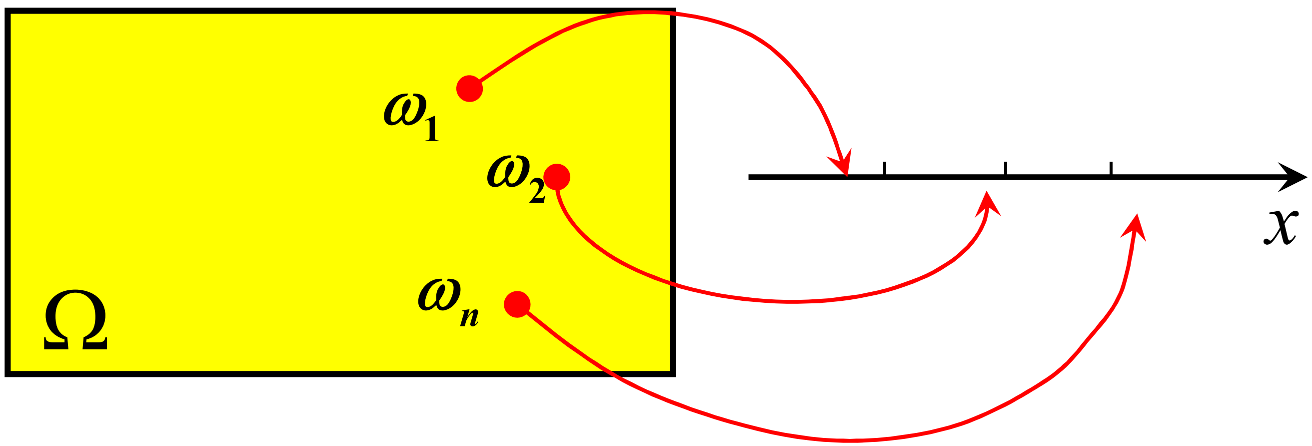
且有　$P\{X=i\} = \dfrac{1}{6}, \quad (i=1,2,3,4,5,6).$

定义1 设$X=X(\omega)$是定义在样本空间Ω上的实值函数，称

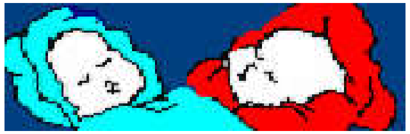
$X=X(\omega)$为随机变量(**random variable**).

随机变量通常用大写字母***X***，***Y***，***Z***，***W***...等表示.

样本点ω与实数$X=X(\omega)$对应的示意图如下



实例2 在有两个孩子的家庭中, 考虑其性别，共有 4 个样本点



$$\omega_1=(男,男),\omega_2=(男,女),\omega_3=(女,男),\omega_4=(女,女).$$

若用 X 表示该家女孩子的个数时，则有

$$X(\omega_1)=0,\quad X(\omega_2)=1,\quad X(\omega_3)=1,\quad X(\omega_4)=2,$$

可得随机变量 $X(\omega)$

$$X(\omega)=\begin{cases}0, & \omega=\omega_1,\\ 1, & \omega=\omega_2,\omega=\omega_3,\\ 2, & \omega=\omega_4.\end{cases}$$

注 (1)随机变量与普通的函数不同

随机变量是一个函数，但它与普通的函数有着本质的差别，普通函数是定义在实数轴上的，而随机变量是定义在样本空间上的（样本空间的元素不一定是实数).

(2)随机变量的取值具有一定的概率规律

随机变量随着试验的结果不同而取不同的值，由于试验的各个结果的出现具有一定的概率，因此随机变量的取值也有一定的概率规律.

(3) 随机变量与随机事件的关系

随机事件包容在随机变量这个范围更广的概念之内.

或者说 ：随机事件是从静态的观点来研究随机现象, 而随机变量则是从动态的观点来研究随机现象.

一般地，设 X 为一随机变量，L为实数集，则

$\{X \in L\}$ 表示一个随机事件.

特别地，如果a, b为实数，则

$$\{a < X \le b\}, \{X < b\}, \{X > a\}, \{X = a\}, \{X \ne a\}$$

等均表示随机事件.

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.

随机变量的四种类型（摘自清华大学公共基础平台课教材）
离散型、连续型、离散连续组合型、奇异性（研究上有意义）

二、分布函数(distribution function)

定义2 设 X 为一随机变量，对于任意实数 x, 称函数 $P\{X \le x\}$ 为 X 的分布函数，记为$F(x)$. 即随机变量 X 的分布函数为$F(x)=P\{X \le x\}, x \in (-\infty,+\infty)$.

【注意】（1）不论随机变量 X 如何取值，其分布函数$F(x)$ 的定义域总是 $(-\infty,+\infty)$.

（2）分布函数$F(x)$的直观意义为随机变量 X 落在区间$(-\infty,x]$ 上的概率.

定理1 对于任意实数a, b ($a<b$), 则有

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

推论1 设 x_0 为任一给定实数，则有

$$P\{X = x_0\} = F(x_0) - F(x_0 - 0)$$

【注】 利用定理1和推论1可得

$P\{X > a\} = 1 - F(a)$，$P\{X \ge a\} = 1 - F(a-0)$，

$P\{X < b\} = F(b-0)$，

$P\{a < X < b\} = F(b-0) - F(a)$

$P\{a \le X \le b\} = F(b) - F(a-0)$，

$P\{a \le X < b\} = F(b-0) - F(a-0)$

分布函数具有下基本性质

1. （有界性） $0 \le F(x) \le 1$

2. （单调不减性）$F(x)$是 x 的单调不减函数

3. （规范性）

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

4. （右连续性）$F(x+0)=F(x)$

【分析】分布函数右连续性

令 $A_n=\left\{X\le x+\dfrac{1}{n}\right\}(n\ge 1)$，则 $A_1\supset A_2\supset\cdots$ 且

$$\bigcap_{n=1}^{\infty}A_n=\{X\le x\}.$$

于是

$$\lim_{n\to\infty}F(x+\frac{1}{n})=\lim_{n\to\infty}P(A_n)=P(X\le x)=F(x).$$

由此知

$$\lim_{\Delta x\to 0^+}F(x+\Delta x)=F(x).$$

例1　已知随机变量 X 的分布函数为$F(x)=a+b\arctan x,$

求：（1）常数 $a, b.$

（2）$P\{-1<X\leq 0\}$ 和 $P\{X\geq\sqrt{3}\}.$

$$a=\frac{1}{2}, b=\frac{1}{\pi}, \frac{1}{4}, \frac{1}{6}$$

例2　在抛一枚硬币的随机试验中，记事件A表示硬币正面向上，令

$$X=\begin{cases}0, & 若A不生\\ 1, & 若A生\end{cases}$$

求随机变量 X 的分布函数$F(x)$，并讨论 $F(x)$ 的连续性.

解　由题意知 $P(A)=P(\bar{A})=\dfrac{1}{2}.$

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

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

当 $x\ge 1$ 时，$F(x)=P\{X\le x\}=P\{\Omega\}=1.$

续解

综上可得 X 的分布函数为

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

故$F(x)$在区间$(-\infty,0),(0,1),(1,+\infty)$内均连续，而点 $x=0$和

$x=1$均为$F(x)$的跳跃间断点（是右连续，不左连续）.

§2.2 离散型随机变量及其分布律

一、离散型随机变量及其分布律的概念

定义1 若随机变量 X 的取值为有限个或可列无限多个，就称 X 为离散型随机变量.

定义2 设 X 为离散型随机变量，其所有可能的取值为 $x_1, x_2, \cdots x_i, \cdots$,且

$$P\{X = x_i\} = p_i, \quad i = 1, 2, \cdots.$$

就称上式为离散型随机变量 X 的分布律或概率分布.

离散型随机变量 X 的分布律或概率分布也记为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

或
$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$$

性质1 设离散型随机变量 X 的分布律为

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{pmatrix}$$

则有 $(1)\ p_i \geq 0, i = 1, 2, \cdots.$ $(2) \sum_i p_i = 1$

【注】 如果实数列 $p_i (i = 1, 2, \cdots)$ 满足上述性质(1)，(2)，则 $p_i \geq 0, i = 1, 2, \cdots.$ 必能构成某离散型随机变量 X的分布律.

结论1 设L为任意实数集合，则 $P\{X \in L\} = \sum_{x_i \in L} p_i.$

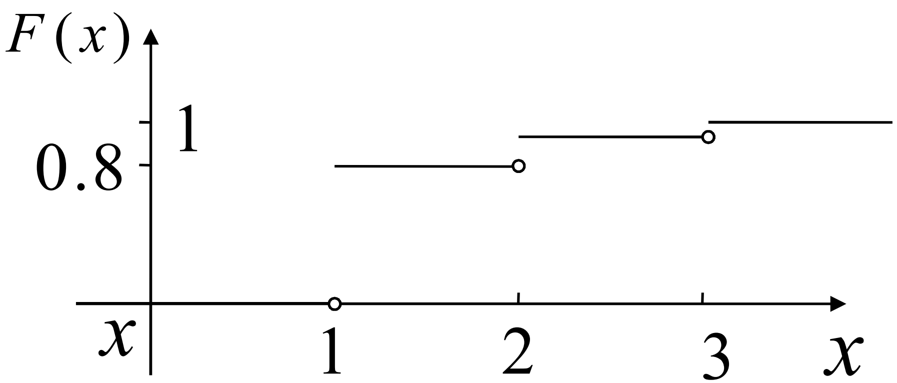
结论2 X的分布函数$F(x) = P\{X \leq x\} = \sum_{x_i \leq x} p_i, -\infty < x < +\infty.$

例1　设盒子中有8个正品和2个次品，现依次不放回地将其逐个取出，记 X 为首次取到正品时的所取产品个数，试求 X 的分布律和分布函数 $F(x)$.

$$X \sim \begin{pmatrix} 1 & 2 & 3 \\ \dfrac{4}{5} & \dfrac{8}{45} & \dfrac{1}{45} \end{pmatrix}$$

利用 $F(x)=P\{X\le x\}=\sum\limits_{x_i\le x}p_i,\ -\infty<x<+\infty,$ 可求得 X 的分布函数为

$$F(x)=\begin{cases}0, & x<1,\\ \dfrac{4}{5}, & 1\le x<2,\\ \dfrac{4}{5}+\dfrac{8}{45}, & 2\le x<3,\\ \dfrac{4}{5}+\dfrac{8}{45}+\dfrac{1}{45}, & x\ge 3\end{cases}=\begin{cases}0, & x<1,\\ \dfrac{4}{5}, & 1\le x<2,\\ \dfrac{44}{45}, & 2\le x<3,\\ 1, & x\ge 3.\end{cases}$$



例2　已知随机变量 X 的分布律为 $P\{X=i\}=\dfrac{k}{2^i}, i=1,2,\cdots$ 试求常数k，以及 X 取奇数的概率.

解　由 $\displaystyle\sum_{i=1}^{\infty}P\{X=i\}=\sum_{i=1}^{\infty}\frac{k}{2^i}=k\times\frac{\frac{1}{2}}{1-\frac{1}{2}}=k$, 以及分布律

的性质可得 $k=1$. 由上可知，X 的分布律为

$$P\{X=i\}=\frac{1}{2^i},\quad i=1,2,\cdots,$$

所以 X 取奇数的概率为

$$\sum_{i=1}^{\infty}P\{\mathrm{X}=2\mathrm{i}-1\}=\sum_{i=1}^{\infty}\frac{1}{2^{2i-1}}=\frac{\frac{1}{2}}{1-(\frac{1}{2})^2}=\frac{2}{3}.$$

二、几种常见的离散型随机变量的分布律

1. 0-1两点分布

定义3　如果随机变量 X 的分布律为

$$P(X=k)=p^k(1-p)^{n-k}, k=0,1$$

即

X	0	1
P	$1-p$	p

就称 X 服从0-1两点分布，记为 $X \sim B(1,p)$.

【注】 一般地，在随机试验*E*中，如果样本空间 Ω

只包含两个样本点 $\Omega=\{\omega_1,\omega_2\}$，且 $X=\begin{cases}0, & \omega=\omega_1,\\ 1, & \omega=\omega_2.\end{cases}$

则 $X \sim B(1,p)$，其中 $p=P\{X=1\}=P(\{\omega_2\})$

- 在现实生活中，**0-1**两点分布有着广泛的应用. 例如某产品合格与不合格；某课程的考试及格与不及格；某事件*A*发生与不发生等许多现象都能刻划成**0-1**两点分布.

家族简介

编辑

在科学史上，父子科学家、兄弟科学家并不鲜见，然而，在一个家族跨世纪的几代人中，众多父子兄弟都是科学家的较为罕见，其中，瑞士的伯努利（也译作贝努力、伯努利）家族最为突出。

伯努利家族3代人中产生了8位科学家，出类拔萃的至少有3位；而在他们一代又一代的众多子孙中，至少有一半相继成为杰出人物。伯努利家族的后裔有不少于120位被人们系统地追溯过，他们在数学、科学、技术、工程乃至法律、管理、文学、艺术等方面享有名望，有的甚至声名显赫。最不可思议的是这个家族中有两代人，他们中的大多数数学家，并非有意选择数学为职业，然而却忘情地沉溺于数学之中，有人调侃他们就像酒鬼碰到了烈酒。

老尼古拉·伯努利（Nicolaus Bernoulli，公元1623~1708年）生于巴塞尔，受过良好教育，曾在当地政府和司法部门任高级职务。他有3个有成就的儿子。其中长子雅各布（Jocob，公元1654~1705年）和第三个儿子约翰（Johann，公元1667~1748年）成为著名的数学家，第二个儿子小尼古拉（Nicolaus I，公元1662~1716年）在成为彼得堡科学院数学界的一员之前，是伯尔尼的第一个法律学教授。

1699年，雅各布当选为巴黎科学院外籍院士；1701年被柏林科学协会（后为柏林科学院）接纳为会员。

许多数学成果与雅各布的名字相联系。例如悬链线问题（1690年），曲率半径公式（1694年），“伯努利双纽线”（1694年），“伯努利微分方程”（1695年），“等周问题”（1700年）等。

雅各布对数学最重大的贡献是在概率论研究方面。他从1685年起发表关于赌博游戏中输赢次数问题的论文，后来写成巨著《猜度术》，这本书在他死后8年，即1713年才得以出版。

【贝努里试验】

由贝努里概率模型，在n重贝努里试验中，记 X表示事件A发生的次数，则 $X \sim B(n, p)$, 其中$p = P(A)$, 此分布称为贝努里分布.

2. 二项分布（贝努里（Bernoulli）分布）

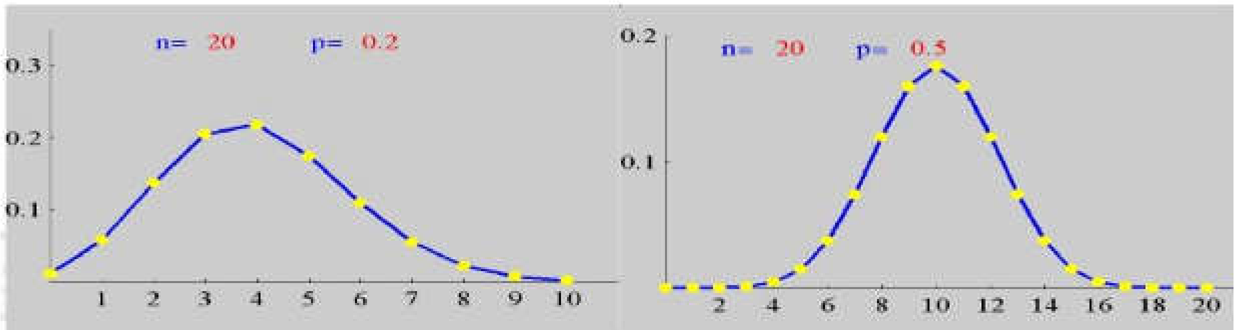
中文名	二项分布	涉及实验	伯努利试验；两点分布
外文名	Binomial Distribution	属　于	概率论与数理统计
提出者	伯努利	应用学科	大气科学；气候学；计算机科学

定义4　如果随机变量 X 的分布律为

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k}, k=0,1,2,\cdots,n.$$

就称 X 服从二项分布，记为 $X \sim B(n,p)$,

其中n为正整数，$0<p<1$.



图形特点

- 当$(n+1)p$不为整数时，二项概率$P\{X=k\}$在$k=[(n+1)p]$时达到最大值；$[x]$为不超过x的最大整数.
- 当$(n+1)p$为整数时，二项概率$P\{X=k\}$在$k=(n+1)p$和$k=(n+1)p-1$时达到最大值；

应用领域

- 在医学领域中,有一些随机事件是只有两种互斥结果的离散型随机事件,称为二项分类变量.如对病人治疗结果的有效与无效,某种化验结果的阳性与阴性，接触某传染源的感染与未感染等.
- 二项分布在心理与教育研究中，主要用于解决含有机遇性质的问题.所谓机遇问题，即指在实验或调查中，实验结果可由猜测而造成的。比如，选择题目的回答，划对划错，可能完全由猜测造成。欲区分由猜测而造成的结果与真结果之间的界限，就要用二项分布来解决。

【注1】 $C_n^k p^k (1-p)^{n-k} > 0, k=0,1,2,\cdots,n.$

$$\sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = [p+(1-p)]^n = 1.$$

满足分布律的性质

【注2】 又 $C_n^k p^k (1-p)^{n-k}$ 为二项式 $[p+(1-p)]^n$ 的展开式中的各项，因此称 X 服从二项分式.

【注3】 当n=1时，二项分布即为**0-1**两点分布.

例3　设随机变量$X \sim B(2, p), Y \sim B(3, p)$, 若$P\{X \geq 1\} = \dfrac{5}{9}$,

求$P\{Y \leq 1\}$.

解　由于$P\{X = 0\} = 1 - P\{X \geq 1\} = \dfrac{5}{9}$ 知,

$$P\{X = 0\} = C_2^0 p^0 (1-p)^2 = (1-p)^2 = \frac{4}{9},$$

所以 $p = \dfrac{1}{3}$, 从而

$$P\{Y \leq 1\} = P\{Y = 0\} + P\{Y = 1\}$$

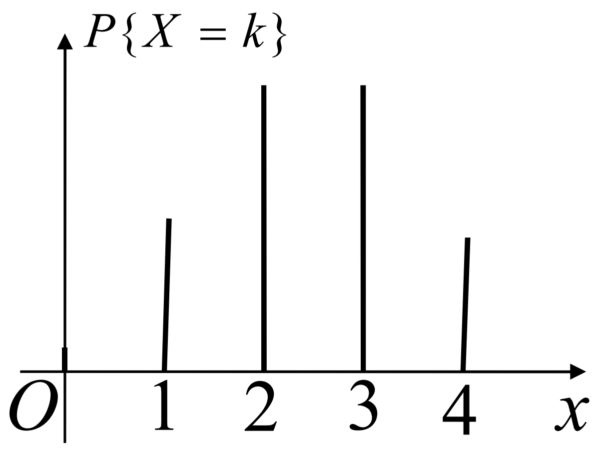
$$= C_3^0 (\frac{1}{3})^0 (1 - \frac{1}{3})^3 + C_3^1 (\frac{1}{3})^1 (1 - \frac{1}{3})^2 = \frac{20}{27}.$$

例4 设某射手独立地向一目标射击4次，每次击中目标的概率为0.6，求该射手在4次射击中，命中目标次数 X 的分布律，并问 X 取何值时的概率最大.

$$X \sim B(4,0.6)$$

$$P\{X=k\}=C_4^k 0.6^k 0.4^{4-k}, k=0,1,2,3,4.$$

$$P\{X=0\}=C_4^0\times0.6^0\times0.4^4=0.0256,$$
$$P\{X=1\}=C_4^1\times0.6^1\times0.4^3=0.1536,$$
$$P\{X=2\}=C_4^2\times0.6^2\times0.4^2=0.3456,$$
$$P\{X=3\}=C_4^3\times0.6^3\times0.4^1=0.3456,$$
$$P\{X=4\}=C_4^4\times0.6^4\times0.4^0=0.1296.$$



例5 设某机械产品的次品率为0.005，试分别求任意1000个产品中恰有10个次品的概率和不多于5个次品的概率.

$$X \sim B(1000, 0.005)$$

$$P(X = 10) = C_{1000}^{10} 0.005^{10} 0.995^{990}$$

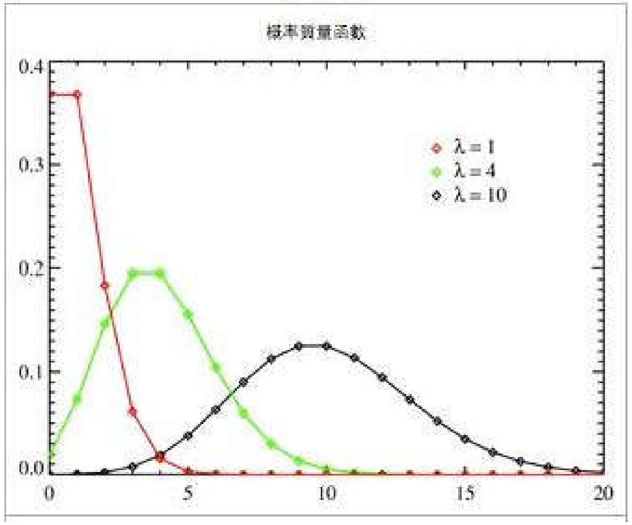
$$P(X \le 5) = \sum_{k=0}^{5} C_{1000}^{k} 0.005^{k} (1 - 0.005)^{1000-k}$$

上面两个计算结果虽然精确，但其计算量都非常大。在后续内容中，将陆续介绍泊松定理（第二章）和中心极限定理（第五章），利用这些定理可以近似计算出它们的值.

3. 泊松分布(Poission分布)

中文名	泊松分布	时　间	1838年
外文名	poisson distribution	台　译	卜瓦松分布
分　类	数学	提　出	西莫恩·德尼·泊松

泊松分布



定义5 如果随机变量 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\cdots,\lambda>0$$

就称 X 服从参数为 λ 的泊松分布，

记为 $X \sim P(\lambda)$.

【注1】 $P(X=k)=p_k>0, k=0,1,2,\cdots,n.$

$$\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{\lambda}\cdot e^{-\lambda}=1.$$

满足分布律的性质.

应用场景

- 在实际事例中，当一个随机事件，例如某电话交换台收到的呼叫次数、某段时间内来到某公共汽车站的乘客数、某放射性物质发射出的粒子、显微镜下某区域中的白血球数等近似服从泊松分布。

- 以固定的平均瞬时速率（或称密度）随机且独立地出现时，那么这个事件在单位时间（面积或体积）内出现的次数或个数就近似地服从泊松分布.

- 泊松分布在管理科学、运筹学以及自然科学的某些问题都占有重要的地位。（在早期学界认为人类行为服从泊松分布，2005年在nature上发表的文章揭示了人类行为具有高度非均匀性。）

泊松过程（请查找资料了解）

【注2】 一般来说，在一定的时间内，“稀有事件”发生的次数 X 服从泊松分布.

Siméon Denis Poisson
1781~1840

- 1837年，《概率论及其重要应用》一书中从数学推演的角度，得到了Poisson分布是二项式分布的极限分布.
- 在该书中并没有继续讨论这种分布的性质，在往后的研究中，Poisson似乎也把它忘掉了。

Bortkiewicz（1868～1931年）

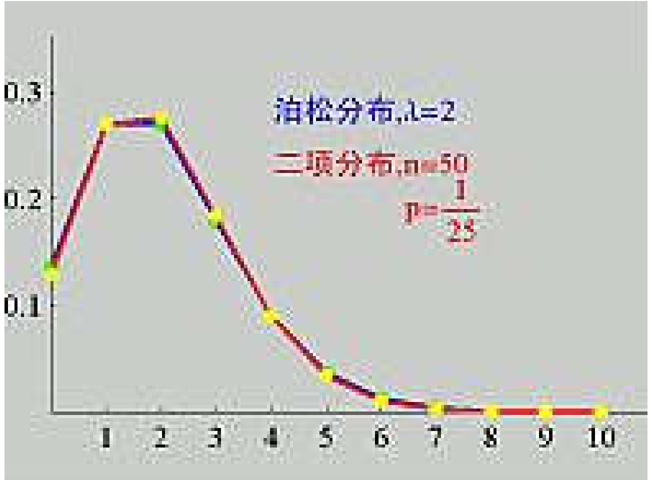
LADISLAUS VON BORTKIEWICZ (1868–1931)

- 直到19世纪末，Bortkiewicz（博尔特科威茨）注意到Poisson分布与某些类型的数据之间也有关系.
- 他不但在理论方面得到了泊松分布的许多性质，关键是介绍了泊松分布在众多领域的应用，从而使得该分布为人重视，并成为统计学中重要的离散型分布之一.

定理1（泊松定理） 在n重贝努里试验中，事件A在每次试验中发生的概率为p_n，其中$0<p_n<1$，且p_n与试验次数n有关，若$\lim\limits_{n\to\infty} np_n=\lambda\ (\lambda>0)$，则对任意非负整数$k$，有

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

k	按 $C_n^k p^k(1-p)^{n-k}$ 计算				按 $\frac{e^{-\lambda}}{k!}\lambda^k$ 计算
	$n=10$ $p=0.1$	$n=20$ $p=0.05$	$n=40$ $p=0.025$	$n=100$ $p=0.01$	$\lambda=np=1$
0	0.349	0.358	0.363	0.366	0.368
1	0.385	0.377	0.372	0.370	0.368
2	0.194	0.189	0.186	0.185	0.184
3	0.057	0.060	0.060	0.061	0.061
4	0.011	0.013	0.014	0.015	0.015
	0.004	0.003	0.005	0.003	0.004



$n\ge 100, np=\lambda\le 10$

n越大，p越小，近似效果越好

百度百科：$n>20, p<0.05$

例5 设某机械产品的次品率为**0.005**，试分别求任意**1000**个产品中恰有**10**个次品的概率和不多于**5**个次品的概率.

$$X \sim B(1000, 0.005)$$

$$P(X=10)=C_{1000}^{10}0.005^{10}0.995^{990}$$

$$P(X\le 5)=\sum_{k=0}^{5}C_{1000}^{k}0.005^{k}(1-0.005)^{1000-k}$$

$n=1000, p=0.005, np=5=\lambda, X \sim P(5)$, 查泊松分布表

$$P(X=10)=C_{1000}^{10}0.005^{10}0.995^{990}=\frac{5^{10}}{10!}e^{-5}\approx 0.018.$$

$$P(X\le 5)=\sum_{k=0}^{5}\frac{5^{k}}{k!}e^{-5}=P(X\le 5)=1-P(X\ge 6)=1-0.384=0.616.$$

例6 设随机变量 $X \sim P(\lambda)$, 且 $P\{X=1\}=P\{X=2\}$,

求 $P\{X=3\}$ 和 $P\{X=2\}$. $P(X >= 1)$

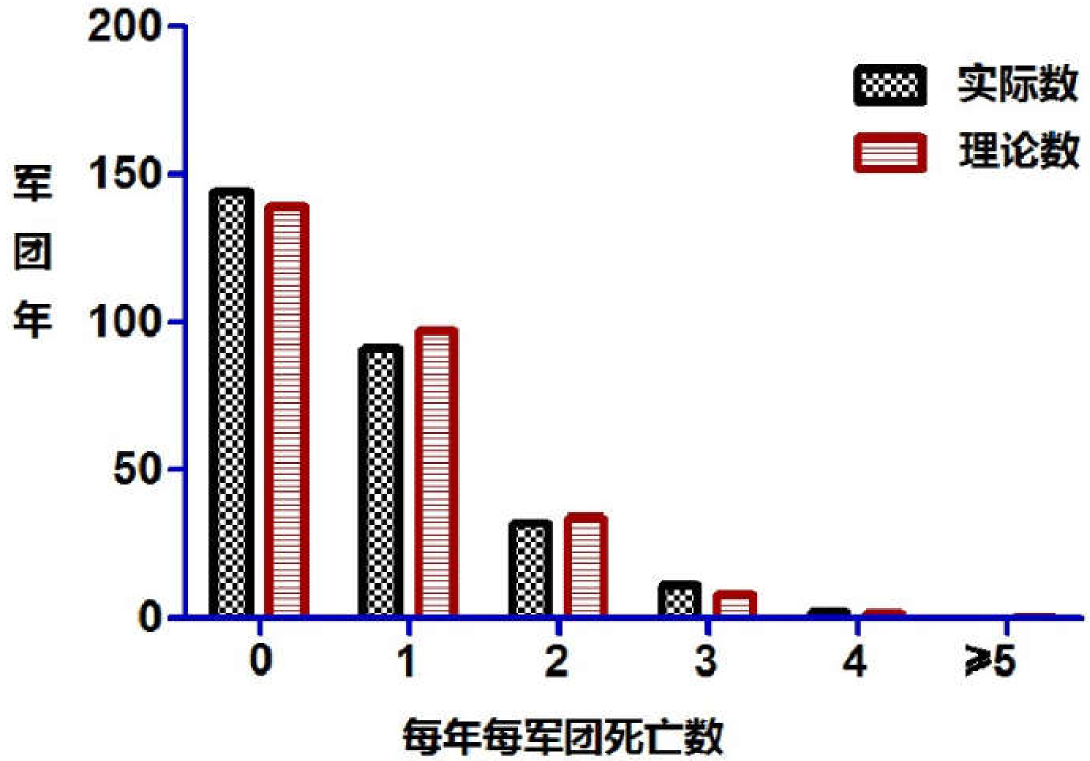
$$\lambda = 2,\ P\{X=3\} = \frac{4}{3}e^{-2},$$

$$P\{X \geq 1\} = 1 - P\{X=0\} = 1 - e^{-2}.$$

■ 从1875到1894年的20年间，德国的14个军团都有士兵被马踢伤因而致死的人数记录.

表　平均计数为0.7的Poisson分布

每年每军团死亡人数	观察数	理论数
0	144	139.0
1	91	97.3
2	32	34.1
3	11	8.0
4	2	1.4
≥ 5	0	0.2
合计	280	280


200
150
100
50
0
军团年
实际数
理论数
0
1
2
3
4
≥5
每年每军团死亡数

■ 2002年韩日世界杯64场比赛中，各队进球数有多有少。大部分是0,1,2个进球，个别队是5个以上进球，最多的是8个进球，平均是1.2578个/场/对.虽然强队大都能进球、赢球（如巴西队），弱队大都不能进球（如中国队）.但宏观上来说，各队进球数服从Poission分布！

每场各队进球数	场次	理论数
0	37	36.39
1	47	45.77
2	27	28.78
3	13	12.07
4	2	3.79
5	1	0.95
≥6	1	0.25
	128	128.00

- 20世纪初罗瑟福和盖克两位科学家在观察与分析放射性物质放出的α 粒子个数的情况时，他们做了2608次观察（每次时间为7.5秒）发现放射性物质在规定的一段时间内，其放射的粒子数 X 服从泊松分布.

放射粒子数的观察频率与概率

放射粒子数 X	观察到次数 μ_k	频率 $\nu_k=\frac{\mu_k}{2608}$	$p_k=\frac{1}{k!}\lambda^k e^{-\lambda}$ $(\lambda=3.87)$
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.201
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
10	10	0.004	0.004
≥11	6	0.002	0.003
总计	2608	0.999	1.000

例7

进行重复独立试验，设每次试验成功的概率为 p，失败的概率为 $q=1-p,(0<p<1)$.

（1）将试验进行到出现一次成功为止，以X表示所需试验次数，求X的分布律.

$X \sim Ge(p)$

几何分布也称“等待分布”

（2）将试验进行到出现 r 次成功为止，以Y表示所需试验次数，求 Y 的分布律.

$Y \sim F(r,p)$

负二项分布也称“帕斯卡分布”

“守株待兔”

4. 几何分布和负二项分布

定义6　如果随机变量 X 的分布律为

$$P\{X=k\}=(1-p)^{k-1}p,\ k=1,2,\cdots,0<p<1.$$

就称 X 服从参数为 p 的几何分布，记为 $X \sim Ge(p).$

定义7　如果随机变量 X 的分布律为

$$P(X=k)=C_{k-1}^{r-1}p^r q^{k-r},k=r,r+1,\cdots$$

就称X服从参数为 p 和 $\boldsymbol{r}$ 的负二项分布，记为

$$X \sim F(r,p).$$

例8 设袋中有10个红球和6个白球，现从中任取5个球，则5个球中恰有 k 个白球的概率？

$$P\{X=k\}=\frac{C_6^k C_{10}^{5-k}}{C_{16}^5}, k=0,1,\cdots,5.$$

5. 超几何分布

定义8 如果随机变量X的分布律为$P\{X=k\}=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$,

其中 $N>1, M\le N, n\le N, \max\{0, M+n-N\}\le k\le \min\{M,n\}$,

就称X服从参数为M,N,n的超几何分布, 记为$X \sim H(M,N,n)$.

自学课本第47页例
2.2.10，　性质2.2.2

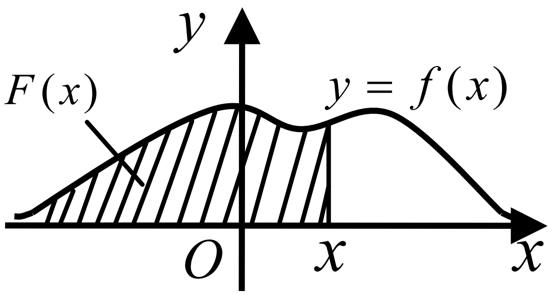
§ 2.3 连续型随机变量及其密度函数

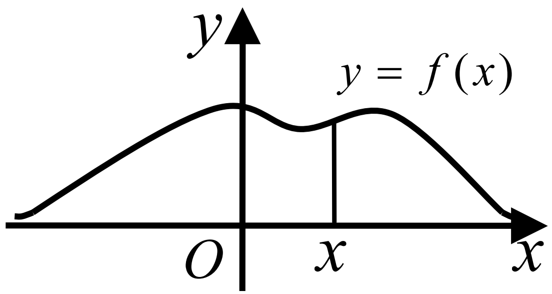
一、连续型随机变量及其密度函数的概念

定义1 设随机变量 X 的分布函数为 $F(x)$, 若存在非负可积函数 $f(x)$, 使对任意实数 x, 均有

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

就称 X 为连续型随机变量, 其中$f(x)$称为X的密度函数或概率密度(probability density function).





性质1　设连续型随机变量 X 的密度函数为 $f(x)$，则

（1）非负性 $f(x)\geq 0, -\infty < x < +\infty$;

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

【注】若函数 $f(x)$ 满足上述性质 (1) (2)，则 $f(x)$ 必为某连续型随机变量 X 的密度函数.

【结论1】 连续型随机变量X的分布函数$F(x)$是连续函数.

且在其密度函数$f(x)$的连续点x处, $F(x)$可导, $F'(x) = f(x).$

【注】 如果分布函数 $F(x)$ 不是连续函数，则随机变量 X 不是连续随机变量.

【注】 如果随机变量X的分布函数 $F(x)$ 是连续函数，并且 $F(x)$ 最多有有限个不可导点，则 X 为连续型随机变量.

且其密度函数 $f(x)$ 可由下列方式确定：

（1）在$F(x)$的可导点x处，取 $F'(x) = f(x)$;

（2）在$F(x)$的不可导点x处，取$f(x)$为任意非负数值,

可验证$f(x)$满足密度函数的性质.

例：随机变量X的分布函数为

$$F(x)=\frac{1}{2}+\frac{1}{\pi}\arctan x,\ -\infty<x<+\infty$$

处处可导，因此X为连续型随机变量，且其密度函数为

$$f(x)=F'(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},-\infty<x<+\infty.$$

例：随机变量X的分布函数为

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

非连续非离散型

由于$F(x)$在点$x=0$处不连续，故X不是连续型随机变量，

由2. 2可知，X也不是离散型随机变量.

密度函数的概率意义？

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

$$=\lim_{\Delta x\to 0}\frac{P\left\{x<X\leq x+\Delta x\right\}}{\Delta x}$$

当 Δx 充分小时，$f(x)\approx\dfrac{P\left\{x<X\leq x+\Delta x\right\}}{\Delta x}$

密度函数即随机变量在单位长度内发生的概率. 反映了连续型随机变量X在点x处附近的概率“密集程度”.

【结论2】 设X为连续型随机变量, x_0 为任意一个实数, 则

$$P\{X=x_0\}=0$$

【注】 由于结论2，所以在有限个点处改变$f(x)$的取值，不影响X的整体分布，表明密度函数$f(x)$的表达式可以不惟一.

【注】 $P(A)=0 \not\Rightarrow A=\varnothing;$

如： 由$P(AB)=0$不能推得 $AB=\varnothing$,不能说明A与B互斥.

$P(A)=1 \not\Rightarrow A=\Omega.$

【结论3】 设X为连续型随机变量，a,b为任意两个实数，且$a<b$，则 $P\{a<X\le b\}=\int_a^b f(x)dx.$

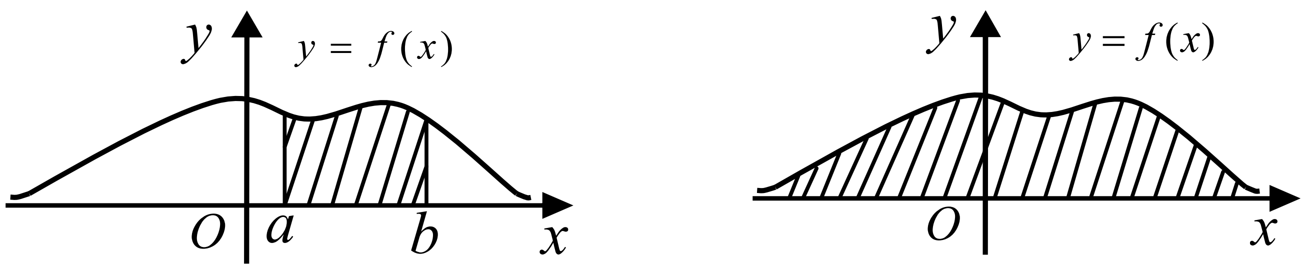
即 $$P\{a<X\le b\}=F(b)-F(a)=\int_a^b f(x)dx$$

【注】当X为连续型的随机变量有

$$P\{a<X<b\}=P\{a\le X\le b\}=P\{a<X\le b\}=\int_a^b f(x)dx.$$

一般地，设 I 为一区间，则有 $P\{X\in I\}=\int_I f(x)dx.$

几何直观：随机变量属于某区间的概率



例1 设随机变量X的密度函数为

$$f(x)=\begin{cases} k\cos x, & |x|<\dfrac{\pi}{2}, \\ 0, & |x|\ge\dfrac{\pi}{2}. \end{cases}$$

（1）求常数 k;

（2）求 X 的分布函数 $F(x)$;

（3）计算 $P\{-\dfrac{\pi}{4}<X<\pi\}$.

基本题型

（一定掌握）

【解】 由 $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} k\cos x dx = 1$, 计算得 $k = \frac{1}{2}$.

即X的密度函数为 $f(x) = \begin{cases} \frac{1}{2}\cos x, & |x| < \frac{\pi}{2}, \\ 0, & |x| \geq \frac{\pi}{2}. \end{cases}$

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

$$F(x) = \begin{cases} 0, & x < -\frac{\pi}{2}, \\ \int_{-\frac{\pi}{2}}^{x} \frac{1}{2}\cos t dt, & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}, \\ 1, & x > \frac{\pi}{2}. \end{cases} = \begin{cases} 0, & x < -\frac{\pi}{2}, \\ \frac{\sin x + 1}{2}, & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}, \\ 1, & x > \frac{\pi}{2}. \end{cases}$$

（3） $P\left\{-\frac{\pi}{4}<X<\pi\right\}=F(\pi)-F(-\frac{\pi}{4})$

$$=1-\frac{\sin(-\frac{\pi}{4})+1}{2}=\frac{2+\sqrt{2}}{4}.$$

或 $P\left\{-\frac{\pi}{4}<X<\pi\right\}=\int_{-\frac{\pi}{4}}^{\pi}f(x)dx$

$$=\int_{-\frac{x}{4}}^{\frac{\pi}{2}}\frac{1}{2}\cos xdx=\frac{1}{2}\sin x\Big|_{-\frac{x}{4}}^{\frac{\pi}{2}}$$

$$=\frac{2+\sqrt{2}}{4}.$$

例2 设随机变量X的密度函数$f(x)$为偶函数, $F(x)$为X的分布函数，证明：

（1） $F(x)+F(-x)=1;$

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

（3） $F(-x)=\dfrac{1}{2}-\int_{0}^{x} f(t)dt.$

证（1）$F(-x)=\int_{-\infty}^{-x} f(t)\,\mathrm{d}t \overset{t=-u}{=} -\int_{+\infty}^{x} f(-u)\,\mathrm{d}u = \int_{x}^{+\infty} f(u)\,\mathrm{d}u$

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

故 $F(x)+F(-x)=1.$

（2） 令（1）中 $\boldsymbol{x=0}$, 得 $F(0)=\int_{-\infty}^{0} f(x)\,\mathrm{d}x=\dfrac{1}{2}$, 且

$$\int_{0}^{+\infty} f(x)\,\mathrm{d}x=\frac{1}{2}.$$

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

$$=1-\left[\int_{-\infty}^{0} f(t)\,\mathrm{d}t+\int_{0}^{x} f(t)\,\mathrm{d}t\right]$$

$$=1-\left[\frac{1}{2}+\int_{0}^{x} f(t)\,\mathrm{d}t\right]=\frac{1}{2}-\int_{0}^{x} f(t)\,\mathrm{d}t$$

二、几种常见的连续型随机变量的概率分布

1. 均匀分布(uniform distribution)

定义2　如果随机变量 X 的密度函数为

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

就称X服从区间$[a,b]$上的均匀分布，记为$X \sim U[a,b]$.

【均匀分布产生的背景】

在估计、计算及 测量引起的随机误差问题中，有一类误差具有“均匀性”，相应的分布叫做均匀分布.

均匀分布的应用 （计算机定点近似的舍入误差）

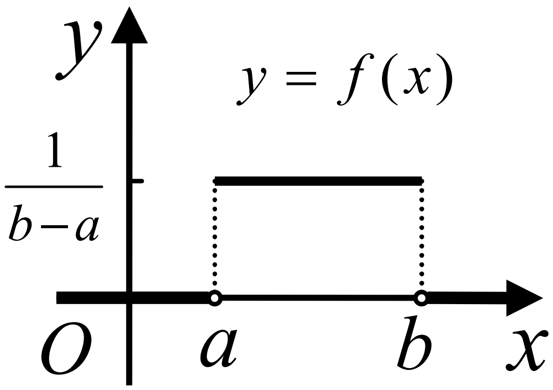
在计算机上，对实数要作近似处理，所谓单精度实数和双精度实数，就是这类处理的规则和结果.

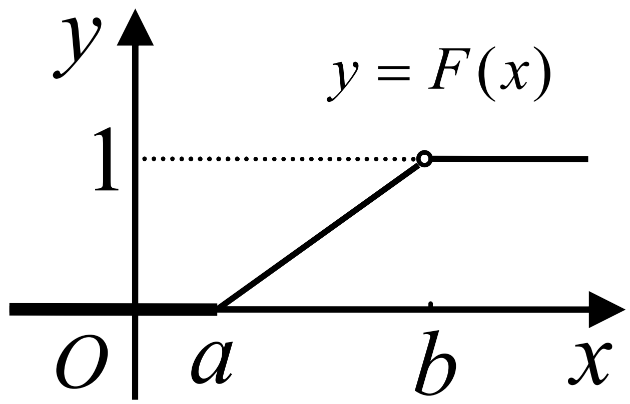
设将实数小数点后第6位四舍五入，则舍入误差X在 $(-0.5\times10^{-5},0.5\times10^{5})$ 内任一点的领域为等可能的，因此可认为X服从此区间上的均匀分布.

况，单精度浮点数占4字节（32位）内存空间，其数值范围为-3.4E38~3.4E+38,；双精度型占8 个字节（64位）内存空间，其数值范围为-1.7E308~1.7E+308。

单精度浮点数的实际有效精度为24位二进制，这相当于 24*log102≈7.2 位10进制的精度，所以平时我们说"单精度浮点数具有7位精度"。（精度的理解：当从1.000...02变化为1.000...12时，变动范围为2^23，考虑到因为四舍五入而得到的1倍精度提高，所以单精度浮点数可以反映2^24的数值变化，即24位二进制精度）

误差





【注】均匀分布与第一章中介绍的几何概型原理相通，适用于一维的几何概型试验. 此时，X 落入某区间 I 内（上）的概率为

$$P\{X \in I\} = P\{X \in I \cap [a,b]\} = \frac{m\{I \cap [a,b]\}}{b-a}$$

例3 如果随机变量 $X \sim U[1,6]$，求方程 $x^2 + Xx + 1 = 0$ 有实根的概率.

解 $P\{\text{原方程有实根}\} = P\{\Delta = X^2 - 4 \geq 0\}$

$$= P\{|X| \geq 2\} = P\{2 \leq X \leq 6\} = \frac{6-2}{6-1} = \frac{4}{5}.$$

2. 指数分布(exponnent distribution)

定义3　如果随机变量 X 的密度函数为

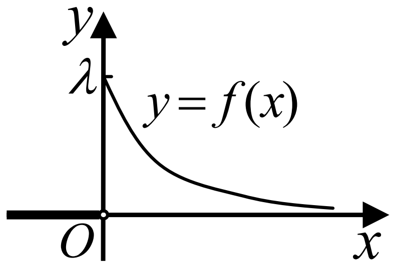
$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0,\\ 0, & x<0.\end{cases}$$

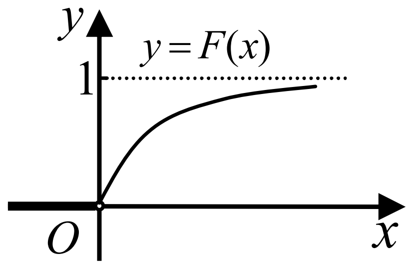
就称 X 服从参数为 λ 的指数分布，记为 $X \sim e(\lambda)$.

λ – 失效率（单位时间内失效的概率）　$\dfrac{1}{\lambda}$ –平均寿命

指数分布是重要的连续型分布之一，常用来作各种“寿命”分布的近似，在产品的可靠性分析、生存分析、随机服务系统中均有重要地位.

【注】 X的分布函数为 $F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0,\\ 0, & x<0.\end{cases}$





性质2 设随机变量 $X \sim e(\lambda)$, 则当 $s,t>0$时

$$P\{X>s+t \mid X>s\}=P\{X>t\}$$

【注1】性质2称为指数分布的无记忆性或“永远年轻”或“无老化”.

【注2】连续型随机变量中仅指数分布具有“无记忆性”;

离散型随机变量中仅几何分布具有“无记忆性”;

证　由题意知，X 的密度函数为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0, \\ 0, & x\le 0.\end{cases}$

则当 $a>0$ 时，有 $P\{X>a\}=\int_a^{+\infty}\lambda e^{-\lambda x}\mathrm{d}x=e^{-\lambda a}$,

所以 $P\{X>s+t \mid X>s\}=\dfrac{P\{X>s+t, X>s\}}{P\{X>s\}}$

$$=\frac{P\{X>s+t\}}{P\{X>s\}}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$=e^{-\lambda t}=P\{X>t\}.$$

定理1 设连续型随机变量 X 取非负值，且满足“无忆性”，则该随机变量必为指数分布.

证明 令 $f(u)=P\{X>u\}, u>0,$ 由 $P\{X>s+t \mid X>s\}=P\{X>t\},$

可知 $f(s+t)=f(s)f(t),$ 于是

$$f(1)=f\left(\frac{1}{n}+\frac{1}{n}+\cdots+\frac{1}{n}\right)=\left[f\left(\frac{1}{n}\right)\right]^n,$$

故
$$f\left(\frac{1}{n}\right)=\left(f(1)\right)^{\frac{1}{n}},$$

$$f\left(\frac{m}{n}\right)=f\left(\frac{1}{n}+\frac{1}{n}+\cdots+\frac{1}{n}\right)=\left[f\left(\frac{1}{n}\right)\right]^m=\left(f(1)\right)^{\frac{m}{n}},$$

故对一切正有理数 r 有 $f(r)=\left(f(1)\right)^r$,

由于$0<f(1)<1$, 且 $f(u)$是 u 的减函数,不难推知,

对一切 $u\geq 0$ 有 $f(u)=\left(f(1)\right)^u$, 令 $\lambda=-\ln f(1)$, 则

$$f(u)=e^{-\lambda u},$$

即 $$P\{X>u\}=e^{-\lambda u}=\int_u^{+\infty}\lambda e^{-\lambda x}dx$$

故 $$P\{X\leq u\}=\int_{-\infty}^{u}\lambda e^{-\lambda x}dx$$

这表明 X 服从指数分布.

例4 某仪器装有三只独立工作的同型号电子元件，其寿命（单位：小时）都服从**参数为 1/600 的指数分布**. 试求在仪器使用的最初 **200** 小时内，至少有一只电子元件损坏的概率？

解 设电子元件的寿命为随机变量X, 则 $X \sim e(\frac{1}{600})$,

三只电子元件中损坏的只数为随机变量 Y，则

$$Y \sim B(3, p_0)$$

其中 $p_0 = P\{X < 200\} = F(200) = 1 - e^{-\frac{1}{3}}$.

故至少有一只电子元件损坏的概率为

$$P\{Y \geq 1\} = 1 - P\{Y = 0\} = 1 - e^{-1}.$$

例5

设某地在任何长为 t（周）的时间内发生地震的次数 $N(t)$服从参数 λt 的泊松分布.（称为泊松流）

若 T表示直到下一次地震发生所需的时间（单位:周），求 T的概率分布.

解 由题意 $N(t) \sim P(\lambda t)$, 即 $P\{N(t)=k\}=\dfrac{(\lambda t)^k}{k!}e^{-\lambda t}.$

则 $P\{T \le t\}=1-P\{T>t\}=1-P\{N(t)=0\}=1-e^{-\lambda t},$

即 $T \sim e(\lambda).$

泊松分布是泊松流在 t=1 时的特殊情形；当某事件在时间段 t 内发生的次数服从泊松流时，则首次发生需要的等待时间服从指数分布，也可理解为两次事件发生之间的时间长度服从指数分布，第 r 次事件发生的时间长度服从 Γ 分布.

收藏 4034

正态分布 编辑

本词条由"科普中国"百科科学词条编写与应用工作项目 审核 。

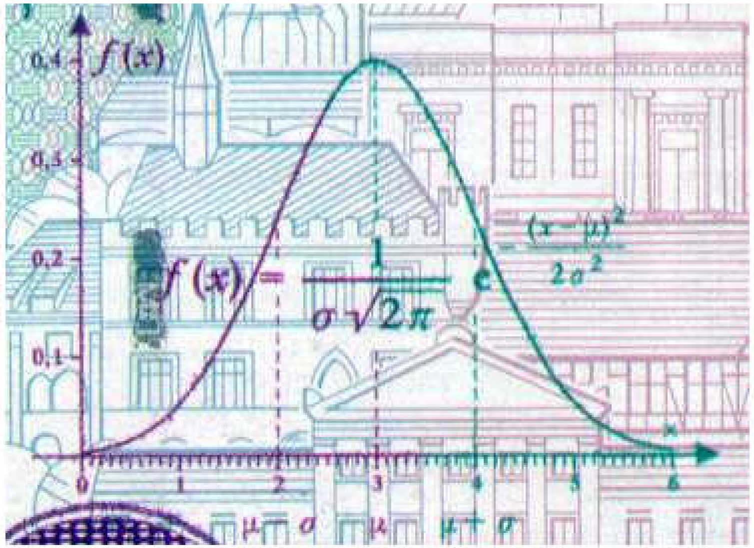
正态分布（Normal distribution），也称"常态分布"，又名高斯分布（Gaussian distribution），最早由A.棣莫弗在求二项分布的渐近公式中得到。C.F.高斯在研究测量误差时从另一个角度导出了它。P.S.拉普拉斯和高斯研究了它的性质。是一个在数学、物理及工程等领域都非常重要的概率分布，在统计学的许多方面有着重大的影响力。

中文名	正态分布	所属学科	概率论
外文名	normal distribution	又　名	高斯分布
发现者	棣莫弗	应用领域	数学、物理及工程等领域

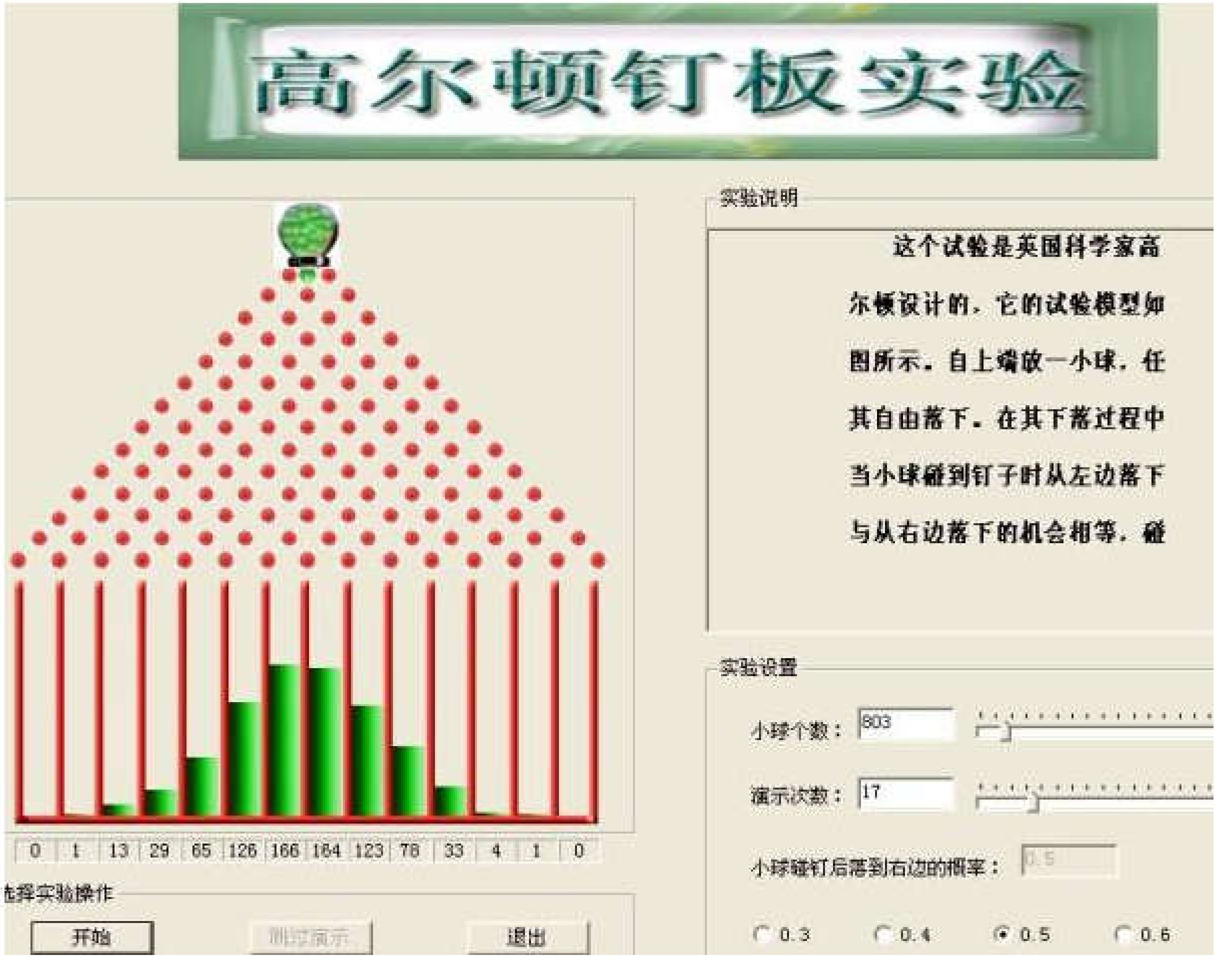
历史发展

编辑

正态分布概念是由德国的数学家和天文学家Moivre于1733年首次提出的，但由于德国数学家Gauss率先将其应用于天文学家研究，故正态分布又叫高斯分布，高斯这项工作对后世的影响极大，他使正态分布同时有了"高斯分布"的名称，后世之所以多将最小二乘法的发明权归之于他，也是出于这一工作。但现今德国10马克的印有高斯头像的钞票，其上还印有正态分布的密度曲线。这传达了一种想法：在高斯的一切科学贡献中，其对人类文明影响最大者，就是这一项。在高斯刚作出这个发现之初，也许人们还只能从其理论的简化上来评价其优越性，其全部影响还不能充分看出来。这要到20世纪正态小样本理论充分发展起来以后。拉普拉斯很快得知高斯的工作，并马上将其与他发现的中心极限定理联系起来，为此，他在即将发表的一篇文章（发表于1810年）上加上了一点补充，指出如若误差可看成许多量的叠加，根据他的中心极限定理，误差理应有高斯分布。这是历史上第一次提到所谓"元误差学说"——误差是由大量的、由种种原因产生的元误差叠加而成。后来到1837年，海根（G.Hagen）在一篇论文中正式提出了这个学说。



德国马克和纪念币上的高斯头像和正态分布曲线



高尔顿钉板实验
实验说明
这个试验是英国科学家高尔顿设计的，它的试验模型如图所示。自上端放一小球，任其自由落下。在其下落过程中当小球碰到钉子时从左边落下与从右边落下的机会相等，碰
0 1 13 29 65 126 166 164 123 78 33 4 1 0
选择实验操作
开始
退出
实验设置
小球个数：
803
演示次数：
17
小球碰钉后落到右边的概率：
0.3
0.4
0.5
0.6

CAPITAL OF STATISTICS

PROFESSION, HUMANITY & INTEGRITY

首页 关于 论坛 投稿 搜索

← 返回首页 2013-01-28 编辑本页 →

统计模型

正态分布的前世今生（上）

靳志辉

关键词：历史; 正态分布

CAPITAL OF STATISTICS

PROFESSION, HUMANITY & INTEGRITY

首页 关于 论坛 投稿 搜索

← 返回首页 2013-01-28 编辑本页 →

统计模型

正态分布的前世今生（下）

靳志辉

关键词：历史; 正态分布; 经典理论

6. 开疆拓土，正态分布的进一步发展

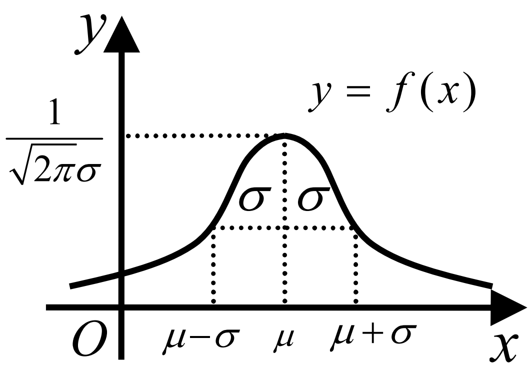
3. 正态分布（高斯分布：**Gauss**）

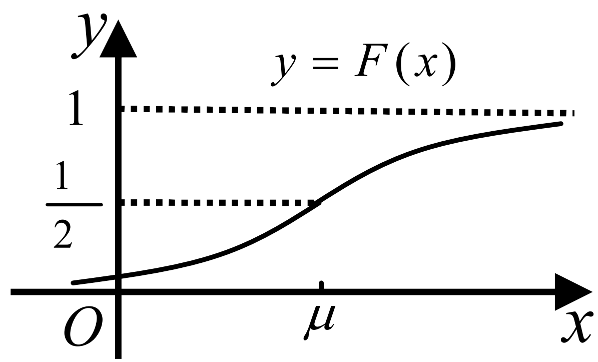
定义4　如果随机变量 X 的概率密度函数为

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

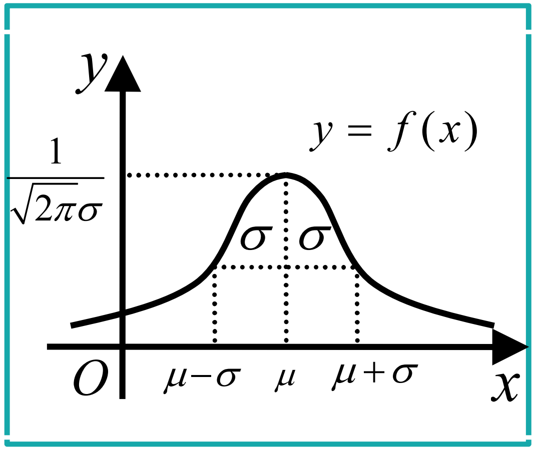
就称 X 服从参数为 μ,σ 的正态分布，记为 $X \sim N(\mu,\sigma^2)$.

其中 $-\infty<\mu<+\infty,\sigma>0$. $N-Normal.$





【1】 $f(x)$关于 $x=\mu$ 对称； $\mu-$ 位置参数



【2】 $f(x)$ 在 $(-\infty,\mu)$ 内单调增加，在 $(\mu,+\infty)$

内单调下降， $\max f(x)=f(\mu)=\dfrac{1}{\sqrt{2\pi}\,\sigma}$.

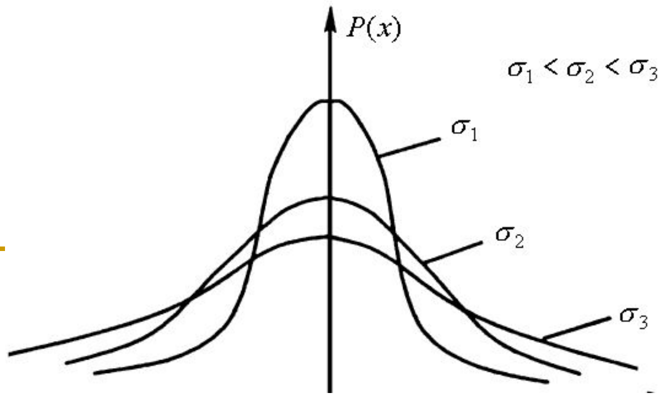
【3】 $\lim\limits_{x\to\infty} f(x)=0$, 所以 x 轴为 $f(x)$ 的水平渐近线.

【4】 $f(x)$的图形是中间大，两头小，呈对称的“钟形”曲线.

σ越大, $f(x)$的图形越平坦, σ越小, $f(x)$的图形越陡峭.

并且 $f(x)$ 在 $x=\mu\pm\sigma$ 处取到拐点.

$\sigma-$ 刻度参数



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

【5】 函数 $F(x)=\dfrac{1}{\sqrt{2\pi}\,\sigma}\displaystyle\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt$ 不可用初等函数表示，因此给 $F(x)$ 以及概率

$$P\{a<X\le b\}=F(b)-F(a)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{a}^{b} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

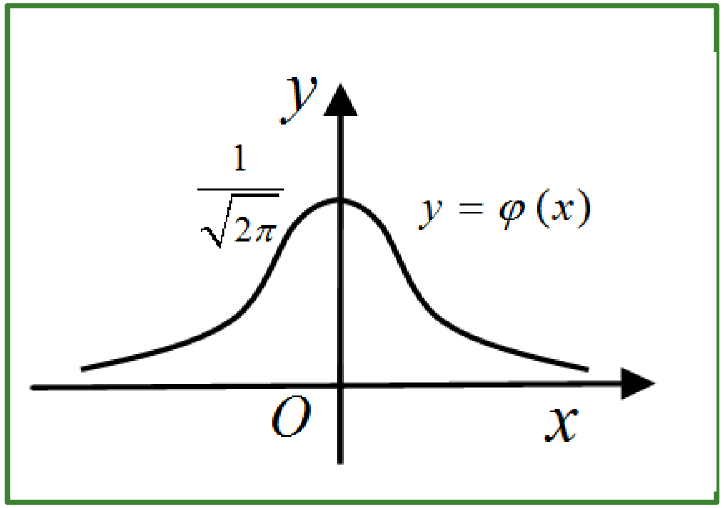
等计算带来不便，为此引入标准正态分布的概念.

【6】当 $\mu=0,\sigma=1$, 即 $X \sim N(0,1)$ 时，称随机变量X服从标准正态分布，其密度函数记为$\varphi(x)$,分布函数记为$\Phi(x)$. 即有 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{x^2}{2}}dx.$

对于给定的$x>0$, $\Phi(x)$可以通过查标准正态分布表近似求得, 且由$\varphi(x)$对称性,

$$\Phi(0)=0.5,$$

当$x<0$时，$\Phi(x)=1-\Phi(-x).$



$\varphi(x)-$ 偶函数

【7】如果 $X \sim N(0,1)$, 则对任意的实数 $a, b(a<b)$, 有

$$P\{a < X \le b\} = \Phi(b) - \Phi(a).$$

【8】如果 $X \sim N(\mu,\sigma^2)$则可得 $\dfrac{X-\mu}{\sigma} \sim N(0,1)$, 从而

$$P\{a < X \le b\} = P\{\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\}$$

$$= \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$$

例6 设随机变量 $X \sim N(1,4)$，分别计算 $P\{X \le 3\}, P\{-1 < X < 5\}$.

解 由题意知 $\mu = 1, \sigma = 2$.

$$P\{X \le 3\} = \Phi(\frac{3-1}{2}) = \Phi(1) = 0.8413.$$

$$\begin{aligned} P\{-1 < X < 5\} &= \Phi(\frac{5-1}{2}) - \Phi(\frac{-1-1}{2}) \\ &= \Phi(2) - \Phi(-1) \\ &= \Phi(2) - [1 - \Phi(1)] \\ &= 0.9772 + 0.8413 - 1 \\ &= 0.8185. \end{aligned}$$

例7 设随机变量$X \sim N(\mu,\sigma^2)$, 计算$P\{|X-\mu|<k\sigma\}$, 其中k分别取1,2,3.

解 $P\{|X-\mu|<k\sigma\}=P\{-k\sigma<X-\mu<k\sigma\}$

$$=P\{-k<\frac{X-\mu}{\sigma}<k\}=\Phi(k)-\Phi(-k)=2\Phi(k)-1$$

因此

$$P\{|X-\mu|<\sigma\}=2\Phi(1)-1=0.6826,$$

$$P\{|X-\mu|<2\sigma\}=2\Phi(2)-1=0.9544,$$

$$P\{|X-\mu|<3\sigma\}=2\Phi(3)-1=0.9974,$$

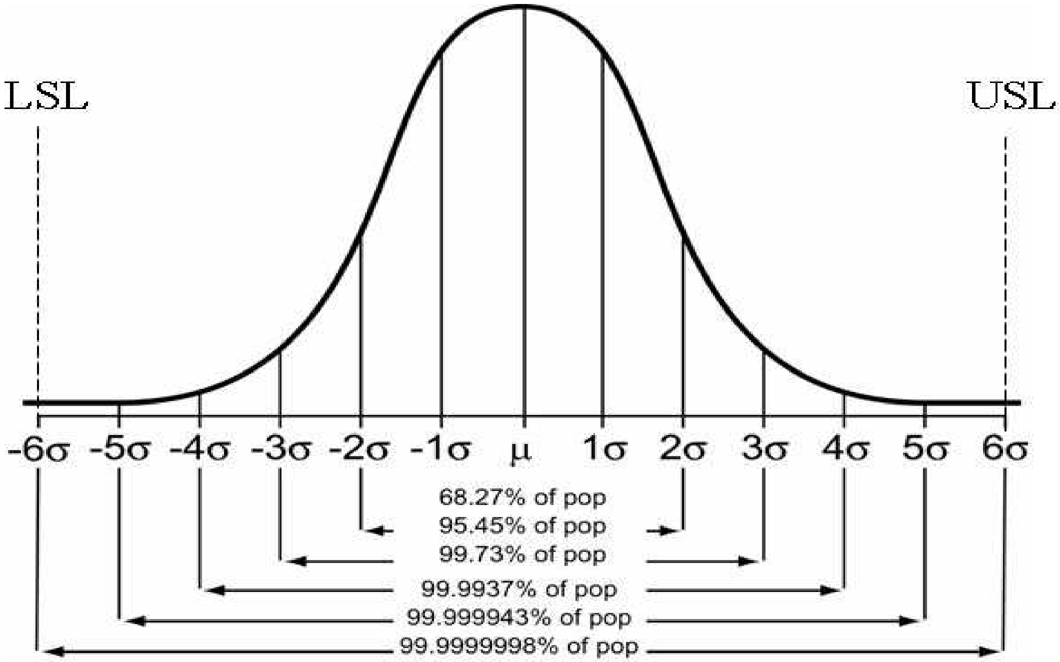
特别地 $P\{|X-\mu|<3\sigma\}=2\Phi(3)-1=0.9974\approx 1,$

这表明X的取值基本上都落在区间$(\mu-3\sigma,\mu+3\sigma)$内,而在其外的可能性很小,因此通常也称为“$3\sigma$”准则.

六西格玛“6σ”管理

六西格玛（Six Sigma,）是20世纪80年代有摩托罗拉公司（Motorola）提出的概念和相应的管理体系，从开始实施的1996年到1999年，公司平均每年提高生产率12.3%，不良率只有以前的1/20，其创建此概念管理，主要在于20世纪60年代，日本从美国引入了质量管理的思想，使产品质量有了大幅度的提升。到了20世纪70年代末，80年代初，日本产品凭借过硬的品质，从美国人手中抢占了大量的市场份额。美国的摩托罗拉公司在同日本组织的竞争中，先后失去了收音机、电视机、半导体等市场，到了1985年公司频临倒闭。面对残酷的竞争和严峻的生存形式，摩托罗拉公司痛定思痛，得出了这样的结论：“摩托罗拉失败的根本原因是其产品质量比日本组织同类产品的质量差很多”，从而最终总结创建了此管理理念.

“六西格玛”管理理念和方法在摩托罗拉、通用电气、戴尔、惠普、西门子、索尼、东芝、华硕等众多跨国企业的实践中被证明是卓有成效的， 被认为是工业界加强质量管理提高质量水平的一件大事。为此，国内一些部门和机构在国内企业大力推广管理工作，引导企业发展管理。现在“六西格玛”作为质量水平的含义已经泛化，不管质量指标能否用正态分布刻画，只要产品的不合格率不超过“6σ”，人们就说质量达到了六西格玛水平。



百万件产品中不合格数不多于4个.

例8 （课本P57例2.3.8）

在电源电压不超过200伏，在200-240伏和超过240伏三种情况下，某种电子元件损坏的概率分别 0.1,0.001, 0.2. 假设电源电压 X 服从正态分布 $N(220, 25^2)$,

试求：(1) 该电子元件损坏的概率 α.

(2) 该电子元件损坏时，电源电压在200−240伏的概率 β.

解 设 $A_1=\{X\le 200\}, A_2=\{200<X\le 240\}, A_3=\{X>240\}$,

B 表示该电子元件损坏. 由于 $X \sim N(220,25^2)$, 因此

$$P(A_1)=P\{X\le 200\}=P\{\frac{X-220}{25}\le -0.8\}$$

$$=\Phi(-0.8)=1-\Phi(0.8)=1-0.788=0.212,$$

$$P(A_2)=P\{200<X\le 240\}=P\{-0.8<\frac{X-220}{25}\le -0.8\}$$

$$=\Phi(0.8)-\Phi(-0.8)=2\Phi(0.8)-1=0.576,$$

$$P(A_3)=P\{X>240\}=P\{\frac{X-220}{25}>0.8\}$$

$$=1-P\{\frac{X-220}{25}\le 0.8\}=1-\Phi(0.8)=0.212.$$

续解

又由题意知 $P(B|A_1)=0.1, P(B|A_2)=0.001, P(B|A_3)=0.2.$

（1）由全概率公式知

$$\alpha=P(B)=\sum_{i=1}^{3}P(A_i)\mathrm{P}(B|A_i)=0.0642;$$

（2）由贝叶斯公式知

$$\beta=P(A_2|B)=\frac{P(A_2)P(B|A_2)}{P(B)}=0.009.$$

2016 数学一、三

设随机变量$X \sim N(\mu,\sigma^2)(\sigma>0)$, 记$p=P\{X\le\mu+\sigma^2\}$,

则（ ）

$(A)\,p$ 随着μ 的增加而增加. $(B)\,p$ 随着σ的增加而增加.

$(C)\,p$随着μ 的增加而减少. $(D)\,p$ 随着σ 的增加而减少.

解 $$p=P\{X\le\mu+\sigma^2\}=\Phi\left(\frac{\mu+\sigma^2-\mu}{\sigma}\right)=\Phi(\sigma)$$

分布函数$\Phi(x)$ 为单调递增函数，故答案为B.

2013 数学一、三

设 X_1, X_2, X_3 是随机变量，且 $X_1 \sim N(0,1), X_2 \sim N(0,2^2)$,

$X_3 \sim N(5,3^2), p_i = P\{-2 \le X_i \le 2\}(i=1,2,3)$，则（　　）

$(A)\ p_1 > p_2 > p_3.$　　　$(B)\ p_2 > p_1 > p_3.$

$(C)\ p_3 > p_1 > p_2.$　　　$(D)\ p_1 > p_3 > p_2.$

解　$p_1 = P\{-2 \le X_1 \le 2\} = 2\Phi(2) - 1, p_2 = P\{-2 \le X_2 \le 2\} = 2\Phi(1) - 1$

$p_3 = P\{-2 \le X_3 \le 2\} = \Phi(-1) + \Phi(7/3) - 1$

由正态分布函数的单调增加知 $\Phi(2) > \Phi(1)$，知 $p_1 > p_2$.

又因为 $p_3 - p_2 = \Phi(-1) - \Phi(1) + \Phi(7/3) - \Phi(1)$

$= \varphi(\xi_1)\cdot(-2) + \varphi(\xi_2)\cdot(4/3),\qquad (-1 < \xi_1 < 1, 1 < \xi_2 < 7/3)$

$< \varphi(\xi_1)(-2 + \frac{4}{3}) < 0.$　故 $p_2 > p_3$.　　答案为 C.

2004 数学一、三

随机变量X服从$N(0,1)$,对给定的α（$0<\alpha<1$),均有$\cdot P(X>u_{\alpha}\}=\alpha$,

若$P(|X|<x)=\alpha$,则$x=?$

§2.4 随机变量函数的分布

一、离散型随机变量函数的分布

例1 设随机变量 $X \sim \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{6} & \frac{1}{4} & \frac{1}{3} \end{pmatrix}$，分别求出 $Y = X^2$ 和 $Z = \max\{X, 1\}$ 的分布律.

解 作下列列表计算

X	-1	0	1	2
P	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$
Y	1	0	1	4
Z	1	1	1	2

作适当合并得

$$Y \sim \begin{pmatrix} 0 & 1 & 4 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix}, Z \sim \begin{pmatrix} 1 & 2 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

设离散型随机变量 X 的分布律为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

则求 $Y = g(X)$ 的分布律的步骤为：

第一步：在X的分布律中添加一行 Y=$g(X)$, 并将计算 $y_i = g(x_i)\,(i = 1, 2, \cdots)$ 的值对应填入该行中.

第二步：对其中 Y 取值相同的项适当进行概率合并, 即得 Y 的分布律.

例2 随机变量X的分布律为$P\{X=k\}=\dfrac{1}{2^k}, k=1,2,\cdots.$

求 $Y=\sin(\dfrac{\pi}{2}X)$ 的分布律.

解 由题意知Y的取值为 -1,0,1, 并且

$$P\{Y=-1\}=\sum_{I=1}^{\infty}P\{X=4i-1\}=\sum_{i=1}^{\infty}\frac{1}{2^{4i-1}}=\frac{2}{15};$$

$$P\{Y=0\}=\sum_{I=1}^{\infty}P\{X=2i\}=\sum_{I=1}^{\infty}\frac{1}{2^{2i}}=\frac{1}{3};$$

$$P\{Y=1\}=\sum_{i=1}^{\infty}P\{X=4i-3\}=\sum_{i=1}^{\infty}\frac{1}{2^{4i-3}}=\frac{8}{15};$$

即 $$Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \dfrac{2}{15} & \dfrac{1}{3} & \dfrac{8}{15} \end{pmatrix}.$$

二、连续型随机变量的函数$Y=g(X)$的分布

例3 设随机变量 $X \sim N(\mu,\sigma^2)$, 求$Y=aX+b$ 的密度函数 $f_Y(y)$, 其中a,b为常数, 且$a\neq 0$.

解 先求$Y=aX+b$的分布函数$F_Y(y)$,再求其密度函数.

因 $F_Y(y)=P\{Y\le y\}=P\{aX+b\le y\}$

当 $a>0$ 时, $F_Y(y)=P\{X\le \frac{y-b}{a}\}=\int_{-\infty}^{\frac{y-b}{a}} f_X(x)dx$

上式关于 y 两边同时求导

$$f_Y(y)=f_X(\frac{y-b}{a})\cdot(\frac{y-b}{a})'=\frac{1}{a}\cdot f_X(\frac{y-b}{a})$$

类似, 当 $a<0$ 时, $f_Y(y)=-\frac{1}{a}\cdot f_X(\frac{y-b}{a})$,

分布函数法

综上, $Y=aX+b$的密度函数为

$$f_X(x)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)=\frac{1}{|a|}\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma|a|}e^{-\frac{(y-(a\mu+b))^2}{2\sigma^2a^2}},\ y\in(-\infty,+\infty).$$

(1) 当 $X\sim N(\mu,\sigma^2), Y\sim N(a\mu+b,a^2\sigma^2)$.

即　正态分布的线性函数仍为正态分布.

(2) 当 $X\sim N(\mu,\sigma^2),\ Y=\dfrac{X-\mu}{\sigma}\sim N(0,1)$.

称为随机变量的标准化.

定理1 设随机变量 X 的密度函数为$f_X(x)$,$y=g(x)$为$(-\infty,+\infty)$内为严格单调函数，其反函数 $x=h(y)$ 具有一阶连续导数，则$Y=g(X)$ 为连续型随机变量，且其密度函数为

$$f_Y(y)=\begin{cases} f_X(h(y))\left|h'(y)\right|, & \alpha<y<\beta, \\ 0, & y\in other. \end{cases}$$

其中 $\alpha=\min\{g(-\infty),g(+\infty)\},\beta=\max\{g(-\infty),g(+\infty)\}.$

【注1】

当 $f_X(x)$ 在有限区间$[a,b]$之外取值为零时，只需假设 $y=g(x)$ 在$[a,b]$上严格单调，反函数具有连续导数，则该结论仍成立，此时

$$\alpha = \min\{g(a), g(b)\}, \beta = \max\{g(a), g(b)\}.$$

【注2】

定理中的“严格单调”可以减弱为“逐段严格单调”，分别在每段区间上使用定理.

例4 设随机变量$X \sim N(0,1)$, 求$Y = X^2$的密度函数.

解 当$y \le 0$时 $F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\} = P\{\phi\} = 0$

当$y > 0$时 $F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\}$

$$= P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx$$

$$f_Y(y) = F_Y'(y) = f_X(\sqrt{y}) \cdot (\sqrt{y})' - f_X(-\sqrt{y}) \cdot (-\sqrt{y})'$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot \frac{1}{2\sqrt{y}} - \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot \left(-\frac{1}{2\sqrt{y}}\right) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}$$

即 $Y = X^2$的密度函数为
$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

$$Y = X^2 \sim \chi^2(1).$$

柯（ke）西分布-Cauchy distribution

柯西分布是一个数学期望不存在的连续型概率分布。当随机变量X满足它的概率密度函数时，称X服从柯西分布。

中文名	柯西分布	参　数	位置参数，尺度参数
外文名	Cauchy distribution	定义域	全体实数
提出者	柯西(Cauchy)	特　点	数学期望，方差，高阶矩均不存在

例5 若$\theta \sim U(-\frac{\pi}{2},\frac{\pi}{2})$, 求$Y=\tan\theta$的密度函数.

解

$$f(\theta)=\begin{cases}\frac{1}{\pi}, & -\frac{\pi}{2}<\theta<\frac{\pi}{2},\\ 0, & \theta\in other.\end{cases}$$

又因为函数$y=\tan\theta$在区间$(-\frac{\pi}{2},\frac{\pi}{2})$ 严格单调且反函数 $\theta=\arctan y$具有一阶连续可导, 由定理1可得Y的密度函数为

$$f_Y(y)=\frac{1}{\pi}\cdot\frac{1}{1+y^2},\quad y\in(-\infty,+\infty).$$

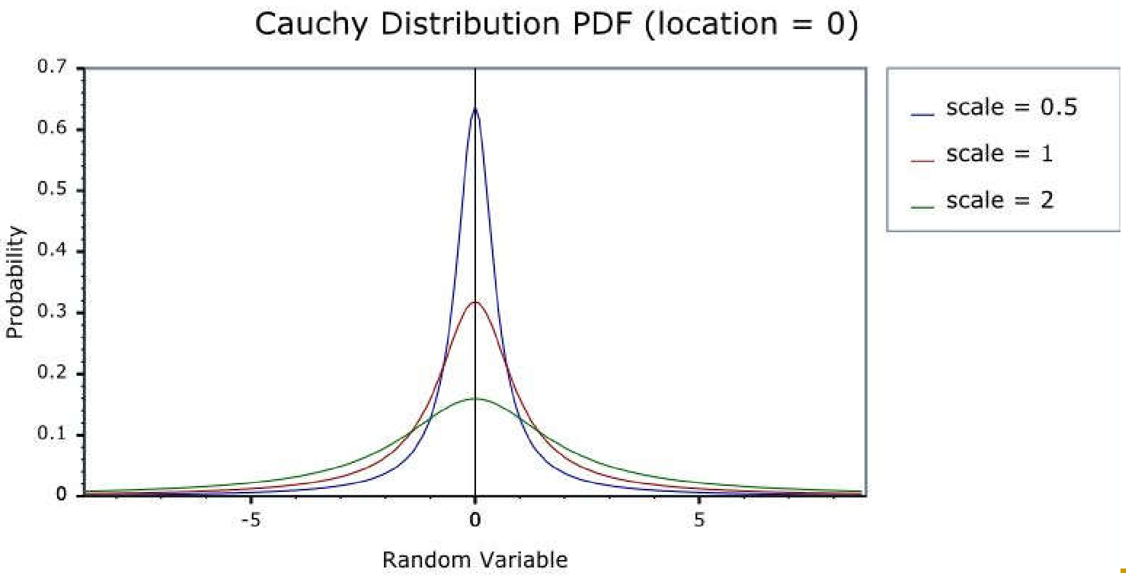
定义 若随机变量X的密度函数为

$$f(x)=\frac{1}{\pi}\cdot\frac{d}{d^2+(x-\mu)^2},\ d>0,-\infty\le\mu\le+\infty.$$

则称 X服从参数为μ,d 的柯西分布, 记为$X\sim C(\mu,d)$.

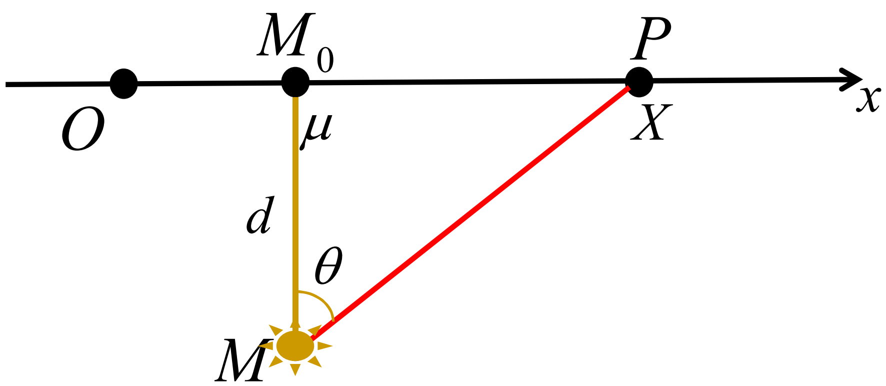
特别当$\mu=0,d=1$时，称为标准柯西分布.

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},$$



柯西密度曲线与正态分布很类似, 但趋于0的速度要慢很多, 即重尾.

柯西分布的直观几何意义：



设在平面上点M处有一束光射向一条直线，与直线的交点为P,M到该直线的垂足为M_0，射线与垂线的夹角为θ,（当P点在M_0之右，规定$\theta > 0$，P点在M_0之左，规定$\theta < 0$）

MM_0的长度为d,设该直线为数轴,原点在O,M_0的坐标为μ,

P点的坐标是X，则

$X = d\tan\theta + \mu$，若$\theta \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$，则 $X \sim C(\mu, d)$.

例6 设随机变量$X \sim U[-1,2]$，求 $Y=\operatorname{sgn}(X)$ 的概率分布.

解

$$P\{Y=-1\}=P\{X<0\}=\frac{1}{3};$$

$$P\{Y=0\}=P\{X=0\}=0;$$

$$P\{Y=1\}=P\{X>0\}=\frac{2}{3};$$

$$Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}.$$

例7 设随机变量 $X \sim e(\lambda)$, 求 $Y=\min\{X,2\}$ 的分布函数.

解 $F_Y(y)=P\{Y\le y\}=P\{\min\{X,2\}\le y\}$

当 $y<0$ 时, $\{\min\{X,2\}\le y\}=\phi, F_Y(y)=0.$

当 $0\le y<2$ 时, $F_Y(y)=P\{\min\{X,2\}\le y\}=P\{X\le y\}$

$$=\int_0^y \lambda e^{-\lambda x}dx=1-e^{-\lambda y},$$

当 $y\ge 2$ 时 $\{\min\{X,2\}\le y\}=\Omega, F_Y(y)=1.$

故 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0,\\ 1-e^{-\lambda y}, & 0\le y<2,\\ 1, & y\ge 2.\end{cases}$$

【注】该例中的随机变量 Y 既不是离散型也不是连续型随机变量.

2013 数学一

设随机变量X的概率密度为 $f(x)=\begin{cases}\dfrac{1}{9}x^2, & 0<x<3,\\ 0, & \text{其他}.\end{cases}$

令随机变量

$$Y=\begin{cases}2, & X\le 1,\\ X, & 1<X<2,\\ 1, & X\ge 2.\end{cases}$$

（1）求Y的分布函数；

（2）求概率 $P\{X\le Y\}$.

解（1）由题设条件知 $P\{1 \le Y \le 2\} = 1$

记 Y 的分布函数为 $F_Y(y)$，则

当 $y < 1$ 时，$F_Y(y) = 0$;

当 $1 \le y < 2$ 时，$F_Y(y) = P\{Y \le y\} = P\{Y = 1\} + P\{1 < Y \le y\}$

$$= \int_2^3 \frac{1}{9} x^2 dx + \int_1^y \frac{1}{9} x^2 dx = \frac{y^2 + 18}{27};$$

当 $y \ge 2$ 时 $F_Y(y) = 1.$

所以 Y 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 1, \\ \dfrac{y^2 + 18}{27}, & 1 \le y < 2, \\ 1, & y \ge 2. \end{cases}$$

（2）$P\{X \le Y\} = P\{X \le 2\} = \int_0^9 \frac{1}{9} x^2 dx = \frac{8}{27}.$

点评

当X为连续型随机变量，函数$Y=g(X)$可能会出现各种情形：Y为离散型，连续型，非离散非连续型等.

当X为连续型随机变量,函数$Y=g(X)$为非离散型随机变量时,此时可以利用分布函数法,先求出分布函数$F_Y(y)$, 若根据结果分析出Y为连续型随机变量，则进一步可求得Y的密度函数为 $f_Y(y)=F_Y'(y), \quad -\infty<y<+\infty.$

分布函数法重点掌握,难点在于:在计算

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{g(x)\le y} f_X(x)dx, -\infty<y<+\infty.$$

的过程中,经常需要对变量 y 进行分段讨论.

例9　课本P69页第34题（1）

设随机变量X的分布函数 $F(x)$，求 $Y=F(X)$ 的密度函数.

解　由于函数 $y=F(x)$ 的取值范围为$[0,1]$, 可知

当 $y<0$ 时, $F_Y(y)=P\{Y\le y\}=P\{F(X)\le y\}=P\{\phi\}=0$,

当 $y>1$时, $F_Y(y)=P\{Y\le y\}=P\{F(X)\le y\}=P\{\Omega\}=1$,

当$0\le y\le 1$时　分布函数单调不减，即函数 $y=F(x)$ 单调不减, 故

$$F_Y(y)=P\{F(X)\le y\}=P\{X\le F^{-1}(y)\}=F(F^{-1}(y))=y,$$

所以Y为连续型随机变量, 其密度函数

$$f_Y(y)=\begin{cases}1, & y\in[0,1]\\ 0, & y\in other.\end{cases}$$

$Y=F(X)$服从$[0,1]$上均匀分布

合肥工业大学
HEFEI UNIVERSITY OF TECHNOLOGY

例5的结论等价于

设 $F(x)$ 是任何分布函数, 若 U 是[0,1]上服从均匀分布的随机变量, 且

$$X = F^{-1}(U)$$

则 X 的分布函数恰好是 $F(x)$.

该结论可用于随机变量的模拟生成, 非常有应用价值.

基本原理

编辑

随机模拟方法以随机模拟和统计实验为手段，是一种从随机变量的概率分布中，通过随机选择数字的方法产生一种符合该随机变量概率分布特性的随机数值序列，作为输入变量序列进行特定的模拟实验、求解的方法。在应用方法时，要求产生的随机数序列应符合该随机变量特定的概率分布。而产生各种特定的、不均匀的概率分布的随机数序列，可行的方法是产生一种均匀分布的随机数字列，然后再设法转换成特定要求的概率分布的随机数字列，以此作为数字模拟试验的输入变量序列进行模拟求解。[5]