02

随机变量及其分布

《概率论与数理统计》淮阴师范学院数科学院

目录/Contents

目录/Contents

2.1 随机变量及其分布

一、随机变量的定义

许多随机试验的结果与实数密切联系，也有些随机试验结果从表面上看并不与实数相联系．下面我们通过几个例子来引入随机变量的概念.

例 1

抛掷一颗均匀的骰子，出现的点数 X的取值样本空间={正面朝上, 反面朝上},样本空间不是一个数集．但是我们可以人为地把试验结果和实数对应起来.令

样本点		X的取值
正面朝上	→	1
反面朝上	→	0

$$\{X=1\}=\{正面向上\},\{X=0\}=\{反面向上\}$$

一、随机变量的定义

定义1 在随机试验E中，Ω是相应的样本空间，如果对Ω中的每一个样本点ω，有唯一一个实数$X(\omega)$与它对应，那么就把这个定义域为Ω的单值实值函数

$X = X(\omega)$称为是（一维）**随机变量**.

随机变量一般用大写字母X, Y, Z表示.

引进随机变量后，随机事件及其概率可以通过随机变量来表达.

一、随机变量的定义

- 随机变量
  - 离散型随机变量
  - 连续型随机变量

A 如果一个随机变量仅可能取有限或可列个值，则称其为**离散型随机变量**、

B 如果一个随机变量的取值充满了数轴上的一个区间（或某几个区间的并），则称其为**连续型随机变量**。

一、随机变量的定义

随机变量的直观解释

随机变量 X 是样本点的函数，这个函数的自变量是样本点，可以是数，也可以不是数，定义域是样本空间，而因变量必须是实数。这个函数可以让不同的样本点对应不同的实数，也可以让多个样本点对应于一个实数。

二、随机变量的分布函数

定义2

给定一个随机变量X，对任意实数$x \in (-\infty, +\infty)$，称函数$F(x) = P(X \leq x)$为随机变量X的**分布函数**.对任意满足条件$-\infty < a < b < +\infty$的实数$a, b$，有

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

（1）分布函数是定义在$-\infty < x < +\infty$上，取值在[0,1]上的一个函数。

（2）任一随机变量X都有且仅有一个分布函数，有了分布函数，就可计算与随机变量X相关的事件的概率问题。

二、随机变量的分布函数

例1 设一盒子中装有10个球，其中5个球上标有数字1, 3个球上标有数字2，2个球上标有数字3。

从中任取一球, 记随机变量X为“取得的球上标有的数字”

（1）写出X的分布函数$F(x)$;

（2）作出分布函数$F(x)$的图像.

解 容易得到X可取1,2,3，由古典概型的计算方法，对应的概率值分别为0.5,0.3,0.2。

由分布函数定义知

若$x < -1$, 则$\{X \le x\}$为不可能事件

$$F(x)=P(X \le x)=0$$

若$1 \le x < 2$, 则$\{X \le x\} = \{X = 1\}$

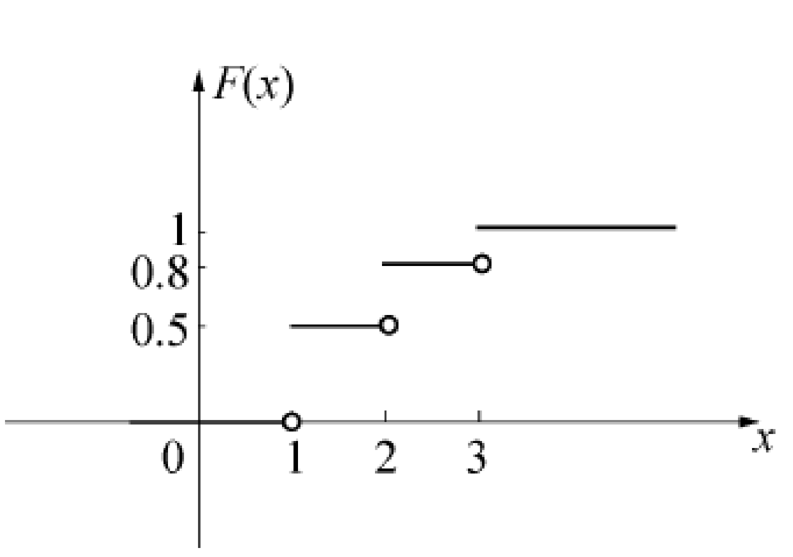
$$F(x)=P(X \le x)=P(X=1)=0.5$$

同理，当 $2 \le x < 3$时，有 $F(x)=P(X \le x)=P(X=1)+P(X=2)=0.8$

当$x \ge 3$时，有 $F(x)=P(X \le x)=P(X=1)+P(X=2)+P(X=3)=1$

二、随机变量的分布函数

综上，随机变量的分布函数为 $F(x)=P(X\le x)$



$$=\begin{cases}0 & x<1\\ 0.5 & 1\le x<2\\ 0.8 & 2\le x<3\\ 1 & x\ge 3\end{cases}$$

分布函数的性质

设$F(X)$是随机变量X的分布函数，则有

1. $0 \le F(x) \le 1;\quad \lim_{x\to-\infty} F(x) = 0,\ \lim_{x\to+\infty} F(x) = 1$
2. 分布函数单调不减;
3. 对任意的$x \in (-\infty,+\infty)$，分布函数右连续;
4. $P(X < x) = F(x^-)$

定义3 设 $\Omega_x = \{x_1, x_2, \ldots x_n \ldots\}$ 且$P(X = x_i) = p_i$

其中 p_i 满足:

（1）非负性 $p_i \geq 0(i = 1, 2, \cdots);$

（2）规范性 $\sum_{i=1}^{\infty} p_i = 1,$

那么称表达式$P(X = x_i) = p_i, i = 1,2, \ldots$ 为随机变量X**的分布律或概率函数.**

三、离散型随机变量及其分布律

换句话说，如果一个随机变量只可能取有限个值或可列无限个值，那么称这个随机变量为（一维）**离散型随机变量**.

一维离散型随机变量的**分布律也可表示为**：

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

例 2 设随机变量X的分布律如下：

X	-1	0	2
概率	0.2	0.4	0.4

.

求 （1）$P(X \le -0.7)$ （2）X的分布函数$F(x)$

解 （1）$P(X \le -0.7) = P(X = -1) = 0.2$，

（2） X 的分布函数 $F(x)$ 求解过程同例 1，可得 $F(x) = \begin{cases} 0, & x < -1 \\ 0.2, & -1 \le x < 0 \\ 0.6, & 0 \le x < 2 \\ 1, & x \ge 2 \end{cases}$

定义4

给定一个连续型的随机变量X，如果存在一个定义域为 $(-\infty,+\infty)$ 的非负实值函数$f(x)$，使得X的分布函数$F(x)$可以表示为

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t \qquad -\infty<x<+\infty$$

那么称$f(x)$为连续型随机变量X的概率密度函数.

四、连续型随机变量及其密度函数

概率密度函数满足下面两个条件:

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

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

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

对照一下离散型随机变量的概率函数所满足的两个条件,

1 $p_i \geq 0$

2 $\sum_i p_i = 1$

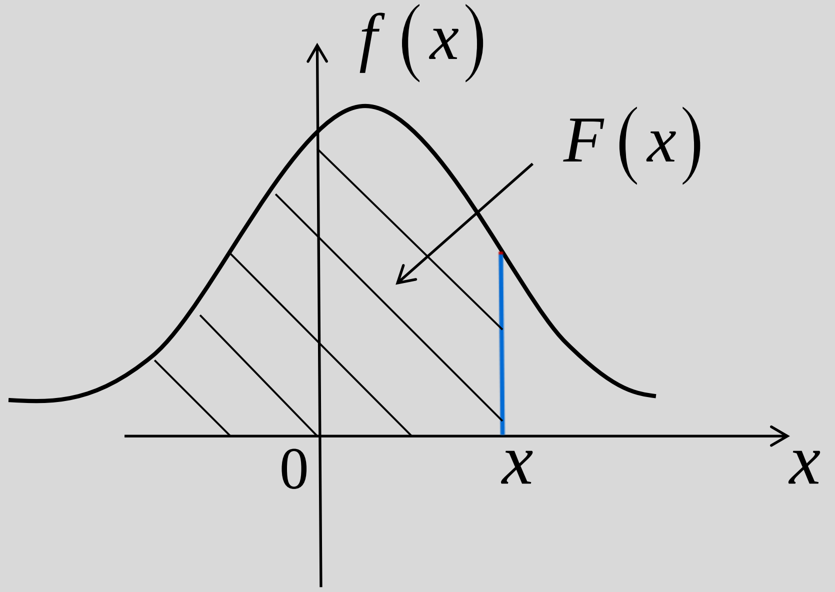
这两个条件同样刻划了密度函数的特征性质，即如果有实值函数具备这两条性质，那么它必定是某个连续型随机变量的概率密度函数.

四、连续型随机变量及其密度函数

分布函数和概率密度函数的关系在几何上的体现:

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

$$-\infty < x < +\infty$$



连续型随机变量的性质

设X是任意连续型随机变量，且$F(x)$与$f(x)$分别是它的分布函数与概率密度函数，则有:

1. $f(x)$是连续函数，且在$F(x)$的连续点处，有

$$F'(x) = f(x);$$

2. 对任意常数$c(-\infty < c < +\infty)$有$P(X = c) = 0$;

结合结论2可知:

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

例 3

设随机变量X的概率密度函数为

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

求

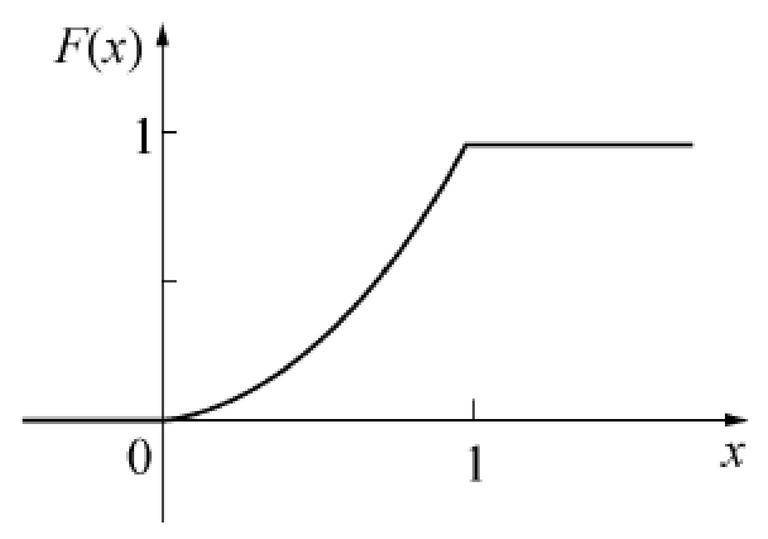
（1）$P(|X|<0.5)$；（2）X的分布函数$F(x)$

解

（1）$P(|X|<0.5)=\int_{-0.5}^{+0.5} f(x)\mathrm{d}x=\int_{0}^{+0.5} 3x^2\mathrm{d}x=0.125$

解 (2)

$$F(x)=\begin{cases}0 & x<0\\ \int_0^x 3t^2\mathrm{d}t & 0\le x<1\\ \int_0^1 3x^2\mathrm{d}x & x\ge 1\end{cases}=\begin{cases}0 & x<0\\ x^3 & 0\le x<1\\ 1 & x\ge 1\end{cases}$$



目录/Contents

目录/Contents

2.2 常用的离散型随机变量

一、二项分布

设对一随机试验 E，我们只关心某个事件 A 发生与否,此时试验的结果可以看成只有两种: A 发生或者 $\bar{A}$ 不发生。那么称这个试验为**贝努利试验.**

在 n 重贝努利试验中，若以X 表示事件 A 在n 次试验中出现的次数.

则 X的取值为$0,1,2,\cdots,n$, 相应的概率为:

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

分布律为

$$\begin{array}{c|cccccccc} X & 0 & 1 & \cdots & \cdots & k & & \cdots & n \\ \hline P & (1-p)^n & C_n^1 p(1-p)^{n-1} & & \cdots & C_n^k p^k(1-p)^{n-k} & & \cdots & p^n \end{array}.$$

一、二项分布

其中p 为事件A 发生的概率. 则称X 服从参数为n,p 的二项分布，记成

$$X \sim B(n,p).$$

在概率论中，二项分布是一个重要的分布. 在许多独立重复试验中，都具有二项分布的形式.

若二项分布$X \sim B(\mathrm{n},\mathrm{p})$中取$n = 1$，相应的分布律为

$$\begin{array}{c|cc} X & 0 & 1 \\ \hline P & 1-p & p \end{array},$$

即随机变量X的取值为 **0, 1**，相应的概率记为

$$P(X=1)=p, \qquad P(X=0)=1-p \,(0<p<1),$$

则又称服从 $0-1$分布（或两点分布）.

例 4

某人向同一目标重复独立射击5次，每次命中目标的概率为0.8，求（1）此人能命中3次的概率；（2）此人至少命中2次的概率。

设 X 表示在5次重复独立射击中命中的次数，则

$$X \sim B(5,0.8)$$

1 $P(X=3)=\binom{5}{3}\times 0.8^3\times 0.2^2=0.2048$

2 $P(X\geq 2)=1-P(X<2)=1-P(X=0)-P(X=1)=0.99328$

设随机变量 X 的概率密度函数为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\qquad k=0,1,2,\ldots;\quad \lambda>0$$

则称 X 服从参数为 λ 的泊松分布, 记为 $X\sim P(\lambda)$.

由无穷级数知识知: $\sum\limits_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=1$

二、泊松分布

泊松分布也是一种常用的离散型分布，它常常与计数过程相联系，例如

01 OPTION

某一时段内某网站的点击量；

02 OPTION

早高峰时间段内驶入高架道路的车辆数；

03 OPTION

一本书上的印刷错误数。

例 5 设随机变量 X 有分布律 $P(X=k)=\dfrac{c\times 3^k}{k!}(k=0,1,2,\cdots)$，求 c 的值，并求解 $P(X\le 2)$.

解

根据分布律的定义有$\sum_{k=0}^{+\infty}\frac{c\times 3^k}{k!}=1\Rightarrow c=e^{-3}$.

事实上，不难看出$X\sim P(3)$，所以$c=e^{-3}$。

$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)$$

$$=\frac{e^{-3}\times 3^0}{0!}+\frac{e^{-3}\times 3^1}{1!}+\frac{e^{-3}\times 3^2}{2!}=\frac{17}{2}e^{-3}。$$

例 6 已知一购物网站每周销售的某款手表的数量X服从参数为6的泊松分布.问周初至少预备多少货源才能保证该周不脱销的概率不小于0.9.假定上周没有库存,且本周不再进货.

解 设该款手表每周的需求量为X，则有$X \sim P(6)$；设至少需要进n块该款手表，才能满足不脱销的概率不小于0.9，即要满足

$$P(X \le n) \ge 0.9$$
$$P(X \le n\text{-}1) < 0.9$$

解得$P(X \le 8)$=0.847237，$P(X \le 9)$=0.916076

所以周初预备 9 块时，能满足 90%的顾客需求而不脱销。

定理（泊松定理）

在 n 重贝努利试验中，记 A 事件在一次试验中发生的概率为 p_n，当 $n \to +\infty$ 时，有 $np_n \to \lambda$

对于任意一个非负整数 k，有

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

泊松定理告诉我们：满足一定条件时，二项概率可以用泊松分布的概率值来近似.

例 7 设某保险公司的某人寿保险险种有1000人投保，每个投保人在一年内死亡的概率为0.005，且每个人在一年内是否死亡是相互独立的，试求在未来一年中这1000个投保人中死亡人数不超过10人的概率.

解 记X 未来一年中这1000个投保人中死亡人数，则有

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

此时可近似看作参数为5的泊松分布，

$$P(X \le 10) = \sum_{k=0}^{10} \binom{1000}{k} \times 0.005^k \times 0.995^{1000-k}$$

$$\approx \sum_{k=0}^{10} \frac{5^k}{k!} e^{-5} = 0.986305$$

三、超几何分布

设有n件产品，其中M件次品．现从中不放回任取n个产品，$(n \le N)$．

则这n个产品中所含的次品X 的分布律为

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

$$k = \max\{0, n+M-N\}, 2, \cdots, \min\{M, n\}.$$

我们称 X 服从参数为n， N 和 M的**超几何分布**.

记 $p=\dfrac{M}{N}$，可以证明，有

$$\lim_{N\to\infty}\frac{\dbinom{M}{k}\dbinom{N-M}{n-k}}{\dbinom{N}{n}}=\binom{n}{k}p^k(1-p)^{n-k}.$$

在实际应用中，当$n \ll N$ 时，即抽取个数n远小于产品总数N时，每次抽取后，总体中的不合格品率 $p=\dfrac{M}{N}$ 改变很微小，所以不放回抽样可以近似地看出放回抽样，这是超几何分布可用二项分布近似。

四、几何分布与负二项分布

如果随机变量 X 的分布律为

X	1	2	3	$\cdots$	k	$\cdots$
概率	p	$p(1-p)$	$p(1-p)^2$	$\cdots$	$p(1-p)^{k-1}$	$\cdots$

则称 X 服从参数为 p 的**几何分布**.

四、几何分布与负二项分布

几何分布也是一种常用的离散型分布，例如

01 OPTION

抛掷一颗均匀的骰子，首次出现 6 点时的投掷次数 $X \sim Ge(\frac{1}{6})$；

02 OPTION

首次投篮命中时投篮的次数 $X \sim Ge(p)$， p 为每次投篮时的命中率；

03 OPTION

任课教师每次上课随机抽取 10%的学生签到，某位学生首次被老师要求签到时总的上课次数 $X \sim Ge(0.1)$.

例 8 设 $X\sim Ge(p)$，则对任意正整数 m 和 n，证明 $P(X>m+n|X>m)=P(X>n)$

证明 可以解得 $P(X>n)=\sum_{i=n+1}^{+\infty}(1-p)^{i-1}p=\frac{p(1-p)^n}{1-(1-p)}=(1-p)^n$

故 $P(X>m+n|X>m)=\frac{P(X>m+n)}{P(X>m)}=\frac{(1-p)^{m+n}}{(1-p)^m}=(1-p)^n=P(X>n)$

这个例题说明，几何分布具有无记忆性的性质.

四、几何分布与负二项分布

负二项分布是几何分布的一个延伸.在伯努利试验中，记每次试验中 A 事件发生的概率 $P(A)=p(0<p<1)$，设随机变量 X 表示 A 事件第 r 次出现时的试验次数，则 X 的取值为 $r,r+1,\cdots,r+n,\cdots$，相应的分布律为:

$$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r},\quad 0<p<1,\quad k=r,r+1,\cdots,r+n,\cdots$$

称随机变量 X 服从参数为 r,p 的**负二项分布**，记为. $X\sim\mathrm{NB}(r,p)$。其中当 r=1时，即为几何分布.

目录/Contents

目录/Contents

2.3 常用的连续型随机变量

一、均匀分布

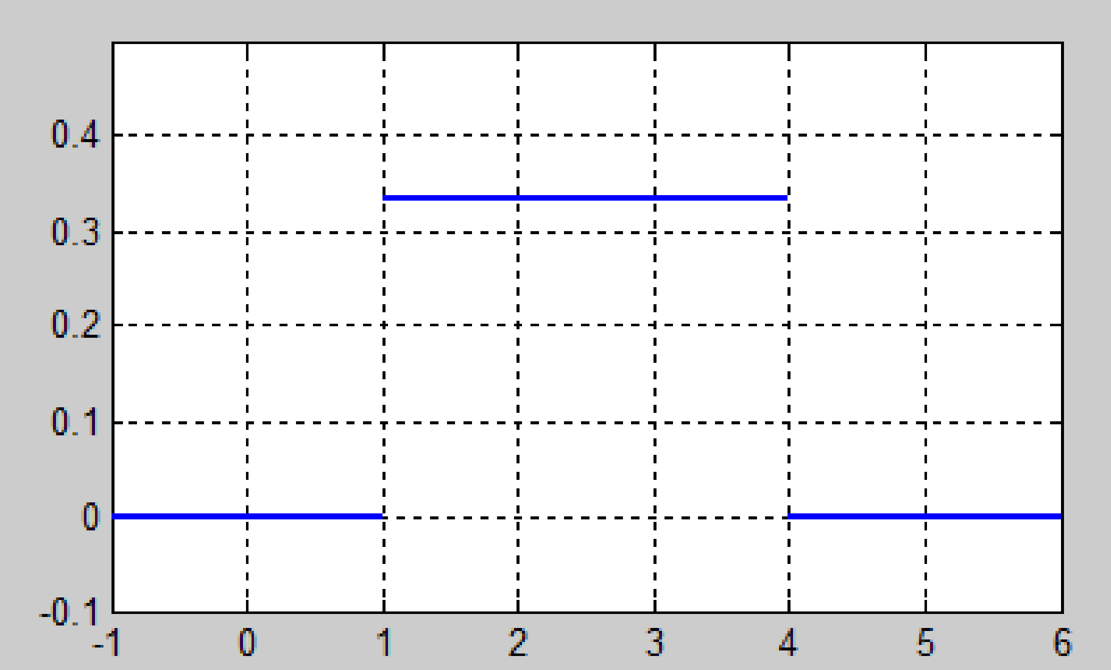
设随机变量 X 的概率密度函数为

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

则称 X 服从区间 (a,b) 上的

均匀分布，记作$X \sim U(a,b)$

密度函数图形如右图:



一、均匀分布

计算可得分布函数为:

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

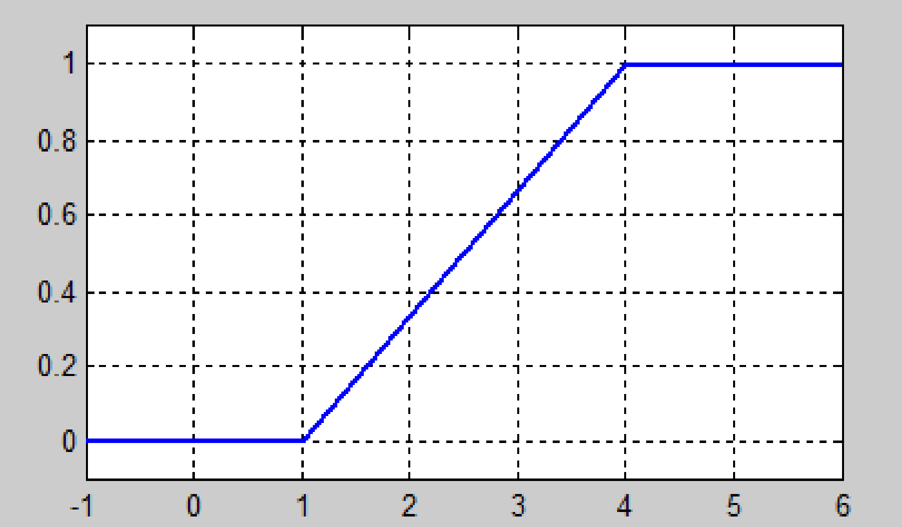
我们考察一下这个分布函数，若 $a<c<c+d<b$，

则 $p(c\le X<c+d)=F(c+d)-F(c)$

$$=\frac{c+d-a}{b-a}-\frac{c-a}{b-a}=\frac{d}{b-a}$$

这恰好是区间 $[c,c+d)$ 和取值总区间的长度比，只与区间长度d有关，与区间位置c无关.

分布函数图形如右所示



例 9 设随机变量$X \sim U(-1,4)$，求（1）事件$\{|X|<2\}$的概率；（2）Y表示对X作3次独立重复观测中事件$\{|X|<2\}$出现的次数，求$P(Y=2)$.

解 （1） X 的概率密度函数为 $f(x)=\begin{cases}\dfrac{1}{4-(-1)} & -1<x<4\\ 0 & \text{其它}\end{cases}$

所以 $P(|X|<2)=P(-1<X<2)=\int_{-1}^{2}\frac{1}{5}dx=\frac{3}{5}$；

（2） Y 表示对 X 作3次独立重复观测中事件$\{|X|<2\}$出现的次数，

故$Y \sim B\left(3,\frac{3}{5}\right)$，所以 $P(Y=2)=\binom{3}{2}\left(\frac{3}{5}\right)^2\frac{2}{5}=\frac{54}{125}$.

二、指数分布

如果随机变量X 的密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & \text{其余}\end{cases}$$

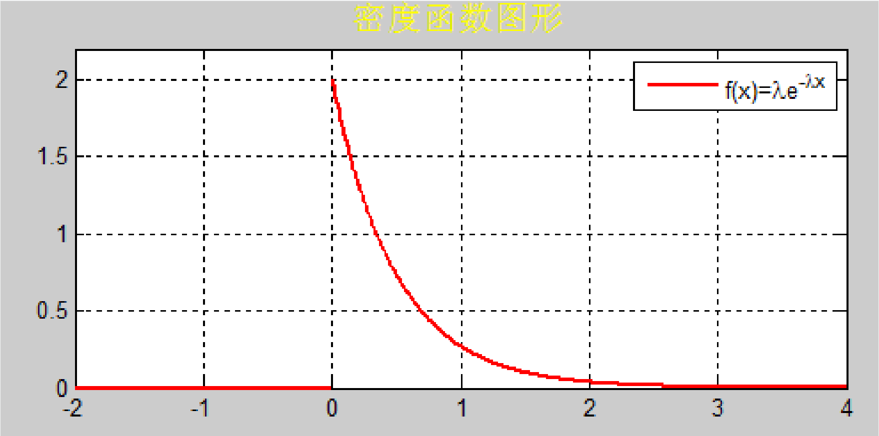
则称 X 服从参数为 λ 的**指数分布**,

记为$X\sim E(\lambda),(\lambda>0)$

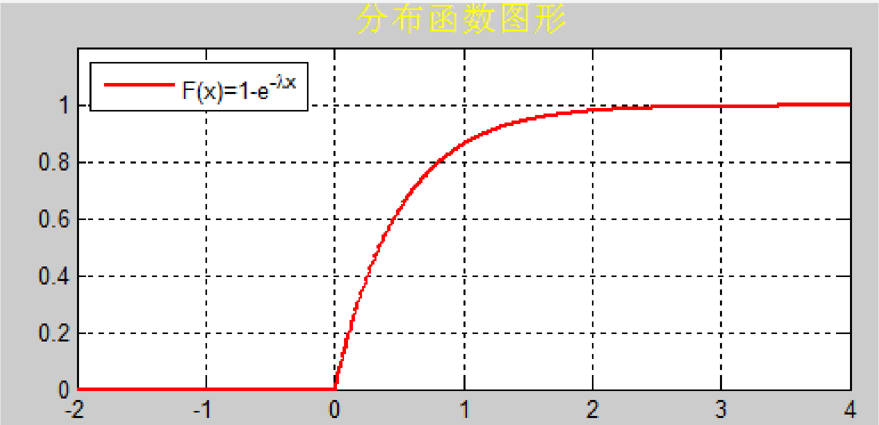
其分布函数为

$$F(x)=\begin{cases}1-e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$$

指数分布的密度函数图形如下:



指数分布的分布函数图形如下:



由定义易得服从指数分布的随机变量的概率计算公式:

设$X \sim E(\lambda), 0 \le a < b$,则$p(a < X \le b) = e^{-\lambda a} - e^{-\lambda b}$

例 10 设随机变量$X \sim E(\lambda)$，则对任意实数$s,t>0$，

证明$P(X>s+t \mid X>s)=P(X>t)$

证明 可以解得$P(X>t)=1-F(t)=e^{-\lambda t}$

故$P(X>s+t \mid X>s)=\dfrac{P(X>s+t)}{P(X>s)}=\dfrac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=P(X>t)$.

设随机变量 X 的概率密度函数为

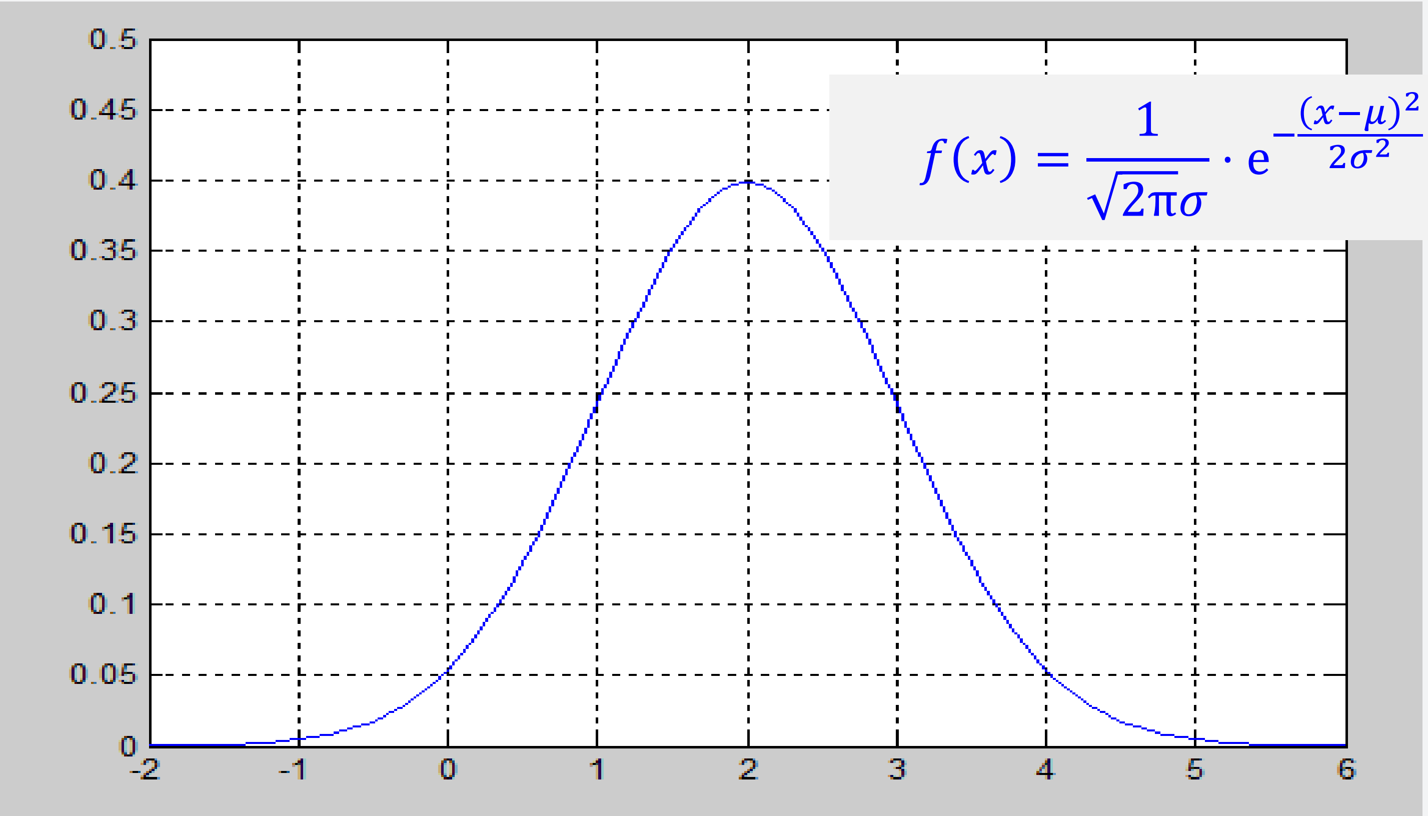
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\cdot \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad x\in R$$

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

服从正态分布的随机变量统称为**正态随机变量**.

三、正态分布

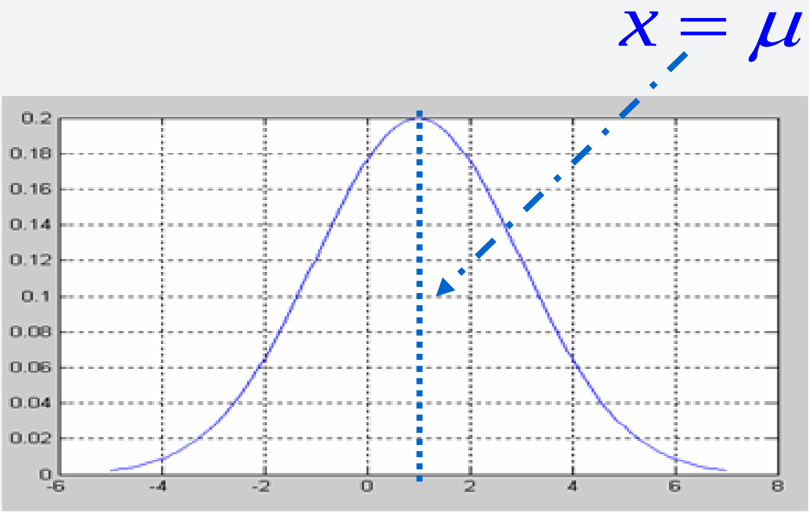
正态分布的密度函数曲线图形



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

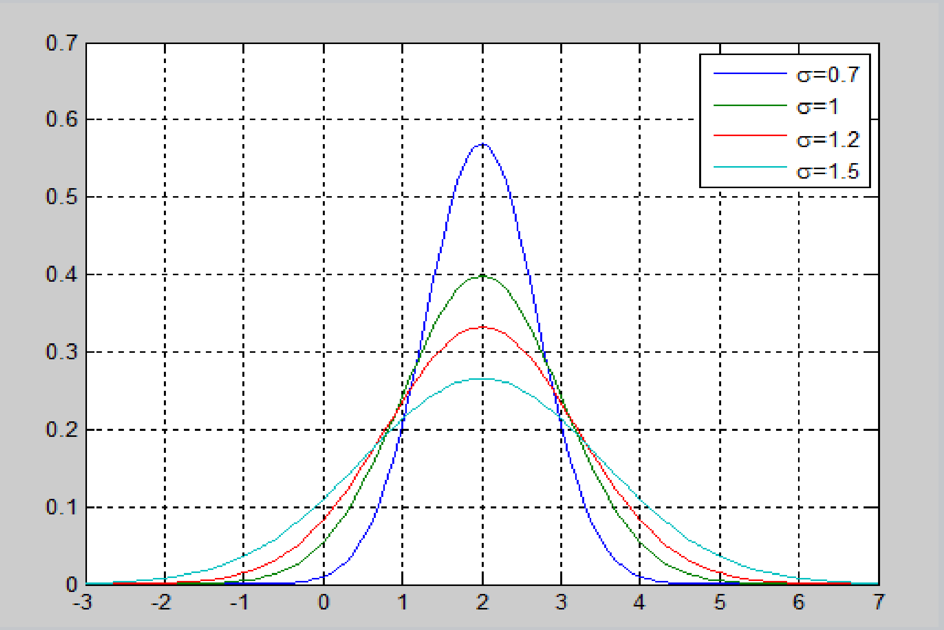
正态分布概率密度函数的曲线特征:

1. 密度函数$f(x)$的图形关于$x=\mu$对称;
2. $f(x)$在$x=\mu$处取得最大值$f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$;
3. 当$|x|\to\infty$时, $f(x)\to 0$;



正态分布概率密度函数的曲线特征:

4 当 σ^2 较大时曲线比较平坦，当 σ^2 较小时曲线比较陡峭.



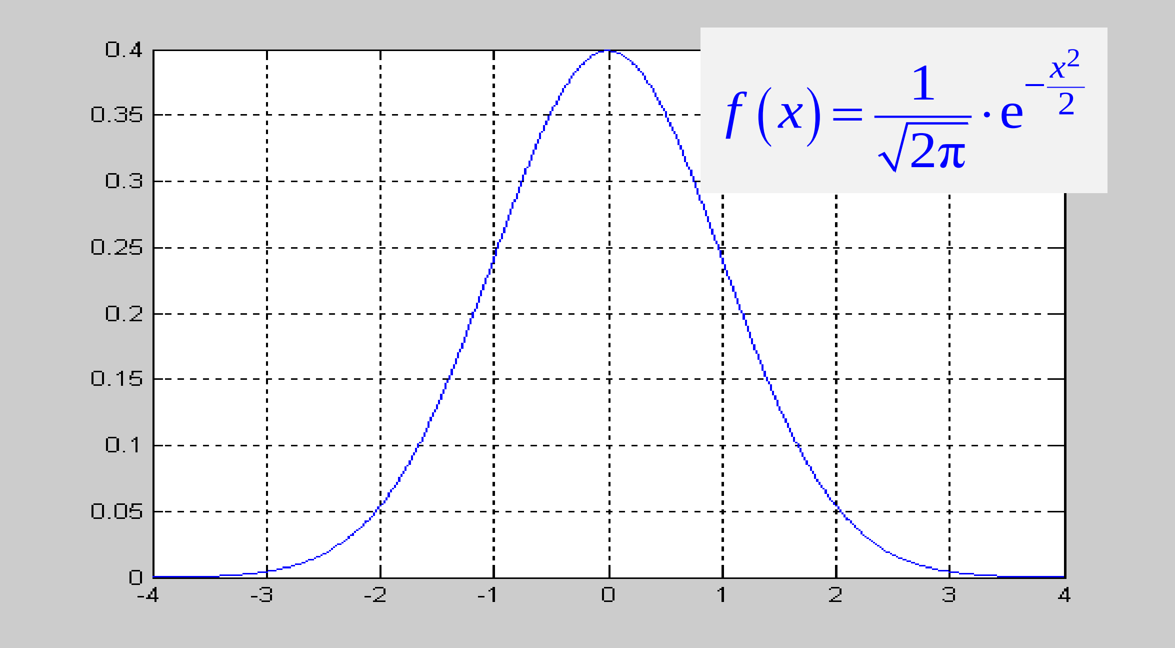
三、正态分布

$\mu=0,\sigma^2=1$时的正态分布称为标准正态分布.其概率密度函数和分布函数分别为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot \mathrm{e}^{-\frac{x^2}{2}}$$

$$\Phi(x)=\int_{-\infty}^{x}\varphi(t)\mathrm{d}t=P(X\le x)$$

$$-\infty<x<+\infty$$



标准正态分布密度函数图形

关于标准正态分布有以下结果:

1 当$x > 0$时, $\Phi(x)$的值可以查概率函数值表得到，且

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

2 当$x < 0$ 时，由密度函数对称性可得 $\Phi(x) = 1 - \Phi(-x)$,

特别地，有$\Phi(0) = \frac{1}{2}$;

3 若$X \sim N(\mu, \sigma^2)$，则$P(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

特别地$P(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$　　$P(X > a) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$

例 11

设随机变量$X \sim N(0,1)$，查表求下列概率值:

(1) $P(-1 < X \leq 1.22)$;　　(2) $P(|X| \leq 1.22)$

解

(1) 查表并计算可得

$$P(-1 < X \leq 1.22) = \Phi(1.22) - \Phi(-1) = \Phi(1.22) - (1 - \Phi(1))$$

$$= 0.8888 - 1 + 0.8413 = 0.7301$$

(2) 同样地

$$P(|X| \leq 1.22) = P(-1.22 \leq X \leq 1.22) = \Phi(1.22) + \Phi(-1.22)$$

$$= 2\Phi(1.22) - 1 = 0.7776$$

一般地，有下列结论:

设随机变量$X \sim N(0,1)$，$c > 0$，则

$$P(|X|<c)=\Phi(c)-\Phi(-c)=2\Phi(c)-1$$

$$P(|X|>c)=1-P(|X|\le c)=2-2\Phi(c)$$

若$X \sim N(\mu,\sigma^2)$，则

$$P(a<X\le c)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

特别地

$$P(X\le b)=\Phi\left(\frac{b-\mu}{\sigma}\right);\quad P(X>a)=1-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

例 12 设$X \sim N(\mathbf{1},\mathbf{4})$，试求概率$P(X \le 3), P(X \le -3)$

解 查表并计算可得

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

$$P(X \le -3) = \Phi\left(\frac{-3-1}{2}\right) = 1 - \Phi(2) = 0.0228$$

例 13

设随机变量X服从标准正态分布$N(0,1)$，c为何值时才能满足

$P(X \le c) = 0.95$

解

由 $P(X \le c) = \Phi(c) = 0.95$ ，查附录4知

$$\Phi(1.645) = 0.95 \Rightarrow c = 1.645.$$

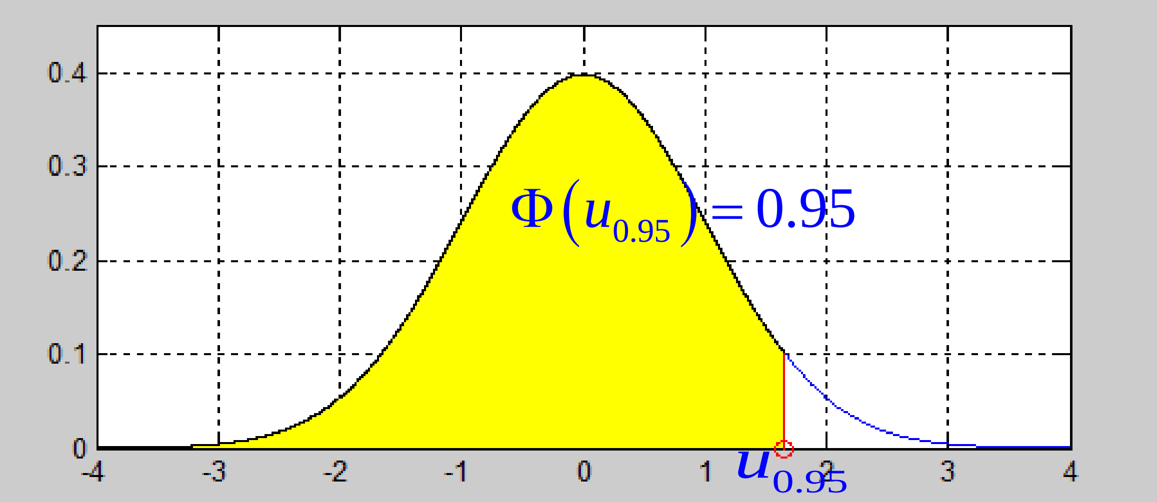
标准正态分布的分位数概念:

设$X \sim N(0,1)$, 对给定的 α，若数u_α满足

$$\Phi(u_\alpha)=\int_{-\infty}^{u_\alpha}\varphi(x)\mathrm{d}x=P(X\le u_\alpha)=\alpha$$

称 $\boldsymbol{u_\alpha}$ 为随机变量X 的α –分位数

右图为分位数的几何意义



例 14 某学校规定划分考生成绩的等级方法如下：考试成绩的实际考分在前10%的为A等，考分在前10%以后但在前50%的为B等，考分在前50%以后但在前85%的为C等，考分在后10%的为D等.某次期末考试中，设考生的成绩X服从正态分布$X \sim N(\mu,\sigma^2)$，经计算可知 $\mu = 73$，σ^2=144，求这次期末考试等级划分的具体分数线。

解 由题意可知$X \sim N(73,144)$, 则

$$P(X \geq a) = 1 - \Phi(\frac{a-73}{12}) = 0.1$$

$$\Rightarrow (\frac{a-73}{12}) = u_{0.9} = 1.282 \Rightarrow a = 88.384 \approx 88$$

又 $P(X \geq b) = 1 - \Phi(\dfrac{a-73}{12}) = 0.5$

$$\Rightarrow (\frac{b-73}{12}) = u_{0.5} = 0 \Rightarrow b = 73$$

又 $P(X \leq c) = \Phi(\dfrac{a-73}{12}) = 0.5$

$$\Rightarrow (\frac{c-73}{12}) = u_{0.1} = -u_{0.9} = -1.282 \Rightarrow c \approx 58$$

综述所求，可知，在此次考试中，分数在88.384以上的，为等级A，分数在73至88.384之间的，为等级B，分数在57.616至73之间的，为等级C，分数在57.616以下的，为等级D。

目录/Contents

目录/Contents

2.4 随机变量函数的分布

设随机变量X，定义一个函数$g(x)$，则$Y=g(X)$是随机变量X的函数，也是一个随机变量。问题：已知X的分布，如何求$Y=g(X)$的分布.

例 15 设随机变量 X 的分布律为

X	-1	0	1	2
概率	0.1	0.2	0.3	0.4

1 求 $Y = X + 2$ 的分布律;

2 求 $W = X^2$ 的分布律。

一、离散型随机变量函数的分布

解 (1) 随机变量$Y = X + 2$的取值为$0,1,2,4$ 且$P(Y = y) = P(X = y - 2)$.

由此得到相应的分布律:

Y	1	2	3	4
概率	0.1	0.2	0.3	0.4

(2) 随机变量$W = X^2$的取值为$0,1,4$,

同理可得对应的分布律为:

W	0	1	4
概率	0.2	0.4	0.4

一、离散型随机变量函数的分布

设离散型随机变量 X 的分布律为

X	x_1	$\cdots$	x_i	$\cdots$
概率	p_1	$\cdots$	p_i	$\cdots$

则 $Y=g(X)$ 的分布律为

$Y=g(X)$	$g(x_1)$	$\cdots$	$g(x_i)$	$\cdots$
概率	p_1	$\cdots$	p_i	$\cdots$

但要注意的是，与 $g(x_i)$ 取相同值对应的那些概率应合并相加。

例 16 **设X服从区间$(1,3)$上的均匀分布，求X^2的密度函数。**

解 随机变量 X 的取值范围$(1,3)$，故随机变量$Y = X^2$的取值范围为区间$(1,9)$，Y仍然是一个连续型随机变量。因此需求解Y的分布函数为$F_Y(y)$和概率密度函数$f_Y(y) = F'(y)$。

根据题意，X的概率密度函数为 $f(x) = \begin{cases} \frac{1}{2} & 1 < x < 3 \\ 0 & \text{其它} \end{cases}$

Y的分布函数$F_Y(y) = P(Y \le y) = P(X^2 \le y)$

二、连续型随机变量函数的分布

Y 的分布函数 $F_Y(y)=P(Y\le y)=P(X^2\le y)$

01 OPTION

当$y<1$时，$F_Y(y)=0$；

02 OPTION

当$1\le y<9$时，$F_Y(y)=P(-\sqrt{y}\le X\le\sqrt{y})=\int_1^{\sqrt{y}}\frac{1}{2}dx=\frac{\sqrt{y}-1}{2}$；

03 OPTION

当$y\ge 9$时，$F_Y(y)=1$。

通过 $F_Y'(y)=f_Y(y)$得到 Y 的密度函数

$$f_Y(y)=F_Y'(y)=\begin{cases}\dfrac{1}{4\sqrt{y}}, & 1<y<9;\\ 0, & \text{其他}.\end{cases}$$

二、连续型随机变量函数的分布

$Y=g(X)$的分布函数与密度函数求解的一般步骤：

1. 由随机变量X的取值范围Ω_X确定随机变量Y的取值范围Ω_Y；

2. 对任意一个$y\in\Omega_Y$，求出$F(y)=P(Y\le y)=P(g(X)\le y)=P\{X\in G_y\}=\int_{G_y} f(x)dx.$

   其中$\{X\in G_y\}$是与$\{g(X)\le y\}$等价的随机事件，而$G_y=\{x: g(x)\le y\}$是实数轴上的某个集合（通常是一个区间或若干个区间的并集）。

3. 按分布函数的定义写出$F_Y(y)$，$-\infty<y<+\infty$

4. 通过对分布函数求导，得到$f_Y(y)=F_Y'(y)$，$-\infty<y<+\infty$。

例 17 设$X \sim N(0,1)$，求随机变量$Y=|X|$的密度函数.

解 易得随机变量$Y=|X|$的取值范围为区间$[0,+\infty)$，Y仍然是一个连续型随机变量。当$y \ge 0$时，Y的分布函数为

$$F_Y(y)=P(Y \le y)=P(|X| \le y)=P(-y \le X \le y)=\Phi(y)-\Phi(-y)$$

直接对上式求导有 $f_Y(y)=F_Y'(y)=\Phi'(y)-\Phi'(-y)=\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$

所以，Y的概率密度函数为

$$f_Y(y)=\begin{cases}\dfrac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}, & y>0,\\ 0, & \text{其它}.\end{cases}$$

定理 1

设连续型随机变量 X 的密度函数为$f_X(x)$，$Y=g(X)$是连续型随机变量，若$y=g(x)$为严格单调函数，$x=g^{-1}(y)$为相应的反函数，且为可导函数，则$Y=g(X)$的密度函数为

$$f_Y(y)=f_X(g^{-1}(y))\cdot\left|\left[g^{-1}(y)\right]'\right|$$

定理 2

设$X\sim N(\mu,\sigma^2)$，则当$k\neq 0$时，$Y=kX+b\sim N(k\mu+b,k^2\sigma^2)$，

特别地，$\frac{X-\mu}{\sigma}\sim N(0,1)$

例 18 设$X \sim N(\mu,\sigma^2)$，求随机变量$Y = e^X$的密度函数.

解 因$y = e^x$的反函数为$x = ln\, y$； 当$y > 0$时单增，$x' = \frac{1}{y}$，所以当$y > 0$时

$$f_Y(y) = f_X(\ln y)(\ln y)' = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}$$

所以$Y = e^X$的密度函数为

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{(ln\, y-\mu)^2}{2\sigma^2}}, & y > 0, \\ 0, & 其它. \end{cases}$$

总结/summary

- 随机变量
  - 分布函数
    - 离散型随机变量、分布律、二项分布、泊松分布、几何分布
    - 连续型随机变量、密度函数、均匀分布、指数分布、正态分布
  - 随机变量函数的分布