概率论与数理统计

主讲教师： 朱丽娜　讲师

研究方向： 智能交通，车联网与智能驾驶

电子邮件： lnzhu@xidian.edu.cn

个人主页： http://web.xidian.edu.cn/lnzhu/

第二章 随机变量及其分布

§ 2.4 连续型随机变量及其概率密度

➲ 连续型随机变量及概率密度 定义:

- 如果对于随机变量X的分布函数$F(x)$，存在非负可积函数$f(x)$，使对于任意实数x有
- $$F(x)=\int_{-\infty}^{x} f(t)dt$$
- 则称X为连续型随机变量，其中函数$f(x)$称为X的概率密度函数，简称概率密度

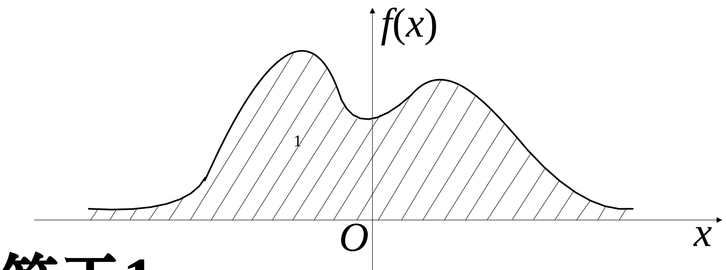
- ※由数学分析的知识，连续型随机变量的分布函数是连续函数
- ※在实际应用中遇到的基本上是离散和连续型随机变量

几种提法:
X的概率分布:
是指分布函数
X为连续型时:
是指概率密度
X为离散型时:
是指分布律

§ 2.4 连续型随机变量及其概率密度

➲ 概率密度函数的性质：

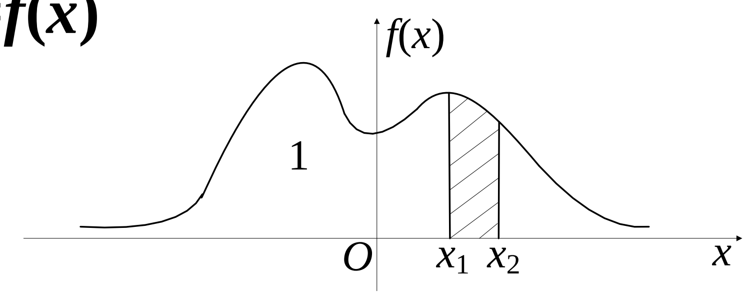
- **1° 非负性：$f(x)\geq 0$. 由定义可知**
- **2° 规范性：** $\int_{-\infty}^{\infty} f(x)dx = F(\infty) = 1$
- **曲线$y=f(x)$与Ox轴之间的面积等于1**



- **3° 对任意实数x_1，x_2 ($x_1\leq x_2$)，有$P\{x_1<X\leq x_2\}=F(x_2)-F(x_1)$**

$$=\int_{x_1}^{x_2} f(x)dx$$

  - **概率$P\{x_1<X\leq x_2\}$等于区间$(x_1,x_2]$上曲线$y=f(x)$之下曲边梯形的面积**
- **4° 若$f(x)$在点x处连续，则有$F'(x)=f(x)$**
  - **比右连续强，一般的是右连续）**



§ 2.4 连续型随机变量及其概率密度

a) 一维随机变量的概率密度的线密度含义：

- 由连续性定义，在连续点x处有
- $$f(x)=\lim_{\Delta x\to 0^+}\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{\Delta x\to 0^+}\frac{P\{x<X<x+\Delta x\}}{\Delta x}$$
- 这正好与物理学中的线密度的定义类似：随机点落在单位区间上概率的大小
- 当Δx充分小时，点x处的曲边梯形可近似为长Δx高$f(x)$的矩形。
- 即$P\{x<X\le x+\Delta x\}\approx f(x)\Delta x$，随机变量X落在$(x，x+\Delta x]$的概率近似等于$f(x)\Delta x$，忽略了高阶无穷小。

§ 2.4 连续型随机变量及其概率密度

- **b) 对于连续型随机变量，单点概率为0**
  - **即若X为连续型随机变量，对任意的实数a，有$P\{X=a\}=0$**
- **证：显然有包含关系$\{X=a\}\subset\{a-\Delta x<X\le a\}$**
  - **$\therefore 0\le P\{X=a\}\le P\{a-\Delta x<X\le a\}=F(a)-F(a-\Delta x)$**
  - **而$F(x)$是连续的（也满足左连续），因为连续型随机变量的分布函数是连续的**
  - **$\therefore$当$\Delta x\to 0$时有$0\le P\{X=a\}\le \lim\limits_{\Delta x\to 0}(F(a-\Delta x)-F(a))=F(a)-F(a)=0$**
  - **即$P\{X=a\}=0$**
- **要注意几点:**
  - **1° 只有当X为连续型时，才一定有$P\{X=a\}=0$，否则不一定**
  - **2° 尽管$P\{X=a\}=0$，但$\{X=a\}$不是不可能事件**
  - **3° 连续型随机变量中区间的端点不影响概率值，即对以下概率不加区分 $P\{a<X\le b\}=P\{a\le X\le b\}=P\{a\le X<b\}=P\{a<X<b\}$**

$P(A)=0$不能$\Rightarrow A=\Phi$
$P(A)=1$不能$\Rightarrow A=S$

§2.4 连续型随机变量及其概率密度

- **例：** **由分布函数求概率密度**

- **设** $F(x)=\begin{cases}0, & x<1\\ \ln x, & 1\le x<e\\ 1, & x\ge e\end{cases}$ **，求$f(x)$**

- **解：先判断$F(x)$在区间端点的连续性(是否有阶跃特性)，如果连续则直接对$F(x)$进行分段求导即可，由于连续型随机变量单点概率为0，不可导的点可直接取右导数即可，也可随意给定。**

- $$f(x)=\frac{dF(x)}{dx}=\begin{cases}0, & x<1\\ \dfrac{1}{x}, & 1\le x<e\\ 0, & x\ge e\end{cases}=\begin{cases}1/x, & 1\le x<e\\ 0, & \text{其它}\end{cases}$$

§ 2.4 连续型随机变量及其概率密度

- **例1：由概率密度求分布函数和概率**

- **设随机变量X具有概率密度** $f(x)=\begin{cases} kx, & 0\le x<3 \\ 2-\dfrac{x}{2}, & 3\le x\le 4 \\ 0, & \text{其它} \end{cases}$
-
- **(1)确定常数*k*；(2)求*F*(*x*)；(3)求*P*{1<X≤7/2 }**

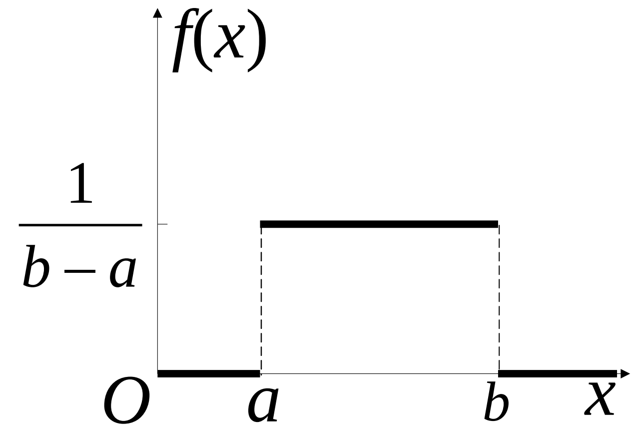
解：(1) 由规范性得 1=$\int_0^3 kxdx+\int_3^4 (2-x/2)dx$ **，解得*k*＝1/6；**

- **(2) 由定义**
- $$F(x)=\int_{-\infty}^{x} f(t)dt=\begin{cases} 0, & x<0 \\ \int_0^x x/6dx, & 0\le x<3 \\ \int_0^3 x/6dx+\int_3^x (2-x/2)dx, & 3\le x<4 \\ 1, & x\ge 4 \end{cases}=\begin{cases} 0, & x<0 \\ x^2/12, & 0\le x<3 \\ -3+2x-x^2/4, & 3\le x<4 \\ 1, & x\ge 4 \end{cases}$$
- **(3) *P*{1<X≤7/2 }＝*F*(7/2)－*F*(1)＝41/48**
- **或者用性质(3)在区间(1,7/2]上对*f*(*x*)积分**

§ 2.4 连续型随机变量及其概率密度

- **三种重要的连续型随机变量**
- **(一)均匀分布**
  - **设连续型随机变量X具有概率密度**

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

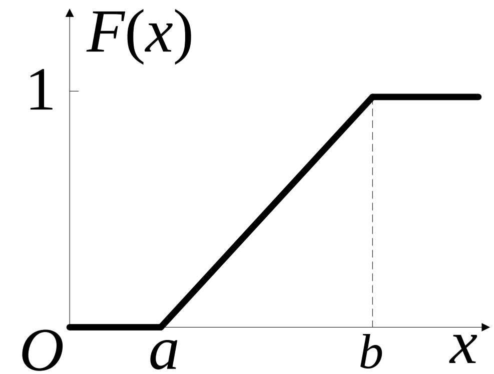


  - **则称x在区间(a,b)上服从均匀分布.记为$X \sim U(a, b)$。**
  - **例如：电子和量子在一定能级上的运动轨迹**
  - **数值计算中，研究四舍五入引起的误差。**
- **概率密度的两个条件显然成立:**
  - **非负性成立$f(x)\geq 0$，且规范性:$\int_{-\infty}^{\infty} f(t)dt = 1$**

§ 2.4 连续型随机变量及其概率密度

- **X的分布函数：**

$$F(x)=\int_{-\infty}^{x} f(t)dt=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b\\ 1, & x\ge b\end{cases}$$



- **$F(x)$为$f(x)$的积分图像**

  - **关于自变量范围：连续型随机变量概率密度函数的自变量一定是在$(-\infty,\infty)$上都有定义，有些区间函数值为0，表明随机变量X落在这些区间的概率为0，或不可能事件。**

  - **均匀分布的含义：代表一种等可能性，即X落在区间(a,b)中任意等长度的子区间内的概率是等可能的，或者说X落在区间(a,b)中的概率只依赖于子区间长度而与子区间位置无关。**

- **证：对任意长度为l的子区间$(c,c+l)$，$a\le c<c+l\le b$，有**

  - $$P\{c<X<c+l\}=\int_{c}^{c+l} f(x)dt=\int_{c}^{c+l}\frac{1}{b-a}dt=\frac{l}{b-a}$$

几何概型

§ 2.4 连续型随机变量及其概率密度

- **例2： 设电阻值R是一个随机变量，均匀分布在900Ω～1100Ω。求R的概率密度及R落在950Ω～1050Ω的概率。X～U(900，1100)**

- **解：首先由题意均匀分布，得到R的概率密度函数为**

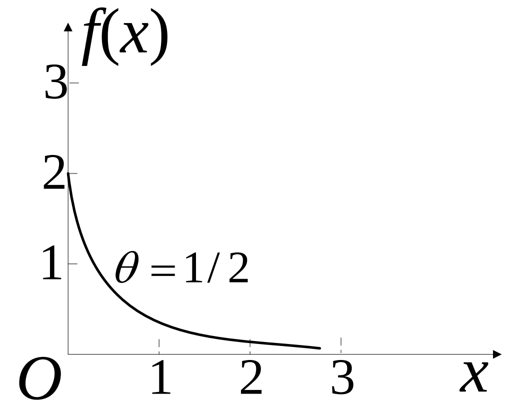
$$f(x)=\begin{cases}\dfrac{1}{1100-900}, & 900<r<1100\\ 0, & \text{其它}\end{cases}$$

- $\therefore P\{950<R\leq 1050\}=\int_{950}^{1050}\frac{1}{200}dt=0.5$

§ 2.4 连续型随机变量及其概率密度

- **(二)指数分布**
  - 设连续型随机变量X的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$$



  - 其中θ>0为常数，则称X服从参数为θ的**指数分布**。也叫**负指数分布**

- **概率密度的两个条件显然成立:**
  - 易知非负性成立$f(x)\geq 0$，且规范性: $\int_{-\infty}^{\infty} f(t)dt=1$

- **X的分布函数:**

$$F(x)=\int_{-\infty}^{x} f(t)dt=\begin{cases}1-e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$$

§2.4 连续型随机变量及其概率密度

- 无记忆性：（指数分布又称“永远年青”的分布）
  - 服从指数分布的X有如下性质：$P\{X>s+t|X>s\}=P\{X>t\}$
- 证：$P\{X>s+t|X>s\}=P\{X>s+t\cap X>s\}/P\{X>s\}=P\{X>s+t\}/P\{X>s\}$
  - $=[1-F(s+t)]/[1-F(s)]$
  - $=e^{-(s+t)/\theta}/e^{-s/\theta}=e^{-t/\theta}$
  - $=P\{X>t\}$
  - 含义：如果X是某一元件寿命，那么已知元件使用了s小时，它还能再使用t小时（即$s+t$小时）的条件概率，等于从开始使用时算起它至少能使用t小时的概率。即元件对它使用s小时没有记忆，符合这样数学模型的随机现象有很多。
- 应用：排队论，可靠性理论
  - 描述衰老作用不明显的寿命分布；θ为寿命X的平均值
  - 衰老作用明显的时候常采用weibull分布

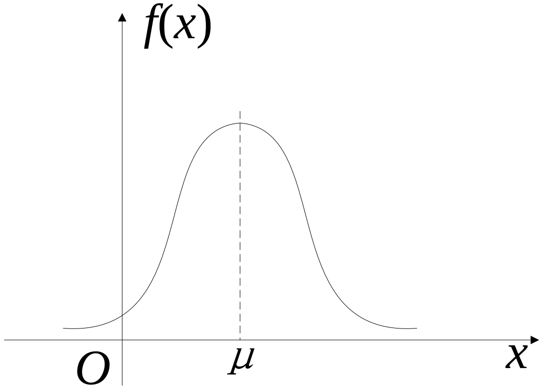
§ 2.4 连续型随机变量及其概率密度

- **(三)正态分布**
- **在自然界中大量的随机现象服从或近似的服从正态分布**
  - **人的身高、体重；农作物的收获量；测量误差；**
  - **产品的尺寸、炮弹弹着点、考试成绩**
  - **器件中的电子热噪声；信道噪声**
  - **若随机变量 X 受到众多相互独立的随机因素的影响，而每一个别因素的影响都是微小的，且这些影响可以叠加，则 X 服从正态分布(大量微小的独立的随机因素综合作用的结果(第五章进一步学习))**
  - **这些现象的分布常具有两头小、中间大，呈左右对称的特点，钟形曲线**

§ 2.4 连续型随机变量及其概率密度

- **定义：设连续型随机变量X的概率密度为**
  - $$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty$$
  - **其中μ，$\sigma(\sigma>0)$为常数，则称X服从参数为μ，σ的正态分布或高斯(Gauss)分布，记为$X\sim N(\mu,\sigma^2)$**
  - **注意，第二个参数是平方的形式**
  - **μ，σ的含义在第四章介绍，它们分别是均值和方差的概念**

- 此分布的数学表达式最早（1718年）由法国哲学、数学家A.DeMoivre提出，以后由（德）数学家C.F.Gauss（1777～1855年）作了系统研究，故又称de Moivre分布或Gauss分布。



§ 2.4 连续型随机变量及其概率密度

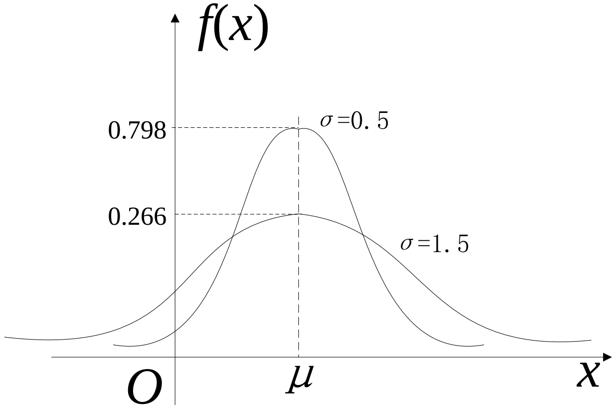
- **概率密度的两个条件：**
  - 易知非负性成立$f(x)\geq 0$，现在证规范性：$\int_{-\infty}^{\infty} f(x)dx = 1$
  - 令$t=(x-\mu)/\sigma$，有$dt=dx/\sigma$，$-\infty<t<\infty$，代入得
    $$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt$$
  - 记 $I=\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt$
  - 则 $I^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{t^2+u^2}{2}} dtdu$
  - 利用极坐标系将它化成累次积分，$dtdu = rdrd\theta$，$0<r=(t^2+u^2)^{1/2}<\infty$，$0<\theta<2\pi$得
  - $I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} re^{-\frac{r^2}{2}} drd\theta = \frac{1}{2\pi} 2\pi = 1$ 即$I=1$
  - 于是$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = I = 1$

§2.4 连续型随机变量及其概率密度

- **正态函数的性质：函数曲线的性质**
- **1° 对称轴：**
  - 曲线关于$x=\mu$对称，对任意的$h>0$有
  - $P\{\mu-h<X\le\mu\}=P\{\mu<X\le\mu+h\}$
  - 函数的对称性即证明，对任意的a有$f(\mu-a)=f(\mu+a)$
  - 证明略。
- **意义：X落在关于$x=\mu$对称的两个区间的概率相等**
- **对称轴平移，曲线形状不变。曲线$f(x)$的位置完全由μ决定，μ称为位置参数**

§2.4 连续型随机变量及其概率密度

- 2° 最大值：当$x=\mu$时取到最大值$f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$
  - 只需令$df(x)/dx=0$，可解出$x=\mu$时取得唯一极值
  - 意义：极大值与σ有关，σ越小，最值越大，X落在对称轴附近的概率越大，x离μ越远$f(x)$的值越小

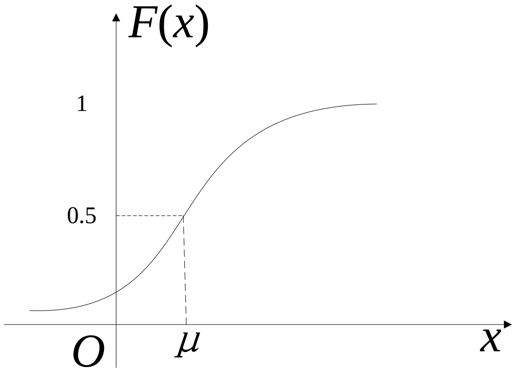


- 3° 拐点：上凸和下凹的分界点
  - 在$x=\mu\pm\sigma$处曲线有拐点。
  - 二阶导异号?
- 4° 渐近线：$y=0$（Ox轴）为渐近线，无限接近但不相交
  - 由 $\lim\limits_{x\to\infty} f(x)=0$ 可得
- **由以上性质可见4个特征基本描述了曲线的特征，$f(x)$的特征均由μ，σ决定**

§ 2.4 连续型随机变量及其概率密度

- 分布函数：不可积

$$F(x)=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$



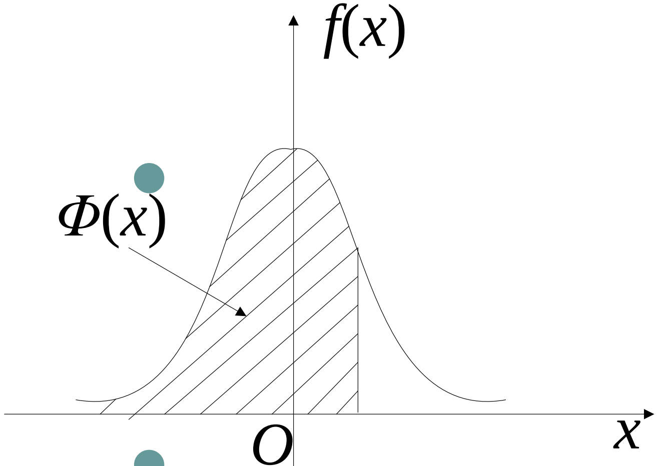
- 标准正态分布：

  - 当$\mu=0$，$\sigma=1$时称X服从标准正态分布，其概率密度和分布函数分别用$\varphi(x)$，$\Phi(x)$，记做X～$N(0，1^2)$
  - $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$，$-\infty<x<\infty$，$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$
  - $\Phi(x)$是超越积分，为此人们编制了一张$\Phi(x)$函数表，以供查用

§ 2.4 连续型随机变量及其概率密度

- **$\Phi(x)$的性质：$\Phi(-x)=1-\Phi(x)$**

- $\Phi(-x)=\int_{-\infty}^{-x}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du$，
- 令$t=-u$，$\Phi(-x)=-\int_{\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\int_{x}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$



- $=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt-\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}d$
- $=1-\Phi(x)$

- **这样可以只作$x\geq0$时的一半的表格**

§2.4 连续型随机变量及其概率密度

- **正态分布随机变量的标准化：线性变换**
- **引理：若X～$N(\mu,\sigma^2)$，则$Z=\dfrac{X-\mu}{\sigma}$～$N(0,1)$**
  - 随机变量Z是随机变量X的函数，更一般情形留到第五节
- **证：先由分布函数的含义，由X的分布函数求随机变量Z的分布函数**
  - $F(z)=P\{Z\le z\}=P\{\dfrac{X-\mu}{\sigma}\le z\}=P\{X\le\mu+\sigma z\}$
  - 由X的分布函数有 $=\displaystyle\int_{-\infty}^{\mu+\sigma z}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$
  - 令 $\dfrac{t-\mu}{\sigma}=u$，则$dt=\sigma du$，积分换限，而由$t=\mu+\sigma u$，及t的范围，得出u的范围
  - $$=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=\Phi(z)$$
  - 所以$Z=\dfrac{X-\mu}{\sigma}$～$N(0,1)$

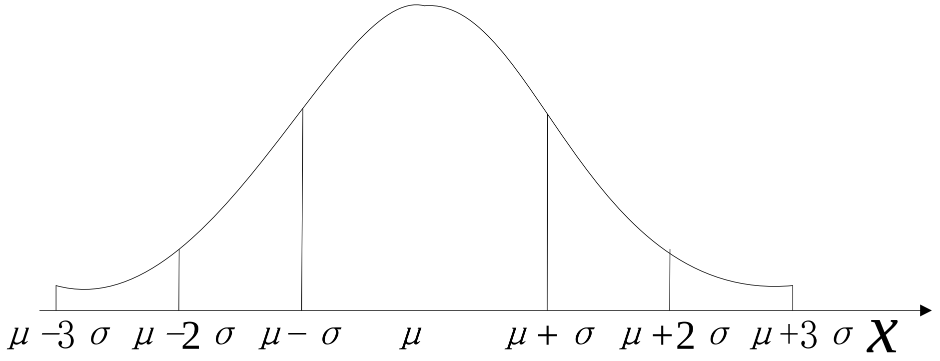
§ 2.4 连续型随机变量及其概率密度

- **对于一般的正态分布随机变量可表示为标准正态分布的形式**
- 于是若X～$N(\mu, \sigma^2)$，则其分布函数
  - $F(x)=P\{X\le x\}=P\{\frac{X-\mu}{\sigma}\le\frac{x-\mu}{\sigma}\}=\Phi(\frac{x-\mu}{\sigma})$
  - 而标准正态函数可查表求值
- 对于任意的区间$(x_1,x_2]$，有
  - $P\{x_1<X\le x_2\}=P\{\frac{x_1-\mu}{\sigma}<\frac{X-\mu}{\sigma}\le\frac{x_2-\mu}{\sigma}\}=\Phi(\frac{x_2-\mu}{\sigma})-\Phi(\frac{x_1-\mu}{\sigma})$
- 例：设X～$N(1, 4)$，求$P\{0<X\le 1.6\}$
- 解：$\mu=1$，$\sigma=2$
  - $P\{0<X\le 1.6\}=\Phi((1.6-1)/2)-\Phi((0-1)/2)=\Phi(0.3)-\Phi(-0.5)$
  - 查表及由标准正态分布函数的性质 $=0.6179-[1-\Phi(0.5)]$
  - $=0.6179-1+0.6915=0.3094$

§ 2.4 连续型随机变量及其概率密度

➲ **3σ准则：**

- **尽管正态变量X～$N(\mu，\sigma^2)$的取值范围是$(-\infty,\infty)$，但X落在$(\mu-3\sigma，\mu+3\sigma)$内几乎是肯定的**
- **1σ: $P\{\mu-\sigma<X\leq\mu+\sigma\}=\Phi(1)-\Phi(-1)=0.6826$**
- **2σ:$P\{\mu-2\sigma<X\leq\mu+2\sigma\}=\Phi(2)-\Phi(-2)=0.9544$**
- **3σ:$P\{\mu-3\sigma<X\leq\mu+3\sigma\}=\Phi(3)-\Phi(-3)=0.9974$**
- **使得正态分布函数值的表只需要做到0到3.49即可，再大都近似为1**
- **企业管理中，经常应用3σ--规则进行质量检查**

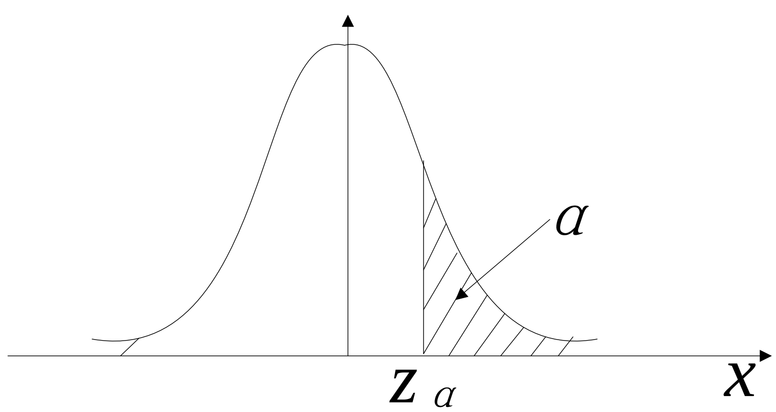


§2.4 连续型随机变量及其概率密度

- 例3：正态分布随机变量X～$N(d，0.5^2)$
  - 若d＝90，求X小于89的概率
  - 若X>80的概率不低于0.99，问d至少为多少?
- 解：(1)此时X～$N(90，0.5^2)$
  - $P\{X<89\}=P\{(X-90)/0.5<(89-90)/0.5\}$
  - $=\Phi((89-90)/0.5)=\Phi(-2)=1-\Phi(2)=1-0.9772=0.0228$
  - (2) $0.99\le P\{X\ge 80\}= P\{(X-d)/0.5\ge(80-d)/0.5\}=1-\Phi((80-d)/0.5)$
  - 即$\Phi((80-d)/0.5)\le 1-0.99=1-\Phi(2.327)=\Phi(-2.327)$
- 以上用了反查表：由概率值求自变量取值，且变换为大于0.5的概率值才可查
  - 又由分布函数的不减性知
  - $(80-d)/0.5\ge -2.327$，故需$d>81.1635$

§2.4 连续型随机变量及其概率密度

- **关于上α分位点：**
  - 为了数理统计中的应用而引入
  - 设X～N(0，1)，若z_α满足条件，$P\{X\geq z_\alpha\}=\alpha$，0<$\alpha$<1，则称点$z_\alpha$为标准正态分布的上$\alpha$分位点。
  - 性质：$z_{1-\alpha}=-z_\alpha$
  - 注意上α分位点与正态分布函数表的关系



§ 2.4 连续型随机变量及其概率密度

- **对数正态分布**

- 一些正偏态资料的变量值，通过对数转换后，由偏态分布转为正态分布。某些正偏态资料，如血铅含量、某些传染病的潜伏期等，经对数变换后可符合正态分布。在金融、工程、医学、生物学等领域有广泛的应用

- 如果 X 是服从正态分布的随机变量，则 e^X 服从对数正态分布；同样，如果 Y 服从对数正态分布，则 $\ln(Y)$ 服从正态分布。

- 如果一个变量可以看作是许多很小独立因子的乘积（与正态的多个很小独立因子的求和相对应），则这个变量可以看作是对数正态分布。一个典型的例子是股票投资的长期收益率，它可以看作是每天收益率的乘积。

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma x}e^{-(\ln x-\mu)^2/2\sigma^2}, & x>0\\ 0, & \text{其它}\end{cases}$$

§ 2.4 连续型随机变量及其概率密度

- Γ函数及其相关的几种重要连续型分布介绍
  - β分布、γ分布、Weibull分布
- Γ函数
  - Γ函数在概率分布中具有极其重要的地位
  - $\Gamma(s)=\int_0^{\infty} x^{s-1}e^{-x}dx$ ，s>0，这是一个广义积分，不可积
  - 重要的性质$\Gamma(s+1)=s\Gamma(s)$，$\Gamma(1/2)=\sqrt{\pi}$ ，
  - 余元公式$\Gamma(s)\Gamma(1-s)=\pi/\sin s\pi$
  - 当$0<s<1$时，$\Gamma(s)$的值可查表求得,

§ 2.4 连续型随机变量及其概率密度

- (1) *Γ*分布
- *Γ*分布可记作：$X \sim \Gamma(\alpha, \beta)$，等待第α件事发生所需要的等待时间
  - $$f_X(x)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}, & x>0\\ 0, & \text{其它}\end{cases}\quad \alpha>0,\beta>0$$
  - *Γ*分布的可加性：$X \sim \Gamma(\alpha_1,\beta)$, $Y \sim \Gamma(\alpha_2,\beta)$，X和Y相互独立
  - 则$Z=X+Y \sim \Gamma(\alpha_1+\alpha_2,\beta)$
  - 其中α是形状参数，决定曲线形状，β是尺度参数决定曲线的伸缩
- *Γ*分布与很多分布都有关系
  - 当α =n/2，β =2时，*Γ*分布即为自由度为n的χ^2分布
  - 当α =1时，*Γ*分布即为参数为β的指数分布
  - 当α =k+1，β =1时，*Γ*分布的表达式与泊松分布表达式的形式是一致的，但参数λ(即x)是连续的

§ 2.4 连续型随机变量及其概率密度

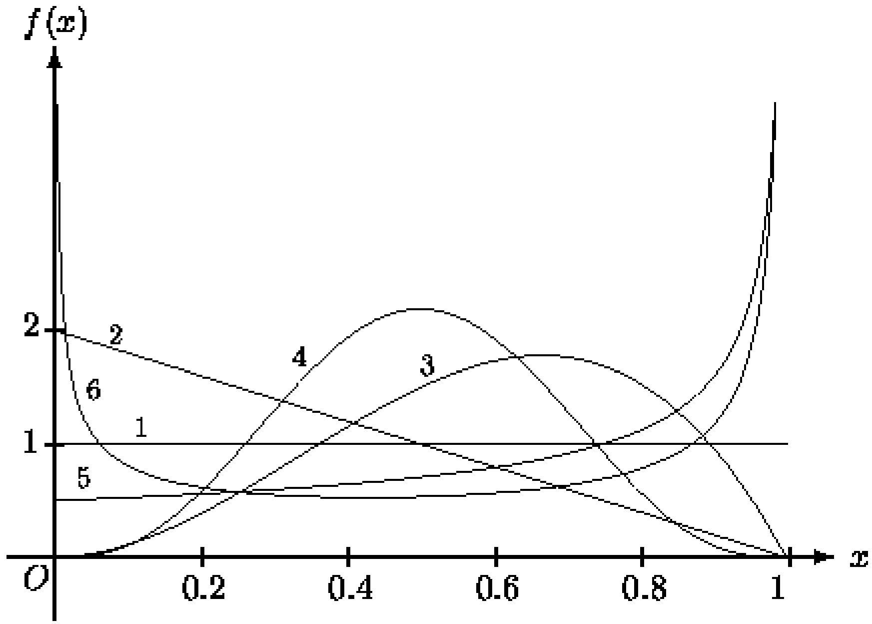
- (2) β分布
- β分布定义在区间(0, 1)，可记作：X～$Beta(\alpha,\beta)$

$$f(x)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1\\ 0, & \text{其它}\end{cases}$$

- $\alpha>0,\beta>0$
- β分布与贝叶斯理论有联系，常用来描述取值在(0, 1)的概率模式，比如描述由后验概率来查看抛硬币中使用硬币的不均匀度(正反面可能不是等概出现)
- 当$\alpha=1$，$\beta=1$时，β分布就是U(0, 1)
- 期望等于$\alpha/(\alpha+\beta)$

§ 2.4 连续型随机变量及其概率密度

- **Beta分布的曲线**



Beta分佈p.d.f.之圖形, 1,2,3,4,5,6分別對應

$(\alpha, \beta) = (1,1), (1,2), (3,2), (4,4), (1,1/2)$, 及$(1/2,1/3)$

§ 2.4 连续型随机变量及其概率密度

- **(3) 韦布尔分布(Weibull分布)**

- ***Weib*(α, β, b)分布:它是可靠性工程中广泛使用的一种连续性分布，常用来描述各种产品、材料疲劳失效的分布规律，如真空管的疲劳失效、机械轴承的疲劳失效等，该分布为三参数分布，其中α为形态参数，β为尺度参数，b为位置参数。当α=1，α＞1，α＜1时分别对应了某些产品（或材料）的偶发期故障(正常随机故障)、耗损期故障(老化严重故障多发期)和早发故障(新设备投入运行)三种形态的失效分布。考虑损耗衰老作用时比指数分布更准确**

- **该分布由瑞典数学家W.Weibull是在对双参数指数分布 $E_1(\lambda_1,b)$ 和Ray(σ)的研究基础上综合推广而成。该分布同样在无线电信号分析和元器件可靠性分析中有重要应用**

§ 2.4 连续型随机变量及其概率密度

- **概率密度表达式**
- $$f(x)=\begin{cases}\dfrac{\alpha}{\beta}(x-b)^{\alpha-1}e^{-\frac{(x-b)^{\alpha}}{\beta}}, & 0<x<1\\ 0, & 其它\end{cases}$$ ***x≥b, b≥0, m*为正整数**
- **令*β=η^α*，带入则可得到概率密度的另一种表达式**
- **当*α*=1时， *Weib*(1,*β*,*b*)即双参数指数分布 $E_1(\beta,b)$**
- **从上面的描述可知，指数分布只适用于浴盆曲线(指产品从投入到报废为止的整个寿命周期内)的底部，但任何产品都有早期故障，也总有耗损失效期。在可靠性工程学中用威布尔分布来描述产品在整个寿命期的分布情况**

本章作业

- P_{59}： 32，33，35，36，37