



信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

第10章 数据统计分析

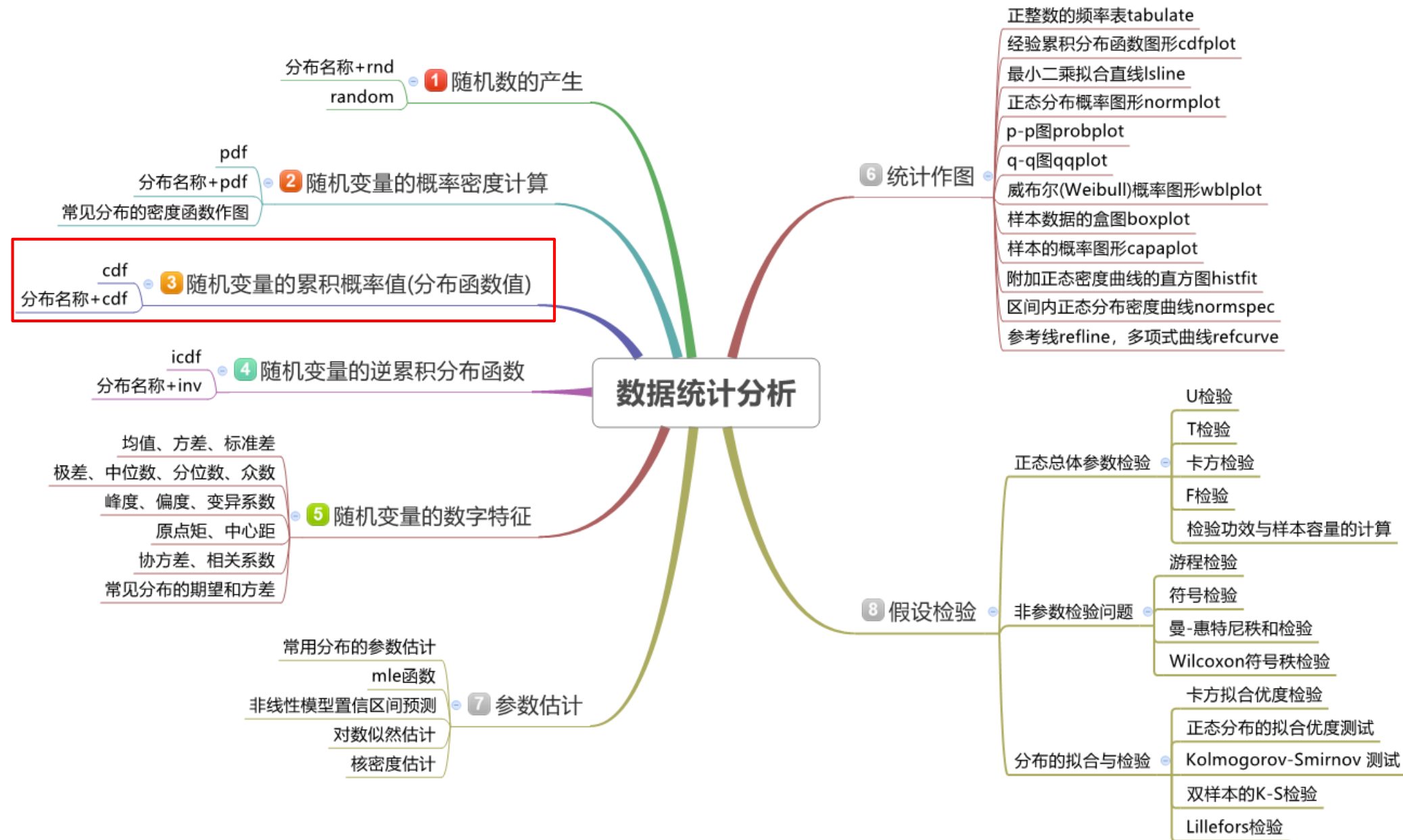


讲授人：牛言涛



日期：2020年4月7日

第10章 数据统计分析知识点思维导图



1. 通用函数计算累积概率值

计算随机变量 $X \leq K$ 的概率之和（累积概率值），函数cdf

- `cdf('name', K, A)`
- `cdf('name', K, A, B)`
- `cdf('name', K, A, B, C)`
- 说明：返回以name为分布、随机变量 $X \leq K$ 的概率之和的累积概率值，name的取值见常见分布函数表。

累积概率函数就是分布函数 $F(x) = P\{X \leq x\}$ 在 x 处的值。

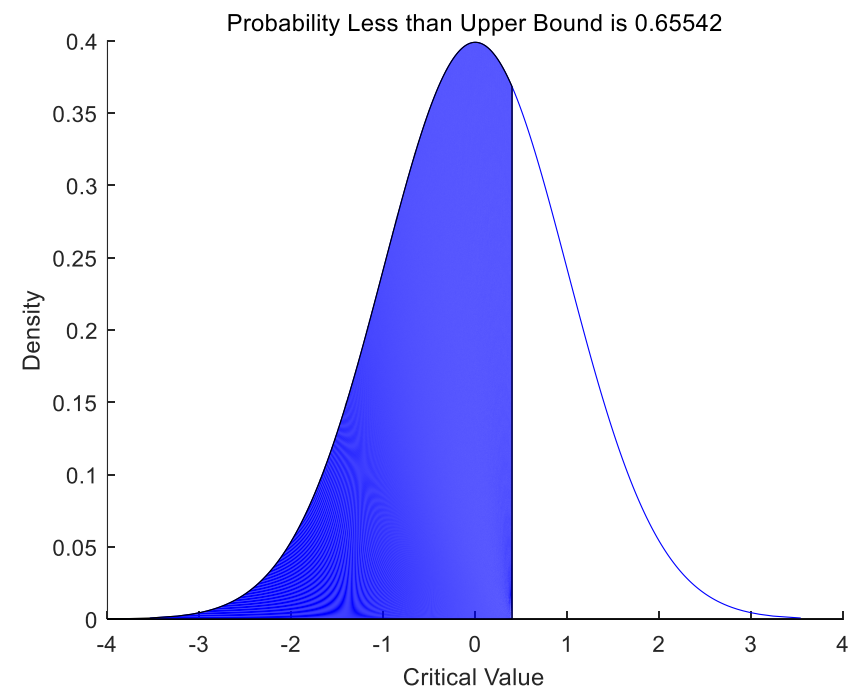
1. 通用函数计算累积概率值

例1: 求标准正态分布随机变量 x 落在区间 $(-\infty, 0.4)$ 内的概率（该值就是概率统计教材中的附表：标准正态数值表）。

```
>>p = cdf('norm',0.4,0,1)
p =
    0.6554
>> normspec([-inf,0.4],0,1)
```

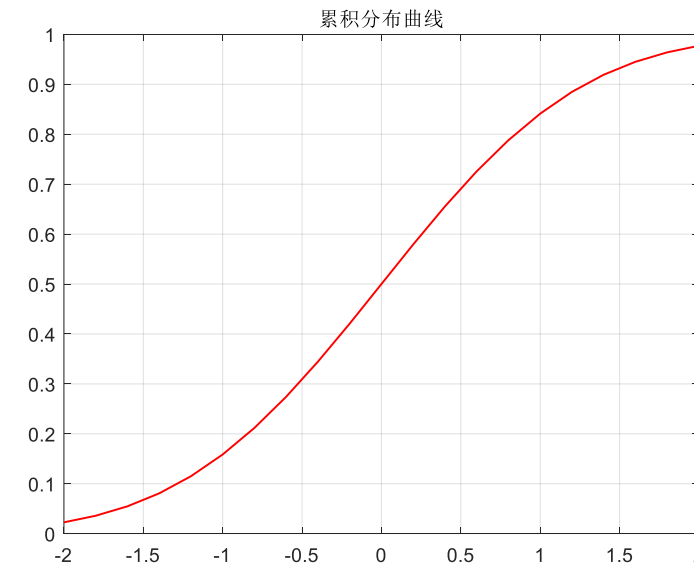
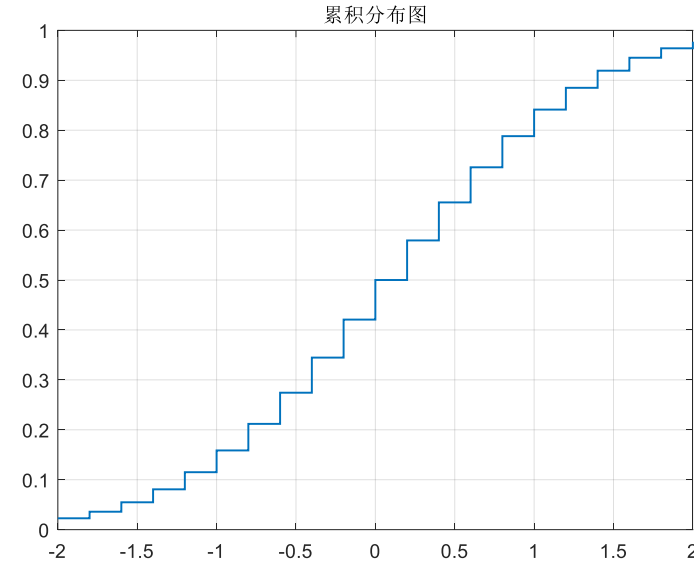
例2: 求自由度为16的卡方分布随机变量落在 $[0, 6.91]$ 内的概率。

```
>>p = cdf('chi2',6.91,16)
p =
    0.0250
```



1. 通用函数计算累积概率值

```
>> mu = 0;  
>> sigma = 1;  
% makedist生成指定分布的对象  
>> pd = makedist('Normal','mu',mu,'sigma',sigma);  
>> x = [-2:0.2:2];  
%cdf第一个参数以makedist对象  
>> y = cdf(pd,x);  
>> stairs(x,y,'LineWidth',1)  
>> title('累积分布图')  
% 实现同样的功能  
>> y2 = cdf('Normal',x,mu,sigma);  
>> plot(x,y2,'r-','LineWidth',1)  
>> title('累积分布曲线')
```



1. 通用函数计算累积概率值

例3：设二维随机变量 (X, Y) 的密度函数为

$$f(x, y) = \begin{cases} ce^{-(3x+4y)}, & x > 0, y > 0 \\ 0, & \text{otherwise} \end{cases}$$

(1) 求常数 c ;

```
>> syms x y c
```

(2) 求 (X, Y) 的联合分布函数;

```
>> fh = c*exp(-3*x-4*y);
```

(3) 求 $P\{0 < X < 1, 0 < Y < 2\}$;

```
>> ccoff = solve(int(int(fh,x,0,inf),y,0,inf) == 1,c) %归一性
```

```
c =
```

(4) $P\{X + Y \geq 1\}$;

```
12
```

(5) 边缘分布 $f_X(x)$

```
>> fh = subs(fh,c,ccoff) %密度函数
```

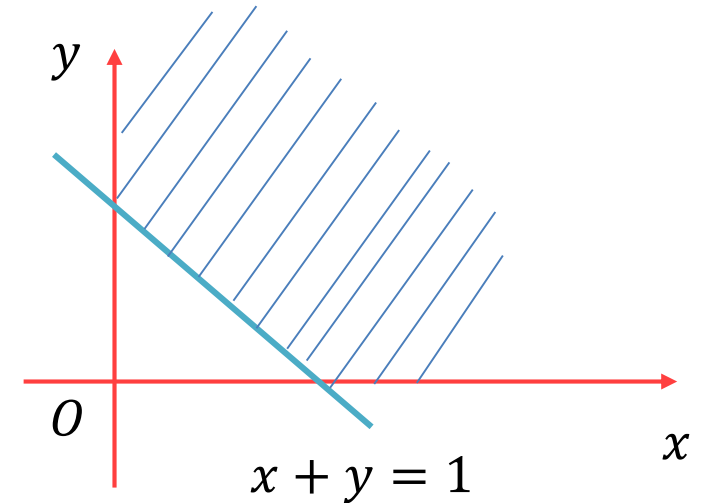
```
fh =
```

```
12*exp(- 3*x - 4*y)
```

1. 通用函数计算累积概率值

```
>> syms u v
>> fuv = subs(subs(fh,x,u),y,v)
fuv =
12*exp(- 3*u - 4*v)
>> Fxy = int(int(fuv,v,0,y),u,0,x) %求分布函数
Fxy =
(exp(-3*x) - 1)*(exp(-4*y) - 1)
>> P3 = int(int(fuv,v,0,2),u,0,1) %积分求解
P3 =
(exp(-3) - 1)*(exp(-8) - 1)
>> P3 = vpa(P3,15)
P3 =
0.949894170705024
```

```
>> P3 = vpa(subs(subs(Fxy,x,1),y,2),15) %直接带入分布函数
P3 =
0.949894170705024
>> P4 = 1 - int(int(fh,x,0,1-y),y,0,1) %求x+y < 1区域积分
P4 =
4*exp(-3) - 3*exp(-4)
>> P4 = vpa(P4,15)
P4 =
0.144201356805253
%求边缘分布密度
>> fx = int(fh,y,0,inf)
fx =
3*exp(-3*x)
```



2. 专用函数的累积概率值函数表

函数名	调用形式	注 释
unifcdf	unifcdf (x, a, b)	[a,b]上均匀分布(连续)累积分布函数值 $F(x)=P\{X\leq x\}$
unidcdf	unidcdf(x,n)	均匀分布（离散）累积分布函数值 $F(x)=P\{X\leq x\}$
expcdf	expcdf(x, Lambda)	参数为Lambda的指数分布累积分布函数值 $F(x)=P\{X\leq x\}$
normcdf	normcdf(x, mu, sigma)	参数为mu, sigma的正态分布累积分布函数值 $F(x)=P\{X\leq x\}$
chi2cdf	chi2cdf(x, n)	自由度为n的卡方分布累积分布函数值 $F(x)=P\{X\leq x\}$
tcdf	tcdf(x, n)	自由度为n的t分布累积分布函数值 $F(x)=P\{X\leq x\}$
fcdf	fcdf(x, n ₁ , n ₂)	第一自由度为n ₁ ,第二自由度为n ₂ 的F分布累积分布函数值
gamcdf	gamcdf(x, a, b)	参数为a, b的 分布 γ 累积分布函数值 $F(x)=P\{X\leq x\}$
betacdf	betacdf(x, a, b)	参数为a, b的 分布 β 累积分布函数值 $F(x)=P\{X\leq x\}$
logncdf	logncdf(x, mu, sigma)	参数为mu, sigma的对数正态分布累积分布函数值

2. 专用函数的累积概率值函数表

函数名	调用形式	注 释
nbincdf	nbincdf(x, R, P)	参数为R, P的负二项式分布概累积分布函数值 $F(x)=P\{X\leq x\}$
ncfcdf	ncfcdf(x, n ₁ , n ₂ , delta)	参数为n ₁ , n ₂ , delta的非中心F分布累积分布函数值
nctcdf	nctcdf(x, n, delta)	参数为n, delta的非中心t分布累积分布函数值 $F(x)=P\{X\leq x\}$
ncx2cdf	ncx2cdf(x, n, delta)	参数为n, delta的非中心卡方分布累积分布函数值
raylcdf	raylcdf(x, b)	参数为b的瑞利分布累积分布函数值 $F(x)=P\{X\leq x\}$
wblcdf	wblcdf(x, a, b)	参数为a, b的韦伯分布累积分布函数值 $F(x)=P\{X\leq x\}$
binocdf	binocdf(x,n,p)	参数为n, p的二项分布的累积分布函数值 $F(x)=P\{X\leq x\}$
geocdf	geocdf(x,p)	参数为 p的几何分布的累积分布函数值 $F(x)=P\{X\leq x\}$
hygecdf	hygecdf(x,M,K,N)	参数为 M, K, N的超几何分布的累积分布函数值
poisscdf	poisscdf(x,Lambda)	参数为Lambda的泊松分布的累积分布函数值 $F(x)=P\{X\leq x\}$

2. 专用函数的累积概率值函数表

例4: 某人向空中抛硬币100次，落下为正面的概率为0.5。这100次中正面向上的次数记为 x ：

(1)试计算 $x = 45$ 的概率和 $x \leq 45$ 的概率；

(2)绘制分布函数图象和分布列图象。

```
>> px=binopdf(45,100,0.5) % 计算x=45的概率值
```

```
px = 0.0485
```

```
>> fx=binocdf(45,100,0.5) % 计算x≤45的累计概率值
```

```
fx = 0.1841
```

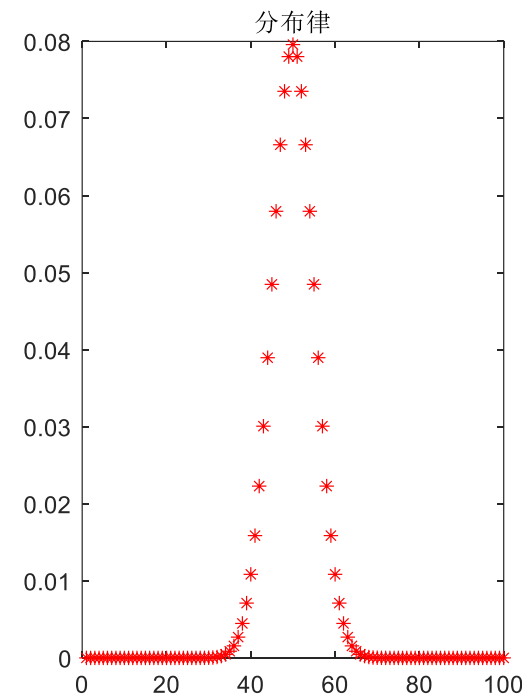
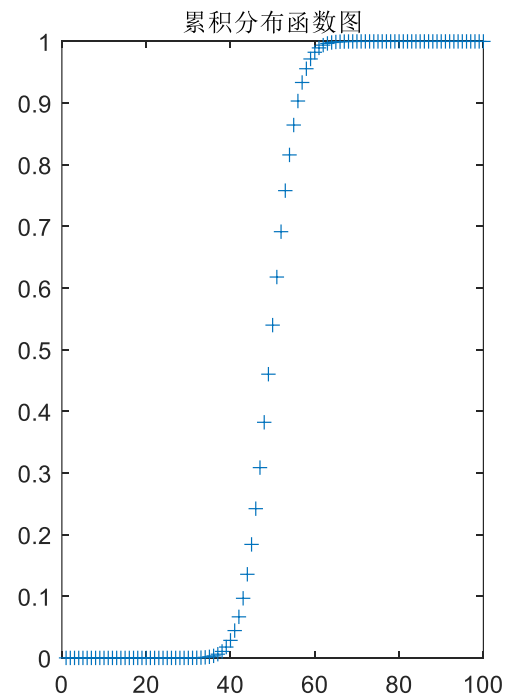
```
>> subplot(1,2,1);
```

```
>> x=1:100;
```

```
>> p1=binocdf(x,100,0.5); plot(x,p1,'+'); title('累积分布函数图')
```

```
>> subplot(1,2,2);
```

```
>> p2=binopdf(x,100,0.5); plot(x,p2,'*r'); title('分布律')
```



2. 专用函数的累积概率值函数表

例5: 设 $X \sim N(2, 0.5)$

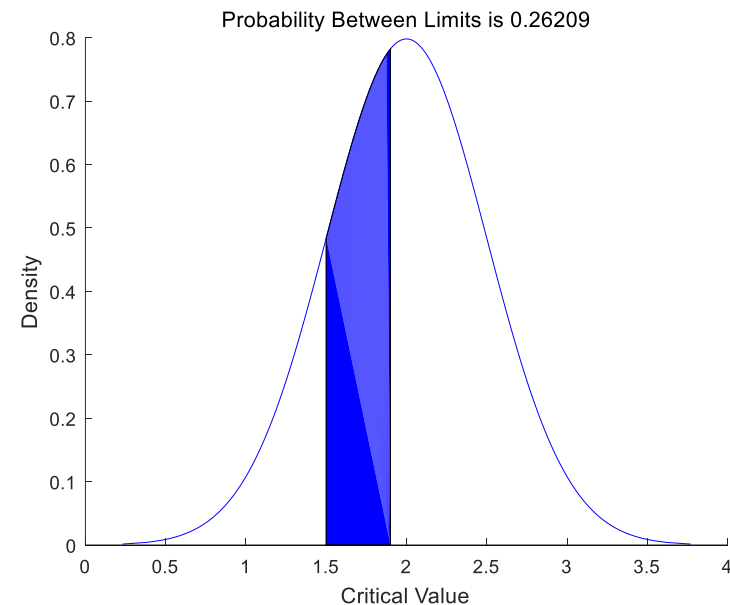
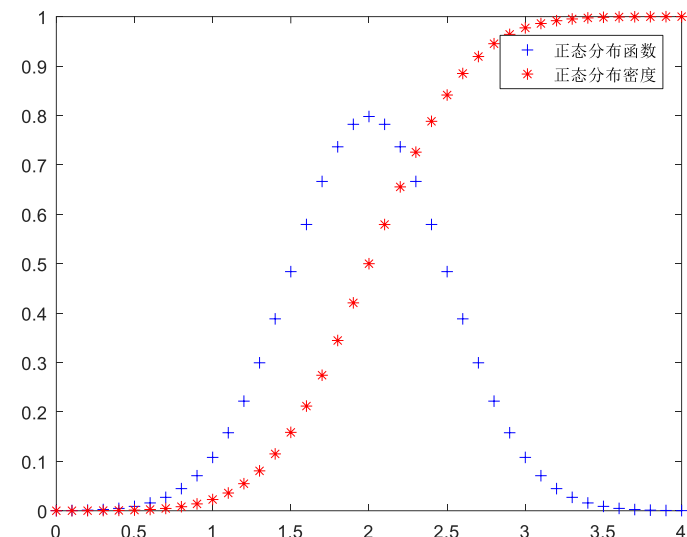
(1) 求概率 $P\{1 < X < 2.5\}$;

(2) 绘制分布函数图象和分布密度图象;

(3) 画出区间 $[1.5, 1.9]$ 上的分布密度曲线下方区域。

```
>> p=normcdf(2.5,2,0.5)-normcdf(1,2,0.5)
p = 0.8186

>> x=0:0.1:4; %绘制分布函数图象和分布密度图象;
>> px=normpdf(x,2,0.5);
>> fx= normcdf(x,2,0.5);
>> plot(x,px,'+b'); hold on;
>> plot(x,fx,'*r');
>> legend('正态分布函数','正态分布密度');
>> specs=[1.5,1.9]; %画出区间[1.5,1.9]上的分布密度曲线下方区域
>> pp=normspec(specs,2,0.5)
```





感谢聆听
