



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

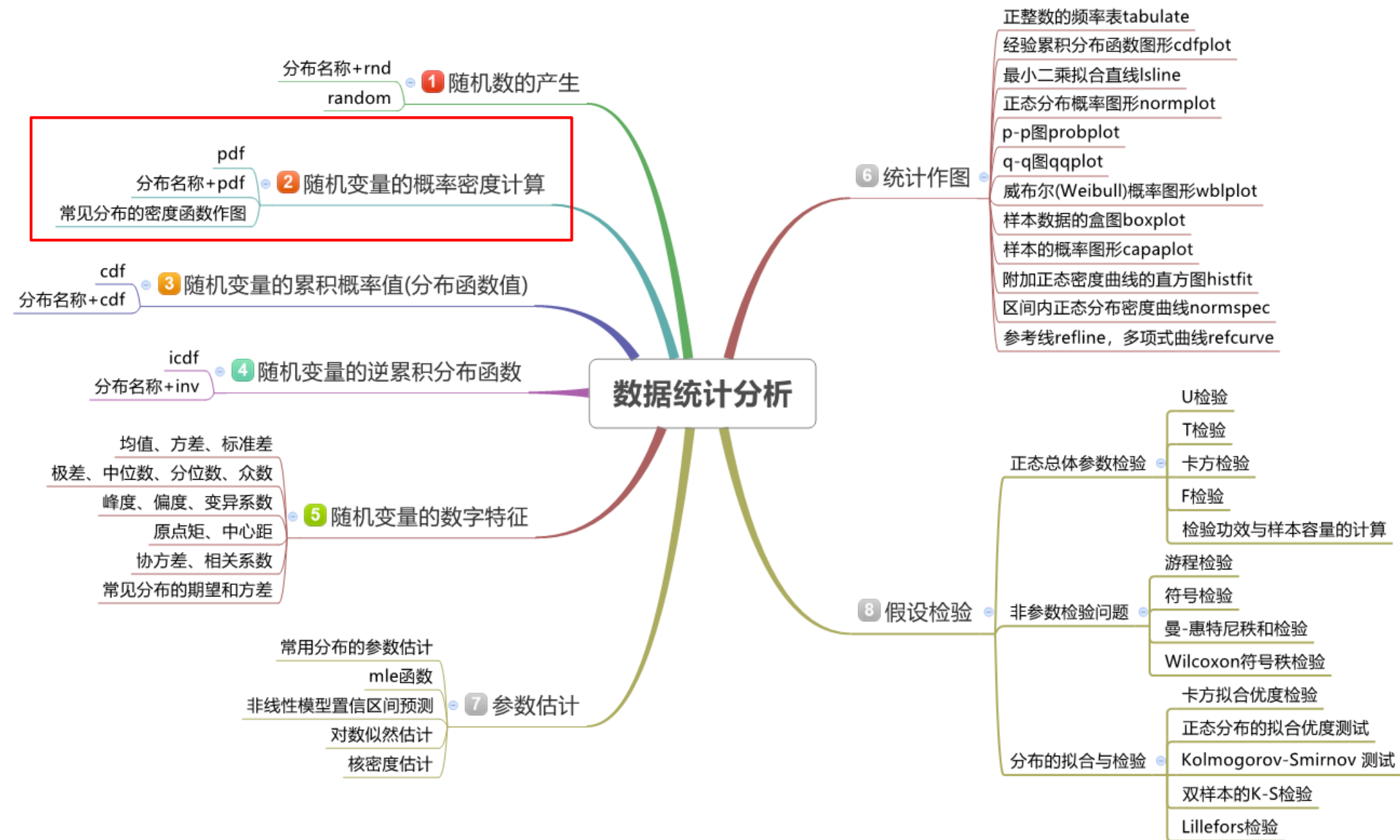


讲授人：牛言涛



日期：2020年4月6日

第10章 数据统计分析知识点思维导图



1. 通用函数计算概率密度函数值

通用函数pdf (probability density function) 计算概率密度函数值

- $Y = \text{pdf}(\text{name}, K, A)$
- $Y = \text{pdf}(\text{name}, K, A, B)$
- $Y = \text{pdf}(\text{name}, K, A, B, C)$

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

$F(x)$ 分布函数, $f(t)$ 概率密度函数

说明: 返回在 $X = K$ 处、参数为 A 、 B 、 C 的概率密度值, 对于不同的分布, 参数个数是不同; name 为分布函数名, 其取值如常见分布函数表。

2. 常见分布函数表

name的取值	函数说明	name的取值	函数说明
'beta'或'Beta'	Beta分布	'ncf'或'Noncentral F'	非中心F分布
'bino'或'Binomial'	二项分布	'nct'或'Noncentral t'	非中心t分布
'chi2'或'Chisquare'	卡方分布	'ncx2'或'Noncentral Chi-square'	非中心卡方分布
'exp'或'Exponential'	指数分布	'norm'或'Normal'	正态分布
'f'或'F'	F分布	'poiss'或'Poisson'	泊松分布
'gam'或'Gamma'	GAMMA分布	'rayl'或'Rayleigh'	瑞利分布
'geo'或'Geometric'	几何分布	't'或'T'	T分布
'hyge'或'Hypergeometric'	超几何分布	'unif'或'Uniform'	均匀分布
'logn'或'Lognormal'	对数正态分布	'unid'或'Discrete Uniform'	离散均匀分布
'nbin'或'Negative Binomial'	负二项式分布	'weib'或'Weibull'	Weibull分布

2. 常见分布函数表

例1: 计算正态分布 $N(0,1)$ 的随机变量 X 在点0.6578的密度函数值。

```
>> np = pdf('norm',0.6578,0,1)
np =
    0.3213
```

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

例2: 自由度为8的卡方分布，在点2.18处的密度函数值。

```
>> cp = pdf('chi2',2.18,8)
cp =
    0.0363
```

$$f(x,k) = \begin{cases} \frac{x^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

3. 专用函数计算概率密度函数表

函数名	调用形式	注 释
unifpdf	unifpdf (x, a, b)	[a,b]上均匀分布(连续)概率密度在 $X=x$ 处的函数值
unidpdf	unidpdf(x,n)	均匀分布（离散） 概率密度函数值
exppdf	exppdf(x, Lambda)	参数为Lambda的指数分布概率密度函数值
normpdf	normpdf(x, mu, sigma)	参数为mu, sigma的正态分布概率密度函数值
chi2pdf	chi2pdf(x, n)	自由度为n的卡方分布概率密度函数值
tpdf	tpdf(x, n)	自由度为n的t分布概率密度函数值
fpdf	fpdf(x, n ₁ , n ₂)	第一自由度为n ₁ ,第二自由度为n ₂ 的F分布概率密度函数值
gampdf	gampdf(x, a, b)	参数为a, b的 分布 γ 概率密度函数值
betapdf	betapdf(x, a, b)	参数为a, b的 分布 β 概率密度函数值
lognpdf	lognpdf(x, mu, sigma)	参数为mu, sigma的对数正态分布概率密度函数值

3. 专用函数计算概率密度函数表

函数名	调用形式	注 释
nbinpdf	nbinpdf(x, R, P)	参数为R, P的负二项式分布概率密度函数值
ncfpdf	ncfpdf(x, n ₁ , n ₂ , delta)	参数为n ₁ , n ₂ , delta的非中心F分布概率密度函数值
nctpdf	nctpdf(x, n, delta)	参数为n, delta的非中心t分布概率密度函数值
ncx2pdf	ncx2pdf(x, n, delta)	参数为n, delta的非中心卡方分布概率密度函数值
raylpdf	raylpdf(x, b)	参数为b的瑞利分布概率密度函数值
wblpdf	wblpdf(x, a, b)	参数为a, b的韦伯分布概率密度函数值
binopdf	binopdf(x,n,p)	参数为n, p的二项分布的概率密度函数值
geopdf	geopdf(x,p)	参数为 p的几何分布的概率密度函数值
hygepdf	hygepdf(x,M,K,N)	参数为 M, K, N的超几何分布的概率密度函数值
poisspdf	poisspdf(x,Lambda)	参数为Lambda的泊松分布的概率密度函数值

2. 专用函数计算概率密度函数值



例3: 某人射击的命中率为0.02，他独立射击400次，试求其命中次数不少于2次的概率。

```
>> x0 = binopdf(0,400,0.02); %二项分布命中0次
```

```
>> x1 = binopdf(1,400,0.02); %二项分布命中1次
```

```
>> x2 = 1 - x0 - x1
```

```
x2 =
```

```
0.9972
```

%泊松定理近似，取 $\lambda = np = 400 \times 0.02 = 8$

```
>> y0 = poisspdf(0,8); %泊松分布，命中0次
```

```
>> y1 = poisspdf(1,8); %泊松分布，命中1次
```

```
>> y2 = 1 - y0 - y1
```

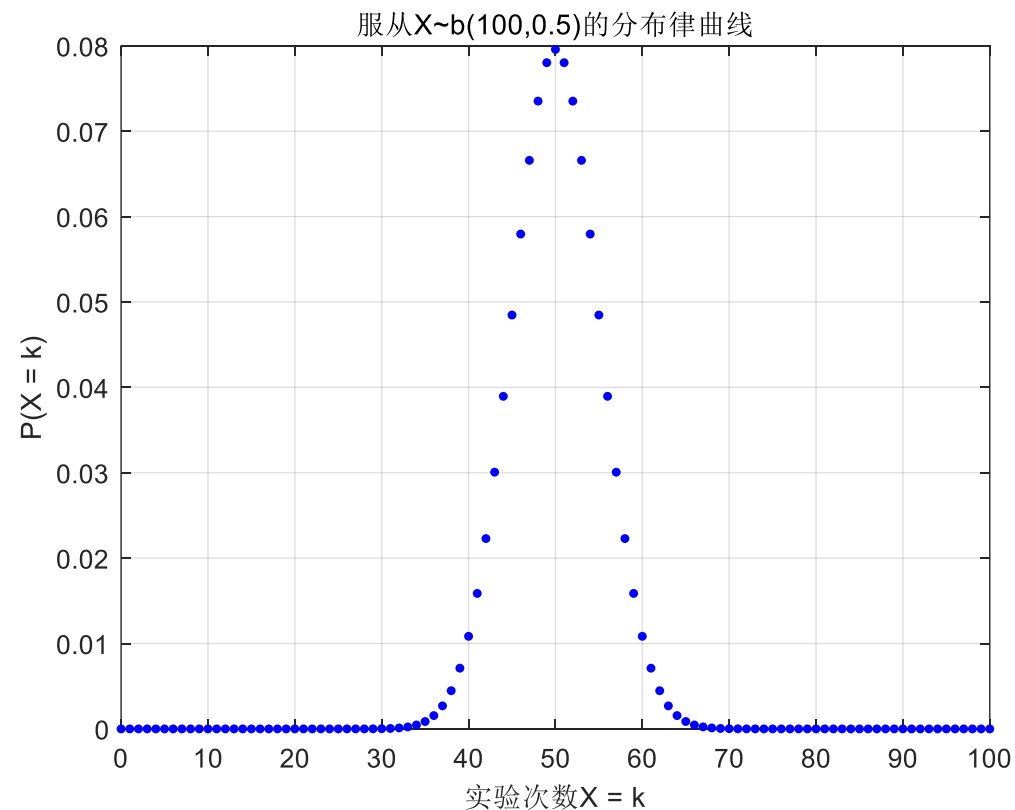
```
y2 =
```

```
0.9970
```


4. 常见分布的密度函数作图

1. 二项分布 $P\{X = k\} = C_n^k p^k (1-p)^{n-k}$

```
>> x = 0:100;  
% 100次实验, 每次实验事件A发生的概率为0.5  
>> y = binopdf(x,100,0.5);  
>> plot(x,y,'b.','MarkerSize',10)  
>> grid on  
>> title('服从X~b(100,0.5)的分布律曲线')  
>> xlabel('实验次数X = k')  
>> ylabel('P(X = k)')
```

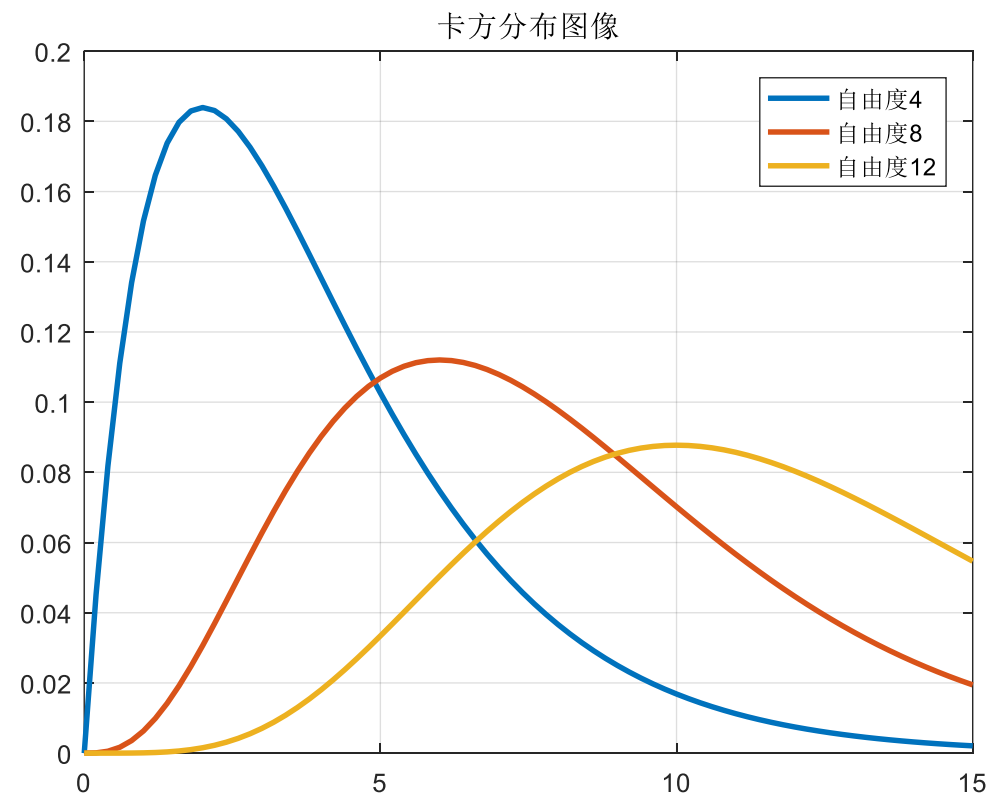


4. 常见分布的密度函数作图

2. 卡方分布

$$f_{\chi^2}(x) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

```
>> x = 0:0.2:15;  
>> plot(x,chi2pdf(x,4),'LineWidth',2)  
>> hold on  
>> plot(x,chi2pdf(x,8),'LineWidth',2)  
>> plot(x,chi2pdf(x,12),'LineWidth',2)  
>> legend('自由度4','自由度8','自由度12')  
>> grid on  
>> title('卡方分布图像')
```

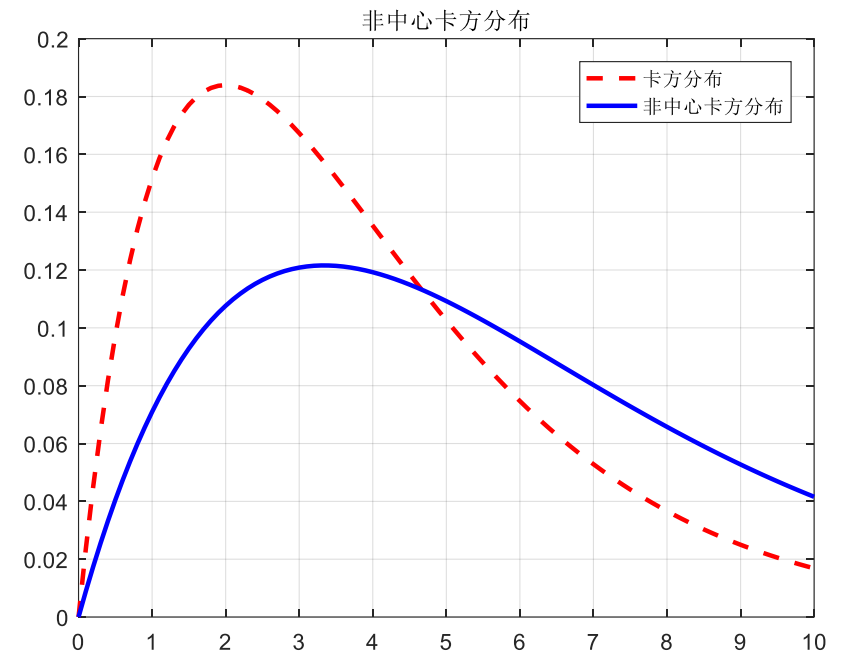


4. 常见分布的密度函数作图

3. 非中心卡方分布

$$f(x; k, \lambda) = \frac{1}{2} e^{-\frac{x+\lambda}{2}} \left(\frac{x}{\lambda} \right)^{\frac{k}{4} - \frac{1}{2}} I_{\frac{k}{2} - 1}(\sqrt{\lambda x})$$

```
>> x = (0:0.1:10)';  
>> p1 = ncx2pdf(x,4,2);  
>> p = chi2pdf(x,4);  
>> plot(x,p,'r--',x,p1,'b-','LineWidth',2)  
>> grid on  
>> title('非中心卡方分布')  
>> legend('卡方分布','非中心卡方分布')
```

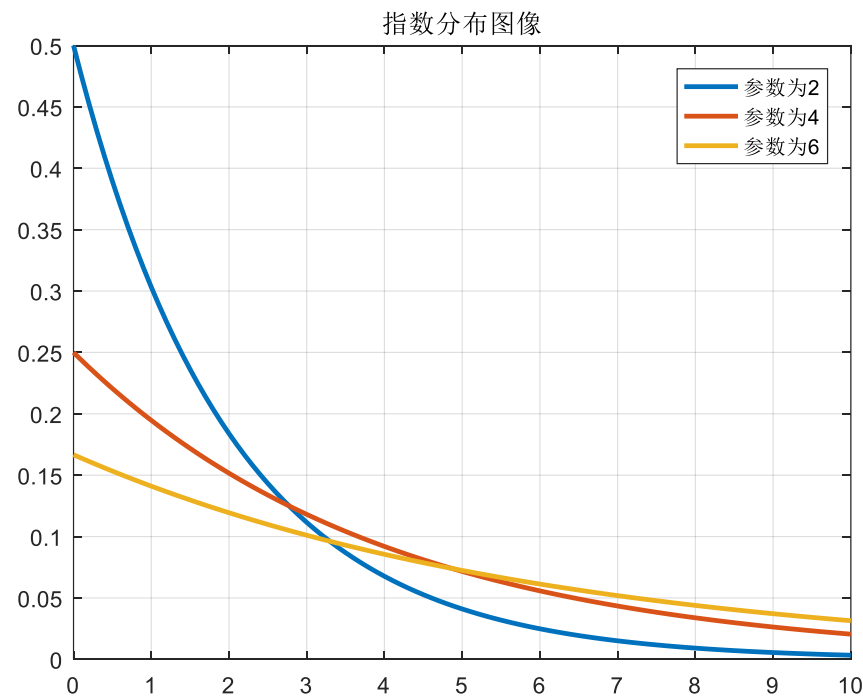


4. 常见分布的密度函数作图

4. 指数分布

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

```
>> x = 0:0.1:10;  
>> plot(x,exppdf(x,2),'LineWidth',2)  
>> hold on  
>> plot(x,exppdf(x,4),'LineWidth',2)  
>> plot(x,exppdf(x,6),'LineWidth',2)  
>> legend('参数为2','参数为4','参数为6')  
>> title('指数分布图像')  
>> grid on
```



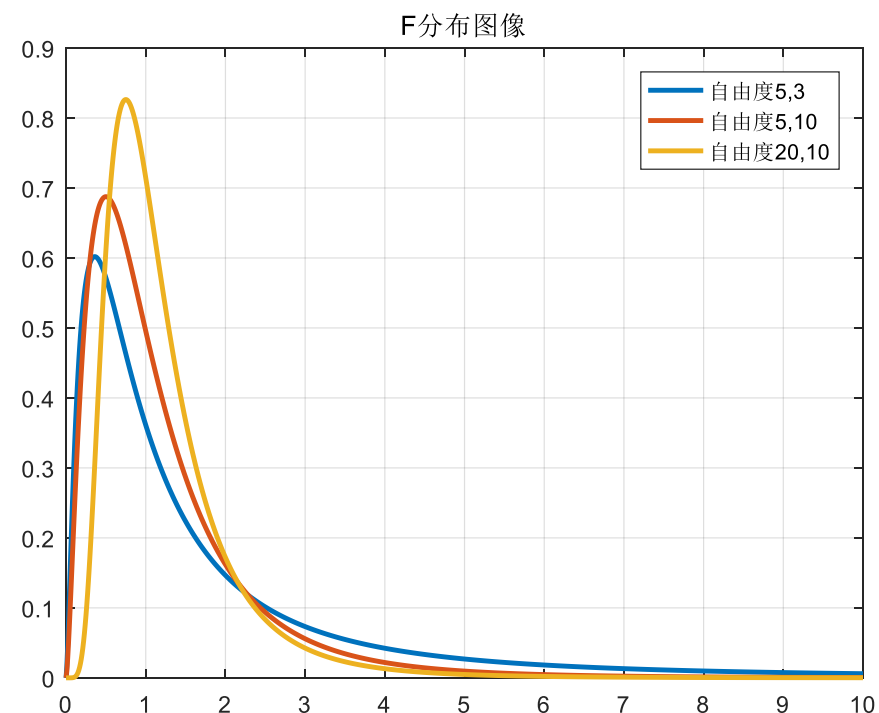
4. 常见分布的密度函数作图



5. F分布

$$f_F(x) = \begin{cases} \frac{\Gamma(\frac{p+q}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{q}{2})} p^{\frac{p}{2}} q^{\frac{q}{2}} x^{\frac{p}{2}-1} (p+qx)^{-\frac{p+q}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

```
>> x = 0:0.01:10;  
>> plot(x,fpdf(x,5,3),'LineWidth',2)  
>> hold on  
>> grid on  
>> plot(x,fpdf(x,5,10),'LineWidth',2)  
>> plot(x,fpdf(x,20,10),'LineWidth',2)  
>> legend('自由度5,3','自由度5,10','自由度20,10')  
>> title('F分布图像')
```

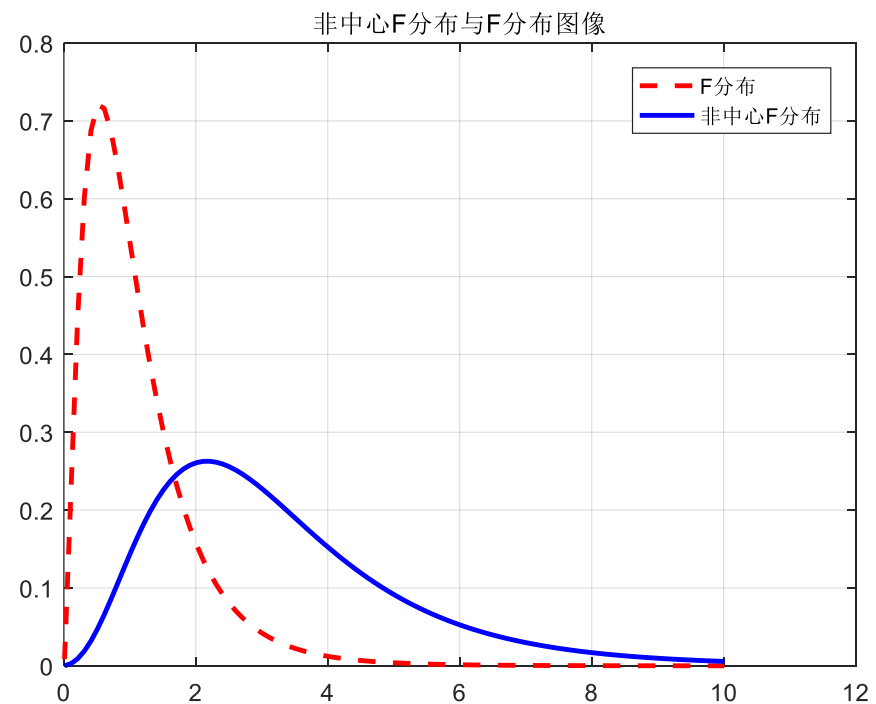


4. 常见分布的密度函数作图

6. 非中心F分布

$$p(x) = \begin{cases} \frac{m^{\frac{m}{2}} n^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} e^{-\frac{\sigma}{2}} x^{\frac{m}{2}-1} \sum_{k=0}^{\infty} \frac{\left(\frac{\sigma m x}{2}\right)^k \Gamma\left(\frac{m+n}{2} + k\right)}{\Gamma\left(\frac{m}{2} + k\right) (m x + n)^{\frac{m+n}{2} + k}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

```
>> x = (0.01:0.1:10.01)';  
>> p1 = ncfpdf(x,5,20,10);  
>> p = fpdf(x,5,20);  
>> plot(x,p,'r--',x,p1,'b-','LineWidth',2)  
>> grid on  
>> legend('F分布','非中心F分布')  
>> title('非中心F分布与F分布图像')
```

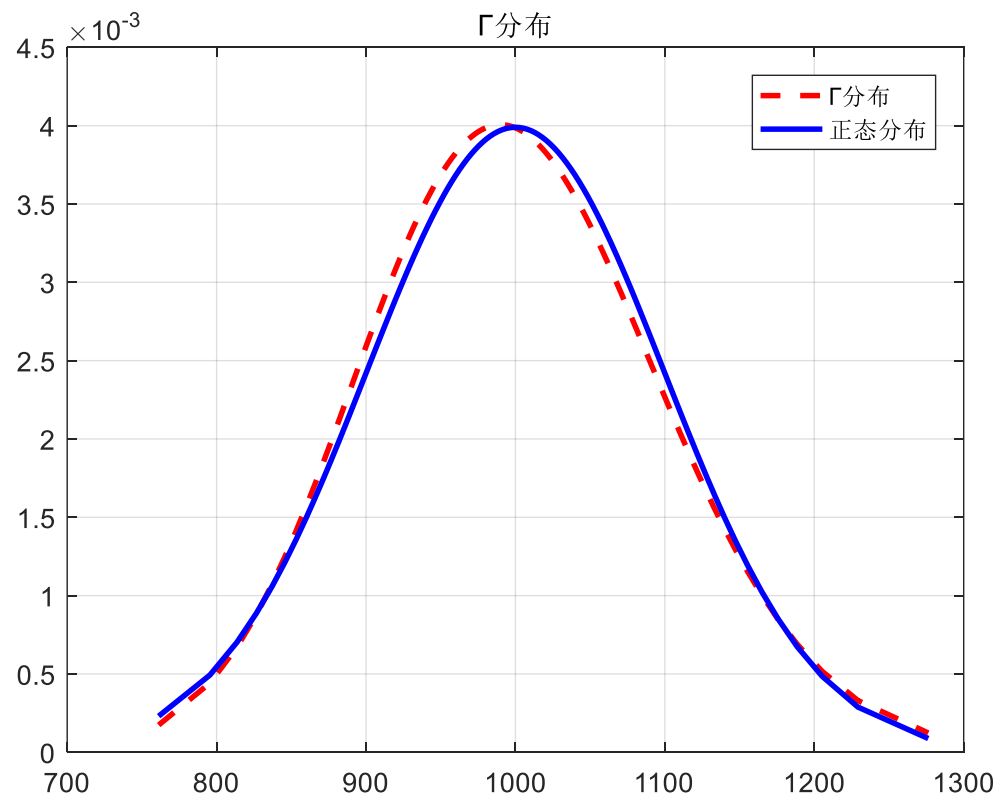


4. 常见分布的密度函数作图

7. 「分布

$$f_{\Gamma}(x) = \begin{cases} \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}, \quad \Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx$$

```
x = gaminv((0.005:0.01:0.995),100,10);  
y = gampdf(x,100,10);  
y1 = normpdf(x,1000,100);  
plot(x,y,'r--',x,y1,'b-','LineWidth',2)  
grid on  
legend('Γ分布','正态分布')  
title('Γ分布')
```



4. 常见分布的密度函数作图

8. 对数正态分布

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

```
x = (10:1000:125010)';
```

```
y = lognpdf(x,log(20000),1.0);
```

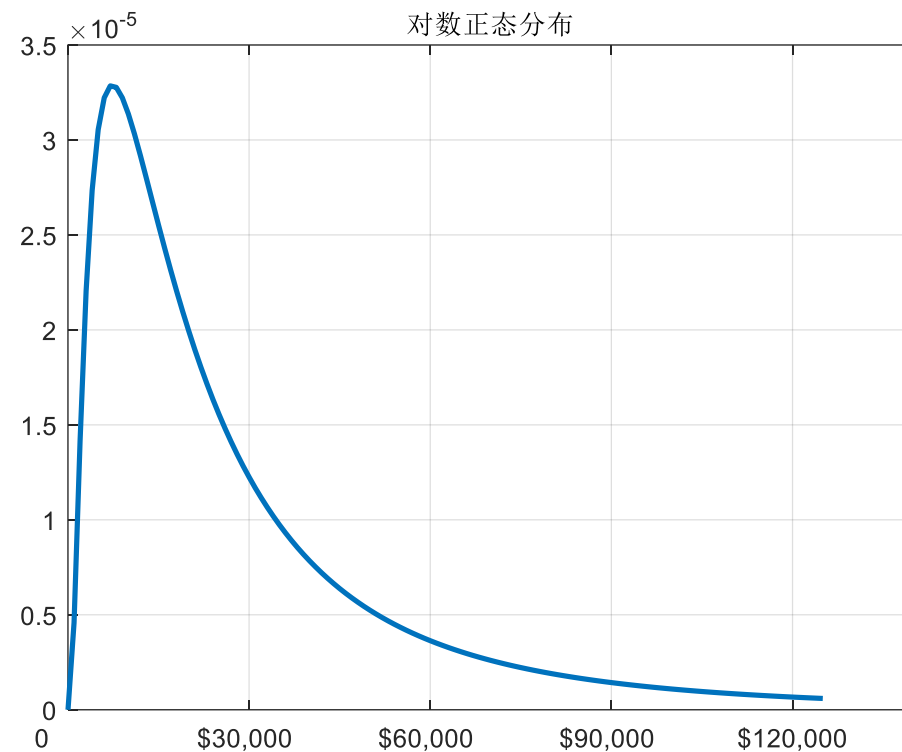
```
plot(x,y,'LineWidth',2)
```

```
grid on
```

```
set(gca,'xtick',[0 30000 60000 90000 120000])
```

```
set(gca,'xticklabel',str2mat('0','$30,000','$60,000','$90,000','$120,000'))
```

```
title('对数正态分布')
```

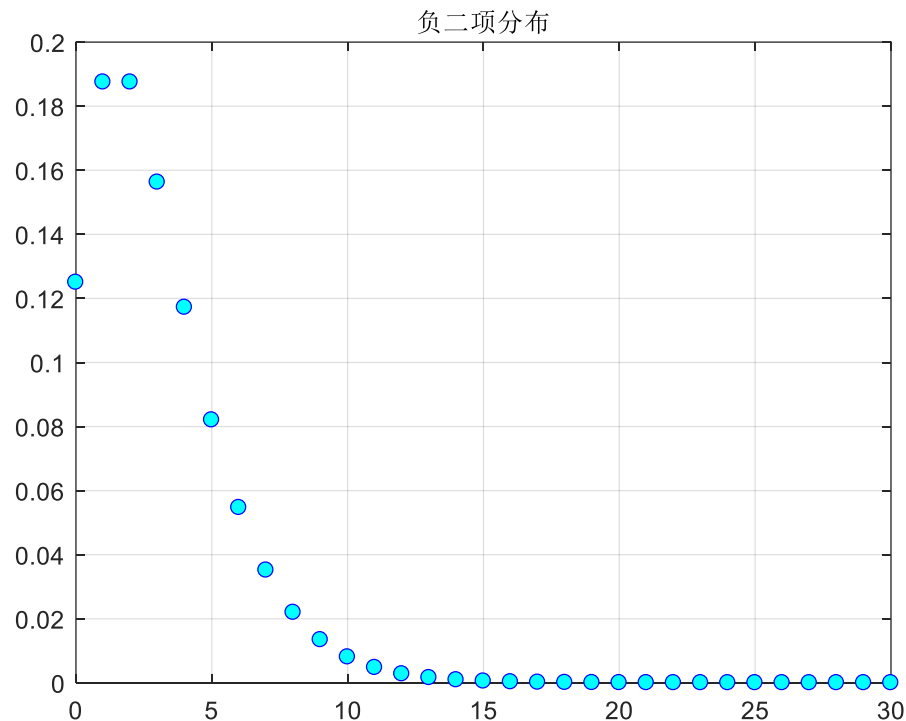


4. 常见分布的密度函数作图

9. 负二项分布

$$f(k; r, p) = \binom{k+r-1}{r-1} p^r (1-p)^k$$

```
>> x = 0:30;  
>> y = nbinpdf(x,3,0.5);  
>> plot(x,y,'bo','MarkerFaceColor','c')  
>> grid on  
>> title('负二项分布')
```



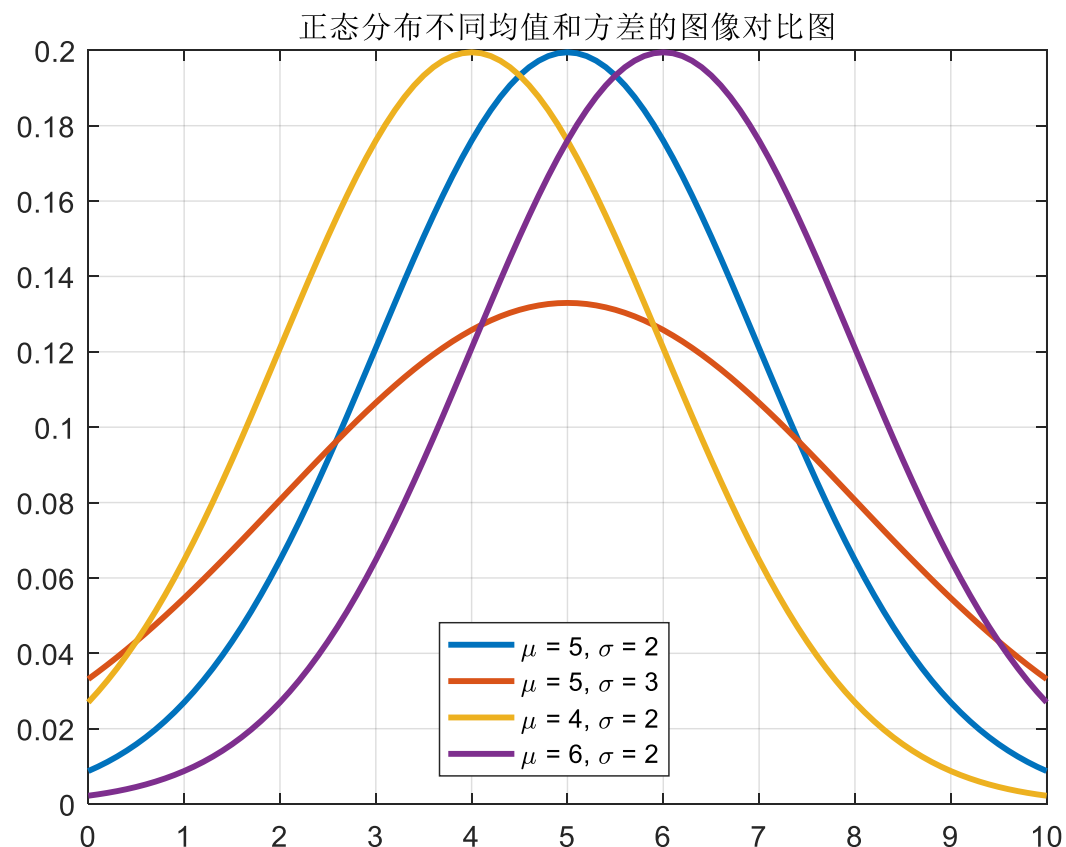
实验包含一系列独立的实验，每个实验都有成功、失败两种结果，成功的概率是恒定的，实验持续到r次成功，r为正整数。

4. 常见分布的密度函数作图

10. 正态分布

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

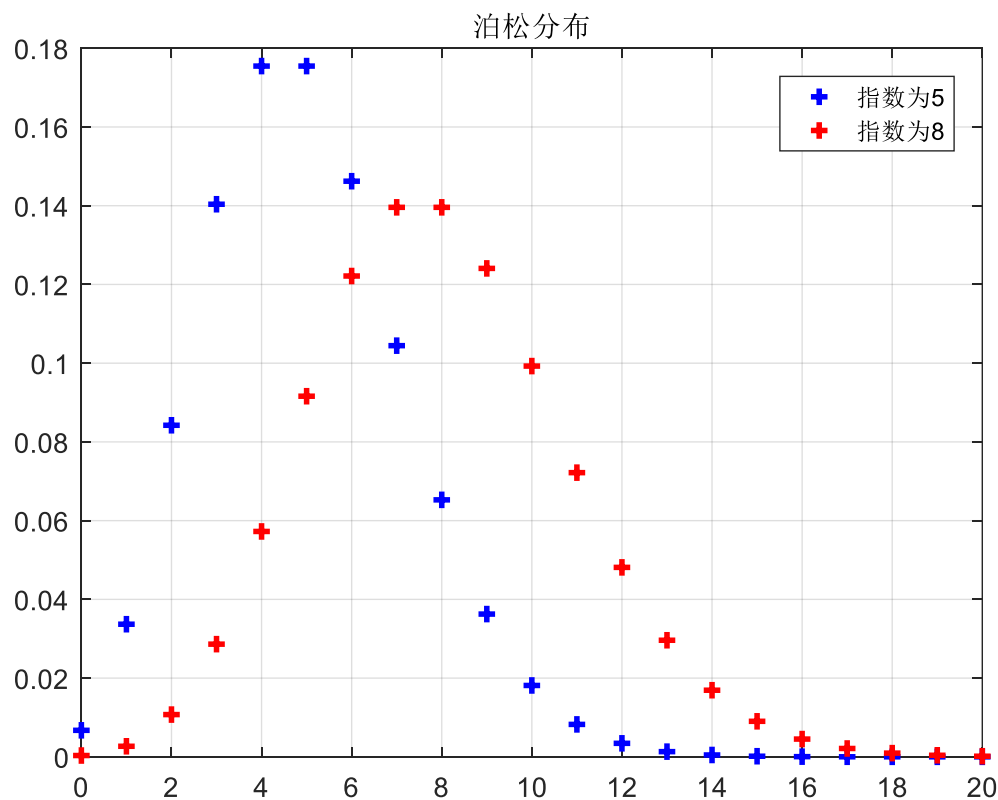
```
>> x = 0:0.1:10;  
>> y1 = normpdf(x,5,2);  
>> y2 = normpdf(x,5,3);  
>> y3 = normpdf(x,4,2);  
>> y4 = normpdf(x,6,2);  
>> plot(x,y1,x,y2,x,y3,x,y4,'LineWidth',2)  
>> grid on;  
>> legend('\mu = 5, \sigma = 2','\mu = 5, \sigma  
= 3','\mu = 4, \sigma = 2','\mu = 6, \sigma = 2')
```



4. 常见分布的密度函数作图

11. 泊松分布 $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$

```
>> x = 0:20;  
>> y1 = poisspdf(x,5);  
>> plot(x,y1,'LineWidth',2)  
>> plot(x,y1,'b+','LineWidth',2)  
>> y2 = poisspdf(x,8);  
>> hold on  
>> plot(x,y2,'r+','LineWidth',2)  
>> grid on  
>> legend('指数为5','指数为8')  
>> title('泊松分布')
```



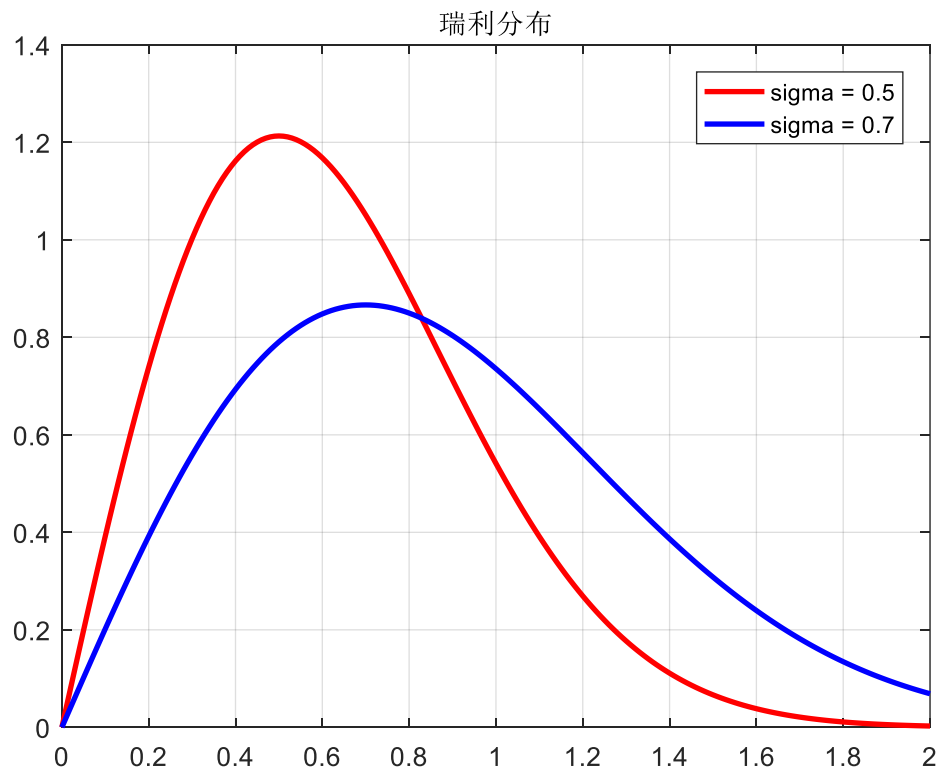
4. 常见分布的密度函数作图

12. 瑞利分布

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0$$

```
>> x = 0:0.01:2;  
>> p1 = raylpdf(x,0.5);  
>> p2 = raylpdf(x,0.7);  
>> plot(x,p1,'r-',x,p2,'b-', 'LineWidth',2)  
>> grid on  
>> legend('sigma = 0.5','sigma = 0.7')  
>> title('瑞利分布')
```

当一个随机二维向量的两个分量呈独立的、有着相同的方差的正态分布时，这个向量的模呈瑞利分布。



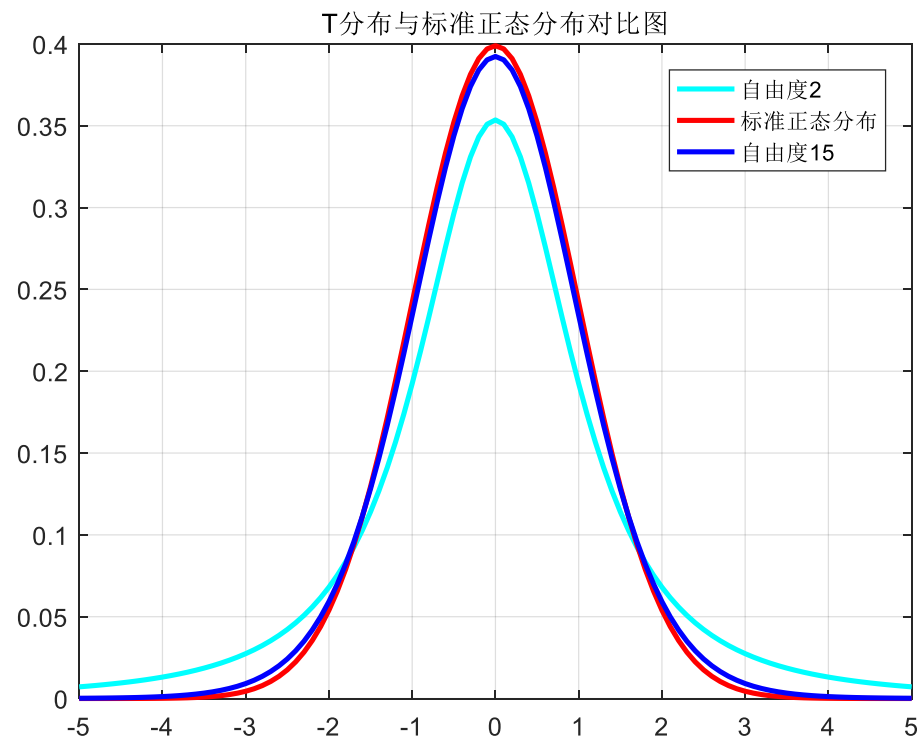
4. 常见分布的密度函数作图

13. T分布

$$f_T(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}$$

```
x = -5:0.1:5;  
y = tpdf(x,5);  
y2 = tpdf(x,15);  
>> z = normpdf(x,0,1);  
plot(x,y,'c-',x,z,'r-',x,y2,'b-','LineWidth',2)  
grid on  
legend('自由度2','标准正态分布','自由度15')
```

t 分布的上 α 分位点可通过附表查得，在 $n > 45$ 时，用**标准正态分布**的上 α 分位点近似



4. 常见分布的密度函数作图

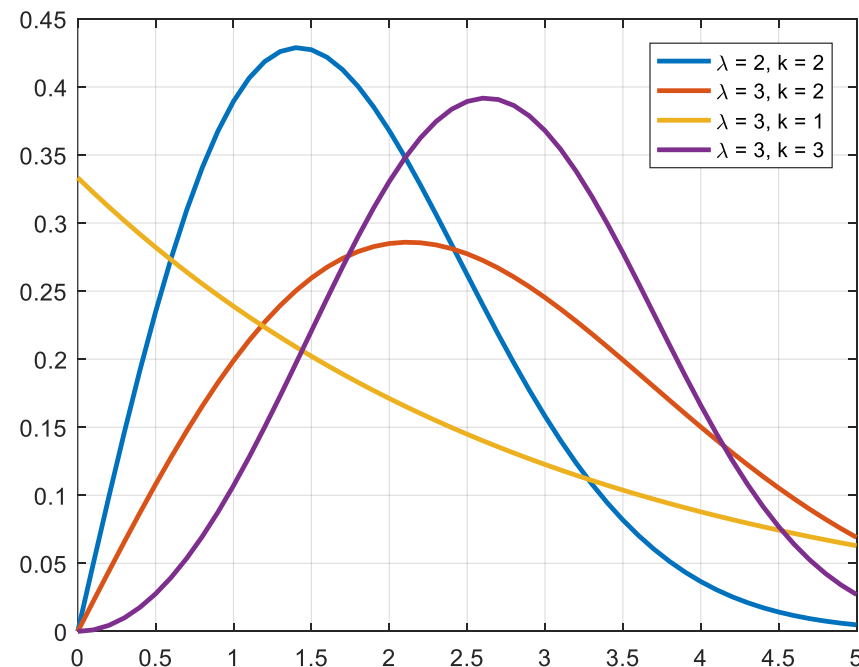


14. 威布尔分布

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda} \right)^k}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

威布尔分布在可靠性工程中被广泛应用，尤其适用于机电类产品的磨损累计失效的分布形式。由于它可利用概率值很容易推断出它的分布参数，被广泛应用于各种寿命试验的数据处理。

```
t = 0:0.1:5;  
y1 = wblpdf(t,2,2); plot(t,y1,'LineWidth',2)  
y2 = wblpdf(t,3,2);  
hold on; grid on  
plot(t,y2,'LineWidth',2)  
y3 = wblpdf(t,3,1); plot(t,y3,'LineWidth',2)  
y4 = wblpdf(t,3,3); plot(t,y4,'LineWidth',2)  
legend('\lambda = 2, k = 2', '\lambda = 3, k = 2', '\lambda = 3, k =  
1', '\lambda = 3, k = 3')
```





感谢聆听
