```
  b = a(1:10^i,1)&a(1:10^i,2)&a(1:10^i,3);
  d(i) = sum(b);
  e(i) = d(i)./c(i);
end
e
e =      0        0  0.0114  0.0266  0.0263  0.0277
e =      0   0.0313  0.0263  0.0290  0.0264  0.0278
e =      0        0  0.0163  0.0272  0.0282  0.0278
e =      0   0.0345  0.0393  0.0273  0.0273  0.0275
e = 0.3333   0.0345  0.0340  0.0278  0.0276  0.0278
a = round(rand(1000000,10) - 0.2);
for i = 1:6
  b = sum(a(1:10^i,:),2) - 3;
  c(i) = sum(~b)/(10^i);
end
c
c = 0.3000  0.2600  0.3000  0.2712  0.2665  0.2666
c = 0.4000  0.2200  0.2660  0.2713  0.2686  0.2674
c = 0.2000  0.1900  0.2470  0.2613  0.2662  0.2665
c = 0.4000  0.2600  0.2840  0.2727  0.2670  0.2666
c = 0.4000  0.2900  0.2580  0.2674  0.2668  0.2669
```

从上述模拟结果可以看出,模拟的结果与理论计算是吻合的.请读者读懂上述两段程序,并写出模拟的思路.

第二节　随机变量的分布及其数字特征

随机变量的统计行为完全决定于其概率分布.按随机变量的取值不同,通常可将其分为离散型、连续型和奇异型三大类.由于奇异型在实际应用中很难碰到,因而这里我们不作讨论,仅讨论离散型和连续型两类随机变量的概率分布及其数字特征.

一、离散型随机变量的分布及其数字特征

如果随机变量 X 的所有可能取值为有限个或无穷可列个,则称 X 为离散型随机变量.设 X 的所有可能值为 $X_1, X_2, \cdots$,并且 X 取这些值的概率为

$$P\{X = X_k\} = p_k, \quad k = 1, 2, \cdots,$$

则称其为随机变量 X 的概率分布.它满足以下性质:

① $p_k \geqslant 0$,　$k=1,2,\cdots$,　　② $\sum_{k=1}^{\infty} p_k = 1$.

称 $F(x)=\sum_{X_k \leqslant X} p_k$ 为累积概率分布.我们研究一个随机变量,就是研究随机变量的概率分布、累积概率分布和分布的数字特征,有时也需要求给定显著概率条件下假设检验的临界值,这实际上是累积分布函数的反函数问题,称为逆累积分布函数.常用的离散型随机变量的分布有:两点分布、二项分布、泊松分布和超几何分布等.在二项分布中,当 $n=1$ 时为两点分布.因此,我们只讨论二项分布、泊松分布和超几何分布.

1. 二项分布

若随机变量 X 的所有可能取值为 $0,1,\cdots,n$,其概率分布为

$$P\{X=k\}=C_n^k p^k q^{n-k},\quad k=0,1,2,\cdots,n,$$

其中 $q=1-p$,则称 X 服从参数为 n 和 p 的二项分布,记作 $X \sim B(n,p)$.显然,两点分布是二项分布的特例.二项分布的数学期望为 $E(X)=np$,方差为 $D(X)=npq$.

在 MATLAB 中提供有二项分布的统计函数:binopdf( )、binocdf( )、binoinv( )、binornd( ) 以及计算二项分布均值和方差的函数 binostat( ),其使用格式为

binopdf(X,N,P)	二项分布的密度函数
binocdf(X,N,P)	二项分布的累积分布函数
binoinv(Y,N,P)	二项分布的逆累积分布函数
binornd(N,P,m,n)	产生服从二项分布的随机数
binostat(N,P)	求二项分布的数学期望与方差

其中 X 为随机变量;N 为独立试验的重复数;P 为事件发生的概率;m 和 n 分别是所产生随机矩阵的行数和列数.若不指定 m 和 n,则返回一个随机数,否则返回一个服从二项分布的 $m \times n$ 阶随机矩阵.可用下述程序描绘出二项分布的概率分布图和累积概率分布图.由于 MATLAB 不提供绘分段函数图象的命令,故可借助如下函数 sstep()描绘累积函数分布图,其用法如下

sstep(Y)或 sstep(X,Y)

其中 X 用于指定画线位置,Y 表示线相对坐标轴的高度,还可增加线的属性.

```
function [xo, yo] = sstep(varargin)
error(nargchk(1, 3, nargin));
sym = [];
if isstr(varargin{nargin}),
```

```
  sym = varargin{nargin};
  [msg,x,y] = xychk(varargin{1:nargin-1},'plot');
  if ~isempty(msg), error(msg); end
else
  [msg,x,y]=xychk(varargin{1:nargin},'plot');
  if ~isempty(msg), error(msg); end
end
if min(size(x))==1, x = x(:); end
if min(size(y))==1, y = y(:); end
[n,nc] = size(y);
ndx = [1:n;1:n];
y2 = y(ndx(1:2*n-1),:);
if size(x,2)==1,
  x2 = x(ndx(2:2*n),ones(1,nc));
else
  x2 = x(ndx(2:2*n),:);
end
x2(2*n)=2*x2(2*n-1)-x2(2*n-3);
y2(2*n)=y2(2*n-1);
if (nargout < 2)
  if isempty(sym),
     for i=1:n
        hold on;plot(x2(2*i-1:2*i),y2(2*i-1:2*i));
     end
     hold off;
  else
    for i=1:n
       hold on;plot(x2(2*i-1:2*i),y2(2*i-1:2*i),sym);
    end
    hold off;
  end
  else
    xo = x2;
    yo = y2;
end
```

然后利用该函数可绘出不同试验重复数 n 和不同概率 p 下二项分布的函数分布图和累积分布函数图(图 5 - 1).程序如下

```
x = 0:70;
y1 = binopdf(x,30,0.67);z1 = binocdf(x,30,0.67);
y2 = binopdf(x,50,0.67);z2 = binocdf(x,50,0.67);
y3 = binopdf(x,80,0.67);z3 = binocdf(x,80,0.67);
subplot(2,2,1);plot(x,y1,'k.',x,y2,'k.',x,y3,'k.');
subplot(2,2,2);sstep(x,z1,'k');
sstep(x,z2,'k');sstep(x,z3,'k');
y1 = binopdf(x,50,0.3);z1 = binocdf(x,50,0.3);
y2 = binopdf(x,50,0.6);z2 = binocdf(x,50,0.6);
y3 = binopdf(x,50,0.9);z3 = binocdf(x,50,0.9);
subplot(2,2,3);plot(x,y1,'k.',x,y2,'k.',x,y3,'k.');
subplot(2,2,4);sstep(x,z1,'k');
sstep(x,z2,'k');sstep(x,z3,'k');
```

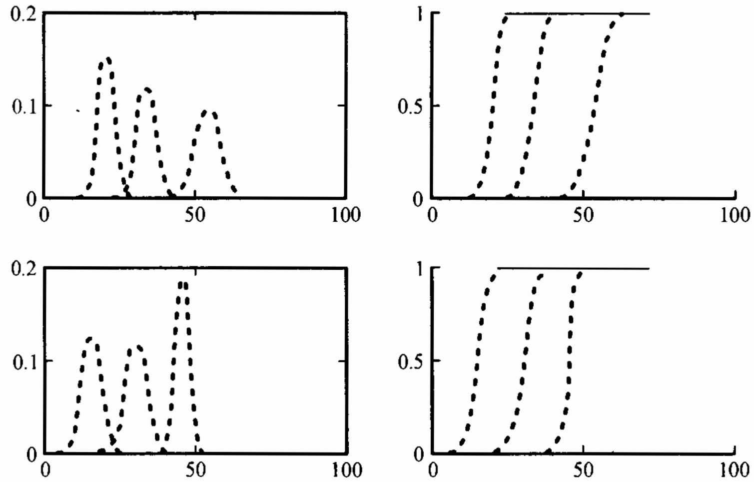

图 5 - 1　二项分布的概率密度与累积概率分布图

还可以利用逆累积概率分布函数求得一定显著概率条件下,二项分布假设检验的临界值,求出的结果均为整数,即应试验的最少重复数.如

```
x = 0:0.1:1
x = 0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.90  1.0
binoinv(x,10,1.7)
```

```
ans =0  5  6  6  7  7  7  8  8  9  10
binoinv(x,10,0.3)
ans =0  1  2  2  3  3  3  4  4  5  10
binoinv(x,50,0.7)
ans =0  31  32  33  34  35  36  37  38  39  50
binoinv(x,50,0.3)
ans =0  11  12  13  14  15  16  17  18  19  50
```

生成一个或多个服从二项分布的随机数.如

```
binornd(10,0.7)                ans =  6
binornd(10,0.7,5,10)
ans = 7  8  6   7  8  8  8  7  9  7
      7  6  9   6  6  5  9  7  9  7
      5  6  5   8  7  7  8  9  8  5
      6  6  8   9  3  4  9  8  7  7
      7  8  7  10  8  5  7  7  9  5
```

求二项分布的数学期望(e)和方差(d)

```
[e d] = binostat(10,0.3)      e =    3    d =    2.1000
[e d] = binostat(20,0.7)      e =   14    d =    4.2000
```

作为二项分布的特例,对于两点分布,可用下述程序绘出两点分布的概率密度图和累积概率分布图,程序及运行结果如下(图 5-2)

```
x = 0:1;
y = binopdf(x,1,0.7);z = binocdf(x,1,0.7);
subplot(1,2,1);plot(x,y,'k.');axis([-0.1,1.2,0,0.8]);
subplot(1,2,2);sstep(x,z,'k');axis([0,2,0,1.1]);
```

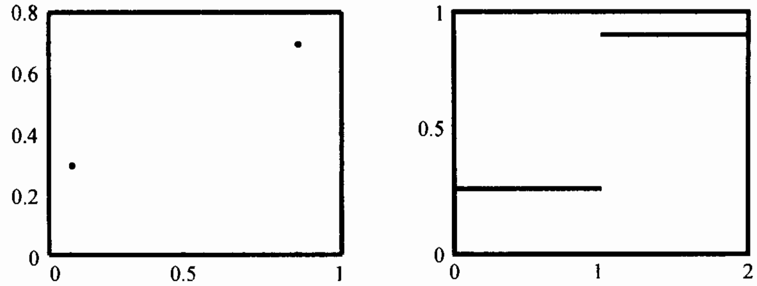

图 5-2　两点分布的概率密度与累积概率分布图

2. 泊松分布

如果随机变量的概率分布为

$$P\{X=k\}=\frac{\lambda^k}{k!}\exp(-\lambda),\quad k=0,1,2,\cdots,$$

其中 $\lambda>0$ 为常数,则称 X 服从参数为 λ 的泊松分布,记作 $X\sim P(\lambda)$,泊松分布的数学期望 $E(X)=\lambda$,方差 $D(X)=\lambda$.

在 MATLAB 中,提供如下有关泊松分布的统计函数,使用格式为

poisspdf(X,LMD)	泊松分布的密度函数
poisscdf(X,LMD)	泊松分布的累积分布函数
poissinv(Y,LMD)	泊松分布的逆累积分布函数
poissrnd(LMD,M,N)	产生服从泊松分布的随机数
poissstat(LMD)	求泊松分布的数学期望与方差

其中 X 为随机变量;Y 为显著概率值;LMD 为参数;M 和 N 为产生随机矩阵的行数和列数.例如类似于二项分布可用下述程序绘出服从泊松分布的密度函数和累积分布函数图(图 5-3)

```
x = 0:5;
y1 = poisspdf(x,0.3);z1 = poisscdf(x,0.3);
y2 = poisspdf(x,0.6);z2 = poisscdf(x,0.6);
y3 = poisspdf(x,0.9);z3 = poisscdf(x,0.9);
subplot(2,1,1);plot(x,y1,'k.',x,y2,'k.',x,y3,'k.');
subplot(2,1,2);sstep(x,z1,'k');
sstep(x,z2,'k');sstep(x,z3,'k');
```

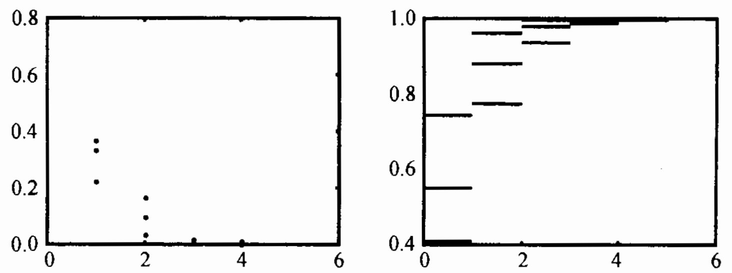

图 5-3　泊松分布的概率密度与累积概率分布图

利用逆累积概率分布函数求一定显著概率条件下,泊松分布假设检验临界值

```
x = 0:0.1:1;
x = 0.0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9  1.0
poissinv(x,5)
ans  = 1  2  3  4  4  5  5  6  7  8  Inf
poissinv(x,10)
```

```
ans =1  6  7  8  9  10  11  12  13  14  Inf
poissinv(x,100)
ans =1  87  92  95  97  100  102  105  108  113  Inf
poissinv(x,1)
ans =1  1  1  1  1  1  1  1  2  2  Inf
```

求服从泊松分布的随机数及数学期望与方差如下

```
poissrnd(1)            ans =  1
poissrnd(5)            ans =  5
poissrnd(5,5,10)
ans = 4  6  9  1  4  10  7  3  3  6
      7  5  4  3  4  5   4  6  5  2
      6  4  3  8  2  6   4  5  5  6
      3  8  6  8  8  8   1  6  3  7
      4  4  10 6  7  4   5  4  2  3
[e,d] = poisstat(5)     e =  5      d = 5
[e,d] = poisstat(10)    e = 10      d =10
```

3. 超几何分布

如果随机变量 X 的所有可能取值为 $0,1,2,\cdots,L(L=\min\{M,N\})$，$X$ 的概率分布为

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}\qquad(k=0,1,2,\cdots,L),$$

其中整数 $M,N>0$，且 $n\leqslant N-M$，则称 X 服从参数为 N,M,n 的超几何分布，记作 $X\sim H(N,M,n)$. MATLAB 中有关超几何分布的统计函数为

hygepdf(M,n,k,N)　　超几何分布的密度函数

hygepcdf(M,n,k,N)　　超几何分布的累积分布函数

hygeinv(P,n,k,N)　　超几何分布的逆累积分布函数

hygestat(n,k,N)　　超几何分布的数学期望与方差

其中各参数的意义与定义相同. 产生随机数的函数为

hygernd(n,k,N,mr,mc)

其中 mr 与 mc 分别为所产生随机矩阵的行数和列数，省略时产生一个随机数. 绘制超几何分布的密度函数和累积分布函数图形的程序如下(图 5－4)

```
x = 0:5;
y1 = hygepdf(x,10,5,6);z1 = hygecdf(x,10,5,6);
y2 = hygepdf(x,15,5,9);z2 = hygecdf(x,15,5,9);
y3 = hygepdf(x,20,8,10);z3 = hygecdf(x,20,8,10);
```

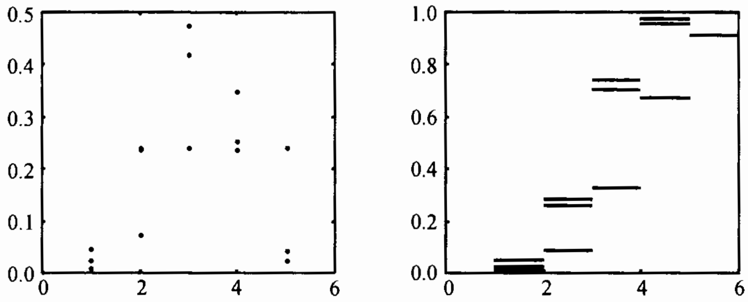

图 5-4　超几何分布的概率密度与累积概率分布图

```
subplot(1,2,1);plot(x,y1,'k.',x,y2,'k.',x,y3,'k.');
subplot(1,2,2);sstep(x,z1,'k');
sstep(x,z2,'k');sstep(x,z3,'k');
```

用逆累积概率分布函数求一定显著概率条件下,超几何分布假设检验临界值如下

```
x = 0:0.1:1;
x = 0.0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9  1.0
hygeinv(x,10,5,6)
ans =0  2  2  3  3  3  3  3  4  4  5
hygeinv(x,15,5,9)
ans =0  2  2  3  3  3  3  3  4  4  5
hygeinv(x,20,8,10)
ans =0  3  3  3  4  4  4  5  5  5  8
hygeinv(x,15,7,9)
ans =0  3  3  4  4  4  4  5  5  5  7
```

求服从超几何分布的随机数及数学期望与方差如下

```
hygernd(15,7,9)            ans = 5
hygernd(15,7,9,5,10)
ans =5  4  4  5  4  5  6  3  4  4
     5  3  4  5  3  6  2  5  5  3
     5  3  5  4  4  3  5  4  6  3
     5  3  3  6  4  3  2  6  4  4
     4  5  5  4  5  4  4  3  4  5
[e,d] = hygestat(15,7,9)        e = 4.2000          d = 0.9600
[e,d] = hygestat(20,8,10)       e = 4               d = 1.2632
```

二、连续型随机变量的分布及其数字特征

设随机变量 X 的分布函数为 $F(x)$，若存在非负函数 $f(x)$，使对任意实数 x，有

$$F(x)=P\{X\leqslant x\}=\int_{-\infty}^{x}f(x)\mathrm{d}x,$$

则称 X 为连续型随机变量，并称 $f(x)$ 为 X 的概率密度，它满足以下性质：

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

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

③ $P\{a<x\leqslant b\}=F(b)-F(a)=\int_{a}^{b}f(x)\mathrm{d}x$；

④ $P\{x=a\}=0$

最后一点和离散型随机变量截然不同，它表明概率为零的事件并不一定是不可能事件. 常用的三种连续型随机变量的概率分布是均匀分布、指数分布和正态分布.

1. 均匀分布

若连续型随机变量 X 的概率密度为

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

则称 X 在区间 (a,b) 上服从参数为 a 和 b 的均匀分布，记作 $X\sim U(a,b)$. MATLAB 提供的有关均匀分布的函数如下

unifpdf(X,A,B)	均匀分布的密度函数
unifcdf(X,A,B)	均匀分布的累积分布函数
unifinv(P,A,B)	均匀分布的逆累积分布函数
unirnd(A,B,m,n)	均匀分布的随机数发生器
unifstat(A,B)	均匀分布的数学期望与方差

其中 X 为随机变量；P 为概率值；A,B 为均匀分布参数；m 和 n 为生成随机数矩阵的行数和列数. 绘制均匀分布的密度函数及累积分布函数图的程序如下(图 5－5)

```
x = 0:10;
y = unifpdf(x,3,7);z = unifcdf(x,3,7);
x1 = [x(1:3) 3 x(4:8) 7 x(9:11)];
y = [y(1:3) 0 y(4:8) 0 y(9:11)];
subplot(1,2,1);plot(x1,y,'k');
```

```
axis([0,10,-0.01,0.3]);
subplot(1,2,2);plot(x,z,'k');
axis([0,10,-0.1,1.1]);
```

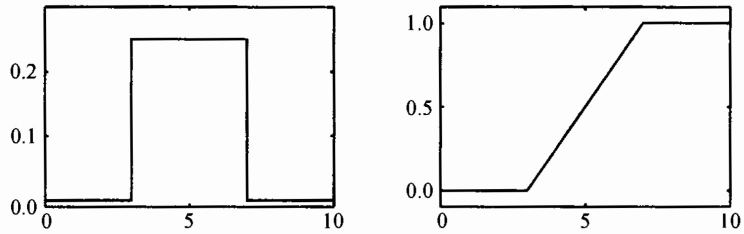

图 5-5 均匀分布的密度函数及累积分布函数图

用逆累积概率分布函数求一定显著概率条件下,均匀分布假设检验临界值如下

```
x=0:0.1:1
x =0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9  1.0
y=unifinv(x,3,7)
y =3.0  3.4  3.8  4.2  4.6  5.0  5.4  5.8  6.2  6.6  7.0
y=unifinv(x,2,8)
y =2.0  2.6  3.2  3.8  4.4  5.0  5.6  6.2  6.8  7.4  8.0
y=unifinv(x,4,6)
y =4.0  4.2  4.4  4.6  4.8  5.0  5.2  5.4  5.6  5.8  6.0
```

求服从均匀分布的随机数如下

```
unifrnd(3,7)              ans = 6.8005
unifrnd(3,7,5,7)
ans =5.7289  4.5135  6.5991  4.3679  4.2372  5.1863  6.8274
     4.2111  6.4400  6.2865  4.1589  6.3540  4.7795  5.0904
     5.1667  6.4146  5.5796  4.3648  5.2723  5.7783  6.5206
     3.6035  5.3743  6.2719  5.1363  4.4817  5.4852  3.6918
     5.7916  4.9862  5.6409  5.9085  5.8110  6.1793  6.9190
```

求均匀分布的数学期望和方差如下

```
[e d]=unifstat(3,7)        e = 5        d = 1.3333
[e d]=unifstat(2,8)        e = 5        d = 3
[e d]=unifstat(1,5)        e = 3        d = 1.3333
```

2. 指数分布

如果随机变量 X 的概率密度为

$$f(x)=\begin{cases}\lambda\exp(-\lambda x), & x\geqslant 0,\\ 0, & x<0\end{cases}$$

其中 λ 为常数,则称 X 服从参数为 λ 的指数分布,记作 $X\sim e(\lambda)$. MATLAB 提供的有关指数分布的函数如下

exppdf(X,L)	指数分布的密度函数
expcdf(X,L)	指数分布的累积分布函数
expinv(P,L)	指数分布的逆累积分布函数
exprnd(X,L,m,n)	产生服从指数分布的随机数
expstat(L)	求指数分布的数学期望与方差

其中 X 为随机变量,L 为参数 λ,P 为显著概率,m 和 n 为随机数矩阵的行数和列数. 绘制指数分布密度函数和累积分布函数图形的程序如下(图 5-6)

```
x = -0.1:0.001:0.4;
y = exppdf(x,0.05);z = expcdf(x,0.05);
subplot(1,2,1);plot(x,y,'k');
axis([-0.1,0.4,-0.1,21]);
subplot(1,2,2);plot(x,z,'k');
axis([-0.1,0.4,-0.1,1.1]);
```

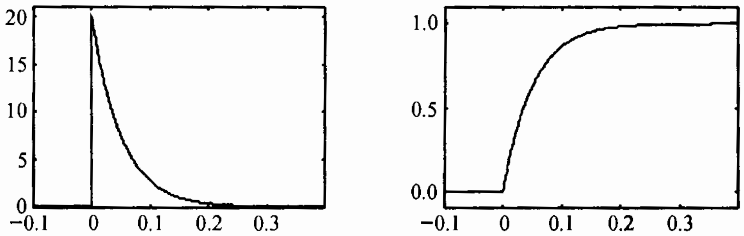

图 5-6　指数分布的密度函数及累积分布函数图

用逆累积分布函数求一定显著概率条件下,指数分布假设检验的临界值如下

```
x = 0:0.1:1
x = 0   0.1000   0.2000   0.3000   0.4000   0.5000   0.6000   0.7000
    0.8000   0.9000   1.0000
expinv(x,0.05)
ans = 0   0.0053   0.0112   0.0178   0.0255   0.0347   0.0458
      0.0602   0.0805   0.1151      0
expinv(x,0.1)
ans = 0   0.0105   0.0223   0.0357   0.0511   0.0693   0.0916
      0.1204   0.1609   0.2303      0
```

```
expinv(x,1)
ans =0   0.1054   0.2231   0.3567   0.5108   0.6931   0.9163
      1.2040   1.6094   2.3026      0
expinv(x,10)
ans =0   1.0536   2.2314   3.5667   5.1083   6.9315   9.1629
      12.0397   16.0944   23.0259      0
expinv(x,0.001)
ans =0   0.0001   0.0002   0.0004   0.0005   0.0007   0.0009
      0.0012   0.0016   0.0023      0
```

求服从指数分布随机数的例子如下

```
exprnd(0.05)              ans = 0.0652
exprnd(0.05,5,7)
ans = 0.0689   0.0056   0.0378   0.0331   0.0137   0.0122   0.1414
      0.0066   0.0807   0.1368   0.0548   0.0318   0.0192   0.0253
      0.0152   0.0604   0.0006   0.0419   0.0223   0.0387   0.1495
      0.0996   0.0207   0.0270   0.0744   0.0783   0.0283   0.0439
      0.2222   0.0629   0.0430   0.0273   0.0484   0.0115   0.0594
```

求指数分布的数学期望与方差的例子如下

```
[e d] = expstat(0.05)       e = 0.0500       d = 0.0025
[e d] = expstat(0.5)        e = 0.5000       d = 0.2500
[e d] = expstat(5)          e = 5            d = 25
```

3. 正态分布

如果随机变量 X 的概率密度为

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

其中 μ 和 σ 均为常数,且 $\sigma>0$,则称 X 服从参数为 μ 和 σ^2 的正态分布,记作 $X\sim N(\mu,\sigma^2)$. 当 $\mu=0,\sigma=1$ 时,称 X 服从标准正态分布,记作 $X\sim N(0,1)$. MATLAB 提供的有关正态分布的函数如下

normpdf(X,M,C)　　正态分布的密度函数
normcdf(X,M,C)　　正态分布的累积分布函数
norminv(P,M,C)　　正态分布的逆累积分布函数
normrnd(M,C,m,n)　　产生服从正态分布的随机数
normstat(M,C)　　求正态分布的数学期望和方差

其中 X 为随机变量,M 为正态分布参数 μ,C 为参数 σ,P 为显著概率,m 和 n 为随机矩阵的行数和列数. 绘制标准正态分布的密度函数及累积分布函数图(图

5－7 上）和一般正态分布的密度函数及累积分布函数图（图 5－7 下）的程序如下

```
x = -4:0.01:4;
y = normpdf(x, 0, 1);z = normcdf(x, 0, 1);
subplot(2, 2, 1);plot(x, y,'k');
axis([-4, 4, -0.1, 0.5]);
subplot(2, 2, 2);plot(x, z,'k');
axis([-4, 4, -0.1, 1.1]);
x = -4:0.01:16;
y1 = normpdf(x, 6, 1);z1 = normcdf(x, 6, 1);
y2 = normpdf(x, 6, 4);z2 = normcdf(x, 6, 4);
y3 = normpdf(x, 6, 0.6);z3 = normcdf(x, 6, 0.6);
subplot(2, 2, 3);plot(x, y1,'k', x, y2,'k', x, y3,'k');
axis([-4, 16, -0.1, 0.8]);
subplot(2,2,4);plot(x, z1,'k', x, z2,'k', x, z3,'k');
axis([-4, 16, -0.1, 1.1]);
```

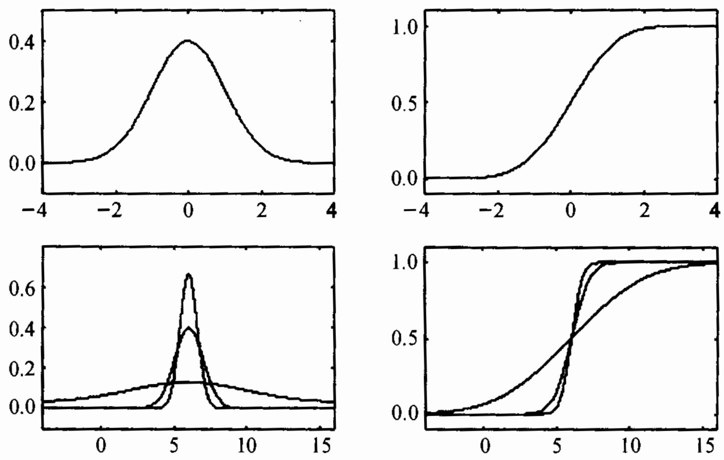

图 5－7　正态分布的密度函数及累积分布函数图

用正态分布的逆累积分布函数求一定显著概率条件下，假设检验临界值如下

```
x = 0:0.1:1
x  = 0   0.1000   0.2000   0.3000   0.4000   0.5000   0.6000
   0.7000   0.8000   0.9000   1.0000
norminv(x, 0, 1)
```

```
ans = - Inf  - 1.2816  - 0.8416  - 0.5244  - 0.2533   0  0.2533
   0.5244  0.8416  1.2816  Inf
normminv(x, 0.1, 0.01)
ans = - Inf  0.0872  0.0916  0.0948  0.0975  0.1000  0.1025
   0.1052  0.1084  0.1128  Inf
norminv(x, 2, 4)
ans = - Inf   - 3.1262   - 1.3665   - 0.0976   0.9866
   2.0000  3.0134   4.0976  5.3665  7.1262  Inf
norminv(x, 2, 9)
ans = - Inf   - 9.5340   - 5.5746   - 2.7196   - 0.2801  2.0000
   4.2801  6.7196  9.5746  13.5340  Inf
norminv(x, 2, 25)
ans = - Inf  - 30.0388  - 19.0405  - 11.1100  - 4.3337  2.0000
   8.3337   15.1100  23.0405  34.0388  Inf
```

产生服从标准正态分布或一般正态分布的随机数的例子如下

```
normrnd(0, 1)          ans =  - 0.4326
normrnd(2, 4)          ans =  - 4.6623
normrnd(2, 4, 5, 7)
ans =   2.5013    1.8495   - 0.3533    2.2371   4.8573   - 4.3749
        5.2625    3.1507     3.3092   10.7327   1.6174     8.4942
      - 3.7639    4.8476   - 2.5859    2.6986   1.4544   - 1.3294
      - 0.7671    4.2846     7.1610    6.7637   1.2532     2.4557
        3.1776    5.4320     0.4005    4.6744   6.7567     4.9032
        6.2671  - 3.3447     7.0160    4.7600   6.7634
normrnd(0, 1, 5, 7)
ans = - 1.2025  - 1.0565  - 0.9219    0.5077  - 1.0091    1.0950
        0.5779  - 0.0198    1.4151  - 2.1707    1.6924  - 0.0195
      - 1.8740    0.0403  - 0.1567  - 0.8051  - 0.0592    0.5913
      - 0.0482    0.4282    0.6771  - 1.6041    0.5287  - 1.0106
      - 0.6436    0.0000    0.8956    0.5689    0.2573    0.2193
        0.6145    0.3803  - 0.3179    0.7310  - 0.2556
```

求正态分布数学期望和方差的例子如下

```
[e d] = normstat(0, 1)      e = 0      d = 1
[e d] = normstat(2, 4)      e = 2      d = 16
[e d] = normstat(2, 0.1)    e = 2      d = 0.0100
```

三、χ^2 分布、t 分布和 F 分布

设二维随机变量(X,Y) 的分布函数为 $F(x,y)$,如果存在非负函数 $f(x,y)$,使对任意实数 x,y,有

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

则称(X,Y)为连续型随机变量,称 $f(x,y)$为二维连续型随机变量(X,Y)的概率密度,或称为 X 与 Y 的联合概率密度.

若二维随机变量最多只能取有限对或无穷可列对值(X_i,Y_j), $i,j=1,2,\cdots$,则称(X,Y)为二维离散型随机变量,称(X,Y)的所有可能取值的概率规律

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots$$

为二维离散型随机变量(X,Y)的概率分布,或称为 X 与 Y 的联合概率分布.

统计分析中常用的分布为正态分布所产生的抽样分布,常用的有 χ^2 分布、t 分布和 F 分布.

1. χ^2 分布

设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,且均服从 $N(0,1)$,令

$$X=X_1^2+X_2^2+\cdots+X_n^2,$$

则 X 的分布称为具有自由度 n 的 χ^2 分布,记作 $X\sim\chi^2(n)$. MATLAB 同样提供有关 χ^2分布的如下五个函数

chi2pdf(X,N)	χ^2 分布的密度函数
chi2cdf(X,N)	χ^2 分布的累积分布函数
chi2inv(P,N)	χ^2 分布的逆累积分布函数
chi2rnd(N,m,n)	产生服从 χ^2 分布的随机数
chi2stat(N)	求 χ^2 分布的数学期望与方差

其中 X 为随机变量,N 为χ^2 分布自由度,P 为显著概率,m 和 n 为产生的随机数矩阵的行数和列数. 可用下述程序绘制 χ^2 分布的密度函数和累积分布函数图(图 5－8)

```
x = 0:0.01:40;
y = chi2pdf(x,1);z = chi2cdf(x,1);
y1 = chi2pdf(x,4);z1 = chi2cdf(x,4);
y2 = chi2pdf(x,10);z2 = chi2cdf(x,10);
y3 = chi2pdf(x,20);z3 = chi2cdf(x,20);
subplot(1,2,1);
plot(x,y,'k',x,y1,'k',x,y2,'k',x,y3,'k');
axis([0,40,0,0.2]);
```

```
subplot(1,2,2);
plot(x,z,'k',x,z1,'k',x,z2,'k',x,z3,'k');
axis([0,40,0,1.01]);
```

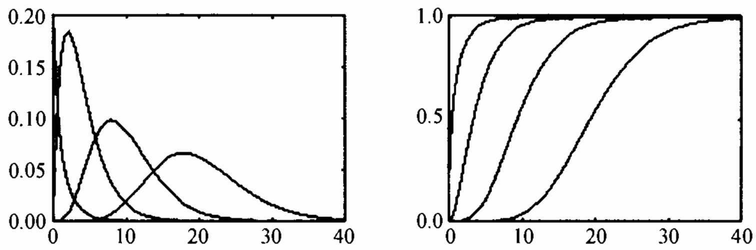

图 5-8 χ^2 分布的密度函数和累积分布函数图

用 χ^2 分布的逆累积分布函数求一定显著概率条件下,假设检验临界值的例子如下

```
x=0:0.1:1
x =0   0.1000   0.2000   0.3000   0.4000   0.5000   0.6000
  0.7000   0.8000   0.9000   1.0000
chi2inv(x,4)
ans =0   1.0636   1.6488   2.1947   2.7528   3.3567   4.0446
    4.8784   5.9886   7.7794   Inf
chi2inv(x,10)
ans =0   4.8652   6.1791   7.2672   8.2955   9.3418   10.4732
    11.7807   13.4420   15.9872   Inf
chi2inv(x,20)
ans =0   12.4426   14.5784   16.2659   17.8088   19.3374   20.9514
    22.7745   25.0375   28.4120   Inf
chi2inv(x,1)
ans =0   0.0158   0.0642   0.1485   0.2750   0.4549   0.7083
    1.0742   1.6424   2.7055   Inf
chi2inv(0.99,1)      ans = 6.6349
chi2inv(0.95,1)      ans = 3.8415
```

产生服从 χ^2 分布随机数的例子如下

```
chi2rnd(4)        ans = 11.4762
chi2rnd(10)       ans = 9.6514
chi2rnd(4,5,7)
```

```
ans =7.7609  1.4901   4.0327  7.8515  3.9358  2.2073  1.6876
     4.8208  3.0873   1.1200  2.4287  3.1461  7.3635  1.2443
     1.5995  5.2958  10.1221  0.5890  5.3455  4.6007  6.3982
     2.1422  9.2001   8.9214  5.6819  2.9725  1.4937  2.2308
     5.8455  4.4831   1.1640  1.4917  1.7174  9.8505  0.3866
chi2rnd(10,5,7)
ans =  9.2207   6.1305  13.4944   4.8023  13.8081  10.2618  12.2339
      12.6050  13.1922  11.2796  16.5278  13.3437   5.0874   4.0443
       8.4511  10.9290   9.5872   7.6733   7.0667   7.2528   7.0863
      10.2110   7.4108  11.8578  12.7922   7.9745  11.8391  21.1212
       2.5985   9.0337   6.7035   8.8331   2.7342   6.6494  10.9755
```

求 χ^2 分布数学期望与方差的例子如下

```
[e d]=chi2stat(4)        e = 4        d = 8
[e d]=chi2stat(10)       e = 10       d = 20
```

2. t 分布

设随机变量 X 与 Y 独立,且 $X\sim N(0,\ 1)$, $Y\sim\chi^2(n)$,令

$$T=\frac{X}{\sqrt{Y/n}},$$

则称 T 为具有自由度 n 的 t 分布,记为 $T\sim t(n)$.MATLAB 提供了有关 t 分布的五个函数如下

tpdf(X,N)	t 分布的密度函数
tcdf(X,N)	t 分布的累积分布函数
tinv(P,N)	t 分布的逆累积分布函数
trnd(N,m,n)	产生服从 t 分布的随机数
tstat(N)	求 t 分布的数学期望与方差

其中 X 为随机变量,N 为 t 分布的自由度,P 为显著概率,m 和 n 为所产生随机数矩阵的行数和列数.用下述程序可绘制 t 分布的密度函数和累积分布函数图.程序及运行结果如下(图 5-9)

```
x=-5:0.01:5;
y=tpdf(x,1);z=tcdf(x,1);
y1=tpdf(x,2);z1=tcdf(x,2);
y2=tpdf(x,10);z2=tcdf(x,10);
subplot(1,2,1);plot(x,y,'k',x,y1,'k',x,y2,'k');
axis([-5,5,0,0.4]);
subplot(1,2,2);plot(x,z,'k',x,z1,'k',x,z2,'k');
```

```
axis([-5,5,0,1.01]);
```

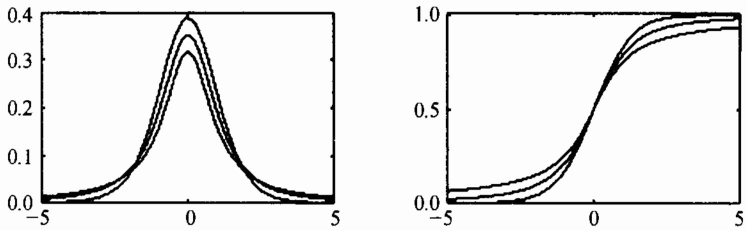

图 5-9　t 分布的密度函数和累积分布函数图

用逆累积分布函数求一定显著概率条件下,服从 t 分布的假设检验临界值如下

```
tinv(0.99,1)          ans = 31.8205
tinv(0.95,1)          ans =  6.3138
tinv(0.99,4)          ans =  3.7469
tinv(0.95,4)          ans =  2.1318
tinv(0.99,10)         ans =  2.7638
tinv(0.95,10)         ans =  1.8125
x=0:0.1:1
x =0      0.1000    0.2000    0.3000  0.4000  0.5000  0.6000
   0.7000 0.8000    0.9000    1.0000
tinv(x,1)
ans = -Inf   -3.0777   -1.3764   -0.7265   -0.3249  0  0.3249
  0.7265  1.3764  3.0777  Inf
tinv(x,4)
ans = -Inf   -1.5332   -0.9410   -0.5686   -0.2707  0  0.2707
  0.5686  0.9410  1.5332  Inf
tinv(x,10)
ans = -Inf   -1.3722   -0.8791   -0.5415   -0.2602  0  0.2602
  0.5415  0.8791  1.3722  Inf
tinv(x,20)
ans = -Inf   -1.3253   -0.8600   -0.5329   -0.2567  0  0.2567
  0.5329  0.8600  1.3253  Inf
```

产生服从 t 分布随机数的例子如下

```
trnd(1)       ans = 0.2597
trnd(4)       ans =  -0.6506
trnd(10)      ans = 0.5607
```

```
trnd(10,5,7)
ans = -0.9257  -0.8351  -0.0974   0.8557  -0.4281  -1.1515
      -1.4456  -1.5288  -1.1224   1.6795  -1.6570   1.1646
      -0.0427  -0.8442  -0.2648  -1.5493   0.8303   0.9546
      -1.0879   0.0940  -0.9870   0.3268   0.1450   0.4123
      -0.3795   0.8932  -0.3527  -1.2183   0.0907   1.7641
       0.4928   1.5947   0.4946   0.5981   0.8435
trnd(4,5,7)
ans = -2.7042  -0.8212  -0.7402   0.3606  -1.2289  -0.4208
       1.0640  -0.2207  -0.0575   0.5193  -0.7456   5.6528
       0.6174  -0.5890  -0.1482  -0.1417   1.0177   7.1297
       0.6014  -1.2989  -1.7904  -1.8013  -1.8423  -0.6154
       0.2114  -0.1626  -1.0069  -0.3111   1.0355  -0.2609
       0.0663  -0.7707   1.0912  -1.4316   0.4298
```

求 t 分布数学期望与方差的例子如下

```
[e d] = tstat(4)      e = 0      d = 2
[e d] = tstat(10)     e = 0      d = 1.2500
[e d] = tstat(20)     e = 0      d = 1.1111
```

3. F 分布

设随机变量 X 与 Y 独立，且 $X\sim\chi^2(m)$，$Y\sim\chi^2(n)$，令

$$F=\frac{X/m}{Y/n},$$

则称 F 服从第一自由度为 m，第二自由度为 n 的 F 分布，记作 $F\sim F(m,n)$. MATLAB 提供的有关 F 分布的函数如下

fpdf(X,M,N)	F 分布的密度函数
fcdf(X,M,N)	F 分布的累积分布函数
finv(P,M,N)	F 分布的逆累积分布函数
frnd(M,N,m,n)	产生服从 F 分布的随机数
fstat(M,N)	求 F 分布的数学期望与方差

其中 X 为随机变量，M 和 N 分别为 F 分布的第一和第二自由度，P 为显著概率，m 和 n 为所产生随机数矩阵的行数和列数. 用下述程序绘制 F 分布的密度函数与累积分布函数图(图 5-10)

```
x = 0:0.001:3;
y = fpdf(x,1,10);   z = fcdf(x,1,10);
y1 = fpdf(x,10,4); z1 = fcdf(x,10,4);
```

```
y2 = fpdf(x,10,20);z2 = fcdf(x,10,20);
y3 = fpdf(x,10,50);z3 = fcdf(x,10,50);
subplot(1,2,1);
plot(x,y,'k',x,y1,'k',x,y2,'k',x,y3,'k');
axis([0,3,0,1]);
subplot(1,2,2);
plot(x,z,'k',x,z1,'k',x,z2,'k',x,z3,'k');
axis([0,3,0,1.01]);
```

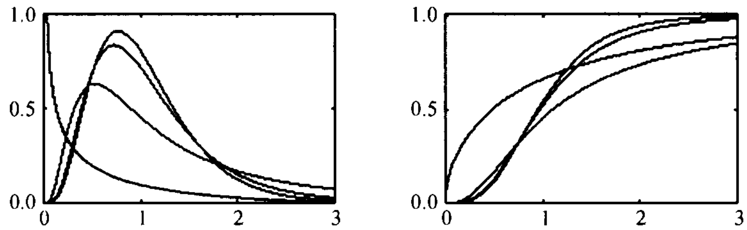

图 5-10 F 分布的密度函数和累积分布函数图

用逆累积分布函数求一定显著概率条件下,服从 F 分布的假设检验临界值如下

```
finv(0.99,10,4)        ans = 14.5459
finv(0.99,10,10)       ans = 4.8491
finv(0.99,10,20)       ans = 3.3682
finv(0.95,10,4)        ans = 5.9644
finv(0.95,10,10)       ans = 2.9782
finv(0.95,10,20)       ans = 2.3479
x = 0:0.1:1
x = 0  0.1000  0.2000  0.3000  0.4000  0.5000  0.6000
  0.7000  0.8000  0.9000  1.0000
finv(x,10,4)
ans = 0  0.3838  0.5469  0.7100  0.8926  1.1126  1.3971
  1.8002  2.4596  3.9199  Inf
finv(x,10,10)
ans = 0  0.4306  0.5775  0.7112  0.8484  1.0000  1.1787
  1.4061  1.7316  2.3226  Inf
finv(x,10,20)
ans = 0  0.4544  0.5944  0.7166  0.8375  0.9663  1.1122
  1.2901  1.5313  1.9367  Inf
```

产生服从 F 分布的随机数的例子如下

```
frnd(10,4)          ans = 2.8855
frnd(10,10)         ans = 0.2757
frnd(10,20)         ans = 1.5200
frnd(2,10,5,7)
ans =1.6613   0.1962   1.8418   0.2085   0.4924   0.2682   0.8976
     0.7198   0.3330   0.7767   1.9925   0.5375   1.0791   0.4058
     0.4973   1.0102   3.3659   0.1749   0.2769   0.9471   2.6272
     5.6393   2.8133   1.5102   3.3269   0.2992   0.0420   0.9972
     0.0888   0.0992   0.1880   0.4350   0.3622   0.2180   0.2901
frnd(10,20,5,7)
ans =0.6071   0.8229   0.8108   0.8462   0.9936   1.5513   1.0835
     1.5606   0.6781   0.9967   0.3176   1.5840   0.2994   1.4440
     1.6310   0.1674   0.7198   1.2456   0.5857   0.2091   0.8690
     0.6606   2.5233   2.1569   0.9866   0.3704   0.9820   0.5452
     0.4959   0.3080   2.8579   2.4121   0.6325   0.4102   1.3503
```

求 F 分布数学期望与方差的例子如下

```
[e d]=fstat(10,4)       e = 2          d = NaN
[e d]=fstat(2,10)       e = 1.2500     d = 2.6042
[e d]=fstat(4,15)       e = 1.1538     d = 1.0288
[e d]=fstat(10,20)      e = 1.1111     d = 0.4321
```

第三节 参数估计与假设检验

一、样本的数字特征

统计推断的基础是通过搜集、整理、加工和分析统计数据，使之系统化、条理化，以显示数据资料的趋势、特征和数量关系.

1. 集中趋势(位置)度量

数据样本集中趋势度量的目的在于对数据样本在数据分布线上分布的中心予以定位，即中心位置的度量. 集中趋势度量包括几何平均值、调和平均值、算术平均值、中位数和修正的样本均值等. 在 MATLAB 中分别用下述函数来求这些度量值，格式如下

geomean(X)　　样本的几何平均值

harmmean(X)　　样本的调和平均值