概率论与随机过程上机实验报告

自动化73 郭王懿 2176112723

同组者：罗中沛，陈子恺

题目一

题目

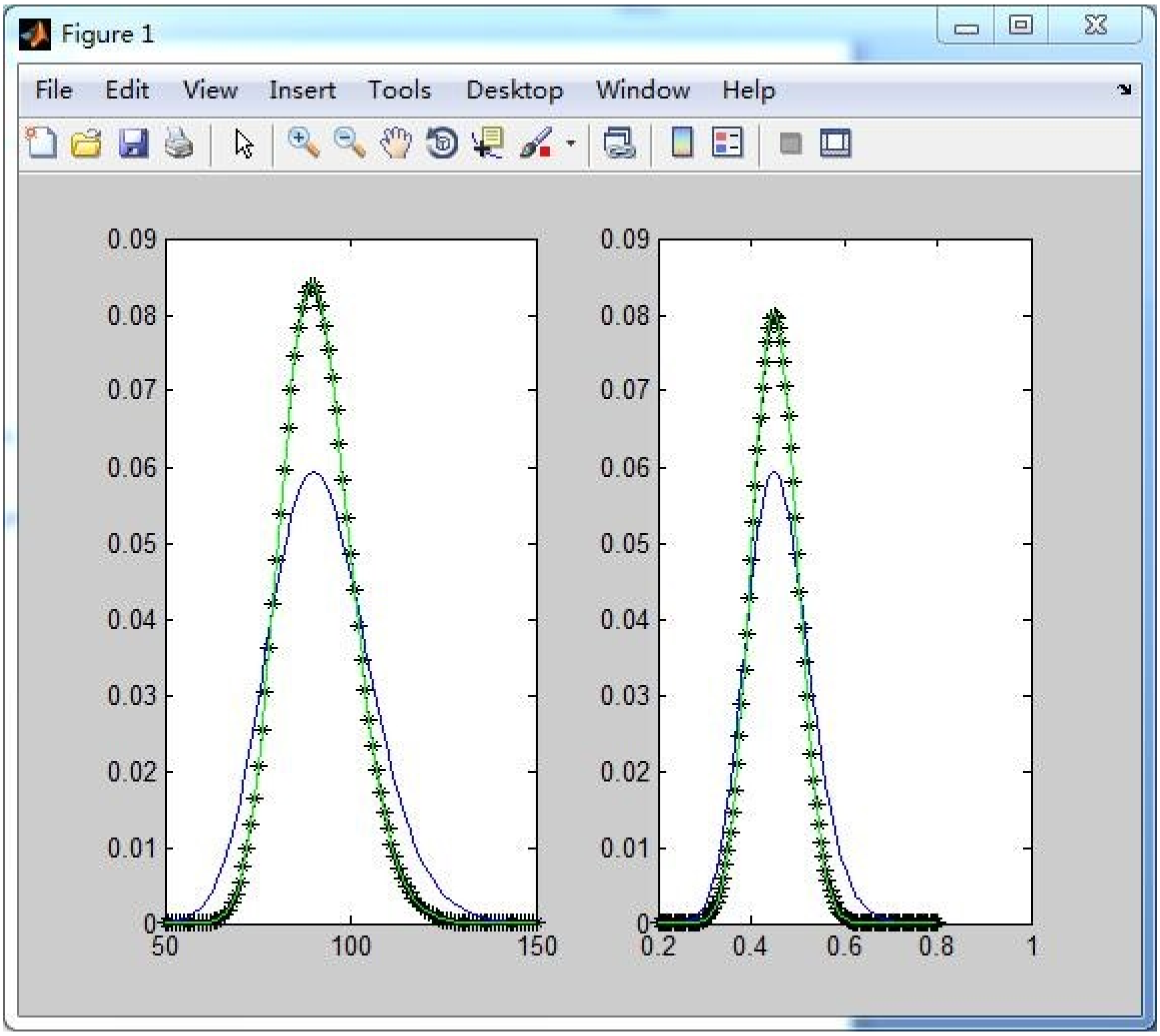
对二项分布事件的概率的精确计算与用泊松分布和中心极限定理的近似计算进行对比。

- P变化n固定，进行比较
- n固定，p变化进行比较。

源代码

```
n=50:150;
p=0.5;
p1=binopdf(45,n,p);
p2=poisspdf(45,n*p);
p3=normpdf(45,n*p,sqrt(n*p*(1-p)));
subplot(1,2,1);
plot(n,p1,'k*-',n,p2,'b-',n,p3,'g-')
n1=100;
pn=0.2:0.005:0.8;
p4=binopdf(45,n1,pn);
p5=poisspdf(45,n1.*pn);
p6=normpdf(45,n1.*pn,sqrt(n1.*pn.*(1-pn)));
subplot(1,2,2);
plot(pn,p4,'k*-',pn,p5,'b-',pn,p6,'g-')
```

运行结果



黑星代表二项分布，

蓝色是泊松分布

绿线是中心极限定理

小结

n变化从50开始到150，中心极限定理的计算方法更加接近二项分布的精确计算，泊松分布于精确计算差距稍微增大但保持原有的变化趋势。p改变时，p=0.5时取最大值，仍然是中心极限定理比泊松分布更加接近二项分布精确计算。

第二题

题目

对正态总体参数的区间估计，进行验证及区间长度的变化情况（注：对一个参数，验证一种情形即可）。

(a)样本容量固定，置信度变化；

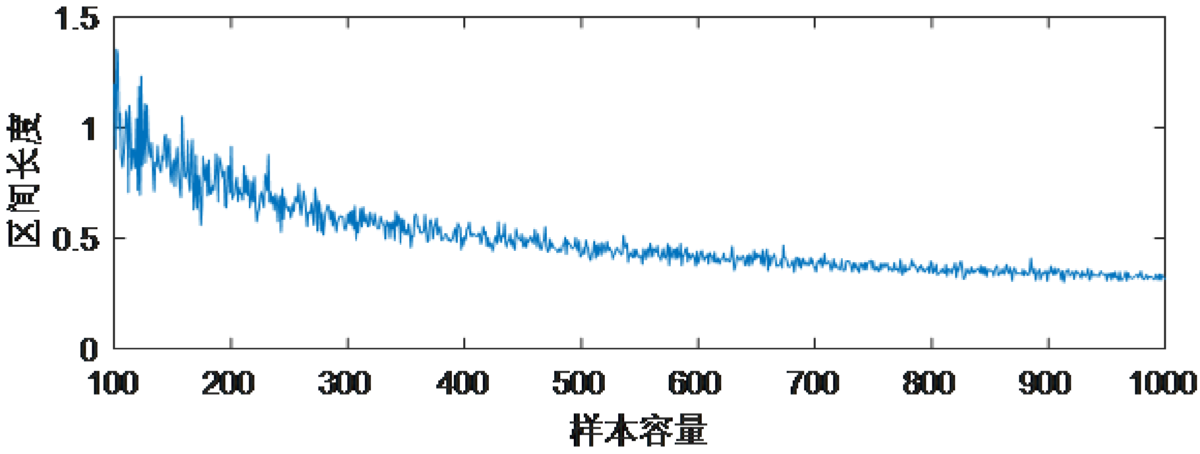
(b) 置信度固定，样本容量变化。

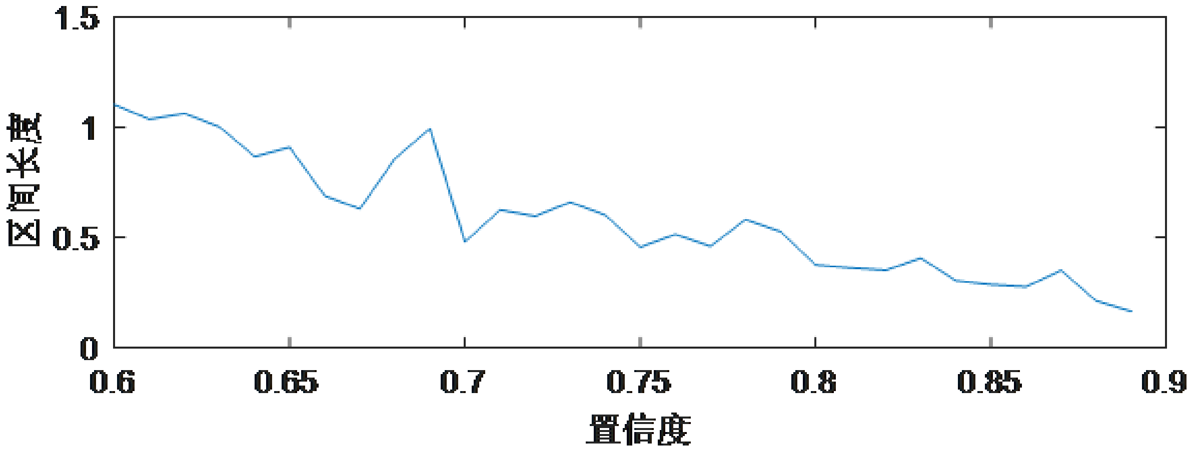
源程序

```
N=100;%样本容量
conf=0.9;%置信度
plot_N=zeros(1,900);
plot_conf=zeros(1,30);
for N=100:1000
ori=normrnd(10,2,N,1)';
    S=var(ori);
plot_N(N-100+1)=2*tinv(0.9,N)*S/sqrt(N);
end
subplot(2,1,1)
plot(100:1000,plot_N)
xlabel('样本容量')
ylabel('区间长度')
%axis([85,115,0,1.5])
N=100;%样本容量
conf=0.9;%置信度
for X=1:30
    conf=0.60+0.01*X;
ori=normrnd(10,2,N,1)';
    S=var(ori);
plot_conf(X)=2*tinv(conf,N)*S/sqrt(N);

end
subplot(2,1,2)
plot(0.89:-0.01:0.60,plot_conf)
xlabel('置信度')
ylabel('区间长度')
```

运行结果





小结

可以看出来，当样本容量不断增加时，区间估计的精度越来越高；同时，当置信度不断提高时，区间估计的精度也越来越高。

第三题

题目

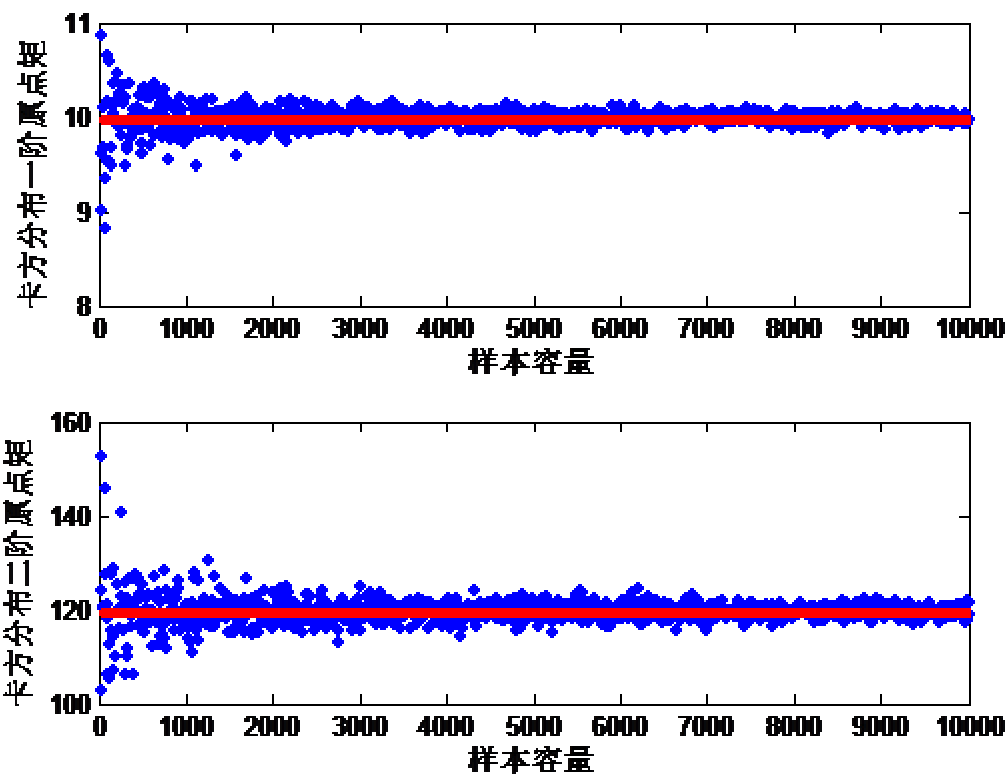
自己选一个总体，验证样本k阶矩的观察值随样本容量的增大与总体k阶矩接近程度（对k=1，2进行验证）

源代码

```
clear
%卡方分布，自由度10
sample_size = 1000;
x = 1:10:sample_size*10;
res_1 = zeros(1,sample_size);
for i=1:sample_size;
```

```
    sample = chi2rnd(10,1,i*10);
    res_1(i) = sum(sample)/(i*10);%样本一阶矩
end
res_2 = zeros(1,sample_size);
for i=1:sample_size;
    sample = chi2rnd(10,1,i*10);
    res_2(i) = sum(sample.*sample)/(i*10);%样本二阶矩
end
one = 10;%总体一阶原点矩
two = 10*10+2*10;%总体二阶原点矩
subplot(2,1,1);
plot(x,res_1,'.')
ylabel('卡方分布一阶原点矩');
xlabel('样本容量')
hold on
plot(x,one,'-r','LineWidth',3);
subplot(2,1,2)
plot(x,res_2,'.')
ylabel('卡方分布二阶原点矩');
xlabel('样本容量')
hold on
plot(x,two,'-r','LineWidth',3);
```

运行结果



小结

使用自由度为10的卡方分布作为研究总体，取样本容量大小从1到10000。图像表明，，随着样本容量的增加，样本观测值的一阶原点矩和二阶原点矩都越来越接近于总体的一二阶原点矩，即10和120。

第五题

题目

自己设计一种情形，当样本至少为多少时，产品的合格率才能符合给定的合格率

源程序

```
clc;
clear;
a = 0.02;%置信度为0.02
u = 50;%假设正态总体的期望
s = 4;%假设正态总体的方差
sample_sum = 0;



mean_j=0;
for mean=1:100
for i = 1:10000
    x(1,i) = normrnd(50,4);
end
for j = 1:10000
    for k = 1:j
        sample_sum = x(1,k) + sample_sum;%计算样本的和
    end
    sample_average = (1/j)*sample_sum;%计算样本均值
    sample_sum = 0;
    U = ((j^(1/2))*(sample_average - u))/(s^(1/2));%构造统计量
    if (abs(U) < normpdf(a/2,0,1))
        mean_j=mean_j+j;%若在接受域，跳出循环
        j;
        break
    end
end
end
mean_j=mean_j/100
```

运行结果

```
mean_j =

  168.2800
```

小结

可以看出来，在给定的强条件下，需要样本容量平均在168左右才可以满足合格率。但具有一定的随机性。

第六题

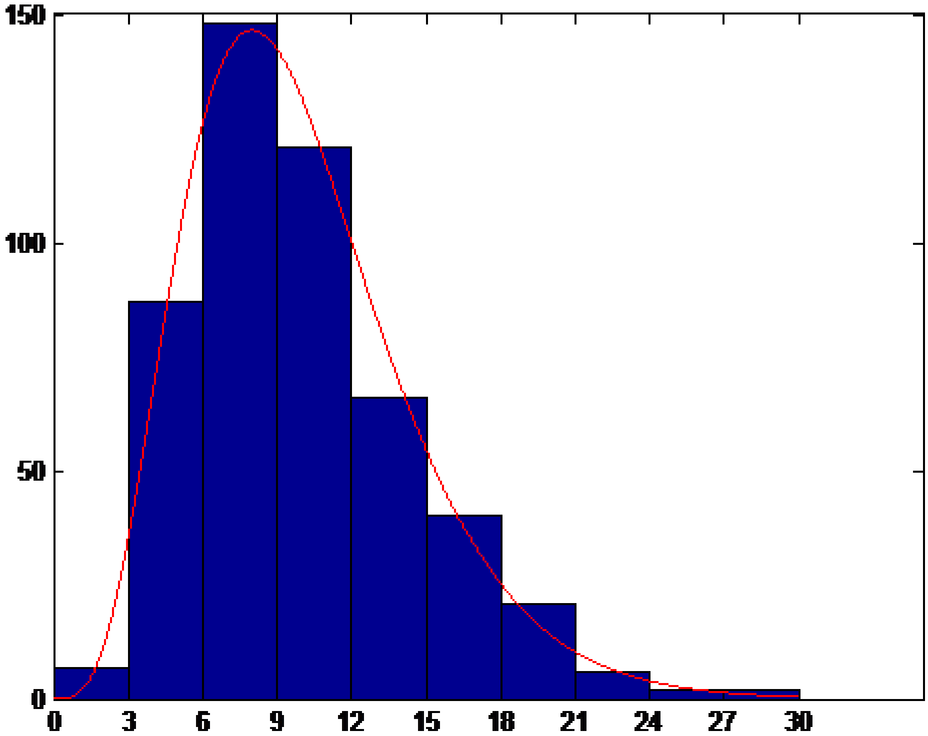
题目

产生卡方分布的随机数，并子一个坐标图中画出该卡方分布的统计直方图与真实卡方分布图形。

源代码

```
clear
rand_num = chi2rnd(10,1,500);
x = 0:0.01:30;
y = chi2pdf(x,10);
index = 0:3:30;
[bincounts]=histc(rand_num,index);
figure
bar(index,bincounts,'histc')
hold on
plot(x,y*1500,'-r');
```

运行结果



小结

观察可知，卡方分布产生的500个随机数的统计直方图的形状与真实卡方分布曲线形状基本拟合。

个人感想

之前大一在进行数学建模的时候通常要用到数理统计的相关知识，但由于没有系统的学习过，始终是一知半解。经过一学期对概率论与随机过程的学习，掌握了很多统计学上的观点以及方法，这对之后的工作或是科研都有着很大的作用。经过这次的上机实验，也能让我们从编程的角度更深入的理解一些方法在实践中的用法，受益匪浅。最后，感谢老师一学期的辛勤教学，也希望老师之后身体健康工作顺利。