概率论与数理统计实验报告

宋德培[1]*

摘要

对由泊松分布与中心极限定理估计二项分布做出了验证；对正态总体的均值的区间估计做了验证并验证了区间长度的变化；从服从泊松分布的样本出发验证了样本一、二阶矩与总体的接近程度；设计了一种实际情形并进行了求解；画出了 χ^2 分布的直方图。

关键词

泊松分布　中心极限定理　大数定律　χ^2 分布

[1] 能动 *A71*　　* **学号**: 2174110112

1. 问题分析与求解

1.1 问题一

题目　对二项分布事件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，分

(a) p 变化，n 固定

(b) n 固定，p 变化。

Listing 1. 源代码

```
clear
clc
p=input('p=');
n=input('n=');
y=[];
y1=[];
y2=[];
for k=0:n
y=[y,binocdf(k,n,p)];
y1=[y1,poisscdf(k,n*p)];
y2=[y2,normcdf((k-n*p)/sqrt(n*p*(1-p)),0,1)];
end
plot([0:n],y)
hold on
plot([0:n],y1,'.-')
plot([0:n],y2,'*-')
legend('二项分布函数','泊松分布函数','\chi^2
    分布函数');
title( { ['\it p = ', num2str(p), ',\it n =
    ',num2str(n)] } )
```

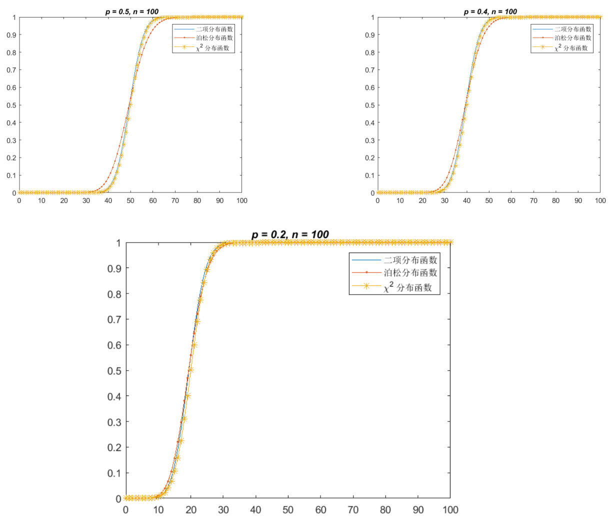

图 1. n 固定，p 变化

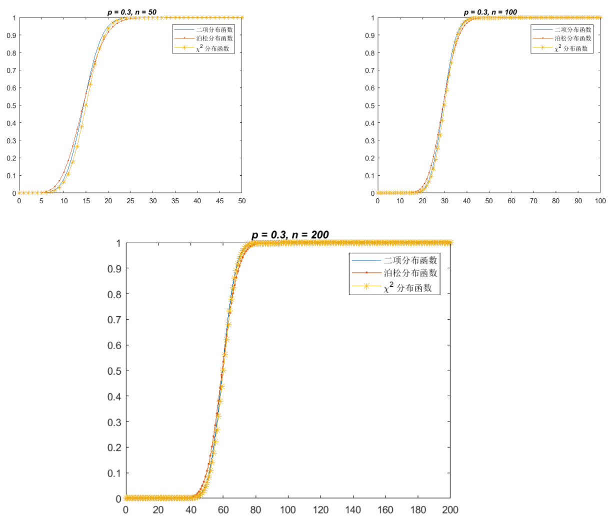

图 2. p 固定，n 变化

小结 通过实验发现当 n 固定，p 越小，泊松分布近似计算的精度越高，当 p 固定化，n 越大，泊松分布与大数定律的近似计算精度都提升。正好与理论相符。

1.2 问题二

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，分

(a) 样本容量固定，置信度变化

(b) 置信度固定，样本容量变化

设置一种情形：正态分布总体 σ 已知为 2，总体均值 μ 为 2，对均值进行区间估计，当输入样本容量及置信度，自动输出置信区间及区间长度。

Listing 2. 源代码

```
clear
clc
n=200;
alpha=0.05;
mu=2; sigma=2;
u=norminv(1-alpha/2);
y=normrnd(mu,sigma,1,n);
fprintf('n=%d, alpha=%3f\n',n,alpha)
fprintf('domain is :[%f ,
    %f]\n',mean(y)-sigma/sqrt(n)*u,mean(y)+
sigma/sqrt(n)*u)
length=2*sigma/sqrt(n)*u;
fprintf('length=%f\n',length)
```

使用双侧区间估计的结果：

$$[\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}]$$

运行结果截图：

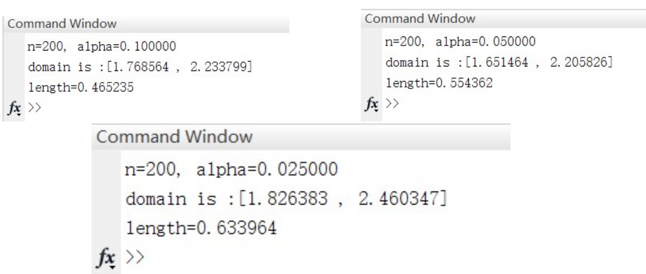

图 3. 样本容量固定，置信度变化

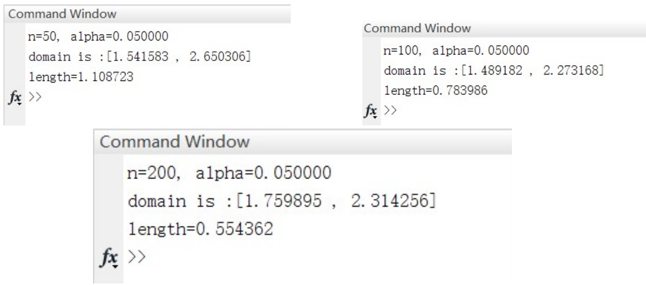

图 4. 置信度固定，样本容量变化

小结 通过实验发现区间估计很好地包括了样本均值，当样本容量固定，置信度越大，区间长度越大；当置信度固定，样本容量越大，区间长度越小；因此为了得到较高的置信度并且较短的区间长度，必须提高样本容量。

1.3 问题三

题目 对于服从泊松分布的总体，验证样本的 k 阶矩的观察值随样本容量的增大与总体 k 阶矩的接近程度。

设置一种情形：服从参数为 $\lambda=5$ 泊松分布总体的均值与方差均为 5，观察一阶原点矩与二阶中心矩随样本容量的变化。

Listing 3. 源代码

```
clear
clc
lambda=5;
r1=[];r2=[];
for n=10:10:200
x=poissrnd(lambda,1,n);
a_1=mean(x);
A_1=lambda;
b_2=var(x);
B_2=lambda;
r1=[r1,abs(a_1-A_1)];
r2=[r2,abs(b_2-B_2)];
end
figure(1)
plot(r1,'*')
xlabel('n');ylabel('|a-A|');legend('一阶原点矩')
figure(2)
plot(r2,'*')
xlabel('n');ylabel('|b-B|');legend('二阶中心矩');
```

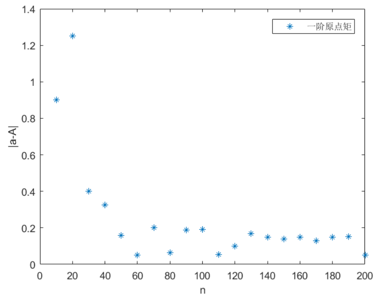

图 5. 一阶原点矩变化

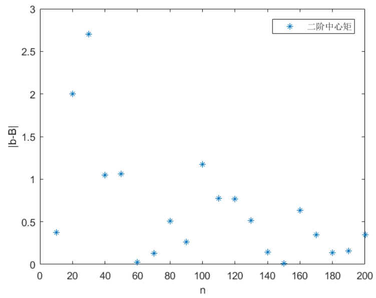

图 6. 二阶中心矩变化

> **小结** 从散点图看出总体矩与样本矩的差总体来看在随样本容量增大而减小。故通过实验发现随样本容量增大，样本一二阶矩都更加接近总体之矩。

1.4 问题四

题目 已知一批产品有 1000 个，次品率为 0.3，现要使抽到的样本次品率小于 15%，则最小样本数为？

理论解：

$$\frac{m-700}{\sqrt{1000\times 0.3\times 0.7}} \geqslant 1.6449$$
$$m \geqslant 723.8$$

即 m 最小为 724.

```
clear
clc
a=ones(2000,1);
a(randsample(1:2000,600))=0;
for i=1:1000
if sum(a(1:i))>=0.95*i;
break
i
end
end
```

经模拟计算，在理论值以下得到合格率 95% 的概率极小，事实上，几次实验中未发现此情形。

> **小结** 通过实验，由中心极限定理得到所需的最小样本数。最小样本的选择在实际问题中很重要也很有价值的，样本数过小通常不具有代表性，过大又增加了工作量，因此研究最小样本问题有很高的实用价值。

1.5 问题五

题目 画出服从 χ^2 分布的统计直方图。

Listing 4. 源代码

```
clear
clc
Degrees_of_Freedom=input('n=');
n=100;
x=chi2pdf([1:n],Degrees_of_Freedom);
figure(1)
plot(x)

y=chi2rnd(Degrees_of_Freedom,1,n);
figure(2)
hist(y)
```

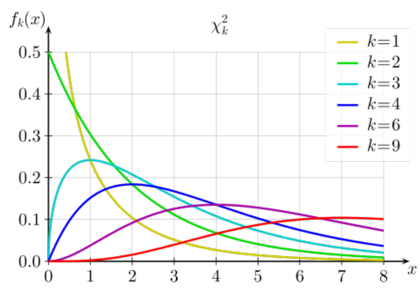

图 7. 理论曲线

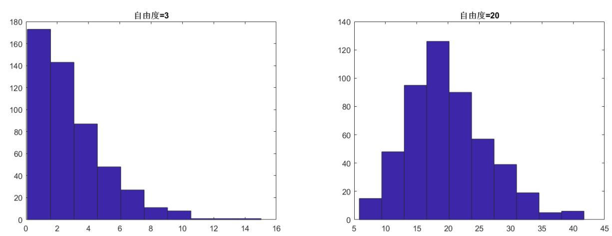

图 8. 直方图

小结　χ^2 分布是 Γ 分布的一个特例，是概率统计领域常用的重要分布函数，同时他与正态分布的关系使得它的一些很好的性质能够在实际问题中被利用。

2. 感想

这门课程给我最深刻的体会就是这门课程很抽象，很难以理解，初学时，就算觉得理解了老师的讲课内容，但是一联系实际也会很难以应用上，简化不出有关所学知识的模型。后来经过老师的生动现实的实例分析，逐渐对这门课程有了新的认识。首先，这门课程给我带来了一种新的思维方式。前几章的知识好多都是高中大学讲过的，接触下来觉得挺简单，但是后面从大数定理及中心极限定理就开始是新的内容了。我觉得学习概率论与数理统计最重要的就是要学习书本中渗透的一种全新的思维方式。统计与概率的思维方式，和逻辑推理不一样，它是不确定的，也就是随机的思想。这也是一个人思维能力最主要的体现，整个学习过程中要紧紧围绕这个思维方式进行。这些都为后面的数理统计还有参数估计、检验假设打下了基础。

通过对概率论的发展史的了解，我对概率论课程中学习的一些知识有了更深层次的理解。此外，概率论在各个学科和金融、保险、生物、医学、经济、运筹管理和工程技术等领域也得到了广泛应用。主要包括：极限理论、随机过程论、数理统计学、概率论方法应用、应用统计学等。概率论方法应用是一个涉及面十分广泛的领域，包括随机力学、统计物理学、保险学、随机网络、排队论、可靠性理论、随机信号处理等有关方面。

比如罗纳德・艾尔・费舍尔爵士是英国统计学家和遗传学家。由于他在统计学方面的工作, 他被描述为” 一个几乎单枪匹马地为现代统计奠定了基础的天才” 和”20 世纪统计中最重要的人物”。在遗传学, 他的工作使用数学结合孟德尔遗传学和自然选择; 这促成了达尔文主义在 20 世纪初被称为现代综合的进化论的修订中的复兴。费舍尔因其对生物学的贡献而被称为” 达尔文最伟大的接班人”。

从 1919 年起, 他在罗斯迈德实验站工作了 14 年; 那里, 他分析了它的巨大的数据从作物实验从 19 世纪 40 年代, 并且开发了方差分析 (anova)。在随后的几年里, 他在那里建立了生物统计学家的声誉。他被称为世界的三位主要创始人之一。

通过伟人们的努力，我们得以看清世界的运行方式：即使量子力学中也认为存在概率波，换句话说，我们的世界归根结底是由概率统治的！概率论提供给我们一种全新的世界观，同时也是更加客观的世界观。

其次，概率是一个分析工具，能够分析利弊，分析一件事情值不值得做，做成功的概率是多少？让你规避做一件事的风险。概率是帮助你选择的工具，通过概率分析，选择最好的决策，让你做出更好的选择。生活中发生的事件是受多因素影响的，根据概率论来说就是一维随机变量，二维随机变量，多维随机变量。但事实上，概率论暂时只讨论到二维，即使你讨论到三维，四维，五维。甚至 N 维，但是你也始终推导不出一件事情发生受到了哪些随机变量的影响，我们只是猜测。但也没事，因为有一个人 [贝叶斯] 发现了一个定理——贝叶斯定理，专门用来事后检验，就是一件事情已经发生了，检验各种因素对此事件发生的概率。这个人非常伟大，因为这个定理对分析事件发生所受到的影响因素的概率大小非常有帮助，对于以后人们预防这类不利的影响因素起到了很大的作用。虽然某些人会认为这个定理有“事后诸葛亮”的意味，其实他们并没有发现这个定理的精髓。总结来说，概率论对于帮助我们过更好的生活有很大的作用，运用贝叶斯定理能够分析出哪些是小概率因素，针对小概率因素采取相应的措施。概率论的神器之处还有很多，需要读者在生活中去发现，有一句话说得好，处处留心皆学问。因此，在生活中，既要谨防小概率事件的有害方面，又要把握小概率事件的有利方面。这样，我们的生活才会更加精彩！