大数定律

教学设计与制作：胥斌

大数

引例 灯泡寿命估计

现有一批来自同一生产线的同型号灯泡，每只灯泡的寿命服从相同的分布，问任取一只灯泡，其寿命X不低于1100小时的概率？

分析 寿命分布未知，$P\{X \geq 1100\}$无法精确求得；

直观想法 测若干只灯泡的寿命，统计寿命不低于1100小时的灯泡数量，求得频率f，用于近似概率$P\{X \geq 1100\}$。

录屏：为验证引例中直观想法，利用Matlab进行模拟实验。

MATLAB 7.6.0 (R2008a)

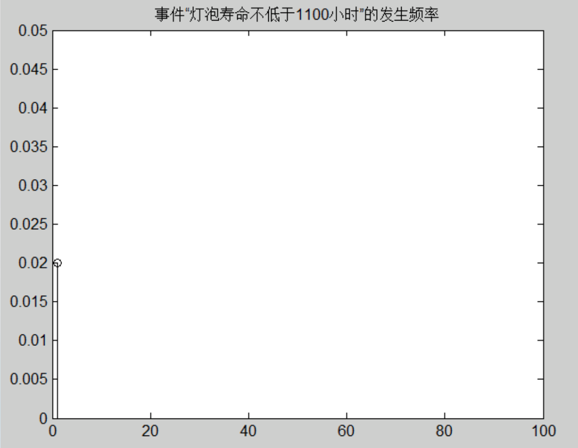
File Edit Debug Parallel Desktop Window Help

D:\Program Files\MATLAB\R2008a\work\Xdataplot.m

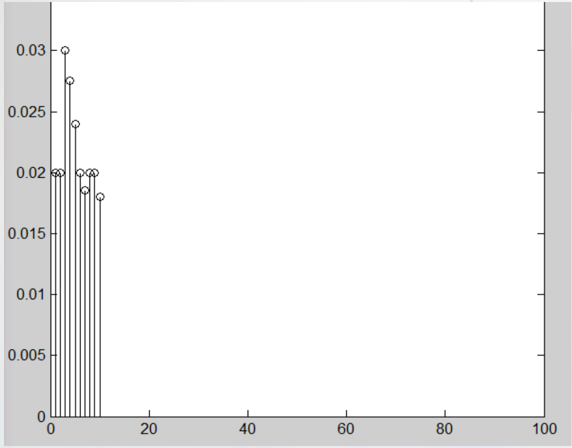
File Edit Text Go Cell Tools Debug Desktop Window Help

```
clear
%载入灯泡寿命数据，共10000只灯泡
load bulbdata
number=0;%number表示事件“寿命超过1100小时”的发生频数
%灯泡分100组，每组100只
for i=1:100
    index=i*100;
for j=1:index
    if data(j)>=1100
        number=number+1;
    end
end
  p(i)=number/index;%p表示事件的发生频率
  number=0;
end
x=1:length(p);
figure;
stem(x(1),p(1),'k');%以杆图形式绘制曲线，前100只的结果
title('事件“灯泡寿命不低于1100小时”的发生频率');
axis([0 100 0 0.05]);
hold on
stem(x(2:10),p(2:10),'k');%前200--1000只的结果
```

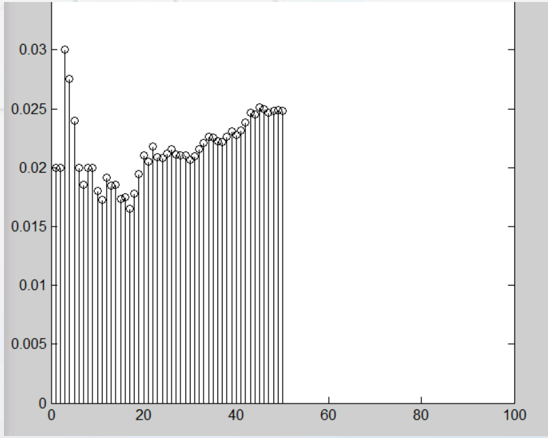
Demos, or read Getting Started.



试验次数100

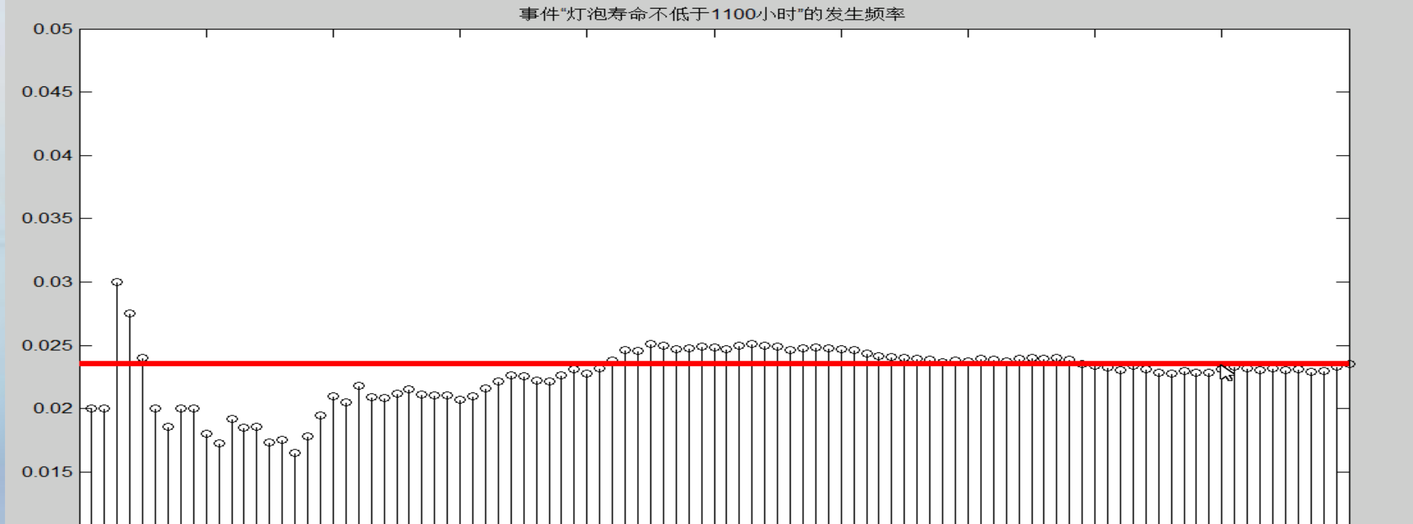


试验次数渐增到1000



试验次数渐增到5000

试验次数
渐增到10000



两个疑问

用“多次试验中事件发生的频率”去估计“一次试验事件发生的概率”**合理吗**?

① **频率是否是稳定的?**

② **频率是否是稳定到概率?**

大数定律对这两个问题作出了肯定的回答，是频率稳定性的理论保证，大数定律是概率论中最著名的成果。

大数定律

马尔科夫不等式

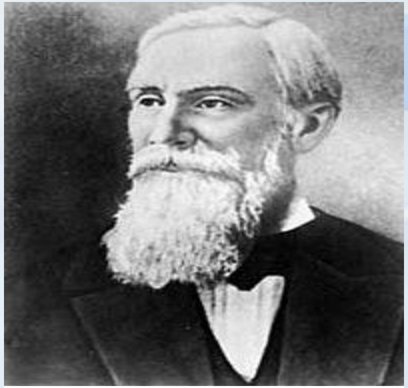
↓

切比雪夫不等式

↓

强大数定律

安德烈·马尔科夫



切比雪夫

马尔科夫不等式

如果 X 是只取非负值的随机变量，则对于任意 $a>0$

$$P\{X \ge a\} \le \frac{EX}{a}$$

简证 在 X 是具有密度 f 的连续情形下给出证明：

$$EX = \int_0^{\infty} xf(x)dx = \int_0^{a} xf(x)dx + \int_a^{\infty} xf(x)dx$$

$$\ge \int_a^{\infty} xf(x)dx \ge \int_a^{\infty} af(x)dx$$

$$= a\int_a^{\infty} f(x)dx = aP\{X \ge a\}$$

证毕

切比雪夫不等式

如果 X 是具有期望 μ 和方差 σ^2 的随机变量，则对于任意 $\varepsilon > 0$

$$P\{|X-\mu| \ge \varepsilon\} \le \frac{\sigma^2}{\varepsilon^2}$$

简证 因为 $(X-\mu)^2$ 是非负随机变量，由马尔科夫不等式得

$$P\{(X-\mu)^2 \ge \varepsilon^2\} \le \frac{E(X-\mu)^2}{\varepsilon^2}$$

等价于

$$P\{|X-\mu| \ge \varepsilon\} \le \frac{E(X-\mu)^2}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

证毕

强大数定律

设$X_1,\cdots X_n$是一系列独立同分布的随机变量，且 $EX=\mu$

则,以概率为1地,

$$\frac{X_1+X_2+\cdots+X_n}{n}\xrightarrow{n\to+\infty}\mu$$

简证 令 $\dfrac{X_1+X_2+\cdots+X_n}{n}=\bar{X}$ 则 $E\bar{X}=\mu, D\bar{X}=\dfrac{\sigma^2}{n}$

由切比雪夫不等式，得

$$P\{|\bar{X}-\mu|\geq\varepsilon\}\leq\frac{D\bar{X}}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}\xrightarrow{n\to+\infty}0$$ **证毕**

强大数定律的结论

设A是随机试验E的一个固定事件，$P(A)$是A发生的概率。独立地将试验E进行n次，

令 $$X_i=\begin{cases}1，若在第i次试验A发生\\0，若在第i次试验A未发生\end{cases}$$

由强大数定律，以概率为1地有

$$\frac{X_1+X_2+\cdots+X_n}{n}\longrightarrow EX=P(A)$$

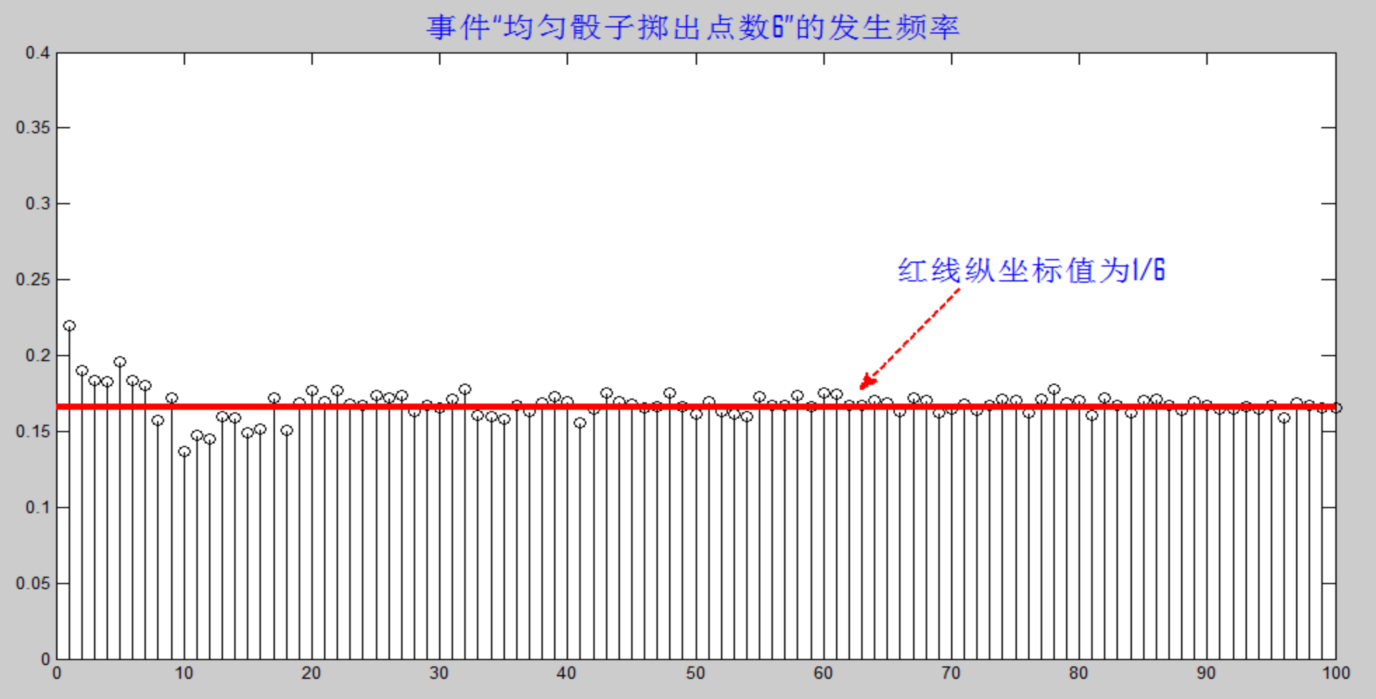
A发生的频率f　　　　A发生的概率P

回到引例

引例中，通过试验模拟，得到灯泡寿命不低于1100小时的频率逐渐稳定在0.0235 左右，由强大数定律，结论“任取一只灯泡，其寿命不低于1100小时的概率$P\{X\geq 1100\}=0.0235$”是合理的，随着试验次数的增大，这个频率趋于真实概率的可能性趋于1。

大数定律的实例验证

掷一枚均匀骰子，“点数6出现”的概率应该为1/6，通过Matlab试验模拟，观测多次掷骰子“点数6出现”的发生频率，体会大数定律。



鲜花和掌声献给切比雪夫、马尔科夫、辛钦等大咖们！