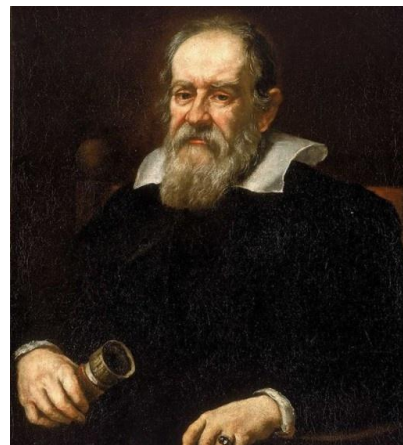


# 概率论与数理统计

Probability and Statistics for Computer Scientists

# 概率论的起源

- 追溯至文艺复兴



伽利略,1632年第一个提出观测误差(随机误差)

- ① 所有观测值都有误差
- ② 观测误差对称分布在0两侧
- ③ 小误差出现的比大误差更频繁



# 概率论的起源

经典问题&概率思想

分析方法:严格证明&标准术语

逻辑基础&公理化概率论的应用



## 组合概率论

- 赌博问题
- Huygens,<论赌博中的计算>
- Jacob Bernoulli,<猜度术>
- A. deMoivre,正态分布和正态分布率
- C. de Buffon,几何概率
- 1809 年, 高斯从误差函数角度, 发现正态分布

## 分析概率论

- Laplace,<分析概率论>
- 俄,彼得堡数学学派:
  - 切比雪夫(大数定律,极限定理)
  - 马尔可夫
  - V.Ya. Buniako vsky的<数学概率理论>

## 测度概率论

- E. Borel, 测度论
- 集合论
- 马尔可夫过程
- 莱维,<随机过程与布朗运动>,独立增量过程
- J. Ville,鞅论
- 日本,伊藤清,随机过程研究

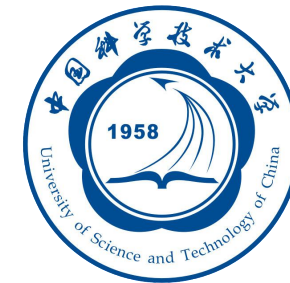
参考来自<简明概率论学术史纲要>



# 概率论的发展

每个公平博弈的参与人愿意拿出经过计算的公平赌注冒险而不愿拿出更多的数量。即赌徒愿意押的赌注不大于其获得赌金的数学期望数。

-----Huygens 《论赌博中的计算》



# 概率论的起源

“后验的去探知我们所无法先验的决定的东西, 从大量同类事例的观察结果中去探知它.”

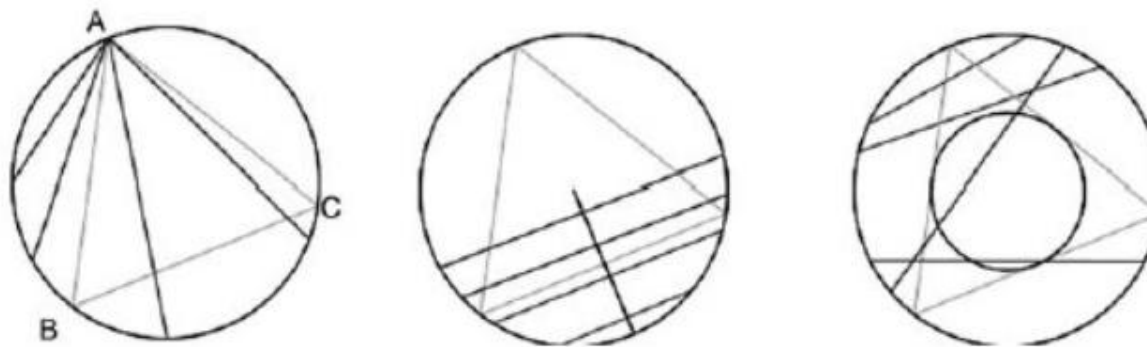
“频率的不稳定性随观察次数的增加而减少.”

-----Bernoulli 《猜度术》

# 概率论的发展

- 悖论问题---贝特朗悖论

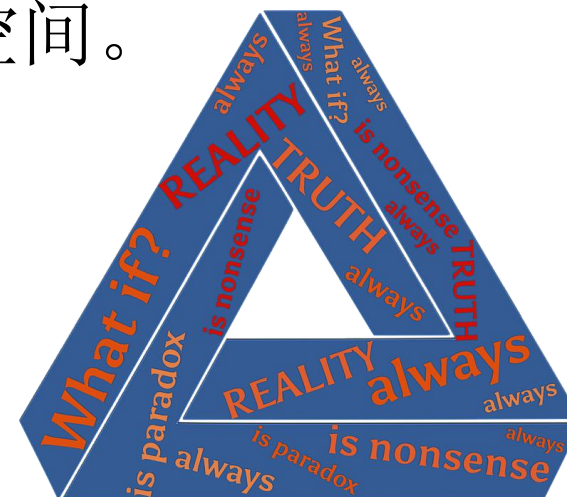
在半径为1的圆内随机地取一条弦，问其长超过该圆内接等边三角形的边长的概率为多少？

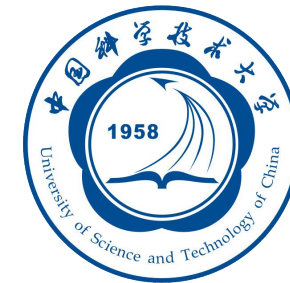




# 概率论的发展

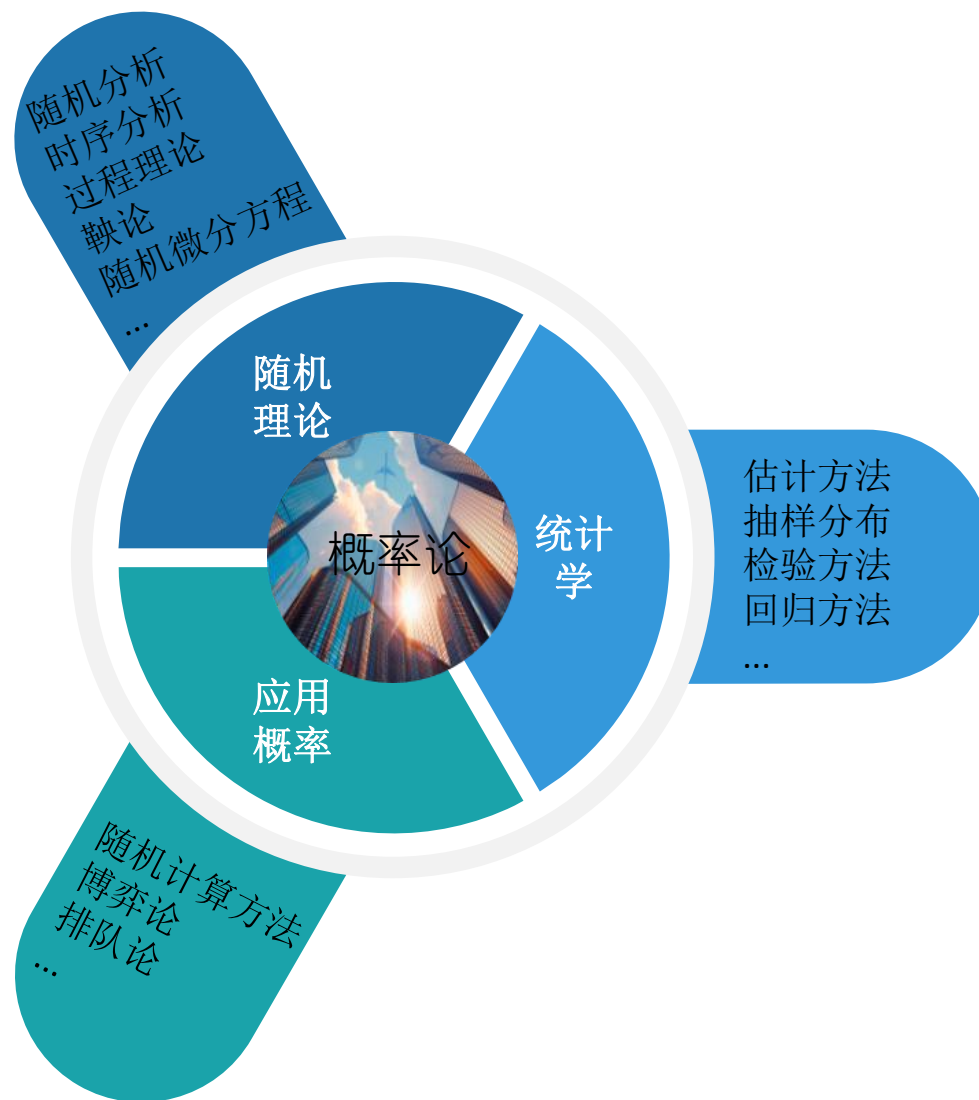
这类悖论说明概率的概念是以某种确定的试验为前提的，这种试验有时由问题本身所明确规定,有时则不然。因此定义概率时要事先明确指出样本空间。



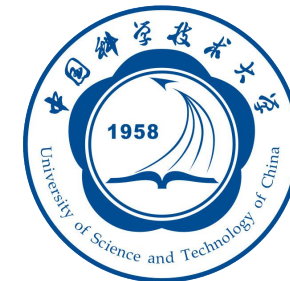


# 概率论的发展

- 当今概率论知识体系





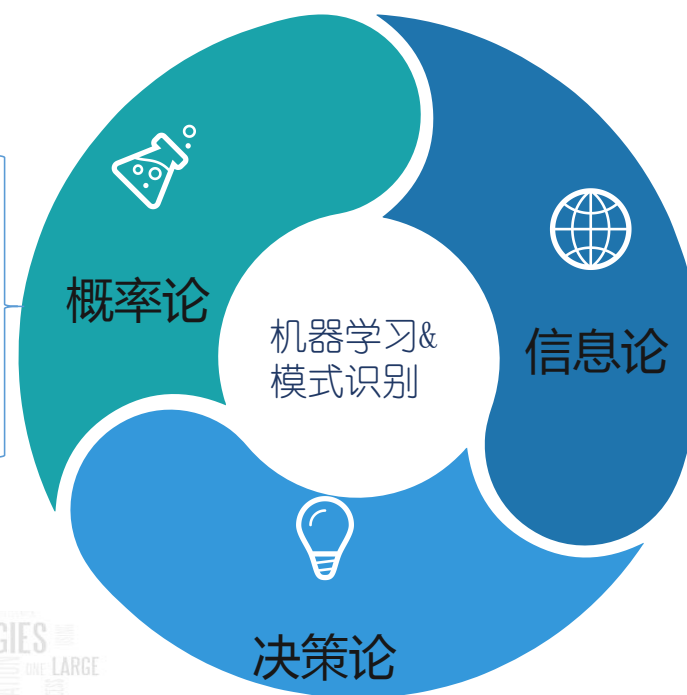


# 概率论的发展

- 概率论的应用

---不确定性 (uncertainty)

二元分类 → 伯努利分布  
多元分类 → 多项分布  
回归 → 高斯分布  
先验概率\后验概率  
贝叶斯定理  
损失函数  
.....

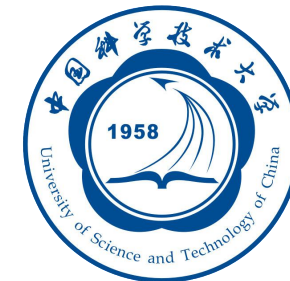


熵  
相对熵和KL散度  
convex function和jensen不等式  
.....

分类问题的决策依据?  
回归问题的决策依据?  
.....



参考来自<https://www.jianshu.com/p/f2b4b06d189a>

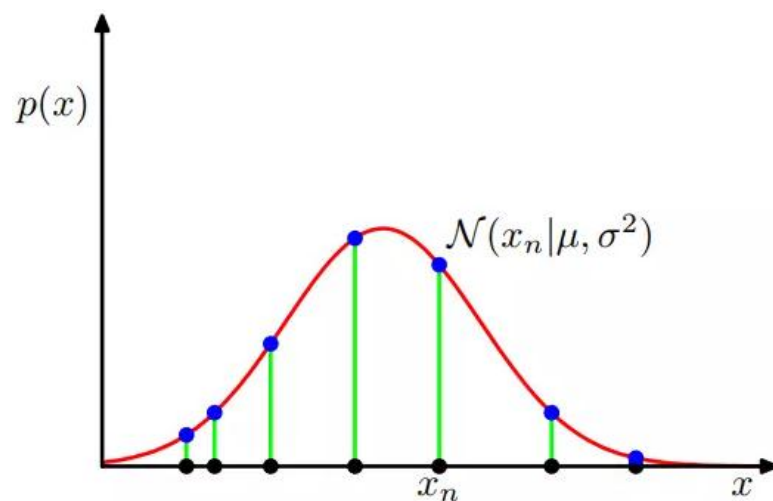


# 概率论的发展

---高斯分布在回归中的应用

$X$  的数据集:  $X = (x_1, x_2, x_3, \dots, x_N)^T$ , 标量  $X$  的  $N$  个观测值。

使用观测到的数据集确定  $\mu = ? \sigma^2 = ?$



数据点独立恒等分布,i.i.d:

$$p(x | \mu, \sigma^2) = \prod_{n=1}^N (x_n | \mu, \sigma^2) \quad (\text{高斯的似然方程})$$

$$\ln p(X | \mu, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{n=1}^N (x_n - \mu)^2 - \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln(2\pi)$$

使上式最大化, 偏分为0:

$$\mu_{ML} = \frac{1}{N} \sum_{n=1}^N x_n \quad \sigma_{ML}^2 = \frac{1}{N} \sum_{n=1}^N (x_n - \mu)^2$$

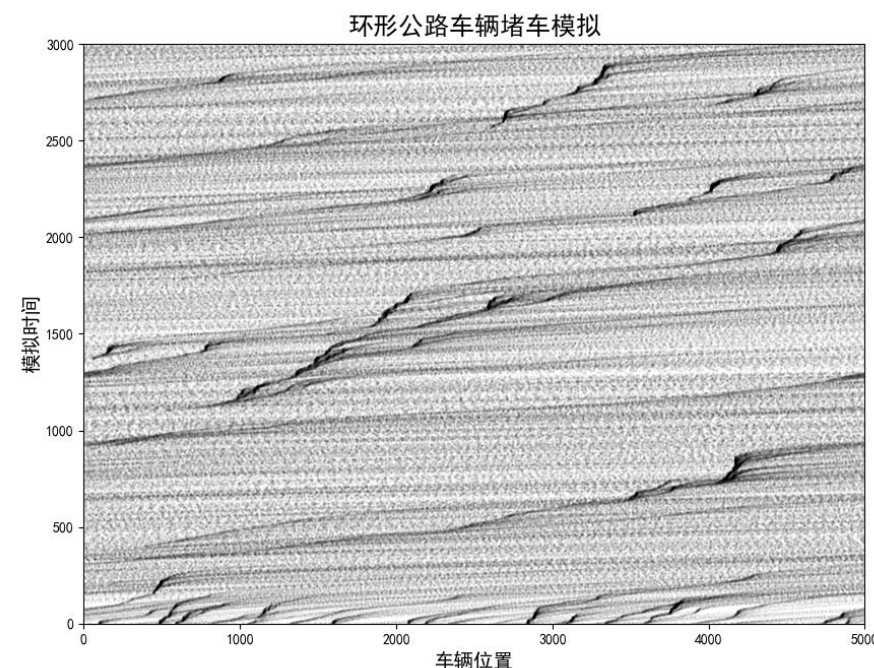


# 概率论的发展

## ---基于概率的模型

### Nagel-Schreckenberg交通流模型

- ✓ 假设某辆车的当前速度是 $v$ 。
- ✓ 若前方可见范围内没车，则它在下一秒的车速 提高到 $v+1$ ，直到达到规定的最高限速。
- ✓ 若前方有车，前车的距离为 $d$ ，且 $d < v$ ，则它下一秒的车速降低到 $d-1$ 。
- ✓ 每辆车以概率 $p$ 随机减速 $v-1$ 。
- ✓ 该公路是环形



堵车概率模型 ( $P=0.3$ )

参考来自[https://blog.csdn.net/weixin\\_44510615/article/details/88921246](https://blog.csdn.net/weixin_44510615/article/details/88921246)

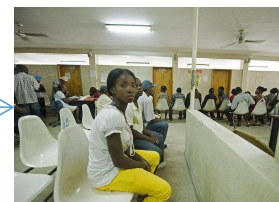
# 概率论的发展

- 数据收集（gathering of data）：  
collected, summarized, reported, and stored for perusal
- 推理统计（inferential statistics）：  
“toolbox”



85%

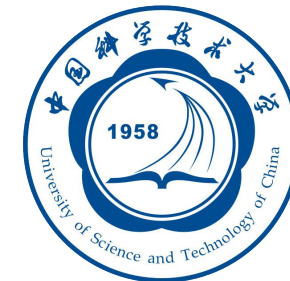
natural variation from study to study



75%

**Decision:** more expensive ,side effects... Should the new drug be adopted?

# 关于课程



- 参考:

- ① 《概率导论(第二版)》, Dimitri P. Bertsekas, John N. Tsitsiklis 著, ISBN: 9787115215444
- ② 《概率论与数理统计(第4版)》, 盛骤, 谢式千, 潘承毅 著, ISBN: 9787040238969

- 成绩: 课上点名+作业+期末考试