# 概率论与数理统计

主讲教师：陈熙霖(阶一5)、王瑞平(教305)
助教：阚美娜、曾加贝(阶一5) 、杨双(教305)

# 主要教材与参考书

- 盛聚、谢式千、潘承毅，概率论与数理统计，高级出版社，2014
- 威廉-费勒，概率论及其应用，人民邮电出版社，2014
- L.沃塞曼 ，统计学完全教程，科学出版社，2008

- 延伸阅读：
  – 拉普拉斯，龚光鲁，钱敏平 译，关于概率的哲学随笔，高等教育出版社， ISBN：9787040378207

# 绪论

- 概率是相对的，部分出于某种无知，部分出于我们的知识所限

  ——拉普拉斯
- 概率论是研究客观世界随机现象数量规律的学科
- 起源：
  - 16世纪意大利学者开始研究掷骰子
  - 17世纪Pascal等解决了“赌徒分金”问题
  - J. 伯努利等奠定了概率论的基础
  - 18世纪拉普拉斯引入分析概率
  - 20世纪，柯尔莫哥洛夫奠定了公理化概率理论体系

- 数理统计是研究怎样有效地收集、整理和分析带有随机性的数据，以及对所考察的问题作出推断或预测，直至为采取一定决策和行动提供依据和建议的学科

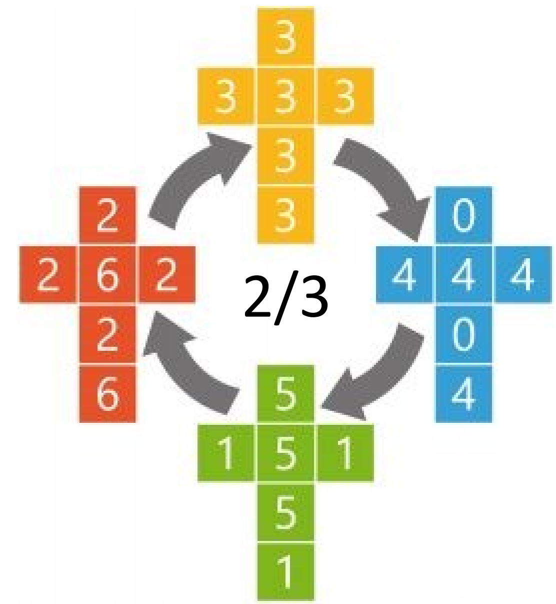
- 概率是数理统计的基础，数理统计是概率的一种应用

# 概率与大数据

- 三类意义下的概率
  - 古典概率/先验概率——每种结果等可能，n种可能，k种发生，则为k/n。典型的掷骰子、抛硬币都是
  - 频率概率或统计概率——具有不同发生概率的事件，通过反复观察，记录某种结果出现的概率
  - 分析概率——以数学分析为基础

# 不一样的骰子——非传递骰子

- 石头-剪子-布一般的骰子
- 你总可以赢




Aka.ms/ntdice

# 概率的应用

- 生活中最重要的问题，其中绝大多数在实质上只是概率的问题

  ——拉普拉斯

- 概率的重要性——现代若干学科的基础
  - 机器学习、模式识别
  - 电子系统、航空、航天——可靠性
  - 制造业——产品抽样
  - 保险业——赔付赔率的设定
  - 医学——治疗的有效性检验
  - 人口调查——抽样
  - 传染病的流行——生灭过程
  - ……

# 一些问题

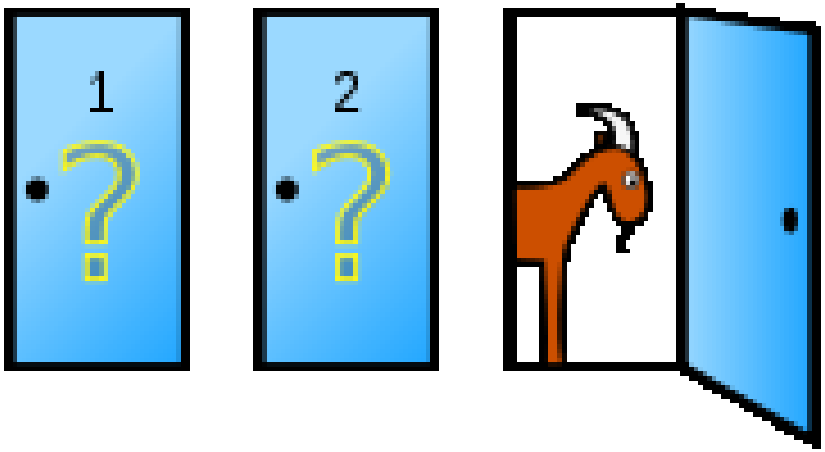
- 问题一(Chevalier de Méré问题)：
  掷一个骰子四次中出现一次“6”的概率与掷一对骰子24次中出现一对“6”的概率哪个大？

- 问题二(公平还是不公平)：
  有ABCD四个人，A坐庄，要在纸片上写下1234四个数中的一个，由他的下家B开始猜，猜之前往杯子里倒酒，倒多少随意，如果猜中，则杯里的酒由B 喝，猜不中的话，C随意倒酒，继续猜，猜中喝酒。如果最后大家都没猜中，那么庄家喝杯里所有的酒。第一轮中谁喝酒，谁就是下一轮的庄家。

  你愿意做那个角色？

# 问题三(Monty Hall问题)

- 三个门，其中有一个门后面是车，其他两个门后面是羊。当游戏者选中了一个门之后，无论怎样，主持人在知道他首次的选取情况的基础之上，从剩下的两个门之中打开一个后面是羊的门。游戏者也知道主持人是故意选的后面没有车的门。 问题：此时，游戏者应当保持自己原先选择的那扇门，还是换选另外一个门，以最大可能地获得汽车？



# 问题四(Chevalier de Méré的“赌徒分金”)

- 梅累和赌友掷骰子，各人下赌注32个金币，约定先赢三局为胜（梅累如果先掷出三次6点，或者赌友先掷出三次4点就算赢了对方）．赌博进行了一段时间后，梅累已经两次掷出6点（赢了两局），而赌友已经掷出一次4点（赢了一局），这时梅累接到通知，要其马上去陪国王接见贵宾，赌博便只好中断了，那么，留下的这64个金币两人应该怎样分才算合理呢？

# 问题五(抛硬币决定命运)

- 通过“正/反”决定与通过“正反反 / 反反正”决定
- 连续抛硬币，直到最近三次结果出现“正反反”或“反反正”为止
- 前两次决定命运，只有“反反”后者才会赢

# 问题六(路边的游戏)

- 两个路边摆摊的艺人在和路人玩两个相似的游戏，一个通过艺人和游戏参与者“**抛硬币**”决定，另一个通过双方“**亮硬币**”决定，如果双方两人的结果**都是正面**，那么艺人付给参与者**3元**，如果**都是反面**，艺人付给参与者**1元**，**剩下的情况**，则参与者需要付给艺人**2元。**”
  这两个游戏的公平性是否一样？你该和谁玩？如何玩？

# 问题七(独生子女)

- 独生子女如果老大是女儿，允许生第二个孩子，是否会破坏男女比例

# 问题八(俄罗斯轮盘赌)

- 左轮手枪中只有一颗子弹决斗，双方轮流对准自己头部扣动扳机，概率如何？
- 如果手枪中有两颗连续的子弹，概率如何？三颗、四颗、五颗子弹又如何？

# 问题九(统计的障眼法)

- 手术的有效性
- 1986 年，为了对比肾结石治疗中两种手术的优劣，有人统计了接受这两种手术的病人数量和成功率。在统计中，按照严重程度把病人分为大块结石和小块结石两组，然后分别统计了两组病人中手术 A 和手术 B 的成功率。得到的结果可以清晰地展示在下面的这个表格中：

| | 手术A | 手术B |
|---|---|---|
| 小块结石 | 93% | 87% |
| 大块结石 | 73% | 69% |

# 计数与空间的骗局

- 困惑的兄弟

# 课程的主要内容

- 第一部分 概率
  - 随机事件与概率
  - 随机变量及其分布
  - 多维随机变量及其分布
  - 随机变量的数字特征
  - 大数定律及中心极限定理

# 课程的主要内容

- 第二部分 数理统计
  - 样本及抽样分布
  - 参数估计
  - 假设检验
  - 方差分析
  - 随机过程及其统计描述
  - 马尔可夫链
  - 平稳随机过程
  - Bayes统计推断

# 考试

- 成绩计算：
  - 作业占总成绩30%
  - 期中考试成绩占总成绩30%
  - 期末考试成绩占总成绩40%

# 第一章 随机事件与概率

# §1. 随机事件及其概率

- 现象——客观世界中存在着两类现象
  - 一类是在一定的条件下必然出现的现象，称之为必然现象
  - 另一类是在一定的条件下可能出现也可能不出现的现象，称之为随机现象
- 随机试验
  1. 可在相同条件下重复进行
  2. 一次试验之前无法确定具体出现哪种结果
  3. 能确定所有的可能结果
- 随机试验常用$E$ (Event)表示

# 样本空间与样本点

- 样本点——随机试验的单个结果或样本空间的单个元素称为样本点
- 由样本点组成的单点集称为基本事件,也记为e.
- 样本空间——随机试验的所有可能结果所组成的集合称为样本空间，一般记为$\Omega=\{e\}$(书中为S)
- 随机事件
  - 定义： 样本空间的任意一个(有限或无限可列个元素构成的)子集称为随机事件, 简称“事件”
  - 随机事件记作A、B、C等
  - 任何事件均可表示为样本空间的某个子集
  - 称事件A发生当且仅当试验的结果是子集A中的元素
  - 两个特殊事件:
    - 必然事件：$\Omega$
    - 不可能事件：$\varnothing$

# 试验与事件

- E1: 抛一枚硬币，观察正面H和反面T的出现情况
- E2: 将一枚硬币连抛三次，观察正反面出现的情况
- E3:掷一颗骰子 现的点数
- E4: 站一分钟内访 的次数
- E5:在一批灯泡中任取一只，测其寿命;

- 对于试验E2，以下A、B、C为三个 件:
  - A为 正面”，
    HH, HHT, HTH, HTH，HTT，THT，TTH}；
  - B为“三次出现同一面”，于是B={HHH,TTT}
  - C为“恰好出现一次正面”，于是C={HTT，THT，TTH}
- 试验E5中，定义D = “灯泡寿命超过1000小时”，于是D =

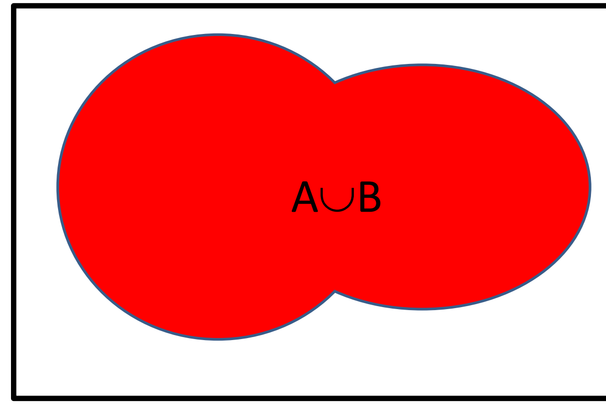
事件是集合，因此有关事件间的关系、运算及运算规则也就按集合间的关系、运算及运算规则来处理

# 样本空间的一些例子

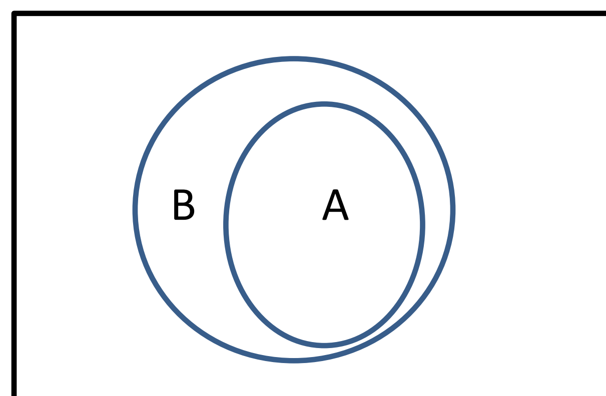
- 观察某电话交换台在一天内收到的呼叫次数，其样本点有可数无穷多个：
  样本空间为={0次，1次，2次, …… }
- 连接射击直到命中为止。为了简洁地写出其样本空间，我们约定以“0”表示一次射击未中，而以“1”表示命中。
  样本空间 ={1, 01, 001, 0001, …… }

# 集合的一些运算

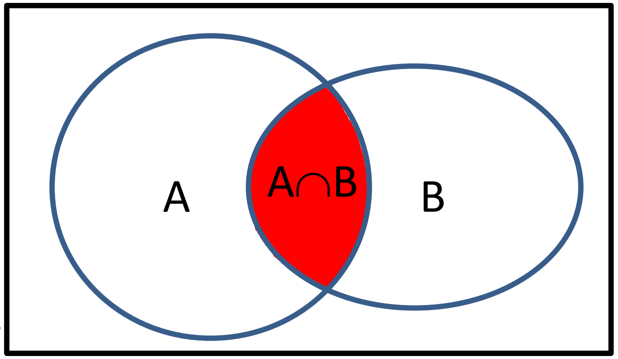
- 事件E的样本空间为Ω，A, B, …是其中的子集
- 集合的补
  - $\bar{\mathbf{A}} = \Omega \backslash \mathrm{A}$



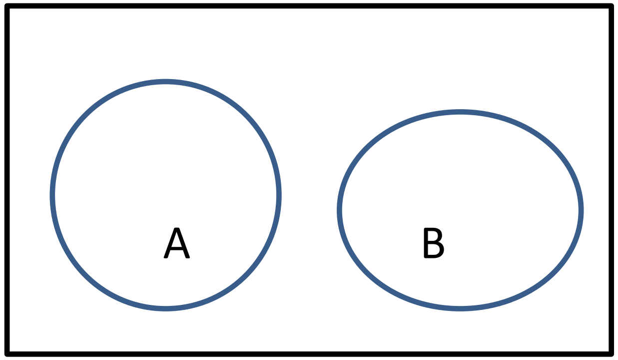

- 集合的并
  - 两个集合的并
    $$\mathrm{A} \cup \mathrm{B} = \{\mathrm{x} | \mathrm{x} \in A \ or \ \ x \in B\}$$
  - 多个集合的并$\bigcup_{k=1}^{n} A_k$,为$n$个事件$A_1, A_2, \dots, A_n$的并事件
  - $\bigcup_{k=1}^{\infty} A_k$,为可列个事件$A_1, A_2, \dots$的并事件
  - 集合并的“事件”含义
    - N个事件至少有一个发生
- 若A⊂B ➔ 称为事件B包含事件A，
  即A发生必导致B发生
  若A⊂B且B⊂A，则A=B

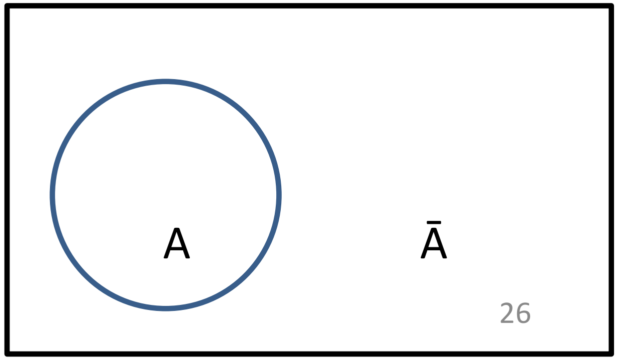


# 集合的一些运算



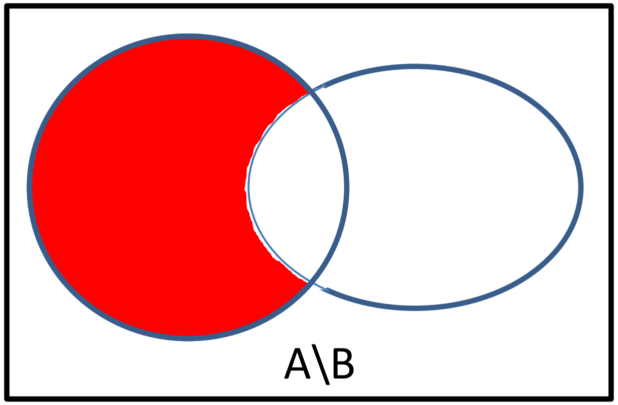

- 集合的交
  - 两个集合的交
    $$A \cap B = \{x|x \in A \text{ and } x \in B\}$$
  - 多个集合的交$\cap_{k=1}^{n} A_k$,为$n$个事件$A_1, A_2, \ldots, A_n$的交事件
  - $\cap_{k=1}^{\infty} A_k$,为可列个事件$A_1, A_2, \ldots$的交事件
  - 集合交的“事件”含义
    - N个事件同时发生
    - 若A∩B=∅，则称A与B不相容或互斥
    - 若A∩B=∅，且A∪B=Ω，则A、B为逆事件或对立事件，A的对立事件又记为$\bar{A}$=Ω\A



# 集合的一些运算

- 集合的差
  - 两个集合的“差”
  - 集合“差”的“事件”含义
    - A事件发生且B事件不发生
    - 何时：A\B=∅
    - 何时：A\B=A



# 集合基本运算

- 交换律
  $A \cup B = B \cup A$,
  $A \cap B = B \cap A$
- 结合律
  $A \cup (B \cup C) = (A \cup B) \cup C$,
  $A \cap (B \cap C) = (A \cap B) \cap C$
- 分配律
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- De Morgan律
  - $\overline{A \cup B} = \overline{A} \cap \overline{B}$
  - $\overline{A \cap B} = \overline{A} \cup \overline{B}$

# 事件运算的例子

- 掷三枚硬币，A、B、C分别表示三枚正面朝上的事件，则可以有事件组合：
  - A1: 至少有一枚正面朝上：$A \cup B \cup C$
  - A2: 恰有一枚正面朝上：$A\overline{B}\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\overline{B}C$
- 掷一个骰子，A表示掷出“6”点事件，B表示掷出“1-5”的事件
  - $A=\Omega-B$