DATA601 UMD 概率论与统计

概率论v.s.统计学：

- 应用：物理、金融、计算机科学
- 典型概率论问题：投掷均匀硬币100次，头朝上大于60次的概率
- 典型统计学问题：投掷硬币100次，头朝上64次，有多少概率认为这个硬币时有偏的
- **概率论**：
  - 已知系统
  - 预测未来
- **统计学**：
  - 已知数据
  - 想要预测系统

样本空间Sample Space：

- 定义：所有基本结果空间，记作$\mathcal{S}$或Ω
- 例如：投掷硬币$\mathcal{S}=\{H,T\}$，投掷骰子$\mathcal{S}=\{1,2,3,4,5,6\}$，投掷三枚硬币$\mathcal{S}=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$，一个人想要通过驾驶证的尝试次数$\mathcal{S}=\{1,2,...,\}$

事件Event：

- 定义：样本空间的子集
- 例如：投掷硬币三次，事件D定义为头超上4次，则$D=\{\}=\emptyset$

事件的运算：

- $A\cup B$ **Union**：A发生或B发生或同时发生
- $A\cap B$ **Intersection**：A和B同时发生
- A^C **Complement**：A没有发生
- $A-B$ **Difference**：发生A但没有发生B
- 如果$A\cap B=\emptyset$，则A和B被称作不相交Disjoint，或相互排斥Mutually Exclusive

恒等式：

- Union和Intersection分别满足交换律
- $E\cap(F\cup G)=(E\cap F)\cup(E\cap G)$
- $E\cup(F\cap G)=(E\cup F)\cap(E\cup G)$
- $P(E)=1-P(E^C)$

DeMorgan's Law：

- $(E\cup F)^C=E^C\cap F^C$
- $(E\cap F)^C=E^C\cup F^C$
- 前两个式子可以表达为多项式形式，如$(U_{i=1}^{n}E_i)^C$

Note：

- 不是每个样本空间的子集都能被称为一个事件

σ-algebra：一个关于$\mathcal{S}$的子集集合$\mathcal{F}$被称为σ-algebra如果它有一下三个性质：

1. $\mathcal{S}\in\mathcal{F}$

2. 如果$E \in \mathcal{F}$，那么$E^C \in \mathcal{F}$
3. 如果$E_1, E_2, ... \in \mathcal{F}$，那么$E_1 \cup E_2 \cup ... \in \mathcal{F}$

- **注意空集**也可以在σ-algebra中

概率：

- ~~一个$\mathcal{F}$到$\mathcal{R}$的映射，代表出现次数#，记作$P(A)$，对于每个S中的事件A~~见书

$$P(E) = \lim_{n \longrightarrow \infty} \frac{n(E)}{n}$$

- $P(A)$：事件A的概率

概率相关公理：

- Unit Interval Axiom：对于任何事件A，$0 \leq P(A) \leq 1$
- 确定性公理：$P(S) = 1$
- 相互排斥事件的加法公理：对于任何相互排斥的事件$A_1, A_2, ...(A_i \cap A_j = \emptyset\ i \neq j)$，则

$$P(\cup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(A_i)$$

子集与不等式：

- 如果$E \subset F$，则$P(E) \leq P(F)$

其它Property：

- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cup B \cup C)$
- $P(A \cup B \cup C \cup D) =$add singletons-pair interstions+triple intersection - quadruple intersection
- $P(A \cap B^C) = P(A) - P(A \cap B)$

含有相同可能性样本空间：

- 对于任意事件E，$P(E) = \frac{\#E\ outcomes}{\#\Omega\ outcomes}$

Combinatonal Analysis：

- 基础计数定理：假设有2个实验，实验一有m个结果，实验二有n个结果，则一共有$m \cdot n$个可能的结果

排列Permutations：不同的排列顺序，顺序重要

- 无重复时，从n个物体中选k个五i的排列是$P(n, k) = \frac{n!}{(n-k)!}$

组合Combination：从一个集合中选择物体，顺序不重要

- 无重复时，从n个物体中选k个的组合是$C(n, k) = \frac{n!}{(n-k)!k!}$

通常，对于相同的n和k，排列的数量大于组合的数量

多项式系数：有$\frac{n!}{n_1!n_2!...n_k!}$个不同的排列对于n个物体，其中$n_i$是一样的。排列是重要的，但是相同物体的排列是不重要的

EX: How many different 8 letter arrangements can be made from the letters NONSENSE?

$$\frac{8!}{3! \cdot 2! \cdot 2!}$$