一、概率

Probability

一、样本空间与事件

1. 基本概念

- **样本空间(sample space)**：样本空间Ω是一个实验中所有可能结果的集合，Ω中的点ω称为**sample outcomes**或者**realizations**。
  - 例如：投两次硬币(实验)的样本空间为$\Omega = \{HH, HT, TH, TT\}$，无限投硬币的样本空间则是无穷集合$\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \ldots), \omega_i \in \{H, T\}\}$。
- **事件(event)**：事件是Ω的子集。
  - 例如：在投两次硬币的实验中，“第一次投正面朝上”就是一个事件，$A = \{HH, HT\}$。

2. 事件的补、并、交、差

- **事件的补**：给定事件A，称$A^c = \{\omega \in \Omega; \omega \in A\}$为事件$A$的补事件。特别地，$\Omega$的补事件为$\emptyset$。
- **事件的并**：事件A和B的并集表示为$A \cup B = \{\omega \in \Omega : \omega \in A \ or \ \omega \in B \ or \ \omega \in both\}$。
- **事件的交**：事件A和B的交集表示为$A \cap B = \{\omega \in \Omega : \omega \in A \ and \ \omega \in B\}$。
- **事件的差**：事件A和B的差集表示为$A - B = \{\omega : \omega \in A, \omega \notin B\}$。

3. 事件互斥和样本空间划分

- **事件互斥**：如果$A_i \cap A_j = \emptyset, i \neq j$，则称事件$A_i$和$A_j$互斥。
- **样本空间划分**：若存在一个互斥事件集合$A_1, A_2, \ldots$且满足$\cup_{i=1}^{\infty} A_i = \Omega$，则称事件集合$A_1, A_2, \ldots$为样本空间$\Omega$的一个划分(partition)。

4. 其他

- **指示函数(indicator function)**：给定一个事件A，定义A的指示函数为

$$I_A(\omega) = I(\omega \in A) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

- **单独递增集合序列**：若集合$A_1 \subset A_2 \subset \ldots$，则称集合序列$A_1, A_2, \ldots$单调递增，并定义$\lim_{n \to \infty} A_n = \cup_{i=1}^{\infty} A_i$。
- **单独递减集合序列**：若集合$A_1 \supset A_2 \supset \ldots$，则称集合序列$A_1, A_2, \ldots$单调递减，并定义$\lim_{n \to \infty} A_n = \cap_{i=1}^{\infty} A_i$。
  - 令$\Omega = \mathbb{R}$且$A_i = [0, 1/i), i = 1, 2, \ldots$；那么有$\cup_{i=1}^{\infty} A_i = [0, 1)$且$\cap_{i=1}^{\infty} A_i = \{0\}$。

二、概率

1. 概率的定义

为每个事件A分配一个实值$\mathbb{P}(A)$，称$\mathbb{P}(A)$为A的概率，$\mathbb{P}$是概率分布或者概率度量。下面是正式定义

> 将每个事件A映射为实值$\mathbb{P}(A)$的函数$\mathbb{P}$满足公理：
>
> - 公理1：对于所有A，$\mathbb{P}(A) \geq 0$;
> - 公理2：$\mathbb{P}(\Omega) = 1$;
> - 公理3：若$A_1, A_2, ...$间互斥，那么有$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$;
>
> 则称$\mathbb{P}$为概率分布(**probability distribution**)或者概率度量(**probability measure**)。

2. 概率的解释

$\mathbb{P}(A)$有两种常见的解释：频率和确信程度。

- 频率：在重复实验的过程中，事件A为true的长期比例为$\mathbb{P}(A)$;
- 确信程度：$\mathbb{P}(A)$衡量观测者对于A的确信程度；

  在统计推断中，这两种观点将导致频率学派和贝叶斯学派。

3. 概率的连续性

> 引理：对于任意事件A和B，$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$;

概率的连续性：在$n \to \infty$情况下，若$A_n \to A$，则$\mathbb{P}(A_n) \to \mathbb{P}(A)$;

三、有限样本空间的概率

若样本空间Ω是有限的并且每种输出是等概率的，那么

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

称为均匀概率分布(uniform probability distribution)。

四、独立事件

1. 独立定义

- **两事件独立**：若$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$，则称事件$A$和$B$是独立的，记为$A \coprod B$。
- **多事件独立**：存在一个事件集合$\{A_i : i \in I\}$，若满足$\mathbb{P}(\cap_{i \in J} A_i) = \prod_{i \in J} \mathbb{P}(A_i)$，则称$\{A_i : i \in I\}$相互独立，其中$J$是$I$的所有可能子集。

  > 具有正概率的互斥事件不一定独立。即$\mathbb{P}(A)\mathbb{P}(B) > 0$，但可能$\mathbb{P}(AB) = \mathbb{P}(\emptyset) = 0$。

2. 确定事件独立

有两种确定事件独立的方式：

- **基于事实**：基于事实观察，直接假设事件独立。
  - 例如投两次硬币，两次投掷间并不会相互影响，因此是相互独立的。
- **基于公式**：通过推导公式$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$来确定事件独立。

五、条件概率

1. 定义

若概率$\mathbb{P}(B)>0$，那么给定B的情况下A的**条件概率**为

$$\mathbb{P}(A|B)=\frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$$

2. 性质

- 对于固定的B，$\mathbb{P}(\cdot|B)$是一个概率，其满足概率的三个公理

$$\mathbb{P}(A|B)\geq 0$$
$$\mathbb{P}(\Omega|B)=1$$
$$\mathbb{P}(\cup_{i=1}^{\infty}A_i|B)=\sum_{i=1}^{\infty}\mathbb{P}(A_i|B),\quad A_1,A_2,\ldots\text{互斥}$$

- 一般来说，$\mathbb{P}(A|B)\neq\mathbb{P}(B|A)$；
- 当且仅当$\mathbb{P}(A|B)=\mathbb{P}(A)$，则$A$和$B$独立。
  - 独立的另一个解释是：知识B的情况下并不能改变概率A。

六、贝叶斯理论

1. 全概率法则

令$A_1,\ldots,A_k$是Ω是一个划分，对于任意事件B，有

$$\mathbb{P}(B)=\sum_{i=1}^{k}\mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

2. 贝叶斯定理

设$A_1,\ldots,A_k$是Ω的一个划分，且每个$\mathbb{P}(A_i)>0$。若$\mathbb{P}(B)>0$，那么每个$i=1,\ldots,k$

$$\mathbb{P}(A_i|B)=\frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_j\mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

其中，$\mathbb{P}(A_i)$称为A的**先验概率(prior probability)**，$\mathbb{P}(A_i|B)$是A的后验概率**(posterior probability)。**