Lecture 03: Conditional Probability and Bayes Formula

Conditional Probability

“事件 A 和 B 同时发生的概率“与” B 发生的条件下，A 发生的概率“不一定相等！

- “事件 A 和 B 同时发生” 是在 Ω 中 寻找 $A \cap B$，$P(A \cap B) = \frac{|A \cap B|}{|\Omega|}$。
- " B 发生的条件下，A 发生"，样本空间退化为 B，$P(A|B) = \frac{|A \cap B|}{|B|}$。

$\boxed{\text{Definition 3.1}}$ 设事件 A, B 满足 $P(B) > 0$，则称 $P(A|B) = \frac{P(AB)}{P(B)}$ 为在事件 B 发生的条件下事件 A 发生的概率。

注：若将 B 视为一个样本空间 Ω_B，则可定义概率空间 $(\Omega_B, \mathscr{F}_B, P_B)$，其中 $\mathscr{F}_B = \{A \cap B | A \in \mathscr{F}\}$，$P_B(A) = P(A|B)$。此时若直接取 P 为 Ω_B 的概率测度，即令 $P_B = P$，则 $P_B(B) < 1$，违反了概率空间的定义 (出现了某种概率损失)，所以应该对其除掉 $P(B)$ 进行一个**归一化**：$P_B(A) = \frac{P(AB)}{P(B)}$。

事实上， P_B 也是 $(\Omega, \mathscr{F})$ 上的一个概率

- $P_B(A) \geq 0$，$P_B(\Omega) = 1$；
- 满足可列可加性。

因此，之前推导的关于交、并的概率公式在条件概率上仍然适用。P_B 的特点是将所有发生的事件聚焦在 B 以内，即对于任意 $A \cap B = \emptyset$，$P_B(A) = 0$。

乘法公式：当 $P(A_{n-1} \cdots A_1) > 0$ 时，

$$P(A_n A_{n-1} \cdots A_1) = P(A_n | A_{n-1} A_{n-2} \cdots A_1) P(A_{n-1} | A_{n-2} \cdots A_1) \cdots P(A_2 | A_1) P(A_1)$$

注：对于任意 $1 \leq m \leq n-1$，$\bigcap_{i=1}^{n-1} A_i \subseteq \bigcap_{i=1}^{m} A_i$，因此 $P(\bigcap_{i=1}^{n-1} A_i) > 0$ 的条件已经为之后所有的条件概率做好了假设。

【例】n 个人抽签，求放回/不放回的情况下，第 k 个人抽中的概率。

> 解：令 A_k 表示事件“第 k 个人抽中“，则要求 $P(A_k \overline{A_{k-1}} \cdots \overline{A_1})$，根据乘法公式，
> $P(A_k \overline{A_{k-1}} \cdots \overline{A_1}) = P(A_k | \overline{A_{k-1}} \cdots \overline{A_1}) P(\overline{A_{k-1}} | \overline{A_{k-2}} \cdots \overline{A_1}) \cdots P(\overline{A_2} | \overline{A_1}) P(\overline{A_1})$。
>
> (a) 当 $k = 1$ 时，$P(A_k) = \frac{1}{n}$，当 $k \geq 2$ 时，$P(\overline{A_1}) = \frac{n-1}{n}$，对于任意 $2 \leq m \leq k-1$，$P(\overline{A_m} | \overline{A_{m-1}} \cdots \overline{A_1}) = \frac{n-m}{n-m+1}$，$P(A_k | \overline{A_{k-1}} \cdots \overline{A_1}) = \frac{1}{n-k+1}$，所以 $P = (\prod_{m=1}^{k-1} \frac{n-m}{n-m+1}) \frac{1}{n-k+1} = \frac{1}{n}$
>
> (b) 当 $k = 1$ 时，$P(A_k) = \frac{1}{n}$，当 $k \geq 2$ 时，$P(\overline{A_1}) = \frac{n-1}{n}$，对于任意 $2 \leq m \leq k-1$，$P(\overline{A_m} | \overline{A_{m-1}} \cdots \overline{A_1}) = \frac{n-1}{n}$，$P(A_k | \overline{A_{k-1}} \cdots \overline{A_1}) = \frac{1}{n}$，所以 $P = \left(\frac{n-1}{n}\right)^{k-1} \frac{1}{n}$。

Total Probability Formula

$\boxed{\text{Definition 3.2}}$ 若事件 $A_1, A_2, \cdots, A_n$ 满足

- 对于任意 $1 \leq i < j \leq n$，$A_i \cap A_j = \emptyset$；
- $\bigcup_{i=1}^{n} A_i = \Omega$。

则称 $A_1, \cdots, A_n$ 为 Ω 的一个**划分/完备事件集**。

$\boxed{\text{Theorem 3.3}}$ (全概率公式) 设 $A_1, \cdots, A_n$ 为 Ω 的一个划分，则

$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i) \, 1\{P(A_i) > 0\}$$

其中最后的部分是示性函数 (indicator function)，表示略过 $P(A_i) = 0$ 的那些部分的条件概率。

> Proof：首先有 $B = B\Omega = B(\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n}(BA_i)$，且 BA_i 彼此互不相容，所以
>
> $$P(B) = P\left(\bigcup_{i=1}^{n}(BA_i)\right) \overset{\text{可列可加性}}{=} \sum_{i=1}^{n} P(BA_i) \qquad \blacksquare$$

注：全概率公式对于无限可列的划分仍然成立。

对于全概率公式的一个感性理解是：**如果要计算事件A的概率，我们可以“分类讨论”A在各种情况下发生的概率，再全部加起来。**

对于一个条件概率 $P(B|C)$，我们可以考虑一个划分 $A_1, \cdots, A_n$ 并将其写为 $P(B|C) = \sum_{i=1}^{n} P(BA_i|C)$。若这仍不好算，仍然可以有以下变形：

$$\sum_{i=1}^{n} P(BA_i|C) = \sum_{i=1}^{n} \frac{P(BA_iC)}{P(C)} = \sum_{i=1}^{n} \frac{P(A_iC)}{P(C)} P(B|A_iC) \, 1\{P(A_iC) > 0\}$$

我们将条件 C 加强到了条件 A_iC，从而可能更容易计算。

【例】有两批灯泡各10支，第一批有一个次品，第二批有两个次品。运输过程中两批各打碎了一个。现从剩余灯泡中抽取一个，抽到次品的概率是多少？

> 解：讨论打碎的两个灯泡的情况：
> $B_1 = (\text{好}, \text{好}), B_2 = (\text{好}, \text{次}), B_3 = (\text{次}, \text{好}), B_4 = (\text{次}, \text{次})$。记 A 为抽到次品这个事件
>
> $$\begin{aligned} P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) + P(B_4)P(A|B_4) \\ &= \frac{9}{10} \cdot \frac{8}{10} \cdot \frac{3}{18} + \frac{9}{10} \cdot \frac{2}{10} \cdot \frac{2}{18} + \frac{1}{10} \cdot \frac{8}{10} \cdot \frac{2}{18} + \frac{1}{10} \cdot \frac{2}{10} \cdot \frac{1}{18} \\ &= \frac{3}{20} \end{aligned}$$

Bayes Formula

在全概率公式 $P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$ 中，B 可以看作发生的结果，A_i 可以看作 B 发生的可能原因。$P(A_i)$ 被称为先验概率，$P(A_i|B)$ 衡量了 B 已经发生的情况下原因发生的可能性，称为后验概率。

$\boxed{\text{Theorem 3.4}}$ (贝叶斯公式) 设 $A_1, \cdots, A_n$ 为 Ω 的一个划分，对于任意 $1 \le k \le n$，$P(A_k) > 0$。则对于 $B \in \mathscr{F}$，我们有

$$P(A_k|B) = \frac{P(A_kB)}{P(\Omega B)} = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

【例】有一种罕见病，用某方法来检测时，一个人患病且被检测出有病的概率是 0.95，一个人没有患病且被检测出没病的概率是 0.9，正常人中罕见病发病率为 0.0004。现有一个人检测出有病，他真正有病的概率是多少？

> 解：令 $C = \{\text{某人患罕见病}\}, A = \{\text{某人被检测出罕见病}\}$，则
>
> $$P(C|A) = \frac{P(C)P(A|C)}{P(C)P(A|C) + P(\overline{C})P(A|\overline{C})} = \frac{0.0004 \cdot 0.95}{0.0004 \cdot 0.95 + 0.9996 \cdot 0.1} \approx 0.003$$

该概率实际上非常小的原因是：人群中不患病的很多，而不患病情况下的误诊率不够小，这导致有很大概率都是这种情况导致的检测结果异常。