第四章. 条件概率. 独立性和贝叶斯定理.

4.1 条件概率.

在已知某个事件的前提下. 我们考察的焦点会落在样本空间的一个子集上

4.1.1 猜测条件概率公式.

1° $P_r(A|B) = F(P_r(A), P_r(B))$

考虑当 $P_r(A) = P_r(B) = P_r(B^c)$ 时

由 $P_r(B) + P_r(B^c) = 1$. 有上述3个事件的概率均为 $1/2$.

所以. $P_r(A|B) = F(\frac{1}{2}, \frac{1}{2}) = F(P_r(B), P_r(B^c))$

$= P_r(B|B^c) = 0$.

而 $P_r(A|B) = F(\frac{1}{2}, \frac{1}{2}) = F(P_r(B), P_r(B))$

$= P_r(B|B) = 1$.

矛盾.

2° $P_r(A|B) = F(P_r(A), P_r(B), P_r(A \cap B))$

考虑到条件概率可以看作对样本空间的调整，将 $Pr(B)$ 作为除数的猜想是很合理的。

我们给出：$Pr(A|B) = \dfrac{\alpha Pr(A) + \beta Pr(A\cap B)}{Pr(B)}$

① 当 $Pr(B) = 1$ 时（B 为必然事件）

有 $Pr(A|B) = Pr(A) = (\alpha+\beta) Pr(A)$

$\therefore \alpha+\beta = 1$.

② 当 $Pr(A) = 1$ 时（A 为必然事件），

有 $Pr(A|B) = Pr(A) = 1$

$$= \frac{\alpha}{P(B)} + \beta.$$

1° $\alpha \neq 0$. 令 $P(B) \to 0$. $\dfrac{\alpha}{P(B)} + \alpha \to +\infty$ (舍)

2° $\alpha = 0$. $1 = \beta$.

故 $\alpha = 0$, $\beta = 1$.

我们得到 $Pr(A|B) = \dfrac{Pr(A\cap B)}{Pr(B)}$.

（习题 4.8.1）.

4.1.2 由期望计数法推导.

表 4-1 关于事件 A 和 B 的可能结果. 如果已知事件 B 发生, 那么只需考察 B 所在行的事件就行了

	A	A^c
B	$A\cap B$	$A^c\cap B$
B^c	$A\cap B^c$	$A^c\cap B^c$

如果只考察关于 A 和 B 的可能结果, 那么这样的结果有 4 个. (注意, A^c 和 B^c 分别表示“A 不发生”和“B 不发生”.) 但是, 如果已知事件 B 发生, 那么我们的考察范围就限定在表 4-1 中 B 所在的行上. 此时, 可能的结果只有 $A\cap B$ 和 $A^c\cap B$. 于是, 条件概率 $\Pr(A|B)$ 由下面的式子给出

$$\Pr(A|B) = \frac{\Pr(A\cap B)}{\Pr(A\cap B)+\Pr(A^c\cap B)} = \frac{\Pr(A\cap B)}{\Pr(B)}.$$

最后一个等号之所以成立是因为 $(A\cap B)\cup(A^c\cap B)=B$, 而且对于两个互不相交的集合的并, 其概率就是两个概率之和 (这是概率公理之一).

概率第三公理.

4.1.3 文氏图法推导.

4.1.4 蒙提霍尔问题 (Monty Hall)

看不懂解析, 但是用期望计数法很好理解.

附有代码验证.

4.2 一般乘法法则.

$$\Pr(A\cap B) = \Pr(A|B)\times \Pr(B)$$

4.2.2 补充问题.

要领：有时条件概率公式很好用，但有时直接从题给信息出发更好.

4.2.3 帽子问题与纠错码.

4.2.4 条件概率的高等注解.

也是没看懂.

4.3 独立性

独立性的暂时定义： 考虑事件 A 和 B. 如果

$$\Pr(A|B) = \Pr(A),$$

那么我们暂时说 A 和 B 是**相互独立的**. 也就是说, 如果事件 B 的发生不会改变事件 A 发生的概率, 那么 A 和 B 就是**相互独立的**.

考虑到：① 此公式不覆盖 $Pr(B)=0$ 的情况.

② 定义应该能体现独立性质的对称性.

我们需要进行拓展.

此, 两个事件独立性的正确定义是

> **独立性 (两个事件)**：如果事件 A 和 B 满足
> $$\Pr(A\cap B)=\Pr(A)\cdot\Pr(B),$$
> 那么 A 和 B 就是独立的.

三个或更多个事件的独立性该如何定义？我们对两个事件独立性的定义可以扩展到三个或更多的事件上.

> **独立性 (三个事件)**：如果事件 A, B 和 C 满足
> (1) $\Pr(A\cap B\cap C)=\Pr(A)\cdot\Pr(B)\cdot\Pr(C)$, 并且
> (2) 其中任意两个事件都是独立的,
> 那么 A、B 和 C 就是**相互独立的**.

> **独立性 (一般情形)**：如果事件 $A_1,\cdots,A_n$ 满足
> (1) $\Pr(A_1\cap\cdots\cap A_n)=\Pr(A_1)\cdots\Pr(A_n)$, 并且
> (2) $\{A_1,\cdots,A_n\}$ 的任意一个非空子集都是相互独立的,
> 那么 $A_1,\cdots,A_n$ 就是相互独立的.

关于三个或更多个事件的独立性, 这里有个重要的警告：当任意两个事件都独立时, 三个或更多事件可能会相互依赖. 例如, 抛掷一枚骰子两次. 设

反例： A = 第一次掷出偶数.

B = 第二次掷出偶数.

C = 两次掷出点数和是偶数.

4.4 贝叶斯定理.

贝叶斯定理：由一般乘法法则可以推出：对于事件 A 和 B, 有

$$\Pr(B|A) \cdot \Pr(A) = \Pr(A|B) \cdot \Pr(B).$$

因此, 只要 $\Pr(B) \neq 0$, 我们就有

$$\Pr(A|B) = \Pr(B|A) \cdot \frac{\Pr(A)}{\Pr(B)}.$$

4.5 划分和全概率法则

有趣的类比：条件概率之间的转换, 类似交换积分、求和和求导的顺序.

目的：构造一般版本的贝叶斯定理.

因为有的时候作为条件的 Pr(A)，
Pr(B)，并不直观. 需要一些额外的条件.

样本空间 S 的一个**划分**就是满足下列条件的可数个集合 $\{A_1, A_2, \cdots\}$.

(1) 如果 $i \neq j$, 那么 A_i 和 A_j 不相交. 我们通常用 $A_i \cap A_j = \varnothing$ 来表示这两个集合的交是空集.

(2) 全体 A_i 的并就是整个样本空间: $\bigcup_i A_i = S$.

不重、不漏

全概率法则: 如果 $\{B_1, B_2, \cdots\}$ 构成了样本空间 S 的一个划分 (分成了至多可数个部分), 那么对于任意的 $A \subset S$, 我们有

$$\Pr(A) = \sum_n \Pr(A|B_n) \cdot \Pr(B_n).$$

对于所有的 n, 都应该有 $0 < \Pr(B_n) < 1$, 否则条件概率就是无定义的 (注意, 如果有一个 B_n 的概率为 0, 那么我们就不需要这个 B_n 了, 因为它会给出因子 $\Pr(B_n) = 0$; 但如果它的概率是 1, 那么其他所有项都是不必要的).

由不相交集合的概率可加性可得.

我们现在已经找到了一个合理替
换贝叶斯定理中 Pr(A)，Pr(B) 的方式.

4.6 回顾贝叶斯定理.

贝叶斯定理： 设 $\{A_i\}_{i=1}^n$ 是样本空间的一个划分, 那么

$$\Pr(A|B) = \frac{\Pr(B|A)\cdot\Pr(A)}{\sum_{i=1}^n \Pr(B|A_i)\cdot\Pr(A_i)}.$$

通常情况下, A 就是某个 A_i.

例.

一个人的血型是由一组基因决定的. 假设一个人遗传某个等位基因的概率完全独立于他遗传的其他等位基因, 并且该基因有三种可能性, 即 A、B 和 O. 如果某人是 AA 或 AO, 那么他就是 A 型血. 如果是 BB 或 BO, 那他就是 B 型血. 如果是 OO, 那就是 O 型血; 如果是 AB, 那么他的血型就是 AB. 不妨设有 45% 的等位基因是 O, 40% 是 A, 而剩下的 15% 是 B.

习题 4.8.19 已知 Justine 是 O 型血, 计算她的父母具有每种血型的概率.

习题 4.8.20 计算 Justine 的哥哥也是 O 型血的概率.

习题 4.8.21 A 型血的人既可以输入 A 型血, 也可以输入 O 型血. 计算能为 A 型血输血的人所占的比例.

4.8.19. J: OO $^{(R)}$

父: AO $^{(a)}$ BO $^{(b)}$

母: AO $^{(c)}$ BO $^{(d)}$

$$\Pr(ac\,|\,R) = \frac{\Pr(R\,|\,ac)\cdot\Pr(ac)}{\Pr(R)}$$

$$= \frac{0.25\times(0.45\times0.4\times2)^2}{(0.45)^2}$$

$$= 0.16$$

4.8.20

B = 哥哥也是 O 型血.

A 为父母血型的情况集合.

$$Pr(B) = \sum_{i=1}^{n} Pr(B \cap A_n)$$

$$= \sum_{i=1}^{n} Pr(B \mid A_i) \cdot Pr(A_i)$$

$$= P(a) \cdot P(c) \cdot \frac{1}{4} + P(a) \cdot P(d) \cdot \frac{1}{4}$$

$$+ P(b) \cdot P(c) \cdot \frac{1}{4} + P(b) \cdot P(d) \cdot \frac{1}{4}$$

$$= \cdots$$

4.8.21

$$Pr = Pr(A) + Pr(O)$$

$$= 0.4 \times 0.45 \times 2 + 0.4 \times 0.4 + (0.45)^2$$

$$= \cdots$$