条件概率 (Conditional Probability)

September 18, 2023

回忆: 例 0.5

回到 Poker Hands 的例子:

- **Gaming**: a one-pair hand that draws 5 cards from 52 cards
- Q1: the count of a one-pair hand is less than point 12
  - $A = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $P(A) = \frac{4}{2{,}598{,}960}$
- Q2: a one-pair hand only contains two S cards

$$P(B) = \frac{\binom{13}{2}\binom{39}{3}}{\binom{52}{5}}$$

- Q3: 现有一副手牌, 如果其总点数小于 12, 那么仅包含两张 S 的概率是多少?
  - 事件 A 发生的条件下, 事件 B 发生的概率

回忆: 例 0.5 的解法

回到 Poker Hands 的例子:

- Q1 和 Q2 都是在整个样本空间上进行, 无任何额外因素或条件
- Q3 并非发生的整个样本空间上, 条件 A 对样本空间进行了限制
  - $\omega_1 = \{2S, 2H, 2C, 2D, 3S\}$
  - $\omega_2 = \{2S, 2H, 2C, 2D, 3H\}$
  - $\omega_3 = \{2S, 2H, 2C, 2D, 3C\}$
  - $\omega_4 = \{2S, 2H, 2C, 2D, 3D\}$
  - 事件 A 发生的条件下, 事件 B 发生的概率 $P(B \mid A) = 1/4$
- 一个随机事件发生的概率可能随着条件的改变而改变
  - **问题:** 事件 $A \to$ 事件 C : “the count of a one-pair hand is less than point 13”
  - $P(B \mid C) =?$

回忆: 例 0.5 的解法

回到 Poker Hands 的例子:

- Q1 和 Q2 都是在整个样本空间上进行, 无任何额外因素或条件
- Q3 并非发生的整个样本空间上, 条件 A 对样本空间进行了限制
  - $\omega_1 = \{2S, 2H, 2C, 2D, 3S\}$
  - $\omega_2 = \{2S, 2H, 2C, 2D, 3H\}$
  - $\omega_3 = \{2S, 2H, 2C, 2D, 3C\}$
  - $\omega_4 = \{2S, 2H, 2C, 2D, 3D\}$
  - 事件 A 发生的条件下, 事件 B 发生的概率 $P(B \mid A) = 1/4$
- 一个随机事件发生的概率可能随着条件的改变而改变
  - **问题:** 事件 $A \to$ 事件 C : “the count of a one-pair hand is less than point 13”
  - $P(B \mid C) = \frac{10}{28}$

条件概率

定义 0.7 设 (Ω, Σ, P) 是一个概率空间, 事件 $A \in \Sigma$ 且 $P(A) > 0$, 对于任意事件 $B \in \Sigma$, 称

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

为事件 A 发生的条件下的事件 B 发生的概率, 简称为条件概率 (Conditional Probability), 读作 “given A, B 的概率.”

Remarks:

- 条件概率是概率, 因此遵循非负性、规范性、可列可加性、容斥原理
- 任何事件的概率可看作必然事件的条件概率 $P(A) = P(A \mid \Omega)$
  - 相对地, 条件概率 $P(B \mid A)$ 也可以看作 A “缩小了” 空间上的某个概率
- 约定: 当我们提及条件概率 $P(B \mid A)$ 时, 默认 $P(A) > 0$.

条件概率的基本计算性质

- 非负性 $P(B \mid A) \geq 0$
- 规范性 $P(\Omega \mid A) = 1$ and $P(A \mid A) = 1$
- 可列可加性, if $B_i \cap B_j = \emptyset$ for any i, j, we have

$$P(\cup_{i=1}^{n} B_i \mid A) = \cup_{i=1}^{n} P(B_i \mid A)$$

- 容斥原理

$$P(B_1 \cup B_2 \mid A) = P(B_1 \mid A) + P(B_2 \mid A) - P(B_1 \cap B_2 \mid A)$$

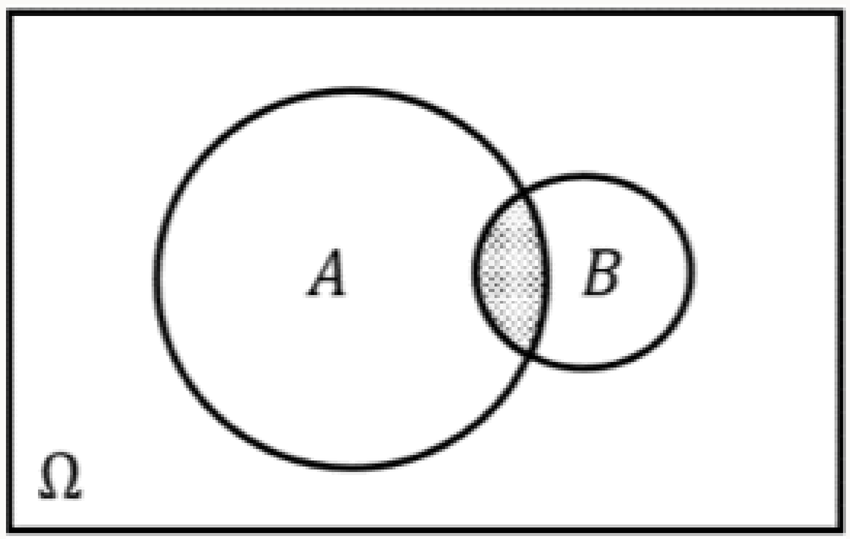
and

$$P(B \mid A) = 1 - P(\bar{B} \mid A)$$

条件概率的本质: 空间缩减

条件概率的本质是缩小了有效的样本空间

- 计算条件概率的方法: 样本空间缩减 $\Omega \to A$, 在 "新" 样本空间 A 下考虑事件 B 的发生



条件概率: 例 0.26

例 0.26 盒子中有 4 只不同的产品, 其中 3 只一等品, 1 只二等品. 从盒子中不放回随机取两次产品. 用 A 表示第一次拿到一等品的事件, 用 B 表示第二次取到一等品的事件, 求条件概率 $P(B \mid A)$.

条件概率: 例 0.27

例 0.27 一个箱子中有 a 个红球和 b 个白球, 依次任意无放回地取出 n 个球 $(n \le a+b)$, 其中包括 k 个白球 $(k \le b)$, 求在此情形下第一次取出白球的概率.

乘法公式

对概率非零的事件 A 和 B, 根据条件概率的定义

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

可得

$$P(AB) = P(A)P(B \mid A) = P(B)P(A \mid B)$$

进而, 我们可以推广到多事件的情况

定理 0.5 对随机事件 $A_1, A_2, \ldots, A_n$, 有

$$P(A_1A_2\ldots A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\ldots P(A_n \mid A_1A_2\ldots A_{n-1})$$

乘法公式: 例 0.28

例 0.28 有 n 把钥匙, 只有一把能打开门, 不放回地随机取出一把开门, 求第 $k(k \leq n)$ 次打开门的概率.

乘法公式: 例 0.29

例 0.29 (匹配问题) 假设有 n 对夫妻参加活动, 被随机分成 n 组, 每组一男一女, 求

- n 对夫妻中, 恰好两两被分到一组的概率
- n 对夫妻中, 至少有一对夫妻被分到一组的概率

乘法公式: 例 0.30

例 0.30 箱子里有 n 个白球和 m 个红球, 随机取出一球后放回, 并加入 l 个与取出球同色的球, 求前两次取出红球、后两次取出白球的概率.

乘法公式: 例 0.31

例 0.31 第一个箱子里有 n 个不同的白球, 第二个箱子里有有 m 个不同的红球, 从第一个箱子任意拿走一球, 再从第二个箱子里任意拿走一球放入第一个箱子, 依次进行, 直至第一、第二个箱子都为空, 求第一个箱子最后一次取走的球是白球的概率.

解答: 例 0.31

问题: 第一个箱子里有 n 个不同的白球, 第二个箱子里有有 m 个不同的红球, 从第一个箱子任意拿走一球, 再从第二个箱子里任意拿走一球放入第一个箱子, 依次进行, 直至第一、第二个箱子都为空, 求第一个箱子最后一次取走的球是白球的概率.

解答:

- 我们注意到, 这个问题里球是不同的, 也就是第一个箱子最后一次取走的白球也是不一样的.
- 球具有唯一识别性, 所以我们先将这个白球选出来 $\binom{n}{1}$ 标记为 a.
- 用 A 表示 "第一个箱子最后一次取走的白球 a" , 求 $P(A)$.
- 这同样说明, 前 $n+m-1$ 次取球没有取到 a.
- 类似于例 0.30, 我们可以用序列来描述事件 A: $B_1 B_2 \ldots B_{n+m-1}$, 其中 B_i 表示第 i 未取到 a 球.
- 我们有 $P(A) = P(B_1 B_2 \ldots B_{n+m-1})$, 遵循乘法公式.

事件独立性 (Event Independency)

September 18, 2023

两事件的独立性

- 在一般情况下, 由条件概率定义知

$$P(B \mid A) = \frac{P(AB)}{P(A)} \neq P(B)$$

即事件 (原因) A 发生对事件 B 的发生有影响.

- 然而在有些情况下, 事件 A 的发生对事件 B 的发生可能没有任何影响, 即

$$P(B \mid A) = P(B) \quad \text{or} \quad P(AB) = P(A)P(B) \qquad \text{独立性}$$

独立性的定义

定义 0.8 若事件 A, B 满足

$$P(AB) = P(A)P(B)$$

称事件 A 与 B 相互独立.

独立性的性质:

- 任何事件与不可能事件 (或必然事件) 相互独立
- 若 $P(A)P(B) > 0$, 则有

$$P(AB) = P(A)P(B) \Leftrightarrow P(B \mid A) = P(B) \Leftrightarrow P(A \mid B) = P(A)$$

- 若事件 A 与 B 相互独立, 则 A 与 $\bar{B}$, $\bar{A}$ 与 B, $\bar{A}$ 与 $\bar{B}$ 都互相独立.

独立性 VS 互斥 (不相容)

- A 与 B **相互独立:** 与概率相关, 反映事件的概率属性

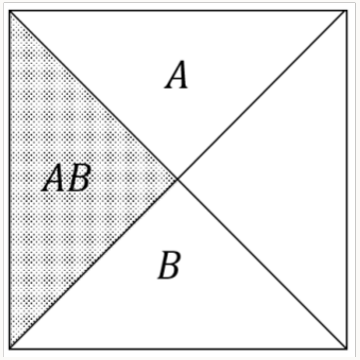
$$P(AB) = P(A)P(B)$$

- A 与 B **互不相容:** 与概率无关, 与事件的运算相关

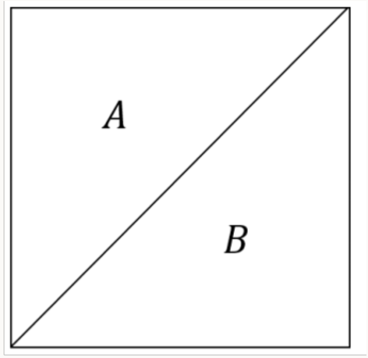
$$A \cap B = \emptyset \Rightarrow P(A \cap B) = 0$$

- A 和 B 独立 $\Rightarrow$ A 和 B 不互斥; A 和 B 互斥 $\Rightarrow$ A 和 B 不独立

A与B独立,
但并不互斥



A与B互斥,
但并不独立



思考: 例子

例 **(1)** 若事件 A 和 B 互斥且 $P(A)P(B) > 0$, 下面哪些说法正确?

1. $P(B \mid A) > 0$
2. $P(A \mid B) = 0$
3. A 和 B 不独立
4. $P(A \mid B) = P(A)$

例 **(2)** 若事件 A 和 B 独立且 $P(A)P(B) > 0$, 下面哪些说法正确?

1. $P(B \mid A) > 0$
2. $P(A \mid B) = P(A)$
3. $P(A \mid B) = 0$
4. $P(AB) = P(A)P(B)$

如何判断独立性?

1. 直接计算判断 $P(AB) \stackrel{?}{=} P(A)P(B)$

2. 根据实际问题判断事件的独立性

- 两人独立射击打靶且互不影响, 因此两人中靶的事件相互独立
- 抛投质量不均匀的硬币
- 从 n 件产品中随机抽取两件, 事件 A_i 表示第 i 件是合格品. 若有放回抽取则事件 A_i 与 A_j 相互独立; 若不放回则不独立
- 从一副扑克 (不含大王、小王) 中随机抽取一张扑克, 用事件 A 表示抽到 10, 事件 B 表示抽到黑色的扑克. 事件 A 与 B 是否独立?
- 机器学习的经典假设是训练数据独立同分布采样 (i.i.d.)

条件独立性

定义 0.9 设 (Ω, Σ, P) 是一个概率空间, 事件 $C \in \Sigma$ 且满足 $P(C) > 0$, 若事件 $A, B \in \Sigma$ 且满足

$$P(AB \mid C) = P(A \mid C)P(B \mid C)$$

或者等价地

$$P(A \mid BC) = P(A \mid C)$$

则称事件 A 和 B 在事件 C 发生的情况下是条件独立的 **(conditional independent)**.

三事件的独立性

定义 0.10 若事件 A, B, C 满足:

- $P(AB) = P(A)P(B), P(AC) = P(A)P(C), P(BC) = P(B)P(C)$
- $P(ABC) = P(A)P(B)P(C)$

则称事件 A、B、C 是相互独立的 **(mutually independent)**.

Question:

$$\text{事件}A, B, C\text{ 相互独立} \qquad \subseteq \quad \stackrel{?}{=} \quad \supseteq \qquad \text{事件}A, B, C\text{ 两两独立}$$

Bernstein 反例

例 0.32 (Bernstein 反例) 一个均匀的正四面体, 第一面红色, 第二面白色, 第三面黑色, 第四面同时有红、白、黑三种颜色. 随意投掷一次, 用 A, B, C 分别表示红色、白色、黑色朝下的事件.

考虑这三事件的相互独立性与两两独立性.

多事件的独立性

定义 0.11 若事件 $A_1, A_2, \ldots, A_n$ 中任意 k $(k \leq n)$ 个事件独立, 即对任意 $k \in [n]$ 有

$$P(A_{i_1} \ldots A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$$

其中 $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n$. 则称事件 $A_1, A_2, \ldots, A_n$ 相互独立.

Notice:

- n 个事件的相互独立性共有 $2^n - n - 1$ 个等式 (思考题)
- 事件 $A_1, A_2, \ldots, A_n$ 的相互独立性与两两独立性存在区别
- 可以类似定义多个事件的条件独立性

独立性: 例 0.33

例 0.33 三人独立破译一份密码, 每人单独能破译的概率分别为

$$1/5, \quad 1/3, \quad 1/4$$

问三人中至少有一人能破译密码的概率.

增加事件数量可以减小极端事件发生的概率.

小概率原理 (Minor Probability Principle)

若 n 个事件 $A_1, A_2, \ldots, A_n$ 相互独立, 其发生的概率分别为 $p_1, p_2, \ldots, p_n$

- 事件 $A_1, A_2, \ldots, A_n$ 中至少有一个事件发生的概率为

$$P(\cup_{i=1}^n A_i) = 1 - P(\cap_{i=1}^n \bar{A}_i) = 1 - \prod_{i=1}^n (1 - p_i)$$

- 事件 $A_1, A_2, \ldots, A_n$ 中至少有一个事件不发生的概率为

$$P(\cup_{i=1}^n \bar{A}_i) = 1 - P(\cap_{i=1}^n A_i) = 1 - \prod_{i=1}^n p_i$$

解读:

- 若每个事件的概率 p_i 都非常小, 但 n 非常大, 则 n 个相互独立的事件中至少有一事件发生或至少有一事件不发生的概率都很大.
- 若事件 A 在一次试验中发生的概率非常小, 但经过多次独立地重复试验, 事件 A 的发生是必然的, 称之为小概率原理.

小概率原理: 例 0.34

例 0.34 冷战时期美国的导弹精度 90%, 苏联的导弹精度 70%, 但苏联的导弹数量特别多, 导弹的数量能否弥补精度的不足?

小概率原理: 例 0.35

例 0.35 假设市场上有 n 种不同类型的邮票, 一位集邮爱好者收集第 i 种邮票的概率为 p_i, 且

$$p_1 + p_2 + \cdots + p_n = 1$$

假设每次集邮都是独立同分布的 (i.i.d.), 若现已收集到 n 张邮票, 用 A_i 表示至少收集到第 i 种类型邮票的事件.

请计算:

- $P(A_i)$
- $P(A_i \cup A_j)$ $(i \neq j)$
- $P(A_i \mid A_j)$ $(i \neq j)$