例 2.12 已知事件A为病人被诊断为肝癌, 事件C为病人患有肝癌, $P(A|C) = 0.95$, $P(\overline{A}|\overline{C}) = 0.9$, $P(C) = 0.0004$. 求$P(C|A)$.

上面的例子仅作为课堂练习题, 这里不再讲解.

例 2.13 (三囚徒问题) 三犯人 a, b, c 均被判为死刑, 法官随机赦免其中一人, 看守知道谁被赦免但不会说. 犯人 a 问看守: b 和 c 谁会被执行死刑? 看守的策略: i) 若赦免 b, 则说 c; ii) 若赦免 c, 则说 b; iii) 若赦免 a, 则以 $1/2$ 的概率说 b 或 c. 看守回答 a: 犯人 b 会被执行死刑. 犯人 a 兴奋不已, 因为自己生存的概率为 $1/2$. 犯人 a 将此事告诉犯人 c, c 同样高兴, 因为他觉得自己的生存几率为 $2/3$. 那么谁错了?

解 用事件 A, B, C 分别表示犯人 a, b, c 被赦免, 由题意可知

$$P(A) = P(B) = P(C) = 1/3.$$

用事件 D 表示看守人说犯人 b 被执行死刑, 则有

$$P(D|A) = 1/2 \quad P(D|B) = 0 \quad P(D|C) = 1.$$

由全概率公式有

$$P(D) = P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C) = 1/2.$$

由贝叶斯公式有

$$P(A|D) = \frac{P(A)P(D|A)}{P(D)} = \frac{1}{3} \quad P(C|D) = \frac{P(C)P(D|C)}{P(D)} = \frac{2}{3}$$

所以犯人 a 的推断不正确, 犯人 c 的推断正确.

与三囚徒类似的问题是如下三门问题, 这里仅给出问题的描述, 求解方案与上面类似.

例 2.14 (三门问题) 在一电视节目中, 参赛者看到三扇关闭的门, 已知一门后面是汽车, 其它两门后面是山羊, 选中什么则获得什么, 主持人知道三门后有什么. 当参赛者选定一扇门但未开启, 此时节目主持人则开启剩下有山羊的一扇门. 问题: 若参赛者允许重新选择, 是否换一扇门?

2.2 独立性

在一般情况下, 由条件概率定义知

$$P(B|A) = P(AB)/P(A) \neq P(B),$$

即事件 A 发生对事件 B 的发生有影响. 然而在很多情况下, 事件 A 的发生对事件 B 的发生可能没有任何影响, 这是本节研究的事件独立性.

2.2.1 两事件的独立性

定义 2.3 若事件 A, B 满足 $P(AB) = P(A)P(B)$, 则称 **事件 A 与 B 相互独立**.

根据上面的定义可知, 对事件 A 和 B 满足 $P(A)P(B) > 0$, 有

$$P(AB) = P(A)P(B) \Leftrightarrow P(B|A) = P(B) \Leftrightarrow P(A|B) = P(A).$$

根据定义还可以发现任何事件与不可能事件 (或必然事件) 相互独立.

性质 2.4 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.

证明 根据事件差公式 $P(A-B) = P(A) - P(AB)$ 有

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

同理可证 $P(\bar{A}B) = P(\bar{A})P(B)$. 利用容斥原理有

$$\begin{aligned} P(\bar{A}\bar{B}) &= 1 - P(A \cup B) = 1 - P(A) - P(B) + P(AB) \\ &= 1 - P(A) - P(B) + P(A)P(B) = (1-P(A))(1-P(B)) = P(\bar{A})P(\bar{B}), \end{aligned}$$

从而完成证明.

如何判断事件的独立性? 根据定义直接计算进行判断:

例 2.15 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?

解 根据问题可知一副扑克 (不含大王、小王) 52张, 黑色扑克 26 张, 4 张 10, 因此

$$P(A) = 4/52 = 1/13, \qquad P(B) = 1/2.$$

另一方面有

$$P(AB) = 2/52 = 1/26 = P(A)P(B),$$

由此事件 A 和 B 相互独立.

也可以根据实际问题判断事件的独立性, 例如

- 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;
- 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_1 与 A_2 相互独立; 若不放回则不独立;

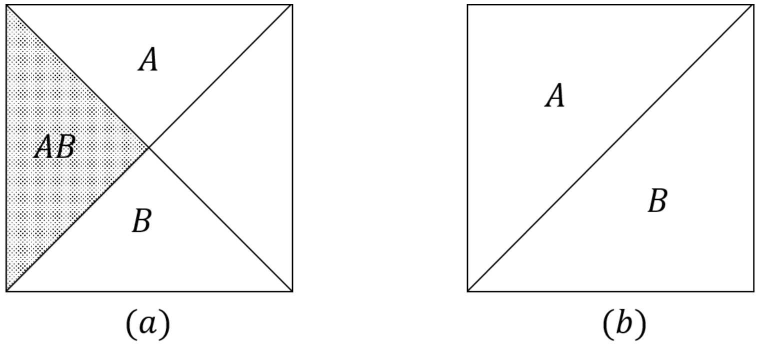


图 2.1 假设落入正方形z 个点的可能性完全相同, 即几何概型

- 机器学习的经典假设是训练数据独立同分布采样.

现在我们讨论独立性与互不相容性 (互斥性) 之间的关系: 事件 A 和 B 独立, 根据定义可知 $P(AB) = P(A)P(B)$, 独立性与概率相关, 反映事件的概率属性;

若事件 A 和 B 互不相容, 根据定义有 $AB = \emptyset$, 互斥性与事件的运算关系相关, 与概率无关. 因此独立与互斥反映事件不同的性质, 无必然联系.

如图 2.1(a) 所示: 事件 A 和 B 独立并不意味着事件 A 和 B 互斥, 如图 2.1(b) 所示: 事件 A 和 B 互斥并不意味着事件 A 和 B 独立. 我们进一步有

性质 2.5 事件 A 和 B 满足 $P(A)P(B) > 0$, 若事件 A 和 B 独立则 A 和 B 不互斥; 若事件 A 和 B 互斥则 A 和 B 不独立.

证明 若事件 A 和 B 独立且 $P(A)P(B) > 0$, 有

$$P(AB) = P(A)P(B) > 0$$

事件 A 和 B 不互斥; 另一方面, 若事件 A 和 B 互斥且 $P(A)P(B) > 0$, 有

$$P(AB) = 0 \neq P(A)P(B)$$

事件 A 和 B 不独立.

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

a) $P(B|A) > 0$, b) $P(A|B) = 0$, c) A, B不独立, d) $P(A|B) = P(A)$.

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

a) $P(B|A) > 0$, b) $P(A|B) = P(A)$, c) $P(A|B) = 0$, d) $P(AB) = P(A)P(B)$.

2.2.2 多个事件的独立性

定义 2.4 若事件 A, B, C 满足 $P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, $P(BC) = P(B)P(C)$ 且 $P(ABC) = P(A)P(B)P(C)$, 则称 **事件 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 相互独立**.

根据定义可知: 事件 A, B, C 相互独立和事件 A, B, C 的两两独立不同, 由事件 A, B, C 相互独立可知事件 A, B, C 两两独立; 反之不一定成立, 还需满足 $P(ABC) = P(A)P(B)P(C)$.

下面定义 n 个事件的独立性:

定义 2.5 若事件 $A_1, A_2, \cdots, A_n$ 中任意 k 个事件独立, 即对任意 $k \in [n]$ 有

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

其中 $1 \leqslant i_1 \leqslant i_2 \leqslant \cdots i_k \leqslant n$, 则称 **事件 $\mathbf{A_1}, \mathbf{A_2}, \cdots, \mathbf{A_n}$ 相互独立**.

同样需要注意 n 个事件的相互独立性与两两独立性的区别. 下面来看一个独立性的例子.

例 2.16 三人独立破译一份密码, 每人单独能破译的概率分别为 $1/5, 1/3, 1/4$, 问三人中至少有一人能破译密码的概率.

解 用事件 A_i 表示第 i 个人破译密码 ($i \in [3]$), 根据题意有

$$P(A_1) = 1/5, \qquad P(A_2) = 1/3, \qquad P(A_3) = 1/4.$$

根据容斥原理和独立性, 三人中至少有一人能破译密码的概率为

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3) \\ &= \frac{1}{5} + \frac{1}{4} + \frac{1}{3} - \frac{1}{20} - \frac{1}{15} - \frac{1}{12} + \frac{1}{60} = 0.6 \end{aligned}$$

我们也可以根据对偶性和独立性来求解该问题, 三人中至少有一人能破译密码的概率为

$$P(A_1 \cup A_2 \cup A_3) = 1 - P(\bar{A}_1\bar{A}_2 \cdots \bar{A}_n) = 1 - P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3) = 1 - \frac{4}{5} \cdot \frac{2}{3} \cdot \frac{3}{4} = 0.6.$$

从上例可知: 尽管每个人能破译密码的概率都小于 $1/2$, 但三人独立进行破译, 则至少有一人破译密码的概率则为 $2/3$, 由此提高了破译密码的概率. 我们可以将类似问题推广到更一般的情况.

若 n 个事件 $A_1, A_2, \cdots, A_n$ 相互独立, 以及其发生的概率分别为 $p_1, p_2, \cdots, p_n$, 则事件 $A_1, A_2, \cdots, A_n$ 中至少有一事件发生的概率为

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - P(\bar{A}_1\bar{A}_2 \cdots \bar{A}_n) = 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n);$$

此外, 事件 $A_1, A_2, \cdots, A_n$ 中至少有一事件不发生的概率为

$$P(\bar{A}_1 \cup \bar{A}_2 \cup \cdots \cup \bar{A}_n) = 1 - P(A_1A_2 \cdots A_n) = 1 - p_1p_2 \cdots p_n.$$

由此可知: 尽管每个事件发生的概率 p_i 都非常小, 但若 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” 的概率可能很大.

定义 2.6 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**.

小概率原理可通过严格的数学证明得到: 若事件 $A_1, A_2, \cdots, A_n, \cdots$ 独立且每事件发生的概率 $P(A_i) = p > 0$ 非常小, 则有

$$P(A_1 A_2 \cdots A_n) = 1 - P(\bar{A}_1 \bar{A}_2 \cdots \bar{A}_n) = 1 - (1-p)^n \to 1 \quad \text{当} \quad n \to \infty,$$

即独立重复多次的小概率事件亦可成立必然事件.

还可以进一步研究: 若独立事件 $A_1, A_2, \cdots, A_n$ 发生的概率 $P(A_i) = p \ (i \in [n])$, 则 n 个事件中恰有 k 个事件发生的概率为 $\binom{n}{k} p^k (1-p)^{n-k}$.

例 2.17 冷战时期美国的导弹精度 99%, 苏联的导弹精度 60%, 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?

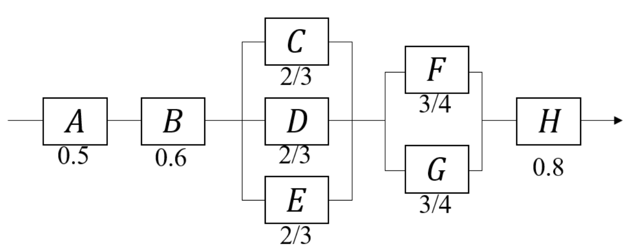
解 假设每次独立发射 n 枚导弹, 用事件 A_i 表示第 i 枚导弹命中目标, 则 n 枚导弹击中目标的概率为

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - (1-0.6)^n \geqslant 0.99 \quad \Rightarrow \quad n \geqslant 5,$$

因此每次独立发射 5 枚导弹, 击中目标的概率高于 99%.

在上例中, 若美国的导弹精度为 90%, 苏联的导弹精度为 70%, 则苏联每次只需独立发射两枚导弹即可达到 91%.

例 2.18 一串电路如下图所示: A, B, C, D, E, F, G 是电路元件, 电路元件各自下方的数字表示正常工作的概率. 若各电路元件之间相互独立. 求电路正常工作的概率.



解 用事件 W 表示电路正常工作, 则有 $W = A \cap B \cap (C \cup D \cup E) \cap (F \cup G) \cap H$. 根据独立性假设有

$$P(W) = P(A)P(B)P(C \cup D \cup E)P(F \cup G)P(H).$$

根据 $P(C \cup D \cup E) = 1 - P(\bar{C})P(\bar{D})P(\bar{E}) = 1 - (2/3)^3 = 19/27$ 和 $P(F \cup G) = 1 - P(\bar{E})P(\bar{G}) = 7/16$, 可得 $P(W) = 133/1800$.

2.3 案例分析

本节研究的问题: 给定矩阵 $A, B, C \in \{0,1\}^{n\times n}$ (n 非常大, 如 $n \geqslant 10000000$), 验证 $AB = C$ 是否成立? 若直接执行矩阵乘法运算、并验证等式是否成立, 计算复杂度为 $O(n^3)$; 若采用分治法, 计算复杂度为 $O(n^{\log_2 7})$, 目前最好的计算复杂度为 $O(n^{2.37})$. 为进一步降低计算复杂度, 可利用独立性验证 $AB = C$ 是否成立?

独立随机产生一个向量 $r \in \{0,1\}^n$, 判断

$$A(Br) = Cr?$$

计算 $A(Br)$ 和 Cr 的复杂度均为 $O(n^2)$. 若 $A(Br) \neq Cr$ 则直接可得 $AB \neq C$; 若 $A(Br) = Cr$ 并不能得出 $AB = C$. 将上述过程独立进行 K 次, 可以证明以较大的概率有 $AB = C$ 成立, 该过程被称为 Freivalds 算法.

- -

Freivalds 算法

- -

Input: A, B, C

Output: Yes/No

- -

For $i = 1 : K$

    Select a random vector $r = (r_1, r_2, \cdots, r_n)$ with $P(r_j = 0) = P(r_j = 1) = 1/2$ ($j \in [n]$)

    Compute $p = A \times (Br) - Cr$

    If $p \neq 0$ then

        Return ‘No’.

    EndIf

EndFor

Return ‘Yes’.

- -

首先发现该算法的计算复杂度为 $O(Kn^2)$, 若 K 比较小则显著降低了计算复杂度. 进一步研究算法的有效性, 若返回 ‘No’, 则必然有 $AB \neq C$; 若返回 ‘Yes’, 然而并不一定有 $AB = C$ 成立, 下面研究当算法返回 ‘Yes’ 时 $AB = C$ 成立的概率.

设 $D = AB - C \neq 0$, 则 D 中必存在一些元素不为 0, 不妨令 $d_{11} \neq 0$. 对任意一轮循环, 不妨设随机向量 $r = \{r_1, r_2, \cdots, r_n\}$, 根据返回 ‘Yes’ 可知 $Dr = 0$, 进一步可得向量 Dr 的第一个元素等于

0, 即

$$\sum_{j=1}^{n} d_{1j} r_j = 0 \quad \Longrightarrow \quad r_1 = -\frac{1}{d_{11}} \sum_{j=2}^{n} d_{1j} r_j$$

无论 $r_2, \ldots, r_n$ 取何值, 等式 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 是否成立由 r_1 的值决定. 根据 $P(r_1 = 0) = P(r_1 = 1) = 1/2$ 可知 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 成立的概率不超过 1/2. 因此在 K 轮独立的循环中, 等式 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 成立的概率不超过 $1/2^K$.

取 $K = \log_2 n$, 则算法 Freivalds 计算复杂度为 $O(n^2 \log n)$, 若算法返回 'No', 则 $AB \neq C$; 若返回 'Yes', 则有

$$P(AB = C) > 1 - 1/n,$$

即至少以 $1 - 1/n$ 的概率有 $AB = C$ 成立.