第 2 章　条件概率与独立性

前面关于事件概率的讨论都是在整个样本空间上进行, 不需考虑其它条件或限制因素. 然而在很多实际问题中, 我们往往关心随机事件在一定附加信息 (条件) 下发生的概率, 即条件概率. 它是概率论中一个非常重要且实用的概念. 可以帮助我们更好地分析和理解复杂的随机事件, 同时也有助于简化复杂事件概率的计算.

2.1 条件概率

2.1.1 条件概率

首先来看一个例子, 随意投掷一枚骰子观察点数, 其样本空间 $\Omega = \{1, 2, \cdots, 6\}$. 用 A 表示观察到奇数点的事件, 则事件 $A = \{1, 3, 5\}$, 根据古典概型有 $P(A) = 1/2$. 用 B 表示观察到 3 点的事件, 根据古典概型有 $P(B) = 1/6$.

现在考虑在事件 A 发生的情况下事件 B 发生的概率, 记为 $P(B|A)$. 由于 $A = \{1, 3, 5\}$ 且每种结果等可能发生, 由此可得条件概率

$$P(B|A) = 1/3 > P(B).$$

用事件 C 表示观察到 2 点, 根据古典概型有 $P(C) = 1/6$, 但在事件 A 发生的情况下事件 C 不可能发生, 因此条件概率

$$P(C|A) = 0 < P(C).$$

由此可知一个随机事件发生的概率可能随着条件的改变而改变, 同时通过观察可以发现

$$P(B|A) = 1/3 = \frac{P(AB)}{P(A)} \quad \text{和} \quad P(C|A) = 0 = \frac{P(AC)}{P(A)}.$$

针对一般情形, 我们将上述关系作为条件概率的定义.

定义 2.1　设 (Ω, Σ, P) 是一个概率空间, 随机事件 $A \in \Sigma$ 且 $P(A) > 0$. 对任意事件 $B \in \Sigma$, 称

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

为事件 A 发生的条件下事件 B 发生的概率, 简称 **条件概率** (conditional probability).

对任意事件 $A \in \Sigma$ 有 $P(A) = P(A\Omega)/P(\Omega) = P(A|\Omega)$ 成立, 因而任何随机事件的概率可以看作必然事件下的条件概率. 根据条件概率的定义有 $P(AB) = P(A)P(B|A)$. 在本书后续章节中, 若出现条件概率 $P(B|A)$, 一般都默认 $P(A) > 0$.

根据条件概率的定义, 不难验证条件概率 $P(\cdot|A)$ 具有以下一些基本性质:

1° **非负性**: 对任意事件 B 有 $P(B|A) \geqslant 0$;

2° **规范性**: 对样本空间 Ω 有 $P(\Omega|A) = 1$;

3° **可列可加性**: 若 $B_1, B_2, \ldots, B_n, \ldots$ 是可列无穷个互不相容的事件, 即 $B_iB_j = \emptyset$ $(i \neq j)$, 有

$$P(B_1 \cup B_2 \cup \cdots \cup B_n \cup \cdots |A) = P(B_1|A) + P(B_2|A) + \cdots + P(B_n|A) + \cdots.$$

从概率和条件概率的定义可知 $P(B|A) = P(AB)/P(B) \geqslant 0$ 以及 $P(\Omega|A) = P(A\Omega)/P(A) = 1$, 由此验证公理 1° 和公理 2°. 若可列个事件 $B_1, B_2, \cdots, B_n, \cdots$ 是两两互不相容的, 则可列个事件 $AB_1, AB_2, \cdots, AB_n, \cdots$ 也是两两互不相容的, 根据分配律有

$$P\left(\bigcup_{i=1}^{\infty} B_i \middle| A\right) = \frac{P\left(A(\bigcup_{i=1}^{\infty} B_i)\right)}{P(A)} = \frac{P(\bigcup_{i=1}^{\infty} AB_i)}{P(A)} = \sum_{i=1}^{\infty} \frac{P(AB_i)}{P(A)} = \sum_{i=1}^{\infty} P(B_i|A)$$

由此可知公理 3° 成立. 由于条件概率满足概率的三条公理, 因此条件概率 $P(\cdot|A)$ 仍然是一种概率.

性质 2.1 (容斥原理) 对随机事件 A, B_1 和 B_2 且满足 $P(A) > 0$, 有

$$P(B_1 \cup B_2|A) = P(B_1|A) + P(B_2|A) - P(B_1B_2|A).$$

证明 由条件概率的定义有

$$P(B_1 \cup B_2|A) = P((B_1 \cup B_2) \cap A)/P(A).$$

再根据随机事件的分配律和容斥原理有

$$P((B_1 \cup B_2) \cap A) = P(AB_1 \cup AB_2) = P(AB_1) + P(AB_2) - P(AB_1B_2),$$

上式两边同时除以 $P(A)$ 即可完成证明.

性质 2.2 对随机事件 A 和 B 且满足 $P(A) > 0$, 有 $P(B|A) = 1 - P(\bar{B}|A)$.

证明 根据容斥原理有

$$1 = P(\Omega|A) = P(B \cup \bar{B}|A) = P(B|A) + P(\bar{B}|A) - P(B\bar{B}|A)$$

再根据事件 B 和 $\bar{B}$ 互不相容有 $P(B\bar{B}|A) = 0$, 从而完成证明.

事件 A 发生的条件下事件 B 发生的概率, 可以将 A 看作新的样本空间、而忽略以前的样本空间 Ω, 由此可以发现: **条件概率的本质是缩小了有效的样本空间**. 这也为计算条件概率提供了一种方法: 空间缩减法, 在新的样本空间 A 下考虑事件 B 发生的概率.

例 2.1 盒子中有 4 个不同的产品, 其中 3 个一等品, 1 个二等品. 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, B 表示第二次取到一等品的事件, 求条件概率 $P(B|A)$.

解 将盒子中 3 个一等产品分别编号为 $1,2,3$, 二等品编号4. 用 i 和 j 分别表示第一、二次抽取的产品的编号, 由此可得

$$\Omega=\{(i,j)\colon i\neq j, i,j\in[4]\}, \qquad A\ =\{(i,j)\colon i\neq j, i\neq 4\},$$
$$B=\{(i,j)\colon i\neq j, j\neq 4\}, \qquad AB=\{(i,j), i\neq j, i,j\in[3]\}.$$

计算可得 $|\Omega|=12$, $|A|=9$, $|B|=9$ 以及 $AB=6$. 根据古典概型有

$$P(A)=\frac{|A|}{|\Omega|}=\frac{3}{4}, \quad P(B|A)=\frac{P(AB)}{P(A)}=\frac{1/2}{3/4}=\frac{2}{3}.$$

也可以采用 **样本空间缩减法** 来求解此问题: 当事件 A 发生后, 剩下 2 只一等品, 1 只二等品, 因此直接得到 $P(B|A)=2/3$.

例 2.2 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 ($n\leqslant a+b$), 其中包括 k 个白球 ($k\leqslant b$), 求在此情形下第一次取出白球的概率.

解 用 A 表示第一次取出白球的事件, 用 B 表示依次取出 n 个球中包括 k 个白球的事件. 根据题意和超几何分布有

$$P(A)=\frac{b}{a+b}, \quad P(B)=\binom{a}{n-k}\binom{b}{k}\Big/\binom{a+b}{n}, \quad P(B|A)=\binom{a}{n-k}\binom{b-1}{k-1}\Big/\binom{a+b-1}{n-1}.$$

所求条件概率

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(A)P(B|A)}{P(B)}=\frac{k}{n}.$$

也可以采用 **样本空间缩减法** 来求解此问题: 在事件 B 发生的情况下, 即选中的 n 个球中有 k 个白球, 由于任何一球被第一次选中的可能性一样, 因此事件 A 发生的概率为 k/n.

2.1.2 乘法公式

随机事件 A 和 B 满足 $P(A)>0, P(B)>0$, 根据条件概率的定义可知

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

将上式进一步推广, 根据条件概率的定义有下面的乘法公式:

定理 2.1 设 (Ω,Σ,P) 是一个概率空间, 若随机事件 $A_1,A_2,\dots,A_n\in\Sigma$, 且满足条件 $P(A_1A_2\cdots A_{n-1})>0$, 则有

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

例 2.3 假设一批灯泡有 100 只, 其中有次品 10 只, 其余为正品. 不放回抽取地每次抽取一只, 求第三次才是正品的概率.

解 用 A_i 表示第 i 次抽到正品的事件 ($i \in [3]$), 事件 B 表示第 3 次才抽到的正品, 则有 $B = \bar{A}_1\bar{A}_2A_3$. 根据乘法公式有

$$P(B) = P(\bar{A}_1\bar{A}_2\ A_3) = P(\bar{A}_1)P(\bar{A}_2|\bar{A}_1)P(A_3|\bar{A}_1\bar{A}_2) = \frac{10}{100} \times \frac{9}{99} \times \frac{90}{98} = \frac{9}{1078}.$$

例 2.4 设 n 把钥匙中只有一把能打开门. 不放回随机取出一把开门, 求第 k 次打开门的概率.

解 用 A_i 表示第 i 次没有打开门的事件, 则第 k 次打开门的事件可表示为 $A_1A_2\cdots A_{k-1}\bar{A}_k$, 根据乘法公式有

$$\begin{aligned}
&P(A_1A_2\cdots A_{k-1}\bar{A}_k)\\
= \ & P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_{k-1}|A_1\cdots A_{k-2})P(\bar{A}_k|A_1A_2\cdots A_k)\\
= \ & \frac{n-1}{n} \times \frac{n-2}{n-1} \times \cdots \times \frac{n-(k-1)}{n-(k-2)} \times \frac{1}{n-(k-1)} = \frac{1}{n}
\end{aligned}$$

也可以根据 **抽签原理** 来求解该问题: 第 k 次打开门的概率与 k 无关, 每次打开门的概率相同, 共 n 把钥匙, 因此第 k 次打开门的概率为 $1/n$.

例 2.5 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 对夫妻恰好两两被分到一组的概率.

解 用 A_i 表示第 i 对夫妻被分到同一组的事件, 则 n 对夫妻恰好两两被分到一组的事件可表示为 $A_1A_2\cdots A_n$. 根据乘法公式有

$$\begin{aligned}
&P(A_1A_2\cdots A_n)\\
= \ & P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})\\
= \ & \frac{1}{n} \times \frac{1}{n-1} \times \cdots \times \frac{1}{1} = \frac{1}{n!}.
\end{aligned}$$

例 2.6 第一个箱子里有 n 个不同的白球, 第二个箱子里有 m 个不同的红球, 从第一个箱子任意取走一球, 再从第二个箱子里任意取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 假设第一个箱子里的白球分别标号为 $1, 2, \cdots, n$, 用 A_i 表示第一个箱子最后取走的是第 i 号白球的事件. 由此可知事件 $A_1, A_2, \cdots, A_n$ 是两两互不相容的, 且第一个箱子最后一次取走的球是白球的事件可表示为 $A_1 \cup A_2 \cup \cdots \cup A_n$, 根据事件的对称性可得其概率为

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) = nP(A_1).$$

若事件 A_1 发生, 则从第一个箱子中取走的 $m+n-1$ 个球均不是第 1 号白球, 用事件 B_j 表示第 j 次从第一个箱子里取走的球不是第 1 号白球, 即 $A_1 = B_1B_2\cdots B_{m+n-1}$. 根据乘法公式有

$$\begin{aligned} P(A_1) &= P(B_1)P(B_2|B_1)\cdots P(B_m|B_1B_2\cdots B_{m-1}) \times P(B_{m+1}B_{m+2}\cdots B_{m+n-1}|B_1B_2\cdots B_m) \\ &= \left(1-\frac{1}{n}\right)^m P(B_{m+1}B_{m+2}\cdots B_{m+n-1}|B_1B_2\cdots B_m) = \left(1-\frac{1}{n}\right)^m \times \frac{1}{n}. \end{aligned}$$

由此可知第一个箱子最后一次取走的球是白球的概率为 $(1-1/n)^m$.

例 2.7 假设箱子里有 m 个红球和 n 个白球, 现随机取出一球后放回, 并加入 c 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.

解 用 A_i 表示第 i 次抽到红球的事件 $(i \in [2])$, 事件 B_i 表示第 i 次抽到红球的事件 $(i = 3, 4)$, 我们有

$$P(A_1) = \frac{m}{m+n}, \qquad P(A_2|A_1) = \frac{m+c}{m+n+c},$$
$$P(B_1|A_1A_2) = \frac{n}{m+n+2c}, \qquad P(B_2|A_1A_2B_1) = \frac{n+c}{m+n+3c}.$$

根据乘法公式有

$$P(A_1A_2B_1B_2) = \frac{mn(m+c)(n+c)}{(m+n)(m+n+c)(m+n+2c)(m+n+3c)}$$

上述例子可用来作为疾病传染的粗略解释, 每取出一球代表疾病的一次传染, 每次传染将增加再传染的可能性.

2.2 全概率公式和贝叶斯公式

本节介绍概率计算中两个重要的公式: 全概率公式和贝叶斯公式.

2.2.1 全概率公式

全概率公式是概率论中最基本的公式之一, 将一个复杂事件的概率计算分解为若干简单事件的概率计算. 具体而言, 将一个复杂事件分解为若干不相容的简单事件之和, 通过分别计算简单事件的概率, 利用概率的可加性得到复杂事件的概率. 首先定义样本空间的一个分割.

定义 2.2 设 $A_1, A_2, \cdots, A_n$ 是样本空间 Ω 的一组事件, 若满足:

i) 任意两个事件是互不相容性的, 即 $A_i \cap A_j = \emptyset$ $(i \neq j)$;

ii) 完备性 $\Omega = A_1 \cup A_2 \cup \cdots \cup A_n$,

则称事件 $A_1, A_2, \cdots, A_n$ 为样本空间 Ω 的一个 **分割**, 亦称 **完备事件组**.

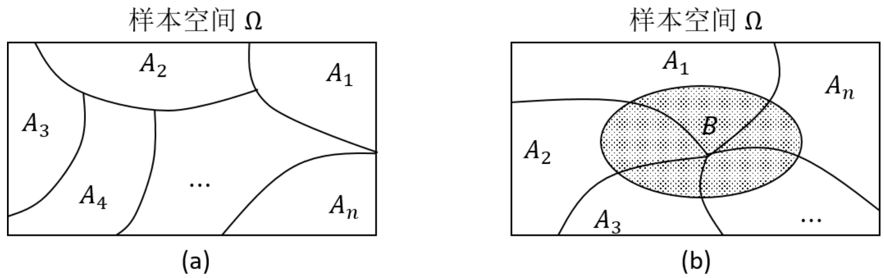

图 2.1 样本空间的分割与事件

如图 2.1(a) 所示, 若 $A_1, A_2, \cdots, A_n$ 为样本空间 Ω 的一个分割, 则每次试验时在 $A_1, A_2, \cdots, A_n$ 中有且仅有一个事件发生. 对任何事件 $A \subseteq \Omega$, 事件 A 与对立事件 $\bar{A}$ 构成样本空间 Ω 的一个分割.

基于样本空间的分割, 下面给出全概率公式:

定理 2.2 设事件 $A_1, A_2, \cdots, A_n$ 是样本空间 Ω 的一个分割, 对任意事件 B 有

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

该公式被称为 **全概率公式** (Law of total probability).

证明 该定理的证明本质上是对加法和乘法事件的综合运用. 首先根据分配律有

$$B = B \cap \Omega = B \bigcap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} BA_i$$

由 $A_i \cap A_j = \emptyset$ 可得 $BA_i \cap BA_j = \emptyset$, 由概率的有限可列可加性有

$$P(B) = P\left(\bigcup_{i=1}^{n} BA_i \right) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$

由于任意事件和其对立事件构成一个分割, 对任意概率非零的事件 A 和 B 有

$$P(B) = P(AB) + P(\bar{A}B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}).$$

如图 2.1(b), 还可以从另一个角度来理解全概率公式: 将事件 B 看作一个结果, 将事件 $A_1, A_2, \cdots, A_n$ 看作产生该结果的若干原因, 针对不同原因事件 B 发生的概率 (即条件概率 $P(B|A_i)$) 各不相同, 而到底是哪一种原因具有随机性. 具体而言, 每一种原因发生的概率 $P(A_i)$ 是已知的, 以及每一种原因对结果 B 的影响 $P(B|A_i)$ 已知, 则可以计算结果 $P(B)$.

下面来看一些例子:

例 2.8 小明参加一次人工智能竞赛, 目前的排名不理想, 分析其原因: 方法不够新颖的概率为 50%, 通过设计新方法后取得理想排名的概率为 50%; 程度代码有误的概率为 30%, 通过纠正代码后

取得理想排名的概率为 60%; 数据不充分的概率为 20%, 通过采集更多数据后取得理想排名的概率为 80%. 求小明最后取得理想排名的概率.

解 用 B 表示小明最后取得理想排名的事件, 用 A_1, A_2, A_3 分别表示方法不够新颖、程度代码有误、数据不充分这三个事件, 根据题意有

$$P(A_1) = 50\%, P(A_2) = 30\%, P(A_3) = 20\%, P(B|A_1) = 50\%, P(B|A_2) = 60\%, P(B|A_3) = 80\%.$$

小明最后取得理想排名的概率

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3) = 59\%.$$

例 2.9 随意抛 n 次硬币, 证明正面朝上的次数是偶数 (或奇数) 的概率为 1/2.

证明 用事件 A 表示前 $n-1$ 次抛硬币正面朝上的次数为偶数, 其对立事件 $\bar{A}$ 表示前 $n-1$ 次抛硬币朝上的次数为奇数, 事件 B 表示前 n 次硬币朝上的次数为偶数. 于是有

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = \frac{P(A)}{2} + \frac{P(\bar{A})}{2} = \frac{1}{2}.$$

方法二: 直接计算概率. 若正面朝上的次数是偶数, 则随意抛 n 次硬币中正面朝上的次数为偶数分别有 $\{0, 2, 4, \ldots, 2k\}$ $(2k \leqslant n)$, 根据概率公式直接计算有

$$\sum_{0 \leqslant k \leqslant n/2} \binom{n}{2k} \left(\frac{1}{2}\right)^{2k} \left(\frac{1}{2}\right)^{n-2k} = \frac{1}{2^n} \sum_{0 \leqslant k \leqslant n/2} \binom{n}{2k} = \frac{1}{2},$$

这里使用公式 $\sum_{0 \leqslant k \leqslant n/2} \binom{n}{2k} = 2^{n-1}$.

例 2.10 假设有 n 个箱子, 每个箱子里有 a 只白球和 b 只红球, 现从第一个箱子取出一个球放入第二个箱子, 第二个箱子取出一个球放入第三个箱子, 依次类推, 求从最后一个箱子取出一球是红球的概率.

解 用 A_i 表示从第 i 个箱子取出红球的事件 $(i \in [n])$, 则 $\overline{A_i}$ 表示从第 i 个箱子取出白球的事件. 则有

$$P(A_1) = b/(a+b) \qquad \text{和} \qquad P(\overline{A_1}) = a/(a+b)\ .$$

根据全概率公式有

$$P(A_2) = P(A_1)P(A_2|A_1) + P(\overline{A_1})P(A_2|\overline{A_1}) = \frac{b}{a+b} \times \frac{b+1}{a+b+1} + \frac{a}{a+b} \times \frac{b}{a+b+1} = \frac{b}{a+b}.$$

由此可知 $P(\overline{A_2}) = a/(a+b)$. 依次类推重复上述过程 $n-1$ 次, 最后一个箱子中取出一球是红球的概率为 $b/(a+b)$.

2.2.2 贝叶斯公式

贝叶斯公式也是概率论中最基本的公式之一, 在结果发生的情况下探讨是由何种原因导致结果. 具体而言, 假设有 $A_1, A_2, \cdots, A_n$ 种原因导致事件 B 发生, 贝叶斯公式研究在事件 B 发生情况下由原因 A_i 导致的概率, 即条件概率 $P(A_i|B)$.

定理 2.3 设 $A_1, A_2, \cdots, A_n$ 是样本空间 Ω 的一个分割, 用 B 表示任一事件且满足 $P(B)>0$. 对任意 $1 \leqslant i \leqslant n$ 有

$$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^n P(A_j)P(B|A_j)},$$

该公式被称为 **贝叶斯公式** (Bayes' formula).

贝叶斯公式由条件概率和全概率公式直接推导可得. 由于任何事件和其对立事件都是样本空间的一个分割, 对任意概率非零的事件 A 和 B 有

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(\bar{A})P(B|\bar{A})}.$$

全概率公式和贝叶斯公式应用的条件是相同的, 但解决的问题不同: 将事件 $A_1, A_2, \cdots, A_n$ 看作事件 B 发生的"原因", 而事件 B 是伴随着原因 $A_1, A_2, \cdots, A_n$ 而发生的"结果". 若知道各种原因 $P(A_i)$, 以及在该原因下事件 B 发生的概率 $P(B|A_i)$, 此时利用全概率公式计算结果事件 B 发生的概率; 若结果事件 B 已经发生, 此时利用贝叶斯公式探讨是由某原因 A_i 导致该结果的概率 $P(A_i|B)$.

贝叶斯公式被应用于生活中的很多决策问题, 与决策理论密切相关, 下面来看一个简单的例子:

例 2.11 小明参加一次人工智能竞赛, 目前的排名不理想, 分析其原因: 方法不够新颖的概率为 50%, 通过设计新方法后取得理想排名的概率为 50%; 程度代码有误的概率为 30%, 通过纠正代码后取得理想排名的概率为 60%; 数据不充分的概率为 20%, 通过采集更多数据后取得理想排名的概率为 80%. 因为时间有限, 小明只能选择三种方案 (设计新方法、纠正代码、采集更多数据) 中一种, 想要取得理想的排名, 小明应该选择哪一种方案.

解 用 B 表示小明最后取得理想排名的事件, 用 A_1, A_2, A_3 分别表示方法不够新颖、程度代码有误、数据不充分这三个事件, 根据题意有

$$P(A_1)=50\%, P(A_2)=30\%, P(A_3)=20\%, P(B|A_1)=50\%, P(B|A_2)=60\%, P(B|A_3)=80\%.$$

小明最后取得理想排名的概率

$$P(B) = P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3) = 59\%.$$

根据贝叶斯公式有 $P(A_1|B) = P(A_1)P(B|A_1)/P(B) = 25/59$, 同理可得 $P(A_2|B) = 18/59$ 和 $P(A_3|B) = 16/59$. 因此小明应该选择设计新方法来获得理想排名的概率更高.

例 2.12 (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

$$P(A|D) = P(A)P(D|A)/P(D) = 1/3 \quad \text{和} \quad P(C|D) = P(C)P(D|C)/P(D) = 2/3,$$

所以犯人 a 的推断不正确, 犯人 c 的推断正确.

贝叶斯公式提出了重要的推理逻辑, 在概率统计以及日常生活中存在多方面的应用. 假定 $A_1, A_2, \cdots, A_n$ 是导致结果事件 B 的“原因”, 概率 $P(A_i)$ 被称为 **先验概率** (prior probability), 反映了各种“原因”的可能性大小, 一般都是根据先前的经验总结而成. 若现在试验产生了事件 B, 这个信息有助于探讨事件发生的“原因”, 条件概率 $\Pr(A_i|B)$ 被称 **后验概率** (posterior probability), 反映了试验之后对各种“原因”发生可能性的新知识.

例如, 医生为诊断病人患了疾病 $A_1, A_2, \ldots$ 中哪一种疾病, 可以对病人进行检查, 确定某个指标 B (如血糖、血脂、血钙等), 从而帮助诊断, 此时可以采用贝叶斯公式来计算相关概率. 根据以往的数据资料确定先验概率 $P(A_i)$, 即人们患各种疾病的可能性; 在通过医学知识确定概率 $P(B|A_i)$, 最后通过贝叶斯公式计算后验概率 $P(A_i|B)$. 在实际应用中, 可能检验多个指标 B, 综合所有的后验概率进行诊断. 在自动诊断和辅助诊断的专家系统中, 这种方法非常实用.

贝叶斯公式使用中最存在争议之处在于先验的选取, 在很多实际应用中往往都根据以往的数据而得出的, 符合概率的频率解释, 但需要以往大量的历史数据, 在实际应用中通常难以满足. 其次, 在很多应用中先验概率可能由某一种主观的方式给出, 例如对未来宏观经济形势 (或对某人诚信度) 的判断, 这种将概率解释为信任程度的做法明显带有主观性, 通常被称为 **主观概率**.

伊索寓言“孩子与狼”讲一个小孩每天到山上放羊, 山里有狼出没, 第一天他在山上喊“狼来了！狼来了！”, 山下的村民们闻声便去打狼, 到了山上发现没有狼; 第二天仍是如此; 第三天狼真来了, 可无论小孩怎么喊叫, 也没有人来救他, 因为前二次他说了谎话, 人们不再相信他了. 我们可以将这个寓言抽象为一个主观概率的例子, 并利用贝叶斯公式来分析这个寓言中村民们的心理活动.

例 2.13 假设村民们对这个小孩的印象一般, 认为小孩说谎话和说真话的概率相同, 均为 1/2. 假设说谎话的小孩喊狼来了时狼真来的概率为 1/3, 而说真话的小孩喊狼来了时狼真来的概率为 3/4. 若第一天、第二天上山均没有发现狼, 请分析村民们的心理活动.

解 用 B_1 和 B_2 分别表示第一天和第二天狼来了的事件, 用 A_1 表示小孩第一天说谎话的事件, 用 A_2 表示在第一天狼没有的情况下小孩第二天说谎话的事件, 根据题意可知

$$P(A_1) = P(\overline{A_1}) = 1/2,\ P(B_1|A_1) = 1/3,\ P(B_1|\overline{A_1}) = 3/4,\ P(B_1|A_2) = 1/3,\ P(B_1|\overline{A_2}) = 3/4.$$

第一天村民上山打狼但没有发现狼, 根据贝叶斯公式可知村民们对说谎话小孩的认识发生了改变, 体现在

$$P(A_2) = P(A_1|\overline{B_1}) = \frac{P(\overline{B_1}|A_1)P(A_1)}{P(\overline{B_1}|A_1)P(A_1) + P(\overline{B_1}|\overline{A_1})P(\overline{A_1})} = \frac{8}{11} \approx 0.7273, \quad P(\overline{A_2}) = \frac{3}{11}.$$

此时, 村民对这个小孩说谎话的概率从 50% 调整到 72.72%.

第二天村民上山打狼还是没有发现狼, 根据贝叶斯公式可知村民们对说谎话小孩的认识又发生了改变, 体现在

$$P(A_2|\overline{B_2}) = \frac{P(\overline{B_2}|A_2)P(A_2)}{P(\overline{B_2}|A_2)P(A_2) + P(\overline{B_2}|\overline{A_2})P(\overline{A_2})} = \frac{64}{73} \approx 0.8767.$$

此时, 村民对这个小孩说谎话的概率从 72.72% 调整到 87.67%.

这表明村民们经过两次上当, 对这个小孩说谎话的概率从 50% 上升到 87.67%, 给村民留下这种印象, 他们听到第三次呼叫时不会再上山打狼.

2.3 事件独立性

前面的例子表明, 在事件 A 发生的情况下事件 B 发生的条件概率 $P(B|A)$, 通常不等于事件 B 发生的概率 $P(B)$ (无任何附加条件), 即 $P(B|A) \neq P(B)$, 也就是说“事件 A 发生通常会改变事件 B 发生的可能性”. 然而在有些特殊情形下, 事件 A 的发生对事件 B 的发生可能没有任何影响, 这就是本节所研究的事件独立性.

2.3.1 两事件的独立性

定义 2.3 设 (Ω, Σ, P) 是一个概率空间, 若事件 $A, B \in \Sigma$ 且满足 $P(AB) = P(A)P(B)$, 则称**事件 A 与 B 是相互独立的**, 简称 **独立**.