

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, 没有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

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 $\Omega = \{1, 2, \dots, 6\}$. 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 $P(B|A) = 1/3 \neq P(B)$.

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

$$P(B|A) = 1/3 = \frac{P(AB)}{P(A)}.$$

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 $P(B|A) \geq 0$;

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

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

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

从概率和条件概率的定义可知 $P(B|A) = P(AB)/P(A) \geq 0$ 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 $B_1, B_2, \dots, B_n, \dots$ 是两两互不相容的, 则可列个事件 $AB_1, AB_2, \dots, AB_n, \dots$ 也是两两互不相容的, 根据分配律有

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

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, 满足概率的性质, 例如,

性质 2.1 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 根据条件概率的定义和集合的运算有

$$\begin{aligned} 1 &= P(\Omega|A) = P(B \cup \bar{B}|A) = P(A(B \cup \bar{B}))/P(A) \\ &= P((AB) \cup (A\bar{B}))/P(A) = (P(AB) + P(A\bar{B}))/P(A) = P(B|A) + P(\bar{B}|A). \end{aligned}$$

性质 2.2 对随机事件 A, B 和 C 且 $P(C) > 0$, 有

$$P(A - B|C) = P(A|C) - P(AB|C) = P(A \cup B|C) - P(B|C).$$

证明 根据条件概率的定义有 $P(A - B|C) = P(A\bar{B}C)/P(C)$, 而

$$P(A\bar{B}C) = P(AC - B) = P(AC) - P(ABC),$$

由此证明 $P(A - B|C) = P(A|C) - P(AB|C)$. 针对另一个等式, 利用集合的性质有

$$\begin{aligned} P(A\bar{B}C) &= P((A\bar{B}C) \cup (A\bar{C}C)) = P(AC(\bar{B} \cup \bar{C})) = P(AC - BC), \\ &= P((AC) \cup (BC)) - P(BC) = P((A \cup B)C) - P(BC), \end{aligned}$$

由此完成证明.

性质 2.3 (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 根据定义有 $P(B_1 \cup B_2|A) = P((B_1 \cup B_2) \cap A)/P(A)$, 利用容斥原理和集合的运算有

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

由此完成证明.

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

$$\begin{aligned}\Omega &= \{(i, j): i \neq j, i, j \in [4]\}, & A &= \{(i, j): i \neq j, i \neq 4\}, \\ B &= \{(i, j): i \neq j, j \neq 4\}, & AB &= \{(i, j): i \neq j, i, j \in [3]\}.\end{aligned}$$

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 盒子中有 a 个红球和 b 个白球, 依次任意无放回地取出 n 个球 ($n \leq a + b$), 其中包括 k 个白球 ($k \leq b$), 在此情形下求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

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

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 .

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

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

可以将上式推广到更一般情形下的乘法公式.

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 $P(A_1 A_2 \cdots A_{n-1}) > 0$, 则有

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

下面来看几个利用乘法公式求解问题的例子.

例 2.3 盒子中有 1 个白球和 $n-1$ 个红球, 不放回随机抽出一球, 求第 k 次抽出白球的概率.

解 用 A_i 表示第 i 次没抽出白球的事件, 则第 k 次抽出白球的事件可表示为 $A_1 A_2 \cdots A_{k-1} \bar{A}_k$, 根据乘法公式有

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

也可根据 **抽签原理** 来求解该问题: 第 k 次抽出白球的概率与 k 无关, 每次抽中白球的概率相同, 共 n 个球, 因此第 k 次抽中白球的概率为 $1/n$.

例 2.4 设盒子里有 m 个红球和 n 个白球, 现随机抽出一球后观察其颜色并放回, 并加入 c 个相同颜色的球, 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 ($i \in [2]$), 以及用 B_i 表示第 i 次抽到红球的事件 ($i = 3, 4$), 于是有

$$\begin{aligned} P(A_1) &= \frac{m}{m+n}, & P(A_2|A_1) &= \frac{m+c}{m+n+c}, \\ P(B_1|A_1 A_2) &= \frac{n}{m+n+2c}, & P(B_2|A_1 A_2 B_1) &= \frac{n+c}{m+n+3c}. \end{aligned}$$

根据乘法公式有

$$P(A_1 A_2 B_1 B_2) = \frac{mn(m+c)(n+c)}{(m+n)(m+n+c)(m+n+2c)(m+n+3c)}.$$

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.

例 2.5 有甲乙两盒子分别装 n 个不同的白球和 m 个不同的红球, 从甲盒中任意取走一球, 再从乙盒里任意取走一球放入甲盒, 依次进行直至两盒子为空, 求甲盒中最后一次取走白球的概率.

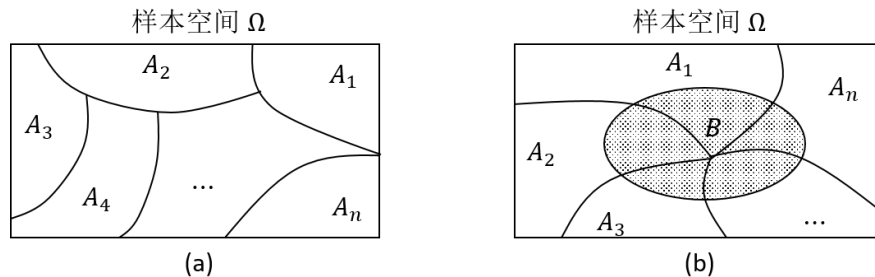


图 2.1 样本空间的分割与事件

解 将甲盒中的白球分别标为 $1, 2, \dots, n$, 用事件 A_i 表示从甲盒中最后取走的是第 i 号白球. 易知这些事件 A_1, A_2, \dots, A_n 是两两互不相容的, 且从甲盒中最后一次取走白球的事件可表示为 $A_1 \cup A_2 \cup \dots \cup A_n$, 根据事件的等可能性有

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

若事件 A_1 发生, 则从甲盒中取走的前 $m+n-1$ 个球均不是第 1 号白球, 用事件 B_j 表示第 j 次从甲盒中取走的球不是第 1 号白球, 即 $A_1 = B_1 B_2 \cdots B_{m+n-1}$. 根据乘法公式有

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

由此可知从甲盒中最后一次取走白球的概率为 $(1 - 1/n)^m$.

2.2 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. 首先介绍样本空间的一个分割.

定义 2.2 设 A_1, A_2, \dots, 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 $\Omega = A_1 \cup A_2 \cup \dots \cup A_n$,

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

如图 2.1(a) 所示, 若 A_1, A_2, \dots, A_n 为样本空间 Ω 的一个分割, 则每次试验时在 A_1, A_2, \dots, 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, \dots, 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

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

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, \dots, 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 2.6 小明在一次 AI 竞赛初期不利. 方法不好的概率为 50%, 改进方法后成功的概率为 50%; 程序有误的概率为 30%, 纠正程序后成功的的概率为 60%; 数据不充分的概率为 20%, 采集更多数据后成功的的概率为 80%. 求小明成功的概率.

解 用 B 表示小明成功的事件, 用 A_1, A_2, A_3 分别表示方法不好、程序有误、数据不充分三个事件, 根据题意有

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

根据全概率公式, 小明成功的概率为

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

例 2.7 随意抛 n 次硬币, 证明正面向上的次数是偶数 (或奇数) 的概率为 $1/2$.

证明 用事件 A 表示前 $n-1$ 次抛硬币正面向上的次数为偶数, 其对立事件 \bar{A} 表示前 $n-1$ 次抛硬币向上的次数为奇数, 事件 B 表示前 n 次硬币向上的次数为偶数. 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. 若正面向上的次数是偶数, 则随意抛 n 次硬币中正面向上的次数为偶数分别有 $\{0, 2, 4, \dots, 2k\}$ ($2k \leq n$), 根据概率公式直接计算有

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

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

例 2.8 设有 n 个盒子, 每个盒子有 a 个白球和 b 个红球, 现从第一个盒子取出一球放入第二个盒子, 第二个盒子取出一球放入第三个盒子, 依次类推, 求从最后一个盒子取出一球是红球的概率.

解 用 A_i 表示从第 i 个盒子取出红球的事件 ($i \in [n]$), 则 \bar{A}_i 表示取出白球的事件. 有

$$P(A_1) = b/(a+b) \quad \text{和} \quad P(\bar{A}_1) = a/(a+b).$$

根据全概率公式有

$$P(A_2) = P(A_1)P(A_2|A_1) + P(\bar{A}_1)P(A_2|\bar{A}_1) = \frac{b}{a+b} \times \frac{b+1}{a+b+1} + \frac{a}{a+b} \times \frac{b}{a+b+1} = \frac{b}{a+b}.$$

由此可知 $P(\bar{A}_2) = a/(a+b)$. 依次类推重复上述过程 $n-1$ 次, 最后一个盒子中取出一球是红球的概率为 $b/(a+b)$.

2.2.2 贝叶斯公式

贝叶斯公式是概率论中另一个重要的公式, 探讨在结果发生的情况下由何种原因导致结果. 具体而言, 假设有 A_1, A_2, \dots, 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, \dots, 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 $1 \leq i \leq n$ 有

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

该公式被称为 **贝叶斯公式** (Bayes' formula).