

第一章 随机事件和概率

1.1 本章导读

概率论研究的不是“完全不知道会发生什么”，而是研究这样一类现象：所有可能结果可以预先说清楚，但在试验发生前，无法确定这一次究竟出现哪个结果。本章先把日常语言中的“可能、同时、至少、条件、互不影响”等说法翻译成事件和概率，再建立后续所有章节都会调用的计算工具。

本章的知识链如下：

随机试验 $\longrightarrow$ 样本空间与事件 $\longrightarrow$ 事件运算 $\longrightarrow$ 概率模型与公理
 $\longrightarrow$ 条件概率 $\longrightarrow$ 独立性 $\longrightarrow$ 全概率与贝叶斯公式。

学习本章时要反复检查三个问题：

1. 题目中的字母事件分别表示什么，事件之间是“或”“且”还是“非”？
2. 使用某个公式前，它的条件是否满足，例如等可能、条件事件概率为正、事件独立或构成完备事件组？
3. 算出的数是否位于 $[0, 1]$ ，是否符合“条件增加后样本空间缩小”等直观判断？

1.2 随机试验、样本空间与随机事件

1.2.1 从不确定现象到可研究的试验

“明天下雨”与“掷骰子得到几点”都不能在事前确定结果，但二者并非毫无规律。概率论首先要把研究对象的范围固定下来。

定义 1.1 (随机试验)

若一个试验满足:

1. 在相同条件下可以重复进行;
2. 每次试验的可能结果不止一个, 而且全部可能结果事先可知;
3. 每次试验前不能确定哪一个结果会出现,

则称它为随机试验, 常记为 E .

第二、三条合起来描述“随机现象”; 再加上第一条, 才能通过大量重复试验研究其稳定规律. 这里的“随机”不等于“结果范围未知”: 骰子的结果不确定, 但所有可能点数仍然是已知的 $1, 2, \dots, 6$.

定义 1.2 (样本点与样本空间)

随机试验 E 的每一个不可再分的可能结果称为**样本点**, 记为 ω ; 所有样本点组成的集合称为**样本空间**, 记为 Ω .

例题 1.3 (写出样本空间)

分别写出下列试验的样本空间:

1. 掷一枚骰子, 记录点数;
2. 观察某视频一小时内的播放量;
3. 记录某种机器的寿命;
4. 同时记录一名学生的身高与体重.

选择方法的原因: 样本空间只回答“所有不可再分的结果有哪些”, 因此先确定一次试验最终记录的是一个数还是一对数.

$$\Omega_1 = \{1, 2, 3, 4, 5, 6\},$$

$$\Omega_2 = \{0, 1, 2, \dots\},$$

$$\Omega_3 = (0, +\infty) \quad \text{或依据实际范围写成} [a, b],$$

$$\Omega_4 = \{(x, y) : x > 0, y > 0\}.$$

结果检查: 一次试验只产生其中一个样本点; 样本空间本身则收集了全部可能结果.

若样本点有限或可列, 称为离散样本空间; 若样本点充满一个区间或区域, 称为连续样本空间. 这个分类会直接发展成第二章的离散型与连续型随机变量.

1.2.2 事件是样本空间的子集

定义 1.4 (随机事件)

样本空间 Ω 的一个子集称为随机事件, 常记为 A, B, C . 一次试验产生的样本点属于 A 时, 称事件 A 发生.

由一个样本点组成的事件称为**基本事件**; 空集 $\emptyset$ 称为**不可能事件**; 整个样本空间 Ω 称为**必然事件**. 不可能事件与必然事件在客观上具有确定性, 但把它们作为随机事件的两个极端情况, 可以使运算规则保持统一.

例题 1.5 (样本点与事件不要混淆)

掷一枚骰子. 令 A 表示“点数为偶数”, 则

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad A = \{2, 4, 6\}.$$

若本次掷出 4, 出现的样本点是 $\omega = 4$, 事件 A 发生. A 不是“这次掷出的结果”, 而是所有能使“偶数”成立的结果集合.

注: “单点概率为零”不等于“该点不可能出现”. 在连续样本空间中, 单个点可能是非空事件, 但其概率仍可为零. 事件是否为空是集合问题, 概率是否为零是度量问题, 二者不能随意反推.

小结

本节把随机现象表示为“样本空间中的一个样本点出现”. 使用时先明确试验记录什么, 再写出 Ω , 最后把题目关心的结果写成 Ω 的子集. 后面所有概率公式都作用在这些事件上.

1.3 事件的关系与运算

1.3.1 把自然语言翻译成集合语言

事件本质上是集合, 所以事件关系就是集合关系; 但做题时还要能把集合式翻译回“发生”的语言.

关系或运算	集合表示	事件语言
包含	$A \subseteq B$	A 发生必然导致 B 发生
相等	$A = B$	两事件含有相同样本点
并	$A \cup B$	A, B 至少一个发生
交	$A \cap B$ 或 AB	A, B 同时发生
差	$A - B = A \cap \bar{B}$	A 发生而 B 不发生
互斥	$AB = \emptyset$	A, B 不能同时发生
对立	$AB = \emptyset, A \cup B = \Omega$	A, B 必有且仅有一个发生

注: 互斥不等于对立. 互斥只说明“两者不能同时发生”, 两者还可能同时不发生; 对立还要求二者并起来覆盖整个样本空间. 对立事件必互斥, 互斥事件未必对立.

例题 1.6 (文字事件的表示)

从一批产品中依次抽取三件, 令 A, B, C 分别表示第一、二、三件为正品. 表示:

1. 第三件为正品;
2. 只有第三件为次品;
3. 恰有一件次品;
4. 至多一件次品.

先把“且”译成交、“或”译成并、“不是”译成对立:

- (1) C ;
- (2) $A\bar{B}\bar{C}$;
- (3) $\bar{A}BC \cup A\bar{B}C \cup AB\bar{C}$;
- (4) $AB \cup AC \cup BC$.

第 (4) 问选择“至少两件正品”来表示, 比逐项枚举“零件或一件次品”更短. 检查时逐

个代入“正正正、次正正”等结果, 确认没有遗漏或重复.

1.3.2 运算律与德摩根律

并与交满足交换律、结合律和分配律:

$$\begin{aligned} A \cup B &= B \cup A, & AB &= BA, \\ A \cup (B \cap C) &= (A \cup B) \cap C, & A(BC) &= (AB)C, \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), & A(B \cup C) &= AB \cup AC. \end{aligned}$$

性质 1.7 (德摩根律)

$$\overline{A \cup B} = \bar{A} \bar{B}, \quad \overline{AB} = \bar{A} \cup \bar{B}.$$

更一般地,

$$\overline{\bigcup_{i=1}^{\infty} A_i} = \bigcap_{i=1}^{\infty} \bar{A}_i, \quad \overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \bar{A}_i.$$

注: 证明思路 (B 级): 以第一式为例, 任取样本点 ω . 它属于 $\overline{A \cup B}$, 当且仅当它既不属于 A 也不属于 B , 即同时属于 $\bar{A}$ 与 $\bar{B}$. 两个集合包含完全相同的样本点, 因此相等. 该方法是“逐点判断集合相等”, 比死记“长横变短横”更可靠.

例题 1.8 (用德摩根律识别对立事件)

事件 A 表示“甲产品畅销且乙产品滞销”. 令 B 表示“甲产品畅销”, C 表示“乙产品畅销”, 则

$$A = B\bar{C}, \quad \bar{A} = \overline{B\bar{C}} = \bar{B} \cup C.$$

所以 $\bar{A}$ 表示“甲产品滞销或乙产品畅销”. 关键转折是先把文字写成事件式, 再求整体的对立, 不能只把句中某一个词反过来.

注: 事件的差不是普通代数减法. 优先使用 $A - B = A\bar{B}$, 再用交、并、对立的运算律化简; 普通代数中的“去括号”规则一般不适用于事件差.

小结

本节解决“怎样把复杂文字条件写成事件式”. 先找“且、或、非”, 再检查互斥与对立是否被混淆. 德摩根律负责把“至少一个”“全部”等互为反面的说法连接起来, 后面计算复杂事件概率时会反复使用.

1.4 古典概型与几何概型

1.4.1 古典概型: 有限个等可能结果

定义 1.9 (古典概型)

若样本空间只含有限个样本点, 且每个样本点出现的可能性相同, 则称该试验为古典概型 (等可能概型). 若 $|\Omega| = n$, 事件 A 含 k 个样本点, 则

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

使用古典概型前必须检查“有限”和“等可能”. 若给样本点编号的方式改变, 分子与分母必须始终采用同一种计数口径.

排列与组合是古典概型的计数工具:

$$A_n^m = \frac{n!}{(n-m)!}, \quad \binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

顺序重要时用排列, 顺序不重要时用组合. 组合数恒等式

$$\binom{n}{m} = \binom{n}{n-m}, \quad \binom{n}{m} + \binom{n}{m-1} = \binom{n+1}{m}$$

用于简化计数; 其完整代数推导属于 C 级, 本阶段会解释含义并正确使用即可.

例题 1.10 (生日问题: 正难则反)

假设一年有 365 天且生日等可能, 求 60 人中至少两人生日相同的概率.

选择方法的原因: 正面需要分类“两人相同、三人相同、多个小组分别相同”, 容易重复; 反面“所有人生日互不相同”只有一种结构.

把 60 个人看成不同的球, 把 365 天看成不同的盒子. 全部生日安排有 365^{60} 种; 生日互

不相同有

$$A_{365}^{60} = 365 \cdot 364 \cdots 306$$

种. 因此

$$P(\text{至少两人生日相同}) = 1 - \frac{A_{365}^{60}}{365^{60}} \approx 0.994.$$

结果检查: 人数增加时概率应上升; 答案接近 1 虽然反直觉, 但没有超出概率范围.

例题 1.11 (平均分班中的特殊元素)

15 名新生中有 3 名竞赛生, 随机平均分到 3 个有编号的班级, 每班 5 人. 求:

1. 每班恰有 1 名竞赛生的概率;
2. 3 名竞赛生在同一班的概率.

选择方法的原因: 限制集中在竞赛生身上, 所以先安排特殊元素, 再补普通学生, 可以避免重复和遗漏. 全部分班方法数为

$$\binom{15}{5} \binom{10}{5}.$$

第 (1) 问, 先将 3 名竞赛生全排列到 3 个班, 再分别补入普通学生:

$$P_1 = \frac{3! \binom{12}{4} \binom{8}{4}}{\binom{15}{5} \binom{10}{5}} = \frac{25}{91}.$$

第 (2) 问, 先选择竞赛生所在班, 再补足该班和另一个班:

$$P_2 = \frac{\binom{3}{1} \binom{12}{2} \binom{10}{5}}{\binom{15}{5} \binom{10}{5}} = \frac{6}{91}.$$

检查: 两个事件互斥, 且 $P_1 + P_2 < 1$; 剩余情形是 3 名竞赛生按 2 + 1 + 0 分布.

1.4.2 几何概型: 连续区域中的等可能

定义 1.12 (几何概型)

若样本点充满一个有界区域 Ω , 随机点落入等度量区域的可能性相同, 则称为几何概型. 事件 A 对应区域 $G \subseteq \Omega$ 时,

$$P(A) = \frac{m(G)}{m(\Omega)},$$

其中 m 可表示长度、面积或体积.

几何概型把古典概型的“等可能”从有限样本点推广到连续区域. 概率只与区域的度量有关, 与位置和形状无关; 它也是第二章均匀分布的直观背景.

例题 1.13 (二维几何概型)

在区间 $(0, 1)$ 内独立任取两个数 X, Y , 求 $X + Y < \frac{6}{5}$ 的概率.

因为一次试验产生一对数 (X, Y) , 样本空间是单位正方形

$$\Omega = \{(x, y) : 0 < x < 1, 0 < y < 1\}.$$

事件区域位于直线 $x + y = \frac{6}{5}$ 下方. 直接求五边形面积较繁, 改求右上角补三角形: 两条直角边均为 $\frac{4}{5}$. 因此

$$P\left(X + Y < \frac{6}{5}\right) = 1 - \frac{1}{2} \left(\frac{4}{5}\right)^2 = \frac{17}{25}.$$

关键转折是把两个数视为平面坐标; 结果检查是最后必须除以总面积, 本题总面积恰为 1.

注: 经典辨析: 古典概型要求“有限且等可能”, 几何概型要求“连续区域内按度量均匀”. 若只有“随机”二字而没有等可能依据, 不能擅自使用“有利情况数除以总情况数”.

小结

本节把概率计算转化为“有利部分占整体的比例”. 古典概型先判断顺序是否重要, 再计数; 几何概型先建立坐标区域, 再求度量之比. 两者都必须先检查等可能条件.

1.5 概率的公理化定义与基本性质

1.5.1 频率提供直觉, 公理给出定义

进行 n 次相同条件下的试验, 事件 A 发生 n_A 次, 称 n_A 为频数,

$$f_n(A) = \frac{n_A}{n}$$

称为频率. 频率是一次实际试验得到的数; 概率 $P(A)$ 是描述事件发生可能性大小的数学量. 大量重复试验中频率的稳定性将在第五章由大数定律精确说明.

公理 1.14 (概率公理)

设 Ω 为样本空间. 若集合函数 P 满足:

1. 非负性: 对任意事件 A , $P(A) \geq 0$;
2. 规范性: $P(\Omega) = 1$;
3. 可列可加性: 若 $A_1, A_2, \dots$ 两两互斥, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

则称 $P(A)$ 为事件 A 的概率.

概率公理是理论起点, 不再证明 (C 级). 古典概型与几何概型都是满足这三条公理的具体模型.

性质 1.15 (由公理推出的常用性质)

对任意事件 A, B :

$$P(\emptyset) = 0, \quad 0 \leq P(A) \leq 1,$$

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

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

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

若 $B \subseteq A$, 则

$$P(A - B) = P(A) - P(B), \quad P(B) \leq P(A).$$

注: 证明思路 (B 级): 所有性质都围绕“把事件拆成两两互斥的部分”. 例如

$$A = (A - B) \cup AB$$

且右侧两部分互斥, 所以由有限可加性得到减法公式; 再将

$$A \cup B = A \cup (B - A)$$

拆成互斥部分, 就得到加法公式. 真正起作用的条件是互斥, 而不是图形长得像面积.

性质 1.16 (有限可加性与容斥公式)

若 $A_1, \dots, A_n$ 两两互斥, 则

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

对任意三个事件,

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(AB) - P(AC) - P(BC) + P(ABC). \end{aligned}$$

例题 1.17 (概率为零能推出什么)

已知 $P(AB) = 0$. 能否断定 A, B 互斥?

不能. $P(AB) = 0$ 只说明交事件的概率为零, 不能推出 $AB = \emptyset$. 在连续模型中, AB 可以是一个非空单点事件而概率为零.

但若 $P(AB) > 0$, 则一定有 $AB \neq \emptyset$, 所以 A, B 能同时发生. 这里使用的是“空事件概率必为零”的逆否命题.

例题 1.18 (概率大小比较)

比较 $P(AB), P(A), P(A \cup B), P(A) + P(B)$ 的大小.

利用包含关系

$$AB \subseteq A \subseteq A \cup B$$

得到

$$P(AB) \leq P(A) \leq P(A \cup B).$$

再由加法公式及 $P(AB) \geq 0$,

$$P(A \cup B) = P(A) + P(B) - P(AB) \leq P(A) + P(B).$$

所以

$$P(AB) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B).$$

方法选择: 概率不等式优先找事件包含, 再使用 $0 \leq P(\cdot) \leq 1$.

小结

本节从三条概率公理得到后续计算的基础工具. 使用加法、减法或拆分公式时, 先寻找两两互斥的分块; 遇到概率不等式时, 先找事件包含. 特别注意: 事件关系可以推出相应概率关系, 反向一般不成立.

1.6 条件概率与乘法公式

1.6.1 条件就是缩小后的样本空间

定义 1.19 (条件概率)

设 A, B 为事件且 $P(B) > 0$, 称

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

为在事件 B 发生的条件下事件 A 发生的条件概率.

已知 B 发生后, 候选样本点从 Ω 缩小到 B ; 在这个新范围中, 真正使 A 发生的是 AB . 因此条件概率本身仍是概率, 对竖线左侧事件满足非负性、规范性、可加性、对立事件公式等性质.

注: 只能直接化简竖线左侧:

$$P(\bar{A} | C) = 1 - P(A | C).$$

而 $P(A | \bar{C})$ 不能写成 $1 - P(A | C)$, 因为右侧是条件, 必须回到定义计算.

定理 1.20 (乘法公式)

若 $P(B) > 0$, 则

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

若 $P(A_1 A_2 \cdots A_{n-1}) > 0$, 则

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

证明: 证明目标: 从条件概率定义推出两事件乘法公式.

出发点: 已知 $P(B) > 0$, 所以 $P(A | B)$ 有定义.

关键转换: 把条件概率定义中的分母 $P(B)$ 乘到等式另一边.

正式推导: 由定义

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

因为 $P(B) > 0$, 等式两边同乘 $P(B)$, 得到

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

这就是乘法公式. 多事件情形对 $P(A_1 \cdots A_n)$ 反复应用两事件公式即可.

证明小结: 真正保证变形合法的是 $P(B) > 0$. 这个公式把“一起发生”拆成“先发生条件事件, 再在该条件下发生另一个事件”, 是全概率公式的直接工具. ■

例题 1.21 (不放回拨号)

某人忘记电话号码最后一位, 随机尝试且错误号码不再重复. 求不超过三次拨通的概率.

选择方法的原因: 正面可以分“第 1、2、3 次拨通”, 反面“三次均错误”更容易用乘法公式连乘.

$$P(\text{不超过三次拨通}) = 1 - \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{7}{8} = \frac{3}{10}.$$

每个分数都带有前次失败给出的新信息, 所以不能把三次失败简单当成独立事件. 检查:

十个号码在试验前地位对称, 前三个位置包含正确号码的概率也应为 $3/10$.

小结

本节解决“获得新信息后怎样重算概率”. 先检查条件事件概率是否为正, 再把样本空间缩小到条件事件. 乘法公式适合沿时间或层次逐步展开, 尤其要判断前一步结果是否改变后一步概率.

1.7 事件的独立性

1.7.1 独立不是互斥

定义 1.22 (两个事件相互独立)

若

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

则称事件 A 与 B 相互独立. 当 $P(B) > 0$ 时, 这等价于

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

即知道 B 是否发生并不改变对 A 的概率判断.

独立描述“概率互不影响”; 互斥描述“不能同时发生”. 它们讨论的层次不同.

命题 1.23 (正概率互斥事件不独立)

若 A, B 互斥且 $P(A) > 0, P(B) > 0$, 则 A, B 不独立.

证明: 证明目标: 在两个事件概率均为正的前提下, 证明互斥排除独立.

出发点: 互斥给出 $AB = \emptyset$; 独立要求 $P(AB) = P(A)P(B)$.

关键转换: 分别计算交事件概率与两个边缘概率的乘积并比较.

正式推导: 由 $AB = \emptyset$,

$$P(AB) = P(\emptyset) = 0.$$

另一方面, 因为 $P(A) > 0, P(B) > 0$,

$$P(A)P(B) > 0.$$

因此 $P(AB) \neq P(A)P(B)$, 不满足独立定义, 故 A, B 不独立.

证明小结: 正概率条件不可删除. 若 $A = \emptyset, B = \Omega$, 二者既互斥又独立, 这个反例说明没有正概率条件时, 不能声称“互斥必不独立”. ■

性质 1.24 (独立性的常用性质)

若 A, B 独立, 则

$$A \text{ 与 } \bar{B}, \quad \bar{A} \text{ 与 } B, \quad \bar{A} \text{ 与 } \bar{B}$$

也分别独立. 概率为 0 或 1 的事件与任意事件独立.

注: 证明思路 (B 级): 例如

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

故 A 与 $\bar{B}$ 独立. 若 $P(A) = 0$, 由 $AB \subseteq A$ 得

$$0 \leq P(AB) \leq P(A) = 0,$$

于是 $P(AB) = 0 = P(A)P(B)$; 概率为 1 的情形可转化为其对立事件概率为 0.

1.7.2 多个事件的独立性

三个事件 A, B, C 两两独立, 是指

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C).$$

它们相互独立还要求

$$P(ABC) = P(A)P(B)P(C).$$

相互独立比两两独立更强; 不能只验证两两乘积就宣称相互独立.

例题 1.25 (至少一人成功)

三人独立猜谜, 猜中的概率分别为 $\frac{1}{5}, \frac{1}{3}, \frac{1}{4}$. 求谜语被猜中的概率.

“至少一人猜中”的反面是“三人都未猜中”. 因为三人的试验独立, 反面事件概率可以相乘:

$$P(\text{至少一人猜中}) = 1 - \frac{4}{5} \cdot \frac{2}{3} \cdot \frac{3}{4} = \frac{3}{5}.$$

方法检查: 若没有“独立”条件, 最后一步不能直接相乘.

小结

本节解决“一个事件的信息是否会改变另一个事件的概率”. 判断独立的最终标准是乘积等式. 互斥与独立不能混用; 涉及多个事件时, 还必须区分两两独立和相互独立.

1.8 全概率公式与贝叶斯公式

1.8.1 先把样本空间不重不漏地划分

定义 1.26 (完备事件组)

若事件 $B_1, \dots, B_n$ 满足

$$B_i B_j = \emptyset \quad (i \neq j), \quad \bigcup_{i=1}^n B_i = \Omega,$$

则称 $B_1, \dots, B_n$ 构成 Ω 的一个完备事件组, 也称样本空间的一个划分. 每次试验中它们必有且仅有一个发生.

识别这类题的信号是: 题目先给出一组“原因、来源或第一阶段”, 它们不重不漏地覆盖全部情况; 随后研究建立在这些情况之上的结果.

定理 1.27 (全概率公式)

若 $B_1, \dots, B_n$ 为完备事件组, 且 $P(B_i) > 0$, 则对任意事件 A ,

$$P(A) = \sum_{i=1}^n P(B_i)P(A | B_i).$$

证明: 证明目标: 利用样本空间的划分, 把 $P(A)$ 写成各类条件下概率的加权和.

出发点: $\bigcup_{i=1}^n B_i = \Omega$ 且各 B_i 两两互斥.

关键转换: 用 B_i 将 A 分割成互斥事件 AB_i , 再分别使用有限可加性与乘法公式.

正式推导:

$$A = A\Omega = A\left(\bigcup_{i=1}^n B_i\right) = \bigcup_{i=1}^n (AB_i).$$

由于 B_i 两两互斥, AB_i 也两两互斥, 因此

$$P(A) = \sum_{i=1}^n P(AB_i).$$

再由乘法公式 $P(AB_i) = P(B_i)P(A | B_i)$, 得到

$$P(A) = \sum_{i=1}^n P(B_i)P(A | B_i).$$

证明小结: 全概率公式真正使用的是“划分 + 互斥可加 + 乘法公式”. 若分类有重叠或遗漏, 就不能直接套用. ■

定理 1.28 (贝叶斯公式)

在全概率公式的条件下, 若 $P(A) > 0$, 则

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{j=1}^n P(B_j)P(A | B_j)}.$$

证明: 证明目标: 在已知结果 A 发生后, 反推它来自第 i 类原因 B_i 的概率.

出发点: 条件概率定义、乘法公式和已经证明的全概率公式.

关键转换: 分子 $P(AB_i)$ 用乘法公式展开, 分母 $P(A)$ 用全概率公式展开.

正式推导:

$$P(B_i | A) = \frac{P(AB_i)}{P(A)}$$

$$= \frac{P(B_i)P(A | B_i)}{\sum_{j=1}^n P(B_j)P(A | B_j)}.$$

证明小结: 分子正是全概率展开中的第 i 项, 分母是所有来源产生结果 A 的总概率. 贝叶斯公式不是孤立的新公式, 而是条件概率定义与全概率公式的组合. ■

例题 1.29 (产品来源: 由因到果, 再由果推因)

甲、乙、丙三厂的市场占有率分别为 $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$, 次品率分别为 2%, 1%, 3%. 随机购买一件产品, 求:

1. 买到次品的概率;
2. 已知买到次品, 它来自甲厂的概率.

令 B_1, B_2, B_3 分别表示产品来自甲、乙、丙厂, 令 A 表示产品为次品. 三厂构成完备事件组.

第 (1) 问是“由来源求结果”, 使用全概率公式:

$$\begin{aligned} P(A) &= \frac{1}{4} \times 0.02 + \frac{1}{4} \times 0.01 + \frac{1}{2} \times 0.03 \\ &= 0.0225. \end{aligned}$$

第 (2) 问是“已知结果反推来源”, 使用贝叶斯公式:

$$P(B_1 | A) = \frac{\frac{1}{4} \times 0.02}{0.0225} = \frac{2}{9} \approx 0.2222.$$

检查: 后验概率约 0.2222, 低于甲厂原有份额 0.25, 因为甲厂次品率低于总体平均次品率 2.25%, 方向合理.

例题 1.30 (通信误码)

发送信息 A, B 的频率之比为 3 : 2. A 被误收为 B 的概率为 0.04, B 被误收为 A 的概率为 0.07. 已知接收到 A , 求原发送信息也是 A 的概率.

令 H 表示“原发送 A ”, R 表示“接收到 A ”. 则

$$P(H) = \frac{3}{5}, \quad P(\bar{H}) = \frac{2}{5}, \quad P(R | H) = 0.96, \quad P(R | \bar{H}) = 0.07.$$

先由全概率公式计算收到 A 的概率:

$$P(R) = \frac{3}{5} \times 0.96 + \frac{2}{5} \times 0.07 = 0.604.$$

再由贝叶斯公式反推:

$$P(H | R) = \frac{\frac{3}{5} \times 0.96}{0.604} = \frac{144}{151} \approx 0.9536.$$

关键转折是自己重新设置不混淆的事件字母, 并区分“原发送”和“接收到”两个层次.

性质 1.31 (抽签原理)

在无放回随机抽取中, 试验开始前, 第 k 个位置抽到某一类对象的概率都等于该类对象在总体中所占的比例, 与抽取顺序无关.

注: 抽签原理的完整一般证明属于 C 级, 本阶段不要求复现. 要注意它描述的是试验开始前各位置的边缘概率相同; 一旦前面抽取结果已经公开, 后续位置的条件概率会随信息改变.

小结

本节解决两层随机过程. 全概率公式把“结果”按所有可能原因分解, 贝叶斯公式在观察到结果后更新各原因的概率. 使用时先检查 B_i 是否构成完备事件组, 再写清先验概率 $P(B_i)$ 、条件概率 $P(A | B_i)$ 和所求方向.

1.9 本章总结

1.9.1 知识结构与方法选择

题目特征	优先方法	先检查的条件
有限个等可能结果	古典概型、排列组合	有限、等可能、计数口径一致
连续区域内均匀取点	几何概型	均匀性、区域与度量
至少一个、至多、全都	对立事件、德摩根律	正面还是反面更易算
已知某事件发生	条件概率	条件事件概率大于零
分阶段同时发生	乘法公式	每一步的条件是否改变
概率互不影响	独立性	乘积等式是否成立
原因不重不漏、求结果	全概率公式	原因构成完备事件组
已知结果、反推原因	贝叶斯公式	分母的总概率大于零

1.9.2 核心公式及适用条件

$$P(A \cup B) = P(A) + P(B) - P(AB),$$

$$P(A | B) = \frac{P(AB)}{P(B)}, \quad P(B) > 0,$$

$$P(AB) = P(B)P(A | B), \quad P(B) > 0,$$

$$A, B \text{ 独立} \iff P(AB) = P(A)P(B),$$

$$P(A) = \sum_i P(B_i)P(A | B_i),$$

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_j P(B_j)P(A | B_j)}.$$

1.9.3 重点辨析与易错点

1. 互斥是交集为空; 独立是概率乘积等式. 正概率互斥事件不独立, 但没有正概率条件时不能直接判断.
2. 互斥未必对立; 对立还要求并集为 Ω .

3. $P(A) = 0$ 未必有 $A = \emptyset$; $P(A) = 1$ 也未必有 $A = \Omega$.
4. 条件概率只对竖线左侧事件使用概率运算律, 右侧条件不能随意取反或拆分.
5. 古典概型的分母不是“看起来所有情况”, 而是与分子采用相同单位划分后的等可能基本结果数.
6. 两两独立不等于相互独立; 多个事件时必须按定义逐级验证.
7. 全概率公式的分类必须不重不漏; 贝叶斯公式的分子是全概率展开中的一项.

1.9.4 证明掌握清单

- A 级: 能从条件概率定义推出乘法公式; 能用事件划分完整推出全概率公式和贝叶斯公式; 能证明正概率互斥事件不独立并给出缺少正概率条件时的反例.
- B 级: 能说明概率基本性质来自“互斥分割”; 能用逐点法解释德摩根律; 能说明独立性对补事件保持以及概率为 0 或 1 的事件为何与任意事件独立.
- C 级: 概率公理作为理论起点不证明; 抽签原理的一般证明和组合恒等式的完整代数推导本阶段不要求复现, 但必须会写条件并正确使用.

1.9.5 自测题

练习 1.32 (概念与事件运算)

设 A, B, C 为事件. 分别表示“恰有两个发生”和“至少两个发生”; 再说明互斥与对立的区别.

练习 1.33 (概率性质)

已知 $P(A) = 0.6, P(B) = 0.5, P(AB) = 0.3$, 求 $P(A \cup B)$ 、 $P(A - B)$ 和 $P(\bar{A}B)$, 并判断 A, B 是否独立.

练习 1.34 (条件概率与独立性)

袋中有 4 个红球、6 个白球, 不放回抽取两次. 求第二次为红球的概率, 并分别用全概率公式和抽签原理解释结果.

练习 1.35 (贝叶斯判断)

某检测对患病者呈阳性的概率为 0.95, 对未患病者误报阳性的概率为 0.02, 人群患病率为 0.01. 求检测阳性者实际患病的概率, 并解释为什么结果不等于 0.95.

本章小结

本章建立了概率论的语言和第一套计算体系: 用样本空间刻画全部可能结果, 用事件运算表达复杂条件, 用概率公理保证计算规则, 再通过条件概率、独立性、全概率公式和贝叶斯公式处理信息、层次与反推问题. 做题时始终先写事件、查条件、再选公式; 这套习惯会直接决定第二章以后能否正确建立随机变量及其分布.