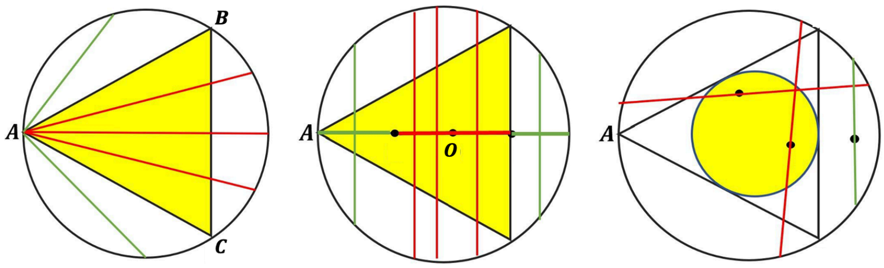

图 1.4: 伯特兰悖论的三种解法: 红色的弦为满足要求的弦

* 伯特兰悖论 (Bertrand's Paradox)

法国数学家约瑟夫·伯特兰 (Joseph Bertrand) 于 1889 年提出这样一个问题:

> "在一个单位圆内随机画一条弦, 问这条弦的长度大于圆内接等边三角形的边长 (即弦长大于 $\sqrt{3}$) 的概率是多少?"

解法一: 固定端点法. 固定弦的第一个端点 A, 第二个端点在圆周上均匀分布. 记以 A 为端点的圆内接等边三角形端点为 A, B, C, 则弦长大于 $\sqrt{3}$ 当且仅当第二个端点落在圆周的 $B - C$ 段圆弧上 (见图 1.4 左). 因此, 所求概率

$$P = \frac{1}{3}.$$

解法二: 径向均匀法. 弦的中点在径向上均匀分布. 弦长大于 $\sqrt{3}$ 当且仅当弦的中点与圆心 O 的距离 d 大于 $1/2$ (见图 1.4 中). 因此, 所求概率

$$P = \frac{1}{2}.$$

解法三: 中点均匀法. 弦的中点在圆内均匀分布. 弦长大于 $\sqrt{3}$ 当且仅当弦的中点落于半径为 $1/2$ 的同心圆内 (见图 1.4 右). 因此, 所求概率

$$P = \frac{\pi(1/2)^2}{\pi \cdot 1^2} = \frac{1}{4}.$$

这三种解法看似都合理, 但结果不同. 纠其原因, 是由于直观描述中的" 随机" 机制不明确导致的. 事实上, 这三种解法是基于三种对" 随机" 的理解做出的, 它们实际上是在解决三种不同的几何概型问题, 对应有三个不同的概率空间

$$\Omega_1 = (0, 2\pi) \quad \sim \quad \text{弦的第二个端点在圆周上均匀分布,}$$
$$\Omega_2 = [0, 1) \quad \sim \quad \text{弦的中点与圆心的距离均匀分布,}$$
$$\Omega_3 = \{(x, y) : x^2 + y^2 < 1\} \quad \sim \quad \text{弦的中点在圆内均匀分布.}$$

由此例可见, 对同一问题的不同表述可能导致不同的概率模型. 因此, 概率有时依赖于主观选择. 在现代概率论中, 一般会按照测度论的方法, 事先明确概率空间与概率的定义, 以避免这个问题.

1.5 条件概率和事件的独立性

本节我们介绍条件概率, 这是概率论中最重要的概念之一. 首先, 在计算某些随机事件的概率时, 往往同时会有某些关于该试验的附加信息, 利用这些信息获得的概率即为条件概率. 其次, 即便没有附加信息可用, 利用条件概率的方法往往能使得概率的计算变得简单.

1.5.1 条件概率

定义 1.17. 设 A, B 是某随机试验的两个事件, 且 $P(B) > 0$. 称

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

为在事件 B 发生的条件下事件 A 发生的条件概率, 简称 A 关于 B 的**条件概率**.

注 1.18. (i) 条件概率 $P(A \mid B)$ 可看成是将样本空间压缩为 B 后 A 发生的概率;

(ii) 条件概率也是概率, 即 $P(\cdot \mid B)$ 满足概率的公理化定义1.5. 因此, 概率所满足的性质条件概率都满足, 例如 $P(\overline{A} \mid B) = 1 - P(A \mid B)$.

例 1.19. 某批产品共有 10 件，其中 3 件次品，7 件正品. 若以不放回抽样从中先后任取两件, 若第一件取得的是次品, 求第二件又取到次品的概率.

解. 记 A=" 第一件取得次品"， B =" 第二件取得次品". 下求 $P(B \mid A)$.

法一: 由于

$$P(A) = \frac{3}{10}, \quad P(AB) = \frac{A_3^2}{A_{10}^2} = \frac{1}{15},$$

由条件概率的定义, 有

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{2}{9}.$$

法二: 由于第一件已抽到次品, 样本空间缩小为"2 件次品, 7 件正品". 从中再取一件取得次品的概率为

$$P(B \mid A) = \frac{2}{9}.$$ □

1.5.2 乘法公式

由条件概率的定义 $P(A|B) = \frac{P(AB)}{P(B)}$ 可推得乘法公式:

$$\boxed{\begin{aligned} P(AB) &= P(B)P(A \mid B) \\ &= P(A)P(B \mid A). \end{aligned}}$$

这是个常用的公式. 例如, 在 3 白 2 黑共 5 个球中随意摸取 2 个球, 记 A_i =" 第 i 次摸到白球", 则

$$P(A_1A_2) = P(A_1)P(A_2 \mid A_1) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}.$$

这正是例 1.15 中 $P(W, W)$ 的计算依据.

乘法公式可推广至多个事件的情形. 下面通过一个例子, 波利亚罐子模型, 来说明这一点. 该模型常用来模拟传染病的传染.

例 1.20 (波利亚罐子模型)**.** 设罐中混装有 a 个白球 b 个黑球. 从中随机抽取一个, 观其颜色后放回罐中, 并再加入 c 个与之同色的球, 重复此过程. 试求第一、二次取到白球且第三、四次取到黑球的概率.

解. 记 W_i =" 第 i 次取到的是白球", B_i =" 第 i 次取到的是黑球", $i = 1, 2, 3, 4$. 由乘法公式,

$$\begin{aligned} P(W_1W_2B_3B_4) &= P(W_1W_2B_3)P(B_4|W_1W_2B_3) \\ &= P(W_1W_2)P(B_3|W_1W_2)P(B_4|W_1W_2B_3) \\ &= P(W_1)P(W_2|W_1)P(B_3|W_1W_2)P(B_4|W_1W_2B_3) \\ &= \frac{a}{a+b} \cdot \frac{a+c}{a+b+c} \cdot \frac{b}{a+b+2c} \cdot \frac{b+c}{a+b+3c}. \end{aligned}$$

□

例 1.21. 在 2016 年欧洲冠军杯四分之一决赛的 8 支球队中, 有 4 支是公认的强队: 巴塞罗那、拜仁慕尼黑、皇家马德里和巴黎圣日耳曼. 假设这一轮的配对是完全随机的, 求在这一轮中没有强队对战的概率.

解. 不妨记四支强队为 T_1, T_2, T_3, T_4, 记 W_i =“球队 T_i 与四支弱队中的一支对战”. 由乘法公式, 有

$$\begin{aligned} P(W_1W_2W_3W_4) &= P(W_1)P(W_2|W_1)P(W_3|W_1W_2)P(W_4|W_1W_2W_3) \\ &= \frac{4}{7} \cdot \frac{3}{5} \cdot \frac{2}{3} \cdot 1 = \frac{8}{35} \approx 23\%. \end{aligned}$$

□

1.5.3　全概率公式与贝叶斯公式

全概率公式提供了分情况讨论某事件发生的概率的方法. 若 $A_1, A_2, \cdots, A_n$ 构成对样本空间 Ω 的一个划分, 即满足

$$A_i \cap A_j = \emptyset, \quad \bigcup_{i=1}^{n} A_i = \Omega,$$

则任意事件 $A \subset \Omega$ 有表示

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

因此, 由概率的有限可加性和乘法公式,

$$\boxed{P(A) = \sum_{i=1}^{n} P(AA_i) = \sum_{i=1}^{n} P(A_i)P(A \mid A_i).}$$

该式称为**全概率公式**.

例 1.22 (三门问题). 美国电视节目《Let's Make a Deal》的主持人蒙提·霍尔（Monty Hall）设计了一个游戏：在你面前有三扇关闭的门：门 A、门 B、门 C, 其中一扇门后是一辆汽车，另外两扇门后是山羊, 你随机选择一扇门 (例如门 A), 将获得门后的奖品。若此时主持人 (知道每扇门后是什么) 打开了另一扇你未选择且后面是山羊的门 (例如打开门 B，展示山羊), 你是否要将选择从门 A 换到剩下的未打开的门 (门 C)? 换门赢得汽车的概率是多少?

许多人认为，" 剩下两扇门中汽车随机分布，换不换门的概率都是 50%，因此无需换门". 事实是否果真如此呢?

解. 记事件 A =" 最初选择的门后是汽车", B =" 换门赢得汽车". 易见,

$$P(A) = 1/3, \quad P(B|A) = 0, \quad P(B|\overline{A}) = 1.$$

于是, 由全概率公式,

$$\begin{aligned} P(B) &= P(A)P(B|A) + P(\overline{A})P(B|\overline{A}) \\ &= \frac{1}{3} \times 0 + \frac{2}{3} \times 1 = \frac{2}{3}. \end{aligned}$$ □

此例中最终的结果与最初的直觉相悖的原因是主持人的行为提供了额外信息, 门后有汽车的概率不再均等. 事实上, 可以计算, 若有 100 扇门, 主持人打开了 98 扇, 换门赢得汽车的概率为 99%! 这给我们一个启示: 不断利用新的信息做出改变是有利于决策的.

例 1.23. 小张从家到公司上班共有三条路可以直达 (如图1.5), 这三条路*不会发生拥堵*的概率分别为

$$P(B_1) = 0.2, \quad P(B_2) = 0.4, \quad P(B_3) = 0.7.$$

结合路况, 小张对不同路线的选择倾向性为

$$P(A_1) = 0.5, \quad P(A_2) = 0.3, \quad P(B_3) = 0.2.$$

那么, 若不为拥堵预留时间, 小张能按时上班的概率是多少?

解. 记 $A=$" 小张按时上班". 由于这三条路线都可能达成按时上班的结果, 由全概率公式,

$$\begin{aligned}P(A)&=\sum_{i=1}^{3}P(A_i)P(A\mid A_i)\\&=0.5\times0.2+0.3\times0.4+0.2\times0.7\\&=0.36.\end{aligned}$$ □

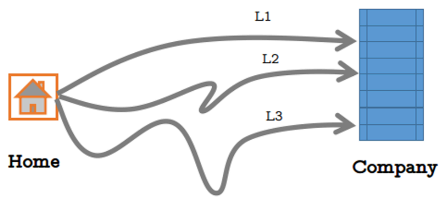

图 1.5: 小张上班可能走的路线

现在考虑反问题: 如果某日小张按时到达了公司, 他走的哪条路的可能性最大? 这类*由果溯因*的问题可由贝叶斯公式来解决.

若 $A_1,A_2,\cdots,A_n$ 构成对样本空间 Ω 的一个划分, 且 $P(A_i)>0$, 则对随机事件 A, $P(A)>0$, 有

$$\boxed{P(A_i|A)=\frac{P(A_iA)}{P(A)}=\frac{P(A|A_i)P(A_i)}{\sum_{k=1}^{n}P(A_i)P(A|A_i)}.}$$

该式称为**贝叶斯公式**.

贝叶斯公式将条件概率 $P(A_i|A)$ 的计算转化为条件概率 $P(A|A_i)$ 的计算. 这在应用中用于将" 由果溯因" 的问题转化为" 由因到果" 的问题.

例 1.24. 若某日例 1.23 中的小张按时上班了, 他是走第一条路线上班的概率是多少?

解. 由贝叶斯公式和例 1.23 中的结果,

$$P(A_1|A)=\frac{P(A_1A)}{P(A)}=\frac{P(A|A_1)P(A_1)}{P(A)}=\frac{0.2\times0.5}{0.36}\approx0.28.$$ □

例 1.25. 2020 年, COVID-19 病毒肆虐全球, 核酸检测是 COVID-19 诊断的"金标准". 据中国疾控中心 (CDC) 发布的 2020 年武汉数据, 核酸检测 (鼻拭子) 对 COVID-19 的检测灵敏度 (感染者被检出) 约为 80%, 假阳性率 (未感染者被错误检出, 通常由试验室污染或误操作所致) 约为 2%. 而据 2020 年 4 月武汉血清抗体调查显示, 约有 4.4% 的武汉居民携带 COVID-19 抗体. 若此间武汉的张先生核酸检测呈阳性, 他感染 COVID-19 的概率约为多少?

解. 记 $A=$" 感染 COVID-19", $B=$" 核酸检测呈阳性", 需计算 $P(A|B)$.

由统计数据, $P(A)=4.4\%$, $P(B|A)=80\%$, $P(B|\bar{A})=2\%$. 于是, 由贝叶斯公式,

$$\begin{aligned}P(A|B)&=\frac{P(B|A)P(A)}{P(A)P(B|A)+P(\bar{A})P(B|\bar{A})}\\&=\frac{80\%\cdot4.4\%}{4.4\%\cdot80\%+95.6\%\cdot2\%}\approx64.7\%.\end{aligned}$$ □

上例的结果是否令你感到意外? 即便核酸检测非常靠谱 (假阳性率只有 2%), 阳性检测结果也只说明有 64.7% 的概率的确感染病毒, 而非想象中的 98%. 简单计算可知, 如果假阳性率提高到 5%, 相应的条件概率即降至 42.4%.

1.5.4　事件的独立性

顾名思义, 事件 A, B 独立是指

$$\boxed{P(A \mid B) = P(A).}$$

但条件概率 $P(A \mid B) = \frac{P(AB)}{P(B)}$ 中暗含了 $P(B) > 0$. 用乘积的形式定义独立可避免此要求.

定义 1.26. 若事件 A, B 满足

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

则称两事件 A, B 相互独立, 简称**独立**.

据此定义, 零概率事件与任何事件独立.

值得指出的是, 三个事件 A, B, C 相互独立是指

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

同时成立. 三个事件 A, B, C 两两独立不等价于 A, B, C 相互独立. 看一个例子.

例 1.27. 设有四张卡片, 一张涂有红色, 一张涂有白色, 一张涂有黑色, 一张涂有红、白、黑三色. 从中任意取一张, 记 A =" 抽出的卡片上出现红色", B =" 抽出的卡片上出现白色", C =" 抽出的卡片上出现黑色". 试分析 A, B, C 的独立性.

解. 易见,

$$P(A) = P(B) = P(C) = \frac{1}{2}, \quad P(AB) = P(AC) = P(BC) = \frac{1}{4},$$

因此, A, B, C 两两独立. 然而,

$$P(ABC) = \frac{1}{4} \neq P(A)P(B)P(C),$$

因此, A, B, C 不相互独立. □

例 1.28 (于北辰飞弹问题)**.** 据台湾退役名将于北辰的消息, 台湾" 天弓" 拦截飞弹的拦截命中率约为 70%. 若三发此款飞弹 (独立) 齐射拦截同一目标, 目标被拦截的概率约是多少?

解. 记 A_i =" 第 i 发飞弹命中目标", $i = 1, 2, 3$, A =" 目标被命中", 则

$$\begin{aligned} P(A) &= 1 - P\left(\overline{A}_1\overline{A}_2\overline{A}_3\right) \\ &= 1 - P\left(\overline{A}_1\right)P\left(\overline{A}_2\right)P\left(\overline{A}_3\right) \\ &= 1 - (30\%)^3 = 97.3\%. \end{aligned}$$

□

注 1.29. 事件独立与互不相容 ($AB = \emptyset$) 是两个不同的概念:

$$\begin{aligned} \text{独立} \quad &\Leftrightarrow \quad P(AB) = P(A)P(B), \\ \text{互不相容} \quad &\Rightarrow \quad P(A \cup B) = P(A) + P(B). \end{aligned}$$

例如, 在掷一枚骰子的试验中, 事件 $A = \{1, 3, 5\}$ 与事件 $B = \{1, 2, 3, 4\}$ 是独立的, 但并非互不相容.

例 1.30. 设 $P(A) > 0, P(B) > 0$. 证明: 若 A, B 互不相容, 则 A, B 必不独立.

证明. 由于 $AB = \emptyset$, 有 $P(AB) = P(\emptyset) = 0$. 而 $P(A) > 0$, $P(B) > 0$, 故 $P(A)P(B) \neq 0$. 因此, $P(A)P(B) \neq P(AB)$, 即 A, B 不独立. □

1.6　*σ-代数与概率测度

设 Ω 是一个样本空间, $\mathscr{F}$ 是由 Ω 的一些子集构成的集合. 若 $\mathscr{F}$ 满足

(i) $\Omega \in \mathscr{F}$;

(ii) 若 $A \in \mathscr{F}$, 则 $\overline{A} \in \mathscr{F}$;

(iii) 若 $A_i \in \mathscr{F}$ $(i = 1, 2, \cdots)$, 则 $\cup_{i=1}^{\infty} A_i \in \mathscr{F}$,

则称 $\mathscr{F}$ 是 Ω 上的一个**σ-代数**或**事件域**, $\mathscr{F}$ 中的元素称为**事件**.

一个样本空间上可以定义不同的 σ-代数. 例如, 在掷一枚骰子的试验中, 样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$,

$$\begin{aligned} &\mathscr{F}_1 = \{A : A \text{ 是 } \Omega \text{ 的子集}\}, \\ &\mathscr{F}_2 = \{\emptyset,\ \{1, 3, 5\},\ \{2, 4, 6\},\ \Omega\}, \\ &\mathscr{F}_3 = \{\emptyset, \Omega\}, \end{aligned}$$

均为 Ω 上的 σ-代数. 其中, $\mathscr{F}_1$ 包含了 Ω 的所有子集, 涵盖了掷骰子试验的所有可能的事件, 是最大的 σ-代数. $\mathscr{F}_3$ 只包含不可能事件 $\emptyset$ 和必然事件 Ω, 是最小的 σ-代数. 若掷骰子只关心奇、偶性, 则 $\mathscr{F}_2$-是合适的 σ-代数: 它既包含了感兴趣的事件, 又没有包含过多的冗余信息.

指定了 σ-代数 $\mathscr{F}$ 的样本空间称为**可测空间**, 简记为 $(\Omega, \mathscr{F})$.

定义 1.31. 设 $(\Omega, \mathscr{F})$ 是一个可测空间. 若 $\mathscr{F}$ 上的实值映射 $P: \mathscr{F} \to [0,1]$ 满足

(i) $P(\Omega) = 1$;

(ii) $P(A) \geqslant 0, A \in \mathscr{F}$;

(iii) 若 $A_i \in \mathscr{F}$ $(i = 1, 2, \cdots)$ 两两互不相容, 则有

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

则称 P 是 $(\Omega, \mathscr{F})$ 上的一个**概率测度**, 简称**概率**. 此时, $(\Omega, \mathscr{F}, P)$ 称为**概率空间**, $\mathscr{F}$ 中的事件也称为 P-可测集, 简称**可测集**.

注意, P 是定义在事件域 $\mathscr{F}$ 上的, 而非直接定义在样本空间 Ω 上的. σ-代数决定了哪些子集被希望是“可测的”(即可以分配概率的事件). 在某些情况下, σ-代数的选择会影响概率测度的存在性与唯一性. 一般来说, 选取的 σ-代数越大, 构造概率测度的难度越高. 这也是引入 σ-代数, 而非一概讨论样本空间的所有子集的原因. Banach-Tarski 分球定理很好地说明了这一点.

*Banach-Tarski 分球定理

1924 年, 波兰数学家斯特凡·巴拿赫 (Stefan Banach) 和阿尔弗雷德·塔斯基 (Alfred Tarski) 提出了 Banach-Tarski 分球定理, 是测度理论和集合论中的一个经典结果:

> *在三维空间中, 存在这样的分解, 它将一个体积为 1 的实心球分解成 5 块, 然后仅通过旋转和平移这些块, 重新组合成两个 (体积和形状) 与原球完全相同的球.*

直观上, 这意味着仅通过切割再旋转组合 (未经挤压或拉伸), 1 个实心球可变为 2 个完全相同的球. 由于旋转不改变体积, 体积凭空增加了一倍! 在现实生活中这是不可能的, 然而数学上却得到了证明, 何处有问题?

实际上, 上述分球理论中分解成的这个 5 子块具有极其复杂的分形结构边界, 传统的" 体积" 概念 (即 Lebesgue 测度, 当限制在体积为 1 的球内时即是一个特殊的概率测度) 对它们无效, 不满足可加性. 换言之, 这 5 个子块不是可测集, 它们无体积 (注意, 无体积和体积为零是两码事)! 无体积的物体在物理世界中并不存在, 因为物理世界中的物体结构是离散的, 不能无限细分. 因此, Banach-Tarski 分球定理与物理的质量守恒律并不矛盾.

这个例子说明数学真理可以超越物理直觉. 它同时也警示我们, 为避免反直觉现象, 在概率论中需排除不可测集, 即应选择合适的 σ-代数而非一概对样本空间的所有子集进行讨论.