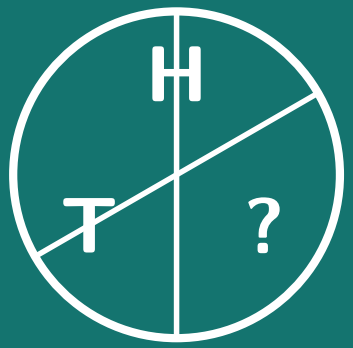


概率论与数理统计

第一章：概率论的基本概念

随机试验　样本空间　概率　条件概率　独立性

叶文杰

yewenjie@fjnu.edu.cn

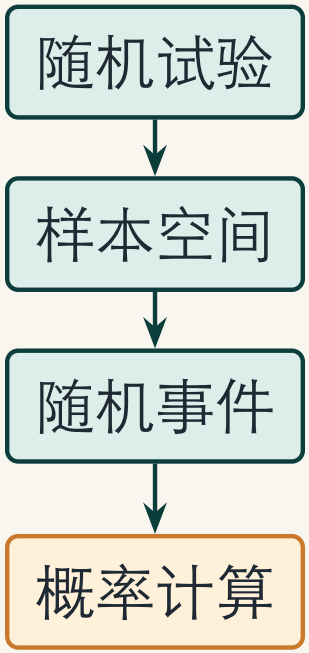
yewenjie.org

2026 年秋季学期

目的：从“不确定”到“可计算”

本章主线

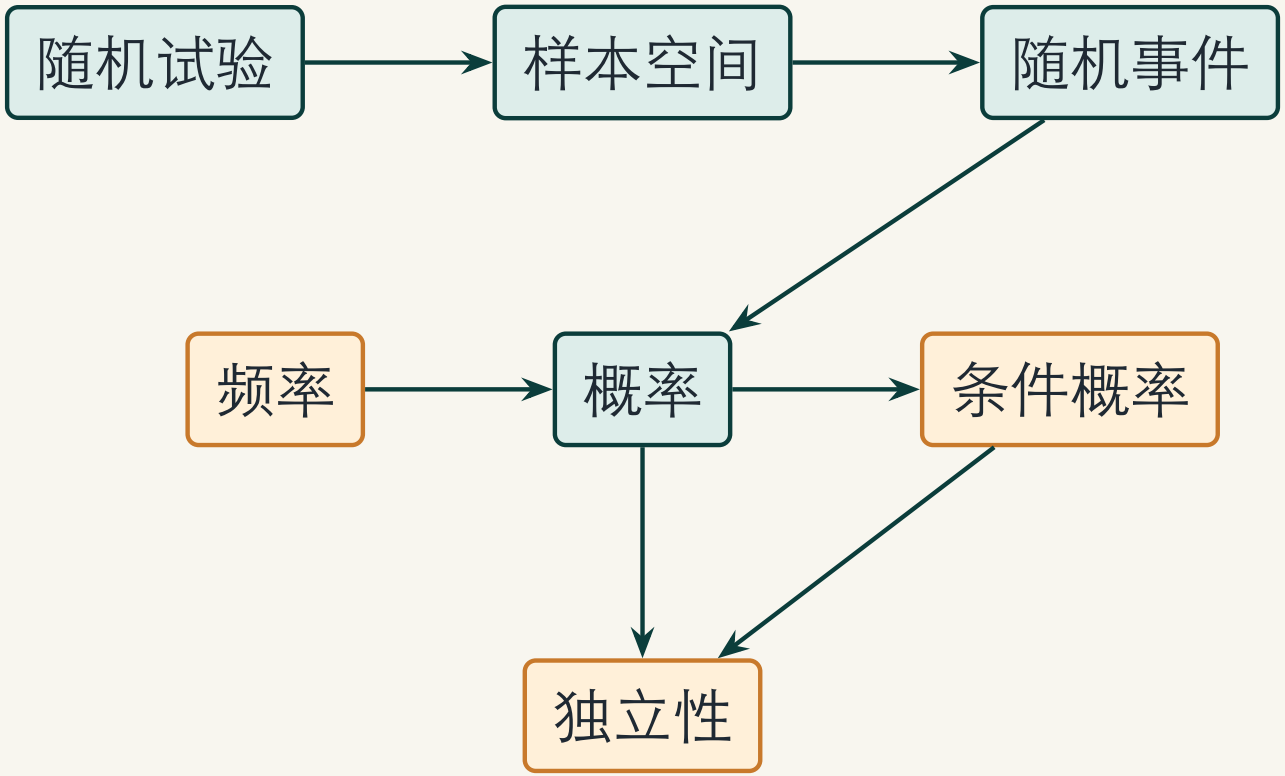
1. 明确一次随机试验会产生哪些结果；
2. 用集合语言描述我们关心的事件；
3. 给事件分配概率，并建立计算规则；
4. 处理“已知信息”和“相互独立”的情形。



提醒

公式不是第一步。第一步通常是：
把问题翻译成样本空间和事件。

第一章内容



本章的目标不是记住术语，而是形成一套稳定的建模语言。

1

随机试验

第一章　概率论的基本概念

随机现象与确定性现象

确定性现象

在给定条件下，结果可以事先确定。

- 太阳东升西落；
- 标准大气压下水加热到 $100°C$ 会沸腾；
- 异性电荷相吸。

随机现象

个别结果不确定，但大量重复后往往呈现规律。

- 抛硬币、掷骰子；
- 电话交换台一分钟内来电数；
- 天气、金融市场、保险风险。

总结

概率论研究的是随机现象背后的**统计规律性**。

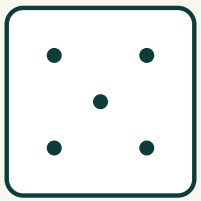
从观察试验开始

试验的广义含义

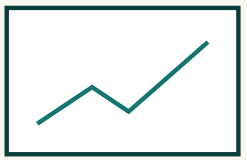
概率论中的“试验”不只指实验室操作，也包括一次观察、一次抽样、一次记录。



抛硬币



掷骰子



记录数据

随机试验的三个特点

定义

具有下列特点的试验称为**随机试验**，通常记为 E。

1. **可重复性**：可以在相同条件下重复进行；
2. **明确性**：每次试验的可能结果不止一个，且结果范围事先明确；
3. **不确定性**：试验前不能确定哪一个结果会出现。

问题

“明天下雨吗?”是不是随机试验？怎样把它说成一个可重复、可观察的试验?

几个典型随机试验

- 抛一枚硬币，观察正面 H 和反面 T；
- 抛一枚硬币三次，观察 H 出现的次数；
- 掷一颗骰子，观察出现的点数；
- 任取一支灯泡，测试它的寿命；
- 记录某电话交换台一分钟内接到的呼叫次数。

观测结果可以有不同形式

类别型
如 H, T

离散数值型
如 $0, 1, 2, 3$ 或 $1, \ldots, 6$

连续数值型
如灯泡寿命 $t \geq 0$

共同点：在试验进行之前，不能确定究竟哪一个可能结果会出现。

2

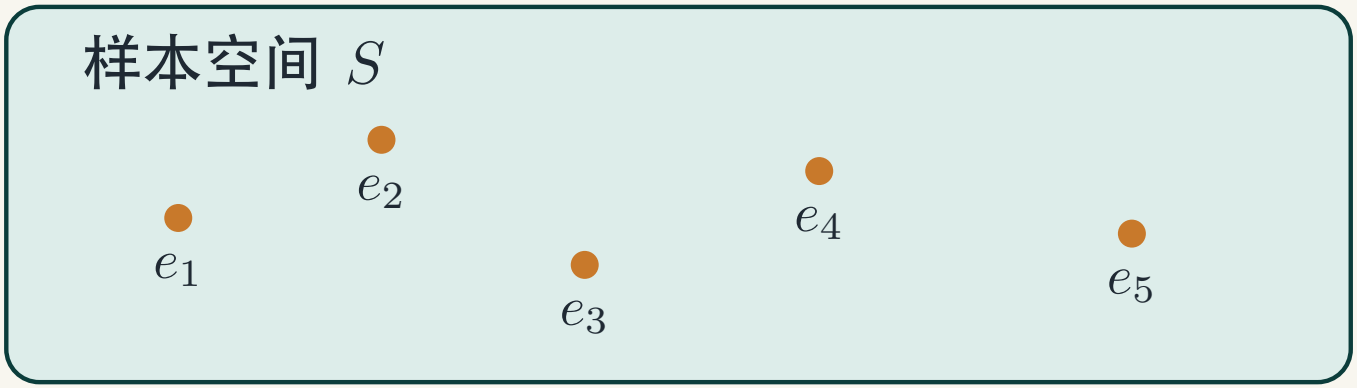
样本空间与随机事件

第一章　概率论的基本概念

样本空间：把结果放进一个集合

定义

随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**，记为 S。样本空间中的元素称为**样本点**。



易错点

样本空间由**观察目的**决定。同一个物理试验，观察内容不同，样本空间也可能不同。

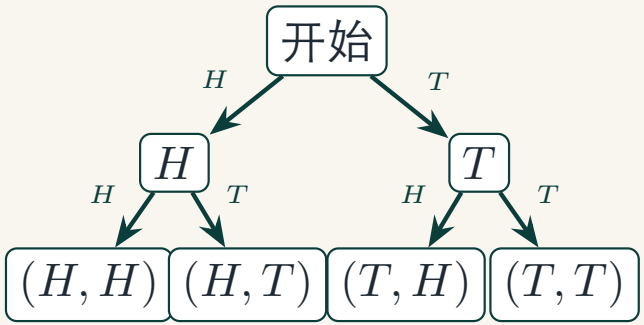
样本空间例子：抛硬币两次

观察每次正反面：

$$S = \{(H,H),(H,T),(T,H),(T,T)\}.$$

观察正面出现次数：

$$S' = \{0,1,2\}.$$



要领

先想“观察什么”，再写 S。

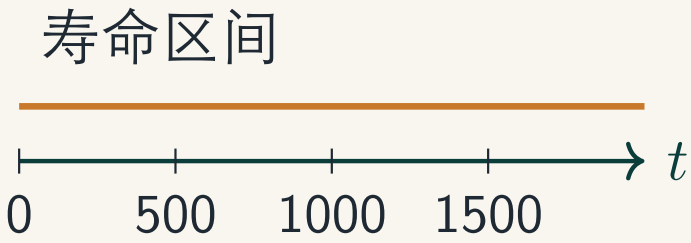
连续型样本空间例子

测试一只灯泡的寿命，样本点可以看作非负实数：

$$S = \{t : t \geq 0\} = [0, \infty).$$

记录一分钟内来电次数：

$$S = \{0, 1, 2, \ldots\}.$$



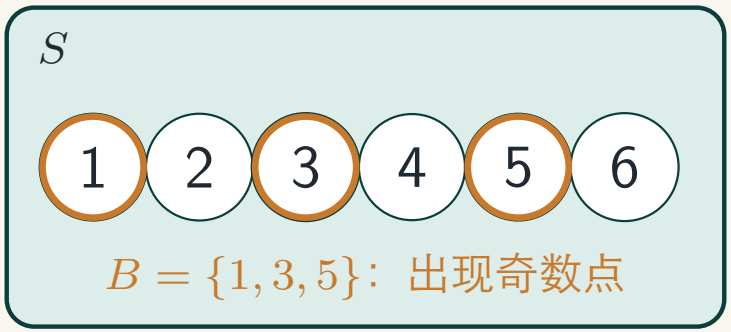
随机事件：样本空间的子集

定义

试验 E 的样本空间 S 的子集称为 E 的**随机事件**，简称事件，常用 A, B, C 表示。

掷骰子，$S = \{1, 2, 3, 4, 5, 6\}$。

$$A = \{1\},$$
$$B = \{1, 3, 5\},$$
$$C = \{5, 6\}.$$



基本事件、必然事件、不可能事件

基本事件

由一个样本点组成，或相对于观察目的不可再分。

$$\{i\}, \quad i = 1, \ldots, 6.$$

必然事件

每次试验中一定发生，记为 S。

$$\{1, 2, 3, 4, 5, 6\}.$$

不可能事件

一次试验中不可能发生，记为 $\varnothing$。

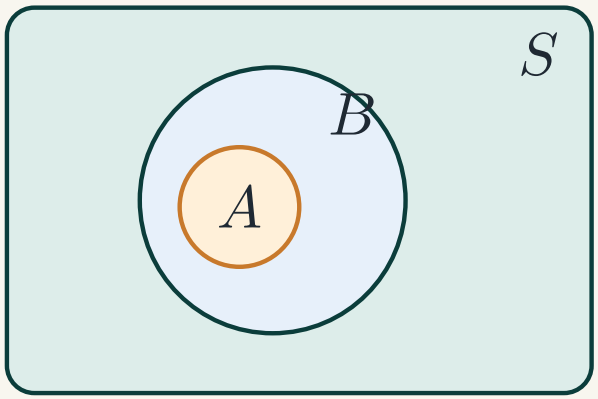
$$\{8\} = \varnothing.$$

注意：基本事件是否“不可再分”，依赖于我们观察的层次。

事件关系：包含与相等

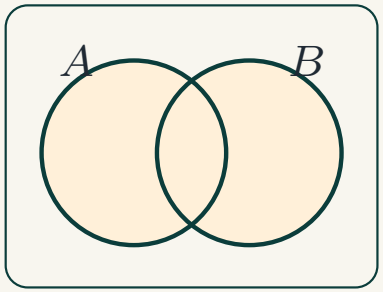
设 $A, B \subset S$。

- 若 A 发生必然导致 B 发生，则 $A \subset B$；
- 若 $A \subset B$ 且 $B \subset A$，则 $A = B$；
- 对任何事件 A，有 $\varnothing \subset A \subset S$。

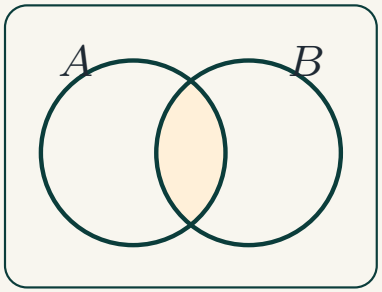


事件运算：和、积、差、补

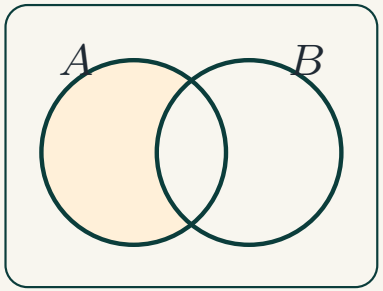
- 和事件：$A \cup B$，表示 A, B 中至少有一个发生；
- 积事件：$A \cap B$，表示 A, B 同时发生，常记为 AB；
- 差事件：$A - B = A \cap \overline{B}$，表示 A 发生而 B 不发生；
- 对立事件：$\overline{A} = S - A$，表示 A 不发生。



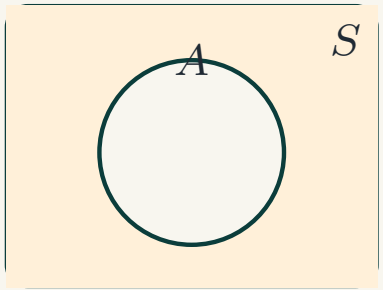
$A \cup B$



$A \cap B$



$A - B$

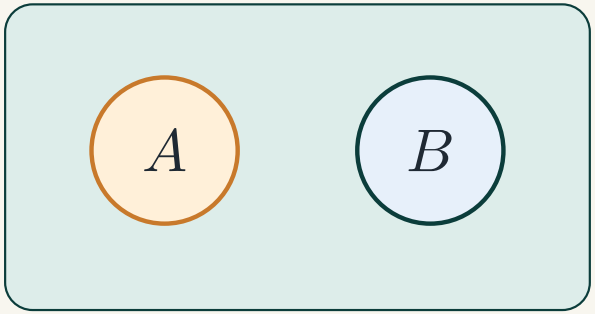


$\overline{A}$

互斥与对立

互斥事件

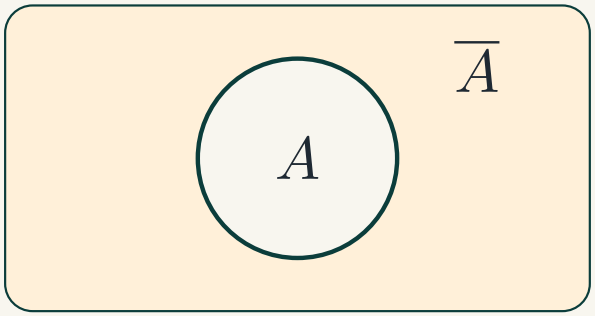
若 $A \cap B = \varnothing$，则 A 与 B 互斥，即不能同时发生。



互斥但未必对立

对立事件

若 $A \cap B = \varnothing$ 且 $A \cup B = S$，则 A 与 B 互为对立事件。



对立一定互斥

事件运算规律

交换律与结合律

$$A \cup B = B \cup A, \qquad AB = BA.$$

$$(A \cup B) \cup C = A \cup (B \cup C).$$

$$(AB)C = A(BC).$$

分配律

$$A(B \cup C) = AB \cup AC.$$

$$A \cup (BC) = (A \cup B)(A \cup C).$$

德 Morgan 公式

$$\overline{A \cup B} = \overline{A} \cap \overline{B}.$$

$$\overline{AB} = \overline{A} \cup \overline{B}.$$

要点

“至少一个发生”的反面是“一个都不发生”；“都发生”的反面是“至少一个不发生”。

例：用 A, B, C 表示复合事件

设 A, B, C 表示三个随机事件。

- A 出现，B, C 不出现：

$$A \cap \overline{B} \cap \overline{C}.$$

- A, B 出现，C 不出现：

$$A \cap B \cap \overline{C}.$$

- 三个事件都出现：

$$A \cap B \cap C.$$

- 三个事件至少一个出现：

$$A \cup B \cup C.$$

- 三个事件都不出现：

$$\overline{A} \cap \overline{B} \cap \overline{C}.$$

- 不多于一个出现：

$$(\overline{A} \cap \overline{B} \cap \overline{C}) \cup (A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C).$$

- 至少两个出现：

$$(A \cap B) \cup (A \cap C) \cup (B \cap C).$$

- 恰好两个出现：

$$(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C).$$

概率论与集合论的对应

概率论语言	集合论语言	记号
样本空间、必然事件	全集	S
不可能事件	空集	$\varnothing$
基本事件	元素或单点集	$e, \{e\}$
随机事件	子集	$A \subset S$
事件 A 的对立事件	补集	$\overline{A}$
A 发生必导致 B 发生	子集关系	$A \subset B$
A 与 B 同时发生	交集	$AB = A \cap B$
A 与 B 至少一个发生	并集	$A \cup B$

本节小结

随机试验 E 给出样本空间 S；事件是 S 的子集；事件运算就是集合运算。

3

频率与概率

第一章　概率论的基本概念

为什么需要概率?

事件 $\leftrightarrow$ 可能性大小

研究随机现象，不仅要知道“会发生哪些事件”，更要知道“它们发生的可能性有多大”。

保险	服务配置	质量控制
估计风险，确定保费与赔付准备。	估计顾客到达量，安排窗口和人员。	抽样验收，控制次品风险。

概率是对随机事件可能性大小的**数学**描述。

频率的定义

定义

在相同条件下重复进行 n 次试验，事件 A 发生 n_A 次，则

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

称为事件 A 在这 n 次试验中的**频率**。

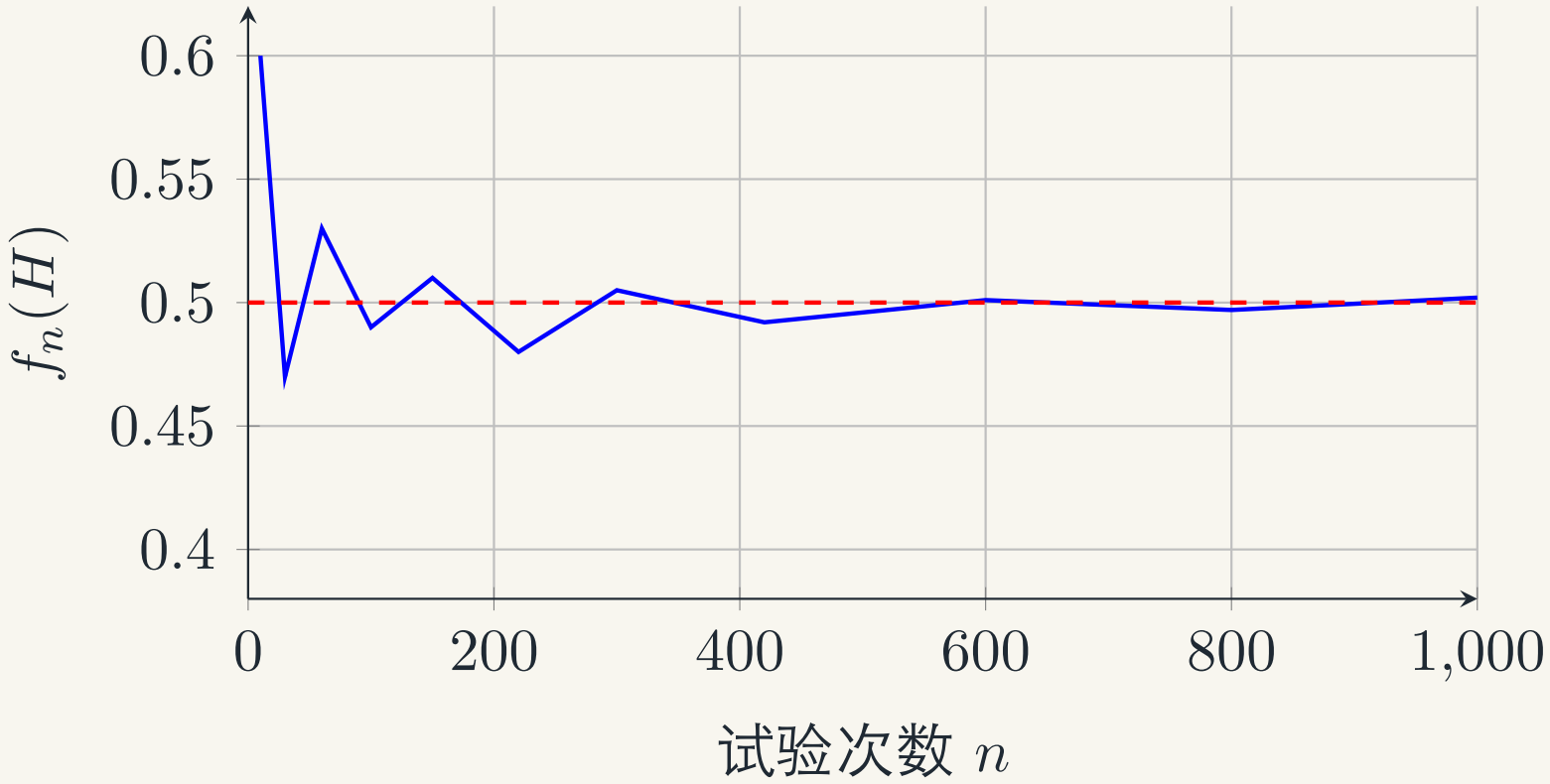
频率的基本性质：

- $0 \leq f_n(A) \leq 1$；
- $f_n(S) = 1$；
- 若 $A \cap B = \varnothing$，则 $f_n(A \cup B) = f_n(A) + f_n(B)$。

直觉

n 很大时，许多随机试验的频率会在某个数附近摆动并逐渐稳定。

频率稳定性：抛硬币的模拟

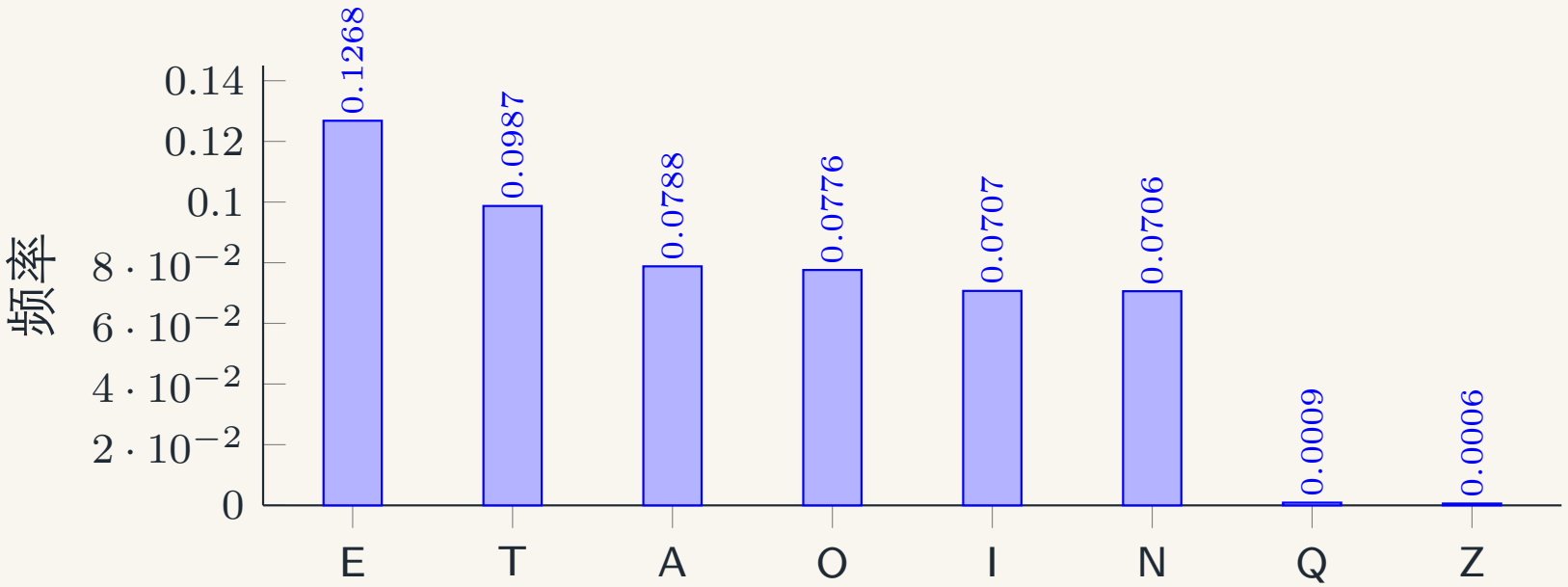


结论

频率不是每次都等于 0.5，但重复次数增加后，它常常围绕 0.5 稳定。

字母频率：随机中也有稳定结构

Dewey 统计了大量英文单词后发现，字母出现频率差异明显。下面选取若干高频与低频字母作比较。



单次观察具有随机性，但大量重复观察中往往会呈现稳定的频率规律。

概率的公理化定义

Kolmogorov 公理化观点

设 S 是样本空间，$\mathcal{F}$ 是事件族。若函数 $\mathbb{P}: \mathcal{F} \to [0,1]$ 满足：

1. 非负性：$\mathbb{P}(A) \geq 0$；
2. 规范性：$\mathbb{P}(S) = 1$；
3. 可列可加性：若 $A_1, A_2, \ldots$ 两两互斥，则

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

则称 $(S, \mathcal{F}, \mathbb{P})$ 为概率空间；$\mathbb{P}(A)$，称为事件 A 的概率，其中 $A \in \mathcal{F}$

公理的直接推论

基本性质

$$\mathbb{P}(\varnothing) = 0, \qquad \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A).$$

若 $A \subset B$，则

$$\mathbb{P}(B - A) = \mathbb{P}(B) - \mathbb{P}(A), \qquad \mathbb{P}(A) \leq \mathbb{P}(B).$$

加法公式

对任意两个事件 A, B，

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

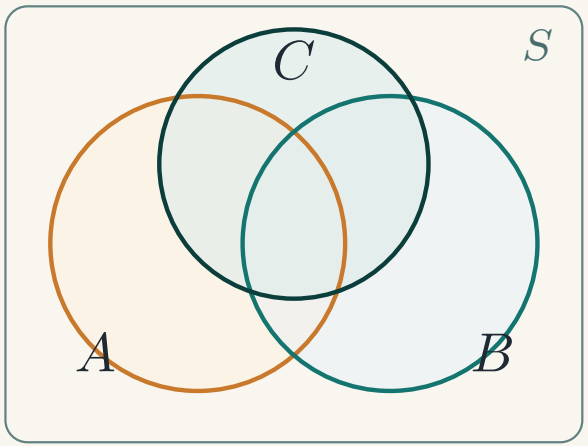
特别地，若 $A \cap B = \varnothing$，则

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

三个事件的加法公式

容斥公式

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\ + \mathbb{P}(A \cap B \cap C).\end{aligned}$$



1. 先计算

$$\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C).$$

两两交集被算了两次，三者公共部分被算了三次。

2. 再减去三个两两交集。此时两两交集恰好算一次，但三者公共部分变成

$$3 - 3 = 0.$$

3. 最后加回

例：订阅报纸和杂志

设 $A =$ “订阅报纸”，$B =$ “订阅杂志”，已知

$$\mathbb{P}(A) = 0.7, \qquad \mathbb{P}(B) = 0.2, \qquad \mathbb{P}(AB) = 0.1.$$

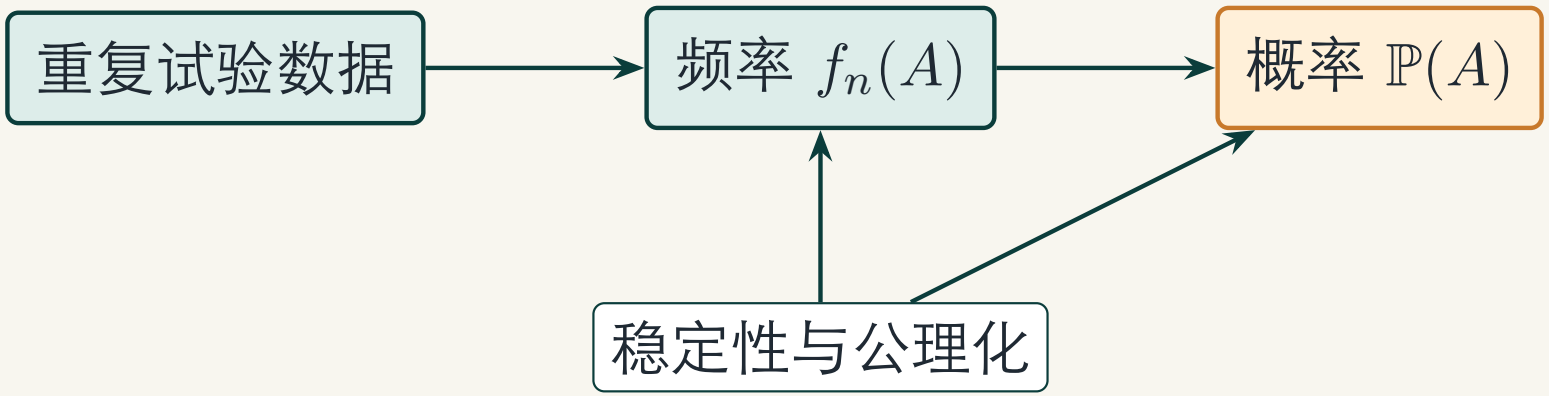
求：只订阅报纸、至少订阅一种、两种都不订阅的概率。

$$\begin{aligned} \mathbb{P}(A\overline{B}) &= \mathbb{P}(A) - \mathbb{P}(AB) = 0.6, \\ \mathbb{P}(A \cup B) &= 0.7 + 0.2 - 0.1 = 0.8, \\ \mathbb{P}(\overline{AB}) &= 1 - \mathbb{P}(A \cup B) = 0.2. \end{aligned}$$

要领

先设事件，再翻译成概率语言，最后代公式。

频率与概率小结



本节要点

概率是 0 到 1 之间的数量指标；概率运算的基础是集合运算；互斥时可以直接相加，不互斥时要减去重复计算的交集。

作业

第六版: P_{28}, 习题一 1, 2
第五版: P_{25}, 习题一 1, 2

4

古典概型与几何概型

第一章　概率论的基本概念

古典概型：有限且等可能

定义

如果随机试验的样本空间只有有限个基本事件

$$S = \{e_1, e_2, \ldots, e_N\},$$

且每个基本事件发生的可能性相同，则称该模型为**古典概型**。

古典概率公式

若事件 A 含有 k 个基本事件，也即 $A = \{e_{i_1} \cdots e_{i_k}\}$，则

$$\mathbb{P}(A) = \frac{k}{N} = \frac{|A|}{|S|}.$$

古典概型的核心：计数

计算 $\mathbb{P}(A)$ 的步骤

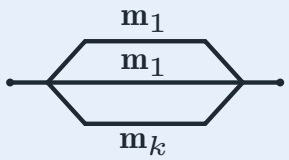
1. 样本空间中共有多少个等可能结果?
2. 事件 A 对应多少个有利结果?

常见错误

没有分清**排列**与**组合**，或者把“不放回”当成“有放回”。

计数原理

加法原理：整个过程，只需要其中任何一步即可完成。



并联

$$m_1 + \cdots + m_k$$

乘法原理：整个过程需要分步完成，每步都是必要的。



串联

$$m_1 m_2 \cdots m_n$$

摸球模型：不放回与有放回

不放回

袋中 4 白 2 黑，依次摸出 2 只，求都是白球的概率：

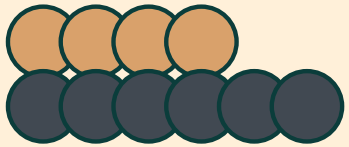
$$\mathbb{P}(A) = \frac{\binom{4}{2}}{\binom{6}{2}} = \frac{2}{5}.$$



有放回

袋中 4 红 6 黑，有放回摸球 3 次，前 2 次黑、第 3 次红：

$$\mathbb{P}(A) = \frac{6}{10} \cdot \frac{6}{10} \cdot \frac{4}{10} = \frac{18}{125}.$$



课堂练习：有序的情况

电话号码问题

在 7 位电话号码中，第一位不能为 0。求数字 0 恰好出现 3 次的概率。

$$\mathbb{P} = \frac{9 \cdot \binom{6}{3} 9^3}{9 \cdot 10^6}.$$

骰子问题

掷 3 颗均匀骰子，求点数之和为 4 的概率。

$$\mathbb{P} = \frac{3}{6^3} = \frac{1}{72}.$$

提示：先说清楚每个基本事件是否等可能。

球放 (选) 杯模型

杯子容量无限

把 4 个不同球放入 3 个杯子，求第 $1,2$ 个杯子各有两个球的概率：

$$\mathbb{P} = \frac{\binom{4}{2}}{3^4} = \frac{2}{27}.$$



每杯最多一个球

把 4 个不同球放入 10 个杯子，每杯最多一个，求第 1 至第 4 个杯子各放一个球的概率：

$$\mathbb{P} = \frac{4!}{10 \cdot 9 \cdot 8 \cdot 7} = \frac{1}{210}.$$

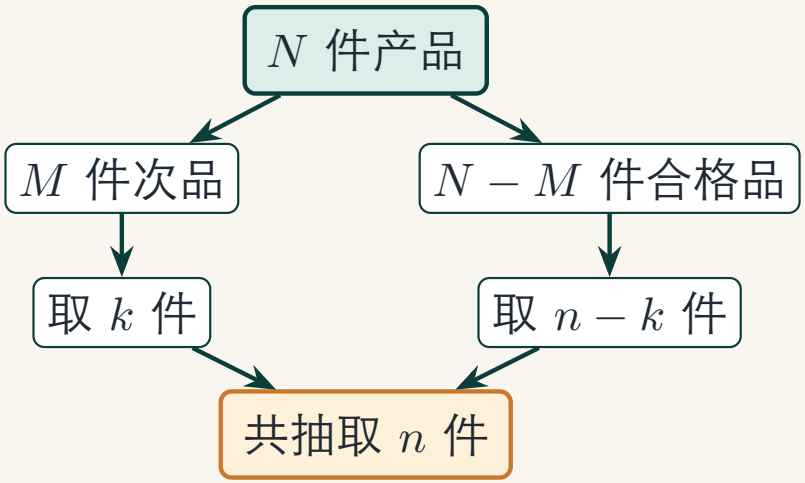


超几何分布：不放回抽样的标准模型

模型

N 件产品中有 M 件次品，从中不放回抽取 n 件。令 X 为抽到的次品件数，则

$$\mathbb{P}(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$



例：整除问题

在 $1 \sim 2000$ 的整数中随机取一个，求取到的整数既不能被 6 整除，也不能被 8 整除的概率。

解

设 $A =$ “能被 6 整除”，$B =$ “能被 8 整除”。则

$$|A| = 333, \qquad |B| = 250, \qquad |AB| = 83.$$

所以

$$\begin{aligned}
&\mathbb{P}(\overline{A} \cap \overline{B}) = 1 - \mathbb{P}(A \cup B) \\
=&1 - \Big(\mathbb{P}(A) + \mathbb{P}(A) - \mathbb{P}(A \cap B)\Big) \\
=&1 - \frac{333 + 250 - 83}{2000} = \frac{3}{4}.
\end{aligned}$$

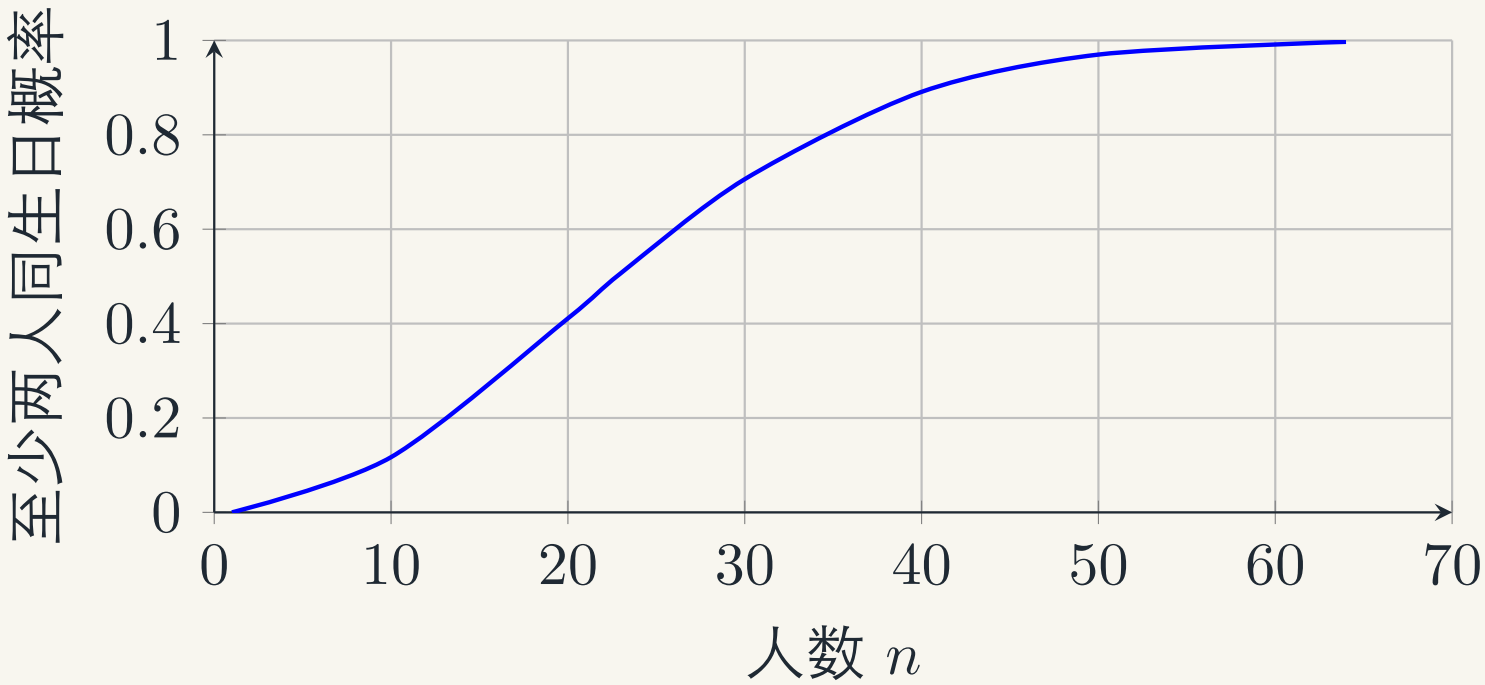
这里同时使用了古典概型与容斥原理。

生日问题:

问题形式

假设每个人的生日独立且等可能落在一年 365 天中。n 个人中至少有两人生日相同的概率为

$$1 - \frac{365 \cdot 364 \cdots (365 - n + 1)}{365^n}.$$



几何概型：连续无穷多个结果

定义

当随机试验的样本空间是某个区域，并且任一点落在该区域中是等可能的，则

$$\mathbb{P}(A) = \frac{\text{事件 } A \text{ 对应区域的度量}}{\text{样本空间 } S \text{ 的度量}}.$$

这里的度量可以是长度、面积或体积。

步骤

引入参数 $\to$ 建立坐标系 $\to$ 写出区域 S 与 A $\to$ 计算度量比。

会面问题:

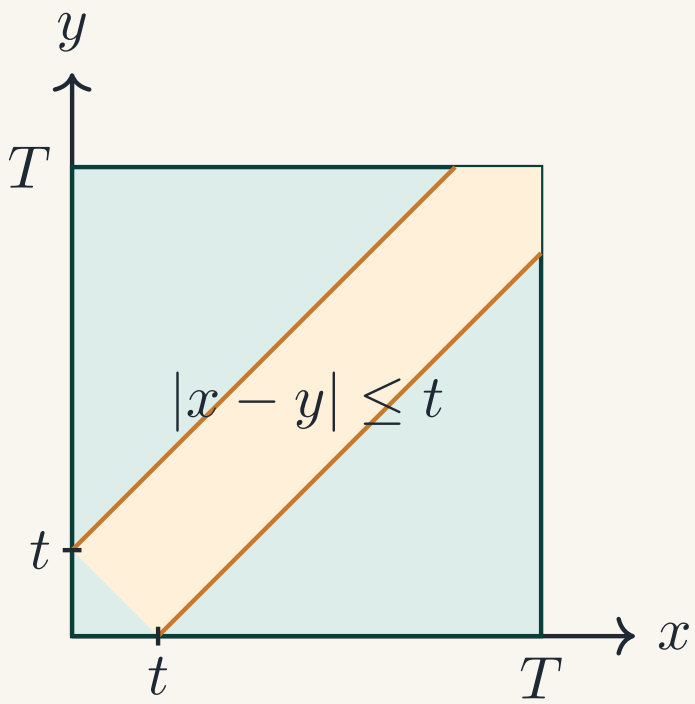
甲、乙两人相约在 0 到 T 这段时间内会面，先到者等候 t 后离开，$0 < t < T$。若两人的到达时刻独立且等可能，求甲乙两人能会面的概率。

设两人到达时刻为 $(x, y) \in [0, T]^2$，甲乙两人能会面当且仅当

$$|x - y| \leq t.$$

因此

$$\mathbb{P}(\text{甲乙能会面}) = 1 - \frac{(T-t)^2}{T^2}.$$

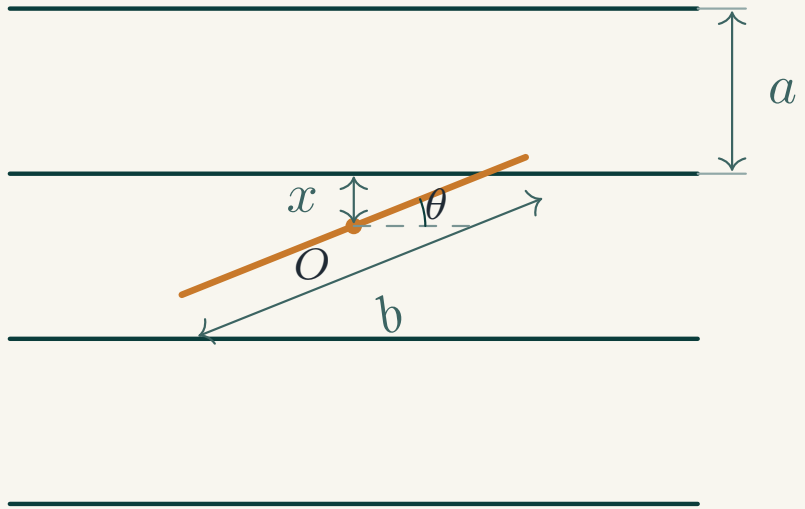


蒲丰投针：从几何概率到 π

平面上画有间距为 a 的平行直线，投一根长度 $b < a$ 的针。求针与某条平行线相交的概率。

用 x 表示针中心到最近直线的距离，用 θ 表示针相对于平行线方向的夹角。可取

$$0 \le x \le \frac{a}{2}, \qquad 0 \le \theta \le \pi.$$



相交条件是

$$x \le \frac{b}{2} \sin\theta.$$

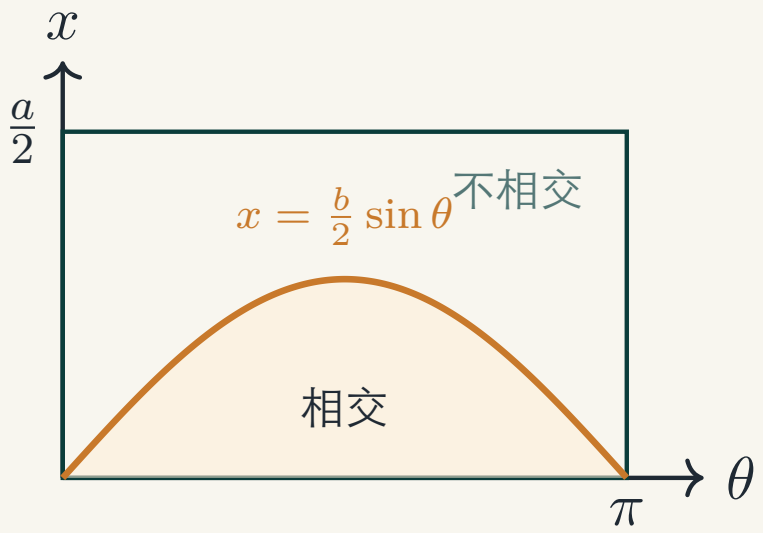
蒲丰投针：相交概率

针与平行线相交当且仅当

$$x \le \frac{b}{2}\sin\theta.$$

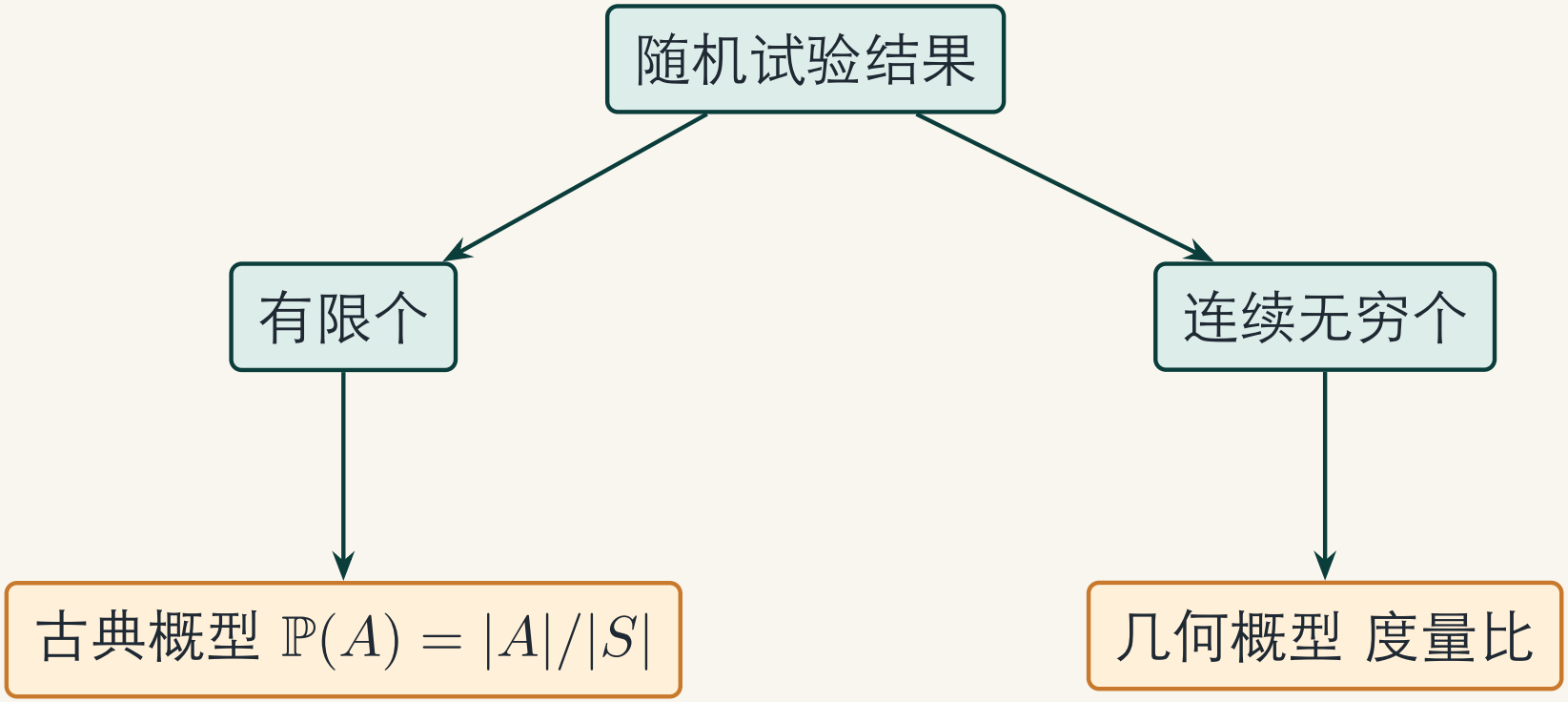
因此

$$\mathbb{P}(\text{相交}) = \frac{\displaystyle\int_0^{\pi} \frac{b}{2}\sin\theta\,\mathrm{d}\theta}{\pi \cdot \frac{a}{2}} = \frac{2b}{\pi a}.$$



概率 = 阴影面积 / 矩形面积

古典概型与几何概型小结



共同点

找到正确的“等可能基本结果”。

作业

第六版: $P_{28}-P_{29}$, 习题一 4, 6 7 11 13

第五版: $P_{25}-P_{26}$, 习题一 4, 6 7 11 13

5

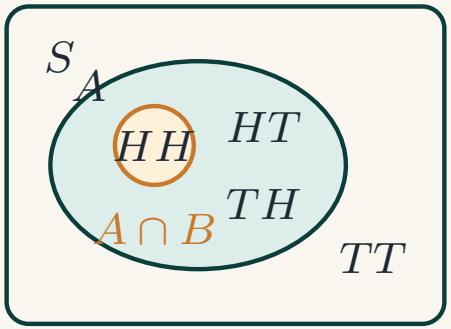
条件概率、全概率公式与贝叶斯公式

第一章　概率论的基本概念

条件概率：信息改变样本空间

将一枚硬币抛掷两次，设

$A = \{\text{至少出现一次正面}\}$，　　$B = \{\text{两次结果相同}\}$.



原样本空间为

$$S = \{HH, HT, TH, TT\}.$$

已知 A 发生后，可能结果缩减为

$$A = \{HH, HT, TH\}.$$

其中事件 B 只剩下结果 HH，因此

$$\mathbb{P}(B \mid A) = \frac{1}{3}.$$

直觉：已知 A 发生，新的样本空间缩小为 A。

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}, \qquad \mathbb{P}(A) > 0.$$

条件概率的定义与性质

定义

若 $\mathbb{P}(A) > 0$，则事件 A 发生条件下事件 B 发生的概率为

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

对固定的 A，$\mathbb{P}(\cdot \mid A)$ 仍满足概率的基本性质：

- $0 \leq \mathbb{P}(B \mid A) \leq 1$；
- $\mathbb{P}(S \mid A) = 1$；
- 若 $B_1, B_2, \ldots$ 两两互斥，则

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i \,\middle|\, A\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i \mid A).$$

例：不放回抽样中的条件概率

一盒产品中有 3 只一等品、1 只二等品。从中不放回抽取两次。令

$$A = \text{“第一次取到一等品”}, \qquad B = \text{“第二次取到一等品”}.$$

计算 $\mathbb{P}(B \mid A)$

解

在 A 已发生的条件下，盒中还剩 2 只一等品和 1 只二等品，故

$$\mathbb{P}(B \mid A) = \frac{2}{3}.$$

若用定义计算：

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)} = \frac{(3/4)(2/3)}{3/4} = \frac{2}{3}.$$

例：三个孩子问题

已知某家庭有 3 个小孩，且至少有一个女孩。求该家庭至少有一个男孩的概率。

设

$$A = \{\text{至少有一个女孩}\}, \qquad B = \{\text{至少有一个男孩}\}.$$

假设每个孩子为男孩或女孩的可能性相同，且彼此独立，则共有 $2^3 = 8$ 个等可能结果。

已知 A 发生，相当于排除 bbb，故

$$|A| = 7.$$

在 A 中，要使 B 发生，只需再排除 ggg，因此

$$|A \cap B| = 6.$$

所以

$$\mathbb{P}(B \mid A) = \frac{|A \cap B|}{|A|} = \frac{6}{7}.$$

注意

“至少有一个女孩”与“某个指定的孩子是女孩”是不同的条件。

乘法定理

由条件概率直接得到

若 $\mathbb{P}(A) > 0$，则

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

更一般地，若相关条件概率有意义，则

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

要点

当所求事件需要**分步完成**时，优先考虑乘法公式。

抽签公平性：先抽后抽一样吗?

问题

5 张同样卡片中只有一张写有“入场券”，5 人依次抽取。第 i 个人抽到入场券的概率是多少?

解

令 $A_i =$ “第 i 个人抽到入场券”。

$$\mathbb{P}(A_1) = \frac{1}{5}.$$

而

$$\mathbb{P}(A_2) = \mathbb{P}(\overline{A}_1)\mathbb{P}(A_2 \mid \overline{A}_1) = \frac{4}{5} \cdot \frac{1}{4} = \frac{1}{5}.$$

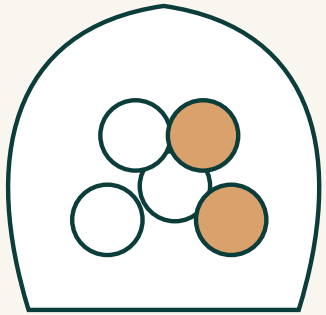
同理每个人抽到入场券的概率都是 $1/5$。

不放回不代表“不公平”；关键要把前面的人没抽到这个条件乘进去。

波里亚罐子模型

罐中有 b 个白球和 r 个红球。每次抽出一球后放回，并加入 c 个同色球。求前两次白、第三四次红的概率。

设 $W_i =$ “第 i 次取出白球”，$R_i =$ “第 i 次取出红球”。



强化效应

$$\mathbb{P}(W_1W_2R_3R_4)$$
$$= \frac{b}{b+r} \cdot \frac{b+c}{b+r+c} \cdot \frac{r}{b+r+2c} \cdot \frac{r+c}{b+r+3c}.$$

直觉

这是一个**自增强模型**：过去发生了什么，会改变下一步发生什么的概率。一次结果的出现，会使同类结果在以后更容易再次出现。

全概率公式:

样本空间的划分

若 $B_1, \ldots, B_n$ 两两互斥，且

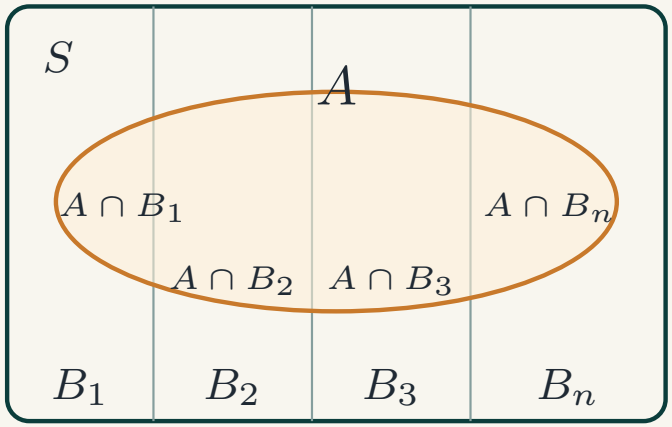
$$B_1 \cup \cdots \cup B_n = S, \qquad \mathbb{P}(B_i) > 0,$$

则 $B_1, \ldots, B_n$ 称为样本空间的一个划分。

全概率公式

对任意事件 A，

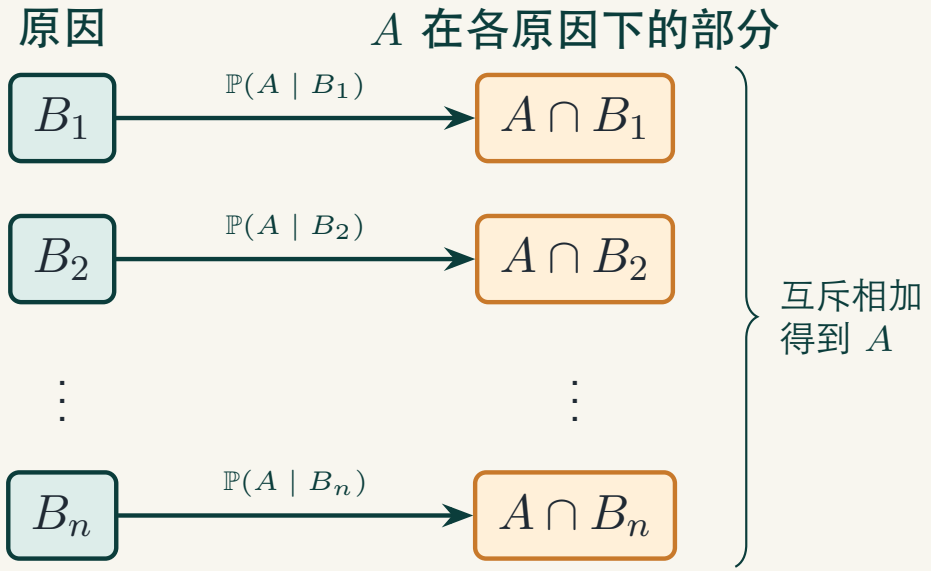
$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(B_i)\mathbb{P}(A \mid B_i).$$



$$A = (A \cap B_1) \cup \cdots \cup (A \cap B_n),$$

$$\mathbb{P}(A \cap B_i) = \mathbb{P}(B_i)\mathbb{P}(A \mid B_i).$$

全概率公式的结构图



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

当 $B_1, \ldots, B_n$ 构成样本空间的划分时，把 A 分到各个 B_i 中分别计算，再相加。

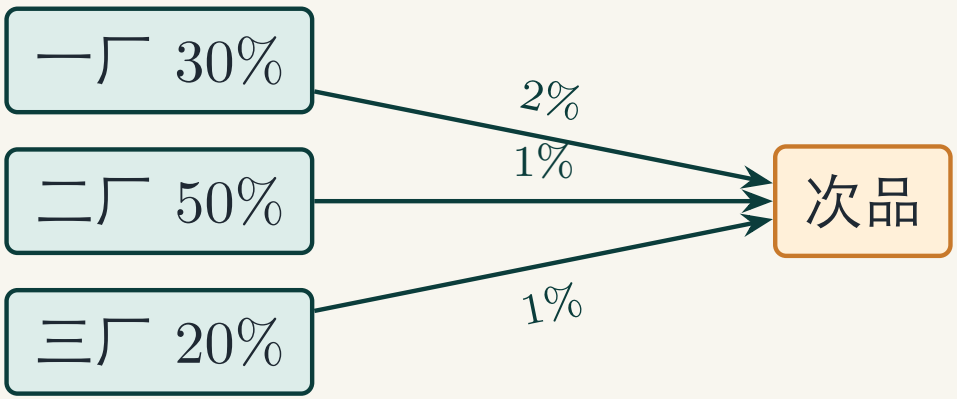
例：来自三家工厂产品的次品率

某批同型号产品由一厂、二厂、三厂生产，比例分别为 $30\%, 50\%, 20\%$，次品率分别为 $2\%, 1\%, 1\%$。求任取一件为次品的概率。

解

设 $A =$ “任取一件为次品”，$B_i =$ “产品由第 i 厂生产”。由全概率公式，

$$\mathbb{P}(A) = 0.3 \times 0.02 + 0.5 \times 0.01 + 0.2 \times 0.01 = 0.013.$$



例：飞机被击落的概率

甲、乙、丙三人同时射击，击中概率分别为 $0.4, 0.5, 0.7$。飞机被 1 人击中而击落的概率为 0.2，被 2 人击中而击落的概率为 0.6，被 3 人击中必定击落。

解

设 $A =$ “飞机被击落”，$B_i =$ “飞机被 i 人击中”。则

$$\mathbb{P}(B_1) = 0.36, \qquad \mathbb{P}(B_2) = 0.41, \qquad \mathbb{P}(B_3) = 0.14.$$

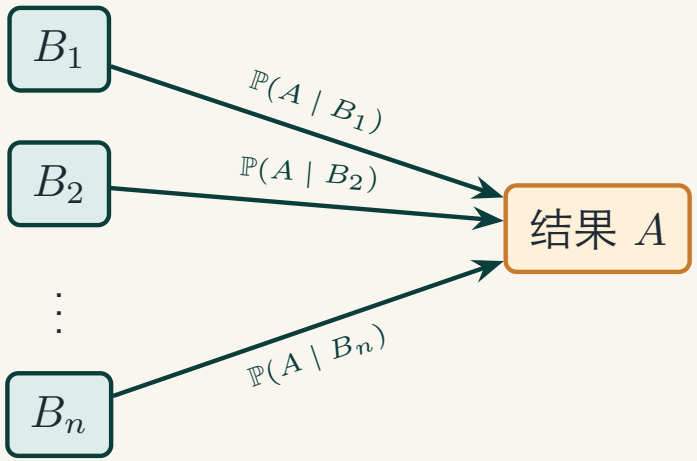
由全概率公式，

$$\mathbb{P}(A) = 0.36 \times 0.2 + 0.41 \times 0.6 + 0.14 \times 1 = 0.458.$$

全概率公式：从原因到结果

设 $B_1, \ldots, B_n$ 构成样本空间的一个划分。若事件 A 可能由这些不同原因产生，则

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

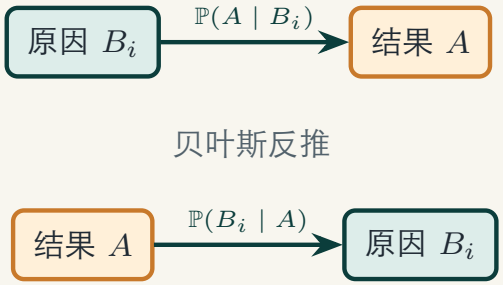


正向问题

已知各种原因 B_i 及产生结果 A 的概率 $\mathbb{P}(A \mid B_i)$，求结果 A 发生的概率。

贝叶斯公式：从结果反推原因

全概率公式描述“原因产生结果”，而贝叶斯公式解决相反的问题：已知结果 A 已经发生，反推它来自原因 B_i 的概率。



由条件概率公式，

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)}.$$

而

$$\mathbb{P}(B_i \cap A) = \mathbb{P}(B_i)\mathbb{P}(A \mid B_i),$$

并且由全概率公式，

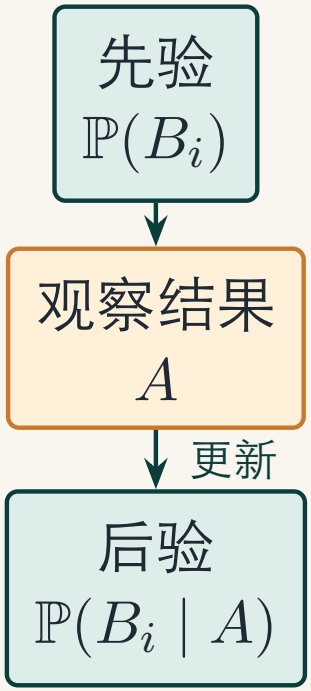
$$\mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(B_j)\mathbb{P}(A \mid B_j).$$

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

贝叶斯公式：从先验到后验

三个关键量

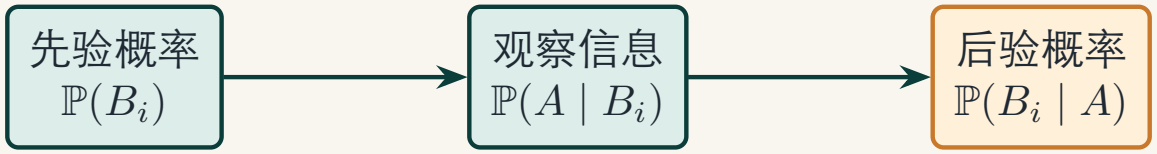
- $\mathbb{P}(B_i)$：原因 B_i 原来的可能性，称为**先验概率**；
- $\mathbb{P}(A \mid B_i)$：在原因 B_i 下观察到结果 A 的可能性；
- $\mathbb{P}(B_i \mid A)$：观察到 A 后，对原因 B_i 重新修正得到的概率，称为**后验概率**。



核心思想

贝叶斯公式不是在计算一个新的“结果概率”，而是在获得新信息 A 后，重新评价各个原因 B_i 的可能性。

先验概率与后验概率



直观

先验概率描述观察结果 A 之前对原因 B_i 的判断；观察到 A 后，利用 $\mathbb{P}(A \mid B_i)$ 提供的信息对这种判断进行修正，得到后验概率 $\mathbb{P}(B_i \mid A)$。

核心

贝叶斯公式把正向模型与反向推断联系起来：

$$\begin{aligned} &\text{正向模型：} && B_i \longrightarrow A, && \mathbb{P}(A \mid B_i), \\ &\text{反向推断：} && A \longrightarrow B_i, && \mathbb{P}(B_i \mid A). \end{aligned}$$

贝叶斯公式改变的是信息更新方向

设

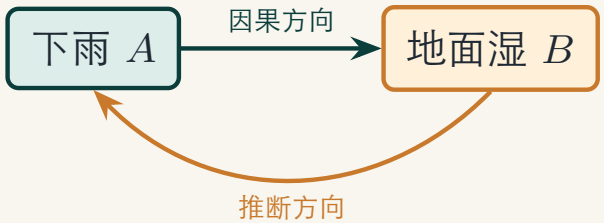
$$A = \{\text{下雨}\}, \qquad B = \{\text{地面湿}\}.$$

信息更新是对称的

由条件概率公式，

$$\frac{\mathbb{P}(B \mid A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \mid B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)}.$$

因此，如果知道“下雨”会使“地面湿”的可信度提高若干倍，那么知道“地面湿”也会使“下雨”的可信度提高同样倍数。



区别

看到地面湿，会提高“下雨”的可信度：

$$\mathbb{P}(A \mid B) > \mathbb{P}(A).$$

但这并不表示地面湿导致下雨。贝叶斯公式改变的是**信息更新方向**，不是**因果方向**。

例：低患病率下的阳性结果

某病在人群中的患病率为 0.5%。检测灵敏度为 95%，假阳性率为 5%。若某人检测阳性，求其实际患病的概率。

解

设 $D =$ “患病”，$A =$ “检测阳性”。则

$$\mathbb{P}(D \mid A) = \frac{\mathbb{P}(D)\mathbb{P}(A \mid D)}{\mathbb{P}(D)\mathbb{P}(A \mid D) + \mathbb{P}(\overline{D})\mathbb{P}(A \mid \overline{D})}.$$

代入得

$$\mathbb{P}(D \mid A) = \frac{0.005 \times 0.95}{0.005 \times 0.95 + 0.995 \times 0.05} \approx 0.087.$$

直观

阳性并不等于几乎一定患病。低基础率会显著影响后验概率。

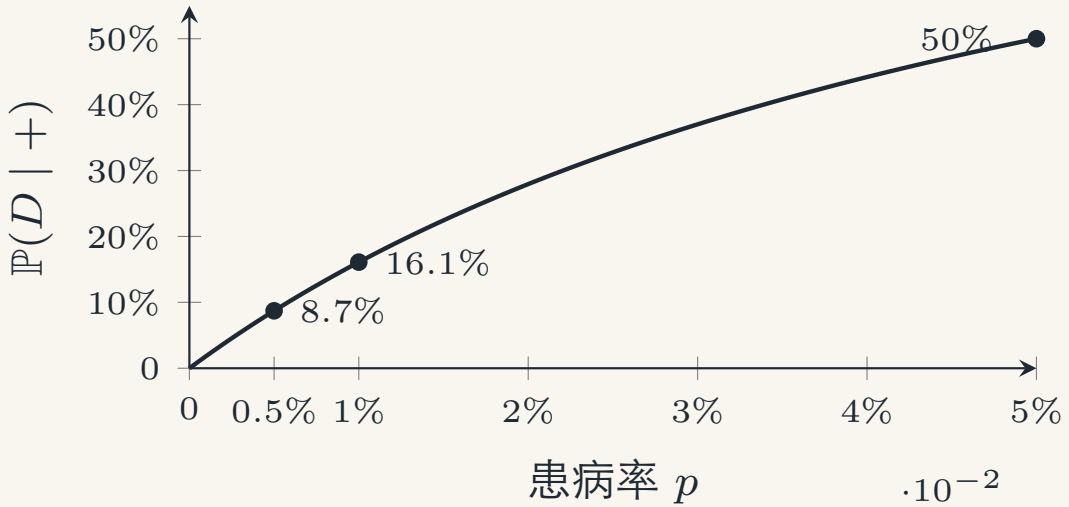
讨论：基础率如何影响后验概率

设患病率为 $p = \mathbb{P}(D)$，并假设

$$\mathbb{P}(+ \mid D) = 0.95, \qquad \mathbb{P}(+ \mid D^c) = 0.05.$$

由贝叶斯公式，

$$\mathbb{P}(D \mid +) = \frac{0.95p}{0.95p + 0.05(1-p)}.$$



观察

- 患病率仅为 0.5% 时，

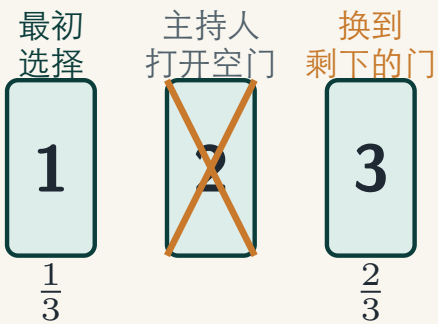
  $$\mathbb{P}(D \mid +) \approx 8.7\%.$$

- 基础率越高，阳性后的患病概率越高。
- 检测结果必须结合基础率解释。

Monty Hall：为什么应该换门?

问题

三扇门中只有一扇门后有大奖。你先选 1 号门，主持人知道大奖位置，并打开一扇你没选且没有大奖的门。问：**坚持原选择，还是换门?**



关键：

$$\text{换门中奖} \iff \text{第一次选错}$$

第一次随便选一扇门，

$$\mathbb{P}(\text{第一次选对}) = \frac{1}{3}, \qquad \mathbb{P}(\text{第一次选错}) = \frac{2}{3}.$$

主持人的作用不是改变最初概率，而是把错误门排除掉；如果你第一次选错，剩下那扇未打开的门就是大奖门。

$$\mathbb{P}(\text{坚持中奖}) = \frac{1}{3}, \qquad \mathbb{P}(\text{换门中奖}) = \frac{2}{3}.$$

结论

应该换门。换门的成功概率等于第一次选错的概率，即 $\frac{2}{3}$。

条件概率部分小结

条件概率

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}, \qquad \mathbb{P}(A) > 0.$$

已知 A 发生，在样本空间 A 中考察 B。

乘法公式

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

适合处理需要分步完成的事件。

全概率公式

若 $B_1, \ldots, B_n$ 构成样本空间的划分，则

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

由各个原因计算结果 A 的概率。

贝叶斯公式

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(B_i)\mathbb{P}(A \mid B_i)}{\mathbb{P}(A)}.$$

已知结果 A，反推原因 B_i。

作业

第六版: $P_{29}-P_{30}$, 习题一 14, 19 22 24

第五版: $P_{26}-P_{27}$, 习题一 14, 19 22 24

6

独立性

第一章　概率论的基本概念

两个事件的独立性

定义

若

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

则称事件 A 与 B 相互独立。

当 $\mathbb{P}(A) > 0$ 时，上式等价于

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

直觉

知道 A 已发生，并不改变我们对 B 发生概率的判断。

有放回与不放回：独立性 例子

袋中有 a 只黑球、b 只白球。每次取出一球，设

$$A = \{\text{第一次取出白球}\}, \qquad B = \{\text{第二次取出白球}\}.$$

有放回

每次取球后放回，因此第二次取球不受第一次影响。

$$\mathbb{P}(A) = \mathbb{P}(B) = \frac{b}{a+b}.$$

并且

$$\mathbb{P}(A \cap B) = \left(\frac{b}{a+b}\right)^2 = \mathbb{P}(A)\mathbb{P}(B).$$

因此 A 与 B 独立。

不放回

若第一次取到白球，袋中只剩 $b-1$ 只白球，因此

$$\mathbb{P}(B \mid A) = \frac{b-1}{a+b-1}.$$

于是

$$\mathbb{P}(A \cap B) = \frac{b}{a+b}\,\frac{b-1}{a+b-1}.$$

一般有 $\mathbb{P}(B \mid A) \neq \mathbb{P}(B)$，故 A 与 B 不独立。

独立性的基本性质

定理

若 A 与 B 相互独立，则下列各对事件也相互独立：

$$A \text{ 与} \overline{B}, \qquad \overline{A} \text{ 与} B, \qquad \overline{A} \text{ 与} \overline{B}.$$

只证明第一个，其他类似

$$\mathbb{P}(A\overline{B}) = \mathbb{P}(A) - \mathbb{P}(AB) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(\overline{B}).$$

独立与互斥：两种不同的关系

互斥：不能同时发生

$$A \cap B = \varnothing.$$

若 A 已经发生，则 B 一定不会发生。

独立：互不影响

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

等价地，当 $\mathbb{P}(A) > 0$ 时，

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

即知道 A 是否发生，不改变 B 的概率。

若 A, B 互斥，并且

$$\mathbb{P}(A) > 0, \qquad \mathbb{P}(B) > 0,$$

则

$$\mathbb{P}(A \cap B) = 0,$$

但是

$$\mathbb{P}(A)\mathbb{P}(B) > 0.$$

因此

$$\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B),$$

所以 A 与 B 不独立。

结论

两个正概率的互斥事件一定不独立。

三个事件：两两独立与相互独立

定义

A, B, C **两两独立**，是指

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \quad \mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C), \quad \mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C).$$

A, B, C **相互独立**，是指在两两独立的基础上，还满足

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

袋中有四个外形相同的球：三个球分别只涂红、白、黑色，另一个球同时涂有红、白、黑三种颜色。任取一球，设

$$A = \{\text{有红色}\}, \qquad B = \{\text{有白色}\}, \qquad C = \{\text{有黑色}\}.$$

此时

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2},$$

并且

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2},$$

所以 A, B, C 两两独立。

但是

$$\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C),$$

因此 A, B, C 不相互独立。

n 个事件相互独立

定义

事件 $A_1, \ldots, A_n$ 相互独立，是指对任意 $1 \leq i_1 < \cdots < i_k \leq n$，$k = 2, \ldots, n$，都有

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

注意

只验证两两独立不够；还必须验证任意多个事件同时发生的概率可以分解。

独立事件中“至少一个发生”的概率

若 $A_1, \ldots, A_n$ 相互独立，则

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = 1 - \mathbb{P}\Big(\bigcap_{i=1}^{n} \overline{A_i}\Big) = 1 - \prod_{i=1}^{n} (1 - \mathbb{P}(A_i)).$$

特别地，若 $\mathbb{P}(A_i) = p$，则

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = 1 - (1-p)^n.$$

直觉

小概率事件在一次试验中不易发生，但在大量独立重复中试验中会发生。

例：30 次掷骰子的双 6 博弈

问题

把一对骰子掷 30 次。押注 $A =$ “至少出现一次双 6”和押注 $B =$ “一次双 6 都没有”，哪个概率更大?

解

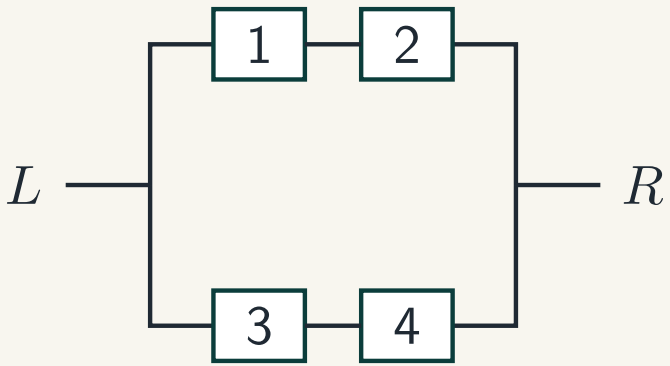
一次掷出双 6 的概率为 $1/36$。因此

$$\mathbb{P}(B) = \left(\frac{35}{36}\right)^{30} \approx 0.430, \qquad \mathbb{P}(A) = 1 - \left(\frac{35}{36}\right)^{30} \approx 0.570.$$

多次独立重复时，先算“一次都不发生”通常更简单。

例：简单系统可靠性

设各元件正常工作的事件相互独立，元件 i 正常工作的概率为 p_i。如下系统由上下两条串联支路并联而成。



上支路正常概率为 p_1p_2，下支路正常概率为 p_3p_4。系统正常概率为

$$1-(1-p_1p_2)(1-p_3p_4).$$

若 $p_i=r$，则可靠性为

$$1-(1-r^2)^2=2r^2-r^4.$$

例：抽样验收中的独立测试

问题

一批 100 件乐器中恰有 4 件音色不纯。随机抽取 3 件分别测试；若至少一件被判为音色不纯，则拒收。已知不纯乐器被检出的概率为 0.95，纯乐器被误判为不纯的概率为 0.01，且三件乐器的测试结果相互独立。求**这批乐器被接受的概率**。

设

$$A = \{\text{抽出的三件均被判为音色纯}\},$$

并令

$$H_i = \{\text{抽出的 3 件中恰有 } i \text{ 件音色不纯}\}, \qquad i = 0, 1, 2, 3.$$

由于是不放回抽样，

$$\mathbb{P}(H_i) = \frac{\binom{4}{i}\binom{96}{3-i}}{\binom{100}{3}}.$$

记

$$p = 0.99, \qquad q = 0.05.$$

在 H_i 发生时，

$$\mathbb{P}(A \mid H_i) = q^i p^{3-i}.$$

由全概率公式，

$$\mathbb{P}(A) = \sum_{i=0}^{3} \frac{\binom{4}{i}\binom{96}{3-i}}{\binom{100}{3}} q^i p^{3-i} \approx 0.8629.$$

独立性小结

本节要点

- 两事件独立：$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$；
- 独立与互斥没有必然联系；
- 两两独立不等同于相互独立；
- 独立重复中，常先算“一个都不发生”。

注意

1. 事件是否说清楚?
2. 样本空间是否等可能?
3. 给定条件后是否改变概率?
4. 独立性是题目给定，还是需要证明?

作业

第六版: $P_{31}-P_{32}$, 习题一 28, 29 34

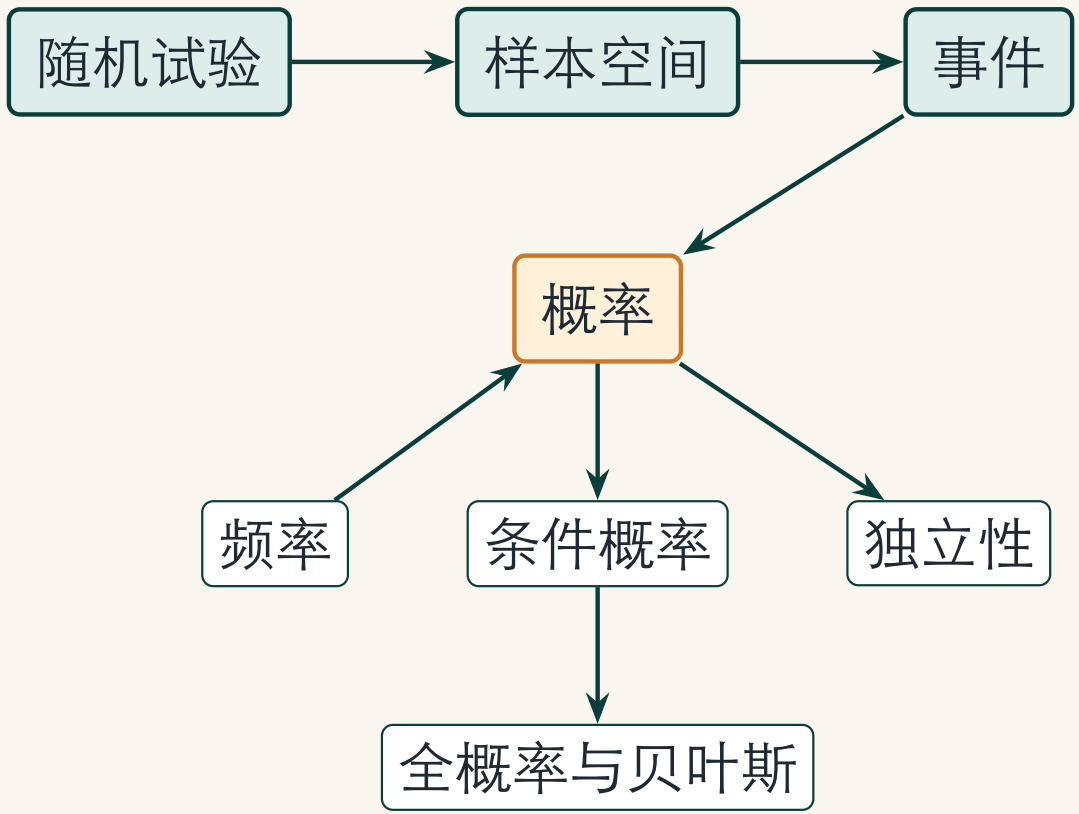
第五版: $P_{27}-P_{28}$, 习题一 28, 29 34

7

本章总结与作业

第一章　概率论的基本概念

第一章总览



要求

能够把自然语言问题转化为事件关系，并选择合适的概率计算工具。

谢谢！

下一章：随机变量及其分布