概率论与数理统计

第一章：概率 (Probability)

李立颖 (lily@sustech.edu.cn)

南方科技大学数学系

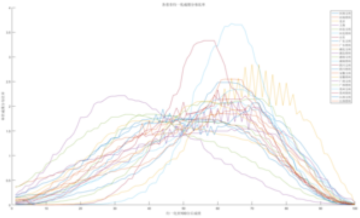
2023 秋季

第一章: 概率 (Probability)

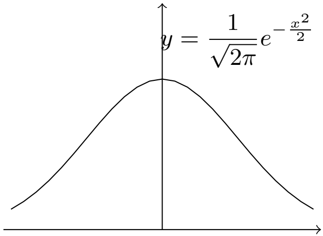
为什么学习概率论与统计?

- 现实生活中的随机现象主要来源于复杂系统:
  - 人口普查
  - 分子热运动
  - 天气预报
  - 量子力学?
- 概率论与统计提供了研究随机现象的工具
  - 概率论: 建立于现代分析学基础上, 主要提供理论基础, 得到的很多结果简洁而深刻, 如大数定律、中心极限定理、遍历定理等
  - 统计学: 解决与复杂系统相关的实际问题 (估计、判断、决策), 如参数估计、假设检验、分类、机器学习等

例 1: 正态分布



(a) 高考成绩 [1]



(b) 标准正态分布

[1] 图片来源: https://blog.csdn.net/HugoChen_cs/article/details/107650258

例 2: 核物理、随机矩阵与黎曼猜想

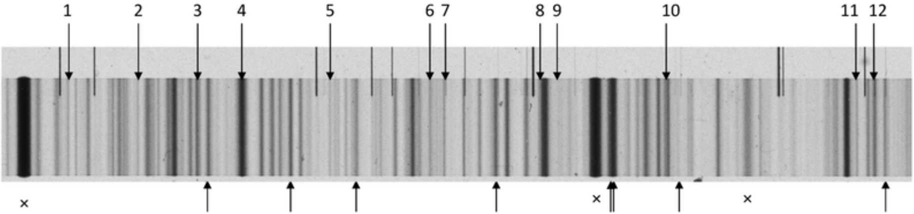


图: 铀原子能量光谱 [2]

- 薛定谔方程

$$i\frac{\partial}{\partial t}\phi = \hat{H}\phi.$$

- 能级 = 特征值

$$\hat{H}\phi = \lambda\phi, \quad \lambda \in \mathbb{R}.$$

- 重原子对应的算子 $\hat{H}$ 类似于一个高维对称矩阵 X

[2] https://www.mdpi.com/2218-2004/5/3/24

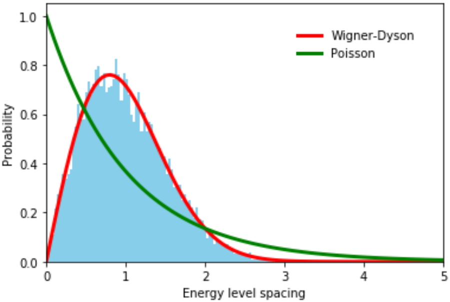
例 2: 核物理、随机矩阵与黎曼猜想, II

- 对称随机矩阵

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \ldots & & \ddots & \ldots \\ x_{N1} & x_{N2} & \cdots & x_{NN} \end{bmatrix},$$

  在满足 $x_{ij}=x_{ji}$ 条件下, 矩阵元素随机选取

- 特征值 λ 满足 $Xv=\lambda v$. 对称矩阵的特征值都是实数.
- 相邻特征值 $\lambda_{i+1}-\lambda_i$ 的分布是什么?



例 2: 核物理、随机矩阵与黎曼猜想, III

- 黎曼 ζ-函数, $\mathrm{Re}\, z > 1$
$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}.$$
以上函数可延拓至 $z \in \mathbb{C} = \{a + b \cdot i : a, b \in \mathbb{R}\}$, $i = \sqrt{-1}$.
- 黎曼猜想: 黎曼函数的全部零点在 $\{a = \frac{1}{2}\}$ 上.
- Montgomery 猜想: $\{a = \frac{1}{2}\}$ 上相邻零点的距离分布由 Dyson–Wigner 分布给出
- 小结:
  - 概率问题: 怎么定义随机性? 怎么刻画极限过程?
  - 统计问题: 怎么比较两个分布?

如何学习概率统计?

- 学思想：概率统计特殊的研究对象包含了许多独特的思维方法和思想方法，特别是如何看待和处理随机规律，这是其他学科中没有的。例如，以比较各种事件出现的可能性的大小进行决策的思想。
- 学方法：定量描述随机现象及其规律的方法，收集、整理、分析数据，从而建立统计模型的方法。
- 学应用：尽可能多地了解各种概念的背景、各种方法和模型的实际应用。不仅要学课程中提及的，也要自己收集、寻找各种实例。
- 课前预习、课中认真听讲、课后多做练习。

定义

试验：科学实验，或者对某一事物的某一特征进行观察

例

- 抛一枚硬币，观察正面 H，反面 T 出现的情况
- 将一枚硬币连抛三次，观察正面 H 出现的次数
- 掷一颗骰子，观察出现的点数
- 从一批产品中抽取 n 件，观察次品出现的数量
- 对某厂生产的电子产品进行寿命测试
- 观察某地区的日平均气温和日平均降水量

问题:

这些实验都有什么特点?

试验前无法预知结果.

随机实验与样本空间

试验的特征

- 试验可以在相同的条件下重复进行
- 试验的结果可能不止一个，但试验前知道所有可能的全部结果
- 在每次试验前无法确定会出现哪个结果

定义

具有上述特征的称为**随机试验**，或简称**试验**

例

实验 E: 掷一颗骰子，观察出现的点数.
E 的结果

- “1 点”、“2 点”、...、“6 点” (基本结果，不可分)
- “出现的点数不超过 3”、“至少出现 4 点” (复合结果，可分解)

定义

- 称基本结果为**样本点、基本事件**
- 称试验的全部样本点构成的集合为**样本空间**.

例

- 离散样本空间:
  - 掷一颗骰子, 观察出现的点数, 其样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - 抛两枚硬币, 观察正、反两面出现的情况, 其样本空间为

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

- 连续样本空间:
  - 记录深圳地区的日平均气温, 其样本空间为 $\Omega = (-60, 60)$
  - 播种飞机对位置为 (x_0, y_0) 的目标进行播种, 观察其所覆盖的范围 (x, y), 其样本空间为

$$\Omega = \{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 < \infty\}.$$

随机事件

基本结果	不可分	样本点、基本事件
复合结果	可分解	随机事件、事件

表: 试验的结果

- 从集合看: 事件是样本空间的子集
- 从试验看: 事件是基本事件的复合

定义

满足一定条件的样本点的集合称为**随机事件**, 简称**事件**. 事件用大写字母 A, B, C, $\ldots$ 等表示.

例

掷一颗骰子, 观察出现的点数, 其样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$

- 事件 A "至少出现 3 点": $A = \{3, 4, 5, 6\}$
- 事件 B "出现最小或最大点的点": $B = \{1, 6\}$

几个特殊事件:

- 基本事件: 一个样本点构成的单点集 $\{\omega\}$.
- 必然事件: 每次试验都总发生的事件 $\Omega \subset \Omega$.
- 不可能事件: 每次试验都不会发生的事件 $\varnothing$ (空集 $\varnothing \subset \Omega$).

定义

记

$$\mathcal{A} = \{A \mid A \subset \Omega,\ \text{A 是事件}\}.$$

称 $\mathcal{A}$ 为试验的事件域, 即试验产生的所有事件为元素构成的集合.

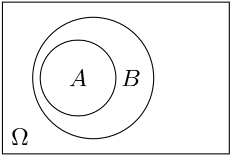
事件语言	基本结果	全体基本试验	事件 A	事件 A 发生
集合语言	样本点 ω	样本空间 Ω	子集 A	$\omega \in A$

表: 小结 — 随机试验的数学描述

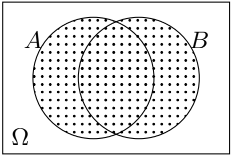
事件间的关系与运算 I

设 A, B 以及 A_k, $k = 1, 2, \ldots$ 为事件

- A 是 B 的 **子集**, 记作 $A \subset B$: A 发生必然导致 B 发生
- 集合 A 与 B **相等**, 记作 $A = B$: $A \subset B$, $B \subset A$
- 集合 A 与 B 的**并**, 记作 $A \cup B = \{\omega \mid \omega \in A$ 或 $\omega \in B\}$: A 发生或 B 发生, 即 A, B 至少有一个发生, 称为事件 A 和 B 的**和**.



(a) $A \subset B$



(b) $A \cup B$

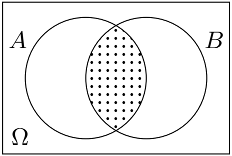
事件间的关系与运算 II

- 集合 A 与 B 的**交**, 记作 $A \cap B = \{\omega \mid \omega \in A,\ \omega \in B\}$: A , B 同时发生, 称为事件 A, B 的**积**, 也记作 AB.
- 类似地, 可以定义 n 个事件或者可列个事件的积

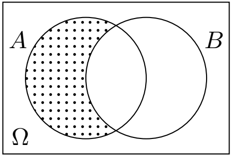
$$\bigcap_{i=1}^{n} A_i = \{\omega \mid \omega \in A_i,\ i = 1, 2, \ldots, n\}$$

$$\bigcap_{i=1}^{\infty} A_i = \{\omega \mid \omega \in A_i,\ i = 1, 2, 3, \ldots\}$$

- 集合 A 和 B 的**差集**, 记作 $A - B = A \setminus B = \{\omega \mid \omega \in A,\ \omega \notin B\}$: A 发生而 B 不发生



(a) $A \cap B$

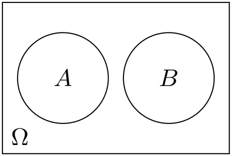


(b) $A \setminus B$

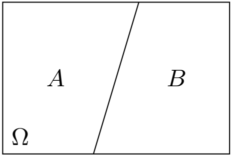
事件间的关系与运算 III

- 若 $A \cap B = \varnothing$, 则称 A, B **互不相容 (互斥)**.
- 若 $A \cup B = \Omega$, 且 $A \cap B = \varnothing$, 则称 A, B 为**逆事件**或**对立事件**, 记为

$$A = \Omega \setminus B = B^c, \quad B = \Omega \setminus A = A^c.$$



(a) $A \cap B = \varnothing$, A 与 B 互斥



(b) A 与 B 为对立事件

事件的运算定律

- 交换律:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A.$$

- 结合律:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C.$$

- 分配律:

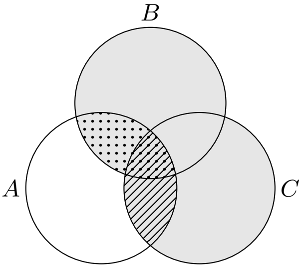
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

- De Morgan 律:

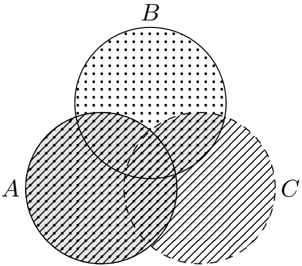
$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c,$$

$$\left(\bigcup_{k=1}^{n} B_k\right)^c = \bigcap_{k=1}^{n} B_k^c, \quad \left(\bigcap_{k=1}^{n} B_k\right)^c = \bigcup_{k=1}^{n} B_k^c.$$

分配律



(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$



(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cap C)$

如何用定义进行证明

命题

$$\left(\bigcup_{k=1}^{n} B_k\right)^c = \bigcap_{k=1}^{n} B_k^c.$$

证明

令 $P = \left(\bigcup_{k=1}^{n} B_k\right)^c$, $Q = \bigcap_{k=1}^{n} B_k^c$. 欲证 $P = Q$, 我们只需分别证明 $P \subset Q$ 与 $Q \subset P$. 我们这里只示范前者的证明. 由包含关系的定义, 对任意的 $\omega \in P$, 我们要推出 $\omega \in Q$. 事实上,

$$\begin{aligned}
\omega \in P = \left(\bigcup_{k=1}^{n} B_k\right)^c &\overset{(\text{补集的定义})}{\Longrightarrow} \omega \notin \bigcup_{k=1}^{n} B_k \\
\overset{(\text{并集的定义})}{\Longrightarrow} \omega \notin B_k,\ \forall k &\overset{(\text{补集的定义})}{\Longrightarrow} \omega \in B_k^c,\ \forall k \\
\overset{(\text{交集的定义})}{\Longrightarrow} \omega \in \bigcap_{k=1}^{n} B_k^c = Q.
\end{aligned}$$

可列

- 可列集: 指一个无穷集 S, 其元素可与自然数形成一一对应, 因此可表为 $S = \{s_1, s_2, ...\}$.
- 至多可列: 指可列或有限
- 可以证明: 可列是 "最小的" 无穷, 即任何一个无穷集合均含有可列子集

作业

- P20: 5, 6
- 补充题:
  1. 设随机事件 A, B 满足条件 $AB = A^cB^c$, 试求 $A \cup B$.
  2. 试把事件 $A_1 \cup A_2 \cup \ldots \cup A_n$ 表示成 n 个两两互不相容事件之并.

概率论公理化的三种学派

1921, J. M. Keynes, “主观概率学派”

凯恩斯主张把任何命题都看作事件, 例如 “明天将下雨”, “土星上有生命” 等等都是事件, 人们对这些事件的可信程度就是概率, 而与随机试验无关, 通常称为**主观概率**.

1928, von Mises, “客观概率学派”

米泽斯定义事件的概率为该事件出现的频率的极限, 而作为公理就必须把这一极限的存在作为第一条公理, 通常称为**客观概率**.

1933, Komolgorov (柯尔莫哥洛夫), “以测度论为基础的概率公理化体系”

目前, 绝大多数教科书都是采用柯尔莫哥洛夫的概率公理化体系.

概率论是研究随机现象的统计规律性的数学学科

问题一：什么是统计规律性？

统计规律性是指在大量试验中呈现出的**数量规律（用频率来刻画）**.

问题二：什么是概率？

概率是指刻划随机事件在一次试验中发生的可能性大小的数量指标，这个数量指标应该满足：

- 它是事件固有的，不随人们主观意愿而改变；可以在相同条件下通过大量重复试验予以识别和检验
- 符合常情：事件发生可能性大，该值就大，反之就小；不可能事件的值最小 (0)；必然事件的值最大 (1)

频率

频率的定义

设 A 为一随机事件, 在相同条件上进行 n 次重复实验. 令

$$n_A = n \text{ 次实验中 } A \text{ 发生的事件}$$
$$f_n(A) = \frac{n_A}{n}$$

称 n_A 为事件 A 的频数, $f_n(A)$ 为事件 A 的**频率 (frequency)**.

频率的一般特性

- 一般地, n 越大, 则 n_A 越大.
- n_A、$f_n(A)$ 的值是 "随机的".
- $0 \le f_n(A) \le 1$.

问题

频率是否有统计规律性?

实例 1: “抛硬币实验”

将一枚硬币连续抛 n 次, 记 $H=\{$出现正面$\}$.

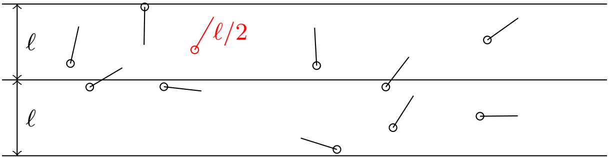
问题

$f_n(H)$ 有什么规律?

实验者	n	n_K	$f_n(H)$
德摩根	2048	1061	0.5181
蒲丰	4048	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

表: 历史上有名的抛硬币实验

实例二: “蒲丰投针实验”



记投针的总数为 n, 针与平行线相交的次数为 n_A, 则

$$\frac{n_A}{n} \approx \frac{1}{\pi}, \qquad 或\ \frac{n}{n_A} \approx \pi.$$

实例 3: 英文字母 (密码破译)

考察英语文章中 26 个字母出现的频率, 当观察次数 n 较大时, 每个字母出的频率呈现稳定性, 下面是 Deway 统计了 438023 个字母得到的统计表.

字母	频率	字母	频率	字母	频率	字母	频率
E	0.1268	R	0.0594	M	0.0244	K	0.0060
A	0.0788	L	0.0394	Y	0.0202	J	0.0010
O	0.0776	D	0.0389	G	0.0187	Q	0.0009
I	0.0707	U	0.0280	P	0.0186	Z	0.0006
N	0.0706	C	0.0268	B	0.0156		
S	0.0634	F	0.0256	V	0.0102		

随机事件的统计规律性

频率的稳定性 (大数定律)

当 n 很大时，事件 A 的频率 $f_n(A)$ 接近一个常数，即有

$$f_n(A) \to p, \quad n \to \infty$$

- 常数 p 就是事件 A 发生的可能性大小，即**概率**.
- 由于频率的取值是 "随机的"，那么上述极限是在什么意义下成立呢？(第五章研究此问题)

实例四: “掷骰子” 实验

问题

记事件

$$A_i = \{\text{ 出现 } i \text{ 点}\}, \quad i = 1, 2, \ldots, 6.$$

将一颗骰子连续掷 n 次, $f_n(A_i)$ 有什么规律?

如果一颗骰子六个面是均匀的, 则当 n 很大时应有

$$f_n(A_i) = \frac{n_{A_i}}{n} \approx \frac{1}{6}, \quad n = 1, 2, \ldots, 6.$$

频率的基本性质

1. $0 \le f_n(A) \le 1$
2. $f_n(\Omega) = 1$
3. (有限可加性) 若 $A_1, \ldots, A_m$ 两两不相容 ($A_i \cap A_j = \varnothing, \forall i \ne j$), 则

$$f_n(\cup_{i=1}^m A_i) = \sum_{i=1}^m f_n(A_i).$$

概率的公理化定义

设 $\mathcal{A}$ 为样本空间 Ω 上的事件域. 对任意 $A \in \mathcal{A}$, 若存在实数 $\mathsf{P}(A)$ 与之对应, 且满足

1. 非负性: $\mathsf{P}(A) \geq 0, \quad \forall A \in \mathcal{A}$;
2. 规范性: $\mathsf{P}(\Omega) = 1$;
3. 可列可加性: 对两两不相容的事件列 $\{A_k\}_{k=1}^{\infty}$, 有 $\mathsf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k)$

则称 $\mathsf{P}(A)$ 为事件 A 的**概率**, 称 $(\Omega, \mathcal{A}, \mathsf{P})$ 为**概率空间 (probability space)**.

概率的公理化定义

设 $\mathcal{A}$ 为样本空间 Ω 上的事件域. 对任意 $A \in \mathcal{A}$, 若存在实数 $\mathsf{P}(A)$ 与之对应, 且满足

1. 非负性: $\mathsf{P}(A) \geq 0, \quad \forall A \in \mathcal{A}$;
2. 规范性: $\mathsf{P}(\Omega) = 1$;
3. 可列可加性: 对两两不相容的事件列 $\{A_k\}_{k=1}^{\infty}$, 有 $\mathsf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k)$

则称 $\mathsf{P}(A)$ 为事件 A 的**概率**, 称 $(\Omega, \mathcal{A}, \mathsf{P})$ 为**概率空间 (probability space)**.

注:

- 1933 年苏联的 Kolmogorov 测度论基础上提出的概率论公理化体系.
- 概率是定义在事件域上的特殊函数.
- 物体的长度、区域的面积都具有“非负性”与“可加性”, 故“概率”实际上是对“事件”发生可能性大小的一种“度量”.

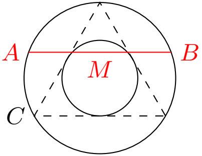
问题 (贝特朗悖论, Bertrand's Paradox)

在半径为 r 的圆 C 内“任意”作一弦. 试求此弦长度 ℓ 大于圆内接等边三角形边长 $\sqrt{3}r$ 的概率 p. (何为正确的概率空间 $(\Omega, \mathcal{A}, \mathbb{P})$?)

解一

作半径为 $r/2$ 的同心圆 C_1. 设弦 AB 的中点 M “任意”落于圆 C 内.
若 M 落于圆 C_1 内, 则 $\ell > \sqrt{3}r$, 于是

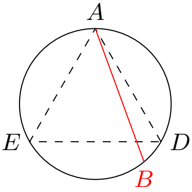
$$p = \frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4}.$$



解二

设弦 AB 的一端 A 固定于圆周上, 另一端任意. 考虑等边 $\triangle$ADE, 如 B 落于角 A 所对应的弧 DE 上, 则 $\ell > \sqrt{3}r$. 于是

$$p = \frac{DE \text{ 的弧长}}{\text{圆周长}} = \frac{1}{3}$$



概率的基本性质

性质 1

$\mathsf{P}(\varnothing)=0$.

证明: 因为 $\varnothing=\varnothing\cup\varnothing$, 所以 $\mathsf{P}(\varnothing)=\mathsf{P}(\varnothing)+\mathsf{P}(\varnothing)$, 故 $\mathsf{P}(\varnothing)=0$.

性质 2 (有限可加性)

若 $A_1, A_2, \ldots, A_n$ 是两两不相容事件, 则 $\mathsf{P}\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n}\mathsf{P}(A_k)$.

证明: 因为 $\bigcup_{k=1}^{n} A_k=A_1\cup\cdots\cup A_n\cup\varnothing\cup\varnothing\cup\cdots$. 由可列可加性,

$$\mathsf{P}\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n}\mathsf{P}(A_k)+\mathsf{P}(\varnothing)+\mathsf{P}(\varnothing)+\cdots=\sum_{k=1}^{n}\mathsf{P}(A_k).$$

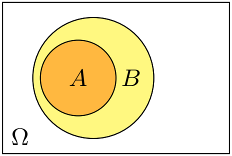
性质 3 (单调性)

若 $A \subset B$, 则 $\mathsf{P}(A) \le \mathsf{P}(B)$.

证明: 因为 $A \subset B$, 故 B 为不相容事件 A 与 $B \setminus A$ 之和. 由有限可加性,

$$\mathsf{P}(B) = \mathsf{P}(A) + \mathsf{P}(B \setminus A).$$

再由非负性, $\mathsf{P}(B \setminus A) \ge 0$. 证毕.



性质 4

$0 \le \mathsf{P}(A) \le 1.$

由 $\varnothing \subset A \subset \Omega$ 及单调性.

性质 5

$\mathsf{P}(A^c) = 1 - \mathsf{P}(A).$

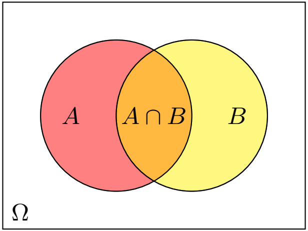
由 $A \cup A^c = \Omega$, $A \cap A^c = \varnothing$ 及有限可加性.

性质 6: 容斥原理 (inclusion-exclusion principle)

两事件容斥原理

对任何事件 A, B 有

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

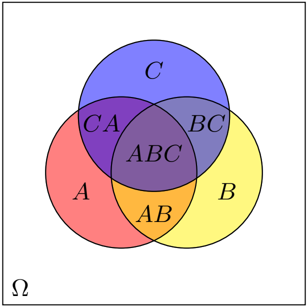


性质 6: 容斥原理 (inclusion-exclusion principle)

三事件容斥原理

对任何事件 A, B, C 有

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



性质 6: 容斥原理 (inclusion-exclusion principle)

多事件容斥原理

对于 n 个事件, 有

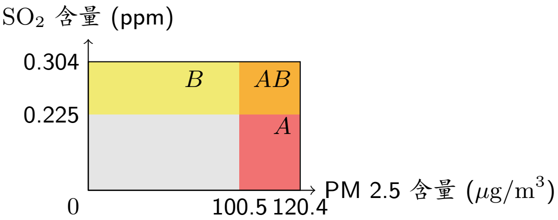
$$\begin{aligned}\mathsf{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = & \sum_{i=1}^{n} \mathsf{P}(A_i) \\ & - \sum_{1 \le i < j \le n} \mathsf{P}(A_i A_j) \\ & \sum_{1 \le i < j < k \le n} \mathsf{P}(A_i A_j A_k) \\ & + \cdots + (-1)^{n-1} \mathsf{P}(A_1 A_2 \cdots A_n).\end{aligned}$$

规律

奇加偶减.

例题

已知空气中 PM2.5 含量一般在 0.0-120.4 ($\mu g/\mathrm{m}^3$) 之间，SO_2 含量一般在 0.000-0.304 (ppm) 之间，假设在上述范围内取值为等可能的. 一般认为，PM2.5 含量在 100.5 $\mu g/\mathrm{m}^3$ 以上或 SO_2 含量在 0.225 ppm 以上为对人体有害. 问空气质量为有害的概率是多少？



解

$$\mathrm{P}(A) = 0.165, \quad \mathrm{P}(B) = 0.260, \quad \mathrm{P}(AB) = 0.043.$$
$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB) = 0.382.$$

或利用对立事件计算: $(A \cup B)^c = A^c B^c$.

作业

- P20: 4, 7
- 补充题:
  1. 设 A, B, C 是三个随机事件, 且 $\mathsf{P}(A)=\mathsf{P}(B)=\mathsf{P}(C)=\frac{1}{4}$, $\mathsf{P}(AB)=\mathsf{P}(BC)=0$, $\mathsf{P}(AC)=\frac{1}{8}$. 求 A, B, C 至少发生一个的概率.
  2. 已知 A, B 两个事件满足条件 $\mathsf{P}(AB)=\mathsf{P}(\bar{A}\bar{B})$, 且 $\mathsf{P}(A)=p$. 求 $\mathsf{P}(B)$.

古典概型

“抛硬币”、“掷骰子”等随机试验的特征:

- 只有有限个基本结果 (样本空间为有限集)
- 每个基本结果的出现是等可能的.

古典概型

设随机试验的样本空间为 Ω. 若

1. Ω 只含有有限个样本点, 即 $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$;
2. 每个样本点的出现是等可能的, 即

$$\mathsf{P}(\omega_1) = \mathsf{P}(\omega_2) = \cdots = \mathsf{P}(\omega_n) = \frac{1}{n},$$

则称该实验为**等可能概型**, 也称为**古典概型**.

问题:

怎样计算古典概型中一般事件的概率?

古典概型的概率计算

设事件 A 含 k 个样本点, 记为 $A=\{\omega_{i_1},\omega_{i_2},\ldots,\omega_{i_k}\}$. 则由有限可加性:

$$\mathsf{P}(A)=\mathsf{P}(\omega_{i_1})+\mathsf{P}(\omega_{i_2})+\cdots+\mathsf{P}(\omega_{i_k})=\frac{1}{n}\times n=\frac{k}{n}.$$

$$\mathsf{P}(A)=\frac{\text{导致 } A \text{ 发生的方式个数}}{\text{所有试验结果个数}}=\frac{k}{n}.$$

$$\mathsf{P}(A)=\frac{A \text{ 的有利场合数}}{\text{样本点总数}}=\frac{k}{n}.$$

例

抛两枚硬币，求出现一个正面一个反面的概率.

解

该实验的样本空间为

$$\Omega = \{HT, HT, TH, TT\}.$$

这是一个古典概型. 事件 $A = \{$一个正面, 一个反面$\}$ 的有利场合是 HT, TH. 因此

$$\mathsf{P}(A) = \frac{2}{4} = \frac{1}{2}.$$

历史趣闻

18 世纪著名法国数学家达朗贝尔取样本空间 $\Omega = \{HH, HT, TT\}$ 得到

$$\mathsf{P}(A) = \frac{1}{3}.$$

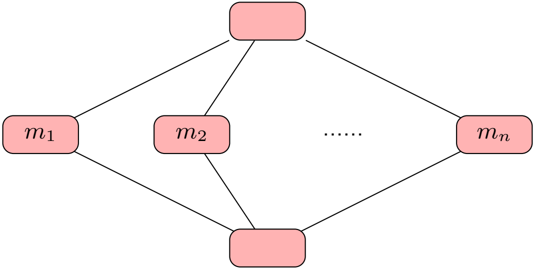
他错在了哪里? **这不是古典概型!**

加法原理 (Addition Principle)

加法原理

做一件事一共有 n 类方法, 第一类有 m_1 种方法, 第二类有 m_2 种方法, ..., 第 n 类有 m_n 种方法, 则完成这件事的方法总数为

$$N = m_1 + m_2 + \cdots + m_n.$$



乘法原理 (Multiplication principle, basic principle of counting)

乘法原理

做一件事有 n 个步骤, 第一步有 m_1 种方法, 第二步有 m_2 种方法, ..., 第 n 步有 m_n 种方法, 则完成这件事的方法总数为

$$N = m_1 \cdot m_2 \cdot \cdots \cdot m_n.$$

排列与组合

选排列 (Arrangement)

从 n 个不同的元素中，任取 $k(\leq n)$ 个元素，按照一定的顺序排成一列，全部排列个数为

$$A_n^k = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1).$$

全排列 (Permutation)

当 $k=n$ 时，称为全排列.

$$A_n^k = n!$$

组合 (Combination)

从 n 个不同的元素中，任取 $k(\leq n)$ 个元素并成一组，全部组合数为

$$\binom{n}{k} = C_n^k = \frac{A_n^k}{A_k^k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

组合数的推广

一般地, 把 n 个球随机地分成 r 组 ($n>r$), 要求第 i 组恰好有 n_i 个球, $i=1,\ldots,r$, 共有分法

$$\frac{n!}{n_1!n_2!\cdots n_r!}=\binom{n}{n_1\ n_2\ \cdots\ n_r}.$$

排列与组合的区别

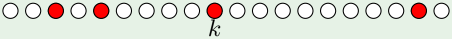
- 排列: 方法数与次序有关
- 组合: 方法数与次序无关.

例题

袋中有 a 只红球, b 只白球, 从袋中随机地将球一个一个取出. 求第 k 次取出的是红球的概率 $(1 \leq k \leq a+b)$.

思路一

假设除颜色外, 球是可区分的, 将取出的球依次排成一行



把每一种排法作为一个样本点. 有利场合为第 k 个位置为红球, 其它位置的 $(a+b-1)$ 个球可任意放置.

解

样本点总数为 $(a+b)!$, 有利场合数为 $C_a^1 \cdot (a+b-1)!$. 故所求概率为

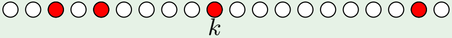
$$p_k = \frac{a(a+b-1)!}{(a+b)!} = \frac{a}{a+b}, \quad 1 \leq k \leq a+b.$$

例题

袋中有 a 只红球，b 只白球，从袋中随机地将球一个一个取出．求第 k 次取出的是红球的概率 $(1 \le k \le a+b)$.

思路二

假设除颜色外，球是不可的，将取出的球依次排成一行，



把每一种红球可能出现的位置作为一个样本点．有利场合为第 k 个位置为红球，其它位置 $(a+b-1)$ 个位置中出现 $(a-1)$ 个红球

解

样本点总数为 C_{a+b}^{a}，有利场合数为 C_{a+b-1}^{a-1}．故所求概率为

$$p_k = \frac{(a+b-1)!/(a-1)!b!}{(a+b)!/a!b!} = \frac{a}{a+b}, \quad 1 \le k \le a+b.$$

总结

结果

$$p_k = \frac{a}{a+b}, \quad 1 \le k \le a+b.$$

- 思路一用了**排列**的思路, 假设除颜色外球是可区分的
- 思路二用了**组合**的思路, 假设除颜色外球是不可区分的
- 确定了样本空间的结构后, 有利场合的构造必须与样本空间结构一致.
- 概率与 k 无关: 抽签结果与顺序无关.
- 很多实际问题都可以归结为 "摸 (或扔) 球模型".

例题

将 n 只 (可区分的) 球随机放入 N $(N \geq n)$ 个 (可区分的) 盒子中去, 试求每个盒子至多有一个球的概率.

分析

任意一只球放进任一盒子是等可能的. 我们把盒子编号为 1, 2, . . . , N, 并把每个球放入的盒子编号 $(a_1, a_2, \ldots, a_n)$ 作为样本点, 并用古典概型求解.

解

我们样本点总数为 N^n. 每个盒子至多有一个球, 意味着 $(a_1, a_2, ..., a_n)$ 互不相同, 因此有利场合数为选排列数 $A_N^n = N(N-1)\cdots(N-n+1)$. 故所求概率为

$$p = \frac{A_N^n}{N^n} = 1 \cdot \left(1 - \frac{1}{N}\right) \cdots \left(1 - \frac{n-1}{N}\right).$$

生日问题

问题

参加某次聚会共 n 个人，求没有两人生日相同的概率.

分析

此即为 n 个球 (人) 放入 365 个盒子中的问题.

解

$$p = 1 \cdot \left(1 - \frac{1}{365}\right) \cdot (1 - \frac{2}{365}) \ldots (1 - \frac{n-1}{365}).$$

n	20	25	30	40	50	55	100
$1-p$	0.41	0.57	0.71	0.89	0.97	0.99	0.999997

例题

从数字 0, 1, ..., 9 中随机地 (可重复) 抽 5 个数字, 则抽出的 5 个数字都不相同的概率是

$$p = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{10^5} = 0.3024$$

无理数的随机特征

考虑无理数 $e = 2.71828\cdots$.
将 e 的前 800 位小数分成 160 组, 每组 5 个数字视为从 $0, 1, 2, ..., 9$ 中随机抽出. 数得 5 个数字都不相同的共有 52 组, 其频率为

$$\frac{52}{160} = 0.325 \approx 0.3024.$$

这仍然只是一个猜想! (Normal numbers)

例题

某接待站在某周接待了 12 次来访，已知这 12 次来访都是在周二和周四进行的. 问是否可以推断接待站的接待时间是有规定的?

解

假设接待站的时间没有规定，且认为来访者每周任一天到达是等可能的. 则

$$\mathsf{P}\left(\{12\text{ 次来访都在周二和周四}\}\right) = \frac{2^{12}}{7^{12}} = 3 \times 10^{-7}.$$

定义

概率非常小的事件，称为**小概率事件**.
实际推断原理: 小概率事件在一次试验中是几乎不可能发生的.

由实际推断原理，可推断接待站的时间是有规定的.

思考题

问题

一根 1 米长的木条随机砍两刀. 问得到的三段中最短一段小于 $1/4$ 米的概率是多少?

古典概型的特点

有限样本点，等可能性

问题

如何将样本空间推广到“无限”个样本点，同时又有某种“等可能性”?

几何概型

定义

随机试验 向平面有界区域投掷一个点

样本空间 Ω

事件 点落在可测量面积的平面区域 A

事件概率 $\mathrm{P}(A) = \dfrac{A \text{ 的面积}}{\Omega \text{ 的面积}}$.

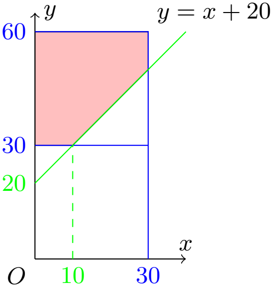
称上述试验为**几何概型**.

注记

- 事件 A 发生的概率与位置无关, 只与 A 的面积有关, 这体现了某种“等可能性”.
- 如果样本空间为有界区间、空间有界区域 (或更高维的几何体), 则 “面积” 改为 “长度”、“体积” (或高维推广).

例题

在一次演习中，某部队 A 接到命令要赶到某小河 D 岸为行进中的 B 部队架设浮桥. 假设 A 部队将于 7 点到 7 点 30 分之间到达 D 岸，架桥需要 20 分钟时间; B 部队将于 7 点 30 分至 8 点到达 D 岸. 试求 B 部队到达 D 岸时能立即过河的概率.



解

用 $x, y \in [0, 60]$ 表示 A 与 B 部队到达 D 岸的时间，则 $(x, y) \in \Omega = [0, 30] \times [30, 60]$. B 部队能立即过河的充要条件是 $x + 20 \le y$. 我们用几何概型求解，所求概率为

$$p = \frac{30^2 - 20^2/2}{30^2} = \frac{7}{9}.$$

例题

三个人去参加某个集合的概率均为 0.4，其中至少有两个参加的概率为 0.3，都参加的概率为 0.05. 求 3 人中至少有一个参加的概率.

解

以 A, B, C 分别记 3 个人参加的事件. 由条件知,

$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = 0.4, \quad \mathsf{P}(AB \cup CA \cup BC) = 0.3, \quad \mathsf{P}(ABC) = 0.05.$$

类似于容斥原理的证明，我们有

$$\mathsf{P}(A\cup B\cup C) = \mathsf{P}(A)+\mathsf{P}(B)+\mathsf{P}(C)-\mathsf{P}(AB\cup BC\cup CA)-\mathsf{P}(ABC) = 0.4\times 3-0.3-0.05 = 0.8$$

例题

设甲、乙两人同时向同一目标进行射击. 已知甲击中的概率为 0.7, 乙击中目标的概率为 0.6, 两人同时击中目标的概率为 0.4. 求

(1) 目标不被击中的概率;

(2) 甲击中目标而乙未击中的概率.

解

记事件 A, B 为甲、乙击中目标的事件. 于是由条件知

$$\mathsf{P}(A) = 0.7, \quad \mathsf{P}(B) = 0.6, \quad \mathsf{P}(AB) = 0.4.$$

于是,

$$\mathsf{P}(\text{ 目标不被击中}) = \mathsf{P}(\bar{A}\bar{B}) = \mathsf{P}(\overline{A \cup B}) = 1 - \left[\mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)\right] = 0.1.$$

$$\mathsf{P}(\text{ 甲击中而乙未击中}) = \mathsf{P}(A\bar{B}) = \mathsf{P}(A) - \mathsf{P}(AB) = 0.7 - 0.4 = 0.3.$$

伯努利信封问题

n 封信随机送到 n 个人手中. 求所有人都没收到自己的信的概率.

解

用 A_i, $i=1,2,\ldots,n$ 表示第 i 个人拿到自己的信的事件. 于是由容斥原理, 至少一个人拿到信件的概率为

$$\mathsf{P}(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}\mathsf{P}(A_i)-\sum_{1\le i_1<i_2\le n}\mathsf{P}(A_{i_1}A_{i_2})+\cdots+(-1)^{n-1}\mathsf{P}(A_1A_2\cdots A_n).$$

由乘法原理, 我们有 $\mathsf{P}(A_{i_1}A_{i_2}\cdots A_{i_k})=\dfrac{(n-k)!}{n!}$. 这样的项一共有 C_n^k 项.
所以前述概率为

$$\sum_{k=1}^{n}(-1)^{k-1}C_n^k\frac{(n-k)!}{n!}=\sum_{k=1}^{n}(-1)^{k-1}\frac{1}{k!}.$$

故没有人拿到自己的信的概率为

$$1-\sum_{k=1}^{n}(-1)^{k-1}\tfrac{1}{k!}=\sum_{k=0}^{n}(-1)^{k-1}\tfrac{1}{k!}\to e^{-1},\quad n\to\infty.$$

作业

- P21: 28, 29
- 补充题:
  1. 从 n 双尺码不同的鞋子中任到 $2r$ $(2r < n)$ 只, 求下列事件的概率:
     1. 所取 $2r$ 只鞋子中没有两只成对;
     2. 所取 $2r$ 只鞋子中只有两只成对;
     3. 所取 $2r$ 只鞋子恰好配成 r 对.
  2. (匹配问题) 将 4 把能打开 4 间不同房门的钥匙随机发给 4 个人. 试求至少有一人能打开门的概率.

三门问题 (Monty Hall Problem)

假如你在参加一个游戏，有三扇门 A、B、C：其中有一扇门后面放着一辆汽车，另外两扇门后面是山羊，你会赢得你选择的那扇门后面的礼物。游戏开始时，你任意选择一扇门，假如为门 A。主持人从剩余两扇门中选择一扇后面不是汽车的门打开，比如为门 B，现在主持人问：为了赢得车，是否要改选门 C（另外一扇没有被打开的门）？

选	B	C
A	羊	C
??	羊	??

改选后概率赢得汽车的概率为 $2/3$.

问题

一个家庭中有两个小孩, 已知至少一个是女孩, 问两个都是女孩的概率是多少? (假设生男生女是等可能的)

解

令事件 A 为一个为女孩的事件, B 为两个都是女孩的事件. 由于事件 A 已经发生, 所有可能的结果只有三种, 而事件 B 包含的基本事件只占其中一种, 所以有

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

$\mathsf{P}(B \mid A)$ 表示在 A 发生的条件下, B 发生的**条件概率**.

条件概率

定义

设 A, B 是两个事件, 且 $\mathsf{P}(B) > 0$. 记

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(AB)}{\mathsf{P}(B)},$$

并称其为事件 B 发生的条件下, 事件 A 发生的**条件概率**.

注

- 当 $\mathsf{P}(B) = 0$ 时, 条件概率 $\mathsf{P}(A \mid B)$ 无意义.
- 条件概率 $\mathsf{P}(A \mid B)$ 的直观理解:

  B 发生带来的“信息”对 A 的“推断”的新认识
- “$A \mid B$”不是一个事件.
- $\mathsf{P}(A \mid B) \geq \mathsf{P}(AB)$.

条件概率 $\mathsf{P}(\cdot \mid B)$ 的基本性质

设 $\mathsf{P}(B) > 0$, 有

- 非负性: 对于任一事件 A 有 $\mathsf{P}(A \mid B) \geq 0$;
- 规范性: 对于必然事件 Ω 有 $\mathsf{P}(\Omega \mid B) = 1$;
- 可列可加性: 设 $\{A_k\}$ 是两两不相容列, 则 $\mathsf{P}\left(\bigcup_{k=1}^{\infty} A_k \,\middle|\, B\right) = \sum_{k=1}^{\infty} \mathsf{P}(A_k \mid B)$

可列可加性的证明

由于 $\{A_k\}$ 两两不相容, 所以 $\{A_k \cap B\}$ 亦两两不相容. 由条件概率的定义及分配律, 有

$$\mathsf{P}\left(\bigcup_{k=1}^{\infty} A_k \,\middle|\, B\right) = \frac{\mathsf{P}\left\{\left(\bigcup_{k=1}^{\infty} A_k\right) \cap B\right\}}{\mathsf{P}(B)} = \frac{\mathsf{P}\left\{\bigcup_{k=1}^{\infty} (A_k \cap B)\right\}}{\mathsf{P}(B)}$$

$$= \frac{\sum_{k=1}^{\infty} \mathsf{P}(A_k B)}{\mathsf{P}(B)} = \sum_{k=1}^{\infty} \mathsf{P}(A_k \mid B).$$

条件概率构成的概率空间

条件概率也是一个概率空间.

性质

设 $\mathsf{P}(B) > 0$. 则条件概率 $\mathsf{P}(\cdot \mid B)$ 定义了如下概率空间:

- 样本空间: $\tilde{\Omega} = \Omega \cap B$.
- 事件域:
$$\tilde{\mathcal{A}} = \{A \cap B \mid A \in \mathcal{A}\}.$$
- 概率测度: 对 $A_1 = A \cap B$,
$$\tilde{\mathsf{P}}(A_1) = \mathsf{P}(A \mid B).$$

乘法公式

乘法公式

$$\mathsf{P}(AB) = \mathsf{P}(A \mid B)\mathsf{P}(B), \quad \mathsf{P}(B) > 0$$
$$\mathsf{P}(AB) = \mathsf{P}(B \mid A)\mathsf{P}(A), \quad \mathsf{P}(A) > 0.$$

证明

$$\mathsf{P}(AB) = \frac{\mathsf{P}(AB)}{\mathsf{P}(B)} \cdot \mathsf{P}(B) = \mathsf{P}(A \mid B) \cdot \mathsf{P}(B), \quad \mathsf{P}(B) > 0.$$

例题

第一个袋中有黑、白球各 2 只, 第二个袋中有黑、白球各 3 只. 先从第一个袋中任取一球放入第二个袋中, 再从第二个袋中任取一球. 求第一、第二次均取到白球的概率.

解

记 A_i, $i = 1, 2$ 为第 i 次取到白球的事件. 易知 $\mathsf{P}(A_1) = \dfrac{1}{2}$, $\mathsf{P}(A_2 \mid A_1) = \dfrac{4}{7}$. 于是由乘法公式得

$$\mathsf{P}(A_1 A_2) = \mathsf{P}(A_2 \mid A_1)\mathsf{P}(A_1) = \frac{1}{2} \cdot \frac{4}{7} = \frac{2}{7}.$$

注

条件概率是定义的, 但是条件概率的值通常要根据实际问题中具体定义确定.

推广

多个事件的乘法公式

若事件 $A_1, A_2, \ldots, A_n$ 满足 $\mathsf{P}(A_1A_2\cdots A_{n-1}) > 0$，则

$$\mathsf{P}(A_1A_2\cdots A_n) = \mathsf{P}(A_n \mid A_1A_2\cdots A_{n-1})\mathsf{P}(A_{n-1} \mid A_1A_2\cdots A_{n-2}) \cdots \mathsf{P}(A_2 \mid A_1)\mathsf{P}(A_1).$$

注

$0 < \mathsf{P}(A_1A_2\cdots A_{n-1}) \leq \mathsf{P}(A_1\cdots A_k),\ \forall k \leq n-1.$

滑坡谬误

例

员工偷懒公司就会损失，公司赚不到钱就要裁员，被裁员的人会没工作，没工作的人为了生计就会无恶不作. 因此，上班偷懒是非常严重的罪恶.

象箸之忧 (韩非子, 说林上)

以前商纣王弄了一双象牙筷子，箕子看到了感到恐惧。他认为说纣王有了象牙筷子，之后就会认为用陶土做的器皿配不上，因此就必然会再想办法得到犀牛角和玉石做的杯子；而在得到象牙筷子和玉石做的杯子后，就必然不会满足用这些东西盛装豆子和豆叶等粗食，而会想办法得到牦牛、大象和豹的胎盘之类的珍馐美馔；在吃到牦牛、大象和豹的胎盘之类的珍馐美馔后，就不会想穿着用粗布裁制的衣服、住在茅草盖的小屋当中，而一定会想改穿九重锦衣，并住在高高的楼台和宽广的房间之内；再这样下去，整个天下的东西都会满足不了他的欲求的。

$$\mathsf{P}(A_n|A_1A_2\cdots A_{n-1})\le 1 \Rightarrow \mathsf{P}(A_1A_2\cdots A_n)\to 0.$$

例题

已知 $\mathsf{P}(A) = 1/4$, $\mathsf{P}(B \mid A) = 1/3$, $\mathsf{P}(A \mid B) = 1/2$. 求 $\mathsf{P}(A \cup B)$, $\mathsf{P}(\bar{A} \mid A \cup B)$.

解

$$\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B \mid A) = \frac{1}{12}.$$

$$\mathsf{P}(B) = \frac{\mathsf{P}(AB)}{\mathsf{P}(A \mid B)} = \frac{1/12}{1/2} = \frac{1}{6}$$

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB) = \frac{1}{4} + \frac{1}{6} - \frac{1}{12} = \frac{1}{3}$$

$$\mathsf{P}(\bar{A} \mid A \cup B) = 1 - \mathsf{P}(A \mid A \cup B) = 1 - \frac{1/4}{1/3} = \frac{1}{4}.$$

全概率公式 (Law of Total probability)

问题

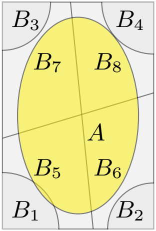
如何将一复杂概率计算问题分解成简单概率计算?

样本空间的分划

设 Ω 为样本空间, 若事件 B_1, B_2, ..., B_n 满足

- B_1, B_2, ..., B_n 两两不相容, 即 $B_iB_j = \varnothing$, $\forall i \neq j$.
- $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$,

则称 $\{B_1, B_2, \cdots, B_n\}$ 为样本空间 Ω 的一个**分划**.



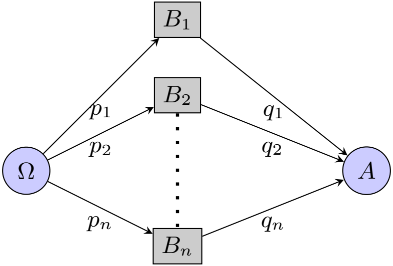
全概率公式

$$\mathsf{P}(A) = \mathsf{P}(AB_1) + \mathsf{P}(AB_2) + \cdots + \mathsf{P}(AB_n) = \sum_{i=1}^{n} \mathsf{P}(A \mid B_i)\mathsf{P}(B_i).$$

全概率公式

$$\mathsf{P}(A)=\mathsf{P}(AB_1)+\mathsf{P}(AB_2)+\cdots+\mathsf{P}(AB_n)=\sum_{i=1}^{n}\mathsf{P}(A\mid B_i)\mathsf{P}(B_i).$$

记 $\mathsf{P}(B_i)=p_i$, $\mathsf{P}(A\mid B_i)=q_i$, $i=1,2,\ldots,n$.



例题 (Polya's Urn)

袋中有 a 只红球和 b 只白球. 先从袋中任取一球, 记下颜色后放回, 同时向袋中放入同颜色的球 1 只, 然后再从袋中取出一球. 求第二次取到白球的概率.

解

记

$$A = \{第\ 2\ 次取到白球\};$$
$$B_1 = \{第\ 1\ 次取到白球\}, \quad B_2 = \{第\ 1\ 次取红球\}.$$

则 B_1, B_2 是 Ω 的一个分划.
由全概率公式有

$$\begin{aligned} \mathsf{P}(A) &= \mathsf{P}(A \mid B_1)\mathsf{P}(B_1) + \mathsf{P}(A \mid B_2)\mathsf{P}(B_2) \\ &= \frac{b+1}{a+b+1} \cdot \frac{b}{a+b} + \frac{b}{a+b+1} \cdot \frac{a}{a+b} \\ &= \frac{b}{a+b}. \end{aligned}$$

注

第二次取到白球的概率与第一次取到白球的概率相等, 与前面放入什么颜色的球无关.

例题

有 10 个袋子，其中甲袋 2 个，每袋中有红球、白球各 2 个；乙袋 3 个，每袋中有红球 3 个、白球 2 个；丙袋 5 个，每袋中有红球 2 个、白球 3 个. 从 10 个袋子中任取一袋，再从袋中任取一球，求取到白球的概率.

解

记 B_1, B_2, B_3 为取到甲、乙和丙袋的事件，A 表示取到白球. 则由全概率公式

$$\begin{aligned}\mathsf{P}(A) &= \sum_{i=1}^{3} \mathsf{P}(A \mid B_i)\mathsf{P}(B_i) \\ &= \frac{2}{10}\cdot\frac{2}{4} + \frac{3}{10}\cdot\frac{2}{5} + \frac{5}{10}\cdot\frac{3}{5} \\ &= \frac{13}{25}.\end{aligned}$$

全概率公式应用: 敏感问题调查

常常会发生被调查者拒绝回答或不真实回答的情况.

随机回答法 (Warner, 1965)

愈少泄漏问题的答案实质, 愈能较好合作.

例

设定两个问题

- Q_1: 你在期末考试中作弊了吗?
- Q_2: 你没在期末考试中作弊吗?

被调查都回答哪个问题随机决定, 且只有被调查者知晓.

贝叶斯 (Bayes) 公式

假设 $B_1, \ldots, B_n$ 为样本空间 Ω 的一个分划. 则

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

讨论

- $B_1, B_2, \ldots, B_n$ 为**原因**. $\mathsf{P}(B_i), i = 1, 2, \ldots, n$ 为**先验概率**.
- A 为结果, $\mathsf{P}(B_i \mid A)$ 为**后验概率**, 即 "由果及因".
- Bayes 方法广泛应用于网络、分类、诊断、估计、检验、判别、推理等方面.

例题

某工厂的一、二和三车间都生产同一产品，产量分别占总产量的 15%, 80% 和 5%. 三个车间的次品率分别为 2%, 1% 和 3%. 现从汇总起来的商品中任到一个，经检查是次品，判断该次品是哪个车间生产的可能性比较大?

解

记 $A=\{\text{取到次品}\}$, $B_i=\{\text{取到的产品是第 } i \text{ 个车间生产的}\}$, $i=1,2,3$.
由全概率公式,

$$\mathrm{P}(A)=\sum_{i=1}^{3}\mathrm{P}(A\mid B_i)\mathrm{P}(B_i)=0.0125.$$

再由 Bayes 公式，有

$$\mathrm{P}(B_1\mid A)=\frac{\mathrm{P}(A\mid B_1)\mathrm{P}(B_1)}{\mathrm{P}(A)}=\frac{0.02\times 0.15}{0.0125}=0.24$$

$$\mathrm{P}(B_2\mid A)=\frac{0.01\times 0.80}{0.0125}=0.64,\quad \mathrm{P}(B_3\mid A)=\frac{0.03\times 0.05}{0.0125}=0.12.$$

因此次品是第二车间生产的可能性最大.

测谎试验

我们用 $+/-$ 表示测谎仪显示受测者撒谎或未撒谎，用 T/L 表示受测者说的是真话或假话. 已知 $\mathsf{P}(+ \mid L) = 0.88$, $\mathsf{P}(- \mid T) = 0.86$, 又设人群中, $\mathsf{P}(T) = 0.99$. 现在若有一个在测谎中显示 “+” 性，问测谎仪出错的概率是多少?

解

由 Bayes 公式，此人说真话的概率为

$$\begin{aligned}\mathsf{P}(T \mid +) &= \frac{\mathsf{P}(+ \mid T)\mathsf{P}(T)}{\mathsf{P}(+ \mid T)\mathsf{P}(T) + \mathsf{P}(+ \mid L)\mathsf{P}(L)} \\ &= \frac{0.14 \times 0.99}{0.14 \times 0.99 + 0.88 \times 0.01} \\ &= 0.94.\end{aligned}$$

注

在普通人群中使用测谎仪有一定程度的危险性.

例题

用某种诊断法诊断癌症，记

$$A=\{\text{判断被检者有癌症}\},\quad C=\{\text{被检者确患有癌症}\}.$$

已知 $\mathsf{P}(A\mid C)=0.95$, $\mathsf{P}(\bar{A}\mid\bar{C})=0.90$. 又设人群中 $\mathsf{P}(C)=0.0004$. 现在若有一个人被诊断患有癌症，问此人真正患有癌症的可能性有多大？

解

由 Bayes 公式，此人真正患有癌症的概率为

$$\begin{aligned}\mathsf{P}(C\mid A)&=\frac{\mathsf{P}(A\mid C)\mathsf{P}(C)}{\mathsf{P}(A\mid C)\mathsf{P}(C)+\mathsf{P}(A\mid\bar{C})\mathsf{P}(\bar{C})}\\&=\frac{0.95\times0.0004}{0.95\times0.0004+0.1\times0.9996}\\&=0.0038.\end{aligned}$$

注

虽然检验法相当可靠，但是真正被诊断患有癌症而真正患有癌症的概率并不大.

例题

一单位有甲、乙两人. 已知甲近期出差的概率为 80%, 若甲出差, 则乙出差的概率为 20%; 若甲不出差, 则乙出差的概率为 90%.

(1) 求近期乙出差的概率.

(2) 若已知乙近期出差在外, 求甲出差的概率.

解

设 A, B 分别为甲、乙出差的事件. 已知 $\mathsf{P}(A) = 0.80$, $\mathsf{P}(B \mid A) = 0.20$, $\mathsf{P}(B \mid \bar{A}) = 0.90$.

(1) 由全概率公式,

$$\mathsf{P}(B) = \mathsf{P}(A)\mathsf{P}(B \mid A) + \mathsf{P}(\bar{A})\mathsf{P}(B \mid \bar{A}) = 0.8 \times 0.2 + 0.2 \times 0.9 = 34\%.$$

(2) 由 Bayes 公式, $\mathsf{P}(A \mid B) = \dfrac{\mathsf{P}(B \mid A)\mathsf{P}(A)}{\mathsf{P}(B)} = \dfrac{0.8 \times 0.2}{0.34} = \dfrac{8}{17}$.

例

一个盒子中有 5 个红球, 4 个白球, 采用不放回抽样. 每次取一个, 取 3 次.

(1) 求前两次中至少有一次取到红球的概率.

(2) 已知前两次中至少有一次取到红球, 求前两次中恰有一次取到红球的概率.

(3) 求第一、二次取到红球第三次取到白球的概率.

解

令 $A_i = \{第\ i\ 次取到红球\}$, $i = 1, 2, 3$. B 为前两次至少有一次取到红球的事件, C 为前两次恰有一次取到红球的事件.

1. $\mathsf{P}(B) = 1 - \mathsf{P}(\bar{B}) = 1 - \mathsf{P}(\bar{A}_1)\mathsf{P}(\bar{A}_2 \mid \bar{A}_1) = 1 - \dfrac{4}{9} \times \dfrac{3}{8} = \dfrac{5}{6}$.
2. $\mathsf{P}(C \mid B) = 1 - \mathsf{P}(\bar{C} \mid B) = 1 - \dfrac{\mathsf{P}(B\bar{C})}{\mathsf{P}(B)} = 1 - \dfrac{\mathsf{P}(A_1A_2)}{\mathsf{P}(B)} = \dfrac{2}{3}$.
3. $\mathsf{P}(A_1A_2\bar{A}_3) = \mathsf{P}(A_1)\mathsf{P}(A_2 \mid A_1)\mathsf{P}(\bar{A}_3 \mid A_1A_2) = \dfrac{5}{9} \times \dfrac{4}{8} \times \dfrac{4}{7} = \dfrac{10}{63}$.

例

某人参加某技能考核. 已知第一次参加能通过的概率为 60%; 若第一次未通过, 经过努力, 第二次能通过的概率为 70%; 若前两次未通过, 则第三次能通过的概率为 80%. 求此人最多 3 次能通过考核的概率.

解

令 $A_i = \{\text{第 } i \text{ 次通过考核}\}$, $i = 1, 2, 3$, $A = \{\text{最多三次通过考核}\}$. 则 $\bar{A} = \bar{A}_1\bar{A}_2\bar{A}_3$. 于是

$$\begin{aligned} \mathsf{P}(A) &= 1 - \mathsf{P}(\bar{A}_1\bar{A}_2\bar{A}_3) \\ &= 1 - \mathsf{P}(\bar{A}_1)\mathsf{P}(\bar{A}_2 \mid \bar{A}_1)\mathsf{P}(\bar{A}_3 \mid \bar{A}_1\bar{A}_2) \\ &= 1 - 0.4 \times 0.3 \times 0.2 = 0.976. \end{aligned}$$

作业

- 教材 P22: 46, 53, 54, 63
- 补充题:
  1. 请用本节所讨论的工具给出 Monty Hall 问题的解答
  2. 据以往资料表明，某一三口之家，患某种传染病的概率有以下规律:
$$\mathsf{P}(\text{孩子得病}) = 0.6,$$
$$\mathsf{P}(\text{母亲得病} \mid \text{孩子得病}) = 0.5,$$
$$\mathsf{P}(\text{父亲得病} \mid \text{孩子得病}) = 0.4.$$
     求母亲及孩子得病但父亲未得病的概率.
  3. 对以往数据分析结果表明，当机器调整得良好时，产品的合格率为 0.98; 而当机器发生某种故障时，产品的合格率为 0.55. 每天早上机器开动时，机器调整良好的概率为 0.95. 试求: 已知某日早上的第一件产品是合格品时，机器调整得良好的概率.

定义

例子

抛甲、乙两枚硬币，观察正反面出现的情况，则样本空间是 $\Omega = \{HH, HT, TH, TT\}$. 记事件 $A = \{$甲出现正面$\}$, $B = \{$乙出现正面$\}$.

从直观上看 A, B 之间是没有任何关系的，它们是“独立”的

从数学上看 $\mathsf{P}(A \mid B) = \mathsf{P}(A)$, $\mathsf{P}(B \mid A) = \mathsf{P}(B)$, 或 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$.

定义

设 A 和 B 是两个事件. 若

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

则称事件 A, B **相互独立**, 简称 A, B **独立**.

与不相容性的关系

问

A, B 独立与 A, B 不相容有什么关系?

分析

A, B 独立说明 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$. A, B 不相容说明 $AB = \varnothing$. 故当 $\mathsf{P}(A) > 0$, $\mathsf{P}(B) > 0$ 时, 二者**不能同时成立**.

问

若 A, B 独立, 问 $\bar{A}$, $\bar{B}$ 是否独立?

分析

若 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$, 则

$$\mathsf{P}(A\bar{B}) = \mathsf{P}(A) - \mathsf{P}(AB) = \mathsf{P}(A)\Big(1 - \mathsf{P}(B)\Big) = \mathsf{P}(A)\mathsf{P}(\bar{B}).$$

因此 A, B 独立可推出 $A, \bar{B}$ 独立. 进而 $\bar{A}$, $\bar{B}$ 也独立. (当然 $\bar{A}$, B 也独立).

例题

从一副不含大小王的扑克牌中任取一张. 记 $A=\{抽到\ \mathsf{K}\}$, $B=\{抽到的牌是黑色的\}$. 问二者是否独立?

解

抽牌适用古典概型计算概率. 因此 $\mathsf{P}(A)=\frac{4}{52}=\frac{1}{13}$, $\mathsf{P}(B)=\frac{26}{52}=\frac{1}{2}$,
$\mathsf{P}(AB)=\frac{2}{52}=\frac{1}{26}$.
于是 $\mathsf{P}(AB)=\mathsf{P}(A)\mathsf{P}(B)$, 故 A, B 独立.

“独立” 与 “不相容” 的区别与联系

习题 1

设 A、B 为不相容事件, 且 $\mathsf{P}(A) > 0$, $\mathsf{P}(B) > 0$. 下面四个结论中正确的是:

1. $\mathsf{P}(B \mid A) > 0$
2. $\mathsf{P}(A \mid B) = \mathsf{P}(A)$
3. $\mathsf{P}(A \mid B) = 0$
4. $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$.

习题 2

设 A、B 为独立事件, 且 $\mathsf{P}(A) > 0$, $\mathsf{P}(B) > 0$. 下面四个结论中正确的是:

1. $\mathsf{P}(B \mid A) > 0$
2. $\mathsf{P}(A \mid B) = \mathsf{P}(A)$
3. $\mathsf{P}(A \mid B) = 0$
4. $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$.

三个事件的独立性

设 A, B, C 是三个事件, 若

$$\begin{aligned} \mathsf{P}(AB) &= \mathsf{P}(A)\mathsf{P}(B) \\ \mathsf{P}(BC) &= \mathsf{P}(B)\mathsf{P}(C) \\ \mathsf{P}(CA) &= \mathsf{P}(C)\mathsf{P}(A) \\ \mathsf{P}(ABC) &= \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C), \end{aligned}$$

则称 A, B, C **相互独立 (独立)**.

n 个事件的独立性

若 n 个事件 $A_1, \ldots, A_n$ 满足

$$\mathsf{P}(A_{i_1}A_{i_2}\cdots A_{i_k}) = \mathsf{P}(A_{i_1})\mathsf{P}(A_{i_2})\cdots\mathsf{P}(A_{i_k}), \quad \forall 1 \le i_1 < \cdots, i_k \le n, \;\; k \ge 2,$$

则称事件 $A_1, A_2, \cdots, A_n$ **相互独立 (独立)**.

两两独立与相互独立

思考

A, B, C 两两独立是否能推出 A, B, C 相互独立?

反例

掷两枚硬币, $A=\{第一次正面\}$, $B=\{第二次正面\}$, $C=\{两次一正一反\}$, 则 $\mathsf{P}(A)=\mathsf{P}(B)=\mathsf{P}(C)=\frac{1}{2}$, $\mathsf{P}(AB)=\mathsf{P}(BC)=\mathsf{P}(CA)=\frac{1}{4}$, 它们两两独立. 但

$$\mathsf{P}(ABC)=0\neq\mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C).$$

思考

问题

- 必然事件 Ω 与任何事件 A 是否独立? **是**
- 不可能事件 $\varnothing$ 与任何事件是否独立? **是**

问题

事件 {甲患感冒} 与 {乙患感冒} 能否认为是独立的?

注

条件概率与事件独立性通常是根据实际意义来确定的.

例题

设每个人血清中含有肝炎病毒的概率为 0.4%, 求混合 100 个人血清中含有肝炎病毒的概率.

解

记 A_i 为第 i 个人血清中含有肝炎病毒的概率, $i = 1, 2, \ldots, 100$. 我们可以认为它们是独立的, 因此

$$
\begin{aligned}
\mathsf{P}\left(\bigcup_{i=1}^{100} A_i\right) &= \mathsf{P}\left(\left(\bigcap_{i=1}^{100} A_i^c\right)^c\right) \\
&= 1 - \mathsf{P}\left(\bigcap_{i=1}^{100} A_i^c\right) \\
&= 1 - 0.996^{100} \approx 0.33.
\end{aligned}
$$

注

可用于设计试验次数 (分组方法).

例题

一支步枪击中目标的概率为 $p=0.001$, 试求 n 支步枪齐射能击中目标的概率.

解

记 A_i 为第 i 支枪击中目标的事件, $i=1,2,\ldots,n$. 我们可以认为 $A_1,A_2,...,A_n$ 相互独立的. 因此所求概率为

$$p_n=1-\mathrm{P}\Big(\bigcap_{i=1}^n A_i^c\Big)=1-(1-p)^n=1-0.999^n.$$

n	1000	2000	3000	4000	5000
p_n	0.632	0.865	0.950	0.982	0.993

注

即使 p 很小, 但只要试验不断进行下去, 小概率事件也会必然发生.

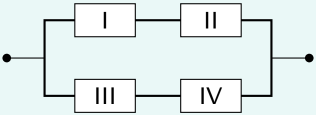
系统可靠性

定义

系统可靠性即“系统正常工作”的概率.

例题

某系统由四个部件 I, II, III, IV 构成, 结构如图. 设每个部件的可靠性均为 p, 且四个部件是独立的. 求整个系统的可靠性.



解

记 A 为整个系统正常工作的事件, A_i 为第 i 个部件正常工作的事件. 由于 I 与 II 串联, III 与 IV 串联, 两组部件再并联, 我们有 $A = (A_1A_2) \cup (A_3A_4)$. 于是系统可靠性为

$$\mathsf{P}(A) = \mathsf{P}(A_1A_2) + \mathsf{P}(A_3A_4) - \mathsf{P}(A_1A_2 \cap A_3A_4) = p^2 + p^2 - p^4 = p^2(2 - p^2).$$

例题

假设第 1、2、3 号高炮同时对飞机进行射击，三门炮击中飞机的概率分别为 0.4、0.5 和 0.7. 飞机被一门炮击中而被击落的概率为 0.2，被两门炮击中而被击落的概率为 0.6，若被三门炮击中，飞机必被击落. 求飞机被击落的概率.

解

令 A 为飞机被击落的事件，A_i，$i=0,1,2,3$，为飞机被 i 门炮击中的事件，B_j，$j=1,2,3$，为飞机被第 j 门炮击中的事件. 由独立性可计算得

$$\begin{aligned} A_1 &= B_1\bar{B}_2\bar{B}_3 \cup \bar{B}_1 B_2 \bar{B}_3 \cup \bar{B}_1\bar{B}_2 B_3, && \mathsf{P}(A_1) = 0.36, \\ A_2 &= B_1 B_2\bar{B}_3 \cup B_1\bar{B}_2 B_3 \cup \bar{B}_1 B_2 B_3, && \mathsf{P}(A_2) = 0.41, \\ A_3 &= B_1 B_2 B_3, && \mathsf{P}(A_3) = 0.14. \end{aligned}$$

由全概率公式有

$$\mathsf{P}(A) = \sum_{i=0}^{3} \mathsf{P}(A \mid A_i)\mathsf{P}(A_i) = 0 + 0.2\times 0.36 + 0.6\times 0.41 + 1\times 0.14 = 0.458.$$

例题

甲、乙两人同时向一目标射击，甲击中率为 0.8，乙击中率为 0.7，求目标被击中的概率.

解

设 $A=\{$甲击中$\}$, $B=\{$乙击中$\}$, $C=\{$目标被击中$\}$. 则 $C=A\cup B$. 由于甲、乙同时射击，其结果互不影响，我们可以认为 A 与 B 相互独立. 由容斥原理及独立性,

$$\mathsf{P}(C)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(AB)=0.7+0.8-0.7\cdot 0.8=0.94.$$

- P24: 68, 71, 74, 77, 79
- 补充题
  1. 设两个独立事件 A 和 B 都不发生的概率为 $1/9$, A 发生 B 不发生的概率与 B 发生 A 不发生的概率相同, 求事件 A 发生的概率.
  2. 设两两相互独立的三事件 A, B, C 事件满足条件: $ABC = \varnothing$, $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C)$, 且已知 $\mathsf{P}(A \cup B \cup C) = 9/16$. 求 $\mathsf{P}(A)$.

1. 样本空间 Ω, 随机事件 $A \subset \Omega$
2. 事件的关系: $A \subset B$, $A = B$
   事件的运算: $A \cup B$, $A \cap B$, $\bar{A}$.
3. 频率: $f_n(A) = \dfrac{n_A}{n}$
   概率的定义: $\mathsf{P}(A) \in [0,1]$, $\mathsf{P}(\Omega) = 1$, 可列可加性
   概率的性质:
   1. $\mathsf{P}(A) = 1 - \mathsf{P}(\bar{A})$
   2. $A \supset B \Rightarrow \mathsf{P}(A) \geq \mathsf{P}(B)$
   3. 容斥原理: $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)$
4. 条件概率:
   1. 定义: $\mathsf{P}(B \mid A) = \dfrac{\mathsf{P}(AB)}{\mathsf{P}(A)} \Leftrightarrow \mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B \mid A)$.
   2. 全概率公式, Bayes 公式
5. 事件独立性