工程概率统计 Probability and Statistics for Engineering

第一章概率论基础 Chapter 1 Basic Ideas in Probability

—

Chapter 1 Basic Ideas in Probability

1.1 基本概念

- 1.2 概率计算
- 1.3 条件概率与独立性

—

1.1 基本概念

- 你对概率的理解是什么？
- 在日常生活中，人们通常将概率解释为对个人所陈述内容的信念程度的度量。

例 1.1
日常生活中的概率：

- 小说《西游记》由吴承恩所写的概率是40%。
- 奥斯瓦尔德单独行动刺杀肯尼迪的概率是0.8。
- 如果你在这门课程中努力学习，你至少有90%的把握能获得A或更高的成绩。
- 在日常生活中这样使用概率概念是可以接受的，但在学术语境中，概率有着更严谨和精确的定义。
- 概率论研究随机事件，但小说《西游记》是否由吴承恩所写并非随机事件，这只是未知而已。

—

1.1 基本概念

- “随机”到底是什么意思？
- “随机”的概念指的是那些即使知道所有可能结果集，也无法确切预测的事件或结果。
- 因此，随机性意味着不确定性，但并非完全不确定。
- 概率论的本质是将个体的随机性转化为整体的确定性。
- 基于对“随机”的理解，我们引入基本概念，如随机试验、样本空间和随机事件。
- 有了这些基本概念，我们就可以用数学方法来描述和研究随机现象。

—

1.1 基本概念

随机试验

一个**随机试验**，通常简称为**试验**，代表一个随机现象的实现或观察，并具有以下特征：

- 它可以在相同条件下重复进行；
- 所有可能的结果都是明确已知的；
- 每次试验恰好发生一个可能结果，但无法预先确定会发生哪一个结果。

样本空间与随机事件

- 随机试验的每个可能的**基本结果**称为一个**样本点**，通常记为 ω。
- 包含试验所有样本点的集合称为**样本空间**，通常记为 Ω。
- 从集合论的角度看，一个**随机事件**，或简称**事件**，是样本空间 Ω 的一个子集，通常用大写字母表示（例如 A, B, C 等）。

- 如果结果是事件 A 中的一个样本点，则我们说事件 A 发生了。

—

1.1 基本概念

例 1.2

一个人开车上班需要经过三个交通灯。在每个交通灯处，他要么遇到红灯（记为0），要么遇到绿灯（记为1）。那么：

- 样本空间是 $\Omega = \{000, 001, 010, 100, 011, 101, 110, 111\}$;
- 事件“三个交通灯中恰好一个是红灯”是 $A = \{011, 101, 110\} \subset \Omega$。

例 1.3

一对夫妇决定要孩子，并计划一直生直到生出一个男孩为止。用B代表男孩，G代表女孩，那么：

- 样本空间是 $\Omega = \{B, GB, GGB, GGGB, ...\}$;
- 这个例子表明样本空间不一定是有限的。

—

1.1 基本概念

- 从集合论的角度看，事件的关系和运算变得非常清晰：
- **包含** $A \subset B$：事件 A 中的所有样本点也在 B 中，即当 A 发生时 B 也发生。
- **和/并** $A \cup B (A + B) := \{\omega | \omega \in A$ 或 $\omega \in B\}$，
A 和 B 中至少一个发生。
- **积/交** $A \cap B (AB) := \{\omega | \omega \in A$ 且 $\omega \in B\}$，
A 和 B 同时发生。

—

1.1 基本概念

- 从集合论的角度看，事件的关系和运算变得非常清晰：
- 差

  $$A - B\,(A \setminus B) := \{\omega|\omega \in A\,\text{且}\,\omega \notin B\},$$

  A 发生而 B 不发生。

- 互斥/互不相容

  $$A \cap B = \emptyset;\, A\,\text{和}\,B$$

  不能同时发生。

- 对立/互补

  $$A \cap B = \emptyset \quad \&\, A \cup B = \Omega;\,\text{要么}\,A\,\text{发生，要么}\,B\,\text{发生},$$

  记为 $B = A^c$ 或 A。

—

1.1 基本概念

- 事件的运算遵循某些规则，类似于集合的规则：
- 交换律:

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

- 结合律:

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

- 分配律:

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

- 德·摩根律:

  $$\overline{\bigcup_{i=1}^{\infty} A_i} = \bigcap_{i=1}^{\infty} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A_i}$$

—

1.1 基本概念

- 定义了样本空间和随机事件之后，我们现在可以讨论事件的概率了。
- 大家同意概率是对随机事件发生可能性的定量描述。但是概率的正式数学定义是什么？

概率测度，或简称概率，

是定义在样本空间 Ω 的子集上的实值函数，满足以下三条公理：

- **非负性**: 对任意事件 $A \subseteq \Omega$，有 $P(A) \geq 0$。
- **规范性**: $P(\Omega) = 1$。
- **可加性**: 对任意**互斥**事件 $A_1, A_2, ...$, 有

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

—

1.1 基本概念

- 根据这三条公理，可以推导出许多有用的概率计算规则。

概率的性质

- $P(\emptyset) = 0$。

  证明: 考虑事件序列 $A_1 = \Omega, A_2 = A_3 = \cdots = \emptyset$。那么，这些事件是互斥的，并且 $\Omega = A_1 \cup A_2 \cup \cdots$。因此，由第三条公理（即可加性）：

$$P(\Omega) = P(\Omega) + \sum_{i=1}^{\infty} P(\emptyset),$$

  这意味着 $P(\emptyset) = 0$。

- **有限可加性**: 对任意有限的互斥事件序列 $A_1, A_2, ..., A_n$，

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

—

1.1 基本概念

概率的性质

- **补事件规则**: $P(\overline{A}) = 1 - P(A)$。

  证明: 令 $A_1 = A$ 和 $A_2 = \overline{A}$，那么根据 $n = 2$ 时的有限可加性，

  $$1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A}) \Rightarrow P(\overline{A}) = 1 - P(A).$$

- **数值界限**: $0 \leq P(A) \leq 1$。

- **单调性**: 如果 $A \subseteq B$，那么 $P(A) \leq P(B)$ 且 $P(B - A) = P(B) - P(A)$。

- **加法定律**: $P(A \cup B) = P(A) + P(B) - P(AB)$。

- **容斥原理**: ($A_1, A_2, ..., A_n$ 不一定互斥)

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

—

1.1 基本概念

例 1.4

设 A, B, C 是三个随机事件，且 $P(A) = P(B) = P(C) = 0.25$，$P(AB) = P(BC) = 0$，$P(AC) = 0.125$。求 A, B, C 中至少有一个发生的概率。

解

根据概率的单调性，由于 $ABC \subseteq AB$，那么 $P(ABC) \leq P(AB) = 0$。
根据非负性公理，$P(ABC) = 0$。

A, B, C 中至少有一个发生的概率表示为 $P(A \cup B \cup C)$。最后，根据容斥原理：

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$$

$$= 0.25 + 0.25 + 0.25 - 0 - 0 - 0.125 + 0 = 0.625.$$

—

Chapter 1 Basic Ideas in Probability

- **1.1 基本概念**
- **1.2 概率计算**
- **1.3 条件概率与独立性**

—

1.2 概率计算

- 概率是对随机事件发生可能性的精确描述。是否有必要精确计算概率？
- 在日常生活中，可能没有必要，因为我们很难区分概率为0.3和0.4的事件。
- 然而，在专业领域，精确的概率测量变得非常重要。例如：
- 在**赌场**中，通过精确的概率测量和设计，庄家只需要比玩家有稍高一点的获胜概率，就能从整体上盈利。
- **保险公司**也是基于概率的业务。它们通过精确计算理赔概率来设计和定价保险产品，确保整体盈利。

—

1.2 概率计算

- 早期的概率计算大多基于一个相对简单的模型，称为**古典概率模型**。

古典概率模型

如果一个随机试验满足：

- 样本空间 Ω 中只有有限个样本点，即

$$\Omega = \{\omega_1, \omega_2, ..., \omega_n\},$$

- 每个样本点发生的可能性相同，即

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

则该试验的概率模型称为**古典概率模型**。

古典概率模型下的概率计算

假设事件 A 中有 k 个样本点，则事件 A 的概率为

$$P(A) = \frac{\text{事件 A 中的样本点数}}{\text{样本空间 }\Omega\text{ 中的样本点总数}} = \frac{k}{n}.$$

因此，在古典概率模型下，概率计算只涉及对样本空间和感兴趣事件中的样本点进行计数。

- 随机事件的概率就是该事件（样本空间的一个子集）在样本空间中所占的比例。

—

1.2 概率计算

例 1.5 如果有两个盒子，每个盒子都装有特定数量的红球和绿球，允许你先选择一个盒子，然后从选中的盒子中随机抽取一个球。如果抽到红球，你就赢得奖品。那么你应该选择哪个盒子？

（图示：几个盒子，标有红球和绿球的数量，例如：5红6绿，3红4绿，6红3绿，9红5绿，11红9绿，12红9绿等）

辛普森悖论
（加州大学伯克利分校在1970年代的性别歧视）
—

1.2 概率计算

- 然而，更多时候，样本点的数量不容易确定，需要更系统的计数方法。
- 我们介绍/回顾两种常用的计数方法：加法原理和乘法原理，以及排列与组合。

加法原理和乘法原理

- **加法原理**: 如果有 n 类方法可以完成一项任务，第一类有 m_1 种具体方法，第二类有 m_2 种具体方法，...，第 n 类有 m_n 种具体方法，那么完成这项任务的具体方法总数为：

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

- **乘法原理**: 如果完成一项任务需要 n 个步骤，第一步有 m_1 种可能方法，第二步有 m_2 种可能方法，$\cdots$，第 n 步有 m_n 种方法，那么完成这项任务的方法总数为：

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

—

1.2 概率计算

排列与组合

- **排列**: 从 n 个不同元素中无重复地随机抽取 k 个元素（$k \le n$）并按顺序排列，则不同排列的数目为

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

- **组合**: 从 n 个不同元素中无重复地随机抽取 k 个元素（$k \le n$），不考虑顺序，其组合数由下式给出

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

- **思考**: 在有重复随机抽取的情况下，排列和组合是什么？

—

1.2 概率计算

例 1.6 假设一个班级有 n 名学生，学校分配了 $m(< n)$ 张音乐会门票给该班。老师决定通过抽签来分发门票。

老师准备了一个帽子，里面有 n 张纸条，其中 m 张纸条上标有“1”，其余的纸条上标有“0”。

学生们轮流从帽中抽取纸条，抽到标有“1”的纸条的学生将得到一张音乐会门票。

如果你是这个班级的一名学生，并且非常想得到一张票，你会选择早点抽还是晚点抽？

—

1.2 概率计算

解

- 核心问题是：假设你是第 l 个抽纸条的人，你得到票的概率是否依赖于 l?
- 要计算概率，我们首先需要弄清楚样本空间是什么？
- 将班级的抽票过程视为一个随机试验，那么抽票过程的每个可能结果就是一个样本点（即一个 n 位数，其中有 m 位是1，其余是0），样本空间就是所有这些样本点的集合。
- 然后，确定样本空间中样本点的数量：C_n^m。
- 令 A 表示你得到票的事件，则事件 A 中的样本点数量为：C_{n-1}^{m-1}。

- 综合起来，事件 A 的概率为

$$P(A) = \frac{\text{事件 A 中的样本点数}}{\text{样本空间 }\Omega\text{ 中的样本点总数}} = \frac{C_{n-1}^{m-1}}{C_n^m} = \frac{m}{n}.$$

- 因此，得到票的概率不依赖于你抽纸条的顺序!

—

1.2 概率计算

例 1.7

（生日问题）一个班级有 $n(n < 365)$ 名学生，至少有两名学生的生日相同的概率是多少?（不考虑闰年）

解

- 每个由 n 个生日组成的序列是一个样本点，样本空间是所有可能的 n 个生日序列的集合。
- 那么，样本空间中样本点的数量是：365^n。
- 令 A 为至少有两名学生生日相同的事件，则 $\overline{A}$ 是每位学生生日都不同的事件。
- 计算 $\overline{A}$ 中的样本点数量更容易：

$$A_{365}^n = 365 \times 364 \times \cdots \times (365 - n + 1)$$

- 因此，事件 A 的概率计算如下：

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{A_{365}^n}{365^n}$$

—

1.2 概率计算

解

- 结果

n	20	25	30	40	50	55	100
$P(A)$	0.41	0.57	0.71	0.89	0.97	0.99	0.999997

- 在一个有50名学生的班级中，至少两人同一天生日的概率高达97%。
- 在一个有100名学生的班级中，几乎可以肯定至少有两人生日相同。
- 这让你感到惊讶吗？

—

1.2 概率计算

例 1.8

（配对问题）

- 一个班级有 n 名学生，学号分别为 $1, 2, ..., n$。
- 在去年春节前，每个人都准备了一份礼物，礼物上标有自己的学号。
- 然后所有的礼物都被放入一个袋子中，每个人从袋子中随机抽取一份礼物。
- 至少有一名学生拿到自己准备的礼物的概率是多少？

解

- $\{1, 2, ..., n\}$ 的每一个排列都是一个样本点，代表学号为 $1, 2, ..., n$ 的学生拿到的礼物编号。
- 所以，样本空间中样本点的数量是：$A_n^n = n!$。
- 令 A 为至少有一名学生拿到自己礼物的事件，直接确定 A 中的样本点数量不那么简单。

—

1.2 概率计算

解

- 思路是将模糊复杂的问题分解为清晰简单的问题。
- 令 A_i 为学号为 i 的学生拿到自己礼物的事件，则

$$\begin{aligned}P(A) &= P(A_1 \cup A_2 \cup \cdots \cup A_n)\\ &= \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \cdots + (-1)^{n-1}P(A_1A_2\ldots A_n)\end{aligned}$$

- 事件 $A_i, A_iA_j, A_1A_2\cdots A_n$ 非常清晰，它们的概率计算很简单：

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{C_n^1}, \quad P(A_iA_j) = \frac{(n-2)!}{n!} = \frac{1}{2!C_n^2}, \cdots$$

- 综合起来，我们有：

$$\begin{aligned}P(A) &= \frac{1}{C_n^1} \times C_n^1 - \frac{1}{2!C_n^2} \times C_n^2 + \frac{1}{3!C_n^3} \times C_n^3 + \cdots + (-1)^{n-1}\frac{1}{n!}\\ &= 1 - \frac{1}{2!} + \frac{1}{3!} + \cdots + (-1)^{n-1}\frac{1}{n!} \approx 1 - e^{-1} \approx 0.632.\end{aligned}$$

1.2 概率计算

- 古典概率模型假设样本点数量有限且等可能。
- 另一种模型，称为**几何概率模型**，是古典模型向无限样本点情况的扩展，同时保持等可能性。

几何概率模型及其概率计算

如果一个随机试验可以表示为将点随机投到一个有界区域 Ω 上，且点落在区域内任何位置的可能性相同，则该试验的概率模型称为几何概率模型。令 A 表示点落在 Ω 的子区域 A 的事件，则事件 A 的概率计算如下：

$$P(A) = \frac{\text{区域 } A \text{ 的长度/面积/体积}}{\text{区域 } \Omega \text{ 的总长度/面积/体积}}.$$

- 无论使用哪种模型，概率本质上都是随机事件（样本空间的子集）在样本空间中所占的比例。

—

1.2 概率计算

例 1.9 罗密欧与朱丽叶计划在晚上会面，两人都会在晚上7点到8点之间到达花园，所有到达时间对都是等可能的。先到者会等待15分钟，如果另一方仍未到达就会离开。他们会相遇的概率是多少？

解

- 以晚上7点为原点，x轴和y轴分别表示罗密欧和朱丽叶的到达时间。
- 那么样本空间是 $\Omega = \{(x, y) \in [0, 1] \times [0, 1]\}$。
- 要使罗密欧与朱丽叶相遇，需要满足 $|x - y| \leq 1/4, 0 \leq x, y \leq 1$。
- 事件 A 表示相遇的区域。总样本空间面积为 $1 \times 1 = 1$。
- 不相遇的区域是两个三角形，每个面积是 $\frac{1}{2} \times \frac{3}{4} \times \frac{3}{4} = \frac{9}{32}$，总不相遇区域面积为 $2 \times \frac{9}{32} = \frac{18}{32} = \frac{9}{16}$。
- 因此，相遇区域 A 的面积为 $1 - \frac{9}{16} = \frac{7}{16}$。
- 所以，$P(A) = \frac{7}{16}$。

—

1.2 概率计算

例 1.10 贝特朗悖论

考虑一个半径为 1 的圆。随机画一条弦，其长度大于圆内接等边三角形边长的概率是多少？

解法 1

- 取圆的一条半径 OA，在半径上随机选择一点 C（所有点等可能），然后过 C 点作垂直于 OA 的弦。
- 根据初等几何，OA 与三角形的中位线相交，记中点为 B。
- 样本空间是 OA 上的所有点。
- 要使弦长大于三角形边长，C 必须落在 OB 上。
- 因此，概率为
$$\frac{\text{线段 } OB \text{ 的长度}}{\text{线段 } OA \text{ 的长度}} = \frac{1}{2}.$$

—

1.2 概率计算

解法 2

- 取圆上一点，例如顶点 V，过 V 作圆的切线。
- 然后在圆上随机选择另一点 D，连接 V 和 D 得到弦，该弦与切线形成一个随机角 Θ。
- 样本空间是 $\Omega = \{\Theta \in [0, \pi]\}$。要使弦长大于内接等边三角形的边长，需要 $\pi/3 < \Theta < 2\pi/3$。
- 因此，概率为
$$\frac{(2\pi/3 - \pi/3)}{\pi} = \frac{\pi/3}{\pi} = \frac{1}{3}.$$

—

1.2 概率计算

解法 3

- 在圆内随机选择一个点 E，以 E 为中点作弦。
- 要使弦长大于内接等边三角形的边长，E 必须落在三角形的内切圆内，该内切圆半径为 $1/2$。

- 样本空间是原始圆盘（面积 $\pi \times 1^2 = \pi$），感兴趣的事件区域是内切圆盘（面积 $\pi \times (1/2)^2 = \pi/4$）。
- 因此，概率为
$$\frac{\pi/4}{\pi} = \frac{1}{4}.$$

—

1.2 概率计算

讨论

- 哪种解法是正确的？
- 三种解法都正确！
- 同一个问题有三种不同答案的原因是，“随机画弦”在问题中没有明确定义。
- 不同的等可能性假设导致了**不同的样本空间**：
  - 解法1假设弦的中点等可能地落在一条半径上，样本空间是半径上的所有点。
  - 解法2假设弦的另一个端点等可能地落在圆周上，样本空间是圆周上的所有点。
  - 解法3假设弦的中点等可能地落在圆盘内，样本空间是圆盘内的所有点。
- **因此，在计算概率时，首先明确定义样本空间至关重要。**

—

Chapter 1 Basic Ideas in Probability

- **1.1 基本概念**
- **1.2 概率计算**
- **1.3 条件概率与独立性**

—

1.3 条件概率与独立性

- 首先看一个曾在美国引起广泛讨论的问题——**蒙提霍尔问题（三门问题）**。

例 1.11

- 该问题源自美国电视节目《Let's Make a Deal》，以主持人蒙提·霍尔的名字命名。
- 该节目在1975年引起公众关注。参赛者会看到三扇关闭的门，其中一扇后面有汽车，另外两扇后面各有一只山羊。如果他们成功选中藏有汽车的门，就能赢得汽车。
- 你首先被要求选择一扇门，比如1号门。
- 然后，主持人（他知道汽车在哪里）打开剩下两扇门中的一扇，露出一只山羊，比如在2号门后面。
- 最后，主持人给你一个改变选择的机会——要么坚持原来的选择（1号门），要么切换到3号门。
- 你会如何选择？

—

1.3 条件概率与独立性

- 学习了条件概率之后，你会发现一个逻辑严谨的方法来解决蒙提霍尔问题。
- 条件概率指的是，在另一个事件发生的条件下，一个随机事件的概率会发生变化。

条件概率

设 A 和 B 是两个随机事件，且 $P(B) > 0$。则在事件 B 发生的条件下，事件 A 的**条件概率**定义为

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

- $P(A|B)$ 的直观理解是，事件 B 的发生提供了新的信息，更新了我们对事件 A 发生可能性的信念。
- 这个定义背后的思想是，如果事件 B 已经发生，那么样本空间就变成了 B 而不是 Ω。

—

1.3 条件概率与独立性

例 1.12

- 你正在玩扑克游戏，发到5张牌，牌面朝下。
- 皇家同花顺是同一花色的A、K、Q、J、10。
- 1. 你拿到皇家同花顺的概率是多少？
- 2. 如果你发到的一张牌牌面朝上，显示是黑桃A，那么现在的概率是多少？

解

- 令 A 表示拿到皇家同花顺的事件，则 A 中的样本点数为4（每种花色一个）。
- 样本空间是所有可能的5张牌组合，样本点总数为 C_{52}^5。
- 所以，事件 A 的概率为：

$$P(A) = \frac{4}{C_{52}^5} = \frac{1}{649,740}.$$

—

1.3 条件概率与独立性

解（续）

- 令 B 表示5张牌中有一张是黑桃A的事件，则 B 中的样本点数为 C_{51}^4。

- $AB = A \cap B$ 是拿到黑桃皇家同花顺的事件，所以 AB 中的样本点数为1。因此，在 B 发生的条件下 A 的概率为

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{1/C_{52}^5}{C_{51}^4/C_{52}^5} = \frac{1}{C_{51}^4} = \frac{1}{249,900}.$$

—

1.3 条件概率与独立性

例 1.13

- 你进行一项罕见疾病的血液检测，该疾病在人群中的发病率为十万分之一。
- 如果你患有该疾病，检测结果有0.95的概率呈阳性（有0.05的概率呈阴性）。
- 如果你没有患病，检测有0.001的概率出现假阳性。
- 如果检测结果显示你患病，你实际患病的概率是多少？

思路

- 首先，我们用概率语言表达例子中给出的条件。
- 令 A 表示随机选择的人检测结果呈阳性的事件，B 表示随机选择的人患有该疾病的事件，则：

$$P(A|B) = 0.95, \quad P(A|\overline{B}) = 0.001, \quad P(B) = 0.00001.$$

- 需要计算的概率表示为 $P(B|A) = P(AB)/P(A)$。
- 如何根据给定条件计算 $P(AB)$ 和 $P(A)$？

—

1.3 条件概率与独立性

为了解决这个问题，我们引入乘法定律（用于计算 $P(AB)$）和全概率公式（用于计算 $P(A)$）：

乘法定律

设 A 和 B 是两个随机事件，且 $P(B) > 0$，则

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

全概率公式

设 A 和 B 是两个随机事件，则（假设若 $P(B) = 0$ 则 $P(A|B) = 0$）

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

更一般地，如果 $B_1, B_2, ..., B_n$ 是 n 个互斥的随机事件，并且 $B_1 \cup B_2 \cup ... \cup B_n = \Omega$，则称 $B_1, B_2, ..., B_n$ 为样本空间 Ω 的一个**划分**，并且

$$\begin{aligned}
P(A) = P(A \cap \Omega) &= P(AB_1 \cup AB_2 \cup ... \cup AB_n) \\
&= P(AB_1) + P(AB_2) + ... + P(AB_n) \\
&= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_n)P(B_n).
\end{aligned}$$

—

1.3 条件概率与独立性

- 就像加法定律有推广到多个事件的一般形式（即容斥原理）一样，乘法定律也有一般形式：（假设 $P(A_1A_2 \cdots A_{n-1}) > 0$）

  $$\begin{aligned}
  P(A_1A_2 \cdots A_n) &= P(A_n|A_1A_2 \cdots A_{n-1})P(A_1A_2 \cdots A_{n-1}) = \cdots \\
  &= P(A_n|A_1A_2 \cdots A_{n-1})P(A_{n-1}|A_1A_2 \cdots A_{n-2}) \cdots P(A_2|A_1)P(A_1),
  \end{aligned}$$

  这称为随机事件的**链式法则**。

例 1.13 的解答

- 根据乘法定律和全概率公式：

  $$\begin{aligned}
  P(B|A) = \frac{P(AB)}{P(A)} &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})} \\
  &= \frac{0.95 \times 0.00001}{0.95 \times 0.00001 + 0.001 \times 0.99999} \approx 0.0094.
  \end{aligned}$$

- 即使检测准确率很高，阳性结果正确的概率也非常小。
- 这与你经验或直觉相悖吗？为什么会这样？
- 原因是该疾病在人群中的发病率非常低。

—

1.3 条件概率与独立性

实际上，在血液检测例子的解答中使用了一个著名的定理，即**贝叶斯定理**。

贝叶斯定理 / 贝叶斯法则

设 $B_1, B_2, ..., B_n$ 是随机事件，且 $B_1, B_2, ..., B_n$ 是样本空间的一个划分。则对于任何满足 $P(A) > 0$ 的事件 A，我们有

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

- $P(B_i|A)$ 是在 A 发生的条件下 B_i 的概率，也称为 B_i 的**后验概率**。
- $P(B_i)$ 称为 B_i 的**先验概率**或**边缘概率**，指的是在没有任何其他先验信息情况下的概率值。

托马斯·贝叶斯 1702-1761
英国数学家

贝叶斯法则在各个领域被广泛应用，尤其是在涉及不确定性和不完全信息下决策的场景。

—

1.3 条件概率与独立性

贝叶斯法则是否被广泛应用？

—

是的，贝叶斯法则在各个领域被广泛应用，尤其是在涉及不确定性和不完全信息下决策的场景。以下是一些关键应用：

1. **医学诊断**: 在医疗保健领域，贝叶斯定理帮助计算患者在检测结果呈阳性的情况下确实患有特定疾病的概率，同时考虑检测的可靠性和疾病的流行程度。

2. **机器学习与人工智能**: 许多算法，如朴素贝叶斯分类器，依赖贝叶斯法则进行预测。它被用于垃圾邮件检测、情感分析和文档分类。

3. **金融**: 贝叶斯定理用于风险评估、投资组合管理和市场预测，随着新信息的出现更新投资策略。

4. **法证与法律推理**: 在法庭案件中，贝叶斯定理可以辅助评估证据，根据现有数据确定犯罪的可能性。

5. **认知科学**: 贝叶斯法则模拟人类的推理和决策过程，特别是人们如何根据新证据更新他们的信念。

总的来说，贝叶斯法则是一个强大的工具，适用于所有需要概率推理和推断的领域，从科学、工程到日常决策。

—

1.3 条件概率与独立性

- 贝叶斯法则应用的一个真实例子——**临床决策支持系统**
- $B_1, B_2, ..., B_n$ 代表各种疾病，A 代表某种症状或指标（例如血液中的转氨酶水平）。
- 先验概率 $P(B_i)$ 可以通过统计方法确定。
- $P(A|B_i)$ 可以通过医学知识确定。
- 应用贝叶斯法则，我们可以计算后验概率 $P(B_i|A)$。
- 将对应较大 $P(B_i|A)$ 的疾病提供给医生，以供进一步的临床诊断。

—

1.3 条件概率与独立性

- 如何理解贝叶斯法则？

- 实际上，我们可以将事件 A 视为结果，事件 $B_1, B_2, ..., B_n$ 视为导致该结果的各种原因。
- 全概率公式根据不同的原因推断结果 A 发生的概率——这是一个**从原因到结果的推理过程**。
- 然而，在我们的日常生活中还有另一个重要场景：我们观察到某种现象，然后反向推理导致它的各种原因的概率。简而言之，就是**从结果到原因的推理**。
- 通过贝叶斯法则得到的条件概率 $P(B_i|A)$ 帮助我们推断观察到的结果 A 是由特定原因 B_i 导致的可能性，为我们后续的决策提供支持。
- 后验概率 $P(B_i|A)$ 是在获得新信息后对先验概率 $P(B_i)$ 的修正。

—

1.3 条件概率与独立性

- 最后，我们来解决蒙提霍尔问题。

例 1.11 的解答

- 不失一般性，假设你选择了1号门，主持人打开了2号门。
- 令 B_1, B_2, B_3 分别表示汽车在1号门、2号门、3号门后面的事件，A 表示主持人打开2号门的事件。则：

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}, \quad P(A|B_1) = \frac{1}{2}, \quad P(A|B_2) = 0, \quad P(A|B_3) = 1.$$

- 感兴趣的概率是 $P(B_1|A)$ 和 $P(B_3|A)$。根据全概率公式，我们有

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = \frac{1}{2}\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3} = \frac{1}{2}.$$

- 最后，根据贝叶斯法则：

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{(1/2)\times(1/3)}{1/2} = \frac{1}{3}, \quad P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{1\times(1/3)}{1/2} =$$

- 因此，你应该切换到3号门!

—

1.3 条件概率与独立性

- 最后，我们来讨论事件之间的**独立性**。你可能之前学过一些关于独立性的定义，例如：
- 两个事件之间的独立性意味着一个事件的发生不影响另一个事件的概率。
- 如果 $P(AB) = P(A)P(B)$，则称事件 A 和 B 是独立的。
- 然而，第一个不是数学上严格的定义，第二个似乎难以直观理解。
- 借助条件概率，可以轻松清晰地定义事件的独立性。
- "A 的发生不影响 B 的概率，反之亦然"可以表示为

$$P(A|B) = P(A), \quad P(B|A) = P(B).$$

- 根据乘法定律，我们因此有

$$P(AB) = P(A|B)P(B) = P(A)P(B), \quad \text{同样有} \quad P(AB) = P(B|A)P(A) = P(A)P(B).$$

—

1.3 条件概率与独立性

例 1.14 设 A 和 B 是两个随机事件，以下两个陈述之间有什么关系？

- A 和 B 是独立的。
- A 和 B 是互斥的。

解

- A 和 B 独立 $\Rightarrow P(AB) = P(A)P(B)$。
- A 和 B 互斥 $\Rightarrow P(AB) = 0$。
- 因此，如果 $P(A) > 0$ 且 $P(B) > 0$，那么"A 和 B 独立"与"A 和 B 互斥"不能同时成立。

- “A 和 B 互斥”意味着如果 B 发生，则 A 永远不会发生，即 B 的发生为 A 的发生提供了新的信息，所以这两个事件不是独立的。

—

1.3 条件概率与独立性

- 我们来看独立性的正式定义。

独立性

设 $A, B, C, A_1, A_2, ..., A_n$ 均表示随机事件。

- 如果 $P(AB) = P(A)P(B)$，则称 A 和 B 是独立的。
- 如果满足以下条件，则称 A, B, C 是**（相互）独立的**：

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

- 如果对于事件的每一个子集 $A_{i_1}, A_{i_2}, ..., A_{i_k} \quad (1 \leq i_1 < \cdots < i_k \leq n, k = 2, ..., n)$，都有

$$P(A_{i_1}A_{i_2}...A_{i_k}) = P(A_{i_1})P(A_{i_2})...P(A_{i_k}),$$

则称 $A_1, A_2, ..., A_n$ 是**（相互）独立的**。

- 上述定义中的条件都可以通过条件概率推导出来。
- 你可能会自然地想知道，为什么多个事件的独立性需要这么多条件。这些条件是否冗余？
- 例如，两两独立是否意味着三个事件相互独立？

—

1.3 条件概率与独立性

例 1.15

- 从一副洗匀的标准扑克牌中抽取三张牌，有放回且每次抽后重新洗牌（即抽一张牌，记录，放回牌堆，洗牌，抽下一张）。
- 令 A 表示“第1张和第2张牌花色相同”的事件，B 表示“第2张和第3张牌花色相同”的事件，C 表示“第1张和第3张牌花色相同”的事件。
- 这三个事件是否相互独立？

解

- 不难得到 $P(A) = P(B) = P(C) = 1/4$。
- 此外，事件 AB, AC, BC, ABC 都指的是“第1、2、3张牌花色都相同”的事件，所以 $P(AB) = P(AC) = P(BC) = P(ABC) = 1/16$。
- 因此，我们有 $P(AB) = P(A)P(B), P(AC) = P(A)P(C), P(BC) = P(B)P(C)$，表明这三个事件是两两独立的。
- 然而，由于 $P(ABC) = 1/16 \neq (1/4)^3 = 1/64$，即 $P(ABC) \neq P(A)P(B)P(C)$，所以这三个事件不是相互独立的。

—

1.3 条件概率与独立性

例 1.16

- 一个系统有4个组件，其连接结构如图所示（通常表示为两个并联的串联组件对）。
- 假设每个组件正常工作的概率为 p，且4个组件相互独立。
- 系统正常工作的概率是多少？

解

- 令 A 表示系统正常工作的事件，A_i 表示组件 i 正常工作的事件（$i = 1,2,3,4$）。则 $P(A_1) = P(A_2) = P(A_3) = P(A_4) = p$。
- 从系统结构可以看出，系统正常工作当且仅当（组件1和2都正常工作）或（组件3和4都正常工作），即 $A = (A_1A_2) \cup (A_3A_4)$。
- 因此，系统的可靠性为

$$
\begin{aligned}
P(A) &= P(A_1A_2 \cup A_3A_4) = P(A_1A_2) + P(A_3A_4) - P(A_1A_2A_3A_4) \\
\text{由独立性} &= P(A_1)P(A_2) + P(A_3)P(A_4) - P(A_1)P(A_2)P(A_3)P(A_4) \\
&= p^2 + p^2 - p^4 = 2p^2 - p^4 = p^2(2 - p^2).
\end{aligned}
$$

—

1.3 条件概率与独立性

- 有时，事件可能不直接独立，但在特定事件发生的条件下是独立的。这就是**条件独立性**的概念。

条件独立性

设 $A_1, A_2, ..., A_n, B$ 均为随机事件。

- 如果对于事件的每一个子集 $A_{i_1}, A_{i_2}, ..., A_{i_k}$ $(1 \le i_1 < ... < i_k \le n, k = 2, ..., n)$，都有

$$P(A_{i_1}A_{i_2}...A_{i_k}|B) = P(A_{i_1}|B)P(A_{i_2}|B)...P(A_{i_k}|B),$$

则称 $A_1, A_2, ..., A_n$ 在给定 B 的条件下是**条件独立的**。

- 条件独立性是一个基础概念，在各个领域被广泛应用，使得对复杂系统的计算和推理变得高效。
- 条件独立性至关重要的关键领域包括**图模型**（例如贝叶斯网络）、**隐马尔可夫模型**、**因果推断中的结构因果模型**、**自然语言处理中的主题模型**（例如潜在狄利克雷分配）等。

—

1.3 条件概率与独立性

例 1.17

- 朴素贝叶斯是一种简单而强大的概率机器学习算法，用于分类任务，特别是在文本分类中，如垃圾邮件过滤、情感分析等。
- 它基于**贝叶斯定理**，并假设用于执行分类的特征之间具有**条件独立性**。
- 例如，在垃圾邮件过滤中，假设一封邮件中有 N 个词，W_i 表示第 i 个词在邮件中出现的事件（$i = 1, 2, ..., N$），S 表示邮件是垃圾邮件的事件，则

$$P(S|W_1 \cap W_2 \cap \cdots \cap W_N) = \frac{P(W_1W_2\cdots W_N|S)P(S)}{P(W_1W_2\cdots W_N|S)P(S) + P(W_1W_2\cdots W_N|\overline{S})P(\overline{S})}.$$

- 通过假设在给定邮件是否为垃圾邮件的条件下，邮件中出现的词是条件独立的，我们有

$$P(W_1W_2\cdots W_N|S) = P(W_1|S)P(W_2|S)\cdots P(W_N|S),$$

$$P(W_1W_2\cdots W_N|\overline{S}) = P(W_1|\overline{S})P(W_2|\overline{S})\cdots P(W_N|\overline{S}).$$

- 这极大地简化了计算。
- 然而，条件独立性假设可能不现实，这就是为什么它被称为“朴素”的原因。

—

1.3 条件概率与独立性

- 最后，我们来讨论现实中的独立性。
- 在数学中，判断事件是否独立并不难，因为它有明确的定义。
- 然而，在现实生活中，辨别事件是否独立可能并不容易。
- 例如，我们听说过著名的蝴蝶效应，“南美洲一只蝴蝶扇动翅膀，最终可能在美国德克萨斯州引起一场龙卷风”。

- 所以，看似不相关的事情可能会带来重大的变化。
- 现实中许多独立性实例是我们为了极大地简化问题而做出的假设，即独立性通常只是我们用来描述随机事件的数学模型。

例 1.18 请从系统开发难度的角度比较淘宝购物系统和12306售票系统。

（注：此例通常引发关于系统复杂性、并发处理、事务一致性、系统耦合度与独立性假设的讨论。淘宝商品之间购买相对独立，而12306票务系统涉及强一致性、库存瞬间竞争，事件间独立性假设更弱，系统开发更复杂。）

—

第一章结束

（本章内容结束）