

试验、事件与概率

1.1 等概率模型

1.1.1 定义

例 1.1.1 硬币的一面称为“头”(H),另一面称为“尾”(T). 抛掷一枚硬币, 落定后, 向上的面有两种可能性: H 或 T(排除硬币立着的可能). 假定硬币是均匀的, 则可以认为两种结果实现的“机会均等”. 如何用数量来表示“机会均等”的概念呢? 一种表示方法是用“机会比”(odds)——H 对于 T 的机会比为 $1:1$. 这种表示机会大小的方式在博彩业中被普遍使用. 但是在数学中使用的表示机会大小的概念是“概率”(probability): 用两个正数分别表示 H 和 T 实现的可能性, 其和为 1. 在机会均等的场合, H 和 T 实现的概率相同, 各为 $1/2$.

例 1.1.1 给出了一个最简单的“等概率模型”. 为以后便于叙述及一般化表示, 我们在此引进几个数学名词. 在例 1.1.1 中, 掷一枚硬币这件事称为一个“试验”(experiment), 可能的结果 H 和 T 称为“样本点”(sample points), 由样本点构成的集合 $\Omega = \{H, T\}$ 称为“样本空间”(sample space).

例 1.1.2 掷一枚骰子, 落定后观察点数. 这个试验的样本点有 6 个, 样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$. 假定骰子是均匀的, 则可以用等概率模型: 点数 $1 \sim 6$ 出现的概率相同, 各为 $1/6$.

假定甲乙两人打赌: 掷一枚骰子, 若出现的点数小于等于 3, 则甲胜; 否则乙胜. 显然甲乙两人取胜的机会均等——概率各为 $1/2$. 此事实数学上可以如此描述: 事件“出现的点数小于等于 3”是样本空间 Ω 中的一个子集 $\{1, 2, 3\}$, 其中包含 3 个样本点, 每个样本点的概率为 $\frac{1}{6}$, 其和为 $\frac{1}{2}$. 与之相互排斥的另一事件是“出现的点数大于等于 4”—— Ω 中的另一子集 $\{4, 5, 6\}$, 其概率也是 $\frac{1}{2}$. 由

1 试验、事件与概率

于这两个事件相互排斥,并且合并起来为整个样本空间 Ω ,因此它们的概率之和为 1.

从以上两个简单例子可以引出概率论中的一些基本概念.

- 一个试验可能产生的结果称为样本点,全体样本点构成的集合 Ω 称为样本空间.
- 事件(events)描述样本点的某些特性,是样本空间的子集. 每个样本点本身也是事件,样本空间 Ω 本身也是事件,称为必然事件. 为数学上的方便,把空集 $\emptyset$ 也定义为事件,称为不可能事件.
- 设一个试验产生的样本空间 Ω 为有限集合,假定对每个样本点赋予相同的概率,且保证所有样本点概率的总和为 1,则为一个等概率模型. 显然,若样本空间 Ω 中有 N 个样本点,则每个样本点的概率为 $\frac{1}{N}$.
- 在等概率模型中,任何一个事件(样本空间的子集)发生的概率为该事件中样本点概率之和. 若事件 A 中有 m 个样本点,则 A 发生的概率 $P(A) = \frac{m}{N}$. 必然事件 Ω 的概率为 1,不可能事件 $\emptyset$ 的概率为 0. 以后常用大写拉丁字母表示事件,任一事件 A 发生的概率用 $P(A)$ 来表示.

等概率模型是一种简单而又非常有用的模型,又称为“古典概率模型”,因为在概率论发展的早期,研究的问题都是等概率模型.

例 1.1.3 将一枚硬币掷两次,则样本点有 4 个: 样本空间 $\Omega = \{HH, HT, TH, TT\}$. 在等概率模型假定下,每个样本点的概率都是 $1/4$. 考虑下面的事件与相应的概率:

$$\text{事件 } A = \text{“第一次掷出现 H”} = \{HH, HT\}, P(A) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

$$\text{事件 } B = \text{“两次掷中至少出现一次 H”} = \{HH, HT, TH\}, P(B) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

问题: 这个试验共有多少个事件?

1.1.2 等概率模型中的计数方法

在等概率模型中,计算事件的概率涉及计数: 样本空间中的样本点总数和事件中的样本点数. 常用的计数方法有: 乘法法则、加法法则、排列数、组合数、牛顿二项式展开系数等. 我们假定读者已经在中学数学中学过这些基本计数方

法,为节省篇幅,不再赘述.先规定两个排列组合数的符号.

- 对于正整数 n , 规定 $n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$ 为 n 个物体各种不同的有序排列的总数. 为方便, 规定 $0! = 1$.

- 对于正整数 n 和 m , 规定 $\binom{n}{m}$ 为从 n 个物体中任意取出 m 个物体的不同取法的总数, 在中学数学教材中通常用 C_n^m 来表示. 显然, 当 $m > n$ 时, $\binom{n}{m} = 0$; 当 $m = 0, 1, 2, \cdots, n$ 时,

$$\binom{n}{m} = \frac{n(n-1)\cdots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}.$$

例 1.1.4 掷 n 枚硬币(或将一枚硬币掷 n 次), 样本空间 Ω 中有 2^n 个样本点, 每个样本点是 H 和 T 在 n 个位置上的排列(允许重复). 在等概率模型中, 每个样本点的概率为 $1/2^n$. 考虑事件 $B_m = \text{“H 恰出现 } m \text{ 次”}$. B_m 中有 $\binom{n}{m}$ 个样本点, 因此

$$P(B_m) = \binom{n}{m} \frac{1}{2^n}, \quad m = 0, 1, 2, \cdots, n.$$

事件 $B_m, m = 0, 1, 2, \cdots, n$ 中任何两个不能同时发生, 且覆盖整个样本空间, 因此它们的概率之和应该为 1. 此事实可以用牛顿二项展开式验证:

$$1 = \left(\frac{1}{2} + \frac{1}{2}\right)^n = \sum_{m=0}^n \binom{n}{m} \frac{1}{2^n} = \sum_{m=0}^n P(B_m).$$

例 1.1.5 掷两枚骰子, 令事件 $A = \text{“两枚骰子的点数都不超过 3”}$, 事件 $B = \text{“两枚骰子的点数之和大于等于 10”}$, 求 $P(A)$ 和 $P(B)$.

解 掷两枚骰子, 每枚骰子可能出现的点数为 $1 \sim 6$, 由乘法原理, 样本点总数为 $6 \times 6 = 36$.

对于事件 A , 每枚骰子可能出现的点数为 $1, 2, 3$, 由乘法原理, A 中样本点数为 $3 \times 3 = 9$, 因此 $P(A) = \frac{9}{36} = \frac{1}{4}$.

对于事件 B , 可以分成 3 个子事件: “两枚骰子的点数之和等于 k ”, $k = 10, 11, 12$. 事件“两枚骰子的点数之和等于 10”中有 3 个样本点: $(4, 6), (5, 5)$ 和 $(6, 4)$, 其概率为 $\frac{3}{36}$. 类似可以得到: 事件“两枚骰子的点数之和等于 11”中有 2 个样本点: $(5, 6)$ 和 $(6, 5)$, 其概率为 $\frac{2}{36}$; 事件“两枚骰子的点数之和等于 12”中

1 试验、事件与概率

有 1 个样本点 $(6, 6)$, 其概率为 $\frac{1}{36}$. 由加法原理, $P(B) = \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{1}{6}$.

下面介绍一些计算古典概率时有用的组合计数公式.

(1) 将 n 个可区别的球投入到 k 个编号 $1 \sim k$ 的盒子中. 此试验共有多少个样本点?

解 n 个球中的任一个投入到 k 个盒子中有 k 种结果, 根据乘法原理, 样本点总数为 k^n .

(2) 将 n 个可区别的球投入到 k 个编号 $1 \sim k$ 的盒子中. 给定非负整数 $n_1, n_2, \dots, n_k$, 满足 $n_1 + n_2 + \dots + n_k = n$, 定义事件 $C =$ “第 i 号盒恰有 n_i 个球, $i = 1, 2, \dots, k$ ”, 事件 C 发生的概率是多大?

解 从 n 个球中先取出 n_1 个球放到第 1 号盒中, 有 $\binom{n}{n_1}$ 种结果. 再从其余 $n - n_1$ 个球中取出 n_2 个球放到第 2 号盒中, 有 $\binom{n - n_1}{n_2}$ 种结果. 再从其余 $n - n_1 - n_2$ 个球中取出 n_3 个球放到第 3 号盒中, 有 $\binom{n - n_1 - n_2}{n_3}$ 种结果. 如此继续, 直到最后 n_k 个球放到第 k 号盒中. 根据乘法原理, 事件 C 中样本点总数为

$$\binom{n}{n_1} \binom{n - n_1}{n_2} \binom{n - n_1 - n_2}{n_3} \cdots \binom{n_k}{n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}.$$

因此 $P(C) = \frac{n!}{n_1! n_2! \cdots n_k!} k^{-n}$.

(3) (多项式展开系数) 将多项式 $(a_1 + a_2 + \dots + a_k)^n$ 展开后, 求项 $a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}$ 的系数, 其中 $n_1 + n_2 + \dots + n_k = n$.

解 $a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k}$ 的系数与问题(2)的结果数相等, 为 $\frac{n!}{n_1! n_2! \cdots n_k!}$.

特别地, 当 $k = 2$ 时为牛顿二项式展开系数: 将二项式 $(a + b)^n$ 展开后, 项 $a^m b^{n-m}$ 的系数为 $\binom{n}{m}$, $m = 0, 1, 2, \dots, n$.

(4) 求 $\sum_{n_1 + n_2 + \dots + n_k = n} \frac{n!}{n_1! n_2! \cdots n_k!}$.

解 将多项式 $(a_1 + a_2 + \dots + a_k)^n$ 展开后, 再令 $a_1 = a_2 = \dots = a_k = 1$, 得到

$$k^n = \sum_{n_1 + n_2 + \dots + n_k = n} \frac{n!}{n_1! n_2! \cdots n_k!}.$$

(5) n 个球在 k 个编号 $1 \sim k$ 的盒子中不同的数目分配共有多少种?

解 将 k 个盒子依顺序排成一排, 将相邻盒子的隔板看成是一个球, 与 n

个球合起来总共有 $n+k-1$ 个球. 从中任选 $k-1$ 个球看成是相邻盒子的隔板, 得到 n 个球在 k 个盒子中的一个数目分配. 因此, 不同的数目分配共有 $\binom{n+k-1}{k-1}$ 种.

(6) 将正整数 n 分成 k 个有序的非负整数之和, 共有多少种可能性?

解 “将正整数 n 分成 k 个有序的非负整数之和”等价于“将 n 个球投入到 k 个编号为 $1 \sim k$ 的盒子中的数目分配”. 因此, 此问题的答案也是 $\binom{n+k-1}{k-1}$.

(7) n 个球在 k 个编号为 $1 \sim k$ 的盒子中, 每个盒子都不空的数目分配共有多少种?

解 将 n 个球任意排成一排, 球之间的空当儿共有 $n-1$ 个, 从中任选 $k-1$ 个空当儿看成是相邻盒子的隔板, 得到 n 个球在 k 个非空盒子中的一个数目分配. 不同的数目分配共有 $\binom{n-1}{k-1}$ 种.

(8) 将正整数 n 分成 k 个有序的正整数之和, 共有多少种可能性?

解 此问题等价于“ n 个球在 k 个编号为 $1 \sim k$ 的盒子中, 每个盒子都不空的数目分配”. 因此, 此问题的答案也是 $\binom{n-1}{k-1}$.

1.1.3 等概率模型的应用

等概率模型虽然简单, 但是其重要性不可轻视. 本小节介绍等概率模型的两个重要应用.

1. 随机抽样

等概率模型最重要的应用莫过于统计学中的随机抽样(random sampling). 例如, 儿童的身高是评估儿童发育状况的重要指标. 假定要研究某地区(例如北京)10岁儿童的身高分布状况. 由于10岁儿童的总体非常庞大, 把每个儿童的身高都测量记录通常是既不现实也不必要的. 统计学方法是从总体中抽选有代表性的一小部分(称为“样本”), 通过对样本的研究分析来推断总体分布的重要特征. 问题是: 如何从总体中抽选有代表性的一小部分呢? 出乎一般人的想象, 统计学认为样本要有代表性, 就应该“闭着眼睛随便抽”. 确切地说, 是要保证总体中的每个个体有同等的机会(概率)被抽到. 这就是随机抽样的概念. 假定总体中有 N 个个体, 如果只抽一个样本, 则应该保证总体中的每个个体有 $1/N$ 的概

1 试验、事件与概率

率被抽到. 这恰好是一个等概率模型.

2. 抓阄的公平性问题(一)

抓阄是在机会稀缺时人们公平获得机会的常用方法,也是等概率模型应用的一个典型案例. 假定有 n 个人要抓阄, n 个阄中只有一个是“中”,其他都是“白”. 首先假定开阄方式是: 每个人先按任意顺序抓一个阄,再同时将 n 个阄打开看. 这种开阄方式可称为“同时开阄”. 在同时开阄方式下, n 个阄按随机的顺序分配到 n 个人. 所谓“随机的顺序”是指 n 个阄对于 n 个人的任意一种排列被等可能地实现. 由于 n 个阄的排列(样本点)总数是 $n!$, 因此每种排列实现的概率是 $1/n!$. 某人抽到“中”的事件中包含多少样本点呢? 假定此人抽到“中”,则其余 $n-1$ 个阄在其余 $n-1$ 个人中的排列共有 $(n-1)!$ 种, 因此, 此人抽到“中”的概率为 $\frac{(n-1)!}{n!} = \frac{1}{n}$. 此概率与人无关, 因此抓阄是公平的.

但是, 这个等概率模型不能直接解释另外一种抓阄方式的公平性. 另一种常见的抓阄方式是这样的: n 个人按任意顺序依次抓阄, 每个人抓完阄后立即打开看. 当某个人抓到“中”时, 整个抓阄过程就结束了, 因为后面的人没有再抓的必要. 这种开阄方式可称为“即时开阄”. 在这种开阄方式中, 排在后面的人往往会感觉比较吃亏, 因为他(她)可能还没有轮到抓阄, 已经有人中了. 上面那种按完全排列计算的等概率模型在即时开阄的场合似乎不适用. 我们将在 1.4 节证明: 在即时开阄方式中, 每个人的机会仍然是均等的.

1.2 试验、样本空间与事件

1.2.1 概念

等概率模型是一类简单的概率模型, 起源于公平博取机会的问题. 在自然和社会中有大量不确定现象, 可以用一般的概率模型来刻画. 在科学或工程试验中常遇到试验结果的不确定性. 例如, 对同一台纺纱机上纺出的同一型号的纱, 截取若干样本(纱段), 测试其抗拉强度, 结果可能不同. 这是因为纱本身不是绝对均匀的, 不同纱段的粗细和致密度不同, 因而抗拉强度不同. 自然或社会的进程(如天气、股市等)往往有多种可能的结果(天气的阴晴、股市的牛熊等), 事先无法肯定哪种结果会发生. 当我们对未来可能发生的自然或社会现象进行预测时, 不能准确无误地预测到实际发生的结果. 以后我们把所有实际的试验和对未来

自然或社会现象的预测统称为试验.

对于一个试验,每个可能发生的结果称为一个样本点,所有样本点所构成的集合称为样本空间,记为 Ω .具有一定特性的样本点子集(包括 Ω 和空集 $\emptyset$)称为事件,用大写拉丁字母表示.

例 1.2.1 硬币的一面称为“头”,记作 H,另一面称为“尾”,记作 T. 掷 n 枚硬币(或将一枚硬币掷 n 次), Ω 中有 2^n 个样本点. 事件 $B_m = \text{“H 恰出现 } m \text{ 次”}$ 中有 $\binom{n}{m}$ 个样本点, $m=0,1,2,\dots,n$.

例 1.2.2 编号 1~8 的 8 名运动员百米比赛结果的名次为 1~8 的一个排列, Ω 中有 $8!$ 个样本点. 事件 $C = \text{“1 号选手得冠军”}$ 中有 $7!$ 个样本点.

例 1.2.3 钓鱼比赛中,某选手钓到鱼的重量(kg)为一个实数, $\Omega = \{x: 0 \leq x < \infty\}$. 事件 $D = \text{“钓到鱼的重量超过 } 10 \text{ kg”}$ $= \{x: 10 < x < \infty\}$.

例 1.2.4 某银行在一小时营业时间内到来的顾客人数是一个非负整数, $\Omega = \{0,1,2,\dots\}$. 事件 $E = \text{“顾客人数在 } 100 \text{ 到 } 200 \text{ 之间”}$ $= \{100,101,\dots,200\}$.

例 1.2.5 股市一日中交易 4 h. 从开盘到时刻 t 的交易量(亿元)为 t 的一个单调增函数 $f(t)$, $\Omega = \{\text{区间}[0,4]\text{上的单调增函数 } f(t)\}$. 事件 $G = \text{“开盘后 } 1 \text{ h 内交易量超过 } 10 \text{ (亿元)”}$ $= \{\text{区间}[0,4]\text{上的单调增函数 } f(t): f(1) \geq 10\}$. 此例为一个“随机过程”模型.

例 1.2.6 一罐中有 6 个红球、3 个黄球和 1 个白球. 从中先后取出两球,则样本空间 Ω 中有 $10 \times 9 = 90$ 个样本点. 事件 $H = \text{“取出两个红球”}$ 中有 $6 \times 5 = 30$ 个样本点,事件 $I = \text{“取出一个红球”}$ 中有 $6 \times 4 + 4 \times 6 = 48$ 个样本点.

例 1.2.7 一罐中有两个红球和一个黄球. 从中先取一球,观察其颜色后,向罐中放回两个同色球;然后再从罐中取一球. 则样本空间 Ω 中有 $3 \times 4 = 12$ 个样本点. 事件 $J = \text{“取出两个红球”}$ 中有 $2 \times 3 = 6$ 个样本点,事件 $K = \text{“取出一个红球”}$ 中有 $2 \times 1 + 1 \times 2 = 4$ 个样本点.

由以上的例题可以看出:有各种各样不同的试验,相应的样本空间也千差万别. 对于一个试验,一定要想清楚其样本空间,不能丢失样本点. 对于等概率模型,也不能有多余的样本点. 例 1.2.6 和例 1.2.7 中的试验模型称为“罐子模型”. 在本书中,我们会遇到各种各样的罐子模型. 这些模型既直观易懂又有很实际的背景.

1.2.2 事件的运算与关系

当我们说到多个事件时,总是指在同一个样本空间上的事件. 由于事件是样

1 试验、事件与概率

本空间的子集,因此事件的运算与关系可以用集合的运算与关系来表示,但是通常使用概率论中特有的语言.

1. 事件的运算

- A 与 B 同时发生(交): $A \cap B$ 或 AB (可推广到多个事件同时发生,即多个事件的交);
- A 与 B 中至少发生一个(并): $A \cup B$ (可推广到多个事件中至少发生一个,即多个事件的并);
- A 发生但 B 不发生(差): $A \setminus B$;
- 事件 A 不发生(求余): $\Omega \setminus A = \bar{A}$ ($A\bar{A} = \emptyset, A \cup \bar{A} = \Omega$).

2. 事件的关系

- A 发生则 B 必不发生;反之亦然(互斥): $AB = \emptyset$ (可推广到多个事件两两互斥);
- A 与 B 中必然且至多发生一个(互余): $AB = \emptyset$, 且 $A \cup B = \Omega$ ($B = \bar{A}, A = \bar{B}$);
- A 发生则 B 必发生(A 含于 B): $A \subset B$ (或称 B 包含 A : $B \supset A$);
- A 发生则 B 必发生;反之亦然(等价): $A = B$ ($A \subset B$ 且 $B \subset A$).

由于事件即样本空间的子集,所以事件的运算与关系即集合的运算与关系,有相应的运算法则,如:

- 并交运算的结合律和交换律:

$$\begin{aligned}A \cup B \cup C &= A \cup C \cup B = C \cup A \cup B = \dots, \\ABC &= ACB = CAB = \dots;\end{aligned}$$

- 并交运算的分配律:

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

- De Morgan 公式: $\overline{\left(\bigcup_i A_i\right)} = \bigcap_i \bar{A}_i$.

事件的运算与关系可以用 Venn 图来直观表示. 我们假定读者已经对集合论的基本概念有一定基础,因此对上述的集合运算法则和 Venn 图不再赘述.

例 1.2.7(续) 一罐中有两个红球和一个黄球. 从中先取一球,观察其颜色后,向罐中放回两个同色球;然后再从罐中取一球. 记事件 A = “第一次取得红球”, B = “第二次取得红球”, C = “恰取得一个红球”, D = “至少取得一个红球”, E = “取得两个红球”, F = “没取得红球”. 用事件 A 和 B 表示其他事件.

解

$$C = (A\bar{B}) \cup (\bar{A}B), \quad D = A \cup B, \quad E = AB, \quad F = \overline{A\bar{B}} = \overline{A \cup B}.$$

若事件 $A_1, A_2, \dots, A_n$ 两两互斥, 且 $\bigcup_{i=1}^n A_i = \Omega$, 则称事件组 $\{A_1, A_2, \dots, A_n\}$ 为完备事件组, 其中 n 可以是无穷. 特别地, $\{A, \bar{A}\}$ 为完备事件组. 若 $\{A_1, A_2, \dots, A_n\}$ 为完备事件组, 则对任一事件 B 有

$$B = \bigcup_{i=1}^n (A_i B).$$

例 1.2.6(续) 一罐中有 6 个红球、3 个黄球、1 个白球. 从中先后取出两球. 记事件 A_1, A_2, A_3 分别为“第一次取出红球”、“第一次取出黄球”和“第一次取出白球”, 则 $\{A_1, A_2, A_3\}$ 为完备事件组. 记事件 B_1, B_2, B_3 分别为“第二次取出红球”、“第二次取出黄球”和“第二次取出白球”, 则 $\{B_1, B_2, B_3\}$ 也是完备事件组.

- 事件“两次都取到红球” $= A_1 B_1$;
- 事件“第二次取出红球” $= B_1 = (A_1 B_1) \cup (A_2 B_1) \cup (A_3 B_1)$;
- 事件“第一次取出红球” $= A_1 = (A_1 B_1) \cup (A_1 B_2) \cup (A_1 B_3)$;
- 事件“至少取到一个红球” $= A_1 \cup B_1 = (A_1 \setminus B_1) \cup (B_1 \setminus A_1) \cup (A_1 B_1)$;
- 事件“没取到红球” $= \overline{A_1 \cup B_1} = \bar{A}_1 \bar{B}_1$;
- 事件“取到两个不同色球” $= (A_1 \bar{B}_1) \cup (A_2 \bar{B}_2) \cup (A_3 \bar{B}_3)$.

例 1.2.2(续) 编号 1~8 的 8 名运动员百米比赛结果的名次为 1~8 的一个排列. 定义事件 $I_{ij} =$ “第 i 号选手获得第 j 名”, $i, j = 1, 2, \dots, 8$. 定义事件 $C_i =$ “第 i 号选手进入前两名”, $i = 1, 2, \dots, 8$. 显然有:

- $I_{ij} I_{i'j} = \emptyset, i \neq i', I_{ij} I_{ij'} = \emptyset, j \neq j'$;
- $C_i C_{i'} = (I_{i1} I_{i'2}) \cup (I_{i2} I_{i'1}), i \neq i'$;
- $C_i C_{i'} C_{i''} = \emptyset, i \neq i' \neq i'' \neq i$.

1.3 事件的概率

1.3.1 概率的公理与性质

在 1.1 节中我们介绍了等概率模型, 归纳了等概率模型下概率的一些性质. 在实际问题中经常会遇到不等概率模型. 对一般的概率模型, 我们需要从数学上规定概率应该满足的若干最基本的性质. 其他性质可以由这些最基本的性质推

1 试验、事件与概率

演出来. 概率最基本的性质有三条, 称为“概率公理”.

定义 1.3.1(概率公理) 在一个试验的样本空间 Ω 中, 记由所有事件构成的集合为 F . 对 F 中的每个事件 A 赋予一个实数 $P(A)$, 称为概率, 满足以下性质:

$$(1) P(A) \geq 0, \forall A \in F;$$

$$(2) P(\Omega) = 1;$$

$$(3) \text{对两两互斥的事件可列无穷序列 } A_1, A_2, \dots, P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

概率公理的前两条是把概率限制在 $[0, 1]$ 之内, 称为概率的“正规性”. 第(3)条性质称为“可列可加性”. 概率的可列可加性与计数、累计长度、累计面积、积分和累计比率等计算的性质一致, 统称为“测度”. 由概率公理可以推出概率的以下性质:

$$(4) P(\emptyset) = 0;$$

$$(5) \text{若有限个事件 } A_1, A_2, \dots, A_n \text{ 两两互斥, 则 } P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i);$$

$$(6) P(A) + P(\bar{A}) = 1;$$

$$(7) P(A) \leq 1;$$

$$(8) \text{若 } A \subset B, \text{ 则 } P(A) \leq P(B);$$

$$(9) \text{若 } A_1, A_2, \dots, A_n \text{ 两两互斥且 } \bigcup_{i=1}^n A_i = \Omega \text{ (其中 } n \text{ 可以是无穷), 则}$$

$$P(B) = \sum_{i=1}^n P(A_i B), \text{ 特别, } P(A) = P(AB) + P(A\bar{B});$$

$$(10) P(A \cup B) = P(A) + P(B) - P(AB).$$

下面给出概率性质(4)、性质(5)和性质(10)的部分证明.

性质(4)的证明:

定义事件的无穷序列: $A_1 = \Omega, A_i = \emptyset, i = 2, 3, \dots$, 根据概率公理的(2)和(3)有

$$\begin{aligned} 1 = P(\Omega) &= P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \\ &= 1 + P(\emptyset) + P(\emptyset) + \dots \end{aligned}$$

再由概率公理的(1)有 $0 \leq P(\emptyset)$, 因此必有 $P(\emptyset) = 0$.

性质(5)的证明:

设 $A_1, A_2, \dots, A_n$ 为两两互斥的有限个事件, 令 $A_i = \emptyset, i = n+1, n+2, \dots$, 则

$$P(A_i) = 0, \quad i = n+1, n+2, \dots$$

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

性质(6)~性质(9)的证明留给读者作为练习.

性质(10)的证明:

由 $A \cup B = A \cup (B \setminus A)$, 且 $A(B \setminus A) = \emptyset$, 根据性质(5)有

$$P(A \cup B) = P(A) + P(B \setminus A).$$

又由 $B = (AB) \cup (B \setminus A)$, 且 $(AB)(B \setminus A) = \emptyset$, 根据性质(5)有

$$P(B) = P(AB) + P(B \setminus A).$$

综合上述两式有

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

上述概率性质都是非常基本的, 务必熟练掌握.

例 1.3.1 掷两枚骰子, 求点数之和在 3~11 之间的概率.

解 掷两枚骰子, 每个骰子可能出现的点数为 1~6, 由乘法法则, 该试验的样本空间 Ω 共有 $6 \times 6 = 36$ 个样本点. 在等概率模型下, 每个样本点的概率为 $1/36$. 记事件 $A = \text{“点数之和在 3~11 之间”}$, A 中样本点总数的计算比较麻烦. 因为根据加法法则, 需要逐一计算事件组 $\{\text{“点数之和为 } k\text{”}, k=3, 4, \dots, 11\}$ 中每个事件所含的样本点数, 再相加. 不过我们可以改变思路.

记事件 $B = \text{“点数之和为 2”}$, 事件 $C = \text{“点数之和为 12”}$. 显然, 事件 A, B, C 两两互斥, 且 $A \cup B \cup C = \Omega$. 因此

$$P(A) = P(\Omega) - P(B) - P(C) = 1 - P(B) - P(C).$$

事件 B 和 C 中各只有一个样本点, 因此 $P(B) = P(C) = 1/36$. 因此 $P(A) = 34/36 = 0.94$.

例 1.3.2(例 1.2.6 续) 一罐中有 6 个红球、3 个黄球、1 个白球. 从中先后取出两球. 记事件 A_1, A_2, A_3 分别为“第一次取出红球”、“第一次取出黄球”和“第一次取出白球”; B_1, B_2, B_3 分别为“第二次取出红球”、“第二次取出黄球”和“第二次取出白球”. 求:

- (1) “两次都取到红球”的概率;
- (2) “第一次取出红球”的概率;
- (3) “第二次取出红球”的概率;
- (4) “至少取到一个红球”的概率;
- (5) “没取到红球”的概率;
- (6) “取到两个不同色球”的概率.

解 此试验的样本空间共有 $10 \times 9 = 90$ 个样本点, 在等概率模型下, 每个

1 试验、事件与概率

样本点的概率为 $1/90$.

(1) 事件“两次都取到红球”= $A_1 B_1$, 其中有 $6 \times 5 = 30$ 个样本点, 因此

$$P(A_1 B_1) = \frac{30}{90} = \frac{1}{3}.$$

(2) 事件“第一次取出红球”= A_1 , 其中有 $6 \times 9 = 54$ 个样本点, 因此

$$P(A_1) = \frac{54}{90} = \frac{3}{5}.$$

(3) 事件“第二次取出红球”= $B_1 = (A_1 B_1) \cup (\bar{A}_1 B_1)$, 因此

$$P(B_1) = P(A_1 B_1) + P(\bar{A}_1 B_1) = \frac{6 \times 5}{90} + \frac{4 \times 6}{90} = \frac{3}{5}.$$

注意 虽然 $P(B_1) = P(A_1)$, 但 $P(B_1)$ 的计算要困难些, 因为第二次取球的结果受到第一次取球结果的影响.

(4) 事件“至少取到一个红球”= $A_1 \cup B_1$, 因此

$$P(A_1 \cup B_1) = P(A_1) + P(B_1) - P(A_1 B_1) = \frac{13}{15}.$$

(5) 事件“没取到红球”= $\overline{A_1 \cup B_1}$, 则

$$P(\overline{A_1 \cup B_1}) = 1 - P(A_1 \cup B_1) = \frac{2}{15}.$$

(6) 事件“取到两个不同色球”= $(A_1 \bar{B}_1) \cup (A_2 \bar{B}_2) \cup (A_3 \bar{B}_3)$, 则

$$P[(A_1 \bar{B}_1) \cup (A_2 \bar{B}_2) \cup (A_3 \bar{B}_3)] = \frac{6 \times 4}{90} + \frac{3 \times 7}{90} + \frac{1 \times 9}{90} = \frac{3}{5}.$$

例 1.3.3

$$\begin{aligned} & P\left(\bigcup_{i=1}^n A_i\right) \\ &= P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus (A_1 \cup A_2)) + \cdots \\ & \quad + P(A_n \setminus (A_1 \cup A_2 \cup \cdots \cup A_{n-1})) \end{aligned}$$

证明 令 $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$, $\cdots$, $B_n = A_n \setminus (A_1 \cup A_2 \cup \cdots \cup A_{n-1})$. 显然, $B_i B_j = \emptyset, i \neq j$, 且 $\bigcup_{i=1}^n A_i = \bigcup_{i=1}^n B_i$. 因此

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

1.3.2 概率的背景

概率就是将事件发生的可能性数量化, 其背景可以从客观和主观两个角度来解释.

概率的客观背景即“频率背景”或“统计背景”. 将一枚均匀的硬币连续掷 n 次, 记录“头”向上的次数 n_H , 则“头”向上的频率为 $\frac{n_H}{n}$. 有人通过试验显示: 当 n 不断增大时, 频率 $\frac{n_H}{n}$ 会不断逼近 $\frac{1}{2}$. 一般, 若在一次试验中某个事件 A 发生的概率为 p , 则在 n 次重复试验中事件 A 发生的频率会随着 n 的不断增大而不断逼近 p . 这就是概率的频率背景.

如同力学定律一样, 概率的频率背景可以看成是自然运行所遵循的数学规律. 从纯粹数学的角度也可以严格证明(第4章): 当试验的重复次数 n 趋向于无穷时, 频率会收敛到概率. 不过首先要规定“收敛”的数学含义, 并假定以下前提: 试验可以独立地重复任意多次. 这里, 独立试验的含义将在 1.5 节中说明, 目前可直观理解为: 各次试验之间互不影响.

概率的频率背景启发统计学家: 若在一次试验中某个事件 A 发生的概率未知, 则可以用 n 次独立重复试验中 A 发生的频率估计概率. 以频率估计概率的思想概括了统计的基本思想.

但是, 现实中很多不确定现象(例如, 自然进程、社会进程等)是无法重复的. 这类不确定现象发生的概率难以用频率来解释. 但人们仍然对这类不确定现象赋予一定的概率, 这种概率具有一定的主观性(取决于人们的知识和个人的偏好). 表面上看主观概率会因人而异, 缺乏客观性. 其实, 如果主观概率是由专家所赋, 则其中包含有经验和知识的信息, 因此有一定程度的客观性.

无论是客观概率还是主观概率, 都必须满足概率公理, 纯粹从数学的角度看是没有区别的.

概率模型是人们用来描述随机现象的数学模型. 例如, 抛掷一枚真实的硬币, 由于硬币本身未必是十分均匀的, 因此两面出现的概率未必完全相同. 另外, 硬币有一定的厚度, 因此存在硬币立着的可能性(概率虽然很小, 但大于 0). 因此, 假定抛掷一枚硬币两面出现的概率各为 $1/2$, 只是一种“理想模型”, 是对现实掷硬币试验的一种近似. 人们对随机现象所赋予的概率模型可以是好的, 也可以是坏的. 在抛掷一枚硬币的试验中, 如果硬币确实是均匀的, 则可以采用等概率模型, 这是一个好的模型. 但是, 如果硬币不是均匀的(可能是被人做了手脚的), 这时头向上的概率与尾向上的概率会不同, 如果再采用等概率模型, 则是一个坏的模型. 如果不知道硬币是否均匀, 则可以假定头向上的概率为 p , 尾向上的概率为 $1-p$. 现在 p 是未知的, 概率论本身不能回答 $p=?$ 的问题. 统计学则提供 p 的估计. 根据“频率估计概率”的思想, 将该硬币掷 n 次, 如果头出现 n_H 次, 可以用频率 $\frac{n_H}{n}$ 来估计 p . 这就是统计与数学的差别: 数学是从原因推导结

果,而统计是从结果反推原因.

1.4 条件概率

1.4.1 条件概率与乘法公式

在不同条件下,事件发生的概率可能不同,因为事先得到的信息可能不同.

例 1.4.1 一号罐中有两个红球,一个白球;二号罐中有一个红球,两个白球.从一号罐中随机取一个球放到二号罐中,再从二号罐中随机取一个球.求:

(1) 从二号罐中取得红球的概率.

(2) 在已知一号罐中取得的是红球的条件下,从二号罐中取得红球的概率.

解 记事件 A = “从一号罐中取得红球”, B = “从二号罐中取得红球”. 样本空间 Ω 共有 $3 \times 4 = 12$ 个样本点,每个样本点的概率为 $1/12$.

(1) 事件 $B = (AB) \cup (\bar{A}B)$, 有 $2 \times 2 + 1 \times 1 = 5$ 个样本点,因此 $P(B) = 5/12$.

(2) 解决方案一: 在已知一号罐中取得的是红球的条件下,二号罐中应该有红、白球各两个,取得红球的概率为 $1/2$.

解决方案二: 在 A 发生的条件下, B 发生意味着 AB 发生,因此在 A 发生的条件下 B 发生的概率应为 $\frac{P(AB)}{P(A)}$. $P(A) = 8/12$, $P(AB) = 4/12$, 故 $\frac{P(AB)}{P(A)} = 1/2$.

例 1.4.1 第(2)问中的两种解决方案得到相同结论,其中第一种解决方案在等概率模型中直观、易懂,而第二种解决方案具有数学上的一般性.从第(1)问和第(2)问的答案中可以看出:一个事件发生的无条件概率与条件概率是可以不同的.

定义 1.4.1 设 A 和 B 为任意两个事件,且 $P(A) > 0$, 在 A 发生的条件下, B 发生的条件概率定义为 $P(B|A) = \frac{P(AB)}{P(A)}$.

上述定义很直观,但缺点是要求 $P(A) > 0$. 为方便使用,规定: 当 $P(A) = 0$ 时, $P(B|A) = 0$. 这样,在任何情况下都有

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

此公式称为乘法公式. 在一些实际问题中,条件概率 $P(B|A)$ 往往根据试验背景容易得到,因此常用乘法公式计算 $P(AB)$.

例 1.4.2 一号罐中有 n_1 个红球, m_1 个白球; 二号罐中有 n_2 个红球, m_2 个白球. 从一号罐中随机取一个球放到二号罐中, 再从二号罐中随机取一个球. 求: 两次都取得红球的概率.

解 记事件 A = “从一号罐中取得红球”, B = “从二号罐中取得红球”, 则

$$P(A) = \frac{n_1}{n_1 + m_1}.$$

在事件 A 发生的前提下, 二号罐中有 $n_2 + 1$ 个红球, m_2 个白球, 因此

$$P(B | A) = \frac{n_2 + 1}{n_2 + m_2 + 1}.$$

由乘法公式有

$$P(AB) = P(A)P(B | A) = \frac{n_1}{n_1 + m_1} \cdot \frac{n_2 + 1}{n_2 + m_2 + 1}.$$

在例 1.4.1(2) 的解决方案二和例 1.4.2 中, 我们实际上采用了这样的思路: 已知事件 A 发生, 相当于把原来的样本空间 Ω 压缩到一个小的样本空间 A 上. 因此, 在原样本空间 Ω 中的任意事件 B 被压缩到小样本空间 A 中的事件 BA . 因此, 条件概率 $P(B|A)$ 应该与 BA 在原样本空间 Ω 上的概率 $P(AB)$ 成正比, 即

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

但是, 在小样本空间 A 上应该有 $P(A|A)=1$, 或者说: 在已知事件 A 发生的条件下, A 发生的概率当然应该为 1. 因此, 应该有

$$1 = P(A | A) = aP(AA) = aP(A).$$

由此得到 $a=1/P(A)$. 这个思路有助于我们直观理解条件概率, 并解决一些实际问题.

但是还可以有一种更具一般性的思路. 即把条件概率看成是在原样本空间 Ω 上定义的一个新的概率函数. 若 $P(A)>0$, 则对样本空间 Ω 上的事件集 F 中的任意事件 B 都定义了条件概率 $P(B|A) = \frac{P(AB)}{P(A)}$. 要证明条件概率 $P(\cdot | A)$ 也可以看成是 F 上的概率函数, 只需证明 $P(\cdot | A)$ 满足三条概率公理 (定义 1.3.1).

定理 1.4.1 设 A 为样本空间 Ω 上的一个事件, $P(A)>0$, 则在 A 发生条件下的条件概率函数 $P(\cdot | A)$ 满足定义 1.3.1 中的概率公理.

定理的证明留给读者作为练习.

根据定理 1.4.1, 条件概率 $P(\cdot | A)$ 是 Ω 上的概率函数. 一个自然的问题就产生了: 对这个概率函数是否还可以再定义条件概率呢? 假定在已知事件 A 发生的条件下, 又知道事件 C 也发生了. 可以有两种等价的方式来定义新的条件

1 试验、事件与概率

概率. 一种方式是: 在条件 A 之上再添加条件 C , 则意味着条件 AC (A 与 C 同时发生), 这时, 若 $P(AC) > 0$, 任意事件 B 发生的条件概率定义为

$$P(B|AC) = \frac{P(BAC)}{P(AC)}.$$

另一种方式是对条件概率 $P(\cdot | A)$ 再定义条件概率. 以下记 $P(\cdot | A) = P_A(\cdot)$. 对于概率函数, 定义已知事件 C 发生的条件概率 $P_A(\cdot | C)$:

$$P_A(B|C) = \frac{P_A(BC)}{P_A(C)}.$$

不难看出: 这两种定义方式是等价的(请验证). 在实际问题中前一种方式更方便易懂. 采用第一种方式, 我们可以不断添加条件. 无论有多少条件, 最终的条件概率是在把所有条件看成是同时发生的一个条件下计算的. 利用这个思想, 不难将关于两个事件同时发生的乘法公式推广到多个事件同时发生的乘法公式.

定理 1.4.2(乘法公式的推广) 设 $A_1, A_2, \dots, A_n$ 为一组事件, 则

$$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1}).$$

证明

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

例 1.4.3 一号罐中有 n_1 个红球, m_1 个白球; 二号罐中有 n_2 个红球, m_2 个白球; 三号罐中有 n_3 个红球, m_3 个白球. 从一号罐中随机取一个球放到二号罐中, 再从二号罐中随机取一个球放到三号罐中, 最后从三号罐中随机取一个球. 求: 三次都取得红球的概率.

解 记事件 A = “从一号罐中取得的是红球”, B = “从二号罐中取得红球”, C = “从三号罐中取得红球”, 则

$$P(ABC) = P(A) P(B | A) P(C | AB) = \frac{n_1}{n_1 + m_1} \frac{n_2 + 1}{n_2 + m_2 + 1} \frac{n_3 + 1}{n_3 + m_3 + 1}.$$

例 1.4.4 52 张扑克牌随机分成四组, 每组 13 张, 求每组中恰有一张 K 的概率.

解 定义下列事件:

A = “黑桃 K 在任一组中”;

B = “黑桃 K 和红心 K 不在同一组中”;

$C = \text{“黑桃 K、红心 K 和草花 K 都不在同一组中”}$;

$D = \text{“四个 K 都不在同一组中”}$.

显然 $P(D) = P(ABCD)$.

$$P(A) = 1, P(B | A) = \frac{39}{51}, P(C | AB) = \frac{26}{50}, P(D | ABC) = \frac{13}{49},$$

$$P(D) = P(ABCD) = \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49} \approx 0.105.$$

例 1.4.5 再论抓阄公平性问题.

我们利用条件概率来解决即时开阄方式的公平性研究. 假定 n 个人按任一顺序依次抓阄, 每个人抓完阄后立即打开看. 当某个人抓到“中”时, 整个抓阄过程就结束. 首先考虑第一人抓中的概率. 由于有 n 个阄, 因此第一人抓中的概率为 $\frac{1}{n}$. 然后考虑第二人抓中的概率. 第二人只有在第一人没抓中时才可能抓中.

第一人没抓中的概率为 $\frac{n-1}{n}$. 在第一人没抓中的前提下, 剩余的 $n-1$ 个阄中有

一个“中”, 此时第二人抓中的条件概率为 $\frac{1}{n-1}$. 因此第二人抓中的无条件概率

为 $\frac{n-1}{n} \cdot \frac{1}{n-1} = \frac{1}{n}$. 再考虑第三人. 前两个人都没有抓中的概率为 $\frac{n-1}{n} \cdot \frac{n-2}{n-1} =$

$\frac{n-2}{n}$. 在此前提下, 第三人抓中的条件概率为 $\frac{1}{n-2}$. 因此, 第三人抓中的无条件

概率为 $\frac{n-2}{n} \cdot \frac{1}{n-2} = \frac{1}{n}$. 依此顺序类推, 用归纳法不难证明每个人抓中的(无条件)

概率都是 $\frac{1}{n}$. 这样就回答了即时开阄方式的公平性问题, 而无需假定所有的

阄都要被抓过.

问题: 上述模型的样本空间是什么? 样本点各具有什么样的概率?

在很多实际问题中, 我们的兴趣或观察往往只集中在一类事件, 人们只对这些事件是否发生有兴趣, 而对事件内样本点的区分不感兴趣或不记录. 例如, 在抓阄问题中, 人们只关注是谁抓中了, 而对其他人抓到的阄不加区分. 这类事件称为“基本事件”. 如果一组基本事件两两不相交, 且覆盖整个样本空间, 则构成一个完备的基本事件组. 在抓阄问题中, n 个人中的每一个人抓中对应一个基本事件, 总共有 n 个基本事件, 每个基本事件的概率为 $1/n$.

由完备的基本事件组可以导出一个经过压缩的新样本空间, 其中每个样本点对应原样本空间中的一个基本事件, 其发生的概率即是原样本空间中该基本事件的概率. 两个概率模型在数学上是等价的, 如果可以分别从两个样本空间中

1 试验、事件与概率

各找到一组完备的基本事件,它们之间有一一对应的关系,且对应的基本事件有相同的概率.

显然,同时开闸的模型与即时开闸的模型在数学上是等价的:可以压缩到同一个样本空间,其中每个样本点对应一个人抓中,概率为 $1/n$. 利用这种等价的概率模型的观点,在实际问题中我们可以根据方便的原则在两个等价的概率模型中选择一个.

遵循上述思路,我们把 n 个闸中只有一个“中”的简单情况推广到有 m 个“中”的一般情况. 现在,用 n 个人抓完所有闸后同时开闸的模型更方便. m 个“中”被 n 个人中的 m 个人抓到的不同结果共有 $\binom{n}{m}$ 种,某人抓中,意味着其余 $m-1$ 个“中”被另外 $n-1$ 个人中的 $m-1$ 个人抓到,不同结果有 $\binom{n-1}{m-1}$ 种. 因此,某人抓中的概率为 $\binom{n-1}{m-1} / \binom{n}{m} = \frac{m}{n}$. 此概率与具体的人无关.

1.4.2 全概率公式与贝叶斯公式

如果事件 A 的发生能够提供事件 B 发生与否的信息,那么事件 $\bar{A}$ (A 不发生)也能提供事件 B 发生与否的信息. 由于 A 与 $\bar{A}$ 构成一个完备事件组,即 $A\bar{A} = \emptyset, A \cup \bar{A} = \Omega$, 因此

$$P(B) = P(BA) + P(B\bar{A}) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}).$$

此公式是“全概率公式”的一个特例. 在此公式中,信息 $P(A), P(B|A)$ 和 $P(B|\bar{A})$ 决定了 $P(B)$.

例 1.4.6 一号罐中有 n_1 个红球, m_1 个白球;二号罐中有 n_2 个红球, m_2 个白球. 从一号罐中随机取一个球放到二号罐中,再从二号罐中随机取一个球. 求第二次取得红球的概率.

解 记事件 $A =$ “第一次取得红球”, $\bar{A} =$ “第一次取得白球”, $B =$ “第二次取得红球”, 则

$$\begin{aligned} P(B) &= P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) \\ &= \frac{n_1}{n_1 + m_1} \cdot \frac{n_2 + 1}{n_2 + m_2 + 1} + \frac{m_1}{n_1 + m_1} \cdot \frac{n_2}{n_2 + m_2 + 1} \\ &= \frac{n_2(n_1 + m_1) + n_1}{(n_1 + m_1)(n_2 + m_2 + 1)}. \end{aligned}$$

在例 1.4.6 中, A 与 $\bar{A}$ 可以解释为两个不同的“原因”, 事件 B 可以解释为

“结果”. 公式 $P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$ 说明: 知道不同的原因发生的概率 $P(A), P(\bar{A})$ (事实上, 由于 $P(A) + P(\bar{A}) = 1$, 只要知道其中一个即可) 和在不同的原因发生的前提下, 结果发生的条件概率 $P(B|A), P(B|\bar{A})$, 就可以得到结果发生的“全概率” $P(B)$. 很多情况下, 原因可能有多个. 设 $A_1, A_2, \dots, A_n$ 为一完备事件组, 即满足: $A_1 \cup A_2 \cup \dots \cup A_n = S$, 且 $A_i A_j = \emptyset, i \neq j$. 完备事件组 $\{A_1, A_2, \dots, A_n\}$ 可解释为各种可能的原因.

全概率公式 任一事件 B 发生的“全概率” $P(B)$ 可表示为

$$\begin{aligned} P(B) &= P(A_1 B) + P(A_2 B) + \dots + P(A_n B) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n). \end{aligned}$$

例 1.4.7 有三个罐子, 第 i 号罐中有 n_i 个红球, m_i 个白球, $i=1, 2, 3$. 先从三个罐中选一个, 以概率 π_i 选中第 i 号罐, $i=1, 2, 3, \pi_1 + \pi_2 + \pi_3 = 1$; 再从选中的罐中随机取一个球. 求: 取得红球的概率.

解 记事件 $A_i =$ “选中 i 号罐”, $i=1, 2, 3; B =$ “取得红球”. 则

$$P(A_i) = \pi_i, \quad i = 1, 2, 3;$$

$$P(B|A_i) = \frac{n_i}{n_i + m_i}, \quad i = 1, 2, 3;$$

$$P(B) = \sum_{i=1}^3 P(A_i)P(B|A_i) = \sum_{i=1}^3 \frac{\pi_i n_i}{n_i + m_i}.$$

数学的逻辑是从原因推演结果. 但是在现实中, 人们常要从结果反推原因. 例如, 警察在侦破刑事案件时, 要根据搜集的证据来确定犯罪嫌疑人. 全概率公式可以解释从原因推演结果. 如何从结果反推原因呢? 设事件 A 为原因, 事件 B 为结果, 我们想得到在结果 B 发生的前提下, 原因 A 发生的条件概率 $P(A|B)$. 由乘法公式, 我们有

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

从而可以解出

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

这个看起来很简单的公式即著名的贝叶斯公式(Bayes formula), 它在统计学中占有非常重要的地位. 统计学思维的基本特征之一就是结果推断原因(另一个基本特征是从部分推断全体). 在贝叶斯公式左端的 $P(A|B)$ 是在已知结果 B 发生的前提下原因 A 发生的条件概率, 称为后验概率(posterior probability). 贝叶斯公式的右端给出获取后验概率所需的信息: $P(A), P(B|A)$ 和 $P(B)$, 其中 $P(A)$ 称为先验概率(prior probability). 这里, “先验”和“后验”的概念与哲学无关, 其中的“验”是指“试验”, “先验”的意思是“试验之前”, 而“后验”的意思

1 试验、事件与概率

是“试验之后”。

一般,我们手头的信息有 $P(A)$, $P(B|A)$ 以及 $P(B|\bar{A})$. 为得到后验概率 $P(A|B)$, 还需要计算出 $P(B)$. 由全概率公式

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

可得到贝叶斯公式的常见形式:

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

例 1.4.6(续) 一号罐中有 n_1 个红球, m_1 个白球; 二号罐中有 n_2 个红球, m_2 个白球. 从一号罐中随机取一个球放到二号罐中, 再从二号罐中随机取一个球. 求: 已知第二次取得红球时, 第一次取得红球的后验概率.

解 前面已解出

$$P(A)P(B|A) = \frac{n_1}{n_1 + m_1} \cdot \frac{n_2 + 1}{n_2 + m_2 + 1},$$

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = \frac{n_2(n_1 + m_1) + n_1}{(n_1 + m_1)(n_2 + m_2 + 1)},$$

从而由贝叶斯公式有

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{n_1(n_2 + 1)}{n_2(n_1 + m_1) + n_1}.$$

例 1.4.8 在信号传输过程中, 发送机只发送 0, 1 两个信号. 假定: 当发送 0 时, 接收机以概率 0.99 收到 0, 以概率 0.01 收到 1; 当发送 1 时, 接收机以概率 0.97 收到 1, 以概率 0.03 收到 0. 又假定发送机以概率 0.75 发送 1, 以概率 0.25 发送 0. 求: 当接收到 0 时, 实际发送的是 1 的条件概率.

解 记事件 0 = “发送 0”, 事件 1 = “发送 1”; 事件 $\hat{0}$ = “接收 0”, 事件 $\hat{1}$ = “接收 1”. 则

$$P(0) = 0.25, \quad P(1) = 0.75;$$

$$P(\hat{0}|0) = 0.99, \quad P(\hat{0}|1) = 0.03;$$

$$P(1|\hat{0}) = \frac{P(1)P(\hat{0}|1)}{P(\hat{0})} = \frac{0.75 \times 0.03}{0.25 \times 0.99 + 0.75 \times 0.03} = 0.0833.$$

因此, 当接收到 0 时, 实际发送的是 1 的条件概率为 0.0833.

当有多个原因时, 用一个完备事件组 $\{A_1, A_2, \dots, A_n\}$ 来表示, 后验概率

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{P(B)}, \quad j = 1, 2, \dots, n,$$

其中 $P(B)$ 用全概率公式计算.

例 1.4.9 有三个罐子, 第 1 号罐中有 1 个红球, 3 个白球; 第 2 号罐中有 3 个红球, 1 个白球; 第 3 号罐中有 4 个红球, 1 个白球. 先从三个罐中选一个, 选中