(3) 弦可被弦的中心点唯一确定, 当且仅当弦的中心点位于半径为 1/2 的同心圆内时, 弦长才会大于 1/3. 半径为 1/2 的小圆面积为 $\pi/4$, 大圆面积为 π, 故所求概率等于 1/4 (如图 1.3-**c**).

同一问题却有三种不同的答案, 其根本原因在于取弦时采用不同的等可能性假定. 第一种方法假定端点在圆周上均匀分布, 第二种方法假定弦的中点在直径上均匀分布, 第三种方法假定弦的中心点在小圆内均匀分布. 这三种方法采用三种不同的随机试验, 对于各自的随机试验而言, 它们都是正确的, 因此概率的计算一定要明确具体的试验.

贝特朗奇论在现代概率的发展中起到过重要作用, 它直接反驳了"等可能性可完全定义概率"的观点. 从此概率论开始向公理化方面发展, 根据满足的基本性质来定义概率, 而不是某些具体事件的概率. 基于此, 希尔伯特于 1900 年在巴黎举行的第二届数学家大会上提出了 23 个著名的数学问题, 其中第六个问题就是如何概率公理化.

1.4 概率公理化

随机事件在一次试验中可能发生也可能不发生, 我们通常关心随机事件发生的可能性究竟有多大, 最好能用介于 0 和 1 之间的一个数来进行刻画. 为此首先引入频率, 用以描述随机事件发生的频繁程度, 在此基础上引入事件的概率, 这就是基于统计的概率.

1.4.1 频率的稳定性

定义 1.5 随机事件 A 在相同条件下重复进行的 n 次试验中出现了 n_A 次, 则称 $f_n(A) = n_A/n$ 为事件 A 在 n 次试验中发生的 **频率**, 并称 n_A 为事件 A 发生的 **频数**.

事件的频率在一定程度上反应了事件发生的可能性, 若事件发生的频率越大, 则事件 A 发生越频繁, 因而事件在一次试验中发生的可能性越大. 根据上面的定义可知频率具有如下性质:

$1°$ 对任意事件 A 有 $f_n(A) \in [0, 1]$;

$2°$ 对必然事件 Ω 有 $f_n(\Omega) = 1$;

$3°$ 对互不相容的事件 $A_1, A_2, \cdots, A_k$ 有 $f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$.

性质 $1°$ 和 $2°$ 根据定义显然成立. 对互不相容的事件 $A_1, A_2, \cdots, A_k$, 并事件 $A_1 \cup A_2 \cup \cdots \cup A_k$ 发生的频数等于每个事件 A_i 发生的频数之和, 由此可知性质 $3°$ 成立.

频率在实际中通常表现出一定的随机性, 例如, 在相同条件下进行两轮 n 次试验, 每轮试验中事件 A 发生的频率往往不同. 其次, 随着试验次数 n 的增加, 事件 A 发生的频率 $f_n(A)$ 会发生一定的变化, 表现出一定的随机性.

尽管频率表现出一定的随机性, 但经过大量重复的试验, 事件的频率通常在一个确定的常数 p 附近摆动, 而且随着试验次数的增大, 摆幅越来越小, 频率也越来越稳定于常数 p, 将这种规律称为 **频率的稳定性**. 例如历史上有多人做过重复投掷硬币的试验, 表 1.2 列出了其中一些试验统计结果.

表 1.2 历史上多人重复投掷硬币的试验结果

实 验 者	投掷总数	正面朝上的频数	正面朝上的频率
德摩根	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
K. 皮尔逊	12000	6019	0.5016
K. 皮尔逊	24000	12012	0.5005

图 1.4 任意投掷硬币, 正面朝上频率的趋势

也可以利用计算机产生随机数对投掷硬币的试验进行模拟仿真, 图 1.4 给出了相应的试验结果. 这些研究结果均表明: 尽管对不同的投掷总数, 正面朝上的频率不并相同, 但随着投掷次数的增加, 正面朝上的频率越来越接近常数 1/2, 即频率逐渐稳定于 1/2. 这种频率的稳定性即通常所说的统计规律性, 是随机事件本身所固有的客观属性, 可用于度量事件发生的可能性大小.

定义 1.6 随机事件 A 在大量重复试验中发生的频率总是稳定地在一个常数 p 附近摆动, 且随着试验次数的增加而摆幅逐渐越小, 则称常数 p 为事件 A 发生的 **概率**, 记为 $P(A)=p$.

该定义又称 **概率的统计定义**, 其概率称为 **统计概率**, 提供了计算随机事件概率的一种方法, 即当试验次数足够多时, 可用频率来给出事件概率的近似值. 概率的统计定义存在着数学上的不严谨性, 在实际中也不太可能每一个事件做大量重复试验来计算频率, 进而近似概率. 受到频率的稳定性及其性质的启发, 下面给出严谨的概率公理化定义.

1.4.2 概率公理化

古典概率和几何概率只适用于等可能的情形, 而统计概率需要大量的重复试验才能得到较准确的概率值, 在数学上不具严谨性. 为克服这些定义的局限性, 20 世纪 30 年代前苏联数学家柯尔莫哥洛夫 (A. Kolmogorov) 提出了概率论的公理化体系, 通过基本的性质给出了概率的严格定义, 建立可媲美于欧氏几何公理化的理论体系.

定义 1.7 (概率公理化) 在可测空间 (Ω, Σ) 上, 若函数 $P: \Sigma \to R$ 满足以下条件:

1° **非负性**: 对任意 $A \in \Sigma$ 有 $P(A) \geqslant 0$;

2° **规范性**: 对样本空间 Ω 有 $P(\Omega) = 1$;

3° **可列可加性**: 若 $A_1, A_2, \cdots, A_n, \cdots \in \Sigma$ 是可列个互不相容的事件, 即 $A_iA_j = \emptyset\ (i \neq j)$, 有 $P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$,

则称 $P(A)$ 为随机事件 A 的 **概率**, 称 (Ω, Σ, P) 为 **概率测度空间** 或 **概率空间** (probability space).

概率 $P(A)$ 是定义在可测空间 (Ω, Σ) 上的实值函数, 满足非负性、规范性和可列可加性三条公理, 该定义简明扼要地刻画了概率的本质, 为现代概率论奠定了基础, 公理化体系是概率论发展历史上的一个里程碑, 从此概率论被公认为数学的一个重要分支.

根据概率公理化的定义, 可以推导出很多概率的性质, 进而形成概率论的学科体系.

性质 1.1　对不可能事件 $\emptyset$ 有 $P(\emptyset) = 0$.

证明　令 $A_i = \emptyset\ (i = 1, 2, \cdots)$, 则有 $\emptyset = \cup_{i=1}^{\infty} A_i$ 且 $A_i \cap A_j = \emptyset\ (i \neq j)$. 根据公理 3° 有

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

再根据公理 1° 可知 $P(\emptyset) = 0$.

不可能事件 $\emptyset$ 的概率为 0, 但概率为 0 的事件并不一定是不可能事件; 同理, 必然事件 Ω 的概率为 1, 但概率为 1 的事件并不一定是必然事件. 例题可参考几何概型.

性质 1.2 (有限可加性)　若 $A_1, A_2, \cdots, A_n$ 是两两不相容的事件, 则

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

证明　令 $A_i = \emptyset\ (i > n)$, 则有 $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i$, 且 $A_1, A_2, \cdots, A_n, A_{n+1}, \cdots$ 是两两互不相容事件. 根据公理 3° 可知

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

性质得证.

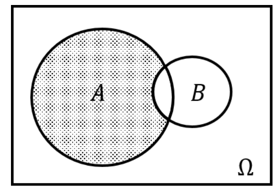
对任意随机事件 A 有 $\Omega = \bar{A} \cup A$ 和 $\bar{A} \cap A = \emptyset$ 成立, 再根据有限可加性和规范性有 $1 = P(\Omega) = P(A) + P(\bar{A})$.

推论 1.1　对任意随机事件 A 有 $P(\bar{A}) = 1 - P(A)$.

性质 1.3 对任意事件 A 和 B, 有

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

证明 如右图所示 $A=(A-B)\cup(AB)$, 以及 $A-B$ 与 AB 互斥, 根据有限可加性则有



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

再根据 $A\cup B=(A-B)\cup B$, 以及 $A-B$ 与 B 互斥, 则有

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

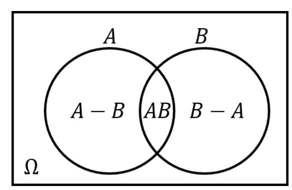
推论 1.2 若事件 $B\subset A$, 则有 $P(A-B)=P(A)-P(B)$ 和 $P(B)\leqslant P(A)$.

性质 1.4 (容斥原理) 对任意随机事件 A 和 B 有

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

证明 如右图所示

$$A\cup B=(A-B)\cup(AB)\cup(B-A),$$



再根据 $A-B$, $B-A$, AB 两两互不相容, 由有限可加性可知

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

再将 $P(A-B)=P(A)-P(AB)$ 和 $P(B-A)=P(B)-P(AB)$ 代入上式即可完成证明.

类似地, 对三个随机事件 A,B,C 有

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

进一步对 n 个随机事件 $A_1,A_2,\cdots,A_n$ 有

$$\begin{aligned}P\left(\bigcup_{i=1}^{n}A_i\right) &= \sum_{i=1}^{n}P(A_i)-\sum_{i<j}P(A_iA_j)+\sum_{i<j<k}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1\cdots A_n)\\ &= \sum_{r=1}^{n}(-1)^{r+1}\sum_{i_1<\cdots<i_r}P(A_{i_1}\cdots A_{i_r}).\end{aligned}$$

性质 1.5 (Union Bound) 对事件 $A_1, A_2, \dots A_n$ 有

$$P(A_1 \cup A_2 \cup \dots \cup A_n) \leqslant P(A_1) + P(A_2) + \dots + P(A_n).$$

证明 可以利用数学归纳法进行证明. 当 $n = 2$ 时, 由容斥原理有

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2) \leqslant P(A_1) + P(A_2). \tag{1.3}$$

假设当 $n = k$ 时性质成立, 对 $n = k + 1$ 有

$$\begin{aligned} P(A_1 \cup \dots \cup A_{k+1}) &= P((A_1 \cup \dots \cup A_k) \cup A_{k+1}) \\ &\leqslant P(A_1 \cup \dots \cup A_k) + P(A_{k+1}) \leqslant P(A_1) + \dots + P(A_k) + P(A_{k+1}), \end{aligned}$$

这里第一个不等式成立是根据式 (1.3), 而第二个不等式成立是根据归纳假设. 完成证明.

Union bounds 又被称为布尔不等式, 在很多研究中经常用到. 类似地还可以得到下列不等式:

推论 1.3 (Bonferroni 不等式) 对事件 $A_1, A_2, \dots A_n$ 有

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &\leqslant \sum_{i=1}^n P(A_i), \\ P\left(\bigcup_{i=1}^n A_i\right) &\geqslant \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i A_j), \\ P\left(\bigcup_{i=1}^n A_i\right) &\leqslant \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k), \\ \cdots \quad & \cdots \quad \cdots \end{aligned}$$

例 1.13 设 $P(A) = p$, $P(B) = q$, $P(AB) = r$, 用 p, q, r 分别表示事件的概率: 1) $P(\bar{A} \cup \bar{B})$, 2) $P(\bar{A}B)$; 3) $P(\bar{A} \cup B)$; 4) $P(\bar{A} \cap \bar{B})$.

解 对问题 1), 根据事件的德摩根律有

$$P(\bar{A} \cup \bar{B}) = P(\overline{AB}) = 1 - r.$$

对问题 2), 根据差事件的定义

$$P(\bar{A}B) = P(B - A) = P(B) - P(AB) = q - r.$$

对问题 3), 根据容斥原理有

$$P(\bar{A}\cup B)=P(\bar{A})+P(B)-P(\bar{A}B)=1-p+q-(q-r)=1-p+r.$$

对问题 4), 根据德摩根律与容斥原理有

$$P(\bar{A}\cap\bar{B})=P(\overline{A\cup B})=1-P(A\cup B)=1+r-p-q.$$

例 1.14 一个人在做事时可能犯 n 种不同的错误，每种错误发生的概率为 p, 求此人不犯错的概率至少是多少.

解 用 A_i 表示此人犯第 i 类错误的事件, 根据题意可知 $P(A_i)=p$. 此人不犯错的事件可以表示为 $\bar{A}_1\cap\bar{A}_2\cap\cdots\cap\bar{A}_n$, 根据 Union bounds 有

$$\begin{aligned}P(\bar{A}_1\cap\bar{A}_2\cap\cdots\cap\bar{A}_n) &= 1-P(A_1\cup A_2\cup\cdots\cup A_n)\\ &\geqslant 1-P(A_1)-P(A_2)-\cdots-P(A_n)=1-np,\end{aligned}$$

若这 n 个事件互不相容, 则上述不等号中的等号成立.

例 1.15 (匹配问题) 有 n 对夫妻参加一次聚会, 现将所有参会人员任意分成 n 组, 每组一男一女, 求至少有一对夫妻被分到同一组的概率.

解 用 A 表示至少有一对夫妻被分到同一组的事件, 以及 A_i 表示第 i 对夫妻 $(i\in[n])$ 被分到同一组的事件, 于是有 $A=A_1\cup A_2\cup\cdots\cup A_n$. 根据容斥原理有

$$P(A)=P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{r=1}^{n}(-1)^{r+1}\sum_{i_1<\cdots<i_r}P(A_{i_1}\cdots A_{i_r}).$$

对任意 $r\in[n]$, 考虑事件 $A_{i_1}\cdots A_{i_r}$ 概率, 若参会人员任意分成 n 组且每组一男一女, 共有 $n!$ 种不同的分法, 若将第 $i_1,i_2,\cdots,i_r$ 对夫妻分别分组, 则有 $(n-r)!$ 种不同的分法. 根据等可能性原则有

$$P(A_{i_1}\cdots A_{i_r})=\frac{(n-r)!}{n!}.$$

而和式 $\sum_{i_1<\cdots<i_r}P(A_{i_1}\cdots A_{i_r})$ 中共有 $\binom{n}{r}$ 项, 由此可得

$$\sum_{i_1<\cdots<i_r}P(A_{i_1}\cdots A_{i_r})=\binom{n}{r}\frac{(n-r)!}{n!}=\frac{1}{r!},$$

于是事件 A 发生的概率

$$P(A)=1-\frac{1}{2!}+\frac{1}{3!}+\cdots+(-1)^{n+1}\frac{1}{n!}.$$

当 n 较大时, 利用泰勒展式 $e^x = 1 + x + x^2/2! + \cdots + x^n/n! + \cdots$ 以及令 $x = -1$ 有

$$e^{-1} = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} + \cdots \approx \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!},$$

由此近似有 $P(A) = 1 - 1/e = 0.632$.

在概率计算的过程中, 有时可适当利用概率的性质来简化计算, 例如,

例 1.16 从 $\{1, 2, \cdots, 9\}$ 数中有放回取 n 个, 试求取出 n 个数的乘积被 10 整除的概率.

解 用事件 A 表示取出 n 个整数的乘积能被 10 整除, 用事件 B 表示取出的 n 个数中有偶数, 用事件 C 表示取出的 n 个数中至少有一个 5, 于是有 $A = BC$. 直接计算事件 B 发生的概率较难, 可以考虑 B 的对立事件的概率

$$P(\bar{B}) = P\big(\text{取出的 } n \text{ 个数中无偶数}\big) = P\big(\{\text{取出的 } n \text{ 个数只包括} 1, 3, 5, 7, 9\}\big) = 5^n/9^n.$$

同理可得

$$P(\bar{C}) = 8^n/9^n \qquad \text{和} \qquad P(\bar{B}\bar{C}) = 4^n/9^n.$$

根据概率的性质有

$$P(A) = 1 - P(\overline{BC}) = 1 - P(\bar{B} \cup \bar{C}) = 1 - P(\bar{B}) - P(\bar{C}) + P(\bar{B}\bar{C}) = 1 - \frac{5^n}{9^n} - \frac{8^n}{9^n} + \frac{4^n}{9^n}.$$

1.5 组合计数*

组合计数研究满足一定条件下计数对象的个数, 概率论中的很多问题都可以通过计算一个事件发生的数目来解决, 如古典概型, 此外组合计数在人工智能、计算机等领域具有广泛的应用, 本节将介绍经典的组合计数: 十二重计数 (The twelvefold way).

十二重计数由著名的组合学家 G.-C. Rota (1932-1999) 提出, 最初的问题表述为研究一定条件下两个集合之间映射的个数. 为了问题的可理解性, 这里采用 D. Knuth 在《计算机程序设计艺术》中的表述: 将 n 个不同或相同的球, 放入 m 个不同或相同的箱子, 在无任何限制、每个箱子至多放一个球、每个箱子至少放一个球这三种条件下, 研究十二种情形中分别有多少种不同的放法. 表 1.3 给出了十二重计数的结果, 相关知识和符号说明将在后续逐一介绍.

1.5.1 排列、环排列、组合与多重组合

在前面 1.2 节中介绍了排列, 即从 n 个不同的元素中取出 r 个元素进行排列, 不仅要考虑取出的元素, 还要考虑他们之间的排列顺序, 共有 $(n)_r = n(n-1)\cdots(n-r+1)$ 种不同的排列方法. 若 $r = n$ 时称全排列, 有 $n!$ 种方法.