

第 1 章 随机事件及概率计算

中学已经学过本章的大部分内容，本章对该部分进行复习，以便于更好地对接本课程其他内容。同时，为将来工作中的数学语言交流，建议注意事件的关系运算以及在求解相关概率计算时的事件设定和运算表示，以便于更深入地理解概率及其计算的理论依据。

学习目标与要求

1. 理解随机现象及统计规律
2. 理解样本空间的概念，理解随机事件概率的定义
3. 掌握事件的关系运算，掌握概率的加法公式
4. 掌握条件概率与乘法公式，掌握全概率公式与贝叶斯公式
5. 理解事件独立的概念，掌握随机事件的独立性
6. 理解伯努利概型，掌握伯努利概型的概率计算

1.1 样本空间与随机事件

1.1.1 样本空间

1. 随机现象及统计规律

自然界和人类社会实践中所发生的现象多种多样，有一类现象是**确定现象**，它在一定条件下必然出现，如在一个标准大气压下将水加热到 100°C 便会沸腾、水往低处流、向上抛一个石子必然下落等。另一类现象是**随机现象**，随机现象情况下，即使完全相同的条件也会出现不同结果，换句话说，条件不能完全决定结果。如重复抛掷一枚硬币，其结果可能是字面向上也可能是花面向上，只有抛掷完成后才能观察到真正的结果；同一工艺条件下生产出的灯泡，其寿命长短参差不齐，一直到用坏才能知道一只灯泡的使用寿命；一批新产品投放市场，可能畅销也可能滞销，到底是畅销还是滞销，也只有经过市场检验后才知道答案；射击运动员用同一支枪向同一目标射击，无论怎样控制设计条件不变，每次弹着点总不尽相同等。

随机现象在一次次的观察或试验中，其各次的结果都呈现出不确定性，但如果在相同条件下进行大量地重复观察或试验，其结果却往往会呈现出某种明显的规律性。例如，多次重复地抛掷同一枚均匀硬币，尽管每次抛掷试验之前都不知道其字面朝上还是花面朝上，但随着重复抛掷次数的增加，出现字面朝上的次数，会随着投掷总次数的增加逐渐趋于总次数的 $\frac{1}{2}$ 。我们把这种在相同条件下进行大量重复试验，随机现象所呈现的规律性称为随机现象的**统计规律性**。一种观点认为，概率论与数理统计就是研究和揭示随机现象所具有的这种统计规律性的一门数学学科。而这一研究则往往从随机试验开始。

2. 随机试验

广义而言，对事物某一特征的观察、检测、测量等都称为试验。而研究随机现象的试验称为随机试验，具体定义如下。

定义 1.1 随机试验

具有如下三个特征的试验叫做**随机试验** (random experiment):

- (1) 可以在同一条件下重复进行;
- (2) 每次试验的可能结果不止一个, 并且在试验之前能明确知道试验的所有可能结果;
- (3) 进行每一次试验之前无法预先确定哪一个结果会出现.

以下是一些随机试验的例子.

例 1.1 E_1 : 抛掷一枚硬币, 观察花面 H 、字面 T 出现的情况.

E_2 : 抛掷一颗骰子, 观察出现的点数.

E_3 : 从一副无王牌的扑克中任抽一张, 记录牌点.

E_4 : 在大街上任找一人, 询问其当前的月收入.

E_5 : 从一班名字中任选一个, 观察是男生还是女生 (或观察其出生年月).

随机试验是研究随机现象的基本手段, 而随机试验中所出现的所有可能结果则是研究随机现象的突破口.

3. 样本空间

通过随机试验来研究随机现象, 尽管在每次试验之前不能预知试验的结果, 但试验的所有可能结果组成的集合是已知的, 而这一集合就是样本空间.

定义 1.2 样本空间、样本点

随机试验 E 的所有可能结果组成的集合称为**样本空间**, 记为 Ω . 样本空间的元素, 即 E 的每个可能的结果称为**样本点**, 记为 ω .

如, 对于随机试验“投掷一枚骰子, 观察点数”, 其样本点为 1、2、3、4、5、6, 样本空间为 $\{1, 2, 3, 4, 5, 6\}$, 如图 1.1 所示; 而对于随机试验“抛一枚硬币, 观察其字面及花面”, 若记 H = “花面”, T = “字面”, 则其样本点为 H, T , 样本空间为 $\{H, T\}$; 对于随机试验“任取一个灯泡, 测其使用寿命”, 其样本空间为 $\{t | t \geq 0\}$, 而样本点则为任意一个大于等于零的实数.



图 1.1 骰子的点数

通常, 不同的随机试验对应不同的样本空间, 而样本空间则构成了研究随机现象的基本数学模型. 应当注意的是, 即使相同的随机试验, 试验目的不同, 对应的样本空间也不同. 如, “投掷一枚骰子”, 若目的是考察点数, 则样本空间为 $\{1, 2, 3, 4, 5, 6\}$; 若考察大小点, 则其样本空间为 {大点, 小点}. 因此, 以下所说的随机试验往往包含相应的试验目的.

1.1.2 随机事件的概念

为对试验结果作更细致的研究, 引入以下定义.

定义 1.3 随机事件

随机试验所形成样本空间的子集称为**随机事件**, 简称事件. 通常用大写英文字母 A, B, C 等表示.

随机事件也可以认为是对随机试验结果的一种预言或概括, 它可能发生也可能不发生.

例 1.2 对于随机试验“投掷一枚骰子, 观察点数”, $A = \{\text{大于 3 点}\}$, $B = \{\text{小于 3 点}\}$, $C = \{\text{等于 6 点}\}$, $D = \{\text{大于 6 点}\}$, $\Omega = \{\text{小于等于 6 点}\}$ 均为随机事件, 因为它们都是这一随机试验对应样本空间 $\{1, 2, 3, 4, 5, 6\}$ 的子集, 事实上, $A = \{4, 5, 6\}$, $B = \{1, 2\}$, $C = \{6\}$,

$D = \emptyset$, $\Omega = \{1, 2, 3, 4, 5, 6\}$. 注意, 这些事件的每一个也同样都是这里的随机试验结果的一种预言.

另外值得注意的是, 有的事件简单, 有的事件复杂. 例如从上述例子的随机事件 $A = \{4, 5, 6\}$, $B = \{1, 2\}$, $C = \{6\}$ 中可以看出, 事件 A 由随机试验的三个可能结果 4、5、6 所构成, 事件 B 由随机试验的两个可能结果 1、2 构成, 它们均由随机试验的多个可能结果联合构成, 而 C 只包含随机试验的一个可能结果 6, 它不能再分解, 且多个这样的事件可以构成比较复杂的事件, 如 A, B . 为细分它们, 我们定义如下基本事件和复合事件.

定义 1.4 基本事件、复合事件

只包含一种随机试验可能结果 (或一个样本点构成) 的事件, 称为**基本事件**. 由若干基本事件组合而构成 (或多个样本点构成) 的事件称为**复合事件**.

基本事件是可以发生但不可能再分解的随机事件.

最后, 上例 1.2 中的事件 $D = \emptyset$, $\Omega = \{1, 2, 3, 4, 5, 6\}$ 是样本空间的两个非常特殊的子集, 在此约定为两个特殊事件, 定义如下.

定义 1.5 必然事件、不可能事件

在一定条件下必然发生的事件称为**必然事件**. 在一定条件下必然不能发生的事件称为**不可能事件**. 必然事件由样本空间的所有样本点构成, 为样本空间, 常用 Ω 表示, 不可能事件为空集, 常用 $\emptyset$ 表示.

1.1.3 随机事件的关系及运算

随机事件是样本空间的子集, 首先是集合, 因此, 集合的相关运算及运算法则、集合的文氏图 (Vn) 表示等都适用于这里的随机事件.

1. 包含关系

若事件 B 发生必然导致事件 A 发生, 则称事件 A 包含事件 B (或称事件 B 包含于事件 A), 记为 $A \supset B$ 或 $B \subset A$. 如图 1.2 所示.

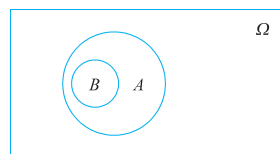


图 1.2 包含关系

2. 相等关系

若 $B \supset A$ 且 $A \supset B$, 即事件 A 发生必导致事件 B 发生, 同时事件 B 发生同样也必导致事件 A 发生, 此时, 则称事件 A 与 B 相等, 记为 $A = B$.

可见, 当且仅当事件 A 与事件 B 完全相同, 或者说它们包含完全一样的基本事件时, $A = B$.

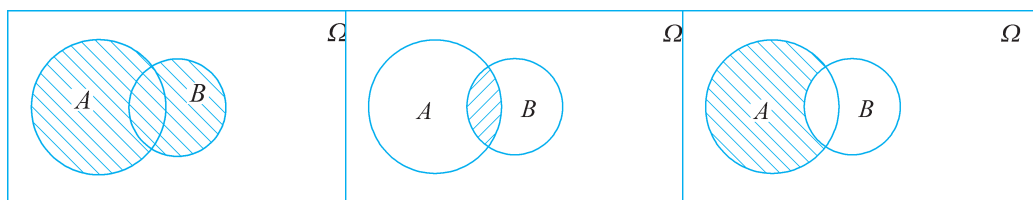
3. 事件的和 (并)

A 与 B 至少有一个发生的事件, 即“事件 A 或事件 B 发生”的事件, 称为事件 A 与事件 B 的和 (并), 记为 $A \cup B$. 如图 1.3 所示.

有限多个事件 $A_1, A_2, \dots, A_n$ 的并记为 $\bigcup_{i=1}^n A_i$; 可列无穷多个事件 $A_1, A_2, \dots, A_n, \dots$ 的并记为 $\bigcup_{i=1}^{\infty} A_i$.

4. 事件的积 (交)

A 与 B 同时发生的事件, 称为事件 A 与 B 的积 (交), 记为 $A \cap B$. 如图 1.4 所示.

图 1.3 并 $A \cup B$ 图 1.4 交 $A \cap B$ 图 1.5 差 $A - B$

有限多个事件 $A_1, A_2, \dots, A_n$ 的积记为 $\bigcap_{i=1}^n A_i$; 可列无穷多个事件 $A_1, A_2, \dots, A_n, \dots$

的积记为 $\bigcap_{i=1}^{\infty} A_i$.

5. 事件的差

A 发生 B 不发生的事件, 称为事件 A 与事件 B 的差, 记为 $A - B$. 如图1.5 所示.

一般地, $A - B \neq B - A$.

6. 互不相容事件 (互斥事件)

若事件 A 与 B 不能同时发生, 即 $AB = \emptyset$, 则称事件 A 与 B 互不相容. 互不相容也称为互斥. 如图1.6 所示.

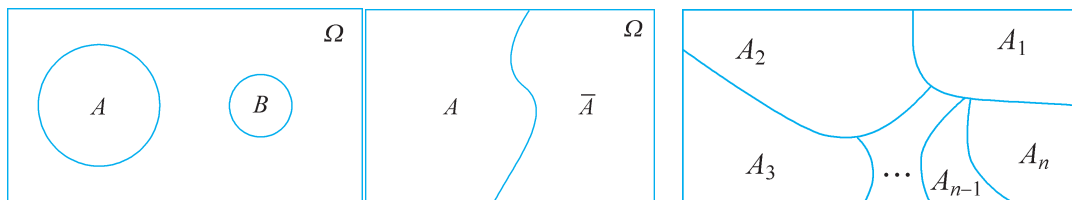


图 1.6 互斥 (不相容)

图 1.7 对立 (互逆)

图 1.8 样本空间的划分

7. 对立事件 (逆事件)

若事件 A 与 B 不能同时发生, 且至少一个必然发生, 即 $AB = \emptyset$, 且 $A \cup B = \Omega$, 则称事件 A 与事件 B 互为对立事件 (或逆事件), 记为 $B = \bar{A}$, 或 $A = \bar{B}$. 如图1.7所示.

8. 完备事件组 (样本空间的划分)

若事件 $A_1, A_2, \dots, A_n$ 两两互不相容, 且 $A_i A_j = \emptyset (i \neq j)$, $\sum_{i=1}^n A_i = \Omega$, 则称事件 $A_1, A_2, \dots, A_n$ 构成一个完备事件组. 如图1.8 所示.

例 1.3 在掷骰子试验中, 目的是考察朝上面的点数, 则有样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$. 记事件 $A = \{\text{点数为 } 2\}$, $B = \{\text{点数为偶数}\}$, $C = \{\text{点数为小于 } 3 \text{ 的偶数}\}$, $D = \{\text{点数为 } 1\}$, $E = \{\text{点数小于 } 4\}$, $F = \{\text{点数为奇数}\}$, 则

$$A = \{2\}, B = \{2, 4, 6\}, C = \{2\}, D = \{1\}, E = \{1, 2, 3\}, F = \{1, 3, 5\},$$

因此, 有如下关系或运算结果:

- (1) $B \supset A$ 或 $A \subset B$;
- (2) $A = C$;
- (3) $A \cup D = \{\text{点数小于 } 3\} = \{1, 2\}$;
- (4) $A \cap B = \{\text{点数为 } 2\} = \{2\}$;
- (5) $B - E = \{\text{点数为 } 4 \text{ 或 } 6\} = \{4, 6\}$, $E - B = \{\text{点数为 } 1 \text{ 或 } 3\} = \{1, 3\}$;
- (6) 事件 B 与 D 为互不相容事件;
- (7) B 与 F 互为对立事件, 它们满足 $B = \bar{F}$, 或 $F = \bar{B}$. 亦即 $B \cap F = \emptyset$ 且 $B \cup F = \Omega$;

(8) 容易看出, 三事件 $E - B$, $B - E$ 与 $\{2, 5\}$ 构成样本空间 Ω 的一个划分. 以下以定理形式, 给出事件所满足的算律.

定理 1.1 设 A, B, C 为事件, 则有以下运算法则成立.

- (1) 交换律: $A \cup B = B \cup A$, $A \cap B = B \cap A$;
- (2) 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$;
- (3) 分配律: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;
- (4) 德摩根律: $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

例 1.4 现从含有 5 件次品的 100 件产品中任取 3 件进行检验. 记 $A_i = \{3 \text{ 件中恰有 } i \text{ 件次品}\} (i = 1, 2, 3)$, 试用 $A_i (i = 1, 2, 3)$ 或其对立事件表示下列事件: $A = \{3 \text{ 件中至少有一件次品}\}$, $B = \{3 \text{ 件中至少有两件次品}\}$, $C = \{3 \text{ 件都是合格品}\}$.

解 由题意分析, 事件 A 等同于三种情况的并: 3 件中有一件次品、3 件中有两件次品、3 件中有 3 件次品. 因此有

$$A = A_1 \cup A_2 \cup A_3,$$

同样, “3 件中至少有两件次品”是指“3 件中有两件次品或有 3 件次品”, 因此

$$B = A_2 \cup A_3,$$

最后, “3 件都是合格品”则表示“3 件都不是次品”, 即

$$C = \overline{A_1 A_2 A_3} = \overline{A_1 \cup A_2 \cup A_3} = \overline{A}.$$

例 1.5 设 A, B, C 是三个事件, 请以 A, B, C 的运算表示以下各事件.

- (1) A 发生, B 与 C 都不发生;
- (2) A, B, C 至少有一个发生;
- (3) A, B, C 恰有一个发生;
- (4) A, B, C 至多一个发生;
- (5) A, B, C 不多于两个发生.

解 以 $D_i (i = 1, 2, 3, 4, 5)$ 表示要求的事件. 注意到事件 A 不发生就是事件 $\overline{A}$ 发生.

(1) 事件 B 与 C 都不发生, 就是事件 $\overline{B}$ 发生, 事件 $\overline{C}$ 也发生, 因此本题事件就是“事件 $A, \overline{B}, \overline{C}$ ”都发生, 因此有 $D_1 = A \cap \overline{B} \cap \overline{C}$.

(2) A, B, C 至少有一个发生, 就是事件 $A \cup B \cup C$, 因此 $D_2 = A \cup B \cup C$.

(3) A, B, C 恰有一个发生, 就是“ A 发生, B, C 都不发生, 或 B 发生, A, C 都不发生, 或 C 发生, B, A 都不发生”, 因此有 $D_3 = (A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$.

(4) A, B, C 至多一个发生, 就是“ A, B, C 恰有一个发生”, 或“ A, B, C 都不发生”, 因此有 $D_4 = D_3 \cup (\overline{A} \cap \overline{B} \cap \overline{C})$.

(5) A, B, C 不多于两个发生, 就是“ A, B, C 中至少有一个不发生”, 或“ $\overline{A}, \overline{B}, \overline{C}$ 中至少有一个发生”, 因此 $D_5 = \overline{A} \cup \overline{B} \cup \overline{C}$ 或 $D_5 = \overline{A \cap B \cap C}$.

1.2 随机事件的概率

人们研究随机现象, 不仅要知道可能出现哪些事件, 更希望知道事件发生的可能性大小, 这就需要一个刻画事件发生可能性大小的度量指标, 这种度量应能反映随机现象所呈现的统计

规律性, 我们称这个刻画事件发生可能性大小的度量指标为事件发生的概率. 为此, 我们从最早出现的古典概型讲起.

1.2.1 概率的定义及计算

1. 古典概型

古典概型中的“古典”表明了这种问题起源的古老, 它源于赌博. 博弈的形式多种多样, 但是它们的前提是“公平”, 即“机会均等”, 而这正是古典定义适用的重要条件: 同等可能. 16 世纪意大利数学家和赌博家卡尔丹 (1501—1576) 所说的“诚实的骰子”, 即道明了这一点. 在卡尔丹以后约三百年的时间里, 帕斯卡、费马、伯努利等数学家都在古典概率的计算、公式推导和扩大应用等方面做了重要的工作. 直到 1812 年, 法国数学家拉普拉斯 (1749—1827) 在《概率的分析理论》中给出的概率古典定义, 成为最早的概率计算问题及概率计算方法. 以下是整理后的古典概型定义.

定义 1.6 古典概型

若随机试验 E 具有如下特征

- (1) 有限性: 试验的样本空间只包括有限多个样本点, 即 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$;
- (2) 等可能: 试验中每个基本事件发生的可能性相等.

则称 E 为古典概型 (也称等可能概型).

设古典概型 E 的样本空间为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, 即样本点总数为 n . 根据等可能性原则, 可知每个基本事件发生的概率均为 $\frac{1}{n}$. 给定事件 $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\}$, 其包含 k 个样本点, 则事件 A 的概率为

$$P(A) = k \cdot \frac{1}{n} = \frac{\text{组成} A \text{的基本事件个数}}{\text{样本空间的基本事件总数}}. \quad (1.1)$$

实际问题 1.1 掷骰子中的概率

依次抛掷两粒骰子, 观察它们出现的点数.

- (1) 写出这一随机试验的样本空间;
- (2) 求事件 $A = \text{“两粒骰子点数相同”}$ 与 $B = \text{“两粒骰子点数之和为 4”}$ 的概率.

解 (1) 记 $(i, j) (i, j = 1, 2, \dots, 6)$ 表示第一粒骰子投 i 点, 第二粒骰子投 j 点, 则样本空间为

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)\},$$

其中共计 36 个基本事件.

(2) 事件“两粒骰子点数相同”出现有 6 种可能, 即 $A = \{(1, 1), (2, 2), \dots, (6, 6)\}$; 事件“两粒骰子点数之和为 4”也有 3 种可能, 即 $B = \{(1, 3), (3, 1), (2, 2)\}$. 问题是古典概型问题, 因此

$$P(A) = \frac{6}{36} = \frac{1}{6}, \quad P(B) = \frac{3}{36} = \frac{1}{12}.$$

实际问题 1.2 抽取产品的概率

一批产品共有 42 件, 其中有一等品 15 件, 二等品 14 件, 三等品 13 件, 今从这批产品中一次任取 3 件 (取后不放回). 求这 3 件产品分属下列情形的概率. (1) $A = \{3 \text{ 件产品都是一等品}\}$; (2) $B = \{3 \text{ 件产品等级相同}\}$; (3) $C = \{3 \text{ 件产品等级不相同}\}$.

解 (1) 由于 A 事件中 3 件产品都是一等品, 故事件 A 的基本事件数为 $M_A = C_{15}^3$, 故

$$P(A) = \frac{C_{15}^3}{C_{42}^3}.$$

(2) 由于 B 事件中 3 件产品等级相同, 故事件 B 的基本事件数为 $M_B = C_{15}^3 + C_{14}^3 + C_{13}^3$, 故

$$P(B) = \frac{C_{15}^3 + C_{14}^3 + C_{13}^3}{C_{42}^3}.$$

(3) 由于 C 事件中 3 件产品等级不相同, 故事件 C 的基本事件数为 $M_C = C_{15}^1 C_{14}^1 C_{13}^1$, 故

$$P(C) = \frac{C_{15}^1 C_{14}^1 C_{13}^1}{C_{42}^3}.$$

实际问题 1.3 接待时间有无规定

某接待站在某一周内曾接待过 12 次来访, 已知所有这 12 次接待都是在周二和周四进行的, 问是否可以推断接待时间是有规定的.

解 假设接待站的接待时间没有规定, 且各来访者在一周的每一天中去接待站都是等可能的. 则一周内 12 次来访的可能方式有 7^{12} 种; 而 12 次接待均在周二和周四进行的有 2^{12} 种, 因此, 记 A 表示事件“12 次接待都是在周二和周四进行”, 则其概率为

$$P(A) = \frac{2^{12}}{7^{12}} \approx 0.000\ 000\ 295\ 9.$$

而小概率事件在实际中几乎是不可能发生的, 从而可知接待时间是有规定的.

实际问题 1.4 超几何分布的概率

一批产品共有 N 件, 其中有 M 件次品, 今从这批产品中任取 n 件. 求这 n 件产品中恰好有 k 件次品的概率.

解 设 $A = \{\text{任取的 } n \text{ 件产品中恰有 } k \text{ 件次品}\}$, 易知样本空间样本点数为 C_N^n , 对于事件 A , 这 n 件产品中的 k 件次品只能在 M 件次品中抽取, 有 C_M^k 种取法, 剩下的 $n-k$ 件正品只能在 $N-M$ 件正品中抽取, 有 C_{N-M}^{n-k} 种取法, 故事件 A 包含的样本点数为 $C_M^k C_{N-M}^{n-k}$, 于是

$$P(A) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \quad (k = 0, 1, 2, \dots, \min\{M, n\}). \quad (1.2)$$

式(1.2)称为超几何分布的概率公式.

实际问题 1.5 有(无)放回抽样的概率

一袋中共有 20 只小球, 其中有 15 只红球 5 只黑球, 今从中任取 3 只, 每次取出一只, 记事件 A 为“取出的 3 只球恰有 2 只是黑球”, 试在下列抽取方式下分别求概率 $P(A)$. (1) 每取一只看颜色后放回, 再继续抽取下一只 (有放回抽样); (2) 每取一只不再放回, 再继续抽取下一只 (不放回抽样).

解 (1) 有放回抽样条件下, 基本事件总数与重复排列有关, 即 $n = 20^3$, 事件 A 包含的基本事件数为 $m = C_3^2 5^2 \cdot 15$, 故所求的概率为

$$P(A) = \frac{C_3^2 5^2 \cdot 15}{20^3} = \frac{9}{64}.$$

(2) 不放回抽样条件下, 基本事件总数与排列有关, 即 $n = A_{20}^3$, 事件 A 包含的基本事件

数 $m = C_3^2 A_5^2 A_{15}^1$, 故所求的概率为

$$P(A) = \frac{C_3^2 A_5^2 A_{15}^1}{A_{20}^3} = \frac{5}{38}.$$

注: 上述小题 (2) 可用组合计算. 基本事件总数为 C_{20}^3 , 事件 A 包含的基本事件数 $m = C_5^2 C_{15}^1$, 故所求的概率为

$$P(A) = \frac{C_5^2 C_{15}^1}{C_{20}^3} = \frac{5}{38}.$$

可见, 古典概型的计算有时排列组合均可用, 但在整个问题求解中要么使用排列, 要么使用组合, 统一即可.

古典概型的概率具有如下性质.

定理 1.2 古典概型的概率性质.

(1) 非负性: 对任意事件 A , 有 $0 \leq P(A) \leq 1$;

(2) 规范性: $P(\Omega) = 1$;

(3) 有限可加性: 若 $A_1, A_2, \dots, A_n$ 互不相容, 则 $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.

古典概型解决了许多实际问题中的概率计算, 但古典概型的概率计算能应用的范围仍然很窄. 这促使雅各布·伯努利 (1654—1705) “寻找另一条途径找到所期待的结果”, 伯努利大数定律就是他在研究古典概率时的另一重要成果. 在古典概型下, 伯努利证实了“当试验次数愈来愈大时, 频率接近概率”. 事实上, 这不仅对于古典概型适用, 人们确信“从现实中观察的频率稳定性”的事实是一个普遍规律, 所以可以“用频率估计概率”.

2. 由频率到概率

先引入事件频率的概念.

定义 1.7 事件的频率

若在同一条件下将试验 E 重复 N 次, 事件 A 发生的次数 m (频数) 与试验次数 N 的比值 $\frac{m}{N}$, 称为事件 A 的频率, 记为 $f_N(A)$, 即

$$f_N(A) = \frac{m}{N}. \quad (1.3)$$

容易验证, 事件的频率 $f_N(A)$ 具有如下性质.

定理 1.3 事件频率的性质.

(1) 非负性: 对任意事件 A , 有 $0 \leq f_N(A) \leq 1$;

(2) 规范性: $f_N(\Omega) = 1$;

(3) 有限可加性: 若 $A_1, A_2, \dots, A_n$ 互不相容, 则 $f_N(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n f_N(A_i)$.

频率在一定程度上代表了概率, 但毕竟它不是概率. 事实上, 当试验次数 N 无限增大时, 事件 A 发生频率的稳定值就是事件 A 的概率.

实际问题 1.6 历史上曾经做过的投币试验

历史上浦丰等人做过的投币试验见表 1.1. 随着试验次数的增加, 发现频率具有稳定性: 当 N 很大时, 频率 $f_N(A)$ 在某个常数附近摆动, 而随着试验次数 N 的增大, 摆动的幅度越来越小. 这个常数就是 A 发生的概率.

表 1.1 历史上投掷硬币试验

人 名	投掷次数 N	出现正面次数 m	频率 $f_N(A)$
德摩根	2 048	1 061	0.518 1
浦 丰	4 040	2 048	0.506 9
费 勒	10 000	4 979	0.497 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

这一试验也充分表明：可以用频率估算概率. 尽管这样不能像古典定义那样确切地算出概率，但是却给出了一个估计概率的方法. 而且，它不再需要“等可能”的条件，因此从应用的角度来讲，它的适用范围更广.

但是从数学理论上讲却是有问题的：在古典概率的场合，事件概率不依赖于事件频率定义——它根本不用诉诸试验，因此必然有一个频率与概率是否可以充分接近的问题，而这导致了伯努利大数定律. 另一方面，在统计定义的场合这却是一个悖论：你如果不从承认大数定律的角度出发，概率就无法定义，因而谈不上频率与概率接近的问题；但是你如果承认大数定律，以便可以定义概率，那大数定律就是你的前提，从而不再是一个需要证明的论断了.

3. 随机事件概率的公理化定义

事件的频率虽然克服了古典概型概率计算的局限性，适合任何随机试验. 但它毕竟还是一种直观的、描述性的估算方法，不具有数学的严谨性与精确性.

古典概型的概率、事件发生的频率都具有“非负性、规范性、可列可加性”三条基本性质，它们却反映了概率的本质特征. 因此，可以用它们作为公理来定义概率.

定义 1.8 随机事件的概率

在一个随机试验中，用来表示一个随机事件 A 发生可能性大小的实数称为事件 A 的概率，记为 $P(A)$ ，并规定：

- (1) 非负性公理：对于任一事件 A ，有 $P(A) \geq 0$ ；
- (2) 规范性公理：必然事件 Ω 的概率 $P(\Omega) = 1$ ；
- (3) 可列可加性公理：对于互不相容的可列个事件 $A_1, A_2, \dots, A_n, \dots$ ，有

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

1900 年，38 岁的希尔伯特（1862—1943）在世界数学家大会上提出了建立概率公理系统的问题，这就是著名的希尔伯特 23 个问题中的第 6 个问题. 1933 年，苏联数学家柯尔莫哥洛夫（1903—1987）在《概率论基础》中，借助实变函数论和测度论来定义概率，形成了概率论的公理化体系，这一公理体系既概括了古典定义、统计定义的基本特性，又避免了各自的局限. 至此，概率论才真正成为了严格的数学分支.

1.2.2 概率的性质

由随机事件概率的定义，可得概率的如下重要性质.

定理 1.4 不可能事件的概率为零，即 $P(\emptyset) = 0$.

证明 令 $A_n = \emptyset$ ($n = 1, 2, \dots$)，则 $\bigcup_{n=1}^{\infty} A_n = \emptyset$ ，且 $A_i A_j = \emptyset$ ($i \neq j, i, j = 1, 2, \dots$) .

由概率的可列可加性得

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

由概率的非负性, $P(\emptyset) \geq 0$, 故由上式可知 $P(\emptyset) = 0$.

定理 1.5 (概率有限可加性) 若事件 $A_1, A_2, \dots, A_n$ 两两互不相容, 则有

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n). \quad (1.4)$$

证明 令 $A_{n+1} = A_{n+2} = \dots = \emptyset$, 即有 $A_i A_j = \emptyset (i, j = 1, 2, \dots)$. 由可列可加性公理得

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \\ &= \sum_{k=1}^n P(A_k) + 0 = P(A_1) + P(A_2) + \dots + P(A_n). \end{aligned}$$

定理 1.6 (概率的单调性) 设 A, B 是任意两个事件, 若 $A \subset B$, 则有

$$P(B - A) = P(B) - P(A), \quad (1.5)$$

$$P(A) \leq P(B). \quad (1.6)$$

证明 由 $A \subset B$ 知 $B = A \cup (B - A)$, 且 $A(B - A) = \emptyset$, 再由概率的有限可加性得

$$P(B) = P(A) + P(B - A),$$

于是, 式(1.5)得证.

又由概率的非负性得 $P(B - A) \geq 0$, 知

$$P(A) \leq P(B).$$

定理 1.7 对任意事件 A , 有 $P(A) \leq 1$, 从而也就有 $0 \leq P(A) \leq 1$.

证明 因为 $A \subset \Omega$, 由定理1.6 得

$$P(A) \leq P(\Omega) = 1.$$

定理 1.8 (对立事件的概率) 对任意事件 A , 有 $P(\bar{A}) = 1 - P(A)$.

证明 因 $A \cup \bar{A} = \Omega$, 且 $A\bar{A} = \emptyset$, 由式(1.4) 得

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}).$$

实际问题 1.7 至少取到一件次品的概率

40 件产品中有 8 件次品, 32 件正品, 现从中任取 3 件, 求至少有 1 件次品的概率.

解 设事件 $A = \{\text{取出的3件中至少有1件是次品}\}$, 则 A 中包含 3 种情况: 取到 1 件次品 2 件正品、或取到 2 件次品 1 件正品、或取到 3 件次品, 而 A 的对立事件 $\bar{A}$ 只包含 1 种情况, 即取出的 3 件都是正品, 于是

$$P(\bar{A}) = \frac{C_{32}^3}{C_{40}^3} = \frac{124}{247}.$$