第一篇　概 率 部 分

第 1 章　随机事件及其概率

1.1　随机事件及其运算

1.1.1　随机现象

概率论与数理统计是研究随机现象统计规律的一门科学. 在科学研究中，随机现象的普遍存在性决定了它的广泛应用性.

在一定条件下，并不总是出现相同结果的现象称为**随机现象.** 例如，掷一枚硬币的结果、一天内进入某超市的顾客数.

从亚里士多德时代开始，哲学家们就已经认识到随机性在生活中的作用，直到 20 世纪初，人们才认识到随机现象亦可以通过数量化方法来进行研究. 概率论就是以数量化方法来研究随机现象及其规律性的一门数学课程，微积分等则是研究确定性现象的数学课程.

在现实世界中，存在只有一个结果的现象，称为**确定性现象或必然现象.** 例如，水在标准大气压下加温到 100℃就会沸腾；同性电荷相互排斥，异性电荷相互吸引；太阳从东方升起.

【例 1.1.1】　随机现象的例子.

（1）掷一颗骰子，出现的点数；

（2）抛一枚硬币，可能出现正面，也可能出现反面；

（3）某种型号电视机的寿命；

（4）一天内进入某商场的顾客数；

（5）某地区 10 月份的平均气温；

（6）一天内自己的手机的来电数.

思考问题：

随机现象具有什么特点？

1.1.2　随机试验

在相同条件下可以重复对随机现象进行的观察、记录、实验称为**随机试验**（Random Trial），**记为** E .

随机试验具有三个特点.

（1）**重复性：**可以在相同的条件下重复地进行；

（2）**明确性**：每次试验的可能结果不止一个，并且事先可以明确试验所有可能出现的结果；

（3）**随机性**：进行一次试验之前不能确定哪一个结果会出现.

【例 1.1.2】 随机试验的例子.

E_1：抛一枚硬币，观察正面 H 和反面 T 出现的情况；

E_2：掷两颗骰子，观察出现的点数之和；

E_3：在一批电视机中任意抽取一台，测试它的寿命；

E_4：某城市某一交通路口，指定一小时内的汽车流量；

E_5：考察某地区 10 月份的平均气温.

在现实生活中，也有很多随机现象是不能重复的. 例如，某场篮球赛的输赢是不能重复的；失业、经济增长速度等经济现象也是不能重复的. 概率论与数理统计主要研究能大量重复的随机现象，但是也研究不能重复的随机现象.

1.1.3 样本空间

随机试验的所有可能结果组成的集合称为**样本空间**（Sample space），记为 $\Omega=\{\omega\}$. 样本空间的元素 ω 称为**样本点**，表示基本结果. 样本点是统计中抽样的最基本单元. 写出随机试验的样本空间是认识随机现象的第一步.

【例 1.1.3】 下面给出随机现象的样本空间.

（1）掷一颗骰子的样本空间 $\Omega_1=\{\omega_1, \omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$，其中 ω_i（$i=1,2,3,4,5,6$）表示出现 i 点，样本空间也可直接写为 $\Omega_1=\{1,2,3,4,5,6\}$.

（2）抛一枚硬币的样本空间 $\Omega_2=\{H,T\}$，其中 H 表示正面，T 表示反面.

（3）某种型号电视机的寿命的样本空间 $\Omega_3=\{t\,|\,t\geqslant 0\}$.

（4）一天内进入某商场的顾客数的样本空间 $\Omega_4=\{0,1,2,\cdots,n,\cdots\}$.

（5）考察某地区 10 月份的平均气温的样本空间 $\Omega_5=\{t\,|\,T_1\leqslant t\leqslant T_2\}$，其中 t 表示平均气温.

（6）从一批含有正品和次品的产品中任意抽取一个产品的样本空间 $\Omega_6=\{N,D\}$，其中 N 表示正品，D 表示次品.

注意：

（1）样本空间中的元素可以是数，也可以不是数.

（2）根据样本空间含有样本点的个数来区分，样本空间可分为有限与无限两类. 例如，样本空间 Ω_1、Ω_2、Ω_6 中的样本点个数为有限个，而样本空间 Ω_3、Ω_4、Ω_5 中的样本点个数为无限个. 无限又分为可列与不可列，这里 Ω_4 中的样本点个数是可列个，Ω_3、Ω_5 中样本点个数是不可列个. 在今后的学习中，我们往往将样本点个数为有限个或可列个的情况归为一类，称为**离散样本空间**，而将样本点个数为不可列、无限个的情况归为另一类，称为**连续样本空间**. 这两类有本质的差别.

思考问题：

样本空间至少有多少个样本点？

1.1.4 随机事件

随机现象的某些样本点组成的集合，即样本空间 Ω 的子集称为**随机事件**（Random

Event），简称事件[①]，通常用大写字母 $A,B,C,\cdots$ 表示. 在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一事件发生. 例如，在掷一颗骰子中，$A=$“出现点数为偶数”是一个事件，若试验结果是“出现 2 点”，则称事件 A 发生.

注意：

（1）任何事件 A 对应样本空间的一个子集. 在概率论中常用一个长方形表示样本空间 Ω，用其中一个圆或其他几何图形表示事件 A，即事件 A 的维恩图，如图 1.1.1 所示.

（2）事件 A 发生是指当且仅当 A 中某个样本点出现了. 例如，当 $\omega_1(\in A)$ 出现时，则 A 发生；当 $\omega_2(\notin A)$ 出现时，则 A 不发生.

（3）事件可以用集合表示，也可用准确的语言描述.

（4）样本空间 Ω 中单个样本点组成的子集称为**基本事件**. 必然发生的事件称为**必然事件**，记为 Ω；不可能发生的事件称为**不可能事件**，记为 $\varnothing$.

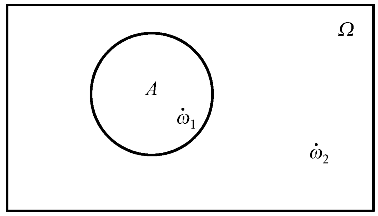

图 1.1.1　事件 A 的维恩图

思考问题：

样本空间的最大子集是什么事件？最小子集是什么事件？

【例 1.1.4】 抛一颗骰子的随机试验 E，样本空间记为 $\Omega=\{1,2,3,4,5,6\}$，请写出下列事件.

（1）事件 $A=$“出现 2 点”；

（2）事件 $B=$“出现点数小于 7”；

（3）事件 $C=$“出现点数大于 6”.

分析：事件 A 是 Ω 的单个样本点“2”组成的事件，是一个基本事件；事件 B 是由 Ω 的所有样本点组成的事件，是必然事件；事件 C 不可能发生，Ω 的任意样本点都不在事件 C 中，事件 C 是不可能事件.

答：（1）$A=\{2\}$；

（2）$B=\{1,2,3,4,5,6\}$；

（3）$C=\varnothing$.

1.1.5　事件间的关系

事件是一个集合，因而事件间的关系可以对比集合之间的关系来处理. 下面的讨论假设在同一样本空间 Ω 中进行，主要有以下几种关系.

（1）包含关系

集合语言：属于集合 A 的元素必属于集合 B.

概率论语言：若事件 A 发生必然导致事件 B 发生，则称事件 A 包含于事件 B（或称事件 B 包含事件 A）.

记为 $A\subset B$（或 $B\supset A$）.

对于任一事件 A，有 $\varnothing\subset A\subset\Omega$.

---

① 严格地说，事件是指 Ω 中满足某些条件的子集. 当 Ω 由有限个元素或无限可列个元素组成时，每个子集都可作为一个事件. 当 Ω 由不可列无限个元素组成时，某些子集必须排除在外. 幸而这种不可容许的子集在实际应用中几乎不会遇到. 今后，我们讲的事件均指它是容许考虑的那种子集.

（2）相等关系

集合语言：属于集合 A 的元素必属于集合 B，而且属于集合 B 的元素必属于集合 A.

概率论语言：若事件 A 发生必然导致事件 B 发生，且事件 B 发生必然导致事件 A 发生（$A\subset B$ 且 $B\supset A$），则称事件 A 与 B 相等（或等价）.

记为 $A=B$.

（3）互不相容

集合语言：若集合 A 与集合 B 没有相同的元素，则称集合 A 与集合 B 互不相容.

概率论语言：事件 A 与事件 B 不可能同时发生.

记为 $A\cap B=\varnothing$.

1.1.6 事件间的运算

事件的运算有并、交、差和对立四种，与集合的并、交、差和补集运算相当.

（1）事件 A 与事件 B 的并

集合语言："由集合 A 与 B 中所有的元素（相同的只计入一次）"组成的新集合.

概率论语言："事件 A 与 B 中至少有一个发生"的事件称为 A 与 B 的并（和）.

记为 $A\cup B$.

如：若 $A=\{1,2,3\}$， $B=\{1,2,5\}$，则 $A\cup B=\{1,2,3,5\}$.

（2）事件 A 与事件 B 的交（积）

集合语言："由事件 A 与 B 中公共的元素"组成的新集合.

概率论语言："事件 A 与 B 同时发生"的事件.

记为 $A\cap B$ 或 AB.

如：若 $A=\{1,2,3\}$， $B=\{1,2,5\}$，则 $A\cap B=\{1,2\}$.

（3）事件 A 与事件 B 的差

集合语言："由属于集合 A 而不属于集合 B 中的元素"组成的新集合.

概率论语言："事件 A 发生而 B 不发生".

记为 $A-B$.

如：若 $A=\{1,2,3\}$， $B=\{1,2,5\}$，则 $A-B=\{3\}$.

（4）对立事件

集合语言："由属于 Ω 而不属于集合 A 中的元素"组成的新集合称为集合 A 的对立事件.

概率论语言：A 不发生，记为 $\overline{A}$，即 $\overline{A}=\Omega-A$.

若 $A\cup B=\Omega$ 且 $A\cap B=\varnothing$，则称事件 A 与事件 B 互为逆事件（对立事件）. A 的对立事件记为 $\overline{A}$， $\overline{A}$ 是由所有不属于 A 的样本点组成的事件，它表示"A 不发生"事件. 显然 $\overline{A}=\Omega-A$.

注意：

1．对立事件是相互的，即 A 的对立事件是 $\overline{A}$， $\overline{A}$ 的对立事件是 A（$\overline{\overline{A}}=A$）.

2．必然事件 Ω 与不可能事件 $\varnothing$ 互为对立事件，即 $\overline{\Omega}=\varnothing$， $\overline{\varnothing}=\Omega$.

思考问题：

对立事件一定是互不相容的事件，互不相容的事件一定是对立事件吗？为什么？

可以验证一般事件的运算满足如下关系：

1．交换律 $A\cup B=B\cup A$， $A\cap B=B\cap A$ 或者 $AB=BA$；

2．结合律　$A\cup(B\cup C)=(A\cup B)\cup C$，

$A\cap(B\cap C)=(A\cap B)\cap C$ 或者 $A(BC)=(AB)C$；

3．分配律　$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$，

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

分配律可以推广到有穷或可列无穷的情形，即

$$A\cap\left(\bigcup_{i=1}^{n}A_i\right)=\bigcup_{i=1}^{n}(A\cap A_i)，\quad A\cup\left(\bigcap_{i=1}^{n}A_i\right)=\bigcap_{i=1}^{n}(A\cup A_i)；$$

$$A\cap\left(\bigcup_{i=1}^{\infty}A\right)_i=\bigcup_{i=1}^{\infty}(A\cap A_i)，\quad A\cup\left(\bigcap_{i=1}^{\infty}A\right)_i=\bigcap_{i=1}^{\infty}(A\cup A_i).$$

4．对偶律（德·摩根公式）

事件交的对立等于对立的并：$\overline{A\cap B}=\overline{A}\cup\overline{B}$

事件并的对立等于对立的交：$\overline{A\cup B}=\overline{A}\cap\overline{B}$

对有限个或可列无穷个 A_i，恒有

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

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

5．概率论与集合论之间的对应关系.

概率论与集合论之间的对应关系如表 1.1.1 所示.

表 1.1.1　概率论与集合论之间的对应关系

记　　号	概　率　论	集　合　论
Ω	样本空间，必然事件	全集
$\varnothing$	不可能事件	空集
ω	基本事件	元素
A	事件	子集
$\overline{A}$	集合 A 的对立事件	集合 A 的余集
$A\subset B$	事件 A 发生导致 B 发生	集合 A 是集合 B 的子集
$A=B$	事件 A 与事件 B 相等	集合 A 与集合 B 的相等
$A\cup B$	事件 A 与事件 B 至少有一个发生	集合 A 与集合 B 的和集
AB	事件 A 与事件 B 同时发生	集合 A 与集合 B 的交集
$A-B$	事件 A 发生而事件 B 不发生	集合 A 与集合 B 的差集
$AB=\varnothing$	事件 A 和事件 B 互不相容	集合 A 与集合 B 没有相同的元素

以上事件的关系及运算可以用维恩图来直观地描述. 若用平面上一个矩形表示样本空间 Ω，圆 A 与圆 B 分别表示事件 A 与事件 B，则 A 与 B 的各种关系及其运算如图 1.1.2～图 1.1.7 所示.

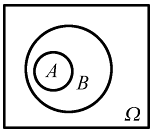

图 1.1.2 $A\subset B$

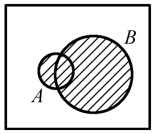

图 1.1.3 $A\cup B$

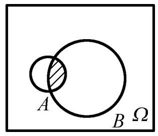

图 1.1.4 $A\cap B$

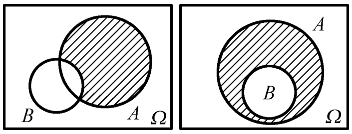

图 1.1.5 $A-B$

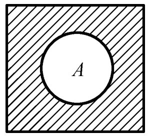

图 1.1.6 $\overline{A}$

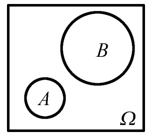

图 1.1.7 $AB=\varnothing$

【例 1.1.5】 甲、乙、丙三人各射一次靶，记 A 表示“甲中靶”，B 表示“乙中靶”，C 表示“丙中靶”，则可用上述三个事件的运算来分别表示下列各事件：

（1）“甲未中靶”： $\overline{A}$

（2）“甲中靶而乙未中靶”： $A\overline{B}$

（3）“三人中只有丙未中靶”： $AB\overline{C}$

（4）“三人中恰好有一人中靶”： $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$

（5）“三人中至少有一人中靶”： $A\cup B\cup C$

（6）“三人中至少有一人未中靶”： $\overline{A}\cup\overline{B}\cup\overline{C}$ 或 $\overline{ABC}$

（7）“三人中恰有两人中靶”： $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$

（8）“三人中至少有两人中靶”： $AB\cup AC\cup BC$

（9）“三人均未中靶”： $\overline{A}\,\overline{B}\,\overline{C}$

（10）“三人中至多有一人中靶”： $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}\,\overline{B}\,\overline{C}$

（11）“三人中至多有两人中靶”： $\overline{ABC}$ 或 $\overline{A}\cup\overline{B}\cup\overline{C}$

注：用其他事件的运算来表示一个事件，方法往往不唯一，如例 1.1.5 中的（6）和（11）实际上是同一事件，读者应学会用不同方法表达同一事件，特别是在解决具体问题时，往往要根据需要选择一种恰当的表示方法.

习题 1.1

1．写出下列随机试验的样本空间及下列事件包含的样本点.

（1）掷一颗骰子，出现奇数点.

（2）掷两颗骰子，$A=$“出现点数之和为奇数，且恰好其中有一个 1 点”，$B=$“出现点数之和为偶数，但没有一颗骰子出现 1 点”.

（3）将一枚硬币抛两次，$A=$“第一次出现正面”，$B=$“至少有一次出现正面”，$C=$“两次出现同一面”.

2．设 A、B、C 为三个事件，试用 A、B、C 的运算关系式表示下列事件：

（1）A 发生，B、C 都不发生；

（2）A 与 B 发生，C 不发生；

（3）A、B、C 都发生；

（4）A、B、C 至少有一个发生；

（5） A 、 B 、 C 都不发生；

（6） A 、 B 、 C 不都发生；

（7） A 、 B 、 C 至多有两个发生；

（8） A 、 B 、 C 至少有两个发生；

（9） A 、 B 、 C 不多于一个发生；

（10） A 、 B 、 C 至少有一个不发生.

3．指出下列各等式命题是否成立，并说明理由：

（1） $A \cup B = (A\overline{B}) \cup B$ ；

（2） $\overline{A}B = A \cup B$ ；

（3） $\overline{A \cup B} \cap C = \overline{A}\overline{B}\overline{C}$ ；

（4） $(AB)(A\overline{B}) = \varnothing$ ；

（5）若 $A \cup C = C \cup B$ ，则 $A = B$ ；

（6）若 $A - C = B - C$ ，则 $A = B$ ；

（7）若 $AC = BC$ ，则 $A = B$ ；

（8）若 $AB = \varnothing$ ，则 $\overline{A}\overline{B} = \varnothing$.

4．化简下列事件：

（1） $(\overline{A} \cup \overline{B})(\overline{A} \cup B)$ ； （2） $A\overline{B} \cup \overline{A}B \cup \overline{A}\overline{B}$ ； （3） $A \cap (B \cup \overline{B})$.

5．证明： $A - B = A\overline{B} = A - AB$.

6．从某学院学生中任选一名学生. 若事件 A 表示该生是男生，事件 B 表示该生是大学一年级学生，事件 C 表示该生是贫困生.

（1）叙述 $AB\overline{C}$ 的意义.

（2）在什么条件下 $ABC = B$ 成立？

（3）在什么条件下 $\overline{A} \subset B$ 成立？

7．一名射击选手连续向某个目标射击三次，事件 A_i 表示第 i （i=1,2,3）次射击时击中目标，试用文字描述下列事件： $A_1 \cup A_2$ ； A_2 ； $A_1A_2A_3$ ； $A_1 \cup A_2 \cup A_3$ ； $A_3 - A_2$ ； $\overline{A_2}A_3$ ； $\overline{A_1 \cup A_2}$ ； $\overline{A_2}\,\overline{A_3}$ ；$(A_1A_2) \cup (A_2A_3) \cup (A_1A_3)$.

8．从 N 件产品中任意抽取 M（ $M \leqslant N$ ）件，设 A 表示“至少有一件次品”， B 表示“至多有一件次品”，则 $\overline{A}$ 、 $\overline{B}$ 及 AB 各表示什么事件？

9．检验某种圆柱形产品时，要求长度与直径都符合要求时才算合格品，A =“产品合格”， B =“长度合格”， C =“直径合格”，试讨论：

（1） A 与 B 、 C 之间的关系；

（2） $\overline{A}$ 与 $\overline{B}$ $\overline{C}$ 之间的关系.

1.2　随机事件的概率

对一个随机事件 A ，在一次随机试验中是否会发生，事先不能确定. 人们常常希望了解某些事件在一次试验中发生的可能性的大小. 为此，首先引入频率的概念，它描述了事件发生的频繁程度，进而引出表示事件在一次试验中发生的可能性大小的数——概率.

1.2.1 频率

定义 1.2.1 若在相同条件下进行 n 次试验，其中事件 A 在 n 次试验中发生了 k 次，则称 $f_n(A)=\dfrac{k}{n}$ 为事件 A 在这 n 次试验中发生的**频率**（Frequency）.

由上述定义容易推知，频率具有下述基本性质：

1．对任一事件 A，有 $0\leqslant f_n(A)\leqslant 1$；

2．对必然事件 Ω，有 $f_n(\Omega)=1$；

3．设 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件，则

$$f_n(A_1\cup A_2\cup\cdots\cup A_n)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_n).$$

特别地，当事件 A、B 互不相容，则

$$f_n(A\cup B)=f_n(A)+f_n(B).$$

事件 A 发生的频率 $f_n(A)$ 表示 A 发生的频繁程度，频率越大，则事件 A 发生得越频繁，在一次试验中，A 发生的可能性也就大. 反之亦然. 因而，直观的想法是用 $f_n(A)$ 表示 A 在一次试验中发生可能性的大小. 但是，由于试验的随机性，即使同样是进行 n 次试验，$f_n(A)$ 的值也不一定相同. 但大量实验证实，随着重复试验次数 n 的增大，频率 $f_n(A)$ 会逐渐稳定于某个常数附近，而偏离的可能性很小. 频率具有“稳定性”这一事实，说明了刻画事件 A 发生可能性大小的数——概率具有一定的客观存在性（严格来说，这是一个理想的模型，因为实际上并不能绝对保证在每次试验时，条件都保持完全一样，这只是一个理想的假设）.

【例 1.2.1】 抛硬币的试验.

历史上有不少人做过抛硬币试验，其结果如表 1.2.1 所示，从表中的数据可以看出：出现正面的频率逐渐稳定在 0.5. 用频率的方法可以说：正面出现的可能性的大小为 0.5.

表 1.2.1 历史上抛硬币试验的若干结果

试　验　者	掷硬币次数	出现正面次数	出现正面的频率
德·摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
费勒	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

每个事件都存在一个这样的常数与之对应，当 n 无限增大时，可将频率 $f_n(A)$ 逐渐趋向稳定的这个常数定义为事件 A 发生的**概率.** 这就是概率的统计定义.

定义 1.2.2 在相同条件下重复进行 n 次试验，若事件 A 发生的频率 $f_n(A)=\dfrac{k}{n}$ 随着试验次数 n 的增大而稳定地在某个常数 p（$0\leqslant p\leqslant 1$）附近，则称 p 为事件 A 发生的概率，记为 $P(A)$.

频率的稳定值是概率的外在表现，并非概率的本质. 据此确定某事件的概率是困难的，但当进行大量重复试验时，频率会接近稳定值，因此，在实际应用时，往往是用试验次数足够大的频率来估计概率的大小，且随着试验次数的增大，估计的精度会越来越高.

为了理论研究的需要，我们从频率的稳定性和频率的基本性质得到启发，给出概率的公理化定义.

1.2.2　概率的公理化定义

定义 1.2.3　设 Ω 为样本空间，A 为事件，对于每一个事件 A 赋予一个实数，记为 $P(A)$，若 $P(A)$ 满足以下三条公理：

1．非负性：对每一个事件 A，有 $P(A) \geqslant 0$；

2．规范性：$P(\Omega)=1$；

3．可列可加性：对于两两互不相容的可列无穷多个事件 $A_1, A_2, \cdots, A_n, \cdots$，有

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

则称实数 $P(A)$ 为事件 A 的概率（Probability）.

在概率论发展史上，概率有很多的定义，如概率的古典定义、概率的几何定义、概率的频率定义和概率的主观定义，但是这些定义只能适合一类随机现象，而不适合所有的随机现象. 1900 年数学家希尔伯特（Hilbert，1862—1943 年）在巴黎第二届国际数学家大会上公开提出要建立概率的公理化体系. 直到 1933 年苏联数学家柯尔莫格洛夫（Kolmogorov，1903—1987 年）在他的《概率论基本概率》（丁涛田译，1952 年）一书中首次提出了概率的公理化定义，这个定义既概括了历史上几种概率定义中的共同特性，又避免了各自的局限性和含混之处，不论是何种随机现象，只有满足该定义中的三条公理，才能说它是概率. 这一公理化体系的出现迅速得到举世公认，为现代概率论发展打下坚实的基础，从此数学界才承认概率论是数学的一个分支. 有了这个公理化的体系之后，概率论得到快速发展. 这个公理化体系是概率论发展史上的一个里程碑. 由概率公理化定义（非负性、正则性和可列可加性）可以导出概率的一系列性质.

1.2.3　概率的性质

性质 1.2.1　$P(\varnothing)=0$.

证明　令 $A_n=\varnothing \ (n=1,2,\cdots)$，则

$$\bigcup_{n=1}^{\infty} A_n=\varnothing，且 A_i A_j=\varnothing \ (i \neq j,\ i, j=1,2,\cdots).$$

由概率的可列可加性得

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

而由 $P(\varnothing) \geqslant 0$ 及上式，知 $P(\varnothing)=0$.

思考问题：

不可能事件的概率为 0，概率为 0 的事件一定是不可能事件吗？

性质 1.2.2（有限可加性）　若 $A_1, A_2, \cdots, A_n$ 为两两互不相容事件，则有

$$P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} P(A_k)$$

证明 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$，则 $A_iA_j=\varnothing$. 当 $i\neq j$， $i,j=1,2,\cdots$ 时，由可列可加性，得

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

性质 1.2.3（对立事件公式） 对于任一事件 A，有

$$P(\overline{A})=1-P(A).$$

证明 因为 $\overline{A}\bigcup A=\Omega$， $\overline{A}\bigcap A=\varnothing$，由有限可加性，得

$$1=P(\Omega)=P(\overline{A}\bigcup A)=P(\overline{A})+P(A),$$

即

$$P(\overline{A})=1-P(A).$$

【例 1.2.2】 一颗骰子掷 4 次，求至少出现一次 6 点的概率.

解 用对立事件进行计算，记 $A=$“至少出现一次 6 点”，则 $\overline{A}=$“一次 6 点都不出现”，于是所求概率为

$$P(A)=1-P(\overline{A})=1-\frac{5^4}{6^4}\approx 0.5177.$$

性质 1.2.4 设 A、B 是两个事件，若 $A\subset B$，则有

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

证明 由 $A\subset B$ 知， $B=A\bigcup(B-A)$ 且 $A\bigcap(B-A)=\varnothing$.

再由概率的有限可加性，有

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

即

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

推论 1.2.1（单调性） 若 $A\subset B$，则 $P(A)\leqslant P(B)$.

推论 1.2.2 对任一事件 A， $P(A)\leqslant 1$.

证明 因为 $A\subset\Omega$，由推论 1.2.1 得 $P(A)\leqslant P(\Omega)=1$.

思考问题：

若 $P(A)\leqslant P(B)$，能否得出 $A\subset B$ 成立?

性质 1.2.5 $P(A-B)=P(A)-P(AB)$.

证明 因为 $A-B=A-AB$，且 $AB\subset A$，所以由性质 1.2.4 得

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

结论得证.

性质 1.2.6（加法公式） 对于任意两个事件 A、B 有

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

证明 因为 $A\bigcup B=A\bigcup(B-AB)$ 且 $A\bigcap(B-AB)=\varnothing$.

由性质 1.2.2、1.2.4 得

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

性质 1.2.6 还可推广到三个事件的情形. 例如，设 A_1、A_2、A_3 为任意三个事件，则有

$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_2A_3)-P(A_1A_3)+P(A_1A_2A_3).$$

一般地，设 $A_1,A_2,\cdots,A_n$ 为任意 n 个事件，可由归纳法证得

$$P(A_1\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).$$

推论 1.2.3（半可加性） 对任意两个事件 A、B 有

$$P(A\cup B)\leqslant P(A)+P(B).$$

对任意 n 个事件 $A_1,A_2,\cdots,A_n$，有

$$P\left(\bigcup_{k=1}^{n}A_k\right)\leqslant\sum_{k=1}^{n}P(A_k).$$

【例 1.2.3】 设 A、B 为两个事件，$P(A)=0.5$，$P(B)=0.3$，$P(AB)=0.1$，求：

（1）A 发生但 B 不发生的概率；

（2）A 不发生但 B 发生的概率；

（3）A、B 至少有一个事件发生的概率；

（4）A、B 都不发生的概率；

（5）A、B 至少有一个事件不发生的概率.

解 （1）$P(A\overline{B})=P(A-B)=P(A-AB)=P(A)-P(AB)=0.5-0.1=0.4$；

（2）$P(\overline{A}B)=P(B-AB)=P(B)-P(AB)=0.3-0.1=0.2$；

（3）$P(A\cup B)=P(A)+P(B)-P(AB)=0.5+0.3-0.1=0.7$；

（4）$P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-0.7=0.3$；

（5）$P(\overline{A}\cup\overline{B})=P(\overline{AB})=1-P(AB)=1-0.1=0.9$.

公理化定义没有告诉人们如何去计算概率. 历史上在公理化定义出现之前，概率的频率定义、古典定义、几何定义和主观定义都在一定的场合下有着各自计算概率的方法，所以在有了概率的公理化定义之后，把它们视为计算概率的方法是恰当的. 前面已经介绍了确定概率的频率方法，虽然用频率的方法确定概率合理，但是不可能一个试验无限次地重复下去，所以获得精确的频率稳定值很困难. 下面分别介绍确定概率的古典方法、几何方法和主观方法.

1.2.4　计算概率的古典方法

定义 1.2.4 若随机试验 E 满足以下条件：

（1）试验的样本空间 Ω 只有有限个样本点，即

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\};$$

（2）试验中每个基本事件的发生是等可能的，即

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n),$$

则称此试验为**古典概型**，或**等可能概型**. 在概率论的产生和发展过程中，它是最早的研究对象，且在实际中也是最常用的一种概率模型. 在古典概型的条件下，我们来推导事件发生的

概率的计算公式. 设事件 A 包含其样本空间 Ω 中 k 个基本事件，则事件 A 发生的概率

$$P(A)=\frac{A\text{包含的基本事件数}}{\Omega\text{中基本事件总数}}=\frac{k}{n}.$$

称此概率为**古典概率**. 这种确定概率的方法称为古典方法. 这就把求古典概率的问题转化为对基本事件的计数问题.

【例 1.2.4】 一个袋子中装有 10 个大小相同的球，其中 3 个黑球，7 个白球，求：

（1）从袋子中任取一球，这个球是黑球的概率；

（2）从袋子中任取两球，刚好取到一个白球和一个黑球的概率，以及两个球全是黑球的概率.

解 （1）10 个球中任取一个，共有 $C_{10}^1=10$ 种可能.

根据古典概率计算，事件 A："取到的球为黑球"的概率为 $P(A)=\dfrac{C_3^1}{C_{10}^1}=\dfrac{3}{10}$.

（2）10 球中任取两球的取法有 C_{10}^2 种，其中刚好一个白球和一个黑球的取法有 $C_3^1\ C_7^1$ 种取法，两个球均是黑球的取法有 C_3^2 种，记 B 为事件"刚好取到一个白球和一个黑球"，C 为事件"两个球均为黑球"，则

$$P(B)=\frac{C_3^1C_7^1}{C_{10}^2}=\frac{21}{45}=\frac{7}{15},\qquad P(C)=\frac{C_3^2}{C_{10}^2}=\frac{3}{45}=\frac{1}{15}.$$

【例 1.2.5】 将标号为 1、2、3、4 的 4 个大小相同的球随意地排成一行，求下列各事件的概率：

（1）各球自左至右或自右至左恰好排成 1、2、3、4 的顺序；

（2）第 1 号球排在最右边或最左边；

（3）第 1 号球与第 2 号球相邻；

（4）第 1 号球排在第 2 号球的右边（不一定相邻）.

解 将 4 个球随意地排成一行有 4!=24 种排法，即基本事件总数为 24.

记（1）、（2）、（3）、（4）的事件分别为 A、B、C、D.

（1）A 中有两种排法，故有 $P(A)=\dfrac{2}{24}=\dfrac{1}{12}$；

（2）B 中有 $2\times3!=12$ 种排法，故有 $P(B)=\dfrac{12}{24}=\dfrac{1}{2}$；

（3）先将第 1、2 号球排在任意相邻两个位置，共有 2×3 种排法，其余两个球可在其余两个位置任意排放，共有 2!种排法，因而 C 有 $2\times3\times2=12$ 种排法，故 $P(C)=\dfrac{12}{24}=\dfrac{1}{2}$；

（4）第 1 号球排在第 2 号球的右边的每一种排法，交换第 1 号球和第 2 号球的位置便对应于第 1 号球排在第 2 号球的左边的一种排法，反之亦然.

因而第 1 号球排在第 2 号球的右边与第 1 号球排在第 2 号球的左边的排法种数相同，各占总排法数的 $\dfrac{1}{2}$，故有 $P(D)=\dfrac{1}{2}$.

【例 1.2.6】 有 n 个人，每个人都以同样的概率 $\dfrac{1}{N}$ 被分配在 N（$n<N$）间房的任一间中，求恰好有 n 个房间，其中各住一人的概率.

解 每个人都有 N 种分法，这是可重复排列问题，n 个人共有 N^n 种不同分法. 因为没有

指定是哪几间房，所以首先选出 n 间房，有 C_N^n 种选法. 对于其中每一种选法，每间房各住一人共有 $n!$ 种分法，故所求概率为

$$p=\frac{C_N^n n!}{N^n}.$$

许多直观背景很不相同的实际问题，都和本例具有相同的数学模型. 比如生日问题：假设每人的生日在一年 365 天中的任一天是等可能的，那么随机选取 n（$n\leqslant 36$）个人，他们的生日各不相同的概率为

$$p_1=\frac{C_{365}^n n!}{365^n},$$

因而 n 个人中至少有两个人生日相同的概率为

$$p_2=1-\frac{C_{365}^n n!}{365^n}.$$

如当 $n=64$ 时，$p_2\approx 0.997$，这表示在仅有 64 人的班级里，“至少有两人生日相同”的概率与 1 相差无几，因此几乎总是会出现的. 这个结果也许会让大多数人惊奇，因为“一个班级中至少有两人生日相同”的概率并不如人们直觉中想象得那样小，而是相当大. 这也告诉我们，“直觉”并不很可靠，说明研究随机现象统计规律是非常重要的.

1.2.5　计算概率的几何方法

上述古典概型的计算只适用于具有等可能性的有限样本空间，若试验结果无穷多，则它显然已不适合. 为了克服有限的局限性，可将古典概型的计算加以推广.

设试验具有以下特点：

（1）样本空间 Ω 是一个几何区域，这个区域大小可以度量（如长度、面积、体积等），并把 Ω 的度量记为 $\mu(\Omega)$.

（2）向区域 Ω 内任意投掷一个点，落在区域内任一个点处都是“等可能的”. 或者设落在 Ω 中的区域 A 内的可能性与 A 的度量 $\mu(A)$ 成正比，与 A 的位置和形状无关.

不妨也用 A 表示“掷点落在区域 A 内”的事件，那么事件 A 的概率可用下列公式计算：

$$P(A)=\frac{\mu(A)}{\mu(\Omega)}$$

称为**几何概率**，这种类型的概率称为**几何概型**.

注意：若样本空间 Ω 为一条线段或一个空间立体，则向 Ω “投点”的相应概率仍可用 $P(A)=\mu(A)/\mu(\Omega)$ 确定，但 $\mu(A)$ 应理解为长度或体积.

【例 1.2.7（会面问题）】 甲、乙两人相约在 7 点到 8 点之间在某地会面，先到者等候另一人 20 分钟，过时就离开. 如果每个人可在指定的一小时内任意时刻到达，试计算两人能够会面的概率.

解　记 7 点为计算时刻的 0 时，以分钟为单位，x、y 分别记甲、乙达到指定地点的时刻，则样本空间为 $\Omega=\{(x,y)\mid 0\leqslant x\leqslant 60, 0\leqslant y\leqslant 60\}$. 设 A 表示事件“两人能会面”，则显然有 $A=\{(x,y)\mid (x,y)\in\Omega, |x-y|\leqslant 20\}$，如图 1.2.1 所示，根据题意，这是一个几何概型问题，于是

$$P(A)=\frac{\mu(A)}{\mu(\Omega)}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

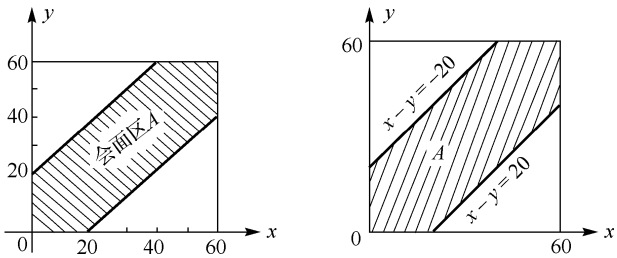

图 1.2.1　例 1.2.7 图

【例 1.2.8】 在区间 (0,1) 内任取两个数，求这两个数的乘积小于 $\dfrac{1}{4}$ 的概率.

解　设在 (0,1) 内任取两个数为 x、y，则

$$0<x<1,\quad 0<y<1$$

即样本空间是由点 (x,y) 构成的边长为 1 的正方形 Ω，其面积为 1.

令 A 表示“两个数乘积小于 1/4”，则

$$A=\left\{(x,y)\,|\,0<xy<\frac{1}{4},0<x<1,0<y<1\right\}$$

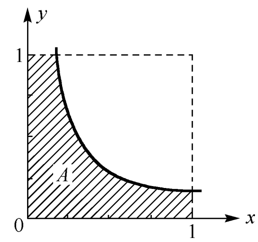

图 1.2.2　例 1.2.8 图

事件 A 所围成的区域如图 1.2.2 所示，则所求概率

$$P(A)=\frac{1-\int_{\frac{1}{4}}^{1}\mathrm{d}x\int_{\frac{1}{4x}}^{1}\mathrm{d}y}{1}=\frac{1-\int_{\frac{1}{4}}^{1}\left(1-\frac{1}{4x}\right)\mathrm{d}x}{1}=1-\frac{3}{4}+\int_{\frac{1}{4}}^{1}\frac{1}{4x}\mathrm{d}x=\frac{1}{4}+\frac{1}{2}\ln 2\,.$$

1.2.6　计算概率的主观方法

在现实世界中，有些随机现象是不能重复的或不能大量重复的，这时有关事件的概率如何确定呢？

统计界的贝叶斯学派认为：一个事件的概率是人们根据经验对该事件发生的可能性所给出的个人信念. 这样给出的概率称为**主观概率**.

这种利用经验确定随机事件发生可能性大小的例子是很多的，人们也常依据某些主观概率来行事.

【例 1.2.9】 用主观方法确定概率的例子.

（1）一位高中班主任根据自己多年的教学经验及对张三、李四两学生的平时学习情况的了解，认为“张三能考取重点大学”的可能性为 90%，“李四能考取重点大学”的可能性为 20%.

（2）一位心脏病医生根据自己多年的临床经验和一位患者的病情，认为“此次心脏移植手术成功”的可能性为 91%.

（3）一位商家根据他多年的经验和目前的一些市场信息，认为“某项新的电子产品在未来市场上畅销”的可能性为 85%.

从以上的例子可以看出：

（1）主观概率和主观臆造有着本质上的不同，前者要求当事人对所考察的事件有透

彻的了解和丰富的经验，甚至是这一行的专家，并能对历史信息和当时信息进行仔细分析，如此确定的主观概率是可信的. 从某种意义上说，不利用这些丰富的经验也是一种浪费.

（2）用主观方法得出的随机事件发生的可能性大小，本质上是对随机事件概率的一种推断和估计. 虽然结论的精确性有待实践和修正，但结论的可信性在统计意义上是有价值的.

（3）在遇到随机现象无法大量重复时，用主观方法去做决策和判断是适合的. 从这点看，主观方法至少是概率方法的一种补充.

另外要说明的是，主观概率的确定除根据自己的经验外，决策者还可以利用别人的经验. 例如，对一项有风险的投资，决策者向某位专家咨询的结果为“成功的可能性为 60%”. 而决策者很熟悉这位专家，认为专家的估计往往是偏保守的、过分谨慎的. 为此决策者将结论修改为“成功的可能性为 70%”. 主观给定的概率要符合公理化的定义.

习题 1.2

1．将一枚硬币抛掷三次，求：

（1）恰有一次出现正面的概率；

（2）至少有一次出现正面的概率.

2．将 3 个球随机放入 4 个杯子中，问杯子中球的个数最多为 1、2、3 的概率各是多少?

3．将 15 名新生（其中有 3 名优秀生）随机地分配到三个班级中，其中一班 4 名，二班 5 名，三班 6 名，求：

（1）每一个班级各分配到一名优秀生的概率；

（2）3 名优秀生被分配到一个班级的概率.

4．对一个 7 人团队小组考虑生日问题：

（1）求 7 个人的生日都在星期日的概率；

（2）求 7 个人的生日都不在星期日的概率；

（3）求 7 个人的生日不都在星期日的概率.

5．六根草，头两两相接、尾两两相接，求成环的概率.

6．n 个人围一圆桌坐，求甲、乙两人相邻而坐的概率.

7．n 个人坐成一排，求甲、乙两人相邻而坐的概率.

8．两颗骰子掷 24 次，求至少出现一次双 6 点的概率.

9．从 $1,2,\cdots,9$ 中返回取 n 次，求取出的 n 个数的乘积能被 10 整除的概率.

10. 观察某地区未来 5 天的天气情况，记 A_i 为事件“有 i 天不下雨”，已知 $P(A_i)=iP(A_0)$，$i=1,2,3,4,5$，求下列各事件的概率：

（1）5 天均下雨；

（2）至少一天不下雨；

（3）至少一天下雨.

11．设 A、B 是两个事件，且 $P(A)=0.6$，$P(B)=0.7$，求：

（1）在什么条件下 $P(AB)$ 取到最大值？

（2）在什么条件下 $P(AB)$ 取到最小值？

12. 设 A、B、C 为三事件，且 $P(A)=P(B)=\frac{1}{4}$，$P(C)=\frac{1}{3}$ 且 $P(AB)=P(BC)=0$，$P(AC)=\frac{1}{12}$，求 A、B、C 至少有一个事件发生的概率.

13．甲、乙两名篮球运动员，投篮命中率分别为 0.8 及 0.7，每人各投了 3 次，求两人进球数相等的概率.

14．从 5 双不同的鞋子中任取 4 只，求这 4 只鞋子中至少有两只鞋子配成一双的概率.

15．某人午觉醒来，发觉表停了，他打开收音机，想听电台报时，设电台每逢正点时报时一次，求他（她）等待时间短于 10 分钟的概率.

16．两人约定上午 11:00～12:00 在某地点会面，求一人要等另一人 20 分钟以上的概率.

17．从 $(0,1)$ 中随机地取两个数，求：

（1）两个数之和小于 $\frac{1}{5}$ 的概率；

（2）两个数之差小于 $\frac{1}{4}$ 的概率.

18．某伞兵夜间在甲、乙两镇之间降落，求他降落的地点离甲镇至少是离乙镇距离的两倍的概率.

19．甲、乙两船驶向同一不能同时停泊两条船的码头，它们在一昼夜内到达的时刻是等可能的，如果甲船的停泊时间是 1 小时，乙船的停泊时间是 2 小时，求它们中任何一船都不需要等候码头空出的概率.

1.3 条 件 概 率

在解决许多概率问题时，往往需要在有某些附加信息（条件）的情况下求事件的概率. 例如，10 个人摸 10 张彩票，其中有 3 张是有奖彩票，已知第 1 个人没摸中，则第 2 个人中彩的概率是多少？已知在某事件发生的条件下求另一事件发生的概率，把求这样事件的概率称为条件概率，条件概率是概率论中的一个基本概念，也是概率论中的一个重要工具，它既可以帮助我们认识更复杂的随机事件，又可以帮助我们计算一些复杂事件的概率.

1.3.1 条件概率的定义

定义 1.3.1 设 A、B 为两个事件，且 $P(A)>0$，则称 $\frac{P(AB)}{P(A)}$ 为在事件 A 发生的条件下事件 B 发生的**条件概率**，记为 $P(B|A)$，即

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

【例 1.3.1】 袋中有 5 个球，有 3 个红球和 2 个白球. 现从袋中不放回地连取两个. 已知第一次取得红球，求第二次取得白球的概率.

解法 1：设 A 表示“第一次取得红球”，B 表示“第二次取得白球”，求 $P(B|A)$.

在 5 个球中不放回连取两球的取法有 P_5^2 种，其中，第一次取得红球的取法有 $P_3^1\ P_4^1$ 种，第一次取得红球同时第二次取得白球的取法有 $P_3^1\ P_2^1$ 种，所以

$$P(A)=\frac{P_3^1P_4^1}{P_5^2}=\frac{3}{5}, \qquad P(AB)=\frac{P_3^1P_2^1}{P_5^2}=\frac{3}{10}.$$

由定义得
$$P(B\mid A)=\frac{P(AB)}{P(A)}=\frac{\frac{3}{10}}{\frac{3}{5}}=\frac{1}{2}.$$

解法 2: 设 A 表示“第一次取得红球”，B 表示“第二次取得白球”，依题意要求 $P(B\mid A)$. 缩减样本空间 A 中的样本点数，即第一次取得红球的取法为 $P_3^1\ P_4^1$，而第二次取得白球的取法有 $P_3^1\ P_2^1$ 种，所以 $P(B\mid A)=\frac{P_3^1P_2^1}{P_3^1P_4^1}=\frac{1}{2}$.

把这种缩小了样本空间来求条件概率的方法称为缩小样本空间法.

思考问题:

这个例子中 $P(AB)$ 与 $P(B\mid A)$ 的区别是什么？

性质 1.3.1　条件概率是概率，即若设 $P(A)>0$，则：

（1）非负性：对于任一事件 B，有 $P(B\mid A)\geqslant 0$；

（2）规范性： $P(\Omega\mid A)=1$；

（3）可列可加性： $P\left(\bigcup\limits_{i=1}^{\infty}B_i\mid A\right)=\sum\limits_{i=1}^{\infty}P(B_i\mid A)$，其中 $B_1,B_2,\cdots,B_n,\cdots$ 为两两互不相容事件.

这说明条件概率符合概率定义的三条公理，故对概率已证明的结果都适用于条件概率. 例如，对于任意事件 A_1、A_2，有
$$P(A_1\cup A_2\mid B)=P(A_1\mid B)+P(A_2\mid B)-P(A_1A_2\mid B).$$
又如，对于任意事件 A，有
$$P(\overline{A}\mid B)=1-P(A\mid B).$$

1.3.2　乘法公式

由条件概率定义 $P(B\mid A)=\frac{P(AB)}{P(A)}$，$P(A)>0$，两边同乘以 $P(A)$ 可得 $P(AB)=P(A)P(B\mid A)$，由此可得以下定理.

定理 1.3.1（乘法公式）　设 $P(A)>0$，则有
$$P(AB)=P(A)P(B\mid A).$$

易知，若 $P(B)>0$，则有
$$P(AB)=P(B)P(A\mid B).$$

乘法定理也可推广到三个事件的情况，例如，设 A、B、C 为三个事件，且 $P(AB)>0$，则有
$$P(ABC)=P(C\mid AB)P(AB)=P(C\mid AB)P(B\mid A)P(A).$$

一般地，设 n 个事件为 $A_1,A_2,\cdots,A_n$，若 $P(A_1A_2\cdots A_n)>0$，则有
$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1A_2\cdots A_{n-1}).$$

事实上，由 $A_1 \supset A_1A_2 \supset \cdots \supset A_1A_2\cdots A_{n-1}$，有 $P(A_1) \geqslant P(A_1A_2) \geqslant \cdots \geqslant P(A_1A_2\cdots A_{n-1}) > 0$.
故公式右边的条件概率每一个都有意义，由条件概率定义可知

$$P(A_1)P(A_2 \mid A_1)\ P(A_3 \mid A_1A_2)\cdots P(A_n \mid A_1A_2\cdots A_{n-1})$$

$$= P(A_1)\cdot\frac{P(A_1A_2)}{P(A_1)}\cdot\frac{P(A_1A_2A_3)}{P(A_1A_2)}\cdots\frac{P(A_1A_2\cdots A_n)}{P(A_1A_2\cdots A_{n-1})} = P(A_1A_2\cdots A_n).$$

【例1.3.2】 一袋中装 10 个球，其中 3 个黑球、7 个白球，先后两次从中随意各取一球（不放回），求两次取到的均为黑球的概率.

分析：这一概率，我们曾用古典概型方法计算过，这里使用乘法公式来计算. 在本例中，问题本身提供了两步完成一个试验的结构，这恰恰与乘法公式的形式相应，合理地利用问题本身的结构来使用乘法公式，往往是使问题得到简化的关键.

解 设 A_i 表示事件“第 i 次取到的是黑球”（$i=1,2$），则 $A_1 \cap A_2$ 表示事件“两次取到的均为黑球”. 由题设知 $P(A_1)=\frac{3}{10}$，$P(A_2 \mid A_1)=\frac{2}{9}$，于是根据乘法公式，有 $P(A_1A_2)=P(A_1)P(A_2 \mid A_1)=\frac{3}{10}\times\frac{2}{9}=\frac{1}{15}$.

【例1.3.3】 设袋中装有 r 只红球和 t 只白球. 每次自袋中任取一只球，观察其颜色后放回，并再放入 a 只与所取出的那只球同色的球. 若在袋中连续取球四次，试求第一、二次取到红球且第三、四次取到白球的概率.

解 以 A_i（$i=1,2,3,4$）表示事件“第 i 次取到红球”，则 $\overline{A_3}$、$\overline{A_4}$ 分别表示事件第三、四次取到白球. 所求概率为

$$P(A_1A_2\overline{A_3}\,\overline{A_4}) = P(A_1)P(A_2 \mid A_1)P(\overline{A_3} \mid A_1A_2)P(\overline{A_4} \mid A_1A_2\overline{A_3})$$
$$= \frac{r}{r+t}\cdot\frac{r+a}{r+t+s}\cdot\frac{t}{r+t+2a}\cdot\frac{t+a}{r+t+3a}.$$

【例 1.3.4】 设某光学仪器厂制造的透镜，第一次落下时打破的概率为 $\frac{1}{2}$，若第一次落下未打破，则第二次落下打破的概率为 $\frac{7}{10}$，若前两次落下均未打破，则第三次落下打破的概率为 $\frac{9}{10}$. 试求透镜落下三次而未打破的概率.

解 以 A_i（$i=1,2,3$）表示事件“透镜第 i 次落下打破”，B 表示事件“透镜落下三次而未打破”. 则 $B=\overline{A_1}\,\overline{A_2}\,\overline{A_3}$，故有

$$P(B)=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(\overline{A_3} \mid \overline{A_2}\,\overline{A_1})=\left(1-\frac{1}{2}\right)\left(1-\frac{7}{10}\right)\left(1-\frac{9}{10}\right)=\frac{3}{200}.$$

1.3.3 全概率公式

定义 1.3.2 设 Ω 为样本空间，$A_1,A_2,\cdots,A_n$ 为 Ω 的一组事件，若满足

（1）$A_iA_j=\varnothing$，$i\neq j$，$i,j=1,2,\cdots,n$；

（2）$\bigcup_{i=1}^{n}A_i=\Omega$.

则称 $A_1,A_2,\cdots,A_n$ 为样本空间 Ω 的**一个划分**或称为样本空间 Ω 的**一个完备事件组.** 如图 1.3.1 所示.

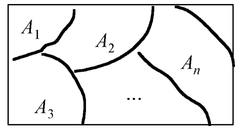

图 1.3.1　样本空间的一个划分

例如，A、$\overline{A}$ 就是 Ω 的一个划分.

若 $A_1,A_2,\cdots,A_n$ 是 Ω 的一个划分，则对每次试验，事件 $A_1,A_2,\cdots,A_n$ 中必有一个且仅有一个发生.

定理 1.3.2（**全概率公式**）设 $A_1,A_2,\cdots,A_n$ 是样本空间 Ω 的一个完备事件组，且 $P(A_i)>0$，$i=1,2,\cdots,n$，则对任一事件 $B\subset\Omega$，有

$$P(B)=P(A_1)P(B\,|\,A_1)+P(A_2)P(B\,|\,A_2)+\cdots+P(A_n)P(B\,|\,A_n),$$

称上述公式为全概率公式.

证明　显然 $B=B\Omega=B\left(\bigcup_{i=1}^{n}A_i\right)=\bigcup_{i=1}^{n}(BA_i)$.

因为 $A_iA_j=\varnothing$（$i\neq j$），所以 $(BA_i)(BA_j)=B(A_iA_j)=B\varnothing=\varnothing$（$i\neq j$），

从而 $P(B)=P\left(\bigcup_{i=1}^{n}(BA_i)\right)=\sum_{i=1}^{n}P(BA_i)$，

又因为 $P(A_i)>0$，由乘法公式有

$$P(B)=P(A_1)P(B\,|\,A_1)+P(A_2)P(B\,|\,A_2)+\cdots+P(A_n)P(B\,|\,A_n).$$

全概率公式表明，在许多实际问题中事件 B 的概率不易直接求得，可根据具体情况构造一个完备事件组 $\{A_i\}$，使事件 B 发生的概率是各事件 A_i（$i=1,2,\cdots$）发生条件下引起事件 B 发生的概率的总和.

由定理 1.3.2 可知，全概率公式最简单的形式为

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

思考问题：

如果定理 1.3.2 中的条件 $A_1,A_2,\cdots,A_n$ 为 Ω 的一个划分改写成 $A_1,A_2,\cdots,A_n$ 互不相容，且 $B\subset\bigcup_{i=1}^{n}A_i$，那么是否有 $P(B)=\sum_{i=1}^{n}P(A_i)P(B\,|\,A_i)$ 成立？

【例 1.3.5】 人们为了解一只股票未来一定时期内价格的变化，往往会去分析影响股票价格的基本因素，比如利率的变化. 现假设人们经分析估计利率下调的概率为 60%，利率不变的概率为 40%. 根据经验，人们估计，在利率下调的情况下，该支股票价格上涨的概率为 80%，而在利率不变的情况下，其价格上涨的概率为 40%，求该只股票将上涨的概率.

解　记 A 为事件“利率下调”，那么 $\overline{A}$ 即为“利率不变”，记 B 为事件“股票价格上涨”. 依题设知

$$P(A)=60\%，\ P(\overline{A})=40\%，\ P(B\,|\,A)=80\%，\ P(B\,|\,\overline{A})=40\%.$$

于是这只股票将上涨的概率为

$$P(B)=P(A)P(B\,|\,A)+P(\overline{A})P(B\,|\,\overline{A})=60\%\times80\%+40\%\times40\%=64\%.$$

【例 1.3.6】 某商店收进甲厂生产的产品 30 箱，乙厂生产的同种产品 20 箱，甲厂每箱装 100 个，废品率为 0.06，乙厂每箱装 120 个，废品率为 0.05，求：

（1）任取一箱，从中任取一个为废品的概率；

（2）若将所有产品开箱混放，求任取一个为废品的概率.

解 记事件 A 、B 分别为甲、乙两厂的产品，C 为废品，则

（1）$P(A)=\dfrac{30}{50}=\dfrac{3}{5}$，$P(B)=\dfrac{20}{50}=\dfrac{2}{5}$，$P(C\mid A)=0.06$，$P(C\mid B)=0.05$，

由全概率公式，得 $P(C)=P(A)P(C\mid A)+P(B)P(C\mid B)=0.056$.

（2）$P(A)=\dfrac{30\times 100}{30\times 100+20\times 120}=\dfrac{5}{9}$，$P(B)=\dfrac{20\times 120}{30\times 100+20\times 120}=\dfrac{4}{9}$，$P(C\mid A)=0.06$，$P(C\mid B)=0.05$.

由全概率公式，得 $P(C)=P(A)P(C\mid A)+P(B)P(C\mid B)\approx 0.056$.

1.3.4 贝叶斯公式

利用全概率公式，可通过综合分析一个事件发生的不同原因、情况或途径及其可能性来求得该事件发生的概率. 下面给出的贝叶斯公式则考虑与之完全相反的问题，即一个事件已经发生，要考察该事件发生的各种原因、情况或途径的可能性. 例如，第 1、2、3 号箱子中分别放有不同数量和颜色的球，现从任一箱中任意摸出一球，发现是红球，求该球是取自 1 号箱的概率. 或问：该球取自哪号箱的可能性最大？

定理 1.3.3 设 $A_1,A_2,\cdots,A_n$ 是一完备事件组，则对任一事件 B，$P(B)>0$，有

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{\sum\limits_{j=1}^{n}P(A_j)P(B\mid A_j)}，\quad i=1,2,\cdots,n.$$

证明 根据条件概率定义及乘法公式，对任意 $i=1,2,\cdots,n$，有

$$P(A_i\mid B)=\frac{P(A_iB)}{P(B)}=\frac{P(A_i)P(B\mid A_i)}{P(B)},$$

这里分母 $P(B)$ 用全概率公式即有

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{\sum\limits_{j=1}^{n}P(A_j)P(B\mid A_j)}.$$

这个公式称为**贝叶斯（Bayes）公式**. 公式的实际背景是：已知出现了试验“结果” B，要求推断是哪一个“原因” A_i 造成结果 B 的可能性大. 方法步骤是：首先计算每一个 $P(A_i)$，它反映了各种“原因”发生的可能性大小，由于它是在试验之前产生的，因此称为**先验概率**；之后计算 $P(B\mid A_i)$，它表示“原因” A_i 发生的条件下结果 B 的概率；再由贝叶斯公式反推出“结果” B 已经发生的条件下，“原因” A_i 发生的概率 $P(A_i\mid B)$，因为它是在试验之后确定的，因此称之为**后验概率**. 比较各个 $P(A_i\mid B)$ 的大小，如果 $P(A_k\mid B)$ 是所有 $P(A_i\mid B)$ 中最大的那一个，那么就说明造成“结果” B 发生的最可能“原因”是 A_k.

特别地，若取 $n=2$，并记 $A_1=A$，则 $A_2=\overline{A}$，于是公式成为

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