概率论与数理统计

郭跃华　主编
陆志峰　钱　峰　赵为华　副主编

科 学 出 版 社
北　京

内 容 简 介

本书介绍概率论与数理统计的基本概念、基本理论和方法, 主要内容包括：随机事件与概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基本概念、参数估计、假设检验、统计分析初步、随机过程等. 全书从直观分析入手逐步过渡到严格的数学表述, 使学生通过学习, 除了掌握概率统计的基本理论外, 还能拓宽知识面, 具备应用计算机解决问题的能力.

本书可作为普通高等院校的理工(非数学)、经管类各专业本科学生教材, 也可供对数学要求较高的高职高专学生使用.

图书在版编目(CIP)数据

概率论与数理统计/郭跃华主编. —北京: 科学出版社, 2007
ISBN 978-7-03-019573-9

Ⅰ.概… Ⅱ.郭 Ⅲ. ① 概率论-高等学校-教材 ② 数理统计-高等学校-教材 Ⅳ.O21

中国版本图书馆 CIP 数据核字(2007) 第 121979 号

责任编辑: 胡华强　李晓鹏　潘继敏 / 责任校对: 陈玉凤
责任印制: 张克忠 / 封面设计: 王　浩

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

中国科学院印刷厂 印刷
科学出版社发行　各地新华书店经销
*
2007 年 8 月第　一　版　开本: B5(720×1000)
2007 年 8 月第一次印刷　印张: 20
印数: 1—5 500　字数: 379 000

定价: 25.00 元

(如有印装质量问题, 我社负责调换〈科印〉)

前　　言

“概率论与数理统计” 是培养学生利用随机思维模式看待和处理随机现象的一门重要的数学基础课程. 由于随机现象存在的普遍性、研究方法的独特性和教学内容的实用性, 这门课越来越受到人们的重视, 高等学校理、工、管类各专业普遍开设了概率论与数理统计课程. 该门课程研究的目的是从大量数据中研究随机现象蕴含的统计规律. 由于理论方法的应用涉及大量的计算, 其中过于深奥的抽象理论、复杂的公式、陈旧的计算手段限制了它的实用性. 为了适应现代化教学手段和计算机软件应用, 结合理、工、管类各专业的教学实际, 我们编写了《概率论与数理统计》这本教材.

理 (非数学)、工、管类大学生的数学素养主要是通过应用数学来体现, 因此我们在教材中突出数学建模的思想, 以加强应用数学的意识和能力的培养为核心, 结合当前学生状况, 在教学内容中删掉一些过于繁琐的推理和完全可以用计算机代替的计算, 压缩一些论述, 增加一些背景知识的介绍, 尽量增加一些与当今社会生活和现代科技密切相关的实际应用问题与综合应用问题. 本书引入了 Excel 进行统计计算和仿真, 为了不占用课堂时间, 我们以实验形式给出了计算和仿真的具体步骤, 可由学生课后自行利用计算机进行, 以提高学生的实际动手能力, 同时减少了花在繁琐计算上的时间, 使学生有更多的时间用在概念的理解和方法的掌握上.

本教材的主要特色如下：

(1) 增加一些概念背景, 由实际案例引进新概念.

(2) 删除一些复杂的理论推导和繁琐的人工计算, 突出重点, 加强方法的应用.

(3) 计算机工具的广泛使用. 通过计算机仿真实验、函数计算及程序调用, 把 Excel 工具广泛使用于概念的引进和数值计算, 帮助学生形象地理解新概念, 直达核心处理思想, 同时提高学生的动手能力.

(4) 贯穿随机思想, 强调实际应用. 通过应用案例分析方法, 引导学生灵活运用所学知识, 掌握随机处理的基本过程, 提高学生分析问题和解决实际问题的能力.

(5) 融入实验教学内容, 注重理论的实践与应用.

(6) 每一节配备适量与本节内容相关的练习题, 便于学生巩固课堂知识; 每一章配综合习题, 便于学生提高知识水平.

本教材分三部分：第一部分为概率论 (第 1~5 章), 是后面应用部分的理论基础; 第二部分为数理统计 (第 6~9 章), 重点阐述了参数估计和假设检验, 并介绍了方差分析和回归分析两个常用的统计方法的最基本部分; 第三部分为随机过程 (第 10、11 章), 介绍了一些基本理论和几个常见的随机过程, 并着重讨论了马尔可夫

链. 数理统计与随机过程内容相互独立, 可根据专业的需要选用. 本教材主要适用于对理论证明要求不高的大学本科各专业的概率统计课程.

全书各章分别由郭跃华 (第 1、9~11 章)、陆志峰 (第 6~8 章)、钱峰 (第 2、3 章)、赵为华 (第 4、5 章) 执笔. 最后由郭跃华统稿.

全书承蒙南京大学博士生导师、江苏省概率统计学会理事长王金德教授审阅, 并提出了许多宝贵意见, 在此表示衷心感谢. 本书得到科学出版社的大力支持, 在此一并表示衷心感谢.

限于编者水平, 有不当和错误之处, 恳请读者和专家批评指正.

郭跃华

2007 年 05 月

目　录

第 1 章　随机事件与概率

艾滋病普查

艾滋病 (acquired immune deficiency syndrome, 获得性免疫缺陷综合症, 缩写为 AIDS) 是现在全球较流行的一种致命的接触性传染病. 艾滋病的受害者几乎没有活过五年的. 到 2004 年全球艾滋病病毒感染者人数总计为 3940 万, 死亡人数为 310 万 (联合国艾滋病规划署和世界卫生组织《艾滋病流行最新报告——2004 年 12 月》). 中国卫生部、联合国艾滋病规划署和世界卫生组织 2006 年 1 月 25 日发布的最新数据显示, 中国现在的艾滋病病毒感染者人数为 65 万, 总感染率估计为 0.05%. 为了阻止这种病的传播, 首先要识别艾滋病病毒的携带者, 对这些人采取适当的预防措施以避免传染给其他人. 科学家一直在寻求识别艾滋病病毒携带者的有效方法. 其中一种方法就是血液试验 (ELISA, enzymelinked immunosorbent assay(酶连接免疫吸附测定) 的缩写), 它能检测出身体中艾滋病病毒的某种抗体的存在性. 尽管这种试验是极其精确的, 但它还是会有两种可能的误诊: 首先, 它可能会对某些真有艾滋病的人作出没有艾滋病的诊断, 这就是所谓的“假阴性”; 其次, 它可能会对某些没有艾滋病的人误诊为患有艾滋病, 这就是“假阳性”.

假设血液试验改进到这样的程度, 它能正确识别患有艾滋病的人中的 95%, 因此 5% 的患有艾滋病的人的血液试验结果将是“假阴性”. 进一步假设接受血液试验的不带艾滋病病毒的人中 99% 的试验结果为阴性. 这就意味着不带艾滋病病毒的健康人中的 1%, 其血液试验结果将是“假阳性”.

美国是艾滋病较为流行的国家之一, 据保守估计大约每 1000 人中就有 1 人受这种病的折磨. 为了能有效地控制和减缓艾滋病的传播, 几年前有人就提议应在申请结婚登记的新婚夫妇中进行艾滋病病毒的血液试验. 该项普查计划一经提出, 立刻就遭到了许多专家学者的反对, 他们认为这是一项既费钱又费力, 同时收效不大的计划. 最终, 此项计划未被通过. 那么, 到底专家的意见对不对? 该普查计划该不该被执行呢? 如果该计划得以实施, 而某人又做了血液试验并且结果是阳性, 他将怎样为之发愁呢? 他真正得了艾滋病的可能性有多大?

在我们的生活中经常会面临许多包含有不确定性的决策问题. 如一笔新投资盈利的可能性有多大? 一项工程按期完成的可能性有多少? 等等.

解决上述这些问题需要运用一些概率论的基本知识. 概率论为解决不确定性

问题提供了最有效的理论和方法. 常言道 "天有不测风云, 人有旦夕祸福", 风险无处不在, 无时不有. 无论在科研、生产、保险、风险投资、管理决策等许多领域, 还是在日常生活中, 人们都会遇到如何在面临不确定性的情况下做出正确决策的问题. 因此, 概率论有着广泛而重要的应用价值. 本章将介绍概率论的基础知识 —— 随机事件与概率.

1.1 随 机 事 件

1.1.1 基本术语

人类社会和自然界发生的现象多种多样. 其中有一类现象, 其规律是只要具备一定的条件, 某确定的现象一定发生 (或一定不发生). 例如：

如果平面图形是三角形 (条件), 则其内角和一定是 $180°$(现象);

在一个标准大气压、温度 $100°$C(条件) 下, 纯水一定沸腾 (现象);

两个同性电荷 (条件) 一定互斥 (现象);

冬天过去 (条件), 春天就会到来 (现象);

在常温 (条件) 下, 铁一定不熔化 (现象);

等等.

定义 1.1.1　在一定条件下, 一定发生 (或一定不发生) 的现象称为**必然现象**.

与必然现象不同, 还存在另一类现象, 其规律是, 在一定条件下, 某种现象可能发生, 也可能不发生. 例如：

抛掷一枚质地均匀的硬币 (条件), 硬币落地后, 带币值的一面朝上 (现象) 可能发生, 也可能不发生;

火炮对坦克射出一发炮弹 (条件), 命中坦克 (现象) 可能发生, 也可能不发生;

同时抛掷两颗质地均匀的骰子 (条件), 点数朝上之和恰好为 6(现象) 可能发生, 也可能不发生;

…….

定义 1.1.2　在一定条件下, 可能发生, 也可能不发生的现象称为**随机现象**.

概率论就是研究随机现象中的数量规律的科学.

定义 1.1.3　在一定条件下, 对某随机现象的一次观察或测量称为**随机试验**(简称**试验**), 记为 E. 随机试验具备以下三条性质：

(1) 试验能在相同条件下重复进行;

(2) 试验结果不止一个, 但能明确所有可能的结果;

(3) 试验前不能预知试验会出现哪个结果.

进行一次试验, 总要有一个观测目的. 试验中可能观测到多种不同的结果. 以抛一枚硬币为例, 如果观测的目的只是看硬币落地后哪一面朝上, 由于硬币不变, 所以掷硬币的行为可以在相同条件下重复实施. 在掷硬币之前, 我们知道出现的结果只能是币值一面或另一面, 但具体是哪一面出现不能确切地预料. 可见这种情形是一个随机试验, 记为 E_1.

下面再看一些随机试验的例子.

E_2: 掷一颗骰子, 观察出现的点数;

E_3: 一枚硬币重复掷三次, 观察其各面出现的情况;

E_4: 观察总机每天 9:00~10:00 接到的电话次数;

E_5: 从一批电子元件中任取一件测量其使用寿命;

E_6: 观察某地区每天的最高温度与最低温度.

定义 1.1.4 设 E 是一试验, 其所有可能出现的结果组成的集合称为试验 E 的**样本空间**, 记为 Ω(或 S). 样本空间的元素, 也就是随机试验的直接结果称为**样本点**.

前面所举的 6 个试验中, 若以 Ω_i 表示试验 E_i 的样本空间, i=1,2,$\cdots$,6, 则

$\Omega_1=\{H, T\}$, 其中 T 表示出现币值一面, H 表示出现另一面;

$\Omega_2=\{1,2,3,4,5,6\}$;

$\Omega_3=\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$;

$\Omega_4=\{0, 1, 2, \cdots\}$, 样本点个数为可数无限多个;

$\Omega_5=\{t|t\geqslant 0\}$, t 表示所抽取的元件寿命为 t 小时;

$\Omega_6=\{(x, y)|\quad T_1\leqslant x\leqslant y\leqslant T_2\}$, 其中 T_1, T_2 分别是该地区的最低与最高温度.

值得注意的是, 对一个随机现象观测的方式和目的有所不同时, 相应的样本空间也可能有所不同. 如在 E_3 中考察的是币面出现次数时, 样本空间表示为 $\Omega_3=\{0,1,2,3\}$.

定义 1.1.5 随机试验的若干个结果组成的集合称为**随机事件**(简称**事件**), 常用大写字母 $A, B, C, \cdots$ 表示. 只含一个试验结果的事件称为**基本事件**.

即事件是样本空间 Ω 的子集, 基本事件就是 Ω 的样本点.

在试验 E_2 中, A_i=“掷出 i 点” (i=1, 2, $\cdots$, 6), B=“掷出偶数点”, C=“掷出的点数不大于 4”, 则 $A_i=\{i\}(i=1, 2, \cdots, 6)$ 都是基本事件, $B=\{2, 4, 6\}$, $C=\{1, 2, 3, 4\}$.

在试验中, 当事件 (集合) 中的一个样本点 (元素) 出现时, 称这一事件发生. 如试验 E_2 中, 当投掷的结果为 “3 点” 时, 事件 A_3, C 均发生.

由于样本空间 Ω 包含了所有的样本点, 且是 Ω 自身的一个子集, 在每次试验中它总是发生的, 所以称 Ω 为**必然事件**. 空集 $\varnothing$ 不包含任何样本点, 它也是样本空间的一个子集, 且在每次试验中总不发生, 所以称 $\varnothing$ 为**不可能事件**.

1.1.2 事件的关系与运算

对于一个随机试验来说, 有很多随机事件, 有的随机事件可能很复杂. 为了从较简单事件的出现规律中寻求复杂事件出现的规律性, 我们需要研究同一试验的各种事件之间的关系和运算. 由于事件实际上是样本空间中的某一个子集, 因此, 事件之间的关系和运算同集合之间的关系和运算是一致的. 对应着集合的关系和运算, 我们定义事件的关系和运算如下：

设 Ω 是试验 E 的样本空间, A, B, C 及 $A_1, A_2, A_3, \cdots$ 都是事件, 即 Ω 的子集.

1. 事件的包含

如果事件 A 的发生必然导致事件 B 的发生, 则称事件 B**包含**事件 A, 或称事件 A**包含于**事件 B, 记作 $A \subset B$. 显然对任何事件 A, 有 $\varnothing \subset A \subset \Omega$.

2. 事件的相等

$$A = B \Leftrightarrow A \subset B \text{ 且 } B \subset A.$$

3. 事件的并 (和)

“两个事件 A, B 中至少有一个发生” 也是一个事件, 称为事件 A 与 B 的**和**(**并**), 记作 $A \bigcup B$ 或 $A + B$.

“事件 $A_1, A_2, \cdots, A_n$ 至少有一个发生” 称为 $A_1, A_2, \cdots, A_n$ 的和事件, 记为 $\bigcup\limits_{i=1}^{n} A_i$.

“事件 $A_1, A_2, \cdots, A_n, \cdots$ 至少有一个发生” 称为 $A_1, A_2, \cdots, A_n, \cdots$ 的和事件, 记为 $\bigcup\limits_{i=1}^{\infty} A_i$.

4. 事件的交 (积)

“两个事件 A 与 B 同时发生” 是一个事件, 称为事件 A 与 B 的**积 (交)**, 记作 $A \bigcap B$ 或 AB.

“事件 $A_1, A_2, \cdots, A_n$ 同时发生” 称为 $A_1, A_2, \cdots, A_n$ 的积事件, 记为 $\bigcap\limits_{i=1}^{n} A_i$.

“事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生” 称为 $A_1, A_2, \cdots, A_n, \cdots$ 的积事件, 记为 $\bigcap\limits_{i=1}^{\infty} A_i$.

5. 事件的差

“事件 A 发生而事件 B 不发生” 是一个事件, 称为事件 A 与 B 的**差**, 记作 $A - B$.

6. 事件的互斥 (互不相容)

如果事件 A 与事件 B 不能同时发生, 即 $AB = \varnothing$, 称事件 A 与 B**互不相容** **(互斥)**. 显然基本事件间是互不相容的.

$A_1, A_2, \cdots, A_n$ 两两互斥 $\Leftrightarrow A_iA_j = \varnothing, \quad i \neq j,\ i,j = 1,\ 2,\ \cdots,\ n.$

$A_1, A_2, \cdots, A_n, \cdots$ 两两互斥 $\Leftrightarrow A_iA_j = \varnothing, \quad i \neq j,\ i,j = 1,\ 2,\ \cdots.$

7. 事件的对立

"事件 A 不发生" 是一个事件, 称为 A 的**对立事件**(或**逆事件**), 记作 $\overline{A}$. 由于 A 也是 $\overline{A}$ 的对立事件, 因此, 称 A 与 $\overline{A}$ 互为对立事件. 由定义知, 两个对立事件一定是互不相容事件; 反之, 两个互不相容事件却不一定是对立事件. 显然有

$$A\overline{A} = \varnothing, \quad A\bigcup\overline{A} = \Omega, \quad \overline{A} = \Omega - A, \quad \overline{\overline{A}} = A.$$

注意 "A 与 B 互相对立" 与 "A 与 B 互斥" 是不同的概念.

事件的关系与运算可用图 1.1 的文氏图直观地来表示, 其中 $A\bigcup B$, AB, $A-B$, $\overline{A}$ 分别为图中的阴影部分.

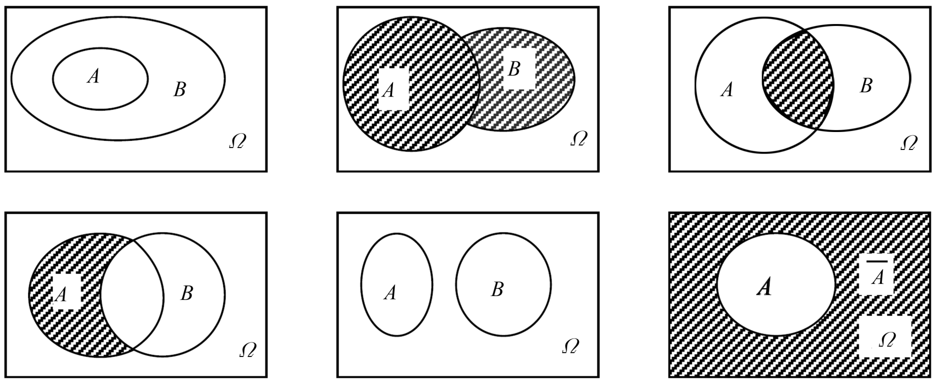

图 1.1 事件的关系与运算图

在进行事件运算时, 要用到如下的事件运算律:

交换律: $A\bigcup B = B\bigcup A$, $AB = BA$;

结合律: $A\bigcup(B\bigcup C) = (A\bigcup B)\bigcup C$, $A(BC) = (AB)C$;

分配律: $A(B\bigcup C) = (AB)\bigcup(AC)$, $A\bigcup(BC) = (A\bigcup B)(A\bigcup C)$;

对偶律 (De Morgan 律): $\overline{A\bigcup B} = \overline{A}\ \overline{B}$, $\overline{AB} = \overline{A}\bigcup\overline{B}$.

对于多个随机事件, 上述运算律也成立. 如 $\overline{A_1\bigcup A_2\bigcup\cdots\bigcup A_n} = \overline{A_1}\ \overline{A_2}\cdots\overline{A_n}$, $\overline{A_1A_2\cdots A_n} = \overline{A_1}\bigcup\overline{A_2}\bigcup\cdots\bigcup\overline{A_n}$ 等.

此外还有

吸收律: $A\bigcup\Omega = \Omega$, $A\Omega = A$, $A\bigcup\varnothing = A$, $A\bigcap\varnothing = \varnothing$;

重余律: $\overline{\overline{A}} = A$;

幂等律: $A\bigcup A = A$, $A\bigcap A = A$;

差化积：$A-B=A\overline{B}=A-AB$.

例 1.1.1　掷两颗骰子, 记事件 A 为 "两颗骰子的点数之和为 5", 事件 B 为 "两颗骰子中最大点数为 3", 求 A, B, $A\cup B$, AB, $A-B$, $B-A$ 的集合表达式.

解　A={(1,4), (4,1), (2,3), (3, 2)};

B={(1, 3), (2, 3), (3, 3), (3, 2), (3, 1)};

$A\bigcup B$ ={(1, 4), (4, 1), (2, 3), (3, 2), (1, 3), (3, 3), (3, 1)};

AB ={ (2, 3), (3, 2) };

$A-B$ ={(1, 4), (4, 1) };

$B-A$ ={(1, 3), (3, 3), (3, 1)}.

例 1.1.2　设 A、B、C 表示三个事件, 试用 A、B、C 表示如下事件：

(1) A 发生且 B、C 至少有一个发生;　(2) A 与 B 发生而 C 不发生;

(3) A、B、C 中至少有两个发生;　(4) A、B、C 中至多有两个发生;

(5) A、B、C 中不多于一个发生;　(6) A、B、C 中恰有一个发生.

解　(1) $A(B\bigcup C)$;　(2) $AB\overline{C}$;　(3) $AB\bigcup AC\bigcup BC$;　(4) A、B、C 中至多有两个发生意味着至少有一个事件不发生, 即 $\overline{A}\bigcup\overline{B}\bigcup\overline{C}=\overline{ABC}$;　(5) $A\,\overline{B}\,\overline{C}\bigcup\overline{A}\,B\,\overline{C}\bigcup\overline{A}\,\overline{B}\,C\bigcup\overline{A}\,\overline{B}\,\overline{C}$;　(6) $A\,\overline{B}\,\overline{C}\bigcup\overline{A}\,B\,\overline{C}\bigcup\overline{A}\,\overline{B}\,C$.

例 1.1.3　在图书馆中随意抽取一本书, 事件 A 表示数学书, B 表示中文书, C 表示平装书, 说明下列各式表示什么事件：$AB\overline{C}$, $\overline{C}\subset B$, $\overline{A}=B$.

解　$AB\overline{C}$—— 抽取的是精装中文版数学书;

$\overline{C}\subset B$—— 精装书都是中文书;

$\overline{A}=B$—— 非数学书都是中文版的, 且中文版的书都是非数学书.

例 1.1.4　化简事件 $\overline{(\overline{A}\,\overline{B}\bigcup C)\overline{A\,C}}$.

解　原式 $=\overline{\overline{A}\,\overline{B}\bigcup C}\bigcup AC=\overline{\overline{A}\,\overline{B}}\,\overline{C}\bigcup AC=(A\bigcup B)\overline{C}\bigcup AC=A\overline{C}\bigcup B\overline{C}\bigcup AC$

$$=A(\overline{C}\bigcup C)\bigcup B\overline{C}=A\Omega\bigcup B\overline{C}=A\bigcup B\overline{C}.$$

练习题 1.1

1. 试写出下列随机试验的样本空间.

(1) 同时抛三颗骰子, 记录三颗骰子点数之和;

(2) 记录一个班级一次考试的平均分数 (以百分制记);

(3) 记录某话务员在一个工作日内接听电话的次数;

(4) 生产产品直到有 10 件正品为止, 记录生产产品的总件数;

(5) 一个口袋中有 5 只外形相同的球, 编号为 1, 2, 3, 4, 5, 从中同时取 3 只球.

2. 将下列事件用事件 A、B、C 表示出来.

(1) A 发生;

(2) 只有 A 发生;

(3) 3 个事件都发生;

(4) 3 个事件中至少有 2 个发生;

(5) 3 个事件中恰好有 2 个发生;

(6) 3 个事件都不发生;

(7) 3 个事件中不多于 2 个发生;

(8) 3 个事件中不多于 1 个发生.

3. 从某班学生中任选一名学生, 设 A ={选出的学生是男生}, B ={选出的学生是数学爱好者}, C={选出的学生是班干部}, 试问下列运算结果分别表示什么事件.

(1) ABC ; (2)$\overline{A}B\overline{C}$; (3)$\overline{A\bigcup C}$; (4)$A-(B\bigcup C)$.

4. 设 $\varOmega=\{1,\ 2,\ \cdots,\ 10\}$, $A=\{2,\ 3,\ 4\}$, $B=\{3,\ 4,\ 5\}$, $C=\{5,\ 6,\ 7\}$, 具体写出下列各式表示的集合.

(1)$\overline{A}\bigcup B$; (2)$\overline{A}B$; (3)$\overline{\overline{A}\ \overline{B}}$; (4)$\overline{A\overline{B\bigcup C}}$; (5)$\overline{A(B\bigcup C)}$.

1.2　事件的概率与性质

在 1.1 节介绍了随机试验, 样本空间和随机事件等基本概念. 在一个随机试验中, 对随机事件发生的可能性大小如何作定量分析研究？以下将要给出的 “概率” 这一概念正是对随机事件发生可能性大小的一种度量. 为此, 先介绍一点预备知识 —— 频率的概念.

1.2.1　频率

定义 1.2.1　设随机事件 A 在 n 次重复试验中发生了 m 次, 则称比值

$$f_n(A)=\frac{m}{n}$$

为事件 A 在 n 次重复试验中发生的**频率**.

由定义, 频率具有如下性质:

(1) 对任何事件 A, $f_n(A)\geqslant 0$(非负性);

(2) $f_n(\varOmega)=1$(归一性);

(3) 若事件 A、B 互不相容, 即 $AB=\varnothing$, 则

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

一般地, 对任意有限多个两两互不相容的事件 $A_1,\ A_2,\ \cdots,\ A_n$, 即 $A_iA_j=\varnothing(i\neq j,\ i,\ j=1,\ 2,\ \cdots,\ n)$, 则有

$$f_n\Big(\bigcup_{i=1}^{n}A_i\Big)=\sum_{i=1}^{n}f_n(A_i)\qquad(\text{有限可加性}).$$

人们经过长期的实践发现, 虽然随机事件在一次试验中可能发生, 也可能不发生, 带有不确定性, 但当重复试验次数 n 充分大时, 随机事件 A 的频率总在一确定的数值附近摆动, 有稳定于确定值的趋势.

频率在一定条件下具有稳定性, 这可以用第 5 章的大数定律加以严格证明, 也可以用试验加以验证. 历史上曾经有一些著名的统计学家进行过抛掷硬币的试验, 其结果如表 1.1 所示.

表 1.1　硬币试验

试验者	投掷次数 n	正面向上次数 m	频率 f_n
蒲丰 (Buffon)	4040	2048	0.5069
皮尔逊 (K.Pearson)	12000	6019	0.5016
皮尔逊 (K.Pearson)	24000	12012	0.5005

模拟试验我们可用 Excel 软件进行仿真.

实验 1.1　抛掷硬币的仿真试验.

具体步骤如下：

第 1 步：进入 Excel 表格界面, 在 Excel 主菜单中依次选择 "工具"、"数据分析";

第 2 步：在 "数据分析" 对话框中选择 "随机数发生器", 单击 "确定";

第 3 步：出现 "随机数发生器" 对话框, 分别输入变量个数 (指试验轮数)、随机数个数 n(试验次数)、分布 (这里选伯努利)、参数 (这里取 0.5)、输出选项, 单击 "确定", 得所产生的 0、1 随机数;

第 4 步：根据随机数是 1 或 0 确定硬币抛掷后出现正面或反面, 随机数求和可得频数 m, 进一步可计算频率 $\dfrac{m}{n}$.

注　如果点击 "工具" 后未出现 "数据分析" 项, 则需点击 "加载宏", 进行该项功能加载.

与物理化学中的实验类似, 在概率统计中模拟仿真的正确使用常可帮助我们验证结论, 建议读者通过上机实验逐步掌握计算机软件在概率统计中的一些重要功能.

从上面的试验记录可以看到, 在多次重复试验中, 事件 "正面向上" 发生的频率虽不完全相同, 但却在一固定的数值 0.5 附近摆动而呈现出一定的稳定性, 频率随着试验次数 n 的增大越来越有接近于数值 0.5 的趋势. 频率的这种稳定性表明一个随机事件发生的可能性有一定的大小可言：频率若稳定于较大的数值, 表明相应事件发生的可能性较大; 频率若稳定于较小的数值, 表明相应事件发生的可能性较小. 而频率所接近的这个固定数值就可作为相应事件发生可能性大小的一个客

观的、定量的度量.

定义 1.2.2 (概率的统计定义) 设随机事件 A 在 n 次重复试验中发生的次数为 m, 当试验次数 n 很大时, 如果频率 $\dfrac{m}{n}$ 稳定在某一数值 p 的附近摆动, 且随作试验次数 n 的增加, 其摆动的幅度越来越小, 则称数值 p 为随机事件 A 的**概率**, 记作 $P(A) = p$.

概率的统计定义有相当直观的试验背景, 易被人们接受. 但定义本身存在很大的缺陷, 即定义中的 “稳定在某一数值 p 的附近摆动” 及 “摆动的幅度越来越小” 等内容模糊不清, 数值 p 无法准确求得, 且在实际问题中, 不可能也不必要对每个事件都要做大量重复试验, 从中得到频率的稳定值. 为了理论研究与实际应用的需要, 由频率的稳定性和频率的性质得到启发, 给出度量随机事件发生可能性大小的概率的公理化定义.

1.2.2 概率的公理化定义

定义 1.2.3 设 Ω 是随机试验 E 的样本空间, 对于 Ω 中的每一个事件 A, 赋于一个满足下列三条公理的实数 $P(A)$:

(1) 非负性: $P(A) \geqslant 0$;

(2) 归一性: $P(\Omega)=1$;

(3) 可列可加性: 若 $A_1, A_2, \cdots, A_n, \cdots$ 两两互不相容, 即 $A_iA_j=\varnothing(i\neq j,\ i, j=1, 2, \cdots)$, 则有

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

称实数 $P(A)$ 为事件 A 的**概率**.

由概率的公理化定义可推得概率的如下性质:

性质 1 不可能事件 $\varnothing$ 的概率为 0, 即 $P(\varnothing) = 0$.

证 因 $\Omega = \Omega \bigcup \varnothing \bigcup \varnothing \bigcup \cdots$, 由定义 1.2.3(3) 可得

$$P(\Omega) = P(\Omega) + P(\varnothing) + P(\varnothing) + \cdots,$$

而 $P(\varnothing) \geqslant 0$, 故由上式知 $P(\varnothing) = 0$.

性质 2 概率具有有限可加性, 即若 $A_iA_j = \varnothing(i \neq j,\ i, j=1, 2, \cdots, n)$, 则

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

证
$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = P(A_1 \bigcup \cdots \bigcup A_n \bigcup \varnothing \bigcup \cdots),$$

由定义 1.2.3(3) 及 $P(\varnothing) = 0$ 得 $P\Big(\bigcup\limits_{i=1}^{n} A_i\Big) = \sum\limits_{i=1}^{n} P(A_i)$.

性质 3 $P(A)=1-P(\overline{A})$.

证 因 $A\bigcup\overline{A}=\Omega$, 且 $A\overline{A}=\varnothing$, 故

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

从而 $P(A)=1-P(\overline{A})$.

性质 4 设 A、B 是两个事件, 若 $A\subset B$, 则有

$$P(B-A)=P(B)-P(A),\qquad P(B)\geqslant P(A).$$

证 因 $A\subset B$, 所以 $B=A\bigcup(B-A)$, 且 $A(B-A)=\varnothing$, 则

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

移项得 $P(B-A)=P(B)-P(A)$. 又因 $P(B-A)\geqslant 0$, 所以 $P(B)\geqslant P(A)$.

性质 5 (加法公式) 对任意事件 A、B, 有

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

证 因为 $A\bigcup B=A\bigcup(B-AB)$, 且 $A(B-AB)=\varnothing$, $AB\subset B$, 故

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

性质 5 还可推广到多个事件的情形. 设 A_1、A_2、A_3 为任意三个事件, 则

$$\begin{aligned}P(A_1\bigcup A_2\bigcup A_3)=&P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)\\&+P(A_1A_2A_3).\end{aligned}$$

一般地, 对任意 n 个事件 $A_1, A_2, \cdots, A_n$ 有

$$\begin{aligned}P\Big(\bigcup_{i=1}^{n}A_i\Big)=&\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)\\&-\cdots+(-1)^{n-1}P\Big(\bigcap_{i=1}^{n}A_i\Big).\end{aligned}$$

例 1.2.1 小王参加某智力游戏, 他能答出甲、乙二类问题的概率分别为 0.7 和 0.2, 两类问题都能答出的概率为 0.1. 求小王 (1) 答出甲类而答不出乙类问题的概率; (2) 至少有一类问题能答出的概率; (3) 两类问题都答不出的概率.

解 事件 A、B 分别表示 “能答出甲、乙类问题”.

(1) $P(A\overline{B})=P(A)-P(AB)=0.7-0.1=0.6$;

(2) $P(A\bigcup B)=P(A)+P(B)-P(AB)=0.8$;

(3) $P(\overline{A}\,\overline{B})=P(\overline{A\bigcup B})=1-P(A\bigcup B)=0.2$.

例 1.2.2 设 A、B 满足 $P(A)=0.6$, $P(B)=0.7$, 在何条件下, $P(AB)$ 取得最大 (小) 值? 最大 (小) 值是多少?

解 由 $P(A\bigcup B)=P(A)+P(B)-P(AB)$, 得

$$P(AB)=P(A)+P(B)-P(A\bigcup B)\geqslant P(A)+P(B)-1=0.3,$$

故最小值在 $P(A\bigcup B)=1$ 时取得.

$$P(AB)\leqslant \min(P(A),P(B))=P(A)=0.6,$$

最大值在$P(A\bigcup B)=P(B)$ 时取得.

思考 若回答当 $A\bigcup B=\Omega$ 时, $P(AB)$ 取得最小值, 是否正确?

1.2.3 概率的古典定义

随机试验的形式多种多样, 我们可根据其不同特点建立不同的数学模型, 从而分类进行研究. 下面先介绍一类最简单的随机试验.

定义 1.2.4 设随机试验 E 具有如下两个特点:

(1) 试验的样本空间 Ω 中只含有限个样本点, 不妨设为 n 个, 并记为 $\omega_1,\omega_2,\cdots,\omega_n$;

(2) 每个样本点出现的可能性 (概率) 相同, 即 $P\{\omega_1\}=P\{\omega_2\}=\cdots=P\{\omega_n\}$.

这类随机试验是概率发展早期的主要研究对象, 称为**古典概型随机试验**, 简称**古典概型**.

下面给出古典概型中事件 A 的概率计算公式. 设随机试验 E 的样本空间为 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$. 由于 E 是古典概型, 故每个样本点出现的概率相同, 即有

$$P\{\omega_1\}=P\{\omega_2\}=\cdots=P\{\omega_n\}.$$

根据概率的定义, 有

$$1=P(\Omega)=P(\{\omega_1\}\bigcup\{\omega_2\}\bigcup\cdots\bigcup\{\omega_n\}).$$

由于$\{\omega_i\}(i=1,\ 2,\ \cdots,\ n)$ 两两互不相容, 因此由概率的有限可加性可得

$$1=P\{\omega_1\}+P\{\omega_2\}+\cdots+P\{\omega_n\},$$

故

$$P\{\omega_i\}=\frac{1}{n},\qquad i=1,\ 2,\ \cdots,\ n.$$

若事件 A 含有 k 个样本点, $A=\{\omega_{j_1},\omega_{j_2},\cdots,\omega_{j_k}\}$, 由于 A 可表示为

$$A=\{\omega_{j_1}\}\bigcup\{\omega_{j_2}\}\bigcup\cdots\bigcup\{\omega_{j_k}\},$$

故由概率的有限可加性得到

$$P(A)=\sum_{m=1}^{k}P\{\omega_{j_m}\}=\frac{k}{n},$$

即

$$P(A)=\frac{A\text{包含的样本点数}}{\Omega\text{中样本点总数}}=\frac{k}{n}.$$

这就是古典概型中随机事件的概率计算公式.

古典概型在概率论中具有相当重要的地位. 一方面由于它简单, 对它的讨论有助于直观理解概率论的许多基本概念; 另一方面, 古典概型的概率计算在产品质量抽样检查等实际问题中以及理论物理问题的研究中, 都有重要的应用. 计算古典概型中事件 A 的概率, 关键是要计算出样本点总数和事件 A 包含的样本点数, 这些数目的计算需要用到排列组合的知识.

以下给出几个古典概型场合概率计算的实际例子.

例 1.2.3(分房模型)　有 n 个不同的粒子, 每个粒子都以同样的概率 $\frac{1}{N}$ 落入 $N(\geqslant n)$ 个格子的每一格子中, 试求下述事件的概率.

(1) $A=\{$指定的 n 个格子中各有一粒子$\}$;

(2) $B=\{$恰有 n 个格子中各有一粒子$\}$.

解　由于每一粒子落入格子的方式都是 N 种, 所以 n 个粒子落入 N 个格子的方式总共有 N^n 种, 此即样本点的总数.

(1) A 所包含的样本点数即为 n 个粒子的全排列 $n!$, 从而

$$P(A)=n!/N^n.$$

(2) 因为 n 个格子可以在 N 个格子中任意选取, 取法有 C_N^n 种, 对选定的 n 个格子, 由 (1) 的讨论有 $n!$ 种落法, 所以 B 包含的样本点数为 $\mathrm{C}_N^n\cdot n!$, 故 B 的概率为

$$P(B)=\mathrm{C}_N^n\cdot n!/N^n.$$

上述例子中, 如将格子解释为粒子的状态空间中的小区域, 就成为统计物理中的模型, 这一模型是由 Maxwell-Boltzmann 提出的. 若将格子看成房子，该模型就称为分房模型. 若题设条件改变, 还可得到其他物理问题的数学模型.

例 1.2.4 (生日问题)　求参加某次集会的 50 人中, (1) 生日全不相同的概率; (2) 至少有两人的生日在同一天的概率.

解　令 $B=\{50$ 人生日全不相同$\}$, $A=\{$至少有两人的生日在同一天$\}$. 我们可把每个人看作上例问题中的粒子, 这里 $n=50$, 而把一年 365 天看作格子, 则 $N=365$, 由上题的解答可知, $P(B)=\mathrm{C}_{365}^{50}\cdot 50!/365^{50}$. 由于事件 A 是事件 B 的对立事件, 因此

$$P(A)=1-P(B)=1-\mathrm{C}_{365}^{50}\cdot 50!/365^{50}\approx 0.97.$$

一般人总认为由于一年 365 天, 50 这个数不怎么大, 50 人中 “至少有两人的生日在同一天” 这件事发生的概率应该比较小, 事实却不是如此, 而是意外之大 ——0.97! 可见, “直觉” 并不很可靠, 科学能纠正不可靠的直觉, 这也有力地说明了研究随机现象统计规律的重要性. 上例是概率论历史上的一个很有名的问题 ——“生日问题”.

实验 1.2　利用 Excel 计算古典概率 (特别是数值较大时).

计算步骤如下：

第 1 步：进入 Excel 表格界面, 将鼠标停留在某一空白单元格 (作为概率值计算结果的输出单元);

第 2 步：在单元格内输入 “= 计算式”, 然后确定即得所求概率.

本例为 “=1−COMBIN(365, 50)× FACT(50)/POWER(365, 50)”.

注　“COMBIN” 表示组合数, “FACT” 表示阶乘, “POWER” 表示某数的乘幂. 所用函数也可直接点击 f_x 进行选择.

例 1.2.5 (抽签原理)　箱中有 a 根红签, b 根白签, 除颜色不同外, 这些签其他方面无区别. 现有 $a+b$ 个人依次不放回地去抽签, 求第 k 个人抽到红签的概率.

记 $A_k=${第 k 个人抽到红签}.

解法 1　把 a 根红签及 b 根白签都看作是不同的 (设想把它们进行编号), 若把抽到的签依次排成一列, 则所有可能的排法相当于把 $a+b$ 个元素进行全排列, 其总数为 $(a+b)!$, 此即样本点的总数. 因为第 k 次抽到红签有 a 种抽法, 而另外 $(a+b-1)$ 次抽签相当于 $a+b-1$ 个元素进行全排列, 有 $(a+b-1)!$ 种抽取方式, 故 A_k 包含的样本点数为 $a\cdot(a+b-1)!$, 从而所求概率为

$$P(A_k)=\frac{a\cdot(a+b-1)!}{(a+b)!}=\frac{a}{a+b},\quad 1\leqslant k\leqslant a+b.$$

解法 2　把 a 根红签看作是没有区别的, 把 b 根白签也看作是没有区别的, 仍然把抽到的签依次排成一列. 因为若把 a 根红签的位置固定下来, 则其他位置必然放白签, 而红签的位置可以有 C_{a+b}^{a} 种放法, 以这种放法作为样本点, 则样本点总数即为 C_{a+b}^{a}. 这时 A_k 包含的样本点数为 C_{a+b-1}^{a-1}, 这是因为由于第 k 次抽得红签, 相应的这个位置必须放红签, 剩下的红签可以在 $a+b-1$ 个位置上任意选取 $a-1$ 个位置, 共有 C_{a+b-1}^{a-1} 种放法. 故所求概率为

$$P(A_k)=\mathrm{C}_{a+b-1}^{a-1}/\mathrm{C}_{a+b}^{a}=\frac{a}{a+b}.$$

两种解法答案是相同的, 其解法不同之处在于选取的样本空间不同. 在第一种解法中把签看作是 “有个性的”(即不同的), 而在第二种解法中则对同色签不加区分. 因此, 在第一种解法中把 $a+b$ 根签的任一全排列作为一个样本点, 而第二种解

法中的每个样本点是第一种解法中相应的 $a!b!$ 个样本点合并而成的. 这个例子告诉我们, 在计算样本点总数及事件包含的样本点数时, 必须对同一个确定的样本空间来考虑, 否则结果就不对了.

既然同一古典概型问题可用不同的样本空间来描述, 因此对同一事件的概率计算也常常有各种不同的方法, 寻求问题的简便解法常常是人们感兴趣的, 本例也还有其他解法.

解法 3 考虑到这是一个不放回抽样, 易知第 k 个人抽到红签这件事仅与前面 $k-1$ 个人抽到签的情况有关, 而与后面的人抽签情况无关. 因此, 要求 A_k 的概率, 只须考虑前 k 个人的抽签情况. 仍如解法 1 把签看作是不同的, 前 k 个人的每一种抽签情况相当于从 $a+b$ 个不同元素中任取 k 个元素的选排列, 故样本点总数为 P_{a+b}^{k}, A_k 包含的样本点数为红签殿后的排列数, 也等于红签打头的排列数, 即等于 $a\cdot\mathrm{P}_{a+b-1}^{k-1}$. 故所求概率为

$$P(A_k)=a\cdot\mathrm{P}_{a+b-1}^{k-1}/\mathrm{P}_{a+b}^{k}=\frac{a}{a+b}.$$

上述结果与 k 无关, 这表明每个人抽到红签的概率与抽签的先后次序无关, 均等于 $\dfrac{a}{a+b}$. 它告诉我们, 体育比赛中进行的抽签分组, 对各队是公平的. 如把签换成阄, 则抓阄游戏等这类活动显然也具有公平合理性. 这些与我们日常的生活经验是一致的.

例 1.2.6 (抽样模型) 某批产品中有 a 件次品, b 件正品, 现分别采用 (1) 有放回抽样和 (2) 不放回抽样方式从中任意抽取 n 件产品, 问正好有 k 件次品的概率各是多少?

解 (1) 有放回抽样情形. 设想把 $a+b$ 件产品进行编号, 由于每次抽取都有 $a+b$ 种可能, 因此从 $a+b$ 件产品中任意抽取 n 件共有 $(a+b)^n$ 种取法, 此即样本点总数. 记 $A=${抽得 n 件产品中恰有 k 件次品}, 则事件 A 包含的样本点数为 $\mathrm{C}_n^k a^k b^{n-k}$, 这是因为若想把抽得的 n 件产品依次排成一列, 每件产品各占据一个位置, 则抽得的 n 件产品的 k 件次品占据这 n 个位置的某 k 个位置, 这种占据方式有 C_n^k 种, 又考虑到每次取后放回, 因此对于每次取得的一件次品都是在 a 件次品中任意抽取, 每件正品也都是在 b 件正品中任意抽取的. 从而事件 A 的概率为

$$P(A)=\mathrm{C}_n^k a^k b^{n-k}/(a+b)^n=\mathrm{C}_n^k\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{n-k},\quad k=\ 0,1,\cdots,n.$$

(2) 不放回抽样情形. 从 $a+b$ 件产品中任意抽取 n 件产品共有 C_{a+b}^n 种取法, 即样本点总数为 C_{a+b}^n, 事件 A 包含的样本点数为 $\mathrm{C}_a^k\cdot\mathrm{C}_b^{n-k}$. 因此所求概率为

$$P(A)=\frac{\mathrm{C}_a^k\mathrm{C}_b^{n-k}}{\mathrm{C}_{a+b}^n}.$$

例 1.2.6 是古典概型概率计算中一个典型问题, 它有着多方面应用, 特别在产品质量检验方面起着重要作用.

1.2.4　概率的几何定义

在实际问题中, 除了前面讨论的属于古典概型的一类随机试验外, 还有许多其他类型的随机试验, 其中一类是：试验的每个可能结果出现的可能性相同, 但所有可能结果却有无穷多个. 例如, 任意向一线段 AB 上投点, 由任意性, 投点落在 AB 上任一点的可能性大小是相同的, 而所有可能结果是线段 AB 上的所有的点为无穷多个. 像这样的随机试验显然已不属于古典概型. 这类试验中事件的概率如何计算呢？现在我们来讨论这个问题.

定义 1.2.5　向某一可度量的区域 G 内投一点, 如果所投的点落在 G 中任意区域 g 内的可能性大小与 g 的度量成正比, 而与 g 的位置和形状无关, 则称这一随机试验为**几何概型试验**, 简称**几何概型**.

上述 "度量" 是指线段长度、可求积平面区域的面积、可求积空间区域的体积等. 几何概型试验的样本点可用 G 中点表示, 因此样本空间就是 $\Omega = G$.

在几何概型试验中事件 $A(A \subset \Omega)$ 的概率与 A 的度量成正比, 即

$$P(A) = K \cdot (A\text{的度量}).$$

再由 $P(\Omega) = K \cdot (\Omega\text{的度量}) = 1$, 知 $K = \dfrac{1}{\Omega\text{的度量}}$, 所以

$$P(A) = \frac{A\text{的度量}}{\Omega\text{的度量}}.$$

例 1.2.7　公共汽车站每隔 5 分钟有一辆公共汽车到站, 乘客到达汽车站的时刻是任意的, 求一个乘客候车时间不超过 3 分钟的概率.

解　令 $A =$\{候车时间不超过 3 分钟\}.

以 x_0 表示该乘客来到车站的时刻, 又假定乘客到达车站后来到的第一辆公共汽车的到站时刻为 a, 则这辆公共汽车的前一辆公共汽车到站时刻为 $a-5$. 由题意知, 乘客是在 $(a-5, a]$ 这段时间来到车站的, 故 $\Omega = \{x{:}a-5 < x \leqslant a\}$, 要乘客候车时间不超过 3 分钟, 必须 $a-3 \leqslant x_0 \leqslant a$, 故 $A = \{x{:}a-3 \leqslant x \leqslant a\}$, 由几何概率得

$$P(A) = \frac{A\text{的度量}}{\Omega\text{的度量}} = \frac{3}{5} = 0.6.$$

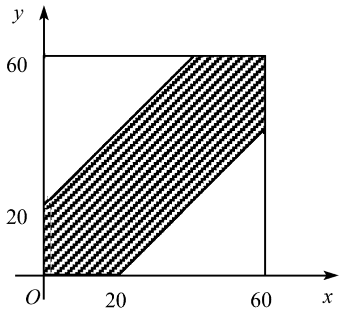

图 1.2

例 1.2.8 (会面问题)　两人约定于中午 12～13 点间在某地会面. 假定每人在这段时间内的每个时刻到达会面地点的可能性是相同的, 且事先约定: 先到者等 20 分钟便可离去. 试求两人能会面的概率.

解　如图 1.2 所示, 用 x, y 分别表示两人到达的时刻 (12 时 x 分, 12 时 y 分), 问题可看作向平面区域

$$\Omega = \{(x,y): 0 \leqslant x \leqslant 60, 0 \leqslant y \leqslant 60\}$$

内投点. 由于两人可分别 "等可能" 的在任何时刻到达, 故问题可看作几何概型试验."两人能会面" 这个事件就是

$$A = \{(x,y): |x-y| \leqslant 20\},$$

因此两人能会面的概率为

$$P(A) = \frac{A\text{的度量}}{\Omega\text{的度量}} = \frac{2000}{3600} = \frac{5}{9}.$$

古典概型和几何概型都强调基本事件的 "等可能" 性, 但在许多情况下常出现基本事件不一定具有相同的概率, 此时就不能直接用上述公式计算概率.

练习题 1.2

1. 已知 $A \subset B$, $P(A) = 0.2$, $P(B) = 0.3$, 求: (1) $P(\overline{A})$; (2) $P(A \bigcup B)$; (3) $P(AB)$; (4)$P(\overline{A}B)$; (5)$P(A-B)$.

2. 有两个电站, 电站 A 正常工作的概率为 0.93, 电站 B 正常工作的概率为 0.92, 两个电站同时工作的概率为 0.898, 求至少有一个电站工作以及只有一个电站工作的概率.

3. 设 A, B 是两个事件且 $P(A) = 0.6, P(B) = 0.7$. 问:

(1) 在什么条件下 $P(AB)$ 取得最大值, 最大值是多少?

(2) 在什么条件下 $P(AB)$ 取得最小值, 最小值是多少?

4. 在房间里有 10 个人, 分别佩戴着从 1 号到 10 号的纪念章, 任意选了 3 人记录其纪念章的号码. (1) 求最小的号码为 5 的概率; (2) 求最大的号码为 5 的概率.

5. 在 1500 个产品中有 300 个次品, 1200 个正品, 任意取出 200 个, 求: (1) 恰有 90 个次品的概率; (2) 至少有 2 个次品的概率.

6. 将 Probability 这 11 个字母分别写在 11 张卡片上, 从中任意连抽 7 张, 求其排列结果为 ability 的概率.

7. 从 5 双不同鞋子中任意取 4 只, 4 只鞋子中至少有 2 只鞋子配成一双的概率是多少?

8. 一副充分洗乱的牌 (含 52 张牌), 求：(1) 任一特定排列的概率; (2) 从中抽取 13 张牌, 所给出的点数都不相同的概率.

9. 在整数 0~9 中任取 4 个数, 能构成 4 位偶数的概率.

10. 某教研室共有 20 名教师, 其中中老年教师 12 人, 年轻教师 8 人, 现要选 4 名优秀教师, 求：(1) 至少有 1 个年轻教师的概率; (2) 有 2 名年轻教师的概率.

11. 某公共汽车线路共有 15 个停车站, 从始发站开车时共有 10 名乘客, 假设这 10 名乘客在各站下车的概率相同, 试求下列事件的概率：

(1) A={10 人各在不同的车站下车};

(2) B={10 人在同一站下车};

(3) C={10 人都在第 3 站下车};

(4) D={10 人中恰有 3 人在终点站下车}.

12. 两艘船都要停泊在同一个码头, 这个码头不能同时停泊两艘船, 它们可能在一个昼夜的任何时刻到达, 设两艘船停靠的时间分别是 1 小时和 2 小时, 求有一艘船要靠位必须等待一段时间的概率.

13. 随机地从区间 $(0,1)$ 取出两个数, 求两数之和大于 3/2 和两数之差绝对值小于 1/4 的概率.

1.3 条件概率

1.3.1 条件概率的定义

在许多实际问题中不仅要计算某一事件 A 的概率, 常常还会遇到求在 "某事件 B 已经发生" 的条件下事件 A 发生的概率问题, 这样的概率称为**条件概率**, 记为 $P(A|B)$. 相对于条件概率 $P(A|B)$ 而言, 把前面所讲的概率 $P(A)$ 称为无条件概率.

由于新增加了条件 "事件 B 已经发生", 所以 $P(A|B)$ 一般与 $P(A)$ 并不相等. 下面来研究如何确定条件概率. 先看一个简单的例子.

例 1.3.1 袋中有 7 只白球, 3 只红球, 白球中有 4 只木球, 3 只塑料球; 红球中有 2 只木球, 1 只塑料球. 现从袋中任取 1 球, 假设每个球被取到的可能性相同. 若已知取到的球是白球, 问它是木球的概率是多少？

解 设 A 表示任取一球, 取得木球; B 表示任取一球, 取得白球, 列表如下.

	白球	红球	小计
木球	4	2	6
塑球	3	1	4
小计	7	3	10

$$P(A|B)=\frac{4}{7}\ \begin{array}{l}\to k_{A\,|\,B}=4=k_{AB}\\ \to n_{\Omega\,|\,B}=7=k_B\end{array}\ =\frac{P(AB)}{P(B)}.$$

从而有

$$P(A|B)=\frac{4}{7}=\frac{k_{AB}}{k_B}=\frac{\dfrac{k_{AB}}{n_\Omega}}{\dfrac{k_B}{n_\Omega}}=\frac{P(AB)}{P(B)}.$$

由此启发我们给出如下定义.

定义 1.3.1 设 A, B 为两事件, 且 $P(B)>0$, 则事件 B 已发生的条件下, 事件 A 发生的**条件概率**$P\ (A|B)$ 定义为

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

可以验证, 条件概率具有概率定义中的三个基本性质:

(1) $P(A|B)=\dfrac{P(AB)}{P(B)}\geqslant 0$;

(2) $P(\Omega|B)=\dfrac{P(\Omega B)}{P(B)}=\dfrac{P(B)}{P(B)}=1$;

(3) 若 $A_1, A_2, \cdots, A_n, \cdots$ 两两互不相容, 则有

$$\begin{aligned}P\Big(\bigcup_{i=1}^{\infty}A_i\,|B\,\Big)&=\frac{P\Big(\bigcup\limits_{i=1}^{\infty}A_i\cdot B\Big)}{P(B)}=\frac{P\Big(\bigcup\limits_{i=1}^{\infty}(A_iB)\Big)}{P(B)}\\&=\frac{\sum\limits_{i=1}^{\infty}P(A_iB)}{P(B)}=\sum_{i=1}^{\infty}\frac{P(A_iB)}{P(B)}=\sum_{i=1}^{\infty}P(A_i\,|B\,).\end{aligned}$$

可见条件概率也是一种概率, 故概率所具有的性质, 条件概率也都具有. 如

$$P\Big(A_1\bigcup A_2|B\Big)=P(A_1|B)+P(A_2|B)-P(A_1A_2|B),$$

$$P(\overline{A}|B)=1-P(A|B),$$

$$P(A_1-A_2|B)=P(A_1|B)-P(A_1A_2|B).$$

当 $B=\Omega$ 时, 条件概率化为无条件概率, 因此 $P(A)$ 可看作 $P\ (A|B)$ 的极端情形.

例 1.3.2 某厂生产的灯泡能用 1000 小时的概率为 0.8, 能用 1500 小时的概率为 0.4 , 求已用 1000 小时的灯泡能用到 1500 小时的概率.

解 令 $A=\{$灯泡能用到 1500 小时$\}$, $B=\{$灯泡能用到 1000 小时$\}$, 所求概率为

$$P(A|B)=\frac{P(AB)}{P(B)}\underset{A\subset B}{\underset{\uparrow}{=}}\frac{P(A)}{P(B)}=\frac{0.4}{0.8}=0.5.$$

1.3.2 乘法公式

由条件概率的定义, 可推得概率的**乘法公式**,

$$P(AB)=P(B)\cdot P(A|B),P(B)>0,\text{或}P(AB)=P(A)\cdot P(B|A),P(A)>0,$$

这个公式在概率计算中有重要作用.

把上述乘法公式推广到任意 n 个事件的积事件的场合, 则有如下公式：

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

$$P(A_1A_2\cdots A_{n-1})>0.$$

利用条件概率的定义很容易验证上式.

例 1.3.3 甲、乙两厂共同生产 1000 个零件, 其中 300 件是乙厂生产的. 而在这 300 个零件中, 有 189 个是标准件. 现从这 1000 个零件中任取一个, (1) 问这个零件是乙厂生产的标准件的概率是多少? (2) 若发现该零件是乙厂生产的, 问该零件是标准件的概率是多少?

解 设 B=$\{$零件是乙厂生产$\}$, $A=\{$零件是标准件$\}$.

(1) $P(AB)=\dfrac{189}{1000}=0.189$;

(2) $P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{0.189}{0.3}=0.63$.

注意 $P(AB)$ 与 $P(A|B)$ 的区别!

例 1.3.4 盒中装有 5 个产品, 其中 3 个一等品, 2 个二等品, 从中不放回地取产品, 每次 1 个, 求：

(1) 取两次, 两次都取得一等品的概率;

(2) 取两次, 第二次取得一等品的概率;

(3) 取三次, 第三次才取得一等品的概率;

(4) 取两次, 已知第二次取得一等品, 求第一次取得的是二等品的概率.

解 令 A_i 为第 i 次取到一等品.

(1) $P(A_1A_2)=P(A_1)P(A_2|A_1)=\dfrac{3}{5}\cdot\dfrac{2}{4}=\dfrac{3}{10}$;

(2) $P(A_2)=P(A_1A_2\bigcup\overline{A}_1A_2)=P(A_1A_2)+P(\overline{A}_1A_2)=\dfrac{3}{5}\cdot\dfrac{2}{4}+\dfrac{2}{5}\cdot\dfrac{3}{4}=\dfrac{3}{5}$, 或直接用古典概率计算得 $P(A_2)=\dfrac{3}{5}$;

(3) $P(\overline{A}_1\overline{A}_2A_3)=P(\overline{A}_1)P(\overline{A}_2|\overline{A}_1)P(A_3|\overline{A}_1\overline{A}_2)=\dfrac{2}{5}\cdot\dfrac{1}{4}\cdot\dfrac{3}{3}=\dfrac{1}{10}$;

(4) $P(\overline{A}_1|A_2)=\dfrac{P(\overline{A}_1A_2)}{P(A_2)}=\dfrac{P(A_2)-P(A_1A_2)}{P(A_2)}=1-\dfrac{\frac{3}{10}}{\frac{3}{5}}=\dfrac{1}{2}$.

例 1.3.5 (Pólya 模型)　一个罐子中包含 b 个白球和 r 个红球. 随机地抽取一个球, 观看颜色后放回罐中, 并且再加进 c 个与所抽出的球具有相同颜色的球. 这种手续进行四次, 试求第一、二次取到白球且第三、四次取到红球的概率.

解　设 W_i={第 i 次取出是白球}, i=1, 2, 3, 4, R_j={第 j 次取出是红球}, j=1, 2, 3, 4, 于是 $W_1W_2R_3R_4$ 表示事件 "连续取四个球, 第一、第二个是白球, 第三、第四个是红球", 用乘法公式容易求出

$$\begin{aligned}P(W_1W_2R_3R_4)&=P(W_1)P(W_2|W_2)P(R_3|W_1W_2)P(R_4|W_1W_2R_3)\\&=\frac{b}{b+r}\cdot\frac{b+c}{b+r+c}\cdot\frac{r}{b+r+2c}\cdot\frac{r+c}{b+r+3c}.\end{aligned}$$

当 $c>0$ 时, 由于每次取出球后会增加下一次也取到同色球的概率. 这是一个传染病模型, 每次发现一个传染病患者, 都会增加再传染的概率.

1.3.3　全概率公式

利用概率的可加性和乘法公式可计算一些较为简单的事件的概率, 但计算比较复杂事件的概率时, 光有这些是不够的. 下面介绍计算复杂事件概率的一种颇为有效的方法, 即全概率公式, 它是将概率的可加性和乘法公式综合所得到的结果. 全概率公式是计算概率的一个非常重要且实用的公式.

定理 1.3.1 (全概率公式)　设 Ω 为随机试验 E 的样本空间, $B_1, B_2, \cdots, B_n$ 为一组事件, 且满足下列条件:

(1) $B_1, B_2, \cdots, B_n$ 两两互斥, 且 $\bigcup\limits_{i=1}^{n}B_i=\Omega$;

(2) $P(B_i)>0,\quad i=1, 2, \cdots, n$.

则对任意的事件 A 有

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A\,|B_i\,).$$

证　由 $\bigcup\limits_{i=1}^{n}B_i=\Omega$, 有

$$A=A\Omega=A\Big(\bigcup_{i=1}^{n}B_i\Big)=\bigcup_{i=1}^{n}(AB_i),$$

且 AB_i 彼此不相容, 由概率的可加性和乘法公式可得

$$P(A)=\sum_{i=1}^{n}P(AB_i)=\sum_{i=1}^{n}P(B_i)P(A|B_i).$$

注 ①满足条件 (1) 的一组事件 $B_1,\ B_2,\ \cdots,\ B_n$ 称为样本空间 Ω 的一个完备事件组 (或划分、分割). ②条件 (1) 也可改为 $B_1,\ B_2,\ \cdots,\ B_n$ 两两互斥, 且 $A\subset\bigcup\limits_{i=1}^{n}B_i$.

例 1.3.6 采购员要购买 10 个一包的电器元件, 他的采购方法是：从一包中随机抽查 3 个, 如这 3 个元件都是好的, 他才买下这一包, 假定含有 4 个次品的包数占 30%, 而其余包中各含一个次品. 求采购员拒绝购买的概率.

解 记 $B_1=$\{取到的是含 4 个次品的包\}, $B_2=$\{取到的是含 1 个次品的包\}, $A=$\{采购员拒绝购买\}, 则 B_1, B_2 构成样本空间的一个完备事件组, 且 $P(B_1)=0.3$, $P(B_2)=0.7$, 又由古典概率计算知

$$P(A|B_1)=1-\frac{\mathrm{C}_6^3}{\mathrm{C}_{10}^3}=\frac{5}{6},\quad P(A|B_2)=1-\frac{\mathrm{C}_9^3}{\mathrm{C}_{10}^3}=\frac{3}{10},$$

从而由全概率公式得

$$P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)=\frac{23}{50}.$$

1.3.4 贝叶斯公式

在实际工作中还常常遇到这样一类问题：已知某个试验结果是由于多种“原因”导致的, 如果人们通过试验确实观察到了这个结果, 于是人们希望通过这个信息来探讨每个“原因”导致这个结果的可能性有多大. 如在病情诊断的问题中, 我们知道某些症状往往是由多种病因导致的, 如果在一次看病中, 某病人有这些症状, 为了确诊病人得的是什么病, 自然, 医生希望知道这种症状由各种病因导致的可能性各有多大. 贝叶斯 (Bayes) 公式正是解决这一计算问题的基本公式.

贝叶斯公式是全概率公式的一个简单变形, 但是由于它被广为应用而十分著名.

定理 1.3.2 (贝叶斯公式) 设 Ω 为随机试验 E 的样本空间, $B_1,\ B_2,\ \cdots,\ B_n$ 为样本空间 Ω 的一个划分, 且 $P(A)>0,\ \ P(B_i)>0,\ i=1,\ 2,\ \cdots,\ n$, 则有

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum\limits_{k=1}^{n}P(B_k)P(A|B_k)},\quad i=1,\ 2,\ \cdots,\ n.$$

证 由条件概率的定义及全概率公式, 得

$$P(B_i|A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum\limits_{k=1}^{n} P(B_k)P(A\,|B_k\,)}, \quad i = 1,\ 2,\ \cdots,\ n.$$

该公式于 1763 年由贝叶斯给出. 它是在观察到事件 A 已发生的条件下, 寻找导致 A 发生的每个原因的概率.

例 1.3.7 某学生把钥匙遗失了, 估计落在宿舍、教室、路上的概率分别为 p_1、p_2、p_3, $(p_1+p_2+p_3 = 1)$, 而在这三种情况下能找到钥匙的概率分别为 α_1、α_2、α_3,

(1) 试求该学生能找到钥匙的概率;

(2) 如果该学生在宿舍里没有找到钥匙, 试问此时他对钥匙落在三个地方的概率可作什么修正 $(0 < p_i < 1,\ 0 < \alpha_i < 1\,,\ i = 1\,,2\,,3)$?

解 (1) 记 B_1、B_2、B_3 分别表示钥匙落在宿舍、教室、路上三个事件, A 为钥匙能找到, 由于 $\{B_1,\ B_2,\ B_3\}$可以构成 Ω 的一个划分, 而且由题设知 $P(B_i) = p_i$, $P(A|B_i) = \alpha_i$, $i = 1\,,\ 2\,,\ 3$. 由全概率公式, 得

$$P(A) = \sum_{i=1}^{3} P(B_i)P(A|B_i) = \sum_{i=1}^{3} p_i\alpha_i.$$

(2) 记 C 表示 “在宿舍里没有找到钥匙”, 则由全概率公式

$$P(C) = \sum_{i=1}^{3} P(B_i)P(C|B_i) = p_1(1-\alpha_1) + p_2 \times 1 + p_3 \times 1 = 1 - p_1\alpha_1.$$

利用贝叶斯公式, 可对钥匙落在三个地方的概率作修正,

$$P(B_1|C) = \frac{P(B_1)P(C|B_1)}{P(C)} = \frac{p_1(1-\alpha_1)}{1-p_1\alpha_1};$$

$$P(B_2|C) = \frac{P(B_2)P(C|B_2)}{P(C)} = \frac{p_2}{1-p_1\alpha_1};$$

$$P(B_3|C) = \frac{P(B_3)P(C|B_3)}{P(C)} = \frac{p_3}{1-p_1\alpha_1}.$$

注意到 $0 < p_1\alpha_1 < 1$, 故 $0 < 1 - p_1\alpha_1 < 1$, 由此得 $0 < p_2 < \dfrac{p_2}{1-p_1\alpha_1}$, 这意味着落在教室里的可能性有所增加, 类似地落选在路上的可能性也会增加, 当然落在宿舍的可能性必有所减少. 这里 $P(B_i)$ 称为先验概率, 它是由以往的经验得到的, 不一定准确. $P(B_i|C)$ 称为后验概率, 它是得到了信息 C 发生, 再对导致 C 发生的原因的可能性大小重新加以修正.

练习题 1.3

1. 已知 $P(A)=\dfrac{1}{4}$, $P(B|A)=\dfrac{1}{3}$, $P(A|B)=\dfrac{1}{2}$, 求 $P(A\bigcup B)$.

2. 已知 $P(\overline{A})=0.3$, $P(B)=0.4$, $P(A\overline{B})=0.5$, 求 $P(B|A\bigcup \overline{B})$.

3. 已知 $P(A)=P(B)=0.4$, $P(AB)=0.28$, 求 $P(A|B)$, $P(A-B)$, $P(A\bigcup B)$.

4. 将两信息分别编码为 A 和 B 传递出去, 接收站收到时, A 被误收作 B 的概率为 0.02, 而 B 被误收作 A 的概率为 0.01. 信息 A 与信息 B 传递的频繁程度为 3:2, 若接收站收到的信息是 A, 问原发信息是 A 的概率是多少?

5. 某商店出售尚未过关的某电子产品, 进货 10 件, 其中有 3 件是次品, 已经出售 2 件, 现从剩下的 8 件产品中任取一件, 求这件是正品的概率.

6. 某工厂有甲、乙、丙三个车间, 它们生产同一种产品, 每个车间产量分别占该工厂总产量的 25%, 35%, 40%, 每个车间的产品中次品的概率分别为 0.05, 0.04, 0.03. 现从该厂总产品中任取一件产品, 结果是次品, 求取出的这件产品是乙车间生产的概率.

7. 某人忘记了电话号码的最后一个数字, 因而随意地拨号, 求他拨号不超过 3 次而接通所需的电话的概率是多少? 如果已知最后一个数字是奇数, 那么此概率是多少?

8. 某商品的商标为 “MAXAM”, 其中有 2 个字母脱落, 有人捡起随意放回. 求放回后仍为 “MAXAM” 的概率.

1.4　事件的独立性

对于任意两个事件 A、B, 若 $P(B)>0$, 则 $P(A|B)$ 有定义. 此时可能有两种情形: $P(A|B)=P(A)$ 和 $P(A|B)\neq P(A)$. 后者说明事件 B 的发生对事件 A 发生的概率有影响, 只是当 $P(A|B)=P(A)$ 时才认为这种影响不存在, 这时自然认为事件 A 不依赖于事件 B, 即 A、B 是彼此独立的. 此时有

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

由此引出关于事件独立性的讨论. 事件独立性是概率论中, 除去概率的概念之外最重要的概念.

定义 1.4.1　设有事件 A、B, 若满足

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

则称事件 A、B 是**相互独立**的.

由上述定义可知必然事件 Ω 和不可能事件 $\varnothing$ 与任何事件都相互独立.

定理 1.4.1　若事件 A、B 相互独立, 则 $P(A|B)=P(A)$.

由条件概率的定义及两个事件的独立性的定义很容易证明该定理. 这一定理表明, 若事件 A、B 相互独立, 则事件 A 关于事件 B 的条件概率等于无