高等教育公共基础课精品系列规划教材

基于 R 的概率论与数理统计

主编 / 欧诗德

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

基于R的概率论与数理统计

主　编　欧诗德

副主编　韦盛学　黄科登

参　编　易亚利　梁　丹　罗中德(百色学院)

北京理工大学出版社

BEIJING INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书主要介绍概率论与数理统计、R统计软件的使用基础，内容包括：随机事件及其概率、随机变量的分布、随机向量及其分布、随机变量的数字特征、大数定律与中心极限定理、样本及其分布、参数估计、假设检验、R统计软件简介以及实验，重点讲明概率论与数理统计的基本概念、定理、性质和计算方法.

本书列举了较多在工程技术、经济管理、社会生活等各方面的实际问题，针对有关计算问题，介绍了如何运用R统计软件解决. 通过对本书的学习，读者可以掌握概率论与数理统计的基本理论和计算方法，学会运用R统计软件解决本专业遇到的有关计算问题.

本书可以作为高等院校相关学科专业的教材.

图书在版编目(CIP)数据

基于R的概率论与数理统计 / 欧诗德主编. —北京：北京理工大学出版社，2019.8
ISBN 978-7-5682-7445-6

Ⅰ.①基…　Ⅱ.①欧…　Ⅲ.①概率论 ②数理统计　Ⅳ.①O21

中国版本图书馆CIP数据核字(2019)第176701号

出版发行 / 北京理工大学出版社有限责任公司
社　　址 / 北京市海淀区中关村南大街5号
邮　　编 / 100081
电　　话 / (010)68914775(总编室)
(010)82562903(教材售后服务热线)
(010)68948351(其他图书服务热线)
网　　址 / http://www.bitpress.com.cn
经　　销 / 全国各地新华书店
印　　刷 / 三河市天利华印刷装订有限公司
开　　本 / 787毫米×1092毫米　1/16
印　　张 / 10
字　　数 / 236千字
版　　次 / 2019年8月第1版　2019年8月第1次印刷
定　　价 / 32.00元

责任编辑 / 多海鹏
文案编辑 / 孟祥雪
责任校对 / 周瑞红
责任印制 / 李志强

图书出现印装质量问题，请拨打售后服务热线，本社负责调换

前　言

概率论与数理统计是高等院校经济、农、医等学科的专业基础课程，是近代数学的重要组成部分，具有极强的应用背景，与其他学科有紧密的联系，广泛应用于工业、农业、军事和科学技术中．随着计算机科学和互联网的迅猛发展，概率论与数理统计理论方法在大数据分析、人工智能领域发挥着极其重要的作用．

随着创新驱动发展、中国制造 2025 等国家重大战略的实施，我国经济结构发生了深刻调整，产业升级加快了步伐，社会对人才培养质量提出了更高的要求，地方行业企业更需要应用型人才．为顺应这一发展趋势，培养高质量的应用型人才，我们编写了本书．通过对本书的学习，学生既能掌握概率论与数理统计的基本理论方法，又能培养数据处理能力．

本书除了讲解概率统计基本理论方法之外，还重视讲授应用 R 统计软件解决计算问题．本门课程，尤其是数理统计部分，涉及的计算量较大，如果只讲明计算法，而不介绍如何使用计算机解决该问题，相当于一个劳动者只了解机器工作原理，而不会操作机器．本书结合 R 统计软件来讲授，就是为了使学生在面对大量数据时能应用计算机解决实际问题．本书并非重点介绍 R 统计软件，只是向学生介绍了学习 R 统计软件的入门知识，以使各专业的学生都能学到运用计算机计算有关概率统计问题的方法．因此，在学习本课程时不学习统计软件知识是很难解决有关实际计算问题的．

本书主要内容包括：随机事件及其概率、随机变量的分布、随机向量及其分布、随机变量的数字特征、大数定律与中心极限定理、样本及其分布、参数估计、假设检验、R 统计软件简介以及实验，共十章．整个编写工作分配如下：第 1、2 章由黄科登编写，第 3 章由易亚利编写，第 4、6 章由韦盛学编写，第 5 章由梁丹编写，第 7、8、9、10 章由欧诗德、罗中德（百色学院）编写，欧诗德负责全书的修改和复查工作．

本课程可在 48 课时内讲完，其中包括 12 课时的上机实验课，数理统计方面的内容可以与上机实验课一起进行．每一章后提供了适量的习题，作业可从习题中选择．为便于学生期末考试复习，理解基本理论方法、性质、概念，本书提供了客观题．

本书的编写得到了我校自编教材建设重点项目（编号：14XJJC002）的资助，对此对学校表示诚挚的谢意．另外，还感谢本教研室的教师潘伟权、陈泊伶、陈丽君、吴荣火提出的修改意见！由于编者水平有限，书中难免有疏漏之处，敬请读者批评指正！

编　者

目　录

第 1 章　随机事件及其概率

纷繁复杂的客观世界中的各类现象可分为两大类:随机现象和确定性现象,也称偶然现象和必然现象.人们用代数、几何、函数理论等这些数学工具对确定性现象进行研究,并掌握了这类现象的许多规律.然而,对随机现象,没有必然规律可言,人们很难用一个确定的公式来描述其变化特征.例如,当我们掷骰子时无法预知朝上的一面出现的点数;在证券市场中,投资者无法预知股票在下一交易日的涨跌;人们无法知道人的寿命长短.但是随机现象也蕴含着规律,人们可通过大量的重复试验和观察发现随机现象的规律.例如,掷骰子,经反复试验我们可发现,各点数出现的频率接近 1/6.投资者根据股票的 K 线图及其各种指标,经过大量的观察发现下一交易日涨跌的可能性.总的来说,随机现象的规律性是在大量重复试验和观察中呈现出来的,我们称之为随机现象的统计规律性.概率论与数理统计就是研究和提示随机现象统计规律的数学学科.

1.1　随机事件及其运算

1.1.1　随机现象

随机现象是指在一定条件下会出现多种可能结果的客观现象.随机现象广泛存在于人们的工作与生活之中.

【例 1.1】

(1)在离光滑桌面上一定高度抛掷一枚均匀的硬币,观察朝上的一面.

(2)在离光滑桌面上一定高度抛掷一颗均匀的骰子,观察出现的点数.

(3)图 1-1 所示为我国 A 股某只股票在某一交易日的走势图.在 10%的涨跌停板制度下,观察该只股票下一交易日的收盘价.

(4)观察某种品牌的同一款手机的寿命时数.

(5)观察测量某物理量(长度、面积、体积)的误差.

我们发现这些现象具有以下两个特点:

(1)会出现多种可能结果.

(2)每一次观察(或试验)会出现哪一种结果,人们事先无法准确知道.

这两个特点就是随机现象的特征.与随机现象相对应的是确定性现象,确定性现象是指在一定条件下,结果只有一个的客观现象.例如,在标准大气压下,把一壶水加热到 100 ℃,结果只有一个,就是沸腾;从一个装有 100 只白色乒乓球的袋子中随机取出一只,结果只有一个,就是一只白色的乒乓球.

在随机问题研究中,为研究其规律性,我们经常要做试验.在相同条件下,可重复随机现象的观察、记录、实验统称为随机试验,简称试验.此外,客观世界也存在不可重复的随机现象,例

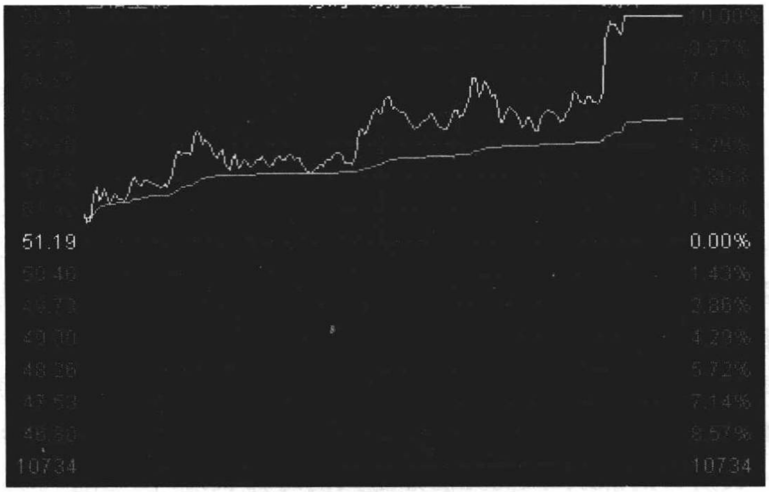

图 1-1　股票分时走势图

如一年一度的欧洲冠军杯决赛的输赢，以及经济领域中的失业现象和某国的 GDP 增长速度等，均为不可重复的随机现象. 概率论与数理统计主要研究可以重复的随机现象的统计规律，但也重视不可以重复的随机现象的研究.

1.1.2　样本空间

研究随机现象时发现存在这样一种事实：任何一种试验无论有多少种可能结果，总可以从中找出一组基本结果，并且：

(1)每进行一次重复试验，必然出现且仅出现其中的一个基本结果；

(2)任何可能结果，都由其中的若干个基本结果组成.

试验的所有基本可能结果组成的集合称为样本空间，记为 Ω. 样本空间的元素称为样本点，一般情况下记为 ω.

样本空间 Ω 本质上是集合，对 Ω 的描述方法有两种：

(1)代表元法.

$\Omega=\{\omega_i \mid \omega_i$ 表示试验的第 i 种的那一种基本可能结果，$i=1,2,\cdots\}$.

(2)枚举(区间)法. 读者可以通过以下例子加以理解.

【例 1.2】　写出“例 1.1”所列 5 种随机现象对应的样本空间 .

解　首先写出样本空间的代表元形式，再写列举(区间)形式如下：

(1)$\Omega_1=\{i \mid i$ 表示掷出结果为 i，$i=0,1$；0 表正面，1 表反面$\}=\{0, 1\}$；

(2)$\Omega_2=\{i \mid i$ 表示掷出 i 点，$i=1,2,3,4,5,6\}=\{1,2,3,4,5,6\}$；

(3)$\Omega_3=\{i \mid i$ 表示该股票下一交易日盘价为 i 元，$i=50.68,50.69,\cdots,61.94\}=\{50.68,50.69,\cdots,61.94\}$；

(4)$\Omega_4=\{t \mid t$ 表示寿命时数为 t，$t\geqslant 0\}=[0,+\infty)$；

(5)$\Omega_5=\{x \mid x$ 表示测量误差为 x，$-\infty<x<+\infty\}=(-\infty,+\infty)$.

关于样本空间的分类：

(1)从含有样本点的个数划分，样本空间可分为有限样本空间与无限样本空间；

(2)从含有样本点是否可数划分，样本空间可分为离散样本空间与连续样本空间.

样本空间中的元素可以是数也可以不是数，样本空间含有的样本点至少有两个，仅含有两个样本点的样本空间是最简单的样本空间.

1.1.3 随机事件

定义与记号：样本空间 Ω 的子集称为随机事件，这是基于样本空间给出的定义. 也可以基于随机现象给出定义，每一个可能的结果称为随机事件，简称为事件，通常用大写英文字母 A，B,C 作为事件的记号.

Venn 图：用图形表示事件，如图 1-2 所示.

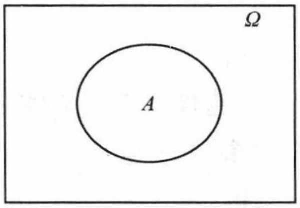

图 1-2　Venn 图

【例 1.2】中(2)的样本空间为：$\Omega_2=\{1,2,3,4,5,6\}$，若要表示事件“掷出奇数点”，则可记为

$$A=\text{“掷出奇数点”，或 } A=\{1,3,5\}.$$

显然，事件 A 为 Ω_2 的子集.

事件定义的进一步解读：

(1)“事件 A 发生”意味着在试验中 A 包含的某个样本点出现了. 反之亦然.

【例 1.2】(2)中，事件 $A=$“掷出奇数点”发生，这表明在一次抛掷中掷出了 1 点或 3 点或 5 点.

(2)事件的描述方法有三种：

方法 1：集合表示法；

方法 2：用明白无误语言(加以双引号)表述法；

方法 3：用随机变量取值表示法.(详见第 2 章)

(3)三种特殊(随机)事件.

基本事件，Ω 的单元素子集；必然事件，Ω 本身(每次试验必然发生的事件)；

不可能事件，记作 Φ (每次试验都不发生的事件).

【例 1.3】　对应于“【例 1.2】(2)”中，若记 $A_i=$“掷出 i 点”，$i=1,2,\cdots,6$. 则

A_1,A_2,A_3,A_4,A_5,A_6 均为基本事件；

记 $B=$“掷出偶数点”，则 B 为随机事件；

记 $C=$“掷出点数小于 7”，则 $C=\Omega_2$ 为必然事件；

记 $D=$“掷出点数大于 6”，则 $D=\Phi$ 为不可能事件.

1.1.4 事件间的关系与运算

1. 事件包含关系

若随机事件 A 发生必然导致随机事件 B 发生，则称事件 A 包含于 B，也称事件 B 包含 A. 记作 $A\subset B$ 或 $B\supset A$.(这是用概率论语言对事件的包含关系的描述)

注：

(1)$A\subset B$ 等价以下定义：

若随机事件 B 不发生，则随机事件 A 必不发生，称随机事件 A 包含于 B.

(2)$A\subset B$ 等价以下定义：(集合论语言对 $A\subset B$ 的描述)

若随机事件 A 包含的样本点全属于随机事件 B，则称随机事件 A 包含于 B.

例如：对于试验 E_2 样本空间为：$\Omega_2=\{1,2,3,4,5,6\}$，若记

$$A=\text{“掷出 4 点”}，B=\text{“掷出偶数点”}，$$

则 $A \subset B$.

2. 事件的等价关系

若随机事件 A 与 B 互相包含，则称事件 A 与 B 具有等价关系，也称随机事件 A 与 B 等价，记为 $A=B$.

对任意事件 A，约定 $\Phi \subset A$，而显然有 $A \subset \Omega$.

例如：【例 1.2】中(2)的样本空间为：$\Omega_2=\{1,2,3,4,5,6\}$，若记

$$A=\text{“掷出非奇数点”}, B=\text{“掷出偶数点”},$$

则 $A=B$.

3. 事件 A 与 B 的并运算

“事件 A 与 B 至少有一个发生”这一事件称为事件 A 与 B 的并(和)，记为 $A \cup B$.

注：

(1)用集合论语言对事件 A 与 B 并的描述：事件 A 与 B 所包含的样本点的并集对应的事件称为事件 A 与 B 的并. 易见，对任意事件 A，有 $A \cup \Omega=\Omega$，$A \cup \Phi=A$.

(2)理解好 $\bigcup\limits_{i=1}^{n} A_i$ 和 $\bigcup\limits_{i=1}^{+\infty} A_i$ 的含义.

4. 事件 A 与 B 的交运算

若事件 A 与 B 同时发生，则称其为事件 A 与 B 的交(积)，记为 $A \cap B$，或 AB.

注：

(1)用集合论语言对事件 A 与 B 的交的描述.

事件 A 与 B 所包含的样本点的交集对应的事件称为事件 A 与 B 的交. 易见，对任意事件 A，有 $A \cap \Omega=A$，$A \cap \Phi=\Phi$.

(2)理解好 $\bigcap\limits_{i=1}^{n} A_i$ 和 $\bigcap\limits_{i=1}^{+\infty} A_i$ 的含义.

5. 事件 A 对 B 的差

若事件 A 发生而事件 B 不发生，则称其为事件 A 对 B 的差，记为 $A-B$.

注：

(1)用集合论语言对事件 A 对 B 的差的描述：从事件 A 的样本点中去除属于事件 B 的样本点后所剩的样本点集对应的事件称为事件 A 对 B 的差.

(2)类似可定义 $B-A$. 而对任意事件 A，有

$$A-A=\Phi, A-\Phi=A, A-\Omega=\Phi.$$

6. 事件的互不相容关系

若随机事件 A 与 B 不可能同时发生，则称随机事件 A 与 B 具有互不相容关系，也称事件 A 与 B 互斥. 记作 $AB=\Phi$.

注：

(1)基本事件之间的关系为互斥关系.

(2)理解好 $\bigcap\limits_{i=1}^{n} A_i=\Phi$ 和 $\bigcap\limits_{i=1}^{+\infty} A_i=\Phi$ 的含义.

7. 事件的对立关系

若随机事件 A 与 B 满足 $A \cup B=\Omega$，且 $AB=\Phi$，则称事件 A 与 B 具有对立关系，也称随机事件 A 与 B 互为对立事件(逆事件).

注：

(1)概率论中记随机事件 A 的对立事件为 $\bar{A}$. 易见，若随机事件 A 与 B 对立，则 $\bar{A}=B$, $\bar{B}=A$.

(2)事件 $\bar{A}$ 是由不属于事件 A 的样本点组成的事件，即"A 不发生"这一事件就是 $\bar{A}$. 显然，若随机事件 A 与 B 对立，则 $B=\Omega-A$(体现出"对立关系"是一种运算——余运算).

(3)每一次试验中，事件 A 与 $\bar{A}$ 有且只有一个发生，所以在一次试验中，若 A 发生，则 $\bar{A}$ 不发生；若 A 不发生，则 $\bar{A}$ 发生；等等.

(4)由定义得：$\bar{\bar{A}}=A$, $\bar{\Omega}=\Phi$, $\bar{\Phi}=\Omega$, $A-B=A\bar{B}=A-AB$, $B-A=B\bar{A}=B-AB$.

(5)由定义容易明白：若随机事件 A 与 B 对立，则 A 与 B 互斥，反之不然. 反例可轻易举出.

事件的关系与运算可用 Venn 图直观表示，如图 1-3 所示.

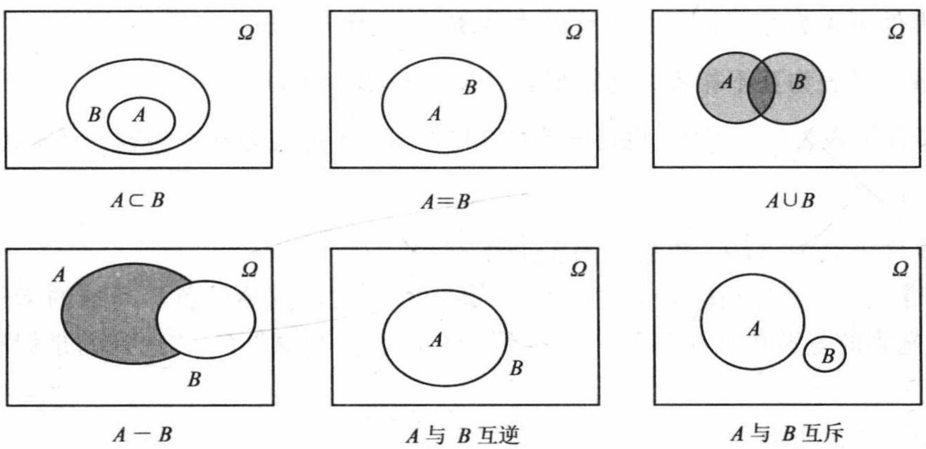

图 1-3 Venn 图

随机事件运算满足以下运算律：

(1)交换律 $A\cup B=B\cup A$, $AB=BA$;

(2)结合律 $(A\cup B)\cup C=A\cup(B\cup C)$, $(AB)C=A(BC)$;

(3)分配律 $(A\cup B)C=AC\cup BC$, $(AB)\cup C=(A\cup C)(B\cup C)$;

(4)对偶律(德莫根公式)

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\bar{A}_i,\quad \overline{\bigcup_{i=1}^{+\infty}A_i}=\bigcap_{i=1}^{+\infty}\bar{A}_i,$$

$$\overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\bar{A}_i,\quad \overline{\bigcap_{i=1}^{+\infty}A_i}=\bigcup_{i=1}^{+\infty}\bar{A}_i.$$

特别地，两个事件运算时的对偶律为

$$\overline{A\cup B}=\bar{A}\,\bar{B},\overline{AB}=\bar{A}\cup\bar{B}.$$

【例 1.4】 设 A、B、C 为同一个 Ω 下的三个事件，请用 A、B、C 及它们的对立事件的运算式子表示以下事件：

(1)"A 发生而 B、C 都不发生"；

(2)"A、B 都发生而 C 不发生"；

(3)"A、B、C 至少有一个发生"；

(4)“A、B、C 至少有两个发生”；

(5)“A、B、C 恰有两个发生”；

(6)“A、B、C 恰有一个发生”；

(7)“A、B 至少有一个发生而 C 不发生”；

(8)“A、B、C 不全发生”.

解　(1)该事件可表为 $A\bar{B}\bar{C}$，也可表为 $A-B-C$，还可表为 $A-(B\cup C)$；

(2)该事件可表为 $AB\bar{C}$，也可表为 $AB-C$；

(3)该事件可表为相容并运算形式 $A\cup B\cup C$，也可表为互不相容并运算形$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC$；

(4)该事件可表为相容并运算形式 $AB\cup BC\cup AC$，也可表为互不相容并运算形式 $AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC$；

(5)该事件可表为互不相容并运算形式 $AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC$；

(6)该事件可表为互不相容并运算形式 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$；

(7)该事件可表为$(A\cup B)\bar{C}$，也可表为$(A\cup B)-C$，还可表为互不相容并运算形式 $AB\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}B\bar{C}$；

(8)该事件可表为 $\bar{A}\bar{B}\bar{C}$，也可表为$\overline{A\cup B\cup C}$.

【例 1.5】　在装着分别标有 1,2,…,100 共 100 个黄白两种颜色乒乓球的袋子中任选一个，若 A=“选到的是黄色乒乓球”，B=“选到的是白色乒乓球”，C=“选到的是标有偶数的乒乓球”.

(1)叙述 $A\bar{C}$ 的含义.

(2)什么条件下有 $BC=C$ 成立？

(3)什么条件下有 $\bar{B}\subset C$ 成立？

解　(1)事件 $A\bar{C}$ 的含义是“选到的是标有奇数的黄色乒乓球”.

(2)由于当 $C\subset B$ 时才有 $BC=C$，因此只有当袋子中标有偶数的乒乓球全是白色乒乓球时，才有 $BC=C$ 成立.

(3)当袋子中的黄色乒乓球都标的是偶数时，$\bar{B}$ 发生就一定导致 C 发生，即 $\bar{B}\subset C$ 成立.

【例 1.6】　记 A=“甲股票上涨，或乙股票下跌”，B=“甲股票上涨”，C=“乙股票上涨或平盘”. 问：$\bar{A}$ 的含义是什么，如何表示？

解　$A=B\cup\bar{C}$，于是

$$\bar{A}=\overline{B\cup\bar{C}}=\bar{B}C,$$

可见 $\bar{A}$ 的含义为“甲股票下跌或平盘，且乙股票上涨或平盘”.

1.2　概率的定义与性质

概率的定义及计算是概率论中最基本的问题. 概率的定义直观通俗的表述是随机事件发

生的可能性大小的度量值. 以下依据概率论发展的历史分别给出讨论,介绍概率的严谨数学定义及计算.

1.2.1　频率

描述随机事件发生的频繁程度的频率是理解概率的基础.

定义 1.1　设在相同条件下,试验重复 n 次. 若随机事件 A 在 n 次试验中发生 k 次,则称比值 $\frac{k}{n}$ 为随机事件 A 在这 n 次试验中发生的频率,记为 $f_n(A)=\frac{k}{n}$.

频率的三个性质　由定义 1.1 容易推知,频率具有以下性质:

(1)(非负有界性)对任意事件 A,有

$$0\leqslant f_n(A)\leqslant 1;$$

(2)(正则性)对必然事件 Ω 有

$$f_n(\Omega)=1;$$

(3)(有限可加性)若 $A_1,A_2,\cdots,A_n$ 为同一样本空间下的互斥事件列,则

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

基于频率导出概率的统计定义:

随机事件 A 发生的频率 $f_n(A)$ 表示随机事件 A 发生的频繁程度,频率越大,随机事件 A 发生就越频繁,换句话说,在一次试验中,随机事件 A 发生的可能性也就越大;反之亦然. 因此,直观的思想就是用频率 $f_n(A)$ 表示随机事件 A 在一次试验中发生的可能性大小. 但是,由于试验的随机性,因此即使是同样进行 n 次试验的两批次试验所得到的 $f_n(A)$ 值也不一定相同. 然而大量试验结果证实,当 n 足够大时,$f_n(A)$ 会稳定在某个常数附近. 这个客观事实,说明描述随机事件 A 发生的可能性大小的数,也就是随机事件 A 发生的概率是客观存在的,它就是 $f_n(A)$ 的稳定值.

关于上面的事实,统计学家做了抛硬币试验. 掷一枚均匀硬币,出现正面的概率的确定.

历史上有三位概率研究学者:德 · 摩根(De Morgan)、蒲丰(Buffon)和皮尔逊(Pearson),他们曾分别进行了大量的掷硬币试验,所得结果如表 1-1 所示.

表 1-1　抛硬币试验

试验者	掷硬币次数	出现正面次数	出现正面的频率
德 · 摩根	2 048	1 061	0.518 1
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

从表 1-1 中的数据可见,出现正面的频率随着试验次数的增大稳定在 0.5 附近,这个 0.5 就反映了出现正面的可能性大小,即事件“出现正面”的概率为 0.5.

定义 1.2　设随机事件 A 在 n 次重复试验中发生的次数为 k,当 n 很大时,频率 $\frac{k}{n}$ 在某一常数 p 附近波动,而随着试验次数 n 的不断增加,发生较大波动的可能性越来越小,则称常数 p 为事件 A 的概率,记为 $P(A)=p$,其中 p 为频率 $f_n(A)$ 的稳定值.

通过频率 $f_n(A)$ 的稳定值 p 来确定事件 A 的概率 $P(A)=p$，这种确定事件概率的方法显然是有局限性的，同时其结果也是粗糙的、不精确的. 但其基本思想是具有很高理论价值的，有学者从中得到启发，引入可用于建构概率论理论体系的概率的公理化定义. 概率的公理化定义于后给出.

1.2.2 古典概型

定义 1.3 古典概率模型简称古典概型，也称等可能概型，或等概率模型. 它是指满足以下两个条件的随机试验模型：

(1)试验模型的样本空间 Ω 包含的样本点总数有限；

(2)试验模型的样本空间 Ω 包含的每一个样本点，即基本事件的出现具有等可能性.

定义 1.4 设 A 为古典概率模型中的随机事件，且 A 包含了该古典概率模型的样本空间 Ω 中全部 n 个样本点中的 k 个，则 A 包含的样本点数与 Ω 包含的样本点数的比值 $\frac{k}{n}$ 称为事件 A 的概率. 这是古典概率模型中事件的概率，在概率论中又常称为古典概率. 记为

$$P(A)=\frac{k}{n}.$$

【例 1.7】 同掷两枚均匀的硬币，求“掷出一个正面一个反面”的概率.

解 设 $A=$“掷出一个正面一个反面”. 易见本问题概率模型的样本空间为

$$\Omega=\{(正,正),(正,反),(反,正),(反,反)\}.$$

该样本空间 Ω 包含的样本点总数 $n=4$(有限)，又由于硬币是均匀的，可判定该 Ω 中的每一个样本点具有等可能性，因而此模型为古典概率模型. 而易见事件 A 包含 Ω 中的(正,反)和(反,正)这两个样本点，即事件 A 包含的样本点数 $k=2$. 于是

$$P(A)=\frac{k}{n}=\frac{2}{4}=0.5.$$

注：对本例，若认为其样本空间为

$$\Omega=\{(二个正面),(二个反面),(一正一反)\},$$

则因其中的样本点不具有等可能性，从而不能用古典概率方法求 $P(A)$. 此外提醒学习者注意解答过程，在确认古典概率模型时，熟练后可写得简洁些.

【例 1.8】 掷一枚均匀的硬币 3 次，求：

(1)恰有 1 次出现正面的概率；

(2)至少有 2 次出现正面的概率.

解 记 $A=$“恰有一次出现正面”，$B=$“至少有 2 次出现正面”.

该问题的随机试验模型为古典概型，且其样本空间 Ω 包含的样本点总数 $n=2^3$. 则

(1)由于事件 A 包含的样本点数为 $k_1=3$，故所求概率为

$$P(A)=\frac{k_1}{n}=\frac{3}{2^3}=\frac{3}{8}.$$

(2)由于事件 B 包含的样本点数为 $k_2=C_3^3+C_3^2=4$，故所求概率为

$$P(A)=\frac{k_2}{n}=\frac{4}{2^3}=0.5.$$

计算古典概率需注意：

(1)公式使用的前提：涉及的概率模型必须是古典概率模型.

(2)公式使用的关键：利用计数、排列、组合，或乘法原理、加法原理等，先计算出古典概率模型样本空间 Ω 包含的样本点总数 n，以及随机事件 A 包含的样本点数 k.

步骤1　在题目没有给出所求概率的事件记号时先明确事件记号(如 A)及其含义，确认问题模型为古典概率模型.

步骤2　用计数、排列、组合，或乘法原理、加法原理等，计算该古典概率模型样本空间 Ω 包含样本点的总数 n，及事件 A 包含的样本点数 k.

步骤3　代入公式

$$P(A)=\frac{k}{n}$$

计算出所求概率 $P(A)$.

【例1.9】 箱中有 a 个白球和 b 个黑球，现无放回抽取，每次取一个.

(1)任取 $m+n$ 个，求恰有 m 个白球，n 个黑球的概率；

(2)求第 k 次才取到白球的概率；

(3)求第 k 次恰好取到白球的概率.

解　考虑将 a 个白球分别标上1到 a 的数字，将 b 个黑球分别标上 $a+1$ 到 $a+b$ 的数字加以区别，从中做无放回随机一个一个抽取，共 $m+n$ 个，经过判断知，这种抽取试验模型为古典概型，且其样本空间包含的样本点总数为

$$n^*=\mathrm{A}_{a+b}^{m+n}.$$

于是

(1)设 A_1="任取 $m+n$ 个，恰有 m 个白球，n 个黑球"，则 A_1 包含的样本点数为 $k_1=\mathrm{C}_{m+n}^m\mathrm{A}_a^m\mathrm{C}_n^n\mathrm{A}_b^n$，所以

$$P(A_1)=\frac{k_1}{n^*}=\frac{\mathrm{C}_{m+n}^m\mathrm{A}_a^m\mathrm{A}_b^n}{\mathrm{A}_{a+b}^{m+n}}.$$

(2)设 A_2="第 k 次才取到白球"，则随机事件 A_2 所包含的样本点数为 $k_2=\mathrm{A}_b^{k-1}\mathrm{A}_a^1\mathrm{A}_{a+b-k}^{m+n-k}$，所以

$$P(A_2)=\frac{k_2}{n^*}=\frac{\mathrm{A}_b^{k-1}\mathrm{A}_a^1\mathrm{A}_{a+b-k}^{m+n-k}}{\mathrm{A}_{a+b}^{m+n}}.$$

(3)设 A_3="第 k 次恰取到白球"，则 A_3 包含的样本点数为 $k_3=\mathrm{A}_a^1\mathrm{A}_{a+b-1}^{m+n-1}$，所以

$$P(A_3)=\frac{k_3}{n^*}=\frac{\mathrm{A}_a^1\mathrm{A}_{a+b-1}^{m+n-1}}{\mathrm{A}_{a+b}^{m+n}}=\frac{a}{a+b}.$$

注：从 $P(A_3)=\dfrac{a}{a+b}$ 可见，结果与 k 无关，理论上说明了抽签的公平性. 但是，有学生说了这样的问题：

"老师，有一次我宿舍为确定一张音乐会优待票的归属权，6位同学一致决定通过抽签解决. 可是在我还没抽时，第一位抽签的同学快速抽取并敏捷打开签后大声说'我中了！'，我连伸手的机会都没有，即使我抽，抽中的概率也是零，抽签真的公平吗?"

如何理解这个问题? 留待1.3节解决.

【例1.10】(盒子模型，也称分房模型)　有 n 个人，每个人都等可能地被分配到 $N(N>n)$ 间房中的任一间住，求恰好有 n 间房子各有一个人住的概率.

解　经过判断易知，该分房模型为古典概率模型，且其样本空间包含的样本点总数为

$$n^*=N^n.$$