重庆文理学院校本教材资助项目（项目编号：XBJC201204）

概率论与数理统计

GAILüLUN YU SHULI TONGJI

主　编　杨树成
副主编　杨春华

西南交通大学出版社

重庆文理学院校本教材资助项目（项目编号：XBJC201204）

概率论与数理统计

主　编　杨树成

副主编　杨春华

西南交通大学出版社

·成　都·

内容简介

本书是普通高等学校数学专业和其他非数学专业的教材. 全书共 8 章，内容包括随机事件与概率、随机变量的分布及其数字特征、多维随机变量的分布及其数字特征、数理统计的基本概念、参数估计、假设检验和回归分析等内容. 各章精选了大量的反映社会实际（如金融保险、投资、六西格玛管理、设备购置、工程设计、交通运输、安全性、质量管理、产品检验、医学检验、科学研究等）的例题和习题，在每一章的最后一节介绍 Excel 进行本章的统计计算. 本书力求在保持体系完整的前提下，弱化理论的推导，强化概率统计的思维训练和知识的实际应用. 全书共 190 多个例题，350 多个习题.

本书可作为高等学校工科、农医、经济、管理等专业的概率统计课程的教材，也可作为实际工作者的自学参考.

图书在版编目（CIP）数据

概率论与数理统计 / 杨树成主编. —成都：西南交通大学出版社，2014.8

ISBN 978-7-5643-3124-5

Ⅰ. ①概… Ⅱ. ①杨… Ⅲ. ①概率论－高等学校－教材②数理统计－高等学校－教材 Ⅳ. ①021

中国版本图书馆 CIP 数据核字（2014）第 129834 号

概率论与数理统计

主编 杨树成

*

责任编辑 张宝华

封面设计 墨创文化

西南交通大学出版社出版发行

成都市金牛区交大路 146 号 邮政编码：610031 发行部电话：028-87600564

http: //www.xnjdcbs.com

成都中铁二局永经堂印务有限责任公司印刷

*

成品尺寸：185 mm × 260 mm 印张：18

字数：448 千字

2014 年 8 月第 1 版 2014 年 8 月第 1 次印刷

ISBN 978-7-5643-3124-5

定价：29.00 元

前言

“概率论与数理统计”是研究随机现象的规律性，并通过有效地收集、整理和分析带有随机性的数据，对所考察的问题作出统计推断或预测的一门数学学科．它是具有高度抽象性和严密逻辑性，且应用广泛的课程，是应用型本科院校许多专业设置的一门重要专业基础课程，甚至是核心课程．

随着精英教育向大众化教育的转变，应用型本科院校的人才培养模式已由传统模式向应用型模式转变，这也给“概率论与数理统计”的教学带了一系列的问题与挑战．概括起来主要表现为：一是课时量的变化．有的专业开设两个学期，总学时为 108 学时；有的专业开设一学期，共 54 学时；还有专业开设一学期，共 48 学时，甚至 36 学时．二是应用性需要加强，理论需要弱化．许多专业应该多开设实验课程，而实验所用的统计软件也多种多样．

面对新的挑战，以前的教材不再适应新的教学要求．**首先**，传统教材过分追求逻辑的严密性和理论体系的完整性，重理论而轻实践，剥离了概念、原理和范例的现实意义，导致教学内容过于抽象，也不利于与其他课程及学生自身专业相衔接，进而造成学生“学不会，用不了”的尴尬局面；传统教材的重书本知识，轻实践操作；重理论推导，轻知识应用，已无法达到应用型人才培养的要求．**其次**，传统教材的教学内容没有和统计软件的应用相结合．概率论与数理统计要进行统计分析，常常需要做大量的数据处理，没有有效的计算方法，就无法满足概率论与数理统计的应用．传统的教学内容只是对基本公式和基本概念进行高等数学试题的练习，无法真正将概率论与数理统计与实际应用联系起来，只有融入现代的计算工具才能更有效地加强概率论与数理统计的应用性．**再次**，传统教材没有设计实验教学内容．传统教材因没有与统计软件的应用相结合，也就没有安排实验项目和实验指导，学生也就没有得到数据处理能力方面的培养．**最后**，传统教材重点不突出，内容面面俱到，教学内容不能适应教学时数减少的趋势要求，其结果往往是学完概率论部分就没有了教学时数，而应用性很强的数理统计部分却无法完成教学．所以教学内容只有进行改革，才能适应新形势下应用型人才的培养．

本书有以下几个特点：**一是注重随机思想方法的培养**．“概率论与数理统计”的首要教学内容是让学生具有“将不确定的现象理解成随机现象，而刻画随机现象的根据是随机变量．随机变量具有自身的分布和数字特征，研究随机变量的分布和数据特征的方法是抽样和统计推断”的随机思想方法．本书力图在不失知识的逻辑结构的前提下，弱化理论的推导过程，强调解决问题的思想方法．**二是压缩概率论部分，增强数理统计部分，强化概率统计方法的应用**．将概率论的教学内容压缩到 40% 左右，降低概率计算的难度，增强应用的广泛性．**三是在内容的叙述上**，不追求系统性和严密性，而是在不失知识内涵和逻辑结构的前提下，对基本概念和方法采用通俗简洁的表述方式，突出通俗性、可读性和应用性，希望这些工作能对读者有所裨益．内容安排上用“案例—概念、公式—应用”的方式进行教学，而不是传统的“概念、公式—验证概念、公式”的方式进行教学．例题、习题都力求减少和降低概念、

公式的验证性习题，增加实际应用型习题．**四是将大家常用的 Excel 统计分析融于教学之中**．概率论与数理统计的教学内容，特别是数理统计中的假设检验、方差分析、线性回归等内容都涉及大量的数据处理，所以，本书有意设计了大量的必须借助统计软件才能解决的习题．同时也为开设实验教学的专业教学提供了试验的素材和项目，通过实验让学生掌握统计软件的操作方法，并真正获得该课程应用的能力．

本书由重庆文理学院数学与财经学院杨树成任主编，杨春华任副主编，参编人员有霍永亮、易文德、骆小琴、李勇、陈晓东、蒲建平、刘静、邬吉波、李云红、贾小勇、付天贵等，全书由杨树成、杨春华老师统稿，由霍永亮等老师审阅．参编教师均是长期从事数学研究和在数学教学上有丰富经验的教师，他们在各自完成的章节中都融入了许多自己深刻的理解和体会，这使本书不仅内容全面、论证严谨，而且深入浅出、易学好教．

本书在编写过程中参考的文献均列在了书末，我们从中获益匪浅，在此一并致谢，读者在学习本书时也可以参考。本书是重庆文理学院校本教材资助项目（项目编号：XBJC201204），重庆文理学院及重庆文理学院数学与财经学院对本书的编写给予了大力支持，在此一并表示深切的谢意！

由于编者水平有限，书中的不足之处，敬请各位专家、同行及读者不吝赐教．

作　者

2014 年 2 月

目　录

1　随机事件与概率

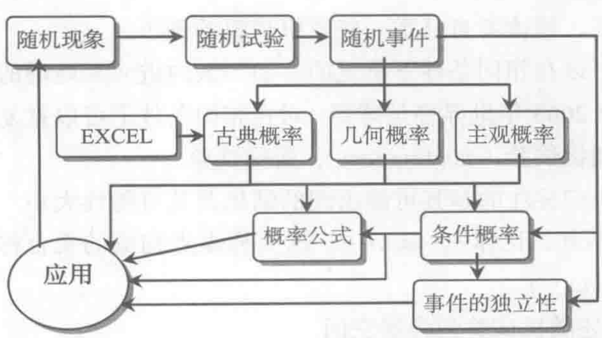
生活中我们会面对大量的不确定性事情，这些事情出现的可能性有大有小，即使发生的可能性非常小，一旦发生其后果可能非常严重．美国航天飞机的每一个部件都经过了极其严格的检验，出现故障的可能性都非常小，不幸的是美国挑战者号在1986年1月28日进行代号STS-51-L的第10次太空任务时，因为右侧固态火箭推进器（Solid Rocket Booster, SRB）上面的一个O形环失效，导致一连串的连锁反应，在升空后73秒时，爆炸解体坠毁，机上7名太空人全数罹难．2003年2月1日哥伦比亚号航天飞机也发生了不幸事故，在代号STS-107的第28次任务重返大气层的阶段中与控制中心失去联系，不久后被发现在德克萨斯州上空爆炸解体，机上7名太空人全数罹难．目前，有关哥伦比亚号失事的直接原因基本确定：超高温空气从机体表面缝隙入侵隔热瓦下部，最终造成航天飞机在返航途中解体坠毁．据介绍，飞机起飞一分钟后，遭遇的风力强度已经接近允许的极限．专家因此认为，原本已开始出现老化的机翼因遭受如此强风吹袭，才在外界异物的撞击下显得“弱不禁风”，从而出现破损，为返航途中的超高温空气入侵打开了“后门”．

挑战者号在升空73秒后爆炸

挑战者号宇航员珍贵照片

又如，有的人不明白福利彩票中奖的可能性大小，一味碰运气，结果是倾家荡产．由此可知，了解不确定事件发生的可能性大小及其后果，对于我们理性看待不确定事件和科学决策是非常有用的．

本章的主要内容有：随机事件、随机事件的运算、事件概率的确定方式、重要的概率公式、条件概率、随机事件的独立性、Excel中计算组合数、排列数、阶乘、数的幂的函数等．



第一章知识结构图

☆ 将某些具有不确定性结果的现象理解为随机现象，研究随机现象的方法是随机试验，随机试验的结果就是随机事件.

☆ 对于随机事件我们关注的焦点是其发生的可能性大小，即概率. 确定概率的方法视其所具有的条件，主要有主观概率、古典概率、几何概率、条件概率等.

☆ 客观世界是极其复杂的，因而随机事件具有复杂性，我们可以利用概率的性质和公式解决复杂事件的概率.

☆ 随机事件可能相互影响，也可能互不影响，互不影响的随机事件称为相互独立，独立性的概念在概率论中具有相当重要的地位，随机事件的独立性可以用条件概率来定义.

☆ 主观概率、古典概率、几何概率、条件概率，以及事件的独立性有其重要的应用. 如条件概率可以用于敏感性调查.

☆ 计算古典概率可能涉及组合数、排列数、阶乘、数的幂的计算，Excel 中内置有这些函数，可用来方便地计算古典概率.

1.1 随机事件

（泡泡糖问题）可怜的琼斯夫人路过一分钱一枚的泡泡糖出售机时，尽量不使她的双胞胎儿子有所察觉. 大儿子："妈妈，我要泡泡糖." 二儿子："妈妈，我也要，我要和比利拿一样颜色的." 分币泡泡糖出售机几乎空了，里面只有 4 粒白色的和 6 粒红色的泡泡糖. 说不准下一粒是什么颜色. 如果琼斯夫人想要满足两个儿子的要求，需要花多少钱呢?

假如大儿子说："妈妈，我想要 1 颗白色的泡泡糖." 二儿子："妈妈，我也要，我要和比利拿一样颜色的." 这时如果琼斯夫人想要满足两个儿子的要求，需要花多少钱呢?

显然，在第一种情况下，琼斯夫人只花 2 分或 3 分钱就够了，而在第二种情况下，琼斯夫人需要花 2 分、3 分、4 分、5 分、6 分、7 分、8 分钱都有可能.

人们所考察的现象可分为两类：一类是在一定条件下必然发生或必然不发生的现象，称之为**必然现象**，如早晨太阳必然从东方升起，掷一颗骰子出现的点数必然小于 7. 另一类是在一定条件下可能发生也可能不发生的现象，称之为**随机现象**，如掷一颗骰子出现的点数，明年某地区的年均降雨量，某只股票明天的收益率等. 随机现象有两个显著的特点：

（1）结果不止一个；

（2）哪个结果出现事先未知.

随机现象到处可见，请读者自己举一些随机现象的例子.

有些随机现象是可以在相同条件下重复的，如一天内进入某商场的顾客数. 有些随机现象是不可以重复的，如 2008 年世界杯足球赛. 对在相同条件下可以重复进行的随机现象的观察、记录、实验称为随机试验（random test），简称试验.

对于随机试验，我们关注的是其可能出现的结果及其可能性大小. 随机试验可能出现的每个基本结果称为样本点，记作 $\omega_1, \omega_2, \cdots$，所有样本点构成的集合称为样本空间（sample space），记作 $\Omega = \{\omega_1, \omega_2, \cdots\}$.

例 1.1.1 写出下述随机试验的样本空间.

E_1：掷一枚骰子，观察出现的点数；

E_2：记录某手机在 1 小时内收到的短信数量；

E_3：在涨停板限制下，某只股票明天的收益率；

E_4：先后掷两枚硬币，观察出现正反面的情况.

解 E_1 的样本空间为 $\Omega=\{1,2,3,4,5,6\}$；

E_2 的样本空间为 $\Omega=\{0,1,2,\cdots\}$；

E_3 的样本空间为 $\Omega=\{r:-10\%\leqslant r\leqslant 10\%\}$；

E_4 的样本空间为 $\Omega=\{$正正，正反，反正，反反$\}$.

随机试验的某个结果出现，表明一些事件发生了，由于这些事件在试验前无法预知其是否会出现，所以我们将随机试验的结果称之为**随机事件**（random event），简称**事件**. 试验的基本结果，即样本点称为**基本事件**（elementary event）. 一定发生的事件称为**必然事件**（certain event），用 Ω 表示；一定不发生的事件称为**不可能事件**（impossible event），用 $\varnothing$ 表示. 习惯上用大写字母 A, B, C 等表示事件.

例如，掷一枚骰子，观察出现的点数. $A=\{$出现 2 点$\}$、$B=\{$出现的点数小于 3$\}$、$C=\{$出现的点数是偶数$\}$都是随机事件，而 $D=\{$出现的点数小于 8$\}$为必然事件，$E=\{$出现的点数小于 0$\}$为不可能事件. 这些事件都是一些样本点构成的集合，换言之都是样本空间的子集（其中不可能事件不包含样本点看作空集）. 显然，当且仅当事件包含的样本点中有一个出现，我们就说该事件发生了.

骰子

随机事件随处可见，请读者自己列举一些试验和相应的随机事件.

随机事件是样本空间的子集，所以事件之间的运算与集合之间的运算是一致的. 概率论中常用矩形表示样本空间 Ω，圆表示随机事件，如下图所示，这类图称为**维恩**（Venn）图. 下面用维恩图直观地介绍一些常用的事件之间的关系和运算.

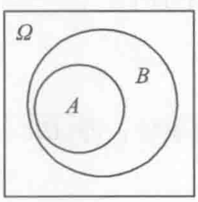


图 1.1.1 $A\subset B$

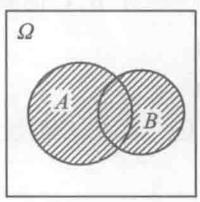


图 1.1.2 $A\cup B$

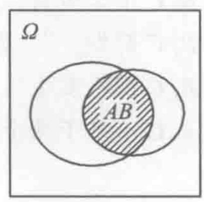


图 1.1.3 AB

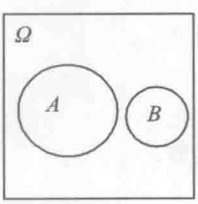


图 1.1.4 $AB=\varnothing$

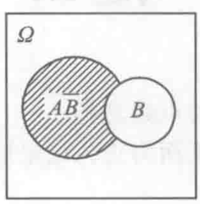


图 1.1.5 $A\overline{B}$

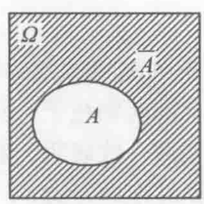


图 1.1.6 $\overline{A}$

如图 1.1.1 所示，如果事件 A 发生必然导致事件 B 发生，则称 B **包含** A，记作 $A\subset B$. 如果有 $A\subset B$ 且 $B\subset A$，则称事件 A 与 B **相等**，记作 $A=B$. 例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点$\}$、$B=\{$出现的点数小于 3$\}$，则 $A\subset B$.

如图 1.1.2 所示，阴影部分表示“事件 A 与 B 中至少有一个发生”，这样的一个事件称作事件 A 与 B 的和事件，记作 $A\cup B$．例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点或 3 点$\}$、$B=\{$出现 3 点或 5 点$\}$，则 $A\cup B=\{$出现 2 点，3 点或 5 点$\}$．

如图 1.1.3 所示，阴影部分是 A 与 B 的公共部分，表示“事件 A 与 B 同时发生”，这样的一个事件称作事件 A 与 B 的积事件，记作 $A\cap B$（或 AB）．例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点或 3 点$\}$、$B=\{$出现 3 点或 5 点$\}$，则 $AB=\{$出现 3 点$\}$．

如图 1.1.4 所示，事件 A 与 B 不能同时发生，即 $AB=\varnothing$，此时称事件 A 与 B 互不相容或互斥．例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点或 3 点$\}$、$B=\{$出现 5 点或 6 点$\}$，则 A 与 B 互不相容，$AB=\varnothing$．

如图 1.1.5 所示，阴影部分表示“事件 A 发生而 B 不发生”，这样的一个事件称作事件 A 与 B 的差事件，记作 $A-B$ 或 $A\bar{B}$．例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点或 3 点$\}$、$B=\{$出现 3 点或 5 点$\}$，则 $A\bar{B}=\{$出现 2 点$\}$．

如图 1.1.6 所示，阴影部分表示“A 不发生”之事件，即 $\Omega-A$，称为 A 的对立事件或逆事件，A 的对立事件记作 $\bar{A}$．显然，有 $A\bar{A}=\varnothing$，$A\cup\bar{A}=\Omega$，$\bar{\bar{A}}=A$．例如，在掷一枚骰子的试验中，事件 $A=\{$出现 2 点或 3 点$\}$，则 $\bar{A}=\{$出现 1 点、4 点、5 点或 6 点$\}$．

例 1.1.2　设 A, B, C 是三个事件，试用 A, B, C 表示下列事件：（1）A 发生，B 和 C 都不发生；（2）A, B, C 是中恰好有一个发生；（3）A, B, C 是中至少有一个发生；（4）A, B, C 都发生；（5）A, B, C 都不发生；（6）A, B, C 不多于两个发生．

解　（1）“A 发生，B 和 C 都不发生”表示为 $A\bar{B}\bar{C}$；

（2）“A, B, C 中恰好有一个发生”，即要么 A 发生而 B 和 C 都不发生，要么 B 发生而 A 和 C 都不发生，要么 C 发生而 A 和 B 都不发生，所以表示为 $A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$；

（3）“A, B, C 中至少有一个发生”是 A, B, C 的和事件，即 $A\cup B\cup C$；

（4）“A, B, C 都发生”，即 ABC；

（5）“A, B, C 都不发生”，即 $\bar{A}\bar{B}\bar{C}$；

（6）“A, B, C 不多于两个发生”是“A, B, C 都发生”的对立事件，即 $\overline{ABC}$．

习题 1.1

1. 写出下列随机试验的样本空间：

（1）同时掷两颗色子，记录出现的点数之和．

（2）重复掷一枚硬币，直到出现正面为止，记录掷硬币的次数．

（3）测算灯泡的寿命．

2. 写出下列事件中的样本点：

（1）一只袋子中有编号为 1, 2, 3, 4, 5 的五个球，从中随机取 3 个球，球的最小号码为 1．

（2）将一枚硬币掷两次，A 表示“第一次出现正面”，B 表示“两次出现不同的面”，C 表示“至少有一次出现正面”．

3. 在先后掷三枚均匀硬币的试验中，用集合表示下列事件：

（1）$A=\{$至少出现一个正面$\}$；（2）$B=\{$最多出现一个正面$\}$；

（3）$C=\{$恰好出现一个正面$\}$；（4）$D=\{$出现三面相同$\}$.

4. 用集合表示下列随机试验的样本空间与随机事件 A：

（1）同时掷三枚骰子，记录三枚骰子的点数之和，事件 A 表示“点数之和大于 10”.

（2）对目标进行射击，击中后便停止射击，观察射击的次数；事件 A 表示“射击次数不超过 5 次”.

5. 以下命题正确的是（　　）.

（A）$AB\cup A\overline{B}=A$　（B）若 $A\subset B$，则 $AB=A$

（C）若 $A\subset B$，则 $\overline{B}\subset\overline{A}$　（D）若 $A\subset B$，则 $A\cup B=B$

6. 若事件 A,B,C 满足等式 $A\cup C=B\cup C$，试问 $A=B$ 是否成立？

7. 指出下列事件等式成立的条件：（1）$A\cup B=A$；（2）$AB=A$.

8. 一个工人生产了 3 个零件，以事件 A_i 表示他生产的第 i 个零件是合格品 $(i=1,2,3)$，试用 $A_i(i=1,2,3)$ 表示下列事件：

（1）$B_1=\{$只有第一个零件是合格品$\}$；

（2）$B_2=\{$三个零件中只有一个合格品$\}$；

（3）$B_3=\{$第一个是合格品，但后两个零件中至少有一个次品$\}$；

（4）$B_4=\{$三个零件中最多只有两个合格品$\}$；

（5）$B_5=\{$三个零件都是次品$\}$；

（6）$B_6=\{$三个零件中最多有一个次品$\}$.

9. 对飞机进行两次射击，每次射一弹，设 $A=\{$恰有一弹击中飞机$\}$，$B=\{$至少有一弹击中飞机$\}$，$C=\{$两弹都击中飞机$\}$，$D=\{$两弹都没击中飞机$\}$. 又设 X 为击中飞机的次数，（1）试用 X 表示事件 A, B, C, D；（2）A, B, C, D 中哪些是互不相容的事件？哪些是对立的事件？

10. 请叙述下列事件的对立事件：

（1）$A=\{$掷两枚硬币，皆为正面$\}$；

（2）$B=\{$射击三次，皆命中目标$\}$；

（3）$C=\{$加工四个零件，至少有一个合格品$\}$.

11. 事件 A 与 B 互不相容，试问 A 与 B 是否对立？反之如何？

12. 设 X 表示某试验的结果，其样本空间为 $\Omega=\{0\leqslant X\leqslant 2\}$，记事件 $A=\{0.5<X\leqslant 1\}$，$B=\{0.25\leqslant X<1.5\}$，写出下列各事件：（1）$\overline{A}B$；（2）$\overline{A}\cup B$；（3）$\overline{AB}$；（4）$\overline{A\cup B}$.

1.2 概率的定义及确定方法

正如英国逻辑学家和经济学家杰文思所说：“它是生活真正的领路人，如果没有对概率的某种估计，我们就寸步难移，无所作为”. 研究随机现象的目的之一是为了确定事件发生的可能性大小，概率（probability）是事件发生的可能性大小的度量. 事件的特征不同，度量其发生的可能性大小的方法也不同，本节介绍古典概率、几何概率、频率和主观概率.

1.2.1 主观概率

Thomas Bayes
(1702—1763)

早在远古时期，人们就注意到了大量偶然事件所表现出的规律性，并将这种规律性应用于实践当中．如在狩猎、捕鱼时根据经验判断猎物或鱼群出现的可能性，选择出现可能性最大的地方去狩猎或捕鱼．数理统计学中的贝叶斯学派认为，一个事件的概率是人们根据对事件发生的可能性所给出的个人信念，根据主观判断来确定事件发生的可能性大小叫做**主观概率**．在现实世界中，有一些随机现象发生的条件经常变化，或不能重复或者不能大量重复，这时可用主观判断来确定事件的概率．

例 1.2.1 用主观方法确定概率的例子：

（1）人们根据经验对下雨的概率做出主观判断．

（2）人们根据经验对股票上涨的概率做出主观判断．

（3）医生根据经验对患者患有某种疾病的可能性进行主观判断．

（4）外科医生根据经验和患者的病情给出手术成功的概率．

（5）体育爱好者根据球队的情况和经验预测球队赢得比赛的概率．

用主观判断来确定事件的概率的例子举不胜举．主观概率是对事件发生的概率的推断和估计，其精确性有待于实践的检验和修正．

1.2.2 概率的频率定义

现实当中许多随机现象的样本点都非常多，各样本点出现的可能性也不相同，而且同样的试验，不同的条件下结果会随时发生变化．如居民的收入、产品的市场占有率等．当样本点不是等可能出现时，或条件会随时发生变化时，事件发生的可能性大小可以通过分析大量试验的结果加以研究．设事件 A 在 n 次重复试验中共发生了 f_n 次，则称 f_n 为事件 A 在 n 次重复试验中发生的**频数**（frequency），$F_n(A)=f_n/n$ 为事件 A 在 n 次重复试验中发生的**频率**．历史上很多数学家和统计学家做过掷硬币的试验，表 1.2.1 是历史上掷硬币试验的一些结果．

表 1.2.1 历史上掷硬币试验的若干结果

试验者	掷硬币次数	出现正面次数	频率
德摩根（De Morgan）	2048	1061	0.5181
蒲丰（Buffon）	4040	2048	0.5069
费勒（Feller）	10000	4979	0.4979
皮尔逊（Pearson）	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

从表 1.2.1 可以看出，随着试验次数的增加，正面出现的频率越来越接近于 0.5．实践表明，频率虽然不是固定的，但当试验次数 n 充分大时，事件 A 的频率就会在一个确定实数 $P(A)$ 处波动．一般地，试验次数越多，频率的波动越小，这叫做**频率的稳定性**．

频率的稳定性在实践中有着重要的应用. Dewey G. 统计了约 438023 个英文单词中各个字母出现的频率，发现英语中 26 个字母出现的频率从高到低依次是 E、T、O、A、N、I、R、S、H、D、L、C、F、U、M、P、Y、W、G、B、V、K、X、J、Q、Z，且频率相当稳定. 字母使用频率的研究，在打字机键盘的设计、印刷铅字的铸造、信息的编码、密码的破译等领域都有重要的应用.

频率在六西格玛管理中也有重要应用. 在六西格玛管理中机会缺陷率 DPO（Defects Per Opportunity），即每次机会中出现缺陷的比率表示每个样本量中缺陷数占全部机会数的比例，由下式计算：

$$DPO=\frac{\text{总的缺陷数}}{\text{产品数}\times\text{机会数}},$$

DPO 又常以百万机会缺陷数 $DPMO$（Defects Per Million Opportunity）表示，即 $DPMO=DPO\times10^6$，或由下式计算：

$$DPMO=\frac{\text{总的缺陷数}\times10^6}{\text{产品数}\times\text{机会数}}.$$

$DPMO$ 对应于过程输出质量特性超出规格限的比率，对应着过程的西格玛水平 z . $DPMO$ 越大说明管理水平越低，六西格玛管理的最高境界是 $DPMO$ 小于 3.4.

例 1.2.2 假定 100 块电路板中，每一个电路板都含有 100 个缺陷机会，若在制造这 100 块电路板时共发现 21 个缺陷，则

$$DPO=\frac{21}{100\times100}=0.0021=0.21\%,$$

$$DPMO=\frac{21\times10^6}{100\times100}=2100.$$

本例的 $DPMO$ 对应的六西格玛水平为 4.3 ~ 4.4，远未达到六西格玛的水准.

当试验次数充分多时，频率稳定地趋向的实数 $P(A)$ 叫做事件 A 的**概率**，这是概率的频率定义，这个定义是奥地利数学家米泽斯于 1919 年提出的. 因此，可以**用充分多次试验中事件 A 的频率作为事件 A 发生的概率的近似值**，这是确定概率的一种有效方法. 大量统计调查获得的某类事件发生的频率即可作为事件的概率的近似值. 如第六次全国人口普查统计数据显示男性人口占 51.27%，女性人口占 48.73%，所以从全国随机抽取 1 人是男性的概率近似为 0.5127. 概率的频率定义使通过计算机大量重复地试验近似求解事件的概率成为可能，此方法称为**蒙特卡罗法**. 应该指出的是，从应用角度看，频率定义可以克服等可能性观点不易解决的某些困难，但从理论上讲，这种定义方法也是不够严谨的.

1.2.3 古典概率

在公元前 2000 年的埃及古墓中已有正方形的骰子，可想而知当时就出现了掷骰子的游戏

P.S.Laplace
(1749—1827)

或赌博．早在16世纪，赌博中的偶然现象就开始引起人们的注意．数学家卡丹诺（G. Cardano，1501—1576）首先觉察到，赌博输赢虽然是偶然的，但较大的赌博次数会呈现出一定的规律性，卡丹诺为此还写了一本《论赌博》的小册子，书中计算了掷两颗骰子或三颗骰子时，在一切可能的方法中有多少方法得到某一点数．

掷一枚均匀的骰子，出现2点的可能性有多大呢？点数小于3的可能性有多大呢？将掷一枚均匀的骰子看作一次随机试验，设事件 $A=\{$出现2点$\}$，事件 $B=\{$出现的点数小于3$\}$．因为样本空间包含6个样本点，可以认为这6个样本点出现的可能性相同，A 只包含1个样本点，B 只包含2个样本点，因此自然想到，A 发生的可能性为 $1/6$，B 发生的可能性为 $2/6=1/3$，换句话说，$1/6$ 度量了 A 发生的可能性大小，$1/3$ 度量了事件 B 发生的可能性大小．

1814年拉普拉斯（P. S. Laplace，1749—1827）出版了《分析概率》，在这部著作中拉普拉斯给出了概率的古典定义：事件 A 的概率 $P(A)$ 等于一次试验中有利于事件 A 的可能的结果数与该试验中所有可能的结果数之比．

定义 1.2.1 设样本空间 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$ 中有 n 个样本点 $\omega_1,\omega_2,\cdots,\omega_n$，且各样本点是等可能发生的．事件 A 中有 m 个样本点，则称 $P(A)=m/n$ 为事件 A 的**古典概率**（classical probability），简称**概率**．

计算古典概率时，样本点的个数可以借助排列组合数来计算．

例 1.2.3 随机地选10个人，试求这10个人生日不在同一天的概率．

解 先计算样本空间 Ω 中样本点的个数．因为随机地选取10个人，所以每个人在一年365天中任何一天出生是等可能的，这样10个人的生日共有 365^{10} 种可能，样本空间中的样本点个数为 365^{10}．10个人生日不在同一天，可以看作是从365天里选出10天进行排列，所以10个人生日不在同一天的样本点个数为排列数 A_{365}^{10}．所以事件 $A=\{$10个人生日不在同一天$\}$ 含有 A_{365}^{10} 个样本点．由古典概率的定义，10个人生日不在同一天的概率为

$$P(A)=\frac{\mathrm{A}_{365}^{10}}{365^{10}}\approx 0.8831.$$

例 1.2.4 从一个班里的36名同学中随机选3名同学当班干部，试求其中甲同学当选的概率．

解 先计算样本空间 Ω 中样本点的个数．从36名同学中随机选3名同学共有 C_{36}^{3} 可能，所以样本空间 Ω 中样本点的个数为 C_{36}^{3}．若要甲同学当选只须从剩下的35名同学中随机选2名即可，即甲同学当选的可能情况有 C_{35}^{2} 种．所以事件 $A=\{$甲同学当选$\}$ 含有 C_{35}^{2} 个样本点．由古典概率的定义有

$$P(A)=\frac{\mathrm{C}_{35}^{2}}{\mathrm{C}_{36}^{3}}\approx 0.0833.$$

例 1.2.5 已知某种产品1000件中有50件次品（次品率为5%）．试求从1000件这种产品中随机抽取100件，其中恰有5件次品的概率．

解 因从 1000 件产品中随机抽取 100 件共有 C_{1000}^{100} 种可能取法，所以样本空间 Ω 中样本点的个数为 C_{1000}^{100}. 又因为是随机抽取，所以这 C_{1000}^{100} 个样本点是等可能的.

随机抽取 100 件产品中恰有 5 件次品共有 $C_{950}^{95}C_{50}^{5}$ 种可能取法，所以事件 $A=\{$随机抽取 100 件产品中恰有 5 件次品$\}$包含的样本点数为 $C_{950}^{95}C_{50}^{5}$. 由古典概率的定义有

$$P(A)=\frac{C_{950}^{95}C_{50}^{5}}{C_{1000}^{100}}\approx 0.1897 .$$

例 1.2.6 盒子中有 100 张彩票，其中只有 10 张有奖，现在把彩票随机地一张张摸出来，摸完后再开奖. 试求第 $k\,(1\leqslant k\leqslant 100)$次摸到奖的概率.

解 把 100 张彩票都看作是不同的（设想它们已编号），若把摸出的彩票依次排列在 100 个位置上，则可能的排列法为 100!，把它们作为样本点全体. 第 k 次要摸到奖，可以设想先在第 k 个位置上放一张有奖彩票，其余 99 个位置上随机地各放一张彩票，则第 k 个位置上有 10 种可能，其余 99 张彩票的排列有 99!种可能. 所以第 k 次要摸到奖共有 $10\times 99!$ 种可能，故事件 $A=\{$第 k 次摸到奖$\}$的概率为

$$P(A)=\frac{10\times 99!}{100!}=\frac{1}{10}.$$

这个结果说明中奖概率与摸彩票先后顺序无关.

1.2.4 几何概率

古典概率要求样本空间是有限的，样本点是等可能发生的. 当样本空间是无限的，样本点是等可能发生的时候，可用几何概率度量事件发生的可能性大小. 1706 年法国数学家蒲丰（George-Louis Leclerc de Buffon，1707—1788）的《偶然性的算术试验》完成，他把概率和几何结合起来，开始了几何概率的研究，他提出的“蒲丰问题”就是采取几何概率的方法来求圆周率 π 的尝试.

G.L.L. de Buffon
(1707—1788)

例如，在区间 [1,5] 内做随机投点试验，假设每次都能投中区间 [1,5]，且 [1,5] 内每个点被投中的可能性相等，则投中区间 [2,4] 的概率是多少？由于区间 [1,5] 的长度为 4，而 [2,4] 的长度为 2，每个点被投中的可能性相同，自然想到 [2,4] 被投中的可能性为 $2/4=1/2$. 长度是一维空间的测度，所以 [2,4] 被投中的可能性是 [2,4] 的测度比 [1,5] 的测度.

定义 1.2.2 若样本空间 Ω 构成 n 维空间的一个有限可测区域 G，且样本空间中的点是等可能出现的，而事件 $A\subset G$ 构成 G 的某一部分可测区域 g，如图 1.2.1 所示，则称

$$P(A)=\frac{g\text{的测度}}{G\text{的测度}} \tag{1.2.1}$$

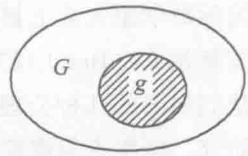


图 1.2.1 几何概率

为事件 A 的几何概率（geometric probability），当 $A=\varnothing$ 时规定 $P(A)=0$.