普通高等教育
十三五
"十三五"规划教材

普通高等教育"十三五"规划教材

Statistical Analysis
统计分析系列

◎ 苏连塔 陈明玉 主编

概率论与数理统计

——基于R

中国工信出版集团

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

普通高等教育“十三五”规划教材　　统计分析系列

概率论与数理统计——基于R

苏连塔　陈明玉　主　编

董会英　杨昔阳
林珊华　黄东兰　副主编

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书是福建省“高等学校教学改革研究项目”的研究成果。本书介绍概率论与数理统计的思想和方法，要求学生在掌握概率论与数理统计的基本概念和理论的同时，初步掌握处理随机现象的基本思想与方法，培养他们运用概率统计方法分析和解决实际问题的能力，实现课程应用型人才的培养目标。全书共九章，主要包括：随机事件与概率、一维/二维随机变量及其分布、随机变量的数字特征、大数定律及中心极限定理、数理统计的基础知识、参数估计、假设检验、R 在概率论与数理统计中的简单应用等。每章前有教学目标，后有实用案例，既保证理论体系严密，又注重可读性。书后附有习题答案，登录华信教育资源网 www.hxedu.com.cn 可下载教学课件、R 源代码。

本书可作为地方本科院校的概率论与数理统计教材，也可满足高职专科院校概率论与数理统计课程的教学需要。

图书在版编目（CIP）数据

概率论与数理统计：基于 R / 苏连塔，陈明玉主编. — 北京：电子工业出版社，2017.10
（统计分析系列）
ISBN 978-7-121-32756-8

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

中国版本图书馆 CIP 数据核字（2017）第 232218 号

策划编辑：秦淑灵
责任编辑：徐　萍
印　　刷：北京京师印务有限公司
装　　订：北京京师印务有限公司
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×1092　1/16　印张：11.75　字数：301 千字
版　　次：2017 年 10 月第 1 版
印　　次：2018 年 8 月第 2 次印刷
定　　价：39.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）88254531。

前　言

概率论与数理统计是研究随机现象的统计规律性的一门数学学科，随着信息化社会的逐步推进与大数据时代的到来，这门学科越来越凸显出其优秀的魅力。在大数据发展的时代背景下，概率论与数理统计是大数据的底层理论之一，概率统计的方法已经应用于以信息和技术为基础的现代社会的各个领域，在现在和未来，管理者、决策者、产品经理、产品运营和开发工程师都需要掌握概率论与数理统计的知识。

本书在编写过程中融入了编者多年从事概率统计教学实践的心得和体会，在结构体系、内容安排、习题选择等方面充分考虑地方本科院校的实际，力求能为地方本科院校所适用，适应地方本科院校的学生学习概率统计的需要，让他们更好地掌握处理随机现象的基本思想与方法，熟悉统计语言并具备统计思想，掌握几种具体的数理统计方法，有效地运用所学的概率统计方法分析和解决实际问题，为今后的专业学习和就业打下坚实的基础。为此，本书努力体现以下几个特色：

第一，突出以培养应用技术型人才为目标，厚基础，重应用，让学生真正学到必需、够用的概率论与数理统计知识。

本书能够满足经济管理、计算机、电子信息类等相关专业在转型时期对概率统计的知识需求，注重介绍概率论与数理统计的思想和方法，结合实际背景和直观形象，从实际问题引入，讲清抽象概念，减少数理论证的过程，注重学生基本运算能力的训练和用随机思想、统计方法分析问题、解决问题能力的培养。

第二，培养学生初步用统计软件 R 处理数据的能力，拓展了公共课的教学环节。

有了本书第九章——R 在概率论与数理统计中的简单应用，本书更富实用性。

本书虽然给出了各种与分布有关的表格，传统的概率统计教学也离不开这些表格，但是使用 R 软件比查表更加方便和可靠，使学生可以从烦琐而又不够精确的查表中解放出来；区间估计、假设检验中繁杂的计算问题使学生对数理统计的学习望而却步，R 软件可以对这些属于简单劳动的计算进行“秒杀”，从而让学生有更多的时间去理解统计方法的思想和原理。初步熟悉了 R 软件，就可以提高学生对概率统计学习的兴趣。对这部分内容的教学可安排在实验室上课，以增加公共课的实践教学环节。

第三，明确提出教学目标在先，附有精选实用案例在后，增强了可读性。

在每章的开头给出了教学大纲中的教学目标和教学的建议学时数，这样既便于老师在教的过程中有的放矢，又有利于学生在学的过程中提纲挈领，不会因知识点一大堆而迷失方向，促使师生双方明白要教什么和要学什么，确定各自的行动方向；在每章的后面，针对该章的内容，以实际问题为背景，精选了典型的案例，通过这些案例，可以引出概率统计的基本概念、基本模型和基本方法，侧重各种方法及其应用，让学生深刻理解知识点，掌握概率统计方法在经济学、管理学及其他学科中的重要应用，学习起来也更有激情和针对性。这些案例大多以故事的形式给出，增强了教材的可读性。

第四，立足地方本科院校，根据长期教学的积累，对教学内容做出了恰当的安排。

本书共有九章，编者根据多年的教学经验，结合地方本科院校本科生的学习能力和要求，

建议将统计部分的教学进行到假设检验为宜，对于回归分析，本书只在第六章的第二节做了初步的铺垫，所给出的统计方法也只放在单个正态总体内讨论，否则会导致学生“消化不良”；另外，本书把所涉及的常用的一维随机变量单独予以介绍，让学生掌握了一般的离散化思想、随机思想后，再来关注典型的情形，这样能做到条理清楚，学习精力不会被分散。

在本书的教学中，第九章的教学内容可分散于前面章节的教学之中，第一、第二两章中有些教学内容可能在高中阶段已讲授过，可统一班级同学的意见，根据实际情况适当删减，减少教学课时数，以便增加统计部分的教学内容，做到不再“重概率轻统计”，让学生在大数据时代多学习必要的统计知识。根据不同专业，若教学计划总时数为36学时，建议整个学期的教学内容覆盖到第七章的第一节；若总时数为54学时，可以使用全部内容。

本书由苏连塔、陈明玉主编，董会英、杨昔阳、林珊华、黄东兰担任副主编。

由于编者水平有限，在编写本书的过程中，书中部分例题或课外读物引自他人著作，在此表示衷心的感谢，加之时间比较仓促，错误或不当之处在所难免，恳请使用本书的教师和同学提出宝贵意见，以便今后改正。

编　者

目　　录

第一章　随机事件与概率

教学目标：

1. 理解随机试验、样本空间、随机事件的概念，掌握事件之间的关系及运算，懂得写出一般随机试验的样本空间和把用概率论语言表达出的事件用集合的形式来表示；

2. 了解概率的统计定义、古典定义、几何定义，理解概率的公理化定义，掌握概率的基本性质，懂得计算一些简单的等可能概型、几何概型问题和利用概率的性质计算简单事件的概率；

3. 理解条件概率的概念，掌握概率的乘法公式、全概率公式和贝叶斯公式，懂得用这些公式计算有关的概率；

4. 理解事件独立性的概念和伯努利概型的定义，懂得用事件的独立性和二项概率公式计算简单事件的概率。

本章共分为 5 节，建议在 7 个课时内完成。

概率论与数理统计是研究随机现象统计规律性的一门学科。本章以研究随机现象统计规律的需要，首先给出了随机试验、样本空间、随机事件等概念，然后在此基础上给出概率的几种定义，奠定了概率的公理化定义在概率论与数理统计中的重要地位，让大家清楚认识到概率的真谛，最后给出了在特定场合下计算概率的一些手段和公式，如乘法公式、全概率公式、贝叶斯公式、事件的独立性、二项概率公式。

第一节　随机事件和样本空间

一、随机试验、样本空间

1. 随机现象

在自然界、生产实践、科学试验和日常生活中，人们观察到的现象大体可以归结为两种类型。一类是在一定条件下必然出现(或者不出现)某种结果的现象。这类现象的一个共同特点就是可以事前预言，即在准确地重复某些条件的情况下，它的结果总是可以肯定的，或是根据它过去的状态，在相同的条件下完全可以预言将来的发展，我们把这类现象称为确定性现象或者必然现象。例如，在标准大气压下，温度达到 100℃的纯水必然沸腾；异性电荷必然相互吸引等。几何、微积分、线性代数都是研究确定性现象的数学工具。另一类现象是不能预言其结果的，即在保持条件不变的情况下，重复实验或观察，或出现这种结果，或出现那种结果，这类现象称为随机现象。例如，投掷一枚骰子时朝上面的点数，某人在公交车站的候车时间，某款产品的使用寿命，天气变化情况，某种股票的涨跌等，它们的变化情况充满了不确定性，个别或少数观察具有偶然性。可以说，随机现象就是在一次观测中其结果不

能事先确定，但进行大量重复观察，这些随机现象会呈现某些规律性。例如，连续把一枚硬币(均匀，指定正反面)抛掷 5 次，则出现正面的次数可能为 0，1，2，3，4，5，出现其中某一个结果带有偶然性。当抛掷的次数很多时，则正面出现的次数占总次数的比例接近 0.5，即正面出现的次数大约占一半，呈现规律性，即统计规律性。

概率论与数理统计就是揭示和应用随机现象统计规律性的一门学科，但是各有侧重。概率论侧重于应用已有的随机现象统计规律性进行计算和推理；而数理统计是以概率论为理论依据，研究如何设计试验并对试验结果进行整理和统计分析，做出统计推断，从而揭示随机现象统计规律性。

2. 随机试验和样本空间

正如前面所述，为了研究随机现象内部存在的数量规律性，就必须对它作一定次数的观察。我们把对随机现象进行的一次观察、测量或实验，称为一次试验，如果其还满足下列三个条件，就称为随机试验(简称试验)，记为 E，E_1 或 E_2。

(1) 试验在相同条件下可以重复进行；

(2) 试验前能明确所有可能的结果；

(3) 试验前不能肯定哪个结果会发生。

因此，随机试验 E 将要出现的结果是不确定的，但其所有可能结果是明确的。我们把随机试验的每一个可能的基本结果，称为 E 的一个样本点，用 ω 表示。随机试验 E 的所有样本点的集合称为 E 的样本空间，记为 Ω。在具体问题的研究中，描述随机现象的第一步就是建立样本空间 Ω。

例 1 抛掷一枚硬币，观察正面和反面出现的情况，将两个基本结果分别记为 ω_1,ω_2，则该试验的样本空间为 $\Omega=\{\omega_1,\omega_2\}$。

例 2 掷一颗骰子，观察朝上面的点数，可能出现的点数分别为 1, 2, 3, 4, 5, 6，则该试验的样本空间为 $\Omega=\{1,2,3,4,5,6\}$。

例 3 从一大批同型号的电子元件中任取一个，测试其使用寿命(单位：小时)，则试验的样本空间为 $\Omega=[0,+\infty)$。

例 4 随机试验为：向一个直径为 50cm 的靶子射击，观察弹着点的位置。为表示出试验的样本空间，设弹着点 ω 的坐标为 (x,y)，则此试验的样本空间为 $\Omega=\{(x,y)\,|\,x^2+y^2\leqslant 25^2\}$。

3. 随机事件

在实际中，我们不仅要全面描述随机试验，也常常关心随机试验出现的某种情况，如例 2 中，我们会关心骰子出现的点数是否为偶数点，满足这一条件的样本点组成其样本空间的子集。一般地，我们称随机试验 E 的样本空间 Ω 的一个子集为一个随机事件，简称事件，常记为 A,B,C 等。任一事件 A 是相应样本空间 Ω 中的一个子集，它是由随机试验 E 的一个或几个样本点组成的集合，在概率论中常用一个长方形示意样本空间 Ω，用其中一个圆表示事件 A，如图 1-1 所示。

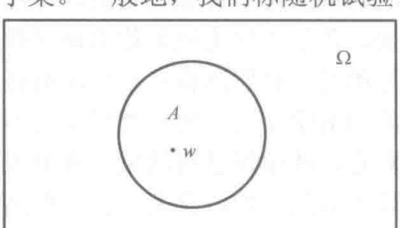

图 1-1

设 $A\subseteq\Omega$，若试验结果 $\omega\in A$，则称在这次试验中事件 A 发生；若试验结果 $\omega\notin A$，则称事件 A 不发生。如例 2 中，可把事件“出现偶数点”表示为 $A=\{2,4,6\}$，若掷

得结果为 4 点，则事件 A 发生；又如例 4 中，若规定弹着点距靶心距离不超过 5cm 为 10 环，则事件“射中 10 环”表示为 $B=\{(x,y)\,|\,x^2+y^2\leqslant 5^2\}$，若射中位置为 $(6,2)$，则事件 B 不发生，即未射中 10 环。

特别，由一个样本点组成的单点集也称为基本事件。样本空间 Ω 也是自身的一个子集，称这个事件为必然事件。不含任何样本点的空集 ϕ 称为不可能事件。虽然这两个事件是随机事件的特殊情况，但在概率论研究中起着重要作用。

二、随机事件的关系和运算

由上所述可理解，集合论是讨论随机试验与事件的合适的数学工具。事件既然用样本空间的子集来表示，事件的关系和运算也就能相应地用集合的关系和运算来表示。下面的讨论在一个固定的样本空间 Ω 中进行，主要根据“事件发生”的含义，给出这些关系和运算在概率中的含义，这有助于将复杂事件分解成简单事件，对以后复杂事件概率的计算是极其有益的。

(1) 事件的包含：若事件 A 发生时事件 B 一定发生，则称事件 B 包含事件 A，记作 $A\subseteq B$，见图 1-2。对于任意事件 A，有 $\phi\subseteq A$。如果 $A\subseteq B$，$B\subseteq C$，则 $A\subseteq C$。

(2) 事件的相等：若事件 A 和事件 B 相互包含，即 $A\subseteq B$ 且 $B\subseteq A$，则称事件 A 和事件 B 相等，记作 $A=B$。

(3) 事件的并（或和）：事件 A 和事件 B 的并事件是指事件 A 和事件 B 中至少有一个发生，记作 $A\cup B$ 或 $A+B$，也称作事件 A 和事件 B 的和事件（见图 1-3）。

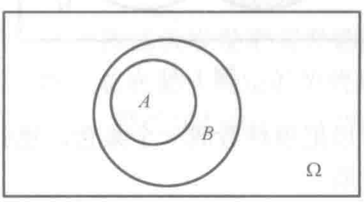

图 1-2

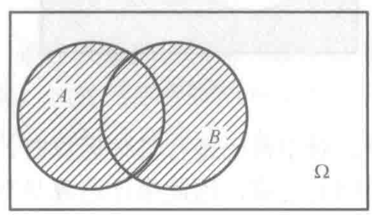

图 1-3

类似地，称“n 个事件 $A_1,A_2,\cdots,A_n$，中至少有一个发生”的事件为 n 个事件 $A_1,A_2,\cdots,A_n$ 的和，记为 $\bigcup\limits_{i=1}^{n}A_i=A_1\cup A_2\cup\cdots\cup A_n$；称“可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 中至少有一个发生”的事件为可列个事件 $A_1,A_2,\cdots A_n,\cdots$ 的和，记为 $\bigcup\limits_{i=1}^{\infty}A_i=A_1\cup A_2\cup\cdots\cup A_n\cup\cdots$。

(4) 事件的交（或积）：事件 A 和事件 B 的交事件是指事件 A 和事件 B 同时发生，记作 $A\cap B$ 或 AB，也称作事件 A 和事件 B 的积事件（见图 1-4）。

类似地，称“n 个事件 $A_1,A_2,\cdots,A_n$ 同时发生”的事件为 n 个事件 $A_1,A_2,\cdots,A_n$ 的积，记为 $\bigcap\limits_{i=1}^{n}A_i=A_1\cap A_2\cap\cdots\cap A_n$；称“可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 同时发生”的事件为可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 的积，记为 $\bigcap\limits_{i=1}^{\infty}A_i=A_1\cap A_2\cap\cdots\cap A_n\cap\cdots$。

(5) 事件的互斥（互不相容）：若事件 A 和事件 B 在同一次试验中不能同时发生，则称事件 A 和事件 B 是互斥（互不相容）的，见图 1-5。按照集合的观点，应有 $AB=\phi$。

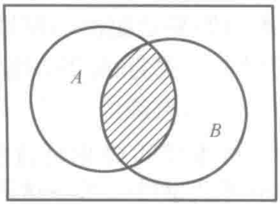

图 1-4

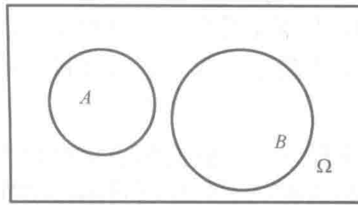

图 1-5

互斥关系可推广到多个事件，事件 $A_1, A_2, \cdots, A_n$ 两两互斥是指对于任意 $1 \leqslant i < j \leqslant n$，$A_i$ 与 A_j 互斥，即 $A_i A_j = \phi$。

(6) 对立事件 $\overline{A}$：称事件 A 不发生为事件 A 的对立事件或补事件，记作 $\overline{A}$，见图 1-6。事件 A 和它的补事件 $\overline{A}$ 满足 $A \cup \overline{A} = \Omega$，$A\overline{A} = \phi$，所以事件 A 和其补事件 $\overline{A}$ 是互斥的特例，同时还有 $\overline{\overline{A}} = A$。

(7) 事件的差：事件 A 与事件 B 的差事件是指事件 A 发生而事件 B 不发生，记作 $A - B$，见图 1-7。根据交事件和补事件的定义，可得 $A - B = A\overline{B} = A - AB$。

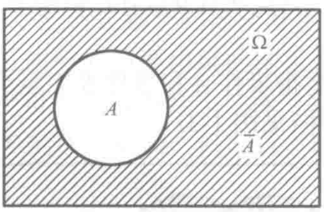

图 1-6

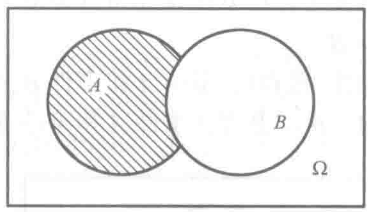

图 1-7

在实际中，对事件间进行具体运算时，完全可以把事件看成一个集合，也就是事件间的运算就是集合间的运算，因此事件运算具有以下性质。

设 A, B, C 为事件，则有：

(1) $A \cup B = B \cup A$，$AB = BA$（交换律）；

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

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

(4) $\overline{A \cup B} = \overline{A}\overline{B}$，$\overline{AB} = \overline{B} \cup \overline{A}$（对偶律）。

更一般的有

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i} \tag{1.1}$$

例 5 一位同学某学期有三门考试课程，以字母 A, B, C 分别表示这三门课程通过考试的三个事件，请用事件的运算表示下列事件：

(1) 三门课程全部通过考试； (2) 三门课程中至少有一门通过考试；

(3) 三门课程中恰好有一门未通过考试； (4) 三门课程中全部未通过考试；

(5) 三门课程中最多有一门通过考试； (6) 三门课程中至少有一门未通过考试。

解 (1) ABC； (2) $A \cup B \cup C$； (3) $\overline{A}BC \cup A\overline{B}C \cup AB\overline{C}$；

(4) $\overline{A}\overline{B}\overline{C}$； (5) $\overline{A}BC \cup A\overline{B}C \cup AB\overline{C} \cup \overline{A}\overline{B}\overline{C}$； (6) $\overline{ABC}$。

第二节 概 率

在研究随机现象时，尽管随机事件在一次试验中可能发生，也可能不发生，但在一次试验中随机事件发生的可能性大小是客观存在的，并且是可以度量的。为此我们称随机试验中事件 A 发生的可能性大小为事件 A 的概率，记作 $P(A)$。但是对于某个特定的事件 A，概率 $P(A)$ 是怎样规定的呢？我们先介绍概率的统计定义，然后再介绍概率的古典定义及几何定义，最后引进概率的公理化定义。

一、概率的统计定义

首先引入频率，它描述了事件发生的频繁程度。

定义 1.1 在相同的条件下，重复 n 次试验，随机事件 A 在 n 次试验中出现的次数 μ_n 称为频数，比值 $\frac{\mu_n}{n}$ 称为事件 A 发生的频率，记为 $f_n(A)=\frac{\mu_n}{n}$。

可以验证，当试验次数 n 固定时，事件 A 的频率 $f_n(A)$ 有如下的性质：

(1) 非负性：$0\leqslant f_n(A)\leqslant 1$；

(2) 规范性：$f_n(\Omega)=1$；

(3) 有限可加性：若 $A_1,A_2,\cdots,A_k$ 互不相容，则 $f_n\left(\bigcup\limits_{i=1}^{k}A_i\right)=\sum\limits_{i=1}^{k}f_n(A_i)$。

尽管事件 A 在一次试验中发生与否是偶然的，但在大量的试验中，事件 A 发生的频率却随着试验次数的增大总在某一确定的数值附近摆动，这种规律性称为频率的稳定性。

例如，在历史上，曾经有人做过大量抛硬币的试验，表 1-1 就记录了掷一枚均匀硬币的结果。

表 1-1

试验者	投掷次数 n	正面出现次数 μ_n	正面出现的频率 $\frac{\mu_n}{n}$
蒲丰(Buffon)	4040	2048	0.5069
德摩根(Demoran)	4092	2048	0.5005
费勒(Feller)	10 000	4979	0.4979
皮尔逊(Pearson)	12 000	6019	0.5016
皮尔逊	24 000	12 012	0.5005
罗曼诺夫斯基	80 640	39 699	0.4923

从表 1-1 可以看出，事件 A 出现的频率总是围绕 0.5 上下波动，且越来越接近 0.5。

这类例子不胜枚举。这一切表明，在大量试验中事件 A 具有频率的稳定性，也就是统计规律性，为此我们引入概率的统计定义。

定义 1.2(概率的统计定义) 在相同条件下重复进行 n 次试验中，事件 A 出现 μ_n 次，当 n 无限增大时，事件 A 发生的频率 $f_n(A)=\frac{\mu_n}{n}$ 的稳定值 p 称为事件 A 的概率，记作 $P(A)$。一般来说，n 越大，摆动幅度越小。

基于频率与统计概率的这种联系，统计概率应满足：

(1) $0 \leqslant P(A) \leqslant 1$；

(2) $P(\Omega)=1$；

(3) 若事件 $A_1, A_2, \cdots, A_n$ 两两互斥，则 $P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} P(A_k)$。

根据概率的统计定义，我们让试验重复大量次数，计算频率 $f_n(A)$，以它来表征事件 A 发生的概率。这样做数学上不够严格且耗费大量的人力、物力，但对两类特殊的随机现象可不必做大量的试验，就能直接求出事件的概率，它们分别是古典概型和几何概型。我们先讨论古典概型。

二、概率的古典定义

若随机试验 E 满足如下特征：

(1) 每次试验的结果(基本事件)只有有限多个，即样本空间是有限集；

(2) 每个基本事件发生可能性相同。

则称 E 为古典概型，古典概型又称等可能概型。

定义 1.3(概率的古典定义)　若样本空间 Ω 的基本事件总数为 $|\Omega|=n$，事件 A 包含的基本事件数为 $|A|=m$，则事件 A 的概率为

$$P(A)=\frac{|A|}{|\Omega|}=\frac{m}{n} \tag{1.2}$$

公式(1.2)是古典概型概率的计算公式，关键在于计算样本空间 Ω 和事件 A 的基本事件个数，这里经常要用到基本计数原理和排列组合的相关知识。

例 6　将一枚均匀的骰子抛掷两次，求两次出现的点数之和等于 8 的概率。

解　显然样本空间 $\Omega=\{(i,j) \mid i,j=1,2,\cdots,6\}$，设 A 表示“两次出现的点数之和等于 8”的事件，则 $A=\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$，于是有 $P(A)=\frac{|A|}{|\Omega|}=\frac{5}{36}$。

例 7　在 100 件产品中有 10 件次品，现从中随机抽取 2 件，求下列事件的概率：

(1) 有放回抽取，先任取 1 件，观察后放回，再任取 1 件，求两件都是次品的概率；

(2) 无放回抽取，先任取 1 件，不放回，再任取第 2 件，求第 1 件是次品、第 2 件为正品的概率。

解　以 A,B 分别表示(1)、(2)两个事件。

(1) 显然样本空间 Ω 的基本事件总数为 100×100，事件 A 包含的基本事件数为 10×10，所以 $P(A)=\frac{|A|}{|\Omega|}=\frac{1}{100}$。

(2) 同理可得 $P(B)=\frac{|B|}{|\Omega|}=\frac{10\times 90}{100\times 99}=\frac{1}{11}$。

例 8　(分房问题)有 n 个人，每个人都以同样的概率 $\frac{1}{N}(n \leqslant N)$ 被分配在 N 间房的任意一间里去住，试求下列事件的概率：

(1) A=“某指定的 n 个房间中各有一人住”；

(2) B=“恰有 n 个房间，其中各有一人去住”。

解　把第一个人分配到 N 间房中之一去有 N 种可能，第二个人分配到 N 间房中之一去有 N 种可能。那么，对于 n 个人来说就有 N^n 种分配方法，即样本空间中包含有基本事件总数为 N^n 个。

(1) 事件 A 即“某指定的 n 个房间中各有一人住”，相当于 n 的一个全排列，即 A 包含的样本点个数为 $n!$，则 $P(A)=\frac{n!}{N^n}$；

(2) 事件 B 中“恰有 n 间房”可以从 N 个房间中任意挑选出来时有 C_N^n 种选法。对每一种选法有 $n!$ 种住法。因此事件 B 所包含的样本点的个数是 $C_N^n \cdot n!$，

则 $P(B)=\frac{C_N^n \cdot n!}{N^n}=\frac{N(N-1)\cdots(N-n+1)}{N^n}$

分房问题是相当广泛的一类问题，很多实际例子都可以归结为它而得到解决。历史上著名的“生日问题”也可以用此模型来解决。

例 9　某次校友聚会有 n 个人（$n \leqslant 365$）。试求下列各事件的概率：

(1) $A=$“所有的人生日全不相同”；

(2) $B=$“至少有两人生日相同”。

解　(1)把一年 365 天看作 365 个房间，假定每个人的生日是一年中任何一天的概率为 $\frac{1}{365}$。那么，事件 A 可看作“恰有 n 个房间，其中各住一人”，则所求的概率为

$$P(A)=\frac{C_N^n \cdot n!}{N^n}=\frac{365 \cdot 364 \cdots (365-n+1)}{365^n}$$

(2) 事件 A 与事件 B 是对立事件，故 $P(B)=1-P(A)=1-\frac{365 \cdot 364 \cdots (365-n+1)}{365^n}$，表 1-2 给出了 $P(A)$ 与 $P(B)$ 的近似值。

表 1-2

n	10	20	22	23	30	40	50	55
$P(A)$	0.88	0.59	0.52	0.49	0.29	0.11	0.03	0.01
$P(B)$	0.12	0.41	0.48	0.51	0.71	0.89	0.97	0.99

从上表可以看出，在仅有 50 人的班级里，“至少有两人生日相同”这一事件的概率与 1 相差无几。因此，如果进行调查，几乎总是会出现的，读者不妨试一试。

三、概率的几何定义

古典概型利用了等可能性计算概率，但要求样本空间必须为有限集，这使得无限样本空间的问题不能适用，于是我们去掉对样本空间为有限集的限制，就得到了几何概型的定义。

若随机试验 E 满足如下两个特征：

(1) 每次试验的结果是无限多个，即样本空间是无限点集，可用某一有限的几何区域(直线、平面或三维空间等)表示，且可以度量该区域的大小(长度、面积、体积等)；

(2) 每次试验的各种结果等可能地出现。

则称 E 为几何概型。

定义 1.4(概率的几何定义)　设几何概型的样本空间可表示成有度量的区域，仍记为 Ω，

若事件 A 所对应的区域仍以 A 表示（$A\subset\Omega$），则定义事件 A 的概率为

$$P(A)=\frac{A\text{的度量}}{\Omega\text{的度量}}=\frac{S_A}{S_\Omega} \tag{1.3}$$

其中，S_A、S_Ω 分别表示 A、Ω 的度量。

公式(1.3)是几何概型概率的计算公式，关键在于选取合适的参数从而把问题几何化，找到样本空间 Ω 和事件 A 对应的几何区域，分别计算它们的度量。当计算面积时，经常要用到定积分的相关知识。

例 10　设某公共汽车站每间隔 5 分钟有一辆公交车到站，乘客到达汽车站的时刻是任意的，求一位乘客候车时间不超过 3 分钟的概率。

解　设一辆公交车 0 时刻到站，则下一辆公交车到达时刻为 5。若 t 表示乘客的等待时间，A 表示“候车不超过 3 分钟”的事件，则样本空间为 $\Omega=\{t\mid 0\leqslant t\leqslant 5\}$，事件 $A=\{t\mid 0\leqslant t\leqslant 3\}$，所求概率为 $P(A)=\dfrac{S_A}{S_\Omega}=\dfrac{3}{5}=0.6$。

例 11　设 x 和 y 是任意两个小于 1 的正数，求 $xy<\dfrac{1}{2}$ 的概率。

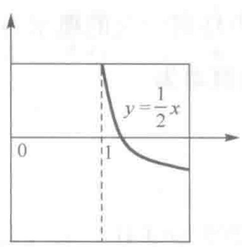

图 1-8

解　本题相当于在平面区域 $\Omega=\{(x,y)\mid 0<x<1,0<y<1\}$ 内任意选一点，求此点落在区域 $A=\left\{(x,y)\mid (x,y)\in\Omega, xy<\dfrac{1}{2}\right\}$ 内（见图 1-8）的概率。

因为平面区域 Ω 是边长为 1 的正方形，其面积为 1，而 A 的面积为 $\dfrac{1}{2}+\displaystyle\int_{1/2}^{1}\frac{1}{2x}\mathrm{d}x=\frac{1}{2}+\frac{1}{2}\ln 2$，

所以 $P(A)=\dfrac{A\text{的面积}}{\Omega\text{的面积}}=\dfrac{\frac{1}{2}+\frac{1}{2}\ln 2}{1}=\dfrac{1}{2}+\dfrac{1}{2}\ln 2$。

四、概率的公理化定义

前面讨论的概率的古典定义和几何定义都具有和统计定义一样的性质，即非负性、规范性和有限可加性，但它们都是在特殊的随机试验 E 下给出的概率的计算方法，具有明显的极限性，不能作为事件概率的严格定义。所以要寻求考虑的一般定义，使其具有普遍性、严密性。苏联概率学家 Kolmogorov 在总结前人的研究成果后，于 1933 年在他的著作《概率论的基本概念》中提出了概率的公理化定义，具体定义如下。

定义 1.5　设随机试验 E 的样本空间为 Ω，若对每一个事件 A，有且只有一个实数 $P(A)$ 与之对应，且满足如下公理：

(1) 非负性：$0\leqslant P(A)\leqslant 1$；

(2) 规范性：$P(\Omega)=1$；

(3) 完全可加性：若对任一列两两互斥事件 $A_1,A_2,\cdots,A_n,\cdots$，有

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n) \tag{1.4}$$

则称 $P(A)$ 为事件 A 的概率。