魏艳华 王丙参 ◎ 编著

概率论与数理统计

GAILüLUN YU SHULI TONGJI

SWJUP 西南交通大学出版社
Http://press.swjtu.edu.cn

概率论与数理统计

魏艳华　王丙参　编著

西南交通大学出版社

内容简介

本书依据最新《普通高中数学新课程标准》及《全国硕士研究生入学统一考试数学考试大纲》对概率论与数理统计的要求，以一种新的体系，系统全面介绍了概率论与数理统计的理论及其应用、软件操作方法. 全书分为 11 章，内容包括事件与概率，随机变量及其分布，随机变量的数学特征，大数定律及中心极限定理，统计量及数据初步处理，参数估计与假设检验，方差分析与试验设计，回归分析，判别分析与聚类分析，随机过程，随机模拟等内容. 在附录中给出了 Matlab、SAS 软件的简明教程，概率论与数理统计的发展历史及普通高中数学新课程标准解读，供读者课外阅读.

本书内容新颖，理论与实际密切结合，将重要的统计原理与上机操作相结合，还将公式的计算与计算机操作方法并列处理，相互对照，通过详尽的典型实例分析，加深对相关理论的理解，尽最大可能增加本书的实用性. 本书数值计算运用 Matlab 软件完成，案例主要运用国内流行的 SPSS 软件完成，部分内容用 Excel 和 SAS 软件完成，所讲述的方法都结合实例介绍软件的实施过程，为广大研究人员和师生对相关理论的应用提供借鉴.

本书可作为高等院校数学与应用数学、统计学、金融数学和保险精算等各专业的本科生教材，也可作为相关专业的参考用书及重点大学经管学院的概率论与数理统计教材.

图书在版编目（CIP）数据

概率论与数理统计 / 魏艳华，王丙参编著. —成都：西南交通大学出版社，2013.8
ISBN 978-7-5643-2324-0

Ⅰ. ①概… Ⅱ. ①魏… ②王… 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（2013）第 119337 号

概率论与数理统计
魏艳华　王丙参　编著
*
责任编辑　张宝华
封面设计　何东琳设计工作室
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://press.swjtu.edu.cn
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：19.75
字数：495 千字
2013 年 8 月第 1 版　2013 年 8 月第 1 次印刷
ISBN 978-7-5643-2324-0
定价：36.00 元

图书如有印装质量问题　本社负责退换

前 言

统计学是一门通过搜索、整理、分析数据等手段，以达到推断所测对象的本质及预测对象未来的一门综合性科学，其中用到了大量数学及其他学科的专业知识，使用范围几乎覆盖了社会与自然科学的各个领域. 调整后的统计学一级学科将原属应用经济学和数学下与统计相关的学科进行了整合，并在一级学科下设有数理统计、社会经济统计、金融统计与风险管理和精算、应用统计等二级学科，各校根据情况授予经济学或理学学位.

概率论与数理统计是全国高等院校数学与统计学院的基础课程，它以丰富的背景、巧妙的思维和有趣的结论吸引着读者，使学生在浓厚的兴趣下学习和掌握基本的概念、理论和方法.

目前，中学对概率论与数理统计的要求越来越高，最新的《普通高中数学新课程标准》给出了详细要求，总体而言，淡化概率论，强化统计学，侧重运用计算机进行实例分析，涉及面之广达到前所未有的程度. 为了满足数学教育专业对概率论与数理统计及培养应用型人才的要求，编者在多次讲授本课程讲稿的基础上，结合同行专家的优秀成果及作者对本课程的研究，经过多次修订和补充写成本书. 本书可作为高等院校数学与应用数学、统计学、金融数学和保险精算等各专业的本科生教材，也可作为相关专业的参考用书及重点大学经管学院的概率论与数理统计教材.

在编写过程中，我们注重基本理论、概念、方法的叙述，关注能力的培养. 在数学工具上，力求简洁、准确，尽量使读者运用浅显的数学工具系统地掌握概率论与数理统计的基本理论和方法. 同时，为了金融数学、保险精算专业尽快进入专业课的学习，本书列举了大量金融、保险方面的例子. 此外我们也参考了最新颁布的《全国硕士研究生入学统一考试数学考试大纲》的要求，力求教材的体系、内容既符合数学学科本科生的特点，又兼顾报考研究生的学生需求. 如何使数理统计的学习者兼为统计分析和数据处理的实践者，是高等学校数理统计教学改革的重要课题，本书反映了作者在这一方面的思考和实践. 为了保持本书的独特之处，本书融合了编者最近对相关理论的研究和教学体会，并理论联系实际，将重要的统计原理与上机操作相结合，还将公式的计算与计算机操作方法并列处理，相互对照，通过详尽的典型实例分析，加深对相关理论的理解，尽最大可能增加本书的实用性. 本书数值计算运用 Matlab 软件完成，案例主要运用国内流行的 SPSS 软件完成，部分内容用 Excel 和 SAS 软件完成，所讲述的方法都结合实例介绍软件的实施过程.

全书共分 11 章. 第 1 章介绍了随机事件和概率，举例讲解了贝叶斯决策；第 2 章详细讲解了随机变量及其分布，重点学习一维、二维随机变量，多维随机变量作为了解；第 3 章论述了随机变量的数学期望，重点讨论了风险型决策；第 4 章重点研究了概率论中的有关变换、大数定律及中心极限定理在保险精算中的应用；第 5 章讨论了统计量的分布及描述统计；第 6 章介绍了参数估计和假设检验，贝叶斯估计作为了解内容；第 7 章为方差分析与试验设计；第 8 章为回归分析，主要讲解一元线性回归，多元线性回归及非线性回归作为了解；第 9 章浅显论述了判别分析与聚类分析；第 10 章讲了随机过程的基本概念，重点讨论了马尔可夫预

测与决策；第 11 章研究了随机数的产生和随机模拟，给出了模拟程序. 每个专题既有相关理论的简单讲解，也配有实用的案例分析，理论与实践相结合，案例既可以作为读者扩宽视野、提高分析水平的学习资料，也可直接作为模版应用于对实际问题的处理.

若概率论与数理统计分两学期开设，每学期 60 学时，本书可以在 120 学时左右讲完. 若作为一门课在一学期开设，可选择部分内容组织教学，略去带*号内容，大致 90 学时讲完. 任课老师可根据实际需要合适安排各章学时.

本书由天水师范学院的魏艳华、王丙参共同编写，具体分工为：第 1、2、3、4、10、11 章及附录 3 由王丙参编写，其余由魏艳华编写，我们经常讨论、切磋写法、选择例题、相互补充，经过反复讨论和修改后由魏艳华定稿. 在本书的编写过程中得到了天水师范学院数学与统计学院的大力支持，统计教研室同事为本书提供了很多宝贵的意见和建议，也得到了西南交通大学出版社张波主任、宋彦博、编辑张宝华及有关各方和同仁的大力支持，特在此一并致以诚挚的谢意！

虽然我们希望编写出一本质量较高、适合当前教学实际需要的教材，但由于编著者水平有限，本书书中难免存在不妥之处，恳切希望读者批评、指正，使本教材不断得以完善. 为方便广大读者，提供支持电子邮箱：wangbingcan2000@163.com.

读者可通过该邮箱与作者取得联系，获取技术支持和教学资料.

魏艳华　王丙参

2013 年 1 月

目　录

第 1 章　随机事件和概率 …… 1
§1.1　随机事件及其运算 …… 1
§1.2　概率的定义及性质 …… 4
§1.3　概率的确定方法 …… 8
§1.4　条件概率 …… 17
§1.5　独立性 …… 24
习题 1 …… 29

第 2 章　随机变量及其分布 …… 32
§2.1　随机变量及其分布 …… 32
§2.2　离散型随机变量 …… 38
§2.3　连续型随机变量 …… 45
§2.4　随机变量函数的分布 …… 51
§2.5　多维随机变量及其分布 …… 54
习题 2 …… 64

第 3 章　随机变量的数字特征 …… 67
§3.1　数学期望与方差 …… 67
§3.2　协方差与相关系数 …… 77
§3.3*　随机向量的数字特征与多元正态分布 …… 81
§3.4　随机变量的其他特征数 …… 83
§3.5　期望效用理论与保险业的存在性 …… 85
习题 3 …… 89

第 4 章　大数定律与中心极限定理 …… 91
§4.1　概率论中常用的几个变换 …… 91
§4.2　大数定律与中心极限定理 …… 98
习题 4 …… 106

第 5 章　统计量及数据的初步处理 …… 108
§5.1　总体与样本 …… 108
§5.2　统计量及其分布 …… 111
§5.3　数据的收集 …… 119
§5.4　数据的整理与显示 …… 123
习题 5 …… 135

第 6 章　参数估计与假设检验 …… 136
§6.1　点估计 …… 136
§6.2　估计量的评价标准 …… 140
§6.3　区间估计 …… 145

§6.4* 贝叶斯估计……151
§6.5 分布参数的假设检验……157
§6.6 分布拟合优度检验……165
§6.7 属性变量的关联性检验……172
习题 6……178
第 7 章 方差分析与试验设计……181
§7.1 单因素方差分析……181
§7.2 两因素等重复试验的方差分析……189
§7.3 正交试验设计的基本方法……195
习题 7……201
第 8 章 回归分析……203
§8.1 变量间的统计关系……203
§8.2 一元线性回归……208
§8.3 多元线性回归模型……214
§8.4 非线性回归……218
习题 8……223
第 9 章 判别分析与聚类分析……225
§9.1 判别分析……226
§9.2 聚类分析……233
习题 9……242
第 10 章 随机过程……244
§10.1 随机过程的基本概念……244
§10.2 几类重要的随机过程……246
§10.3 马尔可夫链的基本概念……250
§10.4 马尔可夫预测与决策……254
习题 10……262
第 11 章 随机模拟……263
§11.1 利用反函数及变换抽样法生成随机数……263
§11.2 蒙特卡罗方法与积分的计算……269
习题 11……276
附录 1 MATLAB 简明教程……277
附录 2 SAS 软件简明教程……285
附录 3 概率论与数理统计简介……290
附录 4 全日制普通高中数学新课程标准解读——概率与统计……295
附录 5 标准正态分布表……300
部分习题解答……302
参考文献……308

第 1 章　随机事件和概率

§ 1.1　随机事件及其运算

1.1.1　随机现象

在自然界与人类社会生活中，存在着两类截然不同的现象：一类是**确定性现象**，例如，早晨太阳必然从东方升起；在标准大气压下，水加热到 100 摄氏度必然沸腾. 这类现象的特点是：在试验之前就能断定它有一个确定的结果，即在一定条件下，重复进行试验，其结果必然出现且唯一. 而另一类是**随机现象**，它是概率论与数理统计的研究对象，例如，某地区的年降雨量；打靶射击时，弹着点离靶心的距离. 因此，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**. 这类现象有两个特点：

（1）结果不止一个；

（2）哪一个结果出现，人们事先不能确定.

例 1.1.1　随机现象到处可见，比如

（1）抛一枚硬币，可能正面朝上，也可能反面朝上；

（2）一天内进入某超市的人数；

（3）明天某时刻天气的温度；

（4）测量某物理量的误差.

概率论与数理统计就是研究随机现象统计规律性的一门学科. 以前由于随机现象事先无法判定将会出现哪种结果，人们就以为随机现象是不可捉摸的，无法研究和预测，但后来人们通过大量实践发现：在相同条件下，虽然个别试验结果在某次试验或观察中可以出现也可以不出现，但在大量试验中却呈现出某种规律性，这种规律性称为**统计规律性**. 例如：在投掷一枚硬币时，既可能出现正面，也可能出现反面，预先作出确定的判断是不可能的，但是假如硬币均匀，直观上出现正面与反面的机会应该相等，即在大量的试验中出现正面的频率应接近 50%，这正如恩格斯所指出的："在表面上是偶然性在起作用的地方，这种偶然性始终是受内部隐藏着的规律支配的，而问题只是在于发现这些规律."

在相同的条件下可以重复的随机现象称为**随机试验**，即一个试验如果满足：

（1）可以在相同的条件下重复进行；

（2）其结果具有多种可能性；

（3）在每次试验前，不能预言将出现哪一个结果，但知道其所有可能出现的结果；

则称这样的试验为随机试验.

简言之，**随机试验是对随机现象的一次观察或试验**，通常用大写字母 E 表示，简称**试验**. 例 1.1.1 中（1）、（4）是随机试验，（2）、（3）由于不能重复进行（历史不可重演），故它们虽然是随机现象，但不是随机试验.

概率论与数理统计主要研究大量重复的随机现象，但也十分注重研究不能重复的随机现象，比如研究天气温度的变化过程.

1.1.2 样本空间与样本点

随机试验的一切可能的基本结果组成的集合称为**样本空间**，用Ω表示；其每个元素称为**样本点**，又称为基本结果，用ω表示.

例 1.1.2 下面给出随机现象的样本空间.

（1）E_1：投一枚均匀硬币两次，观察出现正反面情况，记z为正面，f为反面，则$\Omega_1=\{(z,z),(z,f),(f,f),(f,z)\}$；

（2）E_2：电话总机在单位时间内接到的呼唤次数，则$\Omega_2=\{0,1,2,\cdots\}$；

（3）E_3：测量误差，则$\Omega_3=\mathbf{R}$.

对于样本空间，我们要注意以下几点：

（1）样本空间是一个集合，由样本点构成，表示方法有：列举法、描述法.

（2）样本点可以是一维的，也可以是多维的，可以是有限个，也可以是无限个.

（3）对于一个随机试验而言，样本空间并不唯一. 在同一试验中，当试验的目的不同时，样本空间往往是不同的，但通常只有一个能提供最多信息的样本空间. 例如，在运动员投篮的试验中，若试验的目的是考察命中情况；则样本空间$\Omega=\{中,不中\}$；若试验的目的是考察得分情况，则样本空间$\Omega=\{0分,1分,2分,3分\}$.

在今后数学处理上，我们往往将据样本点个数为有限个或可列个的情况归为一类，称为**离散的样本空间**. 而将样本点为不可列无限多的情况归为一类，称为**连续的样本空间**. 由于两类样本空间有着本质差异，故分别称呼之.

给定集合A,B，若存在A到B上的一一映射，则称A与B**对等**，记作$A\sim B$. 如果两个集合对等，称它们具有相同的**势**. 若$A\sim\mathbf{N}$，其中$\mathbf{N}$为自然数集，则称A为**可数集（可列集）**. 不是可数集的无限集称为**不可数集（不可列集）**.

1.1.3 随机事件

样本空间Ω的某个子集称为**随机事件**，简称**事件**，常用大写字母A,B,C等表示. 随机事件包括基本事件和复合事件. 由一个样本点构成的集合称为**基本事件**；由多个样本点构成的集合称为**复合事件**.

某个事件A发生当且仅当A所包含的一个样本点ω出现，记为$\omega\in A$. 例如，在掷骰子的试验中，事件A：“出现偶数点”，且用ω_i表示“出现i点”，则A包含$\omega_2,\omega_4,\omega_6$这三个样本点，所以$A$是复合事件. “出现2点”就意味着$A$发生，并不要求$A$的每一个样本点都出现.

Ω也是自身的子集，称为**必然事件**，指在每次试验中都必然发生的事件. $\varnothing$称为**不可能事件**，指每次试验都必然不会发生的事件. 严格来讲，必然事件与不可能事件反映了确定性现象，也可以说它们并不是随机事件，但为了研究问题的方便，常把它们作为特殊随机事件来处理，即**退化的随机事件**.

1.1.4　随机事件之间的关系

（1）**事件的包含**：事件 A 发生必然导致事件 B 发生，则称 A 包含于 B 或 B 包含 A，记为 $A\subset B$ 或 $B\supset A$，即 $A\subset B \Leftrightarrow \{若\omega\in A, 则\omega\in B\}$.

反之，$B\supset A \Leftrightarrow$ 若 B 不发生，则 A 必然也不会发生.

注：很多教材对 $\subset$ 与 $\subseteq$ 不加区分，认为二者等价.

显然有：①$\varnothing\subset A\subset\Omega$；② 若 $A\subset B$，$B\subset C$，则 $A\subset C$.

（2）**事件的相等**：若事件 $A\subset B$ 且 $B\subset A$，则称 A 与 B 相等，记为 $A=B$.

（3）**事件的互斥**：若事件 A 与 B 不能同时发生，则称 A 与 B **互斥**（**互不相容**），记为 $AB=\varnothing$. 显然有：① 基本事件是互斥的；② $\varnothing$ 与任意事件互斥.

1.1.5　随机事件的运算

（1）**事件的并**：两个事件 A,B 中至少有一个发生的事件，称为事件 A 与 B 的并（或和），记为 $A\cup B$（或 $A+B$），即 $A\cup B=\{\varpi\,|\,\varpi\in A 或 \varpi\in B\}$.

显然有：① $A\cup A=A$；

② $A\subset A\cup B$，$B\subset A\cup B$；

③ 若 $A\subset B$，则 $A\cup B=B$.

（2）**事件的交**：两个事件 A 与 B 同时发生的事件，称为事件 A 与 B 的交（或积），记为 $A\cap B$（或 AB），即 $A\cap B=\{\omega\,|\,\omega\in A 且 \omega\in B\}$.

显然有：① $A\cap B\subset A$，$A\cap B\subset B$；

② 若 $A\subset B$，则 $A\cap B=A$. 特别地，$A\Omega=A$；

③ 若 A 与 B 互斥，则 $AB=\varnothing$. 特别地，$A\varnothing=\varnothing$.

为直观地表示事件及事件的关系，在概率论中常用长方形表示样本空间 Ω，用一个圆或其他几何图形表示事件 A，点表示样本点 ω_1，见图 1.1.1，这类图形称为维恩（Venn）图.

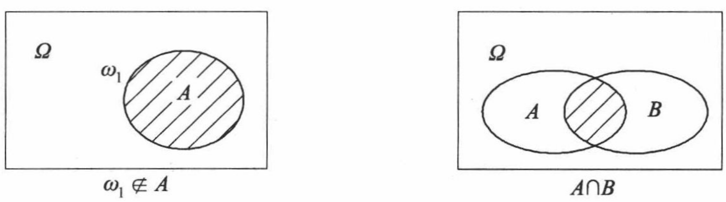

图 1.1.1　事件的维恩图

事件之间的和、积运算可以推广到有限与可列无穷多个事件的情形.

$$\bigcup_{k=1}^{n}A_k=A_1\cup A_2\cup\cdots\cup A_n,\qquad \bigcup_{k=1}^{\infty}A_k=A_1\cup A_2\cup\cdots\cup A_n\cup\cdots;$$

$$\bigcap_{k=1}^{n}A_k=A_1\cap A_2\cap\cdots\cap A_n,\qquad \bigcap_{k=1}^{\infty}A_k=A_1\cap A_2\cap\cdots\cap A_n\cap\cdots.$$

（3）**事件的差**：事件 A 发生而事件 B 不发生的事件，称为事件 A 与事件 B 的差，记为 $A-B$，即 $A-B\Leftrightarrow\{\omega\in A 而 \omega\notin B\}$.

显然有：① $A-B=A-AB$，不要求 $A\supset B$，才有 $A-B$，若 $A\subset B$，则 $A-B=\varnothing$；

② 若 A 与 B 互斥，则 $A-B=A, B-A=B$；

③ $A-(B-C)\neq A-B+C$，$(A-B)\cup B=A\cup B\neq A$.

（4）**事件的逆**：若事件 A 与 B 满足 $A\cup B=\Omega$ 且 $AB=\varnothing$，则称 B 为 A 的逆，记为 $B=\overline{A}$，即 $\overline{A}=\{\omega\mid\omega\notin A,\ \omega\in\Omega\}$，

$A,\overline{A}$ 称为互逆事件或对立事件.

显然有：① $A\cup\overline{A}=\Omega$，$A\cap\overline{A}=\varnothing$；② $A-B=A\overline{B}$；

互逆事件与互斥事件的区别：① 互逆必定互斥，互斥不一定互逆；② 互逆只在样本空间只有两个事件时存在，互斥还可在样本空间有多个事件时存在. 例如，在抛硬币试验中，设 $A=\{$正面$\}$，$B=\{$反面$\}$，则 A 与 B 互斥与互逆等价，而在掷骰子试验中，设 $A=\{1$点$\}$，$B=\{2$点$\}$，则 A 与 B 互斥但不互逆.

1.1.6 事件的运算性质

由前面可知，事件之间的关系与集合之间的关系建立了一定的对应法则，因而事件之间的运算法则与布尔代数中集合的运算法则相同.

（1）交换律：$A\cup B=B\cup A$，$AB=BA$；

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

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

（4）德莫根（对偶）定律：

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i}\ （和的逆=逆的积），\quad \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}\ （积的逆=逆的和）.$$

例 1.1.3 设 A,B,C 为任意三个事件，试用 A,B,C 的运算关系表示下列事件：

（1）三个事件中至少一个发生：$A\cup B\cup C$.

（2）没有一个事件发生：$\overline{A}\,\overline{B}\,\overline{C}=\overline{A\cup B\cup C}$（由对偶律）.

（3）恰有一个事件发生：$A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$.

（4）至多有两个事件发生：

$$(A\overline{B}C\cup AB\overline{C}\cup\overline{A}BC)\cup(\overline{A}\,\overline{B}C\cup\overline{A}B\overline{C}\cup A\overline{B}\,\overline{C})\cup(\overline{A}\,\overline{B}\,\overline{C})=\overline{ABC}=\overline{A}\cup\overline{B}\cup\overline{C}.$$

（5）至少有两个事件发生：$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC\cup ABC=AB\cup BC\cup CA$.

§ 1.2 概率的定义及性质

随机事件在一次试验中，可能发生也可能不发生，具有偶然性，但人们从实践中认识到，在相同的条件下，进行大量重复试验，试验的结果具有某种内在规律性，即随机事件发生可能性的大小是可以比较的，是可以用一个数字进行度量的. 例如，在投掷一枚均匀的骰子试验中，事件 A 表示"掷出偶数点"，B 表示"掷出 2 点"，显然事件 A 比事件 B 发生的可能性要大.

对于一个随机试验，我们不仅要知道它可能出现哪些事件，更重要的是研究事件发生可

能性的大小，也就是事件的概率，从而揭示其内在的规律性.

概率是随机事件发生可能性大小的度量，介于 0 与 1 之间，事件发生的可能性越大，概率就越大. 对于事件 A，通常用 $P(A)$ 来表示事件 A 发生的可能性大小，即 A 发生的概率，但事件的概率是如何进行定义呢？

1.2.1　概率的公理化定义

定义 1.2.1　如果 $\mathcal{F}$ 是样本空间 Ω 中的某些子集的集合，它满足：

（1）$\Omega \in \mathcal{F}$；

（2）若 $A \in \mathcal{F}$，则 $\overline{A} \in \mathcal{F}$；

（3）若 $A_i \in \mathcal{F}\ (i=1,2,\cdots)$，则 $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

则称 $\mathcal{F}$ 为 σ 代数，也称为事件域.

在概率论中，$(\Omega,\mathcal{F})$ 又称为**可测空间**，这里可测指的是 $\mathcal{F}$ 中的元素都具有概率，即都是可度量的.

定义 1.2.2　设 Ω 是一个样本空间，$\mathcal{F}$ 是由 Ω 的某些子集组成的一个事件域，如果对 $\forall A \in \mathcal{F}$，定义在 $\mathcal{F}$ 上的一个实值函数 $P(A)$ 与之对应，且满足：

（1）非负性公理：$0 \leqslant P(A) \leqslant 1$；

（2）正则性公理：$P(\Omega)=1$；

（3）可列可加性公理：设 $A_i \in \mathcal{F}(i=1,2,\cdots)$，且 $A_i \cap A_j = \varnothing (i \neq j)$，有

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

则称 $P(A)$ 为事件 A 的概率，三元总体 $(\Omega,\mathcal{F},P)$ 称为概率空间.

概率的公理化定义刻画了概率的本质，即**概率是集合的函数且满足上述三条公理**. 事件域的引进使我们的模型有了更大的灵活性，在实际问题中可以根据问题的性质选择合适的 $\mathcal{F}$，一般选 Ω 的一切子集为 $\mathcal{F}$. 事件域可以保证事件对交、取余、可列并、可列交等运算封闭，即随机事件经过普通运算后还是随机事件.

1.2.2　概率的性质

利用概率的公理化定义，可导出概率的一系列性质.

性质 1.2.1　$P(\varnothing)=0$.

证明　因为 $\varnothing = \varnothing \cup \varnothing \cup \cdots$，$\varnothing \cap \varnothing = \varnothing$，由可列可加性得

$$P(\varnothing) = P(\varnothing \cup \varnothing \cup \cdots) = P(\varnothing) + P(\varnothing) + \cdots.$$

再由非负性必有 $P(\varnothing)=0$.

请读者思考一下，不可能事件的概率一定为 0，但概率为 0 的事件可能发生么？在概率论中，将概率很小（小于 0.05）的事件称作**小概率事件**.

小概率事件原理，又称为**似然推理或实际推断原理**：在原假设成立的条件下，如果一个

事件发生的概率很小，那么在一次试验中，可以把它看成是不可能事件. 如果在一次试验中，事件发生了，则矛盾，即原假设不正确.

设某试验中出现事件 A 的概率为 p，不管 p 如何小，如果把试验不断独立地重复下去，那么 A 迟早会出现一次，从而也必然会出现任意多次，而**不可能事件是指试验中总不会发生的事件**.

人们在长期的经验中坚持这样一个观点：概率很小的事件在一次试验中与不可能事件几乎是等价的，即不会发生. 如果在一次试验中小概率事件居然发生了，人们会认为该事件的前提条件发生了变化， 或者认为该事件不是随机发生的，而是人为安排的，等等，此即为小概率原理的一个应用. 如果我们把注意仅停留在小概率事件的极端个别现象上，那么就是杞人忧天，就不敢开车，不敢吃饭，一切都不敢做了，事实上，天一定会塌下来的，但在你活着的这段时间内塌下的概率很小，杞人其实是不明白“小概率事件在一次试验中是不可能发生的”.

小概率事件原理是概率论的精髓，是统计学发展、存在的基础，它使得人们在面对大量数据而需要做出分析与判断时，能够依据具体情况的推理来做出决策，从而使统计推断具备严格的数学理论依据.

性质 1.2.2（有限可加性） 设 $A_i \in \mathcal{F}(i=1,2,\cdots,n)$，且 $A_i \cap A_j = \varnothing(i \neq j)$，则

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

证明 令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$，由 $P(\varnothing)=0$ 可得

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

推论 （1）对任意事件 A，有 $P(A)=1-P(\overline{A})$.

（2）对任意两个事件 A,B，有 $P(A-B)=P(A)-P(AB)$；

特别地，当 $A \supset B$ 时，有 $P(A-B)=P(A)-P(B)$ 且 $P(A) \geqslant P(B)$.

证明 （1）因为 $A \cup \overline{A} = \Omega,\ A \cap \overline{A} = \varnothing$，由性质 1.2.2 可得

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

移项即得结论.

（2）因为 $A=(A-B)\cup AB,\quad (A-B)\cap AB=\varnothing$，由性质 1.2.2 可得

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

移项即得结论.

很容易举例说明，若 $P(A) \geqslant P(B)$，无法推出 $A \supset B$. 推论 1 不仅在计算事件的概率时非常有用，而且在今后一些定理的证明或公式的推导过程中也常用到这一性质.

例 1.2.1 抛一枚硬币 5 次，求既出现正面又出现反面的概率.

解 令 $A=\{$既出现正面又出现反面$\}$，则 A 的情况比较复杂，而 A 的对立事件则相对简单，即 $B=\{$全出现正面$\}$ 或 $C=\{$全出现反面$\}$，且 B,C 互不相容，所以

$$P(A)=1-P(\overline{A})=1-P(B\cup C)=1-P(B)-P(C)=1-\frac{1}{2^5}-\frac{1}{2^5}=\frac{15}{16}.$$

性质 1.2.3（加法公式） 对任意两事件 A,B，有

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

证明 因为 $A\cup B=A\cup(B-A)$，且 A 与 $B-A$ 互不相容，由有限可加性得

$$P(A\cup B)=P(A\cup(B-A))=P(A)+P(B-A)=P(A)+P(B)-P(AB)$$

推广 n 个事件和的概率为

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

性质 1.2.4（概率的连续性） 设 $A_1\supset A_2\supset\cdots$，且 $\bigcap_{n=1}^{\infty}A_n=\varnothing$，则 $\lim_{n\to\infty}P(A_n)=0$.

证明 由假设可得 $A_n=\bigcup_{k=n}^{\infty}(A_k-A_{k+1})\ (n=1,2,\cdots)$，而 $A_k-A_{k+1}(k=1,2,\cdots)$ 两两互不相容，由完全可加性得

$$1\leqslant P(A_1)=P\left(\bigcup_{k=1}^{\infty}(A_k-A_{k+1})\right)=\sum_{k=1}^{\infty}P(A_k-A_{k+1}),$$

即无穷级数收敛.

由于 $P(A_n)=\sum_{k=n}^{\infty}P(A_k-A_{k+1})$ 是上面收敛级数的尾项，因而 $\lim_{n\to\infty}P(A_n)=0$.

推论 设 $A_1\supset A_2\supset\cdots$，且 $\bigcap_{n=1}^{\infty}A_n=A$，则 $\lim_{n\to\infty}P(A_n)=P(A)$.

一般称具有推论所述性质的非负实值集函数 P 为**连续的**.

读者自行证明：若设 $A_1\subset A_2\subset\cdots$，且 $\bigcup_{n=1}^{\infty}A_n=A$，则 $\lim_{n\to\infty}P(A_n)=P(A)$.

定理 1.2.1 设 P 为可测空间 $(\Omega,\mathcal{F})$ 上的非负实值集函数，$P(\Omega)=1$，则 P 为完全可加的充分必要条件为：

（1）P 是有限可加的；

（2）P 是连续的.

证明 必要性由性质 1.2.2，1.2.4 已证. 下证充分性.

设 $A_1,A_2,\cdots\in\mathcal{F}$，$A_i\cap A_j=\varnothing(i\neq j)$，令 $A=\bigcup_{i=1}^{\infty}A_i$，$B_n=\bigcup_{i=1}^{n}A_i$，$C_n=A-B_n$，则 A,B_n,C_n（一切 n）均属于 $\mathcal{F}$，且 $A\supset B_n$，同时 $C_1\supset C_2\supset\cdots$，$\bigcap_{n=1}^{\infty}C_n=\varnothing$. 由 P 的连续性假设有

$$\lim_{n\to\infty}P(C_n)=0,$$

即 $$\lim_{n\to\infty}[P(A)-P(B_n)]=0.$$

另一方面，由有限可加性得

$$P(B_n)=\sum_{i=1}^{n}P(A_i),$$

即得 $$P(A)=\sum_{i=1}^{\infty}P(A_i).$$

§1.3 概率的确定方法

公理化定义没有告诉人们如何去确定概率. 然而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，有着确定概率的方法，所以在有了公理化定义后，它们均可以作为概率的确定方法.

1.3.1 排列与组合

排列和组合都是计算“从 n 个元素中任取 r 个元素”的取法总数公式，其主要区别在于：如果不考虑取出的元素间的次序，则用组合公式，否则用排列公式. 所谓考虑元素间的次序，可以从实际问题中得以辨别.

（1）**加法原理**：设完成一件事有 m 种方式，第一种方式有 n_1 种方法，第二种方式有 n_2 种方法，……第 m 种方式有 n_m 种方法，则完成这件事共有 $\sum_{i=1}^{m}n_i$ 种方法.

譬如，某人要从甲地到乙地去，可以乘火车，也可以乘轮船，火车有两班，轮船有三班，那么甲地到乙地共有 $2+3=5$ 个班次供旅客选择.

（2）**乘法原理**：设完成一件事有 m 个步骤才能完成，第一步有 n_1 种方法，第二步有 n_2 种方法，……第 m 步有 n_m 种方法，则完成这件事总共有 $\prod_{i=1}^{m}n_i$ 种方法.

譬如，若一个男人有三顶帽子和两件背心，则他可以有 $3\times2=6$ 种打扮方法.

加法原理和乘法原理是两个很重要的计数原理，它们不但可以直接解决不少具体问题，同时也是推导下面常用排列组合公式的基础.

（1）**排列**：从 n 个不同元素中任取 r 个元素排成一列，若考虑元素先后出现的次序，则称此为一个排列，此种排列的总数记为 P_n^r 或 A_n^r. 由乘法原理可得

$$\mathrm{P}_n^r=n(n-1)(n-2)\cdots\cdots(n-r+1)=\frac{n!}{(n-r)!}.$$

若 $r=n$，则称为**全排列**，记为 P_n. 显然，$\mathrm{P}_n=n!$.

（2）**重复排列**：从 n 个不同元素中每次取出一个，放回后再取下一个，如此连续 r 次所得的排列称为重复排列，此种排列数共有 n^r，r 可以大于 n.

（3）**组合**：从 n 个不同元素中任取 $r\leqslant n$ 个元素并成一组，若不考虑时间的先后顺序，则

称此为一个组合，此种组合的总数记为 C_n^r 或 $\binom{n}{r}$. 由乘法原理可得

$$C_n^r = \frac{P_n^r}{r!} = \frac{n!}{(n-r)!r!},$$

规定 $0!=1$, $C_n^0=1$.

（4）**重复组合：** 从 n 个不同元素中每次取出一个，放回后再取下一个，如此连续 r 次所得的组合称为重复组合，此种组合数共有 C_{n+r-1}^r，r 可以大于 n.

组合系数 C_n^r 又称为**二项式系数**，因为它是二项式展开的系数，

$$(a+b)^n = \sum_{r=0}^{n}\binom{n}{r}a^r b^{n-r}.$$

利用该公式，可得到许多有用的组合公式：

令 $a=b=1$，可得

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n}=2^n.$$

令 $a=-1$, $b=1$ 可得

$$\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n}=0.$$

由 $(1+x)^{m+n}=(1+x)^m(1+x)^n$，运用二项式展开有

$$\sum_{j=0}^{m+n}\binom{m+n}{j}x^j = \sum_{j_1=0}^{m}\binom{m}{j_1}x^{j_1}\sum_{j_2=0}^{n}\binom{n}{j_2}x^{j_2},$$

比较两边 x^r 的系数，可得

$$\binom{m+n}{r}=\sum_{i=0}^{r}\binom{m}{i}\binom{n}{r-i}.$$

将 n 个不同元素分为 k 组，各组元素数目分别为 $r_1,\cdots,r_k$ 的分法总数为

$$\frac{n!}{r_1!r_2!\cdots r_k!},\quad r_1+r_2+\cdots r_k=n,$$

即

$$C_n^{r_1}\cdot C_{n-r_1}^{r_2}\cdot\cdots\cdot C_{r_k}^{r_k}=\frac{n!}{r_1!r_2!\cdot\cdots\cdot r_k!}.$$

在确定概率的古典方法中经常用到上述四种排列组合，但要注意区别有序与无序、重复与不重复.

1.3.2 确定概率的频率方法

定义 1.3.1 在相同的条件下，重复进行了 n 次试验，若事件 A 发生了 n_A 次，则比值 $\frac{n_A}{n}$ 称为事件 A 在 n 次试验中出现的频率，记为 $f_n(A)=\frac{n_A}{n}$.

显然频率具有以下性质：

（1）非负性：对任意 A，有 $f_n(A)\geqslant 0$.

（2）规范性：$f_n(\Omega)=1$.

（3）可加性：若 A,B 互斥，则 $f_n(A\cup B)=f_n(A)+f_n(B)$

在大量的重复试验中，频率常常稳定于某个常数，称为频率的**稳定性**，即随着 n 的增加，频率越来越可能接近于概率. 我们还容易看到，若随机事件 A 出现的可能性越大，一般来讲，其频率 $f_n(A)$ 也越大. 由于事件 A 发生的可能性大小与其频率大小有如此密切的关系，加之频率又有稳定性，故可通过频率来定义概率，这就是概率的统计定义.

定义 1.3.2 在相同的条件下，独立重复地做 n 次试验，当试验次数 n 很大时，如果某事件 A 发生的频率 $f_n(A)$ 稳定地在 $[0,1]$ 上的某一数值 p 附近摆动，而且一般来说随着试验次数的增多，这种摆动的幅度会越来越小，则称数值 p 为事件 A 发生的概率，记为 $P(A)=p$.

概率的统计定义既肯定了任一事件的概率是存在的，又给出了一种概率的近似计算方法，但不足之处是要进行大量的重复试验，而这在有时是不可能实现的.

值得注意的是，概率的统计定义以实验为基础，但这并不等于说概率取决于试验. 事实上，事件发生的概率乃是事件本身的一种属性，先于实验而存在. 例如，抛硬币，我们首先相信硬币质量均匀，那么在抛之前就已知道出现正面或反面的机会均等，所以从概率的计算途径来看，概率的描述性定义是先验的，概率的统计定义是后验的，显然两种定义并非等价.

对于重复试验发现的规律性有：掷一颗均匀的骰子，六点朝上的概率为 $\dfrac{1}{6}$；从装有外形、质量相同而颜色不同的 a 个白球和 b 个黑球的袋子中，任意摸一球，则能摸到白球的概率为 $\dfrac{a}{a+b}$. 诸如此类事件，只要重复无穷次的试验，该事件发生的概率就是事件频率的稳定值，伯努利大数定律给出了严格的证明. 人们把这种有着明确的历史先例和经验的概率称为**客观概率**.

例 1.3.1 一个职业赌徒想要一对灌过铅的骰子，他雇佣一位技工为他制造一对骰子，要求使得掷得两个一点的概率恰好是六十四分之一而不是三十六分之一. 他被告知每一种可能的计算都已做过，并且所需要的概率毫无疑问是六十四分之一. 于是他给技工付了酬金，然而骰子并没有掷过，你会相信技工的话吗？

解 实际中，当概率不易求出时，人们常通过作大量试验，用事件出现的频率去估计概率，只要试验次数 n 足够大，估计精度完全可以满足人们的需求.

职业赌徒可以掷 800 次，如果一点出现的次数在 100 次左右，则骰子完全满足要求；如果一点出现的次数在 133 次左右，则此骰子基本上是正常的骰子.

至于“左右”到底多大，不同的人可取不同的值. 保守、悲观的人可取大点，乐观的人可取小点，这涉及假设检验问题，我们将在第 6 章详细讲解.

例 1.3.2 抛硬币试验.

历史上有不少人做过抛硬币试验，其结果见表 1.3.1. 从表中的数据可看出：出现正面的频率逐渐稳定在 0.5. 用频率的方法可以说，出现正面的概率为 0.5.