刘佩莉　王丙参　牛晓霞／编

概率论与数理统计

GAILÜLUN
YU SHULI TONGJI

SWJTUP
西南交通大學出版社

概率论与数理统计

刘佩莉　王丙参　牛晓霞　编

西南交通大学出版社
·成　都·

内容简介

本书针对高等院校非数学专业教学大纲与《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，系统、全面地介绍了概率论与数理统计的理论及其应用，读者只需具备高等数学基础即可读懂. 全书分为 8 章，主要讲解事件与概率，随机变量（一维与多维）及其分布，随机变量的数字特征，大数定律与中心极限定理，统计量及其分布，参数估计，假设检验. 为方便读者自学，本书给出了详细的习题解答，供大家参考.

本书可作为经济、管理、理工科各专业的本科生教材，也可作为相关专业的参考用书.

图书在版编目（CIP）数据

概率论与数理统计 / 刘佩莉，王丙参，牛晓霞编.
—成都：西南交通大学出版社，2015.5
ISBN 978-7-5643-3898-5

Ⅰ. ①概… Ⅱ. ①刘… ②王… ③牛… 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（2015）第 107947 号

概率论与数理统计

刘佩莉
王丙参 编
牛晓霞

责任编辑 张宝华
特邀编辑 刘文佳
封面设计 墨创文化

印张 14.75 字数 368千
成品尺寸 185 mm × 260 mm
版本 2015 年 5 月第 1 版
印次 2015 年 5 月第 1 次
印刷 成都中铁二局永经堂印务有限责任公司

出版 发行 西南交通大学出版社
网址 http://www.xnjdcbs.com
地址 四川省成都市金牛区交大路146号
邮政编码 610031
发行部电话 028-87600564 028-87600533

书号：ISBN 978-7-5643-3898-5
定价：30.00元

课件咨询电话：028-87600533

前　言

“概率论与数理统计”作为高等学校本科数学基础课程中一门重要的必修课程，是进一步学习许多相关重要应用数学分支（如随机过程、多元统计、抽样技术等）的必备基础. 因此，作为一门 50 ~ 70 学时的基础课，一方面，希望能较好地体现该课程的基本教学要求，保持教学内容的稳定，反映学科的发展和广泛的应用；另一方面，希望能够满足应用型人才培养及《全国硕士研究生入学统一考试数学考试大纲》的要求. 为了把握好这一平衡，我们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了本书，力图使本书能更好地符合经济、管理、理工科各专业对概率论与数理统计的教学需求.

在编写过程中，我们博采百家之长，注重基本理论、概念、方法的叙述，坚持抽象概念形象化的原则，关注应用能力、解题能力的培养. 读者只需具有高等数学基础即可读懂本书. 在每年的考研试题中，数 1、数 3 的概率统计内容占到了 20%以上，因此，本书力求教材的体系、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，书中很多例题直接采用了历年考研真题. 鉴于例题可以加深读者对理论的理解，我们配备了大量例题和习题，难度各异，以满足不同学生的需求. 本书采用了一些经典的例子和段落，在这里对给出这些材料的作者表示感谢.

全书共分 8 章. 第 1 至 5 章为概率部分，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，并探讨了与概率论有关的决策理论；第 6 至 8 章为统计部分，重点讲解统计量及其分布，参数估计，假设检验. 附录给出了 MATLAB 与概率统计，随机模拟，标准正态分布表，供读者查阅并加深对正文的理解. 鉴于计算机软件的普及，我们不再重点讲解查表，而是利用软件直接给出结果. 为了与软件及思维习惯保持一致，书中采用下侧分位数. 最后，为方便读者自学，本书给出了详细的习题解答.

本教材最初草稿是数学与应用数学专业的《概率论与数理统计》(魏艳华、王丙参编著)，我们内部使用多年，学生反映不错. 定稿时，既保留了原书的优点，更正了里面的错误，又做了较大的修改，删除了很多数学专业内容，部分内容重写，同时增加了很多简单、有趣的例题，并且采用了大量的考研真题作为例题. 我们认为，教材内容要比教学大纲的要求多一些，要比教师在课堂上讲授得多一些，这样才能照顾到各类学校各个专业的需要，以满足不同程度学生的学习需要. 超出数 1 要求的内容，我们用*号标注，仅供有精力、有能力的读者参考. 本书可作为经济、管理、理工科各专业的本科生教材，也可作为相关专业的参考用书.

概率论与数理统计，一般每周安排 3~4 学时，即每学期总学时为 54 ~ 72 学时，包括习题课. 但是现在很多院校为了压缩课时，甚至把概率论与数理统计压缩到 36 课时. 如果课时多且学生基础不错，可讲解概念的严格定义，例如概率的公理化定义. 建议按第 1 章到第 8 章的顺序分配学时如下：

（1）[8+8+4+4+2]+[4+4+2]=36 学时；

（2）[8+8+8+8+4]+[6+8+4]=54 学时；

（3）[10+10+10+10+6]+[8+8+6]=68 学时，剩下 4 学时为机动学时.

以上建议仅供参考，任课老师可根据实际需要合理安排各章学时并选择教学重点. 如果是 36 课时，希望任课老师多思考学生的专业要求.

本书由天水师范学院商学院刘佩莉、牛晓霞与数学与统计学院王丙参共同编写，具体分工为：第 1、3、4 章由王丙参编写，第 2、7、8 章及习题解答由刘佩莉编写，第 5、6 章及附录由牛晓霞编写. 最终由王丙参定稿. 在本书的编写过程中得到了学院领导的大力支持；统计教研室魏艳华认真审阅了书稿，并提出了宝贵的修改意见；得到了西南交通大学出版社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，但由于编者水平有限，书中难免存在错误和不妥之处，恳切希望读者批评、指正，以使本教材不断得以完善.

编　者

2015 年 1 月

目　录

1　随机事件与概率

概率论的近代理论是由著名数学家 Kolmogorov（柯尔莫哥洛夫）奠定的，他在《概率论的基本概念》一书中提出了概率论的公理化体系，为概率论的发展提供了逻辑上的坚实基础. 本章运用大量通俗实例介绍这一公理体系，先由随机试验引出样本空间，进而给出概率的公理化定义及其基本性质，并给出概率的计算方法，最后介绍条件概率和独立性.

1.1　随机事件

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，每天早晨太阳必然从东方升起；在标准大气压（压力约为 100 kPa）下，水加热到 100 摄氏度必然沸腾；一个袋子中有 100 只完全相同的白球，从中任取 1 只必然为白球. 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.

另一类是**随机现象**，也称为**偶然现象**，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；某种型号电视机的寿命. 这些例子表明，在可控制的条件相对稳定的情况下，由于影响这类现象的还有大量的、时隐时现的、瞬息万变的、无法完全控制和预测的偶然因素在起作用，致使现象具有随机性. 注意，既然随机性是由大量无法完全控制的偶然因素引起的，那么随着科学的不断发展、技术手段的不断完善，人们可以将越来越多的因素控制起来，从而减少随机性的影响，不过完全消除随机性是不可能的.

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.

随机现象有大量和个别之分. 在相同条件下可以（至少原则上可以）重复出现的随机现象，称为**大量随机现象**；带有偶然性但原则上不能在相同条件下重复出现的随机现象，称为**个别随机现象**，例如，某场足球赛的输赢是不能重复的，某些经济现象（如失业、经济增长速度等）也不能重复.

例 1.1.1　（1）抛一枚硬币，可能正面朝上，也可能反面朝上；

（2）一天内某高速公路的交通事故次数；

（3）明天某时刻天水的温度；

（4）测量某物理量（长度、直径等）的误差.

试验是对现象的观测（观察或测量），而**实验**是根据科学研究的目的，尽可能地排除外界的影响，突出主要因素并利用一些专门的仪器设备，人为地变革、控制或模拟研究对象，使

某些事物（或过程）发生或再现，从而去认识自然现象、自然性质、自然规律.

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，在相同的条件下可以重复的随机现象称为**随机试验**.

随机试验是对随机现象的一次观测或试验，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，而（2）（3）由于不能重复进行（历史不可重演），它们虽是随机现象，但不是随机试验.

人生思考：历史不可重演，人生也不可重过. 因此过去的事情就过去了，不要后悔，因为后悔不仅不能解决问题，反而会给你增加烦恼. 我们要坦然地承受事情的一切结果，不管是好，不管是坏，一切结果都是自己的选择，都是人生经历的一部分. 也许，痛苦也会变成晚年幸福的回忆，因为你经历过，你感受过！

概率与统计是研究随机现象统计规律的一门学科，概率论主要采用演绎方法进行研究，而统计主要采用归纳方法进行研究. 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，预先做出确定的判断是不可能的，但是假如硬币均匀，直观上看，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”. 概率论的任务是要透过随机现象的随机性揭示其统计规律性；统计的任务则是通过分析带随机性的统计数据来推断所研究的事物或现象固有的规律性.

概率与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，比如时间序列分析可以研究气温的变化过程.

1.1.2　样本空间

随机试验的一切可能基本结果组成的集合称为**样本空间**，用 Ω 表示；其中的每个元素称为**样本点**，又称为**基本结果**，用 ω 表示. 样本点是今后抽样的基本单元，认识随机现象首先要列出它的样本空间.

例 1.1.2　下面给出随机现象的样本空间.

（1）投一枚均匀硬币，观察出现正反面情况，记 z 为正面，f 为反面，样本空间 $\Omega_1=\{z,f\}$；

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

（3）测量误差的样本空间 $\Omega_3=\mathbf{R}$.

（4）观察一个粒子在直线 R 上的运动，在时刻 $t=0$ 时，粒子位于直线上某点 x 处，然后粒子开始向左或向右作随机运动. 我们用 $w(t)$ 表示粒子于时刻 t 时在直线上所处的位置，观察结果就是一条定义在时间轴 $[0,\infty)$ 上取值于 $\mathbf{R}$ 上的连续曲线，而这类曲线的全体就构成样本空间：

$$\Omega_4 = \{w(t): w(t)\text{是}[0,\infty)\text{到}\mathbf{R}\text{的连续函数}\}.$$

如例 1.1.2 中（4）的样本点与时间 t 有关，又如股票价格和期权价格都随时间而变，研究它们都要用到这类样本点，这是随机过程的研究对象.

理解样本空间要注意以下几点：

（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. 如例 1.1.2 中（2）的表示方法就是列举法，用描述法表示为：{呼叫次数为自然数}.

（2）样本点可以是一维的，也可以是多维的，可以有限个，也可以无限个.

（3）对于一个随机试验而言，样本空间并不唯一，它由试验目的而定，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=\{$0 分, 1 分, 2 分, 3 分$\}$.

今后在数学处理上，往往将样本点的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**. 由于这两类样本空间有着本质差异，故分别称呼之.

初学者也许会认为无限多都是一样的，其实它们是有本质区别的. 无限多可分为可列无限多和不可列无限多. 下面给出定义：

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）**.

例如，自然数和有理数都是可列集，而无理数是不可列集，它和实数是一样多的. 由于不可列集比可列集要多得多，因此，实数基本上是由无理数构成的，这也许和读者的直觉矛盾.

1.1.3 随机事件

在样本空间 Ω 中，具有某种性质的样本点构成的子集称为**随机事件**，简称**事件**，常用大写字母 A,B,C 等表示. 用集合论语言，随机事件是样本空间 Ω 的子集. 随机事件包括基本事件和复合事件.

（1）由一个样本点构成的集合称为**基本事件**；

（2）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，记作 $\omega\in A$.

例如，在掷骰子的试验中，基本事件有 6 个：出现 1 点、2 点、3 点、4 点、5 点、6 点，而随机事件共有 2^6 个. 假设事件 A 表示“出现偶数点”，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.

任何样本空间 Ω 都有两个特殊子集，即空集 $\varnothing$ 和 Ω 本身，其中空集 $\varnothing$ 称为**不可能事件**，指每次试验一定不会发生的事件；Ω 称为**必然事件**，指在每次试验中都必然发生的事件. 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进行处理，即**退化的随机事件**.

经常会遇到这样的情况，我们感兴趣的是一个较为复杂的事件，但通过种种方法，

可使之与一些较简单的事件联系起来，这时，我们就设法利用这种联系通过简单事件去研究较为复杂的事件.

1）随机事件间的关系

（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

$$A\subset B\Leftrightarrow\{若\omega\in A,则\omega\in B\}.$$

（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**，也称为**互不相容**.

显然有：基本事件是互斥的；$\varnothing$ 与任意事件互斥.

（4）**事件的对立**：称事件 $B=\{A不发生\}$ 为 A 的**对立事件**或**逆事件**，常记为 $\overline{A}$.

作为样本空间的子集，逆事件 $\overline{A}$ 是 A 相对于样本空间 Ω 的**补集**.

对立事件一定互斥，但互斥事件不一定是对立事件. 对立（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. 请读者举出例子.

注：很多教材对 $\subset$ 与 $\subseteq$ 不加区分，认为两者等价，即不区分子集与真子集.

2）随机事件运算

（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**（**和**），记为 $A\cup B$（或 $A+B$），即

$$A\cup B=\{\omega\mid\omega\in A或\omega\in B\}.$$

（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**（**积**），记为 $A\cap B$（或 AB），即

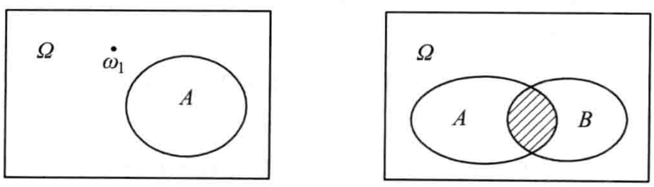
$$A\cap B=\{\omega\mid\omega\in A且\omega\in B\}.$$

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

若 $A\subset B$，则 $A\cap B=A$；

A 与 B 互斥 $\Leftrightarrow AB=\varnothing$.

为直观表示事件及其关系，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**. 在考察事件关系或事件计算中，Venn 图可以起到事半功倍的效果.



$\omega_1\notin A$　　　　$A\cap B$

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和可列无穷多个事件的情形.

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

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

（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 $A-B$，或 $A\setminus B$，即

$$A-B \Leftrightarrow \{\omega \in A \text{ 而 } \omega \notin B\}.$$

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**.

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 Borel 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**：

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

（5）**差积转换律**：$A-B=A\overline{B}$.

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：$\overline{A}\cap\overline{B}\cap\overline{C}=\overline{A\cup B\cup C}$（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　概率的公理化定义

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，可以用一个数字进行度量. 例如，在掷一枚均匀骰子的试验中，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. 所以，对于一个随机试验，我们不仅要知道可能出现哪些事件，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，但事件的概率是如何定义的呢？在概率论的发展史上曾经有过概率的古典定义、几何定义、频率定义和主观定义，这些定义各适合一类随机现象，那么如何才能给出适合一切随机现象的概率的最一般定义呢？

一个随机试验可由一个概率空间 $(\Omega,\mathcal{F},P)$ 所描述，其具体定义由苏联数学家 Kolmogorov 在 1933 年提出，我们称为 Kolmogorov 公理化体系. 它从最少几条本质特性出发刻画概率的概念，既概括了历史上几种概率定义的共同特性，又避免了各自的局限性和含混之处. 这一公理体系迅速获得举世公认. 概率论的发展史表明它是现代概率论的基础，具有里程碑式的意义.

1.2.1 事件域

为了给随机试验提供一个数学模型，我们已经建立了样本空间 Ω，并把 Ω 的一些子集称为事件，介绍了事件的运算. 为了能自由地对有限个或可列个事件进行各种运算，并且运算的结果仍然是事件，我们给出事件域的定义.

设 $\mathcal{F}$ 是样本空间 Ω 的一些子集构成的集类，所谓的集类就是集合的集合，即它的每个元素都是 Ω 的子集. 所谓的“一些”子集，是指不必包含 Ω 的全部子集，只要求此集类对其各种元素封闭，它就是事件域，严格地说有如下定义.

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 域**，或 **σ 代数**、**事件域**.

最简单的 σ 域为 $\{\Omega,\varnothing\}$，称为**平凡 σ 域**.

可见，在事件域 $\mathcal{F}$ 中至多涉及可列个事件的运算，且对差、有限交、有限并、可列交、可列并等运算是封闭的，即在 σ 域 $\mathcal{F}$ 中可以自由地进行有限个或可列个事件的各种运算，这为定义概率和全面研究随机现象奠定了基础. 今后总是给出样本空间 Ω 后，立刻给出 σ 域 $\mathcal{F}$，并把 $\mathcal{F}$ 中的元素称为事件，而不在 $\mathcal{F}$ 中的 Ω 的其他子集皆不是事件，它们不在我们的研究范围之内. 如何确定 σ 域 $\mathcal{F}$，要根据实验类型和需要确定，在一般理论中，样本空间 Ω 和 σ 域 $\mathcal{F}$ 都假定事先给出. 在实际问题中，σ 域 $\mathcal{F}$ 常理解为从随机事件中得到的全部信息.

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.

现在，我们研究直线 $\mathbf{R}$ 上的 σ 域——博雷尔（Borel）域.

定义 1.2.2 设 $\Omega=\mathbf{R}$，$x\in\mathbf{R}$，由所有半无限开区间 $(-\infty,x]$ 生成的最小 σ 域称为 $\mathbf{R}$ 上的 **Borel σ 域**，记为 $\mathcal{B}(\mathbf{R})$，其中的元素称为 **Borel 集合**.

类似地，可定义 $\mathbf{R}^n$ 上的 Borel σ 域 $\mathcal{B}(\mathbf{R}^n)$，那么什么是生成的最小 σ 域呢？

设 C 是样本空间 Ω 的某些子集的全体，称 $\sigma(C)$ 为由 C **生成的最小 σ 域**，如果它满足：

（1）$\varnothing \in \sigma(C)$；

（2）若 $A \in \sigma(C)$，则 $\overline{A} \in \sigma(C)$；

（3）若 $A_i \in \sigma(C), i = 1,2,\cdots$，则 $\bigcup\limits_{i=1}^{\infty} A_i \in \sigma(C)$；

（4）集合 $C \subset \sigma(C)$；

（5）假设 $\mathcal{G}$ 是任意一个包含 C 的 σ 域，则对于集合 $A \in \sigma(C)$，一定有 $A \in \mathcal{G}$.

条件（1）～（3）说明 $\sigma(C)$ 是一个 σ 域. 条件（4）～（5）说明 $\sigma(C)$ 是包含集合 C 的最小 σ 域. 若 A 为事件，则 $\sigma(A) = \{\Omega, A, \overline{A}, \varnothing\}$.

我们来看一下，$\mathcal{B}(\mathrm{R})$ 包含哪些集合. 首先，

$$\mathbf{R} = \bigcup_{n=1}^{\infty} (-\infty, n] \in \mathcal{B}(\mathbf{R}), \quad (a,\infty) = \mathbf{R} \setminus (-\infty, a] \in \mathcal{B}(\mathbf{R}),$$

$$(a,b] = (-\infty, b] \cap (a,\infty) \in \mathcal{B}(\mathbf{R}), (a<b), \quad \{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a \right] \in \mathcal{B}(\mathbf{R}),$$

$$[a,b] = \{a\} \cup (a,b] \in \mathcal{B}(\mathbf{R}), \quad (a,b) = (a,b] \setminus \{b\} \in \mathcal{B}(\mathbf{R}).$$

另外，上述集合的可列并、可列交、有限并、有限交及取逆运算的结果皆在 $\mathcal{B}(\mathbf{R})$ 中，因此，$\mathcal{B}(\mathbf{R})$ 是一个相当大的类，它把实际问题中感兴趣的点全部包含其中了.

1.2.2 概率的定义

首先说明，概率是一个集合函数，因为事件是 Ω 的一个子集，它的概率是一个数. 一个集合对应一个数，称为**集合函数**. 又因事件的全体是 σ 域，所以这个集合函数的定义域为 $\mathcal{F}$，而其值就是区间 $[0,1]$ 中的数. 其实，人们对集合函数的概念并不陌生，如集合元素的个数、区间的长度、区域的面积、物体的体积和质量等都是集合函数，它们在近代数学中统称为测度，所以概率也称为**概率测度**.

上述集合函数的一个重要特征就是可加性. 例如，两个不相交的有限集，其并集的元素个数等于各集元素的个数之和，几个不相交区域并集的长度等于每个区间长度之和等，这种性质称为**可加性**. 自然，概率也应具有可加性.

定义 1.2.3 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 $\mathcal{F}$ 上的一个集合函数 $P(A)$ 与之对应，它满足：

（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，P 称为**概率测度**，简称为**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 $\mathcal{F}$. 事件域可以保证随机事件经过各种运算后仍是随机事件.

事件的概率是事件本身固有的属性，它是一个确定的数，不因在一次具体试验中事件是否发生而改变. 例如，掷硬币时正面出现的概率是 0.5，这是由硬币的形状对称、密度均匀等客观条件决定的. 如果一枚硬币正面是铜，反面是铝，则正面出现的概率就小于 0.5 了.

1.2.3 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.

在概率的正则性中说明了必然事件 Ω 的概率为 1，由此可知，不可能事件 $\varnothing$ 的概率应该为 0. 切记，在数学理论体系中，只有公理、定义、假设不需要证明，其他都要证明.

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

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

再由非负性公理必有 $P(\varnothing)=0$.

在概率论中，将概率很小（小于 0.05）的事件称为**小概率事件**，也称为**实际不可能事件**.

注意：很小是一个模糊概念，没有严格的区分，因人而定，这不属于数学范围之内，在许多情况下，要随试验结果的重要性，具体问题具体分析地加以确定.

小概率事件原理，又称为**实际推断原理**：在原假设成立的条件下，小概率事件在一次试验中可以看成不可能事件. 如果在一次试验中小概率事件发生了，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 迟早必然会出现一次，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. 但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，这是小概率事件原理的一个应用. 如果我们把注意仅停留在小概率事件的极端个别现象上，那我们就是“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

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

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

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

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 $P(A-B)=P(A)-P(AB)$，也称为**减法公式**.

（3）若 $A\supset B$，$P(A-B)=P(A)-P(B)$ 且 $P(A)\geqslant P(B)$，也称为概率的**单调性**.

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

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

移项即得结论.

（3）由（2）显然可知结论成立.

很容易举例说明，若 $P(A)\geqslant P(B)$，无法推出 $A\supset B$. 此推论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非常有用.

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，又由有限可加性得

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

加法公式还能推广到多个事件的情况. 设 A_1,A_2,A_3 为任意三个事件，则有

$$P(A_1\cup A_2\cup A_3)=\sum_{i=1}^{3}P(A_i)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3).$$

利用样本点在等式两端计算次数相等可直观证明这个公式.

一般地，对于任意 n 个事件 $A_1,\cdots,A_n$，可以用数学归纳法证得

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

此式称为**容斥原理**，也称为**多去少补原理**.

1.3　概率的直接计算

事件的概率通常是未知的，但公理化定义并没有告诉人们如何去计算概率. 历史上在公理化定义出现之前，概率的统计定义、古典定义、几何定义和主观定义都在一定的场合下具有计算概率的方法，所以有了公理化定义后，它们均可以作为概率的计算方法.

1.3.1　计算概率的统计方法

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，则称比值 $\dfrac{n_A}{n}$