概率论

与数理统计教程

GAILULUN YU SHULI TONGJI JIAOCHENG

张继昌 编著

浙江大学出版社

概率论与数理统计教程

张继昌　编著

浙江大學出版社

图书在版编目（CIP）数据

概率论与数理统计教程／张继昌编著．—杭州：浙江大学出版社，2003.7

ISBN 7-308-03369-4

Ⅰ.概... Ⅱ.张... Ⅲ.①概率—高等学校—教材②数理统计—高等学校—教材 Ⅳ.O21

中国版本图书馆 CIP 数据核字（2003）第 052992 号

责任编辑 张颖琪
封面设计 俞亚彤
出版发行 浙江大学出版社
（杭州浙大路 38 号 邮政编码 310027）
（网址：http://www.zjupress.com）
（E-mail：zupress@mail.hz.zj.cn）
排　　版 浙江大学出版社电脑排版中心
印　　刷 浙江大学印刷厂
开　　本 787mm×960mm 1/16
印　　张 16.5
字　　数 305 千字
版、印次 2003 年 7 月第 1 版　2003 年 7 月第 1 次印刷
印　　数 0001—4000
书　　号 ISBN 7-308-03369-4/O・294
定　　价 24.00 元

前　　言

“概率论与数理统计”是一门十分重要的、具有广泛应用的基础课. 与其他数学课程相比，这门课程的思想方法具有明显的特殊性，由此导致的结果是该课程不仅难学，而且还不易教好. 30多年来，本人从事于成人教育、自学考试、远程网络教育、专业培训、各类普通高校，直至各种研究生考前复习等不同层次的教学，有成功的经验，也有不尽如人意的体验，还有面对重应用而加强统计和为考试而注重概率这对矛盾时的无奈. 如何让学生轻松地学，让教师自如地教，是我一直思考并孜孜不倦探索的问题. 也许，解决问题的最佳办法是加强基础、培养学生概率论与数理统计的思想方法.

目前，已有很多“概率论与数理统计”的教材. 当然，在众多教材中首推盛骤教授等编著、高等教育出版社出版的《概率论与数理统计》这一力作. 综观这些教材，各有风格和特色，但在内容安排上，要么太深奥，要么太浅显，通用性强的教材则较少见. 针对网络教学，本人曾编写《概率统计四十五讲》讲义，在浙江大学远程教育学院的计算机、工商管理、电子商务和英语等本科专业使用了两届，取得了较好的效果. 为满足一般层次教学的需要，本人根据长期教学和科研的积累，并在充分调查了各类普通高等学校的教学要求的基础上，编写了这本通用性颇强的教材.

本书的编写体现了以下特色：

第一，强调基本概念、基本思想和基本方法，便于教，利于学. 在叙述上力求简明扼要，并用大量的例题说明该课程各部分的内容，并配有一定的应用题. 每章起首提出“教学内容”、“基本要求”、“关键词和主题”；每章结尾给出精练的“小结”，概括“内容提要”、“重点”、“难点”和“深层次问题”；大多数小节后

还加了“说明”，指出学习该节时应特别注意的问题及与该节有关的较深层次的内容. 这样做的优点是既便于在教的过程中有的放矢，又有利于在学的过程中提纲挈领.

第二，浅入深出，通用性强，梯度大. 能较广泛地满足多种专业、不同层次的教学要求，让不同需求的学生都能“吃好、吃饱”. 本书涵盖了本科生“概率论与数理统计”课程教学大纲和全国研究生统考考试大纲的内容，除基本知识外，在内容的叙述、例题和习题的配置上，具有相当的梯度. 如全书有 160 多个例题，420 多个习题. 各节的习题分为 A，B 两类，A 类习题是基本要求；B 类习题是较高要求，大多达到研究生入学统考的要求，供教学要求较高的学校和学有余力或有志于考研的同学选做.

第三，前后概念联系紧密，便于复习巩固，加深理解. 每个概念都以直观的例子引入，并不失时机地与前面的知识点紧密结合，以期对各部分内容反复训练，加深理解，并掌握求解综合题的能力.

第四，便于按不同要求、不同学时组织教学. 每节都是相对完整的一个课时的内容，全书又一气呵成，对教师合理组织、安排教学内容和教学进度带来极大的方便. 考虑到不同层次教学的实际需要和学时的限制，本书包括基本概念、随机变量、多维随机变量、数字特征、极限定理、统计基础、参数估计、假设检验、方差分析和回归分析初步 9 章，共 50 节. 不同院校、不同专业可根据教学大纲要求和学时要求，灵活选用章节，合理组织教学.

本人还为此书制作了多媒体网络课件，可供使用该教材的学校选用.

本教材适用于普通高等院校理、工、医、农、经、管各专业本科，也适用于二级学院、成教学院、远程教育学院相关专业本科，同时也可作为考研学子的复习用书.

本书的出版得到了浙江大学宁波理工学院、浙江工业大学之江学院、浙江大学成人教育学院、浙江大学远程教育学院、浙江大学职业技术教育学院等的大力支持，在此深表感谢.

限于水平，书中不当之处，恳请专家和读者指正.

张继昌

2003 年 4 月于浙江大学

目　　录

课程简介

客观世界中发生的现象可以分为两类：一类是在一定条件下结果是确定的，例如向上抛出的一枚硬币一定下落，水加热到沸点一定沸腾；而另一类是在一定条件下结果是不确定的，例如丢一枚硬币，可能正面朝上，也可能反面朝上. 我们把第一类现象称为确定性现象，第二类现象称为随机性现象.

随机性现象在个别试验中其结果呈现偶然性，而在大量重复试验中其结果具有统计规律性. “概率论”是研究随机性现象统计规律性的学科，是随机性数学的基础课. 由于随机问题的特殊性，在概率论中分析问题、解决问题的思想和方法有别于其他数学课，关键是理解各种概率思想.

“数理统计”讨论概率论的思想和方法在实际问题中的应用. 数理统计的内容非常丰富，随着科学技术的发展，还在不断地充实和提高. 本书仅介绍最基本也是最重要的几个方面.

概率统计方法在自然科学、社会科学等几乎所有的领域都有广泛的应用. 通过该课程的学习，掌握分析、解决随机问题的基本思想和基本方法，掌握几种具体的数理统计方法，也为学习后继课程打好基础.

学习概率论与数理统计需要具备一定的数学基础知识，包括集合论、排列组合、函数的导数、定积分、变上限积分的导数、偏导数和二重积分等.

第 1 章　概率论的基本概念

【教学内容】

本章是概率论的基础，首先应明确概率论所研究的对象是随机试验，随机试验的结果用样本空间和随机事件描述.

研究随机事件的目的是求随机事件的概率，本章给出概率的定义和性质，进一步介绍条件概率、乘法公式、全概率公式、贝叶斯公式和独立性.

本章还讨论一类具体的概率问题：等可能概率问题(古典概型). 讨论这类问题的目的，一是掌握这类问题的计算，二是通过这类问题的讨论熟悉概率的定义和性质.

本章共分为 8 节，在 8 个课时内完成.

【基本要求】

1. 知道随机现象、随机试验、随机事件的概念.

2. 理解样本空间的定义，会写出一般随机试验的样本空间，熟练掌握随机事件的关系及运算，会将一些较复杂的事件用简单事件的运算来表示.

3. 了解频率的概念，理解随机事件概率的定义和性质，熟练掌握利用概率的性质计算有关的概率.

4. 理解"等可能概型"的定义，会求解一些简单的等可能概型问题.

5. 理解条件概率的概念，理解实际问题中哪些是条件概率，条件是什么；掌握概率的乘法公式、全概率公式和贝叶斯公式，会用这些公式计算有关的概率.

6. 理解"事件独立性"的概念，会利用独立性计算概率.

【关键词和主题】

集合，全集和子集，排列，组合，加法原理，乘法原理；

随机试验，随机事件，样本空间，基本事件，随机事件的运算，和事件，交事件，差事件，逆事件；

频率，概率的统计定义，概率的公理化定义，概率的性质；

等可能概率问题(古典概率问题)；

条件概率，乘法公式，完备事件组，全概率公式，贝叶斯公式，随机事件的独立性.

§1.1 随机试验及随机事件

一、随机试验

我们所指的"试验"是一个广义的概念,不仅指各种科学试验,还包括对某个过程的特征的记录,对某一问题的调查等.下面就是几个典型的试验的例子.

例1 (1) 一枚硬币连丢2次,考察出现"正"、"反"的情形;

(2) 丢一颗骰子,观察出现的点数;

(3) 某程控交换机在1分钟时间内接到用户呼叫的次数;

(4) 测量某一零件(如粉笔,标准长度为75毫米)的长度.

以上"试验"具有三个特点:

1. 每次"试验"的可能结果多于一个,且事先能明确试验的所有可能结果;
2. 在"试验"之前,不能确定哪一个结果会发生;
3. 可以在相同条件下重复进行"试验".

这样的"试验"称为随机试验,它是概率论所研究的对象.

二、样本空间

随机试验的研究首先是对它作完整的描述,我们把某随机试验的可能结果的全体构成的集合称为样本空间,常用S(或Ω)表示.

如上例中 (1) $S=\{$(正,正),(正,反),(反,正),(反,反)$\}$;

(2) $S=\{1,2,3,4,5,6\}$;

(3) $S=\{0,1,2,\cdots\}$;

(4) $S=\{L:75-\varepsilon<L<75+\varepsilon\}$($\varepsilon$为长度$L$的允许误差).

样本空间的元素,即随机试验的每一个可能结果,称为样本点.

三、随机事件

例2 甲、乙两人丢一颗均匀的骰子,比点数大小,若甲为4点,则乙面临着三种可能:

"输"$=\{1,2,3\}=A$;

"平"$=\{4\}=B$;

"赢"$=\{5,6\}=C$.

这里,"输"、"平"、"赢"是可能发生也可能不发生的事件,我们称这样的"事件"为随机事件,简称为事件,常用大写字母$A,B,\cdots$表示.

把样本空间看做全集S,则随机事件A是S的子集,我们可用图1.1表示.

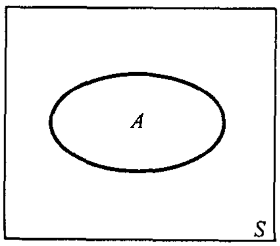

图1.1

例 3　袋中装有编号为 1,2 的两个白球，和编号为 3 的一个黑球，随机地、不放回地取两次，每次取一个球，考察这两个球的编号，试写出样本空间，并写出下列随机事件.

$S=\{(1,2),(2,1),(1,3),(3,1),(2,3),(3,2)\}$；

A："第一次取得黑球"，$A=\{(3,1),(3,2)\}$；

B："第一次取得白球"，$B=\{(1,3),(2,3),(1,2),(2,1)\}$；

C："两次都取得白球"，$C=\{(1,2),(2,1)\}$；

D："第一次取得白球，第二次取得黑球"，$D=\{(1,3),(2,3)\}$；

E："没有取得黑球"，$E=\{(1,2),(2,1)\}$.

在某一随机试验中，一定发生的事件称为必然事件；一定不发生的事件称为不可能事件，用空集$\varnothing$表示. 由一个样本点构成的事件称为基本事件.

说　明

1. 样本空间是指在一定条件下，某随机试验可能结果的全体，不同的随机试验其样本空间是不同的，如习题 1.1(A)第 1 题中的(3)和(4)；表面上看似相同的随机试验，研究的角度不同，其样本空间也是不同的，如习题 1.1(A)第 1 题中的(1)和(2)、(5)和(6).

2. 样本空间是由样本点组成的，样本点是随机试验的那些"不可再分"的最小结果. 如在例 1(1)中，试验是一枚硬币连丢 2 次，(正，正)是试验的一个可能结果，是一个样本点.

3. 随机事件是由样本空间中的某些样本点组成的. 所谓随机事件 A 发生，是指在随机试验时，A 中的某一个样本点出现(严格地说，是由这个样本点组成的基本事件发生)；反之，若 A 中的某一个样本点出现，就称随机事件 A 发生.

4. 必然事件和不可能事件不是随机事件，因为前者一定发生，后者一定不发生.

习题 1.1(A)

1. 写出下列随机试验的样本空间：

(1) 将一枚硬币连丢 3 次，观察正面 H，反面 T 出现的情形；

(2) 将一枚硬币连丢 3 次，观察出现正面的次数；

(3) 袋中装有编号为 1,2 和 3 的三个球，随机地取两个，考察这两个球的编号；

(4) 袋中装有编号为 1,2 和 3 的三个球，依次随机地取两次，每次取一个，不放回，考察这两个球的编号；

(5) 丢甲、乙两颗骰子，观察出现的点数之和；

(6) 丢甲、乙两颗骰子,观察它们出现的点数.

2. 写出下列随机试验中所指出的随机事件:

(1) 丢一颗骰子.

A:出现奇数点;B:点数大于 2.

(2) 一枚硬币连丢 2 次.

A:第一次出现正面;B:两次出现同一面;C:至少有一次出现正面.

(3) 从 1,2,3,4 四个数中随机地取一个,放回,再随机地取一个.

A:其中一个数是另一个数的两倍;B:两数的奇偶性相同.

(4) 10 个零件,其中有 2 个次品,随机地取 5 个.

A:正品个数多于次品;B:正品个数不多于次品.

习题 1.1(B)

1. 记录某程控交换机在 1 分钟内接到的用户的呼叫次数,写出样本空间和下列随机事件.

A:呼叫次数在 10 次以内;B:呼叫次数为奇数.

2. 对一批产品进行检查,合格的记上"1",不合格的记上"0",如查出 2 个次品或已检查 4 个产品,就停止检查,写出样本空间和下列随机事件.

A:检查不超过 3 个;B:只有一个次品.

3. 在单位圆内取一个点,记录它的坐标(x,y),写出样本空间和下列随机事件.

A:x 和 y 都是正的;B:x 是方程 $x^2-x-6=0$ 的根.

§1.2 随机事件的运算

在上一节例 2 中,A="输"$=\{1,2,3\}$,B="平"$=\{4\}$,C="赢"$=\{5,6\}$.

若设 D="不输",则 D 是 B 与 C 之"和",也是 A 的"反". 可见,随机事件之间也必须建立运算.

样本空间 S 看做全集,随机事件是一个集合,是样本空间 S 的子集,因而,随机事件间的关系和运算可按照集合的关系和运算来处理,当然,这种关系和运算需用概率论的语言描述.

一、随机事件之间的关系

若事件 B 发生必导致事件 A 发生,则称事件 A 包含事件 B,记为:$A\supset B$,如图 1.2.

若 $A\supset B$ 且 $A\subset B$,则 $A=B$,称事件 A 与事件 B 相等.

二、随机事件的运算

随机事件的运算与集合运算相同,主要有:和、交、差、逆.

1. 和:事件 $A\cup B=\{x:x\in A$ 或 $x\in B\}$称为事件 A 与 B 的和事件,当且仅

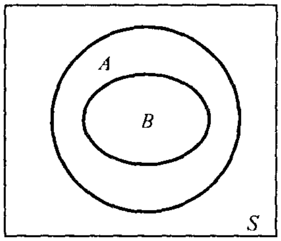

图 1.2

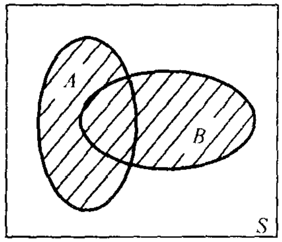

图 1.3

当 A 发生或 B 发生(即 A 与 B 至少一个发生)时,事件 $A \cup B$ 发生.

事件 $A \cup B$ 可用图 1.3 表示,它由三个部分组成,从左到右依次为:A 发生 B 不发生,A 与 B 都发生,B 发生 A 不发生.

2. 交:事件 $A \cap B=\{x: x \in A$ 且 $x \in B\}$称为事件 A 与 B 的交事件(简写为 AB),当且仅当事件 A 与 B 都发生时,事件 AB 发生,如图 1.4.

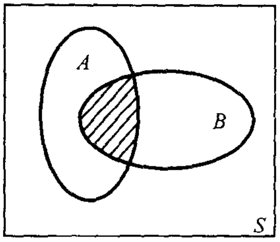

图 1.4

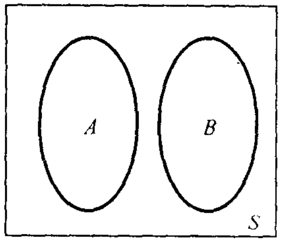

图 1.5

交事件的特殊情形是:$AB=\varnothing$,称事件 A 与事件 B 互不相容(互斥),如图 1.5.

3. 差:事件 $A-B=\{x: x \in A$ 且 $x \notin B\}$称为事件 A 与 B 的差事件.当且仅当事件 A 发生而事件 B 不发生时,事件 $A-B$ 发生,如图 1.6.

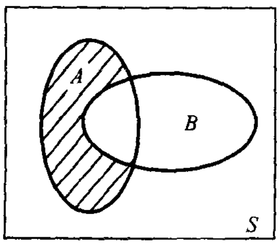

图 1.6

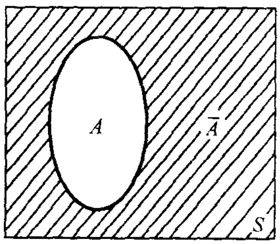

图 1.7

4. 逆事件:事件 $\overline{A}=\{x: x \notin A\}=S-A$ 称为 A 的逆事件,当且仅当事件 A 不发生时,事件 $\overline{A}$ 发生,如图 1.7.

有 $A \cup \overline{A}=S$, $A\overline{A}=\varnothing$,称 A 与 $\overline{A}$ 是互为对立事件.

注意:对立事件是互不相容的,而互不相容事件不一定是对立事件.

可以证明：$A-B=A-AB=A\overline{B}$，

$B-A=B-AB=B\overline{A}$，

于是 $A\cup B=(A-B)\cup(AB)\cup(B-A)$

$=(A\overline{B})\cup(AB)\cup(B\overline{A})$.

三、随机事件运算的运算律

随机事件的运算律与集合运算的运算律完全相同，有以下几条.

交换律：$A\cup B=B\cup A$，$AB=BA$，

但 $A-B\neq B-A$.

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

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

对偶律：$\overline{A\cup B}=\overline{A}\ \overline{B}$，$\overline{AB}=\overline{A}\cup\overline{B}$.

对偶律在事件的表示、概率的计算中很重要，如例 1. 对偶律可推广到多个事件，如$\overline{A\cup B\cup C}=\overline{A}\ \overline{B}\ \overline{C}$，$\overline{ABC}=\overline{A}\cup\overline{B}\cup\overline{C}$.

例 1 设 A 表示"甲乙都成功"，则 $\overline{A}$ 表示什么？

若 B 表示甲成功，C 表示乙成功，则 $A=BC$，于是，$\overline{A}=\overline{BC}=\overline{B}\cup\overline{C}$，即 $\overline{A}$ 表示"甲不成功或乙不成功".

例 2 A,B,C 至少一个不发生表示为：$\overline{A}\cup\overline{B}\cup\overline{C}$；$A,B,C$ 都发生表示为：ABC；而 $AB\cup AC\cup BC$ 表示至少两个发生.

事实上，$AB\cup AC\cup BC=ABC\cup AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$.

例 3 丢一颗骰子，A 表示奇数点，B 表示点数大于 2，则：

$A\cup B=\{1,3,4,5,6\}$，$AB=\{3,5\}$，$A-B=\{1\}$，

$B-A=\{4,6\}$，$\overline{A}=\{2,4,6\}$，$\overline{B}=\{1,2\}$.

说 明

随机事件之间的关系和运算属于集合代数的范畴，集合代数也称逻辑代数、命题代数等，它与普通的代数(称为范氏代数)不同，不能把随机事件之间的关系和运算与数值运算规则等同起来. 虽然两者有相同之处，如交换律、结合律等，但两者有许多本质上的差异. 另外，对随机事件还要特别注意：

(1) $A\cup A=A$，$AA=A$，$A-A=\varnothing$；

(2) 若 $A\supset B$，则 $AB=B$，$A\cup B=A$，$B-A=\varnothing$，$\overline{A}\subset\overline{B}$；

(3) $\overline{A}\ \overline{B}$ 与$\overline{AB}$不能等同，$\overline{A\cup B}$与 $\overline{A}\cup\overline{B}$ 不能等同；

(4) 一般，把 $A-B$ 改写为 $A\overline{B}$ 比较方便；

(5) $A=AB\cup A\overline{B}$，且 AB 和 $A\overline{B}$ 互不相容；

(6) 当 A 发生时 B 一定发生，这表明 $A\subset B$.

习题 1.2(A)

1. 设 A,B,C 为三个事件，用 A,B,C 的运算关系表示下列各事件：

(1) A,B,C 都不发生；　(2) A 与 B 都发生，而 C 不发生；

(3) A 与 B 都不发生，而 C 发生；　(4) A,B,C 中最多两个发生；

(5) A,B,C 中至少两个发生；　(6) A,B,C 中不多于一个发生.

2. 设 $S=\{x:0\leqslant x\leqslant 5\}$，$A=\{x:1<x\leqslant 3\}$，$B=\{x:2\leqslant x<4\}$，具体写出下列事件：

(1) $A\cup B$；(2) AB；(3) $\overline{A}B$；(4) $\overline{A}\cup B$；(5) $\overline{A\,B}$.

3. 已知当 A 发生或 B 发生时，C 一定发生，则不正确的是：

(1) $C\supset A$；(2) $C\supset AB$；(3) $C\supset A\cup B$；(4) $C\subset A\cup B$.

习题 1.2(B)

1. 指出下列命题哪些正确，哪些不正确？

(1) $A\cup B=A\overline{B}\cup B$；　(2) $A=A\overline{B}\cup AB$；

(3) $\overline{A}B=A\cup B$；　(4) $\overline{(A\cup B)}C=\overline{A}\,\overline{B}\,\overline{C}$；

(5) $(AB)(A\overline{B})=\varnothing$；　(6) 若 $A\subset B$，则 $A=AB$；

(7) 若 $A\subset B$，则 $A\cup B=A$；　(8) 若 $A\subset B$，则 $\overline{B}\subset\overline{A}$；

(9) 若 $AB=\varnothing$，则 $\overline{A}\,\overline{B}\neq\varnothing$；　(10) 若 $AB=\varnothing$，则 $\overline{A}\,\overline{B}=\varnothing$.

2. 在某校任选一名学生，令 A 表示男生，B 表示一年级，C 表示计算机专业.

(1) 叙述事件 $AB\overline{C}$ 的含义；　(2) 在什么条件下 $ABC=C$ 成立？

(3) 在什么条件下 $C\subset B$ 是正确的？　(4) 在什么条件下 $\overline{A}=B$ 成立？

§1.3　概率的定义和性质

可能发生、可能不发生的“事件”是随机事件，因此，我们须进一步讨论随机事件在随机试验中发生的可能性大小，即概率.

一、频率

事件 A 在一次试验中，可能发生可能不发生，在相同条件下，进行了 n 次重复试验，在这 n 次试验中，事件 A 发生的次数 k，称为事件 A 发生的频数，称 k/n 为事件 A 发生的频率，记为 $f_n(A)$.

频率反映事件 A 在试验中发生的频繁程度，频率越大，表明事件 A 在一次试验中发生的可能性越大.

例如，多次掷一枚均匀的硬币，考察正面 A 出现的次数，这样的试验在历史上有多人做过，结果有：

$n=2048$，　$k=1061$；　$n=4040$，　$k=2047$；

$$n=12000,\quad k=6019;\qquad n=24000,\quad k=12012.$$

于是,它们的频率分别是:

$$\frac{1061}{2048}=0.5181;\qquad \frac{2047}{4040}=0.5067;$$

$$\frac{6019}{12000}=0.5016;\qquad \frac{12012}{24000}=0.5005.$$

当 $n\to\infty$ 时,频率 $f_n(A)\to 0.5$,这对于现代的人来说,掷一枚均匀的硬币,正面 A 出现的概率是 0.5,这已经是一般的常识.

人们曾用频率的极限来定义概率 P:

$$P=\lim_{n\to\infty}f_n(A).$$

这个定义称为概率的统计定义.它表明了频率的稳定性,将在§5.1中由大数定理在理论上加以证明.

为了实际应用和理论研究的需要,我们从频率的概念得到启发,给出度量事件发生可能性大小即概率的定义.

二、概率的定义

概率的公理化定义:

随机试验的样本空间为 S,对随机事件 A 赋予一个实数 $P(A)$,若满足下列三个公理,则称 $P(A)$ 为 A 的概率.

(1) $P(A)\geqslant 0$;

(2) $P(S)=1$;

(3) A,B 互不相容,则 $P(A\cup B)=P(A)+P(B)$.

所谓公理,是人们在长期的实践中总结出来的完全正确的结论.

对于(1),概率,即可能性大小,总是大于等于零,不可能是负值.

对于(2),S 是样本空间,是随机试验的全体可能结果,其概率肯定是 1.

对于(3),我们注意到概率是一种度量,如面积,是平面区域大小的度量;如长度,是两点之间距离大小的度量;又如电压,是电的一个性质的度量;而概率是可能性大小的度量.不同的度量有各自的特性,但它们也有许多共性,如两块没有重叠的区域,其总面积是两块区域面积之和,这与概率的公理(3)一样.有些度量是比较直观的,如面积和距离;而有些是不直观的,如概率和电压.我们可以用较直观的度量来帮助理解概率,如把概率看做"面积",当然,应注意"总面积"是1.事实上,在第2章的连续型随机变量中就是用面积来度量概率的.

以后,有关概率的所有性质、定理、结论等,都利用公理去证明.

三、概率的性质

由概率的定义,我们可以得到概率的下列性质:

(1) $P(\varnothing)=0$.