高等学校教材

概率论与数理统计

闫在在　主编

高等教育出版社

高等学校教材

概率论与数理统计

Gailülun yu Shuli Tongji

闫在在　主编

高等教育出版社·北京

内容提要

本书遵循教育部高等学校大学数学课程教学指导委员会最新修订的“工科类本科数学基础课程教学基本要求”，并参考教育部考试中心制定的“全国硕士研究生入学统一考试数学考试大纲”编写而成。本书知识体系相对完整，结构严谨，内容丰富，循序渐进，通俗易懂，例题丰富。

全书共分为十章：前五章为概率论部分；后五章为数理统计部分，其中第十章介绍统计软件 R。前九章每章后均附有习题、自测题，第十章附有习题。总复习题大部分选自历年全国硕士研究生入学统一考试数学考试试题，书末提供了详细参考答案或提示。

本书可作为高等学校理工类各专业本科生的概率论与数理统计教材，也可作为学生参加全国硕士研究生入学统一考试的复习参考用书，同时也可供工程技术人员、科技工作者参考。

图书在版编目（CIP）数据

概率论与数理统计 / 闫在在主编 . -- 北京：高等教育出版社，2015.8
ISBN 978-7-04-043454-5

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

中国版本图书馆 CIP 数据核字（2015）第 170051 号

策划编辑 杨 帆　　责任编辑 杨 帆　　封面设计 李树龙　　版式设计 杜微言
插图绘制 杜晓丹　　责任校对 刘春萍　　责任印制 韩 刚

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街 4 号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	涿州市星河印刷有限公司	网上订购	http://www.landraco.com
开　　本	787mm × 960mm　1/16		http://www.landraco.com.cn
印　　张	22.75	版　　次	2015 年 8 月第 1 版
字　　数	410 千字	印　　次	2015 年 8 月第 1 次印刷
购书热线	010-58581118	定　　价	39.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　43454-00

前　言

21 世纪以来, 随着我国经济建设与科学技术的迅速发展, 高等教育已由"精英式"教育模式转变为"大众化"教育模式, 教学观念不断更新, 教学改革不断深入, 办学规模不断扩大。作为高等学校理工科各专业三大基础数学课程之一的概率论与数理统计, 其教学模式、教学方法也发生了很大的变化, 由原来单纯重视概率论与数理统计方法和定理推演的教学模式改变为传统推理教学和统计软件使用相结合的教学模式。特别是对于非统计学专业学生来说, 学习概率论与数理统计的主要目的是应用, 而不使用统计软件很难将概率论与数理统计方法应用到实际问题中。对他们而言, 做过多的公式推导显然意义不大, 学会概率论与数理统计计算不等于能使用概率论与数理统计方法。将统计软件应用到概率论与数理统计教学之中使学生不仅可以从复杂的计算中解脱出来, 也可以在使用计算机解答问题时更好地理解统计思想。现如今, 国内外很多院校在教学和科研活动中使用 R 软件, R 已是主流的统计软件。本教材第十章增加了 R 软件在数理统计中的应用, 以期读者在学习概率论与数理统计课程的同时, 初步掌握统计软件 R 的使用。

本教材沿袭传统理论体系, 注重基本概念和概率思想, 强调实际应用, 力求做到难易适当, 易教好学, 其主要特点如下:

一、理论与实际应用有机结合。大量的实际应用贯穿于理论之中, 体现了概率论与数理统计在各个领域中的广泛应用。

二、紧密结合统计软件 R。介绍了专业统计软件 R 在解决概率论与数理统计问题中的应用, 加强了学生分析问题和解决问题能力的培养。

三、习题安排合理。每一章后面给出简单易算和综合性的习题, 使学生的学习由浅入深, 循序渐进, 各章后面均附有自测题, 便于学生自测。书末精心编制了总复习题, 为学有余力的学生及准备考研的学生复习使用, 并附有答案或解题思路。

四、人物传记叙述了一些统计学家的创造性贡献和他们的故事, 以期励志, 增强读者的学习兴趣。

本教材由闫在在教授担任主编。卢静莉编写了第一、第六章; 彭秀云编写了第二、第三章; 郑丽霞编写了第四、第五章; 洪志敏编写了第七、第八章; 闫在在编写了第九、第十章; 王晓民编写了附录的总复习题、答案和附表。

由于编者水平有限，书中缺点和错误在所难免，恳请读者批评指正。

编　者

2015 年 3 月于内蒙古工业大学

目　录

第一章　事件及其概率

§1.1　随机试验与随机事件

1.1.1　随机现象

在自然界, 在人们的实践活动中所遇到的现象大体上可分为两类:

一类现象是在一定条件下必然会发生. 例如抛掷一枚硬币, 硬币必然要落地; 水在标准大气压下达到 100°C 必然沸腾, 而低于 0°C 必然结冰. 我们把这一类现象称为**确定性现象** (或必然现象).

另一类现象则是在一定条件下具有多种可能结果, 但事先又不能确定究竟会发生哪一种结果. 例如抛掷一枚硬币落地后, 是正面(有币值的一面)向上, 还是反面向上, 在硬币落地之前是不确定的; 一只灯泡在使用之前是无法确定其寿命的; 今年某地区十月份的平均气温是多少, 在十月份结束之前是不能确定的. 我们把这一类现象称为**随机现象**.

随机现象在一次观察或试验中, 具有随机性, 即偶然性. 但是人们在长期实践中发现, 相同条件下, 对随机现象进行大量的观察或试验时, 随机现象的结果会呈现出某种确定的规律性. 例如:

(a) 在相同条件下, 抛掷一枚均匀的硬币, 当抛掷次数很大时, 就会发现正面和反面出现的次数几乎各占一半.

(b) 从物理学的观点来看, 气体分子对器壁的压力是气体分子对器壁碰撞的结果. 由于分子是时刻不停地、杂乱无章地运动着的, 速度和轨道都是随机的, 因而对器壁的碰撞也是随机的. 初看起来器壁所受压力是不稳定的, 可是实验证明, 由于分子数目非常大, 各分子运动所具有的随机性在集体中互相抵消、互相平衡了, 使得器壁所受的总压力呈现一种稳定性. 分子数目越大, 压力就越稳定.

从上述例子可以看出, 看上去其结果具有不确定性的随机现象内部却蕴含着某种确定的规律性. 这种通过大量观察或试验总结出的规律性, 称为**统计规律性**.

概率论与数理统计就是研究随机现象的统计规律性的一门数学学科.

在自然界和人们的生产、生活实践中, 存在着大量的随机现象. 这种随机现象是受其内部蕴含着的规律所支配的. 研究和发现这些规律, 揭示偶然性与必然性之间的内在联系, 是我们学习概率论与数理统计的方向和最终目的. 这也就决

定了概率论与数理统计的思想和方法在自然科学、社会科学、工农业生产及国民经济的各个部门中有着广泛的应用, 并且与其他学科相结合、渗透, 推动和发展了许多边缘学科. 因此, 初步掌握处理随机现象的基本理论和方法, 越来越成为一种基础知识和基本技能了.

1.1.2 随机试验

为了研究随机现象的统计规律性, 我们把各种科学试验和对某一现象的观测统称为试验. 而把具有下述三个特征的试验称为**随机试验**:

(1) 试验可以在相同条件下重复进行;

(2) 每次试验的所有可能结果都是明确可知的, 并且不止一个;

(3) 每次试验之前不能预知将会出现哪一个结果, 但若进行大量重复试验, 其可能结果的出现具有一定的统计规律性.

随机试验简称试验. 今后我们说试验就是指随机试验. 下面给出随机试验的几个例子:

E_1={将一枚硬币抛掷 3 次, 观察其正反面出现的情况};

E_2={将一枚硬币抛掷 3 次, 观察其正面出现的次数};

E_3={掷一颗骰子, 观察其点数};

E_4={同时掷两颗骰子, 观察其点数出现的情况};

E_5={炮击目标直到击中为止, 记录发出的炮弹数};

E_6={一只灯泡, 测试其使用寿命}.

1.1.3 样本空间

我们把随机试验的最基本的、不能再分解的结果叫做**基本事件**, 其特点是: 每次试验必出现一个而且只能出现一个基本事件, 任何两个基本事件在一次试验中不能同时出现. 把一切基本事件的集合 Ω 叫做**样本空间**. 样本空间 Ω 的元素, 即基本事件, 又称为**样本点**, 常用 ω 表示.

对应于上段给出的随机试验, 样本空间分别为:

$\Omega_1 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

(其中 H 表示正面, T 表示反面);

$\Omega_2 = \{0, 1, 2, 3\}$;

$\Omega_3 = \{1, 2, 3, 4, 5, 6\}$;

$\Omega_4 = \{(x, y) | 1 \leqslant x \leqslant 6, 1 \leqslant y \leqslant 6, x \in \mathbf{N}_+, y \in \mathbf{N}_+\}$;

$\Omega_5 = \{1, 2, 3, 4, \cdots\}$;

$\Omega_6 = \{x | 0 \leqslant x < +\infty\}$.

1.1.4 随机事件

定义 1.1 设随机试验 E 的样本空间为 Ω, 称 Ω 的任一子集为随机试验 E 的一个**随机事件**, 简称**事件**, 记为 $A, B, C, \cdots$. 设 A 是 E 的一个事件, 在一次试验中, 当且仅当事件 A 中的某一个样本点出现时, 称事件 A 发生.

只包含一个样本点的集合, 就是基本事件. 显然, 样本空间 Ω 包含多少个样本点, 就包含多少个基本事件. 事件 A 包含几个样本点就相当于包含几个基本事件.

样本空间 Ω 包含所有的样本点, 它是 Ω 自身的子集, 在每次试验中它总是发生的, 称为**必然事件**. 空集 $\varnothing$ 不包含任何样本点, 它也作为样本空间的子集, 它在每次试验中都不会发生, 称为**不可能事件**.

下面举几个事件的例子.

例 1.1 E_4={同时掷两颗骰子, 观察其点数出现的情况}, 对应的样本空间为 $\Omega_4 = \{(x, y) | 1 \leqslant x \leqslant 6, 1 \leqslant y \leqslant 6, x \in \mathbf{N}_+, y \in \mathbf{N}_+\}$.

(1) A_1={两颗骰子点数和为 5}, 则 $A_1 = \{(1,4),(2,3),(3,2),(4,1)\}$;

(2) A_2={两颗骰子点数相同}, 则 $A_2 = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$;

(3) A_3={两颗骰子点数和为 12}, 则 $A_3 = \{(6,6)\}$;

(4) A_4={两颗骰子点数和为 13}, 则 $A_4 = \varnothing$ 为不可能事件;

(5) A_5={两颗骰子点数之差不超过 5}, 则 $A_5 = \Omega_4$ 为必然事件.

1.1.5 事件的关系与运算

由于事件是一个集合, 因而事件之间的关系与事件之间的运算自然按照集合论中集合之间的关系与集合运算来处理. 事件的关系及运算对于研究复杂的事件是十分重要的. 下面给出这些关系和运算在概率论中的提法, 并根据 "事件发生" 的含义, 给出它们在概率论中的含义.

(1) 事件的包含

$A \subset B$, 称为事件 B 包含事件 A, 或者说事件 A 是事件 B 的子事件. 这时集合 A 中的样本点一定属于集合 B. 它表示在一次试验中事件 A 发生必然导致事件 B 发生.

(2) 事件的相等

$A = B$, 称为事件 A 与事件 B **相等**. 其充分必要条件是 $A \subset B$ 且 $B \subset A$. 这时集合 A 与集合 B 中的样本点是相同的. 它表示事件 A 与事件 B 是同一事件.

(3) 事件的和 (并)

$A \bigcup B$, 称为事件 A 与事件 B 的**和(并)事件**. 这个事件是由两个集合 A 与

B 中所有的样本点组成的集合 (两个集合共同的样本点不重复选取). 它表示两个事件 A 与 B 中至少有一个事件发生.

事件的和 (并) 的概念可以推广到有限个或可列无穷多个的情形:

$$A_1 \bigcup A_2 \bigcup \cdots \bigcup A_n \quad \left(\text{简记为 } \bigcup_{i=1}^{n} A_i\right),$$

它表示 n 个事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生.

$$A_1 \bigcup A_2 \bigcup \cdots \bigcup A_n \bigcup \cdots \quad \left(\text{简记为 } \bigcup_{i=1}^{\infty} A_i\right),$$

它表示可列无穷多个事件 $A_1, A_2, \cdots, A_n, \cdots$ 中至少有一个发生.

(4) 事件的积 (交)

$A \bigcap B$, 也记为 AB, 称为事件 A 与事件 B 的**积(交)事件**. 这个事件是由两个集合 A 与 B 中共同的样本点组成的集合. 它表示两个事件 A 与 B 同时发生.

事件的积 (交) 的概念可以推广到有限个或可列无穷多个的情形:

$$A_1 \bigcap A_2 \bigcap \cdots \bigcap A_n = A_1 A_2 \cdots A_n \quad \left(\text{简记为 } \bigcap_{i=1}^{n} A_i\right),$$

它表示 n 个事件 $A_1, A_2, \cdots, A_n$ 同时发生.

$$A_1 \bigcap A_2 \bigcap \cdots \bigcap A_n \bigcap \cdots = A_1 A_2 \cdots A_n \cdots \quad \left(\text{简记为 } \bigcap_{i=1}^{\infty} A_i\right),$$

它表示可列无穷多个事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生.

(5) 事件的差

$A - B$, 称为事件 A 与事件 B 的**差事件**. 它是由属于 A 但不属于 B 的所有样本点组成的集合. 它表示事件 A 发生而事件 B 不发生.

(6) 互斥事件

若 $AB = \varnothing$, 则称事件 A 与事件 B 是**互斥事件**或**互不相容**的. 这时两个集合 A 与 B 没有共同的样本点. 它表示事件 A 与事件 B 不能同时发生. 同一随机试验中, 任意两个基本事件都是互斥的.

(7) 对立事件

若 $A \bigcup B = \Omega$ 且 $AB = \varnothing$, 则称事件 A 与事件 B 为**对立事件**或**互逆事件**. 它表示在一次试验中, 事件 A 与事件 B 必有一个且仅有一个发生. 此时常记 $B = \overline{A}$, 即事件 A 与事件 $\overline{A}$ 互为对立事件. 显然有 $\overline{A} = \Omega - A$, 以及 $A - B = A\overline{B}$.

可以借助于集合关系的维恩图 (Venn) 来直观地表示上述事件间的关系与运算. 见图 1.1. 其中矩形表示样本空间 Ω.

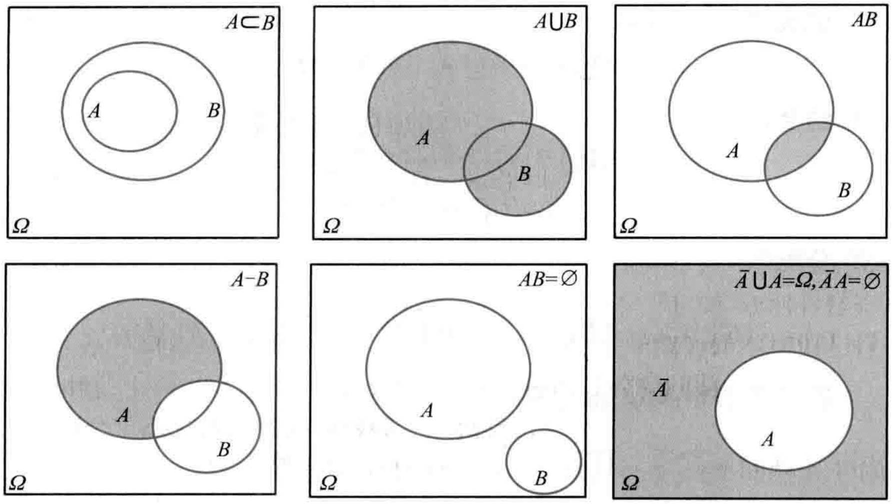

图 1.1

上面利用集合概念描述了事件的关系及运算, 为了将它们与集合论中的相应部分作对照, 列表如表 1.1 所示.

表 1.1 事件在概率论与集合论中含义

符号	概率论	集合论
Ω	样本空间, 必然事件	全集
$\varnothing$	不可能事件	空集
ω	基本事件(样本点)	元素
A	事件	子集
$\overline{A}$	事件 A 的对立事件	A 的余集
$A\subset B$	事件 A 发生必然导致事件 B 发生	A 是 B 的子集
$A=B$	事件 A 与事件 B 相等	A 与 B 相等
$A\bigcup B$	事件 A 与事件 B 至少有一个发生	A 与 B 的并集
$A\bigcap B$	事件 A 与事件 B 同时发生	A 与 B 的交集
$A-B$	事件 A 发生而事件 B 不发生	A 与 B 的差集
$AB=\varnothing$	事件 A 与事件 B 不可能同时发生	A 与 B 没有公共元素

由于事件、事件的关系及运算与集合、集合的关系及运算是相当的, 故根据集合的运算性质可推得事件的运算性质如下:

(1) **交换律:**

$$A \bigcup B = B \bigcup A, \quad AB = BA;$$

(2) **结合律:**

$$(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C),$$
$$(AB)C = A(BC);$$

(3) **分配律:**

$$A(B \bigcup C) = AB \bigcup AC, \quad A \bigcup (BC) = (A \bigcup B)(A \bigcup C);$$

(4) **德摩根 (De Morgan) 律:**

$$\overline{A \bigcup B} = \overline{A}\ \overline{B}, \quad \overline{AB} = \overline{A} \bigcup \overline{B}.$$

德摩根律在事件的运算中经常用到, 它可以推广到更多个事件的情况, 即

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

用语言表述为: 事件和的对立事件等于对立事件的积, 事件积的对立事件等于对立事件的和.

以上运算证明从略.

例 1.2 设 A,B,C 为三个事件, 用 A,B,C 的运算关系表示下列事件:

(1) A 发生, B 与 C 均不发生;

(2) A, B, C 都发生;

(3) A,B,C 都不发生;

(4) A, B, C 中至少有一个发生;

(5) A, B, C 不多于两个发生;

(6) A, B, C 中恰有一个发生;

(7) A, B, C 中至少有两个发生.

解 (1) $A\ \overline{B}\ \overline{C}$ 或 $A\ (\overline{B \bigcup C})$;

(2) ABC;

(3) $\overline{A}\ \overline{B}\ \overline{C}$;

(4) $A \bigcup B \bigcup C$;

(5) $\overline{ABC}$ 或 $\overline{A} \bigcup \overline{B} \bigcup \overline{C}$;

(6) $A\overline{B}\ \overline{C} \bigcup \overline{A}\ B\ \overline{C} \bigcup \overline{A}\ \overline{B}\ C$;

(7) $AB \bigcup AC \bigcup BC$ 或 $A\ B\ \overline{C} \bigcup A\ \overline{B}\ C \bigcup \overline{A}\ B\ C \bigcup\ ABC$.

例 1.3 图 1.2 所示电路. 事件 A 表示信号灯亮, 事件 B,C,D 表示对应开关闭合,

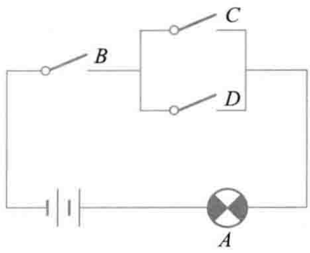

图 1.2

则事件 A,B,C,D 有关系 $BC\subset A, BD\subset A, A=BC\bigcup BD, \overline{B}A=\varnothing$.

例 1.4 袋中有 10 个球, 分别编号为 $1,2,\cdots,10$, 从中任取一球, 设事件

A={取出的球号码为偶数};

B={取出的球号码为奇数};

C={取出的球号码小于 5},

则事件

(1) $A\bigcup B=\Omega$ 为必然事件;

(2) $AB=\varnothing$ 为不可能事件;

(3) AC={取出的球号码为 2 或 4};

(4) $\overline{AC}$={取出的球号码为 $1,3,5,6,7,8,9,10$ 之一};

(5) $\overline{A}\ \overline{C}$={取出的球号码为 $5,7,9$ 之一};

(6) $\overline{B\bigcup C}$={取出的球号码为 $6,8,10$ 之一};

(7) $A-B=A$.

§1.2 频率与概率

对于一个事件(除必然事件和不可能事件外)来说, 它在一次试验中可能发生, 也可能不发生. 我们常常希望知道某些事件在一次试验中发生的可能性有多大. 例如, 为了确定水坝的高度, 就要知道河流在造水坝地段每年最大洪水达到某一高度这一事件发生的可能性大小. 我们希望找到一个合适的数来表征事件在一次试验中发生的可能性大小. 为此, 首先引入频率, 它描述了事件发生的频繁程度, 进而引出表征事件在一次试验中发生的可能性大小的数 —— 概率.