北京大学数学教学系列丛书

本科生
数学基础课教材

概率与统计

陈家鼎 郑忠国 编著

北京大学出版社
PEKING UNIVERSITY PRESS

北京大学数学教学

概率与统计

陈家鼎　郑忠国　编著

北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

概率与统计/陈家鼎,郑忠国编著.—北京:北京大学出版社,2007.8
(北京大学数学教学系列丛书)
ISBN 978-7-301-12114-6

Ⅰ.概… Ⅱ.①陈… ②郑… Ⅲ.①概率论-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(2007)第 063853 号

书　　　名:概率与统计
著作责任者:陈家鼎　郑忠国　编著
责 任 编 辑:刘　勇　曾琬婷
标 准 书 号:ISBN 978-7-301-12114-6/O·0716
出 版 发 行:北京大学出版社
地　　　址:北京市海淀区成府路 205 号　100871
网　　　址:http://www.pup.cn　电子邮箱:zpup@pup.pku.edu.cn
电　　　话:邮购部 62752015　发行部 62750672　理科编辑部 62752021
出版部 62754962
印　刷　者:北京大学印刷厂
经　销　者:新华书店
890 mm×1240 mm　A5　18.75 印张　554 千字
2007 年 8 月第 1 版　2007 年 8 月第 1 次印刷
印　　　数:0001—4000 册
定　　　价:29.00 元

序　言

自1995年以来，在姜伯驹院士的主持下，北京大学数学科学学院根据国际数学发展的要求和北京大学数学教育的实际，创造性地贯彻教育部“加强基础，淡化专业，因材施教，分流培养”的办学方针，全面发挥我院学科门类齐全和师资力量雄厚的综合优势，在培养模式的转变、教学计划的修订、教学内容与方法的革新，以及教材建设等方面进行了全方位、大力度的改革，取得了显著的成效。2001年，北京大学数学科学学院的这项改革成果荣获全国教学成果特等奖，在国内外产生很大反响。

在本科教育改革方面，我们按照加强基础、淡化专业的要求，对教学各主要环节进行了调整，使数学科学学院的全体学生在数学分析、高等代数、几何学、计算机等主干基础课程上，接受学时充分、强度足够的严格训练；在对学生分流培养阶段，我们在课程内容上坚决贯彻“少而精”的原则，大力压缩后续课程中多年逐步形成的过窄、过深和过繁的教学内容，为新的培养方向、实践性教学环节，以及为培养学生的创新能力所进行的基础科研训练争取到了必要的学时和空间。这样既使学生打下宽广、坚实的基础，又充分照顾到每个人的不同特长、爱好和发展取向。与上述改革相适应，积极而慎重地进行教学计划的修订，适当压缩常微、复变、偏微、实变、微分几何、抽象代数、泛函分析等后续课程的周学时。并增加了数学模型和计算机的相关课程，使学生有更大的选课余地。

在研究生教育中，在注重专题课程的同时，我们制定了30多门研究生普选基础课程(其中数学系18门)，重点拓宽学生的专业基础和加强学生对数学整体发展及最新进展的了解。

教材建设是教学成果的一个重要体现。与修订的教学计划相配合，我们进行了有组织的教材建设。计划自1999年起用8年的时间

修订、编写和出版 40 余种教材。这就是将陆续呈现在大家面前的《北京大学数学教学系列丛书》。这套丛书凝聚了我们近十年在人才培养方面的思考，记录了我们教学实践的足迹，体现了我们教学改革的成果，反映了我们对新世纪人才培养的理念，代表了我们新时期的数学教学水平。

经过 20 世纪的空前发展，数学的基本理论更加深入和完善，而计算机技术的发展使得数学的应用更加直接和广泛，而且活跃于生产第一线，促进着技术和经济的发展，所有这些都正在改变着人们对数学的传统认识。同时也促使数学研究的方式发生巨大变化。作为整个科学技术基础的数学，正突破传统的范围而向人类一切知识领域渗透。作为一种文化，数学科学已成为推动人类文明进化、知识创新的重要因素，将更深刻地改变着客观现实的面貌和人们对世界的认识。数学素质已成为今天培养高层次创新人才的重要基础。数学的理论和应用的巨大发展必然引起数学教育的深刻变革。我们现在的改革还是初步的。教学改革无禁区，但要十分稳重和积极；人才培养无止境，既要遵循基本规律，更要不断创新。我们现在推出这套丛书，目的是向大家学习。让我们大家携起手来，为提高中国数学教育水平和建设世界一流数学强国而共同努力。

张 继 平

2002 年 5 月 18 日

于北京大学蓝旗营

前　　言

概率论是研究自然界、人类社会及技术过程中随机现象的数量规律的一门数学. 数理统计学则是以概率论为指导，研究如何有效地收集和分析数据，以对所考察的问题进行推断或预测，直至为采取一定的决策和行动提供依据和建议. 随着现代科学技术的迅速发展和人类生活条件的不断改进，概率论和数理统计学得到了蓬勃的发展. 二者不仅形成了系统的理论，而且在自然科学、人文社会科学、工程技术及经济管理等方面有越来越广泛的应用. 很多院校都开设概率论课、数理统计课或概率统计课.

最近几年我们二人一直担任北京大学数学科学学院为全校本科生开设的基础课程——概率论和数理统计的教学工作. 这两门课各有 60 学时，学生来自文科、理科和医科的多个不同院系. 本书正是在我们讲稿的基础上经过修改、扩充而成的，其中第一章至第五章由陈家鼎编写，第六章至第十章由郑忠国编写.

我们在编写过程中参考了国内外已有的特别是近十年出版的多部优秀教材(见本书的参考文献)，注意吸收这些教材中好的讲法和具体例子. 我们在编写中注意了下面三点.

(1) 恰当处理逻辑严谨性与生动直觉的关系，使学生既有严谨的抽象思维能力又有概率统计的直觉与对随机性的想象力. 通过各方面的例子介绍有关的概念、方法和定理的实际含义，注意引导学生的思维从直觉和想象上升到科学的抽象. 例如，既介绍了概率的“频率定义”和“主观定义”，又介绍了“公理化定义”，说明后者是在前者基础上的科学抽象. 先介绍随机变量的直观含义和直观描述，然后介绍随机变量的严格定义. 在介绍数学期望时先用加权平均的思想介绍离散型随机变量的期望，然后对一般的随机变量用离散型随机变量逼近的办法定义期望. 对每个定理都给出确切的论述，能不用测度论证明的尽量写出证明，但由于教学时数的限制，许多证明打上 * 号或用小字排印，不要求

学生掌握.例如,对“两个随机变量之和的期望等于两个随机变量的期望之和”这一重要定理,我们在正文中只叙述了结论,但其详细证明则放在附录里小字排印.对“中心极限定理”和有关充分统计量的“因子分解定理”则不叙述证明.

(2)认真贯彻理论联系实际的原则.既要使学生掌握概率和统计的基本理论,又要使学生认识这些理论如何灵活运用于实际,从而培养学生解决实际问题的能力.要做到这一点,必须要用心地列举贴近时代生活的,使学生感兴趣的多方面的应用例子.本书努力朝这个方向做.除了叙述日常生活、工业、商业、医学及管理等方面的典型应用例子(包括一些著名例子)外,还介绍一些较复杂的灵活应用例子.例如,第一章中作为独立试验序列的应用,介绍了乒乓球赛制的概率分析;第二章讲述随机变量取值的分散性时,除了“方差”外还介绍了经济学中常用的“基尼系数”;在讲述正态分布的性质之后,介绍了当今工业质量管理工作中广泛关注的“六西格玛”;第九章中作为回归分析的应用介绍了高考作文评分的监控方法,等等.

本书特别注重对理论联系实际的难点进行化解.例如,对“假设检验”,避免单纯从逻辑推理进行论述,着重从多方面的应用实例说明假设检验问题的提法、零假设的设置及两类错误的概率.把实际中的检验问题分成两大类:决策性检验问题和显著性检验问题.有些检验问题强调控制第一类错误的概率(例如第八章例 1.4),有些检验问题则重点在控制第二类错误的概率(例如第八章例 1.5).本书还用一定篇幅介绍 p 值的概念和用法.又如介绍“回归分析”的应用时把自变量分为两类:可控制的和不可控制的,把自变量和反应变量之间的关系分为两类:因果关系和非因果性的相关关系.本书还特别关注数据的来源和变量的性质.

(3)在叙述方法与内容编排上注意基本内容与进一步内容、重点与非重点的界限,力求做到层次分明,便于教和学.我们认为,大学教材应比教学大纲规定的多一些,更应比课堂实际讲授的多一些.这样做有利于教师根据实际情况灵活掌握,有利于学生课外阅读,使有余力的学生可以选学更多的东西.本书中凡打 * 号和小字排印的部分均不是基本内容,不要求学生掌握.有些内容虽未标上 * 号也非小字排印,教师

也可根据实际情况确定为非基本内容.

本教材虽是按两学期的教学安排("概率论"课一学期,"数理统计"课一学期)编写的,但是也可作为一学期的"概率统计"课的教材.作为后者使用时,应选定书中最基本的部分.笔者建议选择下列内容:

第一章(不含§1.7),第二章(不含§2.8),第三章(不含§3.7,§3.8),第四章§4.2和§4.3的部分内容,第五章的§5.1,第六章,第七章的§7.1,§7.5,第八章的§8.1,§8.2,§8.4中关于正态总体参数的检验方法,§8.6中的χ^2检验,第九章的§9.1,§9.2及§9.3至§9.5中方法的应用部分,第十章的§10.1.

北京大学出版社刘勇和曾琬婷同志对本书的出版付出了辛勤的劳动,我们在此向他们表示感谢.

由于我们水平有限,本书一定有不少缺点和谬误,欢迎读者和专家批评指正.

陈家鼎　郑忠国

2007年6月于北京大学数学科学学院

内容简介

本书系统论述概率和统计的概念、方法、理论及其应用，是一部为高等院校本科生学习概率论和数理统计而编写的教材或教学参考书。本书不仅提供了这个学科领域的基本内容，而且叙述了在日常生活、自然科学、技术科学、人文社会科学及经济管理等各方面的应用例子。全书共分十章，内容包括：随机事件与概率，随机变量与概率分布，随机向量，概率极限定理，随机过程，统计学中的基本概念，估计，假设检验，回归分析，统计决策和贝叶斯分析简介。本书恰当处理逻辑严谨性与生动直觉的辩证关系，使学生既有严谨的抽象思维能力，又对随机现象具有直觉想象力；认真贯彻理论联系实际，应用举例贴近时代生活；概率论部分强调了随机现象在社会生活和科学技术中的广泛性及所具有的内在规律，统计学部分则强调了其数据处理的功能，二者都以认识随机性、恰当处理随机性(包括决策和行动)为目标；内容选取上注意对难点进行化解，叙述通俗易懂，结构层次分明，使学生易于理解与掌握。

本书可作为高等学校理工类本科学生的教材或教学参考书，也可供经济管理和财经类等有关专业的研究生和从事统计计算的科技人员阅读。

作者简介

陈家鼎　北京大学数学科学学院教授、博士生导师，1959年毕业于北京大学数学力学系。长期从事数理统计的教学和科研工作，研究方向是不完全数据的统计推断、序贯统计及其在可靠性工程上的应用，发表论文50多篇。曾任北京大学概率统计系系主任、北京大学数学科学学院副院长、中国概率统计学会理事长、中国统计学会副会长。主持完成“序贯分析”、“生存分析与可靠性的若干前沿问题”等多项国家自然科学基金和教育部博士点基金项目。主编的教材《数理统计学讲义》获国家教委优秀教材一等奖(高等教育出版社，1995)。与郑忠国等合作的项目“可靠性评定的数学理论与应用”获北京市科技进步二等奖(2002)。

郑忠国　北京大学数学科学学院教授、博士生导师，1962年毕业于北京大学数学力学系，1965年北京大学研究生毕业。长期从事数理统计的教学和科研工作，研究方向是非参数统计、可靠性统计以及统计计算，发表论文近百篇。主持完成国家自然科学基金项目“不完全数据统计理论及其应用(1999—2001)”，教育部博士点基金项目“应用统计方法研究”和“工业与医学中的应用统计研究”等。研究项目“随机加权法”获国家教委科技进步二等奖。出版的教材有《高等统计学》(北京大学出版社，1995)。

《北京大学数学教学系列丛书》编委会

目　　录

第一章　随机事件与概率

§1.1　随机事件及其概率

在自然科学里，大家都很熟悉的许多定律（或规律）往往可以陈述为这样一种形式：

"在一组条件 S 实现之下，某事件 A 必然发生."

例如：(1) 在标准大气压下，纯水加热至 100℃，则必然开始汽化.

(2) 空中初速为 0 的物体，在重力作用下经过时间 t（秒）下落的距离一定是$\frac{1}{2}gt^2$（这里 g 是重力加速度）.

(3) 在真空中，光的传播速度为定值.

(4) 在室温（10℃ 至 40℃）下，生铁肯定不能熔化.

上述陈述方式是一种"确定性"的判断，它断言在一组条件 S 实现下，某事件 A 必然发生. 我们以后就把这种在 S 实现之下必然发生的事件叫做条件 S 之下的**必然事件**.

但是，我们在日常生活、实际工作和科学研究中，遇到更多的判断不是属于上面那种类型，而具有这样的形式：

"在一组条件 S 实现之下，某事件 A 可能发生也可能不发生."

例 1.1　北京的冬季（条件 S）至少降两次雪（事件 A）.

例 1.2　从一堆含有次品的产品中，随便抽一件，遇到次品.

例 1.3　某射手向远处的小目标射击一发子弹，命中目标.

例 1.4　在桌面上投掷一枚匀称的硬币，出现国徽朝上.

这里的各个事件"至少降两次雪"、"遇到次品"、"命中目标"、"国徽朝上"都不是必然事件. 事情很明显：北京有的年份冬季只降过一次雪；随便抽一件产品，可能遇到正品；一个射手无论他的技术水平多高，对远处的小目标不能保证一定打中；硬币既然匀称，当然也会出现国徽朝下的情形.

这种在条件 S 之下可能发生也可能不发生的事件 A,我们称之为**偶然事件**.

偶然事件也叫偶然现象,是我们经常碰到的.偶然事件的发生与否虽然不能预先知道,但各种偶然事件发生的可能性是有大小之分的.一个偶然事件,其发生的可能性的大小,是这个事件在条件 S 下的固有属性,并不依赖于人们的主观意志.事实上,从我们日常生活和科学实验中已经逐渐积累了这种经验,只是没有严格地给予数量刻画而已.例如上述的例 1.1 至例 1.4 都涉及偶然事件,但我们知道,北京的冬季至少降两次雪的可能性是很大的,比起掷硬币时出现"国徽朝上"的可能性要大.至于例 1.2 和例 1.3 就要具体分析了.如果那堆产品中次品多,则抽到次品的可能性就大;类似地,如果射手水平高,则射中目标的可能性就大.这种体验人皆有之.这表明,偶然事件发生的可能性的大小是客观存在的,是由该事件与条件 S 的内在联系所决定的.怎样刻画可能性的大小呢?我们用区间[0,1]中的一个数来刻画,并把这个数叫做该偶然事件的概率(也叫几率),概率越大表示可能性越大.给定条件 S 及事件 A,怎样定义与计算 A 的概率呢?这个问题的回答不简单,需要考虑两种不同类型的前提条件.

首先考虑条件 S 可以不断重复实现的情形.

定义 1.1 在不变(或基本不变)的一组条件 S 之下,重复做 n 次实验(或观测,下同!),设 μ 是 n 次实验中事件 A 发生的次数.如果对于大量的实验(即 n 很大),发生的频率 $\frac{\mu}{n}$ 稳定地在某一数值 p 左右摆动,而且随着实验次数的增多一般说来这种摆动的幅度变小,则称 A 为**随机事件**(或称有概率的事件),并称数值 p 为随机事件 A 在条件 S 下发生的**概率**,记为 $P(A|S)=p$ 或简写为 $P(A)=p$.

定义 1.1 也可简单地说成:"发生的频率有稳定性的事件叫做随机事件,频率的摆动中心叫做该随机事件的概率."

定义 1.1 中的概率 $P(A)$ 就是 A 在条件 S 下发生的可能性大小的数量刻画.大量实践表明,只要条件 S 能不变(或基本不变)地不断重复实现,事件 A 发生的频率一般都有稳定性,因而事件 A 一般都是随机事件,即是有概率的事件.(见注 1)