大学数学

概率论与数理统计

（第二版）

哈尔滨工业大学数学系
王　勇　主编

高等教育出版社

大学数学

概率论与数理统计

Gailülun yu Shuli Tongji

（第二版）

哈尔滨工业大学数学系

主编　王　勇

编　　方　茹　周永春　李朝艳

　　　田波平　王　勇

高等教育出版社·北京

内容提要

本书是哈尔滨工业大学所编的大学数学系列教材中的一本，全套教材包括《工科数学分析（第四版）（上、下）》、《线性代数与空间解析几何（第四版）》、《概率论与数理统计（第二版）》共 4 本。

本书注重体现工程实际应用背景且注意为现代概率论与数理统计新知识留有接口，同时精简、压缩一些传统内容，淡化计算技巧的训练，加强理论基础的培养；重新组织、精选了例题及习题，使之更有利于培养工科学生利用概率统计方法解决和分析工程实际问题。

本书内容包括随机事件与概率、条件概率与独立性、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与极限定理、数理统计的基本概念、参数估计、假设检验、单因素试验的方差分析及一元正态线性回归等九章，前 6 章配备了拓展例题，对其理论与方法作适当的加深和拓广。附录介绍了如何使用 MATLAB 软件处理概率统计问题。本书适合本科院校工科各专业学生使用，也可作为报考硕士研究生人员及工程技术人员的学习参考书。

图书在版编目（C I P）数据

大学数学．概率论与数理统计／王勇主编．--2 版．--北京：高等教育出版社，2014.12

ISBN 978-7-04-041219-2

Ⅰ．①大… Ⅱ．①王… Ⅲ．①高等数学-高等学校-教材②概率论-高等学校-教材③数理统计-高等学校-教材 Ⅳ．①O13②O21

中国版本图书馆 CIP 数据核字（2014）第 232408 号

策划编辑 张晓丽　责任编辑 张晓丽　封面设计 于文燕　版式设计 马敬茹
插图绘制 黄建英　责任校对 刘春萍　责任印制 毛斯璐

出版发行 高等教育出版社
社　址 北京市西城区德外大街 4 号
邮政编码 100120
印　刷 三河市宏图印务有限公司
开　本 787mm×960mm 1/16
印　张 20.75
字　数 370 千字
购书热线 010-58581118
咨询电话 400-810-0598
网　址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.landraco.com
http://www.landraco.com.cn
版　次 2007 年 7 月第 1 版
2014 年 12 月第 2 版
印　次 2014 年 12 月第 1 次印刷
定　价 30.40 元

物 料 号 41219-00

第二版前言

本书的第一版是普通高等教育“十一五”国家级规划教材，根据多年在哈尔滨工业大学使用本教材的教学实践，并吸取了近年来国内外出版的多部概率统计优秀教材的一些好的方法，同时也认真考虑了使用本教材的相关单位的意见，在此基础上确定了本次修订的框架。本版对第一版的前7章进行了改写与扩充，以便在内容上和编排上更好地适应新时期对“概率论与数理统计”课程教学及应用的需要。

本版教材的特点可概述如下：

(1) 新增了MATLAB软件应用的内容，简单介绍了MATLAB软件及其在概率统计中的应用。

(2) 在前6章增加了“拓展例题”一节，希望扩大知识面，同时激发学生的学习兴趣。这部分主要包括各章涉及的数学实验，现代科技的应用及提高性内容，主要目的是为学生的个性化发展创造良好的条件。

(3) 对第一版的一些教学内容作了补充和修正。并对原有的例题和习题作了一些调整，增加了有关加强对基本概念理解，对重要统计方法掌握的例题和习题及近年来硕士研究生入学考试题。

(4) 在附录中给出了更多的常用统计分布，并给出了实际应用的例子，满足学生进一步学习需要。

参加本书修订工作的教师有方茹，周永春，李朝艳，田波平，王勇。

由于编者水平有限，书中的错误和不妥之处在所难免，恳请广大读者批评指正，以期不断完善。

编　者

2014年2月于哈尔滨工业大学

第一版前言

为适应21世纪高等学校学生和广大工程技术人员对数学的需求,我校作为国家工科数学教学基地之一,多年来在数学教学改革方面进行了一定的探索,已经初见成效。结合这些教学改革成果,我校编写了《大学数学》系列教材,本书就是其中的一本。

概率论与数理统计是研究随机现象统计规律的一门数学学科,已在包括控制、通信、生物、物理、力学、金融、社会科学以及其他工程技术等诸多领域中获得了广泛的应用。学习和掌握概率论与数理统计的基本理论和基本方法并将其应用于科学研究和工程实际中,是社会发展对高素质人才培养提出的必然要求。

概率论与数理统计课程是工科大学的一门应用性很强的必修基础课。针对这门课程的特点,我们编写本书的基本思路是:重概念、重方法、重应用、重能力的培养。本书较详细地阐述了概率论与数理统计的基本概念和基本理论,并叙述了一些主要概念和方法产生的背景和思路。从直观分析入手逐步过渡到严格的数学表述,主线清晰,使初学者易于入门,易于掌握概率论与数理统计的基本理论和方法。书中的例题和习题较丰富,其中包括大量的应用题,有助于培养学生分析问题和解决问题的能力。

本书是编者在多年的教学实践和教学经验的基础上,按照我国现行的工科大学本科数学课程教学基本要求和工学、经济学硕士研究生入学考试大纲编写的。全书内容分六个部分:第一部分为随机事件及其概率(第1章、第2章);第二部分为随机变量及其分布(第3章、第4章);第三部分为随机变量的数字特征与极限定理(第5章);第四部分为数理统计的基本概念(第6章);第五部分为参数估计与假设检验的基本方法(第7章、第8章);第六部分为线性模型的统计分析初步(第9章)。每章后的习题和书末的补充习题较多地收入了历年工学、经济学硕士研究生入学考试题。

本书可作为工科大学本科各专业的教材,也可供财经、理科类某些专业选用,还可以作为报考硕士研究生人员及工程技术人员的学习参考书。

哈尔滨工业大学曹彬教授、许承德教授对本书的编写始终给予了关心与帮助,谨在此致谢。

由于编者水平有限,书中的错误和不妥之处在所难免,恳请广大读者批评指正,以期不断完善。

编　者

2007 年 1 月于哈尔滨工业大学

目　　录

第0章　引　　言

0.1　概率论与数理统计发展简史

概率论与数理统计是研究随机现象及其规律性的一门数学学科.研究随机现象的规律性有其独特的思想方法,它不是寻求出现每一现象的一切物理因素,不能用研究确定性现象的方法来研究随机现象,而是承认在所研究的问题中存在有一些人们不能认识或者根本不知道的随机因素作用下,发生了随机现象.这样,人们既可以通过试验来观察随机现象,揭示其规律性,作出决策,也可以根据实际问题的具体情况找出随机现象的规律,作出决策.

概率论是基于给出随机现象的数学模型,并用数学语言来描述它们,然后研究其基本规律,透过表面的偶然性,找出其内在规律性,建立随机现象与数学其他分支的桥梁,使得人们可以利用已成熟的数学工具和方法来研究随机现象,进而也为其他数学分支和其他新兴学科提供了解决问题的新思路和新方法.数理统计是以概率论为基础,基于有效地观测、收集、整理、分析带有随机性的数据来研究随机现象,进而对所观察的问题作出推断和预测,直至为采取一定的决策和行动提供依据和建议.

概率论的起源与赌博有关.17 世纪中叶,法国数学家帕斯卡(Pascal)、费马(Fermat)及荷兰数学家惠更斯(Huygens)基于排列组合方法,研究利用古典概型解决赌博中提出的一些问题,如“分赌注问题”、“赌徒输光问题”等.到了 18,19 世纪,随着科学的发展,人们注意到社会科学和自然科学中许多随机现象与机会游戏之间十分相似,如人口统计、误差理论、产品检验和质量控制等,从而由机会游戏起源的概率论被应用于这些领域中,同时也大大促进了概率论本身的发展,瑞士数学家伯努利(Bernoulli)作为使概率论成为数学的一个分支的奠基人之一,建立了概率论中第一个极限定理(即伯努利大数定律),阐明了事件发生的频率稳定于它的概率.随后,棣莫弗(De Moivre)和拉普拉斯(Laplace)又导出第二个基本极限定理(即中心极限定理)的原始形式,拉普拉斯在其《分析的概率理论》一书中,明确给出了概率的古典定义,并在概率论中引入了更有力的分析工具,将概率论推向一个新的发展阶段.19 世纪末,俄国数学家切比雪夫(Chebyshev)、马尔可夫(Markov)、李雅普诺夫(Liapunov)等人用分析的方法建

立了大数定律及中心极限定理的一般形式,科学地解释了为什么在实际中遇到的许多随机变量都近似地服从于正态分布.20世纪初,由于大量实际问题的需要,特别是受物理学的刺激,人们开始研究随机过程.爱因斯坦(Einstein)、维纳(Wiener)和列维(Levi)等人对生物学家布朗(Brown)在显微镜下观测到的花粉微粒的无规则运动进行了开创性的理论分析,提出了布朗运动数学模型,并进行了系统的研究;爱尔兰(Erlang)等人则在电话流呼唤中研究了泊松(Poisson)过程,成为排队论的开创者;费勒(Feller)等在生物群体生长模型中提出了生灭过程;克拉默(Cramer)、维纳、辛钦(Khinchin)等人系统研究了平稳过程;柯尔莫果洛夫(Kolmogorov)、费勒和多布(Doob)则开创了更一般的马尔可夫过程和鞅论的系统研究.至今,对于随机过程的研究以及与其他新兴学科的交叉而形成的边缘学科的研究仍在继续.

如何定义概率,如何把概率论建立在严格的逻辑基础上,对这一个问题的探索一直持续了三个世纪.20世纪初,勒贝格(Lebesgue)完成的测度与积分理论,为概率公理化体系的建立奠定了基础.特别是苏联数学家柯尔莫果洛夫于1933年在其著作《概率论基础》一书中首次给出了概率的测度论式的严格定义,归纳总结了事件及事件的概率的基本性质和关系,建立了概率论的公理化体系,柯尔莫果洛夫公理化方法成为近代概率论的基础,使概率论成为严谨的数学分支,对近代概率论的发展起到了积极的作用.

数理统计是随着概率论的发展而发展起来的.只有当人们认识到必须把数据视为来自具有一定概率分布的总体,所研究的对象是这个总体而不能局限于数据本身的时候,数理统计才诞生了.早在19世纪中期之前,数理统计已出现若干重要的工作,特别是高斯(Gauss)与勒让德(Legendre)关于观测数据的误差分析和最小二乘估计方法的研究成果.但是直到20世纪初期,数理统计才发展成为一门成熟的学科,其中皮尔逊(Pearson)与费希尔(Fisher)作出重大贡献.1946年,克拉默发表的《统计学的数学方法》是第一部严谨且比较系统的数理统计著作.

至今,概率论与数理统计的理论与方法已广泛应用于自然科学、社会科学及人文科学等各个领域中,并且随着计算机的普及,概率论与数理统计已成为处理信息、制定决策的重要理论和方法.它们不仅是许多新兴学科,如信息论、控制论、排队论、可靠性理论及人工智能的数学理论基础,而且与其他领域的新兴学科的相互交叉而产生了许多新的分支和边缘学科,如生物统计、统计物理、数理金融、神经网络统计分析、统计计算等.总之,概率论与数理统计作为理论严谨、应用广泛、发展迅速的数学分支正越来越引起广泛的重视.

0.2　概率论与数理统计研究问题的方法

概率论应用随机变量(多维随机变量)与随机变量的概率分布、数字特征及特征函数为数学工具对随机现象进行描述、分析与研究,其前提条件是假设随机变量的概率分布是已知的;而数理统计中作为研究对象的随机变量的概率分布是完全未知的,或者分布类型已知,但其中某些参数或某些数字特征是未知的.概率论研究问题的方法是从假设、命题、已知的随机现象的事实出发,按一定的逻辑推理得到结论的,因此概率论的方法本质上是演绎式的;而统计学的方法是归纳式的,从所研究对象的全体中随机抽取一部分进行试验或观测,以获得试验数据,依据试验数据所获取的信息,对整体作出推断,是"归纳"而得到结论的.例如,统计学家通过大量观测得到的试验数据,按照一定的统计方法得出结论:吸烟与患肺癌有关;吸烟与患支气管炎有关.此结论不是用数学逻辑推理方法证明得到的.因此,掌握统计学的思想与方法对于初学者无疑是很重要的.

第1章　随机事件与概率

随机事件的概率是概率论研究的基本内容.本章将主要介绍概率论中的基本概念——随机事件与随机事件的概率,并进一步讨论随机事件的关系与运算以及概率的性质与计算方法.

1.1　随机事件

1.1.1　必然现象与随机现象

人们在实践活动中所遇到的现象,一般来说可分为两类:一类是**必然现象**,或称**确定性现象**;另一类是**随机现象**,或称**不确定性现象**.

必然现象是指在相同条件下重复试验,所得结果总是确定的现象,只要试验条件不变,试验结果在试验之前是可以预言的.例如,在标准大气压下,将水加热到 100 ℃,水必然沸腾;用手向空中抛出的石子,必然下落;作等速直线运动的物体,如无外力作用,必然继续作等速直线运动,等等,这些现象都是必然现象.

随机现象是指在相同条件下重复试验,所得结果不一定相同的现象,即试验结果是不确定的现象.对这种现象来说,在每次试验之前哪一个结果发生,是无法预言的.例如,新生婴儿,可能是男孩,也可能是女孩;向一目标进行射击,可能命中目标,也可能不命中目标;从一批产品中,随机抽检一件产品,结果可能是合格品,也可能是次品;测量某个物理量,由于许多偶然因素的影响,各次测量结果不一定相同,等等,这些现象都是随机现象.

对随机现象,是否有规律性可寻呢?人们经过长期的反复实践,发现这类现象虽然就每次试验结果来说,具有不确定性,但大量重复试验,所得结果却呈现出某种规律性.例如,

(1) 掷一枚质量均匀的硬币,当投掷次数很大时,就会发现正面和反面出现的次数几乎各占 1/2.历史上,蒲丰(Buffon)掷过 4 040 次,得到 2 048 次正面;皮尔逊掷过 24 000 次,得到 12 012 次正面.

(2) 对一个目标进行射击,当射击次数不多时,对弹孔分布看不出有什么规律性;但当射击次数非常多时,就可发现弹孔的分布呈现一定的规律性,弹孔关于目标的分布略呈对称性,且越靠近目标的弹孔越密,越远离目标的弹孔越稀.

(3) 从分子物理学观点来看,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,速度和轨道都是随机的,因而对器壁的碰撞也是随机的.初看起来器壁所受的压力是不稳定的,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,压力越稳定.

从上述各例可以看到,随机现象也包含着规律性,它可在相同条件下的大量重复试验或观察中呈现出来.这种规律性称为**随机现象的统计规律性**.

概率论与数理统计就是研究随机现象统计规律的一门数学学科.

1.1.2 随机试验与事件、样本空间

对随机现象的研究,总是要进行观察、测量或做各种科学试验(为了叙述方便起见,统称为试验).例如,掷一硬币,观察哪面朝上;向一目标进行射击,观察是否命中;从一批产品中随机抽一产品,检查它是否合格;向坐标平面内任投一根针,测量此针的针尖指向与 x 轴正向之间的交角,等等,这些都是试验.仔细分析,这些试验具有如下的共同特点:

(a) 试验可以在相同条件下重复进行;

(b) 试验的所有可能的结果不止一个,而且是事先已知的;

(c) 每次试验总是恰好出现这些可能结果中的一个,但究竟出现哪一个结果,试验前不能确切预言.

如掷硬币的例子,试验是可以在相同条件下重复进行的.试验的可能结果有两个,即正面和反面,每次试验必出现其中之一.但投掷之前是不可能预言出现正面还是出现反面.

人们将满足上述三个条件的试验称为**随机试验**,简称**试验**,以字母 E 表示.

为了研究随机试验,首先要知道这个试验的所有可能结果是哪些.随机试验的每一个可能结果称为**基本事件**也称为**样本点**,用 e 表示.全体基本事件的集合称为**样本空间**,记为 S.

在讨论一个随机试验时,首先要明确它的样本空间.对一个具体的试验来说,其样本空间可以由试验的具体内容确定.下面看几个例子.

例 1.1.1 掷一均匀对称的硬币两次,观察正反面出现情况,这是个随机试验.

可能结果有四个:(正正)、(正反)、(反正)、(反反).括号内的第一个和第二个字,分别表示第一次和第二次掷的结果.故样本空间

$$S=\{(正正),(正反),(反正),(反反)\}.$$

例 1.1.2 记录某电话交换台在一段时间内接到的呼叫次数,这个试验的

基本事件(记录结果)是一非负的整数,由于难以规定一个呼叫次数的上界,所以样本空间为

$$S=\{0,1,2,\cdots\}.$$

例 1.1.3 从一批灯泡中抽取一只灯泡,测试它的使用寿命.设 t 表示寿命,则样本空间为

$$S=\{t:t\geqslant 0\}.$$

例 1.1.4 观察某地区一昼夜最低温度 x 和最高温度 y.设这个地区的温度不会小于 T_0 也不会大于 T_1,则样本空间为

$$S=\{(x,y):T_0\leqslant x<y\leqslant T_1\}.$$

在试验中可能发生也可能不发生的事情称为**随机事件**,简称**事件**,以字母 $A,B,C,\cdots$ 表示.有了样本空间的概念便可以用集合的语言来定义事件.下面先从一个例子来分析.

例 1.1.5 在例 1.1.1 中,若设事件 A="第一次出现正面",在一次试验中,A 发生当且仅当在这次试验中出现基本事件(正正)、(正反)中的一个.这样可以认为 A 是由(正正)、(正反)组成的,而将 A 定义为它们组成的集合

$$A=\{(\text{正正}),(\text{正反})\}.$$

类似地,事件 B="两次出现同一面",C="至少有一次出现正面",D="第一次出现反面",均可定义为集合 $B=\{(\text{正正}),(\text{反反})\}$,$C=\{(\text{正正}),(\text{正反}),(\text{反正})\}$,$D=\{(\text{反正}),(\text{反反})\}$.

一般地,人们将事件定义为基本事件的某个集合,即样本空间的某个子集,称事件 A 发生,当且仅当 A 中某一基本事件出现.

样本空间 S 和空集 $\varnothing$ 作为 S 的子集也看作事件.由于 S 包含所有的基本事件,故在每次试验中,必有一个基本事件 $e\in S$ 发生,即在试验中,事件 S 必然发生.因此,S 是**必然事件**.又因在 $\varnothing$ 中不包含任何一个基本事件,故在任一次试验中,$\varnothing$ 永远不会发生.因此,$\varnothing$ 是**不可能事件**.常用 S,$\varnothing$ 分别表示必然事件与不可能事件.

必然事件与不可能事件可以说不是随机事件,但为了今后研究的方便,还是把它们作为随机事件的两个极端情形来处理.

1.2 事件的关系与运算

在实际问题中,往往要在同一个试验中同时研究几个事件以及它们之间的联系.详细分析事件之间的关系,不仅可以帮助人们更深入地认识事件的本质,而且可以大大简化一些复杂的事件.