普通高等教育“十二五”规划教材

华中科技大学数学创新教材

应用概率统计

李萍　叶鹰　编

科学出版社

普通高等教育"十二五"规划教材

华中科技大学数学创新教材

应用概率统计

李 萍 叶 鹰 编

科 学 出 版 社
北 京

内 容 简 介

本书分两个部分，第一部分为概率论基础，包括随机事件及其概率的性质与计算、随机变量及其分布、数字特征、常用分布及应用；第二部分为数理统计，包括参数与非参数估计方法、参数与非参数假设检验方法、贝叶斯估计方法、线性统计模型的统计推断方法，在上述统计方法中穿插介绍 Excel 统计函数的应用，并附有 Excel 统计函数检索表方便读者进行统计计算时查询.

本书可作为高等院校理工科大学生基础课教材，也可供希望增强概率基础、拓广统计方法的工程技术人员阅读参考.

图书在版编目(CIP)数据

应用概率统计/李萍，叶鹰编. —北京：科学出版社，2013.3

普通高等教育“十二五”规划教材 华中科技大学数学创新教材

ISBN 978-7-03-037210-9

Ⅰ. ①应… Ⅱ. ①李… ②叶… Ⅲ. ①概率统计-高等学校-教材

Ⅳ. ①O211

中国版本图书馆 CIP 数据核字 (2013) 第 056498 号

责任编辑：杨瑰玉 李梦华 孙翠勤／责任校对：张小霞

责任印制：彭 超／封面设计：苏 波

科 学 出 版 社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

武汉市首壹印务有限公司印刷

科学出版社发行 各地新华书店经销

2013 年 3 月第 一 版 开本：B5(720×1000)

2013 年 3 月第一次印刷 印张：21

字数：402 000

定价：38.00 元

(如有印装质量问题，我社负责调换)

《华中科技大学数学创新教材》丛书编委会

丛书序

随着教学实践的深入进行，现行大学数学教育体系已呈现诸多弊病. 一方面，一些学生反映：数学太抽象，学习数学太枯燥，学完之后仅记得几个数学符号和概念，难以做到学以致用；另一方面，一些高年级本科生和研究生反映：本科阶段所学的数学远远不能满足其专业需求，学懂了的数学用不上，要用的数学没学过. 这一切都说明，现行的“教”与“学”、“学”与“用”严重脱节，现行的数学教学已远远不能满足现代教育及高速发展的科学技术的需要，改革与创新势在必行.

我国的大学数学教育长期以来沿用了前苏联的模式：从课程设置来说，着重于近代数学而较少融入现代数学；从教材内容来说，重理论及其推导而轻知识拓展及其应用. 众所周知，数学是自然科学与工程技术的基础，它已渗透到当代社会科学的众多领域，对于培养和开发学生潜能起着重要作用. 如何构建当代大学数学知识体系，使学生乐而学之、学以致用，是摆在我们每位大学数学教师面前的艰巨任务.

对于非数学专业的数学课程设置问题，我们对国内外高校及我校的开课现状进行了全面调研. 从国外高校情况来看，其课程设置各不相同，没有统一模式，而且他们的非数学专业与数学专业、本科生与研究生的大部分课程是相通的，即非数学专业的学生可修数学专业各层次课程. 对于我们来说，这点是没有可参照性的，因为我国高校的各专业、各层次课程基本上是各自为阵，教育管理部门也有相关规定予以制约. 但从其最基本的数学课程所涵盖的知识内容来看，则大致相同，主要涵盖：微积分、代数、几何、复分析、概率论、微分方程理论、统计学及数值计算等课程. 这些数学课程不但可以给学生以数学、统计、计算等方面知识，而且可以开发学生的逻辑思维能力、空间想象能力及知识应用能力. 因此，基于改革与创新的宗旨，我们拟将大学数学的主体知识内容分解为下列课程：《一元分析学》、《多元分析学》、《解析几何与线性代数》、《应用复分析》、《应用概率统计》、《应用偏微分方程》、《科学计算引论》.

鉴于目前非数学专业教材内容不足以满足专业的需求，且部分内容已经老化，因此我们拟在各科教材中适当更换和增加新的内容. 如以往的《计算方法》教材仅有多项式插值、线性方程组的古典迭代法、数值积分、标量非线性方程数值解及常微分方程初值问题数值解的内容，而当今各专业用于计算机仿真的数值算法有广泛需求，其教材内容显然难以满足诸专业需要，本次改革拟增加线性方程组的 Krylov 子空间法、非线性方程组数值解、常微分方程边值问题数值解及偏微分方程数值解等重要内容. 此外，在教材整体架构方面，《一元分析学》、《多元分析学》、《解析几何

何与线性代数》、《应用复分析》将偏重于数学理论以训练和开发学生的逻辑思维能力、空间想象能力; 而《应用概率统计》、《应用偏微分方程》、《科学计算引论》则力求理论与应用二者兼顾, 以开发学生应用知识的能力. 由上可知我们的教学内容改革不仅仅是名称上的变化, 也不只是知识的重新排列组合, 而是一次由表及里的实质性改革.

本套大学数学系列创新教材是由十余位教学经验丰富、科研基础好的专班教师的编写的, 他们对大学数学课程的改革问题进行了深入研究和探讨, 制定出了科学的编写计划. 本套教材特别注重教学内容各部分知识的系统性和逻辑性, 体现其由浅入深、由易到难、由简单到复杂, 按照逻辑系统和认知理论相结合的思想组织整套书出版工作. 力求突出学生的主体地位, 以学生的发展为本, 充分体现数学知识的趣味性、时代性、可实践性及有用性.

大学数学教材的改革与创新是一项长期而艰巨的任务, 我们愿以本套教材的编写为起点和契机, 继续深入开展教学内容的变革, 以使我们的教书育人工作充分适应时代发展的要求.

由于编者水平所限, 仓促付梓, 书中必有疏漏之处, 诚望读者指正.

编委会

2010 年 5 月 20 日

前　　言

本书《应用概率统计》的目的是作为高等院校非数学专业概率统计课程的教材. 具备基本的高等数学素养的科学工作者、工程技术人员和企业管理者也可以将本书作为自修教材或参考用书.

概率论与数理统计是研究随机现象统计规律性及其应用的一门数学学科, 它不仅具有与其他数学学科一样的严谨性, 更具有广泛的应用性. 本书的编写以“夯实基础, 强调应用”为主导. 全书从框架构建到内容组织各环节进行了改革尝试, 在编写过程中力求体现下述特点.

(1) 强化基础, 深化核心内容. 增加了事件域、事件的极限和概率的连续性等内容, 使概率的概念、随机变量的概念、分布函数的连续性及随机变量序列的收敛性等重要概念能严格建立, 以利于学生透彻掌握.

在系统给出随机变量及其分布理论的基础上, 集中介绍常用分布, 增强对各分布应用背景、参数意义及拓展内容的介绍, 并强调有关分布之间的内在联系; 在统计估计的内容中, 改变传统教材平行介绍估计量的各个评选准则的做法, 以均方误差导出无偏性、有效性和一致性准则, 便于学生更本质地理解这一组概念.

(2) 理论联系应用, 在深度和广度上强化应用. 对抽象的极限理论进行分解, 将中心极限定理放在常用分布一章, 让中心极限定理在正态分布的应用背景下建立, 也使正态分布的广泛应用性有了更明确的理论支撑; 将大数定律放到参数估计一章, 作为建立矩估计法的理论依据, 以使学生在明确其应用意义的情形下更好地掌握该理论.

在数理统计部分从信息量递增的角度介绍非参数估计、参数估计贝叶斯统计这几类统计方法, 揭示它们的层次结构. 将例题和习题尽可能使用真实的统计数据, 以突显教材理论的应用价值. 在教材中编入需要编程计算的习题训练学生应用知识的动手能力.

(3) 调整和优化内容的结构. 在内容上, 对古典概率、几何概率和统计概率的内容进行适当压缩, 并将它们作为不同情形下确定概率的方法引入, 增加在理论与实际中越来越有应用价值的混合分布、条件期望等内容; 在框架结构上, 一元分布与多元分布成一章, 回归分析和方差分析归于线性统计模型、常用分布自成一章, 力求使教材的知识结构更紧凑, 更具层次性和系统性.

(4) 增强统计分量, 引入统计技术内容. 增加了贝叶斯统计、非参数检验的基本内容; 将 Excel 软件的统计应用功能及函数穿插在涉及数值计算的内容中, 以更好

地培养学生应用概率统计知识解决复杂统计问题的综合能力.

本书内容分为 8 章, 前 4 章为概率论, 后 4 章为数理统计. 使用者可以根据专业的需要, 选学其中的部分内容或全部内容.

在此我们感谢华中科技大学概率统计系的领导和老师们, 尤其是概率统计课程组的全体教师, 由于他们的关心、支持和鼓励, 我们才能顺利完成此书, 在课程组教学活动中, 老师们的踊跃讨论以及各位老师的专题报告让我们得到很多启发, 受益匪浅. 我们还要感谢科学出版社武汉分社, 特别是杨瑰玉先生对本书的支持和督促, 由于他们的热心指导和出色编辑, 本书才得以顺利问世.

本书前 4 章由李萍编写, 后 4 章由叶鹰编写, 全书由李萍统稿. 由于编者水平有限, 书中难免存在疏漏或不妥之处, 恳请专家和读者批评指正.

编　者

2012 年 4 月 30 日

目　　录

第 1 章　随机事件与概率

概率论与数理统计的研究对象是随机现象. 对随机现象的统计规律性的探讨是从研究随机事件出发而进行的. 本章先介绍有关随机事件的基本概念与运算, 然后介绍概率的定义、性质以及有关概率计算的方法和公式, 以解决随机事件的概率的计算问题.

1.1　随机事件及其运算

1.1.1　随机现象

在自然界和人类社会的各项活动中, 人们所观察到的现象大致可以分为两类: 确定性现象和随机现象.

在一定条件下, 必然会发生的某种结果的现象称为**确定性现象**. 例如, 冬天过去, 春天一定会到来; 在标准大气压下, 纯水加热至 100 ℃必然开始汽化; 在室温下, 生铁肯定不熔化; 在真空中, 光的传播速度为定值.

在一定条件下, 有多种可能结果, 且事前不能预知哪种结果会发生的现象称为**随机现象**. 例如, 抛一枚硬币, 有可能正面朝上, 也可能反面朝上; 某射手打靶射击一次, 可能命中 0 环, 1 环, 2 环, $\cdots$, 10 环; 在前次标普评级中由 BB+ 级降为 BB 级的某国国债, 在下一次标普评级中可能保持 BB 级, 也可能被升级或者继续降级.

一般地, 人们常将对某种现象观察或者实验统称为试验. 在概率论和数理统计中主要研究能被重复观察的随机现象, 将满足如下条件的试验称为**随机试验**.

(1) 试验在相同的条件下可以重复进行.

(2) 每次试验前不可预知该试验的结果, 但可以明确试验的所有可能的结果.

以后提到的试验均指随机试验, 通常记为 E. 以下是一些试验的例子.

E_1: 掷一枚骰子, 观察出现的点数.

E_2: 甲、乙两队按三局定胜负进行排球比赛, 观察甲队在三局中的胜负结果.

E_3: 记录某路口某个月内交通事故发生的次数.

E_4: 记录某地区某年的年降雨量.

E_5: 某办公楼的每日的用电量为 $10^5 \sim 1.5 \times 10^7$kW·h, 用水量为 20~90t, 观察该楼某日的水电用量.

随机试验是对随机现象的观察实验, 一随机试验的可能结果就是该试验所考

察的随机现象的可能结果. 因此, 从这个意义上说, 随机试验完全代表和反映了随机现象. 在用随机试验研究随机现象的实践中, 人们发现随机现象具有**统计规律性**, 这是指在相同的条件下, 进行大量次的重复试验时, 随机现象的可能结果会呈现出某种规律性.

例 1.1(掷币试验) 表 1.1 中记载了几位数学家投掷硬币试验的结果.

表 1.1 历史上掷硬币试验的若干结果

实验者	掷硬币次数 n	出现正面次数 m	出现正面频率 $\frac{m}{n}$
D.Morgan	2 048	1 061	0.518 1
Buffon	4 040	2 048	0.506 9
Feller	10 000	4 979	0.497 9
Pearson	12 000	6 019	0.501 6
Pearson	24 000	12 012	0.500 5

表 1.1 中的观测数据反映了这样的一个事实: 随着试验次数 n 的增大, 比值 $\dfrac{m}{n}$ 有向 0.5 集中的趋势.

例 1.2(高尔顿板) 设板上钉有如图 1.1 所示排列的钉子, 板的底部有与相邻两钉的水平间距同宽的条形格槽. 自入口处放入一直径略小于两钉间距的小球, 当小球自由下落时, 由于碰到下排的钉子而以相等的机会向左或者向右落下, 于是又碰到在下一排的钉子, 如此继续下去, 小球最后落入底部的一个格槽内. 将大量的这种小球不断地从入口放下后, 观察各个格槽中落入小球的堆积情况. 将这一试验反复进行多次, 发现只要每次放入的小球的数目足够大, 它们在格槽中堆积形成的曲线形状几乎总是如图 1.1 中所示的形状.

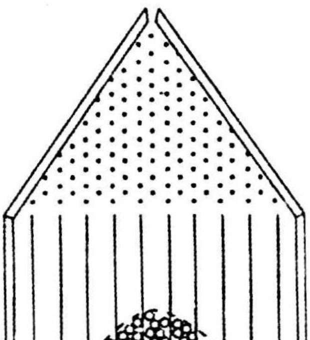

图 1.1 高尔顿板钉

这个试验是由美国生物统计学家 Francis Galton 设计, 称为高尔顿板.

1.1.2 样本空间

随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**, 记为 $\Omega=\{\omega\}$. 样本空间的元素 ω 表示试验 E 的一个可能结果, 称为 E 的**样本点**.

在描述和研究随机现象时, 首先要明确其所有可能结果, 即相应试验的样本空间是怎样的. 例如, 1.1.1 小节中试验 $E_k(k=1,2,\cdots,5)$ 的样本空间 Ω_k 如下:

$\Omega_1=\{1,2,\cdots,6\}$.

$\Omega_2=\{ww,wlw,wll,lww,lwl,ll\}$, 其中, w 表示甲在该局赢, l 表示甲在该局输.

$\Omega_3=\{0,1,2,\cdots\}$.

$\Omega_4=\{x|x\geqslant 0\}$, 其中 x 表示“降雨量为 xmL”.

$\Omega_5=\{(x,y)|10\leqslant x\leqslant 150,20\leqslant y\leqslant 90\}$, 其中 x,y 分别表示用电量和用水量.

需要注意以下两点:

(1) 样本空间可以是数量型, 也可以是非数量型.

(2) 样本点 ω 必须是试验的最基本的结果, 在每次试验中有且仅有一个 $\omega\in\Omega$ 发生.

1.1.3 随机事件

对于随机现象, 人们不只会关心样本空间包含的基本的可能结果, 还会对另外一些可能结果感兴趣. 例如, 在掷骰子试验 E_1 中有如下可能结果:

A: 掷出 6 点.

B: 掷出的点数大于 3.

C: 掷出的点数不大于 5 但也不小于 3.

显然, 结果 A 是 E_1 的基本可能结果 $\omega=6$, 可以用样本空间 Ω_1 的单点子集 $\{6\}$ 表示; 结果 B 要发生必须且仅需基本可能结果“$\omega=4,\omega=5,\omega=6$”之一发生, 故可用 Ω_1 的子集 $\{4,5,6\}$ 完全代表; 而结果 C 发生则当且仅当 Ω_1 的子集 $\{3,4,5\}$ 包含的基本可能之一发生. 一般地, 任何试验的任何可能结果, 必是其样本空间的子集合. 据此, 建立随机事件的概念.

随机试验 E 的样本空间 Ω 的某些元素组成的子集, 称为 E 的**随机事件**, 简称**事件**, 常用大写字母 $A,B,C,\cdots$ 表示.

关于事件这一概念, 要注意以下几点:

(1) 事件可以用集合表示, 也可以用明白无误的语言描述, 在概率论中更侧重于后者.

(2) 由于样本空间 Ω 的单个元素组成的子集称为**基本事件**, 由 Ω 的两个及以上元素组合的子集称为**复合事件**, 而 Ω 的最大子集 (Ω 自身) 称为**必然事件**, Ω 的最小子集 (空集 $\varnothing$) 称为**不可能事件**.