全国高等农林院校“十一五”规划教材

概率论与数理统计

邵崇斌　徐　钊　主编

中国农业出版社

概率论与数理统计

全国高等农林院校“十一五”规划教材

概率论与数理统计

邵崇斌　徐　钊　主编

中国农业出版社

图书在版编目（CIP）数据

概率论与数理统计／邵崇斌，徐钊主编．—北京：中国农业出版社，2007.8
全国高等农林院校“十一五”规划教材
ISBN 978-7-109-11842-3

Ⅰ.概…　Ⅱ.①邵…②徐…　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112999号

中国农业出版社出版
（北京市朝阳区农展馆北路2号）
（邮政编码100026）
责任编辑　朱　雷

北京中兴印刷有限公司印刷　　新华书店北京发行所发行
2007年8月第1版　　2007年8月北京第1次印刷

开本：820mm×1080mm　1/16　印张：23
字数：550千字
定价：33.00元

编　写　人　员

主　编　邵崇斌　徐　钊

副主编　郭满才　李任波　徐雁南　杨俊霞　杨云龙

编写人员（以姓氏笔画为序）

师丽娟（西北农林科技大学）
任争争（西北农林科技大学）
刘　琳（西南林学院）
杜俊莉（西北农林科技大学）
李　青（塔里木大学）
李任波（西南林学院）
李春兰（河北农业大学）
杨云龙（山西农业大学）
杨俊霞（河北农业大学）
邵崇斌（西北农林科技大学）
张远迎（西北农林科技大学）
张雅文（苏州科技学院）
单青松（西北农林科技大学）
徐　钊（西北农林科技大学）
徐雁南（南京林业大学）
郭连红（西北农林科技大学）
郭满才（西北农林科技大学）
魏　宁（西北农林科技大学）
魏杰琼（西北农林科技大学）

前　言

概率论与数理统计是应用数学领域内的一门重要课程,它是由相对独立但又彼此联系的两个部分组成. 概率论研究随机现象的数量规律及其相关理论;数理统计以概率论为基础研究如何有效地收集、整理、分析带有偶然性的数据,从而对所研究的问题做出推断,进而为决策提供科学依据. 随着各学科定量化分析研究的深化,概率论与数理统计的教学愈来愈受到重视,是大学本科生、研究生的许多专业必修的基础课或专业基础课. 特别是随着计算机技术的发展,概率统计与计算机技术的结合,使大量数据的统计处理变得简单易行,是科学研究、试验开发、生产实践处理数据的有力工具,其应用范围越来越广,受到各行各业的重视.

本教材是根据教育部"概率论与数理统计"教学指导委员会课程建设的指导意见和"十一五"规划教材的要求,针对大学本科所设置的生物、农林、医学、经济、管理和其他应用学科专业的特点,结合作者多年从事这方面的教学经验与实践编写而成. 在内容的选取上贯彻少而精的原则,坚持知识的系统性和理论的完整性;既要突出基本概念、原理和方法,又加强了应用领域、应用技巧的介绍和联系实际解决问题的内容. 在结构上保持理论的严谨性和逻辑思维的科学性,也要兼顾通俗性和直观性;既渗透了数学建模的思想,使学生在潜移默化中掌握统计模型的思想和方法,又引入了利用计算机完成实现统计方法的内容,和计算机技术相结合. 在例题和习题的选配上吸收了各行各业的生产与研究中的典型事例和同行专家的科研与教学成果,既注重题目的典型性与广泛性,也适当照顾学生考研究生的需要,以拓宽学生的知识面和实际应用能力. 我们的宗旨是希望把该教材写成学生好读、教师好教的精品教材. 因此形成了自己明显的特点,主要是:

1. 在概率部分增加引入了一些统计的概念和思想,比如提出了条件数学期望和随机变量函数的期望等概念,为统计理论作铺垫.

2. 从理论的完整性考虑,在数理统计中把方差分析和回归分析的理论基础统一在线性模型下,使其在理论上更趋于统一与完整.

3. 结合实际应用的需要,对一些统计方法在应用上的不足作了补充和约定. 比如对于点估计在实际应用中与区间估计有机的结合,归纳出更符合实际合理的方式

方法.

4. 为了培养学生的能力和便于学习、理解和掌握知识,每一章都作了知识系统的小结和典型例题的解析.

本教材是大家分工协作共同努力的结果. 它不仅可以作为教材,也可以作为广大科技工作者的参考书和工具书. 参编的学校有:西北农林科技大学、南京林业大学、河北农业大学、山西农业大学、西南林学院、苏州科技学院和塔里木大学. 由邵崇斌、徐钊任主编,郭满才、李任波、徐雁南、杨俊霞、杨云龙任副主编. 参加编撰的人员还有:刘琳、任争争、师丽娟、李春兰、杜俊莉、单青松、张远迎、张雅文、郭连红、魏杰琼、魏宁和李青. 在编撰的过程中得到了上述学校和中国农业出版社的大力支持和帮助,我们表示衷心的感谢. 书中不当之处,请同行专家及读者不吝指教.

编　者

2007 年 6 月

内容简介

本书是全国高等农林院校“十一五”规划教材，是针对大学本科所设置的生物、农林、医学、经济、管理、地理信息、农业工程、环境类和其他应用学科专业的特点，结合作者多年从事这方面的教学经验与实践编写而成. 它可以作为这些专业的“概率论与数理统计”或“概率论”课程教材和教学参考书. 其涵盖了概率论的基本内容和数理统计中的统计推断（估计与检验）、方差分析与回归分析等基本方法，并介绍了统计方法在Excel下的实现.

本书论述严谨、内容全面、通俗易懂，注重统计思路的阐述及统计方法的实际应用，也可作为相关专业和科技工作者的参考书.

目　　录

前言

第1章　随机事件与概率 …… 1

§1.1　随机试验　随机事件 …… 1

1.1.1　随机现象与随机试验 …… 1

1.1.2　随机事件 …… 2

1.1.3　随机事件的运算 …… 3

1.1.4　随机事件的运算律 …… 5

§1.2　随机事件的概率 …… 6

1.2.1　古典概型中的概率定义 …… 6

1.2.2　几何概型中的概率定义 …… 7

1.2.3　概率的统计定义 …… 8

1.2.4　概率的公理化定义 …… 9

§1.3　概率的性质 …… 10

§1.4　条件概率与事件的独立性 …… 12

1.4.1　条件概率 …… 12

1.4.2　概率乘法公式 …… 13

1.4.3　事件的独立性 …… 14

§1.5　全概率公式和逆概率(Bayes)公式 …… 15

1.5.1　全概率公式 …… 15

1.5.2　逆概率(Bayes)公式 …… 17

§1.6　贝努利概型与二项概率公式 …… 18

小结与例题解析 …… 20

习题1 …… 24

第2章　随机变量(向量)及其概率分布 …… 27

§2.1　随机变量　分布函数 …… 27

§2.2　离散型与连续型随机变量的概率分布 …… 29

2.2.1　离散型随机变量 …… 29

2.2.2　连续型随机变量 …… 31

§2.3　常用的几个随机变量的概率分布 …… 33

2.3.1　离散型随机变量的概率分布 …… 33

2.3.2　连续型随机变量的概率分布 …… 35

§2.4　随机向量及其分布函数　边际分布 …… 41
§2.5　二维离散型与连续型随机向量的概率分布 …… 43
2.5.1　二维离散型随机向量 …… 43
2.5.2　二维连续型随机向量 …… 45
2.5.3　二维均匀分布及二维正态分布 …… 46
§2.6　条件分布　随机变量的独立性 …… 48
2.6.1　条件分布 …… 48
2.6.2　随机变量的独立性 …… 51
§2.7　随机变量函数的概率分布 …… 53
2.7.1　一维随机变量函数的概率分布 …… 53
2.7.2　多维随机向量函数的概率分布 …… 56
小结与例题解析 …… 60
习题2 …… 67

第3章　随机变量的数字特征 …… 71

§3.1　数学期望 …… 71
3.1.1　数学期望的概念 …… 71
3.1.2　常用分布的数学期望 …… 73
3.1.3　随机变量函数的数学期望 …… 75
3.1.4　数学期望的性质 …… 77
3.1.5　条件数学期望 …… 80
3.1.6　中位数 …… 81
§3.2　方差 …… 82
3.2.1　方差的定义 …… 83
3.2.2　常用分布的方差 …… 83
3.2.3　方差的性质 …… 86
3.2.4　随机变量的标准化 …… 87
§3.3　协方差与相关系数 …… 87
3.3.1　协方差与相关系数的定义 …… 87
3.3.2　协方差与相关系数的性质 …… 89
§3.4　矩与协方差阵 …… 92
3.4.1　矩的概念 …… 92
3.4.2　中心矩与原点矩之间的关系 …… 93
3.4.3　偏度和峰度 …… 94
3.4.4　协方差矩阵 …… 94
小结与例题解析 …… 96
习题3 …… 101

第4章　大数定理与中心极限定理 …… 104

§4.1　切比雪夫不等式　随机变量序列的收敛性 …… 105

4.1.1 切比雪夫不等式(Chebychev's Inequality) …… 105
4.1.2 随机变量序列的收敛性 …… 106
§4.2 大数定理 …… 107
§4.3 中心极限定理 …… 109
小结与例题解析 …… 112
习题 4 …… 115

第5章 数理统计的基本概念 …… 118

§5.1 总体 样本 统计量 …… 118
5.1.1 总体与总体特征数 …… 118
5.1.2 样本 …… 120
5.1.3 统计量 …… 122
5.1.4 样本频率分布及直方图 …… 123
§5.2 抽样分布 …… 125
5.2.1 χ^2 分布、t 分布、F 分布 …… 125
5.2.2 抽样分布 …… 131
§5.3 数据简单处理技术 …… 135
5.3.1 Excel 概况 …… 135
5.3.2 常用概率分布的计算 …… 138
5.3.3 样本均值、方差、标准差等的计算 …… 139
5.3.4 频率分布表与直方图 …… 142
小结与例题解析 …… 145
习题 5 …… 148

第6章 参数估计 …… 150

§6.1 参数的点估计 …… 150
6.1.1 矩估计法 …… 150
6.1.2 极大似然估计法 …… 152
6.1.3 估计量的评价 …… 157
6.1.4 点估计的应用 …… 162
§6.2 区间估计 …… 162
6.2.1 区间估计的概念 …… 163
6.2.2 枢轴变量法 …… 163
6.2.3 关于点估计的约定与区间估计 …… 164
§6.3 一个总体均值的估计 …… 165
6.3.1 总体服从正态分布的情况 …… 165
6.3.2 非正态总体情况 …… 167
*6.3.3 大样本方法样本容量的确定 …… 169
*6.3.4 有限总体不重复抽样估计法 …… 170

6.3.5 总体均值 μ 的单侧置信限估计 …… 172
§6.4 一个总体方差与频率的估计 …… 174
6.4.1 正态总体方差 σ^2 的估计 …… 174
6.4.2 正态总体方差 σ^2 的单侧置信限 …… 176
6.4.3 总体频率 p 的估计 …… 177
§6.5 两个总体均值差的估计 …… 181
6.5.1 总体服从正态分布的情况 …… 181
6.5.2 总体分布未知的情况 …… 183
*6.5.3 成对样本的平均数之差 …… 185
§6.6 两个正态总体方差比的估计 …… 186
小结与例题解析 …… 187
习题 6 …… 191

第 7 章 假设检验 …… 195

§7.1 假设检验的概念与原理 …… 195
7.1.1 统计假设 …… 195
7.1.2 小概率原理 …… 196
7.1.3 假设检验的拒绝域和接受域 …… 196
7.1.4 假设检验的两类错误 …… 198
7.1.5 假设检验的一般步骤 …… 198
§7.2 一个总体参数的假设检验 …… 199
7.2.1 正态总体均值的检验 …… 199
7.2.2 正态总体方差的检验 …… 204
7.2.3 非正态总体均值的检验 …… 206
7.2.4 总体频率的检验 …… 208
§7.3 两个总体参数的假设检验 …… 210
7.3.1 两个正态总体均值差异显著性检验 …… 210
7.3.2 两个非正态总体均值差异显著性检验 …… 213
7.3.3 两个正态总体方差的检验 …… 215
*7.3.4 多个正态总体方差齐性检验 …… 218
7.3.5 两个总体频率差异的显著性检验 …… 219
§7.4 适合性检验与独立性检验 …… 221
*7.4.1 拟合优度检验 …… 221
7.4.2 适合性检验 …… 225
7.4.3 独立性检验 …… 226
* §7.5 假设检验的计算机实现技术 …… 228
7.5.1 一个正态总体参数的假设检验 …… 228

：加“”为选学内容.

7.5.2　两个正态总体均值的假设检验 ······ 230
7.5.3　拟合优度检验和独立性检验 ······ 233
*§7.6　检验的功效函数 ······ 236
小结与例题解析 ······ 239
习题7 ······ 242

第8章　方差分析与回归分析 ······ 245

§8.1　方差分析的概念与基本思想 ······ 245
8.1.1　问题的引入 ······ 245
8.1.2　方差分析中的术语 ······ 247
8.1.3　方差分析的基本思想 ······ 248
§8.2　单因素方差分析 ······ 249
8.2.1　单因素等重复试验的方差分析 ······ 249
8.2.2　单因素不等重复试验的方差分析 ······ 258
§8.3　双因素方差分析 ······ 260
8.3.1　双因素无重复试验的方差分析 ······ 260
8.3.2　两因素等重复试验的方差分析 ······ 266
§8.4　回归分析的基本概念 ······ 273
8.4.1　相关关系与回归关系 ······ 273
8.4.2　回归模型与回归方程 ······ 274
8.4.3　一元线性回归模型与回归方程 ······ 275
§8.5　一元线性回归模型的建立与检验 ······ 276
8.5.1　一元线性回归系数 β_0,β_1 的估计 ······ 276
8.5.2　随机误差方差 σ^2 的估计 ······ 279
8.5.3　最小二乘估计量 $\hat{\beta}_0,\hat{\beta}_1$ 的统计性质 ······ 280
8.5.4　一元线性回归的显著性检验 ······ 284
§8.6　预测、控制与残差分析 ······ 289
8.6.1　预测 ······ 289
8.6.2　控制 ······ 292
8.6.3　残差分析 ······ 294
§8.7　可线性化的一元非线性回归 ······ 297
8.7.1　常用的可线性化的非线性回归函数类型 ······ 298
8.7.2　相关指数及可线性化的非线性回归显著性判断 ······ 300
§8.8　多元线性回归 ······ 301
8.8.1　一元线性回归分析的矩阵表示 ······ 301
8.8.2　多元线性回归模型 ······ 303
8.8.3　经验回归方程的建立 ······ 304
8.8.4　多元线性回归系数向量最小二乘估计的统计性质 ······ 306
8.8.5　多元线性回归方程的显著性检验 ······ 307

8.8.6 多元非线性回归 …… 311
小结与例题解析 …… 312
习题 8 …… 323

习题参考答案 …… 326
附表 …… 333
主要参考文献 …… 354

第1章　随机事件与概率

【本章提要】随机事件的概率是概率论研究的基本内容. 本章将介绍随机事件、随机事件的概率，归纳总结概率的基本性质，并通过推导给出计算概率的几个常用公式. 重点要掌握概率的概念、随机事件概率的计算、概率的基本公式及应用. 难点是应用概率的性质和基本公式去解决实际问题.

§1.1　随机试验　随机事件

1.1.1　随机现象与随机试验

在自然界和人类社会中，通常存在着两类不同的现象. 一类是在一定条件下必然会出现惟一结果的现象，我们称之为确定性现象. 这类现象的特点是：条件完全决定结果，告知条件，就可断言其结果. 例如，在标准大气压下，将水加热至100℃，必然会沸腾；矩形面积总是等于底乘高等等. 对于确定性现象的数量规律性，大都是通过几何、代数、微分方程等来研究，物理学和化学中的很多定律也是探讨这类现象的结果. 相对于确定性现象，另一类现象则被称之为非确定性现象. 这类现象由其形成的机理不同又可分为模糊现象和随机现象. 模糊现象是指客观事物在差异的中介过渡状态所呈现的“亦此亦彼”现象，它是由描述事物的概念界限不确定而引起的结果的不确定性. 例如，稠密的森林，高大的山脉等等. 随机现象是指在一定条件下，出现的可能结果不惟一，至少有两个. 可能出现这样的结果，也可能出现那样的结果(条件不完全决定结果)，结果的出现具有一定的偶然性，每次观察前，无法预言其具体结果. 例如，掷一枚硬币，观察出现正反面的情况，结果可能是正面(有图画面)朝上，也可能是反面朝上；取50粒种子做发芽试验，观察发芽的种子粒数，结果可能是0粒，1粒，…，50粒种子发芽；从装有6个白球，4个黑球的袋中任抽一球观察其颜色，结果可能是白球，也可能是黑球等等. 概率论是研究这类现象的学科.

随机现象的观察结果至少有两个，在相同条件下多次观察同一随机现象，尽管其结果不尽相同，但在观察次数足够大时，各种可能结果的出现都具有某种规律性. 为便于深入研究随机现象的这种规律性. 我们约定：对随机现象，在相同条件下可重复进行的观察或试验称为随机试验(random experiment)，简称试验，一般用 E 表示.

【例1.1】抛掷一枚硬币，观察正面朝上的情况.

【例1.2】在分别写有数字1，2，…，10的10张卡片中随意抽取一张卡片，观察其数字.

【例1.3】投掷两枚骰子，观察朝上面的点数.

【例1.4】从一批灯泡中，任抽一只，观察其使用寿命.

从以上4个例子中，可以看出随机试验具有以下特点：

①在相同条件下可重复进行试验；

②试验前由试验条件能明确试验所有可能结果，且所有可能结果至少有两个；

③每次试验前不能预知哪个具体结果出现.

随机试验的这些特点决定了一次试验的结果可能这样也可能那样，具有一定的偶然性，但每次试验的结果绝不可能跳出该随机试验所有可能结果所构成的集合.

随机试验可能出现的每一最基本的结果称为该试验的一个样本点(sample point)，一般用 ω 表示，样本点的全体构成的集合称为该试验的样本空间(sample space)，用 Ω 表示. 例如：在例1.1中 $\Omega=\{$正面朝上，反面朝上$\}$；在例1.2中 $\Omega=\{1, 2, \cdots, 10\}$.

在随机试验中，对同一试验 E，由于试验目的不同，相应地样本空间也常常不同. 如投掷一枚均匀的骰子，若试验目的是观察骰子朝上面的点数，样本空间 $\Omega=\{1, 2, 3, 4, 5, 6\}$. 但若试验目的是观察是否出现偶数点，则 $\Omega=\{$奇数，偶数$\}$. 尽管随机试验在一次或少数几次试验中会出现这样或那样的结果，呈现出偶然性的一面，但若将一个随机试验重复大量次数，就会发现其结果的出现存在一定规律性，称这种规律性为随机现象的统计规律性. 这是随机现象所涉及事物本质特征的一种反映，例如，将抛掷一枚均匀硬币的随机试验重复大量次数，就会发现正面朝上的次数约占试验总次数的1/2. 这恰好说明了硬币的均匀对称性. 揭示随机现象的统计规律性是概率论研究的主要任务.

1.1.2 随机事件

随机试验的可能结果一般称为随机事件(random event)，简称事件，常用英文字母 A，B，C，…表示. 例如，在例1.1中，“正面朝上”就是一个随机事件；例1.2中“出现的数字是3”，“出现的数字是偶数”都是随机事件.

随机事件可分为基本事件和复合事件. 我们把试验最直接的可能结果称为基本事件. 例如，例1.1随机试验中“正面朝上”“反面朝上”就是该试验的基本事件. 而由若干个基本事件共同在一起才能表达的试验结果，称为复合事件. 它由若干个基本事件组合而成. 例如，例1.2随机试验中，“所抽得号码能被4整除”是一个由基本事件“抽得号码为4”、“抽得号码为8”所构成的复合事件. 事实上，对随机事件的理解可以从两个角度去把握. 从本质上来讲，随机事件就是关于随机试验结果的命题；从集合的角度来讲，随机事件是随机试验样本空间的子集，是由一部分样本点构成的集合，某事件发生而且仅当属于它的某一个样本点出现.

在每次试验中必然发生的事件，称之为必然事件，一般用 Ω 表示. 例如，如在例1.2的随机试验中，“抽得的号码不大于10”就是一个必然事件. 在每次试验中必然不会发生的事件称为不可能事件，一般用 $\varnothing$ 表示. 例如，在例1.4的随机试验中，“抽得灯泡的使用寿命是负数”就是一个不可能事件. 这两个事件显然具有确定性，严格讲不是随机事件，但是为了今后讨论问题的方便，我们把它们作为特殊的随机事件看待.

对于一个随机试验 E，样本空间 Ω 就是一个以随机试验的样本点为元素的全集. 而任何随机事件均为样本空间的一个子集，特别是必然事件作为随机事件的特例，用集合表示就是样本空间 Ω；不可能事件作为随机事件的特例，用集合表示就是空集 $\varnothing$.

1.1.3　随机事件的运算

在研究随机事件的有关问题时，有些事件比较简单，有些比较复杂，但它们之间是有联系的，人们总希望用较简单的随机事件表达较复杂的随机事件，以便推导出复杂事件的某些规律性和结论，这就需要研究随机事件的关系和随机事件的运算.

1. 事件的包含关系

设 A, B 是随机试验 E 的事件，若事件 A 发生必然导致事件 B 发生，则称事件 A 包含于事件 B 或事件 B 包含事件 A. 记作 $A \subset B$.

如在例 1.2 中，令 A 表示“抽到能被 4 整除的号码”，B 表示“抽到偶数号码”，显然，$A \subset B$. 事实上，$A = \{$抽到 4，8 号码$\}$，$B = \{$抽到 2，4，6，8，10 号码$\}$.

2. 事件的相等(等价)

设 A, B 是随机试验 E 的两个事件，若 $A \subset B$ 且 $B \subset A$，则称事件 A 与事件 B 相等或等价，记作 $A = B$.

3. 事件的和(并)

设 A, B 是随机试验 E 的任意两个事件，则称“事件 A 与事件 B，至少有一个发生”这一事件为事件 A 与事件 B 的和事件，记作 $A + B$(或 $A \cup B$).

【例 1.5】 接连射击两次，观察各次中靶与否. 设事件 $A = \{$第一次命中$\}$，$B = \{$第二次命中$\}$，则和事件 $A + B = A \cup B = \{$至少命中一次$\}$.

两个事件和的概念可以推广到任意有限多个事件，甚至无穷可列个事件上. 一般地，称“n 个事件 A_1，A_2，…，A_n 至少有一个发生”这一事件为 n 个事件 A_1，A_2，…，A_n 的和事件，记作 $\sum\limits_{i=1}^{n} A_i$ 或 $\bigcup\limits_{i=1}^{n} A_i$. 称“无穷可列个事件 A_1，A_2，…，A_n，…至少有一个发生”这一事件为无穷可列个事件的和事件，记作 $\sum\limits_{i=1}^{\infty} A_i$ 或 $\bigcup\limits_{i=1}^{\infty} A_i$.

4. 事件的积

设 A, B 为随机试验 E 的两个事件，则称“事件 A 与事件 B 同时发生”这一事件为事件 A 与事件 B 的积事件，记作 AB 或 $A \cap B$.

如例 1.5 中，事件 $AB = \{$两次都命中$\}$. 两个事件积的概念可以推广到有限多个事件甚至无穷可列事件上. 一般地，称“n 个事件 A_1，A_2，…，A_n”同时发生这一事件为“A_1，A_2，…，A_n 的积事件”，记作 $\prod\limits_{i=1}^{n} A_i$ 或 $\bigcap\limits_{i=1}^{n} A_i$. 称“无穷可列个事件 A_1，A_2，…，A_n，…都发生”这一事件为无穷可列个事件的积事件，一般用 $\prod\limits_{i=1}^{\infty} A_i$ 或 $\bigcap\limits_{i=1}^{\infty} A_i$ 表示.

5. 事件的差

设 A, B 是随机试验 E 的任意两个事件，称“事件 A 发生而事件 B 不发生”这一事件为事件 A 与事件 B 的差事件，记作 $A - B$.

如在例 1.5 的随机试验中，事件 $A - B = \{$第一次命中而第二次未命中$\}$.