



面向 21 世纪 课 程 教 材
Textbook Series for 21st Century

大学数学

概率论及试验统计 第三版

余家林 朱倩军 主编



高 等 教 育 出 版 社
HIGHER EDUCATION PRESS

面向 21 世纪课程教材
Textbook Series for 21st Century

大学数学

概率论及试验统计

第三版

余家林 朱倩军 主编



高等教育出版社

HIGHER EDUCATION PRESS

内容提要

本书第三版为普通高等教育“十一五”国家级规划教材,第二版是普通高等教育“十五”国家级规划教材,并在2005年获得高等教育国家级教学成果二等奖。内容包括随机事件的概率、随机变量的分布、随机变量的函数、随机变量的数字特征、样本及统计量、总体分布中未知参数的估计、总体分布参数及总体分布的假设检验、方差分析、回归分析与协方差分析等基础理论知识及常用的试验统计分析方法,并且提供了相应的SAS程序。附录有试验统计用表、CASIO fx-3600P型计算器简介、Excel数据分析功能的使用简介和SAS中的英文关键字简介。

本书可作为高等农林类学校各专业概率统计课程的教材,也可作为考研复习的参考用书。

图书在版编目(CIP)数据

大学数学—概率论及试验统计/余家林,朱倩军主编.

3版. —北京:高等教育出版社,2009.2

ISBN 978-7-04-025537-9

I.大… II.①余… ②朱… III.①高等数学—高等学校—教材②概率论—高等学校—教材③试验统计—高等学校—教材 IV.013.021

中国版本图书馆CIP数据核字(2009)第005767号

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街4号	免费咨询	800-810-0598
邮政编码	100120	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
经 销	蓝色畅想图书发行有限公司	网上订购	http://www.landaco.com
印 刷	北京民族印务有限责任公司		http://www.landaco.com.cn
		畅想教育	http://www.widedu.com
开 本	787×960 1/16	版 次	2001年12月第1版
印 张	18.75		2009年2月第3版
字 数	340 000	印 次	2009年2月第1次印刷
		定 价	20.50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。

版权所有 侵权必究

物料号 25537-00



面向21世纪课程教材



普通高等教育“十一五”
国家级规划教材



普通高等教育“十五”
国家级规划教材

第三版前言

本书第二版自2004年出版以来,被很多高等农林院校选作“概率统计”课程的教材或考研复习的参考用书。2005年本书作为《农林类本科大学数学序列课程教材建设》的成果之一,获得湖北省高等学校教学成果一等奖和国家级教学成果二等奖。本书选材适当、逻辑严谨,科学性强;叙述清晰、文字简练,可读性强。

第三版是对第二版进行修订与充实的结果,特别是吸纳了各校师生的意见,有以下六项较大的改动:(1)为适应工程类和经济类学生参加统一命题研究生入学考试的要求,将教材中分布函数的定义由 $P\{X < x\}$ 改为 $P\{X \leq x\}$;(2)精选“选读选讲”的内容,既为学有余力的学生提供“学数学”与“用数学”平台,也为课时充裕的老师引导学生“学数学”与“用数学”准备一些素材,这些内容只供选读与选讲;(3)强调用计算器作统计计算与使用统计分析软件上机计算相同步,为减少用计算器作统计计算的误差,特别在假设检验、方差分析、回归分析的例题中增加了用计算器作连运算的说明;(4)考虑到正态性检验的重要性,编入了正态性检验的偏态峰态检验法和W法及相应的SAS程序;(5)考虑到对多个总体进行方差的同质性检验的重要性,不仅编入了Bartlett的 χ^2 检验,还编入Levene的F检验,前者比较常用,后者的适应性较强;(6)Microsoft Office中的Excel是一件很方便实用的工具,这一版的附录中增加了“Excel数据分析功能的使用简介”。

教育部高等学校统计学专业教学指导分委员会委员、武汉大学刘禄勤教授曾经评介本书,北京大学陈家鼎教授对本书的修订提出了指导性的意见,特在此向他们二位致以谢意。

本书由华中农业大学、湖南农业大学、河南农业大学、湖北农学院及贵州大学农学院合编。第三版继续由余家林、朱倩军任主编。肖枝洪、覃光莲、汪晓银、郭明月、陈晓坤、李新建、朱强参加了编写及修订工作,文凤春、陈华锋提出了宝贵的建议。

在第三版出版之际,编者向关心、爱护本书的师生致以谢意。由于编者的水平所限,第三版的不妥之处仍难以避免,敬请读者和使用本教材的同行学友继续批评指正。

编 者

2008年于华中农业大学

第二版前言

本书是普通高等教育“十五”国家级规划教材、全国教育科学“十五”国家规划课题之子课题“21 世纪中国高等学校农林类专业数理化基础课程的创新与实践”和高等教育出版社“高等教育百门精品课程建设计划”立项研究项目的研究成果。

本书第一版自 2001 年出版以来,很多高等学校将其选作农林、水产、食品、生物、医药、经贸及社会学等专业的教材或考研复习的参考用书。第二版吸纳了各校师生的意见,对第一版进行了修订与充实,特别是删除多处过于详尽的说明,以简单明了的提示取而代之,给教师组织教学、提高教学质量留下更多的余地,给学生学数学、用数学留有更大的空间。

本书第二版继续由华中农业大学、湖南农业大学、河南农业大学、湖北农学院及贵州大学合编。余家林、朱倩军任主编,周建军、曹殿立、刘兴卫、罗文俊任副主编,董锐、肖枝洪参加了编写及修订工作。

在第二版出版之际,编者向关心、爱护本书的师生致以谢意。由于编者的水平所限,第二版的不妥之处仍难以避免,敬请读者和使用本教材的同行学友继续批评指正。

编 者

2003 年于华中农业大学

第一版前言

一、教材编写工作的回顾

1995年原国家教委适时推出了“高等教育面向21世纪教学内容和课程体系改革计划”。1996年,“高等农林院校本科数学(含生物统计)系列课程教学内容和课程体系改革的研究与实践”课题经原国家教委高教司[1996]61号文件批准立项,研究立即展开。华中农业大学制定了相应的改革方案,并在1997年至1999年按方案进行了三轮试验。当时的目标是理顺概率论、数理统计与生物统计、试验设计这三门课程的关系,并设想将各专业生物统计课中带有共性的一部分内容与概率论、数理统计合在一起,既加强生物统计的基础理论,又使概率论、数理统计与应用相融合,同时介绍SAS程序,取名为概率论及试验统计,课堂教学及上机实习共60课时。还设想将各专业生物统计课中不带共性的特色内容及试验设计内容,命名为试验设计及分析,单独开设,并增加一些新的内容,由原来的20课时增加到40课时。这样一来,两门课的总课时便可减少,内涵都有所增加。当时由余家林按上述设想编写校内教材,经过华中农业大学多位教师在三个年级、六个专业近五百名学生中试用,并经过多次修改形成本书的雏形。2000年底,华中农业大学、湖南农业大学、河南农业大学、湖北农学院及贵州大学的代表举行农林教育面向21世纪教材研讨会,决定集中五校的学识与经验,合编一本适合农业大学各专业使用的教材。

编写本教材的思路是:(1)加强概率论及试验统计理论部分的教学,为后续课程的教学及学生从事专业工作奠定坚实的基础,为学生未来的深造作准备;(2)注意概率论、数理统计与应用相融合,以增强学生用数学的意识,培养学生用数学的能力;(3)全文采用第三人称写法,引导学生理解新的概念,发现新的规律,提出新的方法,与学生一起重走概率论及试验统计理论和方法的创新之路;(4)给任课教师课前编写教案、课内因材施教留有余地;(5)对试验统计、生物统计、经济统计、社会统计源自理工科的教材内容体系适当地进行重组,使例题与习题更适合农科学生要求。

本教材由华中农业大学余家林任主编,湖南农业大学周建军、河南农业大学曹殿立、华中农业大学朱倩军、湖北农学院刘兴卫及贵州大学罗文俊任副主编,华中农业大学董锐、肖枝洪参加了编写、审定及试教工作。

本教材由中国农业大学张嘉林教授主审。张教授具有渊博的学识与丰富的

教学经验。上个世纪 80 年代,他参与编写了我国农业大学通用的系列数学教材,我们从中得到许多借鉴。这次又担任本教材的主审,提出许多指导性的意见,为本教材增色不少,特在此向他致谢。

在本教材出版的时候,我们还要向教育部致谢,是教育部推出的“高等教育面向 21 世纪教学内容和课程体系改革计划”为我们提供了一个参与的机会。向支持本教材出版的教育部高教司、高等教育面向 21 世纪教学内容和课程体系改革计划工作协调指导小组、04-6 项目组、华中农业大学、湖南农业大学、河南农业大学、湖北农学院及贵州大学等五校的有关部门、特别是数学教研室的老师们致谢。向华中农业大学谢季坚教授致谢,他是这一套大学数学系列教材的发起人、多校联合编写教材的组织与推动者。

由于编者的水平所限,教材中的不妥之处难以避免,敬请读者和使用本教材的教师批评指正。

二、学科及课程简介

概率论及试验统计是数学的一个分支,研究随机试验与随机事件的数量规律。

概率论起源于 17 世纪中叶,最初是为了解答博弈问题,直到 20 世纪方才建立严密的学科体系。目前,概率论不仅在工业工程、农业工程等自然科学学科中有广泛的应用,在管理科学、社会科学学科中也有广泛的应用。

试验统计的起源更早一些,大概在人类学会数数的同时,也就学会了统计。一开始,试验统计仅仅局限于数据的收集与比较,以后便有了比例、百分率、各种统计图表。随着概率论的出现及其发展,试验统计学的理论与方法也逐步地创新,有了各行各业的统计学,有了更加专门化的分支,例如时间序列分析、可靠性理论等等。统计分析所得到的信息,已经成为工程师、农艺师、银行家、军事指挥人员、部门首长进行推断与决策的根据。

现在,概率论及试验统计作为一门数学课程,已经列入大专院校许多专业的教学计划。某些专业的硕士研究生入学考试、留学国外的资格考试试卷中也有一定比例的概率论及试验统计考题。有些专业的硕士研究生对这门课程还要深入地研读。

本课程讲授概率论的基础理论知识及常用的试验统计分析方法,使理论与应用融合,为工业工程、农业工程、经济管理、社会调查等学科的研究与应用奠定基础。

编 者

2001 年于华中农业大学

目 录

第一章 随机事件的概率	1
§ 1.1 随机试验与随机事件	1
1. 随机试验	1
2. 随机试验的样本空间	2
3. 随机事件	2
4. 随机事件的关系及运算	3
5. 事件的复合	7
6. De-Morgan 对偶定律	8
习题 1.1	9
§ 1.2 随机事件的概率	9
1. 可能性大小的度量	9
2. 经验概率	10
3. 古典概率	11
4. 选读选讲内容	13
5. 几何概率	14
6. 概率的公理化定义	15
习题 1.2	16
§ 1.3 概率的计算公式	16
1. 加法公式	16
2. 条件概率与乘法公式	20
3. 全概率公式	21
4. Bayes 公式	22
5. 选读选讲内容	24
习题 1.3	25
§ 1.4 事件的相互独立	26
1. 两个事件相互独立	26
2. 多个事件相互独立	27
3. 二项概率公式	30
4. Poisson 公式	31

5. 选读选讲内容	32
习题 1.4	34
第二章 随机变量的分布	36
§ 2.1 一维离散型随机变量的分布律	36
1. 随机变量	36
2. 一维离散型随机变量的分布律	37
3. 一维离散型随机变量常用的分布	39
4. 一维随机变量的分布函数	42
5. 选读选讲内容	43
习题 2.1	46
§ 2.2 一维连续型随机变量的分布密度	47
1. 一维连续型随机变量的分布密度	47
2. 一维连续型随机变量常用的分布	48
3. 选读选讲内容	55
习题 2.2	57
§ 2.3 二维随机变量的分布	58
1. 二维离散型随机变量的分布律	58
2. 二维离散型随机变量常用的分布	61
3. 二维随机变量的分布函数	61
4. 二维连续型随机变量的分布密度	63
5. 二维连续型随机变量常用的分布	65
习题 2.3	66
§ 2.4 随机变量相互独立	66
1. 二维离散型随机变量的边缘分布律	66
2. 二维连续型随机变量的边缘分布密度	69
3. 两个随机变量相互独立	72
4. 多个随机变量相互独立	74
5. 两组随机变量相互独立	75
6. 选读选讲内容	76
习题 2.4	79
第三章 随机变量的函数	80
§ 3.1 离散型随机变量的函数	80
1. 函数概念的引入	80
2. 一维离散型随机变量的函数	80
3. 二维离散型随机变量的函数	81

4. 重要的结论	82
5. $\min(X, Y)$ 及 $\max(X, Y)$ 的分布律	84
6. 选读选讲内容	85
习题 3.1	87
§ 3.2 连续型随机变量的函数	88
1. 一维连续型随机变量的函数	88
2. 二维连续型随机变量的函数	89
3. 正态随机变量的线性函数的分布	91
4. 正态随机变量的二次函数的分布	93
5. $\min(X, Y)$ 及 $\max(X, Y)$ 的分布密度	94
6. 选读选讲内容	96
习题 3.2	97
第四章 随机变量的数字特征	99
§ 4.1 数学期望与方差	99
1. 数字特征的意义	99
2. 离散型随机变量的数学期望	100
3. 连续型随机变量的数学期望	101
4. 一维随机变量的方差	101
5. 数学期望与方差的性质	104
6. 常用分布的数学期望与方差	105
7. 选读选讲内容	108
习题 4.1	110
§ 4.2 协方差及相关系数	111
1. 二维随机变量的协方差	111
2. 二维随机变量的相关系数	111
3. 不相关与相互独立	113
4. 选读选讲内容	116
习题 4.2	118
§ 4.3 大数定律与中心极限定理	119
1. Chebyshev 不等式	119
2. Chebyshev 大数定律	120
3. Bernoulli 大数定律	121
4. 中心极限定理	121
习题 4.3	123
第五章 样本及统计量	125

§ 5.1 总体与样本	125
1. 总体、个体与总体容量	125
2. 样本、样本容量与简单随机样本	125
3. 样本的联合分布	126
4. 样本观测值的分布函数	127
5. 样本观测值的频率分布直方图	129
6. 统计分析系统(SAS)简介	131
习题 5.1	132
§ 5.2 样本的数字特征	133
1. 样本总和及均值、离均差平方和	133
2. 样本方差、标准差和变异系数	134
3. 样本常用的简易数字特征	135
4. 计算数字特征的 SAS 程序	137
习题 5.2	138
§ 5.3 χ^2 分布、 t 分布及 F 分布	139
1. χ^2 分布	139
2. t 分布	140
3. F 分布	140
4. t 分布与 F 分布的关系	141
5. 常用分布的分位数	142
习题 5.3	145
§ 5.4 常用的统计量及其分布	145
1. 统计量的定义	145
2. 一个正态总体的常用统计量及其分布	145
3. 两个正态总体的常用统计量及其分布	149
4. 非正态总体的均值的分布	150
5. 顺序统计量及其分布	152
习题 5.4	153
第六章 总体分布中未知参数的估计	155
§ 6.1 未知参数的点估计	155
1. 参数估计的基本概念	155
2. 用矩法求估计量	155
3. 用极大似然法求估计量	156
4. 评价估计量优劣的标准	159
习题 6.1	160

§ 6.2 未知参数的区间估计	161
1. 区间估计的基本概念	161
2. 一个正态总体均值或方差的置信区间	162
3. 选读选讲内容(置信区间的论述)	163
4. 两个正态总体均值差或方差比的置信区间	165
5. 选读选讲内容(置信区间的论述)	166
6. 0-1 分布参数 p 的置信区间	168
习题 6.2	170
第七章 总体分布参数及总体分布的假设检验	172
§ 7.1 总体分布参数的假设检验	172
1. 假设检验的基本概念	172
2. 一个正态总体均值或方差的假设检验	174
3. 两个正态总体均值或方差的假设检验	177
4. 0-1 分布参数 p 的假设检验	182
5. 多个总体同方差的假设检验(一)	185
6. 总体分布参数假设检验的 SAS 程序	186
习题 7.1	186
§ 7.2 总体分布的假设检验	187
1. 总体分布的 χ^2 检验	187
2. 总体分布作 χ^2 检验的 SAS 程序	190
3. 正态性检验	190
4. 正态性检验的 SAS 程序	193
5. 列联表分类标志的独立性检验	193
6. 独立性检验的 SAS 程序	195
习题 7.2	195
第八章 方差分析	197
§ 8.1 单因素试验的方差分析	197
1. 单因素试验及有关的基本概念	197
2. 总离均差平方和的分解	198
3. 总体中未知参数的估计	199
4. 单因素试验的方差分析的步骤	199
5. 组间平均数的多重比较	203
6. 单因素试验方差分析的 SAS 程序	205
7. 多个总体同方差的假设检验(二)	206
习题 8.1	208

§ 8.2 双因素试验的方差分析(一)	209
1. 双因素试验及有关的基本概念	209
2. 总离均差平方和的分解	210
3. 总体中未知参数的估计	211
4. 双因素试验不考虑交互作用的方差分析的步骤	211
5. 双因素试验方差分析(一)的 SAS 程序	214
习题 8.2	214
§ 8.3 双因素试验的方差分析(二)	215
1. 考虑交互作用的双因素试验	215
2. 总离均差平方和的分解	216
3. 总体中未知参数的估计	217
4. 双因素试验考虑交互作用的方差分析的步骤	218
5. 双因素试验方差分析(二)的 SAS 程序	221
习题 8.3	222
第九章 回归分析与协方差分析	223
§ 9.1 一元线性回归	223
1. 一元线性回归的基本概念	223
2. 总体中未知参数的估计	224
3. 线性回归方程的显著性检验	225
4. 利用回归方程进行点预测和区间预测	227
5. 相关系数及其显著性检验	230
6. 线性回归的 SAS 程序	231
习题 9.1	231
§ 9.2 一元非线性回归	232
1. 一元非线性回归简介	232
2. 建立非线性回归方程常用的方法	233
3. 非线性回归方程拟合情况的比较	233
4. 一元非线性回归应用的实例	235
5. 非线性回归的 SAS 程序	236
习题 9.2	236
§ 9.3 统计控制与协方差分析	237
1. 统计控制的基本概念	237
2. 单因素试验的协方差分析	238
3. 双因素试验不考虑交互作用的协方差分析	240
4. 双因素试验考虑交互作用的协方差分析	244

5. 协方差分析的 SAS 程序	249
习题 9.3	250
习题答案	251
附录一 Poisson 分布的数值表(部分)	260
附录二 标准正态分布的分布函数值表	261
附录三 χ^2 分布的分位数表	262
附录四 t 分布的分位数表	263
附录五 F 分布的分位数表	264
附录六 二项分布参数 p 的置信区间表(部分)	268
附录七 新复极差检验的 SSR 数值表(部分)	270
附录八 Q 检验的 Q 数值表(部分)	271
附录九 相关系数检验用表(部分)	272
附录十 CASIO fx-3600P 型计算器简介	273
附录十一 Excel 数据分析功能的使用简介	275
附录十二 SAS 中的英文关键字简介	277
参考文献	280

第一章 随机事件的概率

§ 1.1 随机试验与随机事件

1. 随机试验

概率论及试验统计研究随机试验与随机事件的数量规律,是数学的分支之一.

在概率论及试验统计中,试验是一个含义较为广泛的概念.它包括在某些条件下所作的试验,包括对自然界和人类社会的某个事物及其发展与变化所作的观察或者测量.

不难发现,有一些试验在相同的条件下重复地进行多次,其结果是必然会出现的,另有一些则是不可能出现的.例如,在 $101\,325\text{ Pa}$ 的大气压下将水加热到 100°C 是一个试验,结果水必然沸腾.又如,一个人不借助于任何工具,蹦一蹦,结果就飞上月球,这是不可能的.还有一些试验在相同的条件下重复地进行多次,其结果既不是必然会出现的、也不是不可能出现的,在试验尚未结束之前,不能预知试验的结果,称为**随机试验**.在概率论及试验统计中,以随机试验为研究对象.

现将一些有代表性的随机试验举例如下:

例 1.1 上抛硬币 将一枚硬币垂直上抛,到达最高点后,垂直下落在水平的桌面上,其结果可能是有币值的那一面(正面)朝上,也可能是有图案的那一面(反面)朝上.在硬币还没有落定之前,不能预知试验的结果.因此,上抛硬币是随机试验.

例 1.2 丢掷骰子 将一粒骰子用两手捂住,上下摇动后,丢掷在水平的桌面上,其结果可能是 6 种不同点数中的某一点数朝上.如果骰子还没有停止运动,试验的结果便不能事先知道.因此,丢掷骰子是随机试验.

例 1.3 测量体重 从鸡舍中任取一只鸡测量其体重 w (单位: kg), 则 w 在某个区间内取值.如果体重的最大值小于 3 kg , 那么 w 可以是 0 至 3 之间的某一实数.在任取一只鸡并测量出体重之前,不能预知 w 的数值.因此,测量鸡的体重是随机试验.

这三例说明,随机试验具有以下两个特征:

- (1) 试验可以在相同的条件下重复地进行多次,结果不一定相同;
- (2) 试验前能知道一切可能出现的试验结果,却不能预知哪一个结果会出现.

与例 1.1 类似的试验:(1) 一粒种子作发芽试验,结果是发芽或不发芽;
(2) 某女子乒乓球队与水平相当对手举行团体赛,结果是胜利或失败.

与例 1.2 类似的试验:(1) 一粒种子作发芽试验,结果是发芽期为 8 天、9 天或 10 天;(2) 某女子乒乓球队与水平相当对手举行团体赛,采用三局二胜制,结果是 2:1,2:0,0:2 或 1:2.

与例 1.3 类似的试验:(1) 一粒种子作发芽试验,结果是 15 天的芽高 h (单位:cm),它在某个区间内取某一实数;(2) 某女子乒乓球队与水平相当对手举行团体赛,结果是比赛所用的时间 t (单位:h),它在某个区间内取某一实数.

2. 随机试验的样本空间

在某一随机试验中由可能出现的全部试验结果所组成的集合,称为这个随机试验的**样本空间**,记作 Ω .但是,为研究方便起见,常常限定 Ω 中的试验结果,使它们在一次试验中必然有一个要出现并且只可能有一个出现.或者说,在一次试验中, Ω 中的试验结果不可能都不出现,也不可能有两个以上的试验结果同时出现.

根据这样的限定,在例 1.1 中,若正面朝上这一试验结果记作 H ,反面朝上这一试验结果记作 T ,那么上抛一枚硬币的样本空间 $\Omega = \{H, T\}$.上抛两枚或多枚硬币也是随机试验.如果将两枚硬币区分为甲与乙,用 $(\ast\ast)$ 表示上抛两枚硬币试验的结果,括号中的 $\ast$ 将以 H 或 T 代替,第一个 $\ast$ 表示上抛硬币甲的结果,第二个 $\ast$ 表示上抛硬币乙的结果,那么上抛两枚硬币的样本空间 $\Omega = \{(HH), (HT), (TH), (TT)\}$.在例 1.2 中,若将 i 点 ($i=1,2,3,4,5,6$) 朝上这一试验结果也记作 i ,那么朝上的点数便可能是 1,2,3,4,5,6,丢掷一粒骰子的样本空间 $\Omega = \{i | i=1,2,3,4,5,6\}$ 或 $\Omega = \{1,2,3,4,5,6\}$.丢掷两粒或多粒骰子也是随机试验.丢掷两粒骰子的样本空间 Ω 有 36 个元素,记 $\Omega = \{(i,j) | i,j=1,2,3,4,5,6\}$,其中 (i,j) 表示两粒骰子中某一粒骰子朝上的点数为 i ,另一粒骰子朝上的点数为 j .在例 1.3 中,任取一只鸡测量体重的样本空间 $\Omega = \{w | 0 < w < 3\}$.

正确地确定随机试验的样本空间,是研究随机试验的一个关键.

3. 随机事件

在随机试验中,可能出现、也可能不出现的某一件事情,称为**随机事件**,简称