

新世纪大学数学立体化系列教材

总主编 于义良

概率论 与数理统计

主 编 安建业 张银生

 中国人民大学出版社

新世纪大学数学立体化系列教材

总主编 于义良

概率论与数理统计

主 编 安建业 张银生

副主编 李秉林 鄢 茵

中国人民大学出版社

图书在版编目 (CIP) 数据

概率论与数理统计/安建业, 张银生主编.

北京: 中国人民大学出版社, 2004

(新世纪大学数学立体化系列教材)

ISBN 7-300-05484-6

I. 概…

II. ①安…②张…

III. ①概率论-高等学校-教材②数理统计-高等学校-教材

IV. 021

中国版本图书馆 CIP 数据核字 (2004) 第 033757 号

新世纪大学数学立体化系列教材

总主编 于义良

概率论与数理统计

主 编 安建业 张银生

副主编 李秉林 鄢茵

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室) 010-62511239 (出版部)

010-82501766 (邮购部) 010-62514148 (门市部)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com>(人大教研网)

经 销 新华书店

印 刷 北京东君印刷有限公司

开 本 787×965 毫米 1/16

版 次 2004 年 5 月第 1 版

印 张 14.75

印 次 2005 年 1 月第 2 次印刷

字 数 264 000

定 价 26.00 元 (含光盘)

版权所有 侵权必究

印装差错 负责调换

总 序

随着科学技术的迅猛发展,数量分析已渗透到各个领域,数学的重要性已被整个社会所公认;由于计算机技术的广泛普及与提高,许多繁难的计算和抽象的推理已不再是高不可攀,数学的应用越来越深入;随着人类素质的不断提高,数学素质教育已成为全体公民的必修课,数学的普及越来越广泛。为适应新形势的发展和社会的需要,信息技术与学科课程整合已是教育教学改革的“重中之重”,运用信息技术改造和优化传统学科内容是培养新世纪具有创新能力的高素质人才的必然要求。经过多年的教学研究和实践,我们组织了具有丰富教学经验的第一线教师,编写了这套“新世纪大学数学立体化系列教材”,奉献给大家。

这套系列教材是“21世纪初天津市普通高校教学改革项目《信息技术与经济数学课程整合的研究和实践》”成果的延伸,包括《微积分》、《线性代数》、《概率论与数理统计》、《数学建模》、《运筹学》五本教材和《微积分名师导学》、《线性代数名师导学》、《概率论与数理统计名师导学》三本教学指导书。这套教材力求体现如下特点:

第一,以实用为原则,“教、学、做”融为一体,内容体系整体优化,使读者实现由知识向能力的转化。

第二,以实际为背景,概念阐述简明、通俗化,举例贴近生活,运用多媒体技术使内容直观化、图形化,使读者消除对数学的陌生感、抽象感、恐惧感,激活求知欲,增强学好数学、做好数学的信心。

第三,以计算机为工具,传统内容与信息技术应用融为一体,注重基本知识、基本思想、基本能力的培养,对繁、难、抽象的内容,充分利用当前极为流行的 Mathematica 软件、Excel 软件来实现,比如函数图形描绘、矩阵计算、数据分析等。

第四,每册教材均配有多媒体助学助教光盘,包括课程要求、电子教案、模拟演示、练习详解、单元测试、实例选编、试题分析、名人简介等众多模块,信息量大,使用方便,便于读者更好地理解、掌握、巩固所学知识,并有助于及时检测和提高。

总之,这套系列教材配有光盘,方便教学,信息量大;融入软件,突出技能,实用性强。内容可视化,你不会再为抽象而烦恼;计算软件化,你不会再为繁难而困扰;方法现实化,你不会再因无用而厌学。

2003年冬季，我有幸到澳大利亚 La Trobe 大学学习考察，亲身经历了国外大学数学教育对学生能力、素质培养的实践，它们特别重视数学思想的熏陶和数学知识的应用，“做中学、学中悟、悟中醒、醒中行”做得非常出色。让我们可喜的是即将出版的这套系列教材恰好在这方面做了有益的尝试。

我们期盼这套系列教材能为广大读者带来学数学的轻松、做数学的快乐和用数学的效益。

“新世纪大学数学立体化系列教材”编委员会主任、总主编 于义良

2004年3月

前 言

进入新世纪,人类社会正逐步迈向信息化社会,这个社会所呈现出的两大特点是:计算机技术的迅速发展与广泛应用,以及由之引发的数学知识向各个领域更为广泛和深入的渗透。这两个特点无疑对信息化社会的数学教育提出了新的要求。由于数学应用的普遍性,数学教育质量的高低直接关系到国民素质的高低;由于未来高新技术从某种意义上可归结为数学技术,所以一个国家的数学教育水平在一定程度上也反映了这个国家的综合国力。

《概率论与数理统计》是高等学校本科阶段各专业普遍开设的惟一一门处理随机性现象数量规律性的数学课程。由于随机性现象的普遍存在性、研究方法的独特性和教学内容的实用性,这门课越来越受到人们的重视。但是,长期以来,传统观念的束缚和影响以及计算技术的落后使这门课成了“空中楼阁”,理论一大堆,抽象难懂;公式一大堆,混杂难记;方法一大堆,学了难用;学生学习被动,没兴趣。这些因素使得这门课难以发挥应有的教育功能,同时也制约了学科自身的发展。在素质教育观下,如何将“数学素质教育”贯穿到《概率论与数理统计》课程的教学过程中,便是从事这门课程教学的每个人应该研究的课题。

经过多年的教学实践和认真学习讨论,汲取过去的经验和教训,调查了解国内外数学教育的改革动态,我们认为《概率论与数理统计》课程从内容到教学方法,都应突出这门课的两点特点,即理论

联系实际和计算机技术应用,以“概率适度,统计加强,引入案例”为基本思路,真正使学生的数学实践能力(数学知识、数学建模、数值计算、数据处理)得到培养和提高。

现在奉献给大家的这本《概率论与数理统计》教材,力求基本知识、方法与计算机应用融为一体,统计计算全部由 Excel 实现,使学生感受到所学知识、方法的可用性;通过完成“现实题材——数学问题——数学建模——数学知识与方法——成果释译”的学习过程,帮助学生实现由知识向能力的转化。教学内容贴近实际,让学生边学边提出解决问题的思路和设想,引导学生运用所学知识解决实际问题,培养学生的社会责任感。

参加本书编著的有:于义良、安建业、张银生、李秉林、李乃华、郝茵、张建新、王莉琴、陈萍、梁帮助、李美凤、李爱玲、滕树军。

在本书编著过程中,我们得到了天津市教委高教处、天津商学院教务处的指导,特别是得到天津商学院教材中心、中国人民大学出版社许多同志的大力支持,天津职业技术师范学院何文章教授、天津科技大学张大克教授以及湖南商学院、上海水产大学、天津理工学院、山西大学工程学院、运城学院的老师们提出了非常好的建议,为体现内容的典型性,书中部分例题引自他人著作,在此一并表示衷心的感谢。由于我们水平有限,书中若有不尽如意的地方,敬请广大读者雅正。

编著者

2004年3月18日

目 录

第 1 章 随机变量	1
第 1.1 节 随机事件及其概率	1
1. 随机事件	1
2. 随机事件的概率	4
习题 1.1	7
第 1.2 节 条件概率与独立性及其应用	10
1. 条件概率与独立性	10
2. 应用	13
习题 1.2	18
第 1.3 节 随机变量及其分布	21
1. 随机变量	21
2. 离散型随机变量的概率分布	22
3. 连续型随机变量的概率密度函数	26
习题 1.3	32
第 1.4 节 随机变量的分布函数	36
1. 分布函数	36
2. 正态分布的分布函数及其计算	40
习题 1.4	43
第 2 章 随机向量	47
第 2.1 节 随机向量及其分布	47
1. 随机向量	47
2. 二维离散型随机向量及其联合概率分布、	

	边缘概率分布	48
	3. 二维连续型随机向量及其联合概率密度、 边缘概率密度	52
	习题 2.1	58
第 2.2 节	随机向量的联合分布函数	61
	1. 联合分布函数	61
	2. 随机变量的相互独立性	62
	3. 随机向量函数的分布	67
	习题 2.2	72
第 3 章	数字特征	76
第 3.1 节	随机变量的数字特征	77
	1. 数学期望	77
	2. 方差	86
	习题 3.1	92
第 3.2 节	随机向量的数字特征	95
	1. 随机向量的数学期望和方差	95
	2. 协方差	95
	3. 相关系数	97
	习题 3.2	102
第 3.3 节	大数定律与中心极限定理	104
	1. 大数定律	104
	2. 中心极限定理	107
	习题 3.3	110
第 4 章	统计估值	112
第 4.1 节	数理统计学中的基本概念	113
	1. 基本概念	113
	2. 正态总体下的常用统计量及其分布	114
	3. Excel 实现	117
	习题 4.1	119

第 4.2 节	分布密度的近似求法	122
1.	频率柱形图	122
2.	Excel 实现	123
第 4.3 节	期望与方差的点估计	125
1.	参数估计的基本思想	125
2.	点估计的评选标准	126
3.	数学期望与方差的点估计	129
	习题 4.3	130
第 4.4 节	期望、方差的区间估计及 Excel 实现	131
1.	单正态总体数学期望的区间估计	132
2.	单正态总体方差的区间估计	135
3.	两个正态总体均值差的区间估计	137
4.	两个正态总体方差比的区间估计	140
	习题 4.4	143
第 4.5 节	点估计法	145
1.	矩估计法	145
2.	最大似然估计法	146
	习题 4.5	148
第 5 章	统计检验	150
第 5.1 节	统计检验概要	150
1.	统计检验的基本思想	152
2.	统计检验的实施程序	153
3.	两类错误	153
第 5.2 节	单正态总体的统计检验及 Excel 实现	154
1.	期望的检验	154
2.	方差的检验	160
	习题 5.2	163
第 5.3 节	两正态总体的统计检验及 Excel 实现	164
1.	两总体均值之差的检验	164

	2. 成对数据比较检验法	169
	3. 两总体方差之比的检验	171
	习题 5.3	174
第 5.4 节	两个需要说明的问题	175
	1. 统计检验与区间估计的关系	175
	2. 检验的 p 值	176
第 6 章	方差分析	179
第 6.1 节	方差分析的基本思想	179
第 6.2 节	单因素方差分析	181
	1. 数学模型	181
	2. 方差分析	182
第 6.3 节	双因素方差分析	185
	1. 数学模型	186
	2. 方差分析	186
第 7 章	回归分析	192
第 7.1 节	一元回归分析模型	192
第 7.2 节	回归系数的最小二乘估计	194
第 7.3 节	回归估计的统计推断	198
	1. a, b 的点估计	198
	2. a, b 的点估计的方差	199
	3. σ^2 的点估计和 a, b 的估计标准误差	199
	4. a 和 b 的区间估计	199
	5. $E(y_i)$ 的区间估计	200
	6. y 的样本变差的分解	200
	7. 回归方程的显著性检验	200
	8. 回归分析的表述	201
第 7.4 节	预测	204
	1. 预测值	204
	2. 预测区间	207

第 7.5 节 多元回归分析	208
1. t 检验	213
2. F 检验	214
附表 1 标准正态分布表	216
附表 2 t 分布表	217
附表 3 χ^2 分布表	218
附表 4 F 分布表	219
参考文献	220

第 1 章

随机变量

我们生活在一个日新月异、千变万化的世界中,每个人时时刻刻都要面对生活中碰到的问题.例如:“明天是下雨还是晴天,是否可以出去旅游”;“明天的股市是上涨还是下跌,是买入还是卖出”;“下个月某空调器的销售量是多少,如何组织货源”;“今年夏季长江流域的降雨量有多少,怎样组织抗洪”;“在下届奥运会中我国体育健儿能拿多少块金牌,如何争取金牌总数第一”……这些问题的发生与发展是受诸多因素的影响的,因而这些问题的结果也是不确定的,不可预知的.但是,在许多不确定性问题中往往隐含着一种确定性的规律.也正因为如此,人类才在不断摸索和研究中使许多以前认为不可想像的问题得到解决,人类才取得了辉煌的进步.数学家维纳(N. Wiener)说:“数学的伟大使命是在混沌中发现有序.”

本章的目的就是从基本问题出发,引出随机事件、随机变量的概念,试图从最简单的随机现象(偶然现象)中去探求必然的规律.

第 1.1 节 随机事件及其概率

1. 随机事件

我们先做一个简单的试验:掷一颗质地均匀的骰子,观察出现的点数.显然,这个试验具有如下特点:

- (1) 在相同的条件下可以重复进行;
- (2) 每次试验的可能结果不止一个(出现 1 点, 2 点, $\dots$, 6 点),而究竟

出现哪个结果,在试验之前不能预言;

(3) 试验之前可以预知试验中一切可能的结果(6种结果),每次试验中出现且只出现可能结果中的一个.

我们把具有上述三个特点的试验称为**随机试验**(random experiment),简称**试验**(experiment),记为 E . 本书中所谈的试验,均为随机试验. 我们就是研究随机试验中这些可能结果出现的规律.

试验中的每个可能的结果称为**样本点**(sample),用 ω 表示,全体样本点构成的空间称为**样本空间**(sample space),用 Ω 表示. 从集合论的观点看,样本空间就是针对该试验的全集,而样本点就是构成样本空间的元素. 样本空间 Ω 的子集称为**随机事件**(random event),简称**事件**(event),一般用大写字母 A, B, C 等表示. 我们称一事件在一次试验中出现(发生)了是指该次试验出现的结果(样本点)属于该事件(子集),否则称该事件没有出现(发生). 显然, Ω 在每次试验中均要出现,故称**必然事件**(certainty). 空集 $\emptyset$ 在每次试验中均不出现,故又称**不可能事件**(impossible event).

在上面掷骰子的试验中,我们设:

ω_i 代表出现的点数为 i ($i = 1, 2, \dots, 6$), 则 ω_i 为样本点, 且记 $A_i = \{\omega_i\}$

$B =$ “出现的点数是 2 或者 3”

$C =$ “出现的点数大于 1 小于 5”

$D =$ “出现的点数是偶数”

则样本空间为: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$

$$B = \{\omega_2, \omega_3\}$$

$$C = \{\omega_2, \omega_3, \omega_4\}$$

$$D = \{\omega_2, \omega_4, \omega_6\}$$

显然 A_i ($i = 1, 2, \dots, 6$), B, C, D 均为事件. 而事件“出现的点数小于 7”是必定要出现的, 它就是一个必然事件. 实际上它包括了所有的样本点, 即为 $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. 而事件“出现的点数大于 6”是不可能出现的, 它就是一个不可能事件. 实际上它不包含任何样本点, 即为一个空集 $\emptyset$.

由集合和随机事件之间的关系, 可得到如下结论:

(1) $A \subset B$. A 是 B 的子集. 它表示若事件 A 出现则事件 B 一定出现.

(2) $A \cup B$ (或 $A + B$). A 与 B 的并(和). 它表示一个新事件, 即事件 A 和事件 B 至少有一个出现. 同样, $\bigcup_{i=1}^n A_i$ 表示 $A_1, A_2, \dots, A_n$ 这 n 个事件至

少有一个出现.

(3) $A \cap B$ (或 AB). A 与 B 的交 (积). 它表示一个新的事件, 即事件 A 和事件 B 同时出现. 同样, $\bigcap_{i=1}^n A_i$ 表示 $A_1, A_2, \dots, A_n$ 这 n 个事件同时出现.

(4) $A \cap B = \emptyset$. 表示事件 A 和事件 B 不可能同时出现, 称 A 与 B 为互不相容, 简称互斥 (mutually exclusive events).

(5) $A \cap B = \emptyset$ 且 $A \cup B = \Omega$. 表示事件 A 和事件 B 出现且只出现其中一个, 称为 A 与 B 对立, 并称 B 为 A (或 A 为 B) 的对立事件 (complementary event), 记为 $B = \bar{A}$ (或 $A = \bar{B}$).

(6) $A - B$. A 与 B 的差. 它表示事件 A 出现而事件 B 不出现. 显然, $A - B = A\bar{B}$, 同样 $B - A = B\bar{A}$.

(7) $\overline{A \cup B}$. 它是事件 A 和事件 B 至少出现一个的对立事件. 显然它表示事件 A 和事件 B 都不出现, 即 $\overline{A \cup B} = \bar{A} \cap \bar{B}$. 同样 $\bigcup_{i=1}^n A_i = \bigcap_{i=1}^n \bar{A}_i$.

(8) $\overline{A \cap B}$. 它是事件 A 和事件 B 同时出现的对立事件. 显然它表示事件 A 和事件 B 至少有一个不出现, 即 $\overline{A \cap B} = \bar{A} \cup \bar{B}$. 同样 $\bigcap_{i=1}^n A_i = \bigcup_{i=1}^n \bar{A}_i$.

下面用图将上面的结论直观地显示一下 (如图 1.1.1).

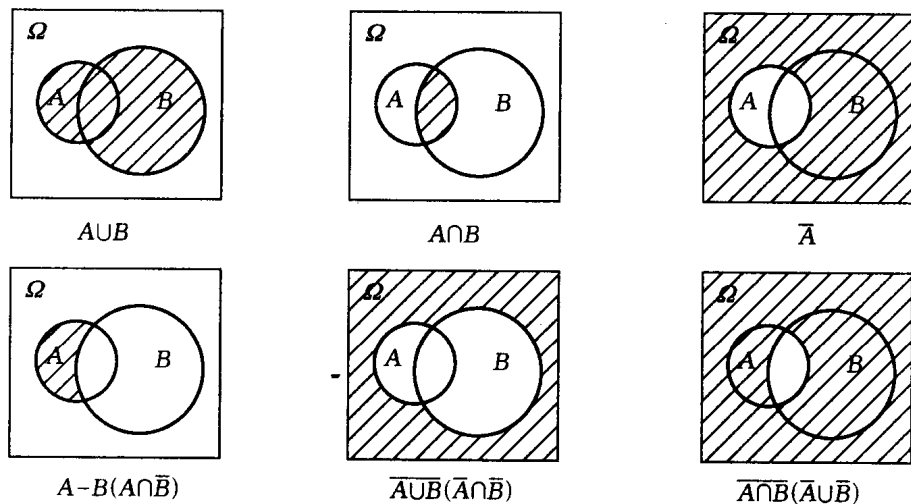


图 1.1.1

从上面的图示中,还可以得到以下一些关系:

$A \cap B \subset A; A - B \subset A$, 且 $A - B = A - AB, AB \subset A; A\bar{B}, B\bar{A}$ 与 AB 两两互斥, $A = A\bar{B} \cup AB, A \cup B = A\bar{B} \cup B\bar{A} \cup AB = A\bar{B} \cup B = B\bar{A} \cup A = A \cup (B - A); \overline{A \cap B} \neq \bar{A} \cap \bar{B}; \overline{A \cup B} \neq \bar{A} \cup \bar{B} \dots\dots$

结合前面的试验以及字母 $\omega_i (i = 1, 2, \dots, 6), B, C, D$ 表示的意义,我们有:

$$\begin{aligned} B &\subset C; & C \cup D &= \{\omega_2, \omega_3, \omega_4, \omega_6\}; \\ C \cap D &= \{\omega_2, \omega_4\}; & \bar{B} &= \{\omega_1, \omega_4, \omega_5, \omega_6\}; \\ \overline{C \cup D} &= \{\omega_1, \omega_5\}; & \overline{C \cap D} &= \{\omega_1, \omega_3, \omega_5, \omega_6\}; \\ C - D &= \{\omega_3\}; & D - C &= \{\omega_6\}. \end{aligned}$$

大家还可以验证其他一些等式或不等式.

上面一些表示法,对我们今后处理某些复杂事件会带来很大方便.

2. 随机事件的概率

下面再回到原来的试验:掷一颗质地均匀的骰子,观察出现的点数.

在这个试验中显然有:

- (1) 所有可能的试验结果是有限个(有限性);
- (2) 每个可能结果在一次试验中出现的可能性相同(等可能性).

我们把具有这两个特点的试验称为**古典概型**(classical probability).

我们知道,在每次试验中出现且只出现样本空间中的一个样本点.所谓事件 A 出现,就是指试验中出现了事件 A 中包含的样本点,因此在古典概型中由于样本点出现的等可能性,事件 A 出现的可能性大小就可以用事件 A 中包含的样本点的个数占样本点总数的比例来度量.

定义 1.1.1 在古典概型中,设样本空间 Ω 中包含有 n 个样本点,则对任意事件 A ,若 A 中含有 k 个样本点,那么事件 A 的概率 $P(A)$ 定义为

$$P(A) = \frac{\text{事件 } A \text{ 中包含的样本点数}}{\text{样本空间 } \Omega \text{ 中样本点总数}} = \frac{k}{n}$$

在掷骰子的试验中,显然有

$$\begin{aligned} P(A_i) &= \frac{1}{6} \quad (i = 1, 2, 3, 4, 5, 6) & P(B) &= \frac{2}{6} \\ P(C) &= \frac{3}{6} & P(D) &= \frac{3}{6} & P(C \cup D) &= \frac{4}{6} \\ P(C \cap D) &= \frac{2}{6} & P(\bar{B}) &= \frac{4}{6} & P(\Omega) &= \frac{6}{6} = 1 \end{aligned}$$

$$P(\emptyset) = \frac{0}{6} = 0$$

在计算古典概型概率中,往往要用到中学所学的排列与组合的知识.

例 1.1.1 箱中装有 10 件产品,其中 1 件是次品,在 9 件合格品中有 6 件是一等品,3 件二等品.现从箱中任取 3 件,试求:

- (1) 取得 3 件产品都是一等品的概率;
- (2) 取得 3 件产品中有 1 件是一等品,2 件是二等品的概率;
- (3) 取得 3 件产品中至少有 2 件是一等品的概率.

解 由于试验中是任取 3 件,显然这个试验是个古典概型.每个样本点就是从 10 件中任取 3 件产品构成的集合,与顺序无关,故样本空间中样本点的总数为 C_{10}^3 .

(1) 设 $A =$ “取得 3 件产品都是一等品”,那么 A 中的样本点个数为 C_6^3 ,所以

$$P(A) = \frac{C_6^3}{C_{10}^3} = \frac{1}{6}$$

(2) 设 $B =$ “取得 3 件产品中有 1 件是一等品,2 件是二等品”,那么 B 中的样本点个数为 $C_6^1 \cdot C_3^2$ (这里用到了乘法原理),所以

$$P(B) = \frac{C_6^1 \cdot C_3^2}{C_{10}^3} = \frac{3}{20}$$

(3) 设 $C =$ “取得 3 件产品中至少有 2 件是一等品”,那么事件 C 显然是由“有 2 件一等品,1 件非一等品”和“3 件都是一等品”两个事件构成,所以 C 中的样本点个数为 $C_6^2 \cdot C_4^1 + C_6^3$ (这里又用到了加法原理),那么

$$P(C) = \frac{C_6^2 C_4^1 + C_6^3}{C_{10}^3} = \frac{2}{3}$$

生活中还有很多问题并不具有古典概型的两个特点,例如投掷的是一颗不均匀的骰子,那么样本空间中的 6 个样本点出现的可能性就不相等了,因此计算其中的一些事件的概率时就不能再用古典概型的公式了.如何解决这类问题呢?最常用也是最直接的办法就是反复掷这颗骰子,例如做了 n 次试验,记下所关心的事件 A 在这 n 次试验中出现的次数,例如出现了 μ 次,比值 $\frac{\mu}{n}$ 随着试验次数 n 的增大而越来越接近某个数值 p ,那么我们就称数值 p 为事件 A 的概率,即 $P(A) = p$. 但是数值 p 是个理论上的数,在有限次的试验中很难求得.一般就用 $\frac{\mu}{n}$ 作为 $P(A)$ 的近似值,即