中国石油大学（华东）远程与继续教育系列教材

概率论与数理统计

（第二版）

Probability Theory and
Mathematical Statistics

常兆光 王清河 曹晓敏 编

石油工業出版社

中国石油大学（华东）远程与继续教育系列教材

概率论与数理统计

（第二版）

常兆光　王清河　曹晓敏　编

石油工业出版社

内 容 提 要

本书是按照普通高等理工院校成人教育《概率论与数理统计》课程基本要求编写的，内容包括随机事件与概率、随机变量及分布、随机变量的数字特征、大数定律和中心极限定理、数理统计的基本概念、参数估计、假设检验、回归分析、方差分析与正交试验设计等。每章均配适量练习。

本书可作为普通高等理工院校成人、网络教育《概率论与数理统计》课程的教材与教学参考书，也可作为工程技术人员的参考书。

图书在版编目（CIP）数据

概率论与数理统计/常兆光等编．—2版．

北京：石油工业出版社，2007.4

（中国石油大学（华东）远程与继续教育系列教材）

ISBN 978-7-5021-5907-8

Ⅰ．概…

Ⅱ．常…

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（2006）第162363号

出版发行：石油工业出版社

（北京安定门外安华里2区1号　100011）

网　址：www.petropub.cn

发行部：（010）64210392

经　销：全国新华书店

印　刷：石油工业出版社印刷厂

2007年4月第2版　2007年4月第6次印刷

787×960毫米　开本：1/16　印张：15

字数：284千字　印数：15301—23300册

定价：17.00元

中国石油大学(华东)
远程与继续教育系列教材编审委员会

总　序

从1995年创办函授夜大学至今，中国石油大学成人教育已经走过了从初创、逐步成熟到跨越式发展的50载历程。50多年来，我校成人教育紧密结合社会经济发展需求，积极开拓新的服务领域，为石油、石化企业培养、培训了10多万名本专科毕业生和管理与技术人才，他们中的大多数已经成为各自工作岗位的骨干和中坚力量。我校成人教育始终坚持“规范管理、质量第一”的办学宗旨，坚持“为石油石化企业和经济建设服务”的办学方向，赢得了良好的社会信誉。

自2001年1月教育部批准我校开展现代远程教育试点工作以来，我校以“创新教育观念”为先导，以“构建终身教育体系”为目标，整合函授夜大学教育、网络教育、继续教育资源，建立了新型的教学模式和管理模式，构建了基于卫星数字宽带和计算机宽带网络的现代远程教育教学体系和个性化的学习支持服务体系，有效地将学校优质教育资源辐射到全国各地，全力打造出中国石油大学现代远程教育的品牌。目前，办学领域已由创办初期的函授夜大学教育发展为今天的集函授夜大学教育、网络教育、继续教育、远程培训、国际合作教育于一体的，在国内具有领先水平、在国外有一定影响的现代远程开放教育系统，成为学校高等教育体系的重要组成部分和石油、石化行业最大的成人教育基地。

为适应现代远程教育发展的需要，学校于2001年9月正式启动了网络课程研制开发和推广应用项目，斥巨资实施“名师名课”教学资源精品战略工程，选拔优秀教师开发网络教学课件。随着由多媒体课件、WEB课件到网络课程的不断充实与完善，建构了内容丰富、形式多样的网络教学资源超市，基于网络的教学环境初步形成，远程教育的能力有了显著提高，这些网上教学资源的建设与研发为我校远程教育的顺利发展起到了支撑和保障作用。相应地，作为教学资源建设的一个重要组成部分，与网络教学课件相配套的纸质教材建设就成为一项愈来愈重要的任务。根据学校现代远程教育发展规划，在“十一五”期间，学校将推进精品课程、精品网络课件和教材建设工作，通过立项研究方式启动远程与继续教育系列教材建设工作，选聘石油石化行业和有关石油高校专家、学者参与系列教材的开发和编著工作，计划用5年的时间，以石油、化工等主干专业为重点，陆续推出成人学历教育、岗位培训、继续教育三大系列教材。系列教材将充分吸收科学技术发展和成人教

育教学改革最新成果，体现现代教育思想和远程教育教学特点，具有先进性、科学性和远程教育教学的适用性，形成纸质教材、多媒体课件、网上教学资料互为补充的立体化课程学习包。

为了保证“远程与继续教育系列教材”编写出版进度和质量，学校成立了专门的远程与继续教育系列教材编审委员会，对系列教材进行严格的审核把关。目前，远程与继续教育系列教材的编写还处于探索阶段，随着我校现代远程教育的进一步发展，新课程的开发、新教材的编写将持续进行，本系列教材的体系也将不断完善。我们相信，有广大专家、学者们的共同努力，一定能够创造出适合现代远程教育教学和学习特点、体系新、水平高的远程与继续教育系列教材。

中国石油大学(华东)远程与继续教育学院

2006 年 10 月

第二版前言

本书出版后，曾收到、听到许多专家及读者的热情指教和建议，现依据高等成人教育《概率论与数理统计》课程教学基本要求，结合我们教学工作的体会，对本书作了如下修改：(1)对部分内容及表达方式进行了修改；(2)对印刷错误作了校正；(3)对前五章原有例题作了适当调整补充，并增加了部分综合性习题，使之题型更全面、更有代表性。

本书再版的修改过程中，得到了王才经教授的大力支持和帮助并审阅了全书，给出了许多指导性意见，在此表示衷心的感谢。

由于作者水平有限，难免仍有错误或不恰当之处，竭诚地希望读者继续对本书提出批评指正。

编者

2006 年 10 月

第一版前言

《概率论与数理统计》是研究和揭示随机现象统计规律性的数学学科，是高等成人教育本、专科各专业的一门重要基础理论课，是基础数学与工程专业课程的桥梁，在各行各业都有广泛的应用。本书是按照普通高等理工院校成人教育《概率论与数理统计》课程基本要求编写的。只要具备高等数学、线性代数的知识，就能顺利地学好其中的内容，达到该课程的基本要求。由于该课程内容比较抽象、思维方式比较独特、知识面涉及广，使自学具有一定的困难，因此本书的编写力求重点突出、深入浅出、通俗易懂，具有较强的实用性。对基本概念、重要公式和定理，注意其实际意义的解释和说明，讲清其实质，并在每节中配有较多的例题。通过各种类型例题的分析、运算，进一步加深和巩固所学的理论与概念，培养学生分析与解决随机现象问题的能力，培养学生的创造性思维能力。

本书可作为普通高等理工院校成人、网络教育《概率论与数理统计》课程的教材与教学参考书，也可作为工程技术人员的参考书。

在教材处理上，部分章、节打 * 号的内容可不作为基本要求，留给读者自学。建议第一章至第五章的参考学时为 40 学时左右。在例题的选择上也分两种类型，一种例题是比较简单且易于为人们所理解的，这种例题主要是用来便于引入或解释概念、定义和定理；另一种例题是提高型的，主要用来培养分析问题和解决问题的能力，进一步加深对概念、定义和定理的理解和掌握。个别打 * 号的例题不作为基本要求。习题中打 * 号的题目是为拓宽知识面之用，也不作为基本要求。

在本书编写过程中，石油大学(华东)成人(网络)教育学院、应用数学系给予了大力支持；王才经教授在百忙之中对本书初稿进行了详细认真的审阅，并提出了许多宝贵意见，在此一并表示感谢。

由于编者水平有限，错误与不足之处在所难免，敬请读者批评指正。

编者

2003 年 5 月

目　　录

第一章　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。本章将主要介绍随机事件、随机事件的概率、概率的基本性质、条件概率以及计算概率常用到的几个重要公式。通过本章的学习，应初步掌握处理随机现象的基本思想和方法，学会运用概率方法分析和解决实际问题。基本要求如下：

(1)理解随机事件和样本空间的概念，掌握事件之间的关系与基本运算。

(2)理解频率的概念和概率的公理化定义及古典概率定义，会进行简单古典概率的计算。

(3)掌握概率的基本性质并会应用这些性质进行概率的计算。

(4)理解条件概率的概念，掌握概率的乘法公式、全概率公式和贝叶斯(Bayes)公式以及应用这些公式进行概率计算。

(5)理解事件的独立性的概念，会应用事件的独立性进行概率计算。掌握贝努利(Bernoulli)概型和二项概率的计算方法。

第一节　随机试验与随机事件

一、必然现象与随机现象

在生产实践、科学实验和日常生活中，人们观察到的现象一般可分为两种类型，一类是确定性现象，即在一定条件下，某些事情一定发生或一定不发生的现象，也称为必然现象。例如：纯净水在一个标准大气压下，加热到 100℃必然沸腾；上抛的物体必然下落等等。早期的科学就是研究这一类现象的规律性，所应用的数学工具如高等数学、几何、代数等是大家熟悉的。但人们逐渐发现另一类现象，它是事前不可预言的。即在一定条件下，可能发生也可能不发生的现象，这一类现象我们称之为偶然现象或随机现象。例如：抛掷一枚均匀硬币，是正面朝上还是反面朝上；从一批含有次品的产品中任取一件，取得的产品是正品还是次品等，这些都是事前不能肯定的。类似的例子还可以举出很多。

必然现象遵循必然性规律，人们根据已知的事实推断它将发生的结果。随机现象具有明显的不确定性(随机性)，就一次试验(或观察)而言，其结果难以确定；但若进行大量重复试验，其结果就会呈现出某种规律性，即所谓统计规律。概率论与数理统计的任务就是要研究和揭示随机现象的这种统计规律性。

二、随机试验

在工农业生产、科学实验和现实生活中，我们遇到过各种各样的试验。在概率论中，我们把试验作为一种广泛的术语，它包括各种各样的科学试验，甚至对某一事物的某一特征的观察也可认为是一种试验。例如(以 E 或 E 加一个下标表示试验)：

E_1：掷一枚均匀硬币，观察正、反面出现的情况(正面记为 H，反面记为 T)；

E_2：将一枚均匀硬币抛两次，观察正、反面出现的情况；

E_3：将一枚均匀硬币抛两次，观察正面出现的次数；

E_4：掷一颗均匀骰子，观察出现的点数；

E_5：一名射手进行射击，直到击中为止，观察射击次数；

E_6：从一批灯泡中任意抽取一只，测其寿命。

不难看出，以上几个试验具有下面三个共同特征：

(1)试验可以在相同条件下重复进行；

(2)每次试验的可能结果不止一个，但能事先明确试验的所有可能结果；

(3)进行一次试验之前，不能确定哪一个结果会出现。

在概率论中，我们把具有以上三个特征的试验称为随机试验，简称为试验，记作 E。

三、随机事件与样本空间

将试验 E 的所有可能出现的结果组成的集合称为 E 的样本空间，记作 Ω。Ω 中的每个元素(可能结果)称为样本点。例如上述试验 $E_i(i=1,2,3,4,5,6)$的样本空间 Ω_i 分别为：

$$\Omega_1 = \{H,T\}$$
$$\Omega_2 = \{HH,HT,TH,TT\}$$
$$\Omega_3 = \{0,1,2\}$$
$$\Omega_4 = \{1,2,3,4,5,6\}$$
$$\Omega_5 = \{1,2,3,4,\cdots\}$$
$$\Omega_6 = \{t \mid t \geqslant 0\}$$

需要指出的是，样本空间中的样本点是由试验目的所确定的。例如 E_2 和 E_3 同是将一枚均匀硬币抛两次，由于试验目的不同，其样本空间也就不一样。

样本空间包含了试验 E 的所有可能结果，将每一个可能的结果称为随机事件(简称为事件)，通常用大写字母 A、B、C 等表示。只包含一个样本点的事件称为基本事件。例如掷骰子试验中，每一个可能出现的点数都是基本事件。而由两个或两个以上的基本事件(样本点)组成的事件称为复合事件。例如将一枚均匀硬币连续掷两次，观察正、反面出现的情况，其样本空间 $\Omega=\{HH,HT,TH,TT\}$，则两次

都出现正面为一基本事件，记作 $A=\{HH\}$；至少有一次出现正面为一复合事件，记作 $B=\{HH,HT,TH\}$。可以看出，B 事件是由三个类似于事件 A 的基本事件所组成的，即这三个事件只要有一个发生就认为 B 事件发生。

在每次试验中，一定发生的事件叫做必然事件，记作 Ω；而一定不发生的事件叫做不可能事件，记作 $\varnothing$。

需要指出的是，无论是必然事件、随机事件还是不可能事件，都是相对“一定条件”而言的。条件发生变化，事件的性质也发生变化。例如抛掷两颗骰子，“出现的点数之和为 5 点”及“出现的点数之和大于 5 点”都是随机事件。若同时抛掷 6 颗骰子，“出现的点数之和为 5 点”则是不可能事件了；而“出现的点数之和大于 5 点”则是必然事件了。为了以后讨论问题方便，通常将必然事件和不可能事件看成是特殊的随机事件。

四、事件及其运算关系

由样本空间的定义知，它是随机试验中所有可能出现的结果（样本点）组成的集合，因此随机事件又可理解为样本空间中具有某种特性的样本点构成的集合，而基本事件可看成此集合的元素。这样一来，集合论中集合之间的运算均可推广到事件之间的运算中来。

若记 Ω：样本空间（必然事件）；$\varnothing$：不可能事件；e：基本事件；$A,B,A_1,A_2,\cdots,A_n$：随机事件，则事件之间的运算关系如下：

(1)包含关系：若 $A\subset B$，则称事件 B 包含事件 A，表示事件 A 发生必导致事件 B 发生，如图 1－1 所示。

对任一事件 A，都有 $\varnothing\subset A\subset\Omega$。

(2)相等关系：若 $A\subset B$ 且 $B\subset A$，则称事件 A 等于事件 B，记作 $A=B$。

(3)和事件：事件 $A\cup B$ 称为事件 A 与事件 B 的和事件，表示事件 A 与事件 B 至少有一个发生，如图 1－2 所示。

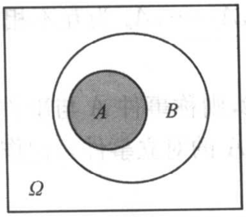

图 1－1

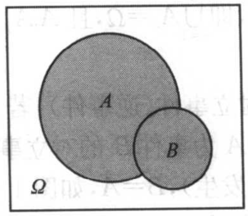

图 1－2

事件 $A_1\cup A_2\cup\cdots\cup A_n\triangleq\bigcup\limits_{i=1}^{n}A_i$ 称为有限个事件的和事件，表示事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生。

类似地，事件 $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots \triangleq \bigcup_{i=1}^{\infty} A_i$ 称为无限多个事件的和事件，表示事件 $A_1, A_2, \cdots, A_n, \cdots$ 至少有一个发生。

(4)积事件：事件 $A \cap B$ 称为事件 A 与事件 B 的积事件，简记为 AB，表示事件 A 与事件 B 同时发生，如图 1-3 所示。

事件 $A_1 \cap A_2 \cap \cdots \cap A_n \triangleq \bigcap_{i=1}^{n} A_i$ 称为有限个事件的积事件，表示事件 A_1，$A_2, \cdots, A_n$ 同时发生。

类似地，事件 $A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots \triangleq \bigcap_{i=1}^{\infty} A_i$ 称为无限多个事件的积事件，表示事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生。

(5)差事件：事件 $A-B$ 称为事件 A 与事件 B 的差事件，表示事件 A 发生而事件 B 不发生，如图 1-4 所示。

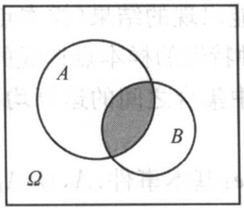

图 1-3

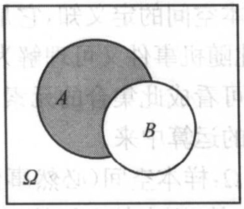

图 1-4

(6)互不相容(互斥)事件：若 $AB=\varnothing$，则称事件 A 与事件 B 互不相容，表示事件 A 与事件 B 不同时发生，如图 1-5 所示。基本事件是互不相容的。

类似地，若 $\forall i \neq j, A_iA_j=\varnothing(i、j=1,2,\cdots)$，，则称 $A_1, A_2, \cdots, A_n, \cdots$ 为两两互不相容的。

若 $A_1, A_2, \cdots, A_n$ 两两互不相容，且在每次试验中事件 $A_1, A_2, \cdots, A_n$ 必发生其中之一，即 $\bigcup_{i=1}^{n} A_i=\Omega$，且 $A_iA_j=\varnothing(i \neq j)$，则称 $A_1, A_2, \cdots, A_n$ 为互不相容事件完备组。

(7)对立事件(逆事件)：若 $AB=\varnothing$ 且 $A \cup B=\Omega$，则称事件 A 与事件 B 互逆，又称事件 A 为事件 B 的对立事件(或事件 B 为事件 A 的对立事件)，记作 $A=\overline{B}$($\overline{B}$ 表示 B 不发生)，$B=\overline{A}$，如图 1-6 所示。

需要指出的是，由(6)、(7)知，若事件 A 与事件 B 互逆，则事件 A 与事件 B 一定互不相容，但反之则不一定。

例 1-1 在试验 E_2 中，若记 $B_1=\{HH\}, B_2=\{HT\}, B_3=\{TH\}, B_4=\{TT\}$，则

A_1 表示“第一次出现正面”，即 $A_1=\{HH, HT\} \triangleq B_1 \cup B_2$；

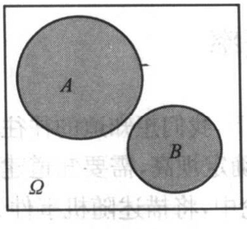

图 1－5

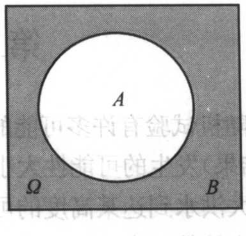

图 1－6

A_2 表示“两次出现同一面”，即 $A_2=\{HH,TT\}\triangleq B_1\cup B_4$；

A_3 表示“只出现一次正面”，即 $A_3=\{HT,TH\}\triangleq B_2\cup B_3$；那么，$A_1\cup A_2=\{HH,HT,TT\}$；$A_1\cap A_2=\{HH\}=B_1$；$A_1-A_2=A_1\overline{A}_2=B_2$。由于 $A_2A_3=\varnothing$，故 A_2 与 A_3 互不相容；又由于 $A_2\cup A_3=\Omega$，所以 A_2 与 A_3 互逆。

(8)德·摩根定律(对偶原理)：

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\overline{A\cap B}=\overline{A}\cup\overline{B}$$

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i},\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i}$$

(9)运算规律：

设 A、B、C 为三个事件，则有

交换律：$A\cup B=B\cup A,A\cap B=B\cap A$

结合律：$A\cup(B\cup C)=(A\cup B)\cup C$

$A\cap(B\cap C)=(A\cap B)\cap C$

分配律：$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$

例 1－2 设 A、B、C 为三个事件，试用事件之间的运算关系表示下列事件：

(1)A 发生而 B 与 C 都不发生；

(2)A、B、C 恰有一个发生；

(3)A、B、C 至少有一个发生；

(4)A、B、C 至多有两个事件发生。

解：(1)A 发生而 B 与 C 都不发生，也就是 A、$\overline{B}$、$\overline{C}$ 同时发生，即 $A\cap\overline{B}\cap\overline{C}$ 或 $A-B-C$；类似可得

(2)$(A\cap\overline{B}\cap\overline{C})\cup(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)$；

(3)$A\cup B\cup C$；

(4)$\overline{A}\cup\overline{B}\cup\overline{C}$ 或 $\overline{ABC}$。

第二节　频率与概率

一个随机试验有许多可能的结果，而在许多情况下，我们想知道的往往是随机事件(可能结果)发生的可能性大小。如建造水坝时，为确定坝高，需要知道建造水坝地段每年最大洪水到达某高度的可能性大小。在概率论中，将描述随机事件 A 发生的可能性大小的数记为 $P(A)$，称 $P(A)$ 为随机事件 A 的概率。那么如何确定随机事件 A 的概率呢？一种方法是通过反复试验来确定，为此先来讨论频率的概念。

一、频率

引例 1－1　将一枚均匀硬币在同样条件下连续掷 n 次，用 A 表示出现正面这一事件，$n(A)$表示 n 次重复试验中 A 出现的次数，则$\frac{n(A)}{n}$在一定程度上能反映事件 A 发生的可能性大小，将其记为

$$f_n(A) \triangleq \frac{n(A)}{n}$$

试验表明(见表 1－1 及表 1－2)，随着 n 的增大，$\frac{n(A)}{n}$在$\frac{1}{2}$附近波动的幅度越来越小，逐渐稳定于$\frac{1}{2}$，$\frac{1}{2}$这个值称为 $f_n(A)$的稳定值，通常把这个稳定值称为事件 A(出现正面)的概率。$f_n(A)$称为 n 次重复试验中 A 出现的频率，$n(A)$称为 n 次重复试验中 A 出现的频数。下面给出频率的一般定义。

表　1－1

实验序号	$n=5$		$n=50$		$n=500$	
	$n(A)$	$f_n(A)$	$n(A)$	$f_n(A)$	$n(A)$	$f_n(A)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

表 1-2

实验者	n	$n(A)$	$f_n(A)$
蒲丰	4040	2048	0.5070
K·皮尔逊	12000	6019	0.5015
K·皮尔逊	24000	12012	0.5005

定义 1-1 设 E 为随机试验，A 为其中任一事件，$n(A)$ 为事件 A 在 n 次重复试验中出现的次数，则称比值 $\frac{n(A)}{n}$ 为 n 次试验中 A 出现的频率，记为

$$f_n(A) \triangleq \frac{n(A)}{n} \tag{1-1}$$

其中，$n(A)$ 称为事件 A 在 n 次重复试验中出现的频数。当 n 增大时，$f_n(A)$ 逐渐稳定于某一个确定值 $P(A)$，称 $P(A)$ 为频率的稳定值。

由频率的定义不难看出，$f_n(A)$ 具有以下性质：

(1)非负性：$0 \leqslant f_n(A) \leqslant 1$；

(2)规范性：$f_n(\Omega)=1$；

(3)可加性：若事件 A 与事件 B 互不相容，则

$$f_n(A \cup B) = f_n(A) + f_n(B)$$

进一步，若 $A_1, A_2, \cdots, A_m$ 两两互不相容，则

$$f_n(\bigcup_{i=1}^{m} A_i) = \sum_{i=1}^{m} f_n(A_i)$$

由掷硬币试验可见，用频率来刻画事件 A 发生的可能性大小比较直观，但它有随机波动的缺陷，即在一定条件下做重复试验，其结果可能是不一样的。因此，用频率的稳定值来刻画事件 A 发生的可能性大小是比较恰当的，从而得出概率的统计性定义如下。

定义 1-2 在不变条件下做大量重复试验，称在重复试验中事件 A 发生的频率的稳定值 p 为事件 A 的概率，记为 $P(A)$。

尽管由频率的稳定值可得概率 $P(A)$，但我们不可能对每个事件都通过做反复试验以求得 $P(A)$。为此，我们以频率的性质和频率的稳定值 $P(A)$ 为背景，采用抽象化方法给出概率 $P(A)$ 的一般定义。

二、概率的公理化定义

定义 1-3 设 E 为随机试验，Ω 为它的样本空间，对 E 中的每一个事件 A 都赋予一个实数，记为 $P(A)$，且满足

(1)非负性：$0 \leqslant P(A) \leqslant 1$；

(2)规范性：$P(\Omega)=1$；