APPLIED PROBABILITY AND STATISTICS

应用概率统计

梁冯珍　宋占杰　张玉环　编

天津大学出版社
TIANJIN UNIVERSITY PRESS

应用概率统计

梁冯珍　宋占杰　张玉环　编

天津大学出版社
Tianjin University Press

内容提要

全书共10章，主要包括随机事件与概率、随机变量及其概率分布、多维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计中的基本概念与抽样分布、参数估计、假设检验、回归分析和方差分析以及MATLAB软件的使用.各章配有大量习题，并在书后附有答案.本书力求使用较少的数学知识，阐述概率统计方法，注重概率统计方法以及MATLAB中的统计软件包在各个领域的应用，例题较多，可操作性强，便于自学.

本书可作为高等院校非数学专业的工、农、医、经济、管理等专业的应用概率统计教材，亦可作为实际工作者的自学参考书.

图书在版编目(CIP)数据

应用概率统计 / 梁冯珍等编.—天津:天津大学出版社，2004.5

ISBN 7-5618-1939-0

Ⅰ.应… Ⅱ.梁… Ⅲ.概率论-高等学校-教材 Ⅳ.O211

中国版本图书馆CIP数据核字(2004)第037688号

出版发行 天津大学出版社
出 版 人 杨风和
地　　址 天津市卫津路92号天津大学内(邮编:300072)
电　　话 发行部:022-27403647　邮购部:022-27402742
印　　刷 昌黎太阳红彩色印刷有限责任公司
经　　销 全国各地新华书店
开　　本 148mm×210mm
印　　张 15.5
字　　数 446千
版　　次 2004年5月第1版
印　　次 2004年5月第1次
印　　数 1-4 000
定　　价 23.00元

前　言

概率统计是现代数学的一个重要分支.近20年来,随着计算机的发展以及各种统计软件的开发,概率统计方法在金融、保险、生物、医学、经济、管理和工程技术等领域得到了广泛的应用.正因为如此,概率统计课程成为高等院校各专业最重要的数学必修课之一.我们在长期的教学实践中感到,要让学生掌握概率统计方法,并在日后解决实际问题时能得到很好的应用,单纯讲授理论是远远不够的.为此,我们在教材的编写中,力图体现以下三个原则.

1.淡化理论

注重概念和理论的直观解释,尽量避免纯数学化的论证,但保持叙述的严谨性.

2.案例丰富

注重概率统计方法的介绍,给出概率统计方法在各个领域中的应用案例,帮助学生正确理解和使用这些方法.

3.结合软件

注重统计软件的应用,充分发挥MATLAB统计软件包的作用.将统计方法和统计软件包相结合,给出各种统计方法的程序编写模式,增加可操作性.

全书共10章,主要内容包括随机事件和概率、随机变量及其概率分布、多维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理.数理统计中的基本概念与抽样分布、参数估计、假设检验、回归分析和方差分析以及MATLAB软件的使用.各章均配有大量习题,并在书后附有参考答案.

本书由梁冯珍担任主编.第1、2、5章由宋占杰编写,第3、4章由张玉环编写,第6、7、8、9章由梁冯珍编写,第10章及各章中的程序由李胜鹏编写,最后由史道济与梁冯珍统稿.

本教材所需教学时数为64~72学时.

本书是在博士生导师史道济教授的指导下编写的，史教授在百忙之中仔细阅读了全书的初稿，并提出了许多宝贵意见，在此表示衷心的感谢.同时也感谢天津大学出版社的大力支持.

由于编者水平有限，错谬之处在所难免，恳请同行和广大读者批评指正.

编　　者

2004年4月

目　录

第 1 章 随机事件与概率

在人类认识自然和改造自然的各种社会活动中，观察到的各种各样的现象大体上可归结为三种类型.第一种类型是确定性现象，我们把事前可以预知结果的，即在某些确定的条件满足时，某一确定的结果必然会发生的现象称为确定性现象.例如，在一个标准大气压下，水加热到 100℃时一定沸腾；两个同性电荷一定互相排斥，两个异性电荷一定互相吸引等，这种现象由确定性数学来研究.第二种类型是模糊现象，我们把客观事物的差异在中介过渡时所呈现的"亦此亦彼"性，称为模糊现象，例如高与低、美与丑等，这种现象由模糊数学来研究.第三种类型是我们将要研究的随机现象，我们把事前能够预知所有可能结果但在每次试验时，不能确定哪一种结果将要出现的现象称为偶然现象或随机现象.例如，抛掷一枚质地均匀的硬币，硬币落地后可能是正面(带国徽的一面)朝上，也可能是反面朝上；同一个射手向同一个目标发射多发炮弹，弹着点却可能落在目标附近的各个位置上；同一个人利用同一台电脑抽奖多次，每一次抽得的号码可能不相同等等.

概率论(probability theory)与数理统计(mathematical statistics)是研究随机现象统计规律性的学科.概率论的主要特点是根据问题首先提出数学模型，然后去研究它们的性质、特征和内在规律性；数理统计的主要特点则是以概率论为基础，利用对随机现象的观察所取得的数据资料来研究数学模型.由于它们定量地刻画了事物发展变化的内在规律性，因而有广泛的应用价值.在工业、农业、交通运输、测量学、地质学、天文学、气象学、物理学、化学、电子技术、信号处理、自动化科学、生物学、医学、经济学，军事科学、管理科学以及许多新兴的高、精、尖技术中都能找到概率论与数理统计的应用背景.

1.1 随机事件

为了定量地把握随机现象内部隐藏的数量规律性,必须对随机现象进行观察、测量或进行科学试验.此类试验称为随机试验(random experiment),一般简称为试验(trial).

例 1.1.1 掷一枚正六面体的骰子,观察出现的点数.

例 1.1.2 抛一枚铝合金硬币和一枚铜质硬币,观察它们正、反面出现的情况.

例 1.1.3 从水泥自动生产流水线上任意抽取一袋水泥,称其质量.

例 1.1.4 一射手打靶,直到击中靶心为止,记录其射击次数.

以上例子有下述共同特点:

(1)可重复性,即在相同条件下试验可以重复进行;

(2)不确定性,即试验的可能结果不止一个(有限个或者无限个),但所有可能结果是预先可知的;

(3)不可预见性,即在试验结束之前不能确定哪一个结果会出现.

我们称具有上述特点的试验为随机试验,记为 E.

在研究随机试验 E 时首先必须弄清这个试验可能出现的所有结果,称每一个可能的结果为样本点(sample point),一般用小写字母 ω 表示.全体样本点构成的集合称为样本空间(sample space),一般用大写字母 Ω 表示.

在例 1.1.1 中,试验出现的所有结果共有 6 个,它们分别是:1 点,2 点,3 点,4 点,5 点和 6 点.将这些样本点简记为

$$\omega_i = i \text{ 点}, i = 1,2,3,4,5,6.$$

则样本空间可记为

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}.$$

在例 1.1.2 中,如果我们先观察铝合金硬币,再观察铜质硬币,则试验出现的所有结果共有 4 个,它们分别是:{正,正},{反,正},{正,反},{反,反}.若将这些样本点记为

$$\omega_1=\{正,正\},\quad \omega_2=\{反,正\},\quad \omega_3=\{正,反\},\quad \omega_4=\{反,反\},$$

则样本空间可记为

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}.$$

在例 1.1.3 中,若用 x 表示"一袋水泥的质量",则 x 的取值围绕一个值变化,若每袋水泥定额为 50 kg,自动流水生产线上最大偏差为 1 kg,则 x 取值范围为[49,51],即样本点有无限多个,它充满了区间[49,51],故样本空间可记为

$$\Omega=\{x:49\leqslant x\leqslant 51\}.$$

在例 1.1.4 中,若用 n 表示"击中目标所需的射击次数",则 n 取正整数 1,2,⋯.若将这些样本点记为

$$\omega_n=直到第\ n\ 次才击中目标, n=1,2,\cdots.$$

则样本空间可记为

$$\Omega=\{\omega_1,\omega_2,\cdots\}.$$

在例 1.1.1 及例 1.1.2 中,样本点的个数分别为 6 个和 4 个,是有限个.例 1.1.3 中样本点为无限多个,这无限多个不能一一列举出来,称它的样本点数为不可列无限多个.例 1.1.4 中样本点也为无限多个,这无限多个能够一一列举出来,称它的样本点的个数为可列无限多个.

1.1.1　随机事件

在一个随机试验中,可能发生也可能不发生的事情称为随机事件(random event),简称事件.每次试验一定发生的事情称为必然事件(certain event);每次试验一定不发生的事情称为不可能事件(impossible event).随机事件常用大写英文字母 A、B、C 等来表示;必然事件用 Ω 表示,不可能事件用 $\varnothing$ 表示.对于一个随机试验,它的每一个可能出现的结果都是一个事件,这种简单的随机事件称为基本事件(elementary event; fundamental event),即由一个样本点构成的集合.由若干个可能结果所组成的事件称之为复合事件(compound event),即若干个基本事件的并集.

在例 1.1.1 中,掷一枚骰子出现 5 点是一个随机事件,记为 $A=\{\omega_5\}$,它是一个基本事件;掷一枚骰子出现偶数点也是一个随机事件,

记为 B，由于偶数点有2点，4点，6点三种情况，所以 $B=\{\omega_2,\omega_4,\omega_6\}$，它不是基本事件，而是复合事件；如果讨论掷一枚骰子出现点数小于7的情形，则它不是随机事件，而是必然事件 Ω 了；如果说掷一枚骰子出现了10点，这是不可能的，也不是随机事件，而是不可能事件 $\varnothing$.

总之，随机事件 A 是样本空间 Ω 的非空真子集，因而是由样本点组成的集合．当且仅当 A 中某一个样本点出现时，称为事件 A 发生．为方便起见，我们通常将必然事件 Ω 和不可能事件 $\varnothing$ 作为随机事件的两个极端情形统一处理．

1.1.2 事件的关系与运算

事件是样本空间的子集，所以事件之间的关系与运算同集合之间的关系与运算完全一致，下面我们采用形象直观的文氏（Venn）图逐一介绍．

1 事件的包含

若事件 A 发生时，必然导致事件 B 发生，即 A 中的每一个样本点都包含在 B 中，则称事件 B 包含（contain）事件 A，或称事件 A 包含于事件 B，记为 $A\subset B$ 或 $B\supset A$（如图1-1-1）.

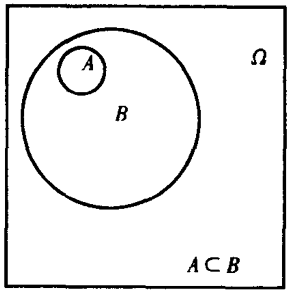

图 1-1-1

在例1.1.1中，掷一枚骰子出现1点或5点构成的事件记为 $A=\{\omega_1,\omega_5\}$；掷一枚骰子出现奇数点构成的事件记为 $B=\{\omega_1,\omega_3,\omega_5\}$；显然事件 B 包含事件 A，事件 A 发生一定导致事件记为 B 发生．

2 事件的相等

若事件 A 包含 B 且事件 B 包含 A，则称事件 A 与事件 B 相等或等价（equivalent），记为 $A=B$（如图1-1-2）.

在例1.1.1中，掷一枚骰子，事件“出现偶数点”$C=\{\omega_2,\omega_4,\omega_6\}$ 与

事件"出现能被 2 整除的点"$D=\{\omega_2,\omega_4,\omega_6\}$所包含的样本点完全相同，即 C 发生必导致 D 发生，D 发生也必导致 C 发生，故 $C=D$.

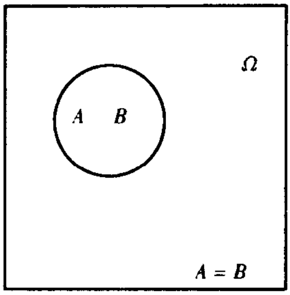

图 1-1-2

3　**事件的积**

若事件 A 与事件 B 同时发生，即由事件 A 与 B 的公共样本点组成的集合，称为事件 A 与 B 的积事件（intersection event），也称为事件 A 与事件 B 的交，记为 $A\cap B$ 或 AB（如图 1-1-3）.

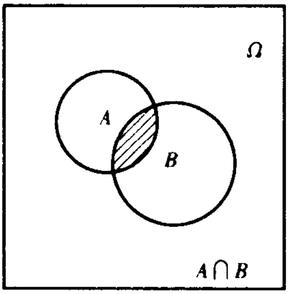

图 1-1-3

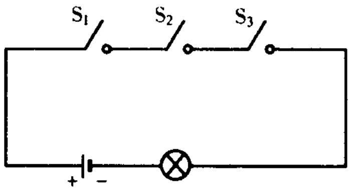

图 1-1-4

例 1.1.5　在图 1-1-4 所示的电路图中，记

$A_1=\{$开关 S_1 闭合$\}$，　$A_2=\{$开关 S_2 闭合$\}$，

$A_3=\{$开关 S_3 闭合$\}$，　$B=\{$灯亮$\}$.

用事件 A_1,A_2,A_3 表示事件 B.

解　由于开关 S_1,S_2,S_3 全部闭合时灯才亮，所以 $B=A_1A_2A_3$. 即 S_1 闭合且 S_2 闭合且 S_3 闭合.

4　**事件的互不相容**

若事件 A 与事件 B 不能同时发生，即所有包含在 A 中的样本点与包含在 B 中的样本点全不相同，则称事件 A 与事件 B 互不相容（exclusive），或称事件 A 与 B 互斥（如图 1-1-5）. 在这种情况下，$A\cap B=\varnothing$.

在例 1.1.1 中，掷一枚骰子出现 1 点和 5 点构成的事件 A 与出现偶数点构成的事件 C 是互不相容的.

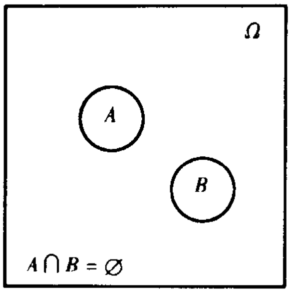

图 1-1-5

5 **事件的并**

若事件 A 与事件 B 至少有一个发生，即由 A 与 B 中的所有样本点组成的集合，称为事件 A 与 B 的并(union)，也称为事件 A 与事件 B 的和，记为 $A\cup B$(如图 1-1-6).

例 1.1.6 在图 1-1-7 所示的电路图中，记事件 A_1,A_2,A_3 及 B 含义同例 1.1.5，用事件 A_1,A_2,A_3 表示事件 B.

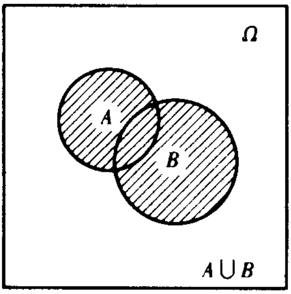

图 1-1-6

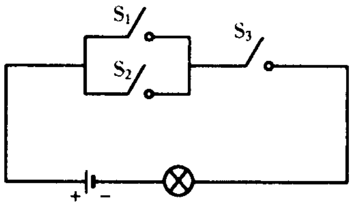

图 1-1-7

解 由于当开关 S_1 或 S_2 闭合，同时 S_3 也闭合时灯才亮，所以 $B=(A_1\cup A_2)A_3$.

事件的积与并都可以推广到有限个事件的情形. 如

$$\bigcap_{i=1}^{n}A_i=A_1\cap A_2\cap A_3\cap\cdots\cap A_n$$

表示事件 $A_1,A_2,\cdots,A_n$ 同时发生；

$$\bigcup_{i=1}^{n}A_i=A_1\cup A_2\cup\cdots\cup A_n$$

表示事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生.

事件的积与并还可以推广到可列多个事件的情形.如

$$\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots$$

表示 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生;

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots$$

表示 $A_1, A_2, \cdots, A_n, \cdots$ 至少有一个发生.

6　事件的对立

若事件 A 不发生,即样本空间 Ω 中所有不包含在 A 中的样本点全体所组成的集合,称为事件 A 的对立(complementation)事件,记为 $\bar{A}$(如图 1-1-8),对立事件有时也称为逆事件.

在例 1.1.1 中,掷一枚骰子出现奇数点构成的事件 B 和出现能被 2 整除的点构成的事件 D 是互为对立事件,即有 $\bar{B} = D$ 及 $\bar{D} = B$,且 $B \cup D = \Omega, B \cap D = \varnothing$.

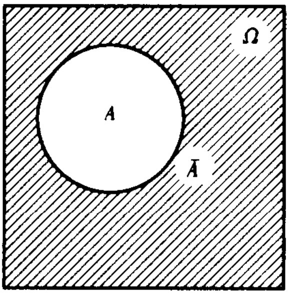

图 1-1-8

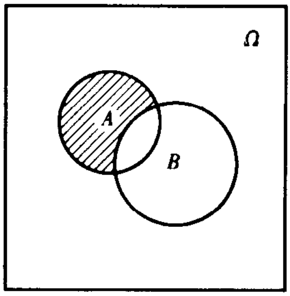

图 1-1-9

7　事件的差

若事件 A 发生而事件 B 不发生,即所有包含在 A 中而不包含在 B 中的样本点全体所组成的集合,称为事件 A 与 B 的差(difference),记为 $A - B$(如图 1-1-9).

在例 1.1.1 中,掷一枚骰子出现不超过 5 点这一事件 $F = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ 与出现偶数点构成的事件 $B = \{\omega_2, \omega_4, \omega_6\}$ 的差

$$F - B = \{\omega_1, \omega_3, \omega_5\}.$$

在进行事件运算时，一般是先进行逆的运算，再进行交的运算，最后再进行并或差的运算．如：$A\bar{B}=A-B$，$A\cup\bar{A}=\Omega$，$A\bar{A}=\varnothing$，$\bar{A}=\Omega-A$，$A\cup B=A\cup(B-A)=A\cup(B-AB)$．

事件运算的性质：

(1)交换律　$A\cup B=B\cup A$，$AB=BA$；

(2)结合律　$(A\cup B)\cup C=A\cup(B\cup C)$，$(AB)C=A(BC)$；

(3)分配律　$(A\cup B)\cap C=AC\cup BC$，

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

(4)幂等律　$A\cup A=A$，$AA=A$；

(5)对偶律　$\overline{A\cup B}=\bar{A}\bar{B}$，$\overline{AB}=\bar{A}\cup\bar{B}$．

对偶律也称德摩根(De Morgan)定理，且对于 n 个事件以至可列个事件，对偶律也成立．特别地，$A\cup\Omega=\Omega$，$A\Omega=A$，$A\cup\varnothing=A$，$A\varnothing=\varnothing$；当 $A\subset B$ 时，$A\cup B=B$，$AB=A$．

在概率论中，很重要的一点是学会用事件之间的关系来表示各种各样的复合事件，它们在求解某些概率问题时，往往能带来很大的方便．

例 1.1.7　设 A、B、C 是随机试验 E 中的三个随机事件，则

(1)事件"A 与 B 发生，C 不发生"可表示成 $AB\bar{C}$，也可以表示成 $AB-C$；

(2)事件"A、B、C 中至少有一个发生"可表示成 $A\cup B\cup C$，有时为运算方便也可表示成

$$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C\cup AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC\cup ABC;$$

(3)事件"A、B、C 中恰好有两个发生"可表示成

$$(AB\bar{C})\cup(A\bar{B}C)\cup(\bar{A}BC).$$

例 1.1.8　某人连续买了五期彩票，设 A_i 表示事件"第 i 期中奖"($i=1,2,3,4,5$)，试用 A_i 及其对立事件 $\bar{A}_i$ 表示下列事件：

(1)五期中至少有一期中奖；　(2)五期都中奖；

(3)五期中恰有一期中奖；　(4)五期都不中奖；

(5)五期中最多有一期中奖．

解　(1)$A=\{$五期中至少有一期中奖$\}$，则诸 A_i 之间是"或"的关

系，也可用五期均不中奖的对立事件来表示，所以

$$A = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 = \overline{\bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4 \bar{A}_5}.$$

(2) $B=\{$五期都中奖$\}$，则诸 A_i 之间是“且”的关系，故

$$B = A_1 A_2 A_3 A_4 A_5.$$

(3) $C=\{$五期中恰有一期中奖$\}$可以表示为第 1、2、3、4、5 期各期单独中奖，其他期未中奖，即

$$C = A_1\bar{A}_2\bar{A}_3\bar{A}_4\bar{A}_5 \cup \bar{A}_1 A_2 \bar{A}_3 \bar{A}_4 \bar{A}_5 \cup \bar{A}_1 \bar{A}_2 A_3 \bar{A}_4 \bar{A}_5 \cup \bar{A}_1 \bar{A}_2 \bar{A}_3 A_4 \bar{A}_5 \cup \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4 A_5.$$

(4) $D=\{$五期都不中奖$\} = \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4 \bar{A}_5$.

(5) $F=\{$五期中最多有一期中奖$\}$，则 F 表示五期中恰有一期中奖或五期都不中奖，即 $F = C \cup D$.

1.2　概率与频率

除必然事件和不可能事件这两种极端情形外，随机事件在一次试验中是否发生是不确定的，但随机事件发生的可能性还是有大小之别，所以人们希望知道某个事件在一次试验中发生的可能性究竟有多大，并希望用一个确定的数值来表示．我们引入“概率”这一概念正是对随机事件发生可能性大小的一种度量．为定量地分析和研究这种可能性的大小，首先引入频率的概念，它初步描述了事件发生的频繁程度．

1.2.1　频率

定义 1.2.1　在相同条件下，随机事件 A 在 n 次试验中发生了 m 次，则称比值

$$f_n(A) = \frac{m}{n}$$

为事件 A 在 n 次重复试验中发生的频率(frequency)．

显然，频率具有如下性质．

(1) 对任一事件 A，有 $0 \leqslant f_n(A) \leqslant 1$.

(2) $f_n(\Omega) = 1$，$f_n(\varnothing) = 0$.