应用概率统计

（第二版）

欧俊豪 王家生 徐漪萍 刘嘉焜 编

天津大学出版社

应用概率统计

（第二版）

欧俊豪　王家生　徐漪萍　刘嘉焜编

天津大学出版社

内 容 提 要

本书是在1990年8月出版的《应用概率统计》第一版的基础上修订而成的,内容包括随机事件与概率、随机变量及其概率分布、随机变量的数字特征、大数定律与中心极限定理、参数估计、假设检验、方差分析、回归分析、统计软件包SAS简介、马尔可夫过程和平稳随机过程.

本书文字流畅、内容适当、例题较多,书末附有习题答案,便于教学,可作为高等工业院校概率统计课教材,也可作为工程技术人员自学参考用.

应用概率统计(第二版)
欧俊豪　等编

责任编辑:杨秀雯
封面设计:谷英卉

出版发行:天津大学出版社(电话:022-27403647)
地　　址:天津市卫津路92号天津大学内(邮编:300072)
印　　刷:天津大学印刷厂
经　　销:新华书店天津发行所

开　　本:850mm×1168mm　1/32
印　　张:16
字　　数:426千
版　　次:1999年1月第2版
印　　次:1999年1月第6次
印　　数:28001~34000
书　　号:ISBN 7-5618-1109-8/O·103
定　　价:20.00元

前　言

本书是在1990年8月出版的第一版的基础上,总结天津大学近年来教学经验修订而成,可作为高等学校工科、理科(非数学专业)概率论与数理统计课程的教材,也可供工程技术人员参考.

本书内容分3部分.第一部分是概率论基础(第1章至第5章),主要内容有随机事件与概率、随机变量及其概率分布、随机变量的数字特征、大数定律与中心极限定理.第二部分是数理统计(第6章至第11章),主要内容有数理统计的基本概念、参数估计、假设检验、方差分析、回归分析、统计软件包SAS简介.第三部分是随机过程(第12章至第14章),主要内容有随机过程的基本概念、马尔可夫链、平稳随机过程.

为培养面向21世纪高层次工程技术人才,提高学生运用高科技手段解决随机问题的能力,本书在精练原有教材成熟部分的基础上,新增加了“随机过程”和“统计软件包SAS”两部分内容.SAS是目前国际最为著名的统计分析软件之一,具有技术先进、功能强大和使用方便的特点.此部分选用了本书的例题作为实例详细介绍了SAS的使用方法.

我们在编写本书时,自始至终注意说明各概念的现实背景和实际意义,对基本原理和方法,除详尽严谨分析外,为便于自学,在叙述上力求通俗易懂,深入浅出.本书还收入大量典型例题,各章配有习题,书末附有全部习题答案及有关概率统计用表.

本书由欧俊豪担任主编.全书各章分别由王家生(第1、5、6、7章),欧俊豪(第2、4、10章),徐漪萍(第3、8、9章),马逢时(第11章),刘嘉焜(第12、13、14章)执笔.

本书是在马逢时教授直接领导下编写而成的,马逢时教授审

阅了全书，并亲自编写了统计软件包 SAS 一章.

全书所需教学时数为 80 学时，其中概率论基础部分为 30 学时，数理统计部分为 34 学时，随机过程部分为 16 学时.

书中不妥之处，恳请读者批评指正.

编者 1998 年 3 月

目　录

第 1 章　随机事件与概率

引　言

在自然界和人类社会活动中经常会遇到各种各样的现象.这些现象大体上可分为两类,一类是确定性现象,另一类是偶然现象,亦称为随机现象.例如,在标准大气压下,水加热到100℃必然沸腾;同性电荷互相排斥,异性电荷互相吸引;向上抛一物体必然下落等等.这类在一定条件下,必然出现某种结果的现象称为**确定性现象**.然而在自然界和人类社会中还有许多现象,它们在一定条件下可能出现这样的结果,也可能出现那样的结果,且不能事先断定出现哪种结果.例如,往桌上掷一硬币,可能正面朝上,也可能反面朝上,而且在掷之前不能预言一定哪一面朝上;检查生产流水线上的一件产品,可能是合格品也可能是不合格品;打靶射击,尽管经过瞄准,弹着点却可能在靶心附近的各个位置等等.这类在一定条件下可能出现的结果不止一个,至于出现哪一个事先无法确定的现象称为**随机现象**.

概率论与数理统计的任务就是要揭示随机现象内部存在的统计规律性.概率论的特点是根据问题先提出数学模型,然后去研究它们的性质、特征和规律性;数理统计则是以概率论的理论为基础,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.

作为数学的一个重要分支,概率论与数理统计大体上17世纪中叶开始形成.在17世纪研究概率论的先驱中,最著名的有惠更斯、巴斯卡、费尔马和J.贝努里等人,后继者中不乏历代的大数学家和科学家.当时,由于赌徒们所提出的一些还未能归入数学范围

的问题,引起了巴斯卡和费尔马的通信讨论,就在这里面逐渐结晶出了概率及数学期望等重要概念.当时研究的模型较简单,就是现在统称的古典概型.

其后,随着生产实践的发展,特别是在射击理论、人寿保险、测量误差等工作中提出的一些概率问题,促使人们在概率论的极限定理方面进行深入研究.起初主要对贝努里试验概型进行研究,其后则推广到更为一般的场合.极限定理的研究在18世纪和19世纪整整200年中成了概率论研究的中心课题.在本世纪初,由于新的更有力的数学方法的引入,这些问题才得到了较好的解决.

虽然概率论的历史悠久,但它的严格的数学基础的建立及理论研究与实际应用的极大发展却主要是本世纪的事情.1933年前苏联著名数学家柯尔莫哥洛夫建立了概率论的公理化体系.这一体系的建立标志着概率论已经成为一门成熟的数学学科.

由于物理学(如统计物理)、生物学及工程技术(如自动电话、无线电技术)发展的推动,概率论与数理统计得到了飞速的发展.理论课题不断扩大与深入,概率统计的思想渗入各个自然科学学科成为现代科学发展的明显标志之一.目前,概率统计在工业、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、通信技术、自动化科学、生物学、医学、经济学、军事科学以及各尖端技术中获得了越来越广泛的应用.概率论与数理统计已经成为最活跃最重要的数学学科之一.

§1.1 样本空间与随机事件

1.1.1 **随机试验**

为了研究随机现象内部存在的数量规律性,必须对随机现象进行观察或试验.今后我们把对随机现象所进行的观察或试验统

称为试验.

例 1 抛一硬币,观察正、反面出现的情况.

例 2 掷一枚骰子,观察出现的点数.

例 3 把一硬币连抛两次,观察正、反面出现的情况.

例 4 一射手进行射击,直到击中目标为止,记录射击次数.

例 5 在同一生产条件下生产的一种电子元件,任意抽取一件测试其寿命.

上面列举的 5 个试验的例子,有以下共同特点.

(1) 试验可以在相同条件下重复进行;

(2) 试验的可能结果不止一个,并且所有可能结果是预先知道的;

(3) 进行一次试验之前不能确定哪一个结果会出现.

我们把具有以上 3 个特点的试验称为**随机试验**,简称试验,记为 E.

1.1.2 样本空间

在一个随机试验 E 中,试验的所有可能结果组成的集合称为随机试验 E 的**样本空间**,通常用字母 Ω 表示. Ω 中的元素,称作样本点,常用 ω 表示.

在上述例 1 中,试验的所有可能结果有两个:正(抛得正面朝上),反(抛得反面朝上),因此样本空间 $\Omega=\{$正,反$\}$. 若记

$$\omega_1=\text{正},\omega_2=\text{反},$$

则样本空间可记为

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

在例 2 中,试验的所有可能结果有 6 个:1 点,2 点,…,6 点. 若记

$$\omega_i=i\text{ 点}\quad,i=1,2,\cdots,6.$$

则样本空间可记为

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}.$$

在例 3 中试验的所有可能结果有 4 个:(正,正),(反,反),(正,反),(反,正).这里记号(正,反)表示"第一次抛得正面,第二次抛得反面"这一结果,其余类似.因此若记

$$\omega_1=(\text{正},\text{正}),\quad \omega_2=(\text{反},\text{反}),$$
$$\omega_3=(\text{正},\text{反}),\quad \omega_4=(\text{反},\text{正}).$$

则样本空间可记为

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

在例 4 中,若用 n 表示"击中目标所需要的射击次数为 n"这一结果,$n=1,2,\cdots$,则样本空间可记为

$$\Omega=\{1,2,3,\cdots\}.$$

在例 5 中,若用 x 表示"电子元件的使用寿命为 x 小时"这一结果,$0\leqslant x<+\infty$,则样本空间为

$$\Omega=\{x:\ 0\leqslant x<+\infty\}.$$

上述例 1、例 2、例 3 各随机试验的样本空间都只有有限个样本点.例 4 的样本空间含有无穷多个样本点,但这些样本点可以依照某种次序排列出来,我们称它的样本点数为可列无穷多个.例 5 的样本空间也含有无穷多个样本点,但它们充满区间$[0,+\infty)$,此时,我们称它的样本点数为不可列无穷多个.

1.1.3 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**.随机事件常用大写英文字母 $A,B,C,\cdots$来表示.如在例 1 中,抛一枚硬币,"正面朝上";例 2 中,掷一枚骰子,"出现的点数小于 3";例 3 中,一枚硬币连续抛两次,"两次都抛得正面朝上"、"仅有一次抛得正面朝上"、"至少有一次抛得正面朝上";例 4 中,"射击次数是 5";例 5 中,"电子元件寿命为 1 000 小时"、"电子元件寿命不超过 2 000 小时"等等,对一次试验而言,它们可能发生,也

可能不发生，因此都是随机事件．这些随机事件可分别记为

$A=\{$正面朝上$\}$；

$B=\{$点数小于 3$\}$；

$C_1=\{$两次都抛得正面朝上$\}$；

$C_2=\{$仅有一次抛得正面朝上$\}$；

$C_3=\{$至少有一次抛得正面朝上$\}$；

$D=\{$射击次数是 5$\}$；

$E_1=\{$元件寿命为 1 000 小时$\}$；

$E_2=\{$元件寿命不超过 2 000 小时$\}$．

对于一个随机试验来说，它的每一个可能结果，显然都是一个随机事件，它们是随机试验中最简单的随机事件，称为**基本事件**．如上述的事件 A,C_1,D,E_1 都是相应随机试验中的基本事件．

在一个随机试验中，除了基本事件外，还有由若干个可能结果所组成的事件，相对于基本事件，称这种事件为**复合事件**．例如，上述随机事件 B,C_2,C_3,E_2 等都是复合事件．

随机试验中的事件，在引入样本空间的概念之后，可以用样本空间 Ω 的子集来表示．例如前面所列举的事件就可以用样本空间的子集表示如下：$A=\{$正$\}$，$B=\{1,2\}$，$C_1=\{($正，正$)\}$，$C_2=\{($正，反$)$，$($反，正$)\}$，$C_3=\{($正，正$)$，$($正，反$)$，$($反，正$)\}$，$D=\{5\}$，$E_1=\{1\ 000\}$，$E_2=\{x:\ 0\leqslant x\leqslant 2\ 000\}$．

由此可见，当一个随机事件 B 用样本空间 Ω 的子集来表示时，B 就是样本点的集合．**当且仅当 B 中某一个样本点出现，称事件 B 发生**．对于基本事件来说，由于它就是某个样本点，所以用以这个样本点为元素的单点集来表示它．对于复合事件来说，由于它是由若干个样本点所组成的，所以用以这若干个样本点为元素的集合来表示它．

在随机试验中，每次试验一定发生的事情称为**必然事件**；每次

试验一定不发生的事情称为**不可能事件**.必然事件用 Ω 表示,不可能事件用 $\varnothing$ 表示.这是因为样本空间 Ω 包含所有的样本点,它是 Ω 自身的子集,在每次试验中,必然有 Ω 中的某一个样本点出现,所以事件 Ω 在每次试验中一定发生,故 Ω 是必然事件.又因为在每次试验中,不可能有 $\varnothing$ 中的样本点出现($\varnothing$ 为空集不含样本点),所以事件 $\varnothing$ 在每次试验中一定不发生,故 $\varnothing$ 是不可能事件.

必然事件 Ω 和不可能事件 $\varnothing$ 本质上不是随机事件.为今后研究问题方便,我们把必然事件和不可能事件作为随机事件的两个极端情形统一处理.

1.1.4 事件的关系和运算

由于事件定义为样本空间的某个子集,因此事件之间的关系与运算和集合论中集合之间的关系与运算是一致的.对应着集合的关系与运算,定义事件的关系与运算如下:

若事件 A 发生时,必然导致事件 B 发生,则称事件 B **包含**事件 A,记为 $A\subset B$,或 $B\supset A$(图 1-1).

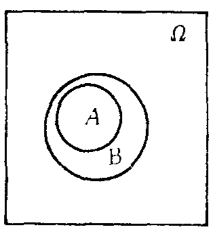

图 1-1 $A\subset B$

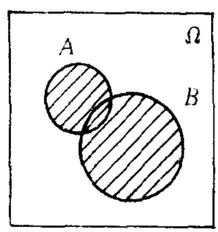

图 1-2 $A\cup B$

若事件 B 包含事件 A,并且事件 A 也包含事件 B,即有 $A\subset B$,且 $B\subset A$,则称事件 A 与事件 B **相等**,记为 $A=B$.

表示事件 A 与事件 B 中至少有一个发生的事件,称为事件 A 与事件 B 的**和事件**,亦称为事件 A 与 B 的**并**,记为 $A\cup B$(图 1-2).

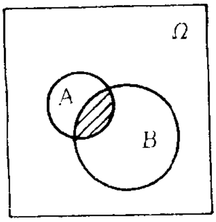

图 1-3　$A\cap B$

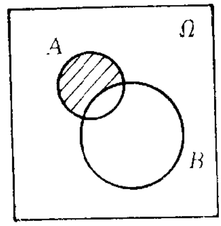

图 1-4　$A-B$

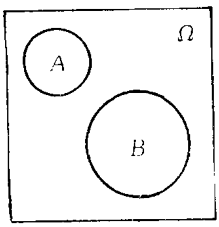

图 1-5　$AB=\varnothing$

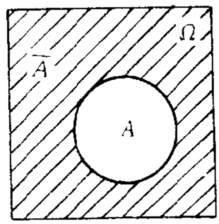

图 1-6　$\overline{A}$

关于事件的并,可以推广到有限个甚至无穷多个事件的情形.

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

表示事件 $A_1, A_2, \cdots, A_n$ 至少有一个发生.

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

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

表示事件 A 与 B 同时发生的事件称为 A 与 B 的**积事件**,亦称为事件 A 与事件 B 的**交**,记为 $A\cap B$ 或 AB(图 1-3).

事件的交,可以推广到有限个甚至无穷多个的情形.

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

表示事件 $A_1, A_2, \cdots, A_n$ 同时发生.

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

表示事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生.

表示事件 A 发生而事件 B 不发生的事件，称为事件 A 与 B 的**差事件**，记为 $A-B$(图 1-4).

若事件 A 与事件 B 不能同时发生，即 $AB=\varnothing$，则称事件 A 与 B 是**互不相容事件**(或称是**互斥事件**)(见图 1-5).

若事件 A 与 B 满足：$A\cup B=\Omega$，且 $AB=\varnothing$，则称事件 A 与 B 互为**逆事件**，亦称事件 A 与 B 互为**对立事件**. 这是指对每次试验来说，事件 A、B 中必有一个发生，且仅有一个发生. A 的对立事件记为 $\bar{A}$，则 $\bar{A}=\Omega-A=B$. 必然事件 Ω 与不可能事件 $\varnothing$ 显然互为对立事件(图 1-6).

在进行事件运算时，经常要用到下述运算规律.

(1) 关于事件和(并)的运算规律

$A\cup B=B\cup A$； (交换律)

$A\cup(B\cup C)=(A\cup B)\cup C$； (结合律)

$A\cup A=A$； (幂等律)

$A\cup\Omega=\Omega$；

$A\cup\varnothing=A$；

若 $A\subset B$，则 $A\cup B=B$.

(2) 关于事件积(交)的运算规律

$AB=BA$； (交换律)

$A(BC)=(AB)C$； (结合律)

$AA=A$； (幂等律)

$A\bar{A}=\varnothing$；

$A\Omega=A$；

$A\varnothing=\varnothing$；

若 $A\subset B$，则 $AB=A$.

(3) 关于事件积与事件和的混合运算规律

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

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