新世纪高校统计学专业系列教材

应用概率统计

YINGYONG GAILÜ TONGJI

王学民 编著

上海财经大学出版社

新世纪高校统计学专业系列教材

应用概率统计

王学民　编著

上海财经大学出版社

图书在版编目(CIP)数据

应用概率统计/王学民编著．—上海:上海财经大学出版社,2005.10
新世纪高校统计学专业系列教材
ISBN 7-81098-495-0/F·448

Ⅰ.应…　Ⅱ.王…　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(2005)第 106780 号

□ 责任编辑　何苏湘
□ 封面设计　优典工作室

YINGYONG GAILÜ TONGJI
应用概率统计

王学民　编著

上海财经大学出版社出版发行
(上海市武东路 321 号乙　邮编 200434)
网　址:http://www.sufep.com
电子邮箱:webmaster@sufep.com
全国新华书店经销
上海崇明裕安印刷厂印刷装订
2005 年 10 月第 1 版　2005 年 10 月第 1 次印刷

890mm×1240mm　1/32　13.5 印张　375 千字
印数:0 001—4 000　定价:25.00 元

前　言

本书可作为财经类、工科类院校本科生“概率论与数理统计”课的教材或教学参考书。本教材编写得较为细致，前后条理清楚，便于自学，因此也可作为一本自学教材。

本教材具有如下一些特点：

(1) 全书在内容的选择上兼顾知识的基础性和实用性，在有限的篇幅和读者有限的学习时间内尽可能地安排对读者更有学习价值的内容，特别是尽量避开数学上较有难度而实用价值又不大的部分。

(2) 本书对读者的数学基础要求不高，在此基础上对概率统计的基础理论、思想方法进行了严谨的论述，同时又突出和展现了概率统计的应用，使读者既能为今后进一步学习数理统计的后继课程打好理论基础，又能切实地学会概率统计的初步应用。

(3) 从第五章开始，本书在每一章后的第一个附录中陆续介绍了SAS9.0(简体中文)软件的菜单系统应用。对这部分内容，即使读者不曾接触过 SAS 软件，一般也能看懂并学会操作。这对读者今后进一步学习 SAS 软件是很有益处的。在介绍 SAS 软件的使用时所涉及的数理统计知识有部分超出了本书正文的范围，这些正文之外的内容在提到时都作了简要的介绍，它们都极具实用价值，可以作为对本书正文内容的重要补充，任课教师可酌情选讲。

(4) 从实用的角度处理了大数定律和中心极限定理的内容。大数定律被安排在了估计量的一致性(见第六章的 6.2.3 部分)中论述。中心极限定理的两个最基本、最常用的应用是，大样本情形下样本均值抽样分布的正态近似和 n 很大时二项分布的正态近似。这两个问题分别

被安排在了第五章的5.2.3部分和第二章的2.3.3部分。这样的安排既降低了数学难度,又突出了知识的应用性。

(5) 本书在数理统计内容开始的部分适当地介绍了有限总体的概念和基于有限总体的抽样、统计量的性质及其抽样分布。在实际应用中,人们经常会遇到有限总体的问题,因此读者了解和熟悉有限总体很有必要。

全书共分十章,前四章是属于概率论的范畴,后六章是属于数理统计的范畴。书中例题和习题的数据(SAS数据文件形式和Excel数据文件形式)读者可以在作者的网页上下载,网址是 http://iclass.shufe.edu.cn/teacherweb/users/wxuemin/。

本书对我国目前"概率论与数理统计"课程的教学改革作了一种新的探索,难免会有不妥之处,敬请读者批评指正,以便今后不断地修正、提高和完善。

王学民
2005年6月

目　录

第一章 随机事件与概率

在日常生活中，我们经常会遇到不确定性的问题，并且希望了解这种不确定性的程度。以下是这方面的一些例子。

(1) 明天下雨的“可能性”有多大？

(2) 股价在未来一周内上涨的“可能”有多大？

(3) 如果提高产品的价格，则销售量下降的“机率”有多少？

(4) 一个家庭中的所有四个孩子都是女孩的“机会”是多少？

(5) 某候选人被当选的“概率”是多大？

所有上述问题中的“可能性”、“机率”和“机会”等都是一个意思，今后我们一般就用规范的术语“概率”来表述。

概率简单地说就是一个数，它总是从 0 到 1 之间取值。小概率(接近 0)的事件很少发生，而大概率(接近 1)的事件则经常发生。例如，某一城市一年内发生七级以上地震的概率很小，而整个地球一年内发生七级以上地震的概率就很大了。

早在 17 世纪，就有零星的有关概率的文章出现。当时的绅士赌徒们试图确定牌和骰子赌博中的赔率，因而对概率发生了兴趣。赌博者希望小概率事件能有较高的赔率，而较高概率的事件则应有较小的赔率。赔率应与事件的概率相对应，这样才能做到公平，也就是投注者既不应最后轻易破产，也不应经常性地有过多的收益。这个问题被提到了当时的数学家面前，这就是概率论的起源。

本章中，我们将介绍随机事件和概率的一些基本概念，这部分内容是后面章节的基础，“概率”的说法将贯穿全书。

§1.1 随机事件及其运算

1.1.1 随机试验与随机事件

在一定条件下，至少有两个可能结果，但不能预先断定出现哪一个结果的现象，称为**随机现象**(random phenomenon)。例如，抛一枚硬币、观察某一地区一天内的交通事故数和随机抽取一批产品所出现的废品率等，出现的结果都具有随机性，都是随机现象。

很多随机现象是可以大量重复的，对这种可重复的随机现象的观察称为**随机试验**(random experiment)，简称**试验**[①]。并规定随机试验必须符合以下条件：

(1) 它可以在相同的条件下重复进行；

(2) 试验的所有可能的结果是事先已知的，并且不止一个；

(3) 每次试验只出现这些可能结果中的一个，但不能预先断定会出现哪个结果。

试验的每一个可能的基本结果称为**样本点**(sample point)，记为 ω。所有样本点的集合称为**样本空间**(sample space)，记为 Ω。从集合论的角度来看，ω 是 Ω 的一个元素，即 $\omega\in\Omega$。

例 1.1.1 掷一颗骰子，试验结果定义为骰子朝上一面的数字。$1,2,\cdots,6$ 为样本点，样本空间 $\Omega=\{1,2,\cdots,6\}$。

例 1.1.2 抽检一件产品，如果把产品分为合格与不合格两类，则可取样本空间 $\Omega_1=\{$合格品，不合格品$\}$；如果把产品分为四个等级，则可取样本空间 $\Omega_2=\{$一等品，二等品，三等品，废品$\}$；如果把所有的产品都区别看待，则样本空间 $\Omega_3=\{$全体产品$\}$。

① 统计学中试验的概念与物理学中试验的概念是稍有不同的。当物理试验在相同的条件下重复进行时，会产生相同的试验结果。而在统计学中，即使试验在相同的条件下重复，结果也是随机确定的，可以得到完全不同的结果。

例 1.1.3　一天内进入某超市的顾客数，其样本空间 $\Omega=\{0,1,2,\cdots\}$。

实际上，上例中的顾客数总有一个上限，不可能有无限多。之所以取成无限是基于这样几点考虑：(1) 顾客数的上限很难确定；(2) 数学处理方便；(3) 顾客数超过一个足够大数目的概率非常小，甚至为零，因而几乎不影响所做的分析。

例 1.1.4　观察电视机的寿命，其样本空间 $\Omega=\{t: t\geqslant 0\}$。

若干个样本点的集合称为**随机事件**(random event)，简称**事件**(event)，记为大写字母 A、B、C 等。可见，事件是样本空间 Ω 的子集。在试验中，如果出现 A 中所包含的一个样本点，则称**事件 A 发生**。随机事件有两种极端的情况，一种是必然会出现的结果，称为**必然事件**(certain event)；另一种是不可能出现的结果，称为**不可能事件**(impossible event)。从样本空间来看，必然事件是由其全部样本点组成的，故可记为 Ω；而不可能事件则不含有任何样本点，故记为空集 $\varnothing$。必然事件和不可能事件实质上都是确定性现象，已经失去了“随机性”。但为方便起见，仍将它们视作随机事件。

例 1.1.5　在例 1.1.1 中，我们研究事件

$$A=\{出现的点数不超过 4\}$$
$$B=\{出现的点数是偶数\}$$
$$C=\{出现的点数不超过 6\}$$
$$D=\{出现的点数为 7\}$$

这些用文字表述的事件也可以表示为样本点的集合，即 $A=\{1,2,3,4\}$，$B=\{2,4,6\}$，$C=\{1,2,3,4,5,6\}=\Omega$，$D=\varnothing$。C 是必然事件，D 是不可能事件。若在试验中出现的结果是 3，则事件 A 发生，事件 B 不发生。

1.1.2　事件间的关系及运算

由于事件可看作是一种集合，因此事件间的关系及运算与集合间的关系及运算是一致的。

随机事件和它们之间的关系可以用直观的几何图形来表示，这类图形称为**文氏图**(Venn diagram)。

(1) 若事件 A 的每一个样本点都包含在事件 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 $B \supset A$ 或 $A \subset B$，如图 1.1.1 所示。$A \subset B$ 表明事件 A 发生必导致事件 B 发生。如果 $A \supset B$ 且 $A \subset B$，则称**事件 A 与 B 相等**，记为 $A=B$，表明事件 A 与 B 是同一个事件。例如，掷一颗骰子，事件"出现奇数点"与事件"不出现偶数点"是相同的。

(2) 由或属于事件 A，或属于事件 B 的一切样本点组成的事件称为**事件 A 与 B 的并**(或和，union)，记为 $A \cup B$，如图 1.1.2 所示。$A \cup B$ 表明事件 A 与 B 中至少发生一个。例 1.1.5 中，$A \cup B=\{1,2,3,4,6\}$。

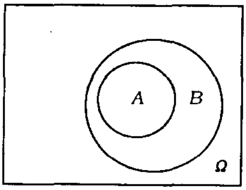

图 1.1.1 $A \subset B$

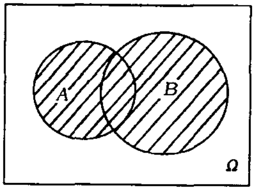

图 1.1.2 $A \cup B$

(3) 由事件 A 与事件 B 中共同的样本点组成的事件称为**事件 A 与 B 的交**(或积，intersection)，记为 $A \cap B$ 或 AB，如图 1.1.3 所示。$A \cap B$ 表明事件 A 与 B 同时发生。例 1.1.5 中，$A \cap B=\{2,4\}$。

(4) 由一切属于事件 A，但不属于事件 B 的样本点组成的事件，称为**事件 A 与 B 的差**(difference)，记为 $A-B$，如图 1.1.4 所示。$A-B$ 表明事件 A 发生而事件 B 不发生。例 1.1.5 中，$A-B=\{1,3\}$。

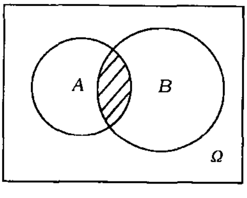

图 1.1.3 AB

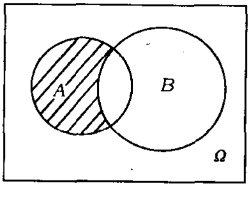

图 1.1.4 $A-B$

(5) 由样本空间中一切不属于事件 A 的样本点所组成的事件称为**事件 A 的对立事件**(complementary events),记为 $\overline{A}$,如图 1.1.5 所示。$\overline{A}$ 表明事件 A 不发生。例 1.1.5 中,$\overline{A}=\{5,6\}$,$\overline{B}=\{1,3,5\}$。$\overline{A}$ 的对立事件就是原来的事件 A,即 $\overline{\overline{A}}=A$。所以对立事件是相互的,必然事件 Ω 与不可能事件 $\varnothing$ 互为对立事件。

(6) 若事件 A 与事件 B 没有共同的样本点,即 $AB=\varnothing$,则称**事件 A 与 B 互不相容**(exclusive),如图 1.1.6 所示。$AB=\varnothing$ 表明事件 A 与 B 不可能同时发生。

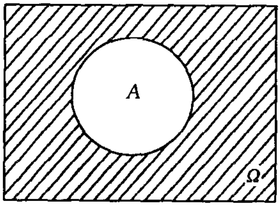

图 1.1.5　$\overline{A}$

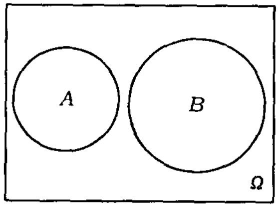

图 1.1.6　$AB=\varnothing$

事件并、交的概念可以推广到 n 个事件的情形。$\bigcup\limits_{i=1}^{n}A_i$ 称为**事件 $A_1,A_2,\cdots,A_n$ 的并**,表明事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生。$\bigcap\limits_{i=1}^{n}A_i$ 称为**事件 $A_1,A_2,\cdots,A_n$ 的交**,表明事件 $A_1,A_2,\cdots,A_n$ 同时发生。

例 1.1.6　设 A、B、C 是 Ω 中的随机事件,则

(1) "A 与 B 不发生,C 发生"可以表示成 $\overline{A}\,\overline{B}C$;

(2) "A、B、C 恰好发生一个"可以表示成 $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$;

(3) "A、B、C 中至少发生两个"可以表示成 $AB\cup AC\cup BC$;

(4) "A、B、C 中至多两个事件发生"可以表示成 $\overline{A}\,\overline{B}\,\overline{C}\cup A\,\overline{B}\,\overline{C}\cup\overline{A}\,B\,\overline{C}\cup\overline{A}\,\overline{B}\,C\cup A\,B\,\overline{C}\cup A\,\overline{B}\,C\cup\overline{A}B\,C$,这个事件的对立事件是"$A$、$B$、$C$ 三个事件皆发生",因此原事件亦可以表示成 $\Omega-ABC$。

例 1.1.7　事件 $A\cup B$ 可以写成互不相容事件的并,即有

$$A\cup B=A\cup(B-A)=A\cup(B-AB)$$