GAILULUN
YU SHULITONGJI

主编\彭美云

概率论与数理统计

GAILULUN YU SHULITONGJI

武汉大学出版社

概率论与数理统计

主　编　彭美云
副主编　程学光　王文祥　钟六一

武汉大学出版社

图书在版编目(CIP)数据

概率论与数理统计/彭美云主编. —武汉：武汉大学出版社，2001.2

ISBN 7-307-03475-1

Ⅰ.概… Ⅱ.彭… Ⅲ.①概率论 ②数理统计 Ⅳ.O21

中国版本图书馆 CIP 数据核字(2000)第 47277 号

责任编辑：龙方明

出版发行：武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件：wdp4@whu.edu.cn 网址：www.wdp.whu.edu.cn)

印刷：武汉市汉桥印刷厂

开本：850×1168 1/32 印张：8.375 字数：212 千字

版次：2001 年 2 月第 1 版 2002 年 12 月第 3 次印刷

ISBN 7-307-03475-1/O·258 定价：12.00 元

前 言

《概率论与数理统计》是高等学校(工科类)本科生必修的数学课程之一.该课程是从数量侧面研究大量随机现象统计规律性的一门应用性较强的数学基础课.近几十年来,伴随着计算机技术的飞速发展,使概率论与数理统计的应用达到了前所未有的规模,其基本理论与方法几乎渗透到自然科学、社会科学、工农业生产、文化教育、医药卫生、经营管理等各个领域之中.《概率论与数理统计》是数学领域中一门颇有特色的分支学科,也是一门既有学习难度又能引起学生学习兴趣的数学课程.

本书是根据高等学校工科数学课程教学指导委员会于1993年审订的《概率论与数理统计课教学基本要求》(以下简称《基本要求》)及原自编教材基础上编写而成.本书的特色之一是内容详略得当,叙述深入浅出,教学中难点处理妥善,例题习题选取兼顾了基础与提高、理论与应用.特色之二是加强了数理统计部分的内容,研制了相应的多媒体教学软件、计算机辅助教学软件(CAI)及数理统计实验课软件并已用于教学中.实验课软件包括:用直方图、比例图、频数分布表等处理数据的方法,参数区间估计,参数假设检验,χ^2拟合检验,单因素试验的方差分析,一元线性回归与二元线性回归等.实践证明,引进多媒体教学与开设数理统计实验课,既可提高学生学习热情与兴趣,又可培养及增强学生的数据处理和分析应用能力.课堂讲授和计算机操作的有机结合,体现了数学教学的创新,对提高学生的数学素质,拓宽学生的视野与思路,都有明显的效果.

讲授本书内容,按《基本要求》定为48课时,其分配如下:第一章6课时,第二章7课时,第三章6课时,第四章6课时,第五章2课时,第六章3课时,第七章6课时,第八章6课时,第九章与第十章共6课时.此外,数理统计实验课共8课时,一般用课外时间.因课时等原因,有些内容(已在有关章节加注了*号)可作为学生自学或深入学习之用.

本书由彭美云主编,程学光、王文祥、钟六一任副主编.其中王文祥编写第一、第四、第五章;彭美云编写第二、第三、第八章;程学光编写第六、第七章;钟六一编写第九、第十章。全书由彭美云执笔修改定稿.

本书可供高等工科院校本科生使用,亦可作为函授、电大、夜大等学生的教材.对应用统计工作者来说,也有很好的参考价值.

限于编者水平,书中不当之处,欢迎批评指正.

编者

1999年3月15日

目　录

第一章　随机事件与概率

日常生活及科学研究活动中人们经常遇到各种各样的现象，例如：

①水在标准气压下加温到100℃变成水蒸汽；

②氢气在空气中燃烧可得到水；

③某工厂用同一种工艺条件生产出一产品的寿命；

④掷10次硬币，出现正面的次数；

⑤用同一仪器测量某项指标，测量所得的数据．等等．

上面提及的是本质完全不同的两类现象．①②是在一定条件下必然发生的现象，这类现象称为确定性现象．而③④⑤这类现象有一个共同的特点就是在基本条件不变的情况下，一系列试验或观察事先无法确定会出现什么结果．时而出现这种结果，时而出现那种结果，这类现象称为随机现象．概率论与数理统计就是研究随机现象统计规律性的一门数学学科，它研究隐藏在现象的偶然性之中的规律性．随着科学技术的发展，这门学科已被越来越广泛地应用到自然科学、技术科学及社会科学的各有关领域中．

第一节　随机试验、样本空间与事件

一、随机试验

在大量的随机现象中，我们将只考虑能够在相同条件下重复试验或观察的随机现象．为了方便起见，对随机现象进行一次观察

就称为一次随机试验，概括地说随机试验是指具有下列几个特点的试验或观察：

①在相同的条件下可任意次重复进行；

②试验可能的结果不止一个，预先不能断定某一结果将在一次试验中出现，但能对试验的一切可能结果进行描述.

特点①是研究规律性的客观前提，特点②表明能对试验的整个情况进行分析.

例 1 观察灯炮厂生产出的一批灯炮的寿命，这就是一个随机试验.

例 2 从装有 10 只红球 10 只白球的盒子中任意摸出 5 只，记录其中的红球数，这也是一个随机试验.

二、样本空间

随机试验 E 的所有可能的结果的全体称为 E 的样本空间，用 S 表示. 样本空间是一个集合，集合中的元素就是试验时可能发生的各种结果.

例 3 掷一枚硬币观察出现什么面的试验中，样本空间 $S=$ {正面，反面}.

例 4 记录某地的气温，样本空间 $S=[a,b]$. 其中 a,b 分别表示该地的最高气温和最低气温.

例 5 记录某电话交换台在某时间内收到的呼叫次数，样本空间 $S=\{0,1,2,\cdots\}$.

对于一个具体的实际问题需要选取一个恰当的样本空间来描述，而在概率论中，一个样本空间形式上可以概括各种实际内容大不相同的问题. 例如只包括两个不同结果的样本空间既可作为掷硬币出现正反面的模型，也能用于产品检验中出现合格品及次品等等. 尽管问题的实际内容不同，但有时却能归结为相同的概率模型，它能使问题的本质更突出，使研究更方便，这是模型的特点，以

后将介绍两个重要的模型(古典概型和贝努利试验概型).

三、随机事件

随机事件是用来描述随机试验所关心的某些可能产生的结果的概念,由于所有可能结果构成样本空间 S,从构成方式看,某些可能结果当然是 S 的子集,因此,随机试验 E 的样本空间 S 的子集称为试验 E 的随机事件,简称事件.

例 6 在掷一次骰子的试验中,若规定 e_i 表示出现 i 点,$i=1,2,\cdots,6$. 则样本空间 $S=\{e_1,e_2,e_3,e_4,e_5,e_6\}$.

$A=\{e_1,e_2,e_3\}$ 表示出现奇数点的事件.

$B=\{e_4,e_5,e_6\}$ 表示出现偶数点的事件.

由定义可知,样本空间 S 中的单个元素(称为样本点)也构成事件,它们称为基本事件. 样本空间 S 本身亦为事件,由于在任一次试验中必然发生,称 S 为必然事件. 相应地空集 $\varnothing$ 亦是 S 的子集,在任一试验中总不会发生,称其为不可能事件. 必然事件和不可能事件发生与否已经失去了“随机性”,规定这种极端情形为事件将会给以后的讨论带来方便.

以后称某事件 A 发生指的是在试验中出现了含在 A 中的某个元素.

四、事件的运算与关系

如果事件 A 中的每一元素都是事件 B 的元素,即作为集合来讲 A 是 B 的子集,则称事件 B 包含事件 A,记作 $A\subset B$ 或 $B\supset A$.

如果 $A\supset B$,$B\supset A$,则称 A 与 B 相等,记作 $A=B$.

如果 e 是集合 A 中的元素,记为 $e\in A$. 用 $e\overline{\in} B$ 表示 e 不是 B 中的元素.

集合 $\{e\mid e\in A$,或 $e\in B\}$ 称为事件 A 与事件 B 之和,记作 $A\cup B$,其实际意义表示在试验中或者 A 发生或者 B 发生.

集合$\{e \mid e\in A$ 且 $e\in B\}$称为事件 A 与事件 B 之积，记作 $A\cap B$或AB，其实际意义表示在试验中事件 A 与事件 B 同时发生.

集合$\{e \mid e\in A$ 但 $e\notin B\}$称为事件 A 与事件 B 之差，记作 $A-B$. 其实际意义表示在试验中事件 A 发生但事件 B 不发生.

如果 A 是事件，S 是样本空间，称 $S-A$ 为事件 A 的对立事件，记为 $\bar{A}$，其实际意义表示在试验中事件 A 不发生.

如果事件 A 与事件 B 之积 $AB=\varnothing$，则称事件 A 与事件 B 是不相容事件. 显然 A 与 A 的对立事件 $\bar{A}$ 是不相容事件.

上面所定义的 $A\cup B$，AB，$A-B$ 和 $\bar{A}$ 等皆为随机事件，事件的这些运算，从集合论的观点来看就是集合中的子集的并、交及差等运算. 由于事件的运算在表达上其实就是集合间的运算，因此关于集合间的一些性质在这里皆成立，兹列举如下：

① 交换律　$A\cup B=B\cup A\quad A\cap B=B\cap A$

② 分配律　$A(B\cup C)=AB\cup AC$

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

③ 结合律　$A\cup(B\cup C)=(A\cup B)\cup C\quad A(BC)=(AB)C$

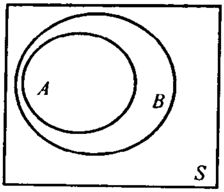
④ 德摩根(De Morgan)律　$\overline{A\cap B}=\bar{A}\cup\bar{B}\quad \overline{A\cup B}=\bar{A}\,\bar{B}$

⑤ $\overline{(\bar{A})}=A$.

事件的和及事件的积可以推广到多个事件上去. 例如设 A_1，A_2，$\cdots$，A_n 是 n 个事件，称$\{e \mid$存在 $1\leqslant i_0\leqslant n$，使 $e\in A_{i_0}\}$为 A_1，A_2，$\cdots$，A_n 之和记为 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$. 称$\{e \mid$对每个 $1\leqslant i\leqslant n$，有 $e\in A_i\}$为事件 A_1，A_2，$\cdots$，A_n 之积，记为 $A_1A_2\cdots A_n$ 或 $\bigcap\limits_{i=1}^{n}A_i$. 此时德摩根律为$\overline{\bigcap\limits_{i=1}^{n}A_i}=\bigcup\limits_{i=1}^{n}\bar{A}_i\quad \overline{\bigcup\limits_{i=1}^{n}A_i}=\bigcap\limits_{i=1}^{n}\bar{A}_i$.

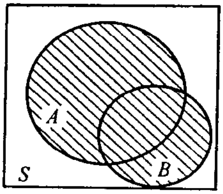
事件的关系及运算用下述直观图形表示(图 1-1 至图 1-6).

例 7　A 及 B 同例 6 中的假定，记 C 表示出现 4 以上的点



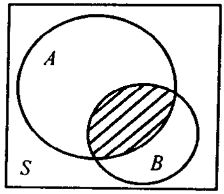
表示A⊂B

图 1-1



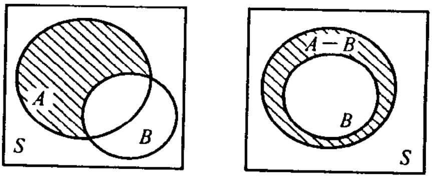
表示A∪B

图 1-2



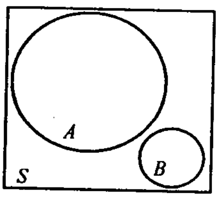
表示A∩B

图 1-3



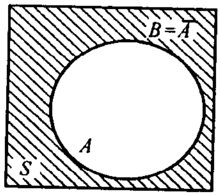
表示A－B或A$\overline{B}$

图 1-4



表示AB=∅

图 1-5



表示Ā

图 1-6

数.则有：

$$A \cup B = S \quad A \cup C = \{e_2, e_4, e_5, e_6\} \quad A - C = \{e_2\}$$

例 8 将一枚硬币抛三次，用 $A_i(i=1,2,3)$表示第 i 次时出现正面，试利用 A_1、A_2、A_3 表示至少出现一次正面，至多出现两次正面等事件，并说明 $\overline{A}_1\overline{A}_2 \cup \overline{A}_2\overline{A}_3 \cup \overline{A}_1\overline{A}_3$ 的实际意义.

解 三次中至少出现一次正面,说明 A_1,A_2,A_3 中至少有一个发生,即可表为 $A_1\cup A_2\cup A_3$.

三次中至多出现两次正面说明 $\bar{A}_1,\bar{A}_2,\bar{A}_3$ 中至少有一个发生,即可表为 $\bar{A}_1\cup\bar{A}_2\cup\bar{A}_3$ 或表为 $\overline{A_1\cap A_2\cap A_3}$.

$\bar{A}_1\bar{A}_2\cup\bar{A}_2\bar{A}_3\cup\bar{A}_1\bar{A}_3$ 表示三次抛掷中至多出现一次正面.

第二节 随机事件的概率

概率论中首先关心的是事件发生的可能性大小,描述事件发生的可能性大小的概念就是概率,下面将通过一个实例来说明事件发生的可能性大小是客观存在且是可描述的.考察掷均匀硬币这个试验,这种试验历史上有人做过,得到如表 1-1 所示的数据,其中 n_H 表示 n 次试验中出现正面的次数.

表 1-1

实验者	次数 n	n_H	n_H/n
德摩根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
K·皮尔逊	12000	6019	0.5016
K·皮尔逊	23000	12012	0.5005

虽然每次试验前不知道将会出现哪一面,但通过上面的试验结果可知,出现正面与出现反面的次数几乎各占总试验次数的一半,这表明出现正面的可能性大小与出现反面的可能性大小应是相等的.因此,出现正面的可能性大小跟已做的试验次数中出现正面所得的比值 n_H/n 有密切关系,从中可以大致了解事件发生的可能性大小.为了更清楚地解释一般事件的概率意义,借助于这种比值给出频率概念.

一、频率

设对随机试验 E 重复独立地进行 n 次，随机事件 A 共发生了 n_A 次，称 n_A/n 为事件 A 的频率，记为 $f_n(A)$，即 $f_n(A)=n_A/n$.

由定义易见频率有以下性质：

① $0\leqslant f_n(A)\leqslant 1$

② $f_n(S)=1$

③ 若 $A_1,A_2,\cdots,A_k$ 是两两不相容事件，则

$$f_n(A_1\cup A_2\cup\cdots\cup A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)$$

表 1-1 中最后一列所列出的数据是出现正面这个事件的频率.对一般事件 A 而言，当 $f_n(A)$较大时，意味着事件 A 发生愈频繁，即在一次试验中发生的可能性较大，说明频率 $f_n(A)$与事件 A 发生的可能性大小有关，事件发生的可能性大小是客观存在可描述的.那么 $f_n(A)$是否就可看作事件 A 发生的可能性大小呢？因为试验结果的随机性，必然导致 $f_n(A)$是变化的且在试验之前无法确定（表 1-1 就说明了 $f_n(H)$的变化），这样不能用 $f_n(A)$来表示在一次试验中事件 A 发生的可能性大小.若以 H 表示出现正面这事件，从表 1 可知，频率 $f_n(H)$上下波动，但随着 n 的增大，$f_n(H)$的波动幅度减小，逐渐稳定于 0.5.大量的实践表明，一般事件的频率 $f_n(A)$具有与抛硬币试验中 $f_n(H)$相同的特性，即 $f_n(A)$随着 n 的增大，稳定于某常数值附近，称这种性质为频率稳定性.频率稳定性的另一例证是测量中的一个问题，我们知道，如果多次测量物体的某指标，其结果虽略有差异，但当测量次数增加时，测量值的平均值在某固定常数附近变化，诸测量值在此常数两旁的分布呈现某种对称性.事实上，对一般情形的事件的频率稳定性已不断地为人类的实践所证实，并且在理论上可以证明（见第五章中贝努利大数定律），在一定的条件下，频率稳定在某常

数附近对任意的随机事件都成立.这样对每一个事件都客观地存在一个数与事件对应,这个数就称为概率,它表征事件在一次试验中发生的可能性大小.

二、概率

在实践中,我们不可能对每一个事件都做大量的试验从中找出频率的稳定值,同时为了理论上的需要,我们从频率的性质得到启发,可以给出如下的概率定义.

定义 设试验 E 的样本空间为 S,如果对于 E 的每一事件 A 都对应一个实数 $P(A)$,且满足

①对每个事件 A,$P(A)\geqslant 0$;

②$P(S)=1$;

③设 $A_1,A_2,\cdots,A_n,\cdots$,是两两不相容的一列事件,即对于 $i\neq j$,$A_i\cap A_j=\varnothing$,$i,j=1,2,\cdots$,则有

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots+P(A_n)+\cdots \tag{1-1}$$

等式(1-1)称为可列可加性.

由概率的定义,可以推得概率的一些重要性质.

性质①$P(\varnothing)=0$

证明 令 $A_n=\varnothing$,$n=1,2,\cdots$显然 $A_iA_j=\varnothing$,$i\neq j$,$j=1,2,\cdots$,由(1-1)得:

$$P(\varnothing)=P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots$$

从而必有 $P(\varnothing)=0$.

性质②设 $A_1,A_2,\cdots,A_n$ 是两两不相容的几个事件,则有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n) \tag{1-2}$$

证明 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$,则对 $A_1,A_2,\cdots\cdots$有 $A_iA_j=$

$\varnothing, i \neq j, i, j = 1, 2, \cdots$. 由(1-1)及性质①得

$$\begin{aligned} & P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1 \cup A_2 \cup \cdots) \\ & = P(A_1) + P(A_2) + \cdots + P(A_n) + P(A_{n+1}) + \cdots \\ & = P(A_1) + P(A_2) + \cdots + P(A_n) \end{aligned}$$

性质③设 A, B 是两个事件,若 $A \subset B$,则

$$\begin{aligned} & P(B - A) = P(B) - P(A) \\ & P(B) \geqslant P(A) \end{aligned} \tag{1-3}$$

证明 由 $A \subset B$ 知:$B - A = B\bar{A}$, $B = A \cup (B\bar{A})$, $A \cap (B\bar{A}) = \Phi$,由(1-2)得:

$$P(B) = P(A) + P(B\bar{A}) = P(A) + P(B - A)$$

即 $P(B - A) = P(B) - P(A)$

由此立得 $P(B) \geqslant P(A)$.

性质④对任一事件有 $P(A) \leqslant 1$.

性质⑤对任一事件 A 有 $P(\bar{A}) = 1 - P(A)$,性质④及性质⑤的证明留给读者完成.

性质⑥对任意两个事件 A、B,有

$$P(A \cup B) = P(A) + P(B) - P(AB) \tag{1-4}$$

证明 因 $A \cup B = A \cup (B\bar{A}) = A \cup (B - AB)$, $A \cap (B - AB) = \varnothing$,由(1-2)及(1-3)得:

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

性质⑥可以推广到任意有限个随机事件之和的情形,即对于任意有限个随机事件 $A_1, A_2, \cdots, A_n$,有:

$$\begin{aligned} & P(A_1 \cup A_2 \cup \cdots \cup A_n) \\ & = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i, j \leqslant n} P(A_i A_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_i A_j A_k) + \cdots + \\ & \quad (-1)^{n-1} P(A_1 A_2 \cdots A_n) \end{aligned} \tag{1-5}$$