21世纪

高等院校工科类数学教材

概率统计

褚宝增 王翠香 主编

北京大学出版社
PEKING UNIVERSITY PRESS

21 世纪
高等院校工科类数学教材

概 率 统 计

褚宝增 王翠香 主编

北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

概率统计/褚宝增,王翠香主编. —北京:北京大学出版社,2010.2
(21世纪高等院校工科类数学教材)
ISBN 978-7-301-15829-6

Ⅰ.概… Ⅱ.①褚… ②王… 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(2009)第173679号

书　　　名:概率统计
著作责任者:褚宝增　王翠香　主编
责 任 编 辑:曾琬婷
标 准 书 号:ISBN 978-7-301-15829-6/O·0800
出 版 发 行:北京大学出版社
地　　　址:北京市海淀区成府路205号　100871
网　　　址:http://www.pup.cn　　电子信箱:zpup@pup.pku.edu.cn
电　　　话:邮购部62752015　发行部62750672　理科编辑部62752021　出版部62754962
印　刷　者:北京大学印刷厂
经　销　者:新华书店
787mm×960mm　16开本　15印张　320千字
2010年2月第1版　　2010年2月第1次印刷
印　　　数:0001—4000册
定　　　价:28.00元

前　言

当前，我国高等教育蓬勃发展，教学改革不断深入，高等院校工科类数学基础课的教学理念、教学内容及教材建设也孕育在这种变革中.为适应高等教育21世纪教学内容和课程体系改革的总目标，培养具有创新能力的高素质人才，我们应北京大学出版社的邀请，经集体讨论，分工编写了这套《21世纪高等院校工科类数学教材》，本册为《概率统计》.

本教材参照2005年教育部数学课程教学指导委员会下发的《工科类本科数学基础课程教学基本要求(修订稿)》，按照“加强基础、培养能力、重视应用”的指导方针，精心选材，力求实现基础性、应用性、前瞻性的和谐统一，集中体现了编者长期讲授工科类概率统计课程所积累的丰富教学经验，反映了当前工科数学教学理念和教学内容的改革趋势.具体体现在以下几个方面：

1. 精心构建教材内容.本教材在内容选择方面，根据工科学生的实际要求及相关专业课程的特点，汲取了国内外优秀教材的特点，对传统的教学内容在结构和内容上作了适当的取舍、补充和调整，为后续课程打好坚实的基础.

2. 内容讲述符合认知规律.以实际的例子导入问题，然后引出相关概念，并在叙述时力求严谨，兼顾直观和抽象，再通过有针对性的例题和习题加深对概念的理解与结论的应用.对重点概念、重要定理、难点内容从多侧面进行剖析，做到难点分散，便于学生理解与掌握.

3. 加强基础训练和基本能力的培养.紧密结合概念、定理和运算法则配置丰富的例题，并剖析一些综合性例题.按节配有适量习题，每章配有总练习题，书末附有参考答案与提示，便于读者参考.

4. 注重学生数学思维的训练.教材自始至终贯穿了从具体到抽象的建模过程和从抽象到具体的应用体验，提高学生的符号演绎能力，力求达到培养学生的数学思想和实际应用能力的双重目标.

全书共分八章.第一章由吴飞编写，第二章由王翠香编写，第三章由褚宝增编写，第四、五、六章由王锡禄编写，第七、八章由郭翠平编写.全书由王翠香统稿，褚宝增审定.

本书的主要特点是：选材取舍精当，行文简约严密，讲解重点突出，服务后续课程，衔接考研思路等.

感谢王祖朝教授、高世臣教授、田东风教授对本书的认真审稿及提出的修改意见.北京大学出版社曾琬婷同志为本教材做了很多细致的工作，在此表示诚挚的谢意.

囿于编者水平及编写时间较为仓促，教材之中难免存在疏漏与不妥之处，恳请广大读者不吝指正.

编　者

2009年10月

目　　录

第一章 概率论的基本概念

概率论源于博弈问题的讨论.在生产实践和物理学、生物学等学科发展的推动下,概率论得到了飞速的发展,理论和实践研究不断深入,其思想和方法已经渗透到各个学科而成为近代科学的明显特征.概率论在可靠性理论、信息论、工程技术和社会科学等很多领域都有重要的应用.本章是概率论的基础,主要介绍概率论的基本概念、基本公式、事件的独立性以及概率的计算.

§1.1 随机现象与随机事件

一、随机现象与随机试验

概率论是研究随机现象统计规律性的一门数学学科.那么什么是随机现象呢?人们通过观察会发现,在自然界和人类社会中存在着两类不同的现象.一类是在一定的条件下必然会发生的现象.例如,在标准气压下,水在100℃时必然会沸腾;太阳从东方升起;等等.这类现象称为**确定性现象**.但在自然界和社会生活中还广泛存在着与确定性现象有着本质区别的另一类现象.例如,在相同的条件下抛掷同一颗骰子,观察出现的点数,其结果可能是1~6点中的任何一个,并且在每次抛掷前无法肯定抛掷的结果是什么;自动车床加工出来的机械零件,可能是合格品,也可能是次品;同一门大炮向同一目标发射多发同种炮弹,因受各种因素的影响,弹落点也不一样;等等.这类现象的一个共同特点是:在基本条件不变的情况下,一系列试验或观察会得到不同的结果.通常称这类现象为**随机现象**.随机现象虽然在个别的试验或观察中会时而出现这种结果,时而出现那种结果,但在大量重复试验或观察下,其结果却呈现出某种规律性.例如,在多次抛掷骰子时会发现,出现各个点数大致各占1/6;同一门大炮向同一目标发射炮弹的弹着点按照一定的规律分布;等等.概率论与数理统计就是研究和揭示随机现象在大量重复试验或观察中所呈现出的统计规律性的一门数学学科.

研究随机现象是通过随机试验来进行的. 概率论中的试验是个含义广泛的术语,包括各种各样的科学实验,甚至对某一事物的某一特征的观察也可认为是一种试验. 我们称具备以下三个特点的试验称为**随机试验**:

(1) 可在相同的条件下重复进行;

(2) 每次试验可出现不同的结果,进行一次试验之前不能确定哪种结果会出现;

(3) 每次试验前知道可能出现的全部结果.

通常我们用 E 表示随机试验. 下面举一些随机试验的例子:

E_1: 抛掷一颗骰子,观察出现的点数.

E_2: 将一枚硬币抛掷 3 次,观察出现正面 H 的次数.

E_3: 将一枚硬币抛掷 3 次,观察正面 H 和反面 T 出现的情况.

E_4: 观察某城市一电话交换台一昼夜接到的呼唤次数.

E_5: 从某厂生产的灯管中任意抽取一根,测试它的使用寿命.

从上述例子可知,随机试验是产生随机现象的过程. 我们是通过研究随机试验来研究随机现象,揭示随机现象的内在规律性的.

二、样本空间和随机事件

样本空间与事件是概率论中最基本的两个概念. 下面我们结合例子来介绍这两个基本概念.

为了研究随机试验,首先需要知道这个试验可能出现的结果. 把随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**,记为 Ω. 样本空间的元素,即 E 的每个结果,称为**样本点**,记为 e. 在具体问题中,给定样本空间是描述随机现象的第一步.

下面给出上一小节中随机试验 $E_k(k=1,2,\cdots,5)$ 的样本空间 $\Omega_k(k=1,2,\cdots,5)$:

$$\Omega_1=\{1,2,3,4,5,6\};\quad \Omega_2=\{0,1,2,3\};$$

$$\Omega_3=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\};$$

$$\Omega_4=\{0,1,2,3,\cdots\};\quad \Omega_5=\{t\mid t\geqslant 0\}.$$

从上面的例子可以看出,随着试验的要求不同,样本空间可以相当简单,也可以相当复杂. 需要注意的一点是:样本空间的元素是由试验的目的所决定的. 例如,在 E_2 和 E_3 中同是将一枚硬币连掷三次,由于试验的目的不一样,其样本空间也不一样.

一般地,把随机试验 E 的样本空间 Ω 的子集称为 E 的**随机事件**,简称为**事件**,常用 A, B, C 等表示. 例如,在 E_1 中,若研究抛出的点数是否为偶数的情形,则满足这一条件的样本点组成 Ω_1 的一个子集: $A=\{2,4,6\}$. A 就是试验 E_1 的一个随机事件.

从上面的定义可以看出，事件为样本点组成的集合. 特别地，由一个样本点组成的单点集，称为**基本事件**. 我们称某事件发生当且仅当它所包含的某一个样本点出现.

若把样本空间作为一个事件，因为在每次试验中必然出现 Ω 中的某个样本点，也即 Ω 必然发生，故称 Ω 为**必然事件**. 类似地，把空集 $\varnothing$ 也作为一个事件，它在每次试验中都不会发生，故称为**不可能事件**. 为了今后研究问题的方便，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来统一处理.

三、事件之间的关系和事件的运算

概率论的一个重要研究课题是从简单事件的概率推算出复杂事件的概率. 为此就需要研究若干个在某些相同条件下的事件以及它们之间的关系和运算. 对这个问题的研究不仅可以帮助我们更深刻地认识事件的本质，而且可以大大地简化一些复杂事件的概率计算.

1. 事件的包含与相等

设 A,B 为两个事件，若事件 A 发生必导致事件 B 发生，即 A 中的每个样本点必在 B 中，则称**事件 A 包含于事件 B**，或称**事件 B 包含事件 A**，记为 $A\subset B$ 或 $B\supset A$. 显然，对任何事件 A，必有 $\varnothing\subset A\subset\Omega$.

设 A,B 为两个事件，若 $A\subset B$ 与 $A\supset B$ 同时成立，则称事件 A 和事件 B **相等**，记为 $A=B$. 从此定义知，对于事件 A,B，当 $A=B$ 时，事件 B 中的样本点都在事件 A 中，且事件 A 中的样本点也都在事件 B 中.

2. 事件的和

“事件 A 或事件 B 至少有一个发生”是一个事件，称为事件 A 与事件 B 的**和**，记为 $A\cup B$ 或 $A+B$.

两个事件的和的定义可推广到有限个和可列无穷多个事件的情形.

一般地，“n 个事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生”是一个事件，称为 n 个事件 $A_1,A_2,\cdots,A_n$ 的和事件，记为 $A_1\cup A_2\cup\cdots\cup A_n$，简记为 $\bigcup\limits_{i=1}^{n}A_i$.

当涉及可列无穷多个事件时，用事件 $\bigcup\limits_{i=1}^{\infty}A_i$ 表示可列无穷多个事件 $A_1,A_2,\cdots$ 中至少有一个发生.

3. 事件的积

设 A,B 为两个事件，“事件 A 和事件 B 同时发生”是一个事件，称为事件 A 与事件 B 的**积**，记为 $A\cap B$ 或 AB.

显然，积事件 AB 的样本点由既属于事件 A 又属于事件 B 的公共样本点组成.

类似于事件的和，可把事件的积的概念推广到 n 个事件及可列无穷多个事件的情形，即："n 个事件 $A_1, A_2, \cdots, A_n$ 同时发生"是一个事件，称为 $A_1, A_2, \cdots, A_n$ 的积，记为 $A_1 \cap A_2 \cap \cdots \cap A_n$ 或 $A_1 A_2 \cdots A_n$，简记为 $\bigcap_{i=1}^{n} A_i$. 对可列无穷多个事件的情形，用 $\bigcap_{i=1}^{\infty} A_i$ 表示可列无穷多个事件 $A_1, A_2, \cdots$ 同时发生.

需要注意的一点是，在概率论中，为解决复杂问题的需要，常将一个事件分解为若干事件之和$\left(A = \bigcup_{i=1}^{\infty} A_i\right)$或若干事件之积$\left(A = \bigcap_{i=1}^{\infty} A_i\right)$.

4. 互不相容事件

设 A, B 为两个事件，若事件 A 与事件 B 不能同时发生，即 $AB = \varnothing$，则称事件 A 与事件 B 是**互不相容事件**或**互斥事件**. 由此定义，两个事件互不相容时，它们没有公共的样本点.

若 n 个事件 $A_1, A_2, \cdots, A_n$ 中的任意两个事件都互不相容，即 $A_i A_j = \varnothing (i \neq j; i, j = 1, 2, \cdots, n)$，则称这 **$n$ 个事件互不相容**.

事件互不相容的概念的引入在某些情况下可大大简化和事件的相关运算，故事件互不相容的概念在概率论的理论研究和实际应用中经常出现，读者应对这一概念认真体会和准确把握.

5. 对立事件

若事件 A 与事件 B 互不相容，且它们的和为必然事件，即 $AB = \varnothing$ 且 $A \cup B = \Omega$，则称事件 B 为事件 A 的**对立事件**或**逆事件**. 事件 A 的对立事件记为 $\overline{A}$.

显然 $\overline{\overline{A}} = A, A\overline{A} = \varnothing, A \cup \overline{A} = \Omega$，即 A 与 $\overline{A}$ 互为对立事件. 需要注意，若事件 A, B 互为对立事件，则事件 A, B 必互不相容；但反之不然. 另外，当事件 A 较为复杂而 $\overline{A}$ 较为简单时，往往通过研究 $\overline{A}$ 来研究 A. 有时为了研究问题的需要，可将一个事件分解为若干个互不相容事件之和. 例如，对于任意的两个事件 A, B，下述分解总是成立的：

$$A = AB \cup A\overline{B}.$$

6. 事件的差

设 A, B 为两个事件，"事件 A 发生而事件 B 不发生"是一个事件，称为事件 A 与事件 B 的**差**，记为 $A - B$.

显然，$A - B$ 表示包含在 A 中而不包含在 B 中的样本点全体，并且有

$$A - B = A\overline{B} = A - AB.$$

这个关系式很重要，在概率论的研究中经常使用.

事件的运算成立下列定律，证明留给读者.

设 A, B, C 为事件，则有

(1) **交换律**：$A \cup B = B \cup A$，$A \cap B = B \cap A$；

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

(3) **分配律**：$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$，$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$；

(4) **德·摩根**(De Morgan)**对偶律**：$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

对 n 个事件，甚至可列无穷多个事件，德·摩根对偶律也成立，即

$$\overline{\bigcup_{i=1}^{n} A_i}=\bigcap_{i=1}^{n}\overline{A_i},\quad \overline{\bigcap_{i=1}^{n} A_i}=\bigcup_{i=1}^{n}\overline{A_i};\quad \overline{\bigcup_{i=1}^{\infty} A_i}=\bigcap_{i=1}^{\infty}\overline{A_i},\quad \overline{\bigcap_{i=1}^{\infty} A_i}=\bigcup_{i=1}^{\infty}\overline{A_i}.$$

另外，易证以下等式成立：

$$A\cup A=A,\quad A\cap A=A,\quad A\cup\varnothing=A,\quad A\cap\varnothing=\varnothing,$$

$$A\cup\Omega=\Omega,\quad A\cap\Omega=A,\quad A\cup B=A\cup B\overline{A}.$$

有时用平面上某矩形及其中的图形来表示事件之间的关系或运算比较直观. 这种表示事件之间关系或运算的图形称为**文氏**(Venn)**图**. 在文氏图中，用平面上的矩形区域表示样本空间 Ω，矩形区域内的点和小圆分别表示样本点和事件. 例如，在图 1 中，两个小圆分别表示事件 A 与事件 B，阴影部分表示事件 A 与事件 B 的各种关系和运算结果.

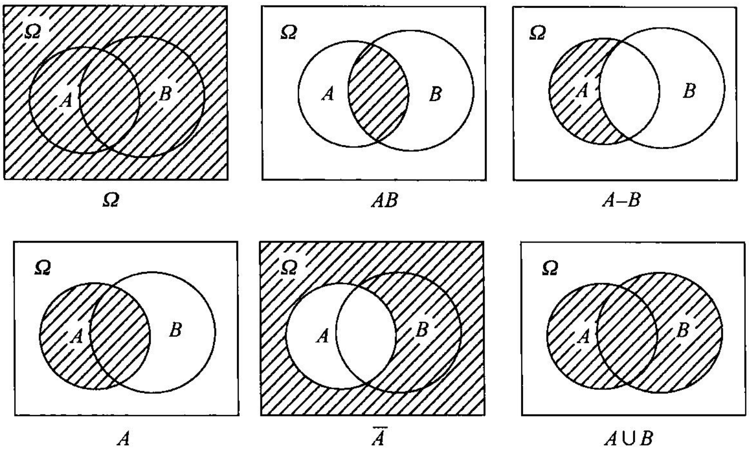

图 1

熟悉集合论的读者会注意到，事件之间的关系及运算与集合论中集合的关系及运算是完全相似的(见表 1)，而且这个相似对建立概率论的严格数学基础非常重要. 在学习时要注意用概率论的语言来解释和描述这些关系及运算，并且学会用这些运算关系来表示一些事件.

表 1

符号	概率论中的含义	集合论中的含义
$\varnothing$	不可能事件	空集
Ω	样本空间;必然事件	全集
$\{e\}$	基本事件	单点集
$e\in\Omega$	样本点	Ω 中的元素
$\overline{A}$	事件 A 的对立事件	集合 A 的余集
$A=B$	事件 A 与事件 B 相等	集合 A 与集合 B 相等
$A\subset B$	事件 A 包含于事件 B	集合 A 包含于集合 B
$A\cup B$	事件 A 与事件 B 的和	集合 A 与集合 B 的并
$A\cap B$	事件 A 与事件 B 的积	集合 A 与集合 B 的交
$A-B$	事件 A 与事件 B 的差	集合 A 与集合 B 的差
$AB=\varnothing$	事件 A 与事件 B 互不相容	集合 A 与集合 B 没有公共元素

例 1 设 A,B,C 是三个事件,可将下列事件用 A,B,C 表示:

(1) 这三个事件都发生:ABC;

(2) 这三个事件都不发生:$\overline{A}\,\overline{B}\,\overline{C}$;

(3) 事件 A 发生而事件 B 与事件 C 都不发生:$A\overline{B}\,\overline{C}$ 或 $A-B-C$ 或 $A-(B\cup C)$;

(4) 事件 A 与事件 B 发生而事件 C 不发生:$AB\overline{C}$或 $AB-C$;

(5) 这三个事件恰好发生一个:$A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C$;

(6) 这三个事件恰好发生两个:$AB\overline{C}+A\overline{B}C+\overline{A}BC$;

(7) 这三个事件中至少发生一个:

$$A\cup B\cup C \quad 或 \quad A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C+AB\overline{C}+A\overline{B}C+\overline{A}BC+ABC.$$

例 2 设一名射手连续向某一目标射击,事件 $A_i(i=1,2,3)$表示该射手第 i 次射击时击中目标,则下列事件可用事件 A_1,A_2,A_3 表示:

(1) 第三次射击未击中目标:$\overline{A_3}$;

(2) 前两次射击中至少有一次击中目标:A_1+A_2;

(3) 三次射击中至少有一次击中目标:$A_1+A_2+A_3$;

(4) 三次射击都击中了目标:$A_1A_2A_3$;

(5) 第三次射击击中但第二次未击中目标:$A_3\overline{A_2}=A_3-A_2$;

(6) 前两次均未击中目标:$\overline{A_1}\,\overline{A_2}=\overline{A_1+A_2}$;

(7) 后两次射击中至少有一次未击中目标:$\overline{A_2}+\overline{A_3}=\overline{A_2A_3}$;

(8) 三次射击中至少有两次击中目标:$A_1A_2+A_2A_3+A_1A_3+A_1A_2A_3$.

习 题 1.1

1. 写出下列随机试验的样本空间：

(1) 投掷两颗匀称的骰子，记录点数之和.

(2) 射击一目标，直至击中目标为止，记录射击次数.

(3) 设袋中装有 4 个白球，6 个黑球. 逐个取出，直至白球全部取出为止，记录取球次数.

(4) 往数轴上任意投掷两个质点，观察它们之间的距离.

2. 设样本空间 $\Omega=\{1,2,3,4,5,6,7,8,9,10\}$，事件 $A=\{2,3,4\}$，$B=\{3,4,5\}$，$C=\{5,6,7\}$，求下列事件：

(1) $\overline{A}\cap\overline{B}$；　(2) $A\cup B$；　(3) $\overline{A\cap B}$；　(4) $A\cap\overline{(B\cap C)}$.

3. 试将事件 $A+B+C$ 表示为互不相容的事件之和.

§1.2 概率的定义

研究随机现象的目的，就是要获得随机现象各种结果发生可能性大小的度量. 这种度量能反映随机现象所呈现出的统计规律性.

一、频率与概率

对一个随机事件 A，用一个数 $P(A)$来表示该事件发生的可能性大小，这个数 $P(A)$就称为随机事件 A 的概率. 因此概率度量了随机事件发生的可能性的大小，是随机事件发生的统计规律性的数量刻画.

为了估算随机事件发生概率的大小及了解概率的性质，下面引入频率的定义及其性质.

定义 1 若事件 A 在 n 次重复试验中出现了 n_A 次，则称 n_A 为**频数**，而称$\frac{n_A}{n}$为事件 A 的**频率**，记做 $f_n(A)$，即

$$f_n(A)=\frac{n_A}{n}. \qquad ①$$

显然，事件 A 的频率 $f_n(A)$具有以下基本性质：

性质 1 (1) **非负性**：对任意的事件 A，有 $0\leqslant f_n(A)\leqslant 1$；

(2) **规范性**：$f_n(\Omega)=1$；

(3) **有限可加性**：若 $A_1,A_2,\cdots,A_k$ 是 k 个互不相容的事件，则

$$f_n\left(\bigcup_{i=1}^{k}A_i\right)=\sum_{i=1}^{k}f_n(A_i).$$

人们在长期的实践中发现，虽然个别随机现象在某次试验或观察中可能出现也可能不

出现，但在大量试验中它却呈现出明显的统计规律性——频率的稳定性.

历史上，曾有一些著名学者为了揭示随机现象的统计规律性进行抛掷硬币的随机试验. 其试验数据如表 1 所示. 由表 1 提供的数据可知，随着抛掷硬币次数的增多，频率$\frac{n_H}{n}$越来越明显地呈现出稳定性. 表 1 最后一列说明，当抛掷硬币的次数充分多时，正面向上的频率在$\frac{1}{2}$这个数的左右摆动. 这些学者所做的这些随机试验，对揭示随机现象的统计规律性具有重要的意义.

表 1

试验者	投掷硬币次数 n	正面出现的次数 n_H	频率$\frac{n_H}{n}$
蒲丰	4040	2048	0.5069
费歇尔	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

高尔顿板是另一个说明频率稳定性的著名随机试验. 它是英国生物统计学家高尔顿(Galton)设计的，试验模型如图 1 所示. 从顶端放入一个小球，任其自由落下，在其下落过程中当小球碰到钉子时，从左边落下与从右边落下的机会相等，碰到下一排钉子又是如此，最后落入底板中的某一个格子. 因此，任意放入一个小球，则该球落入哪个格子，预先难以确定. 但是试验证明，如放入大量小球，则其最后所呈现的曲线，几乎总是一样的. 也就是说，小球落入各个格子的频率十分稳定.

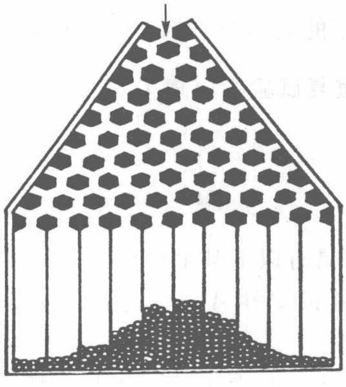

图 1

这些试验中呈现出来的统计规律性，学完第四章大数定律与中心极限定理之后，会有更深刻的理解. 同时这些试验表明，在大量试验中随机事件出现的频率稳定性说明随机事件发