普通高等教育“十一五”规划教材

大学数学教学丛书

概率论与数理统计

主　编　肖玉山
副主编　马秋红　温启军
　　　　韩兆红　董小刚

科学出版社

普通高等教育“十一五”规划教材

大学数学教学丛书

概率论与数理统计

主　编　肖玉山

副主编　马秋红　温启军

　　　　韩兆红　董小刚

科学出版社

北　京

内 容 简 介

本书介绍了概率论与数理统计的基本概念、基本理论与方法. 内容包括: 概率的基本概念; 随机变量与随机向量及其概率分布; 随机变量的数字特征、大数定律与中心极限定理; 数理统计的基本概念; 参数估计; 假设检验; 回归分析. 本书强调直观性, 注重可读性, 突出基本思想, 深入浅出. 每章均配有习题, 并在书末附有习题答案.

本书可用作普通高等院校工科、经济管理类本科专业的概率论与数理统计课程的教材, 也可供相关技术人员参考.

图书在版编目(CIP)数据

概率论与数理统计/肖玉山主编. —北京: 科学出版社, 2011
普通高等教育"十一五"规划教材 / 大学数学教学丛书

ISBN 978-7-03-032793-2

Ⅰ. 概… Ⅱ. 肖… 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(2011) 第 233789 号

责任编辑: 张中兴 于俊杰 / 责任校对: 朱光兰
责任印制: 张克忠 / 封面设计: 迷底书装

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京中新伟业印刷有限公司印刷
科学出版社发行 各地新华书店经销
*
2011 年 12 月第 一 版 开本: 720 × 1000 1/16
2011 年 12 月第一次印刷 印张: 12 1/4
字数: 240 000

定价: 24.00 元

丛 书 序

《大学数学教学丛书》是为普通高等院校本科学生所编写的一套数学教材，是由长春地区五所普通高校具有丰富教学和科研经验的教师联合编写的，是集体智慧的结晶. 本套教材从酝酿到出版经历了近十年的时间，几经修改终于成稿. 在教材内容的编排上，我们一方面借鉴了国内一些品牌教材的先进模式，另一方面结合新形势下的新要求，并根据五所普通高校本科学生的特点，先后编写了逾百万字的教材和讲义，在多年使用过程中不断提炼修订，逐步趋于完善. 应该说，本套教材凝聚了五所高校几代数学教师的心血和汗水，希望能培养出更多的创新型人才.

本套教材包括《微积分(经管类)》、《概率论与数理统计》(两本)、《线性代数》、《计算方法简明教程》、《数学建模》、《复变函数与积分变换》. 编者在取材上着眼于本科生未来的发展和当今世界科学技术的发展，充分反映国内外教学前沿信息和最新学术动态，本着“夯实基础、适当延伸，注重应用、强化实践”的原则，大胆地摆脱了普通高等院校教材编写的传统思路，使这套教材具有很强的实用性、一定的可读性、较高的艺术性和丰富的实践性；同时还保持了数学知识的系统性、严密性、连贯性等特点，内容翔实、清晰易读，便于教学与自学. 另外，本套教材充分考虑了学生课程学习和报考研究生复习的需要，每一本教材都配备了丰富的梯度配置的例题与习题，既具有明显的启发性，又具有典型的应用意义，可供普通高校理工科各专业使用.

本套教材从选题、大纲、组织编写到编辑出版，自始至终得到了科学出版社的支持，同时也得到了长春工业大学、吉林建筑工程学院、长春大学、吉林工程技术师范学院、长春工程学院教务处及数学系各位领导的支持和帮助，在此，我们一并表示衷心的感谢.

《大学数学教学丛书》编委会

2010 年 3 月

前　言

"概率论与数理统计"主要研究和探讨客观世界中随机现象的规律. 它已在包括控制、通信、生物、金融、社会科学及其他工程技术等诸多领域中获得了广泛的应用. 基于该学科应用如此广泛, 它已成为高等院校理工类及经济管理类学生的一门必修基础课程.

本书是面向普通高等院校本科学生编写的教材. 其内容包括随机事件及其概率, 随机变量、随机向量及其概率分布、数字特征, 大数定律及中心极限定理, 数理统计的基本概念, 参数估计, 假设检验, 回归分析. 作为一本入门教材, 本书在编写时, 尽量以实际例子引入概率统计的基本概念、基本方法, 理论推导力求简明, 注重直观性, 紧密联系实际, 突出基本思想, 尽力做到打好基础、够用为度、服务专业、学以致用. 期望本书能使学生增强随机思维能力, 并对培养学生的统计素质有所裨益.

本书共 8 章, 由肖玉山、马秋红、温启军、韩兆红、董小刚参与编写. 全书由肖玉山修改定稿. 本书的编写工作得到了科学出版社和长春大学的大力支持, 在此表示衷心的感谢.

由于编者水平有限, 书中难免有不妥之处, 敬请同行及读者批评指正.

编　者

2011 年 11 月

目　　录

第 1 章　随机事件和概率

1.1　随 机 事 件

1.1.1　随机现象和随机事件

在自然界与人类社会中普遍存在着两类现象：一类为确定性现象, 即在一定条件下必然会发生的现象. 例如, 在标准大气压下, 100°C 的纯净水必然沸腾; 带异性电荷的两个小球一定相互吸引. 微积分、线性代数等学科就是研究确定性现象的有力的数学工具. 另一类现象为随机现象, 即在一定的条件下, 具有多种可能的结果, 但事先无法预知发生哪一种结果的现象. 例如, 在相同条件下抛掷同一枚硬币, 其结果可能是国徽朝上, 也有可能是国徽朝下, 并且在抛掷之前, 无法预知抛掷结果, 这是随机现象表面上的偶然性, 即随机现象的随机性, 但经过多次抛掷时, 就会发现国徽面朝上的次数几乎总是占抛掷次数的 $\frac{1}{2}$ 左右, 这是随机现象内部蕴含着的必然规律, 这种随机现象的必然性, 即为随机现象的统计规律性. 概率论与数理统计就是研究和揭示随机现象统计规律性的一门数学学科.

为了研究随机现象的统计规律性, 需对客观事物进行多次观察或科学试验 (观察或科学试验统称为试验). 如果这种试验满足以下三个条件:

(1) 在相同条件下可重复进行;

(2) 试验的可能结果不唯一, 但其全部结果事先是已知的;

(3) 试验前不能确定哪一个结果发生.

则称其为**随机试验**, 简称**试验**, 记作 E. 随机试验的结果称为**随机事件**, 简称**事件**, 记作 $A, B, C, \cdots$.

下面是随机试验和随机事件的几个例子:

E_1: 掷一颗骰子, 观察出现的点数;

E_2: 记录电话交换台在单位时间内收到的呼唤次数;

E_3: 测试某种型号电子元件的寿命.

在上述试验中, 用 A_1, A_2, A_3, A_4 表示下列事件:

A_1 : 出现点数为 1;

A_2 : 出现奇数点;

A_3 : 单位时间内收到的呼唤次数为 100 次;

A_4 : 元件的寿命大于 1000 小时.

有些事件可以看成是一些事件组合而成的, 如 A_2, A_4, 而有些事件则不能分解成其他事件的组合, 如 A_1, A_3. 我们将不能被分解成其他事件组合的简单事件称为**基本事件**.

在一定条件下一定会发生的事件称为**必然事件**, 记作 Ω; 在一定条件下一定不会发生的事件称为**不可能事件**, 记作 Φ. 例如, 在试验 E_1 中, "点数小于 7" 为必然事件; "点数大于 7" 为不可能事件.

在一次试验中, 所有基本事件的集合称为**基本事件空间或样本空间**, 记作 Ω, 其中的元素, 即基本事件称为**样本点**, 记作 ω. 例如, 在试验 E_1 中, 如用 $\omega_i(i = 1,2,\cdots,6)$ 表示出现 i 点, 则样本空间 $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, 事件 $A_2 = \{\omega_1, \omega_3, \omega_5\}$. 显然, A_2 是 Ω 的一个子集. 实际上, 任何一个事件都是其样本空间的一个子集, 因此某个事件发生当且仅当这个子集中的一个样本点出现.

例 1.1 写出试验 E_1, E_2, E_3 所对应的样本空间, 并用集合表示事件 A_1, A_2, A_3, A_4.

解 $E_1 : \Omega = \{(1\text{ 点}), (2\text{ 点}), \cdots, (6\text{ 点})\}$,

$E_2 : \Omega = \{(0\text{ 次}), (1\text{ 次}), \cdots\}$,

$E_3 : \Omega = \{t\text{小时}\,|t \geqslant 0\}$,

$A_1 = \{(1\text{ 点})\}$,

$A_2 = \{(1\text{ 点}), (3\text{ 点}), (5\text{ 点})\}$,

$A_3 = \{(100\text{ 小时})\}$,

$A_4 = \{t\text{小时}\,|t > 1000\}$.

1.1.2 随机事件间的关系及运算

概率论的重要研究内容之一是希望从简单事件的概率推算出复杂事件的概率, 因此详细地分析事件之间的关系, 不仅帮助人们更深刻地认识事件的本质, 而且可以大大简化一些复杂事件的概率计算. 由于事件是一个集合, 所以事件之间的关系和运算可以用集合的关系和运算来处理, 但要注意事件关系和运算的特有含义.

设 Ω 为某试验 E 的样本空间, $A, B, A_k(k = 1, 2, \cdots)$ 为随机事件.

1. 事件的包含与相等

若事件 A 发生必然导致事件 B 的发生, 则称**事件B包含事件A** 或称 **A是B的子事件**, 记作 $A\subset B$(图 1.1.1). 事件 A 包含于 B, 是指 A 中的元素含在 B 中, 若 A 发生, 当且仅当事件 A 中有元素出现, 该元素一定属于 B, 因此事件 B 也一定发生.

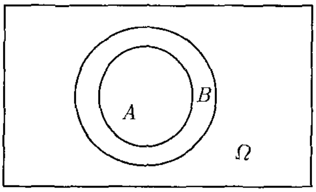

图 1.1.1

若 $A\subset B$ 且 $B\subset A$, 则称**事件A与事件B相等**, 记作 $A=B$, 其直观意义是事件 A 与事件 B 包含的样本点完全相同, 即在一次试验中, 两个事件同时发生或同时不发生.

2. 事件的并与交

事件 A 与事件 B 至少有一个发生, 称为事件 A 与事件 B 的**并或和**, 记作 $A\cup B$(图 1.1.2).

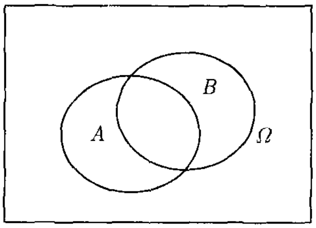

图 1.1.2

类似地, 事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生, 称为 **n个事件的并**, 记作 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$; 可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 中至少有一个发生, 称为**可列个事件的并**, 记作 $A_1\cup A_2\cup\cdots\cup A_n\cdots\cup\cdots$ 或 $\bigcup\limits_{i=1}^{\infty}A_i$.

事件 A 与事件 B 同时发生, 称为事件 A 与事件 B 的**交或积**, 记作 $A\cap B$ 或 AB(图 1.1.3).

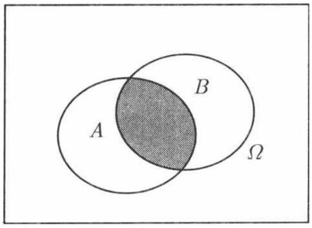

图 1.1.3

类似地, n 个事件 $A_1, A_2, \cdots, A_n$ 的积, 记作 $A_1 \cap A_2 \cap \cdots \cap A_n$ 或 $\bigcap\limits_{i=1}^{n} A_i$; 可列个事件 $A_1, A_2, \cdots, A_n, \cdots$ 的积, 记作 $A_1 \cap A_2 \cap \cdots \cap A_n \cdots \cap \cdots$ 或 $\bigcap\limits_{i=1}^{\infty} A_i$.

3. 事件的互不相容

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**互不相容的**(**或互斥的**)(图 1.1.4).

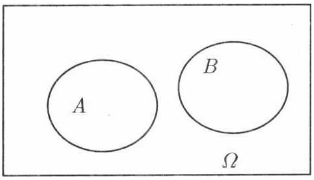

图 1.1.4

例如, 基本事件是两两互不相容的.

4. 事件的逆

对于事件 A, 由不包含在 A 中的所有样本点构成的集合称为事件 A 的**逆**(或称为 A 的**对立事件**), 记作 $\bar{A}$(图 1.1.5).

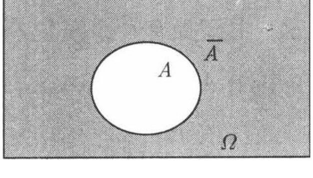

图 1.1.5

如果事件 A 与事件 B 是对立事件, 那么 A 与 B 必然满足 $AB=\varnothing$ 且 $A\cup B=\Omega$.

5. 事件的差

事件 A 发生而事件 B 不发生, 称为事件 A 与事件 B 的**差**, 记作 $A-B$(图 1.1.6), 事件 $A-B$ 是由属于事件 A 但不属于事件 B 的样本点构成.

注意: $A-B=A\bar{B}=A-AB$.

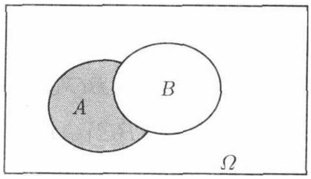

图 1.1.6

6. 样本空间的划分

为了研究某些较为复杂的事件, 常常需要把样本空间 Ω 按样本点的属性, 划分成若干个事件 $A_1,A_2,\cdots,A_n$, 当它们满足:

(1) $A_iA_j=\varnothing(i\neq j,i,j=1,2,\cdots,n)$;

(2) $A_1\cup A_2\cup\cdots\cup A_n=\Omega$,

则称这 n 个事件 $A_1,A_2,\cdots,A_n$ 构成样本空间 Ω 的一个**划分**.

显然, 对于 $A\subset\Omega$, 则 A 与 $\bar{A}$ 构成 Ω 的一个划分.

7. 事件的运算律

与集合的运算一样, 事件间的基本运算 (并、交、逆) 满足下述运算律.

(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: $A\cup(BC)=(A\cup B)\cap(A\cup C),A(B\cup C)=AB\cup AC$;

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

例 1.2 设 A,B,C 为三个事件, 试用 A,B,C 的运算关系表示下列各个事件.

(1) A,B,C 都发生;

(2) A 与 B 都发生而 C 不发生;

(3) A,B,C 至少有一个发生;

(4) A,B,C 恰好有一个发生;

(5) A,B,C 恰好有两个发生;

(6) A,B,C 至少有两个发生;

(7) A, B, C 中不多于两个发生.

解　(1) ABC;

(2) $AB\bar{C}$ 或 $AB-C$;

(3) $A\cup B\cup C$;

(4) $A\bar{B}\bar{C}\cup\bar{A}\bar{B}C\cup\bar{A}B\bar{C}$;

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

(6) $AB\cup AC\cup BC$ 或 $ABC\cup(AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC)$;

(7) $\overline{ABC}$ 或 $\bar{A}\bar{B}\bar{C}\cup(A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C)\cup(AB\bar{C}\cup A\bar{B}C\cup\bar{A}BC)$.

例 1.3　化简下列各式:

(1) $(A\cup B)-(A-B)$;

(2) $(A-\bar{B})(\overline{\bar{A}\cup\bar{B}})$.

解　(1) $(A\cup B)-(A-B)=(A\cup B)(\overline{A-B})=(A\cup B)\overline{A\bar{B}}$

$$=(A\cup B)(\bar{A}\cup B)=A\bar{A}\cup AB\cup B\bar{A}\cup B=B;$$

(2) $(A-\bar{B})(\overline{\bar{A}\cup\bar{B}})=(AB)(\bar{A}\bar{B})=(A\bar{A})(B\bar{B})=\varnothing$.

1.2　概率的定义及计算

在对随机现象进行研究时, 我们不仅关心随机试验可能出现哪些结果, 更关心各种结果发生的可能性大小. 概率就是对事件发生可能性大小的一种数值度量. 本节在给出概率定义的基础上, 讨论一些简单的概率计算问题.

1.2.1　概率的统计定义

在长期的生产实践中, 人们发现虽然个别随机事件在一次试验中可能出现也可能不出现, 但在大量重复试验中它的发生呈现出明显的规律性——频率稳定性.

定义 1.1　在相同条件下, 进行 n 次试验. 若在 n 次重复试验中, 事件 A 发生 m 次, 则称

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

为事件 A 在 n 次试验中发生的**频率**.

由定义 1.1 容易证明事件 A 的频率具有如下性质:

(1) $0\leqslant f_n(A)\leqslant 1$;

(2) $f_n(\Omega)=1$;

(3) 若事件 A, B 两两互斥, 则 $f_n(A\cup B)=f_n(A)+f_n(B)$.

例 1.4 历史上许多统计学家都做过"抛硬币"试验, 若规定均匀硬币某一面为正面, 正面朝上为事件 A 发生. 下面的试验记录反映了抛硬币试验中事件 A 发生的规律性.

试验者	试验次数 n	A 发生的次数 m	频率 $\frac{m}{n}$
摩尔根	2048	1061	0.5181
蒲丰	4040	2048	0.5069
皮尔逊	24000	12012	0.5005
维尼	30000	14994	0.4998

从上表不难看出, 事件 A 的频率具有随机性, 所以用频率来刻画事件发生可能性大小是不合适的, 但是随着试验次数的增加, 事件 A 出现的频率逐渐稳定于 0.5 附近. 这里的 0.5 是大量试验中频率的稳定值, 与频率不同, 它所反映的是事件 A 固有的性质. 这种频率的稳定性就是所谓的统计规律性, 它揭示了隐藏在随机现象中的内在规律, 因此用频率的稳定值来度量事件发生可能性的大小是合适的.

定义 1.2 在 n 次独立重复试验中, 如果事件 A 发生的频率在区间 $[0,1]$ 上的一个确定常数 p 附近摆动. 而且一般情况下, 随着 n 的增加这种摆动的幅度越来越小, 则称常数 p 为事件 A 发生的**概率**, 记作 $P(A)=p$.

依据该定义及频率的性质, 概率也应该具备如下性质:

(1) 对于任意事件 A, $0 \leqslant P(A) \leqslant 1$;

(2) $P(\Omega)=1$;

(3) 若事件 A, B 互斥, 则 $P(A\cup B)=P(A)+P(B)$.

1.2.2 概率的公理化定义

概率的统计定义使我们对概率有了一个直观认识, 也提供了近似计算概率的方法, 即通过计算大量重复试验中的某一事件出现的频率来近似代替该事件出现的概率, 但是在实际问题中, 不可能也没有必要对每个事件都作大量的重复试验, 从中得到频率的稳定值, 另外, 概率的统计定义在数学上也是不严密的. 因此有必要给出概率的一个严格的数学定义. 利用概率的统计定义, 我们得到了概率的三个基本性质, 这三个性质在一定程度上反映了概率的固有属性. 由此, 1933 年, 苏联数学家柯尔莫戈罗夫 (Kolmogrov) 在综合前人成果的基础上, 提出了概率的公理化定义, 明确了概率的严格定义, 使概率论成为严谨的数学分支, 对概率论的发展起到了积极推动作用.

定义 1.3　设随机试验 E 的样本空间为 Ω, 以 E 中所有的随机事件组成的集合为定义域, 定义一个函数 $P(A)$(其中 A 为任意随机事件), 若满足以下公理

(1) 非负性：对每一个事件 A, 都有 $P(A) \geqslant 0$;

(2) 规范性：$P(\Omega) = 1$;

(3) 可列可加性：对于可列多个两两互不相容的事件 $A_1, A_2, \cdots, A_n, \cdots$, 有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots,$$

则称函数值 $P(A)$ 为事件 A 的**概率**.

由概率的定义可以推得概率的如下一些性质：

(1) $P(\varnothing) = 0$, 即不可能事件的概率为零;

(2) 若事件 $A_1, A_2, \cdots, A_n$ 两两互斥, 则

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

(3) 对任何事件 A, 有 $P(\bar{A}) = 1 - P(A)$;

(4) 对事件 A, B, 若 $A \subset B$, 则有 $P(B - A) = P(B) - P(A)$ 且 $P(B) \geqslant P(A)$;

(5) 对任意事件 A, B, 有 $P(A \cup B) = P(A) + P(B) - P(AB)$.

证明　(1) 令 $A_n = \varnothing, n = 1, 2, \cdots$, 且 $A_iA_j = \varnothing, i \neq j$, 于是

$$P(\varnothing) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(\varnothing) = P(\varnothing) + P(\varnothing) + \cdots + P(\varnothing) + \cdots,$$

再由 $P(\varnothing) \geqslant 0$, 得 $P(\varnothing) = 0$.

(2) 令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$, 则 $A_1, A_2, \cdots, A_n, \cdots$ 两两互斥, 由性质 (1) 和概率的可列可加性, 得

$$\begin{aligned}&P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \varnothing \cup \varnothing \cup \cdots)\\ =&P(A_1) + P(A_2) + \cdots + P(A_n) + 0 + 0 + 0 + \cdots = P(A_1) + P(A_2) + \cdots + P(A_n).\end{aligned}$$

(3) 由 $A \cup \bar{A} = \Omega, A \cap \bar{A} = \varnothing$, 所以 $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$, 即 $P(\bar{A}) = 1 - P(A)$;

(4) 由 $A \subset B$ 知, $B = A \cup (B - A)$, 且 $A \cap (B - A) = \varnothing$, 因此 $P(B) = P(A) + P(B - A)$, 即 $P(B - A) = P(B) - P(A)$, 再由 $P(B - A) \geqslant 0$, 得 $P(B) \geqslant P(A)$;