



普通高等教育“十五”国家级规划教材

概率论与 数理统计教程

(第四版)

沈恒范



高等教育出版社

普通高等教育“十五”国家级规划教材

概率论与 数理统计教程

(第四版)

沈恒范

高等教育出版社

图书在版编目(CIP)数据

概率论与数理统计教程/沈恒范编著. —4版. —北京:高等教育出版社,2003.4

ISBN 7-04-009187-9

I. 概... II. 沈... 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(2002)第 096683 号

出版发行	高等教育出版社	购书热线	010-64054588
社 址	北京市东城区沙滩后街 55 号	免费咨询	800-810-0598
邮政编码	100009	网 址	http://www.hep.edu.cn
传 真	010-64014048		http://www.hep.com.cn
经 销	新华书店北京发行所		
排 版	高等教育出版社照排中心		
印 刷	北京印刷三厂		
开 本	787×960 1/16	版 次	2003年4月第4版
印 张	19.5	印 次	2003年4月第1次印刷
字 数	350 000	定 价	20.60 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

第四版序言

2001年我国加入了世界贸易组织,我国与世界各国的国际经济与贸易往来必将日益增多.为了与国际接轨,便于今后的国际交往,我国在近几年中陆续颁布了一系列的国家标准.概率论与数理统计课程由于在自然科学、社会科学、工农业生产、金融、经济等各方面有着广泛的应用,所以本课程在高等学校中的重要性也就更加突出.本书主要是根据上述国家标准及全国自然科学名词审定委员会公布的《数学名词》对概率论与数理统计课程中某些基本概念的定义以及有关的定理、公式等的叙述进行了修订,并对某些数学名词作了修改.

随着21世纪科学技术的迅速发展,现代化的教学设施已在我国高等学校中得到普遍使用.本课程的教学方法和教学手段也必须适应这一新的形势,所以本书特别强调应当尽量利用电子计算器或电子计算机以及有关的软件进行统计计算,这样不仅可以提高计算的精确性,而且可以大大节省计算的工作量.

随着我国高等教育的深入发展,更好地培养高等学校学生的能力和素质,缩减各门课程的教学时数已是必然的趋势.为此,这次修订参照教育部颁发的高等工业学校《概率论与数理统计课程教学基本要求》中的第Ⅱ类(概率少、统计多)教学基本要求,精简了第三版中的若干内容,同时采用了新的课程教学体系.这样,讲授本课程教学要求的全部内容一般只需48学时.此外,本书中还有少量超出本课程教学要求的内容,都已用*号表明,供读者选学时参考.

本书所用的课程教学体系是著者与四川大学王明慈教授于1996年共同拟订,并在以后的教学和教材中付诸实施的;1998年又得到天津大学齐植兰教授和吉林工业大学高文森教授的合作,进一步完善了课程体系,充实了教学内容;对于促进本课程的教学改革,提高本课程的教学质量,都起了重要的作用.所以,著者特别感谢上述三位教授的长期友好的合作,并向他们致以最诚挚的谢意.

本书修订过程中,曾经得到湖北汽车工业学院领导同志的关心和支持,著者谨致以衷心的感谢.黄明副教授、严钦容副教授、迟彦惠副教授以及华中科技大学于寅教授、北京邮电大学王丽霞副教授等都曾经对本书的修订稿提出了很多有益的意见和建议,著者也向他们致以诚挚的谢意.

限于著者的水平,本书难免还存在某些缺点和错误,诚恳希望读者批评指正.

沈恒范

2002年9月

第三版序言

本书是《概率论讲义》(第二版)的修订本,并从这一版起,把书名改为《概率论与数理统计教程》。

这次修订主要是参照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程教学基本要求》(以下简称《本课程教学基本要求》)进行的.根据不同专业的需要,《本课程教学基本要求》分为两种类型:Ⅰ类侧重于概率论,Ⅱ类侧重于数理统计.本书叙述了上述两种类型教学基本要求的所有内容;此外,还有少量超出《本课程教学基本要求》的内容,分别用*号表出,仅供读者参考.

本书讲述数理统计中有关统计量的计算时,强调利用具有统计计算功能的袖珍电子计算器进行计算;为此,尽可能把各种统计量化为便于用电子计算器计算的形式,这样就避免了繁琐的列表计算,节省了计算的工作量.当然,具备一定条件的高等工科院校若能选用或研制适合本课程内容的数理统计计算软件,进行计算机辅助教学,开设数理统计实验课,安排学生利用计算机解题,则本课程的教学手段就更趋于现代化了.

讲授本书全部内容(不包括带*号的部分)约需64学时.注意到《本课程教学基本要求》两种类型的参考学时都不超过52学时,所以使用本书作为教材时应按照Ⅰ类或Ⅱ类教学基本要求适当选讲所需要的内容.

本书编写过程中,得到湖北汽车工业学院数学教研室全体同志的支持与协助,徐希贤副教授、米清河副教授、张传慰、黄明等同志为本书的初稿提供了不少有益的建议.本书由高等学校工科数学课程教学指导委员会审阅,王明慈教授(主审)、富景隆教授、孙家永教授等委员提出了若干宝贵的意见.上述所有建议和意见对提高本书的科学性和教学适用性都起了重要的作用,编者谨致以诚挚的谢意.

限于编者的水平,本书难免还存在某些缺点和错误,诚恳希望读者批评指正.

沈恒范

1993年12月

第二版序言

本书第一版自 1966 年出版以来,曾被很多高等学校选用为教材或教学参考书.这次修订主要是根据各校在教学过程中提出的意见,并参照高等学校工科数学教材编审委员会 1980 年审订的《工程数学教学大纲(草案)》有关概率论部分进行了修改和补充.

考虑到某些专业设置概率论与数理统计课程的需要,适当增加了数理统计的基本内容,把第一版的最后一章补充改写成现在的第六、七、八、九各章.对其余章节也作了必要的修订,同时对第一版中某些习题答案和印刷错误进行了校正.

本书编写和修订过程中,得到吉林工业大学数学教研室全体同志的支持和帮助,罗舜英同志参加了修订工作,黄耀宏、刘承胤、赵忠柏等同志对修订稿提出了不少有益的意见,编者谨致以诚挚的谢意.

限于编者的水平,本书一定还存在不少缺点和错误,希望读者批评指正.

沈恒范

1982 年 4 月

责任编辑	文小西
封面设计	于文燕
责任绘图	黄建英
版式设计	陆瑞红
责任校对	朱惠芳
责任印制	孔源

目 录

第一章 随机事件及其概率	1
§ 1.1 随机事件及其频率·概率的统计定义	1
§ 1.2 样本空间	4
§ 1.3 事件的关系及运算	6
§ 1.4 概率的古典定义	10
§ 1.5 概率加法定理	14
§ 1.6 条件概率·概率乘法定理	17
§ 1.7 全概率公式	19
§ 1.8 随机事件的独立性	23
§ 1.9 独立试验序列	27
* § 1.10 概率论的公理化体系	30
习题一	33
第二章 随机变量及其分布	37
§ 2.1 随机变量的概念	37
§ 2.2 离散随机变量	39
§ 2.3 超几何分布·二项分布·泊松分布	41
§ 2.4 连续随机变量	49
§ 2.5 随机变量的分布函数	52
§ 2.6 连续随机变量的概率密度	56
§ 2.7 均匀分布·指数分布	59
§ 2.8 随机变量函数的分布	61
§ 2.9 二维随机变量的联合分布	66
§ 2.10 二维随机变量的边缘分布	70
§ 2.11 随机变量的独立性	72
§ 2.12 二维随机变量函数的分布	74
习题二	82
第三章 随机变量的数字特征	87
§ 3.1 数学期望	87
§ 3.2 随机变量函数的数学期望	90
§ 3.3 关于数学期望的定理	93

§ 3.4	方差与标准差	95
§ 3.5	某些常用分布的数学期望与方差	99
§ 3.6	原点矩与中心矩	105
§ 3.7	协方差与相关系数	107
§ 3.8	切比雪夫不等式与大数定律	111
	习题三	116
第四章	正态分布	120
§ 4.1	正态分布的概率密度与分布函数	120
§ 4.2	正态分布的数字特征	123
§ 4.3	二维正态分布	126
§ 4.4	正态随机变量的线性函数的分布	131
§ 4.5	中心极限定理	133
	习题四	137
第五章	数理统计的基本知识	140
§ 5.1	总体与样本	140
§ 5.2	样本函数与统计量	145
§ 5.3	数理统计中的某些常用分布	149
§ 5.4	正态总体统计量的分布	154
	习题五	162
第六章	参数估计	165
§ 6.1	参数的点估计	165
§ 6.2	衡量点估计量好坏的标准	171
§ 6.3	正态总体参数的区间估计	174
§ 6.4	两个正态总体均值差及方差比的区间估计	180
* § 6.5	非正态总体参数的区间估计	186
§ 6.6	单侧置信限	190
	习题六	191
第七章	假设检验	196
§ 7.1	假设检验的基本概念	196
§ 7.2	正态总体参数的假设检验	201
§ 7.3	两个正态总体参数的假设检验	204
* § 7.4	非正态总体参数的假设检验	208
§ 7.5	总体分布的假设检验	210
	习题七	215
第八章	方差分析	218

§ 8.1 单因素试验的方差分析	218
§ 8.2 双因素无重复试验的方差分析	223
§ 8.3 双因素等重复试验的方差分析	229
习题八	234
第九章 回归分析	238
§ 9.1 回归分析的基本概念与最小二乘法	238
§ 9.2 线性回归方程	241
§ 9.3 线性相关的显著性检验	243
§ 9.4 利用线性回归方程预测与控制	249
§ 9.5 非线性回归分析	251
§ 9.6 多元线性回归分析	257
习题九	262
习题答案	265
附录	286

第一章 随机事件及其概率

§ 1.1 随机事件及其频率·概率的统计定义

概率论是研究随机现象(偶然现象)的规律性的科学。

人们在自己的实践活动中,常常会遇到随机现象.例如,远距离射击较小的目标,可能击中,也可能击不中,每一次射击的结果是随机(偶然)的.自动车床加工出来的机械零件,可能是合格品,也可能是废品.进行实验时,把得到的实验数据在坐标图纸上用点表示出来,我们可以看到,这些点(假设它们足够多)通常不是位于一条曲线上,而是散布在某一带形区域内,这就是所谓实验点的随机散布.

在事物的联系和发展过程中,随机现象是客观存在的.但是,在表面上是偶然性在起作用的地方,这种偶然性又始终是受事物内部隐藏着的必然性所支配的.

现实世界上事物的联系是非常复杂的,一切事物的发展过程中既包含着必然性的方面,也包含着偶然性的方面,它们是互相对立而又互相联系的,必然性经常通过无数的偶然性表现出来.

科学的任务就在于,要从看起来是错综复杂的偶然性中揭露出潜在的必然性,即事物的客观规律性.这种客观规律性是在大量现象中发现的.

在科学研究或工程技术中,我们会经常遇到,在不变的条件下重复地进行多次实验或观测.抽去这些实验或观测的具体性质,就得到概率论中试验的概念.所谓试验就是一定的综合条件的实现,我们假定这种综合条件可以任意多次地重复实现.大量现象就是很多次试验的结果.

当一定的综合条件实现时,也就是在试验的结果中,所发生的现象叫做事件.如果在每次试验的结果中,某事件一定发生,则这一事件叫做必然事件;相反地,如果某事件一定不发生,则叫做不可能事件.

在试验的结果中,可能发生、也可能不发生的事件,叫做随机事件(偶然事

件)。例如,任意抛掷硬币时,徽花向上是随机事件;远距离射击时,击中目标是随机事件;自动车床加工机械零件时,加工出来的零件为合格品是随机事件;等等。

通常我们用字母 A, B, C, \dots 表示随机事件,而用字母 U 表示必然事件, V 表示不可能事件。

例 已知一批产品共 100 个,其中有 95 个合格品和 5 个次品. 检查产品质量时,从这批产品中任意抽取 10 个来检查,则在抽出的 10 个产品中,“次品数不多于 5 个”这一事件是必然事件 U ;“次品数多于 5 个”这一事件是不可能事件 V ;而事件 A :“没有次品”; B :“恰有 1 个次品”; C :“有 2 个或 3 个次品”; D :“次品数少于 4 个”等等都是随机事件。

用数字表示大量现象中的规律性时,联系到下面的概念。

设随机事件 A 在 n 次试验中发生了 m 次,则比值 $\frac{m}{n}$ 叫做随机事件 A 的相对频率(简称频率),记作 $f_n(A)$;用公式表示如下:

$$f_n(A) = \frac{m}{n}. \quad (1.1)$$

显然,任何随机事件的频率是介于 0 与 1 之间的一个数:

$$0 \leq f_n(A) \leq 1. \quad (1.2)$$

对于必然事件,在任何试验序列中,我们有 $m = n$,所以必然事件的频率恒等于 1:

$$f_n(U) = 1. \quad (1.3)$$

对于不可能事件,我们有 $m = 0$,所以不可能事件的频率恒等于 0:

$$f_n(V) = 0. \quad (1.4)$$

经验表明,当试验重复多次时,随机事件 A 的频率具有一定的稳定性;就是说,在不同的试验序列中,当试验次数 n 充分大时,随机事件 A 的频率 $f_n(A)$ 常在一个确定的数字附近摆动。

例如,我们来看下面的实验结果,表中 n 表示抛掷硬币的次数, m 表示徽花向上的次数, $f_n(A) = \frac{m}{n}$ 表示徽花向上的频率。

实验序号	$n = 5$		$n = 50$		$n = 500$	
	m	$f_n(A)$	m	$f_n(A)$	m	$f_n(A)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506

续表

实验序号	$n = 5$		$n = 50$		$n = 500$	
	m	$f_n(A)$	m	$f_n(A)$	m	$f_n(A)$
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

从上表我们可以看出,当抛掷硬币的次数较少时,徽花向上的频率是不稳定的;但是,随着抛掷硬币次数的增多,频率越来越明显地呈现出稳定性.如上表最后一列所示,我们可以说,当抛掷硬币的次数充分多时,徽花向上的频率大致是在 0.5 这个数的附近摆动.

历史上,曾经有一些著名的统计学家进行过抛掷硬币的试验,得到如下的结果:

试验者	抛掷硬币次数 n	徽花向上次数 m	频率 $f_n(A)$
蒲丰(Buffon)	4 040	2 048	0.506 9
费歇尔(Fisher)	10 000	4 979	0.497 9
皮尔逊(Pearson)	12 000	6 019	0.501 6
皮尔逊(Pearson)	24 000	12 012	0.500 5

所有这些结果都充分验证了上述结论.

类似的例子可以举出很多.这说明随机事件在大量重复试验中存在着某种客观规律性——频率的稳定性.因为它是通过大量统计显示出来的,所以称为统计规律性.

由随机事件的频率的稳定性可以看出,随机事件发生的可能性可以用一个数来表示.这个刻画随机事件 A 在试验中发生的可能性大小的、介于 0 与 1 之间的数叫做随机事件 A 的概率,记作 $P(A)$. 概率的这个定义通常称为概率的统计定义.

当试验次数 n 充分大时,随机事件 A 的频率 $f_n(A)$ 正是在它的概率 $P(A)$ 的附近摆动.在上面的例子中,我们可以认为徽花向上的概率等于 0.5.

直接估计某一事件的概率是非常困难的,甚至是不可能的,仅在比较特殊的情况下才可以计算随机事件的概率.概率的统计定义实际上给出了一个近似计

算随机事件的概率的方法:我们把多次重复试验中随机事件 A 的频率 $f_n(A)$ 作为随机事件 A 的概率 $P(A)$ 的近似值, 即当试验次数 n 充分大时,

$$P(A) \approx f_n(A) = \frac{m}{n}. \quad (1.5)$$

因为必然事件的频率恒等于 1, 所以必然事件的概率等于一:

$$P(U) = 1. \quad (1.6)$$

又因为不可能事件的频率恒等于 0, 所以不可能事件的概率等于零:

$$P(V) = 0. \quad (1.7)$$

这样, 任何事件 A 的概率满足不等式

$$0 \leq P(A) \leq 1. \quad (1.8)$$

应该指出, 随机事件的频率是与我们已进行的试验有关的, 而随机事件的概率却是完全客观地存在着的. 在实际进行的试验中, 随机事件的频率可以看作是它的概率的随机表现. 随机事件的概率表明, 试验中综合条件与随机事件之间有完全确定的特殊的联系, 它从数量上说明了必然性与偶然性的辩证的统一.

还应指出, 随机事件的概率反映了大量现象中的某种客观属性, 这种客观属性是与我们认识主体无关的. 不应该把概率看作认识主体对于个别现象的信念程度. 有时一个人说某事件“可能发生”或“很少可能发生”, 这仅表示说话的人对该事件发生的可能性的一个判断而已. 因为个别现象不是发生, 就是不发生, 所以就个别现象来谈概率是没有任何现实意义的.

§ 1.2 样本空间

为了深入理解随机事件, 我们来叙述试验的样本点与样本空间的概念.

在不变的条件下重复地进行试验, 虽然每次试验的结果中所有可能发生的事件是可以明确知道的, 并且其中必有且仅有一个事件发生, 但是在试验之前却无法预知究竟哪一个事件将在试验的结果中发生.

试验的结果中每一个可能发生的事件叫做试验的样本点, 通常用字母 ω 表示.

试验的所有样本点 $\omega_1, \omega_2, \dots, \omega_n, \dots$ 构成的集合叫做样本空间, 通常用字母 Ω 表示, 于是, 我们有

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

例 1 设试验为任意抛掷一枚硬币, 则有样本点:

ω_1 表示“徽花向上”, ω_2 表示“字向上”.

于是样本空间是由两个样本点构成的集合:

$$\Omega_1 = \{\omega_1, \omega_2\}.$$

例 2 设试验为从装有三个白球(记为 1, 2, 3 号)与两个黑球(记为 4, 5 号)的袋中任取两个球.

(1) 如果观察取出的两个球的颜色, 则有样本点:

ω_{00} 表示“取出两个白球”,

ω_{11} 表示“取出两个黑球”,

ω_{01} 表示“取出一个白球与一个黑球”.

于是样本空间是由三个样本点构成的集合:

$$\Omega'_2 = \{\omega_{00}, \omega_{01}, \omega_{11}\}.$$

(2) 如果观察取出的两个球的号码, 则有样本点:

ω_{ij} 表示“取出第 i 号与第 j 号球”($1 \leq i < j \leq 5$).

于是样本空间是由 $C_5^2 = 10$ 个样本点构成的集合:

$$\Omega''_2 = \{\omega_{12}, \omega_{13}, \omega_{14}, \omega_{15}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{34}, \omega_{35}, \omega_{45}\}.$$

这个例子表明, 试验的样本点与样本空间是根据试验的内容而确定的.

例 3 设试验为观察放射性物质在一段时间内放射的粒子数, 则有样本点:

ω_i 表示“放射 i 个粒子”($i = 0, 1, 2, \dots$).

于是样本空间是由可数无穷多个样本点构成的集合

$$\Omega_3 = \{\omega_0, \omega_1, \omega_2, \dots\}.$$

例 4 设试验为测量车床加工的零件的直径, 则有样本点:

ω_x 表示“测得零件的直径为 x 毫米”($a \leq x \leq b$).

于是样本空间是由不可数无穷多个样本点构成的集合

$$\Omega_4 = \{\omega_x \mid a \leq x \leq b\}.$$

现在我们说明**随机事件与样本空间的关系**.

在例 2 中, 设随机事件 A 表示“取出的两个球都是白球”, 则对于样本空间 Ω'_2 来说, 我们有

$$A = \{\omega_{00}\};$$

而对于样本空间 Ω''_2 来说, 我们有

$$A = \{\omega_{12}, \omega_{13}, \omega_{23}\}.$$

这表明随机事件 A 是样本空间 Ω'_2 或 Ω''_2 的一个子集.

在例 3 中, 设随机事件 B 表示“放射性物质在一段时间内放射的粒子数不超过 10 个”, 则我们有

$$B = \{\omega_0, \omega_1, \dots, \omega_{10}\}.$$

这表明随机事件 B 是样本空间 Ω_3 的一个子集.

由此可见,任一随机事件 A 都是样本空间 Ω 的一个子集,该子集中任一样本点 ω 发生时事件 A 即发生.

因为样本空间 Ω 中任一样本点 ω 发生时,必然事件 U 都发生,所以 U 是所有样本点构成的集合;这就是说,必然事件 U 就是样本空间 Ω . 今后我们就把必然事件记作 Ω .

因为样本空间 Ω 中任一样本点 ω 发生时,不可能事件 V 都不发生,所以 V 不是任何样本点的集合;这就是说,不可能事件 V 是空集 \emptyset . 今后我们就把不可能事件记作 \emptyset .

应该指出,试验的任一样本点 ω 也是随机事件,今后我们把试验的样本点称为试验的基本事件. 显然,基本事件就是样本空间 Ω 的仅由单个样本点构成的子集.

§ 1.3 事件的关系及运算

为了研究随机事件及其概率,我们需要说明事件之间的各种关系及运算.

正如 § 1.2 中已经指出的,任一随机事件都是样本空间的一个子集,所以事件之间的关系及运算与集合之间的关系及运算是完全类似的. 在下面的讨论中,我们叙述事件的关系及运算时所用的符号也是与集合的关系及运算的符号基本上一致的.

(1) 如果事件 A 的发生必然导致事件 B 的发生,则称事件 B 包含事件 A , 或称事件 A 包含于事件 B , 记作

$$B \supset A \quad \text{或} \quad A \subset B.$$

例如,在图 1 中,设事件 A 表示点随机地落在小圆内,事件 B 表示点随机地落在大圆内,则我们有 $A \subset B$.

(2) 如果事件 B 包含事件 A , 且事件 A 包含事件 B , 即

$$B \supset A \quad \text{且} \quad A \supset B;$$

也就是说,二事件 A 与 B 中任一事件的发生必然导致另一事件的发生,则称事件 A 与事件 B 相等, 记作

$$A = B.$$

(3) “二事件 A 与 B 中至少有一事件发生”这一事件叫做事件 A 与事件 B 的并, 记作

$$A \cup B.$$

例如,在图 2 中,设试验是让点随机地落在矩形区域内,事件 A 表示点落在

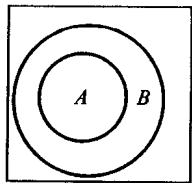


图 1

左边的圆内(图 2(a)), 事件 B 表示点落在右边的圆内(图 2(b)), 则事件 $A \cup B$ 表示点落在任一圆内(图 2(c)).

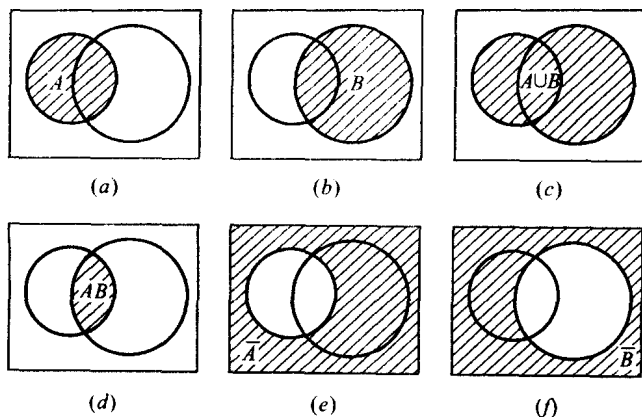


图 2

“ n 个事件 A_1, A_2, \dots, A_n 中至少有一事件发生”这一事件叫做事件 A_1, A_2, \dots, A_n 的并, 记作

$$A_1 \cup A_2 \cup \dots \cup A_n \text{ (简记为 } \bigcup_{i=1}^n A_i \text{)}.$$

(4) “二事件 A 与 B 都发生”这一事件叫做事件 A 与事件 B 的交, 记作 $A \cap B$ 或 AB .

例如, 在图 2 中, 事件 $A \cap B$ 就表示随机点落在二圆的公共部分内(图 2(d)).

“ n 个事件 A_1, A_2, \dots, A_n 都发生”这一事件叫做事件 A_1, A_2, \dots, A_n 的交, 记作

$$A_1 \cap A_2 \cap \dots \cap A_n \text{ 或 } A_1 A_2 \dots A_n \text{ (简记为 } \bigcap_{i=1}^n A_i \text{)}.$$

(5) 如果二事件 A 与 B 不可能同时发生, 即

$$AB = \emptyset,$$

则称二事件 A 与 B 是互不相容的(或互斥的).

通常把两个互不相容事件 A 与 B 的并记作

$$A + B.$$

如果 n 个事件 A_1, A_2, \dots, A_n 中任意两个事件不可能同时发生, 即

$$A_i A_j = \emptyset (1 \leq i < j \leq n),$$

则称这 n 个事件是互不相容的(或互斥的).