21世纪高等院校数学规划教材

GAILÜLUN YU SHULITONGJI

概率论与数理统计

项立群 汪晓云 张伟 梁勇 梅春晖 编著

北京大学出版社
PEKING UNIVERSITY PRESS

21世纪高等院校数学规划教材

概率论与数理统计

项立群　汪晓云　张伟　梁勇　梅春晖　编著

北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书根据高等院校非数学专业概率论与数理统计课程的教学大纲及工学和经济学数学考研大纲编写而成，内容包括：概率论的基本概念、一维和多维随机变量及其分布、随机变量的数字特征、数理统计的基本概念、参数估计、假设检验、方差分析及回归分析初步、数学软件与数学实验等.内容循序渐进，知识由浅入深，图文并茂，例题全面，习题分节设置，方便教学.

本书可作为普通高等院校理、工、经、管(不含数学类)各专业概率论与数理统计课程的教材或参考书，也可供有关技术人员自学或参考.

图书在版编目(CIP)数据

概率论与数理统计/项立群等编著.—北京：北京大学出版社，2011.1
ISBN 978-7-301-18351-9

Ⅰ.①概… Ⅱ.①项… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2010)第 260373 号

书　　名：概率论与数理统计
著作责任者：项立群　汪晓云　张伟　梁勇　梅春晖　编著
责 任 编 辑：潘丽娜
标 准 书 号：ISBN 978-7-301-18351-9/O·0838
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　　电子邮箱：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：河北滦县鑫华书刊印刷厂
经 销 者：新华书店
787mm×980mm　16 开本　13.5 印张　286 千字
2011 年 1 月第 1 版　2011 年 1 月第 1 次印刷
定　　价：26.00 元

前 言

概率论与数理统计是一门从数量方面研究随机现象客观规律性的学科，伴随着社会的发展、科学技术的进步以及研究的深入，它在科学研究、工程技术、国民经济等领域起着越来越重要的作用，是高等学校非数学专业学生的必修课，也是工学和经济学硕士研究生入学考试数学科目的考试内容.

本书是按照21世纪新形势下高等学校数学教材改革的精神，根据教育部“高等院校工科数学教学大纲”的要求，基于编者多年的教学实践专为普通高等院校本科学生而编写的.

本书分三部分：概率论部分(第1章至第5章)作为基础理论知识，是全书的起点和重点；数理统计部分(第6章至第9章)介绍了常用的统计推断的基本方法：参数估计，假设检验，方差分析与回归分析初步；数学软件与数学实验部分主要介绍在概率论和数理统计中如何结合数学软件进行应用的相关方法.

本书体现了编者多年的教学经验和实践成果，注意到了时代的特点.本着丰富背景、加强基础、强化应用、整体优化的原则，力争做到科学性、系统性和可行性相统一，传授数学知识和培养数学素养相统一，先进性和适用性相统一.

为方便教学，本书吸取了国内外同类教材的优点，内容通俗易懂，编排遵循教学规律.书中配置了大量例题，分节设立了数量较多、题型丰富的习题，且配有少量易读、易做的英文题目.书后给出了习题参考答案或提示，以供读者参考.本书可作为普通高等院校非数学类理、工、经管等科各专业“概率论与数理统计”课程教材或教学参考书，也可供有关技术人员自学或参考.

本书由项立群，汪晓云，张伟，梁勇，梅春晖等编写，并请相关教师审阅，由项立群统稿.本书的编写得到许多领导与同事的关心和鼓励，更得到北京大学出版社及潘丽娜编辑的热情支持，编者在此表示衷心的感谢.由于编者水平有限，书中难免有不妥甚至错误之处，恳请读者批评指正，以便我们进一步修改、完善.

编 者

2010年8月

目　　录

第 1 章

随机事件及其概率

概率论与数理统计是一门从数量方面研究随机现象客观规律性并将成果应用于实际的学科，是数学的一个分支. 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、随机事件的独立性以及概率的计算问题.

§1.1 随机试验与概率定义

一、随机现象

现实世界中发生的千变万化，概括起来无非是两类现象：一类是在一定条件下必然出现(或恒不出现)的现象，称为**确定性现象**. 例如，日出东方，水在标准大气压下加热到 100°C 时必定沸腾，三角形的三个内角和为 180°，等等. 我们还可以从物理学、化学等其他学科中举出许多这样的实例. 而另外的一类情况则比较复杂，它是指在一定条件下，可能发生也可能不发生的现象，具有不确定性(或称为**偶然性**或**随机性**). 例如，抛掷一枚硬币，结果可能出现正面向上，也可能出现反面向上，其结果呈现不确定性(图 1.1). 我们称这类现象为**随机现象**. 在我们所生活的世界中充满了这种随机现象，从抛硬币、掷骰子、玩扑克等简单的游戏，到粒子运动、气候变化、流星坠落等自然现象，再到考试成功与否、生男生女、股票价格升降(图 1.2)等社会现象.

从亚里士多德(Aristoteles，公元前 384—前 322，古希腊)开始，哲学家们就已经认识到随机性在生活中的作用，然而他们把随机性看做是破坏生活规律、超越人们理解能力的东西，避之唯恐不及，因而没有去研究随机性.

人们最早对随机性的研究始于博弈，后来慢慢有人研究其他方面的随机性，如帕斯卡(Pascal，法国)、惠更斯(Huygens，荷兰)、伯努利(Bernoulli，法国)、切比雪夫(Chebyshev，俄国)、柯尔莫哥洛夫(Kolmogorov，俄国)等都对随机性进行过研究，但将随机性数量化来尝试回答随机现

象的问题，是直到 20 世纪初才开始的，随之而来的这场革命改变了人们的思维方法，也提供了探索自然奥秘的工具，概率论与数理统计应运而生.

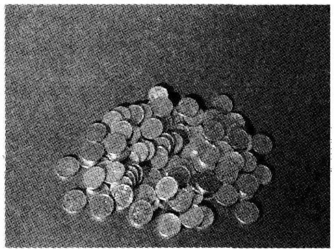
图 1.1

图 1.2

人们发现，研究了大量的同类随机现象后，通常总揭露出一种完全确定的规律，也就是大量随机现象所特有的一种规律性.

例如，抛掷一枚硬币，如果硬币是均匀的，当抛掷次数少时，正面朝上或反面朝上没有明显的规律性，随着抛掷次数增大就会发现，出现正面朝上的次数占总次数的 50%左右. 又如，在射击中，当射击次数不大时，靶上命中点的分布是完全没有规则的，杂乱无章的，没有什么显著的规律性；当射击次数增加时，分布就开始呈现一些规律，射击次数越多，规律性越清楚. 也就是说，个别随机现象可能无规律性，但是大量性质相同的随机现象总存在着统计规律性.

必须指出，随机现象与确定性现象之间，必然性与偶然性之间，并没有不可逾越的鸿沟. 事实上，宇宙中没有哪一种实际的现象不带有某种程度的随机因素，因而在处理任何确定性现象时也不可避免地有随机误差产生. 只是在许多实际问题中，为了处理问题的方便，在所求的精确度的范围内，忽略掉了那些造成随机误差的次要因素，而只考虑起主要和基本作用的那些因素，然后利用数学工具建立数学模型，把它解出来，找出在这组基本条件下现象所具有的主要规律. 例如，过去在船舶设计时，就是根据流体力学的原理列出微分方程，求出船舶在航行中所受的阻力以及它的航行规律. 在某些精确度要求较高的问题中，就必须将一些随机的次要因素也考虑在内，因而还必须应用概率论与数理统计的知识.

一方面，概率论与数理统计的理论与研究方法可以提供良好的数学模型，因而应用范围广泛，几乎遍及所有科学技术领域、工农业生产和国民经济的各个部门，如适航性、可靠工程性、气象、水文、地震预报、自动控制、通信工程、管理工程、金融工程等等.

另一方面，概率论与数理统计的理论与研究方法正在向各基础学科、工科数学、经济数学、金融数学等渗透，产生了许多边缘性的应用学科，如信息论、计量经济学等.

正如一位著名作家所表述的：概率论与数理统计学改变了我们关于自然、心智和社会的看法，这些改变是意义深远而且范围广阔的，既改变着权利的结构，也改变着知识的结构，这些改变使现代科学形成.

二、随机试验

为了方便和确切地研究随机现象，我们把对随机现象进行的一次观测或一次实验统称为它的一次试验，并且用大写英文字母 E 等表示之. 以下是试验的几个例子.

例 1 E_1：掷一颗骰子，记录其出现的点数.

E_2：掷一颗骰子，记录其出现的点数为偶数.

E_3：在某批产品中任取一件，检查其是否合格.

E_4：记录某客运车站一天售票的营业额(最小单位为角).

以上试验具有如下几个共同点：

(1) 在相同条件下可重复进行；

(2) 试验前，全部可能的结果是明确的，而且不止一个；

(3) 事先无法确定试验后哪个结果会出现.

我们称具有上述三个特点的试验为**随机试验**(random experiment)，简称**试验**. 通过研究随机试验，我们可以比较深入地研究随机现象.

随机试验 E 的所有可能的结果的集合称为 E 的**样本空间**(sample space)，简记为 S. 样本空间中的元素，即 E 的每个结果称为**样本点**(sample points)，记为 e，则 $S=S(e)$.

例 1 中每个试验的样本空间分别为 $S_1=\{1,2,3,4,5,6\}$，$S_2=\{2,4,6\}$，$S_3=\{$合格，不合格$\}$，$S_4=\{e|e\geqslant 0, 10e\in\mathbf{N}\}$. 样本空间中的样本点取决于试验的目的. 也就是说，由于试验的目的不同，样本空间的样本点也就不同. 如例 1 中的试验 E_1 和 E_2，同是掷骰子，由于观察的目的不同，样本空间 S_1 和 S_2 也就不同.

但是，无论怎样构造样本空间，作为样本空间中的样本点，一定具备两条基本属性：

(1) **互斥性**：无论哪两个样本点都不会在同一次试验中出现；

(2) **完备性**：每次试验中一定会出现某一个样本点.

三、随机事件

在进行随机试验时，人们常关心的是满足某种条件的样本点的集合. 例如，若规定某一客运站的售票营业额少于 300 元为亏损，则人们最为关心的是不小于 300 元的样本点的集合 A. 显然，A 是随机试验 E 的样本空间 S 的一个子集，即 $A=\{e|e\geqslant 300, 10e\in\mathbf{N}\}$.

我们称试验 E 的样本空间 S 的子集为 E 的**随机事件**，简称为**事件**(event). 事件是概率论中的最基本的概念，一般用大写英文字母 $A,B,\cdots$ 表示事件.

因此，某个事件 A 发生当且仅当这个子集中的至少某一个样本点发生. 如上例中，若某天的营业额为 500 元，则事件 A 发生. 特别地，由一个样本点组成的单点集称为**基本事件**(basic event). 例如，试验 E_1 中有 6 个基本事件：$\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$.

特别地，样本空间 S 当然也是 S 的子集，它包含所有的样本点，在每次试验中它总发生，

称为**必然事件**(certain event).在每次试验中都不发生,称为**不可能事件**(impossible event),空集$\varnothing$也是样本空间 S 的一个子集,它不包含任何样本点.

严格地说,必然事件 S 与不可能事件$\varnothing$都不是随机事件,因为作为试验的结果,它们都是确定的,并不具有随机性,但是为了今后讨论问题方便,我们也将它们当作随机事件来处理.

四、事件间的关系与运算

由于随机事件是样本空间的一个子集,而且样本空间中可以定义不止一个事件,那么分析它们之间的关系不但有助于我们深刻的认识事件的本质,而且还可以简化一些复杂事件的讨论.既然事件是一个集合,我们就可以借助集合论中集合之间的关系及运算来研究事件间的关系与运算.

设试验 E 的样本空间为 S,而 $A,B,C,A_i(i=1,2,\cdots,n)$为 E 的事件.

1. 若事件 A 发生必然导致事件 B 发生,则称事件 B **包含**事件 A,记做 $B\supset A$.若 $B\supset A$ 且 $A\supset B$,则称事件 A 与事件 B **相等**,记做 $A=B$.易知,若 $A=B,B=C$,则 $A=C$.

2. 事件 $A\cup B=\{e|e\in A$ 或 $e\in B\}$称为事件 A 与 B 的**和(或并)事件**.

当事件 A,B 中至少有一个发生时,事件 $A\cup B$ 发生.类似地,当 n 个事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生时,称和事件 $A_1\cup A_2\cup\cdots\cup A_n=\bigcup\limits_{i=1}^{n}A_i$ 发生.

对任一事件 A,有 $A\cup S=S,A\cup\varnothing=A$.

3. 事件 $A\cap B=AB=\{e|e\in A$ 且 $e\in B\}$称为事件 A 与 B 的**积(或交)事件**.

当事件 A,B 都发生时,事件 $A\cap B$ 或 AB 发生.类似地,当 n 个事件 $A_1,A_2,\cdots,A_n$ 都发生时,称积(或交)事件 $A_1\cap A_2\cap\cdots\cap A_n=\bigcap\limits_{i=1}^{n}A_i$ 或 $A_1A_2\cdots A_n$发生.

对任一事件 A,有 $A\cap S=A,A\cap\varnothing=\varnothing$.

4. 事件 $A-B=\{e|e\in A$ 且 $e\notin B\}$称为事件 A 与 B 的**差事件**.

当事件 A 发生而事件 B 不发生时,称差事件 $A-B$ 发生.

5. 对于事件 A,B,若 $AB=\varnothing$,则称事件 A 与 B 是**互不相容事件**,或**互斥事件**.n 个事件 $A_1,A_2,\cdots,A_n$ 被称为互不相容的,是指其中任意两个事件都是互不相容的,即

$$A_i\cap A_j=\varnothing\quad(i\neq j,i,j=1,2,\cdots,n).$$

6. 对事件 A,B,若 $A\cup B=S$,且 $A\cap B=\varnothing$,就是说,无论试验的结果如何,事件 A 与 B 中必有且仅有一个发生,则称事件 A 与 B 为**对立事件**,也称为**互补事件**.A 的对立事件记为 $\overline{A}$,即 $\overline{A}=B=S-A$.

注 对立事件 A 与 $\overline{A}$ 必为互不相容(互斥)事件,但互不相容事件不一定是对立事件.

下面用**文氏图**(即表示事件之间的关系或运算的图形,如图 1.3 所示)来直观的表示一

些事件的关系与运算. 其中矩形部分为样本空间 S,圆形部分分别表示事件 A,B.

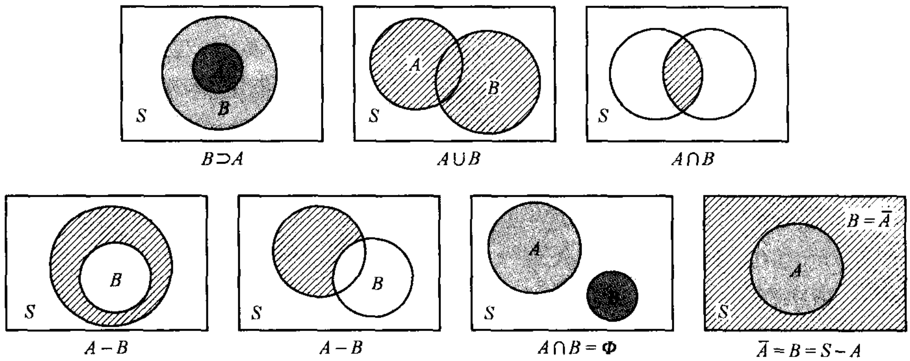

图 1.3

容易验证以下事件的运算规律:

交换律: $A\cup B=B\cup A$, $A\cap B=B\cap A$;

结合律: $A\cup(B\cup C)=(A\cup B)\cup C$, $A\cap(B\cap C)=(A\cap B)\cap C$;

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

德·摩根(De · Morgan)**律**: $\overline{A\cup B}=\overline{A}\cap\overline{B}$,$\overline{A\cap B}=\overline{A}\cup\overline{B}$;

对差事件运算: $A-B=A\overline{B}=A-AB$.

例 2 掷一颗骰子. 设事件 A_1 为"掷出是奇数点",A_2 为"掷出是偶数点",A_3 为"掷出是小于 4 的偶数点",则有 $A_1\cup A_2=\{1,2,3,4,5,6\}$,$A_1\cap A_2=\varnothing$,$A_2\cap A_3=\{2\}$,$A_2-A_3=A_2\cap\overline{A_3}=\{4,6\}$,$\overline{A_1\cup A_3}=\overline{A_1}\cap\overline{A_3}=A_2\cap\overline{A_3}=\{4,6\}$.

例 3 图 1.4 所示的电路中,以 A 表示"信号灯亮"这一事件,以 B,C,D 分别表示"电路接点Ⅰ,Ⅱ,Ⅲ闭合"事件.

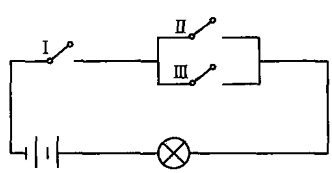

图 1.4

易知 $BC\subset A$,$BD\subset A$,$BC\cup BD=A$,而 $\overline{B}A=\varnothing$,即事件 $\overline{B}$ 与事件 A 互不相容.

例 4 化简下列事件:(1) $AB\cup A\overline{B}$;(2) $A\overline{B}\cup\overline{A}B\cup\overline{A}\overline{B}$.

解 (1) $AB\cup A\overline{B}=A\cap(B\cup\overline{B})=AS=A$;

(2)
$$\begin{aligned}A\overline{B}\cup\overline{A}B\cup\overline{A}\overline{B}&=A\overline{B}\cup\overline{A}B\cup\overline{A}\overline{B}\cup\overline{A}\overline{B}=(A\overline{B}\cup\overline{A}\overline{B})\cup(\overline{A}B\cup\overline{A}\overline{B})\\&=(A\cup\overline{A})\overline{B}\cup(B\cup\overline{B})\overline{A}=S\overline{B}\cup S\overline{A}=\overline{B}\cup\overline{A}=\overline{AB}.\end{aligned}$$

五、概率的定义与性质

由随机事件的特性可知，在一次试验中，它可能发生也可能不发生（除必然事件和不可能事件外）. 我们希望对某一事件出现的可能性的大小给出一个度量，这就是概率. 历史上关于概率的定义有许多，这里只介绍其中的两种.

（1）概率的统计定义

定义 1　在相同条件下，随机事件 A 在进行 n 次独立重复试验中出现的次数称为事件 A 发生的**频数**，记为 n_A. 比值 $n_A/n=f_n(A)$ 称为事件 A 发生的**频率**（frequency）. 当试验次数 n 逐渐增加时，频率在某一常数 p 附近摆动并逐渐稳定于该常数，将频率的这个稳定值 p 称为事件 A 的**概率**，记做 $P(A)=p$.

尽管上述定义非常直观地解释了事件概率的含义，但是人们无法通过将试验无限地重复下去来得到事件的概率，于是就有了概率的古典定义、几何定义和主观定义等，但各有局限性，直到 1933 年俄国数学家柯尔莫哥洛夫提出概率的公理化定义，并得到举世公认.

（2）概率的公理化定义

定义 2　设 E 为随机试验，S 为其样本空间. 对于 E 的每一个事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的**概率**，这里 $P(A)$ 应满足以下三个条件：

（1）**非负性**：对于每个事件 A，有 $P(A)\geqslant 0$；

（2）**规范性（正则性）**：对必然事件 S，有 $P(S)=1$；

（3）**可列可加性**：若 $A_1,A_2,\cdots$ 是两两互不相容的事件，即 $A_iA_j=\varnothing, i\neq j, i,j=1,2,\cdots$，则

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

注　$P(A)$ 可以看做是 A 的函数，与 A 的频率不同.

由概率的公理化定义可以证明概率的一些重要性质.

性质 1　$P(\varnothing)=0$.

证　因为 $\varnothing=\varnothing\cup\varnothing\cup\cdots$，由概率的可列可加性 $P(\varnothing)=P(\varnothing)+P(\varnothing)+\cdots$ 及非负性 $P(\varnothing)\geqslant 0$，证得 $P(\varnothing)=0$.

性质 2（有限可加性）　若 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件，则

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

证　令 $A_i=\varnothing, i=n+1,n+2,\cdots$，由概率的可列可加性及非负性即得.

性质 3　若事件 A,B 满足 $A\subset B$，则有

$$P(B-A)=P(B)-P(A),\quad P(B)\geqslant P(A).$$

证　由条件有 $B=A\cup(B-A)$，且 $A\cap(B-A)=\varnothing$，由性质 2 知 $P(B)=P(A)+P(B-A)$，故有 $P(B-A)=P(B)-P(A)$，又 $P(B-A)\geqslant 0$，故 $P(B)\geqslant P(A)$.

性质 4　对任意事件 A，$P(\bar{A})=1-P(A)$，其中 $\bar{A}$ 为 A 的对立事件.

证 由于 $A\cup\overline{A}=S$ 且 $A\overline{A}=\varnothing$，故 $P(A)+P(\overline{A})=1$，从而 $P(\overline{A})=1-P(A)$.

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

证 由 $A\cup B=A\cup(B-AB)$，又 $A\cap(B-AB)=\varnothing$，且 $AB\subset B$，于是有

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

特别地，当 A,B 互不相容时，$P(A\cup B)=P(A)+P(B)$.

性质5可以推广到多个事件的情形. 比如，对任意三个随机事件 A_1,A_2,A_3，有如下加法公式：

$$\begin{aligned}P(A_1\cup A_2\cup A_3)=&P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)\\&-P(A_2A_3)-P(A_3A_1)+P(A_1A_2A_3).\end{aligned}$$

例5 设 A,B 为两事件，且 $P(B)=0.3$，$P(A\cup B)=0.6$，求 $P(A\overline{B})$.

解 $P(A\overline{B})=P(A-AB)=P(A)-P(AB)$，而 $P(A\cup B)=P(A)+P(B)-P(AB)$，则 $P(A)-P(AB)=P(A\cup B)-P(B)$，从而 $P(A\overline{B})=P(A\cup B)-P(B)=0.6-0.3=0.3$.

例6 (1) 设有两事件 A,B 满足 $P(A)=P(B)=0.5$，证明 $P(AB)=P(\overline{A}\overline{B})$；

(2) 证明对任意两事件 A,B 有 $P(A)+P(B)-1\leqslant P(AB)\leqslant P(A\cup B)$.

证 (1) $P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-[P(A)+P(B)-P(AB)]$

$$=1-[0.5+0.5-P(AB)]=P(AB);$$

(2) 因为 $AB\subset A\cup B$，所以 $P(AB)\leqslant P(A\cup B)$，又 $1\geqslant P(A\cup B)=P(A)+P(B)-P(AB)$，于是 $P(AB)\geqslant P(A)+P(B)-1$，从而 $P(A)+P(B)-1\leqslant P(AB)\leqslant P(A\cup B)$.

习 题 1.1

1. 什么是随机现象？试举出三个随机现象的例子.

2. 什么是随机试验？请举出三个随机试验的实例，并写出它们的样本空间.

3. 用事件 A,B,C 的运算关系表示下列事件：

(1) A 发生，B 与 C 不发生；　(2) A,B,C 中至少有一个发生；

(3) A 与 B 都发生而 C 不发生；　(4) A,B,C 都发生；

(5) A,B,C 都不发生；　(6) A,B,C 中至少有两个发生；

(7) A,B,C 中不多于两个发生.

4. 下列结论中哪些成立，哪些不成立？

(1) $A\cup B=A\overline{B}\cup B$；　(2) $\overline{A}B=A\cup B$；　(3) $(AB)(A\overline{B})=\varnothing$；

(4) $\overline{A\cup B\cup C}=\overline{A}\overline{B}\overline{C}$；　(5) 若 $A\subset B$，则 $A=AB$；　(6) 若 $AB=\varnothing$ 且 $C\subset A$，则 $BC=\varnothing$；

(7) 若 $A\subset B$，则 $\overline{B}\subset\overline{A}$；(8) 若 $B\subset A$，则 $A\cup B=A$.

5. 指出下列结论在什么情况下成立：

(1) $AB=B$；　(2) $A\cup B=B$；　(3) $A\cup B\cup C=B$.