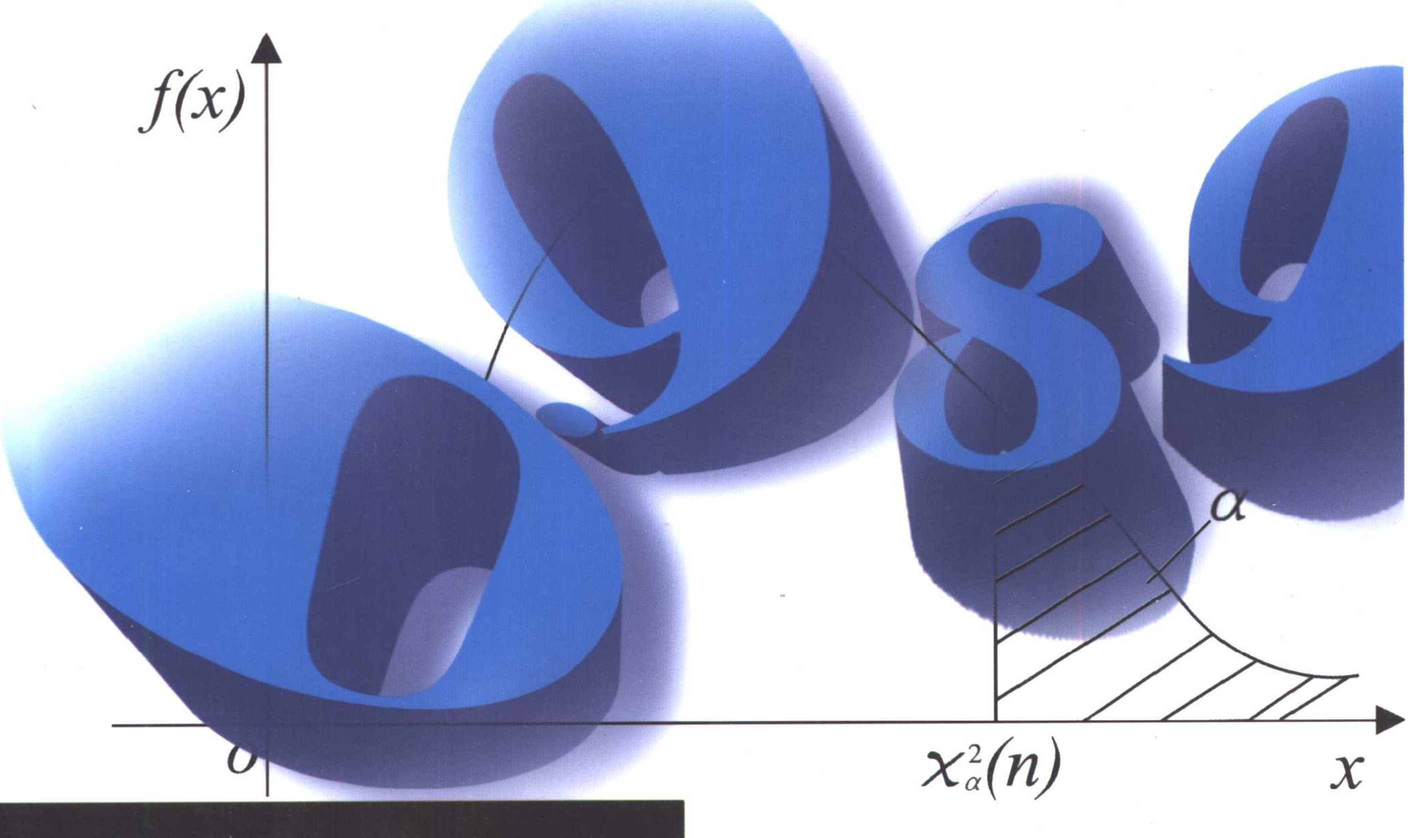
电子·教育

高职高专计算机系列教材

中国计算机学会高职高专教育学会推荐出版

概率与数理统计

季夜眉　吴大贤　等编著



電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
URL: http://www.phei.com.cn

概率与数理统计

高职高专计算机系列教材

概率与数理统计

季夜眉　吴大贤　等编著

電子工業出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

《概率与数理统计》是高职高专工程专科教育的一门技术基础课教材。本书按照高职高专教育对数学课程教学的基本要求，适度精简了概率基础部分的教学内容，在取材上以数理统计为主体。全书在阐述基本的数学思想和原理的基础上，重点介绍在各个领域被广泛运用的参数估计、假设检验、方差分析、回归分析等统计推断方法和数据处理以及正交试验设计方法，突出对学生运用数理统计方法表述和解决实际问题能力的培养。本书增设了课程实验，这些实验可以运用 MATLAB 软件很方便地在计算机上实现。

本书可作为高职高专院校计算机专业基础课教材，对工科其他专业及管理、财经类专业也都适用；还可作为工程技术人员和管理人员参考用书。

图书在版编目（CIP）数据

概率与数理统计/季夜眉等编著 . – 北京：电子工业出版社，2001.9
高职高专计算机系列教材
ISBN 7-5053-6692-0

Ⅰ. 概 ... Ⅱ. 季 ... Ⅲ. ①概率论 – 高等学校：技术学校 – 教材②数理统计 – 高等学校：技术学校 – 教材 Ⅳ.021

中国版本图书馆 CIP 数据核字（2001）第 057697 号

丛 书 名：高职高专计算机系列教材
书　　名：**概率与数理统计**
编 著 者：季夜眉　吴大贤　等
责任编辑：张孟玮
特约编辑：胡国清
排版制作：电子工业出版社计算机排版室
印 刷 者：北京天宇星印刷厂
装 订 者：河北省涿州桃园装订厂
出版发行：电子工业出版社　URL：http：//www.phei.com.cn
　　　　　北京市海淀区万寿路 173 信箱　邮编 100036
经　　销：各地新华书店
开　　本：787×1092　1/16　印张：10　字数：256 千字
版　　次：2001 年 9 月第 1 版　2001 年 9 月第 1 次印刷
书　　号：ISBN 7-5053-6692-0 / G·549
印　　数：6 000 册　　定价：13.00 元

出版说明

高职高专的计算机专业面临着两方面的巨大变化，一是计算机技术的飞速发展，另一方面是高职高专教育本身的改革和重组。

当前，计算机技术正经历着高速度、多媒体网络化的发展，计算机教育特别是计算机专业的教材建设必须适应这种日新月异的形势，才能培养出不同层次的合格的计算机技术专业人才。为了适应这种变化，国内外都在对计算机教育进行深入的研究和改革。美国 IEEE 和 ACM 在推出了《Computing Curricula 2000》之后，立即又推出了《Computing Curricula 2001》。全国高校计算机专业教学指导委员会和中国计算机学会教育委员会在 1999 年 9 月也提出了高等院校《计算机学科教学计划 2000》(征求意见稿)。目前，国内许多院校老师、专家正在研究《Computing Curricula 2001》，着手 21 世纪的中国计算机教育的改革。

高专层次和本科层次的计算机教育既有联系又有区别，高职高专的计算机教育旨在培养应用型人才。自 20 世纪 70 年代末高等专科学校计算机专业相继成立以来，高等专科学校积极探索具有自己特色的教学计划和配套教材。1985 年，在原电子工业部的支持下，由全国数十所高等专科学校参加成立了中国计算机学会教育委员会大专教育学会，之后又成立了大专计算机教材编委会。从 1986 年到 1999 年，在各校老师的共同努力下，已相继完成了三轮高等专科计算机教材的规划与出版工作，共出版了 78 种必修课、选修课、实验课教材，较好地解决了高专层次计算机专业的教材需求。

为了适应计算机技术的飞速发展以及高职高专计算机教育形势发展的需要，中国计算机学会教育委员会高职高专教育学会和高职高专计算机教材编委会于 2000 年 7 月开始，又组织了一批本科高校、高等专科学校、高等职业技术院校和成人高等院校的有教学经验的老师，学习研究参考了高等院校《计算机学科教学计划 2000》(征求意见稿)，提出了按照新的计算机教育计划和教学改革的要求，编写高专、高职、成人高等教育三教统筹的第四轮教材。

第四轮教材的编写工作采取了以招标的方式征求每门课程的编写大纲和主编，要求投标老师详细说明课程改革的思路、本课程和相关课程的联系、重点和难点的处理等。在第四轮教材的编写过程中，编委会强调加强实践环节、强调三教统筹、强调理论够用为度的原则，要求教学计划、教学内容适应高等教育发展的新形势。本套教材的编者均为各院校具有丰富教学实践经验的教师。因此，第四轮教材的特点是体系结构比较合理、内容新颖、概念清晰、通俗易懂、理论联系实际、实用性强。

竭诚希望广大师生对本套教材提出批评建议。

中国计算机学会教育委员会高职高专教育学会

2001 年 1 月

先后参加中国计算机学会教育委员会高职高专教育学会和高职高专计算机教材编委会学术活动的部分学校名单

山西师范大学
河北师范大学
承德石油高等专科学校
河北大学
保定职业技术学院
北京科技大学
北京市机械工业管理局职工大学
北方工业大学
北京船舶工业管理干部学院
海淀走读大学
北京信息工程学院
中国人民大学
北京师范大学
沈阳电力高等专科学校
辽宁交通高等专科学校
吉林大学
吉林职业师范学院
黑龙江大学
哈尔滨工业大学
哈尔滨师范大学
上海理工大学
上海第二工业大学
上海交通大学
上海商业职业技术学院
上海电机技术高等专科学校
上海旅游高等专科学校
金陵职业大学
南京建筑工程学院
南京工程学院
南京师范大学
常州工学院
无锡职业技术学院
苏州市职工大学
空军后勤学院
连云港化工高等专科学校
泰州职业技术学院
潍坊高等专科学校
青岛化工学院
天津轻工业学院
浙江大学
宁波高等专科学校
福州大学
重庆电子职业技术学院
湖南大学
湖南计算机高等专科学校
中国保险管理干部学院
湖南税务高等专科学校
长沙大学
湖南财经高等专科学校
邵阳高等专科学校
江汉大学
中国地质大学
武汉职业技术学院
河南职业技术学院
平原大学
安阳大学
开封大学
洛阳大学
河南大学
广州市财贸管理干部学院
广东轻工职业技术学院
广州航海高等专科学校
韶关大学
佛山科技学技术学院
南宁职业技术学院
广西水利电力职业技术学院
桂林电子工业学院
柳州职业技术学院
成都电子机械高等专科学校
电子科技大学
成都师范高等专科学校
四川师范学院
云南财贸学院
西安电子科技大学
兰州石化职业技术学院
兰州师范高等专科学校

前　言

概率与数理统计是研究随机现象客观规律的数学学科。它的应用非常广泛，几乎存在于人类活动的所有领域中，并且有其独特的思维方式、研究方法和理论。概率与数理统计也因此成为高职高专教育工科各专业的一门必修的技术基础课。

根据《教育部关于加强高职高专教育人才培养工作的意见》，在全国计算机专业教学指导委员会的统一组织下，围绕培养高等技术应用型人才的目标，遵循突出应用性和实践性的原则，结合多年的教学工作实践，我们编写了这本教材。

作为一门技术基础课应该突出基础理论知识的应用和有利于实践能力的培养。因此，我们在教材内容的安排上，采用以数理统计为主干，精简了概率基础部分的教学内容，主要介绍与数理统计方法密不可分的基础知识；对于数理统计部分，也努力做到在讲清基本数学思想和基本数学原理的基础上，以介绍常用的统计推断方法为主。相应地，例题、习题的选配均以帮助学生理解数学概念、理解数学建模思想、掌握常用统计推断方法为目的。

结合计算机应用技术的普及和发展，为了增强学生运用数理统计方法解决实际问题的能力，本教材特意安排了课程实验。这些实验均可以通过运用 MATLAB 软件，很方便地在计算机上实现。

本书以季夜眉、吴大贤为主编著。第 1,2,3,4 章由季夜眉执笔，第 5 章由周家全执笔，第 6,7,8 章由吴大贤执笔，第 9 章由张荣执笔，第 10 章由王宇翔执笔。另外，王思维、黄薇精心地为本书绘制了全部插图。

全书由朱乃立教授主审。

由于对高职高专教育课程和教学内容体系改革的探索还不够，本教材一定有不当之处，恳请同行教师和读者不吝赐教，批评指正。

编著者

2001 年 4 月

目　录

第1章　随机事件与概率 …… (1)

1.1　概率的基本概念 …… (1)

1.1.1　随机事件 …… (1)

1.1.2　事件之间的关系和运算 …… (2)

1.1.3　事件的概率 …… (4)

1.2　概率的运算法则 …… (7)

1.2.1　条件概率与事件的独立性 …… (7)

1.2.2　全概率公式和逆概率公式 …… (9)

习题1 …… (11)

第2章　一维随机变量 …… (13)

2.1　一维随机变量及其分布 …… (13)

2.1.1　随机变量 …… (13)

2.1.2　随机变量的分布函数 …… (13)

2.2　离散型随机变量及其分布律 …… (14)

2.2.1　离散型随机变量及其分布律 …… (14)

2.2.2　常用的离散型随机变量的分布 …… (16)

2.3　连续型随机变量及其概率密度 …… (19)

2.3.1　连续型随机变量及其概率密度 …… (19)

2.3.2　常用的连续型随机变量的分布 …… (21)

2.4　随机变量的函数的分布 …… (26)

2.4.1　离散型随机变量的函数的分布 …… (26)

2.4.2　连续型随机变量的函数的分布 …… (27)

习题2 …… (28)

第3章　二维随机变量 …… (31)

3.1　二维随机变量及其分布 …… (31)

3.1.1　二维随机变量及分布函数 …… (31)

3.1.2　二维离散型随机变量的联合分布律 …… (32)

3.1.3　二维连续型随机变量的联合概率密度 …… (33)

3.1.4　两个常用的二维随机变量的分布 …… (34)

3.2　边缘分布与独立性 …… (35)

3.2.1　二维连续型随机变量的边缘分布 …… (35)

3.2.2　随机变量的独立性 …… (36)

习题3 …… (39)

第4章　随机变量的数字特征 …… (40)

4.1　数学期望 …… (40)

4.1.1 离散型随机变量的数学期望 …… (40)
4.1.2 连续型随机变量的数学期望 …… (41)
4.1.3 随机变量的函数的数学期望 …… (42)
4.1.4 数学期望的性质 …… (43)
4.2 方差 …… (45)
4.2.1 方差的概念 …… (45)
4.2.2 方差的性质 …… (45)
4.3 常用分布的数学期望与方差 …… (46)
4.4 极限定理 …… (48)
4.4.1 大数定律 …… (49)
4.4.2 中心极限定理 …… (49)
习题4 …… (51)
第5章 样本分布与数据处理 …… (54)
5.1 简单随机样本 …… (54)
5.2 样本数据的初步统计分析 …… (54)
5.3 常用统计量及其分布 …… (56)
5.3.1 统计量 …… (57)
5.3.2 常用统计量的分布 …… (57)
习题5 …… (60)
第6章 参数估计 …… (62)
6.1 点估计 …… (62)
6.1.1 数字特征估计法 …… (62)
6.1.2 极大似然估计法 …… (63)
6.2 估计量的评选标准 …… (64)
6.2.1 无偏性 …… (64)
6.2.2 有效性 …… (65)
6.3 区间估计 …… (66)
6.3.1 置信区间 …… (66)
6.3.2 正态总体均值与方差的置信区间 …… (66)
6.3.3 其他 …… (69)
习题6 …… (70)
第7章 假设检验 …… (72)
7.1 假设检验的基本概念 …… (72)
7.1.1 问题的提出 …… (72)
7.1.2 假设检验方法 …… (72)
7.1.3 假设检验中的两类错误 …… (74)
7.2 单个正态总体均值与方差的假设检验 …… (75)
7.2.1 正态总体均值的假设检验 …… (75)
7.2.2 正态总体方差的假设检验 …… (77)
7.3 两个正态总体均值与方差的假设检验 …… (80)

7.3.1 两个正态总体均值相等的假设检验 …………………………………… (80)
7.3.2 两个正态总体方差相等的假设检验 …………………………………… (83)
7.4 总体分布的假设检验 …………………………………… (85)
习题 7 …………………………………… (88)
第 8 章 方差分析与回归分析 …………………………………… (90)
8.1 方差分析 …………………………………… (90)
8.1.1 问题的提出 …………………………………… (90)
8.1.2 单因素方差分析 …………………………………… (91)
8.2 回归分析 …………………………………… (95)
8.2.1 一元线性回归 …………………………………… (95)
8.2.2 线性相关关系的显著性检验 …………………………………… (98)
8.2.3 预测与控制 …………………………………… (99)
8.2.4 可线性化的一元非线性回归 …………………………………… (100)
习题 8 …………………………………… (104)
第 9 章 正交试验设计 …………………………………… (106)
9.1 正交表简介 …………………………………… (106)
9.2 无交互作用的正交试验设计 …………………………………… (107)
9.3 有交互作用的正交试验设计 …………………………………… (109)
习题 9 …………………………………… (111)
第 10 章 课程实验 …………………………………… (113)
10.1 预备实验——MATLAB 使用练习 …………………………………… (113)
10.1.1 MATLAB 软件简介 …………………………………… (113)
10.1.2 MATLAB 基本用法 …………………………………… (113)
10.2 课程实验 …………………………………… (116)
10.2.1 概率分布和数字特征 …………………………………… (116)
10.2.2 参数估计 …………………………………… (122)
10.2.3 假设检验 …………………………………… (125)
10.2.4 方差分析与回归分析 …………………………………… (126)
习题答案 …………………………………… (132)
附录 常用统计数表 …………………………………… (138)
附表 1 泊松分布表 …………………………………… (138)
附表 2 标准正态分布表 …………………………………… (139)
附表 3 χ^2 分布表 …………………………………… (140)
附表 4 t 分布表 …………………………………… (141)
附表 5 F 分布表 …………………………………… (142)
附表 6 相关系数显著性检验表 …………………………………… (147)
附表 7 正交表 …………………………………… (148)

第1章　随机事件与概率

自然现象与社会现象千差万别，但从概率论的观点考察可以分为两大类。一类有明确的因果关系，即在一定条件下必然发生或必然不发生的现象，称为确定性现象。例如，在一个标准大气压下，纯水加热到100℃一定沸腾；上抛的石子一定要落下来等等。另一类称为随机现象，它的因果关系不明显，或根本没有什么因果关系，即是在一定条件下可能发生，也可能不发生的现象。例如：投掷一枚质地均匀的硬币，可能正面（国徽一面）朝上，也可能反面（有数字一面）朝上；远距离射击一个目标，可能击中，也可能击不中等等。随机现象有两个特点：(1) 在一次观察中，现象可能发生，也可能不发生，即结果呈现不确定性；(2) 在大量重复观察中，其结果具有统计规律性。例如，多次重复投掷一枚硬币，就会发现出现正面与反面的次数几乎各占一半。概率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。

1.1　概率的基本概念

1.1.1　随机事件

1. 随机试验

在一定条件下，对随机现象加以研究所进行的观察或实验，称为**试验**。

如果一个试验满足下列三个条件：

(1) 可以在相同条件下重复进行；

(2) 每次试验的可能结果不止一个，并且事先能明确知道试验的所有可能结果；

(3) 在一次试验之前不能确定哪一个结果一定出现。

则称这一试验为**随机试验**，记为 E。

【例1.1】　下列试验都是随机试验。

E_1：掷一枚质地均匀的硬币，观察它出现正面或反面的情形；

E_2：掷一枚骰子，观察其朝上的那一面的点数；

E_3：在一批产品中，任取一件，考察是正品还是次品；

E_4：从一批灯泡中，任取一只，测试其寿命；

E_5：掷两枚质地均匀的硬币，观察它们出现正面或反面的情形。

2. 随机事件与样本空间

在随机试验中，可能发生也可能不发生的结果，称为**随机事件**，简称**事件**。记为 A 或 B 或 C 等。我们将通过考察随机事件研究随机现象。

在每次试验中一定出现的事件称为**必然事件**，记为 Ω。在每次试验中一定不出现的事

件称为**不可能事件**，记为 Ø。必然事件与不可能事件都是确定性的事件，它们不是随机事件，只是为了今后讨论方便，把它们当作一类特殊的随机事件。

在一个随机试验中，不论可能发生的结果有多少，总可以从中找出这样一组基本结果（不能再分的），满足：

(1) 每进行一次随机试验，必然出现而且只能出现其中的一个基本结果；

(2) 任何事件，都是由其中的一些基本结果组成的。

这样的基本结果称为**基本事件**，或称为**样本点**，记为 e。

随机试验 E 的所有基本事件（样本点）组成的集合称为随机试验 E 的**样本空间**，记为 Ω。

在每个随机试验中，确定样本空间是至关重要的。这样，一个基本事件（样本点）e 可视为样本空间 Ω 中的一个元素，即 $e\in\Omega$。一个随机事件 A 由一个或若干个基本事件构成，视它为样本空间 Ω 的一个子集合，有 $e\in A$，$A\subset\Omega$ 等等。

所谓**事件 A 发生**，是指在一次试验中，当且仅当 A 中包含的某个基本事件出现。

样本空间 Ω 包含所有的基本事件，每次试验中它必然发生，它是一个必然事件，故样本空间也用 Ω 表示。不可能事件是样本空间的一个空集，故不可能事件用 Ø 表示。

【例 1.2】 写出例 1.1 中各随机试验的样本空间并举出一些随机事件的例子。

E_1：$\Omega_1=$：{正面，反面}；

E_2：$\Omega_2=\{1, 2, 3, 4, 5, 6\}$；

E_3：$\Omega_3=$ {正品，次品}；

E_4：$\Omega_4=\{t\mid t\geqslant 0\}$；

E_5：$\Omega_5=$ {（正面，正面），（正面，反面），（反面，正面），（反面，反面）}。

E_2 中的 $e=$ "点数为" $6=\{6\}$ 为基本事件；$A=$ "点数为偶数" $=\{2, 4, 6\}$ 与 $B=$ "点数小于 5" $=\{1, 2, 3, 4\}$ 为随机事件，$\varnothing=$ "点数大于 6" 为不可能事件等等；E_4 中 $C=\{t\mid 500\leqslant t\leqslant 550\}$ 表示随机事件"灯泡寿命为（500～550）h"，等等。

1.1.2 事件之间的关系和运算

我们把随机事件定义为样本点的某个集合，就能方便地将集合论的全部知识用来解释事件之间的关系及其运算（见图 1.1），也就是能用简单事件来表示较为复杂的事件。

(1) 包含关系：如果事件 A 发生必然导致事件 B 发生，即 A 中的样本点一定属于 B，则称事件 B **包含**事件 A，也称事件 A **包含于**事件 B；或称事件 A 为事件 B 的**子事件**，记为 $A\subset B$（或 $B\supset A$）。如在例 1.2 中 $\{6\}\subset A$。任何事件都包含于 Ω。

(2) 相等关系：如果 $A\subset B$ 且 $B\subset A$，则称事件 A、B **相等**，记为 $A=B$。

(3) 事件的和（并）：事件 A 与事件 B 中至少有一个发生的事件，称为事件 A 与事件 B 的**和**，也称为事件 A 与事件 B 的**并**，记为 $A\cup B$ 或 $A+B$。事件 A、B 的和是由 A 与 B 的样本点合并而成的事件。

n 个事件的和为 $A_1\cup A_2\cup\cdots\cup A_n$，记为 $\bigcup\limits_{k=1}^{n}A_k$；无穷可列个事件的和记为 $\bigcup\limits_{k=1}^{+\infty}A_k$。

(4) 事件的集（交）：事件 A 与事件 B 同时发生的事件，称为事件 A 与事件 B 的**集**，也称为事件 A 与事件 B 的**交集**，记为 $A\cap B$ 或 AB。事件 A、B 的集是由 A 与 B 的公共样本点所构成的事件。

n 个事件的集为 $A_1 \cap A_2 \cap \cdots \cap A_n$，记为 $\bigcap\limits_{k=1}^{n} A_k$ ；无穷可列个事件的集记为 $\bigcap\limits_{k=1}^{+\infty} A_k$ 。

(5) 事件的差：事件 A 发生而事件 B 不发生的事件，称为事件 A 与事件 B 的**差**，记为 $A-B$。事件 A 与 B 的差是由属于 A 而不属于 B 的样本点所构成的事件。

(6) 事件的互斥（互不相容）：如果事件 A 与事件 B 不能同时发生，即 $A \cap B = \varnothing$，则称事件 A 与事件 B **互斥**，也称事件 A 与事件 B **互不相容**。如果一组事件 $A_1, A_2, \cdots, A_n, \cdots$ 中任意两个都互斥，则称这组事件**两两互斥**。显然，基本事件是两两互斥的。

(7) 对立事件：如果两个事件 A 与 B 满足 $A \cup B = \Omega$，且 $AB = \varnothing$，则称事件 A 与事件 B 互为**对立事件**，亦称事件 A 与事件 B 互为**逆事件**，记为 $B = \overline{A}$ 或 $A = \overline{B}$。事件 A 与事件 B 对立，是指事件 A 与事件 B 既不能同时发生又不能同时不发生，即在每次试验中，A 与 B 中有且仅有一个发生。

由定义可知，对立事件必为互斥事件。反之，互斥的两个事件未必为对立事件。

事件的运算与集合的运算具有完全相同的规则：

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

(4) 德·摩根律：$\overline{A \cup B} = \overline{A}\,\overline{B}$，$\overline{AB} = \overline{A} \cup \overline{B}$（可以推广到任意多个的情形）。

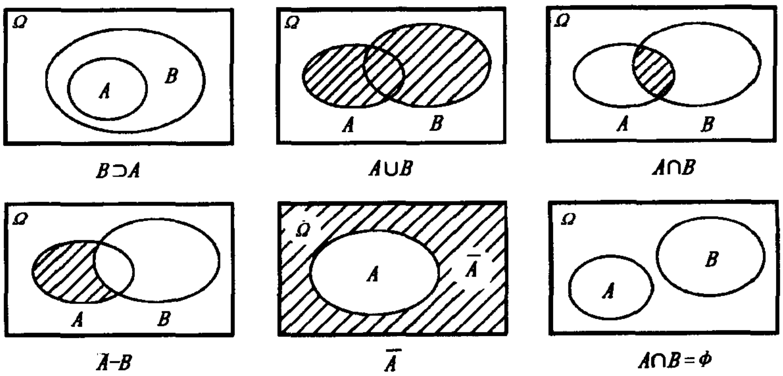


图 1.1

【例 1.3】 设 A，B，C 为三个事件，试用 A，B，C 表示下列事件：

(1) A，B，C 中至少有一个发生；

(2) A，B，C 中恰好有一个发生；

(3) A，B，C 中至多有一个发生；

(4) A 与 B 都发生而 C 不发生；

(5) A，B，C 中恰好有两个发生；

(6) A，B，C 中不多于两个发生；

(7) A 发生而 B 与 C 都不发生；

(8) A，B 中至少有一个发生而 C 不发生。

解

(1) $A \cup B \cup C$，或 $A\overline{BC} \cup \overline{ABC} \cup \overline{AB}C \cup \overline{ABC} \cup \overline{AB}C \cup \overline{A}BC \cup ABC$；

(2) $A\overline{BC} \cup \overline{ABC} \cup \overline{AB}C$；

(3) $A\overline{BC} \cup \overline{A}B\,\overline{C} \cup \overline{AB}C \cup \overline{ABC}$；

(4) $AB\overline{C}$，或 $AB-C$；

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

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

(7) $A\overline{B}\,\overline{C}$，或 $A-B-C$；

(8) $(A\cup B)\overline{C}$，或 $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup AB\overline{C}$。

1.1.3 事件的概率

在实际问题中，知道随机事件是否发生是很重要的，但常常更关心随机事件发生的可能性的大小。随机事件发生的可能性的大小可以用一个确定的数来表示，这个数就叫随机事件 A 的概率，记为 $P(A)$。

1. 频率与概率

定义 1.1 在相同的条件下，进行了 n 次重复试验，若随机事件 A 在这 n 次试验中发生了 n_A 次，则比值 $f_n(A)=\frac{n_A}{n}$ 称为事件 A 发生的**频率**。

频率具有**性质**：

(1) 对任一事件 A，有 $0\leqslant f_n(A)\leqslant 1$；

(2) 对必然事件 Ω，有 $f_n(\Omega)=1$；对不可能事件 $\varnothing$，有 $f_n(\varnothing)=0$；

(3) 对任意事件 A、B，有 $f_n(A\cup B)=f_n(A)+f_n(B)-f_n(AB)$；

特别，若事件 A、B 互斥，则 $f_n(A\cup B)=f_n(A)+f_n(B)$。

对两两互斥事件 A_1，A_2，…，A_n，有 $f_n(\bigcup_{i=1}^{k}A_i)=\sum_{i=1}^{k}f_n(A_i)$。

【例 1.4】 历史上曾有不少著名的科学家分别做一个实验：“大量重复投掷一枚质地均匀的硬币，观察它出现正面或反面朝上的次数。”表 1.1 是他们所做实验结果的部分记录，其中规定均匀硬币正面向上为事件 A 发生。

表 1.1

实验者	投掷次数 n	出现正面的次数 n_A	频率 $f_n(A)=n_A/n$
蒲　丰	4040	2048	0.5069
费　勒	10000	4970	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5006

事件 A 发生的频率表示事件 A 发生的频繁程度，频率越大，说明事件 A 发生得越频繁。但是，从表 1.1 可以看出，频率具有随机波动性，也就是说 $f_n(A)=\frac{n_A}{n}$ 不是一个常数。然而，这种波动也不是无规律可言的，随着 n 逐渐增大，频率的波动性逐渐消失而呈现出稳定趋势，也就是逐渐稳定于某个常数 p。频率的这一性质，称为频率的**稳定性**。正如表 1.1 最后一列所表明的，投掷硬币出现正面的频率 $f_n(A)=\frac{n_A}{n}$ 总在 0.5 附近波动，随着试验次数 n 的增加，它逐渐稳定于 0.5。这个 0.5 反映了投掷硬币正面出现的可能性的大小，也就是上面所说的某个常数 p。每个事件都有这样一个常数 p 与之对应。

因而，我们将事件 A 的频率 $f_n(A)=\frac{n_A}{n}$ 随 n 增大时所逐渐趋于稳定的那个常数 p 定义为事件 A 发生的**概率**，记为 $P(A)=p$。这是**概率的统计定义**。

显然，当 n 相当大时，可以用事件 A 的频率 $f_n(A)$ 作为事件 A 发生的概率 $P(A)$ 的近似值，即 $P(A)\approx f_n(A)$。用频率近似代替概率的好处是便于实际应用。

由概率的统计定义和频率的有关性质，容易得到概率具有下列性质：

(1) 对任一事件 A，有 $0\leqslant P(A)\leqslant 1$; (1.1.1)

(2) $P(\Omega)=1$; $P(\varnothing)=0$; (1.1.2)

(3) 加法定理：

对任意事件 A，B，有 $P(A\cup B)=P(A)+P(B)-P(AB)$; (1.1.3)

特别，若事件 A，B 互斥，则 $P(A\cup B)=P(A)+P(B)$; (1.1.4)

对两两互斥事件 A_1，A_2，…，A_n，有 $P(\bigcup\limits_{i=1}^{k}A_i)=\sum\limits_{i=1}^{k}P(A_i)$; (1.1.5)

(4) $P(A)=1-P(\bar{A})$; (1.1.6)

(5) 若 $A\supset B$，则 $P(A-B)=P(A)-P(B)$; $P(A)\geqslant P(B)$。 (1.1.7)

【例 1.5】 某城市只发行 A、B 两种报纸，订阅 A 报的住户数占总住户数的 70%，订阅 B 报的住户数占总住户数的 50%，同时订阅两报的住户数占总住户数的 30%。求下列事件的概率：

(1) 只订阅 A 报；

(2) 至少订阅一种报纸；

(3) 不订阅任何报纸；

(4) 只订阅一种报纸。

解 设 $A=$ "订阅 A 报"，$B=$ "订阅 B 报"，根据题设有

$$P(A)=0.70,\ P(B)=0.50,\ P(AB)=0.30$$

(1) $P(A\bar{B})=P(A-B)=P(A-AB)=P(A)-P(AB)=0.40$

(2) $P(A\cup B)=P(A)+P(B)-P(AB)=0.7+0.5-0.3=0.9$

(3) $P(\bar{A}\bar{B})=P(\overline{A\cup B})=1-P(A\cup B)=0.1$

(4) $\because$ $A\bar{B}$ 与 $\bar{A}B$ 互斥

$$\begin{aligned}\therefore P(A\bar{B}\cup\bar{A}B) &= P(A\bar{B})+P(\bar{A}B)\\ &=P(A)-P(AB)+P(B)-P(AB)\\ &=0.7-0.3+0.5-0.3=0.6\end{aligned}$$

2. 古典概型

设随机试验 E 满足下列条件：

(1) 试验的样本空间只有有限个基本事件（样本点），$\Omega=\{e_1, e_2, e_3, \cdots, e_n\}$;

(2) 每个基本事件（样本点）的发生是等可能的，即

$$P(e_1)=P(e_2)=p(e_3)=\cdots=P(e_n)$$

则称此类随机试验的数学模型为**古典概型**，也称为**等可能概型**。

在古典概型中，对随机试验 E，若样本空间的基本事件（样本点）总数为 n，事件 A 所包含的基本事件（样本点）数为 m，则由于等可能性，事件 A 的概率的计算公式为

$$P(A)=\frac{m}{n} \tag{1.1.8}$$

【例 1.6】 掷两颗骰子,求两颗同时出现奇数点的概率。

解 这是古典概型。

设 A =“掷两颗骰子,两颗同时出现奇数点”,显然

$$A=\{(1,1),(1,3),(1,5),(3,1),(3,3),(3,5),(5,1),(5,3),(5,5)\}$$

样本空间的基本事件总数为 $n=6^2=36$,事件 A 所包含的基本事件数为 $m=9$(或 $m=C_3^1C_3^1=9$),由(1.1.8)式得

$$P(A)=\frac{m}{n}=\frac{9}{36}=\frac{1}{4}$$

【例 1.7】 有 10 件产品,其中恰好有 7 件正品,3 件次品。试求下列事件的概率:

(1)从这些产品中任取 4 件,其中恰好有 2 件次品;

(2)无放回地连续取 4 件,其中恰好有 2 件次品;

(3)有放回地连续取 4 件,其中恰好有 2 件次品。

解

(1) 设 A =“任取 4 件,其中恰好有 2 件次品”。由于样本空间的基本事件总数 $n_A=C_{10}^4$,事件 A 所包含的基本事件数 $m_A=C_3^2C_7^2$,所以

$$P(A)=\frac{m_A}{n_A}=0.3$$

(2)设 B =“无放回地连续取 4 件,其中恰好有 2 件次品”。由于要考虑顺序,故样本空间的基本事件总数 $n_B=P_{10}^4$,事件 B 所包含的基本事件数 $m_B=C_4^2P_3^2P_7^2$,所以

$$P(B)=\frac{m_B}{n_B}=0.3$$

(3)设 C =“有放回地连续取 4 件,其中恰好有 2 件次品”。由于样本空间的基本事件总数 $n_C=10^4$,事件 C 所包含的基本事件数 $m_C=C_4^2 3^2 7^2$,所以

$$P(C)=\frac{m_C}{n_C}=0.2646$$

【例 1.8】 有 10 件产品,其中 7 件正品,3 件次品,现在从中任取 4 件。试求下列事件的概率:

(1)其中没有次品;

(2)其中只有 1 件次品;

(3)其中最多有 1 件次品;

(4)其中至少有 1 件次品。

解 设 A_i =“任取 4 件,其中恰好有 i 件次品”$(i=0,1,2,3)$;B =“任取 4 件,其中最多有 1 件次品”;C =“任取 4 件,其中至少有 1 件次品”。

(1) $\because n=C_{10}^4=210, m_{A_0}=C_7^4=35$ $\qquad \therefore P(A_0)=\frac{m_{A_0}}{n}=\frac{35}{210}=\frac{1}{6}$

(2) $\because n=C_{10}^4=210, m_{A_1}=C_3^1C_7^3=105$ $\qquad \therefore P(A_1)=\frac{m_{A_1}}{n}=\frac{105}{210}=\frac{1}{2}$

(3) $\because B=A_0+A_1$,且 A_0 与 A_1 互斥

$\therefore$ 由加法定理中公式(1.1.4)可得

$$P(B)=P(A_0)+P(A_1)=\frac{1}{6}+\frac{1}{2}=\frac{2}{3}$$

(4) $\because C=\overline{A}_0$

$\therefore$ 由对立事件计算公式(1.1.6)可得

$$P(C)=P(\overline{A}_0)=1-P(A_0)=\frac{5}{6}$$

1.2 概率的运算法则

1.2.1 条件概率与事件的独立性

1. 条件概率

在实际中,常常会遇到这样的问题:在事件 B 已经发生的条件下,另一个事件 A 发生的概率怎样计算呢? 它与不考虑事件 B 是否发生的情况下事件 A 发生的概率是否相同呢? 它们之间又有什么关系呢? 研究下面的例子。

【例 1.9】 在 100 个零件中,如果长度合格的有 95 个,直径合格的有 94 个,长度与直径都合格的有 92 个,现在从这 100 个零件中任意抽取一个。求:

(1)抽到的该零件长度合格的概率;

(2)抽到的该零件直径合格的概率;

(3)抽到的该零件长度及直径都合格的概率;

(4)抽到的该零件,已知其直径合格,现在长度也合格的概率。

解 设 A ="抽到的该零件长度合格",B ="抽到的该零件直径合格"。

(1)由于 100 个零件中长度合格的有 95 个,按古典概型计算,得

$$P(A)=\frac{C_{95}^1}{C_{100}^1}=\frac{95}{100}$$

(2)由于 100 个零件中直径合格的有 95 个,按古典概型计算,得

$$P(B)=\frac{C_{94}^1}{C_{100}^1}=\frac{94}{100}$$

(3)事件 AB ="抽到的零件长度及直径都合格"。由于 100 个零件中只有 92 件既是长度合格又是直径合格的,按古典概型计算,得

$$P(AB)=\frac{C_{92}^1}{C_{100}^1}=\frac{92}{100}$$

(4)用 $P(A|B)$ 表示事件"抽到的该零件,已知其直径合格,现在长度也合格"的概率。

由于直径合格的零件有 94 个,直径及长度都合格的零件有 92 个,按古典概型计算,得

$$P(A|B)=\frac{C_{92}^1}{C_{94}^1}=\frac{92}{94}$$

比较例 1.9 中(1)和(4),可以看出,求事件 A 发生的概率时,在某一事件 B 没有确定发生之前与事件 B 确定发生之后,其概率是不一定相同的。即事件 A 发生的概率受到事件 B 是否发生的影响。这一类概率称为 A 对 B 的条件概率,记为 $P(A|B)$。同时还可以看出