高等院校“十一五”规划教材

丛书总主编　李海峰　霍振宏

概率论与数理统计

主　编　王宜静　崔宏宇

副主编　戴晓明　陈宝凤　苏婷　刘肖云

中国水利水电出版社

www.waterpub.com.cn

高等院校“十一五”规划教材

概率论与数理统计

主编　王宜静　崔宏宇

副主编　戴晓明　陈宝凤　苏　婷　刘肖云

中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书内容包括随机事件与随机事件的概率、离散型随机变量、连续型随机变量、极限定理、统计学基本概念、点估计与区间估计、假设检验、方差分析、回归分析等。书中简单介绍概率统计的产生和发展，以及在概率统计方面做出重要贡献的著名数学家，并融入数学历史、数学文化教育。

本书强调概率论与数理统计的应用性，力求结合实际的同时又兼顾趣味性，并在设定的数学程度内，力求做到论述严谨。书中精选百余道习题，并在书末附有提示与解答。

本书可作为高等学校非数学系的概率统计课程的教材，也可以作为具有相当数学准备（初等微积分及少量矩阵知识）的读者自修之用。

图书在版编目（CIP）数据

概率论与数理统计 / 王宜静，崔宏宇主编. -- 北京：中国水利水电出版社，2010.4
高等院校“十一五”规划教材
ISBN 978-7-5084-7369-7

Ⅰ. ①概… Ⅱ. ①王… ②崔… 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073221号

策划编辑：杨谷/向辉　责任编辑：宋俊娥　加工编辑：胡海家　封面设计：李佳

书　　名	高等院校“十一五”规划教材 **概率论与数理统计**
作　　者	主　编　王宜静　崔宏宇 副主编　戴晓明　陈宝凤　苏　婷　刘肖云
出版发行	中国水利水电出版社 （北京市海淀区玉渊潭南路1号D座　100038） 网址：www.waterpub.com.cn E-mail：mchannel@263.net（万水） sales@waterpub.com.cn 电话：（010）68367658（营销中心）、82562819（万水）
经　　售	全国各地新华书店和相关出版物销售网点
排　　版	北京万水电子信息有限公司
印　　刷	北京蓝空印刷厂
规　　格	170mm×227mm　16开本　18.25印张　372千字
版　　次	2010年4月第1版　2010年4月第1次印刷
印　　数	0001—3000册
定　　价	32.00元

前　　言

《概率论与数理统计》是我国高校的绝大部分工科、理科专业及管理类专业所开设的一门重要的基础课程。这不仅是因为它们在各个领域中的广泛应用，而且还在于它对人才素质的全面培养是不可或缺的。例如，进入21世纪之后，人们可以通过各种渠道得到越来越多的统计信息，它们传递着自然科学、工程技术、农业生产、经济、医学、金融等各个领域的发展趋势，没有基本的统计知识，就不可能很好地把握这些统计信息的特性，并善加运用。

概率论与数理统计的主要内容包括概率论和数理统计两大部分。概率论是数理统计的理论基础，数理统计是概率论的应用。本书是以这两大部分为主体编写的。

概率论部分的内容侧重以下三个方面：

（1）从集合入手引入随机试验、样本空间、随机事件等概念及其运算和关系。

（2）以古典概型为核心引出条件概率、全概率公式和贝叶斯公式；讨论条件概率、乘法定理之间的关系。

（3）从函数概念入手引入随机变量的概念；较完整地阐述离散型随机变量、连续型随机变量的概念、分布、数字特征等。

数理统计部分的内容侧重如下四个方面：

（1）直观地介绍总体、样本的概念；引出几个重要的统计量，较详细地介绍统计量的分布（抽样分布），为统计推断作基础。

（2）完整地阐述参数估计的基本思想、基本理论与基本方法。

（3）完整地阐述假设检验的基本思想、基本理论与基本方法。

（4）简单地介绍方差分析与回归分析的基本思想、基本理论与方法。

在上述内容的阐述中始终围绕“加强基础，强调应用”八个字，着重培养学生分析问题与解决问题的能力，熟练运用基本概型进行计算的能力，适当训练逻辑思维与推理能力。本教材的每章内容以应用例题引出，以例题选讲结束，符合学生学习知识的心理结构的形成规律，使学生在学习过程中能迅速地构建成自身学习的心理结构。

本书由安阳工学院王宜静、崔宏宇主编，制定编写大纲和方案，把握本书的特点，负责全书的修改和定稿。具体参编的有王宜静（第5章、第8章、概率统计的产生与发展及附录）；崔宏宇（第2章、第4章）；戴晓明（第3章、习题答案）；陈宝凤（第1章、第9章）；苏婷（第7章）；刘肖云（第6章）。

参与编写本教材的老师都具有多年的教学经验，在本书中融入了他们在教学

中发现的诸多问题与注意事项。为了让学生更好地了解概率论与数理统计这门课，在教材中加入概率论与数理统计的产生与发展等诸多内容。

在本教材的编写过程中，得到各级领导的大力支持，在此表示衷心感谢。由于时间仓促及作者水平所限，错漏之处在所难免，望各位读者不吝指正。

编　者

2009 年 5 月

目　　录

概率论与数理统计的产生与发展

一、概率论的产生与发展

概率，又称几率、或然率，指一种不确定的情况出现可能性的大小，例如，投掷一个硬币，“出现国徽”（国徽一面朝上）是一个不确定的情况．因为投掷前，无法确定所指情况（“出现国徽”）发生与否，若硬币是均匀的且投掷有充分的高度，则两面的出现机会均等，也就是说“出现国徽”的概率是 1/2；同样，投掷一个均匀的骰子，“出现 4 点”的概率是 1/6，除了这些以及类似的简单情况外，概率的计算并不容易，往往需要一些理论上的假定，在现实生活中往往用经验方法确定概率，例如，某地区有 N 个人，查得其中患某种疾病者有 M 个人，则称该地区的人患该种疾病的概率为 M/N，这事实上是使用统计方法对发病概率的一个估计．

概率的概念起源于中世纪欧洲流行的用骰子赌博，这一点不难理解，某种情况出现可能性的大小要能够体察并引起研究的兴趣，必须满足两个条件：一是该情况可以在多次重复中被观察其发生与否（在多次重复下出现较频繁的情况有更大的概率），二是该情况发生与否与当事人的利益有关或为其兴趣关注之所在，用骰子赌博满足这些条件．

当时有一个“分赌本问题”曾引起热烈的讨论，并经历了长达一百多年才得到正确的解决．在这过程中孕育了概率论中一些重要的基本概念，举该问题的一个简单情况：甲、乙二人赌博，各出赌注 30 元，共 60 元，每局甲、乙胜的机会均等，都是 1/2．约定谁先胜满 3 局则谁赢得全部赌注 60 元，现已赌完 3 局，甲 2 胜 1 负，而因故中断赌情，问这 60 元赌注该如何分给 2 人才算公平，初看觉得应按 2:1 分配，即甲得 40 元，乙得 20 元，还有人提出了一些另外的解法，结果都不正确，正确的分法应考虑到如在这基础上继续赌下去，甲、乙最终获胜的机会如何，至多再赌 2 局即可分出胜负，这 2 局有 4 种可能结果：甲甲、甲乙、乙甲、乙乙．前 3 种情况都是甲最后取胜，只有最后 1 种情况才是乙取胜，二者之比为 3:1，故赌注的公平分配应按 3:1 的比例，即甲得 45 元，乙得 15 元．

当时的一些学者，如惠更斯、巴斯噶、费尔马等人，对这类赌情问题进行了许多研究，有的出版了著作，如惠更斯的一本著作曾长期在欧洲作为概率论的教科书，这些研究使原始的概率和有关概念得到发展和深化．不过，在这个概率论的草创阶段，最重要的里程碑还是伯努利的著作《推测术》．在他死后的 1713 年才发表，这部著作除了总结前人关于赌情的概率问题的成果并有所提高外，还有一个极重要的内容，即如今以他的名字命名的“大数律”，大数律是关于（算术）

平均值的定理，算术平均值，即若干个数 x_1、x_2、…、x_n 之和除以 n，是最常用的一种统计方法，人们经常使用并深信不疑．但其理论根据何在并不易讲清楚，就是伯努利的大数律要回答的问题，在某种程度上可以说，这个大数律是整个概率论最基本的规律之一，也是数理统计学的理论基石．

概率论虽发端于赌博，但很快在现实生活中找到多方面的应用，首先是在人口、保险精算等方面，在其发展过程中出现了若干里程碑的著作，如《机遇的原理》，其第三版发表于 1756 年，法国大数学家拉普拉斯的《分析概率论》，发表于 1812 年，1933 年苏联教学家柯尔莫哥洛夫完成了概率论的公理体系，在几条简洁的公理之下，又发展出概率论整座的宏伟建筑，又如在欧几里得公理体系之下发展出整部几何．自那以来，概率论成长为现代数学的一个重要分支，使用了许多深刻和抽象的数学理论，在其影响下，数理统计的理论也日益向深化的方向发展．

二、数理统计的产生与发展

数理统计学是研究收集数据、分析数据并据以对所研究的问题作出一定结论的科学和艺术．数理统计学考察的数据都带有随机性（偶然性）的误差．这给根据这种数据所作出的结论带来了一种不确定性，其量化要借助于概率论的概念和方法．数理统计学与概率论这两个学科的密切联系，正是基于这一点．

数理统计学重要源头来自天文和测地学中的误差分析问题．早期，测量工具的精度不高，人们希望通过多次测量获取更多的数据，以便得到对测量对象更高精确度的估计值．测量误差有随机性，适合于用概率论即统计的方法处理，伽利略就做过这方面的工作，他对测量误差的性态作了一般性的描述，法国大数学家拉普拉斯也曾对这个问题进行了长时间的研究，现今概率论中著名的“拉普拉斯分布”，即是他在此研究中的一个产物．在此方面最著名且影响深远的研究成果有两个：一个成果是法国数学家兼天文家勒让德 19 世纪初（1805 年）在研究慧星轨道计算时发明的“最小二乘法”，他在估计过巴黎的子午线长这一工作中，曾使用这个方法．现今著作中把这一方法的发明归功于高斯，但高斯使用这一方法最早见诸文字是 1809 年，比勒让德晚．有一种说法现在逐步取得公认——这项发明是由二人独立作出，看来是比较妥当的．另外一个重要成果是德国大学者高斯 1809 年在研究行星绕日运动时提出用正态分布刻画测量误差的分布．正态分布也常称为高斯分布，其曲线是钟形，极像颐和园中玉带桥那样的形状，故有时又称为“钟形曲线”，它反映了这样一种极普通的情况：天下形形色色的事物中，“两头小，中间大”的居多，如人的身高，太高太矮的都不多，而居于中间者占多数——当然，这只是一个极粗略的描述，要作出准确的描述，须动用高等数学的知识；正是其数学上的特性成为其广泛应用的根据．

正态分布在数理统计学中占有极重要的地位，现今常用的许多统计方法，就是建立在“所研究的量具有或近似地具有正态分布”这个假定的基础上，而经验和理论（概率论中所谓“中心极限定理”）都表明这个假定的现实性，现实世界许

多现象看来是杂乱无章的，如不同的人有不同的身高、体重，大批生产的产品，其质量指标各有差异．看来毫无规则，但它们在总体上服从正态分布．这一点，显然在纷乱中有一种秩序存在，提出正态分布的高斯，一生在多个领域里面有不少重大的贡献，而且在德国10马克印有高斯图像的钞票上，还画有正态曲线，以此可以看出人们对他这一贡献评价之高．

20世纪以前数理统计学发展的一个重要成果，是由19世纪后期的英国遗传学家兼统计学家高尔顿发起的．而后经现代统计学的奠基人之一的K·皮尔逊和其他一些英国学者逐步发展形成了统计相关与回归理论．所谓统计相关，是指一种非决定性的关系如人的身高X与体重Y，存在一种大致的关系，表现在X大（小）时，Y也倾向于大（小），但并非决定性的，因为X并不能决定Y．现实生活和各种科技领域中，这种例子很多，如受教育年限与收入的关系，经济发展水平与人口增长速度的关系等，都是属于这种性质，统计相关的理论是把这种关系的程度加以量化，而统计回归则是把有统计相关的变量，如上文的身高X和体重Y的关系的形式，作近似的估计，称为回归方程，现实世界中的现象往往涉及众多变量，它们之间有错综复杂的关系，且许多属于非决定性质．相关回归理论的发明，提供了一种通过实际观察去对这种关系进行定量研究的工具，对认识现实世界有着重大的实用意义．

三、数理统计的现实意义及作用

数理统计学的理论和方法，与人类活动的各个领域在不同程度上都有关联．因为各个领域内的活动，都得在不同的程度上与数据打交道，都有如何收集和分析数据的问题，因此也就有数理统计学用武之地．这里举几个例子来说明这一点，如在工业中生产一种产品，首先，有设计的问题，包括配方和工艺条件的选定，这要通过从大量可能的条件组合中分析试验结果来选定，可能的条件组合很多，选择哪一部分去做试验是一个很有讲究的问题，在数理统计学中有一个专门分支叫“试验设计”，就是研究怎样在尽可能少的试验次数之下，达到尽可能高效率的分析结果；其次，在生产过程中，存生由于原材料、设备调整及工艺参数等条件可能的变化，而造成生产条件不正常并导致出现废品，在统计学中有一门“工序控制”的学问，通过在生产过程中随时收集数据并用统计方法进行处理，可以监测出不正常情况的出现以便随时加以纠正，避免出大的问题；再次，大批量的产品生产出来后，还有一个通过抽样检验以检验其质量是否达到要求，是否可以出厂或为买方所接受的问题，处理这个问题也要使用数理统计方法，在我国现行的国家标准中有一些就与这个问题有关．

在农业上，有关选种、耕作条件、肥料选择等一系列的问题的解决，都与统计方法的应用有关，在历史上，现行的一些重要的统计设计与分析方法，就是近代最伟大的数理统计学家费歇尔于20世纪20年代在英国一个农业试验站工作时，因研究田间试验的问题而发明的．

医学和生物学是统计方法应用最多的领域之一，统计学是在有变异的数据中研究和发现统计规律的科学，就医学而言，人体变异是一个重要的因素，不同的人的情况千差万别，其对一种药物和治疗方法的反应也各不相同，因此，对一种药物和治疗方法的评价，是一种统计性规律的问题，不少国家对一种新药的上市和一种治疗方法的批准，都设定了很严格的试验和统计检验的要求，又如，许多生活习惯（如吸烟、饮酒、高盐饮食之类）对健康的影响、环境污染对健康的影响，都要通过收集大量数据进行统计分析来研究.

对社会现象的研究大量地使用了统计方法，因为组成社会的单元（如人、家庭、单位、地区等）都有很大的变异性. 如果说在自然现象中还不乏一些（在误差可以允许的限度内）严格的、确定性的规律，在社会现象中这种规律则绝少，因此只能从统计的角度去考察，我们常说，某某措施，某某政策，对大多数人是有利的，这就是一种统计性规律，因为这种“有利”是指对大多数，而非一切人. 在20 世纪初，就有统计学家在英国研究过几种救助贫困方式的效果评估，这都是借助抽样调查并通过复杂的统计分析得出的结果，如今抽样调查已经成为研究社会现象的一种最有力的工具，因为全面调查往往不可行，而抽样调查从其方案的制定到数据的分析，都是以数理统计学的理论和方法为基础.

四、概率统计的发展前景

对于未来统计学将向哪个方向发展或应当向哪个方向发展的问题，有一种得到比较广泛认同的观点是认同统计学研究应努力与其他实用学科结合而形成交叉或边缘学科，这一点目前已有一定的表现，如生物统计、医药统计、工业统计、金融统计等，都是当前发展很快的热点，有的学者认为研究数理统计学必须与另一门专门学问结合，才有可能作出有重要意义的成果. 这一点已在若干成功的学者身上得到印证，有个别走得更远的学者认为，统一的统计学将会因为与其他学科结合发展而分裂成许多并行的学科，好比一个大国分裂成一些小国，并把这称为统计学的巴尔干化——与昔日巴尔干半岛上统一的南斯拉夫如今分裂为一些小国相比. 但是，数理统计学与其他学科结合形成交叉学科这个引人注目的发展，是否将导致“统一的”或“一般的”统计学的消亡或衰落，这一点现在看来并不确定，至少现在多数学者还不这么认为.

图基（美国老一辈著名统计学家）在 1962 年在一篇长文中提出“数据分析”的思想，几十年来得到国际上一些有影响的学者的支持，要全面讲清楚这种观点需要较多的篇幅，这里只就其一个核心的观点来讨论一下，这涉及到对现行的数理统计规范的地位问题，前面曾谈到，由于统计学处理的是带随机误差的数据，由分析这种数据得出的结论就有可能出错或不准确，出错可能性的大小和不准确的程度如何，需要用概率论的概念和方法作定量的刻画，在研究统计问题时，必须把这作为一个目标，朝这个方向努力，这就是现行数理统计学的规范. 数理统计学之所以能被承认为一门有严格理论基础的学科，是与遵守这一规范联系在一

起的．但是，如果我们真的严格遵守这一规范，则以现在的知识水平而言，许多问题将无法下手．于是，学者们只好转向一些人为的、不太复杂的、用现行数学工具可以处理的模型，这种模型往往有“闭门造车”的缺点而缺乏现实性，图基的“数据分析”思想的一个观点是主张淡化这个规范．

这种说法有一定的事实根据，可以说，在实用统计学的领域中，这个规范并不总是得到严格遵守的，现在有一些统计方法，它用起来有较好的效果，但在理论上并没有搞清楚其错误或偏差的可能性或数量有多大；另外，随着科技的发展，不断提出一些更复杂的模型，以现有的知识水平，不可能对之作出完全符合上述规范的处理，而只能退而求其次寻求一种在实用上可行的解法．当然，应当明确，在研究工作中达不到上述规范，与从根本上取消或淡化这个规范是两回事，一门学科必须有其规范或科学的定位（回答这门学科是什么的问题，判定其成果的可信性与意义等，而这不能用笼统的说法，必须用确切的科学语言）．如果用数据分析取代现行的数理统计学，就有一个为数据分析定位的问题，而这至今还没有一个满意的解决方案，以此之故，虽然数据分析的提法获得不少支持且在实际的统计应用中有所反映（如现在媒体中常提及的“数据挖掘”Data Mining）．数据挖掘并不单纯是一个统计学课题，它至今尚未能动摇现行数理统计学的主流地位．

除了上述几种富于原则性的思想外，也有一部分学者致力于在现行统计学的框架下寻求新的生长点，在这一点也有不少的异议和争论，如关于费歇尔的统计学思想和研究成果的再认识、关于数理统计学中的“频率学派”（有时也称为经典学派）与“贝叶斯学派”之间的争论等，因为涉及较多的数学概念，所以就不细谈了．

第 1 章　事件与概率

1654 年，一个名叫梅累的骑士就“两个赌徒约定赌若干局，且谁先赢 c 局便算赢家，若在一赌徒胜 a 局（$a<c$），另一赌徒胜 b 局（$b<c$）时便终止赌博，问应如何分赌本”为题求教于帕斯卡，帕斯卡与费马通信讨论这一问题，于 1654 年共同建立了概率论的第一个基本概念——数学期望.

概率论是数学的一个分支，它研究随机现象的数量规律. 一方面，它有自己独特的概念和方法，另一方面，它与其他数学分支又有紧密的联系，它是现代数学的重要组成部分. 概率论的广泛应用几乎遍及所有的科学技术领域，如天气预报、地震预报、产品的抽样调查；工农业生产和国民经济的各个部门，在通信工程中可用以提高信号的抗干扰性、分辨率等.

1.1　样本空间和随机事件

一、样本空间

在自然界与人的实践活动中经常遇到各种各样的现象，这些现象大体可分为两类：一类是确定的，如“在一个标准大气压下，纯水加热到 100 摄氏度时必然沸腾.”、“同性电荷相斥，异性电荷相吸.”等，这种在一定条件下有确定结果的现象称为**必然现象**；另一类现象是随机的，例如，在相同的条件下，向上抛一枚质地均匀的硬币，其结果可能是正面朝上，也可能是反面朝上，不论如何控制抛掷条件，在每次抛掷之前无法肯定抛掷的结果是什么，这个试验多于一种可能结果，但是在试验之前不能肯定试验会出现哪一个结果. 同样地同一门大炮（使用同一型号的炮弹）对同一目标进行多次射击，各次弹着点可能不尽相同，并且每次射击之前无法肯定弹着点的确切位置，以上所举的现象都具有随机性，即在一定条件下进行试验或观察会出现不同的结果，而且在每次试验之前都无法预言会出现哪一个结果，这种现象称为**随机现象**.

在科学研究和社会生活中，常常要对上述必然现象和随机现象进行实验或观察，统称**试验**. 一般地，称具有以下两个特点的试验为**随机试验**：①试验的所有可能结果是已知的或是可以确定的；②每次试验究竟将会发生什么结果是事先无法预知的. 为方便起见，随机试验也简称为试验，今后讨论的试验都是指随机试验.

例 1.1.1　判断下列试验是否随机试验.

E_1：一盒中有十个完全相同的白球，搅匀后从中任意摸出一球，观察所摸取

球的颜色.

E_2：一盒中有十个完全相同的球，其中 5 个白球，5 个黑球，搅匀后从中任意摸取一球，观察所摸取球的颜色.

E_3：连续投掷两次骰子，观察所掷的点数.

E_4：观察一个飞镖运动员扎中一次红心所需要的投掷次数.

E_5：在一批灯泡中任意抽取一只，测试它的寿命.

E_6：观察北京市建国 100 周年时的国庆节那天是否晴天.

解 对于 E_1 而言，在球没有取出之前，就能确定取出的球必是白球，也就是说在试验之前就能判定它只有一个确定的结果，这种现象就是必然现象. 该实验不符合随机试验定义中的条件②，因此不是随机试验.

对于 E_2、E_3、E_4、E_5、E_6 来说，他们的所有可能结果显然是已知的，而每次试验究竟会发生什么结果则无法事先确定. 如 E_2 的所有可能结果只有黑球或白球，但到底在一次摸球过程中会获得哪一种结果则无法确定. 因此这 5 个试验都是随机试验.

对于随机试验来说，感兴趣的往往是随机试验的所有可能结果. 如掷一枚硬币，关心的是出现正面还是出现反面这两个可能结果. 若观察的是掷两枚硬币的试验，则可能出现的结果有（正、正）、（正、反）、（反、正）、（反、反）四种，如果掷三枚硬币，其结果还要复杂，但还是可以将它们描述出来的，总之为了研究随机试验，必须知道随机试验的所有可能结果.

通常，依据研究目的，将随机试验的每一个可能的结果，称为**基本事件**. 因为随机事件的所有可能结果是明确的，从而所有的基本事件也是明确的，如在抛掷一枚硬币的试验中“出现反面”和“出现正面”是两个基本事件，又如在掷骰子试验中“出现一点”、“出现两点”、“出现三点”、……、“出现六点”这些都是基本事件.

基本事件的全体，称为**样本空间**. 也就是试验所有可能结果的全体是样本空间，样本空间通常用大写的希腊字母 Ω 或大写英语字母 S 表示，Ω 中的点即是基本事件，也称为**样本点**，常用 ω 表示.

在具体问题中，给定试验的样本空间是研究随机现象的第一步.

例 1.1.2 给出上例中随机试验 E_2、E_3、E_4、E_5、E_6 的样本空间.

解 分别设 E_2、E_3、E_4、E_5、E_6 的样本空间为 Ω_2 、Ω_3 、Ω_4 、Ω_5 、Ω_6 ，则

$\Omega_2=\{\omega_1,\omega_2\}$ ，其中 $\omega_1=$“黑色”，$\omega_2=$“白色”；

$\Omega_3=\{(i,j)\mid i=1,2,3,4,5,6;j=1,2,3,4,5,6\}$ ，其中 $i=$“第一次掷出 i 点”，$j=$“第二次掷出 j 点；

$\Omega_4=\{1,2,3\cdots\}$ ，这里飞镖运动员要扎中一次红心所需投掷次数一定是正整数而且很难指定一个数为它的上界，所以把样本空间取为 $\{1,2,3\cdots\}$ ；

$\Omega_5=[0,+\infty)$ ，这里灯泡的寿命很难限定一个上界，所以用取其样本空间为 $[0,+\infty)$

$\Omega_6=\{0,1\}$，其中 0=“不是晴天”，1=“是晴天”.

仔细观察可以发现，样本空间Ω_2、Ω_3和Ω_6只有有限个样本点，是比较简单的样本空间；样本空间Ω_4含有无穷个样本点，但这些样本点可以依照某种顺序排列起来，称它为可列样本空间；样本空间Ω_5含有无穷个样本点，且充满一个区间，称它为无穷样本空间.

从这些例子可以看出，随着问题的不同，样本空间可以相当简单，也可以相当复杂．并且对于一个实际问题或一个随机现象来说，考虑问题的角度不同，其样本空间也可能选择的不同．如，掷骰子这个随机试验，若考虑出现的点数，则样本空间$\Omega=\{1,2,3,4,5,6\}$；若考虑的是出现奇数点还是出现偶数点，则样本空间$\Omega=\{奇数,偶数\}$.

二、随机事件

在E_3中，考虑下面几个事件：

$A=$“第一次掷出 1，第二次掷出 2”；$B=$“第一次掷出 1，第二次掷出偶数”；$C=$“两次掷出的点数和超过 10”

这里，A为一个基本事件，而B与C则是由基本事件所组成的**复杂事件**．如，B发生（出现）必须而且只须下列样本点之一发生：(1,2)、(1,4)、(1,6)，它由三个基本事件组成；同样地，C发生必须而且只须下列三个样本点之一发生：(5,6)、(6,5)、(6,6).

无论基本事件还是复杂事件，它们在试验中发生与否都带有随机性，所以叫做**随机事件**或简称为**事件**，习惯上用大写英文字母A、B、C等表示，在试验中如果发生了A中所包含的某一个基本事件ω，则称作A发生，并记作$\omega\in A$．此时也称ω为事件A的有利事件.

我们知道，样本空间Ω包含全体基本事件，而随机事件不过是由具备某些特征的基本事件组成的．从集合论的角度来看，一个随机事件不过是样本空间Ω的一个子集而已，所以对随机事件的学习和认识可以借助于集合的相关知识.

因为Ω是所有基本事件组成的，因而在一次试验中，必然要出现Ω中的某一基本事件，也就是在试验中Ω必然要发生，称之为**必然事件**．而不含任何样本点的空集Φ，也是样本空间的子集，它在任何一次试验中都不会发生，称之为**不可能事件**，如E_4中，$\{0\}$是一个不可能事件.

实质上必然事件就是在每次试验中都发生的事件，不可能事件就是在每次试验中都不发生的事件，必然事件与不可能事件的发生与否，已经失去了不确定性，即随机性，因而本质上不是随机事件，但为了讨论问题的方便，还是将它们看作随机事件．如，一批产品共 10 件，其中 2 件次品，其余为正品，从中任取 3 件，则：$A=\{恰有一件正品\}$，$B=\{恰有两件正品\}$，$C=\{至少有两件正品\}$都是随机事件；而$\Omega=\{三件中至少有一件正品\}$为必然事件；$\Phi=\{三件中有三件次品\}$为不可能事件．对于随机试验而言，它的样本空间Ω可以包含很多随机事件，概率论的

任务之一就是研究随机事件的规律，通过对较简单事件规律的研究从而掌握更复杂事件的规律，为此需要研究事件之间的关系与运算.

三、事件的关系及运算

随机事件是样本空间的子集，因此事件的关系与运算同集合的关系与运算完全相类似. 以下若无特殊说明，则认为样本空间Ω是给定的，并且还定义了Ω中的一些事件A、B、$A_i(i=1,2,\cdots)$等.

定义 1.1.1 若事件A发生必然导致事件B发生，则称事件B**包含**A或称A是B的**特款**，记作$A\subset B$或$B\supset A$.

比如前面 E3 中，设$A=$"第一次掷出 1，第二次掷出 2"；$B=$"第一次掷出 1，第二次掷出偶数"，则A的发生显然会导致B的发生，所以有$A\subset B$.

可以给$A\subset B$一个几何解释，设样本空间是一个全集（论域）Ω，A、B是两个事件，则A、B是全集Ω的子集，如图 1-1 所示. "A发生必然导致B发生"意味着属于A的样本点在B中. 由此可见，事件$A\subset B$的含义与集合论是一致的.

特别地，对任何事件A，有$A\subset\Omega$，$\Phi\subset A$.

定义 1.1.2 设$A,B\subset\Omega$，若$A\subset B$，同时有$B\subset A$，称A与B**相等**，记为$A=B$，易知相等的两个事件A、B总是同时发生或同时不发生，在同一样本空间中两个事件相等意味着它们含有相同的样本点.

定义 1.1.3 设$A,B\subset\Omega$，称事件"A与B中至少有一个发生"为A与B的**和事件**或**并事件**，记作$A\cup B$.

如，检测某种型号的油桶，如果底面直径和高都合格，那么该油桶合格. 令$A=\{$直径不合格$\}$，$B=\{$高度不合格$\}$，则$A\cup B=\{$产品不合格$\}$.

称$\bigcup\limits_{k=1}^{n}A_k$为$n$个事件$A_1,A_2,\cdots A_n$的和事件；称$\bigcup\limits_{k=1}^{\infty}A_k$为可列个事件$A_1,A_2,\cdots$的和事件.

显然有

（1）$A\cup\Phi=A$，$A\cup\Omega=\Omega$，$A\cup A=A$；

（2）$A\subset A\cup B$，$B\subset A\cup B$；

（3）若$A\subset B$，则$A\cup B=B$.

定义 1.1.4 设$A,B\subset\Omega$，称"A与B同时发生"这一事件为A和B的**积事件**或**交事件**，记作AB或$A\cap B$.

如前面所举检测油桶质量一例中，令$C=\{$直径合格$\}$，$D=\{$高度合格$\}$，则$C\cap D=\{$产品合格$\}$.

称$\bigcap\limits_{k=1}^{n}A_k$为$n$个事件$A_1,A_2,\cdots,A_n$的积事件；称$\bigcap\limits_{k=1}^{\infty}A_k$为可列个事件$A_1,A_2,\cdots$的积事件.