浙江省级重点学科应用数学教学改革与科学研究丛书

信息论与密码学

邸继征　编著

科学出版社

浙江省级重点学科应用数学教学改革与科学研究丛书

信息论与密码学

邸继征　编著

浙江省级重点学科应用数学建设基金资助
浙江工业大学重点教材建设项目基金资助

科学出版社

北京

内 容 简 介

本书概念清晰，推理严密，论证细致，对每部分内容，都展示是什么、为什么和怎么做的全过程，并将基础和应用并重的教育理念融入其中. 全书分 6 章，介绍信息论和密码学的基础知识. 在信息论方面，引入给出信源和信道概念的联合概率空间，并由此给出离散信源的数学模型，介绍信息量、熵和信源编码；给出离散信道的数学模型，介绍互信息、信道容量和信道编码. 在密码学方面，讲述密码学的基础理论，介绍以 DES 系统为代表的分组密码和以 RSA 系统为代表的公钥密码.

本书可作为高等院校数学和应用数学、信息与计算科学专业和信息类、软件类本科生和研究生的信息论与密码学教材和参考书.

图书在版编目(CIP)数据

信息论与密码学/邸继征编著. —北京：科学出版社，2013
（浙江省级重点学科应用数学教学改革与科学研究丛书）
ISBN 978-7-03-037808-8

Ⅰ. ①信… Ⅱ. ①邸… Ⅲ. ①信息论-高等学校-教材 ②密码-理论-高等学校-教材 Ⅳ. ①TN911.2 ②TN918.1

中国版本图书馆 CIP 数据核字(2013)第 125988 号

责任编辑：石 悦 李梦华/责任校对：宣 慧
责任印制：阎 磊/封面设计：华路天然设计工作室

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

北京市文林印务有限公司 印刷

科学出版社发行 各地新华书店经销
*
2013 年 8 月第 一 版 开本：720×1000 B5
2013 年 8 月第一次印刷 印张：12
字数：242 000

定价：27.00 元
（如有印装质量问题，我社负责调换）

“浙江省级重点学科应用数学教学改革与科学研究丛书”
编　委　会

总　　序

近年来，关于数学的各种新观点不断出现.

有一种观点认为，随着数学的发展，数学已经从自然科学中分离出来，成为独立的科学门类——数学科学.

持这种观点的学者的依据是：①从现代数学的发展情况可以看出，数学的许多内容和方法的产生，不再是基于研究自然界中存在的物质运动规律的需要，而是基于数学自身的需要. 例如，6＝3＋3，8＝3＋5，10＝5＋5＝3＋7，等等，即每一个大于等于6的偶数都可以表示为两个奇素数的和，这就是哥德巴赫猜想，至今没有证明. 但是，这样一个在数学中显得十分重要的著名的猜想，其结果的对与否，不会对数学之外的任何学科产生影响，证明它不是自然科学的需要，而仅仅是数学科学的需要. ②数学不仅具有应用功能，而且具有其他学科不能比拟的教育功能. 数学的应用功能表现在：没有数学，现代科技无从谈起；任何一种学科，只有应用了数学，才能成为科学学科. 数学的教育功能表现在：在中国，语文、数学、英语被认为是初等教育中最重要的三门课程；在世界范围内，有不学中文的学生，有不学英语的学生，但没有不学数学的学生.

我们同意这种观点，希望在数学教学改革和科学研究中体现这种观点.

数学教学改革，首先需要的是教材的改革，而教材的改革，涉及的只有两个方面：一是内容，二是方法.

如何在一本数学教材中以数学科学的观点选取内容、介绍方法？

我们的认识是：无论是选取内容方面，还是介绍方法方面，都要关注数学的应用功能和教育功能的展现.

在内容的选取方面，既不是不管数学的教育功能，狭隘地全部以目前生产生活的实际应用为目的，打乱系统，什么“有用”就选什么，什么“没用”就跳过什么；也不是完全从数学的需要出发，一点也不考虑所选取的内容和实际应用的联系. 本套丛书采取有实际应用背景的内容优先选取的原则. 我们的考虑是：没有迹象表明，没有实际应用背景的内容在体现数学的教育功能时强于有实际应用背景的内容，既然如此，后者更有利于同时展现数学的应用功能和教育功能.

在方法的介绍方面，既不完全采用公理化体系的做法，让读者在接受严格数学训练的基础上自然地受到数学科学的熏陶，也不完全摒弃数学特有的推理过程，以急功近利的方式只讲结果，只讲计算公式. 我们知道，公理化体系的做法是将数学的训练目的不直接说出来，而是藏起来，藏在严密的过程背后，让学生不知不觉得到严格的数学训练. 这种体系在介绍内容时，不交代前因后果，一上来就是莫名其妙的定义、公理，然后一步步以极其严密的方式展开讨论. 这种做法在知识门类相对少的过去是有

效的，但在知识爆炸、课程门类不断增加、学生同时要有做学问和实际应用两手准备的现在，没有时间这样做. 训练要有，但训练目的不是藏起来，而是尽可能直接讲出来. 例如，数学书籍中一定会用到归纳法、演绎法、反证法，这些方法不是数学特有的，但可以被数学最为有效地传授给学生，这一事实恰好可以说明数学的教育功能的强大. 但是，如果我们去问一下数学系的毕业生什么是演绎法，恐怕很少有人能说周全，究其原因，是我们的教材没有明确地告诉学生演绎法的基本内容和过程. 本套丛书将致力于改变这种状况.

本套丛书注意到：根据课程和授课对象的不同，数学的应用功能和教育功能的展现需分层次，两种功能的展现要有机配合. 例如，有的数学分支本来就属于应用数学，对这样的课程，在选取内容和介绍方法时必须首先保证应用方面的需要，其次才考虑教育功能的融入；有的授课对象是文科学生，对这些学生，在编写教材时就要充分注意他们的基础、兴趣、思维方式和希望通过数学的学习要达到的目的，因此要首先考虑数学的教育功能，其次才考虑应用功能的融入.

现代化的标志是数字化，也就是要在所有的领域尽最大可能地使用计算机技术，因此，在数学教学中，对数字化的配合和适应是必需的. 为了展现数学的应用功能，在数学教学的每个环节，都应该关注计算机技术，包括有意考虑内容的计算机实现，如算法问题，内容与几个成功的数学软件的结合问题. 我们知道，介绍如何应用数学软件的最好环境，当为相应的数学课程. 因此，本套丛书中的教材，特别注意介绍与主要内容配套的软件的应用，例如，介绍相应的 MATLAB 软件包的使用.

科学研究成果整理成学术著作，可以总结和条理化研究问题，这对传播研究成果、深化研究工作是有利的，这些著作还可以作为研究生教材使用.

本套丛书中学术著作的撰写遵循了如下的原则：

首先，作为介绍学术成果的学术著作要有新内容、新观点，学术系统应是明显的，不是杂乱的、拼凑的，特别是著作中作者的成果应有重要的分量.

其次，本套丛书中的学术著作特别注意内容的系统性、完备性.

再次，也是最重要的，本套丛书中的学术著作，和教材一样注意展现数学的应用功能和教育功能，在必要时，还考虑内容的计算机实现，如算法问题，内容与几个成功的数学软件的结合问题.

最后，在写作细节上，本套丛书要求作者以严格的科学态度对待自己的著作，概念和符号应明确，推导和介绍要细致，避免突然出现翻遍全书都找不到介绍的概念和符号，避免用显然、易知等词语掩盖困难的证明过程.

教学改革涉及的问题很多，有些问题需要一步步解决，有的还需要根据形势的变化调整解决方案. 我们仅做了初步的尝试，加之水平有限，本套丛书中的问题一定很多，迫切希望读者批评指正.

邸继征

2013 年 3 月 8 日

前　言

信息论由美国通信科学家香农(C. E. Shannon)于20世纪40年代创立. 随着计算机技术的发展,信息论的研究内容与计算机的关系越来越密切. 电子计算机的研制,一开始就涉及符号的编码问题,而编码则是信息论的重要研究内容. 现在,计算机已离不开信息论的研究对象,同时,信息论也转为根据计算机特点进行研究.

密码学的历史比信息论长,但随着信息论的创立,密码学的基础理论得到了充实,信息论中的编码理论和技术也为密码学的发展提供了极大的支持. 此外,计算机技术的发展对密码学提出新的研究课题的同时,也为密码编码和解码方法的实现提供了有力的工具. 软件和网络安全问题已经成为现代人必须时刻面对的问题,而这些问题就是密码学问题,密码编码的实现已不可能由人工进行,必须借助计算机. 信息论、计算机和密码学三者,已经到了必须共同发展、互相不能离开的程度.

本书介绍最基本的信息论和密码学内容,为出身数学的读者编写. 因此,时刻注意概念的明确性、推理和论证的严密性. 避免出现无法找到出处的概念和记号,可以用定义形式表达的概念都将其写作定义,不能用定义描述的概念尽量将其符号化. 例如,信息、信道等概念,都用明确的符号加以表示,正如点、集合的概念不能给出具体的定义,但可用 a,E 等符号表示一样. 本书中的有关结论尽量表述为定理,其证明力求严格,证明中避免使用容易引起歧义的描述性语言和"显然"等字眼,不允许用"易知"掩盖实际上困难的证明步骤.

本书的定位是教材,因此,必须在其中融入教育理念. 我们认为,教育应该基础和应用并重,高等院校应该以提高学生的素质和能力为目标,而不是教学生技术和手艺,所以,不能什么"有用"就学什么,什么"没用"就不学什么. 不然,学生学到的知识是零碎的、不成系统的,几年下来,学生会处于被动、应对的状态,丧失联想能力和主动的创造精神. 在这种理念指引下,本书注意解释一些知识的脉络,以开阔学生的眼界,提高学生的联想能力. 例如,在介绍以公理化方法给出的概率空间定义时,用一定篇幅对该方法与其他方面知识的关系做了阐述,以使学生在见到类似结构的问题时,可以借鉴来自不同方面的知识.

虽然书中许多内容来自参考文献,但这些内容都按数学的习惯加以修改、组合,许多定理的证明是新给出的,如定理3.3.1、定理3.4.3、定理3.4.6、定理3.4.7和定理6.2.2的证明. 有些内容是新的,如5.3.3小节.

在内容的选取上,本书不求全,但求精,编写的原则是要么不讲,要么讲透,所有

内容力求展现是什么、为什么、怎么做的全过程，使读者对所学内容学而知理、知而会用，为必要时深入学习信息论和密码学打好基础.

本书的出版得到了浙江工业大学教务处、理学院的大力支持，在此表示感谢.由于作者水平有限，书中难免有不足之处，恳请读者批评指正.

编　者

2013 年 3 月 8 日

目　录

第1章 绪 论

本章有3节，介绍与本书密切相关的部分概念和数学知识. 1.1节介绍与“信息”相关的几个概念、信息论的研究对象和基本内容. 1.2节介绍概率论知识，以公理化方法叙述概率空间与随机变量，讨论事件独立性，建立联合概率空间. 其中，联合概率空间的引入使在同一概率空间上成立的乘法公式、加法公式等以新的角度呈现，为以后建立信源的数学模型做准备. 离散概率空间与信息论的联系最为密切，该节对此加以重点介绍. 随机序列将是离散信源的模型，马尔可夫链是特殊的随机序列，该节中，对本书必要的随机序列和马尔可夫链知识作介绍. 最后介绍的伯努利试验与伯努利大数定律，在以后章节中证明一些结果时用到. 1.3节介绍的詹森(Jensen)不等式，是以后证明一些不等式的重要工具.

1.1 几个概念和信息论的研究内容

信息二字是近30年来世界上出现频率最高的科技词汇之一. 学习信息论，首先要了解什么是信息，要知道不同环境下使用的信息二字的涵义.

1. 信息

正如数学中点、集合等不能加以定义一样，信息也是不能给予明确定义的概念. 有人将信息称为“有用的消息”，但消息是什么，有用又是什么意思？这都无法给出精确的描述.

其实世界上许多事物都是难以给出定义的，这是因为这些事物内涵广泛、性质复杂，无法用简洁的语言加以界定、描述. 例如，“白菜的味道”是一个事物，但人们却无法将这个事物用语言、文字描述清楚.

不能给出定义的事物往往还是事件的主角，为了让别人明白此主角是此事物，不是彼事物，人们只好想各种办法. 一种办法是把事物包含的内容加以列举，让被介绍者自己去领会事物的本质. 例如，无法给出数学的定义，就告诉别人几何学是数学，微积分学是数学，概率论是数学，让别人在了解这几门学问的基础上，自己领会什么是数学.

另一种办法是指出事物的几个具有代表性的表现形式，让别人在看到这些表现形式的基础上，得出此事物正是此事物的结论.

这两种办法是可以用语言、文字表现的，两种办法还可以结合应用. 这两种办法之外，还有无数种办法介绍事物. 例如，用让人尝的办法介绍白菜的味道，让人看的办法介绍风景，让人听的办法介绍声音. 于是，对于信息，可以将前述两种办法结合应用加以介绍.

信息是什么？电话、电视、雷达、信件等传递出来的声、像、磁、文字等能被感知、可以引起人们兴趣的东西都是信息.“电话、电视、雷达、信件等传递出来的声、像、磁、文字等”，正是在列举信息的内容，而“能被感知、可以引起人们兴趣”，则是在展示信息的表现形式.将信息的表现形式细化，有利于让人抓住信息的本质特征，进而建立数学模型，对其展开科学的研究.

细化了的信息的表现形式有：

(1)发出信息的一定是世界上存在的事物，收到声音信息，一定是有东西在发出声音(信源)；

(2)信息必有一个发出、传递、接收的过程(信源、信道、信宿)；

(3)信息带有不确定性(用信息量表示)，正是这种不确定性引起人们对信息的兴趣.如果一个人在接电话前就知道打电话的是谁、说的是什么事，连每个声音细节都知道，那么这个电话对他来说就没有信息价值了.

2.信息论

信息论自然是研究信息的理论，由美国通信科学家香农(C. E. Shannon)于20世纪40年代创立，最初是研究通信信号传输的理论与实现问题.但随着计算机技术的发展，信息论的研究内容与计算机的关系越来越密切.电子计算机的研制，一开始就涉及符号的编码问题，可以说，没有编码就没有计算机.而编码则是信息论的重要研究内容.现在，计算机的应用中除了研究与人工智能和科学计算有关的问题外，许多都是处理信息的传输、压缩、存储和数据挖掘等问题.计算机已离不开这些信息论的研究对象，同时，信息论也转向专门针对计算机特点进行的理论和应用研究.

3.关于“信息”的理解

大概正是由于计算机与信息的关系密不可分，所以人们将与计算机相关的知识、技术、学科，常冠以“信息”二字.例如，许多大学中以计算机科学为学习和研究主体的学院，都冠名信息学院.以致一些人认为，“信息技术”就是计算机技术的代名词，“信息科学”是研究与“信息技术”或即计算机技术相应的基础理论和应用的科学，“信息科学”和“计算机科学”没有什么区别.目前，为数众多的人一见到“信息”二字，立即想到计算机，而不知道“信息”的含义和来源，引起许多概念上的混乱.

那么，“信息”二字究竟应该如何理解？这个问题的回答存在许多争议.我们认为，除了信息论中的信息具有确定的含义，是真正意义上的信息以外，所有见到的“信息”大都兼有计算机和真正意义上的信息两重意思.例如，“信息技术”与“信息科学”中的“信息”就是如此，通常说的“信息时代”中的“信息”也是如此.人类进入“信息时代”意味着：人类进入了一个需要大量使用计算机，同时需要处理浩瀚如海的信息的时代.仅仅认为“信息时代”就是研究信息论的时代是不对的，仅仅认为“信息时代”就是研究计算机的时代也是不对的.

此外，下一段将会看到，信息论中对信息的研究，有明确的内容与方法，其他地方

对信息(真正意义上的信息)的研究,内容庞杂,方法多样,有一些和信息论相同,多数情况下着重于对数据的处理,与信息论迥异.

4. 信息论的研究内容

由于信息具有未知性、不确定性,因而信息论主要以概率论、统计学和随机过程作为研究工具. 有人认为信息论是统计数学的一个分支,将信息论称为统计信息论.

信息论研究的主要内容有:信息的量化问题(信息量、熵);如何将信息符号转化为能有效地利用计算机及其他设备处理和传输的符号问题(编码);信息传输工具(信道)的量化问题(互信息、信道容量);信息的保密问题(密码学的一方面)等.

因此,信息论有非常明确具体的研究内容和研究工具. 不像"信息技术"和"信息科学"那样许多人不清楚其内涵,连专家学者对这些术语的界定都有许多争议.

1.2 概率论相关知识

本节只介绍本书涉及的概率论中的概念和知识.

和许多数学内容一样,对概率论,可以根据读者数学基础的不同,从不同的角度入手展开讨论. 在我国,面对初学概率论的读者,一般书籍总是从古典概型出发,引入概率的概念,展开概率论的内容. 而面向有一定概率论基础的读者的书籍,往往用公理化方法讲述概率论知识. 就后者而言,这样做的原因在于:高深一点的数学内容要体现数学的本质特点. 一种被普遍接受的观点认为,数学是研究空间形式和量化关系的科学,数学的精髓就是呈现这样的形式与关系. 而公理化方法应用于表述空间形式和量化关系,是最为方便的.

空间是赋予一定结构的集合. 例如,实数直线是一维的线性空间,直角坐标平面是二维的线性空间,以至线性代数中任意有限维的线性空间. 拓扑学建立在拓扑空间的基础上,泛函分析以赋范线性空间为出发点,其中有可数无穷维空间、不可数无穷维空间,分形学讨论分数维空间. 在代数学中,与空间相应的集合直接赋予群、环、域等称谓.

在给出空间的基础上,揭示其中量化关系的方便手段是应用公理化方法,该方法的一个显著的特点是:不断用命令的方式,如定义或公理,强加于空间或其元素一些性质,直接以空间本身、其中集合或元素为原料,性质为法则建立起系统,而不对这些性质的背景进行说明.

什么是数学中的背景? 这个背景就是应用. 对于一种数学内容来说,其应用有两种,一是人类实际生活中的应用,二是数学领域中的应用. 由于公理化方法不讲背景,特别是实际生活中的应用,所以,如果一个人缺少必要的数学知识,面对以公理化方法介绍的数学内容,会感到十分不解. 即使对于已经有了这种内容的初步知识,但还没有达到接受公理化方法程度的读者,在学习以公理化方法介绍的内容时,也会感到

困惑. 那些初步知识,不能与后者建立联系,似乎以不同方法介绍的同一东西,完全是不同的东西. 因此有人认为,数学在人类实际生活中的应用与公理化方法介绍的内容之间,存在不可逾越的鸿沟. 还有人认为,这种现象是公理化方法的缺陷. 我们认为,关于两者存在鸿沟的认识是正确的,但不能说此鸿沟是公理化方法的缺陷导致的. 正像真正的山间鸿沟,可以说土石材质的不同产生了鸿沟,不能说处于高处的坚石有缺陷. 有没有简单的办法消除这个鸿沟? 没有. 正像真正的山间鸿沟,不能说句话就可以填补,非得足够的材料不可.

以下用公理化方法介绍概率论知识.

1.2.1　概率空间与随机变量

1. 可测空间

首先介绍可测空间. 顾名思义,可测空间就是可以给予测度的空间.

定义 1.2.1　设 Ω 为一点集,$\mathfrak{F}$ 为 Ω 的一些子集组成的集族,满足

(i) $\varnothing,\Omega\in\mathfrak{F}$;

(ii)若 $A\in\mathfrak{F}$,则 $\overline{A}=\Omega\backslash A\in\mathfrak{F}$;

(iii)若 $A_n\in\mathfrak{F}$,$n=1,2,\cdots$,则 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathfrak{F}$.

其中 $\varnothing$ 为空集. 称 $\mathfrak{F}$ 为一个 σ 代数,$(\Omega,\mathfrak{F})$ 为可测空间.

在代数学中,定义了满足一定关系的一个运算的集合称为群,这个运算可以称为加法,也可以称为乘法. 但一般来说,如果这个运算是可交换的,即甲与乙运算的结果等于乙与甲运算的结果,则这个运算称为加法;否则,如果这个运算是不可交换的,即甲与乙运算的结果不总是等于乙与甲运算的结果,则这个运算称为乘法. 定义了加法的集合称为加法群,定义了乘法的集合称为乘法群. 加法群中有一个元素称为零元,它与加法群中任何一个元素的和还是这个元素,每一个元素有一个负元素,二者的和等于零元,一个元素甲与另一个元素乙的负元素"－乙"的加法称为甲与乙的减法,记为"甲－乙";乘法群中有一个元素称为单位元,它与乘法群中任何一个元素的积还是这个元素. 定义了满足一定关系的两个运算的集合称为环,两个运算一个称为加法,另一个称为乘法,在所述关系中,乘法对于加法的分配率最重要.

我们可以像处理数字之间的加法与乘法运算那样处理环中元素的运算.

注意上述 $\mathfrak{F}$ 是 Ω 的一些子集组成的集族,其元素是 Ω 的子集,也就是说,$\mathfrak{F}$ 是集合的集合,而集合的集合常称为集族.

$\mathfrak{F}$ 是 Ω 的"一些"子集组成的集族,究竟是哪些子集? 是 Ω 的部分子集还是全部子集? 这无关紧要,我们只关心所述三个条件是否满足.

我们将集合之间的并看成加法,交看成乘法,集合的差看成减法,$\overline{A}$ 看成 A 的负元,则 $\mathfrak{F}$ 成为一个环. 这是因为

$$\bigcap_{n=1}^{\infty}A_n=\overline{\bigcup_{n=1}^{\infty}\overline{A_n}}=\overline{\bigcup_{n=1}^{\infty}\overline{A_n}},$$

当 $A_n \in \mathfrak{F}$ 时，由(ii)知 $\overline{A_n} \in \mathfrak{F}$，由(iii)知 $\bigcup_{n=1}^{\infty} \overline{A_n} \in \mathfrak{F}$，再有(ii)知 $\bigcap_{n=1}^{\infty} A_n \in \mathfrak{F}$；对正整数 $m \geqslant 2$，取 $A_{m+1} = A_{m+2} = \cdots = \varnothing$，则 $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{m} A_n$，知 $\mathfrak{F}$ 对有限并封闭，即 $\mathfrak{F}$ 中有限个元的并仍属于 $\mathfrak{F}$；取 $A_{m+1} = A_{m+2} = \cdots = \Omega$，则 $\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{m} A_n$，知 $\mathfrak{F}$ 对有限交封闭，由集合的交并运算规则，可知 $\mathfrak{F}$ 为环.

关于代数，情况有点复杂，不同的书籍有不同的定义. 按照一些书籍的说法，由于 Ω 本身是 $\mathfrak{F}$ 的元素，因此称 $\mathfrak{F}$ 为一代数.

称 $\mathfrak{F}$ 为一个 σ 代数，是因为 $\mathfrak{F}$ 对于可数并运算封闭，因为 σ 是字母 Σ 的小写体，后者常表示求可数和.

2. 概率空间

有了可测空间，就表示可以在其上定义测度了.

定义 1.2.2 设 $(\Omega, \mathfrak{F})$ 为可测空间. $\mathfrak{F}$ 上定义的一个集函数 P 称为一个概率测度，若

(i) $\forall A \in \mathfrak{F}$，$P(A) \geqslant 0$；

(ii) $P(\Omega) = 1$；

(iii)若 $A_n \in \mathfrak{F}$，$n = 1, 2, \cdots$，且两两不相交，则

$$P(\bigcup_{n=1}^{\infty} A_n) = \sigma \sum_{n=1}^{\infty} P(A_n).$$

称 $(\Omega, \mathfrak{F}, P)$ 为一个概率空间，Ω 中点称为基本事件，$\mathfrak{F}$ 中元称为事件，对 $A \in \mathfrak{F}$，$P(A)$ 称为事件 A 的概率. 对 $A, B \in \mathfrak{F}$，常将 $A \cap B$ 记为 AB.

测度是长度、面积、体积这些概念的概括. 在直线上，线段的长度是测度，在平面上，图形的面积是测度，在三维空间中，物体的体积是测度. 因此，测度值是非负的，如(i)，部分的测度和等于整体的测度，如(iii)，后者被称为概率的可数可加性. (ii)被称为概率测度的概率性质，没有这条性质的 $(\Omega, \mathfrak{F}, P)$ 是一般的测度空间，有了这条性质的测度空间就成为概率空间了.

以上定义概率空间的方法，就是公理化方法. 可以看出，依这种方法给出概率空间的概念，十分简洁. 但是，要将这里的因素与古典概型联系，却会发现许多问题. 例如，在古典概型中，现在的 Ω 是什么，其中的元素是什么，都是十分困难的问题.

对一些十分简单的情形，Ω 及其元素还是容易搞清的.

例如，向一张桌面上扔一枚一元人民币硬币，假定该硬币的图形对硬币落下以后哪面向上的结果不产生影响，还假定硬币扔下以后不会立着不倒. 如此，$\omega_1 =$ {有国徽的一面朝上}，$\omega_2 =$ {有花的一面朝上}，$\varnothing =$ {哪面都不朝上}，Ω 为二元点集 $\Omega = \{\omega_1, \omega_2\}$，则 $\mathfrak{F} = \{\varnothing, \{\omega_1\}, \{\omega_2\}, \Omega\}$，其中 $\{\omega_1\}$，$\{\omega_2\}$，分别为仅包含 ω_1，ω_2 的单点集. 可知 $P(\{\omega_1\}) = \frac{1}{2}$，故 $P(\Omega) = P(\{\omega_1, \omega_2\}) = 1$，而 $P(\varnothing) = 0$，是故，$(\Omega, \mathfrak{F}, P)$ 为一个概率空间.

对于复杂情况，如果仔细考虑，Ω 及其元素以至 $(\Omega,\mathfrak{F},P)$ 还是可以搞清楚的，不过，许多情况下我们没有必要将精力花费在这上面，而仅关注所感兴趣的方面.

3. 随机变量

有了空间，要考虑其中的数量关系，还需要将有关因素数量化，对 Ω 中元 ω，将其数量化的不二选择是给其赋值，这样做的最好方式是定义函数 $X(\omega),\omega\in\Omega$.

有了函数，需要将该函数与 $\mathfrak{F}$，P 联系起来. 由多方面的原因，要求限定一个范围，当该函数的值 $X(\omega)$ 处于这个范围时，对应的自变量 ω 构成的集合应该属于 $\mathfrak{F}$. 例如，取区间 $(a,b]$，应有

$$X^{-1}((a,b])=\{\omega\in\Omega:a<X(\omega)\leqslant b\}\in\mathfrak{F}.$$

(集合 $X^{-1}((a,b])=\{\omega\in\Omega:a<X(\omega)\leqslant b\}$ 称为 $(a,b]$ 的逆象，诸如此类的记号是数学中的常规记号).

在所述多方面的原因中，$X(\omega)$ 的 Lebesgue 可积性十分重要，以下以 $X(\omega)$ 为一元函数为例说明 Lebesgue 积分的定义原理.

设一个一元函数 f 定义于 x 轴上的区间 $[s,t]$，其值包含于 y 轴上的区间 $[a,b]$，即

$$f([s,t])=\{y=f(x):x\in[s,t]\}\subset[a,b].$$

将 $[a,b]$ 分割为若干份(不是分 $[s,t]$，在数学分析课程中学过的积分称为 Riemann 积分，该积分分割 $[s,t]$)，记这个分割为 Δ. 对每份 $[a_i,b_i]$，取函数值属于 $[a_i,b_i]$ 的自变量构成的集合 $E_i=f^{-1}[a_i,b_i]=\{x\in[s,t]:f(x)\in[a_i,b_i]\}$ 的测度 $m(E_i)$，作和

$$m_\Delta=\sum a_i m(E_i),\quad M_\Delta=\sum b_i m(E_i),$$

分别称为相应于 Δ 的小和、大和. 当 Δ 变化时，分别得到小和的数集 A 和大和的数集 B，取 A 的上确界，B 的下确界，如果二者相同，则此共同值称为 f 在 $[s,t]$ 上的 Lebesgue 积分，记为 $\int_{[s,t]}f(x)\mathrm{d}x$，此时称 f 在 $[s,t]$ 上 Lebesgue 可积.

可以看出，f 在 $[s,t]$ 上 Lebesgue 可积的第一要素是集合 E_i 有测度，即可测，因此，将对所有实数 $a\leqslant b$，$E=f^{-1}[a,b]=\{x:a\leqslant f(x)\leqslant b\}$ 可测的函数 f 称为可测函数.

研究表明，上述 $[a,b]$ 代之以 $(a,b]$，$[a,b)$，(a,b)，$(-\infty,b)$，$(-\infty,b]$，(a,∞)，$[a,\infty)$，得到的 f 的可测性都是等价的.

总之，从多种角度考虑，取函数 $X(\omega),\omega\in\Omega$ 为可测函数是有利的，这样的函数就是随机变量.

定义 1.2.3　设 $(\Omega,\mathfrak{F})$ 为可测空间，X 为 Ω 上定义的实函数，若对任意实数 a，b，$a\leqslant b$，$X^{-1}((a,b])=\{\omega:\omega\in\Omega,a<X(\omega)\leqslant b\}\in\mathfrak{F}$，即依测度论的观点 X 为 Ω 上的可测函数，则称 X 为可测空间 $(\Omega,\mathfrak{F})$ 上的一个随机变量，或 Ω 上的一个随机变量.