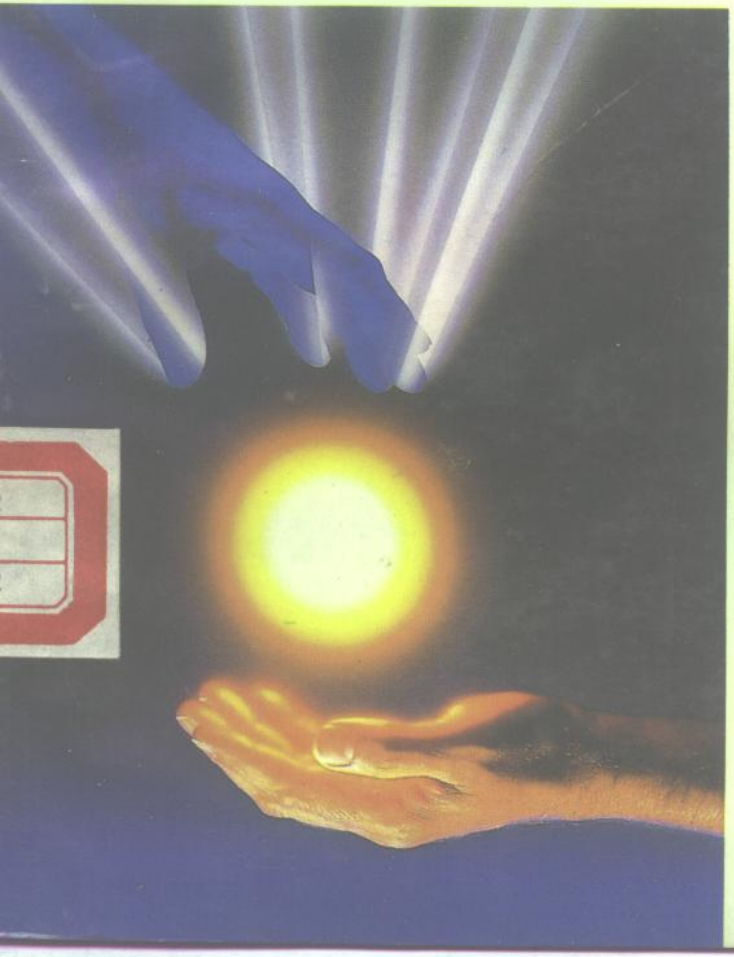
信息的度量及其应用

A MEASURE
OF INFORMATION AND ITS APPLICATION

走向数学丛书

沈世镒　著



走向数学丛书

信息的度量及其应用

沈世镒　著

湖南教育出版社

信息的度量及其应用

A Measure of information and its application

沈世镒 著

Shen Shi Yi

责任编辑：孟实华

湖南教育出版社出版发行（东风路附1号）

湖南省新华书店经销 湖南省新华印刷二厂印刷

787×1092毫米 32开 印张：3.875 字数：80000

1993年12月第1版 1993年12月第1次印刷

ISBN7-5355-1789-7/G·1784

定价：3.70元

"走向数学"丛書

陳省身題

作 者 简 介

沈世镒，男，1939年4月生于上海，1956年9月入天津南开大学数学系，1961年毕业。同年为研究生，功读方向为“信息论”，1965年研究生毕业。1982年至1983年为美国康乃尔大学（Cornell）、斯坦福（Stanford）大学访问学者。1986年为南开大学教授。

主要研究方向为信息论的理论与应用，如信息论的编码问题、多用户信息论。信息统计与人工神经网络系统等。已发表论文40余篇，专著一部。承担国家自然科学基金、教委重点学科基金、“七五”项目等多项工作。

前　言

王　元

从力学、物理学、天文学直到化学、生物学、经济学与工程技术，无不用到数学．一个人从入小学到大学毕业的十六年中，有十三、四年有数学课．可见数学之重要与其应用之广泛．

但提起数学，不少人仍觉得头痛，难以入门，甚至望而生畏。我以为要克服这个鸿沟，还是有可能的．近代数学难于接触，原因之一大概是由于其符号、语言与概念陌生，兼之近代数学的高度抽象与概括，难于了解与掌握．我想，如果知道讨论的对象的具体背景，则有可能掌握其实质．显然，一个非数学专业出身的人，要把数学专业的教科书都自修一遍，这在时间与精力上都不易做到．若停留在初等数学水平上，哪怕做了很多难题，似亦不会有助于对近代数学的了解．这就促使我们设想出一套“走向数学”小丛书，其中每本小册子尽量用深入浅出的语言来讲述数学的某一问题或方面，使工

程技术人员，非数学专业的大学生，甚至具有中学数学水平的人，亦能懂得书中全部或部分含义与内容．这对提高我国人民的数学修养与水平，可能会起些作用．显然，要将一门数学深入浅出地讲出来，决非易事．首先要对这门数学有深入的研究与透彻的了解．从整体上说，我国的数学水平还不高，能否较好地完成这一任务还难说．但我了解很多数学家的积极性很高，他们愿意为“走向数学”撰稿．这很值得高兴与欢迎．

承蒙国家自然科学基金委员会、中国数学会数学传播委员会与湖南教育出版社支持，得以出版这套“走向数学”丛书，谨致以感谢．

引　言

在当代社会中，信息和信息科学这两个名词已为人们所熟知。它们的重要作用也正在被人们所接受。但究竟什么是信息？什么是信息科学？信息到底能否度量？怎样度量？对这些问题人们往往不十分清楚，从最广义的情形来说，这些问题涉及到信息的哲学解释与哲学地位。在本书中，我们不打算讨论这种一般性的问题，而是试图介绍几种关于信息度量的比较成熟的定义形式，以及它们的产生过程、性质与应用。我们希望，通过这些讨论可以帮助人们理解有关信息、信息的度量与信息科学的内容与意义，以及它们的概念特征与应用范围。

由于信息概念的广泛性，试图对信息的一般形式进行度量是十分困难的。本世纪20年代，奈奎斯特(H.Nyquist)与哈特莱(L.Hartley)就已指出了信息度量与通信理论的关系，以及它们与概率、对数函数的联系。1948年在仙农(C.E.Shannon)的著名论文《通信的数学理论》中，对信息的度量、通信中的编码问题给出了一系列确切的分析与论述，这些论述在以后的理论发展中得到了充分的证实与应用。因此，人们往往把信息论的产生与仙农的工作相联系，有关仙农所引进的信息度量被称为仙农熵，而相应的编码理论被称为仙农信息论。

仙农熵在信息度量中的成功不仅在于它具有明确的内在含意与严格的数学表达，而且它能确切地刻划出通信中的一系列本质特征。如由仙农熵确定的信号体积等概念就是信源在通信中的一种本质特征；又如由它派生的交互信息与信道容量反映了通信信道的基本特征。由仙农熵所确定的信息度量单位—“比特”已是通信理论与计算机科学中的一个基本单位。

由于信息概念的广泛性，关于信息度量的推广与一般形式的表达问题的探讨一直是人们密切关心的一个重大问题，任何新的度量形式出现都会导致新的信息科学的学科分支出现。除了由仙农熵及其派生或推广的信息度量之外，还可有与仙农熵有完全不同出发点的其它信息度量。例如算法信息论中的柯尔莫各洛夫(Kolmogorov) 复杂性等。因此，我们对信息度量的理解不仅要从它们的引进来源来理解，更重要的是要从这些量的应用特征来理解，这样就可使我们更好地掌握与应用信息度量的工具。

本书的主要目的就是希望介绍几种主要的信息量及它们所涉一些分支学科，尤其对仙农熵及其有关的信源、信道编码理论作较为完整的叙述。为了适合读者的不同要求，我们对比较专门的内容章、节用星号注明，对此可以省略不读，而不影响对全书基本内容的理解。另外，本书的参考文献将按专题列出，有兴趣的读者可再深入学习了解。

具有初步微分学与概率论知识的读者均可阅读本书。一些微分学与概率论中的基本知识与名词概念，如什么是集合，集合之间的相互关系与交、并、差、积运算，集合之间的映射(或变换)等，又如对数、指数函数，导数与极值，最大、最小值问题，及概率论中的随机试验，随机事件，随机变量，概率，概率分布，概率密度，均值与方差等等，对这些概念名词我们

不再一一解释，希望读者参考有关知识材料。

作者感谢中国数学会传播委员会及湖南教育出版社对本书写作出版的支持与帮助。

沈世镒

1991年12月25日

有关记号

$\boldsymbol{N}=\{1, 2, \cdots, n\}$，$\boldsymbol{M}=\{1, 2, \cdots, m\}$，
$\boldsymbol{M}'=\{1, 2, \cdots, m'\}$：部分自然数集合。
$\boldsymbol{A}=\{0, \pm 1, \pm 2, \cdots\}$，
$\boldsymbol{A}_+=\{0, 1, 2, \cdots\}$：整数与非负整数集合。
$\boldsymbol{GF}(q)=\{0, 1, \cdots, q-1\}$：$q$-值的有限域。
$\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{X}$，$\boldsymbol{Y}$，$\cdots$，$\boldsymbol{U}$，$\boldsymbol{V}$，$\boldsymbol{W}$：集合记号，或称为字母表。
$\boldsymbol{X}\times\boldsymbol{Y}$：$\boldsymbol{X}$，$\boldsymbol{Y}$的乘积空间。
a，b，x，y：分别为集合$\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{X}$，$\boldsymbol{Y}$中的元素，或称为字母。
X，Y，Z：分别取值于$\boldsymbol{X}$，$\boldsymbol{Y}$，$\boldsymbol{Z}$中的随机变量。
$p(x)=P_r\{X=x\}$：随机变量X取值为x的概率。
$q(y)=P_r\{Y=y\}$：随机变量Y取值为y的概率。
$p(x, y)=P_r\{X=x, \ Y=y\}$：随机变量$(X, Y)$取值为$(x, y)$的概率。
$p(x/y)=P_r\{X=x/Y=y\}$：随机变量X在Y取值为y的条件下取值为x的条件概率。
$q(y/x)=P_r\{Y=y/X=x\}$：随机变量Y在X取值为x的条件下取值为y的条件概率。
$p_i=P_r\{X=i\}$：随机变量X取值为i的概率。
$q_j=P_r\{Y=j\}$：随机变量Y取值为j的概率。
$p_{i,j}=P_r\{X=i, \ Y=j\}$：随机变量$(X, Y)$取值为$(i, j)$

的概率。

$p_{i/j}=P_r\{X=i/Y=j\}$：在Y取值为j的条件下，X取值为i的条件概率。

$q_{j/i}=P_r\{Y=j/X=i\}$：在X取值为i的条件下，Y取值为j的条件概率。

$p(\cdot)=(p(x),\ x\in \boldsymbol{X})$或$p^m=(p_1,\ \cdots,\ p_m)$：

在$\boldsymbol{X}$或$\boldsymbol{M}$中取值的离散随机变量X的概率分布。

$q(\cdot)=(q(y),\ y\in \boldsymbol{Y})$或$q^{m'}=(q_1,\ \cdots,q_{m'})$：

在$\boldsymbol{Y}$或$\boldsymbol{M}'$中取值的离散随机变量X的概率分布。

$p(\cdot,\ \cdot)=(p(x,\ y),\ x\in \boldsymbol{X},\ y\in \boldsymbol{Y})$ 或 $p^{m,m'}=(p_{i,j},\ i\in \boldsymbol{N},\ j\in \boldsymbol{N}')$：$\boldsymbol{X}\times\boldsymbol{Y}$或$\boldsymbol{M}\times\boldsymbol{M}'$上的概率分布。

$\boldsymbol{X}^n=\prod_{i=1}^{n}\boldsymbol{X}_i$：$\boldsymbol{X}_1,\ \cdots,\ \boldsymbol{X}_n$的乘积空间，

$\boldsymbol{X}_i=\boldsymbol{X},i=1,\cdots,\ n$.

$x^n=(x_1,\ \cdots,\ x_n)$：$\boldsymbol{X}^n$中的元，或$\boldsymbol{X}$上取值的n－维向量。

$X^n=(X_1,\cdots,\ X_n)$：$\boldsymbol{X}^n$上取值的随机变量，或$\boldsymbol{X}(=\boldsymbol{X}_i)$上取值的$n$－维随机向量。

$p(x^n)=P_r\{X^n=x^n\}$：随机向量X^n取值为x^n的概率。

$\boldsymbol{Y}^n,\ y^n,\ Y^n,\ q(y^{n'})$：与$\boldsymbol{X}^n,\ x^n,\ X^n,\ p(x^n)$类似定义。

$\log(x)$，$\ln(x)$：对数函数，分别以2，e为底数。

$\exp_2(x)$，$\exp(x)$：指数函数，分别以2，e为底数。

$h[p(x)]=-\log[p(x)]$ 或 $h(p_i)=-\log p_i$：$p(x)$或p_i的熵密度。

$$H(X)=-\sum_{x\in X}p(x)\log p(x)\ \text{或}\ H(X)=-\sum_{i=1}^{n}p_i\log p_i$$：

随机变量X(分别在$\boldsymbol{X}$或$\boldsymbol{M}$上取值)的熵。

$h[p(x/y)]=-\log[p(x/y)]$ 或 $h(p_{i/j})=-\log p_{i/j}$：

$p(x/y)$或$p_{i/j}$的条件熵密度。

$$H(X/Y)=-\sum_{x\in X}\sum_{y\in Y}p(x,\ y)\log p(x/y)$$

或

$$H(X/Y)=-\sum_{i=1}^{m}\sum_{j=1}^{m'}p_{i,j}\ \log p_{i/j}$$

：随机变量X关于Y的条件熵。

$h[q(y/x)]=-\log[q(y/x)]$ 或 $h(q_{i/j})=-\log q_{i/j}$：$q(y/x)$或$q_{i/j}$的条件熵密度。

$$H(Y/X)=-\sum_{x\in X}\sum_{y\in Y}p(x,\ y)\log q(y/x)$$

或

$$H(Y/X)=-\sum_{i=1}^{m}\sum_{j=1}^{m'}p_{i,j}\log q_{i/j}$$

：随机变量Y关于X的条件熵。

$I(X;\ Y)=\sum_{x\in X}\sum_{y\in Y}p(x,\ y)\log\frac{p(x,\ y)}{p(x)\cdot q(y)}$ 随机变量X，Y的交互信息。

$i(x;\ y)=\log\frac{p(x,\ y)}{p(x)\cdot q(y)}$随机变量$X$，$Y$的交互信息密度。

$f_E(x)$：$\boldsymbol{X}\to\boldsymbol{U}$，从$\boldsymbol{X}$到$\boldsymbol{U}$的编码函数。

$g_D(v)$：$\boldsymbol{V}\to\boldsymbol{Y}$，从$\boldsymbol{V}$到$\boldsymbol{Y}$的译码函数。

$F(x)$，$P(x)$，$Q(x)$：随机变量X的分布函数，这时

$$F(x)=P(x)=Q(x)=P_r\{X<x\}.$$

$f(x)$，$p(x)$，$q(x)$：连续型分布函数的分布密度，这时

$$f(x)=dF(x)/dx,\qquad p(x)=dP(x)/dx,$$

$$q(x)=dQ(x)/dx.$$

目　录

第一章 概 论

§1.1 仙农信息论的产生、发展与应用

1.信息论的形成与发展

信息的概念是一个十分广泛的概念，有的哲学家甚至把它看作客观世界中有别于物质、能量的第三大要素，并把它看作是推动当今社会文明的主要因素．它的广泛性不仅涉及到人类社会的各个领域，而且在生物世界也都离不开信息的交流．从动物之间的各种动作交流到细胞的遗传生长都有信息的存在与作用．

在早期的人类社会中，人们的信息交流主要在文字、语言乃至动作表示上，当时就有利用各种信号包括符号来传达信息的种种表现(如烽火、鼓乐等等)，但人们大量的利用信息还是在电子通信与计算机技术出现之后．为了提高通信的质量与效率，人们往往从物理与数学两方面出发来进行研究考虑．在物理方面，主要工作是改进通信的物理手段与条件，如采用微波、卫

星、激光等手段使通信方式发生革命性的变化，同时通过对频带与信噪比等指标的改进也可提高通信的数量与质量．而数学的考虑角度则是在信号的设计构造与编、译码的算法上．也就是，在物理设备条件不变的情况下，用数学的方法来改进通信的数量与质量．在一般的通信问题中，数量与质量是两个相互制约且可相互补偿的指标．这样在保证通信质量的前提下，如何传送最多的信息就成为通信理论中的一个基本问题．因此要求我们在数学上解决一系列问题，例如，如何确定通信中信息传递的数量与质量的评估标准，如何在保证通信质量的前提下，提高它的数量问题及相应的计算机的运算实现等问题．

在本世纪20年代，奈奎斯特与哈特莱就提出了解决上述问题的一系列途径，如信息传递的速率与带宽成比例，信息的度量与信号的概率分布、对数函数相联系等等，这些思想为以后仙农信息论的产生打下了基础．

至本世纪40年代，“控制论”的奠基人维纳（N.Wiener）、美国统计学家费希尔(E.Fisher)与仙农几乎同时提出了信息度量的熵的定义形式．这个事实说明了从不同的学科出发，都会导致信息量这个概念的产生．1948年，仙农的著名论文《通信的数学理论》被认为是信息论产生的奠基性工作，因为该论文不仅对这种信息的度量作出了明确的描述，而且成功地利用了这种信息度量解决了信源、信道的编码问题．人们把这种信息的度量称为仙农熵，而相应的编码理论称为仙农信息论．仙农熵与仙农信息论是本书介绍的重点．

2.仙农信息论的发展与应用

自1948年仙农信息论产生以来，由于电子、通信与计算机技术发展的突飞猛进，信息论的发展也十分迅速．四十多年来，