



面向 21 世纪 课程 教材

高等学校
信息管理类专业
核心课教材

信息计量学

邱均平 主编

武汉大学出版社

面向 21 世纪 课程 教材

面向 21 世纪课程教材
高等学校信息管理类专业核心课教材

信息计量学

主 编	邱均平		
副主编	赵蓉英	侯经川	
编著者	邱均平	赵蓉英	侯经川
	黄晓斌	王宏鑫	徐久龄
	马锡敏	朱春艳	

武汉大学出版社

图书在版编目(CIP)数据

信息计量学/邱均平主编. —武汉: 武汉大学出版社, 2007. 1
面向 21 世纪课程教材
高等学校信息管理类专业核心课教材
ISBN 978-7-307-05305-2

I. 信… II. 邱… III. 文献计量学—高等学校—教材
IV. G257

中国版本图书馆 CIP 数据核字(2006)第 133430 号

责任编辑: 严红 谭必勇 责任校对: 程小宜 版式设计: 支笛

出版发行: 武汉大学出版社 (430072 武昌·珞珈山)

(电子邮件: wdp4@whu.edu.cn 网址: www.wdp.com.cn)

印刷: 湖北恒泰印务有限公司

开本: 880×1230 1/32 印张: 19.875 字数: 532 千字 插页: 2

版次: 2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷

ISBN 978-7-307-05305-2/G · 895 定价: 28.00 元

版权所有, 不得翻印: 凡购我社的图书, 如有缺页、倒页、脱页等质量问题, 请与当地图书销售部门联系调换

内 容 提 要

信息计量学是在文献计量学的基础上发展起来的一门重要的新兴学科,也是当前国际学术界研究最活跃的专业领域之一。本书从理论、方法、应用三个角度,首次构建了信息计量学的内容体系,全面、系统地论述了信息计量学的六大基本规律和理论基础,详细讨论了引文分析法、计算机辅助信息计量分析法等主要的定量分析方法,还集中探讨了在信息资源管理、情报学、图书馆学、科学学与科技管理、科学评价与预测等学科领域及其定量管理等方面的具体应用。全书共 14 章,内容包括信息计量学的基本问题、文献信息的增长规律、老化规律和引证规律、布拉德福定律、齐普夫定律、洛特卡定律以及诸定律的共同理论基础、文献信息统计分析法、数学模型分析法、引文分析法、计算机辅助信息计量分析法以及信息计量学在信息资源管理、科技管理与研究中的应用等。最后还阐述了信息计量学的新发展—网络信息计量学的基本理论及应用等。

本书是“教育部面向 21 世纪课程教材”之一,可作为高等院校信息管理与信息系统、管理科学与工程、信息资源管理、电子商务、情报学、图书馆学、档案学、出版发行学、科学学与科技管理、科学评价与预测等专业的教材,也可以作为广大信息工作者、图书情报档案工作者、科研人员、评价人员及有关管理者的业务参考书。

前 言

20 世纪 60 年代以来,在图书馆学、文献学、情报学和科学学领域相继出现了三个类似的术语: Bibliometrics、Scientometrics 和 Informetrics,分别代表着 3 个十分相似的定量性的分支学科,即文献计量学、科学计量学和信息计量学(情报计量学)(简称“三计学”)。经过几十年的努力研究与推动,三计学都不同程度地取得了一定的进展,得到了学术界的广泛承认。在我国,不仅发表了大量的有关论著,开设了相应的大学课程,设立了研究生培养方向,而且还得到了政府管理部门的认定和支持。在 1992 年颁发的国家标准《学科分类与代码》(GB/T13745-92)中,文献计量学、情报计量学和科学计量学作为三级学科名列其中,分别属于相应的二级学科文献学、情报学和科学学。这充分说明三计学已有了自己的学科地位。

“情报计量学”名称最早出自德文 Informetrie,是由德国学者奥托·纳克(Otto Nacke)最先提出来的。与之对应的英文术语为 Informetrics,我国将其译为“情报计量学”,后来改称为“信息计量学”。当时纳克提出“情报计量学”概念,其目的是试图用以概括数学在情报学所有领域的应用。后来,情报(信息)计量学被定义为:采用定量方法来描述和研究情报(信息)的现象、过程和规律的一门学科。它是数学和统计学与情报学广泛结合而形成的情报学的一个新兴的定量性分支学科。

我们曾经提出:信息计量学应分为“广义信息计量学”与“狭义信息计量学”。前者主要探讨以信息论为基础的广义信息的计量问

题,其范围非常广泛。所谓“狭义信息计量学”就是我们通常讲的“信息计量学”(或情报计量学),主要是研究情报信息(或文献情报)的计量问题。它的主要内容是应用数学、统计学等定量方法来分析和处理信息过程中的种种矛盾,从定量的角度分析和研究信息的动态特征,并找出其中的内在规律。信息计量学是在传统文献计量学及科学计量学的基础上扩展和演变而成的,也是信息学(情报学)向定量化方向发展的客观需要和必然产物。它一方面从理论上大大提高信息学的科学性和精确性,另一方面又能为信息管理、图书情报管理等实际工作提供参考依据和指导。同时,信息计量学是在信息时代大背景下,随着信息资源电子化、数字化的发展,适应社会经济信息化和信息产业发展的迫切需要而逐步形成和发展起来的,具有鲜明的时代特征、重要的科学理论价值和实际指导意义。

从 20 世纪 80 年代初开始,我一直在武汉大学从事文献计量学、信息计量学和科学计量学的教学与科研工作。1983 年,我在全国高校中率先开设了“文献计量学”本科课程,并编著了同名教材《文献计量学》,经过油印(1983 年)、铅印(1985 年)和几年的教学使用后于 1988 年在科学技术文献出版社(北京)正式出版。这本书首次从理论、方法和应用相结合的角度构建了文献计量学的内容体系,受到学术界同行的好评和欢迎。它不仅被 10 多所高校采用作为教材,而且其被引率至今一直名列前茅。这“无疑是对我国情报学研究和情报学教育的积极贡献,具有开创性的意义”(著名情报学家杨沛霆语)。但是,由于它出版年代已久,有的内容亟待更新和补充。特别是在当前信息化、电子化、数字化、网络化的新形势下,信息管理领域出现了许多新情况、新问题、新特点、新规律,迫切需要我们以新的视角去研究、去解决,去寻求新的正确答案。文献计量正在向信息计量方向发展。这种科学研究和学科发展的需要是我们撰著本书的出发点之一。其二,是为了满足教学的需要。1999 年,《信息计量学》被选定和列入“教育部面向 21 世纪课程教材”出版计划。在开设“文献计量学”课程的基础上,早在 20 世纪 90 年代初,我就指导和培养计量学方面的硕士研究生,后来又增设了“信息资源计量研究”硕士

生方向和“信息计量与科学评价”博士生专业方向,先后以“文献计量学”、“信息计量学”、“情报计量学研究”、“科学计量学研究”、“网络计量学研究”、“网络信息资源计量研究”等为题,主讲了研究生和本科生课程。在 20 多年的学习、教学和科研工作中,我们积累了大量的心得体会、信息资料和研究成果,现借此机会成书出版,既作为“教育部面向 21 世纪课程教材”之一,又是我们 20 多年来研究“三计学”问题的系统总结及研究著作之一。

我们一直认为文献计量学、信息计量学和科学计量学既有十分密切的联系,又有一定的区别,而且它们之间的相互联系、交叉和重叠大于其间的区别,以致于国外有些人将它们视为同义语而混用,或者主张在不同的场合采用不同的学科名称。本书定名为《信息计量学》,主要考虑到以下几点:第一是与国内学术界将“情报”改为“信息”相对应,也相应地将“情报计量学”改为“信息计量学”;二是信息计量学已有一定的研究基础,也是学术界公认的分支学科;三是信息资源电子化、数字化和网络化的逐渐普及,为信息计量提供了必要的条件和可能,特别是“网络信息计量学”的迅速兴起和发展,使信息计量学面临着难得的发展机遇;四是以期能引起人们更多的关注和重视,旨在促进信息计量学的深入研究和发 展。这主要是面向未来,强调信息计量学是今后的研究重点和发展方向,而从本书的内容来看,必然会同时涉及到三计学,文献计量仍然是基础,讨论的重点主要是文献信息的计量问题。这是目前的研究现状和事实所决定的。但为了叙述的方便,书中有时也采用了“文献计量学”或“情报计量学”的名称。全书共 14 章,主要内容可概括为理论、方法和应用三个部分:一是关于信息计量学的理论问题研究,包括第 1~7 章、第 14 章和第 10 章中的部分内容;二是信息计量学的方法体系研究,包括第 8~11 章的内容;三是信息计量学的应用,主要在第 12、13 章中集中论述,同时在前面的有些章节中也论及了各定律的应用问题。特别是对计算机辅助信息计量分析法、网络信息计量学等新方法、新领域作了重点论述和探讨。但限于篇幅,作者所收集和编写的许多实例和各章的复习思考题未能收入本书中,这将在教学中予以补充

和使用。特别值得说明的是,本书中保留了原来《文献计量学》中的一些典型的应用实例,因为这些实例属于经典案例,至今仍然能够说明问题,目前又没有更好更有代表性的新的案例来替换,因此,这些内容未予更新。这样处理不应该影响本书的新颖性。在撰著过程中,我们试图从理论、方法、应用三个角度全方位地构建信息计量学的学科体系,注重理论与实践相结合,继承与创新相结合,兼顾传统的统计工具与新兴的信息技术方法,力求使全书的思路清晰、结构合理、论述全面、内容丰富、观点新颖、资料翔实,既反映和吸收国内外三计学的最新进展,又融入我们自己的研究成果,使之具有较强的科学性、创新性、系统性和实用性;既适合高等院校的信息管理与信息系统、管理学、信息资源管理、电子商务、情报学、图书馆学、档案学、出版发行学、科学学与科技管理、科学评价与预测等专业作为教材使用,也可供广大信息工作者、知识工作者、科研人员、评价人员和管理人员学习参考。

本书由邱均平主编,首先在邱均平编著的《文献计量学》一书的基础上提出了详细的撰著大纲;然后由8位作者分头撰写初稿;最后,邱均平、赵蓉英作了部分增删、修改或补充,并完成了统稿工作。各章的撰稿人是:第1、4章:邱均平、马瑞敏;第2、5章:侯经川;第3、7、9、12章:邱均平、王宏鑫;第6章:徐久龄;第8章:邱均平、朱春艳;第10、13章:赵蓉英;第11、14章:黄晓斌。本书是《文献计量学》的继承和创新,基本上是在《文献计量学》的基础上修改、补充、更新和扩展而成的。本书的出版得到了教育部“信息管理类专业教学内容和课程体系改革研究”项目总负责人康仲远教授、教育部高等教育司、武汉大学出版社和有关院、系领导的支持和帮助;责任编辑严红等编校人员为之付出了辛勤劳动。在此,我们谨向以上单位和个人表示最诚挚的谢意!

由于多人分头执笔,书中不妥之处乃至错误在所难免,恳请读者批评、指正。

邱均平 于路珈山

2006年7月26日

目 录

1 绪论	1
1.1 信息计量学的由来和发展	1
1.1.1 信息计量学的由来	1
1.1.2 信息计量学的产生背景	3
1.1.3 信息计量学的发展	5
1.2 信息计量学的概念和内容体系	14
1.2.1 信息计量学的研究目的和意义	14
1.2.2 信息计量学的研究对象	16
1.2.3 信息计量学的概念	17
1.2.4 信息计量学的内容体系	19
1.3 信息计量的工具和方法	20
1.3.1 信息计量的数据来源	20
1.3.2 信息计量的工具和应用软件	22
1.3.3 信息计量的方法体系	26
1.4 信息计量学与相关学科	28
1.4.1 信息计量学的相关学科	28
1.4.2 信息计量学与数学和统计学	29
1.4.3 信息计量学与文献计量学	31
1.4.4 信息计量学与科学计量学	32
1.4.5 信息计量学与网络信息计量学	34
1.4.6 信息计量学与科学评价学	34

2 文献信息增长规律	36
2.1 文献信息流的特性及增长规律研究的意义	36
2.1.1 文献信息流的特性	36
2.1.2 文献信息增长的影响及对策	37
2.1.3 文献信息增长规律的研究及意义	39
2.2 科学知识量的增长与科学文献的增长	40
2.2.1 科学知识量的增长规律	41
2.2.2 科学知识量的增长与科学文献增长的关系	44
2.3 文献信息的指数增长规律	45
2.3.1 文献信息量度指标与方法	45
2.3.2 文献信息指数增长模型	46
2.3.3 文献信息指数增长规律的分析	48
2.4 文献信息的逻辑增长规律	50
2.4.1 文献信息逻辑增长模型	50
2.4.2 文献信息逻辑增长规律的分析	52
2.4.3 文献信息逻辑增长模型的修正	54
2.5 文献信息增长的其他数学模型	56
2.5.1 线性增长模型	56
2.5.2 分级滑动指数模型	56
2.5.3 超越函数模型	58
2.5.4 舍-布增长模型	59
2.6 文献信息增长机理的分析	60
2.6.1 文献信息数量增长的原因	60
2.6.2 文献信息增长规律的解释	62
2.7 文献信息增长规律的应用	64
2.7.1 在科学学和科技史研究中的应用	64
2.7.2 在情报研究中的应用	65
2.7.3 在文献信息管理中的应用	65

3 文献信息老化规律·····	67
3.1 文献信息老化的概念和量度指标·····	67
3.1.1 文献老化与情报老化的概念·····	68
3.1.2 文献老化的量度指标·····	71
3.2 文献信息老化的研究方法·····	75
3.2.1 文献管理统计分析方法·····	75
3.2.2 引文分析方法·····	76
3.2.3 数学方法·····	77
3.2.4 综合分析方法·····	78
3.3 文献信息老化的数学模型及老化指标·····	79
3.3.1 经典数学模型及老化指标·····	79
3.3.2 灰色动态模型(GM)及老化指标·····	87
3.4 文献信息老化机理的研究与分析·····	90
3.4.1 文献信息老化的几种类型·····	90
3.4.2 文献信息老化的几种情形·····	91
3.4.3 文献信息老化的影响因素·····	92
3.5 文献信息老化规律的应用·····	94
3.5.1 在文献情报管理中的应用·····	95
3.5.2 在科学学与科技史研究中的应用·····	95
4 文献信息集中与离散分布规律——布拉德福定律·····	97
4.1 布拉德福定律的产生背景·····	97
4.1.1 布氏定律的创始人——布拉德福·····	97
4.1.2 布氏定律的产生背景·····	98
4.2 布拉德福定律的形成·····	101
4.2.1 布氏定律的提出·····	101
4.2.2 布氏定律的确立·····	104
4.3 布拉德福定律的基本内容·····	105
4.3.1 布氏定律的基本阐述·····	105
4.3.2 布氏定律理论原理与实际的一致性·····	109

4.4 布拉德福定律的发展	112
4.4.1 布氏定律的发展过程	112
4.4.2 维克利对布氏定律的推论	114
4.4.3 莱姆库勒对布氏定律的发展	117
4.4.4 布鲁克斯对布氏定律的描述	120
4.4.5 斯马利科夫的统一方程	121
4.4.6 布氏分布理论及发展趋势	123
4.5 布拉德福定律的应用	125
4.5.1 布氏定律应用的基本方法	125
4.5.2 布氏定律应用的主要领域	127
4.5.3 布氏定律应用的条件与局限	131
5 文献信息词频分布规律——齐普夫定律	132
5.1 齐普夫定律的理论基础——最省力法则	132
5.1.1 什么是“最省力法则”	132
5.1.2 最省力法则与词频分布规律	133
5.2 齐普夫定律的形成和确立	134
5.2.1 频率词典的出现	134
5.2.2 艾思杜的发现	135
5.2.3 贡东的公式	135
5.2.4 齐普夫的研究和齐普夫定律的确立	137
5.3 齐普夫定律的基本内容	139
5.3.1 齐普夫定律的文字表述	139
5.3.2 齐普夫定律的图像描述	140
5.3.3 齐普夫定律的一般数学形式	141
5.3.4 齐普夫定律的适用性	142
5.4 齐普夫定律的发展	143
5.4.1 朱斯的双参数公式	143
5.4.2 芒代尔布罗的三参数公式	144
5.4.3 低频词分布规律——齐普夫第二定律	145

5.5 齐普夫定律的应用	147
5.5.1 在文献标引和词表编制中的应用	148
5.5.2 在情报检索中的应用	151
5.5.3 在科学评价中的应用	152
6 文献信息作者分布规律——洛特卡定律	158
6.1 洛特卡定律的产生背景	159
6.1.1 洛氏定律的创始人——洛特卡	159
6.1.2 洛氏定律的产生背景	159
6.2 洛特卡定律的形成和基本内容	161
6.2.1 洛氏定律的形成	161
6.2.2 洛氏定律的内容	167
6.2.3 广义洛特卡定律	167
6.3 洛特卡定律的发展	169
6.3.1 洛特卡定律的验证工作	169
6.3.2 弗拉奇的贡献	174
6.3.3 洛特卡定律在我国的发展	177
6.3.4 合作者问题的研究	184
6.4 普赖斯定律及其他的作者分布	192
6.4.1 普赖斯定律	192
6.4.2 其他的作者分布	195
6.5 洛特卡定律的应用	197
6.5.1 洛特卡定律的作用	197
6.5.2 应用中要注意的问题	202
7 信息计量学的理论基础	203
7.1 信息计量学的基础与结构模型	203
7.1.1 信息计量学的基础	203
7.1.2 信息计量学的结构	207
7.2 文献信息分布特征、规律及其理论解释	210

7.2.1	文献信息集中与离散分布规律	210
7.2.2	文献信息分布规律的理论解释	211
7.3	信息计量学定律的数学理论研究	213
7.3.1	关于齐普夫定律的数学模型研究	213
7.3.2	关于洛特卡定律的数学研究	219
7.3.3	关于布拉德福定律的信息维分析	224
7.4	信息计量学分布规律的理论体系研究	231
7.4.1	布-齐-洛定律的一致性研究	231
7.4.2	信息计量学分布的一般数学模型	241
7.4.3	信息生产过程(IPP)	248
8	文献信息统计分析法	254
8.1	文献信息统计的意义和一般概念	254
8.1.1	文献信息统计及其意义	254
8.1.2	文献信息统计中的一般概念	256
8.2	文献信息统计的原则和指标	259
8.2.1	文献信息统计的原则要求	259
8.2.2	文献信息统计的指标体系	260
8.2.3	信息资源管理中的统计指标	262
8.3	文献信息统计的类型和基本步骤	263
8.3.1	文献信息统计的主要类型	263
8.3.2	文献信息统计分析法的基本步骤	265
8.4	文献信息统计分析法的应用	268
8.4.1	在信息资源管理中的应用	268
8.4.2	在信息用户和文献信息利用研究中的应用	269
8.4.3	在文献信息规律研究中的应用	269
8.4.4	在学科发展规律研究中的应用	270
8.5	数理统计方法及其应用	273
8.5.1	数理统计方法概述	273
8.5.2	数理统计方法的应用	274

9 数学模型分析法	276
9.1 数学方法及其意义	276
9.1.1 数学方法的概念	276
9.1.2 数学方法的特征	276
9.1.3 数学方法应用的意义	277
9.2 数学模型法原理	278
9.2.1 客观基础	279
9.2.2 科学背景	281
9.2.3 实际需要	281
9.3 数学模型的类型	282
9.3.1 模型及简化模型	282
9.3.2 数学模型的类型	282
9.4 数学模型的建立	284
9.4.1 数学模型法的基本步骤	284
9.4.2 建立数学模型举例	286
9.5 数学模型的检验	288
9.5.1 参数估计及其应用	288
9.5.2 假设检验及其应用	292
9.6 数学模型分析法的应用	300
9.6.1 在信息资源管理中的应用	300
9.6.2 在信息检索中的应用	302
9.6.3 在文献信息规律研究中的应用	303
9.6.4 在学科动态研究中的应用	304
9.7 回归分析法及其应用	305
9.7.1 回归分析原理	305
9.7.2 一元线性回归分析法及应用	309
9.7.3 非线性回归分析法及应用	312

10 引文分析法	315
10.1 引文分析的基本概念和方法	315
10.1.1 引文分析的基本概念	315
10.1.2 引证行为与引证动机	317
10.1.3 引文分析的基本类型和步骤	319
10.2 引文分析的主要工具	321
10.2.1 美国《科学引文索引》(SCI)	321
10.2.2 美国《基本科学指标》(ESI)	336
10.2.3 国内引文分析的主要工具	359
10.3 引文分布规律及主要指标分析	362
10.3.1 引文结构及意义	362
10.3.2 引文量的分布规律	363
10.3.3 加菲尔德引文集中定律	365
10.3.4 引文测度的主要指标分析	367
10.3.5 科学文献的自引分析	373
10.4 科学期刊的引文分析	375
10.4.1 期刊文献的分散与集中规律	375
10.4.2 评价期刊的主要测度指标	377
10.4.3 《期刊引证报告》(JCR)	385
10.5 引文网络与聚类分析	397
10.5.1 引文耦合与同被引的概念	397
10.5.2 耦合分析	404
10.5.3 同被引分析	408
10.5.4 引文的聚类分析	412
10.6 引文分析法的应用	418
10.6.1 引文分析法的应用领域	419
10.6.2 引文分析法的应用举例	423
10.6.3 引文分析法的局限性	426

11 计算机辅助信息计量分析法	428
11.1 计算机辅助信息计量分析的意义	428
11.1.1 提高信息计量分析的效率	428
11.1.2 提高信息计量分析的可靠性	430
11.1.3 促进信息计量研究的精确化	430
11.1.4 拓展信息计量学的研究领域	431
11.2 计算机辅助信息计量分析的可行性	432
11.2.1 计算机和网络技术的发展为其奠定了基础	432
11.2.2 文献数字化的发展为其创造了条件	433
11.2.3 信息计量学的发展为其提供了理论原理	434
11.2.4 国外的研究进展为其提供了经验	434
11.3 计算机辅助信息计量分析的基本原理	435
11.3.1 计算机辅助信息计量分析的主要方式	435
11.3.2 计算机信息计量分析系统的结构与功能	437
11.3.3 计算机辅助信息计量分析的步骤	438
11.4 引文数据库的建立与数据挖掘分析	439
11.4.1 引文分析数据库	439
11.4.2 引文分析系统的设计	440
11.4.3 中文社会科学引文索引网络版设计案例分析	443
11.4.4 引文数据的挖掘分析方法	446
11.5 计算机辅助信息计量分析法的应用	450
11.5.1 在科学研究中的应用	451
11.5.2 在信息资源领域的应用	452
11.5.3 在竞争力分析中的应用	453
11.6 计算机辅助信息计量分析的发展方向	453
11.6.1 向深度和广度发展	453
11.6.2 向实用化方向发展	454
11.6.3 向集成化方向发展	454
11.6.4 向模型化方向发展	455
11.6.5 向智能化方向发展	455