

向桂林 吴金艳 辛希孟

采访质量控制数学模型研究

摘要 为实现图书馆采访质量控制,可以建立数学模型。该模型粒度细、易操作。它对采访工作的指导作用是:调整采访政策,按价值构成调整购书经费分配,用最少的钱购回读者最需要的文献,兼顾馆藏建设。公式1。表5。参考文献2。

关键词 采访质量控制 馆藏建设 数学模型

分类号 G253

ABSTRACT In this paper, the authors propose a mathematical model for the quality control of library acquisitions, which has the following functions: regulating acquisitions policy, controlling budgets and monitoring collection development. 1 formula. 5 tabs. 2 refs.

KEY WORDS Quality control of acquisitions. Collection development. Mathematical modeling.

CLASS NUMBER G253

1 采访质量控制的重要性

采访工作的一个重要方面是采访质量控制。采访质量控制的重要性体现在两个方面。

(1)采访质量控制保证馆藏建设的原则、方针得以正确实施。每个馆一般都有自己的馆藏原则。比如买什么,不买什么;藏什么,不藏什么;剔除什么,保留什么等。馆藏建设包括购入与剔除两个互逆的动态过程,目的是形成一种有特色的和有生命力的馆藏^[1]。其中,购入方面,就是采访工作。提高馆藏质量,就得提高采访质量,理想状态下是保证每种书都符合馆藏规则。事实上,这是很难办到的。这既有客观因素,比如馆藏原则本身描述不清楚,很难把握;预定书目内容不详细,无法了解每一种书究竟有什么内容,这些内容是否与馆藏原则相关;也有主观因素,比如采访工作人员素质偏低,既不了解馆藏体系,也不了解读者需求,采访回来的书,肯定很少有读者借阅,效益不高。有了采访质量控制数学模型,把馆藏原则作为一项因素考虑到数学模型中去,借助于计算机,以量化的方式来控制采访质量。

(2)采访质量控制是提高文献信息服务质量的保证。在前一条的基础上,可以采访回来与馆藏原则相关,同时又兼顾读者需求的文献,满足读者需求方面,就容易了。如果采访质量没有控制,采访回一堆于馆藏无益,于读者无用的文献,形成“读者需要什么,馆藏没有什么”的局面,这是很可悲的,这不仅

是对购书经费的浪费,更是造成读者流失的原因。从理论上讲,这是馆藏利用率下降的主要原因。

2 采访质量控制数学模型

在谈采访质量控制之前,先要说明一个概念:书的价值。这个问题比较复杂,可以从社会效益上讲,也可以从经济效益上讲;作者、出版商、读者各自的认识都不一样。但是,什么样的图书是有价值的?这个问题的回答是肯定的:利用率越高,而书价越低的图书是有价值的。用于指导工作就是:用最少的钱,买回读者最爱读的书。正如刘仲英指出:目前图书馆经费紧缺,文献价格不断上涨的情况下,如何使用有限的资金发挥最大效益,值得关注^[2]。

2.1 图书价值的构成因素

一种书的价值可以由8个因素来刻画:(1)被借次数。这是最基本的因素。如果没有任何读者借过这种书,这种书的价值肯定不高。(2)拒借次数。读者来借某种书,而这种书的复本被借完(暂不考虑装订、修补等情况),称为拒借。(3)预约次数。在实际工作当中,预约次数几乎等于拒绝次数,因为读者被拒借之后,第一反应是预约。我们用其中一个因素就可以。(4)复本量:不同种的书,其复本量可能不同,这就有可能造成被借次数不一样。(5)读者级别。同一种书,被教授读过一次和被本科生读过一次是不一样的,即体现的价值是不一样的:教授在选书时,已经进行了去粗取精的过程,但仍要借这种

书,说明这种书的价值高,而一般读者在借书时,没有对比的概念,读者本人也不清楚这种书与同类书的其他种有何区别,属于“随便借一本”。本文后面的算法实例中,提出一个读者级别赋分表,就是来体现不同读者对图书的贡献度。(6)书价。对出版商而言,书价越高,越有价值,但是对图书馆而言,情况不会是这样,有可能出现低价格、高质量的书;也有可能是高价值、低质量的书,比如某些烫金的礼品书就是这种情况。(7)入藏时间。不同种书的入藏时间有早有晚,这可能造成被借次数不一样。(8)馆藏特色。每个馆都有自己的藏书建设原则,要维持一定的购买方向。本文后面的算法实例中,提出一个馆藏特色相关度赋分表,就是来体现各类书与馆藏特色的贴近程度。

2.2 各因素与图书价值的关系

(1)同样的入藏时间内,被借次数越多,书的价值越大;被借次数越少,书的价值越小。书的价值与被借次数呈正相关关系。(2)被借次数相同,入藏时间越长,书的价值越小;入藏时间越短,书的价值越大。书的价值与入藏时间呈负相关关系。(3)相同入藏时间内,同样被借次数情况下,复本量越大,书的价值越小;复本量越小,书的价值越大。书的价值与复本量呈负相关关系。(4)我们在强调购书经费短缺时,可认为书的价值与书的价格呈负相关关系,如果购书经费十分充足,可以不考虑价值因素。(5)馆藏特色这个因素,可以用馆藏特色相关度来描述:与馆藏原则贴合的书的馆藏特色相关度为1.0,不太贴合的书的馆藏特色相关度取小于1.0的值。书的价值与馆藏特色相关度呈正相关关系。(6)考虑到本数学模型的可操作性,暂不考虑预约次数因素。许多图书馆自动化系统并不提供某种书的被预约历史查询/输出功能。如果有读者感兴趣,可以认为书的价值与被预约次数呈正相关关系。(7)读者级别。本数学模型主要依据图书借还日志来进行计算,如果图书馆自动化系统能提供某种书的被借历史查询,就可使用本算法。而某种书的被借历史查询的记录中,记录的是某借书证号(读者)在某年某月某日借过某种书的某个复本,而读者证号一般表达了该读者的级别。综上所述,一种书的价值可以表达为:

$$\text{Value} = (\text{Count} \times \text{Relative}) / (\text{Period} \times \text{Copy} \times \text{Price})$$

其中,

Value:某种书的价值;

Count:该书的被借次数,且已经考虑了读者级别;

Relative:该书的馆藏特色相关度;

Period:入藏时间,即从此书入藏到计算时的时间,一般按年度计算;

Copy:该书的复本量;

Price:该书的价格。

2.3 进行采访质量控制

按上述算法,得出的是每一种书的价值。用这样的数据来指导工作,显然是不合适的:因为量太大,每个馆都有上万种,甚至十几万种图书。正常的做法是按大类来指导工作。比如,科学院图书馆,一般会说,科技图书有多少种多少册;古籍有多少种、多少册;石油大学图书馆,一般会说,石油石化类图书有多少种多少册。因为这样称呼(划分)能代表自己的馆藏特色,指导工作时,操作性很强。因为不同的馆,其馆藏特色不同,所以各馆的大类划分也不一样,比如,有的馆按少儿图书、社科、港台图书来划分;有的馆按文学、政治法律、经济来划分,不一而足。各馆的大类划分都可以归宗于“中图法”,只需建立一种对应关系,比如,文学类含“中图法”的I类和K8类;石油石化类含“中图法”的P类和TE类。本文后面的算法实例中,提出了馆藏图书分类表,就是来表达上述情况。我们在计算出单种图书的价值后,可以通过求平均值办法,得出每个大类的价值。通过计算不同时间段内图书的价值,如果发现某类书的价值增大或变小,可以据此调整采访政策,按价值构成来调整购书经费的分配,力争用最少的钱,买回读者最喜爱读的书,而且兼顾馆藏建设。这就是本数学模型对采访工作的指导作用。

3 采访质量控制数学模型的算法实例

本实例中的数据全部来自北京石油管理干部学院图书馆,图书馆自动化系统为IALS5.0。为了节省篇幅,在此列出每种书的被借历史时,因为借出记录与还回记录成对出现,故删去了还回记录。另外也略去了经济类(F)、文学类(I,K)和外语类(H)各随机选取的2种图书的价值计算表。

表1 读者级别赋分表

读者类型	赋分	证号特征
教师	10	以'0'打头
学生	5	以'1'打头

注:分值是相对值,所以,最后的书的价值也是一个相对的值,即,这种书相对于其它书的价值。

表 2 馆藏特色相关度赋分表
(表达某类书与馆藏特色的相关程度)

中图法分类号	赋分
D	1.0
F	1.0
I	0.5
K	0.5
H	0.8

表 3 馆藏图书分类表

类名	包含的中图法类
政治、法律	D
经济	F
文学	I、K
外语	H

表 4 随机选取的 2 种政治、法律类图书

第 1 种	题 名	责任者	索取号	价格	入藏日期	复本量	累计借数
基本信息	现代人事考评技术及其应用	萧鸣政著	D035.2/11	24.50	19980319	4	4
被借情况	年 月 日	时:分:秒	条码号	流通地点	状态	证号	读者名称
	19990122	09:36:06	0105237	clc	借出	000180	王大伟
	19990312	10:29:01	0105239	clc	借出	000195	郭勇刚
	19990322	15:06:18	0105238	zsh	借出	100543	戴志坚
	19990419	15:38:35	0105236	ttyp0	借出	100646	邓逐耀
被借次数	count = 10 + 10 + 5 + 5 = 30 从读者级别赋分表中查出读者等级并累计						
价 值	value = 30 × 1.0 / (2001 - 1998) × 4 × 24.5 = 0.10204 使用前述数学模型计算						
第 2 种	题 名	责任者	索取号	价格	入藏日期	复本量	累计借数
基本信息	中国可以说不是: 冷战后时代的政治与情感抉择	宋强等著	D822.371.2/1	19.80	19970109	6	3
被借情况	年 月 日	时:分:秒	条码号	流通地点	状态	证号	读者名称
	19990408	15:22:51	0100260	ttyp0	借出	100727	任仁
	19990527	10:00:12	0100260	zsh	借出	100353	柳松
	19990608	15:10:35	0100260	zsh	借出	100784	王永军
被借次数	count = 5 + 5 + 5 = 15						
价 值	value = 15 × 1.0 / (2001 - 1997) × 6 × 19.8 = 0.03156						

表 5 采访经费预算表(假设购书经费 100 万元)

类 名	包含的中图法类	本类图书的价值	比 重 (%)	应分配经费 (万元)
政治法律	D	value _{政治法律} = (0.10204 + 0.03156) / 2 = 0.0668	5.7	5.7
经 济	F	value _{经济} = (0.14029 + 0.29461) / 2 = 0.21745	18.5	18.5
文 学	I、K	value _{文学} = (0.83333 + 1.20192 + 0.02824) / 3 = 0.68783	58.7	58.7
外 语	H	value _{外语} = (0.29166 + 0.08438) / 2 = 0.18802	16.9	16.9

刘可静

WTO 与图书情报业及国际图联的立场

摘要 《服务领域贸易总协定》(GATS)和《与贸易有关的知识产权协议》(TRIPS)使图书情报业的地位、生存和发展产生问题。我国图书情报界应根据国际图联的立场,对入世后面临的诸多问题进行研究。参考文献6。

关键词 世界贸易组织 国际图联 服务领域贸易总协定 与贸易有关的知识产权协议

分类号 G251

ABSTRACT GATS (General Agreement on Trade in Services) and TRIPS (Trade-Related Intellectual Property Rights) challenge the position, survival and development of library and information services. In this paper, the author thinks that we should be in accordance with the position of IFLA and make studies on various issues occurring after China's entrance into WTO. 6 refs.

KEY WORDS WTO. IFLA. GATS. TRIPS.

CLASS NUMBER G251

1 世界贸易组织与图书情报业的关系

伴随着国际贸易活动形式的巨大变化,世界贸

易组织结束了关贸总协定(GATT)只涉及物质商品贸易的历史,揽括了物质商品、服务、投资规则和知识产权等几乎所有当今世界与贸易有关的领域,并

4 结束语

本数学模型的特点在于粒度细、可操作性强,从最原始的图书借还记录,推导出采购经费预算方案,每一步都是可信的,建立在真实的数据之上,几乎没有个人经验掺杂进来。而且,由于图书馆的读者群相对固定,在一定的时限内,读者的需求也是有规律可循的,其间根本没有投机的因素,因而寻求定量分析的方法是可行的。在本算法提出之前,大部分采购人员陷入一种怪圈,比如花大量的钱买计算机类图书,因为他到书库一看,大部分读者在看计算机类的图书,所以采购时,又狠买计算机类的图书。显然,采购人员眼中只有宏观上的利用率,根本不考虑浪费问题 and 价值问题。这就是对图书价值的理解不完整造成的,只抓住了图书价值构成要素中的一两个方面,而且没有量化,以此来指导工作,势必要出差错。

本算法除了用于采访质量控制外,还可以用于以下方面:(1)对采购人员的工作进行评估。比如规定采购回来的书,其价值必须达到某一水平(值)。(2)确定各种图书在藏书布局中的位置。比如把质量高的书,集中放在读者易接触的地方。(3)为藏书复选和剔除提供依据。现在很多馆在搞藏书剔旧时,没有量化标准,一切凭感觉和经验,这是很不科学的。(4)为阅读指导提供参考。把质量好的书推荐给读者,在藏书建设和读者需

求之间,建立一种良性互动关系。

本文提出的数学模型还需进一步研究的问题:

- (1)影响图书价值的因素是否考虑完整。比如,由于不同类的书的出版种数不同,这必然影响到采访工作中的选书问题,是否应设置一个出版影响因子。
- (2)图书价值公式的表达是否完美。首先,因素之间是否都是与线性无关的,如果因素之间存在着相关关系,如何进一步做函数分析。其次,因素之间是否可用加(减)连接,比如:

$$\text{Value} = (\text{Count} + \text{Relative}) / (\text{Period} + \text{Copy} + \text{Price})$$

对于这些问题,笔者也未深入研究,还请广大读者共同讨论。

参考文献

- 1 鞠文红,辛希孟.网络环境下信息资源建设的新理念及其实践和发展前景.中国图书馆学报,2001(4)
- 2 刘仲英.试论高校图书馆文献资源开发与利用.河南高校图书馆情报工作,1994(4)

向桂林 中国科学院文献情报中心2001级博士生.通讯地址:北京市北四环西路33号.邮编100080。

吴金艳 北京石油管理干部学院图书馆自动化系统管理员.通讯地址:北京西三旗黄土店15号.邮编100096。

辛希孟 中国科学院文献情报中心研究馆员,博士生导师.通讯地址:北京市北四环西路33号.邮编100080。

(来稿时间:2001-12-25)