

高校图书馆读者教育培训现状的统计与预测

邓 盛¹, 侯奇焜¹, 陈敏哲¹, 杨宇舟², 赵夏苓³

(1. 湖南师范大学外国语学院, 湖南 长沙 410081)

(2. 厦门大学经济学院, 福建 厦门 361005)

(3. 波士顿大学经济学院, 美国 波士顿 02215)

[摘要] 基于中国知网(CNKI)近 15 年相关高校图书馆培训文献数据的统计,探讨了高校图书馆读者教育培训现状的特点,并对其发展进行了预测,得出图书馆读者教育培训的特点;进一步通过分析文献成果与读者调研数据,可以有效预测未来几年高校图书馆读者教育培训的发展趋势。

[关键词] 高校图书馆, 教育培训, 读者, 预测, 统计分析

[中图分类号] C812 **[文献标志码]** A **[文章编号]** 1001-4616(2022)02-0112-06

Statistics and Prediction of Reader Education and Training in University Library

Deng Sheng¹, Hou Qikun¹, Chen Minzhe¹, Yang Yuzhou², Zhao Xialin³

(1. Foreign Studies College of Hunan Normal University, Changsha 410081, China)

(2. School of Economics, Xiamen University, Xiamen 631005, China)

(3. School of Economics, Boston University, Boston 02215, USA)

Abstract: Based on the statistics of the training literature data of relevant university libraries in the recent 15 years by CNKI, this paper discusses the characteristics of the current situation of reader education and training in university libraries, predicts its development, and obtains the characteristics of reader education and training in university libraries. Furthermore, through the analysis of literature results and reader survey data, we can effectively predict the development trend of reader education and training in university libraries in the next few years.

Key words: university library, educational training, reader, prediction, statistical analysis

新形势下的高校图书馆读者教育培训工作主张“一切以读者发展为中心”,通过“高校图书馆读者教育培训”为主题词检索中文最大数据库 CNKI,可以从过去 15 年的 620 篇文献成果中清晰地看到中国大陆高校图书馆关于读者教育培训工作的发展。随着读者培训内容的不断拓展以及培训形式的多样化,高校图书馆读者教育培训工作针对培训项目内容的设计、调研、以及设置的形式都会根据各自特点开展。在教育培训期的宣传、培训形式的多样性结合也进一步吸引读者的注意力,使他们对培训课程产生浓厚的兴趣。针对培训效果,部分高校图书馆也会针对性地做一些读者调研或是网上培训随堂测。这些工作都会对高校图书馆实施培训教育的工作者予以及时反馈:如对培训项目的需求性是否强烈,对培训内容的知识点是否了解掌握,对施教者本次讲评或课程设计的建议,希望获得或偏重哪些方面的培训等,并以此来提高读者教育培训工作的实际效果。

目前,高校图书馆已进入数字化时代,关于读者教育培训服务的内容和方式也正发生着深刻变化。培训者们似乎更注重教育培养解决问题能力,提高个人在工作中必须具备的知识技能与态度,以胜任某个特定的职务或执行其所从事的一切活动^[1]。通过近 15 年文献成果可以看到中国大陆高校图书馆在读者教育培训工作中取得了很大的进步,尤其是教育培训的形式:从面对面讲座的教授口口相传到简单的 PowerPoint 演示文档,发展到 MOOC、微课、手机视频,甚至还有少部分研究者提出的 lab.3.0 这类虚拟培训。

收稿日期:2021-07-15.

基金项目:湖南省自然科学基金面上项目(2020JJ4054)、国家社会科学基金一般项目(17BYY184)。

通讯作者:邓盛,助理馆员,研究方向:数据整理研究. E-mail:1970696568@qq.com

Web3.0 时代下学生获取文献检索课的行为既表现出与泛在学习相同的特点:随时、随地、集成地获得相关知识和服务,又表现出极强的参与性与合作学习的特点^[2]。高校图书馆的读者对教育培训服务不再停留在对原始教育培训模式上的需求,他们的需求越来越高,要求教育培训者将更加专业化的培训信息转向更加个性化的服务。现在的大学生读者都是在数字化的环境下成长,互联网和数字化信息的使用已经变成常态,在这样的环境下,高校图书馆也应随着读者的受教行为发生变化,需要更好地对读者的教育培训进行分析。选择合适的读者教育培训形式,结合读者受教过程中所要求的专业性、个性化服务而衍生的读者预约式培训活动。针对学科特点,尤其是学科馆员需要结合所属专业的特点,设置适合本专业信息资源的应用,而这些形式在 CNKI 检索文献中都有较好的体现。

图书馆作为高校教学和科研最为重要的补充基地,读者的教育培训内容与培训模式也十分重要。通常情况下,高校图书馆读者教育培训工作主要含有:新生入馆教育、文献检索引用以及专题讲座、多媒体实操辅导、数字参考咨询、知识竞赛和有奖问答等。基于“图书馆是唯一集中的场所,它提供了一个以用户为中心、服务为手段的环境,可以让日新月异的最新信息技术与传统的知识资源相结合”^[3]。这些具体的教育培训模式都会涉及不同学科类的读者,而交流的有效性则很大程度取决于读者行为的参与度。美国 Donald Beagle 教授认为“以数字化信息资源为背景构建一个信息供需双方协同的工作空间,使用户与馆员、用户与用户之间可以进行显性和隐性知识的交流和建模,这种多维的动态交互模式不仅能够为师生提供更好的信息服务,而且通过用户们相互的信息交流可拓展知识背景和进行知识创新”^[4]。从技术上说,将信息共享空间描述为“共同的学习场所、多媒体工作站、高科技教室以及小组研究空间的组合”^[5]是可以实现的。随着手机端口的进一步开发应用,使得读者可以便捷检索文献及使用电子数据库。随着科技的进一步发展,教育培训的推送也更简洁便利,培训教师们可通过 RSS 工具聚合各学科国内外的博客学术动态信息推送给读者,减少用户信息搜索和获取的成本^[6]。

1 数据处理与分析

1.1 数据来源

数据收集最近 15 年中文数据库 CNKI 以“高校图书馆读者教育培训”为关键词的 620 篇论文,并对它们进行年份整理与分类整理;同时收集南京师范大学图书馆官网上披露的近几年读者培训链接记录,通过发布时间进行时段性分类整理。由于读者培训实践与学术文献的内容无法一一对应,为了客观反映图书馆读者教育培训相关学术研究的内容与高校图书馆读者培训实践主要内容的大致对应关系,我们将图书馆培训内容分类与论文分类对应起来,以 1-6 为标志,构成了一个类别变量(见表 1)。

值得注意的是:从数据整理分析中可以看到每年文献内容对当年读者培训规划内容的影响——前 9 个月的文献已经占了当年学术文献的

多数。这是由于网页链接无法获取图书馆每年规划读者培训的时间,而导致我们无法精确区分发表文献对当年读者培训规划的影响。但我们可以发现,在接下来的几年里,不同的内容都在增加,其中关于读者培训的论文数量增长最快(见图 1)。

1.2 不同类别文献发表次数的估计

由于文献数量影响着图书馆的读者培训,所以为了预测读者培训次数的变化,首先需要预测不同类别文献数量在未来的变化。在这种情况下,本文采用普通最小二乘法(OLS)、负二项回归(NBREG)、广义线性模型(GLM)分别预测文献数量,其中 GLM 又根据数据的分布不同,分为负二项 GLM、泊松 GLM 与伽马 GLM 3 种情况^[7-11]。

表 1 分类变量设定

Table 1 Set classification variables

类别变量 classid	图书馆培训内容分类	论文内容分类
1	多媒体利用	新媒体的运用
2	开题文献调研 科研创新 论文查找 论文提交文档编排 文献管理 数据库 资源检索 资源利用	资源文献利用
3	预约讲座	学科嵌入式预约
4	阅读	阅读
5	论文写作与期刊投稿	职业素养
6	新媒体新生入馆教育	读者培训工作

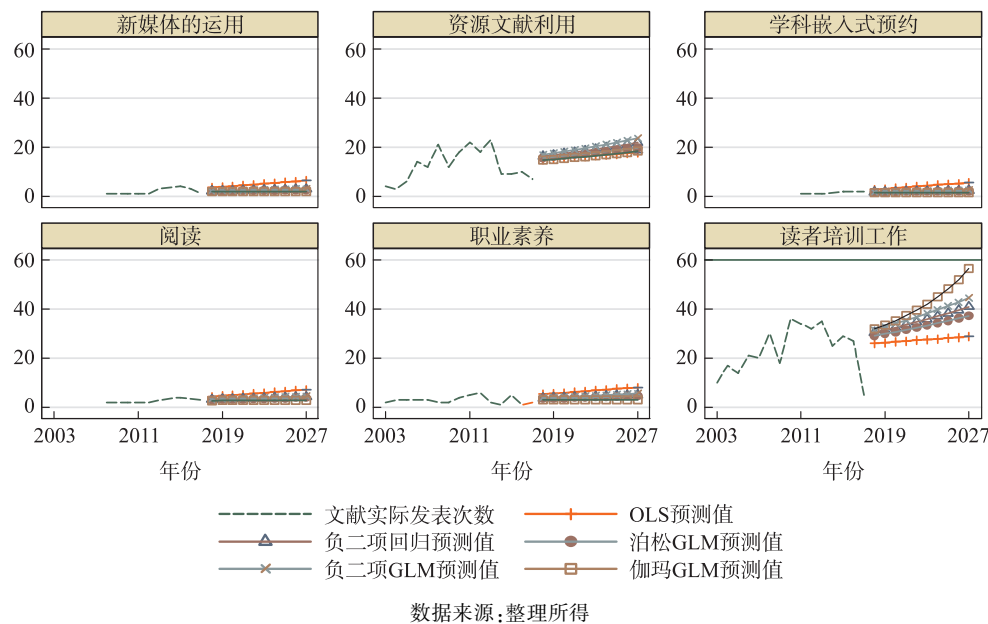


图 1 不同论文内容下的论文发表数预测

Fig. 1 Under different thesis paper presented for prediction

根据文献[12-14]中的方法与工具,可以得到关于不同类别文献发表次数的拟合与预测值.因此,OLS 模型设定如下:

$$cnkinum_{it}=c_{it}+\beta_{1t}year_t+\sum_{j=1}^6\gamma_jclassid_j+\varepsilon_{it},$$

而负二项模型与 GLM 模型设定如下:

$$cnkinum_{it}=\varepsilon_{it}\exp\left(c_{it}+\beta_{1t}year_t+\sum_{j=1}^6\gamma_jclassid_j\right),$$

式中, $cnkinum_{it}$ 是类别 i 在第 t 年的文献数量, $classid_j$ 是第 j 种论文内容的类别变量, $year_t$ 是年度时间趋势项.为此,针对不同论文内容,可以得到 5 种不同回归的论文发表次数预测值,时间为 2019-2028 年, ε_{it} 是残差项.

这些预测值并没有考虑时间滞后项,因为数据太短,在考虑时间滞后的情况下,无法得到较长时间的预测值;并且年度时间趋势项 $year_t$ 已经吸收了时间效应,为了减少样本量的损失,这里不设立时间虚拟变量.

1.3 不同类别图书馆读者培训的拟合

根据类别变量 $classid$,我们可以得到 t 年图书馆读者培训类别 j 的次数 $libnum_{j,t}$.假设读者培训次数受到同一类别文献数量的影响,那么设计图书馆培训课题的相关内容就会受到之前文献的影响,也就是说,读者培训受到的影响是来自于文献发表量的滞后项中.因此,就可以构造简单的回归模型:

$$libnum_{j,t}=\alpha+\sum_{i=0}^n\beta_{t-i}cnkinum_{j,t-i}+\sum_{i=1}^6\gamma_i classid_i+\varepsilon_{jt}+u_{jt}.$$

这个模型中, n 是时间滞后的最大项数, α 是截距项, β_t 与 γ_i 是对应变量的参数. $classid_i$ 是根据 $classid$ 变量构造出来的虚拟变量,用来控制不同类别关联中不随时间变化的影响.在虚拟变量的设定中,我们发现,由于数据过少,实际上只有类别 2 与 5 的虚拟变量可以被构造出来,其他的类别虚拟变量都因为相应类别在样本中的数量太少,所以无法明显地影响结果,故而被程序忽略掉了.因论文内容类别相同的数据中,最少只有 5 个观测值,故时间滞后项不会超过 5 期,以防止出现数据完全缺失.

表 2 PCSE 与 POISSON 模型的参数估计

Table 2 PCSE and POISSON model parameter estimation

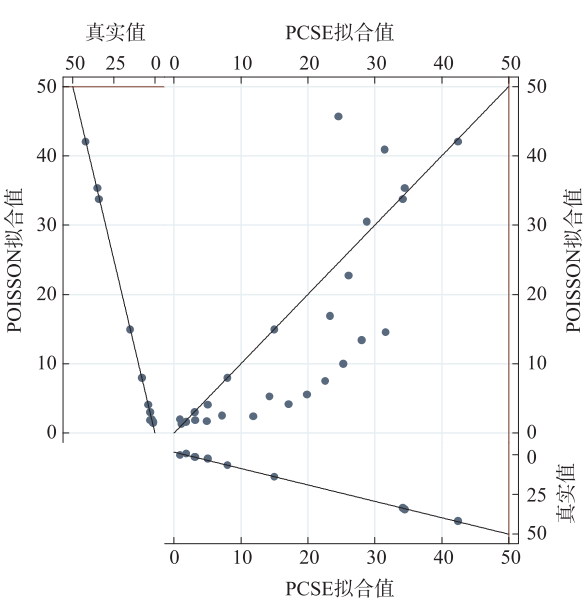
	PCSE	POISSON
cnki _{1,t}	0.341 ** (11.30)	-0.058 ** (-118.26)
cnki _{1,t-1}	1.801 ** (74.28)	-0.055 ** (-36.42)
cnki _{1,t-2}	1.832 ** (66.32)	-0.134 ** (-539.99)
cnki _{1,t-3}	0.927 ** (21.89)	-0.005 ** (-85.75)
cnki _{1,t-4}	1.726 ** (36.55)	-0.026 ** (-131.10)
cnki _{1,t-5}	2.102 ** (41.41)	-0.005 ** (-4.62)
2.classid	-118.300 ** (-56.75)	2.210 ** (422.77)
5.classid	-27.460 ** (-76.62)	
xbols_pcse		0.140 **
_cons	-1.412 ** (-10.73)	1.354 ** (2298.80)
N	11	11

注:括号中的是 t 值 * $p<0.05$, ** $p<0.01$, *** $p<0.001$. 数据来源:整理所得.

由于该数据属于长面板数据,根据其数据特征,同时实验了面板校正标准误差(panel-corrected standard error,即 PCSE)模型、解决组内自相关的 FGLS 模型、全面的 FGLS 模型、OLS 模型、泊松回归、负二项回归等,发现只有 PCSE 模型与泊松回归模型能够保证不同的预测结果稳健显著,并且不同预测值最终的回归结果类似,故本文只展示两种不同回归的结果。

系数表明,2 年前的学术研究与 5 年前的学术研究对读者培训的影响分别达到了一个顶峰,相关内容的文献每增加 1 篇,读者培训增加 1.832 次与 2.102 次。其他年份也对读者培训内容产生明显的影响,但效果随着时间的提前,出现先增后减再增的现象。利用这个模型结果,就可以比较估计值与真实值之间的差异。如图 2 所示,PCSE 的拟合值与泊松回归的拟合值存在一个系统的偏差,这意味着需要在 PCSE 拟合值与泊松回归的拟合值之间进行挑选。

如表 3 所示,可知从具体的估算数值来看,泊松回归的预测结果中,异常值更多——不论是从标准差、最大值还是从均值来看,都太大了。所以本文后续内容均采用 PCSE 拟合值。而表 3 中,PCSE 拟合值的差别并不大,为了保证数据预测结果的稳健性,根据 Stock & Watson^[15] 的建议,本文采用 OLS 估算数据下生成的 PCSE 预测值。



数据来源:整理所得
说明:图中直线为 45°线,反映了横纵坐标轴对应数据相等的地方;并且为了保证数据的可信度,图中删除了小于 0 大于 50 的数据。

图 2 真实值与拟合值的散点图

Fig. 2 Scatter diagram of real value and fitted value

表 3 不同文献发表次数估计值情况下的图书馆培训预测值基本统计量

Table 3 Basic statistics of library training prediction values under different estimated values of literature publication times					
变量名	样本量	样本量	标准差	最小值	最大值
PCSE 估计值					
OLS	77	71.831	93.854	-46.157	281.692
负二项	77	72.981	106.205	-46.157	323.492
泊松 GLM	77	69.942	103.482	-46.157	298.612
负二项 GLM	77	75.716	108.876	-46.157	346.345
伽玛 GLM	77	72.692	116.323	-46.157	397.001
泊松回归					
OLS	77	1.59e+12	9.15e+12	.0011124	7.57e+13
负二项	77	7.07e+13	4.30e+14	.0011124	3.59e+15
泊松 GLM	77	7.62e+12	3.48e+13	.0011124	2.72e+14
负二项 GLM	77	6.69e+14	4.49e+15	.0011124	3.83e+16
负二项 GLM	77	6.81e+16	5.55e+17	.0011124	4.87e+18

2 数据分析

为了预测数据,首先需要预测未来几年的文献发表数量。根据现有数据,按照类似的回归模型,易得不同类型文献在未来几年的发表数量。采取聚类标准误稳健异方差 OLS 预测了发表文献数据之后,再预测出来 2019–2028 年的不同类型图书馆培训的相应结果。根据上一节中得到的结果,可以得到如下预测数据(见表 4)。

由于数据缺失,并不能得到 2023 年之后所有结果的预测值。但当所有图书馆培训的数据预测都来源于文献数量预测值的情况下,就可以得出完整的预测结果。在预测结果中,几乎所有类别的读者培训数量都是上升的,这说明培训总量呈现增长趋势。其中,类别 2 的预测值在 2024 年之前为负数,这是因为真实数据太少,导致了回归截距项小于 0,但这种预测误差并不影响其后的趋势分析。为了保证分析的有效性,

需要将存在负值的年份 2019–2023 年删除,然后考虑其平均变化.

此外为了保证预测值与实际数据的可比性,我们去除没有实际数据的“读者培训工作”类别及只有两年数据的“学科嵌入式预约”类别.在此基础上,我们考察了新媒体利用、资源文献利用、阅读及职业素养的实际数据与预测值的份额变化.

表 4 不同类型图书馆培训次数的预测值
Table 4 Predicted values of training times for different types of libraries

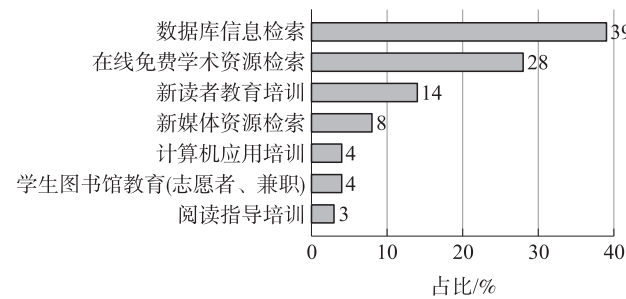
年份	classid = 1	classid = 2	classid = 3	classid = 4	classid = 5	classid = 6
2019		-11.440			-11.046	209.576
2020		-30.871			-2.001	191.278
2021	27.979	-16.539			7.216	215.043
2022	26.028	-10.926			4.869	193.603
2023	27.752	-1.837	24.563	37.504	14.265	185.189
2024	35.179	17.121	28.209	42.485	23.353	231.477
2025	37.906	19.847	30.936	45.212	26.080	234.204
2026	40.633	22.574	33.663	47.938	28.807	236.931
2027	43.359	25.301	36.389	50.665	31.534	239.658
2028	46.086	28.028	39.116	53.392	34.261	242.385

发现短期中新媒体数量有所减少,但它的份额会不断增加.资源文献利用的份额不断减少.阅读的比例也会逐年上升,而职业素养培训的份额几乎保持不变.这种份额变化会直到 2024 年之后变得稳定下来,在这 4 种可比较的类别中,新媒体利用的份额接近 30%,资源文献利用的份额降低到低于 20%,阅读份额足足增加到 40%,职业素养的份额则保持在 10%的水平上基本不变.

3 关于培训方式的预测

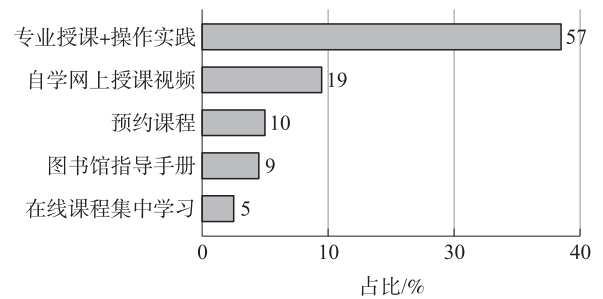
培训效果通常取决于培训内容与培训方式,在互联网越来越发达的今天,培训者通常采用集中讲座交流或借助新媒体多种方式进行培训.针对现有培训种类,根据南京师范大学读者进行问卷单项形式的选择统计,在近 60%以上的培训栏目中,读者勾选的是“碎片化”形式数据库信息检索技巧培训及网上学术资源的免费查找等(见图 3).

在最受欢迎的培训形式中,“专业授课+操作实践”是首选,超过 50%;“自学网上授课视频”紧随其后,接近 20%;剩下的培训形式并没有超过 10%(见图 4).



数据来源:整理所得
图 3 不同培训类型的占比

Fig. 3 Proportion of different training types



数据来源:整理所得
图 4 不同培训形式的占比

Fig. 4 Proportion of different training forms

可见短时间、有效性、简便性更成为年轻一代的选择.通过以上的文献成果调查和对比分析可以看出,高校图书馆在读者培训讲座上呈现出专业化、定制化和虚拟化的特征,碎片化的新媒体趋势也越来越成为现阶段及未来一段时间培训的主要手段.因此,高校图书馆在开展读者培训设置时,更应该积极拓展线上信息素养教育、完善学科服务平台建设,并加强培训的宣传力度及培训师资梯队建设.

4 结论

从文献成果与读者调研的数据分析中可以得到一个有趣的现象:理论性是早于读者教育培训实践发生的,特别通过文献成果的统计,可以有效预测未来几年高校图书馆读者教育培训的发展趋势.利用有限

的图书馆员从事更加专业化、多元化和复杂化的工作和服务,有效促进图书馆与馆员在新媒体环境中规划地结合读者教育培训工作,切实提高高校图书馆教育培训服务效果。

(1)加强培训讲座师资队伍建设和,要求培训者具备一定的学科背景的同时,更需要强烈的信息意识,具备扎实的图书馆信息基础知识和网络化的信息服务技能。图书馆读者培训也要考虑内容新颖性、培训方式的多样化,了解读者的信息需求、增强与读者之间的互动,尝试与学生社团合作,开发更多的有效培训课程。

(2)在读者教育培训设计开发和培训施教的过程中,要积极做好课题调研与主题宣传培训目标,使广大读者在整个培训中能有效的参与,达到设计和选择满意的读者教育培训形式和功能。在培训过程中,要注意营造教育培训的良好氛围,以此获得读者的尊重与信任。

(3)在高校图书馆读者教育培训中需要可持续性发展原则,在与具体软件开发商签订有效的合同时,要注意新媒体功能的开发、维护和升级,保障其稳定地为图书馆读者工作。新媒体功能模块的使用,更多的读者趋向选择“碎片化”新媒体培训形式,所以高校图书馆读者教育培训关于新媒体应用与实操将成为图书馆教育培训的重要发展课题。

[参考文献]

- [1] MCGEHEE W, THANYER P. Training in business and industry[M]. New York: Wiley, 1961.
- [2] SPARROWE R T, LIDEN R C, WAYNE S J, et al. Social networks and the performance of individuals and groups[J]. Academy of management journal, 2001, 44(2): 316-325.
- [3] FREEMAN G T. The library as place: changes in learning patterns, collections, technology, and use in library as place: rethinking roles, rethinking space[C]// Washington, D.C: Council on Library and Information Resources, 2005.
- [4] 孙瑾. 国外 Information Commons 的发展现状及相关研究[J]. 图书馆杂志, 2006(4): 57-60.
- [5] 陈维. 数字图书馆特色资源共享与服务研究[M]. 杭州: 浙江工商大学出版社, 2015.
- [6] DUNAWAY M. Web2.0 and critical information Literacy[J]. Public services quarterly, 2011(3-4): 149-157.
- [7] MCCULLAGH P, NELDER J A. Generalized linear models[M]. 2nd ed. London: Chapman & Hall/CRC, 1989.
- [8] HOFFMANN J P. Generalized linear models: an applied approach[M]. Boston: Pearson, 2004.
- [9] NELDER J A, WEDDERBURN R W M. Generalized linear models[J]. Journal of the royal statistical society, 1972, 135: 370-384.
- [10] WEDDERBURN R W M. Quasi-likelihood functions, generalized linear models, and the gauss-newton method[J]. Biometrika, 1974(61): 439-447.
- [11] MCCULLAGH P. Quasi-likelihood functions.[J]. The annals of statistics, 1983(11): 59-67.
- [12] LONG J S, FREESE J. Regression models for categorical dependent variables using stata[M]. College Station, Texas: Stata Press, 2006.
- [13] CAMERON A C, TRIVEDI P K. Regression analysis of count data[M]. Cambridge: Cambridge University Press, 2013.
- [14] DENG S, YANG Y Z, ZHU Q X. Analysis and estimation of learning style in library[J]. Journal of hunan university of arts and science(science and technology), 2017, 29(4): 18-22, 44.
- [15] STOCK J, WATSON M. Introduction to econometrics[M]. 3rd Edition. Boston: Addison-Wesley Press, 2011.

[责任编辑: 陆炳新]