

基于百度指数的宗教统计数据测量新方法及应用研究

旷 芸¹, 梁宗经²

¹桂林师范高等专科学校图书馆, 广西 桂林

²广西师范大学经济管理学院, 广西 桂林

Email: zjliang@mailbox.gxnu.edu.cn

收稿日期: 2021年7月26日; 录用日期: 2021年8月24日; 发布日期: 2021年8月31日

摘 要

目的: 研究百度指数与调查问卷数据的相关性。本研究针对传统宗教数据统计存在的问题, 提出应用百度指数可作为宗教统计数据的替代变量, 提出一种新的宗教数据统计方法。方法: 应用相关系数法和回归分析法, 研究对象为互联网佛教、基督教、民间信仰和伊斯兰教信息以及相对应的调查问卷数据, 研究数据来源于2013年度的网络搜索指数(百度指数)及全国社会综合调查数据, 结论: 百度指数可作为宗教统计数据的替代变量, 基于百度指数的回归模型性能优于基于调查问卷模型, 通过建立以互联网变量、宏观经济变量和微观心理变量作为验证参数的回归模型所得结论, 证明百度指数可以有效地应用于实时宗教问题研究。本研究结论可为宗教社会学研究提供新的宗教数据源、为宗教社会管理提供实证参考资料。

关键词

宗教统计, 百度指数, 回归分析, 相关性

Research on the New Method and Application of Religious Statistics Data Measurement Based on Baidu Index

Yun Kuang¹, Zongjing Liang²

¹Guilin Teachers College Library, Guilin Guangxi

²School of Economics and Management, Guangxi Normal University, Guilin Guangxi

Email: zjliang@mailbox.gxnu.edu.cn

Received: Jul. 26th, 2021; accepted: Aug. 24th, 2021; published: Aug. 31st, 2021

Abstract

Purpose: To study the correlation between Baidu index and questionnaire data. This research proposes that the Baidu index can be used as a substitute variable for religious statistics, and a new method of religious data statistics is proposed. **Methods:** Using correlation coefficient method and regression analysis method. The research objects are Internet information of the Buddhism, Christianity, folk beliefs and Islam and the corresponding questionnaire data. The research data comes from the 2013 Internet Search Index (Baidu Index) and the national society Comprehensive survey data. **Conclusion:** Baidu index can be used as a substitute variable for religious statistics and the performance of the addition model based on Baidu index is better than the model based on questionnaire. According to the establishment of a regression with Internet variables, macroeconomic variables and micropsychological variables as verification parameters, the conclusion of the model proves that the Baidu Index can be effectively applied to real-time religious research. The conclusions of this research can provide new religious data sources for religious sociology research and empirical reference materials for religious social management.

Keywords

Religious Statistics, Baidu Index, Regression Analysis, Correlation

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

互联网宗教活动关系社会的和谐、稳定乃至国家安全。一方面,随着国家“一带一路”战略的实施,我国迎来了新的重大发展机遇,社会的和谐、稳定和发展,关系到国家的核心利益。另一方面,随着互联网的发展,网上宗教活动呈现出不断增长的趋势,并已引起社会研究学者及政府管理部门的高度关注[1]。此外,与我国接壤及“一带一路”沿线的国家大都具有悠久的宗教传统,其宗教活动势必影响我国。国内外互联网宗教活动的兴起和活跃,已经并将越来越深刻地影响我国社会秩序,乃至国家安全,成为政府关心、民众关切的重要问题,成为互联网对社会的变革性影响的重要方面,亟待系统深入研究。

宗教是现代文明的重要组成部分,它构成了现代文明的伦理基础[2]。随着当代科学技术的快速发展,互联网正被用作宗教信息传播的常用手段[3]。互联网宗教活动是互联网与传统宗教活动相结合的产物,它随着互联网的广泛应用而诞生。互联网宗教活动是指以互联网作为工具,在虚拟空间中进行宗教信息的传播及从事宗教活动的总称。互联网宗教活动近年来发展迅速,根据搜索引擎查询可知,截止到2015年12月,仅仅涉及世界第一大宗教基督教的英文网页就达2700万页(Google搜索,2015)、中文网页560万页(百度搜索,2015),并且呈逐年增长趋势。互联网宗教活动已成为网络环境下宗教社会学研究的新领域。

为了更好地了解及管理互联网宗教活动,实时掌握互联网宗教定量数据成为了当前宗教管理、宗教研究的极为迫切需要解决的问题。传统宗教人数统计方法是采用实地调查法及问卷调查法,因宗教信仰涉及个人隐私,传统统计方法所采集数据具有客观性不足的缺点,与此同时,问卷调查法还存在样本数量少、时间滞后等缺陷。为了解决宗教传统统计方法的不足,本研究拟以网络搜索指数(百度指数)作为宗

教数据替代变量, 从而解决互联网时代宗教数据的采集问题。具体研究过程是首先进行文献综述, 然后提出网络宗教数据采集方法, 并将之与统计数据相比较, 然后将实测数据和百度指数的有效性进行检验, 最后提出政策建议。

2. 研究现状

2.1. 国内研究现状

国内研究主要集中在两个方面。① **定性研究**: 在经济贸易交往中建立全方位反宗教渗透网络研究[4], 研究互联网宗教对国际关系的影响[5], 研究宗教在大学生中传播的特点[6], 作为宗教信息的载体, 宗教网站分析呈现出其必要性[7], 现有研究表明中青年已成为网上宗教活动的主要成员[8], 对网络宗教活动的形态分析[9]等。② **宗教网站研究**: 研究发现各类宗教网站数量剧增, 互联网宗教正日益影响人类社会[10]。剖析国内伊斯兰教网站特点[11]。

2.2. 国外研究现状

国外研究可以分为定性研究与定量研究两类。① **定性研究**: 从社会学角度分析亚洲国家复杂宗教状况[12], 认为互联网宗教研究是一个重要的研究内容[13], 网络宗教研究方法跟不上互联网技术快速发展的步伐[14], 提出互联网宗教是西方文化传播的载体[15], 网络宗教与传统宗教无论在仪式、符号和教条均存在差异[16], 通过研究网络宗教可以为研究整个社会变迁提供一个微观途径[17]。② **定量研究**: 提出互联网宗教四种研究方法[18], 互联网宗教扩散情况研究[19], 以实证方法研究网络宗教论坛的作用及其价值所在[20], 采用地理数据可视化技术展现网络宗教空间分布[21], 宗教现象数量分析[22], 宗教与经济的相关性研究[23] [24], 搜索引擎方法能够有效研究网络社区用户的搜索习惯[3], 应用谷歌指数研究宗教问题[2]等。

2.3. 宗教研究方法的演变

传统宗教活动的社会调查方法主要是问卷调查或实地调查, 此方法存在工作量大、样本数少、准确性差、数据滞后等缺点, 比如, 实地调查数据从采集到汇总再到部门统计, 往往需要数周甚至一年的时间。互联网时代宗教活动的发展, 对传统宗教调查方法提出了挑战, 互联网宗教活动具有变化大、实时性强的特点, 因此目前研究互联网宗教的最大困难是现有研究方法不能与快速发展和变化的互联网宗教活动保持同步[14]。相对传统社会学数据采集方法, 搜索指数具有成本低、真实性好、实时性强、可免费获取的特点, 将之应用于数理社会学研究将具有潜在优势[25]。基于此, 本文研究拟采用搜索指数分析法研究网络宗教活动, 该方法能够有效克服传统统计数据存在的时间滞后问题, 能实现宗教活动的实时跟踪研究。搜索指数分析法是基于搜索引擎指数的一种定量分析方法, 其原理是将搜索指数与实际数据进行比较, 通过建立数理模型分析它们之间的相关性。但现有的中外宗教研究学者少有将此方法应用于宗教活动分析。

2.4. 研究动态

国内研究主要集中于互联网宗教活动的定性分析, 少有定量研究成果。国外互联网宗教活动研究对象主要是西方发达国家, 少有针对发展中国家的研究成果。

基于网络搜索数据的搜索指数分析法已在多个领域得到成功应用, 并表现出其可靠性及可行性, 但未见有将此方法应用于研究网络宗教活动的成果发表。

针对当前国内外研究的不足, 本研究创新地将百度指数作为宗教数据的替代变量, 进行相应研究:

1) 原始数据采集。以佛教、基督教、民间信仰和伊斯兰教作为关键词, 提取对应的百度指数。与此同时, 通过中国社会综合调查数据库, 计算出宗教信仰人数。2) 将实测宗教信仰人数与百度指数进行相关分析, 分析它们之间的相关特性。3) 为验证百度指数在宗教研究中的有效性, 通过构建回归分析模型, 研究宗教与网络变量、宏观经济变量、微观个人心理变量之间的相关性。最后提出政策建议。

3. 数据来源及研究方法

3.1. 百度指数

百度指数为当前全球最大中文搜索引擎百度公司发布的网络搜索指数。百度指数是以互联网中文用户海量搜索数据作为基础, 通过数据变换计算搜索关键词的搜索热度指数, 其值越大, 则表明相应关键词的网络关注度越高。通过百度指数可以实时了解相应关键词在全国各地的关注程度, 该指数的发布为研究宗教传播提供了客观、实时的定量数据。数据提取方法是在百度指数网站中输入相应关键词, 然后系统就会绘出相应曲线。因百度指数仅提供时间序列数据, 不提供原始数据下载, 而为了进行相应的数据分析, 则必须将图形数据转换为时间序列数据, 本研究的方法是采用图形-数字软件将百度指数曲线转换为时间序列数据。

3.2. 实测数据

为了检验及比较宗教百度指数与实际宗教信仰人数之间是否存在相关性, 本研究下载中国综合社会调查(Chinese General Social Survey, CGSS)数据库, 根据调查数据中的各宗教信仰人数比例乘以人口总数作为实际信教人数, 将之与百度指数进行比较。通过 CGSS 数据库采集各宗教信仰人数比例及微观心理数据, 它们分别为参加宗教活动频率、感觉身体很健康、阶级认同和生活非常幸福比例。

宏观经济数据源于国家统计局及国家互联网信息中心, 采集的数据为宏观经济数据, 它们分别为大专以上学历人口、人均 GDP、一般公共服务、人均可支配收入、平均受教育年限、城镇化率、网民数、互联网普及率等。

3.3. 相关分析及回归分析

为了比较宗教实际信教人数与宗教百度指数之间的关系, 本研究拟采用传统的相关分析与回归分析法构建模型, 定量研究变量间的相关特性。

4. 实证研究

本研究所采集原始数据包括实测宗教信仰人数、各宗教的百度指数。使用的数据时间范围为 2013 年, 地域范围为全国 28 个省(根据数据可得性选择): 安徽省、北京市、福建省、甘肃省、广东省、广西壮族自治区、贵州省、河北省、河南省、黑龙江省、湖北省、湖南省、吉林省、江苏省、江西省、辽宁省、内蒙古自治区、宁夏回族自治区、青海、山东省、山西省陕西省、上海市、四川省、天津市、云南省、浙江省和重庆市。

4.1. 原始数据

本研究原始数据来源于 2013 年中国社会综合调查数据库、百度指数网站、国家统计局及中国互联网信息中心。具体指标变量及相应数据来源如表 1 所示。

通过数据采集, 可得本研究所用的原始数据如表 2 所示, 其中 N 为省份数量, 各宗教人数单位为万人。

Table 1. Original data sources
表 1. 原始数据来源

指标变量	数据来源	
佛教人数比例	2013 年中国社会综合调查(CGSS): http://cgss.ruc.edu.cn/	
基督教人数比例		
民间信仰人数比例		
伊斯兰教人数比例		
参加宗教活动频率		
感觉身体很健康		
阶级认同		
生活非常幸福比例		
佛教百度指数		百度指数网站: https://index.baidu.com/v2/index.html#/
基督教百度指数		
民间信仰百度指数		
伊斯兰教百度指数		
大专以上人口	国家统计局: http://www.stats.gov.cn/ , 中国互联网信息中心: http://www.cnnic.net.cn/	
人均 GDP		
一般公共服务		
人均可支配收入		
平均受教育年限		
城镇化率		
网民数(万人)		
互联网普及率		

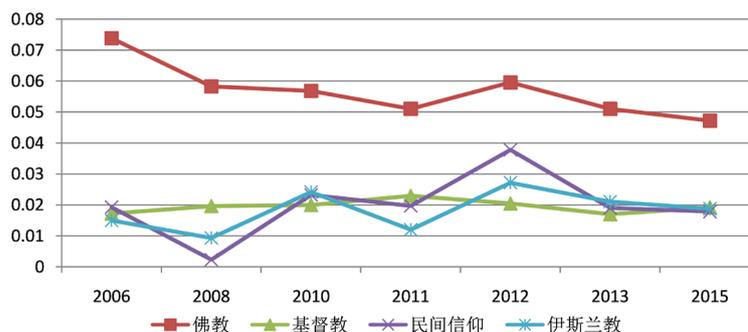
Table 2. Original data table
表 2. 原始数据表

	N	最小值	最大值	平均值	标准误差
佛教人数	28	0.06	53.97	9.3032	11.22515
基督教人数	28	0.00	51.03	3.8043	9.57846
民间信仰人数	28	0.00	51.18	4.1443	10.90660
伊斯兰教人数	28	0.00	28.08	2.7168	6.30327
佛教百度指数	28	101	520	299.71	129.253
基督教百度指数	28	80	600	244.14	122.658
民间信仰百度指数	28	27	168	97.5	31.881
伊斯兰教百度指数	28	41	191	106.46	32.442
网民数(万人)	28	274	6992	2147.79	1406.309
互联网普及率	28	32.60%	75.20%	46.71%	12.18%
参加宗教活动频率	28	9	133	57.5	35.199
感觉身体很健康	28	0.11	0.503333	0.251137	0.099398

Continued

阶级认同	28	0.085	0.42	0.2	0.097
生活非常幸福比例	28	0.333333	0.88	0.731642	0.110905
大专以上人口	28	553	8462	4110.82	2188.575
人均 GDP	28	22921.67	99607	48567.04	21265.89
一般公共服务	28	64.1499	996.4544	432.8705	222.3348
人均可支配收入	28	18964.8	43851.4	26146.06	6388.696
平均受教育年限	28	7.84	12.03	9.1764	0.84469
城镇化率	28	0.38	0.9	0.5614	0.13156

为了直观展示宗教信仰人比例历年来的变化情况, 经过对原始数据进行处理, 得出相应变化图, 如图 1 所示。



注: 数据来源于全国社会综合调查历次统计数据并经数据转换后得出。

Figure 1. The proportion of the number of religious Believers in China from 2006 to 2015

图 1. 2006 年至 2015 年中国宗教信仰人数比例图

从图 1 中可以看出, 2006 年至 2015 年中国宗教信仰人数比例图包括佛教、基督教、民间信仰和伊斯兰教, 横坐标表示为年限, 纵坐标为全国人口百分比, 其中信仰佛教的人数比例最高, 并且呈现出逐年下降的趋势, 从 2006 年 7.4% 降至 2015 年 4.7%。而基督教、民间信仰和伊斯兰教的信教人数则呈现出波动状态, 并且显示出在 2% 水平振荡变动。

4.2. 相关性分析

为了检验问卷实地调查信教人数与百度指数之间的相关性, 本研究拟应用相关系数分析法进行检验。首先绘出信教人数与百度指数散点图, 然后绘出拟合线性曲线, 最后得出相应的拟合回归方程, 从而了解两者之间的数量关系(原始数据经过标准化处理)。

佛教人数与百度指数拟合方程: $y = 0.41x - 1.48E-7$ 。从拟合方程可见, 佛教信教人数与百度指数之间呈现出正相关状况, 百度指数变量系数为 0.41, 即表示百度指数每增加一个百分单位, 佛教人数比例就增加 0.41 个计量单位。

基督教人数与百度指数拟合方程: $y = 0.69x - 1.45E-6$ 。从拟合方程可见, 基督教信教人数与百度指数之间呈现出正相关状况, 百度指数变量系数为 0.69, 即表示百度指数每增加一个百分单位, 基督教人数比例就增加 0.69 个计量单位。

民间信仰人数与百度指数拟合方程： $y = 0.45x - 1.95E-7$ 。从拟合方程可见，民间信仰信教人数与百度指数之间呈现出正相关状况，百度指数变量系数为 0.45，即表示百度指数每增加一个百分单位，民间信仰人数比例就增加 0.45 个计量单位。

伊斯兰教信教人数与百度指数拟合方程： $y = 0.31x + 8.26E-7$ 。伊斯兰教信教人数与百度指数之间呈现出正相关状况，百度指数变量系数为 0.31，即表示百度指数每增加一个百分单位，伊斯兰教人数比例就增加 0.31 个计量单位。

总结信教人数与百度指数之间的拟合方程，可以看出它们之间存在的相同点是信教人数与百度指数均存在正相关关系，不同点是百度指数对信教人数的影响程度及拟合精度则存在差别，其中百度指数对信教人数影响最大的是基督教，其变量系数为 0.69。

4.3. 回归分析

前述部分研究了宗教实测人数与网络百度指数之间的相关性，研究结果表明双方具有线性正相关特性，证明百度指数可以作为宗教信仰人数的替代变量，即通过百度指数可以预测宗教信仰人数比例，但没有说明百度指数及信教人数受什么因素影响。基于此，下面将应用线性回归分析法研究影响百度指数及信教人数的影响因素，为了便于全面了解影响宗教信仰及百度指数的影响因素，拟将影响因素划分为网络因素、宏观经济因素及微观心理因素三类，分别构建以百度指数和信教人数为因变量、以网络因素、宏观经济因素及微观心理因素为自变量的线性回归方程，实证分析各自变量对因变量的影响程度。即通过利用回归分析证明百度指数是否更好地预测社会、经济招标[2]。

为了便于比较模型的有效性或预测精度，本研究拟以回归分析模型中的 R^2 、AIC 信息准则即 Akaike information criterion、施瓦兹准则 SC (Schwarz Criterion) 作为判断模型的拟合精度及模型有效性。根据统计学原理可知，拟合参数 R^2 的取值范围为 0~1 之间，其值越大，则相应模型的拟合度越好。AIC 参数和 SC 参数则是数值越小越好，其值越小则表明相应模型的预测精度越好、模型自变量对因变量的解释力越强。

4.3.1. 互联网影响变量

根据网络信息计量学及经济学原理可知，影响网络搜索指数的互联网影响因素包括网民数及互联网普及率，因此，为了了解影响百度指数的影响因素，建立以百度指数作为因变量、网民数及网络普及率为自变量的回归模型，并比较模型的拟合精度及变量解释力。计算结果如表 3 所示。

Table 3. Regression results of Baidu index, the number of believers and network variables

表 3. 百度指数、信教人数与网络变量的回归结果

因变量	自变量	标准化系数	拟合度 R^2	AIC	SC
佛教百度指数	(常量)	1.07E-16			
	网民数	0.682*	0.72	1.75	1.9
	网络普及率	0.356*			
佛教人数	(常量)	-8.41E-07			
	网民数	0.694123**	0.51	2.30	2.44
	网络普及率	0.068105			

Continued

	(常量)	-7.08E-18			
基督教百度指数	网民数	0.67**	0.56	2.19	2.34
	网络普及率	0.203			
	(常量)	-7.48E-07			
基督教人数	网民数	0.338588	0.13	2.88	3.02
	网络普及率	-0.230663			
	(常量)	-2.71E-16			
民间信仰百度指数	网民数	0.759**	0.74	1.68	1.82
	网络普及率	0.252*			
	(常量)	-1.30E-06			
民间信仰人数	网民数	0.611594*	0.48	2.36	2.51
	网络普及率	0.203283			
	(常量)	-1.91E-16			
伊斯兰教百度指数	网民数	0.231	0.21	2.77	2.91
	网络普及率	0.346*			
	(常量)	7.07E-07			
伊斯兰教人数	网民数	0.267048	0.08	2.93	3.07
	网络普及率	-0.1949			

注: **. 在 0.01 级别(双尾)相关性显著。*. 在 0.05 级别(双尾)相关性显著。

从计算结果表 3 可以得出以下结论:

首先对于佛教模型,因变量为佛教百度指数的网民数及网络普及率标准化系数分别为 0.682 和 0.356, 统计学意义上显著, 相应的拟合度指数 R^2 为 0.72, 而对应的佛教人数比例相应系数分别为 0.69、0.07 和 0.51, 拟合度指标比百度指数的小, 因此可以认为基于百度指数的回归模型相对佛教人数比例的回归模型拟合程度更优。而相应的模型优劣判断参数, 相应参数百度指数模型为 1.75、1.9, 信教人数模型为 2.3、2.44, 基于百度指数的相应参数均小于信教人数模型, 因此可以认为因变量为百度指数的网民数及网络普及率变量可以更好地解释因变量的变化原因。

其次对于基督教模型, 因变量为基督教百度指数的网民数及网络普及率标准化系数分别为 0.67 和 0.203, 其中 0.67 统计学意义上显著, 相应的拟合度指数 R^2 为 0.56, 而对应的基督教人数比例相应系数分别为 0.34、-0.23 和 0.13, 拟合度指标比百度指数的小, 因此可以认为基于百度指数的回归模型相对基督教人数比例的回归模型拟合程度更优。而相应的模型优劣判断参数, 相应参数百度指数模型为 2.19、2.34, 信教人数模型为 2.88、3.02, 基于百度指数的相应参数均小于信教人数模型, 因此可以认为因变量为百度指数的模型更优。

再次对于民间信仰模型, 因变量为民间信仰百度指数的网民数及网络普及率标准化系数分别为 0.759 和 0.252, 统计学意义上显著, 相应的拟合度指数 R^2 为 0.74, 而对应的民间信仰人数比例相应系数分别为 0.61、0.20 和 0.47, 拟合度指标比百度指数的小, 因此可以认为基于百度指数的回归模型相对民间信仰人数比例的回归模型拟合程度更优。而相应的模型优劣判断参数, 相应参数百度指数模型参数为 1.68、1.82, 信教人数模型为 2.36、2.51, 基于百度指数的相应参数均小于信教人数模型, 因此可以认为百度指

数模型更优。

最后对于伊斯兰教模型, 因变量为伊斯兰教百度指数的网民数及网络普及率标准化系数分别为 0.231 和 0.346, 其中 0.346 统计学意义上显著, 相应的拟合度指数 R^2 为 0.21, 而对应的伊斯兰教人数比例相应系数分别为 0.26、-0.19 和 0.08, 拟合度指标比百度指数的小, 因此可以认为基于百度指数的回归模型相对伊斯兰教人数比例的回归模型拟合程度更优。而相应的模型优劣判断参数, 相应参数百度指数模型为 2.77、2.91, 信教人数模型为 2.93、3.07, 基于百度指数的相应参数均小于信教人数模型, 因此可以认为百度指数模型的变量解释力更好、其建立的模型更优。

综合而言, 对于互联网影响因素, 实证结论表明, 基于百度指数的回归分析模型比较基于信教人数的回归分析模型的拟合程度更高, 而影响因素网民数及网络普及率对百度指数因类变量的模型解释力更好。

4.3.2. 宏观经济影响变量

根据经济学和社会学原理可知, 影响民众信仰宗教的因素很多, 其中包括个人原因及社会原因, 而涉及宏观经济变量的分别为大专以上学历人口、人均 GDP、一般公共服务、人均可支配收入、平均受教育年限、城镇化率。回归分析结果如表 4 所示。

Table 4. Baidu index, the number of believers and the regression results of macroeconomic variables
表 4. 百度指数、信教人数与宏观经济变量回归结果

因变量	大专以上学历人口	人均 GDP	一般公共服务	人均可支配收入	平均受教育年限	城镇化率	拟合度 R^2	AIC	SC
佛教百度指数	0.296426	0.278708	0.406082*	0.396846	0.431418	-0.632198	0.78	1.78	2.12
佛教人数	-0.58389	-0.081985	1.096853**	0.747331**	-0.1548	-0.02661	0.68	2.14	2.47
基督教百度指数	0.221021	0.203437	0.470756	0.455057	0.285003	-0.563686	0.66	2.22	2.56
基督教人数	-3.96E-07	-0.014095	-0.14812	0.399464	0.031782	0.302949	0.17	3.11	3.45
民间信仰百度指数	0.001728	0.207522	0.704455***	0.383749	0.377258	-0.44777	0.7	2.08	2.42
民间信仰人数	1.26E-06	-0.81504	0.380555	1.136965	0.317653	-0.46484	0.58	2.42	2.76
伊斯兰教百度指数	1.06504**	-0.273029	-0.524387	-0.256153	-0.67386	1.204483**	0.48	2.66	3
伊斯兰教人数	9.66E-07	0.127528	0.142412	0.053186	0.427164	-0.15058	0.13	3.16	3.49

注: **. 在 0.01 级别(双尾)相关性显著。*. 在 0.05 级别(双尾)相关性显著。

根据实证结果表 4 可得出如下结论:

1) 宏观经济变量对因变量的影响程度分析

影响因素大专以上学历人口变量, 在所建立的 8 个模型中, 因变量为百度指数的模型系数均为正值, 说明文化程度越高, 对因变量百度指数的影响越大, 这与实际情况相符, 因为根据国家互联网信息中心发布的季度统计报告可知, 网民的占比与所接受的文化程度成正比, 接受文化教育程度越高, 越关注网络信息。对于变量人均 GDP, 人均 GDP 为衡量地区经济发达程度的核心指标, 人均 GDP 越高, 说明相应地区的生活水平越高, 实证结果该变量系数如果为正, 说明生活水平越高的地区, 民众越关注网络宗教信息, 与此同时, 也说明生活水平越高的地区, 实际信教人数也越多, 这说明随着人们生活的不断提高, 人们对于精神层面的需求也越高。对于宏观变量: 一般公共服务、人均可支配收入、平均受教育年限, 它们的系数大部分情况下都为正值, 说明这些宏观经济变量对因变量的作用为正向的促进作用。而对于

城镇化率变量, 其系数均为负值, 说明该变量对因变量的影响响度为负向影响作用, 从中也说明城镇化程度越高, 即城市人口越多, 信教人数与对宗教信息的关注度越低, 间接说明信教人数存在城乡差别, 相对而言农村的信教人数更高一些。

在宏观经济变量影响模型中, 因变量为佛教、民间信仰和基督教的模型各系数中, 系数值均为正值, 说明模型中的宏观经济变量对百度指数因变量、信教人数因变量的影响均为正向影响, 各变量数值越高, 相应因变量的数值也越高。但对于因变量为伊斯兰教的模型中, 各变量系数值多数情况下为负, 说明相应自变量对因变量的影响程度为负相关关系, 其内部原因很复杂, 可能与该宗教的性质存在关联。

2) 回归计算结论比较

模型拟合度比较。比较因变量为百度指数和信教人数的回归模型拟合程度 R^2 , 因变量为佛教百度指数、基督教百度指数、民间信仰百度指数、伊斯兰教百度指数的 R^2 值分别为 0.78、0.66、0.7、0.48, 相对应的信教人数模型的 R^2 值分别为 0.68、0.17、0.58、0.13, 百度指数模型相应参数高于相比信教人数模型, 这说明因变量为百度指数的回归模型拟合程度相比因变量为信教人数的回归模型要高, 即百度指数模型更符合实际情况。

变量解释力比较。比较因变量为百度指数和信教人数的回归模型优劣度指标 AIC 和 SC, 因变量为佛教百度指数、基督教百度指数、民间信仰百度指数、伊斯兰教百度指数的 AIC 和 SC 值, 从表 5 中可以看出, 百度指数模型相应参数相低于信教人数模型, 这说明因变量为百度指数的回归模型优于信教人数的回归模型, 从而说明相应宏观经济变量更好地了百度指数内在影响原因。

4.3.3. 微观心理影响变量

前述分析分析了宏观经济变量对宗教搜索及信教人数的影响, 接着从微观心理角度研究影响宗教网络搜索及信教人数的内在影响原因。结合网络信息理论及社会心理学理论, 本研究采取以下参数: 参加宗教活动频率、感觉身体很健康、生活非常幸福比例、阶级认同中等以上四个参数。回归结果如表 5 所示。

Table 5. Regression results of Baidu index, number of believers and micro psychological variables
表 5. 百度指数、信教人数与微观心理变量回归结果

因变量	参加宗教活动频率	感觉身体很健康	生活非常幸福比例	阶级认同中等以上	拟合度 R^2	AIC	SC
佛教百度指数	-0.086	0.505***	0.545**	0.507**	0.39	2.67	2.91
佛教人数	3.79E-07	0.498208	0.111537	0.548251	0.60	2.24	2.48
基督教百度指数	0.098	0.529***	0.54**	0.475**	0.43	2.59	2.83
基督教人数	-4.67E-07	0.211803	0.235993	0.191132	0.12	3.03	3.27
民间信仰百度指数	0.01	0.557***	0.542**	0.312	0.42	2.62	2.85
民间信仰人数	-2.54E-07	0.48662	0.065017	0.302689	0.39	2.67	2.90
伊斯兰教百度指数	-0.126	0.368*	0.417	0.482*	0.25	2.87	3.11
伊斯兰教人数	7.25E-07	0.005797	0.034357	0.019181	0.01	3.15	3.39

注: ***. 在 0.01 级别(双尾)相关性显著。**. 在 0.05 级别(双尾)相关性显著。*. 在 0.10 级别(双尾)相关性显著。

根据实证结果表 5 可得出如下结论:

1) 微观心理变量对因变量的影响程度分析

回归结果表明, 微观心理变量参加宗教活动频率、感觉身体很健康、生活非常幸福比例、阶级认同中等以上的变量系数大部分为正值, 说明微观心理变量对宗教网络搜索及宗教信仰具有正向影响, 但各变量的影响程度及显著性则存在差异, 其中感觉身体健康变量的系数存在显著性。

2) 回归计算结论比较

模型拟合度比较。从表 5 可见, 8 个回归模型中, 除因变量为佛教百度指数的回归模型拟合度 R^2 比因变量为信教人数模型的拟合度低以外, 其它模型的拟合度参数值, 因变量为百度指数的模型均高于因变量为信教人数的回归模型。

变量解释力比较。通过比较 AIC 和 SC 参数值, 可知, 除了因变量为佛教百度指数的 AIC 和 SC 参数值比因变量为信教人数模型大以外, 其它情况下百度指数模型的参数值小于信教人数参数值。因此可以认为百度指数模型优于信教人数模型。

5. 结论

本研究针对传统宗教统计数据存在的数据客观性差、数据发布滞后、样本数少的问题, 提出以百度指数作为宗教信教人数的替代变量, 解决了互联网时代宗教信仰人数问题。本研究根据研究需要采集了互联网影响变量、宏观经济变量和微观心理变量, 使用相关系数和回归分析法, 构建基于百度指数和基于实测数据的宗教信仰回归模型。经过实证分析, 得出以下结论:

1) 百度指数可以作为实际宗教数据的替代变量。

经过实证研究, 发现佛教、基督教、民间信仰和伊斯兰教的实际信教人数与对应的百度指数之间存在正相关关系, 从而说明通过建立回归模型就能根据百度指数预测实际信教人数, 实际信教人数因存在数据采集困难、宏观性差的缺陷, 而百度指数则可以免费实时提取, 这就为实时研究宗教问题提供了极大的便利, 因此可以认为百度指数可作为宗教数据的替代变量。

2) 百度指数回归模型综合性能优于信教人数回归模型。

为了验证百度指数模型相对问卷调查模型的优劣特性, 本研究选取了互联网变量、宏观经济变量和微观心理变量三类验证参数, 分别构建基于百度指数及调查问卷的回归模型, 研究表明, 无论是模型拟合度还是变量解释力, 百度指数模型均优于调查问卷回归模型。研究结果证明百度指数可应用于宗教问题研究。

本研究不足之处:

1) 百度指数局限性。

因为不是每一位信徒都使用百度指数进行宗教信息搜索, 比如一些年老信徒就不常用网络搜索, 再有就是一些边远山区民众, 因通讯条件所限, 不能正常上网, 这也导致百度指数与实际情况存在差异。

2) 实测数据局限性。

本研究所用的实测数据为全国社会综合调查数据, 该调查数据采用抽样方式, 不是全样本抽样, 因此实测数据与实际数据也存在差距。

未来研究方向。尽管本研究方法存在一定的局限性, 并且所构建的回归方程拟合程度不高, 但本研究仍不失为解决宗教实时数据采集提供了一个新的研究视角及研究方法, 未来将对该方法存在的问题进一步研究, 拟通过增加新的数据源, 提高模型拟合精度, 为新时代宗教社会学发展提供新的研究范式。

基金项目

2016 年国家社会科学基金项目“互联网宗教活动的社会影响及对策研究”(项目编号: 16BSH018)。

参考文献

- [1] Bainbridge, W.S. (2012) Internet-Based Research in the Social Science of Religion (ARDA Guiding Paper Series). The Association of Religion Data Archives, The Pennsylvania State University, State College.
- [2] Yeung, T.Y.-C. (2019) Measuring Christian Religiosity by Google Trends. *Review of Religious Research*, **61**, 235-257.
- [3] Wan-Chik, R., Clough, P. and Sanderson, M. (2013) Investigating Religious Information Searching through Analysis of a Search Engine Log. *Journal of the American Society for Information Science and Technology*, **64**, 2492-2506. <https://doi.org/10.1002/asi.22945>
- [4] 郭培清. 反宗教渗透综论[J]. 中央社会主义学院学报, 2007(6): 39-42.
- [5] 黄平. 互联网, 宗教与国际关系——基于结构化理论的资源动员论观点[J]. 世界经济与政治, 2011(9): 52-66.
- [6] 左鹏, 厉彦龙. 基于互联网的基督教传播: 以大学生为对象[J]. 中国青年研究, 2012(1): 97-102.
- [7] 唐名辉. 宗教网站: 一种新型的宗教场所[J]. 世界宗教文化, 2013(2): 89-93.
- [8] 赵冰. “网上宗教”现象在中国的现状, 特征与影响[J]. 世界宗教文化, 2015(1): 100-105.
- [9] 金勋. 互联网时代世界宗教的新形态[J]. 中国宗教, 2015(4): 24-27.
- [10] 刘金光. 国际互联网与宗教渗透[J]. 中国宗教, 2003(8): 27-29.
- [11] 周爱华, 丁明俊. 国内伊斯兰网站的发展及存在问题初探[J]. 西北第二民族学院学报: 哲学社会科学版, 2008(3): 105-108.
- [12] Goh, R.B.H. (2005) The Internet and Christianity in Asia: Cultural Trends, Structures and Transformations. *International Journal of Urban and Regional Research*, **29**, 831-848. <https://doi.org/10.1111/j.1468-2427.2005.00624.x>
- [13] Campbell, H. (2005) Making Space for Religion in Internet Studies. *Information Society*, **21**, 309-315. <https://doi.org/10.1080/01972240591007625>
- [14] Helland, C. (2005) Online Religion as Lived Religion. Methodological Issues in the Study of Religious Participation on the Internet. *Online-Heidelberg Journal of Religions on the Internet*, **1**.
- [15] Campbell, H.A. (2012) Understanding the Relationship between Religion Online and Offline in a Networked Society. *Journal of the American Academy of Religion*, **80**, 64-93. <https://doi.org/10.1093/jaarel/lfr074>
- [16] Drecun, S. and Martinoli, A. (2012) Internet Based Religions: Alternative and Deregulated Systems of Beliefs in New Media Environment. PhD Dissertation, University of Arts in Belgrade.
- [17] Campbell, H.A. (2013) Religion and the Internet: A Microcosm for Studying Internet Trends and Implications. *New Media & Society*, **15**, 680-694. <https://doi.org/10.1177/1461444812462848>
- [18] Bainbridge, W.S. (2007) Expanding the Use of the Internet in Religious Research. *Review of Religious Research*, **49**, 7-20.
- [19] Kawabata, A. and Tamura, T. (2007) Online-Religion in Japan: Websites and Religious Counseling from a Comparative Cross-Cultural Perspective. *Journal of Computer-Mediated Communication*, **12**, 999-1019. <https://doi.org/10.1111/j.1083-6101.2007.00360.x>
- [20] McKenna, K.Y.A. and West, K.J. (2007) Give Me That Online-Time Religion: The Role of the Internet in Spiritual Life. *Computers in Human Behavior*, **23**, 942-954. <https://doi.org/10.1016/j.chb.2005.08.007>
- [21] Shelton, T., Zook, M. and Graham, M. (2012) The Technology of Religion: Mapping Religious Cyberscapes. *The Professional Geographer*, **64**, 602-617. <https://doi.org/10.1080/00330124.2011.614571>
- [22] Clark, A. and Lelkes, O. (2005) Deliver us from Evil: Religion as Insurance. *Papers on Economics of Religion*, **603**, 1-36.
- [23] Dawson, L.L. and Cowan, D.E. (2013) Religion Online: Finding Faith on the Internet. Routledge, London.
- [24] Bettendorf, L. and Dijkgraaf, E. (2011) The Bicausal Relation between Religion and Income. *Applied Economics*, **43**, 1351-1363. <https://doi.org/10.1080/00036840802600442>
- [25] Scheitle, C.P. (2011) Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research. *Social Science Quarterly*, **92**, 285-295. <https://doi.org/10.1111/j.1540-6237.2011.00768.x>