Online Information Review

A study on statistical methods used in six journals of library and information science

Jin Zhang Yuehua Zhao Yanyan Wang

Received 30 July 2015
Revised 8 January 2016
Accepted 11 January 2016

A study on statistical methods used in six journals of library and information science

Jin Zhang, Yuehua Zhao and Yanyan Wang
School of Information Studies, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA

Abstract

Purpose – Quantitative methods, especially statistical methods, play an increasingly important role in research of library and information science (LIS). For different journals, the uses of statistical methods vary substantially due to different journal scopes and aims. The purpose of this paper is to explore the characteristics of statistical methodology uses in six major scholarly journals in LIS.

Design/methodology/approach – Research papers that used statistical methods from the six major journals were selected and investigated. Content analysis method, descriptive statistical analysis method, and temporal analysis method were used to compare and analyze statistical method uses in research papers of the investigated journals.

Findings – The findings of this study show that there was a clear growth trend of statistical method uses in five of the investigated journals; statistical methods were used most in *The Journal of the Association for Information Science and Technology* and *Information Processing & Management*; and the top three most frequently used statistical methods were *t*-test, ANOVA test, and χ^2-test.

Originality/value – The findings can be used to better understand the application areas, patterns, and trends of statistical methods among the investigated journals and their statistical methodology orientations in research studies of LIS.

Keywords Methodology, Quantitative research, Journal, Statistical method

Paper type Research paper

1. Introduction

The growth of scientific and technical knowledge has resulted in a proliferation of studies in a variety of disciplines. Journals are primary channels for researchers to publish their research findings. Research methodology determines at least to some degree the quality of a research study. Most modern research studies adhere to a well-defined research design based on the appropriate selection of research methodology. The application of both quantitative research methodology and qualitative research methodology may vary in different fields and research problems. A wide variety of statistical methods has been applied to analyze the data obtained.

Statistics is a branch of mathematics which can be applied to analyze various data. The applications of statistical techniques promote the establishment of relationships between the quantitative data and the quantitative research questions, and to interpret the complicated resultant data pattern. To speak of statistical methodology, inferential statistics, and predictive statistics are widely utilized in almost all research domains. The inferential statistical methods, such as *t*-test and ANOVA, concentrate on hypothesis testing, while the predictive statistical methods focusses on correlation analysis and regression analysis. Both inferential statistics and predictive statistics play an extremely important role in quantitative research.

Library and information science (LIS) is a broad discipline that has adopted a wide variety of continuously evolving research strategies and techniques (Hider and Pymm,

Online Information Review
Vol. 40 No. 3, 2016
pp. 416-434

1468-4527
DOI 10.1108/OIR-07-2015-0247

2008). Like other disciplines such as mathematics, biology, and physics, LIS is a developing and expanding field. It is particularly necessary to understand the methodology changes over time in LIS research. Many previous evaluations of research studies in LIS have been conducted. They include the methods of content analysis and co-citation analysis on journal articles (e.g. Järvelin and Vakkari, 1993) and dissertations (e.g. Sugimoto *et al.*, 2011). The majority of these studies address research topics in LIS.

Although studies have made substantial contributions to journal comparison and evaluation, prior studies have rarely explicated statistical research method orientations of the journals. Research methodology orientation of journals provides a unique way to assess the journals and their characteristics. In this study, we chose six major scholarly journals from 1999 to 2013 in LIS and investigated their statistical methodology.

This study sheds light on the statistical method use patterns in the major LIS journals and temporal change trends of the statistical method uses in these journals; helps people better understand the statistical method orientations of these journals; and assists researchers in selecting proper publication journals for their research papers.

2. Problem statement

Because of the importance of applying statistical methodology to research studies, and the lack of an overview of statistical method uses in LIS, this study aims at analyzing and comparing the characteristics of statistical methodology used in the major scholarly journals in LIS. The overarching research problem of this study is stated as follows: what is the overall status of statistical method uses in research studies published in the six major scholarly journals of LIS between 1999 and 2013?

The specific research questions of this study:

RQ1. Are statistical methods used differently in the major scholarly journals in LIS?

RQ2. Are the temporal change patterns of the statistical method uses in the major scholarly journals in LIS different?

The objectives of this study are to investigate the use of statistical methods in the six major scholarly journals in LIS; ascertain the application areas of statistical methods; compare the uses of statistical methods among the six journals; and reveal the characteristics and temporal patterns of the applications of statistical methods in the investigated journals.

3. Literature review

The studies of scholarly journals have attracted researchers for decades. The longitudinal and systematic study of core journals sheds light on the understanding of the development and evolution of a certain discipline. Research studies on journal publications in a domain field can reveal the research topic trends, leading researchers, and research methods in that field. The distributions and relationships of research topics in LIS have been investigated by Järvelin and Vakkari (1993). The authors sampled full-length research articles published in core LIS journals and identified library and information-service (L&I) activities as well as information storage and retrieval (IS&R) as main research topics. In addition, Pettigrew and Mckechnie (2001) reported the use of theory in LIS by investigating six journals.

In addition to the research topics of the journal articles, research methodology is also critical to research studies in any disciplines. There are three categories of widely used

research methodologies: quantitative, qualitative, and mixed research methodologies. In addition to having to decide the research question, the researcher must determine whether it will be quantitative or qualitative in nature (Connaway and Powell, 2010). Quantitative research applies a measurement to phenomena that can be expressed in terms of quantity (Kothari, 2004). Qualitative research tends to apply a more holistic and natural approach to the resolution of a problem than does quantitative research (Connaway and Powell, 2010). The mixed methodology is designed to answer the increasingly complex and multifaceted research questions using both quantitative and qualitative measurements (Tashakkori and Teddlie, 2003). In spite of the different focusses, any of these three types of the methodologies share common components such as research design, sampling methods, data collection methods, and data analysis techniques (Kothari, 2004).

The full spectrum of research methods and the essential role of research methodology in different domains have been presented by a variety of researchers. Busha and Harter (1980) provided practical and theoretical insights, as well as discussed technical skills needed to plan research projects and analyze research data effectively for library science. Using the domain of physics, Rajasekar *et al.* (2006) indicated that training in choosing research methods, materials, and scientific tools plays an important role to students and new researchers.

Journal articles are widely used to identify the application status of research methods. Perry and Kraemer (1986) selected *Public Administration Review* (1975-1984) to examine research methods in public administration. They found that the essential research methodologies applied are logical argumentation, legal briefs, and empirical analysis. Through a content analysis of three management journals, Scandura and Williams (2000) reported the research strategies in management research mainly included literature reviews, sample survey, laboratory experimentation, and computer simulation. Bernard (2011) pointed out that anthropologists developed some of today's widely used methods, such as questionnaire survey method and experimentation.

Previous studies employed content analysis and statistical methods to investigate the method applications and research topics in LIS. Research methodologies in a series of studies in LIS were evaluated by using content analysis (Jarvelin and Vakkari, 1990, 1993). They developed a classification scheme with seven variables for the investigated papers. Using this scheme, Hider and Pymm (2008) classified different research strategies, data collection methods, and data analysis methods for the papers in 20 high-profile LIS journals in 2005. In a recent article, Tuomaala *et al.* (2014) applied content analysis and descriptive statistical methods to reveal the trends of research topics and methodologies used in LIS papers from 1965 to 2005. Williams and Winston (2003) examined 119 articles from five LIS journals and evaluated the importance of research methods in academic librarian's decision making, research, and publication. Their findings show that the research topics shifted from library management and information organization (IO), to end users and system development during the past several decades, and the use of empirical research strategies rose over time.

Content analysis is also applied to reveal the subjects and themes of LIS studies. By reviewing a total of 2,664 articles from 91 journals published in 2001, Koufogiannakis *et al.* (2004) identified information access and retrieval as the most productive topic in this field. Through the comparisons of LIS research in six countries: Australia, China, Finland, Spain, Turkey, and the UK, Rochester and Vakkari (2003) also reported that IS&R as well as L&Is were the most popular topics

in most of the studied countries. In a later study, Davarpanah and Aslekia (2008) investigated a total of 894 articles from 56 LIS journals during the years of 2000-2004, and indicated that communications and information technology were the most prevalent topics over the five years. Moreover, other studies also demonstrated that the new technologies in information science had a tremendous impact on the research trends in LIS (Blessinger and Frasier, 2007; Aharony, 2012). In a study, Tuomaala *et al.* (2014) conducted a longitudinal analysis of the distributions of the research topics in LIS over the 40 years, and defined the quantitative core of LIS as the following four research areas: IS&R, scientific communication, L&I activities, and information seeking.

Quantitative methodology is an umbrella term for a broad range of specialized topics and approaches (Somekh and Lewin, 2005). Within the scope of quantitative methodology, inferential statistics covers the techniques which allow researchers to explore in-depth relationships between variables and provides the most useful and powerful tools in data analysis. In the late 1970s and early 1980s, all the social science, including LIS, witnessed a range of "new" approaches to research methodology that influenced the way that research studies were conducted (Somekh and Lewin, 2005). Since then, statistics and statistical methodology have increasingly been used in all disciplines of social sciences, as well as in several areas of humanities (Medhi, 1992). Wallace (1985) compared the use of statistics in 99 journals from four subject areas (LIS, education, social work, and business) and suggested that the use of inferential statistics in the literature of a field can be an indicator of its scientific identity (Wallace, 1985).

Lack of the solid research methodology for LIS has been criticized. Martyn and Lancaster (1981) reported that much of the earlier research in library science was heavily based on opinion rather than the experimental design or scientific analysis. Powell (1997) stated that in a variety of research studies in library science by public librarians research methodologies were limited. Järvelin and Vakkari (1993) distinguished research strategies, methods of data collection, and methods of data analysis in LIS research. They found that there was very little research on methodology (1-8 percent) and concluded a loss of interest in methodology from 1965 to 1985. Later on, Hildreth and Aytac (2007) conducted a methodological study which showed that few of the studies in the field of LIS adopted correlational or inferential statistical analysis methods (Hildreth and Aytac, 2007).

One of the debates in journal publications and books on LIS is whether the field really qualifies as a science (Wallace, 1985) even though the use of quantitative methodology, especially statistical methods, does appear to be a meaningful indicator of inquiry in LIS. Thus, providing an overview of quantitative research methods used in research articles from the scholarly journals is desired. Although some studies have been conducted on research articles which employed statistical methods, few have concentrated on a thorough and systematic analysis of statistical methodology uses in the top-ranking journals and the temporal change analysis of statistical method applications in LIS.

4. Method

The journal factor has six levels, which are *The Journal of the Association for Information Science and Technology* (*JASIST*), *Journal of Documentation* (*JD*), *Journal of Information Science* (*JIS*), *The Library Quarterly* (*LQ*), *Library and Information Science Research* (*LISR*), *Information Processing & Management* (*IPM*). The basic

criteria in this study for journal selection were: these journals are regarded as top journals in LIS; they are research-oriented journals rather than practical-oriented ones; they have a relatively long history so that sufficient data can be collected for the temporal analysis; and the full texts of research papers in these journals can be acquired in order to analyze the application areas and statistical methods used in these studies. Five of these journals (*JASIST*, *JD*, *LQ*, *LISR*, and *IPM*) were also selected to investigate the use of theory in LIS by Pettigrew and Mckechnie (2001).

The criteria for determining top journals were based on both the perceptions of domain experts and citation data such as the impact factor in the Institute for Scientific Information's Journal Citation Reports (Nisonger and Davis, 2005). Manzari (2013) investigated 232 full-time faculty of ALA-accredited programs in LIS. Respondents were asked to rate a list of journals on a scale from 1 (low) to 5 (high) based on each journal's importance to their research and teaching, and also list the five most prestigious journals. Table I illustrates the rankings of the investigated journals on the list of "journals by average ratings" and "journals listed among the top five prestigious" from the study of Manzari (2013), and results from Journal Citation Report (Thomson Reuters, 2013) based on the five-year impact factors.

The statistical method factor includes all statistical methods used in the field except descriptive statistical method, which only offers basic statistical characteristics of a sample from a population (like means and standard deviation) rather than draws an inferential conclusion about the population. Three distinctive levels of the statistical method factor were defined: inferential statistical method, predictive statistical method, and other statistical methods.

The third factor is the application areas of statistical methods. It addresses where the statistical methods were utilized in research papers. Six levels of this factor were determined based on the six phases of the information cycle. They are information creation (IC), information selection and control (ISC), IO, information retrieval (IR), information dissemination (ID), and information use (IU) (Hodge, 2000). This factor reveals the application focus of statistical methods and displays which research areas attracted statistical methods the most.

After useful information was extracted from each of the collected papers, a content analysis followed. The content analysis included the following two aspects. First, classifying it into a proper application area of LIS. The application areas are IC, ISC, IO, IR, ID, and IU. These six areas were used in this study because they reflect a complete information cycle, and every professional activity or research study in LIS can be associated with an application area. The six application areas are mutually exclusive, which is important for the classification processing for the involved papers in the study. Second, categorizing the statistical methods used in the paper into meaningful

Table I. Rankings of the selected journals on the three journal ranking lists

Rankings	Journals by average ratings	Journals listed among the top five prestigious	Five-year impact factors
JASIST	1	1	19
LQ	2	2	37
JD	4	4	28
IPM	7	3	26
LISR	6	5	67
JIS	16	13	27

statistics groups. This categorization process was conducted based on a statistics schema (Gravetter and Wallnau, 2009). The schema covers a variety of statistical methods used in LIS such as *t*-test, ANOVA test, χ^2-test, Mann-Whitney, Wilcoxon, Kruskal-Wallis, Friedman, etc. These methods belong to inferential statistical method category, predictive statistical method category, or statistical method category. These categories are also mutually exclusive and logically meaningful.

It is recognized that a study in a paper can involve multiple statistical methods. It happened when a study addressed a sophisticated issue. In the content analysis process, if a paper employed multiple statistical methods, all the statistical methods were specified and recorded. The specified statistical methods were classified into corresponding categories, respectively.

4.1 Data collection and analysis

In this study, the researchers investigated all the research papers published from 1999 to 2013 in the six scholarly journals, identified above and selected only the research papers that used statistics for analysis. Extracted from the bibliographic information and the full texts of each of these selected research papers were the article title, the publication year, the journal title, the journal issue number, the subject terms, the application areas, and the statistical methods applied. The application areas of the investigated research papers were summarized and classified by content analysis.

A temporal analysis was applied to application areas, and statistical methods of the six investigated journals, respectively.

5. Results and analysis

5.1 Descriptive statistical analysis

5.1.1 Journals. In this study, six journals were selected according to the procedure discussed above. The total number of the investigated issues of these journals from 1999 to 2013 was 589 and the number of the qualified research papers from these issues equaled 5,175. The detailed data on the six journals are listed in Table II. The number of the research papers from these issues using statistical methods was 1,497. They accounted for 28.93 percent of the total qualified research papers. Among these papers, there were 769 (34.33 percent) papers from *JASIST*, 99 (17.97 percent) papers from *JD*, 160 (22.66 percent) papers from *JIS*, 25 (9.73 percent) papers from *LQ*, 123 (31.54 percent) papers from *LISR*, and 321 (31.14 percent) papers from *IPM*. *JASIST* got the highest percentage of the research papers using statistical methods among the six journals, while *LQ* had the lowest percentage (9.73 percent).

Journals	Number of the qualified papers	Number of the papers using statistical methods	% of the papers using statistical methods
JASIST	2,240	769	34.33
JD	551	99	17.97
JIS	706	160	22.66
LQ	257	25	9.73
LISR	390	123	31.54
IPM	1,031	321	31.14
Total	5,175	1,497	28.93

Table II. Research paper distribution in the six journals

The means and standard deviations of the frequency of the papers using statistical methods per issue (FPUSMPI) in the journals are listed in Table III. The means indicate the frequency of statistical method utilization. The standard deviations suggest the stability of statistical method utilization in all the issues from 1999 to 2013 of each journal. In Tables II and III, it is obvious that *JASIST* published the most research papers using statistical methods and obtained the highest mean among the six journals ($\mu = 3.845$, $n = 769$). The value of μ stands for the mean of the FPUSMPI of a specific journal. The value of n represents the total number of the papers using statistical methods in a specific journal. The mean of *IPM* was the second highest, and the number of research papers using statistical methods in this journal was also the second highest among the six journals ($\mu = 3.489$, $n = 321$). The mean of *JIS* ranked the fourth place, while its number of the papers using statistical methods ranked the third place ($\mu = 1.778$, $n = 160$). The mean of *LISR* occupied the third position, and its number of the papers using statistical methods was ranked the fourth place ($\mu = 2.033$, $n = 123$). The means and numbers of the papers of *JD* and *LQ* were the least among the six journals.

5.1.2 Application areas. Table IV and Figure 1 display the descriptive statistical data on the six application areas in the six journals. The application areas are IC, ISC, IO, IR, ID, and IU. Table IV demonstrates the means and standard deviations of the FPUSMPI in the six application areas, while Figure 1 displays the total numbers of

Table III. Means and standard deviations of the FPUSMPI in the six journals

Journals	Number of issues	Mean	SD	95% confidence interval for mean		Minimum	Maximum
				Lower bound	Upper bound		
JASIST	200	3.845	2.649	3.476	4.214	0.000	12.000
JD	89	1.112	1.027	0.896	1.329	0.000	4.000
JIS	90	1.778	1.467	1.471	2.085	0.000	6.000
LQ	60	0.383	0.666	0.211	0.555	0.000	3.000
LISR	60	2.033	1.414	1.668	2.399	0.000	5.000
IPM	90	3.489	2.777	2.907	4.071	0.000	17.000
Total	589	2.525	2.426	2.328	2.721	0.000	17.000

Table IV. Descriptive data of the six application areas in the six journals

Application area	Mean	SD	Mean	SD	Mean	SD
	JASIST		*JD*		*JIS*	
IC	0.135	0.409	0.022	0.149	0.089	0.323
ISC	1.125	1.318	0.337	0.521	0.411	0.652
IO	0.365	0.635	0.067	0.252	0.233	0.425
IR	1.205	1.019	0.449	0.584	0.500	0.723
ID	0.295	0.686	0.011	0.106	0.222	0.469
IU	0.720	0.903	0.225	0.420	0.322	0.537
	LQ		*LISR*		*IPM*	
IC	0.033	0.181	0.250	0.474	0.189	0.421
ISC	0.067	0.252	0.417	0.561	0.211	0.530
IO	0.017	0.129	0.100	0.303	0.422	0.670
IR	0.033	0.181	0.317	0.567	2.056	1.951
ID	0.017	0.129	0.017	0.129	0.078	0.308
IU	0.250	0.474	0.933	0.972	0.611	1.168

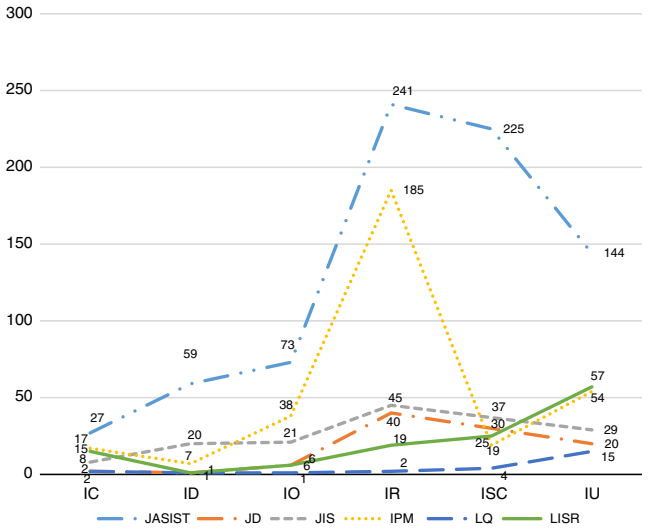

Figure 1.
Numbers of the application papers in the six journals

the research papers using statistical methods on these areas in each journal. In Figure 1, its *X*-axis represents the application area, and the *Y*-axis represents the frequency of the research papers using statistical methods.

In the IR area, *JASIST* ($\mu = 1.205$, $n = 241$), and *IPM* ($\mu = 2.056$, $n = 185$) attracted more research papers using statistical methods than *JIS* ($\mu = 0.500$, $n = 45$), *JD* ($\mu = 0.449$, $n = 40$), or *LISR* ($\mu = 0.317$, $n = 19$). The value of μ stands for the mean of the numbers of the research papers which used statistical methods in a certain application area per issue of a specific journal, and the value of n stands for the total number of the papers using statistical methods in a certain application area of a specific journal. *LQ* ($\mu = 0.033$, $n = 2$) attracted the least number of research papers using the statistical methods among the six journals in the IR area.

In the IO area, *JASIST* ($\mu = 0.365$, $n = 73$), *IPM* ($\mu = 0.670$, $n = 38$), and *JIS* ($\mu = 0.233$, $n = 21$) filled the top three positions. *JD* ($\mu = 0.067$, $n = 6$) and *LISR* ($\mu = 0.100$, $n = 6$) published the same number of research papers that used statistical methods. *LQ* ($\mu = 0.017$, $n = 1$) achieved the sixth place among the six journals.

JASIST ($\mu = 0.295$, $n = 59$) reached first place in the ID area. *JIS* ($\mu = 0.222$, $n = 20$) attracted more research papers using the statistical methods than the remaining four journals. The third place was occupied by *IPM* ($\mu = 0.078$, $n = 7$). *JD* ($\mu = 0.011$, $n = 1$), *LQ* ($\mu = 0.017$, $n = 1$), and *LISR* ($\mu = 0.017$, $n = 1$) attracted the same number of the research papers from 1999 to 2013.

LQ ($\mu = 0.033$, $n = 2$) and *JD* ($\mu = 0.022$, $n = 2$) had the least number of research papers using statistical methods among the six journals in the IC area. *JASIST* ($\mu = 0.135$, $n = 27$), *IPM* ($\mu = 0.189$, $n = 17$), *LISR* ($\mu = 0.250$, $n = 15$), and *JIS* ($\mu = 0.089$, $n = 8$) received the first, second, third, and fourth positions.

In the IU area, *JASIST* ($\mu = 0.720$, $n = 144$) still ranked in first place. *LISR* ($\mu = 0.933$, $n = 57$) placed second, and *IPM* ($\mu = 0.189$, $n = 54$) ranked third in this category. Meanwhile, *JIS* ($\mu = 0.089$, $n = 29$), *JD* ($\mu = 0.022$, $n = 20$), and *LQ* ($\mu = 0.033$, $n = 15$) clenched the last three positions.

In the ISC area, *JASIST* ($\mu = 1.125$, $n = 225$) still reached first place. *JIS* ($\mu = 0.411$, $n = 37$) edged out the other four journals. *JD* ($\mu = 0.337$, $n = 30$), *LISR* ($\mu = 0.417$, $n = 25$), and *IPM* ($\mu = 0.211$, $n = 19$) occupied the third, fourth, and fifth positions. *LQ* ($\mu = 0.067$, $n = 4$) still received the last place.

Among all the six journals, *JASIST* topped all the other journals in all the application areas, while *LQ* published the fewest research papers that used statistical methods in all the application areas.

5.1.3 Statistical methods. Table V and Figure 2 illustrate the descriptive statistical data of the investigated statistical methods for the six journals. In Figure 2, the *X*-axis stands for the type of statistical methods, and the *Y*-axis stands for the frequency of the uses of statistical methods. In terms of the use of inferential statistical methods, *JASIST* ($\mu = 3.340$, $n = 668$) surpassed all the other journals. The value of μ stands for the mean of the number of the papers using a specific type of statistical methods in a specific journal. The value of n represents the total number of the papers using a specific type of statistical methods in a journal. *IPM* ($\mu = 3.478$, $n = 313$), *JIS* ($\mu = 1.311$, $n = 118$), *LISR* ($\mu = 1.683$, $n = 101$), and *JD* ($\mu = 0.820$, $n = 73$) ranked in second, third, fourth, and fifth places respectively. *LQ* ($\mu = 0.317$, $n = 19$) attracted the least research papers using statistical methods among the six journals.

Table V.
Descriptive data of statistical methods in the six journals

Statistical methods	Mean	SD	Mean	SD	Mean	SD
	JASIST		*JD*		*JIS*	
Inferential statistical method	3.340	2.730	0.820	1.018	1.311	1.363
Predictive statistical method	2.725	2.478	0.663	0.865	1.011	1.311
Other method	1.675	2.612	0.169	0.406	0.756	1.327
	LQ		*LISR*		*IPM*	
Inferential statistical method	0.317	0.651	1.683	1.578	3.478	3.216
Predictive statistical method	0.067	0.312	0.683	1.172	0.556	1.133
Other method	0.183	0.504	0.867	1.171	1.111	1.449

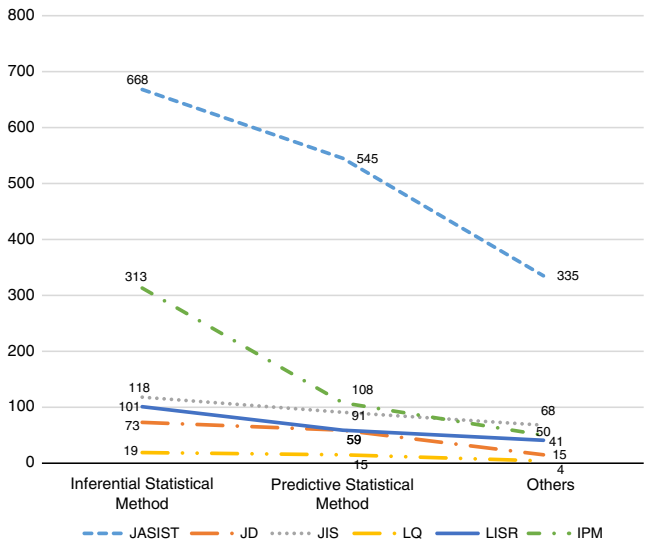

Figure 2.
Comparison of statistical methods across the six journals

In regards to the research papers employing predictive statistical methods, *JASIST* ($\mu = 2.725$, $n = 545$) published much more research papers than the other five journals. *IPM* ($\mu = 0.556$, $n = 108$) and *JIS* ($\mu = 1.011$, $n = 91$) were ranked in second and third places, respectively, while *LISR* ($\mu = 0.683$, $n = 59$) and *JD* ($\mu = 0.663$, $n = 59$) both ranked in fourth place among the six journals. *LQ* ($\mu = 0.067$, $n = 15$) occupied the last place.

When comparing the six journals in terms of their use of other statistical methods, *JASIST* ($\mu = 1.675$, $n = 335$) still edged out the other journals. *JIS* ($\mu = 0.756$, $n = 68$) received more research papers than *IPM* ($\mu = 1.111$, $n = 50$) and they were placed in the second and third positions among the six journals. *LISR* ($\mu = 0.867$, $n = 41$), *JD* ($\mu = 0.169$, $n = 15$), and *LQ* ($\mu = 0.183$, $n = 4$) published the least research papers in this category.

The data also demonstrate that inferential statistical methods were applied most frequently in all the six journals. *JASIST* employed statistical methods more than other journals across all the three categories, while *LQ* that published only 38 papers using statistical methods and clenched the last position among the six journals.

5.2 Temporal analysis

Two journals, *JASIST* and *IPM*, clearly showed increasing trends in their use of statistical methods in research papers between 1999 and 2013. Figure 3 presents the total number of research papers that used statistical methods in the six journals during the 15-year span. Its *X*-axis represents publication years from 1999 to 2013, while its *Y*-axis represents the number of total research papers using statistical methods. In this figure, *JASIST* published the most papers using statistical methods among the six journals in every individual year. In 2009 and 2012, it reached its peak and published 87 research papers in each of the two years, while in 1999 it only published 16 research papers. *IPM* ranked second in most years except 2003, 2004, and 2010. In 2003, *IPM* published seven research papers that used statistical methods, which was the least in the 15 years. In total, 31 of research papers in *IPM* were published in 2007. Slight increases of the papers using statistical methods were found in *JD*, *JIS*, and *LISR*, respectively. The largest increase of research papers was 13 in *JD*. In *JIS* the largest

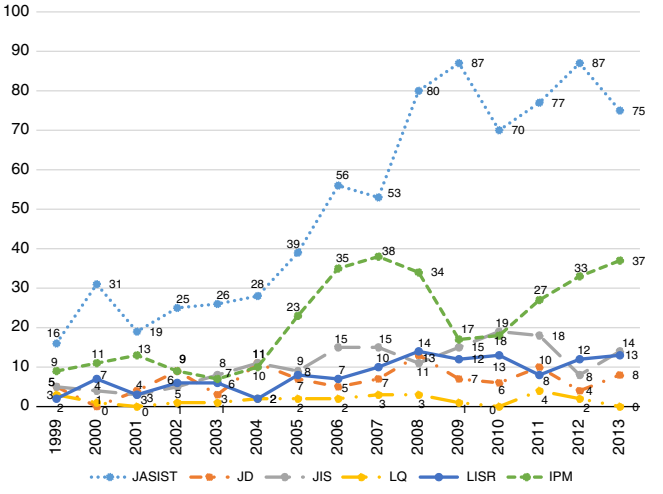

Figure 3.
Temporal analysis of the papers using statistical methods across the six journals

increase was 16, and in *LISR* the number was 12. However, according to Figure 3 no obvious growth in *LQ* was found. This journal published the least number of papers using statistical methods among the six journals from 2001 to 2013, and ranked in fifth place in 1999 and 2000.

5.2.1 Temporal analysis for the application areas. Figure 4, including six charts, displays the temporal patterns of the papers using statistical methods among the six application areas in all the six investigated journals. In these charts the *X*-axis represents publication years from 1999 to 2013, while the *Y*-axis represents the number of total research papers using statistical methods. Each curve denotes an application area. Using this detailed temporal analysis of the application areas of all the six journals we can determine the areas that drove the growth of the statistical methods applications.

Chart (a) in Figure 4 displays the temporal pattern of the papers using statistical methods among the six application areas in *JASIST*. And chart (b) provides the corresponding description of *IPM*. Although, throughout this period, the total numbers of papers adopting statistical methods in both the journals increased, all six areas show the proliferations for *JASIST*. Three areas (ISC, IR, and IU) contributed the most for the total growth for *JASIST*. Studies in these areas included researcher and institution

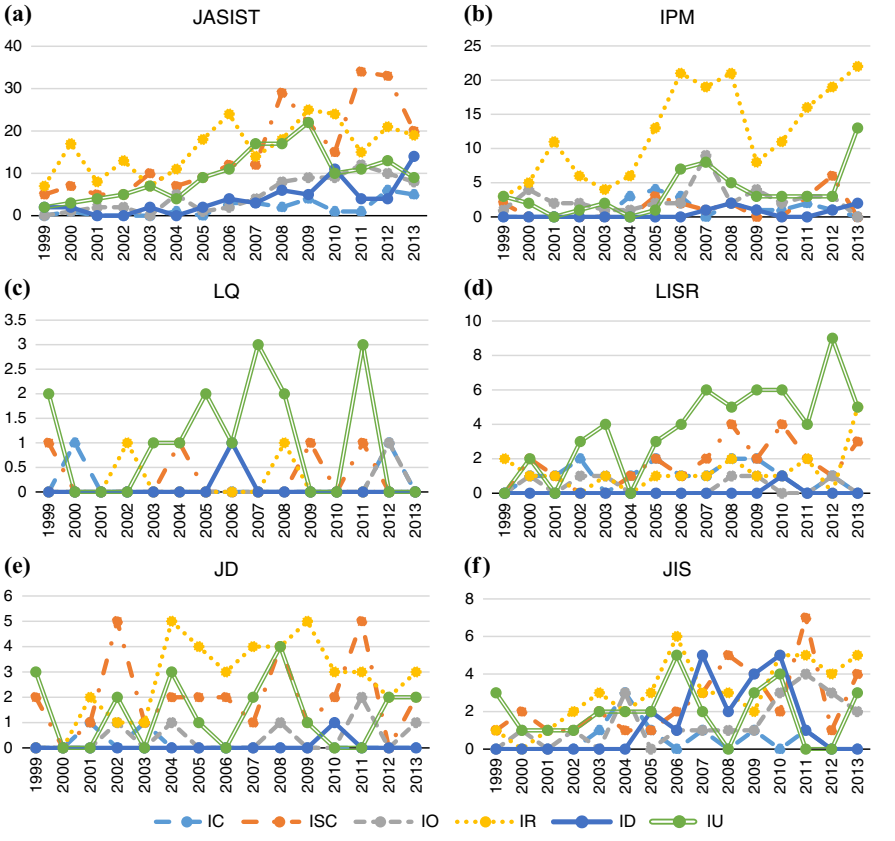

Figure 4.
Temporal analysis of the application areas for six journals

evaluation (e.g. Bornmann *et al.*, 2013), publication evaluation (e.g. Leydesdorff *et al.*, 2011), data mining (e.g. Bae and Lee, 2012), search behavior (e.g. Chung and Chen, 2009), and information evaluation indicators (e.g. Zhao *et al.*, 2012). These findings reflected the aims and scope of *JASIST*, focussing on the production, discovery, recording, storage, representation, retrieval, presentation, manipulation, dissemination, use, and evaluation of information and on the tools and techniques associated with these processes. Interestingly, the scope of *JASIST* includes the six application areas and each area corresponded to a large number of papers using statistical methods.

Notice that the IR area and IU area indicated clear expansion in the use of statistical methods in *IPM*. The IR area obviously dominated the adoption of statistical methodology in *IPM*. Studies within this area included data mining (e.g. Ozmutlu *et al.*, 2002), retrieval algorithm and theory (e.g. Qin *et al.*, 2008), and text summarization (e.g. Ouyang *et al.*, 2013). According to the aims and scope of *IPM*, this journal is devoted to the basic and applied research in information science, computer science, and cognitive science. The orientation of *IPM* seems more focussed on the information techniques. Thus, it is reasonable that the IR area significantly employed statistical methods among its published papers.

LQ and *LISR*, as shown in (c) and (d), did not adopt as many statistical methods as the other four journals. For example, *LQ* just had one paper applying the statistical methods to the ID research in 2000 (Kwon and Zweizig, 2006), and *LISR* only had three papers that adopted the statistical methods in the IO area (e.g. Stvilia and Jörgensen, 2009). Some applications areas did not use statistical methods in some years in these journals. Both *LQ* and *LISR* focussed more on the IU area. The most dominant topic that utilized statistical methods was library service and usage for both *LQ* (e.g. Hider, 2008) and *LISR* (e.g. Aabø *et al.*, 2010).

In chart (e) and (f), the fluctuation of the curves identifies the variation of the applications of statistical methods from 1999 to 2013 for *JD* and *JIS*. Both of these journals concentrated more on the *ISC* and *IR*. *JD* published more papers using statistical methods in publication evaluation (e.g. Nicholas *et al.*, 2005) and search behavior (e.g. Heinström, 2005) while *JIS* published more papers using statistical methods in information quality control (e.g. Surman and Bath, 2013) and search behavior (e.g. Wei *et al.*, 2005).

5.2.2 The temporal analysis for statistical research methods. Figure 5 displays the temporal analysis of statistical methods in each journal. The *X*-axis of each chart in Figure 5 represents publication years from 1999 to 2013 and the *Y*-axis represents the number of total papers using statistical methods in a corresponding year. Each curve denotes an individual category of statistical methods. The temporal analysis of statistical methods applied in the six journals show that all the six investigated journals adopted more inferential statistical methods than the predictive statistical methods. Only in a few years did the predictive methods outnumber the inferential methods. Chart (a) and (b) show that from 1999 to 2013, all three categories of statistical methods had clear growth in *JASIST*, and only inferential methods increased in *IPM*. The inferential statistical methods in *IPM* were used to test the performance of proposed systems (Schedl *et al.*, 2011), algorithm (e.g. Lee and Kim, 2013), and methods (e.g. Lee and Croft, 2013). More specifically, within all the statistical methods, *t*-test (99), ANOVA (55) and χ^2-test (51) were utilized most in *IPM*; *t*-test (187), correlation (163) and ANOVA (132) were employed most in *JASIST*; χ^2-test

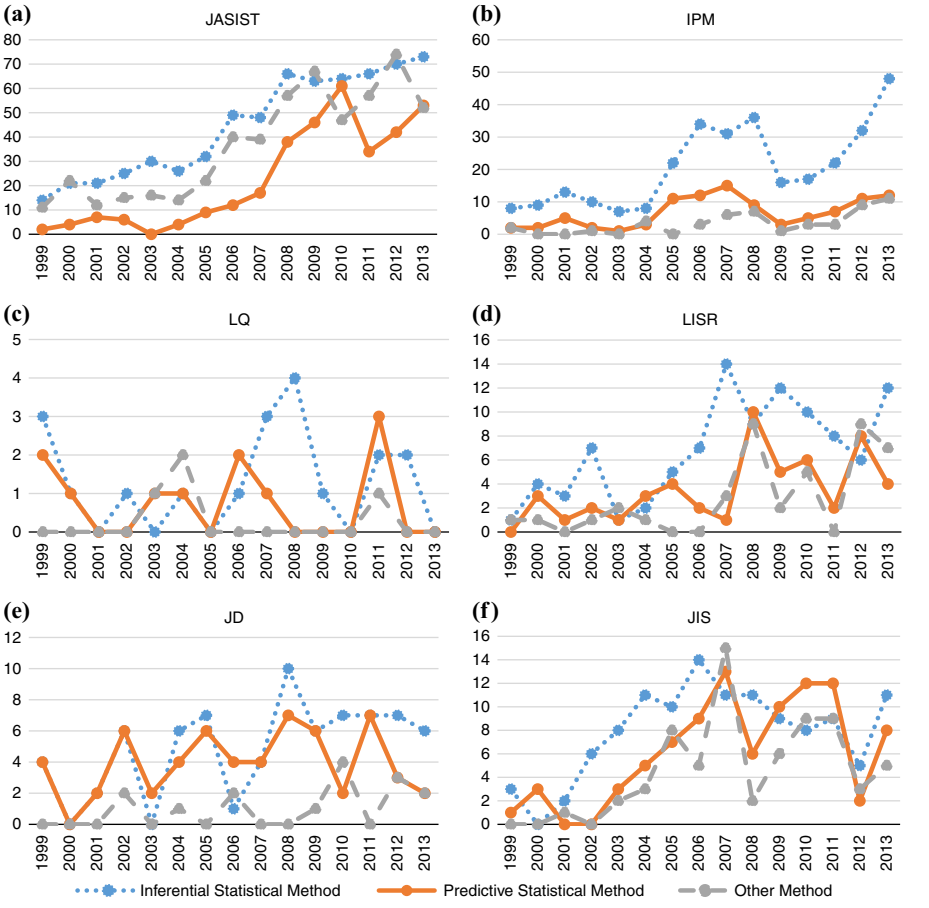

Figure 5.
Temporal analysis of statistical methods for six journals

(9), correlation (6) and regression (5) were employed most in *LQ*; χ^2-test (38), ANOVA (31) and *t*-test (22) were used most in *LISR*; χ^2-test (25), Pearson's correlation (18) and correlation (13) were found most in *JD*; and χ^2-test (31), *t*-test (30) and correlation (28) were applied most in *JIS*.

Among the three defined types of statistical methods: inferential statistical methods, predictive statistical methods, and other methods, the amount of the papers which employed other methods was comparatively small. This category, other methods, included more sophisticated statistical applications, such as the factor analysis, the inter-rater agreement tests, and a variety of *post hoc* tests. From (a), we noticed that *JASIST* applied more of these other methods than the predictive methods. From 2004 to 2012, the employments of these more sophisticated statistical methods clearly rose. In 2000, 2009, and 2012, the applications of other methods even outnumbered the inferential methods. For *JASIST*, the most used methods within the other methods category were the Cronbach's α, the Cohen's κ, and the Tukey test. For instance, the Cronbach's α method was used for the reliability of data (Cortese and Lustria, 2012), the Cohen's κ for the inter-rater agreement for qualitative study (Thelwall *et al.*, 2010), and

the Tukey test in conjunction with an ANOVA sample mean differences (Li, 2009). In contrast, for *LISR* and *JIS*, the curve of predictive statistical method and other methods intertwined with each other throughout the investigated time period. Papers published in *JIS* also applied the confirmatory factor analysis to identify factors based on their experiment data, such as defining the motivational factors which would facilitate the knowledge transfer (Kang *et al.*, 2010). These kinds of sophisticated methods did not frequently appear in the other three journals (*IPM*, *LQ*, and *JD)*.

5.3 Application areas analysis of the journals

The composition patterns varied in the investigated journals. In Figure 6, each pie chart represents the percentage distributions of the six application areas in a journal. *JASIST*, *JIS*, and *JD* showed similar proportional structure: the IR area dominated, followed by ISC area and IU area. *LISR* and *LQ*, however, were dictated by the IU area, whereas *IPM* was governed by the IR area. The findings show that in *IPM*, research on the IR system and algorithm, multimedia retrieval, search behavior, query expansion, cross-language and bilingual retrieval, relevance judgment, information system and

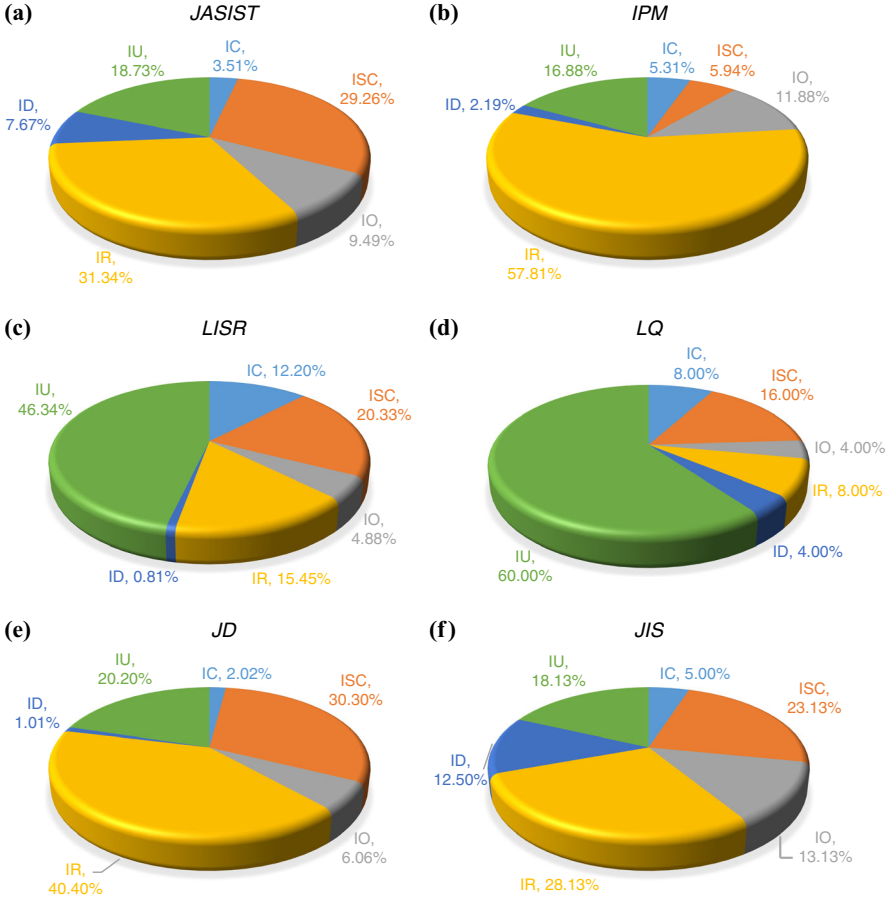

Figure 6.
Percentage distributions of the application areas in the journals

performance evaluation, information literacy, text summarization, retrieval algorithm and theory, web search, natural language processing, and data mining usually incorporated statistical methods.

Regarding the most frequently applied area, the journals varied. As discussed before, *JASIST*, *JIS*, and *JD* concentrated the most on search behavior research, whereas *IPM* had greater emphasis on information system and performance evaluation, and *LQ* and *LISR* adopted statistical methods to deal with the library service and usage. According to the aims of the journals, *LQ* stands as the journal best positioned to chronicle evolution of libraries and the related intersections of information, community, and policy, and *LISR* focusses on the research process in a library. Thus, the findings show that the application areas are consistent with the scopes of the journals.

6. Conclusion

With the growth of research studies in LIS, the quality of research studies is taken more and more seriously by researchers. Research questions, research methodology, and implications of research papers are widely recognized to determine the quality of research papers. Among these factors, research methodology is one of the most important factors. Since scientific and technical methods are applied more and more in LIS research, quantitative methods, especially statistical methods, play an increasingly important role in research. However, statistical methods vary substantially probably due to different journal scopes and aims.

This study explored the characteristics of statistical methodology use in major scholarly journals in LIS. The characteristics of using statistical methods were analyzed in terms of the application areas and the temporal changes. Six major journals (*JASIST*, *IPM, LQ*, *JIS*, *LISR*, and *JD*) were investigated. These journals focus on different research topics and have various scopes but are all top journals in this field. The research papers from 1999 to 2013 in these journals were examined, the total investigated issues were 589 and the total number of the investigated research papers was 1,497.

One of the major findings of this study is that clear increase trends of research papers that used statistical methods were discovered from five of the scholarly journals (*JASIST*, *IPM*, *JIS*, *LISR*, and *JD*). Among the five journals, the growth of research papers that used statistical methods was the largest in *JASIST*, while the second largest growth was revealed in *IPM*. However, no clear increase was found in *LQ* from 1999 to 2013.

Differences were discovered among the six journals in terms of statistical methods uses. Also, different temporal patterns of statistical method uses and distributions of the application areas were found in different journals. In *JASIST* and *IPM*, the two journals that utilized statistical methods the most among the six journals, the numbers of the papers adopting statistical methods increased in all the six application areas, especially *ISC*, *IR*, and IU areas. For *LQ* and *LISR*, since they did not adopt as many statistical methods as the other four journals, some application areas such as ID even did not appear in several years because they did not apply statistical methods in that application area during those years. Both *JD* and *JIS* focussed more on ISC, and IR.

At the same time, differences were also found among the six journals according to their use of each category of statistical methods. However, the most frequently used statistical methods were similar in different journals. For example, in both *JASIST* and

IPM, the top three most frequently used statistical methods were *t*-test, ANOVA test, and χ^2-test.

Generally speaking, the more journals in LIS were included, the clearer the picture of the application of statistical methods in the field is. However, if all journals related to LIS were selected in the study, it was difficult to define the criteria of a related journal to LIS.

If non-statistically based qualitative methods, and mixed qualitative and quantitative methods were examined in this study, they could be compared with statistically based quantitative methods in the contents of the six application areas in LIS. It would provide a different perspective for the investigated journals.

The implications of this study include: first, that the study offers a unique way to evaluate academic journals from the perspective of statistical research method uses. It helps readers better understand the research method orientations of the key journals in LIS and select appropriate journals for their research papers; second, that research methods are essential for research studies in a field. The investigation results from the key journals in LIS assist researchers in gaining insight into current status, patterns, and trends of statistical research method uses in the field; and finally, that the findings of the study provide readers with information on the relationships between statistical research methods and their application areas. It may help them select a proper statistical research method for a specific research topic.

The limitations of this study are: first, that only the research papers published in the six journals during the past 15 years were examined. If more journals and issues were investigated, the results obtained would be more reasonable and reliable; and second, that all the investigated research papers in this study are the ones that apply statistical methods. In other words, all the conclusions were in a statistical context. One of the future research directions may be to investigate the impact of statistically based papers vs non-statistically based papers over a certain period of time in LIS.

References

Aabø, S., Audunson, R. and Vårheim, A. (2010), "How do public libraries function as meeting places?", *Library & Information Science Research*, Vol. 32 No. 1, pp. 16-26.

Aharony, N. (2012), "Library and information science research areas: a content analysis of articles from the top 10 journals 2007-2008", *Journal of Librarianship and Information Science*, Vol. 44 No. 1, pp. 27-35.

Bae, Y. and Lee, H. (2012), "Sentiment analysis of twitter audiences: measuring the positive or negative influence of popular twitters", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 12, pp. 2521-2535.

Bernard, H.R. (2011), *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, Rowman Altamira, Lanham.

Blessinger, K. and Frasier, M. (2007), "Analysis of a decade in library literature: 1994-2004", *College & Research Libraries*, Vol. 68 No. 2, pp. 155-169.

Bornmann, L., de Moya Anegón, F. and Mutz, R. (2013), "Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings?", *Journal of the American Society for Information Science and Technology*, Vol. 64 No. 11, pp. 2310-2316.

Busha, C.H. and Harter, S.P. (1980), *Research Methods in Librarianship: Techniques and Interpretation*, Academic Press, New York, NY.

Chung, W. and Chen, H. (2009), "Browsing the underdeveloped web: an experiment on the Arabic medical web directory", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 3, pp. 595-607.

Connaway, L.S. and Powell, R.R. (2010), *Basic Research Methods for Librarians*, 5th ed., Libraries Unlimited, Santa Barbara, CA.

Cortese, J. and Lustria, M.L.A. (2012), "Can tailoring increase elaboration of health messages delivered via an adaptive educational site on adolescent sexual health and decision making?", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 8, pp. 1567-1580.

Davarpanah, M.R. and Aslekia, S. (2008), "A scientometric analysis of international LIS journals: productivity and characteristics", *Scientometrics*, Vol. 77 No. 1, pp. 21-39.

Gravetter, F.J. and Wallnau, L.B. (2009), *Statistics for the Behavioral Sciences*, Wadsworth Cengage Learning, Belmont, CA.

Heinström, J. (2005), "Fast surfing, broad scanning and deep diving", *Journal of Documentation*, Vol. 61 No. 2, pp. 228-247.

Hider, P. (2008), "Using the contingent valuation method for dollar valuations of library services", *Library Quarterly*, Vol. 78 No. 4, pp. 437-458.

Hider, P. and Pymm, B. (2008), "Empirical research methods reported in high-profile LIS journal literature", *Library & Information Science Research*, Vol. 30 No. 2, pp. 108-114.

Hildreth, C.R. and Aytac, S. (2007), "Recent library practitioner research: a methodological analysis and critique", *Journal of Education for Library & Information Science*, Vol. 48 No. 3, pp. 236-258.

Hodge, G.M. (2000), "Best practices for digital archiving: an information life cycle approach", *D-Lib Magazine*, Vol. 6 No. 1.

Jarvelin, K. and Vakkari, P. (1990), "Content analysis of research articles in library and information science", *Library and Information Science Research*, Vol. 29 No. 4, pp. 395-421.

Järvelin, K. and Vakkari, P. (1993), "The evolution of library and information science 1965-1985: a content analysis of journal articles", *Information Processing & Management*, Vol. 29 No. 1, pp. 129-144.

Kang, M., Kim, Y.G. and Bock, G.W. (2010), "Identifying different antecedents for closed vs open knowledge transfer", *Journal of Information Science*, Vol. 36 No. 5, pp. 585-602.

Kothari, C.R. (2004), *Research Methodology: Methods and Techniques*, New Age International, New Delhi.

Koufogiannakis, D., Slater, L. and Crumley, E. (2004), "A content analysis of librarianship research", *Journal of Information Science*, Vol. 30 No. 3, pp. 227-239.

Kwon, N. and Zweizig, D.L. (2006), "Use of community information and communication technologies (ICTs): explaining the use of community networks with demographic factors, psychological factors, and alternative service accessibility", *Library Quarterly*, Vol. 76 No. 1, pp. 81-106.

Lee, C. and Kim, H. (2013), "Automatic Korean word spacing using Pegasos algorithm", *Information Processing & Management*, Vol. 49 No. 1, pp. 370-379.

Lee, K.S. and Croft, W.B. (2013), "A deterministic resampling method using overlapping document clusters for pseudo-relevance feedback", *Information Processing & Management*, Vol. 49 No. 4, pp. 792-806.

Leydesdorff, L., Bornmann, L., Mutz, R. and Opthof, T. (2011), "Turning the tables on citation analysis one more time: principles for comparing sets of documents", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 7, pp. 1370-1381.

Li, R. (2009), "The representation of national political freedom on web interface design: the indicators", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 6, pp. 1222-1248.

Manzari, L. (2013), "Library and information science journal prestige as assessed by library and information science faculty", *Library Quarterly*, Vol. 83 No. 1, pp. 42-60.

Martyn, J. and Lancaster, F.W. (1981), *Investigative Methods in Library and Information Science: An Introduction*, Information Resources Press, Arlington, TX.

Medhi, J. (1992), *Statistical Methods: An Introductory Text*, Wiley, New York, NY.

Nicholas, D., Huntington, P. and Rowlands, I. (2005), "Open access journal publishing: the views of some of the world's senior authors", *Journal of Documentation*, Vol. 61 No. 4, pp. 497-519.

Nisonger, T.E. and Davis, C.H. (2005), "The perception of library and information science journals by LIS education deans and ARL library directors: a replication of the Kohl-Davis study", *College & Research Libraries*, Vol. 66 No. 4, pp. 341-377.

Ouyang, Y., Li, W., Zhang, R., Li, S. and Lu, Q. (2013), "A progressive sentence selection strategy for document summarization", *Information Processing & Management*, Vol. 49 No. 1, pp. 213-221.

Ozmutlu, H.C., Spink, A. and Ozmutlu, S. (2002), "Analysis of large data logs: an application of Poisson sampling on excite web queries", *Information Processing & Management*, Vol. 38 No. 4, pp. 473-490.

Perry, J.L. and Kraemer, K.L. (1986), "Research methodology in the 'public administration review', 1975-1984", *Public Administration Review*, Vol. 46 No. 3, pp. 215-226.

Pettigrew, K.E. and McKechnie, L.(E.F.) (2001), "The use of theory in information science research", *Journal of the American Society for Information Science and Technology*, Vol. 52 No. 1, pp. 62-73.

Powell, R.R. (1997), *Basic Research Methods for Librarians*, Greenwood Publishing Group, Santa Barbara, CA.

Qin, T., Zhang, X.D., Tsai, M.F., Wang, D.S., Liu, T.Y. and Li, H. (2008), "Query-level loss functions for information retrieval", *Information Processing & Management*, Vol. 44 No. 2, pp. 838-855.

Rajasekar, S., Philominathan, P. and Chinnathambi, V. (2006), "Research methodology", available at: http://arxiv.org/abs/physics/0601009 (accessed January 12, 2015).

Rochester, M.K. and Vakkari, P. (2003), "International library and information science research: a comparison of national trends", Professional Reports No. 82, IFLA, The Hague.

Scandura, T.A. and Williams, E.A. (2000), "Research methodology in management: current practices, trends, and implications for future research", *Academy of Management Journal*, Vol. 43 No. 6, pp. 1248-1264.

Schedl, M., Widmer, G., Knees, P. and Pohle, T. (2011), "A music information system automatically generated via web content mining techniques", *Information Processing & Management*, Vol. 47 No. 3, pp. 426-439.

Somekh, B. and Lewin, C. (2005), *Research Methods in the Social Sciences*, Allyn & Bacon, Needham Heights, MA.

Stvilia, B. and Jörgensen, C. (2009), "User-generated collection-level metadata in an online photo-sharing system", *Library & Information Science Research*, Vol. 31 No. 1, pp. 54-65.

Sugimoto, C.R., Li, D., Russell, T.G., Finlay, S.C. and Ding, Y. (2011), "The shifting sands of disciplinary development: analyzing North American library and information science dissertations using latent Dirichlet allocation", *Journal of the American Society for Information Science and Technology*, Vol. 62 No. 1, pp. 185-204.

Surman, R. and Bath, P.A. (2013), "An assessment of the quality of information on stroke and speech and language difficulty web sites", *Journal of Information Science*, Vol. 39 No. 1, pp. 113-125.

Tashakkori, A. and Teddlie, C. (2003), *Handbook of Mixed Methods in Social & Behavioral Research*, Sage Publications, Thousand Oaks, CA.

Thelwall, M., Wilkinson, D. and Uppal, S. (2010), "Data mining emotion in social network communication: gender differences in MySpace", *Journal of the American Society for Information Science and Technology*, Vol. 61 No. 1, pp. 190-199.

Thomson Reuters (2013), "2012 Journal Citation Reports®", available at: http://thomsonreuters.com/journal-citation-reports/ (accessed January 5, 2015).

Tuomaala, O., Järvelin, K. and Vakkari, P. (2014), "Evolution of library and information science, 1965-2005: content analysis of journal articles", *Journal of the Association for Information Science and Technology*, Vol. 65 No. 7, pp. 1446-1462.

Wallace, D.P. (1985), "The use of statistical methods in library and information science", *Journal of the American Society for Information Science*, Vol. 36 No. 6, pp. 402-410.

Wei, C.Y., Evans, M.B., Eliot, M., Barrick, J., Maust, B. and Spyridakis, J.H. (2005), "Influencing web-browsing behavior with intriguing and informative hyperlink wording", *Journal of Information Science*, Vol. 31 No. 5, pp. 433-445.

Williams, J.F. and Winston, M.D. (2003), "Leadership competencies and the importance of research methods and statistical analysis in decision making and research and publication: a study of citation patterns", *Library & Information Science Research*, Vol. 25 No. 4, pp. 387-402.

Zhao, S.X., Tan, A.M. and Ye, F.Y. (2012), "Distributive h-indices for measuring multilevel impact", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 10, pp. 2074-2086.

Corresponding author
Jin Zhang can be contacted at: jzhang@uwm.edu