Evidence Based Library and Information Practice

Quantitative Methods and Inferential Statistics: Capacity and Development for Librarians

Lise Doucette

Aller au sommaire du numéro

Découvrir la revue

Evidence Based Library and Information Practice

Commentary

Quantitative Methods and Inferential Statistics: Capacity and Development for Librarians

Lise Doucette
Assessment Librarian
Western Libraries
Western University
London, Ontario, Canada
Email: ldoucet@uwo.ca

Introduction

"Librarianship and statistics have always had an uneasy relationship." So begins Dilevko's (2007, p. 209) article investigating the extent to which inferential statistics are used in journals read by academic and public librarians. Uneasiness is an interesting and apt word. In informal conversations locally and at conferences, I have heard librarians acknowledge that statistics may in fact be useful, but then liberally use adjectives such as *intimidating* and *boring*. Personally, I love math and statistics, perhaps to what others might consider an unreasonable or evangelical degree. I do not expect all librarians to become devoted mathophiles (though I would welcome that situation). However, I do contend that increasing our understanding of statistics individually and collectively will lead to better research, better evidence, better assessment, and better library and information practice.

In this paper, I will discuss my observations of the current relationship between academic librarianship and statistics, the utility of and case for statistics, and a number of different ways to learn more about statistics. My presentation at the 2016 C-EBLIP Fall Symposium (upon which this paper is based) was descriptive and subjective in nature, and I did not discuss the wide range of research that relates to the use of statistics by librarians. To round out my personal reflections with more academic work, I will add an analysis of that research in the next section of this paper.

I would like to introduce a few terms and definitions. *Quantitative methods* are used to analyze phenomena that can be measured and expressed in numerical format. Examples of *quantitative data* are type of student or researcher (nominal data), a ranking of library services by students (ordinal data), a rating of satisfaction

with the library on a scale from 1-10 (interval data), and a student's score on an assignment (ratio data). *Descriptive statistics* describe or summarize the data and could include a simple table of numbers of each type of student or researcher, the number of students ranking each library service most highly, range of satisfaction scores, and average of students' assignment scores. *Inferential statistics* go much further, and they allow researchers to test hypotheses about relationships among the data and to make conclusions based on statistical evidence. Questions that inferential statistics might help answer include: Do students who receive information literacy sessions score higher on assignments and by how much? Do graduate students who publish in open access journals have supervisors who also publish in open access journals?

Current Relationship Between Librarianship and Statistics

A number of factors contribute to the uneasy relationship between librarianship and statistics. Students arriving in MLIS programs are generally coming from a humanities background and have little previous experience with statistics. LIS researchers (Stephenson, 1990; Dilevko, 2000; Park, 2003) have found that while the majority of MLIS programs do have required Research Methods courses, most of those courses do not involve a strong statistical analysis component or an applied research project. Statistics anxiety is also an issue and has been extensively studied in psychology, education, and statistics. A study of graduate students in an educational research methods course found that statistics anxiety is one of four key forms of anxiety experienced during the research proposal writing process (Onwuegbuzie, 1997). Although the author does not make any explicit connection to students in MLIS programs, the background of the graduate students and the types of projects undertaken in the educational research methods course are quite similar to those of MLIS students or librarians.

Published research in librarianship does not use inferential statistics to a great degree, and a number of studies address differences based on type of statistics (inferential or descriptive) and role (LIS faculty and practising librarians). Dilevko (2007) found that there was an increase in the use of inferential statistics in his study period of 2001-2005 (14.5% of articles in the journals he studied) as compared to earlier studies from the 1970s and 1980s, where researchers found that 0.5-13.3% of articles used inferential statistics. Dilevko also found that 38.5% of articles use only descriptive statistics, with 46.3% of articles using no statistics.

In 1999, Hernon (then editor of *The Journal of Academic Librarianship*) reflected on research in LIS based on the manuscripts he receives and published literature generally. Among the concerns he expressed about librarianship research are failure to select a confidence interval to guide data interpretation for inferential statistics methods, misuse or limited use of statistics, and inappropriate or incorrect use of statistical language (in particular, *significance*). He also quotes an earlier article (Hernon, Smith, & Coxen, 1993) reviewing ten years of *College & Research Libraries* manuscripts that names "poor use of statistical methods" as one issue with submitted and published research.

I also see these issues in published papers. Given the uneasiness regarding statistics in our discipline, I wonder if reviewers of papers are themselves knowledgeable enough about statistical methods such that they can critically review papers that use inferential statistics. Within librarianship, there is neither a strong culture of critical discussion of research nor a culture of reproducibility and replicability. As an example of a journal in a discipline with such a culture, the *American Journal of Political Science* requires that authors of quantitative papers submit their data as well as the code used to analyze the data. After the regular peer review process, the analysis and results of each conditionally accepted paper are then

independently verified before publication. Developing a strong culture of critically evaluating others' statistical work will strengthen our research and our ability to have good research conversations.

Why Quantitative Methods and Inferential Statistics?

Quantitative methods have long been essential to social sciences research. Research by librarians uses a variety of humanities and social sciences methodologies, and the evidence used in evidence based practice takes many forms. Quantitative methods often complement qualitative methods. Understanding quantitative methods allows librarians to expand their capacity to develop and answer research questions and develop evidence for informing practice, and it also allows them to read, understand, and critically evaluate research results and evidence created by others.

Curiosity is key to evidence based practice and research, and learning a different way of understanding and measuring phenomena can expand your ability to think about all of the interactions in the world around you. Inferential statistics are used to study differences or variance and to explore factors causing that variance. Do students who take a library module on academic integrity change their citing behaviours in future essays? How do their behaviours differ from those who do not take the library module? What factors influence physical library usage? (Qualitative methods might help answer more of the *why* questions and allow for a deeper understanding of, for example, why faculty publish in open access journals.) Inferential statistics also allow for exploration of the degree of difference, the confidence that there is in fact a difference (from a mathematical not a personal perspective), the factors that might be influencing the measurement (such as interactions between different variables), and the ability to which a generalization (inference) or prediction can be made about certain research results. Descriptive

statistics provide a useful overview of your data but can only summarize your results.

Quantitative methods also give you a new language (shared with researchers around the world) to be able to describe phenomena appropriately and to draw appropriate conclusions. Byrne (2007) highlights an example of an apparent difference between two groups when looking at descriptive statistics and then shows that when an inferential statistics test is applied, there is in fact no statistically significant difference. Hernon (1999) highlights librarians' tendencies to use the statistical terms *significant* and *not significant* without applying statistical tests.

Generally, rigorous quantitative methods should be both reliable (consistently reproducible within the sample you choose) and valid (measuring what you say you are measuring). These standards require a great deal of critical thinking and planning and are reliant as much on good research design as they are on appropriate statistical analysis.

Learning About Statistics

"We know accurately only when we know little; doubt grows with knowledge." —Goethe

The above quote captures the joy and frustration of learning for me—the more I learn about statistics, the more I realize how much I really don't know, and the more motivated I am to keep learning. A recent article (Berg and Banks, 2016) highlighted librarians' capacity to grow and evolve as researchers, advocating for a shift away from identifying and attempting to achieve specific research competencies. This resonated with me, as I do not think it would be helpful to have a list of specific statistical tests or statistical knowledge that all librarians should know. I believe librarians will explore statistics as interest, research responsibilities, and professional practice requires, and I certainly agree with Berg and Banks that librarians have a great capacity for learning.

It can be difficult to identify where your current knowledge of quantitate methods fits in when the landscape of the topic is difficult to identify. Additionally, statistics anxiety is very real and can be a barrier to learning. In my experience, there is no shortcut for understanding quantitative methods. I know of a number of librarian researchers who are currently undertaking research projects that require quantitative methods—some have been learning how do to so on their own, and others are working with librarian or other university colleagues who already have this knowledge. (In 2014, I worked with an educational researcher at my university to help me refresh my statistics knowledge.) There are many different ways to learn depending on your existing knowledge, your available time, and what you want to learn. I'll also point out a few examples of what I've done to further my knowledge.

Self-Directed Learning

If you are interested in learning on your own, these options may work well for you:

- Read research articles (you are already doing this!) and deliberately include articles that use quantitative methods. Read their research questions and methodologies carefully, and look up new terms.
- Read introductory articles, such as Gillian Byrne's introduction to statistics (https://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/168). Many universities (and of course Wikipedia!) also have good explanations of many statistical topics. Try a quick search for *introduction regression* and add in *video* if that is your preferred learning method.

- Read (or skim) a research methods or statistics text, such as the following:

  - *Statistics for People Who (Think They) Hate Statistics* by Neil Salkind (includes the chapter "Statistics or Sadistics? It's Up to You!")
  - *Statistical Methods for the Information Professional: A Practical, Painless Approach to Understanding, Using, and Interpreting Statistics* by Liwen Vaughan
  - *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* by John W. Creswell

Structured Courses

If you are looking for more structure, some of these options may work well:

- Look at your university's undergraduate statistics courses in disciplines such as psychology and sociology. Talk to the professor about sitting in on lectures, auditing, or taking the course for credit. (I sat in on the lectures for Western's PSYC3800: Psychological Statistics Using Computers).
- Consider short courses on statistics or statistical software. Your university's statistics department may offer some of these or may be willing to if you express interest. (I took a full-day course in the statistical software R and a half-day review called "Crash Course in Introductory Inferential Statistics.")
- Look at individual courses or sequences of online courses. For example, Coursera has many courses and certificates in inferential statistics, research methods, and data science.

- Consider longer, focused programs (try searching for *summer institute statistics*), such as these courses:

  o The Inter-university Consortium for Political and Social Research (ICPSR) has a Summer Program in Quantitative Methods for Social Research, which is a four- or eight-week program that includes introductory and advanced statistics courses, computer software and math courses, and evening research lectures. (I attended the 2016 eight-week program with courses in regression, categorical data analysis, data management, and various software applications.)
  o ICPSR and many universities offer one-week introductory and advanced courses.

The important thing with any of this is to apply what you are learning. Think of related research applications and try analysing some of your own data. If possible, try this with colleagues; the mutual support and ability to discuss and ask questions will be beneficial. There may be times when you need more substantial support. Find a colleague who knows more about the topic, look at consultancy options at your university (many statistics departments offer this service), or search for a published paper that uses a similar method and contact the authors. Remember that you are the person who cares the most about your data and your research; external support is great but at the same time, you want to ensure that you understand the analysis and would be able to answer questions at a conference presentation.

Conclusions

We all have a limited amount of time in our professional lives, with different priorities and areas of focus. I certainly understand that increasing knowledge of quantitative methods will not be of interest to everyone. However, I

would challenge you to consider the benefits of including or increasing quantitative methods in your own research and practice and to deliberately take on one small learning opportunity (personally or perhaps with colleagues).

By collectively broadening our knowledge of certain types of methodologies, we broaden the types of research questions we can conceive of and address. While there are methods to increase your own knowledge, there may also be larger systemic structures or solutions within MLIS programs or for practising librarians. Our profession has more exploring to do of how and why librarians do not often use inferential statistics; if this is a priority for our community, we can investigate ways to enact change.

References

American Journal of Political Science. (n.d.). Guidelines for accepted articles. In *American Journal of Political Science*. Retrieved 22 January 2016 from https://ajps.org/guidelines-for-accepted-articles/

Berg, S. A., & Banks, M. (2016). Beyond competencies: Naming librarians' capacity for research. *The Journal of Academic Librarianship*, *42*(4), 469-471. http://dx.doi.org/10.1016/j.acalib.2016.06.002

Byrne, G. (2007). A statistical primer: Understanding descriptive and inferential statistics. *Evidence Based Library and Information Practice*, *2*(1), 32-47. http://dx.doi.org/10.18438/B8FW2H

Dilevko, J. (2000). A new approach to teaching research methods courses in LIS programs. *Journal of Education for Library and Information Science*, *41*(4), 307-329. http://dx.doi.org/10.2307/40324048

Dilevko, J. (2007). Inferential statistics and librarianship. *Library & Information Science Research*, *29*(2), 209-229. http://dx.doi.org/10.1016/j.lisr.2007.04.003

Hernon, P. (1999). Editorial: Research in library and information science—Reflections on the journal literature. *Journal of Academic Librarianship*, *25*(4), 263-266. http://dx.doi.org/10.1016/S0099-1333(99)80025-1

Hernon, P., Smith, A., & Croxen, M. B. (1993). Publication in *College & Research Libraries*: Accepted, rejected and published papers, 1980-1991. *College & Research Libraries*, *54*(4), 303-321. http://dx.doi.org/10.5860/crl_54_04_303

Onwuegbuzie, A. J. (1997). Writing a research proposal: The role of library anxiety, statistics anxiety, and composition anxiety. *Library & Information Science Research*, *19*(1), 5-33. http://dx.doi.org/10.1016/S0740-8188(97)90003-7

Park, S. (2003). Research methods as a core competency. *Journal of Education for Library and Information Science*, *44*(1), 17-25. http://dx.doi.org/10.2307/40323939

Stephenson, M. S. (1990). Teaching research methods in library and information studies programs. *Journal of Education for Library and Information Science*, *31*(1), 49-65. http://dx.doi.org/10.2307/40323727