

University at Albany, State University of New York

## Scholars Archive

---

University Libraries Faculty Scholarship

University Libraries

---

2-14-2020

# Citation Analysis of Mathematics and Statistics Dissertations and Theses from the University at Albany

Kathleen H. Flynn

University at Albany, State University of New York, [kflynn@albany.edu](mailto:kflynn@albany.edu)

The University at Albany community has made this article openly available.

**Please share** how this access benefits you.

Follow this and additional works at: [https://scholarsarchive.library.albany.edu/ulib\\_fac\\_scholar](https://scholarsarchive.library.albany.edu/ulib_fac_scholar)



Part of the [Collection Development and Management Commons](#)

---

### Recommended Citation

Flynn, Kathleen H., "Citation Analysis of Mathematics and Statistics Dissertations and Theses from the University at Albany" (2020). *University Libraries Faculty Scholarship*. 157.

[https://scholarsarchive.library.albany.edu/ulib\\_fac\\_scholar/157](https://scholarsarchive.library.albany.edu/ulib_fac_scholar/157)

### License

This Article is brought to you for free and open access by the University Libraries at Scholars Archive. It has been accepted for inclusion in University Libraries Faculty Scholarship by an authorized administrator of Scholars Archive.

Please see [Terms of Use](#). For more information, please contact [scholarsarchive@albany.edu](mailto:scholarsarchive@albany.edu).

## **Citation Analysis of Mathematics and Statistics Dissertations and Theses from the University at Albany**

This study analyzed dissertations and theses completed from 2009-2019 in the Mathematics and Statistics department at the University at Albany to investigate resource type, citation age, publishers, and journal title dispersion. Students cited journal articles 57.0% of the time, while books and book chapters combined accounted for 30.5% of the citations. Publisher usage fit the Pareto Distribution; however, there was a high journal title dispersion. Preprints represented 4.6% of the citations, an increase from previous studies. The amount of graduate student citation quality issues and preprint citations shows a need for instruction and discussions in these areas.

Keywords: citation analysis, mathematics, statistics, dissertations, theses

### **Introduction**

As library collections grow, it can become increasingly difficult to gauge how appropriate and thorough they are for the library's patrons and their needs. When a library's print periodical collection shrinks and online journal subscription prices increase yearly, it creates a challenge to ensure the budget supports both journals and monographs. To further complicate matters, many faculty and students may be less dependent on books compared to journals, depending on the subject area. A common assumption is that researchers in the sciences rely more on journals, while the humanities prefer monographs. One method of evaluating a collection is a citation analysis, which allows the librarian to analyze the resources that faculty or students are citing in their work. This information can reveal the format of resources preferred, popular journals, the ages of resources that patrons cite, and more.

With this data, the librarian can prioritize subscribing to well-used journals, especially if the librarian may have dismissed them as non-essential based on bibliometrics, such as journal

impact factor. Additionally, traditional usage statistics, such as the number of full-text downloads, do not always correspond to the local citation analysis counts. Faculty may download articles once and cite them in multiple papers. Furthermore, they may give previously downloaded articles to their students for them to cite. Without conducting a citation analysis, a journal with low usage numbers may be cancelled despite being well-cited by graduate students and faculty. Therefore, using a combination of collection development methods is preferable to assuming what resources the community needs and wants.

The University at Albany, part of the State University of New York (SUNY) system, offers over 150 graduate programs in nine schools and colleges and has an enrollment of over 17,000 students. The Mathematics and Statistics department, known as the Mathematics department until 2012, currently has over 100 graduate students enrolled. As of 2019, the department has 24 faculty members engaged in research focusing on algebra, analysis, combinatorics, geometry, topology, and probability and statistics. In an effort to better understand the collection and the needs of these faculty and graduate students, the author, as the new mathematics librarian, sought information about the resources these graduate students used in their dissertations and theses via a citation analysis.

## **Literature Review**

Gross and Gross (1927) noted that since a list of journals a librarian deems essential would reflect the librarian's bias, a better strategy is to evaluate which resources are actually being used by the library's community. Their study, seen as the first citation analysis of this type, consisted of an analysis of references in one volume of one well-regarded chemistry journal. Two years later, Allen (1929) published a similar article, this time evaluating citations in issues from nine mathematical journals from the previous year. Unlike Gross and Gross, Allen decided that the

mathematics field did not have one journal that dominated or represented all areas. Hence, nine journals were needed. Allen cautioned against basing collection decisions solely on these findings as the research areas of faculty will influence which journals are considered core journals. Furthermore, some journals feature longer articles than others, so a patron's citations to them may be less frequent, but more substantial (Allen 1929).

Johnson's (1996) two-year citation analysis of statistics faculty publications at Texas Tech University involved 394 citations. His study, conducted during the earlier years of online resources, showed that 46.7% of faculty citations were to articles, while 36.9% were to books. The average age of a citation, calculated by subtracting the resource publication date from the citation date, can show when resources become obsolete for patrons. In his study, the average age was 12.3 years. That information was used to purchase the resources that received the most use.

Sinn (2005) conducted an analysis of Bowling Green State University dissertation citations in mathematics and statistics from 1980-2002. Her study was prompted by a faculty member's assertion that mathematical journals should not be cancelled due to low usage because the field lacks a core few journals like many of the other sciences. Sinn's study analyzed 2,145 citations from 67 dissertations and discovered that articles accounted for 65.9% of the citations, while books were 27.0%. She noted that a study of mathematics and statistics citations should be repeated in later years to measure the impact of MathSciNet, a mathematics database, and the increase in the number of online journals.

In recent years there have been other studies that have shown similar results. Dotson and Franks' (2015) study of mathematics dissertations from 1998-2012 revealed that citations to peer-reviewed articles and books were 58.1% and 27.1% respectively. Additionally, Kelly's

(2015) analysis of a sample of statistics dissertations from 2005-2012 showed that students cited articles 59.0% of the time and books/book chapters 25.0% of the time.

Sinn's (2005) study also evaluated the journal title dispersion, or what percentage of the journal titles were needed to account for 80.0% of the article citations, and concluded that when compared to the other sciences, a larger percentage of journal titles were needed to account for 80.0% of the article citations. This calculation, commonly known as the 80/20 rule, is infrequently reported in these types of studies (Sinn 2005).

Coined by Vilfredo Pareto, a political economist and sociologist in the early 1900s, the 80/20 rule is frequently called the Pareto Distribution or Pareto Principle (Nisonger 2008). Trueswell's study from 1969 (as cited in Nisonger 2008) was one of the earliest depictions of the rule in libraries as it showed that only 20.0% of the collection was needed to account for 80.0% of circulations. Nisonger (2008) found 19 studies where the rule was roughly accurate for print serials, but the percentage of "core" titles was as low as 10.0% in one study and as high as 47.0% in another. Studies on database usage have revealed more varied results. Nisonger also compiled a list of studies pertaining to citations to journals. The results found that some of the studies adhered to the rule, but there were three studies of psychology theses where 50.0%, 47.6%, and a whopping 80.0% of journals were needed to account for 80.0% of the citations. If more than 20.0% of the titles are required for 80.0% of the citations, this is considered a high title dispersion. However, focusing on the top 20.0% alone can have drawbacks. It has been stressed that libraries do not just serve those who wish to use the core 20.0% of journals, but that they also serve niche communities who will appreciate access to the other 80.0% (Nisonger 2008).

Since the earlier mathematical citation analyses were completed, there has been an increased adoption in the mathematical community of the eprint database known as arXiv. Much

of the material in arXiv consists of preprints. A preprint is a version of a document prior to peer-review and publication in a journal. Kurtz et. al (2005), speculated that there are three reasons for researchers publishing their work as a preprint: they believe it will be cited more if there is free access to it, they believe rapid dissemination will result in more citations, and that they only choose to publish their best work as preprints. The arXiv database, run by Cornell University, has surged in popularity since its inception in 1991 (arXiv 2019a), with 140,000 submissions in 2018 (arXiv 2019c) and over 14,000 in October of 2019 alone (arXiv 2019b). It is used primarily by physicists, mathematicians, and computer scientists, and the number of submissions per year for the mathematics and mathematical physics categories have increased since 2002 (arXiv 2019c). According to arXiv (2019c), those categories accounted for 23.8% of arXiv submissions for 2018, the third highest percentage. However, despite the fact that preprints have not been peer-reviewed yet, some authors are citing them.

A study by Noruzi (2016) investigated the number of citations in Scopus to arXiv publications from 1990 to 2017. During that period, there were 49,915 citations in mathematics, the second highest total. With the increase in preprint popularity in mathematics, one wonders how often faculty and students use and cite them, particularly as students may not be able to evaluate their quality as well as their professors. This question and the following were the focus of the present study.

### *Questions*

- Which resource types are graduate students using?
- Is there a higher journal title dispersion in mathematics and statistics compared to the other sciences?
- How popular are preprints among graduate students in mathematics and statistics?

- How often do graduate students cite their advisors?
- How many of the cited journals does the University at Albany provide access to?
- Which publishers are most represented, and are there differences by resource type?
- How old are the citations, and are there differences by resource type?

## **Methods**

The University at Albany has access to ProQuest's Dissertations and Theses @ SUNY Albany database. The database was used to identify 48 dissertations and theses authored by University at Albany Mathematics and Statistics Department (formerly the Mathematics Department) graduate students from 2009 to 2019. As those 48 papers were the entirety of the papers from those departments in the database as of October 2019, they were all selected for the study. The metadata were exported to an Excel spreadsheet and the papers themselves were stored in Zotero. The metadata provided by the database included some redundant or non-useful columns, which were removed. The original Excel spreadsheet was preserved separately. Furthermore, several columns were added to the edited sheet, including degree date, number of pages, advisor, department, degree, number of references, number of references to advisor, and ID. The IDs were assigned one through 48.

A second Excel spreadsheet was created to contain the reference data. Each reference was assigned a reference ID (RefID) which related to the paper ID (ID). For each reference the following information was recorded in this sheet: RefID, ID, resource type, publication year, journal title, publisher, citation age in years, and dissertation publication year. Publishers were identified by locating the journal's homepage or the book's page at the publisher's website. The citation age was calculated by subtracting the reference's publication date from the dissertation, or thesis, publication date. If a reference did not have one of the elements, an NA was entered in

that field. The resource type field consisted of the following categories: article, book, book chapter, preprint, web, dissertation, conference proceedings, thesis, and other.

Occasionally, a preprint, particularly those found in arXiv, will include a note such as “accepted for publication in *journal*” or “to appear in *journal*” to indicate that the paper has been accepted for publication but has not yet been published. If the student’s citation included this information, the reference was simply recorded as a regular article and not a preprint. The “other” category included references such as duplicates, software documentation, private communication, and incomplete citations. A reference was deemed incomplete when the resource was unable to be located at all. The Web category contained resources that could *only* be found online and were not part of a journal or a book. These included online lecture notes, websites, news articles, reports, blog entries, slideshows, and technical reports.

Whenever possible, the information was taken from the student’s citation. However, some citations were either slightly incomplete or incorrect, making it challenging to locate the resource. In these instances, the correct or missing data was entered, if found. Since one of the objectives of the study was to see which publishers currently have the rights to each resource, the current publisher was recorded, when possible. Subsidiaries were recorded as the parent company, e.g. “Springer.”

The completed Excel spreadsheets were then converted to csv files and loaded into R, where data analysis took place. Packages used in R for the analysis included sqldf and plyr. The tables were also created in R before being exported to a csv file.



## Results and Discussion

### *Resource Types*

The 48 student papers consisted of 43 PhD dissertations and 5 master's theses. In total, there were 1,249 citations with an average of 26.0 per paper. This is slightly fewer than the 32.0 citations per paper that Sinn (2005) found. One paper had zero citations while the maximum number of citations found in a paper was 150. The number of pages per paper ranged from a minimum of 29 to a maximum of 145, with an average of 67.3.

The most used resource type was articles with 57.0% of the citations, followed by books and book chapters with a combined 30.5% of the citations, as shown in Table 1. This 30.5% book usage percentage is similar to those found in other studies which evaluated dissertations published as recently as 2012, as shown in Table 2. However, the 57.0% article usage is similar to recent studies with the exception of the 65.9% Sinn (2005) found in dissertations published as recently as 2002. The increase in preprint and web resource usage in the present study compared to Sinn's may account for some of the difference.

These numbers continue to match Sinn's (2005) findings that mathematics and statistics papers cite articles less often than biology and chemistry. While this study found 57.0% article citations, Barnett-Ellis and Tang's (2016) study of biology master's theses found 75.0% article citations. Flaxbart's (2018) 50 chemistry dissertations included 91.5% article citations. The study from Cole et al. (2018) investigating engineering dissertation citations from seven decades found that the percentage of article citations varied over time from over 50.0% to over 70.0%.

**[Insert Table 1 here]**

**[Insert Table 2 here]**

### *Citation Age*

The average age of all citations was 19.9 years, with a minimum of zero years and a maximum of 115 years, as shown in Table 3. Articles and books had similar average citation ages of 20.6 years and 22.2 years, respectively. Preprints and web resources were younger than articles and books, with average citation ages of 8.4 years and 6.0 years, respectively. The article and book ages were older than Sinn's (2005) 16.6 years for articles and 14.5 years for books. They were also both older than Johnson's (1996) 12.3 years and Kayongo and Helm's (2012) 14.9 years, which were the average ages of all mathematics citations from their studies. Since the citation ages from this study were higher than previous studies, it is possible that the average mathematics and statistics citation age is increasing. Alternatively, the collections available to students, the focus of the student's field, and their advisor's input may also be influencing the ages.

Biology master's theses appear to cite younger resources, with an average resource age of 16.0 years old (Barnett-Ellis and Tang 2016). Flaxbart's (2018) study of chemistry dissertations selected from five years between 2006-2015 showed that the median age of books increased over time from 10.0 to 14.5 years old, and the Pareto distribution was 21 years old. Kayongo and Helm (2012) discovered the average age of all resources for the sciences and mathematics was 10.8 years.

**[Insert Table 3 here]**

### *Preprint Use*

Table 1 shows that 4.6% of the citations were to preprints while web resources were 3.4%. These are higher than Sinn's (2005) 2.6% for prepublications and 0.0% for web resources. The years studied may explain part of these differences as this study was for papers from 2009-2019,

while Sinn (2005) analyzed papers from 1980-2002. Dotson and Franks' (2015) mathematics dissertations from 1998-2012 found preprints were cited 4.2% of the time, but they were frequently the third most cited category in each year studied. While preprints existed prior to arXiv, it was the source of nearly all the cited preprints in the present study. Given that preprints have not been peer-reviewed, it is worth investigating their citation patterns within the discipline.

Lariviere et al.'s (2014) analysis of the Web of Science (WoS) discovered that the percentage of arXiv papers that were eventually published and indexed by WoS was 64.0%; however, the percentage was 45.0% for mathematics, quantitative finance, and statistics papers. The authors also found that in 2011, 1.0% of WoS cited references were to mathematics preprints. Ferrer-Sapena et al. (2018) found that 67.5% of arXiv papers were eventually published in a WoS journal, but the number was 64.3% for mathematics, statistics, and nonlinear sciences, higher than the 45.0% Lariviere et al. (2014) discovered. Additionally, there was an average of 2.5 years between the publication of a mathematics, statistics, or nonlinear science arXiv preprint and publication in a WoS journal (Ferrer-Sapena et al. 2018).

The increase in arXiv preprints could also be a response to funding requirements that publications be made freely accessible. A study that investigated author compliance with the open access mandates of 12 funding agencies found that results varied by funding agencies and discipline (Larivière and Sugimoto 2018). At best, roughly 90.0% of U.S. National Institutes of Health (NIH) funded research had open access copies available, while at the low end, 23.0% of publications funded by the Social Sciences and Humanities Research Council of Canada were available for free (Larivière and Sugimoto 2018). Prior to 2016, when the National Science Foundation (NSF) only recommended researchers make their research freely available, 47.0% of their funded research was made open access (Larivière and Sugimoto 2018). Since it is now a

requirement, it could explain part of the increase in papers deposited in arXiv. These free versions may then be more easily discoverable to the students and faculty who go on to cite them.

In this study, preprints had an average age of 8.4 years, but the youngest was zero, the year of the student's dissertation or thesis, and the oldest was 27 years old. Considering most mathematics and statistics arXiv papers will eventually be published in a journal, usually within 2.5 years of arXiv publication, it is surprising that students would cite an 8 or even 27-year-old preprint. One possibility is that the students were given papers by their advisors, either because they were the authors or because they approved of the work. Students cited their advisors an average of 19.4% per paper, with a minimum of 0.0% and a maximum of 66.7%. If advisors sanction the use of preprints, ideally they would also evaluate all of the student's citations. However, 12 of the citations in this study were placed in the "other" category, which included incomplete and unusable citations. Furthermore, there were many citations that were incorrect or incomplete but that could be identified after further investigation.

Johnson's (2014) study, conducted using New Mexico State University dissertations, included interviews with a group of the faculty. These professors reported that some dissertation committees evaluated the student's citations more critically than others. Only one professor reported that the citations were evaluated for accuracy, but most checked to ensure the journals and authors were relevant and of sufficient quality. Johnson reported that although the faculty did not always provide students with expectations regarding references, some stated that websites were prohibited for their students. In these interviews with advisors, all revealed that they gave or recommended articles and/or authors to their students.

### ***Publisher Analysis***

There were 173 publishers cited for all resources with 54 unique book/book chapter publishers and 79 article publishers. However, there were 111 publishers, 64.2% of the total, with only one citation. Twenty publishers, 11.6% of the total, had ten or more citations. The top three book/book chapter publishers, Springer, Cambridge University Press, and the American Mathematical Society, comprised 54.6% of the total book/book chapter citations. The top ten constituted 18.5% of the publishers and were responsible for 81.4% of the citations, as shown in Table 4. The top three article publishers, Elsevier, Springer, and the American Mathematical Society, made up 49.2% of the total article citations. Table 5 shows that to reach 80.5% of the article citations required the top 17 article publishers, or 21.5% of the total. Therefore, the book and journal publisher dispersion both roughly fit the 80/20 rule.

A study by Franks and Dotson (2017) evaluated the science dissertations from Ohio State University students published between 2003 and 2012. The 127 mathematics dissertations, compared to the other sciences investigated, had citations to the largest variety of publisher types, such as commercial, university press, or society. The top five book publishers were Springer, Cambridge University Press, American Mathematical Society, Elsevier, and Wiley. These top five constituted 69.2% of the total (Franks and Dotson 2017). These top five book publishers match those found in the present study, which found that they were 68.8% of the total.

Phillips (2018) conducted a study of science dissertations from 2008-2017 from the City of New York (CUNY) system and concluded that their top five mathematics book publishers were Springer, Wiley, Elsevier, Pearson, and Oxford University Press. These five publishers were equal to 62.9% of the total. However, the books evaluated in Phillips' (2018) study did not include monographs, which may explain some of the differences from the present study.

[Insert Table 4 here]

[Insert Table 5 here]

### *Journal Analysis*

There were 229 journal titles cited, however, 122 of them, or 53.3% of the total, only had one citation. Only 13 journal titles, 5.7% of the total, had 11 or more citations, as shown in Table 6. Sinn (2005) composed a list of the top 20 journal titles based on those which received 12 or more citations. That list shared only two similarities with the top 13 journal titles from this study, the *Proceedings of the American Mathematical Society* and *Transactions of the American Mathematical Society*. Sinn's (2005) list contained mostly statistics focused journals, while this study's list did not. The top journals from this study were in publication during the years analyzed for Sinn's study. The research interests of the faculty and students may explain some of these differences. As of 2019, the University at Albany provides access to 81.5% of the 229 journal titles cited, including five that are open access but not currently in the Libraries' discovery service. The top 13 titles from Table 6 are all available.

To account for 80.1% of the article citations, 42.4% of the journal titles were needed, or 97 out of 229 titles. Although this clearly does not conform to the 80/20 rule, it reaffirms Sinn's (2005) findings where mathematics and statistics showed a higher journal title dispersion than the other sciences. Sinn's (2005) analysis showed that 30.8% of the journal titles were necessary for 80.0% of the article citations. It was noted that geoscience has shown a high title dispersion as opposed to a low dispersion in chemistry (Sinn 2005). Other sciences, such as biology and marine science, follow the predicted 80/20 title dispersion rule (Sinn 2005).

[Insert Table 6 here]

## Conclusion

The results of this study support those found in previous studies that the mathematics and statistics field does not have a small number of core journals like some of the other sciences. Additionally, as books constituted 30.5% of the citations, they are still important for mathematicians and worthy of purchase. As the average age of all materials was 19.9 years old, librarians should exercise caution when weeding and not be afraid to purchase older materials. Librarians will also need to consider the high title dispersion in the field when purchasing and cancelling titles. Furthermore, it may be necessary to create a separate list of the most cited journals in statistics as opposed to mathematics, as the most cited journals from this study were almost entirely different from the statistics focused list from another.

The results demonstrated that the University at Albany provides access to the majority of the journals cited. Due to the strain a greater number of journal subscriptions may place on a library's budget, it would be beneficial to analyze interlibrary loan statistics to determine how many requests were made for materials not owned. However, it is possible some students found articles from a source other than the library and did not check for access through the University Libraries.

One consideration for librarians is the increased popularity of arXiv preprints among students, especially as the average preprint age cited was 8.4 years old. It is possible the faculty are accepting of their use, but if not, it could be an opportunity for a conversation between the faculty and the librarian. Furthermore, the incomplete and incorrect citations found show a need to instruct graduate students in citation creation and their importance in scholarly communication. This could be remedied by working with faculty to arrange instruction and also

offering assistance outside of the classroom. It would be worth investigating if these preprints are advisor written and/or recommended, which could explain their use.

Acknowledgements: nothing to report

Funding: nothing to report

Conflict of interest: nothing to report



## References

- Allen, E. S. 1929. Periodicals for mathematicians. *Science* 70, (1825): 592–94.  
<https://doi.org/10.1126/science.70.1825.592>.
- arXiv.org. 2019a. About arXiv. Accessed November 11, 2019. <https://arxiv.org/about>.
- arXiv.org. 2019b. Arxiv monthly submission rates. Accessed November 11, 2019.  
[https://arxiv.org/stats/monthly\\_submissions](https://arxiv.org/stats/monthly_submissions).
- arXiv.org. 2019c. Arxiv submission rate statistics. Accessed November 11, 2019.  
[https://arxiv.org/help/stats/2018\\_by\\_area/index#math\\_yearly](https://arxiv.org/help/stats/2018_by_area/index#math_yearly).
- Barnett-Ellis, P. and Y. Tang. 2016. User-Centered collection development: A citation analysis of graduate biology theses. *Collection Management* 41 (1): 3–22.  
<https://doi.org/10.1080/01462679.2016.1145088>.
- Cole, C., A. R. Davis, V. Eyer, and J. J. Meier. 2018. Google Scholar's coverage of the engineering literature 10 years later. *Journal of Academic Librarianship* 44 (3): 419-425.  
<https://doi.org/10.1016/j.acalib.2018.02.013>.
- Dotson, D. S. and T. P. Franks. 2015. Winning the popularity contest: Researcher preference when selecting resources for civil engineering, computer science, mathematics and physics dissertations. *Issues in Science & Technology Librarianship* 81.  
<https://doi.org/10.5062/F4N29TZK>.
- Ferrer-Sapena, A., R. Aleixandre-Benavent, F. Peset, and E. A. Sanchez-Perez. 2018. Citations to arXiv preprints by indexed journals and their impact on research evaluation. *Journal of Information Science Theory and Practice* 6 (4): 6–16.  
<https://doi.org/10.1633/JISTAP.2018.6.4.1>.
- Flaxbart, D. 2018. Analysis of citations to books in chemistry PhD dissertations in an era of transition. *Issues in Science & Technology Librarianship* 88.  
<https://doi.org/10.5062/F4DV1H4T>.
- Franks, T. P. and D. S. Dotson. 2017. Book publishers cited in science dissertations: Are commercial publishers worth the hype? *Science & Technology Libraries* 36 (1): 63-76.  
<https://doi.org/10.1080/0194262X.2016.1263172>.
- Gross, P. L. K. and E. M. Gross. 1927. College libraries and chemical education. *Science* 66 (1713): 385–89. <https://www.jstor.org/stable/1651803>.

- Johnson, B. 1996. Citation analysis of the Texas Tech University's statistics faculty: A study applied to collection development at the university library. *LIBRES: Library and Information Science Research* 6 (3). <https://www.libres-ejournal.info/992/>.
- Johnson, P. C. 2014. Dissertations and discussions: Engineering graduate student research resource use at New Mexico State University. *Collection Building* 33 (1): 25-30. <https://doi.org/10.1108/CB-09-2013-0037>.
- Kayongo, J. and C. Helm. 2012. Relevance of library collections for graduate student research: A citation analysis study of doctoral dissertations at Notre Dame. *College & Research Libraries* 73 (1): 47-67. <https://doi.org/10.5860/crl-211>.
- Kelly M. 2015. Citation patterns of engineering, statistics, and computer science researchers: An internal and external citation analysis across multiple engineering subfields. *College & Research Libraries* 76 (7): 859-882. <https://doi.org/10.5860/crl.76.7.859>.
- Kurtz, M. J., G. Eichhorn, A. Accomazzi, C. Grant, M. Demleitner, E. Henneken, and S. S. Murray. 2005. The effect of use and access on citations. *Information Processing & Management* 41 (6): 1395–1402. <https://doi.org/10.1016/j.ipm.2005.03.010>.
- Larivière, V., C. R. Sugimoto, B. Macaluso, S. Milojević, B. Cronin, and M. Thelwall. 2014. ArXiv e-prints and the journal of record: An analysis of roles and relationships. *Journal of the Association for Information Science and Technology* 65 (6): 1157–69. <https://doi.org/10.1002/asi.23044>.
- Larivière, V. and C. R. Sugimoto. 2018. Do authors comply with mandates for open access? *Nature* 562 (7728): 483-486. <https://doi.org/10.1038/d41586-018-07101-w>.
- Nisonger, T. E. 2008. The '80/20 Rule' and core journals. *The Serials Librarian* 55(1–2): 62–84. <https://doi.org/10.1080/03615260801970774>.
- Noruzi, A. 2016. ArXiv popularity from a citation analysis point of view. *Webology* 13 (2): 1–7. <http://www.webology.org/2016/v13n2/editorial22.pdf>.
- Phillips, R. L. 2018. Book citations in PhD science dissertations: An examination of commercial book publishers' influence. *Library Trends* 67 (2): 286-302. <https://doi.org/10.1353/lib.2018.0037>.
- Sinn, R. N. 2005. A local citation analysis of mathematical and statistical dissertations. *Science & Technology Libraries* 25 (4): 25–37. [https://doi.org/10.1300/J122v25n04\\_03](https://doi.org/10.1300/J122v25n04_03).

Table 1. Citations by Resource Type

<b>Resource Type</b>	<b>Number of Citations</b>	<b>Percentage of Citations</b>
Article	712	57.0
Book	316	25.3
Book Chapter	65	5.2
Preprint	57	4.6
Web	42	3.4
Dissertation	21	1.7
Conference Proceedings	19	1.5
Other	12	1.0
<b>Total</b>	<b>1249</b>	<b>100</b>

Table 2. Results of Previous Studies of Mathematics and Statistics Citations

<b>Paper</b>	<b>Percentage of Book Citations</b>	<b>Percentage of Article Citations</b>	<b>Years Studied</b>	<b>Subject</b>
Johnson (1996)	36.9	46.7	1993-1994	Statistics faculty publications
Sinn (2005)	27.0	65.9	1980-2002	Mathematics & statistics dissertations
Kayongo and Helm (2012)	34.0	51.0	2005-2007	Mathematics dissertations
Kelly (2015)	25.0	59.0	2005-2012	Statistics dissertations
Dotson and Franks (2015)	27.1	58.1	1998-2012	Mathematics dissertations

Table 3. Citation Ages in Years by Resource Type

<b>Resource Type</b>	<b>Number of Citations</b>	<b>Mean Citation Age</b>	<b>Min Citation Age</b>	<b>Max Citation Age</b>
Article	712	20.6	0	115
Book	316	22.2	0	115
Book Chapter	65	19.5	0	58
Conference Proceedings	19	17.6	3	66
Dissertation	21	18.8	1	50
Other	4	13.0	3	37
Preprint	57	8.4	0	27
Thesis	5	15.6	2	42
Web	32	6.0	0	25

Table 4. Top 10 Cited Book Publishers

<b>Book/Book Chapter Publisher</b>	<b>Number of Citations</b>	<b>Percentage of Citations</b>	<b>Cumulative Percentage of Citations</b>
Springer	100	26.2	26.2
Cambridge University Press	56	14.7	40.9
American Mathematical Society	52	13.6	54.6
Elsevier	29	7.6	62.2
Wiley	25	6.6	68.8
Pearson	17	4.5	73.2
Taylor and Francis	10	2.6	75.9
Dover Publications	8	2.1	78.0
Princeton University Press	7	1.8	79.8
McGraw-Hill	6	1.6	81.4

Table 5. Top 17 Cited Journal Publishers

<b>Article Publisher</b>	<b>Number of Citations</b>	<b>Percentage of Citations</b>	<b>Cumulative Percentage of Citations</b>
Elsevier	148	20.8	20.8
Springer	130	18.3	39.0
American Mathematical Society	72	10.1	49.2
Cambridge University Press	31	4.4	53.5
De Gruyter	21	2.9	56.5
Duke University Press	20	2.8	59.3
Oxford University Press	19	2.7	61.9
Taylor and Francis	17	2.4	64.3
Princeton University Press	16	2.2	66.6
Theta Foundation	15	2.1	68.7
Mathematical Sciences Publishers	14	2.0	70.6
International Press of Boston	13	1.8	72.5
The Institute of Mathematical Statistics	13	1.8	74.3
Wiley	13	1.8	76.1
AIP Publishing	11	1.5	77.7
World Scientific	11	1.5	79.2
London Mathematical Society	9	1.3	80.5

Table 6. Top 13 Journal Titles Cited

<b>Journal Title</b>	<b>Number of Citations</b>	<b>Percentage of Citations</b>	<b>Cumulative Percentage of Citations</b>
<i>Journal of Algebra</i>	29	4.1	4.1
<i>Proceedings of the American Mathematical Society</i>	29	4.1	8.1
<i>Advances in Mathematics</i>	24	3.4	11.5
<i>Transactions of the American Mathematical Society</i>	22	3.1	14.6
<i>Integral Equations and Operator Theory</i>	20	2.8	17.4
<i>Journal für die reine und angewandte Mathematik</i>	20	2.8	20.2
<i>Journal of Functional Analysis</i>	17	2.4	22.6
<i>Annals of Mathematics</i>	15	2.1	24.7
<i>Journal of Operator Theory</i>	15	2.1	26.8
<i>International Mathematics Research Notices</i>	14	2.0	28.8
<i>Inventiones Mathematicae</i>	13	1.8	30.6
<i>Journal of Pure and Applied Algebra</i>	12	1.7	32.3
<i>Duke Mathematical Journal</i>	11	1.5	33.8