Bibliometrics Laws

Prof. Tom Kwanya

Lotka's Law

- Also known as the Law of Scientific Productivity.
- It deals with predicting the number of authors publishing a given number of papers.
- Developed by Alfred Lotka, a mathematician, in 1926.
- He says the number of authors making **x** contributions in a given period is a fraction of the number making a single (1) contribution.

Lotka's Law…Ctd

- The number of authors publishing a certain number of articles is a fixed ratio to the number of authors publishing a single article.
- As the number of articles published increases, authors producing that many publications become less frequent.
- He saw that the number of persons making 2 contributions is about one fourth of those making one (1) contribution.

Lotka's Law…Ctd

- The number making 3 contributions is about one ninth (1/9).
- Thus, the number of authors making **n** contributions is about $\mathbf{1/n^2}$ of those making one (1) contribution.
- The proportion of all contributors making a single (1) contribution is about 60% of all the contributors.

Lotka’s Law…Ctd

- Generally, Lotka's Law is an inverse square law that for every 100 authors contributing one article, 25 will contribute 2, 11 will contribute 3, and 6 will contribute 4 each.
- There is a general decrease in performance among a body of authors following $1:\mathbf{n}^2$
- The pattern is reflected in the table in the following slide.

Lotka's Law…Ctd

Portion of articles written	Number of authors writing that number of articles
10	$100/10^2 = 1$
9	$100/9^2 \approx 1$ (1.23)
8	$100/8^2 \approx 2$ (1.56)
7	$100/7^2 \approx 2$ (2.04)
6	$100/6^2 \approx 3$ (2.77)
5	$100/5^2 = 4$
4	$100/4^2 \approx 6$ (6.25)
3	$100/3^2 \approx 11$ (11.111…)
2	$100/2^2 = 25$
1	100

Bradford's Law

- Also known as the Law of Scattering.
- Developed by Samuel Clement Bradford, a British librarian/mathematician, in 1934.
- Dr. Bradford was concerned about the scattering of articles on specific subjects in various journals leading to some of them being left out by indexing agencies.

Bradford's Law…Ctd

- Three types of scattering have been identified by scholars who support Bradford's Law:
  - **Lexical scattering**: this involves scattering of words in collections of text;
  - **Semantic scattering**: involves scattering of concepts in collections of text;
  - **Subject scattering**: the scattering of items useful in solving specific problems.

Bradford's Law…Ctd

- Bradford argued that every scientific field is related somehow – however remotely – to all other fields.
- He stated that in any bibliography on a subject, there is always a small group of core journals which account for a substantial (1/3) proportion of the articles on that subject or discipline.

Bradford's Law…Ctd

- Then there is a larger group which account for another 1/3 and even a much larger group accounting for the other 1/3.
- He stated that if all journals in a field are sorted according to the number of articles into three groups with each group having a third of the articles, the number of journals will be proportional to **$1:n:n^2$** where **n** is Bradford's multiplier (a constant).

Bradford's Law…Ctd

- A small core of, for example, journals have as many papers on a given subject as a much larger number of journals, **n**, which again has as many papers on the subject as $\mathbf{n^2}$ journals.
- A growth in the number of articles in a subject requires a growth in the number of journals or information sources in the subject.

Bradford's Law…Ctd

- Bradford's Law has been used as an argument about:
  - how to build collections,
  - how to select journals to be indexed in bibliographies,
  - how to measure the coverage of bibliographies,
  - how to solve practical problems related to information seeking and retrieval

Zipf's Law

- Also known as the Law of Word Occurrence developed by linguist George Kingsley Zipf in 1935. The law counts the frequency with which specific words occur in scientific publications.
- He analysed the words (29,899) used in a novel known as *Ulysses* written by James Joyce.
- He ranked all the words used in the book in order of the frequency of their occurrence with the highest ranked 1 (r=1) and the least ranked last (r=29,899).

Zipf's Law…Ctd

- He noticed that by multiplying the numerical value of the rank (r) with the frequency of occurrence (f), he got a product, C, which was constant throughout the entire list of words.
- Based on the above observation he developed a formula ***rf=C*** which became known as Zipf's Law.
- The frequency of any word is inversely proportional to its rank in the frequency table.

Zipf's Law…Ctd

- Thus, the most frequent word will occur approximately twice as often as the second most frequent word, three times as often as the third most frequent word, etc.
- Therefore, the word in the position **n** appears **1/n** times as often as the most frequent word.
- In this type of distribution, frequency declines sharply as the rank number increases.

Zipf's Law…Ctd

- So, a small number of items appear very often, and a large number rarely occur.
- The most common word in English is *"the,"* which appears about one-tenth of the time in a typical text; the next most common word (rank 2) is *"of,"* which appears about one-twentieth of the time.
- Studies show that this phenomenon also applies in nearly every language.

Garfield's Law

- Also known as Garfield's Law of Concentration.
- It points out that for any field of science, articles are concentrated within the same highly cited or multidisciplinary journals.
- He also opined that the core literature for all scientific literature involves a group of not more than 1,000 journals and could be as few as 500 journals.

Garfield's Law

- This observation implies that a good general science library need not have more journals than a special library holding a specialised collection covering a specific discipline.
- He also argued that an index does not have to include all journals in a discipline as long as it covers the core journals.