

Brookes B C. The derivation and application of the Bradford-Zipf distribution.
J. Documentation 24:247-65, 1968.
[School of Librarianship and Archives, University College London, England]

The exact form of the Bradford law having been recently clarified, the paper reverts to Bradford's application of the law to provide bibliographical estimates. Simple graphical techniques are described and standard forms are proposed to facilitate comparisons. [The *Science Citation Index*® (SCI®) and the *Social Sciences Citation Index*® (SSCI®) indicate that this paper has been cited in over 65 publications, making it the most-cited paper ever published in this journal.]

Bertram C. Brookes
Department of Information Science
The City University
London EC1V 0HB
England

November 4, 1985

In 1966 University College London established a readership in information science and offered the new post to me. I demurred: I could not discern the science I would be expected to profess. Information studies? Yes. I accepted, hoping to help in founding the new science.

Recalling that Lord Kelvin could "understand only what he could measure," I scanned my new field for "measurabilia." My first three papers were on the Swets model of information retrieval,¹ on obsolescence as measured by citation,² and, ultimately the most intriguing, on the Bradford law.

Bradford had regarded his bibliographical data to be exact for the specialist journals but incomplete for those on the periphery. The law could be used to estimate the number of journals and papers yet to be found by extrapolating from the initial rising linearity of the cumulative ranked graph. This idea had been challenged by O.V. Groos,³ who argued that the plot of any real

bibliography fell increasingly below the predicted linearity. By close observation of the compiling of an exhaustive bibliography of vitamins (by Susan Wright of the British Medical Association), I noted that the initial "Groos droop" (as I called it) was gradually rectified as the search patiently continued.

My simple graphical techniques were widely taught in the library schools. Campus bookstores reported sudden sellouts of logarithmic graph papers. "Bibliographs" rained on me from the five continents. But, sadly, few were usefully applied. All too often, the fact that another set of data was found to conform with the Bradford law was hailed as a significant discovery.

Since writing that paper, I have continued to explore the variant forms of Bradford-Zipfian data from many sources. I have tried to relate the laws to information theory, to perspective, to a mental information space, and to categorization. In 1982 Haitun^{4,6} satisfied me that the "Zipfian" distributions found only in the social domain lie beyond the reach of the "Gaussian" statistical theory, which serves the physical domain so well. Supported by a Leverhulme Emeritus Fellowship, I began a book on Bradford-Zipf distributions.

The book was almost complete when Sichel⁷ appeared to refute Haitun's thesis by demonstrating that these distributions, at least in their frequency forms, could be captured by the Inverse Gaussian-Poisson distribution. I have had to revise the arguments of my book.

So far, however, the simple frequency-rank forms introduced by Bradford and Zipf have not been captured by Sichel's penetrating but intractable Bessel functions. Nor, I believe, are they likely to be. The ranked forms retain empirical information that frequency distributions discard. A new statistics for the analysis of the particularities of the social sciences is taking shape.

1. Brookes B C. The measures of information retrieval effectiveness proposed by Swets. *J. Documentation* 24:41-54, 1968.
2. Obsolescence of special library periodicals; sampling errors and "utility contours." *J. Amer. Soc. Inform. Sci.* 21:320-9, 1970.
3. Groos O V. Bradford's law and the Keenan-Atherton data. *Amer. Document.* 18:46, 1967.
4. Haitun S D. Stationary scientometric distributions. Part I. Different approximations. *Scientometrics* 4:5-25, 1982.
5. Stationary scientometric distributions. Part II. Non-Gaussian nature of scientific activities. *Scientometrics* 4:89-104, 1982.
6. Stationary scientometric distributions. Part III. The role of the Zipf distribution. *Scientometrics* 4:181-94, 1982.
7. Sichel H S. A bibliometric distribution which really works. *J. Amer. Soc. Inform. Sci.* 36:314-21, 1985.