e-Pathshala पाठशाला
A Gateway to All Post Graduate Courses

MHRD Govt. of India

An MHRD Project under its National Mission on Education throught ICT (NME-ICT)

Subject: Library and Information Science

Production of Courseware
e-Content for Post Graduate Courses

Paper No : 10 Informetrics and Scientometrics

Module : 05 Classical Law of Bibliometrics

Development Team

Principal Investigator & Subject Coordinator	Dr. Jagdish Arora, Director INFLIBNET Centre, Gandhinagar
Paper Coordinator	Dr I K Ravichandra Rao Retd Professor, Documentation Research and Training Centre
Content Writer	Prof B K Sen Retd Professor, National Institute of Science Communication and Information Resources
Content Reviewer	Dr I K Ravichandra Rao Retd Professor, Documentation Research and Training Centre

Unit 5

Classical Law of Bibliometrics

I. Objectives

After going through this study material you will

- know the classical laws of Bibliometrics;
- learn the formulas pertaining to the laws;
- be able to apply the laws in library environment; and
- educate yourself as to how these laws are verified.

II. Learning Outcome

After completion of this module, you will be familiar with three classical laws of Bibliometrics and their applications. Particularly, you have learnt Bradford's law, Lotka's law and Zipf's law.

III. Module Structure

1. Introduction

Classical laws in Bibliometrics include Bradford's law of scatter [1, 2, 3], three laws of Zipf [28, 29, 30] and Lotka's law [16]. These are purely scientific laws which are well-established formulas and the concepts embedded in them are less likely to change with time. Moreover, their validity can be verified scientifically.

2. Bradford's Law of Scattering

The law was propounded by Samuel Clement Bradford (1878-1948), a British librarian, mathematician and document lists at the Science Museum in London after a laborious study of scientific literature in mid-1930s. After examining the distribution of scientific literature in periodicals and their coverage in abstracting and indexing periodicals he realized that the distribution of literature follow a particular pattern [1, 2]. He opined that 'the nucleus of periodicals devoted to the given subject must contain, individually, more articles on that subject than periodicals dealing with related subjects' [3]. 'In consequence, it is possible to arrange periodicals in zones of decreasing productivity, in regard to papers on a given subject, and the numbers of periodicals in each zone will increase as their productivity decreases' [3]. He described a scattering pattern of journals in the area of applied geophysics and lubrication. He plotted the partial sums of references against the natural logarithm of the partial sum of numbers of journals, and he noticed that the resulting graph is a straight line. On the basis of this observation, he suggested the following linear relation to describe a scattering phenomenon [2]

$$F(x) = a + b \log x.$$

F(x) is the cumulative number of references contained in the first x most productive journal; a and b are constants. The following figure is a hypothetical, but typical, log-linear curve (as described by Bradford) showing aggregates of articles on a given subject corresponding to the number of journals. This type of a curve is usually called a Bradford curve; it is shown in page 3.

P_1 in the figure is the point at which the straight line part of the curve begins. Draw Y_1P_1, Y_2P_2, and Y_3P_3 such that they are parallel to the X-axis and $OY_1 = Y_1Y_2 = Y_2Y_3$. Draw P_1X_1, P_2X_2, and P_3X_3 such that they are parallel to Y-axis. Since P_1P_3 is a straight line and since $Y_1Y_2 = Y_2Y_3$, X_1X_2 and X_2X_3 are equal, say r units. Let the distance between O and X is s units. Thus, if α, β and γ are the positive real numbers corresponding respectively to the logarithmic abscissa OX_1, OX_2 and OX_3, we have, $\log \alpha = s$, $\log \beta = s + r$, and $\log \gamma = s+2r$. That is,

$$\alpha = 10^s,\ \beta = 10^{r+s} = 10^s.10^r,\ \text{and } \gamma = 10^{s+2r} = 10^s.10^{2r}$$

Substituting $n = 10^r$, we see that the natural numbers α, β, and γ are related to each other as $1 : n : n^2$. On the basis of this relationship and also since OX_1 represents a number of periodicals in a subject area Bradford stated his law as t "If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to subject, they and several groups or zones"

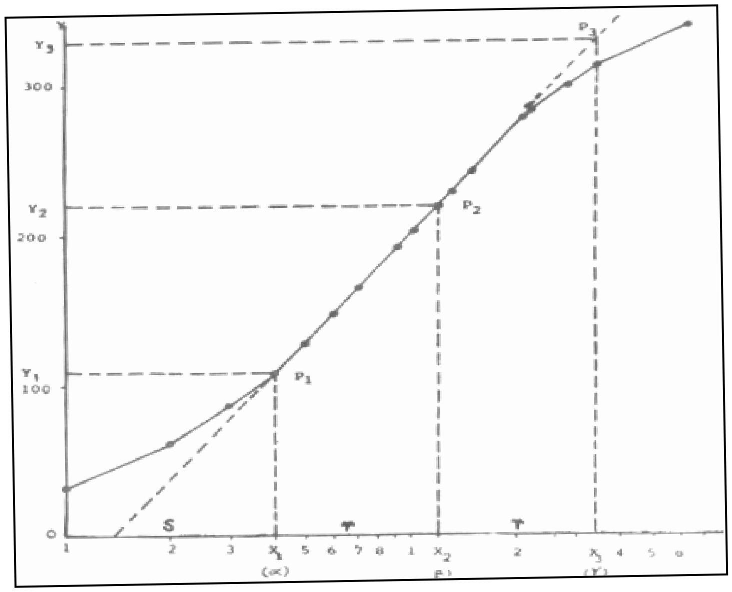

Fig. 1: Bradford Bibliograph

In X-axis: Partial sum of Journals (in log scale)

In Y-axis: Partial sum of articles contained in X top most journals (in linear scale)
Containing the same number of articles as the nucleus, when the zones will be $1:n:n^2$ ….”

This is called Law of Scattering or Bradford's law. He has obtained the partial sums of the journals only after the journals are ranked according to the number of articles they publish. Bradford in 1948 summarized his earlier observations in a book which contained a theoretical derivation of his law of scattering.

Interestingly, Bradford in 1948 [3] derived his law in a different approach. In this approach, he assumed that the collection of journals is ranked (or arranged) in decreasing productivity. Productivity of a journal is implicitly defined in terms of the number of articles, in a given subject, it contains. Divide these journals into k zones. Let m_k be the number of journals in the k^{th} zone. Let r_k be the average number of articles per journal in the k^{th} zone. m_1r_1, m_2r_2,….. and m_kr_k are the total productivity of 1^{st}, 2^{nd}, 3^{rd}, … k^{th} zone respectively. Zones are formed such that:

$$m_1r_1, = m_2r_2 = m_3r_3 = \ldots\ldots\ldots = m_kr_k \qquad \text{(A)}$$

Since

$$r_1 > r_2 > r_3 > \ldots.. > r_k \text{ and } m_1r_1, = m_2r_2 = m_3r_3 = \ldots\ldots\ldots = m_kr_k$$

We have

$$m_1 < m_2 < m_3 < \ldots\ldots\ldots < m_k$$

From (A), we have

$$m_i = \frac{r_1}{r_i} \qquad i = 2,3,4,\ldots\ldots k.$$

That is,

$$m_i = \frac{r_1}{r_2}\frac{r_2}{r_3}\frac{r_3}{r_4}\ldots\ldots\ldots\frac{r_{i-2}}{r_{i-1}}\frac{r_{i-1}}{r_i}m_1$$

Defining

$$n_{i-1} = \frac{r_{i-1}}{r_i} \qquad i = 2,3,4, \ldots\ldots k$$

We have

$$m_i = n_1 n_2 n_3 \ldots\ldots n_{i-1} m_i \prod_{j=1}^{i-1} n_j m_1$$

Bradford, in his analysis considered only three zones (k = 3). He thus had

$$m_2 = n_1 m_1 \text{ and } m_3 = n_1 n_2 m_1$$

He then stated, " … we know no reason why n_1 and n should differ and the simple supposition we could make is that they are equal." Say, $n_1 = n_2 = n$. (Bradford (1948)). Thus, $m_3 = n^2 m_1$. Therefore, he argued that the ratio of the zone sizes will be as 1 : n: n^2 which is again the relationship developed in his earlier work.

Since then many have worked in this area and came out with different approaches as well as different explanations. The author of this Unit discusses some of these works in Unit 6. Vickery [27], Bookstein [4], Brookes [5,6,7], Hubert [11,12,13], Leimkulher [15], Egghe [10], Naranan [17] and many others have given different explanations of this Law; Some of these works were also discussed by Ravichandra Rao [21]; he also observed in one of his papers that Baradford's assumption of $n_1 = n_2 = n$ is unlikely correct.

2.1 Explanation

Bradford conducted his study with scientific periodicals only. In certain cases the law holds good for non-scientific periodicals also. The law will be applicable if the articles pertain to a specific subject. If articles of several subjects are taken together for testing the Bradford law, the test is likely to fail. Though the definition speaks of several zones, the formula given in the definition takes care of three zones only. In certain cases the law may be applicable beyond three zones even. Bradford called the first zone as nucleus and other zones as succeeding zones. We may call the first zone as Core, second zone as Allied, and the third zone as Alien, to identify each zone specifically. Let us take a concrete example to understand the law properly.

Suppose on a subject there are 600 articles. If these articles are equally divided into three zones, then each zone will have 200 articles. Obviously, the nucleus will also have 200 articles. If these articles pertain to five journals, the articles of the next group numbering 200, may pertain to 10 journals, 15 journals or more. If the article of the next zone (allied zone) pertain to 15 journals, then it is likely that the article of the 3^{rd} zone (alien zone) will pertain to 45 journals. Now, we find that the articles of the three succeeding zones belong to 5, 15, and 45 journals. If the numbers are divided by 5, we get 1, 3 and 9, that is 1, 3, and 3^2 which represents Bradford's law, i.e. 1:n:n^2. In Bradford's study, we generally get an approximate value of n. In reality, the articles in the three zones are only roughly equal.

In a situation where the data ideally suits Bradford formulation, the set of journals, say, numbering 310, would be divided as 10, 10x5, and 10x5x5. The first 10 journals will belong to the core zone, the next 50 to the allied zone, and the last 250 to the alien zone. The number of papers in each zone will be the same. In reality, we do not get such data that perfectly satisfies the law. In most cases, the data satisfies the law roughly.

2.2 Verification

It involves the following steps:

i. A comprehensive bibliography on a narrow subject is to be taken. An already compiled bibliography may be taken as did Bradford. Otherwise a bibliography can be compiled. While compiling the bibliography, care should be taken to include only journal articles. Articles presented in a conference, in festschrift volumes, in collected works, etc. are to be excluded.

   The term 'comprehensive bibliography' needs a bit explanation. A question often asked as to the optimum number of journal articles in the bibliography that would be suitable for verifying Bradford's law. The answer is hard to come by. In Bradford's studies in one bibliography there were 1332 journal articles, and in the other 395 journal articles. If the bibliography contains around 500 articles, it is likely to be a good sample.

   If a compiled bibliography is taken, then only the journal articles are to be included leaving aside all other items such as books, conference documents and so on.

ii. If the work is done manually, then cards are to be prepared for the entries required for the purpose. Otherwise the required entries may be entered in a computer in Excel sheet or any other suitable software.

iii. The cards or entries are to be sorted according to periodicals.

iv. Now, the periodicals are to be sorted according to the frequencies. The periodical with the highest frequency is to be ranked 1st, followed by others with descending frequencies.

v. The data is to be tabulated as per the Table 1 given below.

vi. The validity is to be tested according to the formula given above. The data in Table 1 is Bradford's data.

Table 1 needs some explanations. The periodical having the largest number of articles, i.e. 93, ranks first followed by periodicals accounting for descending number of articles. In the ranking you will notice that sometimes two or more periodicals were responsible for the same number of articles. For example, four periodicals ranked 13th with 16 articles each. In such a case, all are given the 13th rank, because if we give all the periodicals 10th rank, then it would seem as if there are 10 periodicals, in fact there are 13 periodicals. By seeing the last rank in the Table, e.g. 326, we easily know the total number of periodicals contributing to the literature in the field. In actual studies the name of the periodical is also given along with the rank number. When the number of periodicals is very high, to shorten the size of the Table, the names of the periodicals are given, say up to 50th rank or even less.

Rank	No. of Periodical/s	No. of Articles	Cumulative Total of Articles
1	1	93	93
2	1	86	179
3	1	56	235
4	1	48	283
5	1	46	329
6	1	35	364
7	1	28	392
8	1	20	412
9	1	17	429
13	4	16	493
14	1	15	508
19	5	14	578
20	1	12	590
22	2	11	612
27	5	10	662
30	3	9	689
38	8	8	753
45	7	7	802
56	11	6	868
68	12	5	928
85	17	4	996
108	23	3	1,065
157	49	2	1,163
326	169	1	1,332

Table 1- Productivity-wise Distribution of Periodicals
[Based on Current Bibliography of Applied Geophysics 1928-1931(1)]

To test the law, we shall now divide the number of articles in three equal parts. If we do so, then each part will have 1332÷3 = 444 articles. This gives us the indication that each zone will have 444 articles. In Table 1 the number 444 is missing. It comes in between 429 and 493. The number of articles up to 9^{th} rank is 429. Four periodicals occupy the 10^{th} to 13^{th} rank with 16 articles each. Adding 16 to 429, we get 445 which is quite close to 444. Thus, the first 10 periodicals accounting for 445 articles form the first zone. It is to be noted that the number of periodicals as well as articles should always be in whole numbers and not infractions. To decide the number of periodicals in the second zone, 444 is multiplied by 2, and we get 888. In Table 1 the cumulative number 888 is missing. It follows 868. The first 56 periodicals account for 868 articles. The next 11 periodicals are having 6 articles each. That means, the first 59 periodicals are accounting for 886 articles which is close to 888. Hence, in the second zone we have 59-10 = 49 periodicals and 886-445 = 441 articles. In the first two zones, there are 10+49 = 59 periodicals and 445+441 = 886 articles. Hence, in the third zone, there will be 326 – 59 = 267 periodicals, and 1332- 886 = 446 articles.

Zone	No. of Periodicals	No. of Articles
1st Zone	10	445
2nd Zone	49	441
3rd Zone	267	446

Table 2 – Distribution of Articles according to Zones

From the data, we get the value of n as 49÷10 = 4.9. Multiplying 49 by 4.9, we get 240 which is close to 267. The ratio of the periodicals is 10:49:267 which is roughly $1:5:5^2$.

2.3 Bradford's Bibliograph

The Bibliograph [Fig. 1] can be obtained by plotting R(n) on a linear scale along the Y axis and n on a logarithmic scale along the x axis. It is best drawn on semi-log graph paper. The graph is J-shaped and turns into a straight line from a particular point.

2.4 Uses

i. It depicts the spread of the journal literature of the subject. The subjects of the journals harboring the literature shows how enormous are the spread and the journal variety.

ii. It helps identify the core periodicals in a subject. Also it throws light on the allied and alien periodicals pertaining to the subject.

iii. It produces a ranked list placing the journals according to their productivity in descending order. It helps a great deal in the selection of periodicals.

iv. Due to budget crunch when certain periodicals need to be deleted, usually the lower ranked periodicals are deleted.

3. Zips's Laws

Zipf's laws [28, 29, 30] are related to the usage of words by individuals. Interestingly, the phenomenon follows a scientific law. Most probably, the French stenographer, Jean-Baptiste Estoup (1858-1950), is the first individual to observe the hyperbolic nature of the frequency of word usage. Estoup recorded his study in his book Gammes Stenographiques published from Paris in 1916. George Kingsley Zipf (1902–1950), the American linguist, saw the 4th edition of Estoup's book and worked further on it and arrived at his laws in 1935.

3.1 First - Definition

"If the number of different words occurring once in a given sample is taken as x, the number of different words occurring twice, three times, four times, n times in the same sample , is respectively $1/2^2$, $1/3^2$, $1/4^2$, ...$1/n^2$ of x, up to, though not including, the few most frequently used words; that is , we find an unmistakable progression according to the inverse square, valid for well over 95% of all the different words used in the sample" [13, Preface, p. vi]. Basing this phenomenon, Zipf developed the formula $ab^2 = k$, where a is the number of words occurring b times, and k is a constant.

It is observed that the equation satisfies quite well with the less frequently used words of the sample, which are much more in number compared to more frequently used words in the sample [30, p.47].

3.1.1 Explanation

Suppose in a particular writing, the number of words occurred only once is 1440. In that case the number of words occurring 2, 3 and 4 times is likely to be $1440 \div 2^2$, $1440 \div 3^2$, and $1440 \div 4^2$, that is 360, 160, and 90. Now, using this data and the formula, we can determine the value of k. We know, 1440 words have occurred only once, that means a = 1440, b = 1. Therefore, $k = ab^2 = 1440 \times 1^2 = 1440$. Now, using the value of k, we can find out how many words have occurred 5 times, 6 times or any other number of times. In this case, the formula will be $a = k/b^2$. Let us try to find out how many words have occurred 6 times. We know k = 1440, and b = 6. Putting these values in the aforesaid equation, we get $a = 1440/6^2 = 40$; around 40 words are likely to occur 6 times each. It is to be noted that the values will change with every piece of writing, literary or non-literary.

3.1.2 Verification

The manual verification of the law is highly laborious and time consuming. The best way to verify the law is to take the help of information technology as detailed under the second law.

3.1.3 Uses

The use of this law in library and information science (LIS) is still negligible, maybe because it is not generally taught in our LIS courses. It will be useful in identifying the style of writing of an author. The only case I know about its application is that of a student of the University of Calcutta who is trying to find out to what extent the writing of a noted Bengali scientist has been influenced by a foreign lady. The details will be known when the thesis is submitted and degree awarded.

3.2 Second Law - Definition

Zipf defined the second law as ‘The conspicuousness or intensity of any element of language is inversely proportionate to its frequency. Using X for frequency and Y for conspicuousness (rank) ‘ the law can be mathematically expressed as

$$y \quad \alpha \quad \frac{1}{x}$$

$\Rightarrow \quad xy = n$, where n is a constant. [30]

3.2.1 Explanation

The words conspicuousness or intensity can be simply termed as rank, and element can be simply taken for word. The definition can be stated in simpler terms as the rank of any word of a language is inversely proportional to the frequency of its usage. The lower the rank, the higher will be the frequency.

3.2.2 Verification

The verification of the Law came through the use of Miles L Hanley’s Index of Words for James Joyce’s Ulysses. Zipf found that the rank frequency word distribution ‘approximate the simple equation of an equilateral hyperbola: r x f = C’` [14, p.24], where r indicates rank and f frequency”. C in the equation is the constant. The Product in Table 3 represents constant. The constant in the case of this Law is not a fixed number, but close to it. A part of the result of Zipf’s experiment in reproduced below in Table 3.

Rank (r)	Frequency (f)	Product (c)
10	2653	26530
20	1311	26220
30	926	27780
40	717	28680
50	556	27800
100	285	28500

Table 3: Distribution of Words in James Joyce’s Ulysses

A manual verification of the Law is highly laborious and time consuming. In the absence of Hanley’s Index of Words for James Joyce’s Ulysses Zipf’s experiment would have been really difficult. The advent of information technology has given us devices with which we can verify the Law without much difficulty. The steps involved in the verification are given below [24].

- Take a piece of writing in English containing not less than 5,000 words. The writing can be an article, a short story, a novelette, a part of a novel, even a technical writing.

- Scan the selected writing with an optical character recognition (OCR) software viz. OmniPage Pro.

- Save the file in a suitable software package such as Microsoft Word for Windows.

- Check the file with the original to ascertain accuracy.

- Consider only the textual portion of the writing and remove the names of the authors, author affiliations, abstract, keywords, alpha-numeric expressions like 2nd and F10, alpha-symbolic expressions like au=, and su=, abbreviations such as FDT and ISO, numbers written with digits (eg 324), serial numbering such as a), b) etc., formulas, punctuation marks, intra- and extra-textual references, tables, figures, and appendices.

- The rationale behind the exclusion of the names of authors and author affiliations is obvious because they don't represent the author's style of writing and as such cannot be used for word counting. The keywords are sometimes chosen consulting a thesaurus where the author has little choice and sometimes added by the editor. Hence, it was felt safe to exclude them. An abstract is the condensed version of an article and does not necessarily represent the normal style of writing of an author. Moreover, sometimes the abstract is prepared by someone other than the author. Therefore, it was not considered. Alpha-numeric as well as alpha-symbolic expressions, abbreviations, numbers written with digits, serial numbering with a, b, c, etc., and formulas are not words, hence excluded. The references comprise certain fixed elements such as author, year, title of the article, and other bibliographical details which are not the creation of the author. Tables and figures were to be excluded for the ease of sorting. Moreover, at times a table may contain numerous YESes and Noes or Ys and Ns representing the answers of respondents. Appendices usually are also not reflective of the style of writing of an author. Consider only the textual part of the article with the above exceptions since that part seemed to be the best part for judging the word use pattern of an author.

In the case of hyphenated words, follow the following rules.

- If the hyphen joins a prefix, such as co-ordination, remove the hyphen, and consider the word as coordination
- In other cases such as short-term. Remove the hyphen, and consider the combination as two different words
- Convert the punctuation marks into spaces,

- Convert all spaces into line breaks.

- Convert all upper case letters into lower case letters
- Make sure that the file has converted into a pure word file
- Sort the file alphabetically and save it as a text file
- Run the file through a small program that can count the frequencies of word occurrence.. The program will take the text file as its input and reproduce the text file with the corresponding frequencies as its output.
- Convert the file into a table
- Sort the table according to descending frequencies
- Add one column to the table for rank
- Compute rank x frequency
- If the product of rank and frequency is found to be more or less the same, the Law is verified.

3.2.3 Uses

This law is taught in some LIS courses in India. It has been the subject of study in Master's level dissertations as well as theses. Ray [22, 23]in his thesis has applied this law of Zipf basing the words of Gitanjali. The ranking of the words help generating automated abstracts as well as keywords.

3.3 Third Law - Definition

The length of a word is very closely related to the frequency of its uses. The greater the frequency, the shorter the words.

3.3.1 Explanation

This law relates to the length of words. In the case of this law, word length means the number of letters present in a word. For example, the word length of ‘equation’ is 8 as there are eight letters. Here the unit of measurement is a letter and not millimeter or centimeter.

You might have noticed that words are of various lengths. The words of shortest length such as a, I, etc are few. The number of words with greater and greater lengths goes on increasing up to a certain point beyond which the number comes down and finally becomes zero [30]. In Table 4 we see that the number of words of length 3 is maximum. After which as the world length increases, the number of words decreases. Beyond 14 word length, the frequency is found to be zero. In the English language there are words whose length go up to 20 or a bit beyond. But, they are very few in number.

Word Length	Frequency
1	36
2	309
3	362
4	252
5	213
6	179
7	198
8	141
9	187
10	91
11	65
12	27
13	4
14	8
15	0
16	0
17	0
18	1
19	0
20	0

Table 4: Word Length vs. Frequency of Occurrence

- Take a piece of writing in English containing not less than 5,000 words. The writing can be an article, a short story, a novelette, a part of a novel, even a technical writing.
- Scan the selected writing with an optical character recognition (OCR) software viz. OmniPage Pro.
- Save the file in a suitable software package such as Microsoft Word for Windows
- Check the file with the original to ascertain accuracy.

3.3.2 Verification

Follow the step nos. i to xii as given under the 2nd Law. Thereafter, a short program is to be written to count the word lengthwise. Once that is done, prepare the Table as given above. The Table itself will provide the indication as to the verification of the Law.

3.3.3 Use

In many software packages like CDS/ISIS, it is necessary to specify the word length. The study of word lengths gives us an idea of the maximum length of words in various languages. Accordingly word lengths can be specified in computerization processes of various LIS activities like database creation, index generation and so on.

4. Lotka's Law

Bradford’s law deals with the scatter of journal literature devoted to a particular subject wherefrom we get an idea of journal productivity as well. Zipf’s laws are devoted to the study of words from various angles using statistical methods. Lotka’s law studies the author productivity. In 1926, Alfred J. Lotka, a statistician of the Metropolitan Life Insurance Company, USA, became engrossed with the idea of determining, ‘if possible, the part which men of different caliber contribute to the progress of science’. For this purpose, he used the index of Chemical Abstracts for the years 1907-1916 and developed a listing of A and B names [i.e. the names starting with the letter A and B] and the corresponding number of papers each author produced. The same procedure was applied to Auerbach’s Geschichtstafeln der Physik till the year 1900 using complete coverage [16]. After studying the productivity of the authors, he was surprised to see that the productivity of the authors can be expressed with a simple equation.

4.1 Equation

The equation derived from the study is $x^n y = c$ where x stands for the number of contributions, y for the number of authors, and c is constant. Lotka found the value of n as 2. Since then, many have worked in this area. References to these works may be seen in Egghe (9,10), Ravichandra Rao (2o); also Egghe and Ravichandra Rao (8) has developed a model to explain the distribution of productivity based on the fractional counting method. Lotka's law was derived based on the simple counting methods as explained above.

4.2 Explanation

It has been observed that more number of authors contribute less number of papers. If you go through the author index of Indian Library Science Abstracts 1992-1999 or 2000-2005 [14], you will notice that the largest number of authors have contributed only 1 paper, less number have contributed 2 papers, still less number contributed 3 papers and so on. If you take a count you may find that around 100 authors have contributed 1 paper each, and only about one or two authors have contributed 10 papers or more. From the formula the productivity of authors producing 1, 2, or more articles, can be estimated as below.

No. of Papers (x)	No. of authors (y)
1	1024
2	256
3	114
4	64
5	41
6	28
7	21
8	16
9	13
10	10

Table 5: Distribution of Papers according to the Number of Authors

4.3 Verification

For the verification of the law, you are to follow the following steps.

- Take an author index of an abstracting or indexing periodical. You can take the whole A to Z index if the number of entries is manageable. Otherwise you can take part of the index.
- Count the number of entries against each author, and write down the number against his name as given in Table 6. If this data is entered into a computer, sorting of the data in column 3 will be very easy.
- From the sorted data in column 3, find out the number of authors who have contributed one article, two articles, and so on.
- Tabulate the data as in Table 5.
- Taking the value of n as 2, check whether the data fit into Lotka's law. If it does not, you are to try with different values of n. It may be more than 2 or less than 2. By trial and error method you may arrive at a value that will bring the figures quite close to the actual values. Otherwise you may follow Sen's method [24] or Pao's method [18] to find out the value of n. Pao's method involves a great deal of mathematical calculation and students with very good mathematical background can apply the same.

Authors	Entry No/s. of Contributions	No. of Contributions
Abbahu (G E P)	2599	1
Abbas (S M)	2738	1
Abbas Ibrahim	0538	1
Abbulu (G E P)	0917	1
Abdella (Woinshet)	3036. 3037	2
Abdul Azeez (T A)	1215	1
Abdul Jaleel (T)	1119	1
Abdul Majeed Baba	0595. 0956, 0929	3
Abdul Rashid	0311, 0312, 0743, 1477, 1684	5
Abdur Rauf Meah (Md.)	2511	1
Abid (Abdelaziz)	2453	1
Abideen P (Sainul)	0461	1
Abidi (Syed A H)	2504	1
Abifarin (Abimbola)	2233	1
Abraham (Deborah V H)	0423	1
Abraham (J)	0424, 0466, 0744, 1015, 1726, 2031, 3011	7

Table 6: Author Productivity*

* Columns 1 &2 reproduced from Indian Library Science Abstracts 1992-1999 [4]

4.4 Use

It helps to determine the highly productive, moderately productive and less productive authors in a subject. It needs to be mentioned here that the book ‘Power Laws in the Information Production Process Lotkaian Informetrics’ brought out by Academic Press in 2005 under the editorship of Leo Egghe dwells on the subject at length. It presents informetric results from the viewpoint of Lotkaian size frequency functions. The theory has been developed in the framework of Information Production Processes (IPPs). Its relationship with the law of Zipf applications has also been indicated in a number of fields.

5. Summary

Three classical laws of bibliometrics have been discussed above. Of all the laws Bradford’s law has found wider application. We shall discuss more about this in the Conclusion section of Bradford Distributions (Unit 6). Of the three laws, it is generally seen that the largest number of papers are appearing on Bradford’s law followed by Lotka’s law and Zipf’s laws. Compared to

Bradford law and Lotka’s law, conducting studies with Zipf’s laws is time consuming and tiring as the laws involve thousand and thousands of words. Nowadays these laws are being studied using computers. The uses of Lotka’s law and Zipf’s laws have so far remain limited. From India, a large number of articles and theses appear on bibliometrics. Hardly any of them generate a new bibliometric law, find a new use of the laws, provide a new method or mathematical formulation of the laws. It is heartening to note that a student of the University of Calcutta is applying one of the Zipf’s laws in authorship attribution. The outcome of his study is encouraging. To this day there was no good course material on the laws. It is hoped that with these course materials they will be able to grasp the laws and work with the laws with ease.

6. References

1. Bradford, Samuel Clement. Wikipedia <en.wikipedia.org>. Web. 1.5.2013.

2. Bradford, S. C. (1934) Sources of information on specific subjects. Engineering, 26: 85-86.

3. Bradford, S.C. (1953) The documentary chaos. **In** Bradford S C. Documentation. Crossby Lockwood. London. Ch. IX, p. 144-59.

4. Bookstein, Abraham. (1976) Bibliometric distributions. Library Quarterly, 46(4): 416-23.

5. Brookes, B.C. (1969a) The complete Bradford-Zipf bibliography, Journal of Documentation. 25: 58-60.

6. Brookes, B.C. (1969b) Bradford's law and the bibliography of science. Nature. 224: 953-6.

7. Brookes, B.C. (1977) Theory of the Bradford laws. Journal of Documentation, 33: 180-209.

8. Egghe, L and Ravichandra Rao, I.K. (2002) Duality revisited: Construction of fractional frequency distributions based on two dual Lotka law. Journal of the American Society for Information Science and Technology, 53(10): 789-801.
9. Egghe, L. (1993) Consequence of Lotka’s law in the case of fractional counting of authorship and of author count. Math Computing Modeling, 18(9): 63-77.

10. Egghe, L. and Rousseau, R. (1990) Introduction to informetrics. Elsevier. Amsterdam.

11. Hubert, John J. (1976) On the Naranan interpretation of Bradford's law. Journal of the American Society for Information Science, 27: 339-41.

12. Hubert, J.J. (1977) Bibliometric models for journal productivity. Social Indicators Research, 4: 441-73.

13. Hubert, J.J. (1978) Relationship between the forms of Bradford's law. Journal of the American Society for Information Science, 29(3); 159-61.

14. Indian Library Science Abstracts. (1967-) Kolkata: IASLIC.

15. Leimkuhler, Ferdinand F. (1967) Bradford distribution. American Journal of Documentation. 23(3): 197-207.

16. Lotka, A.J. (1926) Statistics - the frequency distribution of scientific productivity. Journal of the Washington Academy of Science, 16: 317- 25.

17. Naranan, S. (1970) Bradford's law of bibliography of science: an interpretation. Nature, 227: 631-2.

18. Pao, M.L. (1985) Lotka’s law – a testing procedure. Information Processing and Management. 21(4): 305-320.

19. Ravichandra Rao, I.K. (1983). Quantitative methods for library and information science. Wiley Eastern Limited. New Delhi.

20. Ravichandra Rao, I.K. (2010). Growth of literature and measures of scientific productivity: scientometric models. Ess Ess Publications (published for Sarada Ranganathan Endowment for Library Science, Bangalore), New Delhi.

21. Ravichandra Rao, I.K. (1998). An analysis of Bradford multipliers and a model to explain the law of scattering. Scientometrics. 41(1): 93-100.

22. Ray, Partha Pratim and Sen B.K. (2012). Gitanjali (Song Offerings): a bibliometric study. SRELS Journal of Information Management. 49(4): 343-358.

23. Ray, Partha Pratim. (2012) Tagore in print: publication of Tagore’s literature five years before and after expiry of copyright. Concept Publishing. New Delhi. 2012**.**

24. Sen, B.K.; Che Azlan bin Taib and Mohd, Faris. (1996) “Library and information science literature and Lotka’s law”. Malaysian Journal of Library and Information Science, 1(2): 89-93.

25. Sen, B.K.; Khong Wye Keen; Lee Soo Hoon; Lim Bee Ling; Mohd Rafae Abdullah; Ting Chang Nguam and Wee Siu Hang. (1998) Zipf’s law and writings on LIS. Malaysian Journal of Library and Information Science, 3(1); 25-42

26. Sen, B.K. (2010) Lotka’s law – a viewpoint. Annals of Library and Information Studies. 57(2): 166-168.

27. Zipf’s law. Wikipedia <en.wikipedia.org>. Web. 1.5.2013

28. Zipf, G.K. (1935). The psycho-biology of language: an introduction to dynamic philology: Riverside Press, Houghton Mifflin, Boston.

29. Zipf, G.K. (1949) Human behaviour and the principle of least effort; an introduction to human cology. Addison -Wesley, Reading, Mass.

30. Vickery, B.C. (1948) Bradford's law of scaterring. Journal of Documentation, 28(2): 122-20.