Chapter Five

Scientometrics / Bibliometrics laws
5.1 INTRODUCTION AND HISTORICAL BACKGROUND

Communication is the exchange of information between individuals by means of a common signal system. Scientific communication can be defined as the combined processes of presentation, delivery and receipt of scientific information in human society. To study scientific communication, techniques and measures are needed. The family of techniques used to identify relationships in the published literature is collectively referred to as "Bibliometrics".

Bibliometrics is the generic term used to describe indicators to measure the output of scientific and technological research through data from scientific literature and patents.

Bibliometric data enables us to explore trends, both within a country and worldwide, as well as track patterns of collaboration among countries, institutions and among individual researchers. It provides information such as: trends of a country's share in the worldwide activities in science and technology, 'impact' of a country in the scientific community (national and international), information on international collaboration, international programmes, and the structures of scientific disciplines and their relations with others, etc.

The word "Bibliometrics" first appeared in print in 1969 in Alan Pritchard's article "Statistical Bibliography or Bibliometrics?" (Pritchard, Alan, 1969)¹ Pritchard's article was the result of his judgment that the expression "statistical bibliography" should be replaced with a better
term. The term statistical bibliography is clumsy, not very descriptive, and can be confused with statistics itself or bibliographies on statistics. Pritchard suggested that the word “BIBLIOMETRICS, i.e. the application of mathematics and statistical methods to books and other media of communication” be substituted for “statistical bibliography”.

The term ‘Bibliometrics’ is a combination of two non-English words, that is, ‘biblio’ and ‘metrics’. The first ‘biblio’ is derived from the Greek/Latin word biblion, which means book, while the other word ‘metrics’ indicates to the science of meter, that is, measurement. The later has also been adopted from a Greek/Latin term ‘metrics’ or ‘metrikons’, both having the same meaning (Webster’s, 1981). Therefore, bibliometrics is adopted into English and connotes the science of measurement pertaining to books, research papers, journals or any sort of documents.

Alan Pritchard, who first used the word “bibliometrics,” described it as the “application of mathematics and statistical methods to books and other media of communication” (Pritchard, Alan, 1969). This was paraphrased by Robert A. Fairthome as ‘quantitative treatment of the properties of recorded discourse and behavior appertaining to it. In a later article, “Bibliometrics and Information Transfer,” Pritchard explained bibliometrics as the “metrology” of the information transfer process and its purpose is analysis and control of the process. He based his interpretation upon the fact that measurement is “the common theme through definitions and purposes of bibliometrics” and the things that we are measuring when we carry out a bibliometric study are the process variables in the information transfer process. (Pritchard, Alan, 1972)
The British Standard Glossary of Documentation, of Terms explained bibliometrics as the study of the use of documents and patterns of publications in which mathematical and statistical methods have been applied, (British Standards Institution 1976)^5 definition is basically the same as Pritchard’s original definition.

William Gray Potter, editor of the issue of Library Trends devoted to bibliometrics, followed suit with “bibliometric is simply put, the study and measurement of the publication patterns of all forms of written communication and their authors. (Potter, W.G., 1981)^6 In the same issue, Alvin M. Schrader said it even more simply, “Bibliometrics, is the scientific study of recorded discourse. (Schrader, A.M., 1981)^7.

Bibliometrics, called a quantitative science, is divided into two areas, descriptive and evaluative.

In one of these classes is included the study of the number of publications in a given field, or productivity of literature in the field for the purpose of comparing the amount of research in different countries, the amount produced during different periods, or the amount produced in different subdivisions of the field. This kind of study is made by a count of the papers, books and other writings in the field, or often by a count of those writings which have been abstracted in a specialized abstracting journal. The other includes the study of the literature used by research workers in a given field. Such a study is often made by counting the references cited by a large number of research workers in their papers.

The two areas may also be divided as follows:

1. Productivity Count (descriptive)
a. Geographic (Countries)

b. Time periods (Eras)

c. Disciplines (Subjects)

2. Literature Usage Count (evaluative)

a. Reference

b. Citation (Stevens, R. 1953)

Nicholas and Ritchie divided the two groups as “those describing the characteristics or features of a literature [descriptive studies] and those examining the relationships formed between components of a literature [behavioral studies].”

Although all the descriptive studies are not evaluations, all the evaluative analyses are first descriptive with the evaluative analyses taking the data one step further, providing “data on the condition or character of the literature as a whole” (Nicholson, David; Maureen Ritchie, 1978).

5.2 DEFINITIONS OF THE BIBLIOMETRICS/SCIENTOMETRICS

(i) (Raising, 1962) has defined bibliometrics as “the assembling and interpretation of statistics relating to books and periodicals. Use of books and journals to ascertain its many local situations the generals use of books and journals”.

(ii) (Pritchard, 1969) has defined bibliometrics as “the application of mathematical and statistical methods to books and other media of communication”, which is the most commonly quoted
definition and it has served as the foundation stone and basis for many investigation.

(iii) According to (Fairthorne, 1969)\(^{12}\) bibliometrics is the “Quantitative treatment of properties of recorded discourse and behavior appertaining to it.”

(iv) (British Standard Institution, 1976)\(^{13}\) defines the term bibliometrics as: “the study of the use of documents and patterns of publication in which mathematical and statistical methods have been applied”.

(v) (Hawkins, D.T, 1977)\(^{14}\) in his on line bibliometric study interpreted bibliometrics as “Quantitative analysis of the bibliographic feature of a body of literature.”.

(vi) (Nicholas and Ritchie, 1978)\(^{15}\) opined that: “Bibliometrics...Provide information about the structure of knowledge, and how it is communicated.” They further added that “bibliometric studies fall mainly into two broad groups – those describing the characteristic or features of a literature (descriptive studies), and those examining the relationship formed between components of a literature (behavioral studies)”.

(vii) (Potter, W.G.1981)\(^{16}\) in his introduction to bibliometrics in Library Trends he has stated that: “Bibliometrics is, simply put, the study and measurement of the publication patterns of all forms of written communication and their authors.”

(viii) (Sengupta, I.N, 1985)\(^{17}\), in his study has covered the meanings of bibliometrics as given by different authorities.
Taking into account all the facts of bibliometrics he defined bibliometrics more explicitly as: “Organization classification and quantitative evaluation of publication patterns of all macro- and micro-communications along with their authorships by mathematical and statistical calculus”.

5.3 RELATED TERMS OF BIBLIOBETRICS:

Many sciences and social sciences are ending by suffix metrics and bibliometric is just one of them. Many scientists have been using the term ‘bibliometrics’ under different names, but the concepts anyway are the same. Some of these well known terms are Psychometrics, Econometrics, Informetrics. Scientometrics and Bibliometrics are closely interrelated where mathematical and statistical calculus have been systematically applied to study the problems of library science (Mahashwarappa, 1997)\(^\text{18}\). Some of these terms are described here under:

5.3.1 Scientometrics as a term was first defined and suggested in 1969 by two Russian scholars, namely, Nalimov and Mulchinko, accordingly, scientometrics is a complex of quantitative methods that are used to investigate and study science’s process. But scientometrics was introduced and came into prominence with the founding of the journal called “Scientometrics”, by T. Braunin in 1977, originally published in Hungary and currently from Amsterdam. On the other hand, the term scientometrics is derived from the Russian 'naukometria' and was used mainly in the East and is defined as the study of the measurement of scientific and technological progress (Egghe, 1988b)\(^\text{19}\).

Scientometrics is "the study of the measurement of scientific and technological progress" (Garfield, 1979b)\(^\text{20}\) its origin is in the quantitative study of science policy research, or the science of science,
which focuses on a wide variety of quantitative measurements, or indicators, of science at large. Typically input and output of science programs correspond to two major categories of indicators, the first, scientometrics deals mainly with science policy applications, and the second, scientometric research has a strong application-oriented tradition (Garfield, 1979b)\(^21\), (Raan, 1997)\(^22\). For example, scientometric studies may help governments and private sectors, identify their competitive edges, make strategic plans for future research areas and allocate research funding to key research areas, identified several publications appeared in the 1970s and contributed to the development of scientometrics (Garfield, 1979b)\(^23\).

5.3.2 Informetrics: this term was suggested by German scientists A. Blackert and S. Z. Zygel in 1982, as a newly formed branch of science. Informetrics studies all quantitative aspects of mathematical, statistical, and probabilistic of information and information flow, including the modeling of library management, in the latter sense, the research area usually is called bibliometrics. Therefore, (Potter, W.G., 1981)\(^24\) states that: “Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists”. We harbour no strong feelings about the vagueness of the term, but would like to ask the following quest: Do chemists and physicists quarrel about chemical physics? Is it important to determine whether a certain paper should be considered as a mathematical paper, an econometric one, or even an economics paper? Every new field has vague boundaries and even established fields such as physics and chemistry cannot be separated in a clear way. So, is not informetrics simply that which informetricians do? Therefore, many library scientists have advocated the use of the term informetrics, a term which takes cognizance of the fact that modern technology has imposed on us new
non-documentary forms of knowledge representation and of its transmission and dissemination.

5.3.3 **Librametrics**, this term belongs to the family of metrics, the word (Librametry) historically has appeared in the year 1948. The terminology was put forward by the great Indian library scientist (Ranganthan, 1969) He suggested using of mathematical and statistical techniques for the purpose of analyzing the library activities and library resources. As a matter of fact, in developing countries, no body is concerned or bothered about research output, and subsequently, the term ‘Librametry’ did not take its place in library and information science and was ignored and forgotten for many years, until later on it was rediscovered and called librametrics.

5.3.4 **Cybermetrics**: Recently, a new growth area in bibliometrics has been in the emerging field of webometrics, or cybermetrics as it is often called. Webometrics can be defined as study of quantitative aspects of webpages or nodes as using of bibliometric techniques in order to study the relationship of different sites on the world wide web. Such techniques can be utilized to map out scientific mapping in traditional bibliometric research areas of the Web that appear to be most useful or influential, based on the number of times they are hyperlinked to other web sites.

5.4 **SCOPE AND PURPOSE:**

Bibliometrics, which is the application of mathematics and statistical method to books and other media (Pritchard, 1969) has recently drawn the attention of serious researches in library and information science both theoretical and empirical. Theoretically, it is the quantitative features of the properties of recorded discourse. Quantitative characterization is the setting forth of probabilistically true
ideas about selected phenomena. Thus, the objective of bibliometrics as a scientific study is to produce ideas, that is, theory about recorded discourse and its various important properties. Since the practice is always backed by theory, bibliometrics also promises to provide practical knowledge benefits. This brings bibliometrics within domain of a scientific discipline.

The studies of bibliometrics are falling mainly into two groups, one describing the literature features (descriptive studies), and the others examining the relationship of literature, that is known as behavioral studies and sometimes referred to as citation studies. Descriptive studies are more interested in the condition of past literatures, specially in areas where the previous literatures of the on going research studies, are concerned with representing faithfully the characteristics present in the current literature. Behavioural studies are those which confirm themselves to describing the features of a document. Behavioural studies explore the full bibliometrics field and examine the relationships, the nature and strength of each relationship. (Nicholas and Ritchie, 1978)27.

The pioneer work by (Hulme,1923)26 entitled “Statistical Bibliography in Relation to Growth of Modern civilization", clearly states the purpose of bibliometrics. Hulme sought to identify the shape and period of movements in the development of science by examining patterns in the literature (Nicholas and Ritchie, 1978)29. Pritchard defines the purpose of bibliometrics, “To shed light on the process of written communication and of the nature and course of development of a descriptive (in so far as this is displayed through written communication) means of counting and analyzing the various facets of written communication.” (Pritchard, 1969)30.
Here, we summarize the purposes of bibliometrics as follows:

(i) To shed light on the processes of written communication and of the nature and course of development of a discipline (in so far as this is displayed through written communication), by means of counting and analyzing the various facets of written communication.

(ii) The assembling and interpretation of statistics relating to books and periodicals...to demonstrate historical movements, to determine the national or universal research of books and journals, and to ascertain in many local situations the general use of books and journals.

(iii) To identify and find out forms of literature, as bibliometrics provides a unique identification.

(iv) Preparation of ranked journals and make comparison between them.

(v) To identify the country with a vast and greater literary output.

(vi) To find out the chronological scattering order of all cited literature.

(vii) To identify the authorship and its trends in documents of different subjects, since the author's influence on the publication is significant and obvious.

(viii) To ascertain the utilization of the range of languages employed in the published output on a given topic.
5.5 IMPORTANCE OF BIBLIOMETRICS / SCIENTOMETRICS:

A bibliometric environment encourages awareness of thinking, planning a sharing platform of citation and research among the teaching faculty, students, library media specialists and researchers of many field. In the creation of a bibliometric environment, faculty members (as in the present study) and researchers who cite many works, may monitor and apply their knowledge deliberately modeling bibliometric methods to assist them in becoming aware of many if not all the research work in the field of bibliometric and its related metrics family. Problem solving and research activities in all subjects provide opportunities for developing citations that help in further scientific works in the future. A researcher, say a faculty member needs to focus the attention on how tasks are accomplished. Process goals, in addition to content goals must be established and evaluated with bibliometric citation so as to discover that understanding and transferring thinking process and improving results.

In this rapidly changing world, the challenge of scientific research is to help in many aspects of life and assist in developing skills which will not become obsolete. Based on much extant research on bibliometric, a scheme for Science and Technology showing the importance of our study may also be developed. Bibliometric studies offer an interesting view of the scientific activity of any country, being basic research to obtain information for decision-making by those responsible for formulating scientific policy. The most important objectives of bibliometrics are the study of size, growth and distribution of scientific documents, and also investigation of the structure and dynamics of the concerned groups that produce and use these scientific
data and documentation. Therefore, the study of production circulation, consumption and impact of publications is the focus of bibliometrics.

Bibliometrics has come forward as an essential and well-established technique that covers a wide and huge area of knowledge, which provides for a more practical task. In the present era, Bibliometrics is attaining more sophistication and complexity having national regional and international interdisciplinary character. It has established itself as a variable and distinctive research technique of studying Science of Science based on bibliographic information. The backbone of Bibliographic technique lies in its sound theoretical foundation laid by some great pioneers, namely, (Hulme, 1923)\textsuperscript{31}, (Lotka, 1926)\textsuperscript{32} (Bradford, 1934)\textsuperscript{33} (Zipf, 1949)\textsuperscript{34} (Pritchard, 1969)\textsuperscript{35}, (Garfield, 1971)\textsuperscript{36} (Sengupta, 1985)\textsuperscript{37} and many others librarians or library scientists, belonging to different fields of knowledge.

The data analysis of both citations and volume of publications can be useful in planning bibliographies through statistical and tabular methods. As a matter of fact, bibliometrics provides the researchers with the needful information about the structure of knowledge, its classification studies put forward the information that related to the subject, languages, and country-wise relationship, based on literary warrant. Therefore, bibliometrics is very useful in any field of research study.

5.6 BIBLIOMETRIC INDICATORS:

Bibliometric indicators can give signals about what is going on in the research system, but results have to be interpreted to the complexity of the studied environment. Indicators compare the quantity, quality and visibility of research implemented by various individuals or institutions.
5.6.1 Sources:

The collection of necessary data (publications or patents) is the basis for the construction of bibliometric indicators. The data could be obtained by making one's own database (by using CD-ROM, etc) or by (commercial) database such as chemical abstracts. Science Citation Index, Medline, Derwent for patents etc.

5.6.2 Publications:

Publication counts give a rough measure of the quantity of work produced by a research team or facility.

5.6.3 Citations:

Authors cite each other's papers for many reasons, but analysts generally assume that rates of citation provide some measure of the quality, relevance or interest, the impact of the cited paper. Thus, citation counts are an indicator of the influence a research has had on the larger scientific community.

5.6.4 Mapping (clustering):

Bibliometric maps provide an instrument which can be used optimally in an electronic environment. Moreover, there is a large amount of detailed information 'behind the maps'. Hence it is of crucial importance that this underlying information, particularly about research performance, can be retrieved in an efficient way, to provide the user with a possibility of exploring the fields and of judging the usefulness of maps against the user's own expertise. Advanced internet based user interface facilities are necessary to enable this further exploration of the maps and of the data 'behind the maps' (Noyons, 1999). Thus bibliometric maps and their internet based user-facilities will enable
users to compare the scientific performance of groups/institutes with other ‘benchmark’ institutes. Likewise, the maps can be used for the selection of benchmark institutes, for instance institutes chosen by the experts.

Co-citation analysis provides an alternative type of mapping, but it unavoidably depends on the availability of citation (reference) data and thus its applicability is less general than concept similarity mapping. Co-citation maps are based on the number of times two particular articles are cited together in other articles. The development of this analytical technique is based on the pioneering work of (Small, H., 1973). When aggregated to larger sets of publications, co-citation maps indicate clusters of related scientific work (i.e. based on the same publications, as far as reflected by the cited literature). These clusters can often be identified as ‘research specialties’ (McCain, 1990). Their character may, however, be of a different kind compared with co-word based clusters; because they are based on citation practices they may reflect cognitive as well as social networks and relations (Braam et al, 1991). Moreover, citations only reflect a part of the intellectual structure, and they are subject to certain, often field specific, time lag.

Mapping is used to visualize the structure of research fields or specialties (sub-fields) of science, and linkages of countries in the international arena. Following are some of the methodologies, which enable mapping or clustering:

5.6.5 Co-citation analysis

Co-citation analysis monitors the number of times (two) papers are cited together in single articles or patents in a particular field of S/T. the resultant co-citation clusters reflect the subject similarity and depict the most important works in that specialty (measured by citation). This
map enables one to monitor how specialties (or sub-fields) evolve over time.

5.6.6 Co-word analysis:

Co-word analysis monitors the number of times key words are mentioned together in publication or patents in a particular field of science or technology. A map of these paring co-words describes the structure of a research field.

5.6.7 Interviews:

Interviews monitor the numbers of times key words are mentioned by Interviewed scientists during an interview on a specific subject. A map of these key words describes the structure and sub-structure of a research field.

5.7 BIBLIOMETRIC LAWS:

The bibliometric laws in this chapter are the equivalents to Newton's laws of physics. The classics of their field, built upon by others, but not perfectly correct in certain situations. Like physical laws, they seek to describe the working of a system by mathematical means. Even though bibliometric scholarship is mature, there is little evidence of any Einsteinian breakthroughs that prove bibliometric laws are concrete laws. However, they are incredibly useful in developing general theories about information and provide data to study further.

Three regularities occur in bibliometrics to which have been given the name "law": Lotka's law of Scientific Productivity (authors publishing in a certain discipline); Bradford's law of scattering (scattering of articles); Zipf's Law of Word Occurrence (ranking of word frequency). (Booksten, A., 1979) claims, that the three laws are basically the same.
One of their differences lies in the type of data. Lotka's law dealt with authors publishing and the number of papers published; Bradford observed the scattering of articles on specific subjects in various journals; Zipf counted frequencies of words. (O'Conner, D; Voos, H., 1981)\textsuperscript{43}.

The three most commonly used laws which laid the foundation of bibliometrics are:

(i) Lotka's law of scientific productivity,

(ii) Bradford's law of scatter, and

(iii) Zipf's law of word occurrence,

We explain them in detail as follows:

5.7.1 LOTKA'S LAW:

Alfred J. Lotka was a mathematician, supervisor of mathematical research in the Statistical Bureau of the Metropolitan Life Insurance Company from 1924 to 1933 (Word Who's Who in Science, 1968)\textsuperscript{44}. It was during this time, that his definitive work, later called Lotka's law was produced. His investigation was a productivity analysis (described in the preceding section). Counting names and the number of publications listed for each, the coverage was for only A and B names in Chemical Abstracts from 1907 to 1916 and for Auerbach's Geschichtstafeln der Physik from its beginning through 1900. The data were tabulated and plotted, from which Lotka developed a "general formula for the relation... between the frequency y of persons making x contributions" as

\[ X \times Y = a \text{ constant} \]
Finding the value of the constant when \( n = 2 \), he observed that: the number of persons making 2 contributions is about one fourth of those making one; the number making 3 contributions is about one-ninth, etc; the number making \( n \) contributions is about \( \frac{1}{n^2} \) of those making one, and the proportion, of all contributors, that make a single contribution, is about 60 per cent. (Lotka, A.J., 1926)\(^{45}\)

Lotka's observation deals with the least number of productions.

Since the publication of Lotka's original article in 1926 much research has been done on author productivity in various subject fields. The publications arising from this research have come to be associated with Lotka's work and are often cited as proving or supporting his findings.

Lotka's law, in its generalized form, seems applicable.

(i) When we consider the publications of authors in one periodical and

(ii) When we consider all the publications of the authors in various journals, the observed values deviate considerably from the predictions of the law. This may be due to the differences between pure and applied science. (Radhakrishnan, T.; Kernizan, R., 1979)\(^{46}\).

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. We see a general decrease in performance among a body of authors following \( 1:n^2 \). This ratio shows that some produce much more than the average which seems agreeably true for all kinds of content creation. However, Lotka doesn't take impact into
account, only production numbers. Furthermore, in 1974, Voos found that in Information Science, the ratio was currently.

\[ 1:n. \]  
\( (Voos, 1974) \)

Thus, we can say that Lotka's Law may not be constant in value, but in following inverse square. Our challenge will then be to find the correct exponent in different mediums and fields.

Table: 5.1 RANKING OF AUTHORS

<table>
<thead>
<tr>
<th>No. of authors</th>
<th>No. of articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

5.7.2 BRADFORD'S LAW OF SCATTERING:

Samuel Clement Bradford was a very determined and dedicated person. His thinking was undoubtedly influenced by Paul Otlet and Henri LaFontaine, who organized the First International conference on bibliography held in Brussels in 1895. The theme of the conference was the need for international cooperation to develop a universal index which would recognize the requirement of a standard subject classification to be backed by a central universal library. This idea was staunchly supported by Bradford, as shown in the following paragraphs.
In 1927 “Bibliography by cooperation” appeared in Library Association Record (LAR) wherein Bradford indicated his concern for the accelerated accumulation of “useful scientific and technical literature”. He included in his paper a list of science subjects and the number of bibliographic references for each to alert readers as to what the science Library, of which he was deputy keeper, was doing to develop into an “information service covering the whole field of Science and Technology. At that time, Bradford claimed the number of references assembled, covering many aspects of Science and Technology, as 1, 212, 00. (Bradford, S.C.,1927)\textsuperscript{48}.

This article was followed by his another article titled “The necessity for the Standardization of bibliographical Methods” in which Bradford explained the Science Library’s method of classification and again stressed the need for cooperation. By the time of the introduction of the paper, a form of interlibrary loan had already been practiced for 2 years by the Science Library, with requests having been received from a number of other countries. (Bradford, S.C.,1928)\textsuperscript{49}.

One of Bradford’s hypotheses was that “references are scattered throughout all periodicals with a frequency approximately related inversely to the scope. On this hypothesis, the aggregate of periodicals can be divided into classes according to relevance of scope to the subject concerned, but the more remote classes will, in the aggregate, produce as many references as the more related classes.

Observations of the tables evinced three “rough” zones or groupings which Bradford’s graded as “1) Those producing more than 4 references a year. 2) Those producing more than 1 and not more than 4 a year. 3) Those producing 1 or less a year. Bradford found that “The groups thus produce about the same proportion of references in each
case, and the number of constituents increases from group to group, by a multiplier which though by no means constant, approximates fairly closely to the number 3, especially for the two larger groups.

From his data, Bradford constructed two graphs, plotting the logarithms of cumulated number of journals in relation to the cumulated number of references for each, geophysics and lubrication. He noted that “the later portion of each curve is remarkably close to a straight line,” and observed that “the aggregate of references in a given subject, apart from those produced by the first group of large producers, is proportional to the logarithm of number of sources concerned, when these are arranged in order of productivity. with this observation in mind, Bradford constructed a second graph or diagram. This diagram he used to develop an algebraic relation, but only for the “straight” part of the curve noted originally, from this he deduced his “law”.

The law of distribution of papers on a given subject in scientific periodicals may thus be stated;

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 the subject and several groups or zones containing the same number of articles as the nucleus, when the numbers of periodicals in the nucleus and succeeding zones will be as :

$$1: n: n^2$$

(Bradford, S.C., 1934)
S.C. Bradford also plotted graph of the cumulative number of source items R(n) verses the logarithm values of the cumulative number of journals (log n) such a graph, is sometimes called as Bradford's Bibliograph.

The graph being as a rising curves API, and then continues as a straight line. The rising part of the graph represents the nucleus of highly productive journals. The points P1, P2 and P3 on the bibliographs are the boundaries of three equi-productive zones in which the same number of years.

This was the first statement of what was later to be called the verbal part of Bradford, Law, Usually referred to as Bradford's law of distribution of Scattering, "sometimes this regularity is also called the law of dispersal of publications" (Yablonsky, A.I., 1980).
Contribution of B.C. Vickery and B.C. Brookes

Bradford’s work was extended by many, notably Vickery and Brookes. It was (Vickery, B.C., 1948) who realized that Bradford in fact formulated two laws, in that his verbal description is dissimilar from his graphical interpretation. The former describes his theory while the other traces his observations. Brookes’ work (Brookes, B.C. 1969) has been on the graphical version of the distribution; and the work produced by this method appears to fit empirical data more accurately than the mathematical versions of Bradford’s verbal theory.

The production of a bibliography involves the use of semi-logarithmic scales, which turn the hyperbolic curves of linear scales into straight lines.

Journals are ranked in decreasing order of the number of relevant papers. The cumulative number of papers is then derived. The rank of journal is plotted on the horizontal semi-log axis, and the cumulative number of papers is plotted on the vertical linear axis.

Once the bibliograph has started to form a straight line it can be continued by rule, and an estimate made of the end point.

The plot will have two distinct sections, the rising curve and then the linear section. The demarcation of the core is which is the section prior to the linearity, and the intercept of the linear section with the horizontal axis.

5.7.3 ZIPF'S LAW OF WORD OCCURRENCE

George Kingsley Zipf (1902-1950) was a Professor of Linguistic at Harvard. He developed and extended an empirical law, which was named as “Zipf’s Law of Word Occurrence”. This law is of the
observation that frequency of words within a text or occurrence of some event (P), as a function of the rank (i) when the rank is determined by the above frequency of occurrence, is a power-law function Pi ~1/i^α with the exponent α close to unity.

The most famous example of Zipf's law is the frequency of English words. The second example Zipf showed in his book was the population of cities (or population of communities). The population of the city as plotted as a function of the rank (the most popular city is ranked number one, etc) is a power-law function with exponent close to '1'. The income or revenue of a company as a function of the rank is also an example of the Zipf's law (also in Zipf's book). This should also be called the Pareto's law because Pareto observed this at the end of the last century. As it has been observed, governing a relation between the rank of a word and the frequency of its appearance in a long text if 'r' is the rank of a word and 'T' is its frequency, then mathematically Zipf's law can be stated as follows:

\[ rf = c, \quad \text{Where 'c' is a constant,} \]

More precisely the Law states that in a relatively lengthy text, if you "list the words occurring within that text in order of decreasing frequency, the rank of a word on that list multiplied by its frequency will equal a constant".

The equation for this relationship is:

\[ r \times f = k \] (where r is the rank of the word, f is the frequency, and k is the constant)" (Potter, 1988).^{54}

Zipf illustrated his law with an analysis of James Joyce's Ulysses. "He showed that the tenth most frequent word occurred 2,653 times, the
hundredth most frequent word occurred 265 times, the two hundredth word occurred 133 times, and so on. Zipf found, then that the rank of the word multiplied by the frequency of the word equals a constant that is approximately 26,500".

The following table shows distribution of words inversely proportional to the frequency of occurrence of the word:

Table: 5.2 RANKING OF WORD OCCURRENCE

<table>
<thead>
<tr>
<th>Rank (r)</th>
<th>Frequency (f)</th>
<th>Product (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>133</td>
<td>399</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>400</td>
</tr>
</tbody>
</table>

Zipf's Law, again, is not statistically perfect, but it is very useful for indexers. His law states that in a long textual matter if words are arranged in their decreasing order of frequency, then the rank of any given word of the text will be inversely proportional to the frequency of occurrence of the word.

Thus, these three laws are respectively based on

(i) Number of authors contributing in a discipline or other field;
(ii) Distribution of articles in a set of journals; and
(iii) Ranking word frequency in a particular set of documents. Other, albeit more specific, laws can be used for specific purposes. When applying these laws to bibliometrics and scientometrics analysis, implementation details can be easily imagined.
There are many other laws of bibliometrics and its related family, which have been put forward by many researchers and librarians the world over. It will therefore, be worthwhile to highlight some of them as an accepted matter of fact, as follows:

**Square Root Law of Price States:** "Half of the scientific papers are contributed by the square root of the total number of scientific authors" (Price, (Derek J de Solla, 1971).\textsuperscript{55}

**Garfield in his Law of Concentration argues:** "A basic concentration of journals is the common core or nucleus of all fields." (Garfield, Eugene, 1971).\textsuperscript{56}

**Sengupta** In his new bibliometric law, comments: "During phases of rapid and vigorous growth of knowledge in a scientific discipline, articles of interest to that discipline appear in increasing number in periodicals distant from that field." Mathematically this law stands in the following form:

\[ F(x + y) = a + b \log(x + y) \]

Where, \( f(x + y) \) stands for the cumulative number of the references in the first \((x + y)\) most productive journals, \( 'x' \) indicates number of journals in the same discipline and \( 'y' \) stands for the journals of related discipline, while, \( (y > x) \), and \( 'a' 'b' \) are parameters which stand for constants.

Bibliometrics offers the library and information field much and a lot to come. The research work by Lotka, Bradford and Zipf, indeed, is valuable in helping librarians and researchers who assess patterns of authorship, identify core collections and design better retrieval systems (Sengupta, I.N., 1973).\textsuperscript{57}
5.8 APPLICATIONS OF BIBLIOMETRIC LAWS

One of the main areas in bibliometric research concerns the application of bibliometric laws. The exponential growth of literature and rapid development of libraries generated several evolutionary studies about the effectiveness and efficiency of information services. These studies led to the identification and application of appropriate quantitative measuring techniques known as bibliometrics. Libraries and information managers all over the world began to use bibliometrics techniques in their day to day administration. These bibliometrics studies throw light on the pattern of growth of literature, inter-relationship among different branches of knowledge, productivity and influence of authors, pattern of collection build up, their use and so on. Day by day bibliometrics is attaining inter disciplinary character and sophistication.

Another useful idea is, suppose there is an endemic outbreaks occurring from time to time. Research and reference can follow cycles of use, this is very similar to ideas like memes and paradigm shifts. All of the above bibliometric techniques can be used to test for this theory. Software can be used to study document representations, references, and citations. We can then compare a matrix of values to find correlations. We can also use a technique like this to help design future documents by using like terms, references, and citations to establish stronger linkage relationships among documents.

(i) To quantify research and growth of different area of knowledge:

(ii) To estimate comprehensiveness of secondary periodicals.
(iii) To identify users and authorship of documents of various subjects;

(iv) To measure usefulness of ad hoc and retrospective SDI services:

(v) Experimental models correlating or bypassing the existing models;

(vi) Identification of core journals in different disciplines to formulate a need based acquisition policy within the limited budgetary provision without detriment to the research interest of the parent organization;

(vii) To initiate effective multilevel network system;

(viii) To regulate inflow of information and the communication; and also

(ix) To develop norms of standardization; and so on.

5.9 LIMITATIONS IN APPLICATION OF THE LAWS

The studies of bibliometrics' family left no doubt of their usefulness in achieving wonderful results, here, bibliometric studies are very helpful to library and information science research users, but there are many short-comings. These limitations are also in their application. Here, we highlight some of them. Bibliometrics, scientometrics data are likely to under represent the level of international collaborative research, a further limitations exists in determining just how much work an individual author contribute to a coauthored publication, Wilson and Osareh, (2003)\(^{58}\). It has been argued by O'connor and Voos (1981)\(^{59}\) that bibliometrics: "is merely a method or if it meets the test of a theory
in its ability to explain and predict phenomena. This paper examines the properties of bibliometric distributions in a non-technical manner", and they have mentioned that bibliometrics has largely been used to describe bibliography and is not yet able to explain and predict them, then, it is merely a method and not a theory.

5.10 CONCLUDING REMARKS

This chapter has examined the potential benefits of applying bibliometrics to finding out more about scientometrics usage. As we gain wisdom about behavior of these laws, in chapter six, we are going to apply these laws in science and information technology at the faculty teaching level of the Jordanian Universities. (O'connor and Voos,1981)\(^6\) have mentioned that bibliometrics has offered us by: “the promise of providing a theory that will resolve many practical problems. It is claimed that patterns of author productivity, literature growth rates and related statistical distributions can be used to evaluate authors, assess disciplines and manage collections”. Therefore, information discovered through the application of bibliometrics laws and techniques has the potential of saving, time money and other things, provide more appropriate programmes, and meet more of the users and researchers needful information resources collections.

Bibliometric and scientometrics along with the related other terms can provide data to justify the difficult decisions and funding requests by library administrators, and also this family of metrics can inform the processes and products of knowledge management that have grown in importance within contemporary organization and educational institutions.

Scientometrics/bibliometrics terms are the analysis of the structure of literature using various tools, counting, rank-frequency
distributions, and vitiation analysis, although the structure of literature is basic to all disciplines. The studies of metrics family, specifically bibliometrics and scientometrics have left no doubt of their usefulness in achieving wonderful results. Here we believe that bibliometrics studies are very helpful to library and information science research users, but there are many limitations in the applications of their laws.

Though most of the research studies sided with the utilization of Bradford scattering distribution, some others researchers could not find satisfactory outputs. Some researcher found that Lotka's law is perfect in most of the cases of especially to sciences however, the value of the index 'n' was found to vary for different groups of scientists; these limitations have left many researchers questioning these laws, (Vijaykumar, 1997).
REFERENCES:


15. Nicholson, David and Mauren Ritchie (1978), ibid., p. 9


21. Ibid, p. 5


27. Nicholson, David and Mauren Ritchie (1978) ibid., p. 15

29. Nicholson, David and Mauren Ritchie (1978) ibid., p. 15
38. Noyons, E (1999)"Bibliometric mapping as a science policy and research management tool (Ph.D. Theses: Leiden University)


60. Ibid, p10.