CHAPTER 4
CITATION ANALYSIS: AN OVERVIEW

4.0 Introduction
4.1 Definition
4.2 Scope and purpose of citation analysis
4.3 Functions of citation analysis
4.4 Bibliometric Laws
4.4.1 Lotka’s inverse square law of scientific productivity
4.4.2 Bradford’s law of scattering
4.4.3 Zifp’s law of word occurrence
4.5 Important terms and their meaning
4.6 Limitations of citation analysis
4.0 Introduction:
Citation pattern or the analysis of citation pattern is one of the major and popular branches of bibliometrics. Bibliometrics has established itself as a viable and distinctive research technique for studying science of science based on citation data. Analysis of citation pattern or citation analysis is one of the most important bibliometric techniques involving analysis of the references forming part of primary communication. Citations are the formal explicit linkages between publications that have particular points in common.

The word “Bibliometrics” is although of not so old origin but bibliometric studies were performed much earlier since the beginning of the 20th century. Recently bibliometrics is applied as a popular technique by the information scientists and the librarians to evaluate various kinds of library services and to study the characteristics of subjects and nature of citations in different citations in different branches of knowledge. This technique was evolved from the efforts of earlier 20th century documentalists to apply mathematical and statistical analysis to bibliometric units. The validity of data obtained through bibliometric studies continues to be debated despite general acceptance of the accuracy and objectivity of bibliometric techniques. The most basic bibliometric technique still in use involves counting and categorizing the publications as regard to their type, country of origin and ranking journals on the basis of their use by the scientists in a specific field of study.

The first recorded study on bibliometrics was presented in 1917 by F J Cole & N B Eales in Science progress. This project analyses publications in comparative anatomy by simply counting the number of titles, both books and journal articles. They mainly
studies the fluctuations of interest and distribution of literature among countries. This kind of study is named as “Statistical analysis of the literature.” In 1932 E W Hulme introduced the term “Statistical bibliography.” This is considered as second reported work on bibliometrics, where Hulme analysed the journal articles in “English International Catalogue of Scientific Literature” and derived ranking of countries by their productivity.

However, 1972, the first recorded study of citation data (being the third bibliometric study) was carried out by Gross & Gross. Finally Alan Pritchard is actually credited with introduction of the term “Bibliometrics:” in 1969 to replace the earlier word “Statistical bibliography” used in some concept. This was coined by Pritcherd in his paper “Statistical bibliography or bibliometrics” published in the Journal of Documentation (1969) (Mahapatra, 2000).

In the meanwhile, Dr. S R Ranganathan introduced the term “Librametry” in 1948 at the Aslib Conference held in Leamington Spa. He described the term to be used to study the various library operations by applying statistics. Russian concept “Scientometrics” and FID’s “Informetrics” were also applied almost for the same concept. In 1070’s B C Brooks used the term “Informetrics”, partly because of increasingly digitalization of scholarly communications and library like activities seemed increasingly too restrictive and partly as a candidate term to subsume both “Scientometrics” and “Bibliometrics”. Still then it was expected that the creation of the International Society for Scientometrics and Informatics (ISSI), at the fourth International Conference in Berlin in September 1993 would perhaps serve as the catalyst to standardize the terminology on “Scientometrics” and “Informetrics” (Guha,
1983).

After Pritchard’s initial work, bibliometrics was expanded in two dimensions that is quantitative and qualitative study of bibliometric phenomena. The quantitative analysis of bibliometrics, developed three basic bibliometric laws that is Lotka’s law (1926) predicting the productivity distribution of various authors, Zipf’s Law (1933) describing the word-frequency ranking and the Bradford’s law of scattering (1934) explaining distribution of documents (usually journals) in specific discipline or problem area. Some general objectives of quantitative analysis in numerical aspects of bibliometrics are as follows:

- Design of more economic information systems and networks
- Improvement of efficiency rates of information handling processes
- Identification and measurement of deficiencies in present bibliographical services
- Prediction of publishing trends
- Discovery and elucidation of empirical laws that can provide basis for developing a theory of information science

Qualitative applications of bibliometrics on the other hand emphasize practical utilization of research findings, such as:

- Identifying a core literature
- Ranking publications in zones of diminishing importance
- Establishing a transition point between zones of higher and lower utility
- Tracing the spread of ideas as a study of epidemics
- Classifying segments of literature through interconnection of citations
It is also reported that the discovery of empirical laws of bibliometrics has resulted in combination of the quantitative dimensions that has satisfied research needs in both areas (Mahapatra, 2000).

Bibliometrics is defined variously by different authors. No universally accepted definition still exists for this. As mentioned earlier Hulme (1923) first of all discussed the purpose of statistical bibliography where the bibliographic concepts were almost proposed. The purpose of statistical bibliography according to him, is to shed light on the process of written communication and of the nature and course of development of a discipline, by means of counting and analyzing the various facets of written communication.

On the other hand, statistical bibliography is the assembling and interpretation of statistics relating to books and periodicals to demonstrate historical movements, to determine the national or universal research use of books and journals. According to Pritchard, “Bibliometrics” is the application of mathematics and statistical methods to books and other media of communication. Fairthorne defined the same as quantitative treatment of properties of recorded discourse and behaviour appearing to it. Bibliometrics is also named as quantitative analysis of the bibliographic features of a body of literature. Nicholas and Ritche accepted the definition of bibliometrics as the statistical or quantitative description of a literature. According to Potter, bibliometrics is the study and measurement of publication pattern of all forms of written communication and their authors. Recently Egghe explained bibliometrics as the development and application of mathematical models and techniques to all aspects of communication.
“Scientometrics” which is sometimes also used for bibliometric like phenomena is defined as “quantitative evaluation and intercomparison of scientific activity, productivity and progress”. Brusilovsky explained “Scientometrics” in a more general way as “the study of the measure of scientific and technological progress”. The term “Scientometrics” which is primarily used in the USSR, deals with quantitative analysis of science viewed as a information process. The “Scientometrics” is further considered as bibliometric measurement for evaluation of scientific progress, level of scientific development, social relevance and impact of application of science and technology etc. The term “Scientometrics” was introduced and came into prominence with the founding of the journal named “Scientometrics” by T. Braunin in 1977, originally published in Hungary and currently from Amsterdam. The term “Scientometrics” was used to mean the application of quantitative methods to the history of science but it is now generally used as a generic term for a variety of research approaches within the study of science that a quantifiable aspect of science can be utilized to assess the characteristic of science. Marton and Garfield have defined it as the field of enquiry given over to the quantitative analysis of science and scientific field. “Informatrics” another recently used term “formalizes and consolidates measurement studies which focus on information productivity. It integrates information technology and complex intersections of information theory, cybermatrics, decision theory etc.” According to Brooker the term “informetrics” was first proposed by Otto Nacke of West Germany in 1979. It focused on information productivity. However, “Scientometrics” and “Informatrics” are included under the domain of “Bibliometrics” and comprises several very closely related and highly overlapping research areas. In general “Bibliometrics” therefore may be defined as the quantitative analysis of the characteristics, behaviour and productivity of all aspects of
written communication, library staff and information users.

4.1 Definition:

Studies based on citations are being generally termed as citation analysis. They may be termed as cito-analytical studies. Though citation based studies have been going on in the world since long, such studies, however, did not receive the necessary momentum till the advent of citation indexes. The advent of Science Citation Index in 1963, followed by Social Science Citation Index in 1973, and the Arts and Humanities Citation Index in 1978, considerably widened the vistas of citation-based studies and brought forth a variety of cito-analytical products. Over the past few decades bibliometricians all over the world also busied themselves in analyzing citations from numerous angles and thereby revealed certain laws and trends not known before.

Now citation analysis has become a powerful tool to objectively evaluate a research periodicals; determine the use of documents in a library; find out the active life of literature and also the important periodicals in various fields of knowledge. These help us to weed out the unused or least used documents and guide us to procure the important ones, and thereby develop a balanced collection filling up the yawning gaps. The systemic and logical analysis of citations depicts the impact a scientist or an institution has made in the world, identifies the workers or teams active in a field, locates the significant contributions, and so on (Sen, 1994).

Citation analysis is mathematical analysis of references or citations appended at the end of each scientific communication as an essential part of it. The author(s) of a paper customarily presents a bibliography or references as authentic source of information having research value or to substantiate the point of view of ideas.
expressed in the cited papers. Analysis of cited papers is used as a measure of impact of individual articles, periodicals, authors, etc. and has become an accepted practice in almost all scientific communications and a well established part of information research. A quantitative approach to the description of documents is gaining ground both in research and in practice. Citation analysis is a tool to identify the core references in a subject by counting the citations appended at the end of each scientific articles. It is basically a technique, which involves the process of collection, counting, analysis and interpretation of citations given in research writing and thereby helping in identification of significant sources of information (Kumar & Kumar, 2008)(Gopinath, 1985)(Nkiko & Adetoro, 2007)(Pichappan, 1990).

4.2 Scope and purpose of citation analysis:

Citation analysis reflects two major themes that is use of citations as tools for librarians to evaluate the library collection and services and use of citations as tool to analyse research activity. Citation analysis is concerned with the following phenomena:

- Which authors are most cited
- Which journals are most cited
- What linkages exist between the citing and cited works (i.e., self-citation)
- Language of documents selected for use as citation
- Subject distributions and how quickly the literature on some subject become out-of-date, i.e., obsolescence study
The basic purpose of acquiring, processing, maintaining and serving all kinds of reading materials is to satisfy the information requirements of users in any library or information resource centre. Systematic studies of information gathering habits and behaviour of the user community is important for improving the communication system. Such studies were termed by different names like ‘Statistical Bibliography’ and ‘Bibliography’, Reference counts popularly known as the technique of citation analysis has been employed as a bibliometric technique, for ascertaining the pattern of literature use in a case of large community. There are several applications of citation analysis, viz –

- to show pattern of relationship between subjects, language and countries.
- to shed light on the transmission of knowledge
- to show how the use of literature changes over time
- to establish ‘pecking orders’ of countries, languages or individual items especially periodicals
- to serve as a guide to coverage of literature by secondary services and retrospective bibliographies
- to aid librarians, documentalists and information workers in the selection of representative reading materials for acquisition
- to aid librarians in the identification of items (especially periodicals) for withdrawals

Period-wise distribution of citation and its findings may be relevance to librarians and information workers in framing sound acquisition policy objectively. It may also be of importance in planning and providing current and retrospective bibliographies and documentation lists. It may also provide librarians a rational ground against which
live collection and outdated collection of reading materials could be identified and separated for the sake of their maintenance services. (Mubeen, 1996) (Srivastava, 2002) (Gooden, 2001)

4.3 Functions of citation analysis:

Citation analysis has two broad areas of applications. One of them is suitable for library and information science professionals where citations are used as a tool to find out various features of literature use like type of documents, languages, age, country of origin, subject distribution, highly cited authors, journals, citation ratio, authorship pattern etc. The other kind of study of citations is mainly useful to derive the research activity in a field of study. Several applications of citation analysis are:

- “Literature of” study - Citations in a particular subject area are to be studied to describe the pattern of citation. The source for this kind of analysis may be limited to a single journal in a field, or they may cover many sources, including types of materials in addition to journals. The characteristics of cited materials are frequently examined in citation analysis which includes types, age, highly cited authors, journals, languages, countries of origin and subject of distribution, etc.

- “Type of literature” studies – Citation analysis can be used to gauge the dissemination of results reported in certain types of literature such as government documents, dissertations or the exchange of literature of regional scientific societies. The source of citations used for analysis clearly can determine the generality of one’s conclusions in this type of study.
User studies – The application of citation analysis technique in determining user needs are very much useful for collection development and design of library services. Here the term papers, thesis/dissertations or technical reports submitted to the organisations are verified and references are collected in order to determine type of materials, age of materials, subject, language and finally determines whether available locally in their own library or not. Citation analysis can also be used to compare user behaviour today with user behaviour several years ago.

Historical studies – Historical research using citation analysis is based on a literary model of scientific process. In this literary model scientific work is represented by papers written and published to report it and relationships between discrete pieces of work are represented by references in papers. The citations can be used to trace the chronology of events, relationships among them and their relative importance. The citation analysis may be used to elucidate the historical development of a subject.

Communication pattern – Citations may also indicate the scientific communication pattern and the kind of research carried in a field of study like solo or joint research etc. The analysis of citations reveal the problem of communications like linguistic isolation, limitations in dissemination of new ideas and barriers between basic and applied sciences or between specialists and the public in general.

Evaluative bibliometrics – In these studies, citation analysis is defined as the evaluation and interpretation of the citations received by articles, scientists, institutions and countries, This is used to measure the productivity and scientific influences.
Information retrieval – It is considered that “use of citation relation is perhaps having the greatest impact on information retrieval where citations have been used to augment more traditional approaches to literature searching.” It is also confirmed experimentally that citations are useful supplements as keywords in identifying relevant documents. Citation relations are also seen to be useful in developing document representations in automatic classification. The advantages of using citation as a retrieval tool are that they are not affected by changing terminology, they provide access to interdisciplinary literature and they show papers that are relevant to a subject not found by using conventional indexes. Researches were also done on the use of co-citation as a search strategy to retrieve documents relevant to a given topic using commercially available search systems and citation index database.

Collection development – Citation analysis is applied primarily to develop journal collections in a library and information centre. The citation analysis will help in deciding titles of journals to be acquired, whether to continue or discontinue subscription and to weed or not weed the back volumes of journals. Various ranked lists of journals have been derived from citation counts. Therefore, citation analysis technique can be used as one of the indicators in providing guidelines for collection development in a library and information centre (Mahapatra, 2000).

The primary function of citation is to provide a connection between two documents, one which cites and other which cited. The possible reasons are:

- Paying homage to pioneers
- Giving credit for related work
- Providing background reading
- Correcting one’s own work
- Correcting work of others
- Criticizing previous work
- Identifying original publications in which an idea or concept was discussed
- Describing the nature and scope of the printed document in which the statement is found
- Disclaiming work or ideas of others
- Disputing priority claims of others

Importance of citations may be stated as:

- Identification of key documents and creation of core list of journals
- Study of the coverage of primary journals and other materials in secondary service
- Clustering of documents according to common references and citations
- Study of attributes of literature including growth rate, obsolescence, citation practices
- Study of the structure of scientific literature according to language, country of origin, age, subject, form, authorship or any combination of these attributes
- Study of historical and sociological aspects of scholarly communication in science and technology
Besides these, citations are also used successfully as reading list and in preparation of bibliographies. The value of citations are well observed through Science Citation Index, Social Science Citation Index, Art and Humanities Citation Index. The citation of document implies the use of that document by the citing author. It assumes that the author refers to all, or at least to the most important document(s) in the preparation of his work and all the documents listed were indeed used. Therefore, citations are the indicator of literature use (Mahapatra, 2000).

The result of citation analysis would appear to be of great potential value in the management of library journal collections. Measures of citation frequency and impact factor should be helpful in determining the optimum of both special and general collections. Analysis of chronological distribution of items cited can serve as a guide in determining the optimum size of back files and since the data give a detailed view of each journal’s citation history, binding and retention schedules can be rationally established journal by journal, rather than for groups of journals. Another application, which harried librarians may welcome, is correlation of data on citation frequency and impact with subscription costs. Such a correlation can provide a solid basis for cost-benefit analysis in the management of subscription budget.

Individual scientists also face the problem of selecting journals to read and keep, as well as compiling reference and reading lists for themselves and their students. Although each of the relatively few journals that are very useful in a given discipline or specialty may be well known, it can be difficult to gauge the merits of the other journals in that discipline or specialty and to decide what journals to get and how long to keep them. It should be noted in this connection that analysis of citation frequency
and impact factor can be tailored to the specific interests and requirements of the individuals by restricting the number of citing journals to a small groups of familiar titles.

Editors and editorial boards of scientific and technical journals may also find citation analysis helpful. As it is, those who formulate editorial policies have few objectives and timely measures of their success. A wrong policy, or a policy wrongly implemented, may have serious effects on revenue and prestige, and the work of regaining readers and reputation can be difficult and expensive. Editors can find useful indicators of a journal’s performance in the extent of self citation, the number of time cited per year, and the distribution of citations among citing journals within and outside the specialty literature.

Perhaps the most important application of citation analysis is in studies of science policy and research evaluation. Soviet information scientists had used citation data to evaluate the implementation of science policy in the then USSR (Gerfield, 1972).

The evaluation of the use of library collection is a fundamental tool for the development of a relevant and cost-effective collection. Bibliometrics offers several methods to measure the level of use of collections. The management of journals in libraries is a difficult task, as subscription costs increase year after year, library budget are tight, new journals appear and new modalities of subscription are becoming available (Gomez, 2002).
Because of rapidly growing boom in the publication of literature like journals, books, govt. publication, proceedings, bulletins, etc. in the different fields it may not be feasible for a library to procure all the required materials for its potential readers within the limited resources. Therefore, the knowledge of comparative and up-to-date details of the literature is of great significance and to acquire this knowledge citation analysis is very much important (Biswas & Haque, 2008).

Cytoanalytical study of doctoral dissertations are the products of research activity form an important source of information. Such studies may be useful for acquisition of materials, provision of better services to patrons and knowing the location of materials. What part of literature is cited most, or how long the literature remains useful to readers and languages of most cited publications – knowledge of all these provides guidance to collection development policies, individual item selection and retention and binding decisions (Deo, Mohal & Survey, 1995) and the analysis plays a promontory role for easy identification of earlier research. It is one of the important areas of research in the field of library and information science.

Citation analysis is very often fruitfully applied to derive the following benefits:

- to lead the reader to further studies in the field
- for the preparation of bibliographies
- to study the use pattern of different types of documents
- to find out the relative use of different languages
- to study the use of literature from different countries
- to study the scattering of subjects
- to decide the obsolescence rate of documents in different subjects
- to determine the interdependence and lineage of subjects
to prepare the rank list of documents

➢ to study the rate of collaborative research

➢ for the analysis of scientific journals

• citation rate of a journal

• impact factor

• self-citing rate

• self-cited rate

• immediacy index

4.4 Bibliometric laws:

The three fundamental laws which laid the formation of bibliometrics are:

• Lotka's Inverse Square Law of Scientific Productivity

• Bradford’s Law of Scattering of Scientific Papers

• Zipf's Law of Word Occurrence

4.4.1 Lotka's Inverse Square Law of Scientific Productivity:

In 1962, Alfred J. Lotka proposed his inverse Square Law correlating contributors of scientific papers to their number of contributions. His law provided fundamental theoretical base for bibliometric studies involving authorships. He was interested in determining "the part which men of different calibre contribute to the progress of science." For this he checked the decennial index of 'Chemical Abstract' 1907-1916 and counted the number of names against which appeared 1,2,3, etc., entries. He tabulated the data for 6,891 names, beginning with letter 'A' and 'B'. Similarly the data from the Auerbach's Geschietftafeln der physik was also collected for the 1325 physicists. Lotka then plotted the graph on a logarithmic scale, the number of authors
against the number of contributions made by each author and he found that in each case the points were closely scattered about a straight line, having a slope of approximately two to one. On the basis of these data, Lotka deduced a general equation, for the relation between the frequency 'y' of persons making 'x' contributions as follows : $x^n \cdot y = \text{constant}$ and for the special case $n=2$, the constant is 0.6079. Further he summarised the result as follows:

In the case examined it is found that number of persons making two contributions is about one-fourth of those making one contribution, the number making 'n' contributions is about $1/n^2$ of those making one and the proportion of all contributions is about 60 per cent.

In otherwords, for every 100 authors contributing one article, 25 will contribute two articles, about 11 will contribute three articles and 6 will contribute four articles and so on. Though, the law was based on the study of chemistry and physics literature later it has generated much interest and attracted the attention of researchers and it has been applied and tested in many other fields.

4.4.2 Bradford’s Law of Scattering of scientific papers:

Samual Clement Bradford, one of the pioneers of bibliometrics, should be considered for his classic paper “Sources of Information on Specific Subjects”, which is the first paper published on observations on scattering. He examined two bibliographies prepared in the Science Library (Britain) on Applied Geophysics (1931-32) and Lubrication (1931-32) and he prepared lists of journals arranged by decreasing order of source items contributed by the journals to the bibliographies. He noticed that in
each subject there were a few very productive sources, large number of sources which were moderately productive and still a large number of sources of constantly diminishing productivity. The whole range of periodicals was thus seen as “a family of successive generations of diminishing kinship, each generation being greater in number than the preceding, and each constituent of generation producing inversely according to its degree of remoteness.”

In the list of periodicals ranked by diminishing productivity, Bradford identified three groups of periodicals that produced approximately the same number of articles on the subject, but the number of periodicals in these three equi-productive zones increased by a constant factor. Based on this he stated his law as follows : If scientific periodicals are arranged in order of decreasing productivity of articles on a given subject that 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 number of periodicals in the nucleus of the succeeding zones will be as : $1 : n : n^2$

Bradford also plotted graphs of the cumulative number of source items $R(n)$ versus the logarithm of the cumulative number of journals (log $n$). The resulting graphs for Applied Geophysics and Lubrication were similar to the graph shown in Fig 4.1. Such a graph, is sometimes called as Bradford bibliograph.
The graph begins as a rising curve $AP_1$ and then continues as a straightline. The rising part of the graph represents the nucleus of highly productive journals. The points $P_1$, $P_2$ and $P_3$ in the bibliography are the boundaries of three equi-productive zones in which the same number of articles as the nucleus (represented by $OY_1 = Y_1Y_2 = Y_2Y_3$) derived from an increasingly larger number of journals (represented by $OX_1$, $X_1X_2$ and $X_2X_3$). The law attracted the attention of many researchers in the field has been the main topic of many articles in the literature (Maheswarappa, 1997).

### 4.4.3 Zipf's Law of Word Occurrence

Zipf developed and extended an empirical law, as observed by Estoup, 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 $f$ is its frequency, then mathematically Zipf's law can be stated as: $rf = c$, where $c$ is a constant.
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 (Maheswarappa, 1997).

4.5 Important terms and their meaning:

4.5.1 Ranked List of Journals by Citations:

In the ranked list, journals are arranged according to the number of citations received by a journal over a particular period, may be one year or more. It is generally accepted that the larger the sample of study the more reliable is the rank likely to be. The most cited ones figure at the top followed by the less cited ones. This method, in certain cases, especially in the case of younger journals, may fail to provide the correct rank. This is because of the fact that apart from the quality the age and the number of source items published in a journal per year, also play an important role in deciding the rank. One can determine the ranked list analyzing the citations figured in the important journals of a subject being received in a library without depending on a database. Even this type of local study provides valuable indications as to the important journals in the field and thereby help in the acquisition of better periodicals (Sen, 1994).
4.5.2 Citation half-life and obsolescence:

The concept of half-life has been borrowed from the field of nuclear physics to illustrate the journal obsolescence. The citation half-life provides a quantitative measure of the rate at which the scientific papers become obsolete. The number of journal literature in science and technology grows in much faster rate. There also high probability of use of journal literature in the immediate years after their publication but their subsequent uses decreases rapidly as these journals become old. The half-life in its original context refers to the “time required for disintegration of one half of the currently cited literature was published. In an another view, half-life of literature is the time by which one half of the currently published literature become obsolescent. It is also related to the growth of literature. It is derived that the half-life in citation decay is same as the half-life for the growth of literature (Mahapatra, 2000).

4.5.3 Self-citation:

Self-citation is a common practice in the literature of all kinds. The self-citation of an author refers to those cited references which have been authored or co-authored by same citing author. Most of the researchers believe that these types of self-citations study should be included in the final tabulation of the bibliometric studies of self-citation. Now information scientists have paid more interest towards the study of the self-citation. Mainly there are three types of self-citation. They are the cases where (i) the citing author and the cited author the same, named as author’s self-citations (ii) reference to articles published in the same journal in which the citing article appears, named as journal self-citations (iii) it is used to indicate similarity of subject matter between citing and cited article.
4.5.4 Authorship studies:

Authorship studies mainly deal with kind of authors, nature and degree of collaboration among them and impact of citation rate on collaborative trend of authors. The authorship studies can be carried out both for citing and cited articles. The various kinds of authors appeared in publications are grouped as personal author works, corporate body and anonymous publications. Collaboration is a major area of authorship study, which indicates how authors work together to bring out a publication. An author may publish a work independently or he/she may participate with one or more number of authors to bring out a publication.

To conduct such authorship study, the authors of publications are the main element of study. The researchers after completion of a project or on completion of an area of study, write the result in research papers and publish them in form of articles in different journals and books. The data may further be sub-grouped as single author, two author cases, three author cases and more than three author cases. A grouping can also be made on the basis of their country of origin and organizations where they may be working.

4.5.5 Collaborative research:

Scientific writing is essentially a tool for information exchange. It is also considered as a medium for promoting self-interest on the road to recognition and for enhancing professional reputation in one's field of specialization. Pursuit of research in modern science requires inputs from various branches of knowledge. Thus interdisciplinary interaction amongst scientists is on the increase. The post-world war era has witnessed a phenomenal increase in organized collaborative research activities. The increase in multiple authorship and collaboration between researchers is an indication
of growing professionalism in different fields of science. Such collaborative efforts, in turn, are a natural reflection of the complexities, scale and cost of modern investigations.

4.5.6 Repetitiveness of citations and bibliographic coupling:
M M Kessler introduced the concept of bibliographic coupling. It is the number of common references cited in two documents that indicates the degree of similarity of contents of the citing papers. Two source documents containing a large number of common references are said to have a high coupling strength and are likely to be on the same topic. It is observed that the concept of relationship has certain drawbacks and not seem to be a valid unit of measurement because if two papers are citing a third paper, they may or may not be citing an identical piece of information of third paper being cited. Two documents are bibliographically coupled if their reference lists share one or more of the same cite documents. There are two criteria of coupling. The first was related to coupling strength and was measured by the links to a given test paper. Second was measured by links to one other.

4.5.7 Co-citation:
The concept of co-citation was for the first time suggested independently by Small and Marshakova almost simultaneously in 1973 and later developed by Small who proposed a new method of analyzing citations to generate clusters of related papers. The number of times two papers are cited together in subsequent literature determines the co-citation strength of two cited papers. Co-citation is dynamic measure in that co-citation strength of cited papers can be studied over a period of time as they continue to be cited together in subsequent literature. But one of the disadvantages of co-
citation technique is that it requires comprehensive citation data.

4.5.8 Subject dispersion:
Subject dispersion of citations means the scattering of subjects of citations. The authors refer to publications of various aspects. These may belong to their own subject or to some other subject, which are closely related to it. Therefore, analysis of these citations shows various sub-fields of a subject which are scattered over the citations.

To find out subject dispersion in a particular area, the citations appended in journal articles for some specific period are to be collected on slips of papers. Then each of the citations are to be classified as per their subject. Finally the groups are grouped together and counted.

4.5.9 Language-wise Distribution:
It is great significance for researcher’s to know the language in which their area of specialization is published. The analysis of the languages used to transmit the subject literature is useful not only to understand the coverage patterns, but also gives an indication of the foreign language problem likely to be faced by the users.

4.6 Limitations of citation analysis:
Review papers, on an average, are cited more than the research papers. Scientists writing a sizable number of review articles are likely to be cited more than those not writing, or writing a smaller number of review articles. Hence, this fact needs to borne in mind while evaluating a scientists, or comparing the performance of one scientist with others of his field.
A paper giving rise to any controversy is likely to be cited more, and there may tend to give false indications in case the number is large. It requires a few years, say five, for a paper to receive its reasonable quota of citations. Hence the importance of a paper cannot be correctly judged if the paper is less than five years of age. Similar is the case with new journals (Sen, 1994).

Sometimes the rate of self-citation is found very high; so what is the reason behind this abnormally high rate of self-citation? Is it characteristic of parochialism, eccentricity, mediocrity? Does it indicate that a particular field of study has as yet no basis for interaction with other fields?

What is significance of a wide, multidisciplinary spread of titles cited in the references of a given journal or group of journals? Is it a measure of multidisciplinary activity? And what is the significance of a wide disparity between the number of journals cited by a given journal and the number that cite it? (Gerfield, 1972).

The very logic of depending on citations which are given according to the whims and fancies of authors has been questioned by many researchers. They believe that any result that will obtained through a study of citation leads to erroneous conclusions. The common arguments labelled against citation analysis are the following:

- Negative citation, i.e., citing a paper just to repudiate it. The result is that controversial papers will get more citations than really worthwhile papers.
- Too much of self-citation and in-house citation.
- Practice or citing only to get the favours of the powerful or to appease others.
- Citations given just to dress up the paper.
- Variation of citation rate during the life time of the paper.
- Variation of citation rate with type of paper and speciality.

It is a fact that there are extraneous considerations in giving citations. But that does not totally undermine the value of them. As Sengupta argues, "the number of scientists who are capable of doing such malpractices are not significantly high to make citation studies unworkable or misleading". Further, it is too much to question the honesty and integrity of the majority of scientists.