

## Subject: Library and Information Science

Production of Courseware  
 -Content for Post Graduate Courses



**Paper No : 10** Informetrics and Scientometrics

**Module : 10** Citation Analysis



### 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

K C Garg  
Scientist and Head ISTAG, CSIR-NISTADS

Content Reviewer

Dr I K Ravichandra Rao  
Retd Professor, Documentation Research and  
Training Centre

## **Unit 10**

### **Citation Analysis**

#### **I. Objectives**

- What is citation analysis?
- Tools for citation analysis;
- Classification of citations according to their nature and context;
- Reasons for citations;
- Bibliographic coupling and co-citations; and
  
- Applications of citation analysis and their limitations.

#### **II. Learning Outcome**

At the end of this module you have gained knowledge in citation analysis, including the merits and demerits of citation data. History of citation studies was also covered. You have also learnt the methodology of citation analysis.

#### **III. Module Structure**

1. Introduction
2. Citation Index
3. Nature of citations
4. Classification of citations
5. Reasons for citations
6. Bibliographic coupling
7. Co-citation
8. Applications
9. Limitations
10. Summary
11. References

## **1. Introduction**

Citation analysis is the major thrust area of bibliometric research. It deals with the analysis of the bibliographic references which generally appear at the end of the scientific communication. When an author cites a paper (say X) in his paper (say Y), then X is called the cited paper and Y is the citing paper. Thus, citation analysis presents a connection between cited and citing documents. Analysis of cited and citing papers can provide valuable information about the existing and emerging knowledge on a subject. The more an article is cited, the more significant becomes the paper. Citation analysis can be used to study the influence of the research output of a country on world science. Science policy makers these days are using citation analysis to identify most significant papers or authors or institutions in a discipline. Citation counts help a research administrator to assess the impact of the research output of each individual scientist of his organization, but also that of his organization as a whole. Citation analysis helps to identify those earlier researchers whose concepts, methods, apparatus, etc. were used by the authors in the preparation of his article.

Citations are very field dependent and the number of citations per paper varies from one discipline to another. Citations are high in the field of biomedical sciences. A relatively small, isolated field will attract fewer citations than either a more general field or research within a narrow field that has a wider focus of interest. A paper published in an obscure journal, or publication with small readership usually has a low citation rate. About 80% of the references appended in a paper are to journal articles. Review journals have relatively higher citation rate, because most review papers are long, contain many references, and are cited quite heavily; however, they are not necessarily very different in citations per page when differences are made for length of the paper. Citation analysis also tells about what an author is recognized for (past work, methods, concepts or reviews). Authors of highly cited papers constitute the elite or leading scientists of the specialty. Citation analysis involves counting of the number of citations to a particular paper for a period of years after its publication. The number of years for which the citations are counted is known as the citation window which may vary from one field to another.

## **2. Citation Index**

The practice of citation analysis received a considerable stimulus with the evolution of Science Citation Index (SCI) developed by Eugene Garfield and brought out by Institute of Scientific Information (now Thomson Reuters), USA. Till recently, it was the only database available for citation analysis. However, for the past few years researchers have started using SCOPUS database of the Elsevier for the purpose of citation analysis. Both these database are multidisciplinary international databases. SCI has grown from 600 journals in 1964 to more than 10,000 scientific and technical

journals in most science disciplines now. Thomson Reuters also publishes Social Science Citation Index (SSCI) and Arts & Humanities Citation Index (A&HCI). SCOPUS originated in the year 2004 and claims to index about 15000 periodicals in science, technology and social sciences. International conferences are also indexed by these two databases. Both these indexes are now available online. Besides, the availability of these two commercial databases for citation analysis, Google Scholar is also being used to track the citations of individual authors, which is freely available on the Web.

A citation index is an ordered list of cited articles each of which is accompanied by a list of citing articles. The citing article is identified by a source citation, the cited article by a reference citation. The index is arranged by reference citations. Arrangement by author is favoured in the citation index. Citation indexes have several problems like the cost, various spellings of author names, authors with the same name, incorrect citing information, and other human errors. However, advantages outweigh disadvantages. SCI was also criticized for having national bias in its coverage of journals as the number of journals originated from developing countries and countries publishing in language other than English were less covered. However, in recent years this problem has been solved to a great extent by Thomson Reuters. SCOPUS took note of these problems before its launch and did not face such criticism.

### 3. Nature of citations

Based on the nature as who is the citer, citations have been classified into three categories (Folly et al 1981). These are:

- **Self citations:** Self citation may be said to occur when at least one of the authors of a cited document is the same person as one of the authors of the citing document. The author self citation rate of an individual may be calculated by dividing the number of self citations by the total number of all citations made by the individual. Self citation may artificially inflate citation rate;
- **Co-operational citations:** The cited author under study and one citing author were co-authors prior to the publication of cited paper and criterion (i) does not hold; and
- **Independent citations:** No detectable relations between cited and citing author is observed.

### 4. Classification of citations

On the basis of citation context analysis, citations have been classified into four classes (Murugesan and Moravcsik 1978). These are as follows:

- i. **Conceptual vs. Operational (theory vs. method):** If a concept or theory of the cited paper is used directly or indirectly in the citing paper in order to lay foundation to build on it or to contribute to the citing paper, then the citation is conceptual one. In contrast, the definition of an operational reference involves the situation when a concept or theory is referred to as a tool to substantiate the authors claim (e.g. the author may compare his results based on his theory with the result of another theory) or to indicate alternative

approaches. In addition, a reference is also called operational when it borrows mathematical or physical techniques, results, references, or conclusions from the cited paper.

- ii. Organic vs. Perfunctory (essential vs. non essential):** “Organic” references are those from which concepts or theories are taken to lay the foundations of the citing paper, or papers from which certain results (including numerical ones) are taken to develop the ideas in the citing paper, or papers which help to better understand certain concepts in the citing paper. In contrast “perfunctory” references are those which describe alternative approaches which are not utilized in the citing paper, references which are used to compare certain results or conclusions, references which are used to indicate the fact that a certain method employed is routine in the literature, and references which merely contribute to the chronological context of the citing paper. In short, perfunctory citations are not really necessary for the development of the citing paper.
- iii. Evolutionary vs. Juxtapositional (development of idea vs. contrasting idea):** A reference is called evolutionary if it provides a concept or theory to build on, or a mathematical technique to use, or results of an analysis which is used in the development of the citing paper. In other words, evolutionary references deal with material which contributes directly to the logical development of the subject of the citing paper. In contrast, a reference is called juxtapositional, if it refers to alternative approaches, gives mere references to works using the same general approach but which have branched off so that they do not contribute to the development of the citing paper, refers to other analysis used in the citing paper only to make comparisons, refers to other works which may help to clarify some ideas but do not contribute to the development of the citing paper, or refer to a paper only for references given in the later.
- iv. Confirmative vs. Negational (supports findings vs. opposes findings):** A reference is confirmative if the author of the citing paper considers the paper referred to as correct. Almost all evolutionary papers are also confirmative. In contrast, a negational reference describes the situation when the author of the citing paper is not certain about the correctness of the cited paper. There are two types of negational citations. In the first type the author of the citing paper claims that the cited paper is incorrect. In the other type of negational citation the author of the citing paper disputes the cited paper but cannot come to a definite conclusion, because the issue is still being tested experimentally or theoretically.

## 5. Reasons for citations

Besides scientific merit various other factors influence the choice of references an author cites. It may be familiarity with particular papers or prominence of an author in a field, its availability or language, controversiality of the topic, reprint dissemination, and coverage by secondary services etc. Papers more than 5 years old are cited less frequently, because these are superseded by books or reviews. Only a classic paper is cited when it is old. A well known established author is more frequently cited than a new researcher in the field. The number of citations is influenced by the number of people interested in a field and the number of people who have previously published in the field. Specific reasons for citations are as follows (Weinstock 1971)

- Paying homage to pioneers;
- Giving credit for related work;
- Identifying methodology, equipment etc.;
- Providing background reading;
- Correcting one's own work;
- Correcting the work of others;
- Criticizing previous work;
- Substantiating claims;
- Alerting researchers to forthcoming work;
- Providing leads to poorly disseminated, poorly indexed, or uncited work;
- Authenticating data and classes of fact – physical constants etc.;
- Identifying original publications in which the idea or concept was discussed;
- Identifying original publications or other work describing an eponymous concept or term;
- Disclaiming work or ideas of others; and
- Disputing priority claims of others.

## 6. Bibliographic coupling

Bibliographic citations in scientific papers have been used by a variety of researchers to establish relationship among documents. Two important measures based on this concept which have received considerable attention in the literature are bibliographic coupling suggested by (Kessler 1963) and co-citation suggested by (Small 1973). Bibliographic coupling deals with sharing of one or more references by two documents. Two citing documents are said to be bibliographically coupled if they cite the same publication(s). The strength of the coupling between citing documents depends upon the total number of citations that they have in common. If the number of common reference is one, then the coupling strength is one. If there are two common references, then the coupling strength is two. In other words, the more the number of common references, the more the coupling strength of the papers. For example, let us consider two documents I and II. Document I cites document A, B, C, D; and document II cites C, D, E, and F. Hence, documents I and II are bibliographically coupled; because they both side C and D, and the coupling strength is two. Bibliographic coupling is a fixed and permanent relation and does not change with time.

## 7. Co-citation

It is the frequency with which two items of earlier literature are cited together by the later literature. The number of times that papers are cited together by a new or later document gives the strength of co-citation. The more the frequency of occurrence together, the stronger is the relationship between the two. Co-citation patterns change as the interests and intellectual patterns of the field change with time. Co-citation can be used to map out in detail the relationship between key ideas. It can also be used to depict a literature specialty core or cluster. Co-citation provides a tool for monitoring the development for scientific fields, and for assessing the degree of interrelationship among specialties. Co-citation helps in locating network of frequently cited papers. Co-citation patterns are found to differ significantly from bibliographic coupling patterns. Like document co-citations, these can also be developed for authors and journals.

## 8. Applications

Citation analysis has number of applications. These have been described below.

- **Finding out relative use of different types of documents:** Citation analysis can be used to identify relative use of different kind of documents such as books, periodicals, technical reports, conference proceedings, and theses etc. For instance, in several of the studies reported in literature, it has been observed that citations to articles in journals constitute a major proportion of total cited items among all types of literature sources. Similarly, one can also analyze and identify the language of the cited papers.
- **To trace the historical development of a field:** Citation studies have been used for creating historical descriptions for scientific fields. Employing network of scholarly communications one can establish the history of the field.
- **To guide a reader:** Each citation is a message from the author of a document to his readers. By reading the cited document, the reader can obtain more information on a point or check the authenticity of a particular view, finding or method etc.
- **For compilation of reading lists or bibliographies:** Citations are an excellent tool for the compilation of reading lists or bibliographies, especially in an emerging field of knowledge or in subjects where areas are not clearly demarcated and the literature is scattered. In such fields, perhaps citations are

the only source of information which may be of help for compiling reading lists or bibliographies. The quality of such a list would be excellent as references provided in it have been evaluated by the researchers, who are expert of the discipline.

- **For preparing ranked list of journals:** Based on the assumption that citations are indicative of the literature use patterns, these provide methodology to rank scientific journals according to their relative degree of importance. Librarians are using it as an authentic tool for streamlining their serials acquisition policy by adding and discarding the most and less cited journals respectively. It also helps the librarian in other aspects of library management like binding policy and locating most used periodicals at a place convenient to the readers.
- **Finding useful life of documents:** Material acquired in the library becomes out-of-date with the passage of time. This is known as obsolescence of the published material i.e. slow reduction in the use of the published material. Analysis of citations by age of the cited documents can indicate the useful life of documents especially earlier volumes of periodicals. This is based on the assumption that a document would be cited if it is used. Thus, citation analysis can be used to decide the obsolescence rate of documents in different subjects.
- **Finding out relatedness and interdependence of subjects:** Citation analysis can act as an important tool in establishing important clues about the relatedness and interdependence of one field on another field by identifying how the journals of one field cite the journals of other fields.
- **To study the scatter of literature:** Citation analysis helps in identifying how the literature of one discipline is scattered in the journals of other disciplines. How often the literature related to technology or applied science cite the literature of basic sciences or vice-versa. The technique can also be used to identify the countries and impact factor of journals which are being cited in the literature.
- **Citation analysis as a tool in research evaluation:** Citation analysis is an important tool in evaluative bibliometrics. Citation counts may be used directly as a measure of utilization or influence of a single publication or all publications of an individual, a grant contract, department, university, funding agency, or country. This is based on comparison of citation counts of one research group to number of citations received by similar documents by the other research group. It is then assumed that the document with higher relative numbers of citation counts have more impact than those with lower citation counts. During the last two decades several studies have been published in the international journal *Scientometrics* on the citation analysis of scientific research institutions, research groups and individual scientists.



## **9. Limitations**

Citation analysis has been criticized on certain grounds. Some researchers point out that all the references appended by the author in the paper might have not been consulted by him. He has cited these references from another paper on a similar subject. This makes citation data unreliable to some extent. Another pitfall of citation counting arises due to negative citations. Authors of scientific papers often cite papers of others with a view to contradict claims, or findings of their predecessors because of incorrect results drawn by them in their publications. These citations have a negative value and thus should get negative weight instead of positive weight in citation analysis. Also methodological papers are among highly cited papers.

## **10. Summary**

Citation analysis is the hot area of bibliometric research and it received a considerable stimulus with the evolution of Science Citation Index (SCI). Citation analysis involves counting of the number of citations to a particular paper for a period of years after its publication. It can provide valuable information about the existing and emerging knowledge on a subject and the influence of the research output of a country on world science. Citation counts help a research administrator to assess the impact of the research output of each individual scientist of the organization as well as the organization as a whole. It helps to identify those earlier researchers whose concepts, methods, apparatus, etc. were used by the authors in the preparation of his article. Citations are very field dependent and the number of citations per paper varies from one discipline to another. Review journals have relatively higher citation rate. Authors of highly cited papers constitute the elite or leading scientists of the specialty.

Based on the nature as who is the citer, citations have been classified as self citations, co-operational citations and independent citations. On the basis of citation context analysis these are classified as conceptual vs. operational (theory vs. method) organic vs. perfunctory (essential vs. non essential), evolutionary vs. juxtapositional (development of idea vs. contrasting idea) and confirmative vs. negational (supports findings vs. opposes findings). Besides scientific merit various other factors influence the choice of references an author cites. Two important concepts that have been developed on the basis of citation analysis are bibliographic coupling and co-citation. Citation analysis have extensively been used to find out relative use of different types of documents, historical development of a field, for compilation of reading lists or bibliographies, preparing ranked list of journals, studying obsolescence of literature, relatedness and interdependence of subjects, scatter of literature and most widely in research evaluation.

## 11. References

1. Folly, G.; Haytman, B., Nagy, J.I. and Ruff, I. (1981), Some methodological problems in ranking scientists by citation analysis, *Scientometrics*, 2: 135-147
2. Kessler, M.M. (1963) Bibliographic coupling between scientific papers, *American Documentation*, 14: 10-25
3. Murugesan, P. and Moravcsik, M.J. (1978) Variation of the nature of citation measures with journals and scientific specialties, *J. of the American society for Information Science*, 29: 141-147
4. Small, H.G. (1973) Co-citation in the scientific literature: A new measure of the relationship between two documents, *Journal of the American Society for Information Science*, 24: 265-269
5. Weinstock, M. (1971), *Citation Indexes (in)* *Encyclopedia of Library and Information Science*, Marcel Dekker, New York, 5: 16-40.