INFORMETRICS 87/88
L. Egghe and R. Rousseau (Editors)

SYNCHRONOUS AND DIACHRONOUS CITATION DISTRIBUTIONS

Hideshiro NAKAMOTO

The Integrated Researches for Information Science, Nissei Building, Nishi-Shimbashi 2-5-10, Minato-Tokyo, Japan

Abstract

Using the data of the Science Citation Index, it is confirmed that the decrease in citation by age is exponential and independent of the number of source publications. It is further shown that synchronous and diachronous percentage distributions highly coincide. As such, data of synchronous distribution can replace data of diachronous distribution in literature studies.

1. INTRODUCTION

The concept of "obsolescence of literature" was theoretically discussed by the author [1] in 1964 (see also [2,3,4]). He estimated that synchronous distribution and diachronous change of citations should coincide in their characteristics under the following conditions.

1) The average of the synchronous distribution is equal to the average of the diachronous change in citations.
2) Both the synchronous and the diachronous distribution tend to zero at their ends.

Applying this result, we made prediction for literature obsolescence in het case of users and literature growth. We concluded that the life of the literature would become shorter with an increase of new literature and longer with an increase of users. However, we developed this idea on a theoretical basis but were not able to prove it in a practical environment, because of the unavailability of chronological data in those days. In a real situation, the second condition is usually satisfied. The data to verify the first condition, however, are very difficult to obtain. We often have the data for an synchronous distribution but cannot easily find the data for a diachronous distribution. Only data which are traced for long years are useful for the planning and design of systems.

The Institute for Scientific Information, U.S.A., has been providing citation data in the Science Citation Index for over a quarter of a century. This paper now describes how obsolescence is shown in the data of citation statistics, and that synchronous and diachronous distributions have similar characteristics.

2. CITATION DATA

The Science Citation Index reported figures for its change in numbers of source publications, the issues and total citations from 1961 to the present. The increase is not only due to a growth of the literature in science and technology, but also is due to an increase in the activity of I.S.I. These figures are plotted on a semi-logarithmic scale in Fig. 1, which shows a parallel change in the three curves. The graph indicates that the number of issues per publication and citations per issue did not change so much during the period under consideration. It is clearly understood that citation totals depend a lot upon the number of source

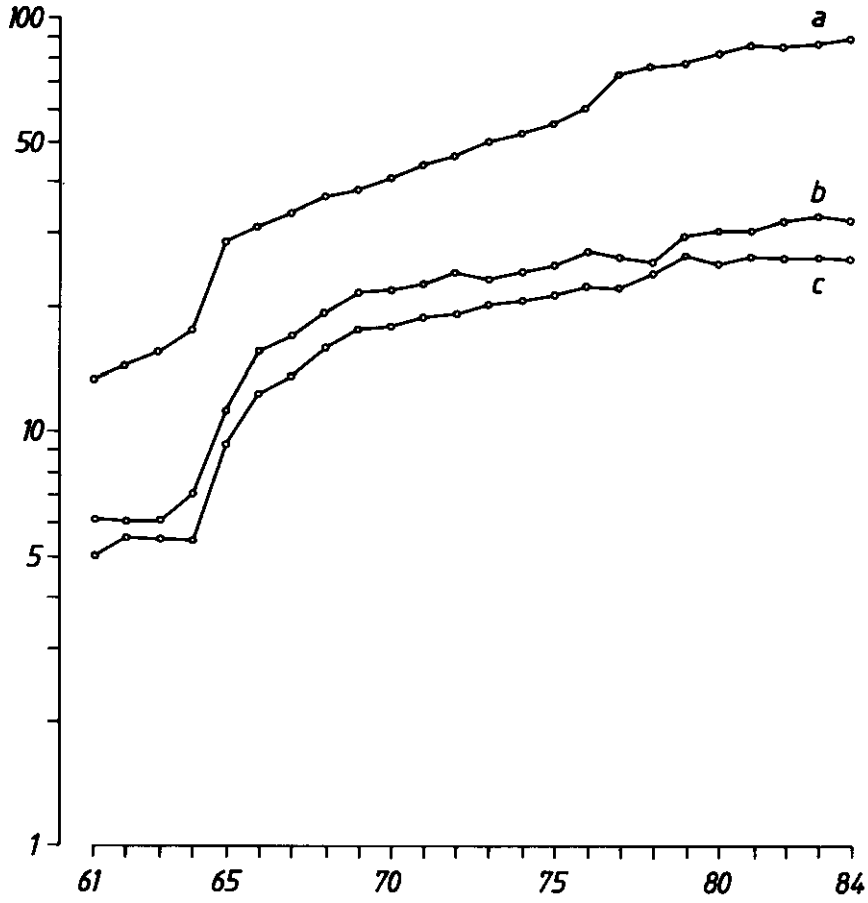

Fig. 1 : Change in numbers of the S.C.I.
a) Total number of citations (in hundreds of thousands)
b) Number of source publications (in hundreds)
c) Number of source issues (in thousands)

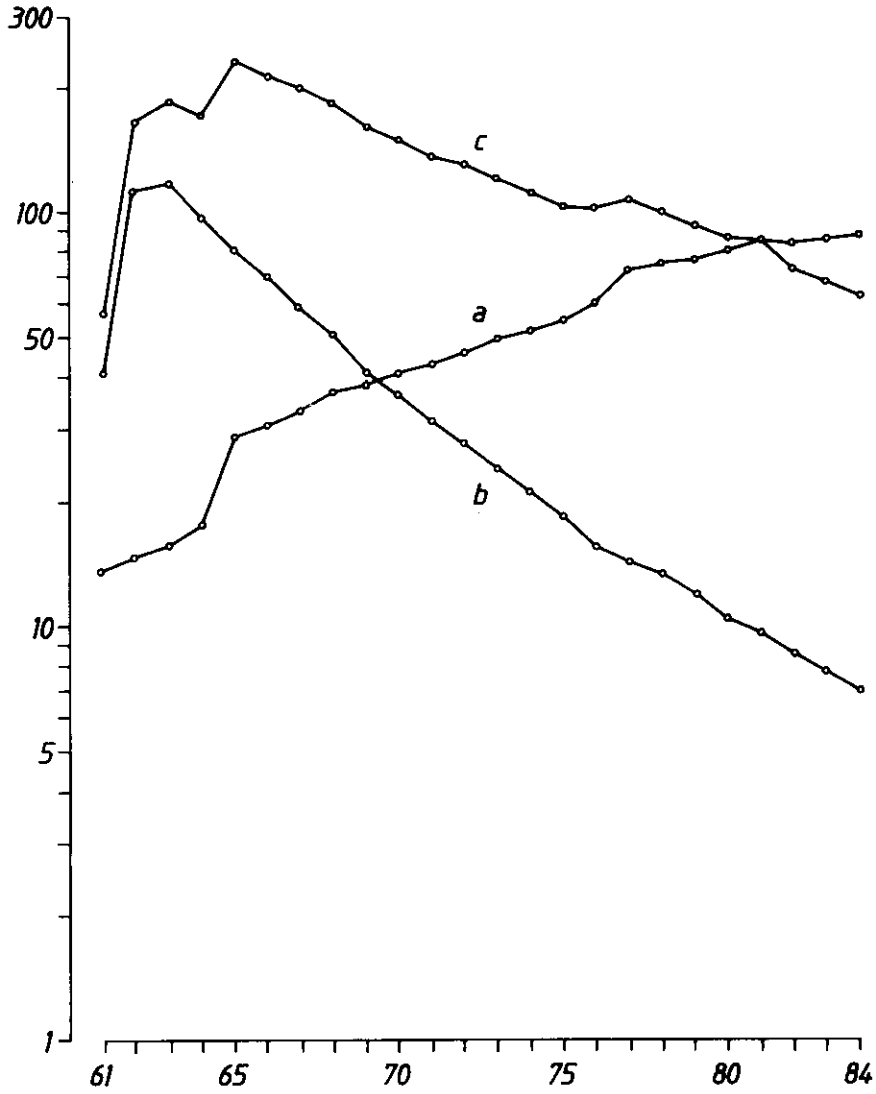

Fig. 2 : Citation Change of 1961 publications over 24 years
a) Total number of citations (in hundreds of thousands)
b) Diachronous distribution of 1961 publications (in per thousands of the annual total of citations)
c) Number of citations to 1961 publications (in thousands)

publications. A change over the years of citations to 1961 publications in shown in Fig. 2.
The ratio of total citations to total source items, calculated for each year is shown in Table 1. From 1961 to 1984, the ratios have remained between 12 and 16. The fact that the synchronous average has not changed much during 24 years, indicates that we have here a process with a stationary average.

Table 1 : Total number of citations, source publications and ratio of total citations to total source items between 1961 and 1984.

Year	Total Citations	Source Publications	Total citations/Total source items
61	1,370,000	631	12.1
62	1,485,635	605	12.0
63	1,558,800	610	12.1
64	1,789,753	700	11.8
65	2,924,940	1,146	12.4
66	3,074,006	1,573	11.2
67	3,387,139	1,711	11.1
68	3,698,715	1,968	12.0
69	3,849,715	2,180	11.3
70	4,107,947	2,192	11.4
71	4,379,705	2,277	12.0
72	4,659,115	2,425	12.3
73	5,017,420	2,364	12.3
74	5,231,710	2,443	13.1
75	5,535,968	2,540	13.2
76	6,176,553	2,717	13.7
77	7,398,026	2,655	14.9
78	7,597,336	2,572	15.2
79	7,769,960	2,993	15.0
80	8,260,543	3,067	15.9
81	8,667,736	3,068	16.1
82	8,490,049	3,246	15.5
83	8,737,642	3,327	15.4
84	8,911,676	3,281	15.7

3. EXPONENTIAL DECREASE OF CITATIONS

In order to go further with my argument, an analysis was made on characteristics of citations in the Science Citation Index 1984, which also shows distribution patterns over 20 year from 1961 on (See Table 2). The data for 1981 are shown in Fig. 3 on semi-logarithmic scales and can be seen to be approximately a straight line.
The lines for other years all have similar slopes and values. We point out that the rates of decrease of citation percentages are unchanged over these 20 years.

4. SYNCHRONOUS AND DIACHRONOUS DISTRIBUTIONS

Taking 0 for the present, we denote the number of citations given t' years ago to papers published t years ago by $V(t',t)$. If $t < t'$, then $V(t',t) = 0$.

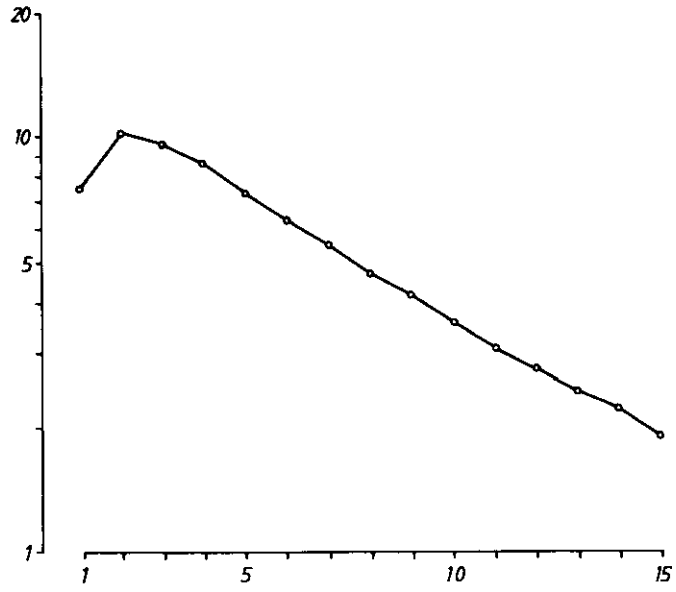

Fig. 3 : Synchronous percentage distribution in 1981; (age on the x-axis).

So, we obtain the following matrix :

0	-	-	0	0	0	V(0,0)
0	-	-	0	0	V(1,1)	V(0,1)
0	-	-	0	V(2,2)	V(1,2)	V(0,2)
0	-	-	V(3,3)	V(2,3)	V(1,3)	V(0,3)
.....						
V(t,t)	-	-	V(3,t)	V(2,t)	V(1,t)	V(0,t)
....						
	-	-	-	-	-	-

The rows of this matrix are time series, showing changes of citations of a fixed publishing year, and columns show the synchronous distribution of citations in a fixed citing year.
Summing citations per colums gives the series :

V(t,n), ..., V(3,n), V(2,n), V(1,n), V(0,n)

Dividing each figure in a column by the respective sum, and multiplying by 100, one gets a synchronous distribution of percentages, denoted $P(t',t)$.
The Science Citation Index of 1985 shows synchronous distributions between 1961 and 1984, corresponding to this matrix (see Table 2). We can trace up then entries for each year, say 1961. Both the synchronous distribution of citations in 1982 and the diachronous change of citations of 1961 are drawn on semi-logarithmic paper (see Fig. 4).

Table 2 : Science Citation Index 1961-1984
Chronological Distribution of Citations to authored items (non-patents)
(taken from the Science Citation Index 1985).

Reference

Percentage of Total Citations

Year	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
1984																								1.80
1983																							1.88	7.65
1982																						1.88	7.76	10.19
1981																					1.49	7.73	10.34	9.62
1980																				1.43	7.13	10.57	9.86	8.62
1979																			1.43	7.16	10.38	9.70	8.47	7.33
1978																		1.50	7.32	10.68	9.90	8.51	7.33	6.33
1977																	1.53	7.58	10.68	10.13	8.58	7.37	6.33	5.50
1976																1.73	7.59	10.86	9.98	8.71	7.35	6.31	5.43	4.72
1975															1.81	8.25	11.13	10.06	8.73	7.51	6.41	5.53	4.81	4.18
1974														1.88	8.82	11.02	10.15	8.68	7.44	6.40	5.48	4.70	4.08	3.58
1973													1.97	9.23	11.37	10.17	8.84	7.52	6.46	5.53	4.79	4.11	3.55	3.14
1972												2.14	9.14	11.51	10.21	8.84	7.60	6.46	5.57	4.81	4.19	3.62	3.16	2.79
1971											2.44	9.38	11.35	9.99	8.65	7.43	6.45	5.49	4.76	4.11	3.62	3.13	2.74	2.44
1970										2.47	9.48	11.40	10.03	8.52	7.33	6.32	5.49	4.76	4.13	3.61	3.22	2.81	2.50	2.23
1969									2.98	9.90	11.37	10.13	8.68	7.32	6.29	5.45	4.75	4.11	3.59	3.14	2.80	2.44	2.16	1.92
1968								2.67	10.38	11.58	10.20	8.71	7.45	6.35	5.46	4.75	4.16	3.62	3.19	2.78	2.49	2.18	1.92	1.73
1967							2.75	10.60	11.82	10.17	8.74	7.44	6.43	5.48	4.72	4.15	3.63	3.21	2.82	2.48	2.23	1.95	1.73	1.55
1966						2.65	10.90	12.05	10.05	8.49	7.20	6.19	5.35	4.57	3.95	3.48	3.05	2.69	2.39	2.11	1.90	1.67	1.48	1.34
1965					4.14	11.50	12.26	10.59	8.64	7.33	6.27	5.42	4.70	4.06	3.55	3.13	2.75	2.44	2.17	1.91	1.75	1.54	1.37	1.24
1964				4.55	13.60	12.45	10.58	8.82	7.22	6.14	5.30	4.58	3.97	3.44	3.00	2.67	2.35	2.08	1.87	1.65	1.51	1.33	1.19	1.08
1963			3.36	14.15	12.50	10.41	8.83	7.38	6.05	5.19	4.49	3.90	3.40	2.94	2.58	2.29	2.02	1.81	1.63	1.43	1.32	1.16	1.05	0.94
1962		3.28	11.26	12.39	10.12	8.55	7.17	6.06	4.99	4.34	3.79	3.30	2.88	2.50	2.18	1.94	1.72	1.54	1.39	1.23	1.14	1.01	0.91	0.82
1961	4.14	11.18	11.82	9.70	8.06	7.00	5.95	5.05	4.16	3.66	3.16	2.79	2.44	2.13	1.87	1.67	1.48	1.33	1.21	1.06	0.98	0.87	0.79	0.71
1960	10.90	12.28	10.11	8.00	6.88	6.05	5.22	4.30	3.73	3.29	2.86	2.54	2.20	1.93	1.71	1.52	1.35	1.21	1.10	0.97	0.90	0.81	0.73	0.66
1959	11.07	9.83	8.01	6.34	5.60	4.94	4.20	3.64	3.04	2.68	2.32	2.06	1.81	1.58	1.42	1.27	1.14	1.03	0.93	0.84	0.77	0.69	0.63	0.57
1958	9.14	8.01	6.62	5.29	4.69	4.20	3.61	3.13	2.65	2.36	2.07	1.84	1.61	1.42	1.27	1.13	1.01	0.91	0.82	0.74	0.69	0.61	0.55	0.51
1957	7.51	6.66	5.63	4.44	4.01	3.61	3.13	2.75	2.34	2.08	1.81	1.62	1.43	1.27	1.12	1.01	0.90	0.82	0.75	0.67	0.62	0.55	0.50	0.46
1956	6.43	5.65	4.77	3.81	3.46	3.16	2.73	2.41	2.03	1.83	1.62	1.45	1.28	1.14	1.01	0.92	0.82	0.75	0.69	0.61	0.58	0.52	0.47	0.43
1955	5.57	4.78	4.13	3.35	3.04	2.77	2.40	2.12	1.80	1.62	1.43	1.28	1.13	1.00	0.89	0.81	0.73	0.66	0.60	0.55	0.51	0.45	0.42	0.38
1954	4.74	4.01	3.46	2.78	2.52	2.31	2.03	1.79	1.53	1.38	1.22	1.09	0.96	0.85	0.77	0.70	0.63	0.57	0.52	0.46	0.43	0.39	0.36	0.33
1953	4.16	3.56	3.11	2.48	2.26	2.11	1.83	1.61	1.40	1.29	1.14	1.01	0.91	0.80	0.71	0.65	0.58	0.54	0.49	0.44	0.42	0.37	0.34	0.31
1952	3.52	3.02	2.67	2.09	1.95	1.85	1.56	1.39	1.19	1.09	0.97	0.86	0.78	0.68	0.61	0.55	0.51	0.46	0.42	0.38	0.36	0.32	0.30	0.27
1951	3.06	2.62	2.30	1.86	1.70	1.65	1.43	1.28	1.13	1.03	0.96	0.87	0.79	0.74	0.69	0.64	0.58	0.55	0.51	0.47	0.45	0.42	0.39	0.36
1950	2.66	2.26	1.99	1.67	1.47	1.39	1.22	1.09	0.95	0.86	0.77	0.68	0.61	0.55	0.50	0.45	0.41	0.37	0.34	0.31	0.29	0.26	0.24	0.22

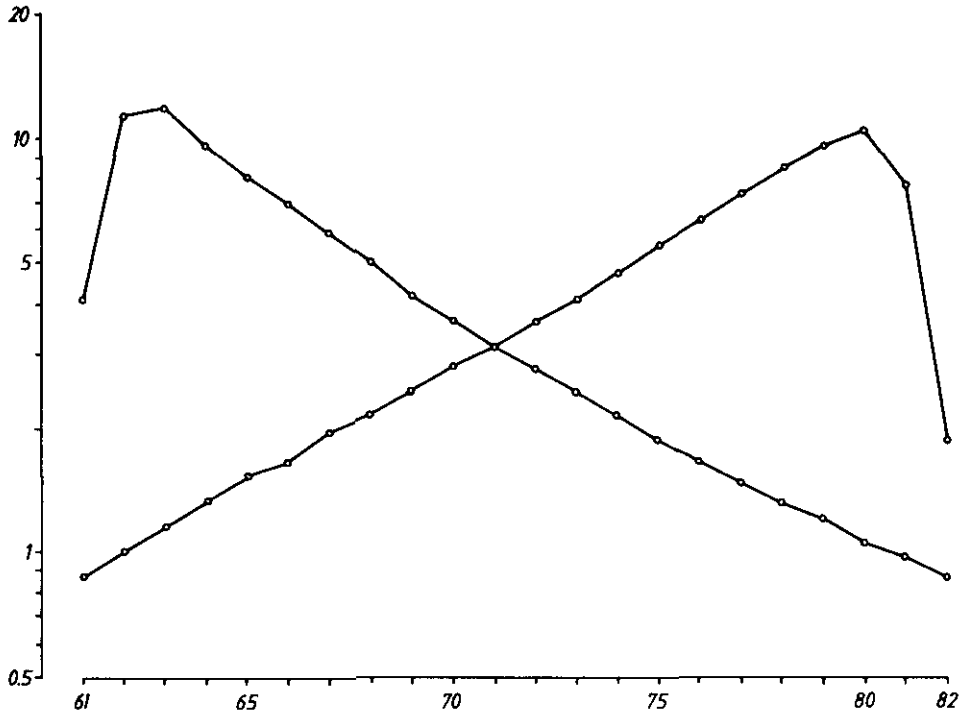

Fig. 4 : Diachronous distribution for 1961 publications and synchronous distribution for 1982 citations.

Note the symmetry of the curves. This indicates that $P(t',t) = P(0,t-t')$, or that citation decrease expressed by percentages, depends only on the age of the literature.
This study suggests that synchronous and diachronous distributions present similar curves. Therefore, we can use the data of synchronous distribution instead of diachronous for literature studies. However, this agreement is not perfect, therefore further investigations are necessary.

5. CONCLUSION

Using the data of the Science Citation Index, it is confirmed that the decrease in citation by age, is exponential and independent of the number of source publications. The coincidence of the synchronous and diachronous percentage distributions has been shown. We conclude that the data of synchronous distributions can replace the data of diachronous distributions in literature studies.

REFERENCES

[1] Nakamoto, H., Obsolescence of literature, Dokumentesyon-kenkyu, (1964) 14(6), pp. 123-128.
[2] Nakamoto, H., Obsolence model of literature, Dokumentesyon-kenkyu, (1965) 16(1) pp. 16-20.
[3] Nakamoto, H., Lifeness of literature, Jyohokanri (1966) 8(15), pp. 32-35.
[4] Burton, R.E. and Keller, R.W., The "half-life" of some scientific and technical literatures, American Documentation (1960) 11, pp. 18-22.