International Core Journal of Engineering Volume 8 Issue 6, 2022
ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

Research on the Model of the Number of Papers with Citation Times Greater Than or Equal to H in the Scientific Research Work of Researchers

Haili Li

St. Paul University Philippines, Tuguegarao City, Cagayan 3500, Philippines

Abstract

How to evaluate academic influence in a more objective and fair way has always been an important and concerned issue. The proposal of the H- index broke the original status quo of taking the impact factor as a single evaluation index, aroused extensive attention and discussion, and produced dozens of derivative indexes in the following years, which greatly promoted the field of scient metrics in-depth research and development. This paper introduces the H- index and its three representative derived indicators in detail. From the perspective of the citation distribution curve of scholars or units, the essence of these indicators is discussed, and the advantages and disadvantages of each are discussed. Finally, the research on the impact evaluation of academic output is prospected.

Keywords

Scientific Research Work; Number of Citations; H Index; Model of the Number of Papers.

1. Introduction

Scholars' evaluation is an important part of modern scientific research activities and an important part of scientific evaluation. On the one hand, there are huge differences in the research level of different scholars. In the case of limited scientific resources, only after evaluating scholars can the rational allocation of academic resources be ensured, so that the limited resources can be fully and reasonably utilized; On the one hand, scholars themselves also hope that their scientific research results can be judged fairly, and the fair recognition of scientific research results can ensure the academic enthusiasm of scholars. In addition, by evaluating scholars, we can also find leaders and active scholars in a certain field, and by observing their academic achievements, we can have a clear and accurate understanding of the research trends in the field. Since the American physicist J.E. Hirsch proposed the h-index for evaluating the individual scientific research level of scholars in 2005, the h-index has been widely used in various fields of scientific evaluation because of its impartiality and simplicity. However, the h-index also has some shortcomings, such as ignoring high-cited papers, only rising but not falling, etc. These shortcomings have a certain impact on the application of the h-index, in order to make the h-index better in scientific evaluation [1]. For the application, it is necessary to improve the h-index. However, most academic influence evaluation systems only use citation data as the basis for analysis and calculation. For example, impact factors, characteristic factors and H- index are all based on citation data. Therefore, first of all, it needs to be clarified here that all H- series evaluation indicators only evaluate the influence of a scholar's academic output, not the scholar's academic influence (because it also includes the scholar's ability to preside over scientific research projects). experience and his/her social influence in academia, etc.). The former is purer, the latter is broader and more comprehensive. This article will take the scholar's citation distribution curve as a unified perspective and tool to introduce and analyse the H- index and its two representative derived indicators, and point out the problems of some indicators.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

2. Definition and Calculation Process of H- bar Index

2.1 Definition of H- bar Index

Hirsch defines the h- index as: the number of citations of a scientist's h paper is greater than or equal to h times, then the scientist has an h- index, and papers with a number of citations greater than or equal to h are in the h- core set. Based on the concept of h- index, the h- bar index is defined as follows: The h paper of a scientist is cited more than or equal to h times and greater than or equal to the h- index of his collaborators, which is the h-bar index value. The h- bar value is usually less than or equal to the h- index value. The h- bar value will be equal to the h value only if a scientist has only independently signed and authored the results of the paper. At the same time, the fact that the h- bar value is equal to the h value does not mean that the scientist only publishes the paper independently [2]. Compared with the h- index, which is constant once it is determined, the h- bar index value changes as the scientists themselves and their collaborators increase the h-index.

2.2 Calculation Example

In Figure 1, all papers of a scholar are arranged in descending order of citations, and then their serial numbers are mapped to the X- axis, and the number of citations of each paper is mapped to the Y- axis to draw the scholar's citation distribution curve [3]. At this point, the area enclosed by the curve and the coordinate axis is equal to the total number of references. The side length of the largest square with (0,0) as an endpoint in this area is the scholar's H- index, which is also equal to the coordinate value corresponding to the intersection of the curve and the straight-line y=x. Of course, there are also some problems with the H- index. There are four main points: 1) The H- index ignores the role of highly cited papers, which are often representative works. For example, the H- index of scholar a and scholar b in Table 1 are both 4, but the influence of scholar a is significantly higher than that of scholar b; 2) The H- index does not consider the value of papers after the h, even if their citation counts very close to h. For example, the H- index of scholar c and scholar d in Table 1 are both 4, but the influence of scholar d is slightly larger; 3) The value of the H- index is limited by the total number of published papers, that is, the H- index will not exceed the number of published papers 4) The discriminative ability of H-index is weak. The H- indices of the four people in Table 1 are all the same, but their influence varies greatly.

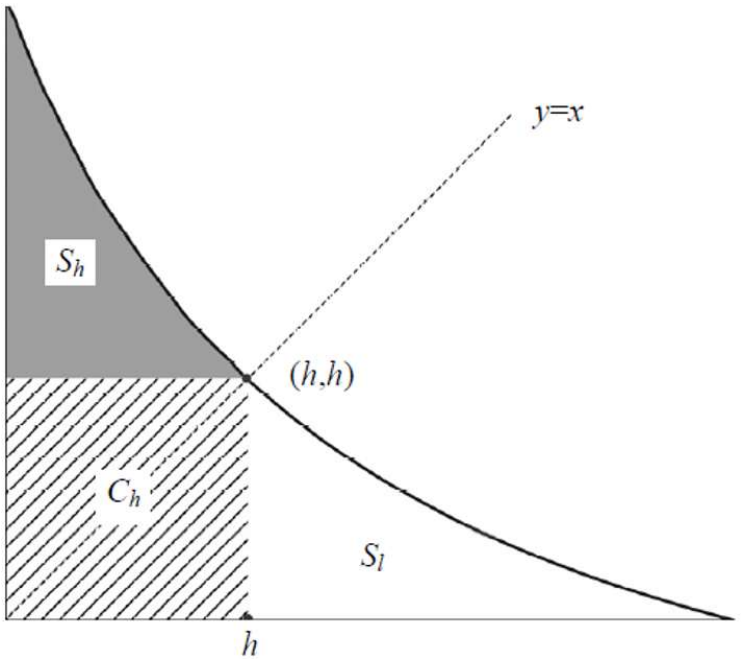

Figure 1. H- Index

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

Table 1. Examples of disadvantages of H- index

Paper sorting	Paper citations			
	Scholar a	Scholar b	Academic c	Scholar D
1	100	5	21	21
2	50	5	18	18
3	7	5	7	7
4	5	5	5	5
5	3	3	0	4
6	3	3	0	4
7	1	1	0	4
8	0	0	0	4

The papers in the h core are dynamically changing. Papers in the h-core move in and out of the h-core with changes in citation frequency, but the h- index does not decrease over time. Papers with only h citations at any one time will likely be eliminated by papers with higher citation rates and no longer contribute to the calculation of the personal h- index. Papers that were initially excluded from the calculation of the h- index have the potential to re-contribute to the calculation of the h- index, which usually happens with papers called "Sleeping Beauty." From observations of many physicists' citation records, the following conclusions can be drawn (Table 2):

(1) When m≈1, after 20 years of scientific research activities, the scientist's h- index is 20. At this time, the scientist can be considered a successful scientist. (2) When m≈2, after 20 years of scientific research activities, the scientist's h- index is 40. At this time, the scientist can be considered as a scientist with outstanding achievements. These scientists are likely just those working in top universities or key laboratories. (3) When m≈3 or greater, the h- index of the scientist after 20 or 30 years of scientific research activities is 60-90 respectively, and the scientist can be considered as a real scientific elite.

Table 2. Physicist h- index and m value

Physicist	h-index	m value	physicist	h-index	m value
Witten	111	3.89	Parisi	73	2.15
Cohen	94	2.24	Louie	70	2.33
Anderson	91	1.88	Jackiw	69	1.92
Fisher	88	1.91	Wilczek	68	2.19
Weinberg	88	1.76	Vafa	66	3.3
Cardona	86	1.87	Maple	66	1.94
deGennes	79	1.75	Gross	66	1.69
Bahcall	77	1.75	Hawking	62	1.59
Fisk	75	2.14	Dresselhaus	62	1.41
Scalapino	75	1.88			

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

2.3 Application Limitations

The driving force for the change of the h- index comes only from outside the performance core, and any performance change inside the performance core has no effect on the h- index. This forms the problem of partial missing power source for h-index changes, resulting in h- index insensitivity to performance changes. The performance core threshold increases with the increase of the performance core. Due to the rising characteristics of the h- index threshold, it is determined that papers other than the performance core need to accumulate more citations as the threshold increases, so the time required is also longer [4]. In other words, the larger the h- index, the longer it takes for the h- index to rise. In reality, not all papers end up contributing to the calculation of the h- index. Papers with fewer citations consistently had no effect on the researcher's h- index, especially those that were published later by the author after the h- value had been calculated.

3. G- index

In order to emphasize the value of highly cited papers, scholars put forward the g- index: the sum of the citations of g papers in N papers published by a scholar is greater than or equal to g^2, and the sum of the citations of any (g+1) papers and both are less than $(g+1)^2$. As shown in Figure 2, we arrange the papers in descending order of the number of citations and draw their citation distribution curve. In Figure 2, S_h and S_h' satisfy the relationship $S_h \geq S_h'$. The g- index expresses all citations of the first g papers by a square S_g with the most approximate area. There is $S_h \approx S_h'$ in general; $S_h = S_h'$ holds only when $\Sigma_{i=1}^{g} c_i = g^2$, where c_i is the number of citations of the i paper. The g-index is the side length of the square S_g (slashed line domain) to represent a scholar's academic output influence [5]. The g- index takes into account the value of highly cited papers, and also breaks the limit of the total number of papers. For example, the number of papers of scholar e in Table 3 is 3, and the H- index is also 3, but the g- index is 10. In addition, it is more beneficial for scholars with fewer papers but high citations. For example, in Table 2, although scholar f has 8 papers, the H- index is 5, which is greater than that of scholar e, but the g- index is 6, which is smaller than that of scholar e. Although the g- index has the above advantages over the H- index, it is susceptible to high-cited papers and is less stable than the H- index.

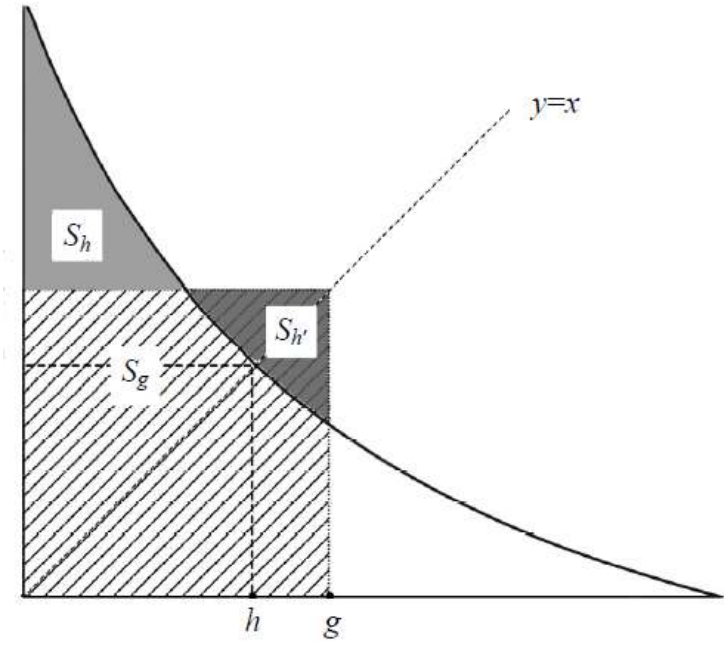

Figure 2. g- index

Table 3. Comparison of H- index and g- index

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

Scholar e				Academic f			
Sort r	r*r	Paper citations	Cumulative citations	Sort r	r*r	Paper citations	Cumulative citations
1	1	100	100	1	1	10	10
2	4	9	109	2	4	9	19
3	9	7	116	3	9	7	26
4	16	0	116	4	16	5	31
5	25	0	116	5	25	5	36
…	…	…	…	6	36	5	41
10	100	0	116	7	49	3	44
11	121	0	116	8	64	2	46

4. R- index

The main function of the R- index is to measure changes in the influence of papers within the performance kernel. Under the premise of not changing the shape of the h- core, the R- index can measure the performance cores with the same h-index and different intensities by calculating the square root of the total citation frequency of each paper in the performance core. It is the main basis for distinguishing the same value h-index [6]. The R- index is the square root of the total number of citations of papers in a scholar's H-core, ie $R=\sqrt{\Sigma_{i=1}^{h}c_i}$. As can be seen from Figure 3, the R- index is the side length of a square with side length r) to represent the academic influence of a scholar's scientific research output, $S_h = S_h'$ that is, the area of the square is equal to the total number of citations of papers in the H- core. The relationship $R=\sqrt{A\bullet h}$ is satisfied among the three A- index, R-index and H- index.

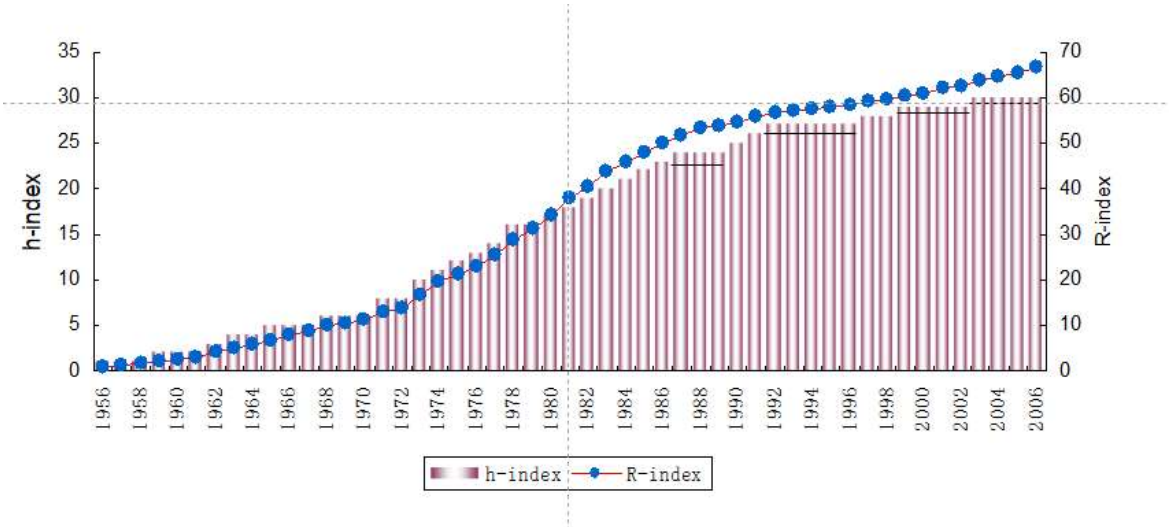

Figure 3. h- index and R-index of GARFIELD

5. AR Index

The AR index has a regulating effect on the trend of the h- index. The AR- index is defined as the square root of the sum of the average annual citations of each paper in a scholar's H- kernel, $AR=\sqrt{\Sigma_{i=1}^{h}\left(c_i/a_i\right)}$. where a_i is the age of the i paper—that is, the current year minus the year in which the paper was published. It can be seen that the H- index only rises and does not fall, while the AR- index rises and falls with the number of citations and age of each paper. If the annual number of

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

citations of the paper decreases year by year, the AR value will decrease, indicating that the academic influence of the scholar's research output will decrease. A-index=citation frequency/ (statistic year-published year-0.5). Calculate the square root of the A-index sum: the square root of 106.61034 is 10.3252. Then AR index = 10.3252 (shown in Table 4).

Table 4. Example analysis table

Paper	Year of publication	Citations	Year of publication	A-index
1	1972	652	1972	18.8986
2	1996	146	1996	13.9048
3	1999	100	1999	13.3333
4	1980	156	1980	5.8868
5	1983	135	1983	5.7447
6	1955	278	1955	5.3981
7	1981	134	1981	5.2549
8	1986	106	1986	5.1707
9	1985	111	1985	5.1628
10	1984	109	1984	4.8444
11	1979	132	1979	4.8000
12	1977	140	1977	4.7458
13	1986	94	1986	4.5854
14	1987	88	1987	4.5128
15	1982	107	1982	4.3673

6. Conclusion

Although the combination of h- index and R-index or AR- index has the value orientation of scientific research quality in the evaluation of individual performance of scientists, as an evaluation index, h-index and R- index or AR- index can only be used at best. as a means of peer review. In the face of complex scientific research activities, we cannot simply evaluate the individual scientific research performance of scientists based on one or two indicators, but should focus on establishing a scientific and reasonable evaluation mechanism and evaluation system based on peer review, various quantitative the setting of indicators is only to provide relevant information for peers to draw fair and reasonable evaluation opinions.

References

[1] Jin Bihui, Rousseau Ronald. R index, AR index: complementary indicators for the expansion of h index function. Science Observation, 2(3) (2022) 89-94.

[2] Ye Ying. Mechanism analysis and empirical research guidance of h-index and h- like index. Journal of University Libraries, 44(5) (2019) 49-54.

[3] Jiang Chunlin, Liu Zeyuan, Liang Yongxia. H-index and G-index-new indicators for the evaluation of academic influence of journals. Library and Information Service, 50(12) (2021) 9-18.

[4] Qiu Junping, Miao Wenting. Application of h-index in talent evaluation: Taking Chinese scholars in the field of library and information science as an example. Science Observation, 2(3) (2019) 67-72.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0005

[5] Zhao Jiming, Qiu Junping, Huang Kai, et al. A Review of a New Scientometric Indicator--h Index and Its Application. Chinese Science Foundation, 11(2) (2020) 147-152.

[6] Zhao Xing. An analysis of the correlation between the h-index and the number of papers. Information and Documentation Services, 7(1) (2020) 55-59.