

Letter to the Editor

Note on a Possible Decomposition of the h-Index

Dear Sir,

In a recent Letter to the Editor (Bartolucci, 2012), the author advocates the introduction of an index that measures the degree of concentration of citations over papers (e.g., of an author). The definition is as follows: Let C be the total number of citations and let h be the h-index of the system (Hirsch, 2005). Then the proposed index is

$$r = \frac{h}{m} \quad (1)$$

where $m = \lceil \sqrt{C} \rceil$, where $\lceil x \rceil$ denotes the largest integer that is smaller than or equal to x . The argument for this index is as follows. If we have C citations in total, we obtain the maximal possible h-index by taking $\lceil \sqrt{C} \rceil$ papers each with $\lceil \sqrt{C} \rceil$ citations. Then, obviously from the definition of the h-index, $h = \lceil \sqrt{C} \rceil$ and hence, by Equation 1, $r = 1$, its maximal value. When all citations are concentrated on one paper, $h = 1$ obviously and, by Equation 1, $r = 1/\lceil \sqrt{C} \rceil$, its minimal value. Therefore, r is a kind of opposite measure of concentration (in biology one calls this a measure of diversity), because it is minimal for the perfect concentration and maximal for the perfect spread of citations over papers.

Concentration is defined exactly in econometrics (Lorenz, 1905; see also Egghe, 2005). For this reason, we will, temporarily, use the econometric terminology, replacing papers by persons and received citations by money they own. Let us have two situations with the same number of people. Then we say that the first situation is more concentrated than the second one if the first situation can be retrieved from the second one via a finite number of elementary transfers. An *elementary transfer* is the action of taking

some money away from a “poor” person and giving it to a “rich” person. Example: (5,2) is obtained from (4,3) by taking one unit away from the second person and adding it to the amount of the first person. The vector (6,1) is obtained by executing two elementary transfers on (4,3): taking two units from the second person and giving them to the first. Indeed, the h-index measures this: $h = 1$ for the situation (6,1) and $h = 2$ for the situation (4,3); the h-index is larger for the less concentrated situation (i.e., for the case of more spread out). Here $r = 1/2$ in the first case and $r = 1$ in the second case. However, the h-index does not always satisfy this property. Let the first situation be (3,3,2,1) and apply an elementary transfer from the fourth article to the third one (I reuse the papers-citations terminology) to obtain (3,3,3,0). Now, $h = 2$ in the first case, but $h = 3$ in the more concentrated second case. Now, $r = 2/3$ in the first case, and $r = 1$ in the second case, although the first case is more spread out. These examples show that the h-index is not a good measure of diversity or concentration. This, in turn, has negative consequences for the validity of Equation 1 as a good measure of diversity (measuring the spread of citations over papers).

Leo Egghe

Universiteit Hasselt, Campus Diepenbeek, Agoralaan B-3590 Diepenbeek,

Belgium

E-mail: leo.egghe@uhasselt.be

References

- Bartolucci, F. (2012). Letter to the editor. On a possible decomposition of the h-index. *Journal of the American Society for Information Science and Technology*, 63(10), 2126–2127.
- Egghe, L. (2005). *Power laws in the information production process: Lotkaian informetrics*. Oxford, United Kingdom: Elsevier.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science of the United States of America* 102, 16569–16572.
- Lorenz, M.O. (1905). Methods of measuring concentration of wealth. *Journal of the American Statistical Association* 9, 209–219.