Journal of Informetrics 10 (2016) 43–47

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Expected number of citations and the crown indicator☆

Lawrence Smolinsky*

Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA

ARTICLE INFO

Article history:
Received 30 July 2015
Accepted 28 October 2015
Available online 17 December 2015

Keywords:
Bibliometric indicator
Crown indicator
Field normalization
Research performance assessment

ABSTRACT

The mean normalized citation score or crown indicator is a much studied bibliometric indicator that normalizes citation counts across fields. We examine the theoretical basis of the normalization method and, in particular, the determination of the expected number of citations. We observe a theoretical bias that raises the expected number of citations for low citation fields and lowers the expected number of citations for high citation fields when interdisciplinary publications are included.

1. Introduction

The mean normalized citation score (MNCS) is a bibliometric indicator of research performance developed at The Centre for Science and Technology Studies, Leiden University. Bibliometric indicators are important as they are used to compare research performance for individuals, research groups, institutions, and countries. A recent survey of bibliometric indicators is included in Waltman (2015). The MNCS bibliometric indicator normalizes citation counts for differences between fields while keeping year and document type fixed. For a set of n publications, it is defined as

$$\mathrm{MNCS} = \frac{1}{n}\sum_{i=1}^{n}\frac{c_i}{e_i} \tag{1}$$

where c_i is the number of citations to the ith publication and e_i is the expected number of citations to the ith publication.

The MNCS is an indicator that normalizes by dividing by an expected value. The determination of these normalizing variables has been heavily discussed for both the MNCS indicator and its predecessor, the mean field citation score/citations per publication (CPP/FCSm). Leydesdorff and Opthof (2011) criticize the MNCS while Lundberg (2007) and Opthof and Leydesdorff (2010) included a discussion in the context of the CPP/FCSm. There was a response by van Raan, van Leeuwen, Visser, van Eck, and Waltman, 2010 and a proposal in Waltman, van Eck, van Leeuwen, Visser, and van Raan (2011). Much of the discussion deals with the issues of validity of field classification and a proper reference set for the normalization. Although there was a suggestion to ignore fields and use reference counts, citations counts were generally accepted. Ultimately the discussion by Waltman et al. of the MNCS indicator used field classification and citation counts.

The citation counts c_i in Eq. (1) are well-defined as variables by choosing to use the values from the Web of Science or Scopus. This is not the case for the expected number of citations e_i. Waltman et al. (2011) write, "We also determine for each publication its expected number of citations. The expected number of citations of a publication equals the average number of citations of all publications of the same document type (i.e., article, letter, or review) published in the same field and in

☆ The peer review process of this paper was handled by Vincent Larivière, Associate Editor of *Journal of Informetrics*.
* Tel.: +1 225 578 1570.

http://dx.doi.org/10.1016/j.joi.2015.10.007
1751-1577/

the same year." Translating this intent into mathematics is not straightforward. This note addresses the expected number of citations from a theoretical perspective and, in particular, critiques the methodology of Waltman et al., when there are overlapping fields. Their method introduces a bias that hurts researchers in low citation fields and benefits researchers in high citation fields. We also discuss that altering the normalizing variables does not violate the uniqueness statement of Waltman et al.'s Theorem 1.

2. The expected value and properties of the MNCS

Waltman et al. show that MNCS is a bibliometric indicator of average performance of a set of publications that satisfies two properties: consistency and homogeneous normalization. Both the properties of consistency and homogeneous normalization involve comparisons between sets of publications of the same size, say n, at a time. The property of consistency is also known as Independence (Bouyssou & Marchant, 2011).

We use the notation from Waltman et al., where a publication is represented as an ordered pair of numbers (c,e), c is the number of citations to the publication and e the expected number of citations. A collection of n publications is then represented as a set of n ordered pairs and a bibliometric indicator is a nonnegative function on the sets of ordered pairs.

A bibliometric indicator of average performance is said to have the property of consistency if $f(S_1) \geq f(S_2) \Leftrightarrow f\left(S_1 \cup \left\{(c,e)\right\}\right) \geq f\left(S_2 \cup \left\{(c,e)\right\}\right)$ for all sets of publications S_1 and S_2 of the same size and all publications (c,e) in the complement of $S_1 \cup S_2$. In other words, the relative ranking of two sets of publications does not change by the addition of the same publication to both sets.

It is important that the sets are of equal size. For example, consider Eq. (1) for two sets of publications all with the same e_i value e. A set S_1 of one publication with 1 citation will be evaluated with a lower value than a set S_2 of 10 publications where one publication has two citations and the nine others have one citation. However, the addition of a publication c with 20 citations to both sets will switch the evaluation having a more dramatic effect on the smaller group, i.e., $f(S_1) < f(S_2)$ but $f(S_1 \cup \{(c,e)\}) > f(S_2 \cup \{(c,e)\})$.

The second property of homogeneous normalization precisely defines the indicator in the case that every publication in the set has the same e coordinate. If $S = \{(c_1,e), (c_2,e), \ldots, (c_n,e)\}$, then homogeneous normalization requires

$$f(S) = \frac{1}{e} \sum_{i=1}^{n} \frac{c_i}{n}$$

The set of publications is considered homogeneous since they all have the same expected number of citations. The indicator is required to average the citations and then weights them by dividing by e.

Waltman et al. furthermore state (Theorem 1) that MNCS is the unique bibliometric indicator of average performance to satisfy these conditions. However, a little caution is required. The proof of Theorem 1 uses the notion 'bibliometric indicator' simply as collection of functions (parameterized by n) of two variables without meaning assigned to the two variables. There is no requirement of the meaning of each variable or a formula for computing a variable. In other words, the statement of Theorem might more properly be stated as: Eq. (1) is the unique function f: $(\mathbf{N}_0 \times \mathbf{R}_+)^n \to \mathbf{R}$ satisfying the properties of homogeneous normalization and consistency of the average performance.

The c_i's may be taken as the number of citations to an article on the Web of Science or Scopus. However, the e_i's may have a number of variations each of which define a function that satisfies Theorem 1. In fact, if the e_i's are arbitrary positive numbers with $e_i = e_j$ whenever publications i and j are in the same field, then the collection of e_i's will allow for an MNCS indicator that satisfies the properties of homogeneous normalization and consistency. For example, if e_i is set equal to 1 for all i, then the indicator is just the average number of citations for articles in the collection and satisfies homogeneous normalization and consistency. Similarly, one may replace the expected number of citations with the median number of citations for each field and obtain another indicator satisfying homogeneous normalization and consistency.

In addition to the properties of homogeneous normalization and consistency, Waltman et al. discuss one further property: "A nice property that we would like the MNCS indicator to have is that the indicator has a value of one when calculated for the set of all publications published in all fields." This condition will be included in later computations, but was not a requirement of Theorem 1. We will call it the unity property.

3. Computing the e_i's in the example from Waltman et al. (2011)

Waltman et al. point out that a reasonable definition of e_i is straightforward if each article in the determining set has a uniquely defined field, i.e., just the average. We look at their example in their Section 6 (*How to handle overlapping fields*) when articles do not have a single classification and determination of e_i is not straightforward. We reproduce their Table 9 (Overview for each publication of the field in which it has been published and the number of citations it has received) as our Table 1.

The expected number of citations is determined by the number of citations in the field as determined by the chart. In any method we consider, $e_4 = 6$. Publication 4 is the only Publication in the table in field Z. Also note $e_1 = e_2$ since both these publications are only in field X.

Table 1
Waltman et al.'s example.

Publication	Field	Citations
1	X	2
2	X	3
3	Y	8
4	Z	6
5	X and Y	5

We look at three methods for computing e_i's for potential MNCS indicators. In each method the MNCS indicator does have a value of one for the entire collection. Before proceeding to the methods, we give a word about the harmonic mean as it figures into the computations, which we illustrate in the following primary school problem. Painters X and Y paint at different rates. Painter X can paint a house in s days and painter Y can paint a house in r days. How long does it take them to paint a house, if they work together? The answer is $1/(1/r+1/s)$ or half the harmonic mean. Let us go slightly deeper into the analysis and to one analogous to citation counts. Painter X can paint $A=1/s$ houses in a day and painter Y can paint $B=1/r$ houses in a day. Consider the following math problems.

1. If X and Y painted for a total of one man-day and they each painted for an equal amount of time, then how many houses did they paint?Answer: $(A+B)/2$ since each worked half a day, i.e., the mean of A and B.
2. If X and Y painted for a total of one man-day and they each painted an equal number of houses, then how many houses did they paint?Answer: The painters did not paint an equal amount of time, but their time totaled one day. X painted $B/(A+B)$ days and Y painted $A/(A+B)$ days. They each painted $AB/(A+B)$ houses. Together they painted a total of

$$\frac{AB}{A+B}+\frac{AB}{A+B}=\frac{2}{\left(1/A\right)+\left(1/B\right)}, \tag{2}$$

the harmonic mean of A and B. It would be a conceptual flaw to impose that X and Y spend both equal time painting and paint an equal number of houses. That outcome would require them to paint at the same rate even though $r \neq s$. Compare splitting a paper and its citations between fields to the painter problem. Splitting the credit of the paper is analogous to splitting the one man-day the two painters worked. Splitting the citation count is analogous to splitting the painting accomplished. The analogy is as follows.

	Painting	Authoring
Agents	Painters	Fields
Effort	Time painting	Credit on article
Accomplishment	# houses painted	# of citations

Method M1, Waltman et al. Waltman et al. use the following method in computing the e_i's in their formulas (10) and (11). They are

$$e_1=e_2=\frac{\left(2+3+5/2\right)}{\left(1+1+1/2\right)} \tag{3}$$

and

$$e_3=\frac{\left(8+5/2\right)}{\left(1+1/2\right)}. \tag{4}$$

Waltman et al. credit both half of the citations of publication 5 and half of publication 5 itself to field X. Then e_5 is then computed as the harmonic mean of e_1 and e_3. The solutions are in Table 2.

The notion of assigning a distribution of credit to fields for a publication or journal with several fields of classification may seem elusive. However, it was implicitly done in the denominators of Eqs. (3) and (4) as the "1/2," thereby assigning

Table 2
Positive solutions.

Method	e_1	e_3	e_5
M1	3	7	4.2
M2	2.48	7.87	5.17
M3	2.72	8.44	4.12

half of the publication to each of X and Y. Similarly, half of the citations from Publication 5 are assigned to each of X and Y as the "5/2" in the numerator.

Using this same example, we give two alternative computations for the expected number of citations: M2 and M3.

Method M2, Equal credit. If you assume that Publication 5 is half X and half Y, then the citations to Publication 5 are not evenly attributed to X and Y. Rather $e_1/(e_1+e_3)$ of the citations to Publication 5 are credited to X and $e_3/(e_1+e_3)$ to Y. The expected number of citations to Publication 5, then is the mean $(e_1+e_3)/2$. (Note the analogous formulas apply if there are more than two fields.)

This yields the following equations

$$e_1\left(\frac{2+3+5\left(e_1/(e_1+e_3)\right)}{\left(1+1+1/2\right)}\right)$$

$$e_3=\left(\frac{8+5\left(e_3/(e_1+e_3)\right)}{\left(1+1/2\right)}\right)$$

$$e_5=\frac{(e_1+e_3)}{2}$$

Computing e_5 is analogous to the painter problem 1 above. There is only one positive solution, which is given in Table 2. To solve the system one must solve the first two equations, which are quadratic in two variables.

Method M3, Equal citation. If, on the other hand, the citations to Publication 5 are evenly attributable to areas X and Y, then Publication 5 is not equally attributable to areas X and Y. For the sake of discussion, suppose Publication 5 is α part X. Then

$$e_1=\frac{\left(2+3+5/2\right)}{(1+1+\alpha)}$$

$$e_3=\frac{\left(8+5/2\right)}{(1+(1-\alpha))}$$

$$e_5=\alpha e_1+(1-\alpha)e_3$$

where $\alpha=e_3/(e_1+e_3)$ and $(1-\alpha)=e_1/(e_1+e_3)$

and e_5 is the harmonic mean of e_1 and e_3. It is analogous to the painter problem 2 above. There is only one positive solution, which is given in Table 2. To solve the system one must solve the first two equations, which are quadratic in two variables.

4. Discussion

Each example method yields an MNCS indicator that satisfies the conditions of homogeneous normalization, consistency, and has a value of one when calculated for the set of all publications published.

Method M1 has the conceptual weakness illustrated in the discussion of the harmonic mean: It credits half of the citations and also credits half of the classification of Publication 5 to field X. Since $e_1 < e_3$, Field 1 will receive a disproportionate share of Publication 5's citation (relative to Field 3). It raises the expected value e_1 by artificially crediting it with the same assumption it applies to the larger citation field (here field Y). It burdens the researchers in the lower citation field (here field X) with a higher expected value and conversely unburdens the higher citation field with a smaller expected value. If a publication is classified in several fields then each field is weighted equally and the citations are divided equally. Then the expected number of citations of lower cited fields is artificially increased.

Consider for example the lowest cited field in the WOS classification. Whenever a contribution to its expected number of citations is derived from a multiply classified article, its expected citation count will be raised relative to the other fields in that article's classification. If a field is classified in the middle of the WOS, then the times it is artificially high or low might average. If a field is at the high end of the WOS, then its expected count will be reduced in comparison to other fields. Hence M1 has a theoretical bias towards leveling expected citation counts across fields. In M1, the bias is introduced in Eqs. (3) and (4) in the computation of e_1 and e_3. The harmonic mean is then required by the unity property.

On the other hand, compared to Methods M2 and M3, M1 has an important advantage of being computationally simple. Only arithmetic is required.

Method M2 assigns Publication 5 equally to fields X and Y by assumption, without knowledge if there is a primary field. An assumption of equality is made in the absence of particular knowledge. This assumption is also made in Method M1. As a result Method M2 credits more of the Publication 5's citations to the higher cited field. M2 finds the e_i's in a manner consistent with the assumption of assigning Publication 5 equally to fields X and Y.

While M2 is conceptually appealing, it is computationally complex. Suppose an evaluation involves m individual fields, i.e., in the total collection of articles under consideration there are m fields some of which may occur in the classification of individual articles, interdisciplinary articles, or both. Then the evaluation by M2 would result in a system of m polynomial equations in m variables that theoretically may be up to degree 2^{m-1}. These m variables are the expected number of citations to the involved individual fields. Interdisciplinary expectations are computed as means.

For Method M3 the assumption is made that the citations are divided evenly between each field even if one field has a significantly lower citation rate. This assumption is the second assumption made in Method M1. Publication 5 will have a primarily classification in the lower citation field. This method always biases the classification. Like Method M2, a computation including m fields would result in a system of m polynomial equations in m variables that may be up to degree 2^{m-1}. Again, the m variables are the expected number of citations to the involved individual fields. Interdisciplinary expectations are computed as harmonic means.

The purpose of these computations is to elucidate the theoretical underlying assumptions being made in the calculation of the expected number of citations in the MNCS discussed in Waltman et al. (2011). Every bibliometric indicator is flawed and the computational simplicity of M1 is important. Nevertheless, for the purposes of research performance assessment, one should be aware that M1 disadvantages researchers working in lower citation fields while benefiting researchers working in high citation fields.

Acknowledgments

The author is grateful to Andrew McDaniel for helpful discussions and Aaron Lercher and Andrew McDaniel for reading a preliminary version of the article. The author is further grateful to Ludo Waltman and the reviewers whose suggestions improved the paper and sharpened its the focus.

References

Bouyssou, D., & Marchant, T. (2011). Bibliometric rankings of journals based on Impact Factors: An axiomatic approach. *Journal of Informetrics*, *5*, 75–86.

Lundberg, J. (2007). Lifting the crown—Citation *z*-score. *Journal of Informetrics*, *1*(2), 145–154.

Leydesdorff, L., & Opthof, T. (2011). Remaining problems with the "New Crown Indicator" (MNCS) of the CWTS. *Journal of Informetrics*, *5*, 224–225 [Letter to the editor].

Opthof, T., & Leydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance. *Journal of Informetrics*, *4*(3), 423–430.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*, 37–47.

Waltman, L. (2015). *A review of the literature on citation impact indicators*. (arXiv:1507.02099 [cs.DL]. Retrieved August 12, 2015, from the arXiv database).

van Raan, A. F. J., van Leeuwen, T. N., Visser, M. S., van Eck, N. J., & Waltman, L. (2010). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics*, *4*, 431–435.