Abstract

We discuss the “rate of averages” versus the “average of rates” in the case of the impact factor. Synchronous as well as diachronous journal impact factors are sensitive to adding non-cited articles (to the denominator). This is a consequence of basic properties of elementary arithmetic. Our findings provide a rationale for not taking uncitable publications into account in impact factor calculations, at least if these items are truly uncitable, that is, are never cited.

Keywords: synchronous and diachronous impact factors; consistency, non-cited items

Introduction

During recent years our field has gone through a paradigm shift (Opthof & Leydesdorff, 2010; van Raan et al., 2010a and b). Should one calculate ratios of averages (RoA) or
averages of ratios (AoR)? Discussions centred mainly on the so-called Leiden Crown Indicator (although proposed first by colleagues from the Hungarian Academy of Science (Schubert & Braun, 1986)), which used to be calculated as CpP / FCSm (Moed et al., 1995). For a research group this means: the average number of citations received (during a given publication and citation window) divided by the weighted (according to the number of publications in each field) average number of citations received in the corresponding field.

Opponents have come to a consensus that calculating an average of ratios is the proper approach (Gingras & Larivière, 2011). The most convincing reason for this conclusion is the requirement that such indicators should be consistent (Waltman et al., 2011). Simply stated this means that when a performance indicator I is determined for actors A and B resulting in I(A) < I(B) then it should not be possible that the operation of adding the same article with the same number of citations to both sets, leads to a reversal of the order in the mutual ranking: I(A) > I(B). It could be shown that this contradictory behaviour is possible for the traditional crown indicator (a ratio of averages), while the so-called Karolinska-Amsterdam indicator (an average of ratios) (Lundberg, 2007) does not have this problem.

In this note we elaborate on the differences between rates of averages and the average of rates for the case of the impact factor as a quotient. We show that, due to the arithmetic operations involved in the determination of synchronous and diachronous journal impact factors, such indicators have a counterintuitive behaviour. Concretely, it is possible that the impact factor of journal J is larger than the impact factor of journal J' and that adding the same number of non-cited articles to both sets reverses the mutual order.

Journal impact factors

We recall the definitions of the synchronous and the diachronous journal impact factor. The n-year synchronous impact factor of journal J in year Y is defined as (Rousseau, 1988):

\[ IF_n(J, Y) = \frac{\sum_{i=1}^{n} CIT_{i}(Y, X - i)}{\sum_{i=1}^{n} PUB_{i}(Y - i)} = \frac{1}{n} \sum_{i=1}^{n} CIT_{i}(Y, X - i) \]

(1)

In this formula the number of citations received by journal J (from all members of the pool of sources under consideration) in the year Y, by articles published in the year X, is denoted as CIT,J(Y, X), where, for simplicity we have not included the index J. Similarly, PUB(Z) denotes the number of articles published by this same journal in the year Z. We made it clear in equation (1) that the standard synchronous journal impact factor is a ratio of averages. Hence we will denote it as RAIF. When n = 2 one obtains the classical Garfield (1972) journal impact factor. Since a few years also the 5-year journal impact factor is provided in Thomson Reuters’ Web of Science. The term ‘synchronous’
refers to the fact that the citation data used to calculate it are data collected in the same year. We next recall the definition of the diachronous impact factor.

The n-year diachronous impact factor of journal J for the year Y, denoted as IMP\(_n(J,Y)\), is defined as

\[
IMP\(_n(J,Y) = \frac{\sum_{i=0}^{s} Cit(Y+i,Y)}{Pub(Y)}
\]

where \( s = 0 \) or \( 1 \), depending on whether one includes the year of publication or not. The term 'diachronous' refers to the fact that the data that are used to calculate this impact factor derive from a number of different years with a starting point somewhere in the past and encompassing subsequent years (Ingwersen et al., 2001).

**Z-consistency**

We consider the following form of consistency. If a performance indicator \( I \) is calculated for journals J and J' and \( I(J) < I(J') \) then, if we add the same number of publications with zero citations, we require that also for the new situation \( I(J) < I(J') \). As this notion of consistency differs from the one introduced by Waltman and van Eck (Waltman & van Eck, 2009; Waltman et al., 2011) or by Marchant (2008), under the name of independence, we refer to it as Z-consistency (Z for zero citations). Recall that, for good reasons, the notion of consistency as defined by Waltman and van Eck (Waltman & van Eck, 2009; Waltman et al., 2011) refers to cases where the number of publications (in the denominator) is the same for both journals. We do not add this requirement, but focus on increasing the denominator resulting in Z-inconsistency. Consider the following example (see Table 1a).

Table 1a: Data for the calculation of the Garfield impact factor (RoA case) for the year Y

|          | J   | J'  |
|----------|-----|-----|
| Pub(Y-1) | 10  | 30  |
| Pub(Y-2) | 10  | 30  |
| Cit(Y,Y-1)| 30  | 60  |
| Cit(Y,Y-2)| 30  | 60  |

On the basis of Table 1a, the Garfield impact factors of journals J and J' are \( IF_2(J,Y) = 3 \) and \( IF_2(J',Y) = 2 \), so that \( IF_2(J,Y) < IF_2(J',Y) \). However, adding 25 non-cited publications, yields the new impact factors: \( IF_2(J,Y) = 60/45 = 1.33 \) and \( IF_2(J',Y) = 120/85 = 1.41 \), so that for the new situation the relation between the impact factors reverses.

A simple variation of Table 1a shows that also the diachronous impact does not satisfy this property either, see Table 1b.
Table 1b: Data for the calculation of the Garfield impact factor (RoA case) for the year Y

|          | J       | J'      |
|----------|---------|---------|
| Pub(Y)   | 20 (+25)| 60 (+25)|
| Cit(Y,Y) | 10      | 20      |
| Cit(Y,Y+1)| 20     | 40      |
| Cit(Y,Y+2)| 30     | 60      |

With \( s = 0 \), we have \( \text{IMP}_3(J,Y) = 60/20 = 3 \) and \( \text{IMP}_3(J',Y) = 120/60 = 2 \). Adding 25 non-cited publications yields the new diachronous impact factors: \( \text{IMP}_3(J,Y) = 60/45 = 1.33 \) and \( \text{IMP}_3(J',Y) = 120/85 = 1.41 \).

**AoR versus RoA**

We have shown that the standard synchronous impact factor is of the RoA-form and that it does not satisfy Z-consistency. Let us analyze whether perhaps an AoR-form of the synchronous impact factor behaves better in this respect. First we define the ARIF as:

\[
\text{ARIF}_n(J,Y) = \frac{1}{n} \sum_{i=1}^{n} \frac{\text{Cit}(Y,Y-i)}{\text{Pub}(Y-i)}
\]

(3)

However, it turns out that the AoR-form behaves even worse with respect to Z-consistency. Indeed, consider the case of a two-year impact factor (ARIF_2) and assume that journals J and J' have each year the same number of publications. If then \( \text{RAIF}_2(J,Y) = \text{IF}_2(J,Y) < \text{RAIF}_2(J',Y) = \text{IF}_2(J',Y) \), this means that journal J received more citations than journal J' (in the year Y). Adding the same number of zero-cited publications to both, does not change the total number of citations received, and hence J's standard impact factor stays smaller than J' (of course the two impact factors decrease). The same argument holds for the n-year synchronous impact factor (RA-case). This, however, does not hold for ARIF. Consider the example shown in Table 2.

Table 2. Data for the calculation of a two-year synchronous impact factor (AoR case) for the year Y

|          | J       | J'      |
|----------|---------|---------|
| Pub(Y-1) | 30 (+10)| 30 (+10)|
| Pub(Y-2) | 20      | 20      |
| Cit(Y,Y-1)| 10     | 120     |
| Cit(Y,Y-2)| 80     | 10      |

Based on the data shown in Table 2, we have: \( \text{ARIF}_2(J,Y) = (0.5) \cdot (10/30+80/20) = 2.17 \) and \( \text{ARIF}_2(J',Y) = (0.5) \cdot (120/30+10/20) = 2.25 \) so that \( \text{ARIF}_2(J,Y) < \text{ARIF}_2(J',Y) \). However, adding 10 publications in the year Y-1 yields the new impact factors: \( \text{ARIF}_2(J,Y) = \)
(0.5)(10/40 + 80/20) = 2.13 and ARIF₂(J',Y) = (0.5)(120/40 + 10/20) = 1.75, so that for the new situation ARIF₂(J,Y) > ARIF₂(J',Y).

It is easy to find similar examples for any n-synchronous impact factor calculated in the AoR way.

**Conclusion**

Clearly, as a consequence of simple arithmetic Z-inconsistency is a normal mathematical property related to taking ratios. For the calculation of synchronous impact factors, the standard RoA approach is to be preferred above the AoR approach. Our findings provide a rationale for not taking uncitable publications into account in impact factor calculations, at least if these items are truly uncitable, that is, are really never cited.

**References**

Garfield, E. (1972). Citation Analysis as a Tool in Journal Evaluation. *Science 178* (Number 4060), 471-479.

Gingras, Y., & Larivière, V. (2011). There are neither "king" nor "crown" in scientometrics: Comments on a supposed "alternative" method of normalization. *Journal of Informetrics, 5*(1), 226-227.

Ingwersen, P., Larsen, B., Rousseau, R. & Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin, 46*(6), 524-528.

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

Marchant, T. (2009). An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors. *Scientometrics, 80*(2), 325-342.

Moed, H. F., De Bruin, R. E., & Van Leeuwen, T. N. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics, 33*(3), 381-422.

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.

Rousseau R. (1988). Citation distribution of pure mathematics journals. In: *Informetrics 87/88* (Egghe L. & Rousseau R., eds). Amsterdam: Elsevier, pp. 249-262.

Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics, 9*(5), 281-291.

van Raan, A. F. J., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & Waltman, L. (2010a). The new set of bibliometric indicators of CWTS. Paper presented at the 11th International Conference on Science and Technology Indicators, Leiden, September 9-11, 2010.

van Raan, A. F. J., van Leeuwen, T. N., Visser, M. S., van Eck, N. J., & Waltman, L. (2010b). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics, 4*(3), 431-435.
Waltman, L. & van Eck, J.N. (2009). A taxonomy of bibliometric performance indicators based on the property of consistency. In: ISSI 2009 (B. Larsen & J. Leta, eds.). Sao Paulo: BIREME & Federal University of Rio de Janeiro, pp. 1002-1003.
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(1), 37-47.