Journal of Informetrics
Volume 16, Issue 4, November 2022, 101340

Relations among the h-, g-, ψ-, and p-index and offset-ability

Author links open overlay panel
Shelia X.Wei[a] TongTong[a] RonaldRousseau[bc] WanruWang[d] Fred Y.Ye[a]

[a] School of Information Management, Nanjing University, Nanjing, China

[b] Centre for R&D Monitoring (ECOOM) and Department MSI, KU Leuven, Belgium

[c] Faculty of Social Sciences, University of Antwerp, Belgium

[d] School of Information Management and Artificial Intelligence, Zhejiang University of Finance and Economics, Hangzhou, China

Received 20 April 2022, Revised 21 September 2022, Accepted 28 September 2022, Available online 5 October 2022, Version of Record 5 October 2022.

https://doi.org/10.1016/j.joi.2022.101340Get rights and content

Abstract

We show that the h-index, g-index, ψ-index, and p-index, are related through the inequalities: $h \leq p \leq g \leq \psi$. Moreover, this relation is proved theoretically in the mathematical framework of Lotkaian informetrics and is verified empirically by using two datasets from the Web of Science in the fields of electrochemistry and gerontology. For quantifying their relations, we estimate the g-index, ψ-index, and their cores and ratios of cores via a second-order Taylor series when the e-index, h-index, and C_1 (the maximum number of citations received by

a paper) are known. Then we find for the two empirical cases, that ratios of cores and average citations are approximately stable. Compared with the g-index, the offset-ability of the h-index decreases by 20% but the average citations increase by 20%. A similar observation holds for the comparison of the g-index and ψ-index. To explore the possible applications of cores of different indices, we apply them to extract the core structure of a network. The h-core is the most efficient, while the ψ-core includes more nodes with high betweenness.

Introduction

Since the h-index was introduced in 2005 (Hirsch, 2005), it has widely been studied in informetrics as an academic measure (Egghe, 2010). Typical mathematical theories of the h-index include Hirsch's original approach (Hirsch, 2005), the Egghe-Rousseau model (Egghe & Rousseau, 2006), and the Glänzel-Schubert model (Glanzel, 2006; Schubert & Glänzel, 2007). These three models have led to a concise meaningful unification of publications and citations (Ye, 2009, 2011). As a robust indicator for measuring both impact (related to quality) and output (related to quantity) of publications, the h-index has been applied to academic assessment (Iglesias & Pecharroman, 2007; Lillquist & Green, 2010; Vanclay, 2007) and network analysis (Schubert et al., 2009; Zhang et al., 2018; Zhao et al., 2011, 2014). It can moreover also be applied to forecast the future academic performance of scholars (Hirsch, 2007). However, the h-index ignores most citations of highly-cited publications.

Other h-type indices were introduced to improve the properties of the h-index (Alonso et al., 2009, 2010; Bihari & Tripathi, 2017; Fassin, 2020; Singh, 2022; Todeschini, 2011; Wu, 2010; Zhang, 2009). The first, really different, h-type index is the g-index proposed by Egghe (2006). It inherits the advantages of the h-index and considers all citations of top-cited papers. One may say that it has offsetting power or offset-ability for the less cited papers. Here the

term offset-ability refers to how an indicator allows papers with a high number of citations to compensate for those with a lower number of citations. However, the g-index does not reflect the maximal ability of higher citations to offset the less cited papers. Thus, the ψ-index was created to present this ability to the greatest extent (Lathabai, 2020). The offset-ability can be investigated using the core of the h-, g- and ψ-index (Lathabai, 2020). The concept of the h-core was defined by Rousseau (2006), while Ye & Rousseau (2010) introduced the core-tail ratio. After the e-index was proposed by Zhang (2009, 2013a), it can be seen that the h-core consists of the e-area and h-area, corresponding to the t-area (Zhang, 2013b) in the core-tail plane. Along with the relations between the h-index, the g-index, and the e-index (Zhang, 2010), the core-tail-zero ratios were investigated (Liu et al., 2013; Ye & Rousseau, 2010). Moreover, the ratio of the h-core vs. the h-tail has also been the center of attention (Chen et al., 2013). In addition, some studies focus on the average citation rate of papers in the h-core (Jacso, 2011), statistical features of papers in the h-core (Sanz-Casado et al., 2016), and the shape centroids of the h-core and h-tail areas to quantify research and innovation performance (Kuan et al., 2011). According to the definitions of the h-index, g-index, and ψ-index, we know that these indices are proposed because the h-index largely ignores highly cited papers in the h-core. Based on the h-core, g-core, and ψ-core, it is significant to investigate how higher citations can offset the less cited papers, and this for the h-index, g-index, and ψ-index.

In addition, the p-index (Prathap, 2010a, 2010b) is not only sensitive to the number of citations in the core, but also considers the long tail. Its relation to the total number of papers and total citations was put forward based on the Glänzel-Schubert model (Schubert & Glänzel, 2007). There is a strong correlation between the h-index and $(CP)a/(1+a)P1/(1+a)$ where C is the total number of received citations, P is the total number of publications and a is related to the Lotka exponent in the Lotkaian framework (Radicchi &

Castellano, 2013; Schubert & Glänzel, 2007). The latter formula is the p-index when the constant a is equal to 2. Therefore, the p-index is highly related to the h-index (Bertoli-Barsotti & Lando, 2017; Ding et al., 2020; Malesios, 2015; Prathap, 2010b, 2017). When the basic unit of energy is defined as the energy that a single paper gathering a single citation possesses, the p-index can be interpreted based on the energy-exergy-entropy paradigm, which is a creative attempt of using the notion of entropy in scientometrics (Lin et al., 2020; Prathap, 2011).

With the emergence of various indices, it is worth comparing different features of the above indices. Egghe, 2008a, 2008b) investigated how the h-index, g-index, the R-index (Jin et al., 2007), and the h_w-index (Egghe & Rousseau, 2008) are affected by the transformation of information production processes. Van Eck & Waltman (2008) introduced two new indices, h_α-index, and g_α-index, generalizing the h-index and g-index, respectively, and analyzed the relationship between these new indices and other ones. Zhang (2010) calculated an estimation of the g-index based on Taylor's formula and gives the relation with the e-index (Zhang, 2009), h-index, and g-index. Chen et al., (2021) investigated the sensitivity of some selected h-like indexes to citations, namely the h-index, the g-index, the A-index (Jin et al., 2007), the R-index, and the e-index, via optimal growth path models. Schreiber et al., (2012) examined and compared the performance and properties of the h-index, 17 h-type variants, and some traditional bibliometric indicators by factor analysis. Prathap (2010b) analyzed the correlation among several indices and found out that the main contenders are the h-, g- and p-index.

Hence, our study tries to compare the h-, g-, ψ-, and p-index and investigates if there exists an order relation between them. The reasons why we choose these four indices are as follows. Firstly, all of them are classical. The h-index, g-index, and ψ-index have similar definitions and geometric meanings (Lathabai, 2020) and the p-index also has a thermodynamic explanation which suggests a physical

interpretation (Prathap, 2011). Secondly, as effective bibliometric indicators, they are the main contenders (Prathap, 2010b) because they all integrate both quality and quantity assessment into a single number (Prathap, 2012). As shown above, some statistical analyses show that there is a strong correlation between the h-index and the p-index. Moreover, the p-index is also highly correlated with the g-index (Ding et al., 2020; Prathap, 2010b). Thirdly, the h-, g-, ψ-, and p-index all have the same dimensions of publications (Prathap, 2014, 2018, 2021b), which means that they can rightfully be compared. Finally, the relations among the h-, g-, and ψ-index were proved according to the Egghe-Rousseau model (Egghe, 2006; Lathabai, 2020), and the relation between the h- and p-index was given based on the Glänzel-Schubert model (Schubert & Glänzel, 2007), so a unified relation among the h-, g-, ψ-, and p-index is expected. If the relations among the h-, g-, and ψ-index are clarified, some other indices related to the Glänzel-Schubert model can be introduced into these relations easily. Therefore, in this study, we focus on the h-index, g-index, ψ-index, and p-index and try to find out the relation of these indices and explore the features and usefulness of their cores.

Section snippets

Methodology

First, we recall the definitions of the h-index, g-index, ψ-index, and p-index.

The h-index (Hirsch, 2005): A set of papers has an h-index h if h of these papers have at least h citations each, and the other papers have at most h citations each.

The g-index (Egghe, 2006): A set of papers has a g-index g if g is the highest rank such that the top g papers have at least g^2 citations in total.

The ψ-index (Lathabai, 2020): A set of papers has a ψ-index ψ if ψ is the unique largest number such that

Empirical studies and results

In this section, four parts are presented, namely the size relation of four indices, the regression result of the constant c in Eq. (3), the estimated values of different cores, and their average numbers of citations.

Discussion

In this section, we attempt to introduce a different dataset and some indices to expand the relations theoretically and empirically.

Conclusion

This study contains two main contributions. First, we selected four significant indices, namely the h-index, g-index, ψ-index, and p-index, and determined their unified relation in the mathematical framework of Lotkaian informetrics (the Lotkaian exponent $\alpha \in]2,3[$): $h \le p \le g \le \psi$. The h-, g-, and ψ-index have geometric meanings, the p-index has a thermodynamic explanation, and we verified their relations theoretically and empirically, which means that we found a unified mechanism of those mathematical

Author contributions

S.X.W collected and processed data and wrote the paper, T.T. assisted data processing, R.R. did the analysis, checked the research, and wrote the paper, W.R.W. wrote the paper, and F.Y.Y initiated the idea, designed the research, and wrote the paper.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgments

We acknowledge the financial support from the National Natural Science Foundation of China Grant Nos. 71673131 and 72204213. We thank the anonymous reviewers for their constructive suggestions.

References (71)

- S.X. Zhao *et al.*
  h-Degree as a basic measure in weighted networks

  Journal of Informetrics
  (2011)
- L. Zhai *et al.*
  The bi-directional h-index and B-core decomposition in directed networks

  Physica A-Statistical Mechanics and Its Applications
  (2019)
- R.W. Wang *et al.*
  Extracting a core structure from heterogeneous information network using h-subnet and meta-path strength

  Journal of Informetrics
  (2021)
- N.J. van Eck *et al.*
  Generalizing the h- and g-indices

  Journal of Informetrics
  (2008)
- A. Schubert *et al.*
  A systematic analysis of Hirsch-type indices for journals

  Journal of Informetrics
  (2007)
- M. Schreiber *et al.*
  Exploratory factor analysis for the Hirsch index, 17h-type variants, and some traditional bibliometric indicators

  Journal of Informetrics
  (2012)
- J.Q. Liu *et al.*

Ratios of h-cores, h-tails and uncited sources in sets of scientific papers and technical patents

Journal of Informetrics
(2013)

- S.S. Li *et al.*
  Characterizing core-periphery structure of complex network by h-core and fingerprint curve

  Physica A-Statistical Mechanics and Its Applications
  (2018)
- H.H. Lathabai
  ψ-index: A new overall productivity index for actors of science and technology

  Journal of Informetrics
  (2020)
- C.-.H. Kuan *et al.*
  Positioning research and innovation performance using shape centroids of h-core and h-tail

  Journal of Informetrics
  (2011)

View more references

Cited by (0)

Recommended articles (0)

View full text