A formal approach to assessing relevance of scientific publications building on citation count

Ricardo Alvira Baeza

Introduction. *The need to evaluate the relevance or value of scientific research for the scientific community and society using objective criteria has led to a generalization of the use of indexes - directly or indirectly- based on accounting for the number of times that published research results are cited by other researchers. However, the consistency of these indexes is challenged by an increasing number of experts. This paper reviews the existing indices, highlighting several contradictions, and explains two alternative indexes which provide a more consistent assessment of relevance based on the received citation.*

Method*. A formal or axiomatic method is followed, stating a system of axioms based on contributions in the framework of economic theory. In addition, an empirical test is carried out by reviewing two samples of journals and comparing the results of the most commonly used indexes against the proposed indexes.*

Analysis & Results*. While no currently used index fulfils the proposed axioms, two indexes are explained that satisfy all the axioms. Besides, the empirical review shows that proposed indexes have high correlations with the most relevant variables and most currently used indexes.*

Conclusions*. The axiomatic approach allows us to highlight some important shortcomings of the indices currently used for citation evaluation, as well as the greater consistency of the two herein proposed indices.*

Keywords

Citation Assessment; Research Evaluation; Bibliometrics; Impact

1 Introduction

The need to efficiently allocate scarce economic resources, often from totally or partially public institutions, to the different agents involved in scientific research and the subsequent dissemination of the results obtained, is at the basis of the creation, since the beginning of the 20th century, of various quantitative indices that seek to assign relevance both to the research itself and to the authors and institutions that make its production possible (Hirsch, 2005; Bollen, et al., 2006).

These indexes, which were initially linked to the decision of which scientific journals or books to acquire for university libraries, have progressively acquired a more central position as a criterion for allocating funding or even assigning jobs and awards (Priem, et al., 2010; D’Antonio 2018; Heckman & Moktan 2020; Park, et al., 2023). In economic terms, this implies that the indexes have been progressively transformed into *collective utility functions*.

However, this use of these indices clashes with the fact that the most widely used indices can lead to significantly different valuations of authors and institutions, and fail to meet some basic conditions of collective utility functions. To help address this issue, this article explains an index (which can be interpreted as two interdependent indices), from a perspective that addresses some of the shortcomings of current indices when used for collective decision-making.

To this end, firstly a review is made of current indexes for assigning relevance to scientific authors based on the citation of their work, as well as the different criticisms raised by authors and experts. Subsequently, the parallelism of the characterisation sought with an economic term is explained: the equivalent (mean) income (Dalton, 1920; Atkinson, 1970).

A system of axioms that such a formulation must comply with is enunciated, approaching the concept of *collective utility function*. While no currently in-use index satisfies this formal system an index is explained that fulfils every axiom, providing thus a more consistent assessment. The empirical review shows a high resemblance with currently used indexes, but also some relevant differences.

2 Brief Historical Review of Citation Indexes

The systematic review of the quantitative assessment of the relevance of scientific publications building on their citation begins in the first quarter of the 20th century. To clarify, we differentiate three approaches:

- Those built on linear approaches (total count or average values)
- Those built on non-linear approaches based on:
  - … analysis of centrality from graphs/networks theory
  - … analysis of the distribution of the citation

Below we review each of these three approaches.

2.1 Linear Approaches

An early contribution towards the *quantitative assessment* of the *relevance* of different contributions to science was made in 1917 by Cole & Nellie, who sought to detect the *relative importance* of each thematic area and country in the overall development of anatomical science. They reviewed the scientific production (books, monographs and articles) published between 1543 and 1860, and concluded that the number of contributions was independent of the importance of such contributions. Therefore, *accounting for an author's publications does not inform us of his/her significance*.

In 1927, Gross & Gross addressed the issue with the aim of proposing a method that allowed the libraries of Chemistry faculties to select which journals to acquire. The authors sought to define an *objective criterion* to 'measure the desirability of acquiring each particular journal', eliminating the *subjectivity* of entrusting such selection to an expert. To do so, they suggested *assessing each journal's actual use by scientists by counting the number of times the articles published in each journal were referenced in other articles.*

The authors stated that by counting all the citations an article received it was taking into consideration the assessment of the article by the other scientific authors; i.e., *the objectivity of the method came from the fact it was the complete scientific community who valued each article.* This method was soon accepted as an objective method of deciding scientific journal subscriptions by a growing number of US universities.

Almost thirty years later, Garfield (1955) made a contribution with great influence on the future evolution of the assessment of the significance of scientific works. His goal was to design a commercial tool, Science Citation Index (SCI), that he conceived as a tool to help both in the study of science and in the evaluation of the relevance of different scientific journals.

Garfield designed the index taking into account the reduced technological capacity of the moment (information was stored in punched cardboards) and making reasonable use of human resources. This led him to greatly restrict the sample of journals and articles whose citation was assessed as well as the sample of articles whose references were accounted for.

Some years later, Garfield & Sher (1963) observed that the number of times a journal was referenced in the literature was usually linked to the number of published articles. Consequently, counting the total citations of each journal penalized those publishing fewer articles. Therefore, they proposed that *a more meaningful measure of importance* was obtained by dividing total citations by the number of published articles. Garfield & Sher (1963) designated this parameter as *Journal Impact Factor (JIF)*, and they stated that the ordering of journals according to their JIF allowed small journals to compete on equal opportunities with large journals.

Subsequently (Garfield, 1970) showed the probability of receiving the Nobel Prize was higher for most cited scientists, so he stated that *the ranking according to the number of received citations was an optimal tool for granting scholarships and awards and personnel selection processes, since it allowed to both assess/predict past/future scientists' performance.*

In 1972, Garfield launched the Journal Citation Reports (JCR) which included a ranking of journals from highest to lowest JIF according to SCI data. Garfield stated there was a direct link between citations and scientific relevance: *low-cited journals were irrelevant for original scientific ideas communication.* Likewise, Garfield formalized the JIF calculation methodology: the number of times that articles published during a period of two years were cited in articles published during the following year was accounted for.

Garfield's proposal was soon widely accepted/used to quantify the current and future scientific relevance of journals and individuals. A new paradigm emerged from which if a high value of JIF implied a highly relevant journal, then an article published in a high JIF journal was a highly relevant article. The JIF and whether the journals where an author had published were in the JCR listings became increasingly important. *Scientists' relevance was increasingly assessed by counting his/her number of published articles assigning each one a value according to the journal where it had been published.*

This paradigm that links the relevance of research to the JIF of the journal where it was published was adopted by the first universities rankings: in 2003, the Academic Ranking of World Universities (ARWU) and in 2004 the Times Higher Education (THE). This in turn led to more and more countries adapting their research assessment policies prioritizing publication in journals indexed in JCR (Web of Science, WoS) or Scopus over any other criteria.

However, some authors have raised several methodological criticisms of the above paradigm:

- Several authors (Redner, 1998;...) pointed out that the two-year interval for counting citations is unjustifiably short, and penalises certain areas of knowledge in which the citation of articles takes longer to occur. Price (1986, p. 112), showed that for many fields of knowledge, the two-year interval left more than 70% of citations unaccounted for.
- Several authors (e.g., Seglen, 1997;...) stressed that a high JIF value does not imply that all articles published in that journal receive a similar number of citations and that JIF does not characterise the expected citation of an article in a journal.

Another methodological criticism made to JIF is that in some fields of knowledge, articles are higher cited than in others, so the JIFs of different journals often are not comparable. To allow comparison between journals from different fields of knowledge, Moed (2009) proposed the Source Normalized Impact for Paper (SNIP):

- Firstly, the field of knowledge of each journal was determined from all published articles that cited an article published in the journal.
- Secondly, the median of references included in the publications of said field of knowledge was calculated and considered as said field of knowledge *citation potential*.
- Thirdly, the SNIP of each journal was calculated by dividing the average number of citations received per published article by the citation potential of the journals in the field of knowledge of the journal.

Finally, among the linear approaches, we must include various alternative metrics or *Altmetrics* that count different types of interactions on the internet (articles visualization, downloads, mentions on blogs ...). Advocates of these measures highlight that they provide updated information which complements that obtained when accounting for references in indexed journals, standing as filters that help decide which part of the abundant scientific production interests us (Priem, et al., 2010).

2.2 Non-linear approaches

We have reviewed proposals which follow a linear approach (based on total citations count or average values), but other authors have adopted different approaches characterized by their nonlinearity. We review them below differentiating two groups: those built on graph theory and those built on the distribution curve.

2.2.1 Citation from Graph/Network Theory

An early contribution was made by Pinski & Narin (1976, p. 299) who proposed that the citations analysis goal was to "describe the interactions between a group of publishing entities (journals, institutions, individual authors, fields of knowledge, geographical subdivisions…)". From this perspective, the authors stated three motives that challenged the validity of JIF to inform about the *influence as* the *importance* of scientific publications:

- Longer articles (e.g., review articles) usually receive more citations.
- JIF assigns the same value to every citation regardless of the journal where it is made.
- JIF does not take into account different citation patterns in different areas of knowledge.

In order to solve these deficiencies, the authors proposed three alternative measures to characterize scientific journals (Pinski & Narin, 1976):

1. The *influence weight* of a journal is the ratio between the number of citations it receives from other journals and the number of references it makes to other journals.
2. The *influence per publication* of the journal is the weighted sum of citations received by the journal (each citation is multiplied by the *influence weight* of the citing journal), divided by the number of journal publications during a year.
3. The *total influence* of the journal is the *influence per publication* multiplied by the total number of publications.

In 1998 Brin and Page (1998) proposed the *Google Page Rank* (GPR) algorithm, intending to rank the internet pages/documents according to their *relevance* for each search made by a user. The algorithm incorporated three steps:

1. First, the set of relevant pages was reduced by selecting both the descriptors of each page and the text in the link of each page, i.e., the algorithm assessed both the reasons why each page claimed to be relevant and the reasons why other pages actually considered the page to be relevant.
2. Relevance was assigned to each page according to two factors:
    a. The first factor assigned each web page a position in the ranking according to the expected number of times each page would be visited by a *random walker*. A value was assigned to each link that each page made to other pages, by dividing the web page position in the ranking by the number of links it included. Assuming an equal initial position in the ranking for all pages, convergent values were achieved through successive iterations (50 or 100 iterations, were considered sufficient).
    b. The second factor introduced in each step was a *randomness coefficient* (d=0.15) that valued the fact that *walkers* do not always follow the links. Sometimes they *randomly jumped* to other pages.
3. Pages geographically closest to the user were prioritized.

It is worth noting that GPR introduced some important differences regarding rankings according to JIF. For example, GPR provided an ordering of *web pages*, and not of the entire websites, *as the individual nodes of the network*, and it did not restrict the set of web pages whose links were assessed. Furthermore, the second step of the algorithm implied *a probability assignment*, while the third step of the algorithm implied encouragement of the local because it was considered that *proximity usually was relevant for the user*.

As consequence, GPR is *a ranking in order of decreasing expected utility for the user who makes the search*; the first links shown by the algorithm should be the ones Google *expect* to provide the user with the *highest utility*. I.e., they are the ones which allow the user *to maximize his/her expected utility*.

Subsequently, Chen, et al. (2006) used the GPR algorithm to review the citation of scientific articles, suggesting a higher value for *d* when reviewing scientific citations ($d = 0.50$) so the shared citation phenomenon was taken into account. The authors found that the algorithm allowed the detection of some articles of high relevance that had gone unnoticed.

That same year, Bollen, et al. (2006) combined the perspective of Pinski & Narin (1976) with the GPR algorithm, suggesting the *Weighted Page Rank* algorithm. This algorithm implied a modification regarding Chen, et al. (2006) analysis. The *nodes* of the analyzed network no longer are the articles (equivalent to web documents) but the scientific journals (equivalent to complete websites), whose centrality is assessed by counting the interactions between individual articles. The above approach was subsequently adopted by ScimagoLab (2007) for the *Scimago Journal Rank (SJR)*, and by Bergstrom (2007) for the *Eigenfactor* of a journal as a sum of the weighted value of citations received by published articles by the journal in one year, and the *Article Influence* dividing the *Eigenfactor* by the number of published articles.

A few years later, West, et al. (2013) adapted the *Eigenfactor* to also assess authors' relevance, proposing that the citations received by each article should be double-weighted:

- The citations each publication receives are weighed by dividing them by the number of authors (this idea was already advanced in Price, 1981)
- Each citation is assigned a value inversely proportional to the number of references included in the citing article, so the fact that some articles include a very high number of references is taken into account.

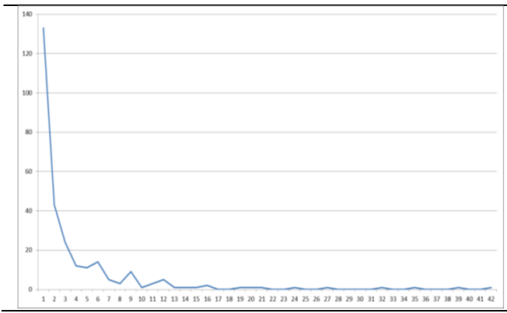
Besides these two weights, also the calculation of each node (author) centrality in the network of authors implies a third weighting of the citations that each author receives.

2.2.2 Citation assessment from Distribution Curve analysis

McAlister (1879) reviewed several types of data sets and concluded that the arithmetic mean did not correctly characterize some distributions. When a data set is characterized by a value we expect that when randomly choosing a subset (or an element) the *error by excess or default is similar.* However, for skewed distributions the arithmetic mean fails to meet this condition.

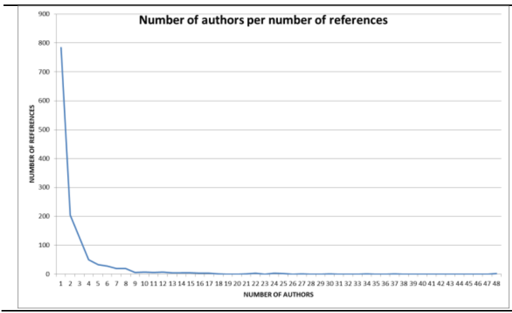
In the following years, several interesting contributions were made to the field of economics. In 1897, Pareto reviewed the income distribution in societies and detected that by dividing the population into income steps, he obtained skewed distributions. Later on, some authors proposed that it is possible to approximately characterize these distributions by logarithmic laws (Pigou, 1912; Dalton, 1920).

In 1922, Dresden detected this type of distribution in bibliometric analysis. He reviewed the number of communications presented at annual meetings, to assess the relevance of the scientific contributions of the members of the Chicago Section of the *Mathematical American Society* and he found a *skewed distribution*. The author stated that characterizing that distribution through the mean value was a conceptual error.



Number of communications by each author of the Chicago section of the AMS. Graph: Own elaboration from Dresden, 1922: 304

In 1926, Lotka reviewed the number of each scientist’s publications referenced in two prestigious yearbooks, finding another skewed distribution. He found out that it is possible to approximately relate *the number N of authors with n contributions N(n)* with the number of authors with a single contribution N(1) by means of a logarithmic law:



Lotka (1926) proposed approximately characterizing the curve as:

$$N(n) \sim N(1) * \frac{1}{n^2}$$

Graph: Own elaboration from Lotka, 1926: 319

The title of Lotka's article ‘*The frequency distribution of scientific production’* is meaningful because *a stable distribution of frequencies implies a probability assignment*. I.e., Lotka continues the paradigm initiated by McAlister, Pigou and Dalton.

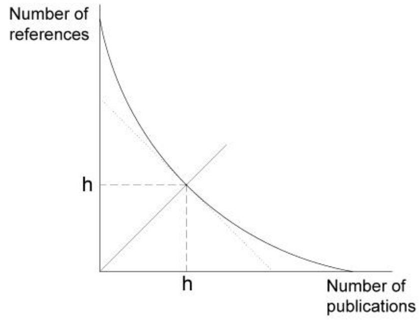
This perspective was subsequently continued by Simon (1955), who analyzed the possible causes by which several different nature phenomena (one of them being the distribution of scientists according to the number of published articles) presented a similar skewed distribution. The author proposed there is a *non-linear probability structure underlying these phenomena* that can be approximately described by the Beta function.

In 1965, Price reviewed the structure of the network of scientific citations finding also a skewed distribution in the citations of articles. In 1976 he continued with the ideas advanced by Lotka and Simon and showed that the Beta function also explains the distribution of citations, thus linking this distribution to an underlying probability function. The Beta function thus acquires a predictive character, and it would be the reason why ‘success generates success’; what Price designated as ‘Cumulative Advantage Distribution’. In 1986, he proposed that the distribution of citations can be approximately characterized as $1/n^3$.

In 1998, Redner reviewed the citations of scientific articles, also finding a $1/n^3$ distribution that led him to state that “the citation distribution provides (…) a much more complete measure of (each journal) popularity than the average or total number of citations” (ibid., p. 4). The author found that almost half of the articles indexed in the ISI had never been cited, which challenged the validity of the JIF for assigning quality to articles published by a journal: almost 50% of the time the JIF would be an incorrect characterization.

Besides, Redner showed that the number of citations in the four years after publication was sometimes similar for articles that would later become classics (hence, read for many years after) and for articles that were immediately forgotten so never cited again. As consequence, Redner asserted *it was necessary to increase the periods of analysis of citations.*

In 2005, Hirsch proposed the first characterization of the *impact and relevance* of a scientist's publications derived from the logarithmic and power laws enunciated by previous authors; the h-index. "A scientist has index h if h of his or her Np papers have at least h citations each and the other Np-h papers have ≤ h citations each" (Hirsch, 2005, p. 16569)



Graphic description of the h-index (Hirsch, 2005, p. 16,570) showing its relation to the distribution of citations. Hirsch index depends on both the number of published works and the number of citations received by each publication.

Hirsch also proposed using the h-index to characterize faculties, departments and research laboratories. However, he warned us that the h-index was distorted if many works were co-authored.

Another challenge to the h-index is that if scientists' distributions of citations are *similar, h is then a function of the number of published works* (Glänzel, 2006). As consequence, the h-index assigns a higher value to scientists who have published more works, who generally are those who have been publishing for a longer period.

To conclude, in recent years, several proposals have been made that adapt the h-index to make it more sensitive to the citation distribution curve (e.g.; Zhang, 2013; Lando & Bertoli, 2014), many of which seek to assign greater value to the group of most cited works.

2.3 Recap

We have reviewed a large number of indexes to assess the relevance of scientific publications, authors, journals and institutions building on the citation count of their published works. A brief recap is convenient to clarify some issues.

First, it is possible to group reviewed indices into two types of measures:

- The first group of indexes seek to characterize the complete set of work published by a journal, author or Institution. They are *Scale Dependent Indexes (SDI)*: the more published works (or the longer the scientist's academic career), the higher the measure may be (although not necessarily it is so).
- A second group of indexes seek to characterize the expected citation of the next published item (or a randomly chosen item among those published) by a journal or scientist. They are *Scale Independent Indexes (SII)*. A greater number of published works or a longer academic career, do not imply a greater value of this measure.

Combining above differentiation of measures with the three approaches used to review the indexes, it is possible to group the indexes into the following categories:

TABLE 1. TYPES OF CITATION INDEXES

	LINEAR APPROACHES	NON-LINEAR APPROACHES	
		DISTRIBUTION CURVE	CENTRALITY IN GRAPH/ NETWORKS THEORY
SDI Scale Dependent	Nº Publications (Au, Jr) (1) Impact (Au, Jr)	h-index (Au, Jr) hl-norm (Au)	Total Influence (Jr) Eigenfactor (Au, Jr)
SII Scale Independent	JIF (Jr) SNIP (Jr) Altmetrics (Au, Jr, art)	hIa-norm (Au)	Influence per publication (Jr) Article Influence (Jr, art) SJR (Jr)

Notes: Codes mean the following: (Au) It can be used to assess authors; (Jr) It can be used to assess Journals; (art) It can be used to assess articles. University rankings mostly adopt a linear approach; the number of publications (total or normalized), and Nº citations (total or normalized citations/researcher) and in some cases normalizing citations according to the usual citation in each knowledge area (a similar approach to SNIP). In all cases, many publications are excluded: ARWU and U-multirank do not take into account publications not indexed by WoS, while THE and QS do not take into account those not indexed by Scopus.

Secondly, the review has shown different authors state they are actually measuring different concepts: the *desirability of acquisition* (Gross & Gross, 1927); *impact* as *relevance* (Garfield, 1963); *influence* as *importance* (Pinski & Narin, 1976; Bergstrom, 2007 ...), *influence* as *quality* (Bollen, et al., 2006), *popularity* (Redner, 1998), *relevance* and *impact* (Hirsch, 2005), *attention* (Kortelainen, et al., 2017)…

This makes it necessary to answer the question: what should the indices measure? To answer, it is necessary to consider the indexes' goal, which is *being a rational criterion for allocating scarce collective resources*. It may be the money spent by a university library to acquire journals; the time spent by a researcher reading other papers to get informed, etc. Therefore, the indices should measure utility in compliance with the requirements of utility functions. We, therefore, arrive at the need to quantitatively characterize skewed distributions, by means of functions that satisfy the conditions required for collective decision-making functions. This brings us closer to the field of economic theory.

3 An economic approach to the characterization and comparison of skewed distributions: the equivalent (mean) income

Citation indexes have an operational objective; i.e., their *raison d'être* is to serve as a criterion for collective decision-making. Assuming that the citation of each published work informs of the utility said work has provided to the community, the indexes must interpret and translate this variable (the number of citations) referred to different works into a measure of the collective utility provided by a set of works.

To do so, the indexes must satisfy the conditions required of (expected) collective utility functions. The problem lies in, given the usually highly skewed distribution of citations received by a set of articles, characterizing it by an average value which can be interpreted in terms of the collective utility generated.

It is worth remembering that, since Pareto (1897) tried to assess the optimality of the highly skewed curves of economic resource distribution among the inhabitants of several European regions/nations, the characterizing of skewed distributions have received much attention in the field of economics. The problem of comparing distributions required defining a way of characterizing these distributions that would allow comparison. Economists have adopted two ways of dealing with this problem:

- Using indexes to characterize the distribution/concentration of economic resources. We find the proposals of Lorenz (1905), Gini (1914), Pietra (1915), Herfindahl Hirschmann (1949), Theil (1969), and Atkinson (1970).
- By characterizing the distribution with a 'representative' value, which can be interpreted in terms of 'equivalent income' or generated welfare.

The latter is the relevant approach for the present text, which we review below: the characterization of *the equivalent average (equally distributed or per capita) income*.

3.1 The equivalent (mean) income.

As noted above, an early contribution to the characterization of skewed distributions was made by McAllister (1879), who stated that the geometric mean did characterize these distributions better than the arithmetic mean. Later, Pigou (1912) proposed to approximately characterize them by logarithmic laws.

In 1920 Dalton specified that the purpose was not to measure the inequality of the distribution of economic resources, but the effects of such unequal distribution on the total amount of generated economic welfare. The objective was to determine the *equivalent per capita (or average) income*; that which is obtained by taking into account the effects that the skewed distribution of economic resources had on the generation of welfare.

Dalton asserted that the maximum welfare would be obtained in a completely equal distribution, thus any deviation from complete equality implied a *reduction in the obtained welfare*. The author considered the Gini coefficient as the best available inequality indicator and *suggested using the harmonic or geometric mean of the individuals' income/wealth, to characterize the welfare generated by each particular distribution.*

Later authors (Westwood, 1939; Aigner & Heins, 1967) continued this paradigm. In 1970 Atkinson defined the *equally distributed equivalent income level* as "the level of income per head which if equally distributed would give the same level of social welfare" as the revised distribution (ibid., p. 250)[i]. This definition implies characterizing the *average equivalent income* as:

$$r_{eq} = \bar{r} * (1 - I) \tag{1}$$

I_ Inequality; r_{eq}_ Per capita income that generates equivalent welfare if equally distributed; $\bar{r}$_ Average income of the distribution.

Therefore:

$$I = 1 - \frac{r_{eq}}{\bar{r}} \tag{2}$$

Atkinson introduced a coefficient e for inequality aversion. If e equals 1, then the *average equivalent income* is the *geometric mean* of incomes, agreeing with earlier proposals by McAllister, Pigou, and Dalton.

Subsequently, other authors (e.g. Alkire & Foster 2010) have continued this paradigm, incorporating the diminishing marginal utility of economic resources.

3.2 The concept of equivalent income applied to scientific citation

The above contributions can be translated to the field of scientific citation evaluation. Citation distributions of sets of scientific works are highly skewed. Characterizing these distributions in terms of the (collective) utility scientific works provide to society is a problem formally identical to that faced by economists in characterizing the (collective) welfare generated by skewed distributions of economic resources.

Therefore, the concentration of citations in a few scientific works implies a wasted potential for collective utility creation; i.e., *the same set of scientific works would generate the same collective utility having fewer citations, if these were equally distributed among the works.*

Inversely expressed, given a set of scientific works with a total number of citations and a distribution of citations, citation indexes should allow us to determine the number of citations that, equally distributed among all publications, would imply the same *total utility* and *average utility*. These numbers of citations are what we respectively designate as *equivalent citation* and *equivalent average citation*. Thus, we arrive at the two citation indexes types reviewed above:

- The concept of *equivalent citation index*, which expresses the *total utility* obtained if citations were evenly distributed among all works, corresponds to that of *Scale-Dependent Indices (e.g., H-index).*

- The concept of the *average equivalent citation index*, which expresses the average utility obtained if citations were distributed uniformly among all works, corresponds to that of the *Scale-Independent Indexes (e.g., JIF)*.

This equivalence between these citation indexes and their economic counterparts (equivalent income and equivalent average income), allows us to characterize both indexes as:

SDI

$$C_{eq} = C * (1 - I) \tag{3}$$

SII

$$\overline{c_{eq}} = \bar{c} * (1 - I) \tag{4}$$

I_ Inequality of citation distribution; Ceq_ Equivalent Citation; C_ total number of citations received by all works; $\overline{c_{eq}}$ _ Average Equivalent Citation; $\bar{c}$_ Average Citation

Consequently, the relationship between both indexes is also defined:

$$C_{eq} = n * \overline{c_{eq}} \tag{5}$$

Where n is the number of works in the set.

3.2.1 An axiomatic framework for equivalent average citation indexes

Economic theory has not yet developed a system of axioms for equivalent income indexes. Thus to define the system of axioms these indices should satisfy we must resort to the axiomatic for inequality measurement. The fact that *equivalent citation is related to total citation through the inequality of the distribution implies that the characterization of citation indices must comply, with* some adaptations, *with inequality indices' axiomatic.*

We list below a compilation of axioms proposed for inequality indices.

3.2.2 System of axioms for inequality indexes

We detail below a formal framework for inequality indices that summarizes the proposals of previous authors (Schwartz and Winship 1979: 6ff; Neves & Perez-Duarte 2019: 10). There are some prerequisites and three axioms that inequality indices must fulfil:

- Properties shared by all concentration/inequality indices:
  - Their value is zero if income is equally distributed and positive otherwise approaching 1 as concentration increases.
  - They are impartial since they do not depend on who owns what income.
- **Axiom 01. Principle of transfers (Pigou, 1912; Dalton, 1920).** Inequality is reduced if we transfer income from a richer person to a poorer person. The transfer should not be so large that the recipient becomes richer than the donor.
- **Axiom 02. Population symmetry (Dalton, 1920).** If two populations are equal in size and income is identically distributed the inequality of each population is identical, and equal to the inequality of the combined population. In other words: the index remains unchanged if the distribution is replicated a finite number of times.
- **Axiom 03. Scale invariance (Atkinson, 1970).** If the incomes of individuals increase in the same proportion, income inequality does not change. In other words, the size of the pie is not relevant in determining inequality; what is relevant is the relative share that each person receives. It implies the units used to measure wealth or income can be ignored.

The three axioms above imply the Lorenz criterion; if when representing two distributions, the curves do not intersect, the outermost one implies a more unequal distribution of money.

However, Lorenz curves often intersect. In such cases, two or more formulas may satisfy the above axioms, while differently arranging the distributions (Atkinson 1970).

- **Ax.04. Principle of diminishing transfers.** A measure of inequality should take into account the differential impact of transfers between different points in the distribution. Consider two persons with incomes of X and Y, with X < Y. The principle of diminishing transfers states that the reduction in inequality attributable to a transfer from the person with income X to another person with income X - C (where 0 < C < X) is greater than the reduction attributable to an equal transfer from the person with income Y to someone with income Y - C.

Noteworthy, although these properties ensure that the indicators behave reasonably, they are not sufficient to single out a measure or a family of measures, and additional properties must be incorporated to ensure the uniqueness of the measure.

3.2.3 Axiom system for equivalent citation

From the above system of axioms, substituting in the equations relating to Inequality with both Equivalent Citation and Equivalent Mean Citation it is possible to state the conditions and axioms that these indexes must fulfil. In addition, it is necessary to consider three issues not usually contemplated by the existing inequality indexes:

- **Monotonicity.** If a work increases its number of citations, the equivalent citation value and equivalent average citation should increase. Conversely, if a work reduces its number of citations, the equivalent citation value and equivalent average citation should be reduced.
- The citation of many works is zero. Therefore, *the index should be computable if there are works with no citations*, and should only have zero value if no work has citations.
- **Additive Invariance (Atkinson, 1970).** If each work in the set increases its citations by k, the *Average Equivalent Citation* increases by *k* (and consequently, the *Equivalent Citation* increases by *n*k*, where *n* is the number of works).

Based on the above contributions, we can propose six axioms that the quantitative characterization of Equivalent Citation (C_{eq}) and Mean Equivalent Citation ($\overline{c_{eq}}$) must comply with:

AXIOMS:

Ax. 00: Limits

There are minimum and maximum limits to the value of the indexes.

- 01. If no published item *i* of the set of works *I* receives at least one citation, the value of the index is zero, and conversely, the value of the index is zero iff no published item receives any citation.

$$C_{eq} = 0 \wedge \overline{c_{eq}} = 0 \leftrightarrow \forall i \in I: C_i = 0 \tag{6}$$

- 02. (SDI): The maximum value of the C_{eq} index is the sum of the number of citations c_i received by each publication:

$$C_{eq} \leq \sum_{i=1}^{n} C_i \tag{7}$$

- 03. (SII): The maximum value of the c_{eq} index is the sum of the number of citations c_i received by each publication divided by the number of publications, i.e., the maximum value is the average citation:

$$\overline{c_{eq}} \leq \frac{1}{n} * \sum_{i=1}^{n} C_i \tag{8}$$

Ax.01: Monotonicity.

The indexes should be monotonic with respect to the total number of received citations.

- If no publication reduces its number of citations and at least one of them increases it, the index value increases.
- If no publication increases its number of citations and at least one of them decreases it, the index value decreases.

Ax. 02: Population Symmetry (Dalton, 1920):

If the distribution is replicated a finite number of times *j*, the Mean Equivalent Citation does not change its value, while the Equivalent Citation is multiplied by *j*.

Ax.03: Principle of Transfers (Pigou, 1912; Dalton, 1920):

If an incorrect assignment of *t* citations to a work *j* with citation number c_j is corrected by assigning it to a work i with citation number c_i such that $c_i - t > c_j + t$, both the *mean equivalent citation* and the *equivalent citation* increase.

This axiom implies a weak principle of diminishing marginality of citation. If $c_i > c_j$ then an additional citation of work j increases the *equivalent citation* and *mean equivalent citation* of the set to a greater value than an additional citation of work i.

Ax. 04: Scale Invariance (Atkinson, 1970):

If the total number of citations for each of the works is multiplied by a factor *k*, the resulting *equivalent citation* and *mean equivalent citation* values are multiplied by *k*.

Ax. 05: Additive Invariance (Atkinson 1970).

If each work in a set of *n* works increases its number of citations by a value *p*, the *mean equivalent citation* increases by that amount *p*, and the *equivalent citation* increases *n*p*.

3.3 Compliance with the axioms by indexes

A summary table indicating axiom compliance for each index is included below:

TABLE 3. COMPLIANCE WITH AXIOMS BY CITATION INDEXES

		Ax.00. Limits			Ax.01 Monotonicity	Ax.02. Symmetry Population		Ax. 03 Principle of Transfers	Ax. 04. Scale Invariance	Ax.05. Additive Invariance
		01	02	03		01	02			
TYPE 01 DSM	Number of published works (Dresden, 1922)	NO	NO	-	NO	YES	-	NO	NO	NO
	Impact (Gross & Gross, 1927)	YES	YES	-	YES	YES	-	NO	YES	YES
	Total Influence (Pinski & Narin, 1976)	YES	YES (1)	-	YES	YES	-	NO	YES (1)	YES (1)
	H-index (Hirsch, 2005)	YES	NO	-	NO	NO	-	NO	NO	NO
	Square root of total citations	YES	NO	-	YES	NO	-	NO	NO	NO

TABLE 3. COMPLIANCE WITH AXIOMS BY CITATION INDEXES

		Ax.00. Limits			Ax.01 Monotonicity	Ax.02. Symmetry Population		Ax. 03 Principle of Transfers	Ax. 04. Scale Invariance	Ax.05. Additive Invariance
	Eigenfactor (Bergstrom, 2007)	YES	YES (1)	-	YES	YES	-	NO	YES	YES (1)
TYPE 02 ISM	JIF (Garfield, 1962)	YES	-	YES	YES	-	YES	NO	YES	YES
	Influence per publication (Pinski & Narin, 1976)	YES	-	YES (1)	YES	-	YES	NO	YES	YES (1)
	Article Influence (Bergsrtom, 2007)	YES	-	YES (1)	YES	-	YES	NO	YES	YES (1)
	SNIP (Moed, 2009)	YES	-	YES (1)	YES	-	YES	NO	YES	YES

NOTES: A blank cell with a dash implies that the criterion does not apply to the reviewed type of measure.

(1) It is met when considering the received citation after weighting.

(2) Noteworthy, the H-Index breaches several axioms due to its non-monotonicity.

As shown above, no index currently in use meets all the axioms. To move forward, two indexes that satisfy all the axioms are explained below.

3.4 Formulation of two consistent citation indexes

Given a set *I* of *n* scientific publications; with a total of C citations received by all the publications in the set, C_i being the citations received by each scientific publication i. We designate as N-Index the *equivalent citation* and Nr-Index the *mean equivalent citation*. The procedure for calculating the indices is:

1. A weighting coefficient k_i is calculated for each publication i:

$$k_i = 1 + \left(\frac{1}{n} * \sum_{j=1}^{n} \frac{c_j - c_{\min}}{c_{\max} - c_{\min}}\right) - \frac{c_i - c_{\min}}{c_{\max} - c_{\min}} \tag{9}$$

2. Two different formulations are calculated:
    a. Nr-index (Type 02 -SII): to characterize the *equivalent mean citation*

$$Nr = \left(\frac{1}{n} * \sum_{i=1}^{n} (c_i - c_{\min}) * k_i\right) + c_{\min} \tag{10}$$

    b. N-index (Type 01 - DSI): to characterize the *equivalent citation*.

$$N = n * Nr \tag{11}$$

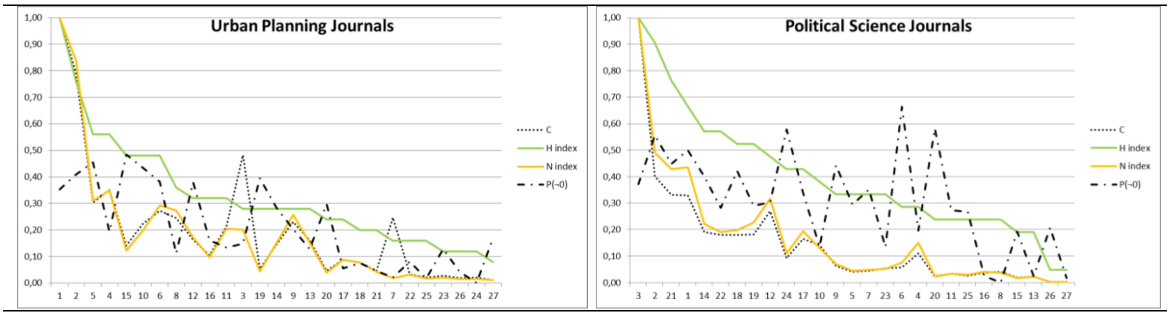
Therefore, using these two indexes, it is possible to characterize the *equivalent citation* (N-index) and the *mean equivalent citation* (Nr-index) of a set of published works.

4 *Review of a Sample of Journals*

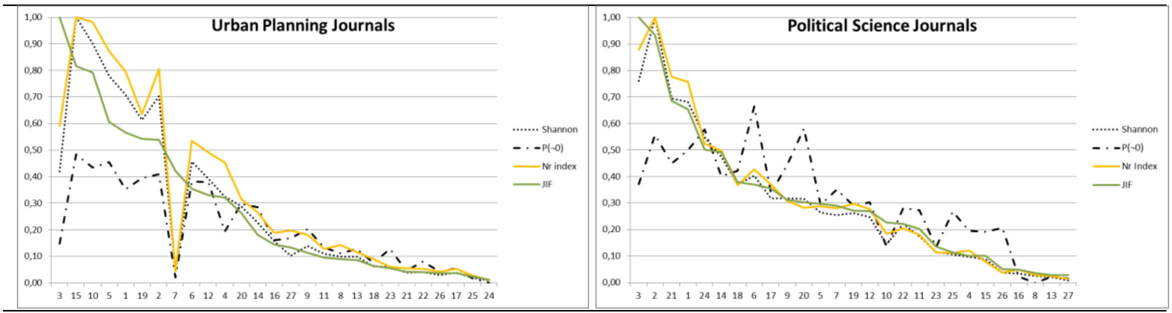
To provide empirical contrast a set of 27 Urban Planning and a set of 27 Political Science journals are reviewed below. The citations of the articles of each journal have been counted using the Publish or Perish® program on Google Scholar. First, the values of the most currently used are compared graphically with herein proposed N and Nr indexes:

GRAPHICAL COMPARISON OF MOST COMMONLY USED AND PROPOSED INDEXES

SCALE-DEPENDENT INDEXES



SCALE-INDEPENDENT INDEXES



Source: Own elaboration. The values of the represented indexes have been normalized in the range 0-max. P (¬0) is the percentage of articles that receive at least one citation. Scale-dependent indexes are ordered according to the decreasing value of the H-index. Scale-independent indexes are ordered according to the decreasing value of JIF.

The graphs on the upper row confirm the high dependence of the H index on the total number of citations C and its independence from the percentage of articles without citations P(0). Moreover, since the H-index is a scale of natural numbers, we observe journals with similar citation structures but different H values, while other journals with more different citation profiles receive the same value of H. In contrast, the proposed N-index shows greater independence from the total number of received citations and is more sensitive to the differences between journals.

Likewise, the graphs on the lower row highlight the major shortcoming of the JIF. It can give a high score to a journal that has published one or two highly cited articles even if the rest of its articles have few or no citations (e.g., urban planning journals 3 and 7). This confirms the JIF does neither inform us of the predicted (expected) citation of a typical journal article; nor is it a measure of equivalent citation. On the other hand, the greater sensitivity of the Nr-index to the percentage of articles that do not receive any citation is observed, which anticipates its greater descriptive character of each element of the set.

It is also relevant to review the existing correlations between the indexes /parameters:

TABLE 5. CORRELATIONS BETWEEN INDEXES

		SCALE-DEPENDENT SDI					**SCALE-INDEPENDENT SII**				
		n	C	H-index	N-index	P(0)	JIF	n*Gini	n*Shannon	n*IHH	Nr-index
SDI	No publications (n)	-	0,35	0,26	0,38	0,36	-0,19	-0,28	-0,18	-0,17	-0,16
	No Citations (C)		-	0,87	0,97	-0,36	0,68	0,52	0,62	0,62	0,66
	H-index			-	0,92	0,73	0,73	0,67	0,81	0,80	0,83
	N-index				-	-0,41	0,62	0,50	0,65	0,64	0,69
	P(0)					-	-0,70	-0,80	-0,84	-0,82	-0,82
SII	JIF						-	0,91	0,92	0,93	0,94
	n* Gini Coefficient							-	0,92	0,92	0,91
	n* Shannon Entropy								-	0,99	0,99
	n*Herfindahl Hirschman									-	0,98
	Nr-index										-

Source and Notes: Own elaboration. The mean value of the correlations between indices calculated separately for both samples of journals is shown. P (0) is the Percentage of Articles that do not receive any citation, i.e., which have 'zero impact'.

The analysis of the correlations between indexes in the two sets of journals analysed shows several relevant issues:

- The H index shows high correlations (R=0.73) with the percentage of articles receiving no citations P(0) and (R=0.87) with the number of citations C, suggesting that *H is mostly a function of the number of citations*, but it fails to detect situations where most articles have never been cited.
- *The N index* shows a very high correlation with C (0.97) and a significant and negative correlation with P (0) (-0.41), suggesting that *it is a balanced function of both variables*. If the number of citations increases, we expect the N index to be higher, but if the percentage of non-cited published works increases, we expect the N index to be lower. If citations are equally distributed among all publications, N equals C.
- A similar correlation is observed between JIF-C and Nr-C (0.68 and 0.66), but a higher negative correlation of the latter with P(0) (-0.82 versus -0.70), showing some *higher predictive quality of the Nr index versus JIF*.
- Finally, the high correlation between H and N (0.92) shows that they transmit a large amount of shared information. The same holds for JIF and Nr (0.94). If citations are equally distributed among all publications, Nr is equal to JIF.

The analysis supports the higher optimality of the two herein indexes proposed here, N and Nr, compared to the currently used H and JIF indexes. Finally, it is useful to plot the journals according to JIF and Nr without normalization:

SCALE-INDEPENDENT MEASURES

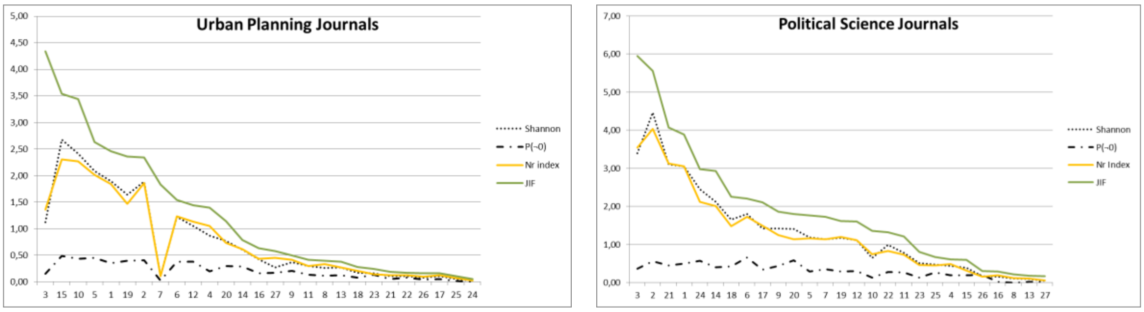


Image: Own elaboration

The graph shows several interesting issues:

- The *average equivalent citation* is lower (37% for urban planning journals and 34% for political science journals) than the JIF suggests, and *the more skewed the distribution of citation between the articles, the greater the reduction between the average citation of articles and the equivalent citation of a typical journal article is.*
- There is a very high resemblance between the value of the proposed average equivalent citation index and the value we obtain by multiplying the Shannon relative entropy by the average citation. *This allows us to conceptualize the Nr Index as the average citation value once we eliminate the uncertainty in the message.*

Precisely, the high resemblance between the Nr Index and Shannon's relative entropy indicates that *Nr accounts also for the uncertainty involved in the prediction of future citation of the articles; it is also an estimate of their expected citation*[ii]. If we have to choose a journal to publish an article, Nr is a better indicator than JIF of which journal provides the highest 'expected article citation'. Given the linkage between N and Nr, this predictive nature is also a feature of the N-index.

5 Recap, conclusions and some pending issues

We have reviewed the quantitative characterization of scientific agents (authors, departments, faculties, journals....) according to the number of times their published works are cited by other scientists. The different characterization proposals try to assess two different issues leading to two types of measures:

- *Scale-dependent measures:* their goal is to assess the overall relevance as a *total equivalent citation of the set of publications* of an author, research laboratory, department, faculty or journal.
- *Scale-independent measures:* their objective is to assess the *mean equivalent* (i.e., that which best characterizes a randomly chosen publication) *or the expected citation* (of the next publication) of an author, research laboratory, department, faculty or journal.

The objective of these measures is to be an objective parameter for the allocation of collective resources. This is an essential objective of economic theory: to model the collective utility provided by different 'states of the world' to facilitate the allocation of scarce resources.

The high bias and nonlinearity of the distribution of citations make the use of linear measures such as the arithmetic mean (e.g., the JIF) incorrect. Thus, we have adopted the concepts of equivalent income and mean equivalent income, specifically developed to characterize skewed distributions. Building on these concepts, two indexes have been explained: a scale-dependent index (N) and a scale-independent index (Nr).

A formal and empirical review of the existing and proposed indexes has been undertaken. The *formal review* shows that N and Nr indexes satisfy a set of six axioms or self-evident conditions and several prerequisites that no currently existing index satisfies. The *empirical review* shows some issues that challenge the validity of most currently used indexes: H-index and JIF:

- JIF does neither report the equivalent citation of a typical article published by the journal nor the expected citation of the next published article (which might be approx. 35% lower) and may lead to inaccurate orderings in terms of journals' relevance.
- The H-index shows an unacceptable correlation with the percentage of un-cited articles; an increase in the number of un-cited works should reduce the value of the index.

On the other hand, the empiric review supports the optimality of N and Nr indexes, which provide consistent values in all cases. Besides, the high correlation between H-N indexes and JIF-Nr indexes confirms that *herein proposed indexes provide similar information to a great extent to the currently most used indexes*. That is, *herein proposed indexes measure the phenomena that the H and JIF indices attempt to measure, but they do so in a manner consistent with the mathematical conditions required of collective utility functions.*

Lastly, an advantage of the N-index over currently used indexes is that it decouples scientists' relevance assessment from the number of published works, it *prioritizes quality over quantity* moving towards a paradigm where the progress of science is measured based on the actual relevancy, and not on the number, of published works[iii]. The N-index enables this transition while allowing a fairer assessment of social and humanities scientists, who tend to publish books and monographs instead of articles in journals.

6 References

Alvira, R. (2014) *A mathematical theory of sustainability and sustainable development.* Amazon: United Kingdom. ISBN-13: 978149757814

Alvira, R. (2018) 'Medir la Brecha Retributiva de Género', comunicación en *I Congreso Internacional sobre Equidad, Educación y Género,* Murcia 19-21 septiembre de 2018.

Bhattacharya, J., & Packalen, M. (2020). Stagnation and Scientific Incentives. *PSN: Science & Technology* (Topic).

Bergstrom, C. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *School & Research Library News,* 68(5), 314-316. Retrieved from https://www.eigenfactor.com

Bollen, Johan; Rodriguez, Marko A. & Van de Sompel, Herbert. 2006. Journal Status, *Scientiametrics*, 69(3), 669–687.

Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, 33 (1998), 107–117.

Chen, P.; Xie, H., Maslov, S. & Redner, S. (2006). Finding Scientific Gems with Google. Retrieved from https://arxiv.org/abs/physics/0604130

Cole, F. J., & Nellie, B. E. (1917). The history of comparative anatomy: Part I.—A statistical analysis of the literature, *Science Progress* (1916-1919), 11(44), 578–596

Dalton, H. (1920). The measurement of inequality of incomes, *The Economic Journal*, 30(119), 348-361

DORA (2012). San Francisco Declaration on Research Assessment, Putting science into the assessment of research. Retrieved from https://sfdora.org/

Dresden, A. (1922). A report on the scientific work of the Chicago Section, 1897-1922, *Bulletin American Mathematical Society*, 28(6), 303-307

Garfield, E. (1955). Citation Indexes for Science, A New Dimension in Documentation through Association of Ideas, *Science*, 122 (3159), 108-111

Garfield, E. (1963). *Science Citation Index 1961*

Garfield, E. y Sher, I. H. (1963). New Factors in the Evaluation of Scientific Literature, Through Citation Indexing, *American Documentation*, 14(3)

Garfield, E. (1970). Citation Indexing for Studying Science, *Nature,* 227, 669-671

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

Glänzel, W. (2006). On the H-index – A Mathematical approach to a new measure of publication activity and citation impact, *Scientiametrics*, 67(2), 315–321. Retrieved from https://www.researchgate.net/

Gross, O.L.K. & Gross, E.M. (1927). College libraries and Chemical Education, *Science*, New Series, 66(1713), 385-389

Harzing, A. (2007). 'Publish or Perish', http://www.harzing.com/resources/

Harzing, Anne-Wil; Alakangas, Satu & Adams, David. 2014. hIa: An individual annual h-index to accommodate disciplinary and career length differences, *Scientiametrics*, 99(3), 811–821. Retrieved from https://harzing.com/publications

Heckman, James J.& Moktan, Sidharth. 2020. Publishing and Promotion in Economics: The Tyranny of the Top Five. *Journal of Economic Literature* 2020, 58(2), 419–470

Hicks, Diana; Wouters, Paul; Waltman, Ludo; de Rijcke Sarah & Rafols, Ismael. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520, 429–431 (2015). https://doi.org/10.1038/520429a

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output, *PNAS*, November 15, 2005, 102(46), 16.569–16.572

Kaltenborn, K-F & Kuhn, K. (2004). The journal impact factor as a parameter for the evaluation of researchers and research, *Rev Esp Enferm Dig*, 96, 460-476.

Kortelainen, T.; Katvala, M. & Länsman, A.S. (2017). Attention and Altmetrics. Proceedings of the Ninth International Conference on Conceptions of Library and Information Science, Uppsala, Sweden, June 27-29, 2016. *Information Research*, 22(1)

Lando, T. & Bertoli-Barsotti, L. (2014). A New Bibliometric Index Based on the Shape of the Citation Distribution, *Pols one*, 9(12): e115962

Lawrence, P. A. (2007). The mismeasurement of science, *Current Physiology,* 17(15): 583-5

Lawrence, P. A. (2008). Lost in publication: how measurement harms science, *Ethical Science and Politics*, 8, 9–11

Lawson, S.; Gray, J. & Mauri, M. (2016). Opening the Black Box of Scholarly Communication Funding, *Open Library of Humanities*, 2(1), 1–35

Leydesdorff, L. & Opthof, T. (2010). Scopus's Source Normalized Impact per Paper (SNIP) versus a Journal Impact Factor based on Fractional Counting of Citations, *JAS&IST*, 61(11), 9-11

Lotka, A. J. (1926). The frequency distribution of scientific productivity, *Journal of the Washington Academy of Sciences*, 16(12), 317-323

Merton, R. K. (1968). The Matthew Effect in Science, *Science*, 159(3810), 56-63

Moed, H. F. (2009). Measuring contextual citation impact of scientific journals, *Journal of Informatics*, 4(3), 265-277

Park, M.; Leahey, E. & Funk, Russell J. 2023. Papers and patents are becoming less disruptive over time. *Nature*, vol. 613, 5 January 2023,

Pigou, A. C. (1912). *Wealth and Welfare*. London: Macmillan.

Pinski, G. & Narin, F. (1976). Citation Influence for Journal Aggregates of Scientific Publications?, *Information Processing & Management*, 12, 297-312

Price, Derek J. De Solla (1986). *Little Science Big Science … and Beyond*, New York: Columbia University Press

Price, Derek J. De Solla (1965). Networks of scientific papers, *Science*, 149(3683), 510-515

Price, Derek J. De Solla (1976). A General Theory of Bibliometric and Other Cumulative Advantage Processes, *Journal of the American Society for Information Science* 27 (5-6), 292-306

Priem, J.; Taraborelli, D.; Groth, P. & Neylon, C. (2010). Altmetrics: A manifesto. Retrieved from http://altmetrics.org/manifesto

Redner, S. (1998). How Popular is Your Paper? An Empirical Study of the Citation Distribution, *Eur. Chem. J. B*, 4(2), 131-134

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research, *BMJ*, 314, 498–502

Simon, H. (1955). On a class of skew distribution functions, *Biometrika*, 42(3/4), 425-440

West, J. D.; Jensen, M. C.; Dandrea, R. J.; Gordon, G. J. & Bergstrom, C. T. (2013). Author-Level Eigenfactor Metrics, *JAS&IST*, 64(4), 787-801. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1636719

Zhang C-T (2013). The h' Index, Effectively Improving the h-Index Based on the Citation Distribution, *Pols One*, 8(4): e59912. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3614896/

Annex 1: Indeterminacy of the formula, and its resolution.

The formula for calculating the weighting coefficients can lead to indeterminacy in the unlikely but possible case that the citation values of all works are equal, which implies that the maximum value minus the minimum value equals zero.

$$\forall i \in n, c_i = k \rightarrow \frac{C_i - min[C_i]_{i=1}^{n}}{max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n}} = \frac{0}{0} \tag{12}$$

However, it is possible to solve this indeterminacy by means of l'Hôpital's rule, taking into account that the denominator is always greater than the numerator, so the quotient is always less than or equal to 1:

$$\begin{aligned} C_i - min[C_i]_{i=1}^{n} \leq max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n} \\ \rightarrow \frac{C_i - min[C_i]_{i=1}^{n}}{max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n}} \leq 1 \end{aligned} \tag{13}$$

Thus…

$$\begin{gathered} max(ci) - min(ci) \rightarrow 0 \\ C_i - min[C_i]_{i=1}^{n} \leq max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n} \rightarrow \end{gathered} \tag{14}$$

Therefore, as all ci values approach equality, the value of the denominator gets closer to the numerator, therefore, the quotient tends to reach the value 1:

$$C_i \rightarrow max[C_i]_{i=1}^{n} \rightarrow \tag{15}$$

$$\frac{C_i - min[C_i]_{i=1}^{n}}{max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n}} = \frac{max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n}}{max[C_i]_{i=1}^{n} - min[C_i]_{i=1}^{n}} = 1 \tag{16}$$

This can be easily incorporated in the formula for calculating the coefficients, by establishing the condition that if all c_i values are equal, then all k_i values are equal to 1, therefore $Nr=c_i$ and $N = n*c_i$

Annex 2: Simplification of the formula when there are non-cited works (thus, there are ci values equal to zero)

When there are works without citations in a set the value min(C_i)=0, which simplifies the calculation formulas as follows:

1. Citations received by each publication C_i are normalized in the interval 0-max as c_i:

$$c_i = \frac{C_i}{max(C_i)_{i=1}^{n}} \tag{17}$$

2. The weighting coefficient k_i is calculated for each publication i:

$$k_i = 1 + \left[\frac{1}{n} * \sum_{i=1}^{n} c_i\right] - c_i \tag{18}$$

1. The above parameters are used in the formulations of the two indexes:

    a. Nr index (Type 02 Independent of Scale Measure):

$$N_R = \frac{1}{n} * \sum_{i=1}^{n} C_i * k_i \tag{19}$$

    b. N index (Type 01 Dependent of Scale Measure):

$$N = n * N_R \ = \sum_{i=1}^{n} C_i * k_i \tag{20}$$

In this case, the formula for Nr agrees with an earlier proposal by the Author (2018).

---

[i] These proposals build on certain transdisciplinarity. By importing concepts from the theory of decision under uncertainty, Atkinson (1970:251) drew a parallel in which the *equivalent level of equally distributed income would be the analogue of income eliminating uncertainty*. This paradigm also underlies Theil's use of Shannon's Entropy to measure economic inequality. If total income is equally distributed among individuals, we *know with certainty* what the income of any individual is: the mean value. Thus, the concentration coefficient is the uncertainty as to the message introduced by the unequal distribution of economic resources.

[ii] Atkinson (1970:251) defined the *equally distributed equivalent income* as the analogue of income eliminating uncertainty. Thus, Shannon's Relative Entropy is the distribution of citations while the concentration (the redundancy or one minus the relative entropy, Shannon, 1949) is the uncertainty introduced by the unequal distribution of citations.

[iii] Besides, the N-index: 1) Disincentives "salami publications" (Colquhoun, 2007) since authors assessment does not improve if an investigation is divided into several small publications instead of published as a single publication 2) It allows overcoming the current *Publish or Perish* paradigm, which is causing an overload to researchers, reviewers and journals, and increasing costs for public institutions which, according to Lawson, et al, 2016, could be spending almost 10,000 MM $ / year on subscriptions to journals (reducing the current publication volume would allow releasing part of this money to finance more research) and 3) By slightly modifying current incentives to scientific production, it could help to increase its efficiency, reducing the stagnation detected in the last decades (Bhattacharya & Packalen 2020; Park, et al. 2023)