

Special Features of the Author–Publication Relationship and a New Explanation of Lotka’s Law Based on Convolution Theory

L. Egghe

LUC, Universitaire Campus, B-3590 Diepenbeek, Belgium* and UIA, Universiteitsplein 1, B-2610 Wilrijk, Belgium

This article makes the obvious but rather unexploited remark that there is a structural difference between author–publication systems and, for example, journal–article systems, in the sense that articles are published in one journal but that papers can have several authors. This difference is then studied mathematically, using convolutions in order to derive the several-author case from the case of a single author per paper.

We show that Lotka’s law $\varphi(i) = C/(i + 1)^\alpha$, where $i \geq 0$ is approximately stable for all $\alpha = 2, 3, 4, \dots$, meaning that if Lotka’s law is valid in systems in which every article has one author then it is approximately valid (in a mathematical strong sense) (with the same α) in the general systems, where more than one author per paper is possible. We also show that the same is true (but in an exact way) for the geometric distribution. Hence, this theory provides intrinsic explanations of the Lotka and geometric functions.

Introduction

In Egghe (1989, 1990) the notion of “Information Production Process” (IPP) is introduced and studied, being generalized source–item relationships. Examples of IPPs are: a classical bibliography of journals and articles in them (on a certain topic), a situation in which sources are authors and items are articles published by these authors, a situation in which books are sources and their borrowings in a library are the items, or a situation in which sources are articles and items are references in (or citations to) these articles, to give just a few examples.

Let us consider the first two examples in more detail. The first example describes a “classical” bibliography consisting of a collection of journals and a selection of articles in them, which are dealing with (or relevant to) a certain topic. This situation was described by Bradford (1934),

and the underlying rank-frequency law is nowadays called “Bradford’s law.” The second example describes a situation in which one studies the production of a group of authors. The historical law involved with this situation is the so-called law of Lotka (Lotka, 1926), stating that, if $\varphi(i)$ denotes the fraction of the authors with i publications ($i \geq 1$), then

$$\varphi(i) = \frac{C}{i^\alpha}, \quad (1)$$

where C is a constant and α is usually larger than or equal to one. Most classically, the value $\alpha = 2$ is put in front, especially in theoretical models. There is a good reason for that: it is well-known (see, e.g., Egghe & Rousseau, 1990) that Bradford’s law is equivalent (in the mathematical sense) with Lotka’s law for $\alpha = 2$. This equivalence and other equivalencies are highlighted in Egghe (1989, 1990) based on duality theory: the duality (i.e., “interchangeability”) of sources and items. This duality is the common link between all the IPPs given above as examples (and many others). In general, in Egghe (1989, 1990) we define an IPP as a “generalized” bibliography of sources (the objects that produce) and items (the objects that are produced by these sources). From this viewpoint it was not only evident from a mathematical perspective that equivalencies between certain informetric laws exist, but this evidence was also present from a conceptual perspective.

This study goes deeper into the duality problem and adds a new dimension to the study of general IPPs. The first observation is very simple (we keep on comparing journal–article and author–publication situations): articles appear in one journal while publications can be written by several coauthors. This conceptual difference should open the possibility for different types of informetric laws. Nevertheless, as stated above, the historic findings of Bradford (for journal–article systems) and of Lotka (for author–publication systems) are the same. This does not seem to be evident.

The mind is puzzled when making this simple remark. Mathematically, it is indeed true that Bradford’s and Lotka’s law ($\alpha = 2$) are equivalent but this should only be

*Permanent address.

Received April 14, 1993; revised October 4, 1993; accepted November 30, 1993.

true when looking at IPPs of the “same type”—for example (as in the case of journal–articles), in which every item has exactly one source.

So let us consider, for the author–publication system, that a paper is written by exactly one author (artificial situation). Based on the above introduction, we can assume that we have a frequency function as in (1), for $\alpha = 2$: indeed, we accept Bradford’s law in this case of items having exactly one source and since this law is equivalent with Lotka’s for $\alpha = 2$, we can assume the validity of this function.

Is there a mechanism to deduce the frequency function in the general situation in which papers can have several authors, from the frequency function in the (artificial) situation in which papers have exactly one author? If so, then we wonder if we can show that function (1) ($\alpha = 2$, or more general), valid in this artificial situation, is also valid in the general author–publication situation. Only then the above-mentioned problem is solved and a new, conceptual explanation of Lotka’s law is given. This topic is discussed.

The main idea of solution is the following: let us denote by φ_1 the function, for $i \geq 0$

$$\varphi_1(i) = \text{the fraction of the authors with } i \text{ publications} \\ \text{conditional to: all the papers have only} \\ \text{one author.} \quad (2)$$

and by $\varphi(i)$ the same but in the general case that papers can have several authors. Let us also define, for $j \geq 2$, $j \in \mathbf{N}$,

$$\varphi_j(i) = \text{the fraction of the authors with } i \text{ publications} \\ \text{conditional to: all the papers have exactly} \\ j \text{ authors.} \quad (3)$$

In the general case, the set A_j of papers with j authors ($j = 1, 2, 3, \dots$) forms a subset of the total set of papers and it is clear that they are disjoint for different j and that their union is the total set again. Hence, by the principle of total chance

$$\varphi(i) = \sum_{j=1}^{\infty} \varphi_j(i) \psi(j), \quad (4)$$

where the above sum is finite in practice, and where

$$\psi(j) = \text{the fraction of the papers that have } j \text{ authors} \\ (j = 1, 2, 3, \dots). \quad (5)$$

We show that if φ_1 is a geometric distribution, then so is φ (although none of the $\varphi_2, \varphi_3, \dots$, are) and we also show that if φ_1 is a power law as in (1), then φ is approximate to a power law with the same exponent, as long as $\alpha = 2, 3, 4, \dots$, meaning that $\varphi_1(i) = C/i^\alpha$ implies $\varphi = O'(1/i^\alpha)$ where $f = O'(g)$ means:

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = A, \quad (6a)$$

where $0 < A < \infty$. This is in fact stronger than what is usually defined as $O(\cdot)$ —see Rankin (1963). If the above limit does not exist, we then require

$$O < \liminf \frac{f(x)}{g(x)} < \limsup \frac{f(x)}{g(x)} < \infty, \quad (6b)$$

which boils down to (6a) when the limit exists.

This explains the fact that, although the author–production system is different from the systems in which every item has only one source, Lotka’s law with $\alpha > 1$ is acceptable from a model-theoretic point of view.

General Theory

As above, let us consider a system in which every paper has exactly j authors ($j = 1, 2, 3, \dots$). Let us define, φ_j being the density function of the distribution F_j where, for all $i \geq 0$

$$F_j(i) = P \text{ (author has } i \text{ or less publications)} \\ \text{each paper has } j \text{ authors} \quad (7)$$

and for φ the same with respect to F :

$$F(i) = P \text{ (author has } i \text{ or less publications)} \quad (8)$$

Let ψ be the function

$$\psi(j) = P \text{ (a paper has } j \text{ authors)} \quad (9)$$

[P in (7) and (8) is in the author probability space; P in (9) is in the paper probability space].

The technique of calculating φ from φ_1 (given function), is to go over φ_2 , then over φ_3 , and so on. Hereby we use convolutions, an idea introduced in Egghe (1993), but here we use a dual framework. We repeat the results on convolutions * that we need here (cf. Chung, 1974, pp. 144–146). For reasons of dealing with calculable convolution formulae we use i as a continuous variable.

Theorem II.1

Let X_1 and X_2 be independent random variables with distribution functions F_1 and F_2 , respectively. Then $X_1 + X_2$ has the distribution function $F_1 * F_2$, where

$$(F_1 * F_2)(x) = \int_{-\infty}^{\infty} F_1(x - y) dF_2(y) \quad (10)$$

is the convolution of the two distribution functions.

When the distribution functions have densities, we have the following theorem.

Theorem II.2

The convolution of two distribution functions with densities g_1 and g_2 is a distribution function with density $g_1 * g_2$.

In a dual way, as in Egghe (1993), we can now state and prove the following results.

Theorem II.3

For every $i \geq 0$ and $j = 2, 3, \dots$,

$$\varphi_j(i) = \underbrace{(\varphi_1 * \dots * \varphi_1)}_{j \text{ times}}(i) \quad (11)$$

Proof: Let $j = 2$ and fix $i \geq 0$. We adapt the following model: the IPP of papers, all with two authors, is considered as a merged system of two IPPs of papers with one author. The first one consists of the papers with only the first author as single author and the second IPP consists of the papers with only the second author as single author. So, author scores in the IPP of papers with two authors are sums of author scores in the two single-author IPPs. Indeed, the score of i papers for an author in systems where every paper has two authors is obtained from a score of y papers as first author and $i - y$ papers as second author; here, $y \geq 0$ arbitrarily. Hence, we are in the case of Theorems II.1 and II.2, yielding (only positive values are possible):

$$\begin{aligned}\varphi_2(i) &= (\varphi_1 * \varphi_1)(i) \\ \varphi_2(i) &= \int_0^i \varphi_1(i - y)\varphi_1(y) dy\end{aligned}\quad (12)$$

Let $j \in \mathbf{N}$, $j \geq 3$ be arbitrary and fix $i \geq 0$. The score of i papers for an author is obtained from a score of y papers when the author is the first, second, ..., or $(j - 1)^{\text{th}}$ author and a score of $i - y$ papers where the author is the j^{th} author. Hence:

$$\begin{aligned}\varphi_j(i) &= (\varphi_{j-1} * \varphi_1)(i) \\ \varphi_j(i) &= \underbrace{(\varphi_1 * \dots * \varphi_1)}_{j \text{ times}}(i),\end{aligned}$$

by the associative property of convolutions. \square

Corollary II.4 For every $i \geq 0$:

$$\varphi(i) = \sum_{j=1}^{\infty} \underbrace{(\varphi_1 * \dots * \varphi_1)}_{j \text{ times}}(i)\psi(j)\quad (13)$$

Note 1: Clearly, it is not so that an IPP in which all papers have two authors is, in practice, a merged system of two IPPs in which all papers have one author. The model developed in the above proof only uses this framework for the purpose of calculating author scores. For this, it is clearly so that scores of authors in the former IPP are sums of scores of authors in the latter IPPs and hence, convolution theory applies.

Note 2: Formula (13) is a dual version of formula (11) in Egghe (1993), where it was used to model fractional author counting when authors have coauthored several papers.

We repeat the main problem of this paper: φ_1 is the "natural" frequency function in the case that every item has only one source (as, e.g., in the journal–article systems—i.e., classical bibliographies) and φ is the "natural" frequency function in the general case that every item can have several sources (as, e.g., in the author–publication systems). Historically, in the first case, Bradford's law is found, hence (for $i = 1, 2, \dots$):

$$\varphi_1(i) = \frac{C}{i^2}\quad (14)$$

being equivalent with Bradford's law (cf. Egghe & Rousseau, 1990) and

$$\varphi(i) = \frac{C}{i^2}\quad (15)$$

is found in the second case (the classical law of Lotka, cf. Lotka, [1926]) (or, more generally with the exponent 2 replaced by α).

This article tries to explain how (15) can follow from (14), at least in an approximate way. This is done in the next section and even for general exponents $\alpha > 1$. An exact "closed circuit" of functions is obtained for the geometric distribution: $\varphi_1(i) = pq^i$ implies $\varphi(i) = ca^i$ exactly, for certain values of c and a . Such an exact closed circuit cannot be obtained for the power functions (1). But in the O' -sense we obtain an approximate closed circuit (and even for the same exponent α , contrary to the case of the geometric distribution), if $\alpha > 1$. This gives new insight into Lotka's law and the geometric distribution and in the validity of general informetric laws in general IPPs.

Stability Study of Lotka's Law

One of the main tools in the sequel are convolutions, as introduced in the previous section. We could have used discrete variables for the functions but then, taking discrete convolutions is very difficult and hard to evaluate. Therefore, we restrict ourselves to functions of a continuous variable. Because with function (1) we will have difficulties with divergent integrals, we will study the function:

$$\varphi_1(i) = \frac{C}{(i + 1)^\alpha}\quad (16)$$

for $i \in [0, \infty[$ (we could use (1) and use $i \in [1, \infty[$ but then we have trouble in defining the convolution integrals).

We can now state and prove our main theorem:

Theorem III.1

Let φ_1 be as in (16). Then the function φ satisfies the stability property

$$\varphi = O'\left(\frac{1}{i^\alpha}\right) = O'(\varphi_1)\quad (17)$$

if $\alpha > 1$. This result is true for any function ψ such that $\psi(1) \neq 0$. Hence, (17) means that:

$$0 < \lim_{i \rightarrow \infty} \frac{\varphi(i)}{\varphi_1(i)} < \infty,$$

meaning that the asymptotic behavior of φ and φ_1 are the same (up to a constant).

Proof: Because the proof is rather intricate and only interesting for mathematical readers, it is given in the Appendix.

So we have that $\varphi_1(i) = C/(1 + i)^\alpha$ for $i \geq 0$ implies $\varphi = O'(1/i^\alpha) = O'(\varphi_1)$ if $\alpha = 2, 3, 4, \dots$. We conjecture that this will be false for $\alpha = 1$; in this case, $\varphi_2(i) = 2C^2 \ln(i + 1)/(i + 2) \neq O'(1/i)$, but an exact proof is not known. For $\alpha = 0$ we have an exact proof of the failure of Theorem III.1. In this case, $\varphi_1(i) = C$ (we limit i to bounded values here) and, for every $j \in \mathbf{N}$:

$$\varphi_j(i) = C^j \frac{i^{j-1}}{(j - 1)!}\quad (18)$$

as is easily seen from (11). Hence, from (13) and supposing (as an example for the failure of Theorem III.1), for $j \in \mathbf{N}$:

$$\Psi(j) = \beta\gamma^j, \quad (19)$$

we have:

$$\varphi(i) = \sum_{j=1}^{\infty} \frac{C^j j^{j-1} \beta\gamma^j}{(j-1)!}$$

$$\varphi(i) = \beta\gamma C \sum_{j=0}^{\infty} \frac{(Ci\gamma)^j}{j!}$$

$$\varphi(i) = \beta\gamma C e^{Ci\gamma}$$

which is an exponential function if $\gamma \neq 0$. Hence $\varphi \neq O'$ (constant) and even worse: $\varphi \neq O'(1/i^\alpha)$ for any α . Note also that no φ_j ($j = 2, 3, \dots$) is O' (constant) nor O' (exponential). This negative result will be better understood when studying functions φ_1 of the exponential type. This will be done in the next section. There we will be able to prove that the class of exponential functions is stable for the transform: φ_1 into φ .

Study of $\varphi_1(i) = pq^i$

The function

$$\varphi_1(i) = pq^i \quad (20)$$

now represents the geometrical distribution, if we use discrete variables $i = 0, 1, 2, \dots$ and is often encountered in library circulation data. In this section we could as well work with discrete convolutions but, to continue the continuous model in this article, we will work with (20) for $i \geq 0$. We have the following exact stability result for the geometric distribution.

Theorem IV.1

Let

$$\varphi_1(i) = pq^i \quad (20)$$

for $i \geq 0$. Then:

$$\varphi(i) = ca^i \quad (21)$$

for $i \geq 0$, if ψ is also of the form:

$$\psi(j) = \beta\gamma^j \quad (22)$$

Proof: The proof is easy. By (12):

$$\varphi_2(i) = p^2 q^i i$$

More generally, for every $j = 2, 3, \dots$

$$\varphi_j(i) = \frac{p^j q^i i^{j-1}}{(j-1)!}$$

By (13) and (22),

$$\begin{aligned} \varphi(i) &= \sum_{j=1}^{\infty} \frac{p^j q^i i^{j-1} \beta\gamma^j}{(j-1)!} \\ &= p\beta\gamma q^i \sum_{j=0}^{\infty} \frac{(pi\gamma)^j}{j!} \\ &= p\beta\gamma q^i e^{pi\gamma} \\ &= ca^i \end{aligned}$$

where $c = p\beta\gamma$ and $a = qe^{p\gamma}$. \square

Note from III.1 that the constant function also belongs to this class (and not to any other class). This explains the note at the end of the previous section.

Problem: Determine other function classes such that φ_1 and φ are within the same class, possibly in an approximate way (O , O' or stronger).

Problem: Prove Theorem III.1 for all values $\alpha \geq 1$, that is, also noninteger α . We conjecture this to be true.

Conclusions

In this article we studied the fact that papers can be written by several authors, a source-item relation different than what is usually found (e.g., papers [= items] are published in one journal [= source] only). Supposing certain frequency functions for the latter source-item relation, we wonder if they are stable in the former source-item relationship. We prove that this is the case for Lotka-type functions of order $O'(1/i^\alpha)$ (which is stronger than the classical O -relation and is, up to a positive constant, the \sim -relation) if $\alpha = 2, 3, \dots$ and disprove this if $\alpha = 0$. We show that the geometric distribution is perfectly stable in this sense.

The author-publication relationship, as studied here, is rather unique among the source-item relations. As pointed out, papers can have several authors but articles have only one journal in which they are published. In the source-item relationship of books and their borrowings it is also clear that a borrowing of a book is strictly related to this book only!

Also in linguistics, the use (= item) of a word in a text is strictly linked with this word (= source) only, in an obvious way. Also in demography and econometry, the source-item relation is one-to-one: an inhabitant lives in one city or village (few exceptions exist however) and a dollar that is earned by an employee belongs to this employee only!

A source-item structure as in the case of author-publication is, however, also encountered in the case that sources are articles and items are references. Indeed, here an item can have different sources too! The same is true for the relation articles-citations (to these articles). The above theory is, therefore, also applicable to these situations.

One could go even further and see if other, less-known cases of multiple sources per item exist. One could argue that patents can have several countries of application (= sources). One could also consider a borrowing in a library as the "package" of several books that one requests to check out (in one time). Here a borrowing (= item) refers to several books (= sources). Informetric studies of such source-item relations envisage the use (or value) of books, when related to other books on similar topics.

It is interesting to formulate other new multiple sources-item relationships and study their value for

informetrics and beyond. In all these cases, this article applies, showing certain stability of informetric laws.

Acknowledgments

The author is grateful to R. Rousseau for interesting discussions on the problems studied in this article.

Appendix

Before proving Theorem III.1, we need a lemma:

Lemma: For all i and y and $\alpha \in \mathbb{N}$:

$$\frac{1}{y^\alpha(i-y+2)^\alpha} = \sum_{m=1}^{\alpha} \left(\frac{C_{m,\alpha}}{(i+2)^{2\alpha-m}y^m} + \frac{D_{m,\alpha}}{(i+2)^{2\alpha-m}(i-y+2)^m} \right) \quad (\text{A1})$$

Proof: The proof is easy but technical. We use complete induction on α . Let $a = 1$. Then:

$$\frac{1}{y(i-y+2)} = \frac{1}{(i+2)y} + \frac{1}{(i+2)(i-y+2)} \quad (\text{A2})$$

proving (A1) for $a = 1$. Let us now suppose (A1) valid for α and we take $\alpha + 1$; now:

$$\frac{1}{y^{\alpha+1}(i-y+2)^{\alpha+1}} = \sum_{m=1}^{\alpha} \left(\frac{C_{m,\alpha}}{(i+2)^{2\alpha-m}y^m} + \frac{D_{m,\alpha}}{(i+2)^{2\alpha-m}(i-y+2)^m} \right) \cdot \left(\frac{1}{(i+2)y} + \frac{1}{(i+2)(i-y+2)} \right),$$

by (A1) and (A2)

$$= \sum_{m=1}^{\alpha} \frac{1}{(i+2)^{2\alpha-m+1}} \left[\frac{C_{m,\alpha}}{y^{m+1}} + \frac{C_{m,\alpha}}{y^m(i-y+2)} + \frac{D_{m,\alpha}}{(i-y+2)^m y} + \frac{D_{m,\alpha}}{(i-y+2)^{m+1}} \right] \quad (\text{A3})$$

We now use that:

$$\frac{1}{y^m(i-y+2)} = \sum_{k=1}^m \frac{1}{y^k(i+2)^{m-k+1}} + \frac{1}{(i+2)^m(i-y+2)} \quad (\text{A4})$$

and

$$\frac{1}{(i-y+2)^m y} = \sum_{k=1}^m \frac{1}{(i-y+2)^k(i+2)^{m-k+1}} + \frac{1}{(i+2)^m y} \quad (\text{A5})$$

(A4) and (A5) in (A3) yields:

$$\frac{1}{y^{\alpha+1}(i-y+2)^{\alpha+1}} = \sum_{n=1}^{\alpha} \frac{1}{(i+2)^{2\alpha-m+1}} \left[\frac{C_{m,\alpha}}{y^{m+1}} + \frac{D_{m,\alpha}}{(i-y+2)^{m+1}} + C_{m,\alpha} \left(\sum_{k=1}^m \frac{1}{y^k(i+2)^{m-k+1}} + \frac{1}{(i+2)^m(i-y+2)} \right) + D_{m,\alpha} \left(\sum_{k=1}^m \frac{1}{(i-y+2)^k(i+2)^{m-k+1}} + \frac{1}{(i+2)^m y} \right) \right] \quad (\text{A6})$$

When working out (A6), term by term and when regrouping according to the factors $1/y^m$ and $1/(i-y+2)^m$, $m = 1, \dots, \alpha + 1$, we find:

$$\frac{1}{y^{\alpha+1}(i-y+2)^{\alpha+1}} = \frac{1}{y} \frac{\sum_{m=1}^{\alpha} (C_{m,\alpha} + D_{m,\alpha})}{(i+2)^{2\alpha+1}} + \frac{1}{y^2} \frac{\sum_{m=1}^{\alpha} C_{m,\alpha}}{(i+2)^{2\alpha}} + \dots + \frac{1}{y^\alpha} \frac{C_{\alpha-1,\alpha} + C_{\alpha,\alpha}}{(i+2)^{\alpha+2}} + \frac{1}{y^{\alpha+1}} \frac{C_{\alpha,\alpha}}{(i+2)^{\alpha+1}} + \frac{1}{i-y+2} \frac{\sum_{m=1}^{\alpha} (C_{m,\alpha} + D_{m,\alpha})}{(i+2)^{2\alpha+1}} + \frac{1}{(i-y+2)^2} \frac{\sum_{m=1}^{\alpha} D_{m,\alpha}}{(i+2)^{2\alpha}} + \dots + \frac{1}{(i-y+2)^\alpha} \frac{D_{\alpha-1,\alpha} + D_{\alpha,\alpha}}{(i+2)^{\alpha+2}} + \frac{1}{(i-y+2)^{\alpha+1}} \frac{D_{\alpha,\alpha}}{(i+2)^{\alpha+1}}$$

which is of the form:

$$\sum_{m=1}^{\alpha+1} \left(\frac{C_{m,\alpha+1}}{(i+2)^{2(\alpha+1)-m} y^m} + \frac{D_{m,\alpha+1}}{(i+2)^{2(\alpha+1)-m} (i-y+2)^m} \right)$$

where $C_{m,\alpha+1}$ and $D_{m,\alpha+1}$ are constants not dependent on i and y . This proves the lemma. \square

Proof of Theorem III.1: We use induction on j in φ_j . By (11):

$$\varphi_2(i) = C^2 \int_0^i \frac{dy}{(y+1)^\alpha(i-y+1)^\alpha}$$

$$\varphi_2(i) = C^2 \int_1^{i+1} \frac{dy}{y^\alpha(i-y+2)^\alpha} \quad (\text{A7})$$

for $i \in [0, \infty[$. We invoke the above lemma: for all $\alpha \in \mathbf{N}$, we have:

$$\frac{1}{y^\alpha(i-y+2)^\alpha} = \sum_{m=1}^{\alpha} \left(\frac{C_{m,\alpha}}{(i+2)^{2\alpha-m}y^m} + \frac{D_{m,\alpha}}{(i+2)^{2\alpha-m}(i-y+2)^m} \right), \quad (\text{A8})$$

where the $C_{m,\alpha}$ and $D_{m,\alpha}$ are positive constants, independent of i and y . Hence, (A7) and (A8) imply, after some calculation,

$$\begin{aligned} \varphi_2(i) = & \frac{C^2}{(i+2)^\alpha} \left[\frac{A_1 \ln(i+1)}{(i+2)^{\alpha-1}} \right. \\ & + \frac{A_2}{(i+2)^{\alpha-2}} \cdot \frac{i}{i+1} \\ & + \dots + \frac{A_{\alpha-1}}{(\alpha-2)(i+2)} \left(1 - \frac{1}{(i+1)^{\alpha-2}} \right) \\ & \left. + \frac{A_\alpha}{\alpha-1} \left(1 - \frac{1}{(1+i)^{\alpha-1}} \right) \right] \quad (\text{A9}) \end{aligned}$$

where $A_m = C_{m,\alpha} + D_{m,\alpha}$ (we deleted the α in A_m , for simplicity), for $m = 1, \dots, \alpha$. Hence,

$$\varphi_2 = O'\left(\frac{1}{i^\alpha}\right) = O(\varphi_1)$$

for all $\alpha > 1$ and $\alpha \in \mathbf{N}$; therefore,

$$\lim_{i \rightarrow \infty} i^\alpha \varphi_2(i) = \frac{A_\alpha C^2}{\alpha-1} \in]0, \infty[.$$

Now (A9) implies:

$$0 < \varphi_2(i) < C \frac{C}{(1+i)^\alpha} \cdot \sum_{m=1}^{\alpha} A_m =: N_2 \varphi_1(i). \quad (\text{A10})$$

where $0 < N_2 < \infty$, for all $i \geq 0$. Inductively, suppose that for $k = 2, 3, \dots, j$ and all $i \geq 0$:

$$0 \leq \varphi_k(i) < N_k \varphi_1(i) \quad (\text{A11})$$

for a certain constant $N_k \in]0, \infty[$. Then:

$$\varphi_{j+1}(i) = \int_0^i \varphi_j(y) \varphi_1(i-y) dy$$

Hence, by (A11) and then (A10):

$$\begin{aligned} 0 \leq \varphi_{j+1}(i) & < N_j \int_0^i \varphi_1(y) \varphi_1(i-y) dy \\ & = N_j \varphi_2(i) < N_j N_2 \varphi_1(i) \\ & =: N_{j+1} \varphi_1(i) \end{aligned}$$

where $0 < N_{j+1} < \infty$. Hence, (A11) is proved for all $j = 2, 3, \dots$ and all $i \geq 0$. (A11) implies:

$$0 \leq \lim_{i \rightarrow \infty} \frac{\varphi_j(i)}{\frac{1}{i^\alpha}} =: B_j \leq N_j \lim_{i \rightarrow \infty} \frac{\varphi_1(i)}{\frac{1}{i^\alpha}} = N_j C < \infty \quad (\text{A12})$$

for all $j = 2, 3, \dots$ and note that for $j = 1$:

$$0 < \lim_{i \rightarrow \infty} \frac{\varphi_1(i)}{\frac{1}{i^\alpha}} = C < \infty \quad (\text{A13})$$

Let M be the highest possible j . Then, by (13): for all $i \geq 0$

$$\varphi(i) = \sum_{j=1}^M \varphi_j(i) \psi(j).$$

Hence:

$$\lim_{i \rightarrow \infty} \frac{\varphi(i)}{\frac{1}{i^\alpha}} = \sum_{j=1}^M \lim_{i \rightarrow \infty} \frac{\varphi_j(i)}{\frac{1}{i^\alpha}} \psi(j) = A$$

where:

$$A = C\psi(1) + \sum_{j=2}^M B_j \psi(j) \in]0, \infty[,$$

by (A12) and (A13) and since $\psi(j) \in [0, \infty[$ for all $j = 1, 2, \dots$ and $\psi(1) \neq 0$. So:

$$\varphi = O'\left(\frac{1}{i^\alpha}\right) = O'(\varphi_1) \quad \square$$

Note: Remark that (A12) does not imply that $\varphi_j = O'(1/i^\alpha)$, for $j = 2, 3, \dots$! But, $\varphi_j = O(1/i^\alpha)$.

References

- Bradford, S. C. Sources of information on specific subjects. *Engineering*, 137, 85–86 (1934). Reprinted in: *Collection Management*, 1, 95–103 (1976–1977). Also reprinted in: *Journal of Information Science*, 10, 148 (facsimile of the first page) and 176–180 (1985).
- Chung, K. L. (1974). *A Course in probability theory*. New York: Academic Press.
- Egghe, L. (1989). *The duality of informetric systems with applications to the empirical laws*. PhD thesis, City University, London, UK.
- Egghe, L. (1990) The duality of informetric systems with applications to the empirical laws. *Journal of Information Science*, 16, 17–27.
- Egghe, L. (1993). Consequences of Lotka's law in the case of fractional counting of authorship and of first author counts. *Mathematical and Computer Modelling*, 18(9), 63–77.
- Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics. Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier. (1990).
- Lotka, A. J. (1926). The frequency distribution of scientific productivity, *Journal of the Washington Academy of Sciences*, 16, 317–323.
- Rankin, R. A. (1963). *An introduction to mathematical analysis*. Oxford: Pergamon.