Zipfian and Lotkaian continuous concentration theory

Peer-reviewed author version

EGGHE, Leo (2005) Zipfian and Lotkaian continuous concentration theory. In: JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 56(9). p. 935-945.

DOI: 10.1002/asi.20186
Handle: http://hdl.handle.net/1942/742

Zipfian and Lotkaian continuous concentration theory

by

L. Egghe. Limburgs Universitair Centrum (LUC), Universitaire Campus, B-3590 Diepenbeek, Belgium[1]
and
Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Wilrijk, Belgium

ABSTRACT

This paper studies concentration (i.e. inequality) aspects of the functions of Zipf and of Lotka. Since both functions are power laws (i.e. they are – mathematically the same) it suffices to develop one concentration theory for power laws and apply it twice for the different interpretations of the laws of Zipf and Lotka.

After a brief repetition of the functional relationships between Zipf's law and Lotka's law, we prove that Price's law of concentration is equivalent with Zipf's law. The major part of the paper is devoted to the development of continuous concentration theory, based on Lorenz curves. We calculate the Lorenz curve for power functions and, based on this, calculate some important concentration measures such as the ones of Gini, Theil and the variation coefficient.

We also show, using Lorenz curves, that the concentration of a power law increases with its exponent and we interpret this result in terms of the functions of Zipf and Lotka.

[1] Permanent address

Key words and phrases: Zipf, Lotka, Price, Lorenz, concentration theory, power law

I. Introduction

The historical law of Lotka (Lotka (1926)) is the basis of modern informetrics and is expressed as follows: the number of authors with n (n=1,2,3,...) publications is proportional to $\frac{1}{n^{\alpha}}$, where $\alpha > 0$. In other words, there is a constant C>0 such that

$$f(n) = \frac{C}{n^{\alpha}} \tag{1}$$

where f(n) denotes the number of authors with n publications. More generally, f(n) can denote the number of sources (authors, journals, word types,...) with n items (publications, articles, word occurrences, ... respectively) and we will, henceforth, use this dual source/item terminology – see also Egghe (1989, 1990), Egghe and Rousseau (1990a).

A totally different informetric formulation, originating from linguistics (in terms of word types and word occurrences) is given by the law of Zipf: if we rank the sources according to their number of items (starting with the source with the highest number of items, hence giving to this source the rank r=1) then the number of items in the source on rank r (r=1,2,3,...) is proportional to $\frac{1}{r^{\beta}}$, where $\beta > 0$. In other words, denoting by g(r) this number of items in the source on rank r, there exists a constant D>0 such that

$$g(r) = \frac{D}{r^{\beta}} \ . \tag{2}$$

Although their informetric definitions are different (they are dual in the sense that, in f and g, the roles of sources and items are interchanged), the functions (1) and (2) are mathematically the same, namely decreasing power laws. A power law is the most occurring regularity in informetrics and far beyond (e.g. also found in economics, sociology incl. the description of social networks, see e.g. Egghe and Rousseau (2003) and references therein) and has the characterising property (see Roberts (1979)) that, if the argument (say x) is multiplied by a constant, say k, we obtain the same power law with the same exponent: for $f(x) = \frac{C}{x^\alpha}$ we have $f(kx) : \frac{1}{(kx)^\alpha} = \frac{1}{k^\alpha x^\alpha} : \frac{1}{x^\alpha} : f(x)$ (– denotes “is proportional to”). This self-similar property explains its widespread occurrence in real life examples and also its use in the description of power type informetrics in terms of self-similar fractals (see Feder (1988)), but this aspect will not be further discussed here.

In the sequel we will use both functions f and g in the continuous setting, since we will then be able to execute concrete calculations of the relation between f and g and of the continuous concentration theory: we will be able to evaluate integrals which is not possible with discrete sums. Therefore we reformulate the functions of Lotka and Zipf as follows (see also Egghe and Rousseau (1990a), Egghe (1989, 1990)), respectively:

$$f(j) = \frac{C}{j^\alpha} \tag{3}$$

j0[1,Δ_m] and

$$g(r) = \frac{D}{(1+r)^\beta} \tag{4}$$

r0[0,T].

<u>Intuition</u>: f(j) denotes the density of the sources with item per source density j0[1,Δ_m] (Δ_m denotes the maximal item per source density) and g(r) denotes the density of the items per source in the source density r0[0,T] (T denotes the total number of sources).

In the above mentioned references one finds the following general (i.e. independent of the specific form of f and g e.g. as in (3) and (4)) relations between f and g:

$$r = g^{-1}(j) = \int_j^{\rho_m} f(j')dj' \tag{5}$$

hence

$$f(j) = -\frac{1}{g'(g^{-1}(j))} \tag{6}$$

$j0[1,\Delta_m]$. Relations (5) or (6) can also be used as the defining relation for g, respectively f, when the other function is given, hereby presenting a mathematically formal approach to the classical informetric functions. Formulae (5) and (6) are the basis of the following well-known theorem (see again Egghe (1985, 1989, 1990), Egghe and Rousseau (1990a), Rousseau (1990)).

Theorem I.1: The following assertions are equivalent for $\alpha > 0, \quad \alpha \neq 1$:

(i)

$$f(j) = \frac{C}{j^{\alpha}} \tag{7}$$

$j0[1,\Delta_m]$, C>0, i.e. Lotka's law with exponent ∀

(ii)

$$g(r) = \frac{D}{(1+Er)^{\beta}} \tag{8}$$

r0[0,T], where

$$D = \rho_m \tag{9}$$

$$E = \frac{\alpha - 1}{C\rho_m^{1-\alpha}} \tag{10}$$

$$\beta = \frac{1}{\alpha - 1} \tag{11}$$

Function (8) is called the law of Mandelbrot with general exponent ∃ (cf. Mandelbrot (1977))

Proof:

(i)Ψ(ii) Formula (5) yields

$$r = g^{-1}(j) = \int_j^{\rho_m} \frac{C}{j'^{\alpha}} dj'$$

$$= \frac{C}{1-\alpha}\left(\rho_m^{1-\alpha} - j^{1-\alpha}\right)$$

hence, since j=g(r):

$$g(r) = \frac{\rho_m}{\left(1 + r\frac{\alpha - 1}{C\rho_m^{1-\alpha}}\right)^{\frac{1}{\alpha - 1}}}$$

from which (8)-(11) follow.

(ii)Ψ(i) Formula (8) yields

$$g'(r) = \frac{-\beta DE}{(1+Er)^{\beta+1}}$$

and also (j=g(r), hence $r=g^{-1}(j)$)

$$j = \frac{D}{(1+Eg^{-1}(j))^{\beta}}$$

hence

$$g^{-1}(j) = \frac{\left(\left(\frac{D}{j}\right)^{\frac{1}{\beta}} - 1\right)}{E}$$

from which it follows that, by (6)

$$f(j) = -\frac{1}{g'(g^{-1}(j))}$$

$$= \frac{D^{1+\frac{1}{\beta}}}{\beta DE} \cdot \frac{1}{j^{1+\frac{1}{\beta}}}$$

proving (7). ~

The following easy corollary on the characterisation of Zipf's law has been formulated explicitely in Rousseau (1990):

<u>Corollary I.2 (Rousseau)</u>: The following assertions are equivalent for $\alpha > 1$:

(i)

$$f(j) = \frac{C}{j^{\alpha}}$$

$j \in [1,\Delta_m]$ with

$$\rho_m = \left(\frac{C}{\alpha - 1}\right)^{\frac{1}{\alpha - 1}} \tag{12}$$

, i.e. the general Lotka law but with parameters restricted to (12)

(ii)

$$g(r) = \frac{D}{(1+r)^{\beta}} \tag{13}$$

r0[0,T], i.e. Mandelbrot's law with E=1 (and further (9) and (11) remain valid). Formula (13) is the same as (4), hence the law of Zipf.

Proof: We have (13) (i.e. Zipf's law) if and only if E=1 from which it follows from (10) that (12) and $\alpha > 1$ are valid. For the equivalency, the previous theorem applies. ~

Note I.3: Corollary I.2 shows that Zipfian informetrics is part of Lotkaian informetrics (the complete Lotkaian informetrics being covered by the law of Mandelbrot as proved in Theorem I.1). Which part? There is the restriction (12). From this it follows that

$$\alpha = 1 + C\rho_m^{1-\alpha} > 1 \tag{14}$$

The limitation to $\alpha > 1$ (which is not needed in Theorem I.1) is not a restriction since it is equivalent with $\beta > 0$, which is obvious (since g(r) must decrease). So Zipfian informetrics covers the complete Lotkaian informetrics for $\alpha > 1$ but limited to the constants as expressed by (12). The latter limitation does not play any role in concentration theory since there one works with normalized values (distributions – see further). Also, (12) always produces a $\Delta_m > 1$ which is required. Indeed: (5) implies

$$T = \int_1^{\rho_m} f(j')dj'$$

$$T = \frac{C}{1-\alpha}\left(\rho_m^{1-\alpha} - 1\right)$$

hence, by (10)

$$E = \frac{\alpha - 1}{C\rho_m^{1-\alpha}} = \frac{1}{\frac{C}{\alpha-1} - T} = 1$$

So

$$\frac{C}{\alpha - 1} = T + 1 > 1$$

since $T \geq 0$. Now (12) implies Δ_m>1.

We can also prove the following important result: If Zipf's law applies then we have:

$$T \to \infty \text{ if and only if } \rho_m \to \infty \ .$$

Proof: Formulae (3) and (5) (for r=T, hence j=1) yield

$$T = \frac{C}{\alpha - 1}\left(1 - \rho_m^{1-\alpha}\right)$$

applying (12) on this yields

$$T = \rho_m^{\alpha-1}\left(1 - \frac{1}{\rho_m^{\alpha-1}}\right)$$

and by the above: $\alpha > 1$. Hence $T \to \infty$ iff $\rho_m \to \infty$. So also in this sense, functions (3) and (4) are the same type of power law (not the same exponent – see (11)): they both have a bounded domain or both have an unbounded domain: [1,Δ_m] respectively [1,T+1] (both starting in 1): indeed, remark that (13) (or (4)) can be reformulated as

$$g(r) = \frac{D}{r^{\beta}} \tag{15}$$

r0[1,T+1], resembling more the "classical" Zipf function (2). It is this form that we will use henceforth since it is the same (up to notation) as (3). This will enable us to develop concentration theory for functions of the form (3) <u>and</u> (15) at the same time (although their informetric meaning is different).

In the next section we will show that Zipf's law is equivalent with the validity of Price's law of concentration (see section II for a formulation).

In the third section we will develop Lorenz concentration theory for power functions which we will then apply to the functions f and g as in (3) and (15): the continuous Lorenz curve in this setting is defined and calculated for power functions. We show that, the higher the exponent in such a power function, the higher the Lorenz curve, hence the more concentrated (unequal) the situation. We prove explicite formulae for the concentration measures of Gini, Theil and the variation coefficient, in the case of power functions. We also describe the informetric relation between concentration theory of Lotka's law and the one of Zipf's law (being mathematically the same but having different informetric interpretations). As a consequence, some (known) results on the general 80/20 rule (also defined in section III) for power laws will be refound.

II. A Characterisation of Price's law of concentration

Price'law (Price (1976)) can be formulated as follows: let there be T sources. Starting with the most productive sources and for any 20]0,1[we have that the T^2 top sources produce a fraction 2 of all the items. A special value is $\theta = \frac{1}{2}$, hence the $\sqrt{T}$ top sources produce 50% of all the items: if T=100, then the top 10 sources yield this 50% of all items. It is clear that Price's law expresses the skewness of the production process (in other words the degree of production inequality between the sources). Skewness is also expressed by the laws of Zipf and Lotka. We can now wonder whether or not these inequalities are related. This section will show that, although their formulations are different, Price's law is equivalent with the law of Zipf. The following proposition is a continuous extension of a result in Egghe and Rousseau (1986). For reasons explained in Note I.3 we will take T+1 as the highest rank (because of the continuous setting).

<u>Proposition II.1</u>: Denote by G(r) the cumulative number of items produced by the sources in the interval [0,r]. Then the following assertions are equivalent:

(i)

$$G(r) = B \log r$$

r0[1,T+1], where B is a constant

(ii) The law of Price is valid: for every 20]0,1[, the top (T+1)^{2} sources produce a fraction 2 of the items.

<u>Proof</u>:

<u>(i)Ψ(ii)</u> Since we have the top (T+1)^{2} (20]0,1[) sources, their cumulative production is given by (definition of G and since (T+1)^{2}0[1,T+1]):

$$G\left((T+1)^{\theta}\right) = B\log\left((T+1)^{\theta}\right)$$

$$G\left((T+1)^{\theta}\right) = \theta G\left(T+1\right)$$

But G(T+1) denotes the total number of items, hence the top (T+1)^{2} sources produce a fraction 2 of the items.

<u>(ii)Ψ(i)</u> Let r0[1,T+1] be arbitrary. Hence there exists 20[0,1] such that r=(T+1)^{2}. By (ii) and the definition of G we have that

$$G\left((T+1)^{\theta}\right) = \theta G\left(T+1\right)$$

for 20]0,1[. This is also true for 2=1 and for 2=0 (since we work with source densities). But r=(T+1)^{2} implies logr=2log(T+1). Hence

$$G(r) = \frac{\log r}{\log(T+1)} G(T+1)$$

$$G(r) = B \log r$$

where $B = \dfrac{G(T+1)}{\log(T+1)}$, a constant. ~

<u>Corollary II.2</u>: Price's law is equivalent with Zipf's law for ∃=1:

$$g(r) = \frac{D}{1+r}$$

r0[0,T].

<u>Proof</u>: This follows from the fact that g and G relate as

$$G(r) = \int_0^r g(r')dr'$$

(hence G^N=g). From this it follows that G(r)=Blogr, r0[1,T+1] is equivalent to

$$g(r) = \frac{B}{r}$$

r0[1,T+1], hence to

$$g(r) = \frac{B}{1+r}$$

r0[0,T], which is Zipf's law. The result then follows from Proposition II.1. ~

<u>Note II.3</u>: Combining corollaries I.2 and II.2 we see that Price's law is valid if Lotka's law is valid but restricted to (12) and ∀=2. For ∀=2 this yields Δ_m=C, a result that was found in

Allison, Price, Griffith, Moravcsik and Stewart (1976) after a long approximative calculation (and limited to Price's law for $\theta = \frac{1}{2}$). This paper apparently (see editors' note) grew out of lengthy and frequently heated correspondence between these authors on the validity of Price's square root law (i.e. Price's law for $\theta = \frac{1}{2}$). We hereby show that a long debate on this issue is not necessary and that Lotka's law (any $\forall$) together with (12) (i.e. restricted to the validity of Zipf's law) shows Price's law in general, for any 2, containing the approximate result in Allison, Price, Griffith, Moravcsik and Stewart (1976).

We refer the reader to Glänzel and Schubert (1985) for a discrete characterization of Price's (square root, i.e. $\theta = \frac{1}{2}$) law. Further discrete calculations on Lotka's law in the connection of Price's law can be found in Egghe (1987). Practical investigations on the validity of Price's law can be found in Berg and Wagner-Döbler (1996), Nicholls (1988) and Gupta, Sharma and Kumar (1998).

<u>Note II.4</u>: Price's law of concentration can be considered as a geometric way of expressing concentration (see also Egghe and Rousseau (1990a)). An arithmetic way of expressing concentration goes as follows: A fraction x0]0,1[of the top sources produce a fraction 2 of the items. This is a generalisation of the well-known 80/20-rule i.e. where 20% of the top sources produce 80% of the items (hence x=0.2 and 2=0.8). We will determine the relationship between 2 and x in case of the law of Lotka and of Zipf. This will – however – be a straightforward consequence of the more general study of Lorenz concentration theory which we will develop now.

III. Lorenz concentration theory

III.1 Discrete case

In order to better understand the continuous case, which is the main topic of this paper, we will briefly repeat the well-known discrete case (see e.g. Egghe and Rousseau (1990b)).

Discrete Lorenz concentration theory is a model to describe the inequality (concentration) of a vector $X=(x_1, x_2, ..., x_N)$. Here we assume all $x_i \geq 0$ and that the vector is decreasing. Examples are given by the law of Zipf or Lotka but where the arguments (n in (1), r in (2)) are the natural numbers 1,2,3,... . The Lorenz curve of X, denoted by L(X) is constructed by linearly connecting the points (0,0) and

$$\left(\frac{i}{N}, \sum_{j=1}^{i} a_j\right) \tag{16}$$

i=1,…,N, where

$$a_j = \frac{x_j}{\sum_{i=1}^{N} x_i} \tag{17}$$

Note that the last point (for i=N) is (1,1). Since X decreases we have that L(X) is a concave polygonal curve. Its form is depicted in Fig. 1

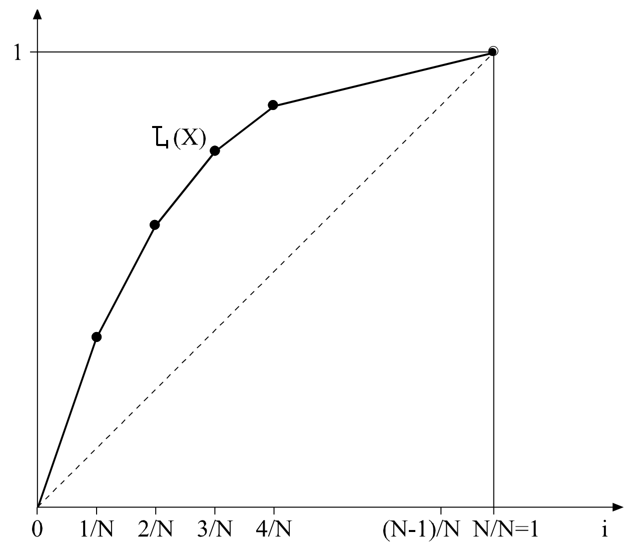

Fig. 1 General form of a discrete Lorenz curve.

The power of the Lorenz curve L(X) lies in the fact that the higher this curve, the more concentrated (unequal) the vector X is (see Egghe and Rousseau (1990b, 1991)). It is then clear that any function C on such vectors $X=(x_1, x_2, ..., x_N)$ and $XN=(xN_1, xN_2, ..., xN_N)$ such that L(X)<L(XN) implies C(X)<C(XN) can be considered as a good concentration measure. Examples are: the Gini index (Gini (1909))

$$G(X) = 2\left\{\text{area under } L_X\right\} - 1 \tag{18}$$

Theil's measure (Theil (1967))

$$\mathrm{Th}(X) = \ln N + \sum_{i=1}^{N} a_i \log a_i \tag{19}$$

(log=ln) and the coefficient of variation V, where

$$V^2(X) = \frac{\sigma^2}{\mu^2} \tag{20}$$

$$V^2(X) = N\sum_{i=1}^{N} a_i^2 - 1 \tag{21}$$

and where Φ^2 and : are the variance and the average of X.

It is clear that, taking for X the discrete values of the Lotka or Zipf function, we can never evaluate these formulae since it is not possible to evaluate (in terms of simple functions) discrete sums. Therefore a continuous theory is needed and exists in econometrics (see e.g. Gastwirth (1971, 1972) and Atkinson (1970)). Continuous Lorenz curves for power functions (such as the ones of Lotka and Zipf) will be constructed here and concentration measures are (analytically) calculated, also showing here the crucial role of the exponents $\forall$ and $\exists$.

III.2 Continuous case

We do not suppose yet the power type function as in (3) or (4) for which we want to establish Lorenz concentration theory but we note that both (3) and (4) are positive decreasing continuous functions h=h(x), where x belongs to an interval of the form $[1,x_m]$, where $x_m>1$. For the moment we will use this general setting. The continuous extension of the Lorenz curve for h(x), $x \in [1,x_m]$ is the curve

$$\left(\frac{x-1}{x_m-1}, \frac{\int_1^x h(x')dx'}{\int_1^{x_m} h(x')dx'} \right) \tag{22}$$

In other words, putting $y=\frac{x-1}{x_m-1} \in [0,1]$, hence $x=y(x_m-1)+1$, the Lorenz curve of the function h is the function L(h), where

$$L(h)(y) = \frac{\int_1^{y(x_m-1)+1} h(x')dx'}{\int_1^{x_m} h(x')dx'} \tag{23}$$

This approach for defining the continuous Lorenz curve is – although equivalent to the definitions in Gastwirth (1971, 1972) and Atkinson (1970) – more direct (more explicite – see (23)) and allows for finite x_m. It also resembles more the discrete case.

It follows that

$$L(h)'(y) = \frac{x_m-1}{\int_1^{x_m} h(x')dx'} h(y(x_m-1)+1) \tag{24}$$

and that

$$L(h)''(y) = \frac{(x_m-1)^2}{\int_1^{x_m} h(x')dx'} h'(y(x_m-1)+1) \tag{25}$$

hence L(h) is a concavely increasing function from (0,0) to (1,1) (since h>0, hN<0). Its general form is depicted in Fig. 2. Since L(h)N is continuous we have that L(h) is a C^1 (i.e. a smooth) function (see e.g. Apostol (1957)).

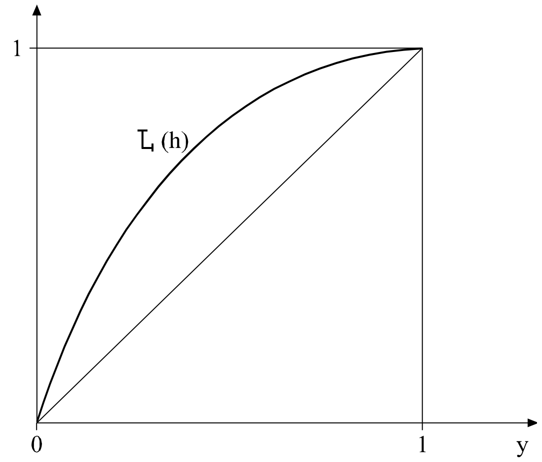

Fig. 2 General form of a continuous Lorenz curve

As in the discrete case such a Lorenz curve can be used to measure the inequality in the (continuous) set of values h(x), x0[1,x_m]. As in the discrete case, we have that a function C on such functions h and hN such that L(h)<L(hN) implies that C(h)<C(hN) is a good measure of concentration. In Egghe (2002a) we proved that the following measures are good concentration measures being the continuous extensions of the measures V^2, Th and G introduced in the discrete case (h is any function as described above):

$$V^2(h) = \int_0^1 [L(h)']^2(y)dy - 1 \tag{26}$$

$$Th(h) = \int_0^1 L(h)'(y)\log(L(h)'(y))dy \tag{27}$$

and, of course,

$$G(h) = 2\int_0^1 L(h)(y)dy - 1 \tag{28}$$

being twice the area under the Lorenz curve of h, minus 1. In fact (26) and (27) follow from a result proved in Egghe (2002a), (based on a general result of Hardy, Littlewood and Pólya (1928)) which says that any measure of the type

$$C(h) = \int_0^1 \varphi(L(h)'(y))dy\,, \tag{29}$$

where ν is a continuous convex function, satisfies the implication L(h)<L(hN)ΨC(h)<C(hN) and hence is a good measure of concentration in the continuous setting.

We will now examine L(h) for h=f and h=g (formulae (3) and (4)) and prove a basic theorem.

III.3 Lorenz curves for power functions

We will now use the function

$$h(x) = \frac{F}{x^{\gamma}} \tag{30}$$

x0[1,x_m], (>0, F>0, representing both the functions f (x_m=Δ_m, (=∀, F=C) and g (x_m=T+1, (=∃, F=D) – see (3) and (4). From (23) we have now

$$L(h)(y) = \frac{(y(x_m - 1) + 1)^{1-\gamma} - 1}{x_m^{1-\gamma} - 1} \tag{31}$$

$(\gamma \neq 1)$ and

$$L(h)(y) = \frac{\log\left[y(x_m - 1) + 1\right]}{\log x_m} \tag{32}$$

((=1), for y0[0,1].

For x_m=4 we can only define the Lorenz curve of h for (<1: by (31):

$$\lim_{x_m \to \infty} L(h)(y) = \lim_{x_m \to \infty} \frac{\left[y\frac{x_m - 1}{x_m} + \frac{1}{x_m} \right]^{1-\gamma} - \frac{1}{x_m^{1-\gamma}}}{1 - \frac{1}{x_m^{1-\gamma}}} = y^{1-\gamma} .$$

Hence we put

$$L(h)(y) = y^{1-\gamma} \tag{33}$$

for h as in (30) with x_m=4, if (<1.

Note that (33), interpreted for f is not applicable since (33) implies ∀<1 while the only admissible ∀-values if Δ_m=4 are the values ∀>2 (see the Lotka existence theorem in case Δ_m=4 in Egghe (2002b)) but (33), interpreted for g is important since this requires ∃<1 (for Zipf's law with T=4). By (11) this requires Lotka's ∀>2 which are the only possible values since T=4 is equivalent with Δ_m=4 (Note I.3) and the latter implies ∀>2 (see Egghe (2002b)). Hence, by (11), also ∃<1 is implied by the condition T=4 so that (33) represents the most general case.

In the latter interpretation (∃<1 hence ∀>2), formula (33) is the key for reproving a result (on the generalized 80/20-rule) which is implicite in Burrell (1992) (p.22) and Gastwirth (1972) (p.307, Table 1).

Proposition III.3.1: Let ∃<1, hence ∀>2. Let : denote the average number of items per source. Then

$$L(g)(y) = y^{\frac{1}{\mu}} . \tag{34}$$

In words: A fraction y0[0,1] of the top sources produce a fraction $y^{\frac{1}{\mu}}$ of the items, i.e. the generalized 80/20-rule for power laws.

Proof: Using (33) we only have to show that

$$1-\beta=\frac{1}{\mu}$$

This, because of the fact that (see (11)) $\beta=\frac{1}{\alpha-1}$, reduces to showing that

$$\mu=\frac{\alpha-1}{\alpha-2} \qquad (35)$$

This is known – see Egghe (2002b) in case Δ_m=4 and this is so because (see Note I.3) T=4 which in turn is so since ∃<1 (see Egghe (2002b)). ~

We can now state and prove the following basic result.

Theorem III.3.1: L(h) is a strictly increasing function of (.

Proof: Since we keep x_m fixed here it suffices to show (by (22)) that $(_1<(_2$ implies that

$$\int_1^x \frac{dx'}{x'^{\gamma_1}} \int_1^{x_m} \frac{dx'}{x'^{\gamma_2}} < \int_1^x \frac{dx'}{x'^{\gamma_2}} \int_1^{x_m} \frac{dx'}{x'^{\gamma_1}}$$

for all x0[1,x_m]. This is equivalent with

$$\int_1^x \frac{dx'}{x'^{\gamma_1}} \int_x^{x_m} \frac{dx'}{x'^{\gamma_2}} < \int_1^x \frac{dx'}{x'^{\gamma_2}} \int_x^{x_m} \frac{dx'}{x'^{\gamma_1}}$$

For this it suffices to show that, for all xN0[1,x] and all yN0[x,x_m]

$$\frac{1}{x'^{\gamma_1}}\frac{1}{y'^{\gamma_2}} < \frac{1}{x'^{\gamma_2}}\frac{1}{y'^{\gamma_1}}$$

or

$$\left(\frac{y'}{x'}\right)^{\gamma_1}\left(\frac{x'}{y'}\right)^{\gamma_2}<1 \tag{36}$$

Since $(_2>(_1$ we can denote $(_2=(_1+\gamma$ where $\gamma>0$. In this notation, inequality (36) reads

$$\left(\frac{x'}{y'}\right)^{\varepsilon}<1$$

which is trivial since xN<yN always. ~

The same proof applies for the analogous result for discrete functions as in (1), (2), by replacing integrals by discrete sums.

<u>Important remark III.3.2</u>:

It is clear that the above theorem applies to Lotka’s law

$$f(j)=\frac{C}{j^{\alpha}}$$

j0[1,Δ_m] and to Zipf's law

$$g(r)=\frac{D}{r^{\beta}}$$

r0[1,T+1]. So we have that L(f) is strictly increasing in ∀ (i.e. the inequality increases with ∀): hereby one examines the inequality in the numbers f(j), j0[1,Δ_m]. Much more important, from an informetric point of view (but mathematically the same) is the study of the inequality in the numbers g(r), r0[1,T+1] since these numbers express the source-productivity. So, theorem III.3.1 gives that L(g) is strictly increasing in ∃, hence the inequality increases with ∃. The latter result can be interpreted in function of ∀. Note that (11) gives $\beta=\frac{1}{\alpha-1}$. So we have the result that L(g) decreases in ∀. Since this is a source of confusion (see Wagner-Döbler and Berg (1995), Rao (1988)) we state these results (in ∀) explicitely as a corollary.

Corollary III.3.3:
Let f and g be as above. Then

(i) L(f) strictly increases in ∀
(ii) L(g) strictly increases in ∃
(iii) L(g) strictly decreases in ∀

It is clear from Rao (1988) that Rao refers to expression (i), so the criticism in Wagner-Döbler and Berg (1995) is not in order, especially since they discuss the inequality in Lotka's law (hence (i)) and not the one in Zipf's law (being (ii)).

This dual interpretation of inequality is typical in informetrics: one can consider the size-frequency function f but also the rank-frequency function g, hereby interchanging the role of sources and items. In this sense corollary III.3.3 is not surprising and similar differences will be found in the next subsection where we will give explicite formulae for the concentration measures G, V^2 and Th for the dual power functions f (Lotka) and g (Zipf). Note that, by using the general function h as in (30), we are able to develop the Lotka- as well as the Zipf-type concentration theory.

III.4 Concentration measures for power functions

III.4.1 Calculation of the Gini index

The easiest measure to calculate is the Gini index G (formula (28)) being twice the area under the Lorenz curve, minus 1. Let $\gamma \neq 1$ and $x_m < \infty$. Then (31) implies for the Gini index of h, denoted G(h),

$$G(h) = 2\int_0^1 L(h)(y)dy - 1$$

$$G(h) = \frac{2}{x_m^{1-\gamma} - 1}\int_0^1 \left(\left(y\left(x_m - 1\right) + 1\right)^{1-\gamma} - 1\right)dy - 1$$

$$G(h) = \frac{2}{(x_m - 1)(x_m^{1-\gamma} - 1)}\left[\frac{1}{2-\gamma}(x_m^{2-\gamma} - 1) - (x_m - 1)\right] - 1 \qquad (37)$$

The limiting value for $x_m \to \infty$ is

$$\lim_{x_m \to \infty} G(h) = \frac{2}{2-\gamma} - 1 = \frac{\gamma}{2-\gamma} \qquad (38)$$

but here (is restricted to (<1. This is in accordance with the calculation of G(h) using (33) which is also restricted to (<1 and which also yields (38).

So, for Lotka's function f we have

$$G(f) = \frac{2}{(\rho_m - 1)(\rho_m^{1-\alpha} - 1)}\left[\frac{1}{2-\alpha}\left(\rho_m^{2-\alpha} - 1\right) - \left(\rho_m - 1\right)\right] - 1 \qquad (39)$$

and for Zipf's function g we have

$$G(g) = \frac{2}{T\left(\left(T+1\right)^{1-\beta} - 1\right)}\left(\frac{1}{2-\beta}\left(\left(T+1\right)^{2-\beta} - 1\right) - T\right) - 1 \qquad (40)$$

G(g) in function of ∀ yields (using $\beta = \frac{1}{\alpha - 1}$)

$$G(g) = \frac{2}{T\left(\left(T+1\right)^{\frac{\alpha-2}{\alpha-1}} - 1\right)}\left(\frac{\alpha - 1}{2\alpha - 3}\left(\left(T+1\right)^{\frac{2\alpha-3}{\alpha-1}} - 1\right) - T\right) - 1 \qquad (41)$$

For the limiting values we have, based on (38), (applicable only on g as explained above but this is the most important case)

$$\lim_{T\to\infty} G(g) = \frac{\beta}{2-\beta} \tag{42}$$

$$= \frac{1}{2\alpha - 3} \tag{43}$$

(∃<1, equivalently, ∀>2). Result (43) was already proved in Burrell (1992), where only this limiting case (and ∀>2) is considered. Note that (42) and (43) trivially confirm corollary III.3.3 as it should. This is also the case for formulae (39), (40), (41); the proof is left to the reader.

We leave it also to the reader to calculate G(h) for (=1, using formula (32).

<u>III.4.2 Calculation of the variation coefficient</u>

We present two methods: one based on (26) for general h and one based on the formula $V = \frac{\sigma}{\mu}$, which only yields V for Zipf's function g.

<u>III.4.2.1 First method</u>

By (26):

$$V^2(h) = \int_0^1 [L(h)']^2(y)dy - 1 \ ,$$

where L(h) is given by (31) $(\gamma \neq 1)$ and (32) $(\gamma = 1)$. If $\gamma \neq 1$, we have

$$L(h)'(y) = \frac{(1-\gamma)(x_m - 1)}{(x_m^{1-\gamma} - 1)(y(x_m - 1) + 1)^{\gamma}} \tag{44}$$

If $\gamma = 1$, we have

$$L(h)'(y) = \frac{x_m - 1}{(\log x_m)(y(x_m - 1) + 1)} \tag{45}$$

Let now $\gamma \neq 1, \quad \gamma \neq \frac{1}{2}$. Then

$$V^2(h) = \int_0^1 \frac{(1-\gamma)^2 (x_m - 1)^2}{\left(x_m^{1-\gamma} - 1\right)^2 \left(y\left(x_m - 1\right) + 1\right)^{2\gamma}} dy$$

$$V^2(h) = \frac{(1-\gamma)^2 (x_m - 1)\left(x_m^{1-2\gamma} - 1\right)}{\left(x_m^{1-\gamma} - 1\right)^2 (1 - 2\gamma)} - 1 \tag{46}$$

If $\gamma = \frac{1}{2}$, then we have

$$V^2(h) = \frac{(x_m - 1)\log x_m}{4\left(\sqrt{x_m} - 1\right)^2} - 1 \tag{47}$$

If $\gamma = 1$, then (45) implies

$$V^2(h) = \frac{(x_m - 1)^2}{(\log x_m)^2 x_m} - 1 \tag{48}$$

As always, $V^2(f)$ (f: Lotka) is then calculated replacing x_m by Δ_m and (by (∀) in the above formulae and $V^2(g)$ (g: Zipf) is then calculated replacing x_m by T+1 and (by $\beta = \frac{1}{\alpha - 1}$. So we have, by (46)

$$V^2(f) = \frac{(1-\alpha)^2 (\rho_m - 1)\left(\rho_m^{1-2\alpha} - 1\right)}{\left(\rho_m^{1-\alpha} - 1\right)^2 (1 - 2\alpha)} - 1 \tag{49}$$

$$V^2(g) = \frac{(1-\beta)^2 \, T\left((T+1)^{1-2\beta}-1\right)}{\left((T+1)^{1-\beta}-1\right)^2 (1-2\beta)} - 1 \tag{50}$$

which takes the following form in $\alpha > 1, \ \ \alpha \neq 2, \ \ \alpha \neq 3$

$$V^2(g) = \frac{(\alpha-2)^2}{(\alpha-1)(\alpha-3)} \cdot \frac{T\left((T+1)^{\frac{\alpha-3}{\alpha-1}}-1\right)}{\left((T+1)^{\frac{\alpha-2}{\alpha-1}}-1\right)^2} - 1 \ . \tag{51}$$

III.4.2.2 Second method

This method only applies to the calculation of $V^2(g)$ since we use the formula $V^2(g) = \frac{\sigma^2}{\mu^2}$,

where Φ^2 and : are the variance and average of the Zipf function g, which can be calculated using the Lotka function $\frac{f}{T}$ as the weight function:

$$\mu = \int_1^{\rho_m} j \frac{f(j)}{T} dj \tag{52}$$

$$\sigma^2 = \int_1^{\rho_m} (j - \mu)^2 \frac{f(j)}{T} dj$$

$$\sigma^2 = \int_1^{\rho_m} j^2 \frac{f(j)}{T} dj - \mu^2 \tag{53}$$

So

$$V^2(g) = \frac{T}{A^2} \int_1^{\rho_m} j^2 f(j) dj - 1 \tag{54}$$

where $\mu = \frac{A}{T}$, the average number of items per source. We have, if $\alpha > 1, \ \ \alpha \neq 2, \ \ \alpha \neq 3$:

$$T = \int_1^{\rho_m} \frac{C}{j^{\alpha}} dj$$

$$T = \frac{C}{1-\alpha}\left(\rho_m^{1-\alpha} - 1\right) \tag{55}$$

$$A = \int_1^{\rho_m} j f(j) dj$$

$$A = \int_1^{\rho_m} \frac{C}{j^{\alpha-1}} dj$$

$$A = \frac{C}{2-\alpha}\left(\rho_m^{2-\alpha} - 1\right) \tag{56}$$

and

$$\int_1^{\rho_m} j^2 f(j) dj$$

$$= \int_1^{\rho_m} \frac{C}{j^{\alpha-2}} dj$$

$$= \frac{C}{3-\alpha}\left(\rho_m^{3-\alpha} - 1\right) . \tag{57}$$

(55), (56) and (57) in (54) now gives

$$V^2(g) = \frac{(\alpha-2)^2}{(\alpha-1)(\alpha-3)} \cdot \frac{\left(\rho_m^{1-\alpha} - 1\right)\left(\rho_m^{3-\alpha} - 1\right)}{\left(\rho_m^{2-\alpha} - 1\right)^2} - 1$$

$$V^2(g) = \frac{(\alpha-2)^2}{(\alpha-1)(\alpha-3)} \cdot \frac{\left(\rho_m^{\alpha-1} - 1\right)\left(\rho_m^{\alpha-3} - 1\right)}{\left(\rho_m^{\alpha-2} - 1\right)^2} - 1 . \tag{58}$$

We now have the task of proving that (51) and (58) are the same. We have, by Zipf's law, that E=1, hence by (10)

$$\rho_m^{1-\alpha} = \frac{\alpha - 1}{C}$$

This and (55) give

$$T + 1 = \frac{C}{\alpha - 1} = \rho_m^{\alpha-1} \tag{59}$$

(59) is the link between (58) and (51). We leave it to the reader to consider the cases ∀=2 and ∀=3.

<u>III.4.2.3 Limiting case</u>

For $x_m \to \infty$, (46) gives if $\gamma < \frac{1}{2}$

$$\lim_{x_m \to \infty} V^2(h) = \frac{(1-\gamma)^2}{1-2\gamma} - 1 = \frac{\gamma^2}{1-2\gamma} \tag{60}$$

As said before, only $V^2(g)$ can be used here (but this is the most important case) and $\beta = \gamma < \frac{1}{2}$ implies ∀>3. Only then we have

$$\lim_{T \to \infty} V^2(g) = \frac{\beta^2}{1-2\beta} = \frac{1}{(\alpha-1)(\alpha-3)} \tag{61}$$

using that $\beta = \frac{1}{\alpha - 1}$. Note again that (61) confirms corollary III.3.3, as it should.

Formula (61) also follows from (33) by direct calculation (again we have to restrict ourselves to the case $\gamma = \frac{1}{2}$).

III.4.3 Calculation of Theil's measure

We have to evaluate (27) with L(h)N as in (44) $(\gamma \neq 1)$ or as in (45) $(\gamma = 1)$. We will limit ourselves to the general case $\gamma \neq 1$, leaving the other calculation to the reader. The calculation is straightforward but tedious. We have

$$\mathrm{Th}(h) = \int_0^1 L(h)'(y) \log\big(L(h)'(y)\big)\,dy$$

$$= \int_0^1 \frac{(1-\gamma)(x_m-1)}{\left(x_m^{1-\gamma}-1\right)\left(y(x_m-1)+1\right)^{\gamma}} \log\left[\frac{(1-\gamma)(x_m-1)}{\left(x_m^{1-\gamma}-1\right)\left(y(x_m-1)+1\right)^{\gamma}}\right] dy$$

$$= \frac{(1-\gamma)(x_m-1)}{x_m^{1-\gamma}-1} \log\left[\frac{(1-\gamma)(x_m-1)}{x_m^{1-\gamma}-1}\right] \int_0^1 \frac{dy}{\left(y(x_m-1)+1\right)^{\gamma}}$$

$$+ \frac{(1-\gamma)(x_m-1)}{x_m^{1-\gamma}-1} \int_0^1 \left(y(x_m-1)+1\right)^{-\gamma} \log\left[\left(y(x_m-1)+1\right)^{-\gamma}\right] dy$$

$$= \frac{(1-\gamma)\log\left[\frac{(1-\gamma)(x_m-1)}{x_m^{1-\gamma}-1}\right]}{x_m^{1-\gamma}-1} \int_1^{x_m} \frac{dz}{z^{\gamma}} - \frac{\gamma(1-\gamma)}{x_m^{1-\gamma}-1} \int_1^{x_m} z^{-\gamma} \log z\,dz$$

using the substitution $z=y(x_m-1)+1$.

Now we use the following integral results which can readily be checked

$$\int \frac{dz}{z^{\gamma}} = \frac{z^{1-\gamma}}{1-\gamma}$$

$$\int \frac{\log z}{z^{\gamma}} dz = -\frac{1}{z^{\gamma-1}}\left(\frac{\log z}{\gamma-1}+\frac{1}{(\gamma-1)^2}\right).$$

We then obtain

$$\mathrm{Th}(h) = \log\left[\frac{(1-\gamma)(x_m-1)}{x_m^{1-\gamma}-1}\right]+\frac{\gamma\left(x_m^{\gamma-1}-(\gamma-1)\log x_m-1\right)}{(\gamma-1)\left(1-x_m^{\gamma-1}\right)} \tag{62}$$

The measure Th(h) for $x_m = \infty$ (restricted to $\gamma < 1$, see (33)) can be obtained in two ways: using (33) and noting that $L(h)'(y) = (1-\gamma)\,y^{-\gamma}$ and reperforming the above calculation of Th(h), or, more simply, take $\lim_{x_m \to \infty} \mathrm{Th}(h)$ in (62). Both ways give the formula (for $\gamma < 1$)

$$\mathrm{Th}(h) = \ln(1-\gamma)+\frac{\gamma}{1-\gamma} \tag{63}$$

Note, as before, that (63) can only be applied to the case of Zipf's function (15), the most important case. Hence $(\beta < 1)$

$$\mathrm{Th}(g) = \log(1-\beta)+\frac{\beta}{1-\beta}\ . \tag{64}$$

In terms of Lotka's exponent α we have $\left(\beta = \frac{1}{\alpha-1}\right)$ (hence $\alpha > 2$):

$$\mathrm{Th}(g) = \log\left(\frac{\alpha-2}{\alpha-1}\right)+\frac{1}{\alpha-2} \tag{65}$$

Note also that Th(g) increases in β and decreases in α as predicted by Corollary III.3.3.

<u>Note III.4.4</u>:

Theil's measure is (as the other measures G and V^2) calculated on the Lorenz curve, hence on normalized data in terms of abscissa and ordinate (see Figs. 1 and 2). It is different from the classical formula for the entropy of a distribution. To see this difference, let us first look at the discrete case. Theil's measure of the vector $X=(x_1,...,x_N)$ is given by (19), where the a_js are given by (17). In the same notation, the entropy of the vector X is defined as

$$\overline{H} = \overline{H}(X) = -\sum_{i=1}^{N} a_i \log a_i \tag{66}$$

Hence we have the relation

$$\mathrm{Th}(X) + \overline{H}(X) = \log N \tag{67}$$

or

$$\mathrm{Th}(X) = -\overline{H}(X) + \log N \tag{68}$$

We notice two differences between Th and $\overline{H}$: although they are linearly related, the relation is decreasing. In other words, $\overline{H}$ increases if and only if Th decreases. Now Th is a good measure of concentration (inequality), hence $\overline{H}$ is a good measure of dispersion (i.e. of equality), used e.g. in biology to measure diversity (see Rousseau and Van Hecke (1999)). Further $|\mathrm{Th}| \neq |\overline{H}|$ due to the fact that Th is calculated on the normalized Lorenz curve while $\overline{H}$ is not.

In the same way Th and $\overline{H}$ are different in the continuous case. We calculated already Th above. In the same way Lafouge and Michel (2002) calculate $\overline{H}$ (in fact, our proof the Th formula was based on this proof) via the formula

$$\overline{H}(f) = -\int_1^{\infty} f(j)\log(f(j))dj \tag{69}$$

(i.e. for $\rho_m = \infty$, hence $\alpha > 2$ necessarily, by Egghe (2002b)):

$$\overline{H}(f) = -\log(\alpha - 1) + \frac{1}{\alpha - 1} + 1 \tag{70}$$

(note that $\overline{H}(f)$ decreases with $\forall$, in agreement with Corollary III.3.3 since $-\overline{H}(f)$ is a good concentration measure). The result, apparently was first stated by Yablonsky (1980). For Zipf's function g we have (for $T = \infty$, hence (Note I.3) $\rho_m = \infty$ hence $\alpha > 2$, hence $\beta < 1$) we find that $\overline{H}(g) = \infty$. We leave it to the reader to calculate $\overline{H}(f)$, $\overline{H}(g)$ for $\rho_m, T < \infty$, in the same way as was done for the calculation of Th.

Since it is more important to study the concentration of g than of f, the above shows the importance of the "normalized" calculation of Th via the Lorenz curve, i.e. via formula (27): this yields a finite value (65) for Th(g) even in the case $T = \infty$, contrary to $\overline{H}(g)$.

III.5 Summary

We have indicated how Zipf's law relates to the one of Lotka: for $\alpha > 1$ they are equivalent provided that (12) is valid. This known result is used to prove that $T \to \infty$ iff $\rho_m \to \infty$ which is the basis for the equivalent treatment of both power laws. Then we showed that Price's law of concentration is equivalent with Zipf's law. As a consequence we could prove, if $\alpha = 2$, that Price's law is only valid if $\rho_m = C$, a result that was proved in an approximate way in Allison, Price, Griffith, Moravcsik and Stewart (1976).

The main part of the paper discusses continuous Lorenz theory applied to the laws of Lotka and Zipf. The general continuous Lorenz curve is introduced and applied to power laws. Also for the limiting case $(\rho_m = T = \infty)$ a Lorenz curve is defined (here the exponent in the power law, necessarily, is restricted to values in]0,1[. For a general power law we prove that the Lorenz curve is strictly increasing with the exponent in the power law.

Based on results in Egghe (2002a) we calculate explicite formulae for the Gini index, the variation coefficient and Theil's measure (for the general case $\rho_m, T < \infty$ and also for the limiting case $\rho_m = T = \infty$). In the latter case we refind a result of Burrell (1992) for the Gini index and of Gastwirth (1972) for the generalized 80/20 rule. The results on Theil's measure are related with entropy.

References

P.D. Allison, D. De Solla Price, B.C. Griffith, M.J. Moravcsik and J.A. Stewart (1976). Lotka's law: a problem in its interpretation and application. Social Studies of Science 6, 269-276, 1976.

T.M. Apostol (1957). Mathematical Analysis. Addison-Wesley, Reading (Massachusetts), USA, 1957.

A.B. Atkinson (1970). On the measurement of inequality. Journal of Economic Theory 2, 244-263, 1970.

J. Berg and R. Wagner-Döbler (1996). A multidimensional analysis of scientific dynamics. Part I. Case studies in mathematical logic in the 20th century. Scientometrics 35(3), 321-346, 1996.

Q.L. Burrell (1992). The Gini index and the Leimkuhler curve for bibliometric processes. Information Processing and Management 28(1), 19-33, 1992.

L. Egghe (1985). Consequences of Lotka's law for the law of Bradford. Journal of Documentation 41(3), 173-189, 1985.

L. Egghe (1987). An exact calculation of Price's law for the law of Lotka. Scientometrics 11 (1-2), 81-97, 1987.

L. Egghe (1989). The Duality of informetric Systems with Applications to the empirical Laws. Ph. D. Thesis, City University, London (UK), 1989.

L. Egghe (1990). The duality of informetric systems with applications to the empirical laws. Journal of Information Science 16(1), 17-27, 1990.

L. Egghe (2002a). Construction of concentration measures for general Lorenz curves using Riemann-Stieltjes integrals. Mathematical and Computer Modelling 35, 1149-1163, 2002.

L. Egghe (2002b). The source-item coverage of the Lotka function. Preprint, 2002.

L. Egghe and R. Rousseau (1986). A characterization of distributions which satisfy Price's law and consequences for the law of Zipf and Mandelbrot. Journal of Information Science 12, 193-197, 1986.

L. Egghe and R. Rousseau (1990a). Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science. Elsevier, Amsterdam, 1990.

L. Egghe and R. Rousseau (1990b). Elements of concentration theory. In: Informetrics 89/90 (L. Egghe and R. Rousseau, eds.). Selection of papers submitted for the second international conference on Bibliometrics, Scientometrics and Informetrics, 97-137, Elsevier, Amsterdam, 1990.

L. Egghe and R. Rousseau (1991). Transfer principles and a classification of concentration measures. Journal of the American Society for Information Science 42(7), 479-489, 1991.

L. Egghe and R. Rousseau (2003). Size-frequency and rank-frequency relations, power laws and exponential relations: a unified approach. Progress of Natural Science, to appear, 2003.

J. Feder (1988). Fractals. Plenum, New York, 1988.

J.L. Gastwirth (1971). A general definition of the Lorenz curve. Econometrica 39(6), 1037-1039, 1971.

J.L. Gastwirth (1972). The estimation of the Lorenz curve and Gini index. The Review of Economics and Statistics 54(3), 306-316, 1972.

C. Gini (1909). Il diverso accrescimento delle classi sociali e la concentrazione della richezza. Giornale degli Economisti, serie 11, 37, 1909.

W. Glänzel and A. Schubert (1985). Price distribution. An exact formulation of Price's "square root law". Scientometrics 7(3-6), 211-219, 1985.

B.M. Gupta, L. Sharma and S. Kumar (1998). Literature growth and author productivity. Patterns in Indian physics. Information Processing and Management 34(1), 121-131, 1998.

G.H. Hardy, J.E. Littlewood and G. Pólya (1928). Some simple inequalities satisfied by convex functions. Messenger of Mathematics 58, 145-152, 1928.

T. Lafouge and C. Michel (2001). Links between information construction and information gain. Entropy and bibliometric distributions. Journal of Information Science 27(1), 39-49, 2001.

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

B.B. Mandelbrot (1977). The fractal Geometry of Nature. Freeman, New York, 1977.

P.T. Nicholls (1988). Price's square root law: empirical validity and relation to Lotka's law. Information Processing and Management 24(4), 469-477, 1988.

D.J. De Solla Price (1976). A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science 27, 292-306, 1976.

I.K. Ravichandra Rao (1988). Probability distributions and inequality measures for analyses of circulation data. In: Informetrics 87/88 (L. Egghe and R. Rousseau, eds.). Select Proceedings of the first International Conference on Bibliometrics and theoretical Aspects of Information Retrieval, 231-248, Elsevier, Amsterdam, 1988.

F.S. Roberts (1979). Measurement Theory with Applications to Decisionmaking, Utility and the social Sciences. Addison-Wesley, Reading (Massachusetts), USA, 1979.

R. Rousseau (1990). Relations between continuous versions of bibliometric laws. Journal of the American Society for Information Science 41(3), 197-203, 1990.

R. Rousseau and P. Van Hecke (1999). Measuring biodiversity. Acta Biotheoretica 47, 1-5, 1999.

H. Theil (1967). Economics and Information Theory. North-Holland, Amsterdam, 1967.

R. Wagner-Döbler and J. Berg (1995). The dependence of Lotka's law on the selection of time periods in the development of scientific areas and authors. Journal of Documentation 51(1), 28-43, 1995.

A.I. Yablonsky (1980). On fundamental regularities of the distribution of scientific productivity. Scientometrics 2(1), 3-34, 1980.