Solution by step functions of a minimum problem in $L^2[0,T]$, using generalized h- and g-indices

¹Leo Egghe and ^{2,3}Ronald Rousseau

¹University of Hasselt, Belgium
leo.egghe@uhasselt.be

² University of Antwerp, Faculty of Social Sciences, B-2020 Antwerpen, Belgium, ronald.rousseau@uantwerpen.be

³ Centre for R&D Monitoring (ECOOM) and Dept. MSI, KU Leuven, Belgium, ronald.rousseau@kuleuven.be

ABSTRACT

In this article we solve a minimum problem involving step functions. The solution of this problem leads to an investigation into generalized h- and g-indices. This minimum problem and the related generalized h- and g-indices are studied in a general context of decreasing differentiable functions as well as in the specific case of Lotkaian informetrics. The study illustrates the use of h-and g-indices and their generalizations in a context which bears no relation to the research evaluation context in which these indices were originally introduced.

Keywords: decreasing differentiable functions; generalized h-indices; generalized g-indices; Lotkaian informetrics; power laws; step functions;

1. Introduction

Let $Z(r)$ be a positive, non-constant, decreasing, differentiable function defined on the interval $[0,T]$:

$$Z: [0,T] \rightarrow \overline{\mathbb{R}^+}$$

and let for each $a \in \mathbb{R}^+$, $0 \leq a \leq T$, φ_a be the step function defined on $[0,T]$

$$\text{such that } \varphi_a(r) = \begin{cases} a & \text{if } r \in [0, a] \\ 0 & \text{if } r \notin [0, a] \end{cases}.$$

We note that the function φ_a is equal to a times the characteristic function of $[0,a]$.

In this article we study the following research problem (RP): find the value a in $[0, T]$, or equivalently the function φ_a , such that the L^2 -distance between Z and φ_a reaches a minimum.

2. Generalized h- and g-values

We consider for every $\theta > 0$, the function M_θ defined as $M_\theta(r) = \theta r$, where also here $r \in [0, T]$.

Proposition 1. If $Z(T) \leq \theta T$, then there exists a unique point h_θ such that $Z(h_\theta) = M_\theta(h_\theta) = \theta h_\theta$.

Proof. Define $k(x) = Z(x) - \theta x$. The function $k(x)$ is continuous and strictly decreasing on $[0, T]$, $k(0) > 0$ and $k(T) \leq 0$. By the intermediate value theorem the function k takes all values on the interval $[k(T), k(0)]$ (where we have used the fact that k is decreasing). Hence there exists a point $h_\theta \in [0, T]$ such that $k(h_\theta) = 0$. Consequently, $Z(h_\theta) = \theta h_\theta$. This value h_θ is unique as the function k is strictly decreasing on $[0, T]$.

Corollary 1. A characterization of h_θ

We note that: $r < h_\theta$ if and only if (in short: iff) $Z(r) > \theta r$, and $r > h_\theta$ iff $Z(r) < \theta r$. If $Z_\theta(r) = Z(r)/\theta$, then h_θ is the unique fixed point of Z_θ .

If $\theta = 1$, $h_\theta = h_1$ is the continuous h-index as introduced in (Egghe & Rousseau, 2006). The general h_θ has been introduced in (van Eck & Waltman, 2008) under the name of h_α -index. Comments on it can be found in (Gagolewski, 2009). It is clear that Proposition 1 and its corollary also hold if $Z(r)$ is a constant function.

Besides $Z(r)$ we define its integral function $Y(r) = \int_0^r Z(s)ds$, with $r \in [0, T]$. The function Y is increasing. Then, similarly to Proposition 1, we have the following proposition.

Proposition 2. If $Y(T) \leq \theta T^2$, then there exists a unique point g_θ in $[0, T]$ such that $Y(g_\theta) = \theta (g_\theta)^2$.

Proof. Based on (Egghe, 2006). Define for all r in $]0, T]$

$$\Gamma(r) = \frac{Y(r)}{r}$$

with, moreover, $\Gamma(0) = \lim_{r \rightarrow 0} \Gamma(r) = Z(0)$. We first prove that, as $Z(r)$ is not constant, Γ is decreasing on $[0, T]$. Indeed: $\Gamma'(r) = \frac{r Z(r) - Y(r)}{r^2} \leq 0$, since $r Z(r) \leq Y(r) = \int_0^r Z(s) ds$ for all r in $]0, T]$. Because of the definition of $\Gamma(0)$, Γ is decreasing on $[0, T]$. Now we apply Proposition 1 for the function $\Gamma(r)$ instead of $Z(r)$ and find a unique point g_θ such that $\Gamma(g_\theta) = \theta g_\theta$ or $Y(g_\theta) = \theta (g_\theta)^2$. Recall that in order to apply Proposition 1 we have the requirement $Z(T) \leq \theta T$, which now, applied to the function Γ , becomes: $Y(T) \leq \theta T^2$.

Corollary 2. A characterization of g_θ

We note that $r < g_\theta$ iff $Y(r) > \theta r^2$, and $r > g_\theta$ iff $Y(r) < \theta r^2$. We further recall from the proof that: $\frac{1}{g_\theta} Y(g_\theta) = \theta g_\theta = M_\theta(g_\theta)$. If $Y_\theta(r) = Y(r)/(r\theta)$, then g_θ is the unique fixed point of Y_θ .

If $\theta = 1$, then $g_\theta = g_1$ is the continuous version of the g -index as introduced in (Egghe, 2006). Again, the general g_θ -index has been introduced in (van Eck & Waltman, 2008).

Finally, we define the function $C(r)$ on $[0, T]$ as $\frac{1}{2} \left(Z(r) + \frac{1}{r} Y(r) \right)$ and f_θ as the unique point such that $\frac{1}{2} \left(Z(f_\theta) + \frac{1}{f_\theta} Y(f_\theta) \right) = C(f_\theta) = M_\theta(f_\theta) = \theta f_\theta$. The existence of f_θ follows from Proposition 1 with $Z(r)$ replaced by $\frac{1}{2} \left(Z(r) + \frac{1}{r} Y(r) \right)$. We note that, in particular:

$$C(f_{0.75}) = 0.75 f_{0.75}$$

This particular value for θ plays an important role further on. Similarly as before, we denote $C_\theta(r)$ as $\frac{1}{2\theta} \left(Z(r) + \frac{1}{r} Y(r) \right)$ and observe that f_θ is the unique fixed point of $C_\theta(r)$ on $[0, T]$.

We note that the function C is defined as an average of two functions, but, of course we do not expect that for each $\theta > 0$ the number f_θ is the average of the corresponding numbers h_θ and g_θ . This will be illustrated further on.

From the developments above we recall the following definitions.

Definitions

Given the positive, non-constant, decreasing, and differentiable function $Z(r)$ defined on the interval $[0, T]$, we define the functions $h_\theta(Z)$, $g_\theta(Z)$ and $f_\theta(Z)$ as the unique solutions of the following equations:

$$Z(h_\theta(Z)) = \theta h_\theta(Z) \quad (1)$$

$$\int_0^{g_\theta(Z)} Z(s) ds = \theta (g_\theta(Z))^2 \quad (2)$$

$$\frac{1}{2} \left(Z(f_\theta(Z)) + \frac{1}{f_\theta(Z)} \int_0^{f_\theta(Z)} Z(s) ds \right) = \theta f_\theta(Z) \quad (3)$$

If it is clear about which function Z we are talking, the symbol Z is removed from the notations $h_\theta(Z)$, $g_\theta(Z)$ and $f_\theta(Z)$. Given $Z(r)$ and T Eqs. 1, 2 and 3 are implicit definitions of the functions h_θ , g_θ and f_θ . As the function Z is differentiable, also the functions h_θ , g_θ and f_θ are differentiable (in θ). We further recall that given $Z(r)$ there may be restrictions on the values of T and θ for which Eqs. (1), (2) and (3) apply. Concretely: given $Z(r)$ and θ , equation (1) only applies for $Z(T) \leq \theta T$; equation (2) only applies for $Y(T) \leq \theta T^2$ with $Y(r) = \int_0^r Z(s) ds$ and equation (3) only applies if $\frac{1}{2} \left(Z(T) + \frac{1}{T} Y(T) \right) \leq \theta T$. The co-domain of $h_\theta(Z)$, $g_\theta(Z)$ and $f_\theta(Z)$ is always a subset of $]0, T]$.

Proposition 3. If Z is strictly decreasing then, for all $\theta > 0$: $h_\theta < g_\theta$

Proof. If $a = h_\theta$ then $Z(a) = \theta a$. As $Z(r)$ is strictly decreasing this implies that for all $r \in [0, a[$: $Z(r) > \theta a$. Integration over $[0, a]$ leads to $\int_0^a Z(r) dr = Y(a) > \theta a^2$. Consequently, by Corollary 2, $a = h_\theta < g_\theta$.

Proposition 4. If Z is strictly decreasing then, for all $\theta > 0$: $f_\theta \in] h_\theta, g_\theta [$

Proof. As Z is not constant and strictly decreasing we have for all r in $]0, T]$: $Z(r) < \frac{1}{r} \int_0^r Z(s) ds$ iff $Z(r) < \frac{1}{2} \left(Z(r) + \frac{1}{r} Y(r) \right) < \frac{1}{r} \int_0^r Z(s) ds$. With $r = f_\theta$, this becomes: $Z(f_\theta) < f_\theta = \frac{1}{2} \left(Z(f_\theta) + \frac{1}{f_\theta} Y(f_\theta) \right) < \frac{1}{f_\theta} \int_0^{f_\theta} Z(s) ds$.

From the first inequality we derive that $h_\theta < f_\theta$ and from the second one that $f_\theta < g_\theta$. This proves Proposition 4.

Remark

It is obvious that if $Z_1(r) \leq Z_2(r)$ on $[0, T]$, and denoting $h_\theta(Z_1)$ for the h_θ -value of the function Z_1 (and with similar notations for the other cases) then for each $\theta > 0$, $h_\theta(Z_1) \leq h_\theta(Z_2)$, $g_\theta(Z_1) \leq g_\theta(Z_2)$ and $f_\theta(Z_1) \leq f_\theta(Z_2)$.

3. The main result

RP asks for the minimum value of $d^2(Z, \varphi_a) = \int_0^T (Z(r) - \varphi_a(r))^2 dr$.

Theorem 1. The L^2 -distance between Z and φ_a reaches a minimum for $a = f_{0.75}$.

Proof. Put $F(a) = d^2(Z, \varphi_a)$. Then: $d^2(Z, \varphi_a) = \int_0^T (Z(r) - \varphi_a(r))^2 dr = \int_0^a (Z(r) - a)^2 dr + \int_a^T (Z(r))^2 dr$. This is equal to $\int_0^a (Z(r))^2 dr - 2a \int_0^a Z(r) dr + a^3 + \int_a^T (Z(r))^2 dr = \int_0^T (Z(r))^2 dr - 2a \int_0^a Z(r) dr + a^3$.

Now $F'(a) = 3a^2 - 2Y(a) - 2aZ(a)$, which is equal to zero iff:

$$2aZ(a) + 2Y(a) = 3a^2 \text{ iff } \frac{1}{2} \left(Z(a) + \frac{1}{a} Y(a) \right) = \frac{3}{4} a \text{ iff } a = f_{0.75}.$$

Corollary 3. The L^2 -distance $d(Z, \varphi_a)$ reaches a minimum for a value a in the interval $]h_{0.75}, g_{0.75}[$, namely for $f_{0.75}$.

4. Properties of the generalized h- and g-indices

We first consider the special case of a constant function $Z_0(r) = K \neq 0$, defined on $\mathbb{R}^+$. Then by definition $Z_0(h_\theta) = K = \theta h_\theta$. Hence $h_\theta(Z_0) = K/\theta$.

Further: from $Y(g_\theta(Z_0)) = \int_0^{g_\theta(Z_0)} K ds = \theta (g_\theta(Z_0))^2$ we derive that $K \cdot g_\theta(Z_0) = \theta (g_\theta(Z_0))^2$.

Hence $g_\theta(Z_0) = K/\theta$. We see that, contrary to Proposition 3, in the case of a constant function $h_\theta = g_\theta$, for all $\theta > 0$. We further observe that $h_\theta (= g_\theta) = K/\theta$ is a convex decreasing function of θ .

We will from now assume that the function $Z(r)$ is positive, non-constant, decreasing, differentiable and, moreover, convex. If such a function is defined on $\mathbb{R}^+$:

$$Z: \mathbb{R}^+ \rightarrow \overline{\mathbb{R}^+}$$

this implies that we are free to choose the endpoint T of the interval $[0, T]$. Definitions (1), (2) and (3) also apply to such functions Z .

Next we show the following proposition.

Proposition 5. The function h is a convex decreasing function.

Proof. From the definition $Z(h_\theta) = \theta h_\theta$ we derive that $Z'(h_\theta) \cdot h'_\theta = h_\theta + \theta h'_\theta$. Hence, $h'_\theta = \frac{h_\theta}{Z'(h_\theta) - \theta} < 0$ (as the numerator is strictly positive and the denominator is negative). This shows that the function h is strictly decreasing.

Next we show that the function h is convex: we will show that h''_θ is strictly positive. Taking the derivative of h'_θ yields:

$$h''_\theta = \frac{(Z'(h_\theta) - \theta)h'_\theta - h_\theta(Z''(h_\theta)h'_\theta - 1)}{(Z'(h_\theta) - \theta)^2}$$

Rewriting the numerator we see that all four terms are positive: $Z'(h_\theta)h'_\theta - \theta h'_\theta - h_\theta Z''(h_\theta)h'_\theta + h_\theta$. As the denominator is also positive we see that h''_θ is strictly positive and hence the function h is convex decreasing.

Also the function g_θ is convex decreasing. This will be shown in a number of steps.

Lemma 1. Denoting $Y(r)/r$ by $\Gamma(r)$ we show that $g_\theta(Z) = h_\theta(\Gamma)$.

Proof. $h_\theta(\Gamma(r))$ is defined through the equality: $\Gamma(h_\theta(\Gamma)) = \theta h_\theta(\Gamma)$, or equivalently:

$$\frac{\int_0^{h_\theta(\Gamma)} Z(s) ds}{h_\theta(\Gamma)} = \theta h_\theta(\Gamma) \Leftrightarrow \int_0^{h_\theta(\Gamma)} Z(s) ds = \theta h_\theta^2(\Gamma) \Leftrightarrow h_\theta(\Gamma) = g_\theta(Z)$$

Remark. This lemma is actually an important result as it shows that g_θ can be studied through h_θ .

Lemma 2. The function $\Gamma(r)$ is positive, non-constant, convex decreasing and differentiable.

Proof. From the definition of $\Gamma(r)$ (lemma 1) we see that $\Gamma'(r) = \frac{rZ(r)-Y(r)}{r^2}$. As $Z(r)$ is decreasing we know that $Y(r) \geq r.Z(r)$, hence $\Gamma'(r) \leq 0$. This proves that $\Gamma(r)$ is decreasing.

Now, $\Gamma''(r) = \frac{r^2(Z(r)+rZ'(r)-Z(r))-(rZ(r)-Y(r))2r}{r^4} = \frac{r^3Z'(r)-2r^2Z(r)+2rY(r)}{r^4}$. This expression is larger than or equal to zero iff $r^3Z'(r) - 2r^2Z(r) + 2rY(r) \geq 0$.

If $s < r$, then the point $(s, Z(s))$ is situated above (or on) the tangent on Z in the point $(r, Z(r))$. This means: $Z(s) \geq Z(r) + Z'(r).(s-r)$. Taking the integral from 0 to r yields: $Y(r) = \int_0^r Z(s)ds \geq \int_0^r Z(r)ds + \int_0^r Z'(r)(s-r)ds = rZ(r) + \frac{r^2}{2}Z'(r) - r^2Z'(r)$. Now, $r^3Z'(r) - 2r^2Z(r) + 2rY(r) \geq r^3Z'(r) - 2r^2Z(r) + 2r^2Z(r) + r^3Z'(r) - 2r^3Z'(r) = 0$. This proves that $\Gamma(r)$ is convex decreasing.

Proposition 6

The function g_θ is convex decreasing.

Proof. From Lemma 1 we know that $g_\theta(Z) = h_\theta(\Gamma)$. Now, as Γ is a positive, non-constant, decreasing, differentiable, convex function defined on $\mathbb{R}^+$ (by Lemma 2) we derive from the fact that the function h is convex decreasing that also g is convex decreasing.

Next, we illustrate the theory by two examples.

Example 1: a power function

We consider the case $Z_p(r) = \frac{B}{r^\beta}$ with $B > 0$; $0 < \beta < 1$, defined on $[0, T]$.

We have shown the existence of h_θ , g_θ and f_θ in the case that Z is a function defined on the interval $[0, T]$. As the function $h_\theta(Z_p)$ is defined through the equality: $\frac{B}{(h_\theta)^\beta} = \theta h_\theta$, with Z_p strictly decreasing and $y = \theta x$

($\theta > 0$) strictly increasing this leads to a unique: $h_\theta = \left(\frac{B}{\theta}\right)^{1/(1+\beta)}$, if this point is not larger than T . Expressing this requirement leads to:

$$\left(\frac{B}{\theta}\right)^{1/(1+\beta)} \leq T$$

In the Lotkaian context $B = T^\beta$ (Egghe, 2005), thus

$$\frac{B}{\theta} \leq T^{(1+\beta)} \Leftrightarrow \frac{T^\beta}{\theta} \leq T^{(1+\beta)} \Leftrightarrow \frac{1}{\theta} \leq T$$

Hence, if T is given we have to restrict the possible values of θ (they must be at least equal to $1/T$), and if θ is given we may only consider functions Z_p defined on $]0, T]$ with T larger than or equal to $1/\theta$. In particular if $\theta = 1$ is given (for the standard h-index) then T must be larger than or equal to 1, which is always assumed (as $\beta < 1$).

Taking $\theta=1$ leads to : $h = h_1 = h(Z_p) = (B)^{(1/(1+\beta))}$, which is a well-known result from (Egghe & Rousseau, 2006). From this we obtain:

$$h_\theta = h. (\theta)^{-(1/(1+\beta))}. \quad (4)$$

Now, the function $g_\theta(Z_p)$ is defined through the equality: $\int_0^{g_\theta} \frac{B}{(s)^\beta} ds = \theta(g_\theta)^2$. This leads to: $B \frac{1}{1-\beta} g_\theta^{1-\beta} = \theta(g_\theta)^2$ or $g_\theta = \left(\frac{B}{\theta(1-\beta)}\right)^{(1/(1+\beta))}$.

Again, we have to formulate the condition under which g_θ exists. This is:

$$\left(\frac{B}{\theta(1-\beta)}\right)^{(1/(1+\beta))} \leq T \Leftrightarrow \frac{T^\beta}{\theta(1-\beta)} \leq T^{(1+\beta)} \Leftrightarrow \frac{1}{\theta(1-\beta)} \leq T,$$

which is equivalent to: T

$$\frac{1}{\theta} \leq (1 - \beta) \quad (5)$$

Taking $\theta=1$ gives: $g = g_1 = g(Z_p) = \left(\frac{B}{1-\beta}\right)^{(1/(1+\beta))}$, leading to:

$$g_\theta = g. (\theta)^{-(1/(1+\beta))}. \quad (6)$$

As the exponent $1/(1+\beta)$ is positive, $h_\theta(Z_p)$ and $g_\theta(Z_p)$ are convex decreasing in θ .

Finally, we also determine f_θ . Let $x = f_\theta$, then $x \neq 0$, is determined by:

$$Z_p(x) + \frac{1}{x} \int_0^x Z_p(s) ds = 2\theta x$$

$$\begin{aligned} \Leftrightarrow 2\theta x^{\beta+1} &= B \left(\frac{2-\beta}{1-\beta} \right) \\ \Leftrightarrow x = f_\theta &= \left(\frac{B}{2\theta} \left(\frac{2-\beta}{1-\beta} \right) \right)^{(1/(\beta+1))} \end{aligned} \quad (7)$$

As for all $\theta > 0$: $f_\theta \in] h_\theta, g_\theta [$ (Proposition 4) we have already a requirement for the existence of f_θ . Yet, we can refine this as follows:

$$\begin{aligned} \left(\frac{T^\beta}{2\theta} \left(\frac{2-\beta}{1-\beta} \right) \right)^{(1/(\beta+1))} &\leq T \Leftrightarrow \left(\frac{1}{2\theta} \left(\frac{2-\beta}{1-\beta} \right) \right)^{(1/(\beta+1))} \leq T^{(1/(\beta+1))} \\ \Leftrightarrow \frac{1}{2\theta} \left(\frac{2-\beta}{1-\beta} \right) &\leq T \Leftrightarrow_{\beta < 1} (2-\beta) \leq 2\theta(1-\beta)T \Leftrightarrow \beta \leq \frac{2\theta T - 2}{2\theta T - 1} \end{aligned}$$

Otherwise, we may state this requirement as: $\frac{1}{\theta} \leq \frac{2(1-\beta)T}{2-\beta}$.

As expected, this is a weaker requirement than the one for g_θ . For the solution of our minimum problem we need $\theta = 0.75$. In that case the requirement on β becomes:

$$\beta \leq \frac{3T-4}{3T-2}.$$

With $\beta < 1$, it suffices to take T large enough.

Example 2. The function $Z_1(r) = 1-r$, defined on $[0,1]$.

Let $x = h_\theta$, then x is determined by:

$$Z_1(x) = \theta x \Leftrightarrow 1 - x = \theta x \Leftrightarrow x = \frac{1}{\theta + 1}$$

Let $y = g_\theta$, then $y \neq 0$, is determined by: $\int_0^y (1-s) ds = \theta y^2 \Leftrightarrow y - \frac{y^2}{2} = \theta y^2 \Leftrightarrow 1 = \left(\frac{1}{2} + \theta \right) y \Leftrightarrow y = \frac{1}{\theta + \frac{1}{2}}$.

Let $z = f_\theta$, then z is determined by:

$$\begin{aligned}
Z_1(z) + \frac{1}{z} \int_0^z Z_1(s) ds &= 2\theta z \\
\Leftrightarrow 2\theta z - (1 - z) - \frac{1}{z} \int_0^z (1 - s) ds &= 0 \\
\Leftrightarrow 2\theta z - (1 - z) - \frac{1}{z} \left(z - \frac{z^2}{2} \right) &= 0 \\
\Leftrightarrow 2\theta z - (1 - z) - 1 + \frac{z}{2} &= 0 \\
\Leftrightarrow z &= \frac{1}{\theta + \frac{3}{4}}
\end{aligned}$$

Clearly f_θ is situated between h_θ and g_θ . However:

$$\frac{1}{2}(h_\theta + g_\theta) = \frac{1}{2} \left(\frac{1}{\theta+1} + \frac{1}{\theta+\frac{1}{2}} \right) = \frac{2\theta+\frac{3}{2}}{2(\theta+1)(\theta+\frac{1}{2})} > \frac{1}{\theta+\frac{3}{4}}. \quad (8)$$

We know that the solution of RP is obtained for $a = f_\theta$, with $\theta = 0.75$. Hence this is $a = 2/3$.

5. Lotkaian theory

Within the so-called Lotkaian theory (Egghe, 2005) there exist corresponding rank-frequency and size-frequency functions, $Z_1(r) = \frac{B}{r^\beta}$, $0 < \beta < 1$, defined on $[0, T]$ and $L(x) = \frac{c}{x^\alpha}$; $\alpha > 2$, defined on $[1, +\infty[$. The relation between the exponents α and β is given by $\beta = \frac{1}{\alpha-1}$. Moreover, $B = T^{\left(\frac{1}{\alpha-1}\right)}$. In this theory, the function Z_1 is referred to as the Zipf function, while L is referred to as the corresponding Lotka function. Using the results from the previous example, we can express h and g and their generalizations h_θ and g_θ using the Lotka parameter α :

$$h_\theta = \left(\frac{B}{\theta} \right)^{\left(\frac{1}{1+\beta} \right)} = \left(\frac{T^{\left(\frac{1}{\alpha-1} \right)}}{\theta} \right)^{\left(\frac{(\alpha-1)}{\alpha} \right)} = \theta^{\left(\frac{(1-\alpha)}{\alpha} \right)} T^{\left(\frac{1}{\alpha} \right)} = \theta^{\left(\frac{(1-\alpha)}{\alpha} \right)} h \quad (9)$$

Similarly:

$$g_\theta = \left(\frac{B}{\theta(1-\beta)} \right)^{(1/(1+\beta))} = \left(\frac{T^{(1/(\alpha-1))}}{\theta^{(\frac{\alpha-2}{\alpha-1})}} \right)^{((\alpha-1)/\alpha)} = \left(\frac{\alpha-1}{\theta(\alpha-2)} \right)^{((\alpha-1)/\alpha)} T^{(1/\alpha)} = \theta^{((1-\alpha)/\alpha)} g. \quad (10)$$

The requirement for the existence of g_θ (Eq. 5), derived previously, can in the Lotkaian context be reformulated in terms of the α -parameter. This leads to:

$$\begin{aligned} \frac{1}{\theta} \leq \frac{\alpha-2}{\alpha-1} T &\Leftrightarrow_{\alpha>1} \frac{\alpha-1}{\theta} \leq (\alpha-2)T \Leftrightarrow \alpha \left(\frac{1-\theta T}{\theta} \right) \leq \frac{(1-2\theta)T}{\theta} \\ &\Leftrightarrow_{\frac{1}{\theta} < T} \alpha \geq \frac{2\theta T - 1}{\theta T - 1} \end{aligned}$$

In the case of the standard g-index ($\theta=1$) this requirement becomes $\alpha \geq \frac{2T-1}{T-1}$.

If $\alpha > 2$, this inequality can always be satisfied by taking T large enough as the right-hand function is a decreasing function in T ($T>1$) with limit 2.

Finally, we also determine f_θ in the Lotkaian case. Let $x = f_\theta$, then $x \neq 0$, is determined by:

$$x = f_\theta = \left(\frac{B}{2\theta} \left(\frac{2-\beta}{1-\beta} \right) \right)^{(1/(\beta+1))} \quad \text{see (7)}$$

In terms of the corresponding Lotka parameters, we have:

$$f_\theta = \left(\frac{T^{(1/(\alpha-1))}}{2\theta} \left(\frac{2\alpha-3}{\alpha-2} \right) \right)^{((\alpha-1)/\alpha)} = \left(\frac{1}{2\theta} \left(\frac{2\alpha-3}{\alpha-2} \right) \right)^{((\alpha-1)/\alpha)} T^{1/\alpha}. \quad (10)$$

As for all $\theta > 0$: $f_\theta \in] h_\theta , g_\theta [$ (Proposition 4) we have already a requirement for the existence of f_θ . Yet, we can refine this:

$$\begin{aligned} \left(\frac{1}{2\theta} \left(\frac{2\alpha-3}{\alpha-2} \right) \right)^{((\alpha-1)/\alpha)} T^{1/\alpha} \leq T &\Leftrightarrow \left(\frac{1}{2\theta} \left(\frac{2\alpha-3}{\alpha-2} \right) \right)^{((\alpha-1)/\alpha)} \leq T^{(\frac{\alpha-1}{\alpha})} \\ \Leftrightarrow \frac{1}{2\theta} \left(\frac{2\alpha-3}{\alpha-2} \right) &\leq T \Leftrightarrow_{\alpha>2} (2\alpha-3) \leq 2\theta(\alpha-2)T \Leftrightarrow \alpha \geq \frac{2\theta T - \frac{3}{2}}{\theta T - 1} \end{aligned}$$

Again we see that this is a weaker requirement than the one for g_θ . For the solution of our minimum problem we need $\theta = 0.75$. In that case the requirement on α becomes:

$$\alpha \geq \frac{\frac{3}{2}T - \frac{3}{2}}{\frac{3}{4}T - 1} = \frac{2T - 2}{T - \frac{4}{3}}$$

With $\alpha > 2$, it suffices to take T large enough.

Proposition 7

In the Lotkaian case we have for every $\theta > 0$:

$$f_\theta > \frac{1}{2} (h_\theta + g_\theta)$$

Proof. We have to show that:

$$\left(\frac{1}{2\theta} \left(\frac{2\alpha - 3}{\alpha - 2} \right) \right)^{((\alpha-1)/\alpha)} T^{1/\alpha} > \frac{1}{2} \left(\left(\frac{1}{\theta} \right)^{\frac{\alpha-1}{\alpha}} + \left(\frac{\alpha - 1}{\theta(\alpha - 2)} \right)^{((\alpha-1)/\alpha)} \right) T^{(1/\alpha)}$$

Or equivalently:

$$\left(\frac{1}{2} \left(\frac{2\alpha - 3}{\alpha - 2} \right) \right)^{((\alpha-1)/\alpha)} > \frac{1}{2} \left(1 + \left(\frac{\alpha - 1}{\alpha - 2} \right)^{((\alpha-1)/\alpha)} \right)$$

Or:

$$\left(\alpha - \frac{3}{2} \right)^{((\alpha-1)/\alpha)} > \frac{1}{2} \left((\alpha - 2)^{((\alpha-1)/\alpha)} + (\alpha - 1)^{((\alpha-1)/\alpha)} \right)$$

This inequality is correct if the function $y(x) = (\alpha - x)^{((\alpha-1)/\alpha)}$ is strictly concave on the interval $]1, 2[$.

Now, $y'(x) = \frac{\alpha-1}{\alpha} (\alpha - x)^{((\alpha-1)/\alpha)-1} (-1) = -\frac{\alpha-1}{\alpha} (\alpha - x)^{(-1/\alpha)} < 0$ and

$$y''(x) = \frac{\alpha-1}{\alpha^2} (\alpha - x)^{(-1/\alpha)-1} (-1) = -\frac{\alpha-1}{\alpha^2} (\alpha - x)^{(-1+\alpha)/\alpha} < 0.$$

This proves this proposition.

This proposition shows that in the Lotkaian case f_θ is situated closer to g_θ than to h_θ . Yet, this does not hold for any continuous function $Z(r)$. Indeed, from the earlier example $Z_1(r) = 1-r$, we know that $h_\theta = \frac{1}{\theta+1}$; $g_\theta = \frac{1}{\theta+\frac{1}{2}}$ and $f_\theta = \frac{1}{\theta+\frac{3}{4}}$. Clearly f_θ is situated between h_θ and g_θ , but we know already (Eq. 8) that $\frac{1}{2}(h_\theta + g_\theta) = \frac{1}{2}\left(\frac{1}{\theta+1} + \frac{1}{\theta+\frac{1}{2}}\right) > \frac{1}{\theta+\frac{3}{4}}$.

Comment. Although f_θ is defined, in words, as the average of the formulations of h_θ and g_θ , its value is not equal to the average of the values of h_θ and g_θ .

Next we consider RP in detail for the case $Z_1(r) = \frac{B}{r^\beta}$: for which a-value does $d^2(Z_1, \varphi_a) = \int_0^T (Z_1(r) - \varphi_a(r))^2 dr$ reach a minimum? Note that we have to exclude the case $\beta \geq 0.5$ as the above integral does not exist in that case.

Theorem 2. The L^2 -distance between Z_1 and φ_a reaches a minimum for $a = \left(\frac{2B(2-\beta)}{3(1-\beta)}\right)^{\left(\frac{1}{1+\beta}\right)}$. In terms of the corresponding Lotka function this becomes: $a = \left(\frac{2(2\alpha-3)}{3(\alpha-2)}\right)^{\left(\frac{(\alpha-1)}{\alpha}\right)} T^{\left(\frac{1}{\alpha}\right)}$.

Proof. We know that the minimum is reached for $f_{0.75}$. As $f_\theta = \left(\frac{B}{2\theta} \left(\frac{2-\beta}{1-\beta}\right)\right)^{\left(\frac{1}{(\beta+1)}\right)}$ this leads to $a = \left(\frac{2B}{3} \left(\frac{2-\beta}{1-\beta}\right)\right)^{\left(\frac{1}{(\beta+1)}\right)}$ or, in terms of the corresponding Lotkaian parameters: $f_{0.75} = \left(\frac{2}{3} \left(\frac{2\alpha-3}{\alpha-2}\right)\right)^{\left(\frac{(\alpha-1)}{\alpha}\right)} T^{1/\alpha}$.

Corollary 4. In this case it is possible to express the minimum a-value as a g_θ -value with known Lotka-parameters, namely $a = g_\theta$ with $\theta = \frac{3(\alpha-1)}{2(2\alpha-3)} = \frac{3}{2(2-\beta)}$.

Proof. From the previous example, we know that in the Lotkaian case

$$g_\theta = \left(\frac{\alpha-1}{\theta(\alpha-2)} \right)^{((\alpha-1)/\alpha)} T^{(1/\alpha)}. \text{ Hence, the obtained minimum } a\text{-value,}$$

$$\text{i.e. } \left(\frac{2}{3} \left(\frac{2\alpha-3}{\alpha-2} \right) \right)^{((\alpha-1)/\alpha)} T^{1/\alpha}, \text{ corresponds to } \theta = \frac{3(\alpha-1)}{2(2\alpha-3)} = \frac{3}{2(2-\beta)}.$$

6. Some remarks

6.1 Another minimum problem

Assume that, instead of functions φ_a we use functions $\psi_a(r) = a$, $r \in [0, T]$ and want to find the minimum value of $d^2(Z, \psi_a) = \int_0^T (Z(r) - \psi_a(r))^2 dr$, what would then be the outcome? This is easy to answer as shown in the following proposition.

Proposition 8

$$d^2(Z, \psi_a) \text{ reaches a minimum for } a = \mu = \frac{\int_0^T Z(r) dr}{T}$$

$$\text{Proof. } F^*(a) = d^2(Z, \psi_a) = \int_0^T (Z(r) - \psi_a(r))^2 dr = \int_0^T (Z(r) - a)^2 dr = \int_0^T (Z(r))^2 dr - 2a \int_0^T Z(r) dr + a^2 T.$$

Now, $(F^*)'(a) = -2 \int_0^T Z(r) dr + 2aT$. This expression is negative, pointing to a decreasing function, iff $a \leq \frac{\int_0^T Z(r) dr}{T} = \mu$. If $a = \mu$, $(F^*)'(a) = 0$. As $(F^*)''(a) = 2T > 0$ this shows that $a = \mu$ is a minimum.

6.2 A remark related to the sum of two functions

Let $Z_1(r)$ and $Z_2(r)$ be positive, decreasing and differentiable functions defined on the interval $[0, T]$, and let $\theta > 0$. Assume further that $Z_j(T) \leq \theta T$, $j=1,2$, and that the corresponding integral functions satisfy: $Y_j(T) \leq \theta T^2$, $j= 1,2$, then

$$\text{a) } h_{2\theta}(Z_1+Z_2) \geq \min(h_\theta(Z_1), h_\theta(Z_2))$$

$$\text{b) } g_{2\theta}(Z_1+Z_2) \geq \min(g_\theta(Z_1), g_\theta(Z_2))$$

Proof. a) We know that $Z_1(h_\theta(Z_1)) = \theta h_\theta(Z_1)$ and similarly: $Z_2(h_\theta(Z_2)) = \theta h_\theta(Z_2)$. If now $x = \min(h_\theta(Z_1), h_\theta(Z_2))$ then $Z_1(x) \geq \theta x$ and $Z_2(x) \geq \theta x$. Consequently: $(Z_1+Z_2)(x) \geq 2\theta x$. This shows that $x = \min(h_\theta(Z_1), h_\theta(Z_2)) \leq h_{2\theta}(Z_1+Z_2)$.

b) We know that $Z_1(g_\theta(Z_1)) = \theta (g_\theta(Z_1))^2$ and similarly: $Z_2(g_\theta(Z_2)) = \theta (g_\theta(Z_2))^2$. If now $y = \min(g_\theta(Z_1), g_\theta(Z_2))$ then $Y_1 \geq \theta y^2$ and $Y_2 \geq \theta y^2$. Consequently: $(Y_1 + Y_2)(y) \geq 2\theta y^2$. This shows that $y = \min(g_\theta(Z_1), g_\theta(Z_2)) \leq g_{2\theta}(Z_1+Z_2)$.

6.3 A remark on a multiple of a function

Let $Z(r)$ be a positive, decreasing and differentiable function defined on the interval $[0, T]$, and let $\theta > 0$. Assume further that c is a strict positive constant and assume that $Z(T) \leq \theta T$ (hence also $cZ(T) \leq c\theta T$) and that the corresponding integral function satisfies: $Y(T) \leq \theta T^2$, (hence also $cY(T) \leq c\theta T^2$) then

$$a) h_{c\theta}(cZ) = h_\theta(Z)$$

$$b) g_{c\theta}(cZ) = g_\theta(Z)$$

Proof. a) We know that $Z(h_\theta(Z)) = \theta h_\theta(Z)$, hence $(cZ)(h_\theta(Z)) = (c\theta) h_\theta(Z)$. This shows that $h_{c\theta}(cZ) = h_\theta(Z)$.

b) We know that $Z(g_\theta(Z)) = \theta (g_\theta(Z))^2$, hence $(cZ)(g_\theta(Z)) = (c\theta)(g_\theta(Z))^2$. This shows that $g_{c\theta}(cZ) = g_\theta(Z)$.

6.4 In the Lotkaian context, with $\alpha > 2$, the standard h-index is a special case of g_θ

We know that $g_\theta = \left(\frac{\alpha-1}{\theta(\alpha-2)}\right)^{(\alpha-1/\alpha)} T^{(1/\alpha)}$. As $h = T^{1/\alpha}$ we see, by taking $\theta = \frac{\alpha-1}{\alpha-2}$ (> 0 if $\alpha > 2$), that $g_\theta = h$.

7. Conclusion

Considering an $L^2[0, T]$ minimum problem led us to the study of a set of generalizations of the continuous h- and g-indices. This illustrates the use of the h- and g-indices and their generalizations in a context which bears no relation to the context in which these indices were originally introduced (Hirsch, 2005; Egghe, 2006a). As such this article shows the

usefulness of the h-index and g-index construction as mathematical constructions.

References

Egghe, L. (2005). *Power laws in the information production process: Lotkaian informetrics*. Amsterdam: Elsevier.

Egghe, L. (2006a). An improvement of the h-index: The g-index. *ISSI Newsletter*, 2(1), 8-9.

Egghe, L. (2006b). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.

Egghe, L., & Rousseau, R. (2006). An informetric model for the Hirsch-index. *Scientometrics*, 69(1), 121-129.

Gagolewski, M. (2009). A remark on limit properties of generalized h- and g-indices. *Journal of Informetrics*, 3(4), 367-368.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569-16572.

van Eck, N.J., & Waltman, L. (2008). Generalizing the h- and g-indices. *Journal of Informetrics*, 2(4), 263-271.