



COMENTARIO

¿Sigue la producción de artículos ISI de los ecólogos chilenos (sensu lato) la ley de Lotka (1926)?

Does the production of ISI papers by Chilean ecologists (sensu lato) fit Lotka's (1926) law?

JAIME R. RAU

Laboratorio de Ecología, Departamento de Ciencias Básicas & Programa IBAM, Universidad de Los Lagos, Casilla 933, Osorno, Chile
email: jrau@ulagos.cl

RESUMEN

Recientemente se propusieron dos nuevos indicadores bibliométricos aplicados a la productividad de artículos ISI para una muestra de 120 ecólogos chilenos (sensu lato). El propósito de esos índices fue realizar una corrección del índice-h tanto para el número de coautores como para el número de autocitas, y efectuar una estandarización para la edad científica. Sin embargo, un análisis de los residuos de la regresión del logaritmo decimal del número de alocitas (correlacionado con el índice-h) demostró un bajo efecto del número de coautores y autocitas y demostró que el examen de los residuales es un indicador bibliométrico más robusto. Pese a ello, utilizo esa base de datos y, a través de un análisis de regresión log-log y el empleo del programa computacional LOTKA, demuestro que los ecólogos chilenos no siguen la ley epistemométrica de Lotka (1926), la cual postula que es solo una minoría la que produce la mayoría de los artículos publicados en una determinada área disciplinaria.

Palabras clave: epistemometría, índice-h, ley cuadrática, productividad científica, programa computacional LOTKA.

ABSTRACT

Recently, two new bibliometric indicators were proposed to gauge the productivity of ISI papers and applied to a sample of 120 Chilean ecologists (sensu lato). The purpose of these indexes was to make a correction of the h-index for the number of coauthors and the number of self-citations as well as to a scientific age-standardization of that index. However, an analysis of residuals from the regression of the logarithm of the number of alocitations (correlated with the h-index) showed little effect of both, the number of coauthors and of self-citations and showed that the examination of the residuals is a more robust bibliometric indicator. Nevertheless, in this comment I use that database and, through a log-log regression analysis and the use of the LOTKA computer program, I show that Chilean ecologists do not follow the scientometric law of Lotka (1926), which postulates that only a minority of authors produces most of the articles published in a particular subject area.

Key words: h-index, LOTKA computer program, quadratic law, scientific productivity, scientometrics.

INTRODUCCIÓN

En líneas generales, la distribución del número de artículos publicados por n autores en un campo del conocimiento no es uniforme. Por el contrario, en la mayoría de los casos se ha observado que es solo en unos pocos autores donde se concentra el mayor volumen de la producción científica, mientras que la mayoría de ellos es la que genera el menor volumen de esa productividad. Tempranamente (Lotka 1926), se encontró que esa distribución se ajustaba a la ley potencial. Matemáticamente, se la denota como $A_n = A_1/n^2$, donde A_1 es el número de autores con un solo artículo

publicado. Verbalmente se la enuncia expresando que: "el número de autores, A_n , que publican n artículos científicos en un área disciplinaria es inversamente proporcional a su cuadrado". Esto significa, aproximadamente, que es solo el 5 % de los autores el que produce la mitad (50 %) de los artículos publicados (Bunge 2001). En epistemometría, a esta generalización se la denomina la ley cuadrática inversa de Lotka (Price 1963). Una implicancia de esta ley es que los autores más productivos corresponderán a la raíz cuadrada del número total de autores. Para corroborar si esta ley se cumple, se hace una regresión log-log entre el número de autores que ha

producido n publicaciones versus el número de publicaciones producidas por autor. Se considera que los datos siguen la ley si la pendiente es cercana a -2 (Lotka 1926) y el coeficiente de determinación cercano a un 100 % (Zar 1974). Alternativamente, los valores observados pueden compararse con los esperados de acuerdo a la ley potencial y obtenerse los dos parámetros (β y C) que la caracterizan (Rousseau & Rousseau 2000).

A pesar de su estatus, la ley de Lotka no es universal y es dependiente tanto del período de tiempo considerado como de la amplitud en la definición de la comunidad de autores (Potter 1981), los cuales no constituyen observaciones independientes entre sí. A mi modo de ver, una aplicación práctica del cumplimiento o no de esta ley, podría traducirse en la asignación de fondos gubernamentales para el desarrollo de un área científica. En Chile, la agencia gubernamental encargada de proveer fondos para centros avanzados de investigación ha privilegiado a la élite. Es decir, a los autores más productivos. Con esto, implícitamente, puede pensarse que las autoridades han supuesto que la ley de Lotka (1926) se cumple. Pero si la ley de Lotka (1926) no se cumpliera, la política gubernamental debiese ser la asignación de fondos equivalentes tanto a los grandes como a los pequeños productores (sensu Price 1963). Según lo planteado por Molina-Montenegro & Gianoli (2010), en Chile los primeros se concentrarían en la capital del país, mientras que los segundos desarrollarían sus líneas de investigación principalmente en las universidades regionales.

Recientemente se ha publicado, como Material Complementario del trabajo publicado por Molina-Montenegro & Gianoli (2010), una base de datos entregando el número de artículos ISI publicados a febrero de 2008 por 120 investigadores del área de las ciencias ecológicas (sensu lato), la cual cubre más de cuatro décadas (véase http://rchn.biologiachile.cl/suppmat/2010/2/MC_Molina-Montenegro_&_Gianoli_2010.pdf).

La intención de esos autores fue corregir el índice-h de Hirsch (2005, 2007) por el número de coautores y autocitas (su índice I) y estandarizarlo por la edad científica (su índice I_s). George-Nascimento (2010) criticó la base de datos por carecer de criterios explícitos

para incluir y excluir investigadores chilenos en ecología, y a los propios índices recién propuestos, demostrando que el índice I estuvo asociado con el logaritmo decimal del número de alocitas¹ y, por lo tanto, fue poco afectado por el número de coautores y autocitas. Además, él corroboró la marcada asimetría del índice I_s al disminuir con la edad científica y maximizar su valor solo en el caso de los investigadores jóvenes (i.e., hasta 40 años). Un aporte interesante de George-Nascimento (2010) fue demostrar la utilidad bibliométrica del análisis de residuos de la regresión logarítmica entre el número de alocitas y la edad científica².

Esa base de datos incluyó ecólogos que publicaron al menos 10 artículos ISI. Para el análisis que efectué, importé esos datos a una planilla Excel, ordené a los autores de acuerdo a su productividad creciente y los grafiqué en escala logarítmica (Fig. 1). En total, esos investigadores publicaron 4177 artículos, lo que es equivalente a una media aritmética de casi 35 artículos por autor. De acuerdo a la ley aquí analizada, se esperaría que 11 (5 %) de esos autores generase 2089 (50 %) de los artículos publicados. Sin embargo, ellos solo publicaron 1309 correspondientes al 31.3 %. Claramente, los datos no siguen la ley cuadrática inversa y, como se verifica en la Fig. 1, la pendiente es casi tres veces menor que -2 (= -0.67; la planilla Excel no entrega el error estándar de la pendiente), siendo el coeficiente de determinación solo un 51 % (i.e., un 49 % de la varianza de la regresión no es explicada por el modelo).

Para corroborar si los valores observados se ajustaron a los esperados por la ley, utilicé además el programa computacional LOTKA (ver Rousseau & Rousseau 2000). En este programa las variables de entrada se denominan “production” y “sources” y corresponden al número de artículos publicados

¹ De manera análoga, el logaritmo decimal del número de citas ha sido utilizado como una medida de solidez del impacto bibliométrico de un autor (cf. Price 1963).

² Coincidentemente, hay 61 residuales positivos en la Fig. 3 de George-Nascimento (2010) y ese es el mismo número que corresponde a los socios de la Sociedad de Ecología de Chile, Socecol (<http://www.socecol.cl>). Naturalmente, esto no significa necesariamente que los residuales positivos coincidan con la identidad de los socios.

y sus autores, respectivamente. La salida del programa entrega los parámetros β y C , el valor de la desviación máxima para el estadígrafo de Kolmogorov-Smirnov (K-S) y los niveles críticos de significación estadística. Si hay un buen ajuste a la distribución potencial, el parámetro β debe variar entre 1.27 y 3.29 y el valor calculado del estadígrafo K-S (D_{\max}) superar al tabulado para el nivel crítico deseado. En este último caso, usé un 5 % como nivel de confiabilidad estadística. Este programa se basa en el modelo discreto de la ley de Lotka (1926). No obstante, se ha demostrado que esta versión es también equivalente al modelo continuo de la misma ley (cf. Egghe 2005).

Como era esperable, para $P = 0.05$ encontré con el programa LOTKA que $\beta = 0$ (i.e., < 1.27) y $C = 1$ y un valor calculado para el estadígrafo $D_{\max} = 0.1193$ ($< 1 =$ valor tabulado). Concluyo entonces que una muestra de 120 ecólogos chilenos no sigue la ley de la productividad científica de Lotka (1926). Como señalé antes, la Fig. 3 del trabajo de George-Nascimento (2010) muestra en su análisis de residuos de la regresión log-log entre el número de aloctas y la edad científica, que al menos 61 de 120 (ca. 51 %) ecólogos chilenos (sensu lato) debiesen ser contratados, promovidos o premiados, si ello de verdad guardase alguna concordancia con su productividad científica.

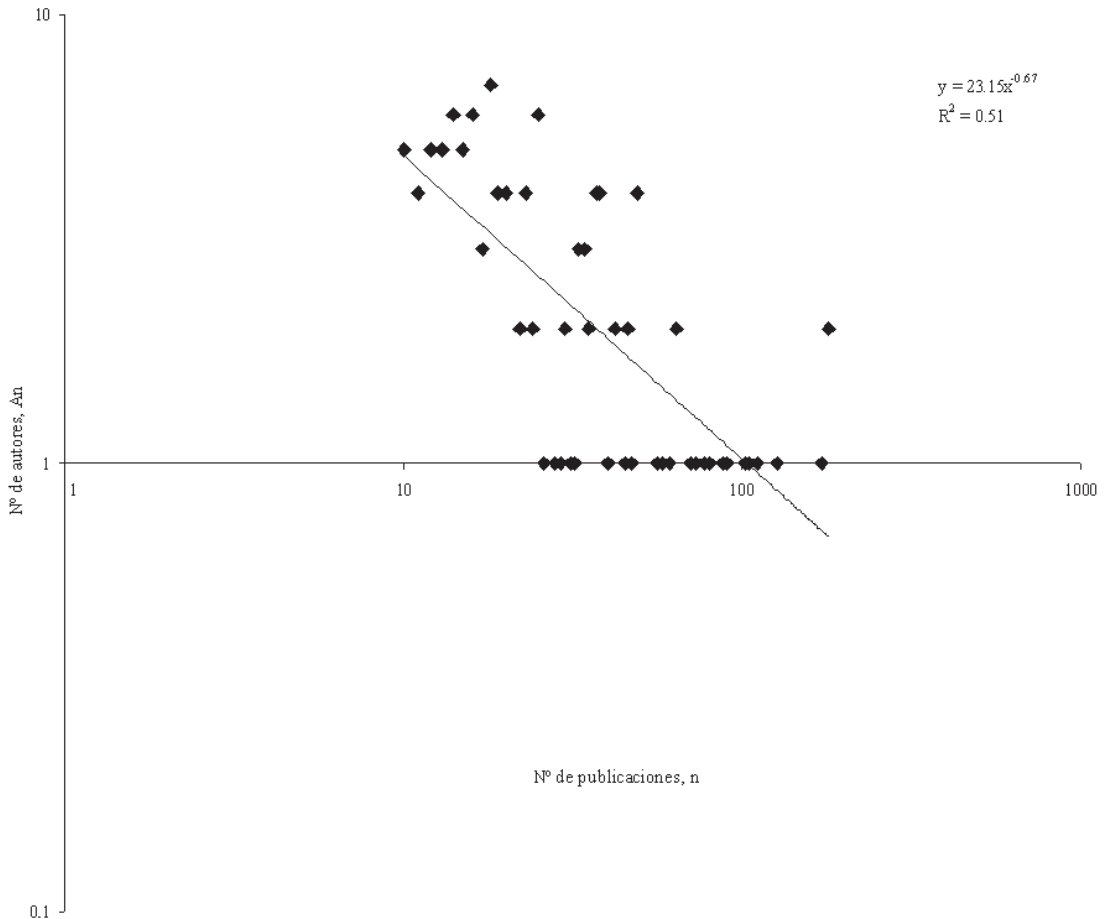


Fig. 1: Relación log-log entre el número de publicaciones ISI publicadas por autor (n) y el número de autores que ha producido n publicaciones ISI (A_n), en el caso de los ecólogos chilenos.

Log-log relation between the number of ISI papers published by author (n) and the number of authors that have produced n ISI papers (A_n), in the case of Chilean ecologists.

Por lo tanto, en mi opinión personal y para concluir este comentario, concluyo que los ecólogos chilenos son muy parejos en su productividad científica, independientemente de su localidad geográfica.

AGRADECIMIENTOS: Agradezco la ayuda de Mateo Rau C., en las diferentes fases necesarias para escribir este comentario y a la Universidad de Los Lagos por financiar sus costos de publicación.

LITERATURA CITADA

- BUNGE M (2001) El efecto San Mateo. *Polis* 1: 1-3.
- EGGHE L (2005) Relation between the continuous and the discrete Lotka power function. *Journal of the American Society for Information Science and Technology* 56: 664-668.
- GEORGE-NASCIMENTO M (2010) Una evaluación de los índices bibliométricos I e I_s de Molina-Montenegro & Gianoli aplicada a investigadores en ciencias ecológicas en Chile. *Revista Chilena de Historia Natural* 83: 229-235.
- HIRSCH JE (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* 104: 19193-19198.
- HIRSCH JE (2007) Does the h-index have predictive power? *Proceedings of the National Academy of Sciences* 49: 19193-19198.
- LOTKA AJ (1926) The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences* 16: 317-323.
- MOLINA-MONTENEGRO MA & E GIANOLI (2010) El índice-I, un nuevo estimador de impacto de la productividad científica: Los ecólogos de Chile como caso de estudio. *Revista Chilena de Historia Natural* 83: 219-227.
- POTTER WG (1981) Lotka's law revisited. *Library Trends* 31: 21-39.
- PRICE DJS (1963) *Little science, big science*. Columbia University Press. New York, USA.
- ROUSSEAU N & R ROUSSEAU (2000) LOTKA: A program to fit a power law distribution to observed frequency data. *Cybermetrics* 4. <http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html> (accedido Abril 10, 2011)
- ZAR JH (1974) *Biostatistical analysis*. Prentice-Hall, Inc. Englewood Cliffs, New Jersey, USA.

Editor Asociado: Mario George-Nascimento

Recibido el 6 de septiembre de 2010; aceptado el 12 de mayo de 2011