

MODELO MATEMÁTICO DE LEY DE POTENCIAS APLICADO AL MARATÓN

MATHEMATICAL MODEL OF POWER LAW APPLIED TO THE MARATHON

ANDRÉS FERNÁNDEZ REVELLES

Facultad de Ciencias del Deporte.
Universidad de Granada.

RESUMEN

Objetivo: Aplicar a una prueba de maratón, un modelo matemático de ley de potencias para la distribución de las marcas y comprobar su nivel de ajuste. Método: Aplicación de dos modelos al maratón femenino de Londres de 2010 en todas sus categorías, con las variables tiempo, modelo creciente y a la velocidad media modelo decreciente. Resultados: Las correlaciones obtenidas en todas las categorías han sido muy significativas mostrándose en el coeficiente de correlación ($r \geq 0,980$; $P < 0,000$) y en el coeficiente de determinación lineal ($R^2 \geq 0,9737$). Conclusiones: La aplicación de un modelo matemático de ley de potencias a la prueba de maratón puede ser útil y viable, y el ajuste de los datos al modelo ha sido bastante preciso.

Palabras claves: Ranking; Ley de Potencias; Marcas; Distribución Espacio – Tiempo.

ABSTRACT

Objective: To apply a marathon a mathematical model of power law for the distribution of records and check their level of fit. Method: Application of two models at London Women's Marathon 2010 in all categories, with the variable time, increasing pattern and decreasing the average speed pattern. Results: The correlates obtained in all categories have been highly significant regarding the correlation coefficient ($r \geq 0.980$, $P < 0.000$) and the linear coefficient of determination ($R^2 \geq 0.9737$). Conclusions: The application of a mathematical model of power law to the marathon can be useful and feasible, and the fit of the data to the model was fairly accurate.

Key words: Ranking; Power Law; Records, Distribution Space-Time

1. INTRODUCCIÓN

El punto de partida de este trabajo es el análisis de la prueba de maratón en función de las marcas en tiempo y velocidad obtenidas en una prueba. Para tal fin buscamos un modelo matemático que nos pueda facilitar su análisis y estudio.

Los científicos siempre han intentado interpretar la realidad de su alrededor a través de modelos matemáticos. Así, desde finales del siglo XX e inicio del siglo XXI han aumentado las investigaciones dedicadas a estudiar el comportamiento de sistemas y redes complejas de datos utilizando diferentes leyes de potencias para poder explicar y dar sentido a esos fenómenos (Albert & Barabasi, 2002; Barabasi & Albert, 1999; Watts & Strogatz, 1998).

En un gran número de casos al enfrentarse a datos reales, la tendencia del comportamiento al aplicarle la ley de potencias es válida solo para los valores de rangos intermedios, aparecen rupturas en los datos que pertenecen a los valores de los rangos de los extremos pertenecientes a las colas de distribución (Amaral, Scala, Barthelemy, & Stanley, 2000; Newman, 2005). De este modo proliferan numerosas investigaciones que se centran en el estudio de los valores de los extremos o colas de distribución (Hong, Ha, & Park, 2007; Laherrere & Sornette, 1998; Naumis & Cocho, 2007, 2008)

La ley de potencias es utilizada cuando los modelos más convencionales de análisis no dan una solución apropiada para el estudio de esos fenómenos como ocurre en la investigación de las redes sociales (Suvakov, Mitrovic, Gligorijevic, & Tadic, 2013), o en el análisis de las redes de trabajo (Albert & Barabasi, 2000, 2002; Amaral, et al., 2000; Barabasi & Albert, 1999; Suvakov, et al., 2013; Watts & Strogatz, 1998), sobre todo por la ingente cantidad de datos utilizadas en este tipo de estudios.

Ya sea por la gran cantidad de datos utilizados en los estudios o por su adaptación a estos fenómenos, la ley de potencias es ampliamente utilizada en diferentes áreas (Alvarez-Martinez, Martinez-Mekler, & Cocho, 2011; Laherrere & Sornette, 1998; Machado, Pinto, & Lopes, 2013; Martinez-Mekler et al., 2009; Newman, 2005).

En biología para el estudio de la secuencia de nado de algunos animales (Schmitt, Seuront, Hwang, Souissi, & Tseng, 2006) o el estudio de la de la cantidad de biomasa producida (Feng & Zhang, 2012).

En lengua y lingüística estudiando la frecuencia y distribución de determinadas palabras (Popescu, 2003), analizando la popularidad y frecuencia de los nombres puestos

a los recién nacidos (Li, 2012), o la frecuencia y distribución de letras y palabras en discursos de presidentes de gobierno (Li & Miramontes, 2011) Mexico City 04510, DF, Mexico.Li, WT (reprint author).

En música (del Rio, Cocho, & Naumis, 2008) relacionando la distribución de las notas en diversas tendencias musicales, rock, jazz, etc..., o el análisis de obras de Bach, Albeniz, etc..., con la interpretación de esas secuencias a nivel de las redes neuronales del cerebro y asociando los exponentes de la ley de potencias a la naturaleza armónica de la obra y otro componente de la ley al espectro acústico. En un tópico de investigación en el que la utilización de la ley de potencias es muy adecuada es en la descripción de la auto-organización de fenómenos críticos (Newman, 2005), tales distribuciones de fenómenos pueden encontrarse en organismos biológicos y en tecnologías avanzadas de ingeniería, llamando a ese sistema o mecanismo que cumple con esa distribución "Highly Optimized Tolerance" HOT, o sistemas con una Tolerancia Altamente Optimizada (TAO) (Alvarez-Ramirez & Rodríguez, 2006; Carlson & Doyle, 1999, 2000, 2002; Doyle & Carlson, 2000; Garcia-Manso, Martin-Gonzalez, Davila, & Arriaza, 2005; Hillers, Carlson, & Archuleta, 2009; Newman, 2005).

Sin embargo la línea de investigación en la que más producción científica se encuentra la ley de potencias es en estudios bibliométricos. Investigaciones en las que el método de trabajo se inicia jerarquizando los datos recogidos produciendo una distribución en forma de ranking o rango (Campanario, 2010a, 2010c, 2010d, 2011a, 2011b; Coile, 1977; Edwards & Collins, 2011; Egghe, 2010b, 2011b, 2012; Garfield, 1980).

El deporte y áreas afines no han quedado al margen en la utilización de la ley de potencias, así existen numerosos estudios que abordan los records deportivos mundiales y su análisis, para ello se buscan modelos matemáticos que se ajusten a las actuales marcas y así poder estudiar su posible evolución e incluso pronosticar o predecir los próximos records (Carbone & Savaglio, 2001; Katz & Katz, 1999; Savaglio & Carbone, 2000), aunque hay estudios que no utilizan leyes de potencias y también han estudiado la predicción de próximos records (Joyner, Ruiz, & Lucia, 2011).

Así encontramos estudios sobre gestión deportiva que utilizan leyes de potencias (Aidt, Leong, Saslaw, & Sgroi, 2006; Ichinomiya, 2006), pero aquellos estudios que nos han parecido mas relacionados con nuestro estudio son los que analizan la distribución jerárquica de las marcas o tiempos en diferentes carreras (Alvarez-Ramirez & Rodríguez, 2006; Alvarez-Ramirez, Rodríguez, & Dagdug, 2007), o los que estudian o intentan predecir el rendimiento deportivo (Garcia-Manso, Martin-Gonzalez, Vaamonde, & Da Silva-Grigoletto,

2012; Malacarne & Mendes, 2000; Nevill, Markovic, Vuce-
tic, & Holder, 2004; Nevill, Stewart, Olds, & Holder, 2004; E.
Wesfreid & Billat, 2009; Eva Wesfreid, Billat, & Meyer, 2005).

Pero no siempre se utiliza la misma ley de potencias, en cada caso se utiliza la más adecuada para explicar cada tópico de investigación. En nuestro caso que queremos analizar la distribución temporal de los tiempos o marcas en una prueba de maratón utilizando un modelo matemático basado en la ley de potencias. Creemos que se podría utilizar la ley de potencias de "rango Beta", ley que mejora la ley de Zipf (del Rio, et al., 2008). Ley también llamada "b-like function" (Naumis & Cocho, 2007, 2008) por tener dos parámetros, "S-shape" (Egghe, 2009a, 2011a, 2013) por su forma o "semi-log rank-order distribution" (Campanario, 2010d; Mansilla, Koppen, Cocho, & Miramontes, 2007) por ser una ley que al realizar una gráfica de dispersión de puntos con la distribución, el eje x tiene escala lineal y el eje y tiene una escala logarítmica.

Finalmente los objetivos de nuestro estudio son los siguientes:

- Objetivo 1: Aplicar a una prueba de maratón un modelo matemático de ley de potencias para la distribución de las marcas.
- Objetivo 2: Comprobar el nivel de ajuste del modelo matemático a las marcas del maratón.

2. MÉTODO

2.1. Muestra

La muestra está formada por todas las participantes que finalizaron el maratón

La muestra está formada por todas las participantes que finalizaron el maratón femenino de Londres de 2010, y estos datos se obtuvieron en enero de 2012 de Marathon Guide (MarathonGuide, 2000). Fue seleccionada esta prueba por acumular en su historia las 200 mejores marcas realizadas en la prueba de maratón femenino en los últimos 30 años, el 21% lo que supone 42 marcas realizadas en la ciudad de Londres (Association of Road Racing Statisticians, 2012).

Las categorías de edad en las que se divide el maratón femenino de Londres son 18 a 39 años, 40 a 44 años, 45 a 49 años, 50 a 54 años, 55 a 59 años, 60 a 64 años, 65 a 69 años y 70 ó mas años.

Para nuestro estudio se han utilizado todas las categorías con los datos de las participantes que finalizaron la prueba, siendo estos 12136 participantes que obtuvieron al finalizar la prueba 7492 tiempos o marcas diferentes.

En el análisis global se incluyeron todas las categorías, pero para el análisis pormenorizado por categorías se excluyeron a partir de los 60 años por el reducido número de participantes.

2.2. Análisis de los datos

El procesamiento de los datos comenzó con la tabulación de los datos descargados en un libro de Microsoft Excel 2007, en el que se obtuvo para cada marca la correspondencia en segundos y en km/h tanto a mitad de prueba como al finalizar la prueba.

Para ordenar las marcas realizadas de una forma jerárquica se utilizó una tabla dinámica con Microsoft Excel 2007 y se estableció un ranking o rango con las marcas ordenadas de mejor marca a peor marca, es decir de menor tiempo a mayor tiempo, siguiendo la metodología ampliamente utilizada en estudios de distribución de rankings (Campanario, 2010c).

Con los datos estructurados en variables y utilizando el programa estadístico SPSS (Versión 18.0, Chigago, IL, USA) se establecieron las diferentes variables. Estas variables fueron la variable orden o ranking que establece el orden jerárquico de las marcas y las variables para la marca tiempo en segundos, variable velocidad en km/h, repitiendo este proceso para el estudio global y para cada una de las categorías estudiadas.

Se realizó una exploración gráfica con una dispersión de puntos que relacionaba el ranking con las marcas de velocidad media al final de la prueba y con las marcas en tiempo al final de la prueba, con el programa SPSS (Versión 18.0, Chicago, IL, USA).

Seguidamente, se utilizó la ecuación de ley de potencias (ver Ecuación 1), que se adapta mejor a este tipo de curvas (del Rio, et al., 2008; Mansilla, et al., 2007) puesto que los parámetros a y b están asociados al ajuste de las colas de la izquierda y de la derecha respectivamente:

$$f(r) = K \frac{(N + 1 - r)^b}{r^a}$$

Ecuación 1. Ley de potencias para curvas con distribuciones crecientes.

Esta ecuación es una mejora de la ecuación de ley de potencias utilizada para el análisis del factor de impacto de las revistas (Lavalette, 1996; Popescu, 2003). Además esta

ecuación ha sido ampliamente utilizada para curvas con valores decrecientes (Alvarez-Martinez, et al., 2011; Campanario, 2010a, 2010b, 2010c, 2010d; del Rio, et al., 2008; Egghe, 2009a, 2009b, 2011a; Mansilla, et al., 2007; Naumis & Cocho, 2007, 2008; Waltman & van Eck, 2009).

Sin embargo no conocemos publicaciones en las que aparezca la curva con los datos crecientes, para ello transformamos la ecuación para que responda a nuestras necesidades quedando de esta forma (Ecuación 2):

$$f(r) = K \frac{r^a}{(N + 1 - r)^b}$$

Ecuación 2. Ley de potencias para curvas con distribuciones de crecientes.

Como se puede observar los elementos y parámetros de ambas ecuaciones de la ley de potencias que conforman esta función lineal son los mismos pero con diferente orden siendo:

f : la función

r : el ranking o posición en el ranking

$f(r)$: la función en función del ranking o posición en el ranking

K : una constante

N : el número de datos de la muestra

a y b : son los parámetros de la función lineal.

Ahora procedemos al cálculo del modelo de curva a partir de los datos obtenidos en el maratón. Para ello seguimos el método descrito por Juan Miguel Campanario (2010a, 2010c, 2010d), así realizamos los cálculos previos en Microsoft Excel 2007 transcribiendo la ley de potencias a ecuaciones en Microsoft Excel 2007. Con los cálculos previos realizados, como hay que calcular la constante K y los parámetros a y b de la ecuación acudimos a un servicio web gratuito de cálculo masivo llamado "zunzun" (Phillips, 2010), y nos proporciona los parámetros citados anteriormente, y el coeficiente de determinación lineal R^2 .

Con el valor de estos parámetros obtenido, aplicamos la ley de potencias a cada dato del ranking para obtener el valor de la función en cada punto, lo hacemos con Microsoft Excel 2007, y así nos resultan los valores de nuestro modelo matemático. Con el fin de tener datos sobre el grado de relación entre los datos obtenidos directamente del mara-

tón y los datos de nuestro modelo, además de contar con el coeficiente de determinación lineal R^2 calculado anteriormente, calculamos la correlación de Pearson entre datos de maratón y modelo calculado.

Finalmente se realiza una gráfica de dispersión de puntos con datos originales y modelo calculado para observar visualmente la similitud entre ambos, apareciendo en el eje x el ranking y en el eje y la velocidad o el tiempo de la marca, ya sea del dato del maratón o del modelo. Con el fin hacer una gráfica que aparezca más extendida sobre el eje y , se le aplica una escala logarítmica tal y como indican en el método diversos autores (Alvarez-Martinez, et al., 2011; Campanario, 2010a, 2010b, 2010c, 2010d; del Rio, et al., 2008; Egghe, 2009a, 2009b, 2011a; Mansilla, et al., 2007; Naumis & Cocho, 2007, 2008; Waltman & van Eck, 2009).

Este proceso es repetido para el estudio completo y pormenorizado por categorías de las marcas en tiempo y velocidad media, utilizando la ley de potencias adecuada en cada caso.

2.3. Resultados

2.3.1. Resultados preliminares

Los resultados de la exploración gráfica atendiendo a la velocidad media y a las marcas en tiempo al final de la prueba, se presentan respectivamente en las figura 1 y 2. Analizando la curva observamos que la figura 1 nos muestra una curva con 2 colas en los datos extremos, quedando a la izquierda y mas cercana al eje de ordenadas la cola con los datos mas altos en el eje y , y quedando a la derecha la cola con los datos mas bajos en el eje y . La figura 2 nos muestra una curva con 2 colas en los datos extremos, quedando a la izquierda y mas cercana al eje de ordenadas la cola con los datos mas bajos en el eje y , y quedando a la derecha la cola con los datos mas altos en el eje y .

Por tanto las colas muestran cierta simetría, esto es debido a que utilizando el ranking de orden de llegada, las mejores clasificadas son las que han obtenido una mayor velocidad y han invertido menor tiempo. Así de este modo la gráfica de la distribución de las velocidades es decreciente, y la gráfica de la distribución temporal de las marcas en función del ranking es creciente.

Objetivo 1: Aplicación del modelo matemático a la prueba de maratón

En nuestro estudio las características de la distribución de las marcas en función del ranking han provocado que tengamos que aplicar dos modelos matemáticos. Uno para los rankings cuyos valores son decrecientes como ocurre

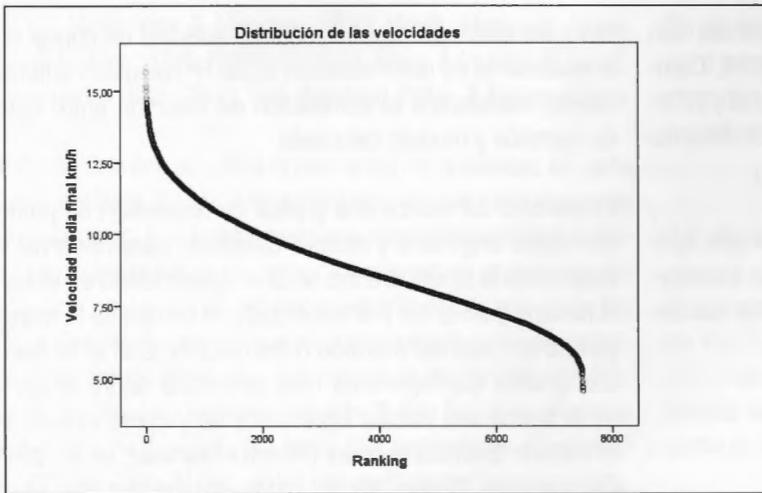


Figura 1. Distribución de velocidades en función del ranking, utilizando todas las categorías de edad del maratón femenino de Londres 2010.

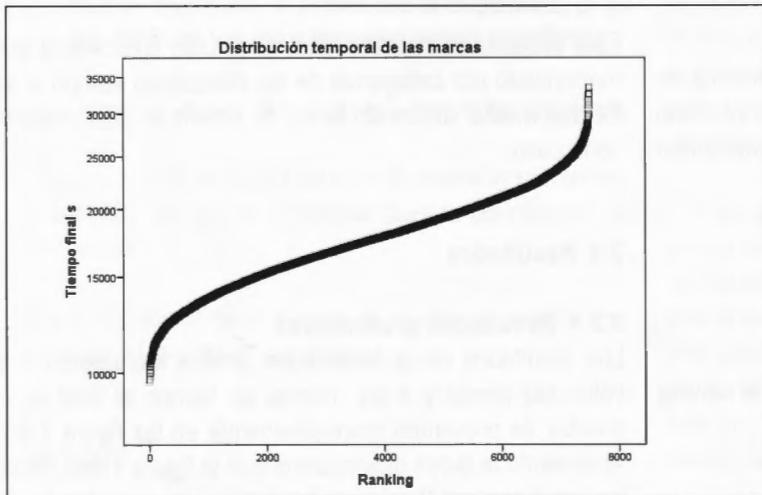


Figura 2. Distribución temporal de las marcas en función del ranking, utilizando todas las categorías de edad del maratón femenino de Londres 2010.

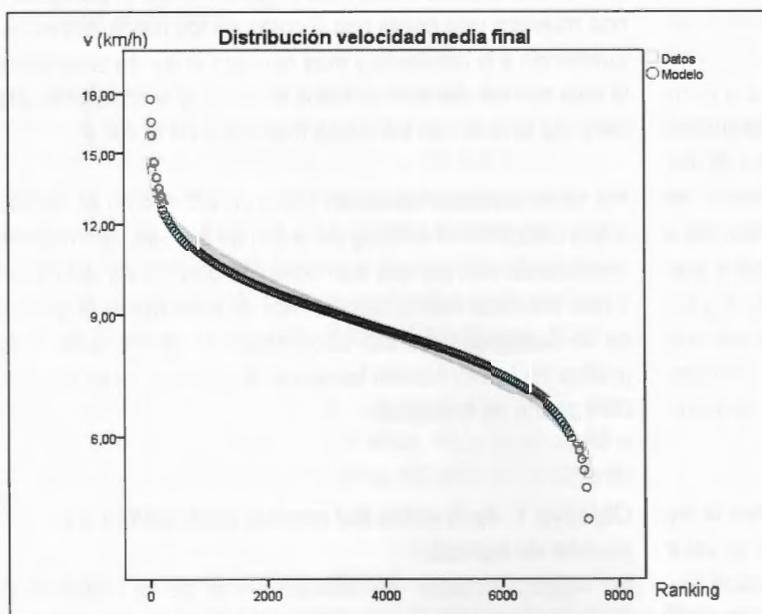


Figura 3. Distribución temporal de la velocidad media final en todas las categorías

en el caso de las marcas de velocidad media y otro modelo para los rankings cuyos valores son crecientes como ocurre en el caso de las marcas temporales.

Con la aplicación del modelo matemático para valores decrecientes, aplicado a los datos de la velocidad media (Ecuación 1), hemos obtenido (Tabla 1 y Figura 3) valores negativos para el parámetro a de la ecuación y valores positivos para el parámetro b . Siendo los valores absolutos de ambos parámetros muy similares, casi sin diferencias, encontramos que para el grupo llamado "Todas" y "Todas*" en el que se incluyen todas las categorías, y en el grupo de 18 a 39 años, el valor absoluto de los parámetros es $a > b$. Ocurriendo al contrario en las otras categorías con el valor absoluto de los parámetros $a < b$.

De este modo con la aplicación del modelo matemático para valores crecientes, aplicado a las marcas temporales (ver Ecuación 2), hemos obtenido (Tabla 2 y Figura 4) valores positivos para el parámetro a y negativos para el b . Igualmente en este caso hemos obtenido valores absolutos de ambos parámetros similares. En este caso el valor absoluto de los parámetros es $a > b$, en el grupo de todas las categorías unidas y en el grupo de 18 a 39, y en las demás categorías ocurre al contrario valor absoluto de los parámetros es $b > a$.

Objetivo 2: Comprobación del ajuste del modelo matemático a la prueba de maratón

Los correlaciones obtenidas en todas las categorías, tanto en la velocidad media a la que se ha aplicado el modelo matemático para valores decrecientes, y en las marcas temporales a las que se ha aplicado el modelo matemático para valores crecientes son muy significativas con ($r \geq 0,980$; $P < 0,000$), y con un coeficiente de determinación lineal que indica un ajuste muy preciso ($R^2 \geq 0,9737$) (Tabla 3 y Tabla 4).

3. DISCUSIÓN Y CONCLUSIONES

El principal resultado es que el modelo matemático de la ley de potencias, en ambas variantes, es aplicable, puesto que el nivel de ajuste en nuestro estudio (en cualquiera de las categorías $R^2 \geq 0,9737$) es similar a investigaciones precedentes con otros tópicos de investigación.

Similares resultados a los nuestros se encuentran en el estudio sobre la universalidad de las distribuciones de rango con un coeficiente de determinación lineal ($R^2 \geq 0,978$) (Naumis & Cocho, 2007), en estudios del lenguaje ($R^2 \geq 0,964$) (del Río, et al., 2008), y en estudios bibliométricos sobre el cambio en el factor de impacto de las revistas ($R^2 \geq 0,9460$) (Campanario, 2010a). Aunque en nuestro caso el coeficiente de determinación lineal mayor encontrado ha sido en la categoría de 50 a 54 años tanto en velocidad media como en marcas temporales ($R^2 = 0,9894$), no se ha llegado en ningún caso a resultados como los descritos que llegan a ($R^2 \geq 0,99$) (Campanario, 2010c).

La interpretación y manipulación de los parámetros de la ecuación es lo que nos va a servir para poder sacar rendimiento al modelo matemático en el estudio de la prueba del maratón. Este modelo matemático expresado de forma gráfica nos muestra una zona central y una curva a cada lado, inicio y final, dependiendo la cola de la izquierda al parámetro a y la cola de la derecha al parámetro b (Mansilla, et al., 2007). La igualdad o similitud del valor absoluto en los parámetros a y b proporciona unas colas con angulación muy similares (Campanario, 2010c), lo que significa en nuestro caso una acumulación de marcas muy similar en ambas colas puesto que la diferencia en valor absoluto entre ambos parámetros es siempre inferior a 0,013.

En el caso haber obtenido un parámetro a en valor absoluto mucho menor que b la cola del lado izquierdo sería mas corta que la de lado derecho como ocurre en el análisis del Cuarteto Opus 131 de Beethoven ($a=0,20$ y $b=1,81$) (del Río, et al., 2008), o viceversa como ocurre en el estudio sobre población en municipios españoles ($a=0,86$ y $b=0,40$) (Alvarez-Martínez, et al., 2011). Así aplicado a nuestro estudio el que la cola izquierda, el parámetro a , fuese mucho mayor que el parámetro b significaría una mayor acumulación de mejores marcas que están muy por encima de la media de la prueba.

Con sólo la visualización de los valores de los parámetros de la ecuación es muy difícil realizar una interpretación de los datos. Sin embargo, en nuestros resultados, llama la atención las colas y la marca en la que se produce un

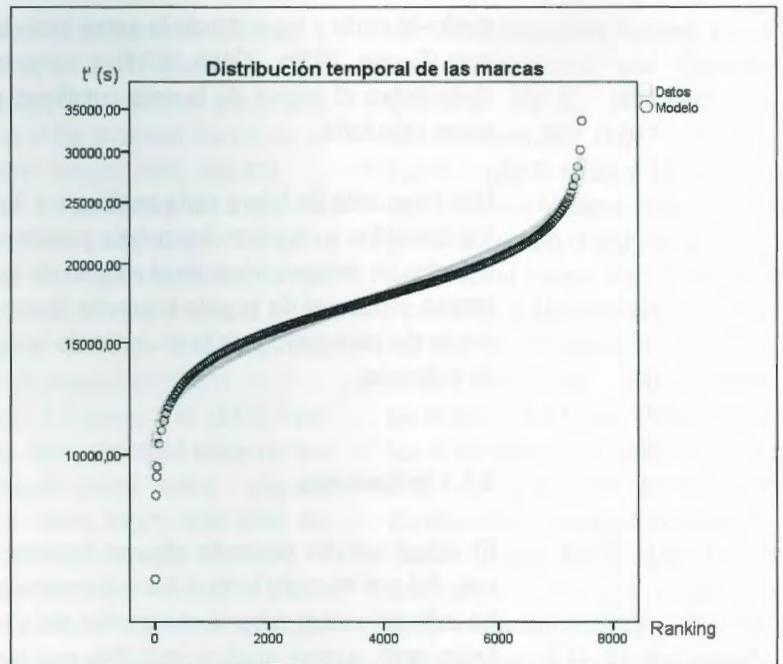


Figura 4. Distribución temporal de las marcas en todas las categorías

Tabla 1.
Parámetros obtenidos con el modelo matemático para valores decrecientes, en este caso para la velocidad media.

Categorías	Magnitud	N	K	b	a	R ²
Todas	v(s)	7492	8,8236	0,1118	-0,1136	0,9804
Todas*	v(s)	4804	9,5362	0,1122	-0,1167	0,9737
18-39 a	v(s)	5413	8,9056	0,1074	-0,1100	0,9853
40-44 a	v(s)	1801	8,3422	0,1076	-0,0976	0,9885
45-49 a	v(s)	1298	8,1744	0,1100	-0,0962	0,9894
50-54 a	v(s)	683	8,0730	0,1091	-0,0962	0,9827
55-59 a	v(s)	331	8,3217	0,1102	-0,1082	0,9818

*Datos calculados a partir de la velocidad media a mitad del maratón.

Tabla 2.
Parámetros obtenidos con el modelo matemático para valores decrecientes, en este caso para las marcas temporales.

Categorías	Magnitud	N	K	b	a	R ²
Todas	t(s)	7492	17215,47	-0,1118	0,1136	0,9804
18-39 a	t(s)	5413	17056,95	-0,1074	0,1100	0,9853
40-44 a	t(s)	1801	18208,89	-0,1076	0,0976	0,9885
45-49 a	t(s)	1298	18582,63	-0,1100	0,0962	0,9894
50-54 a	t(s)	683	18816,08	-0,1091	0,0962	0,9827
55-59 a	t(s)	331	18253,68	-0,1102	0,1082	0,9818

punto de corte y lugar donde la curva inicia la cola (Egghe, 2009a, 2010a, 2013), y surge la duda sobre el origen de la cola curvilíneo y como calcularlo.

Una propuesta de futuro sería analizar los datos obtenidos y visualizar las curvas; próximos estudios se deben centrar en el estudio de los rangos y marcas de la cola izquierda dependiente del parámetro a de la ecuación de la ley de potencias.

3.1. Limitaciones

El actual estudio presenta algunas limitaciones. Así por ejemplo la muestra seleccionada ha sido adecuada para la realización del estudio, pero sugiere realizar estudios con las marcas de maratón masculinas, y realizar el estudio con otros maratones. La falta de investigaciones precedentes que hayan utilizado

la misma ley de potencias para el análisis de pruebas de maratón no enriquece la discusión. Estas limitaciones nos incitan a seguir investigando con este método para generar más resultados que nos puedan llevar a conclusiones más relevantes.

3.2. Conclusiones

La aplicación de un modelo matemático de ley de potencias a la prueba de maratón puede ser útil y viable, y el ajuste de los datos al modelo ha sido bastante preciso.

4. REFERENCIAS BIBLIOGRÁFICAS

- Aidt, T. S., Leong, B., Saslaw, W. C., & Sgroi, D. (2006). A power-law distribution for tenure lengths of sports managers. *Physica a-Statistical Mechanics and Its Applications*, 370(2), 697-703. doi: 10.1016/j.physa.2006.03.016
- Albert, R., & Barabasi, A. L. (2000). Topology of evolving networks: Local events and universality. *Physical Review Letters*, 85(24), 5234-5237.
- Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. [Review]. *Reviews of Modern Physics*, 74(1), 47-97.
- Alvarez-Martinez, R., Martinez-Mekler, G., & Cocho, G. (2011). Order-disorder transition in conflicting dynamics leading to rank-frequency generalized beta distributions. *Physica a-Statistical Mechanics and Its Applications*, 390(1), 120-130. doi: 10.1016/j.physa.2010.07.037
- Alvarez-Ramirez, J., & Rodriguez, E. (2006). Scaling properties of marathon races. *Physica a-Statistical Mechanics and Its Applications*, 365(2), 509-520. doi: 10.1016/j.physa.2005.09.066
- Alvarez-Ramirez, J., Rodriguez, E., & Dagdug, L. (2007). Time-correlations in marathon arrival sequences. *Physica a-Statistical Mechanics and Its Applications*, 380, 447-454. doi: 10.1016/j.physa.2007.03.008
- Amaral, L. A. N., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), 11149-11152.
- Association of Road Racing Statisticians. (2012). ARRS Retrieved 1-02-2012, from <http://www.rrs.net/>
- Barabasi, A. L., & Albert, R. (1999). Emergence of

Tabla 3.
Correlaciones entre datos con las velocidades medias y modelo matemático

Categorías	N	r	P	R ²
Todas	7492	0,985*	0,000	0,9804
Todas*	4804	0,980*	0,000	0,9737
18-39 a	5413	0,988*	0,000	0,9853
40-44 a	1801	0,992*	0,000	0,9885
45-49 a	1298	0,993*	0,000	0,9894
50-54 a	683	0,988*	0,000	0,9827
55-59 a	331	0,990*	0,000	0,9818

*Correlación significativa a nivel $p < 0,001$
 *Datos calculados a partir de la velocidad media a mitad del maratón.

Tabla 4.
Correlaciones entre datos con las marcas temporales y modelo matemático

Categorías	N	r	P	R ²
Todas	7492	0,987*	0,000	0,9804
18-39 a	5413	0,991*	0,000	0,9853
40-44 a	1801	0,992*	0,000	0,9885
45-49 a	1298	0,993*	0,000	0,9894
50-54 a	683	0,990*	0,000	0,9827
55-59 a	331	0,988*	0,000	0,9818

*Correlación significativa a nivel $p < 0,001$

- scaling in random networks. *Science*, 286(5439), 509-512.
- Campanario, J. M. (2010a). Distribution of changes in impact factors over time. *Scientometrics*, 84(1), 35-42. doi: 10.1007/s11192-009-0094-y
 - Campanario, J. M. (2010b). Distribution of ranks of β -decay half-lives [physics.gen-ph]. *arXiv:1011.5390v1*.
 - Campanario, J. M. (2010c). Distribution of Ranks of Articles and Citations in Journals. *Journal of the American Society for Information Science and Technology*, 61(2), 419-423. doi: 10.1002/asi.21238
 - Campanario, J. M. (2010d). Self-Citations That Contribute to the Journal Impact Factor: An Investment-Benefit-Yield Analysis. *Journal of the American Society for Information Science and Technology*, 61(12), 2575-2580. doi: 10.1002/asi.21439
 - Campanario, J. M. (2011a). Empirical study of journal impact factors obtained using the classical two-year citation window versus a five-year citation window. *Scientometrics*, 87(1), 189-204. doi: 10.1007/s11192-010-0334-1
 - Campanario, J. M. (2011b). Large Increases and Decreases in Journal Impact Factors in Only One Year: The Effect of Journal Self-Citations. *Journal of the American Society for Information Science and Technology*, 62(2), 230-235. doi: 10.1002/asi.21457
 - Carbone, V., & Savaglio, S. (2001). Scaling laws and forecasting in athletic world records. *Journal of Sports Sciences*, 19(7), 477-484.
 - Carlson, J. M., & Doyle, J. (1999). Highly optimized tolerance: A mechanism for power laws in designed systems. *Physical Review E*, 60(2), 1412-1427. doi: 10.1103/PhysRevE.60.1412
 - Carlson, J. M., & Doyle, J. (2000). Highly optimized tolerance: Robustness and design in complex systems. *Physical Review Letters*, 84(11), 2529-2532. doi: 10.1103/PhysRevLett.84.2529
 - Carlson, J. M., & Doyle, J. (2002). Complexity and robustness. [Article; Proceedings Paper]. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 2538-2545. doi: 10.1073/pnas.012582499
 - Coile, R. C. (1977). Lotka frequency-distribution of scientific productivity. *Journal of the American Society for Information Science*, 28(6), 366-370.
 - del Rio, M. B., Cocho, G., & Naumis, G. G. (2008). Universality in the tail of musical note rank distribution. *Physica a-Statistical Mechanics and Its Applications*, 387(22), 5552-5560. doi: 10.1016/j.physa.2008.05.031
 - Doyle, J., & Carlson, J. M. (2000). Power laws, highly optimized tolerance, and generalized source coding. *Physical Review Letters*, 84(24), 5656-5659. doi: 10.1103/PhysRevLett.84.5656
 - Edwards, R., & Collins, L. (2011). Lexical Frequency Profiles and Zipf's Law. *Language Learning*, 61(1), 1-30. doi: 10.1111/j.1467-9922.2010.00616.x
 - Egghe, L. (2009a). Mathematical derivation of the impact factor distribution. *Journal of Informetrics*, 3(4), 290-295. doi: 10.1016/j.joi.2009.01.004
 - Egghe, L. (2009b). A Rationale for the Hirsch-Index Rank-Order Distribution and a Comparison With the Impact Factor Rank-Order Distribution. *Journal of the American Society for Information Science and Technology*, 60(10), 2142-2144. doi: 10.1002/asi.21121
 - Egghe, L. (2010a). The distribution of the uncitedness factor and its functional relation with the impact factor. *Scientometrics*, 83(3), 689-695. doi: 10.1007/s11192-009-0130-y
 - Egghe, L. (2010b). A New Short Proof of Naranan's Theorem, Explaining Lotka's Law and Zipf's Law. *Journal of the American Society for Information Science and Technology*, 61(12), 2581-2583. doi: 10.1002/asi.21431
 - Egghe, L. (2011a). The impact factor rank-order distribution revisited. *Scientometrics*, 87(3), 683-685. doi: 10.1007/s11192-011-0338-5
 - Egghe, L. (2011b). Mathematical relations of the h-index with other impact measures in a Lotkian framework. *Mathematical and Computer Modelling*, 53(5-6), 610-616. doi: 10.1016/j.mcm.2010.09.012
 - Egghe, L. (2012). Study of rank- and size-frequency functions and their relations in a generalized Naranan framework. *Mathematical and Computer Modelling*, 55(7-8), 1898-1903. doi: 10.1016/j.mcm.2011.11.047
 - Egghe, L. (2013). Study of the rank- and size-frequency functions in the case of power law growth of sources and items and proof of Heaps' law. *Information Processing & Management*, 49(1), 99-107. doi: 10.1016/j.ipm.2012.02.004
 - Feng, Y. J., & Zhang, J. (2012). Universal Biomass and Energy Flow Distribution in Weighted Food Webs. In Z. Yang & B. Chen (Eds.), *18th Biennial Isem Conference on Ecological Modelling for Global Change and Coupled Human and Natural System* (Vol. 13, pp. 818-828). Amsterdam: Elsevier Science Bv.
 - Garcia-Manso, J. M., Martin-Gonzalez, J. M., Davila, N., & Arriaza, E. (2005). Middle and long distance athletics races viewed from the perspective of complexity. *Journal of Theoretical Biology*, 233(2), 191-198. doi: 10.1016/j.jtbi.2004.10.014
 - Garcia-Manso, J. M., Martin-Gonzalez, J. M., Vaamonde, D., & Da Silva-Grigoletto, M. E. (2012). The limitations of scaling laws in the prediction of performance in endurance events. *Journal of Theoretical Biology*, 300, 324-329. doi: 10.1016/j.jtbi.2012.01.028
 - Garfield, E. (1980). Bradford law and related statistical patterns. *Current Contents*(19), 5-12.
 - Hillers, G., Carlson, J. M., & Archuleta, R. J. (2009). Seismicity in a model governed by competing frictional weakening and healing mechanisms. [Review]. *Geophysical Journal International*, 178(3), 1363-1383. doi: 10.1111/j.1365-246X.2009.04217.x
 - Hong, H. S., Ha, M., & Park, H. (2007). Finite-size scaling in complex networks. *Physical Review Letters*, 98(25). doi: 25870110.1103/PhysRevLett.98.258701
 - Ichinomiya, T. (2006). Power-law distribution in Japanese racetrack betting. *Physica a-Statistical Mechanics and Its Applications*, 368(1), 207-213. doi: 10.1016/j.physa.2005.12.027
 - Joyner, M. J., Ruiz, J. R., & Lucia, A. (2011). Last Word on Viewpoint: The

- two-hour marathon: Who and when? [Letter]. *Journal of Applied Physiology*, 110(1), 294-294. doi: 10.1152/jappl-physiol.01265.2010
- Katz, J. S., & Katz, L. (1999). Power laws and athletic performance. *Journal of Sports Sciences*, 17(6), 467-476.
 - Laherrere, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales. *European Physical Journal B*, 2(4), 525-539.
 - Lavalette, D. (1996). Facteur d'impact: impartialité ou impuissance? Report, INSERM U350. Orsay, France: Institut Curie-Recherche, Bât, 112, Centre Universitaire, 91405.
 - Li, W. T. (2012). Analyses of baby name popularity distribution in US for the last 131 years. *Complexity*, 18(1), 44-50. doi: 10.1002/cplx.21409
 - Li, W. T., & Miramontes, P. (2011). Fitting Ranked English and Spanish Letter Frequency Distribution in US and Mexican Presidential Speeches. *Journal of Quantitative Linguistics*, 18(4), 359-380. doi: 10.1080/09296174.2011.608606
 - Machado, J. A. T., Pinto, C. M. A., & Lopes, A. M. (2013). Power Law and Entropy Analysis of Catastrophic Phenomena. [Review]. *Mathematical Problems in Engineering*, 10. doi: 10.1155/2013/562320
 - Malacarne, L. C., & Mendes, R. S. (2000). Regularities in football goal distributions. *Physica a-Statistical Mechanics and Its Applications*, 286(1-2), 391-395.
 - Mansilla, R., Koppen, E., Cocho, G., & Miramontes, P. (2007). On the behavior of journal impact factor rank-order distribution. *Journal of Informetrics*, 1(2), 155-160. doi: 10.1016/j.joi.2007.01.001
 - MarathonGuide. (2000, 01-02-2012). Web Marketing Associates/MarathonGuide Retrieved 1 Febrero, 2012, from <http://www.marathonguide.com/>
 - Martinez-Mekler, G., Martinez, R. A., del Rio, M. B., Mansilla, R., Miramontes, P., & Cocho, G. (2009). Universality of Rank-Ordering Distributions in the Arts and Sciences. *Plos One*, 4(3). doi: e479110.1371/journal.pone.0004791
 - Naumis, G. G., & Cocho, G. (2007). The tails of rank-size distributions due to multiplicative processes: from power laws to stretched exponentials and beta-like functions. *New Journal of Physics*, 9. doi: 28610.1088/1367-2630/9/8/286
 - Naumis, G. G., & Cocho, G. (2008). Tail universalities in rank distributions as an algebraic problem: The beta-like function. *Physica a-Statistical Mechanics and Its Applications*, 387(1), 84-96. doi: 10.1016/j.physa.2007.08.002
 - Nevill, A. M., Markovic, G., Vucetic, V., & Holder, R. (2004). Can greater muscularity in larger individuals resolve the 3/4 power-law controversy when modelling maximum oxygen uptake? *Annals of Human Biology*, 31(4), 436-445. doi: 10.1080/03014460410001723996
 - Nevill, A. M., Stewart, A. D., Olds, T., & Holder, R. (2004). Are adult physiques geometrically similar? The dangers of allometric scalling using body mass power laws. *American Journal of Physical Anthropology*, 124(2), 177-182. doi: 10.1002/ajpa.10351
 - Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. [Review]. *Contemporary Physics*, 46(5), 323-351. doi: 10.1080/00107510500052444
 - Phillips, J. R. (2010, 19-01-2011). ZunZun.com Online Curve Fitting and Surface Fitting Web Site Retrieved 1-02-2012, 2012, from <http://www.zunzun.com/>
 - Popescu, I. (2003). On a Zipfs Law extension to impact factors. *Glottometrics*, 6(83-93).
 - Savaglio, S., & Carbone, V. (2000). Human performance - Scaling in athletic world records. *Nature*, 404(6775), 244-244.
 - Schmitt, F. G., Seuront, L., Hwang, J. S., Souissi, S., & Tseng, L. C. (2006). Scaling of swimming sequences in copepod behavior: Data analysis and simulation. *Physica a-Statistical Mechanics and Its Applications*, 364, 287-296. doi: 10.1016/j.physa.2005.09.041
 - Suvakov, M., Mitrovic, M., Gligorijevic, V., & Tadic, B. (2013). How the online social networks are used: dialogues-based structure of MySpace. *Journal of the Royal Society Interface*, 10(79), 12. doi: 10.1098/rsif.2012.0819
 - Waltman, L., & van Eck, N. J. (2009). Some comments on Egghe's derivation of the impact factor distribution. *Journal of Informetrics*, 3(4), 363-366. doi: 10.1016/j.joi.2009.05.004
 - Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.
 - Wesfreid, E., & Billat, V. (2009). Power law scaling behavior of physiological time series in marathon races using Wavelet Leaders and Detrended Fluctuation Analysis. In F. Truchetet & O. Lalignat (Eds.), *Wavelet Applications in Industrial Processing VI* (Vol. 7248). Bellingham: Spie-Int Soc Optical Engineering.
 - Wesfreid, E., Billat, V. L., & Meyer, Y. (2005). Multifractal analysis of heartbeat time series in human races. *Applied and Computational Harmonic Analysis*, 18(3), 329-335. doi: <http://dx.doi.org/10.1016/j.acha.2004.12.005>