Estudio analítico de las PC en las empresas en el Uruguay

Juan Grompone

Resumen: No existen estadísticas confiables acerca del número de computadoras personales en Uruguay. Tampoco se conoce la manera como estas computadoras se distribuyen en las empresas. En este trabajo se aplica un modelo de crecimiento para estimar el número total de PC. Se aplica la ley de Zipf para estimar la distribución en las empresas. El resultado puede ser de utilidad para análisis de mercado.

Palabras claves: Ley de Zipf, computadoras personales, Uruguay.

Abstract: Uruguay has not reliable statistics concerning to the number of personal computers in use. Neither we have information of how the personal computers are distributed among the enterprises. In this paper we use a simple grow model to estimate the total number of PC in the country. The enterprise distribution is estimated using the Zipf's law. In spite of the fact that the results obtained are highly speculative, can be useful for market analysis.

Key words: Zipf's law, personal computers, Uruguay.

1. Introducción

Uruguay es un país que tiene muy poca información estadística confiable, especialmente en el área análisis de mercados. No escapa a este hecho un indicador de gran relevancia para el área de la informática: la cantidad de computadoras personales y su distribución en las empresas del país. Por este motivo, es

interesante construir modelos matemáticos que permitan sustituir la información estadística de la cual se carece.

En el presente trabajo [1] se estudian dos modelos acerca del mercado de computadoras personales (en adelante, PC). El primer modelo es dinámico y analiza el crecimiento de las ventas y del parque instalado. El segundo modelo es estático y analiza la distribución de las PC en las empresas.

2. El modelo de crecimiento

En [1] se presentaba un modelo para el crecimiento del número de computadoras en el mundo. Las ecuaciones del modelo eran:

$$V_n = r \,.\, V_{n-1}$$
$$E_n = E_{n-1} + V_n - E_{n-5}$$

Donde V_n son las ventas de PC en el año n, E_n son los PC existentes el año n y r es el factor de crecimiento anual del mercado. Como puede observarse se eliminaban por obsoletas todas las computadoras con 5 años de antigüedad. El valor que mejor ajustaba con las cifras disponibles era un crecimiento anual r = 1,30 (o sea un crecimiento del 30% anual).

A los efectos de analizar el crecimiento de las PC en el Uruguay se puede aplicar el mismo modelo que para el mundo, con los mismos parámetros de ajuste. Es notorio que la conducta del mercado uruguayo sigue –al menos en forma cualitativa– las tendencias generales del mercado mundial. A estas hipótesis agregamos la *estimación* de la cifra de ventas de 50 mil PC vendidos en 1996. Utilizamos un período de obsolescencia de *6 años* que parece reflejar mejor las condiciones del país. Con estas hipótesis podemos elaborar el Cuadro 1 para la última década.

[1] Este artículo fue preparado para el "Segundo Congreso Nacional de Informática", Montevideo, 1999. Fue publicado en 1999 por la Oficina de Apuntes del CECEA como *Perspectivas y tendencias de la Informática en el Uruguay*.

Cuadro 1: Estimación de PC en Uruguay

año	parque miles	ventas miles
1990	36	10
1991	47	14
1992	61	18
1993	79	23
1994	102	30
1995	133	39
1996	173	50
1997	225	66
1998	293	85
1999	380	111
2000	495	144

Es posible que las cifras presentadas en el Cuadro 1 sean mayores a las efectivamente encontradas en la realidad debido a que el crecimiento anual en el Uruguay sea menor que el promedio mundial. Pueden incidir en este hecho algunos factores coyunturales propios de la región o del país. De todas maneras las cifras presentadas son un punto de partida para otras investigaciones sobre el tema.

3. Las empresas en Uruguay y la ley de Zipf

Existen pocos estudios estadísticos acerca de las empresas en el Uruguay. En el presente trabajo hemos tomado un estudio de una consultora que se distribuye en forma profesional [2]. De este estudio nos interesa destacar un aspecto importante: la distribución de las empresas según su facturación.

En la Figura 1 se presenta, en un diagrama doblemente logarítmico, las cifras de facturación anual de las empresas,

[2] El estudio es una encuesta realizada en 1998 por *M.C. Consultores*. Comprende algo más de 1300 empresas. Se encuentran todas la mayores empresas del país y comprende todos los sectores de la actividad.

ordenadas en forma decreciente. Se agrega una línea de tendencia que corresponde a la ley de Zipf. Como puede apreciarse, existe un excelente ajuste entre esta ley empírica y las cifras del informe.

La ley de Zipf establece que, luego de ordenar los elementos en forma decreciente, la empresa número *n* tiene una cifra de ventas anuales de *2400/n* millones de dólares [3]. En el diagrama esta ley representa una recta.

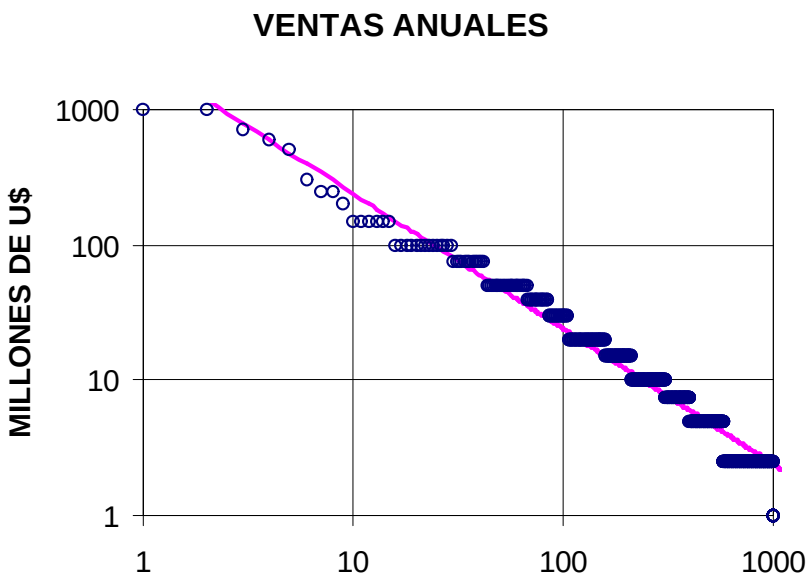

Figura 1: Facturación de las empresas en Uruguay, ordenadas por valores decrecientes. La línea indica los valores de la ley de Zipf.

El excelente ajuste de la medida económica del tamaño de las empresas en el Uruguay con la ley de Zipf nos permite suponer que puede usarse esta ley empírica como herramienta para el

[3] Es habitual que la ley de Zipf no se cumpla para los primeros valores. Es notorio que no existe en Uruguay una empresa que facture 2.400 millones de dólares anuales. Como puede verse, el ajuste es muy bueno a partir de la tercera empresa.

análisis de algunos casos de interés acerca de los cuales carecemos de información directa.

4. Sobre la ley de Zipf

La ley de Zipf fue propuesta hace medio siglo [3] si bien existen referencias anteriores en la literatura económica clásica [2]. Es una ley empírica cuyo campo de aplicación tiene una amplitud sorprendente.

En su forma más general se aplica a un conjunto de elementos –cuyas condiciones todavía no se han podido identificar con precisión– que poseen un valor numérico asociado. Este conjunto se puede ordenar de diversas maneras (una de las más clásicas es el histograma). La ordenación de Zipf consiste en disponerlos por valores *decrecientes*. La ley establece que el valor numérico enésimo A_n (con la posible excepción de unos pocos valores iniciales) posee la siguiente expresión:

$$A_n = \frac{A_1}{n}$$

Obsérvese que un conjunto de valores aleatorios, distribuidos uniformemente en un intervalo posee una expresión de ordenación de Zipf que es lineal en n. No cabe duda que la distribución de Zipf exige condiciones estructurales al conjunto de valores, si bien todavía no existe acuerdo de cuáles son estas condiciones.

Originariamente Zipf [3] la consideró como una “ley natural” y la vinculó a un principio de “mínimo esfuerzo”. No obstante esto, por el hecho que Zipf fuese un lingüista, usualmente se la consideró una ley aplicable solamente a las lenguas naturales para describir la frecuencia de uso de las diferentes palabras. Es de esta manera que la usó Shannon [4] en un célebre trabajo sobre la información que contiene una palabra o una letra en inglés. La aplicación a la lingüística suele se aceptada con bastante generalidad, ver por ejemplo [8] y [13].

La aplicación de la ley de Zipf, sin embargo, va más allá de la lingüística. Se la ha aplicado a diversos campos del conocimiento, dispares entre sí, tales como la economía [2], [5] y [10]; la estructura de la información genética [7]; la popularidad o el uso de los sitios Web [14], [15] y [16]; o la población urbana [11], [12] y [17].

La ley de Zipf aparece como un caso particular de una ley potencial del tipo $1/n^a$, donde $a \geq 1$. Se han dado muchas razones por las cuales podría ocurrir una distribución de este tipo en diversos fenómenos de la naturaleza. Dentro de estas "explicaciones" se encuentran [6], [8] o [9].

En este artículo tomamos la ley de Zipf como una ecuación experimental de la ya que hemos verificado su aplicación mediante la Figura 1. Dejaremos de lado las argumentaciones acerca de su validez o sus campos de aplicación.

5. Las computadoras personales por empresa

La ley de Zipf permite construir una estimación acerca de las computadoras personales que se emplean en las empresas en el Uruguay.

Podríamos pensar que basta con tomar el primer valor de la serie como el número de computadoras personales de la empresa más grande del país. Sin embargo, esta manera de estimar este valor no es precisa porque sabemos que existe un apartamiento de la ley en los primeros valores de la serie. Un mejor método para estimar el coeficiente de proporcionalidad consiste en ajustar el número total de computadoras personales que estimamos que existen en el país. A estos efectos, tomamos la distribución por uso de las computadoras personales que aparece en el Cuadro 2.

Cuadro 2: Distribución de los PC por uso, en 1998.

destino	%	cantidad miles	notas
PC empresas	55%	161	[4]
PC hogares	35%	103	[5]
fuera de uso	10%	29	[6]
total	100%	293	

Tomemos a 1998 como año base para el cálculo. El Cuadro 2 estima, para este año, un total de 293.000 computadoras personales existentes en el país. Con esta cifra podemos estimar en 161.000 la cantidad de computadoras personales existentes en las empresas y, consecuentemente, estimar el coeficiente de proporcionalidad de Zipf de modo de obtener este total [7]. La estimación de cantidad de computadoras personales para la empresa *n*, en 1998 es, aproximadamente de:

$$\frac{16000}{n}$$

Como resulta de inmediato, hay 16.000 empresas con computadoras, desde las mayores empresas del país hasta las pequeñas empresas que poseen solamente una computadora

[4] Esta cifra se obtiene por diferencia de las otras dos.

[5] El censo de 1996 detectó más de 50.000 computadoras en los hogares, pero esta cifra puede tener alguna imprecisión por el significado de la palabra “computadora”. Las últimas encuestas de hogares indicaban, en 1998, un 17% de hogares con computadoras personales. Para 600.000 hogares, esto indica del orden de 100.000 computadoras. Estas dos cifras tienen una razonable coherencia. El porcentaje estimado es el correspondiente a un parque total de 300.000 computadoras personales existentes en 1998 según cifras del Cuadro 3.

[6] Esta cifra proviene de estimar que la vida útil completa de las computadoras personales es de 6 años.

[7] Este punto es muy interesante. La suma de la serie armónica es divergente. Por esta razón Shannon [4] introdujo una técnica de análisis del problema que se presenta en el Apéndice de este trabajo.

personal. Con esta ley de distribución podemos armar el Cuadro 3 que presenta la distribución de computadoras personales en las empresas de Uruguay.

Cuadro 3: Distribución de empresas en el Uruguay según la cantidad de PC. Cifras estimadas para 1998 [8].

rango de empresa	cantidad de PC
1	4,000
2	4,000
3	4,000
4	4,000
5	3,200
6	2,666
7	2,285
8	2,000
9	1,777
10	1,600

rango de empresa	cantidad de PC
20	800
50	320
100	160
200	80
500	32
1000	16
2000	8
5000	3
10000	1
16000	1

Los resultados del Cuadro 3 permiten aproximar la distribución de las computadoras personales en el sector empresas y pueden ser una herramienta útil para estudios de mercado u otros tipos de estudios económicos o comerciales.

Se pueden extraer algunas conclusiones del Cuadro 3. Hay más de 2000 redes de computadoras, si consideramos que toda empresa con más de 8 PC debe tener un servidor y una red. Las 500 mayores empresas del país poseen 32 o más computadoras

[8] Para las primeras tres empresas se han tomado las mismas cifras que para la cuarta. Este comienzo “plano” es típico de las cifras reales en una distribución de Zipf. El resultado parece más ajustado que el que predice la distribución para los primeros valores.

personales. Se estima en 16, las empresas que poseen más de 1.000 computadoras personales.

6. Conclusiones

La ley de Zipf se nos presenta como una ley empírica, acerca de la cual todavía no hay justificaciones aceptadas. No obstante esto, puede ser empleada como herramienta de análisis de algunos problemas sociales o físicos acerca de los cuales hay información insuficiente.

La aplicación reiterada a casos concretos y su análisis crítico permitirá, en el futuro, precisar los alcances de su validez y su rango de acción.

7. Apéndice: ajuste de parámetros de la ley de Zipf

El caso clásico de aplicación de la ley de Zipf lo ha realizado Shannon, [4]. El problema que se planteaba era disponer de la distribución de frecuencias de las palabras inglesas, tema que era el motivo directo de la ley de Zipf original. La frecuencia (que puede tomarse como probabilidad de ocurrencia) de las palabras, ordenadas por frecuencia decreciente, es $0{,}1/n$, donde n es el número de orden de la palabra.

El punto interesante en el análisis de Shannon consiste en observar que la serie armónica es divergente, de modo que no puede aceptarse que el número de palabras en inglés sea infinito. Por el contrario, este número se puede estimar imponiendo una condición suplementaria: la probabilidad de ocurrencia de cualquiera de las N palabras inglesas debe ser 1 y de allí que se tenga que cumplir la ecuación:

$$0{,}1\, H_N = \sum_{1}^{N} \frac{0{,}1}{n} = 1$$

Mediante esta ecuación, derivada de la ley de Zipf, puede calcularse la cantidad de palabras inglesas *de uso corriente*. Se

indica, como es habitual, con H_i la suma de los primeros i términos de la serie armónica. La solución de Shannon de la ecuación es 8.727 [9].

Esta manera de proceder está implícita en el uso de la ley de Zipf. Shannon descubrió indirectamente que, en general, la ley de Zipf exige dos parámetros y no uno como puede aparecer al principio. Por la divergencia de la serie armónica es necesario imponer una condición adicional que permita calcular el rango máximo de aplicación de la ley.

En el caso de la estimación de la cantidad de PC en las empresas, en este trabajo, el valor A_1 nos es desconocido y también lo es el rango máximo que se debe usar: no sabemos cuántas empresas existen en el Uruguay que posean computadoras. Tenemos dos parámetros que se debe definir y dos condiciones a cumplir:

- la cantidad total de las computadoras personales del país se estima en 161.000.
- la ley de Zipf da valor fraccionarios que puede interpretarse tomando la parte entera del número, puesto que el número de computadoras es un numero entero: es notorio que no tienen sentido las cantidades menores de 1.

Esta condiciones nos indican que si N es el número de empresas que poseen computadoras, entonces se debe cumplir las dos condiciones:

$$\frac{A_1}{N} = 1$$

[9] Es muy curioso, pero para mi programa de cálculo la solución de esta ecuación es 12.366. Es claro que este número es tan aceptable como el de Shannon para la consideración del problema.

$$\sum_{1}^{N} \frac{A_1}{n} = 161.000$$

La solución de estas dos ecuaciones da N=15.722. Hemos tomado, por sencillez, 16.000 como aproximación fácil de recordar.

8. Referencias

[1] Grompone, Juan. *Smaller, faster, cheaper*. Primer Congreso Nacional de Informática, Montevideo, 1998.

[2] Pareto, V. *Cours d'Economie Politique*. 1897.

[3] Zipf. G. K. *Human Behaviour and the Principle of Least Effort.* Cambridge, Mass.; Adison-Wesley, 1949.

[4] Shannon, Claude E. *Prediction and Entropy of Printed English*. Bell Sys. Tech. Journal, 50-64, Jan. 1951.

[5] Mandelbrot, Benoit B. *New methods in statistical economics*. Journal of Political Economy, 71, 421-440, 1963.

[6] Mandelbrot, Benoit B. *The Fractal Geometry of Nature*. W. H Freeman, 1982.

[7] Mantegna, R. N. et al. *Linguistic features of non-coding DNA Sequences*. Phys. Rev. Lett. 73, 23, Dec. 1994.

[8] Kessler, David; Kanter, I. *Markov Processes: Linguistics, Zipf's Law and long-range Correlations*. Phys. Rev. Lett. 74, 4559, 1995. http://dave.ph.biu.ac.il/kessler/my_research.html

[9] Laherrère, Jean. *Distributions de type "fractal parabolique" dans la Nature*. Comptes Rendus de l'Académie des Sciences, Série II a: Sciences de la Terre et des Planètes. T.322, Série IIa N. 7, p. 535- 541, 4 Avril 1996. http://hubbertpeak.com/laherrere/fractal.htm

[10] Krugman, Paul. *The Self-Organizing Economy*. Cambridge, Mass.; Blackwell Publishers, 1997.

[11] Gabaix, Xavier. *Zipf's Law for Cities: An Explanation*. MIT Department of Economics. 1998.

[12] Gabaix, Xavier. *Zipf's Law and the Growth of Cities*. MIT Department of Economics. 1999.

[13] Turner, G. R. *Relationship Between Vocabulary, Text Length and Zipf's Law.* http://www.btinternet.com/~g.r.turner/ZipfDoc.htm

[14] Nielsen, Jakob. *Zipf Curves and Website Popularity*. http://www.useit.com/alertbox/zipf.html

[15] van Till, Jaap. *Fractanomics - The Issue of Scale in the Network Economy*. http://huizen.dds.nl/~vantill/fractanomics.html

[16] Crovella, Mark E. et al. *Characteristics of WWW Client-based Traces*. http://www.cs.bu.edu/faculty/crovella/papers.html

[17] Martinotti, Guido. *The new social morphology of cities*. http://www.unesco.org/most/martinot.htm