Sobre el tamaño de las ciudades en España. Dos reflexiones y una regularidad empírica

Goerlich, Francisco José and Mas, Matilde

Instituto Valenciano de Investigaciones Económicas, Universitat de València

June 2008

Sobre el tamaño de las ciudades en España. Dos reflexiones y una regularidad empírica*

Francisco J. Goerlich y Matilde Mas
Universidad de Valencia e Instituto Valenciano de Investigaciones Económicas (Ivie)

RESUMEN

El trabajo propone dos reflexiones previas al estudio del tamaño de las ciudades, destacando su importancia en la contrastación de una de las regularidades empíricas más profusamente visitadas por la literatura, la ley de Zipf. La primera reflexión se refiere a la relevancia que tiene para el análisis la utilización de bases de datos depuradas. La segunda, entra en la problemática no resuelta de qué debe entenderse por ciudad. Las dos reflexiones son tenidas en cuenta en el contraste que se realiza de la ley de Zipf basado en una generalización de la distribución de Pareto. El contraste permite rechazar dicha ley, para el concepto de ciudad adoptado, en todos los años considerados.

Palabras clave: Población, Municipios, Censos, Localización, Áreas Metropolitanas.

Clasificación JEL: J10, J11

ABSTRACT

This paper proposes two considerations prior to the study of city size and emphasizes their importance in the testing of one of the empirical regularities most widely used by the literature, Zipf's law. The first one refers to the importance of using a consistent database for analysis. The second looks at the unresolved problem of what exactly is meant by city. These two considerations are taken into account when carrying out a contrast of Zipf's law, based on the generalization of the Pareto distribution. This contrast permits the rejection of this law for the assumed concept of city in all the years considered.

Key Words: Population, Municipalities, Census, Agglomeration, Metropolitan Areas**.**

JEL Classification: J10, J11

---

* Correspondencia: Matilde Mas, Instituto Valenciano de Investigaciones Económicas, C/. Guardia Civil, 22, Esc. 2, 1º, 46020 Valencia. E-mail: matilde.mas@ivie.es.

Agradecimientos: Los autores agradecen la ayuda prestada por Pilar Chorén en el tratamiento de la información y la financiación de los proyectos del Ministerio de Ciencia y Tecnología/FEDER, SEC2005-02776, y del programa de investigación Fundación BBVA-Ivie.

I. Introducción

Cualquier estudio sobre la localización de la población sobre el territorio debe plantearse, como paso previo, la unidad de análisis que va a ser tomada de referencia. La opción más adecuada dependerá de cuál sea el problema que se desea abordar, pero es probable que muy pronto se tropiece, de una forma u otra, con la necesidad de definir el concepto de *ciudad*. De acuerdo con el Diccionario de la Real Academia *ciudad* "es el conjunto de edificios y calles, regidos por un ayuntamiento, cuya población densa y numerosa se dedica por lo común a actividades no agrícolas". Por su parte, la Wikipedia la define de forma similar como "una entidad urbana con alta densidad de población en la que predominan fundamentalmente la industria y los servicios".

Un intento de precisar el concepto anterior lo proporciona la Conferencia Europea de Estadística de Praga[1] que considera como ciudad una aglomeración de más de 2.000 habitantes siempre que la población dedicada a la agricultura no exceda del 25% sobre el total. A partir de 10.000 habitantes, todas las aglomeraciones se consideran ciudades, siempre que éstos se encuentren concentrados, generalmente en edificaciones colectivas y en altura, se dediquen fundamentalmente a actividades de los sectores secundario y terciario (industria, comercio y servicios). También se tiene en cuenta la densidad de población de dicho país o ciudad.

Aunque el concepto relevante de *ciudad* fuera comúnmente aceptado, que no lo es, continuaría planteándose el problema de definir su contrapartida empírica. Para ello haría falta disponer de la información estadística de forma tal que pudiera establecerse una relación unívoca entre el concepto y su medida. La información estadística suele tomar como referente los municipios y aunque los *Nomenclátores* descienden a mayores niveles de desagregación el seguimiento temporal de la información plantea problemas. Por otra parte, las divisiones administrativas –fruto de la historia- en ocasiones tratan como entes individualizados aglomeraciones de población geográficamente indistinguibles entre sí.

Adicionalmente, los lindes administrativos municipales no se han mantenido constantes, sino que han sufrido continúas transformaciones como resultado de fusiones, agregaciones parciales y segregaciones. Este hecho plantea el problema de la homogeneidad de las series –normalmente disponibles a partir de los censos de población- a lo largo de periodos dilatados de tiempo.

[1] Capel (1975) ofrece una detallada revisión de las distintas acepciones del concepto de ciudad.

Las líneas anteriores han puesto de manifiesto la necesidad de clarificar, como paso previo en cualquier análisis de localización de la población: 1. qué se entiende por *ciudad;* y 2. qué uso se va a hacer de la información estadística disponible. Sólo cuando estas dos cuestiones hayan sido precisadas pueden abordarse con garantías cuestiones como las planteadas por la ley de Zipf. Aunque la relación rango-tamaño, a la que la ley de Zipf (1949) hace referencia, no es el único aspecto, ni tampoco el más relevante, en relación a los procesos de urbanización y el tamaño de las ciudades, sí es cierto que aparece con frecuencia en la literatura, tanto teórica (Richardson 1973; Gabaix 1999, Brakman, Garretsen, Van Marrewijk y van den Berg 1999; Duranton 2002), como fundamentalmente empírica (Rosen y Resnick 1980; Carroll 1982; Smith 1990; Eaton y Eckstein 1997; Soo 2002; Ioannides y Overman 2003).

Este tema concreto ha sido poco estudiado en el caso español, donde los estudios sobre urbanización de la población se han centrado en casos particulares relativamente recientes (Artís, Romaní y Suriñach 1998; Feria 2000; Castañer, Vicente y Boix 2000; Trullén y Boix Doménech 2003; Ajenjo y Sabater 2004; o Boix Domènech 2004), o bien en análisis históricos de corte descriptivo sobre concentración de la población en núcleos o municipios por encima de un determinado umbral (Reher 1986, 1994; Gómez y Luna 1986; Correas 1988; Valero 1989; Camps 1990; Vinuesa 1996; Tafunell 2005). La relación rango-tamaño, y la ley de Zipf, aparece en el caso español en de Vries (1984, capítulo 6), en relación al desarrollo histórico de los procesos de urbanización en comparación con otros países europeos, en Lasuén, Lorca y Oria (1967), en Capel (1972) y más recientemente en Esteve y Devolver (2004). Sin embargo, en este trabajo tomaremos de referencia el de Lanaspa, Perdiguero y Sanz (2004).

La estructura del trabajo se organiza entorno a los tres temas mencionados. El apartado II aborda los problemas planteados por las bases de datos disponibles, explicitando la opción tomada por este trabajo. El epígrafe III ofrece unas breves reflexiones en torno al concepto de ciudad y presenta la acepción por la que nos decantamos. El epígrafe IV discute la ley de Zipf (1949) y presenta un contraste estadístico de la misma, no basado en la tradicional regresión rango-tamaño. Finalmente, en el apartado V se ofrecen unas breves conclusiones.

II. Series históricas de población municipal: fuentes alternativas

Aunque un análisis del tamaño de las ciudades debería partir idealmente de la población asentada sobre núcleos recogidos en los *Nomenclátor de las ciudades, villas,* *aldeas, lugares y otras entidades de población*, que desde 1877 se realizan al mismo tiempo que los censos, es cierto que desde un punto de vista histórico los asentamientos recogidos en los nomenclátores son muy cambiantes; no disponen de superficie asignada; su ubicación geográfica (georeferenciación) no está disponible con generalidad; y, adicionalmente, el crecimiento de ciertos núcleos (normalmente las grandes áreas urbanas, que son las que más interés despiertan entre geógrafos y demógrafos) es tal en algunos casos que, en realidad, no estamos hablando de varios núcleos, sino de una sola *ciudad*, aunque por cuestiones de organización estadística los datos sobre núcleos de población se mantengan separados.

En consecuencia, la fuente de información más utilizada para los estudios sobre localización de la población han sido las poblaciones municipales, o agregaciones urbanas construidas a partir de ellas (Vinuesa 1997; Zoido y Arroyo 2004; Lanaspa, Perdiguero y Sanz 2004; De Cos y Reques 2005, Goerlich, Mas, Azagra y Chorén 2006).[2] De esta forma es frecuente identificar el concepto de *ciudad* con el de población municipal por encima de un determinado umbral, aunque esto no deje de ser una aproximación al fenómeno que queremos medir (Lanaspa, Perdiguero y Sanz 2004). Volveremos brevemente sobre esta cuestión en el epígrafe siguiente.[3]

A la hora de obtener las poblaciones municipales la fuente primaria de información son los censos. El primer censo que presenta el conjunto completo de municipios que cubren el territorio español es el llamado *Censo de la Matrícula Catastral*, fechado en 1842. Dicho censo fue realizado por el procedimiento de imputaciones y, en consecuencia, carece de rigor y fiabilidad en sus cifras. Por ello se considera como primer censo moderno el de 1857. El manejo directo de las fuentes censales presenta dos problemas fundamentales.

En primer lugar, hasta que el INE abrió recientemente una sección en su *web* denominada *Alteraciones de los municipios en los Censos de Población desde 1842*, la consulta de dichos datos debía realizarse (para los censos anteriores al de 1991), directamente en papel, y proceder a su informatización.[4] Dicha sección respeta escrupulosamente la información censal original de cada uno de los censos,[5] y ofrece

---

[2] Existen no obstante notables excepciones de carácter histórico que tratan consistentemente de utilizar datos sobre núcleos de población (Luna 1988; Reher 1994; Esteve y Devolver 2004).

[3] Véase, no obstante, la discusión en Tafunell (2005).

[4] Esta fue la vía seguida en Goerlich, Mas, Azagra y Chorén (2006).

[5] Algunas anomalías detectadas en dicha sección en el transcurso de este trabajo se ofrecen en el apéndice.

adicionalmente una valiosa información sobre alteraciones municipales y cambios de denominación.[6]

En segundo lugar, los cambios en la estructura municipal, fusiones, agregaciones parciales, segregaciones,… han sido notables en España a lo largo de los siglos XIX y XX. Por ejemplo, en el censo de 1900 el número de municipios existentes era de 9.267, mientras que 100 años más tarde, en el censo de 2001, tan sólo aparecen 8.108 municipios. La situación es todavía mucho peor si consideramos los municipios existentes en los censos del siglo XIX.[7]

Estos dos problemas parecen tener una solución sencilla si nos restringimos al siglo XX. La razón es que el Instituto Nacional de Estadística ofrece en su *web*, desde hace tiempo, unas *Series históricas de población de hecho municipal para el periodo 1900 – 1991* obtenidas directamente de los *Censos de Población*. Estas poblaciones pueden descargase con facilidad a nivel provincial en formatos accesibles y además parecen ser homogéneas, es decir toman "*como referencia la relación de municipios del Censo de 1981*" (INE, *web*, nota en la descarga de las series). El investigador se encuentra así con una matriz de datos de poblaciones de hecho municipales con (aparentemente) tantas filas como municipios existían en 1981 (8.022 según el censo de dicho año) y 10 columnas, una por cada censo entre 1900 y 1991. La facilidad en la disponibilidad de esta información ha propiciado su uso, de hecho esta es la fuente de información utilizada por Lanaspa, Perdiguero y Sanz (2004).

Esta base de datos enmascara ciertas peculiaridades y errores que es necesario tener en cuenta. Si estas poblaciones toman "*como referencia la relación de municipios del Censo de 1981*" dos preguntas surgen de forma inmediata: ¿Qué sucede con los municipios que desaparecen (por fusión o incorporación a otro municipio) antes del 1 de marzo de 1981 y después de dicha fecha?,[8] y ¿Qué sucede con los municipios de nueva creación con anterioridad al censo de 1981 y con posterioridad a dicho censo?. Es decir, como se han tratado las alteraciones municipales si la pretensión es *congelar* la estructura municipal en un momento dado del tiempo (censo de 1981) y proyectarla al pasado y al futuro.

---

6 Aún así, la información de dicha sección no está pensada para su descarga, sino simplemente para su consulta *on-line*. Además de esta información, el INE ha puesto a disposición de los usuarios las versiones originales en papel de los censos de 1900 a 1970, en formato *pdf*.

7 Así lo demuestra la codificación de todos los municipios que han existido alguna vez en el periodo 1842 – 2001 realizada por el INE.

8 1 de marzo es la fecha de referencia del censo de 1981.

Parte de la respuesta a estas preguntas la proporciona el propio INE que, para cada provincia, consigna un registro de "*Población en municipios desaparecidos*" y nos indica (de forma algo críptica) que: "*La población reflejada de los municipios desaparecidos está comprendida entre los Censos de 1900 a 1970*" (INE, *web*, nota en la descarga de las series). En consecuencia, los municipios que desparecen con anterioridad al censo de 1981 pasan a engrosar las cifras de un registro de "*Población en municipios desaparecidos*". Los volúmenes de población *desaparecida* no son despreciables, oscilando entre un valor mínimo de 270.660 personas en 1970, hasta un valor máximo de 943.626 personas en 1930 (Goerlich, Mas, Azagra y Chorén 2006, Cuadro 2.10). Por su parte, la inspección de los datos revela que los municipios de nueva creación en este periodo simplemente aparecen de la nada y no se les asigna población con anterioridad a su existencia. La base de datos tiene, pues, un buen número de municipios en los que en algunos años la población es nula.

¿Qué sucede con las alteraciones que tienen lugar con posterioridad al 1 de marzo de 1981? Los municipios de nueva creación entre los censos de 1981 y 1991 vuelven a aparecer de la nada y únicamente tienen población asignada en el año 1991.[9] Este es el caso de, por ejemplo, Los Alcázares (30902), en Murcia, que aparece por primera vez en 1991 resultado de la fusión de dos entidades locales, una de San Javier (código 30035) y otra de Torre Pacheco (30037). Estos municipios sólo tienen un dato de población en todo el siglo XX en la base de datos del INE.[10] Por su parte, los municipios que desaparecen entre los censos de 1981 y 1991 simplemente son eliminados de la base de datos municipal,[11] pero como en 1981 no se consigna "*Población en municipios desaparecidos*", tenemos la paradoja de que en 1981, que se supone es el año de referencia de la base de datos de poblaciones de hecho municipales,

[9] Obsérvese que estos municipios no deberían aparecer si el propósito era ofrecer una base de datos con la estructura municipal vigente en la fecha del censo de 1981.

[10] Cuando un municipio que no aparece en el censo de 1981, pero sí existió anteriormente en otros censos, vuelve a aparecer en el censo de 1991 entonces la casuística es más variopinta. Así por ejemplo, tenemos casos como Emperador (código 46117), en Valencia, o Almodóvar del Pinar (16017), en Cuenca, que no existen como tales municipios sólo en el censo de 1981, pero sí en todos los demás del siglo XX. En estos casos sólo se consigna como no disponible en la base de datos del INE la población en 1981, pero sí aparece la correspondiente población en el resto de años. Por el contrario, tenemos casos como los de Mazaroleja (40903), en Segovia, que solamente deja de aparecer en los censos de 1970 y 1981, pero sí es municipio independiente en el resto de censos del siglo XX, y sin embargo sólo aparece con población en la base de datos del INE en 1991. Otro ejemplo de esta tipología es Torremolinos (29901), en Málaga, que aunque existe en los censos de 1900, 1910, 1920 y 1991, sólo se le asigna población en la base de datos del INE en 1991. En el resto de años la población de estos municipios pasa a engrosar el registro de "*Población en municipios desaparecidos*", cuando la realidad es que en algunos años no existía tal desaparición.

[11] Obsérvese de nuevo que estos municipios no deberían eliminarse si el propósito es ofrecer una base de datos con la estructura municipal vigente en la fecha del censo de 1981.

las cifras agregadas de población de la base de datos del INE no cuadran con los datos censales originales.

En concreto, hay ocho provincias cuya cifra de población de hecho en los censos originales no coincide con la publicada para 1981 en la base de datos del INE en su *web*. Estos casos son los siguientes.

1. En Almería, se ha eliminado el municipio de Beninar (04025) con 294 habitantes en 1981, que no desparece hasta el censo siguiente porque se incorpora a Berja (04029).[12]

2. En Burgos la diferencia es de 649 personas porque no se ha considerado la población de Villorobe (09475), 267 habitantes, que se incorporaría a Villasur de Herreros (09463) en 1991; la de Castil Carrias (09080), 1 habitante, que se agrega a Belorado (09048) en 1991; y la de Junta de Río Losa (09187), 183 habitantes, y Junta de San Martín de Losa (09188), 198 habitantes, que se fusionan en 1991 en un único municipio, Valle de Losa (09908).[13]

3. En Cuenca la diferencia asciende a 1.282 habitantes, que coincide con la población del municipio de Almodóvar de Monte Rey (16907) que surge en 1981 como resultado de la fusión de cinco municipios: Chumillas (16981), Monteagudo de las Salinas (16131), Olmeda del Rey (16141), Solera del Gabaldón (16199) y Almodóvar del Pinar (16017), y desaparece en 1991 al recuperar la independencia dichos municipios. Estos cinco municipios no tienen asignada población en 1981 en la base de datos de INE, a pesar de ser perfectamente identificables en el nomenclator.

4. En Guadalajara la diferencia es de 122 habitantes que resulta de haber imputado a Zorita de Canes (19335) 204 habitantes más que lo que indica el censo original de 1981 y de no haber tenido en cuenta a Alcorlo (19012), que se incorporaría a La Toba (19269) en 1991, y a Torrecilla del Ducado (19276), que también desaparece en el censo siguiente pasando a formar parte de Sienes (19256).

[12] La incorporación a Berja (04029) de Beninar (04025) en 1991 toma como fuente el Índice de Municipios del INE (1986, págs.- 9 y 98), sin embargo según las *Alteraciones de los municipios en los Censos de Población desde 1842* Beninar (04025) se incorpora a Darrícal (04039), que en el censo de 2001 se integra en Alcolea (04007).

[13] De esta forma Valle de Losa (09908) sólo tiene asignada población en la base de datos del INE en 1991, cuando hubiera sido trivial construir una población homogénea para este municipio, simplemente sumando retrospectivamente la población de los dos municipios que lo formaron.

5. La siguiente divergencia se encuentra en Guipúzcoa donde el municipio de Albaltzisketa (20001) presenta 204 habitantes menos que en el censo original.[14]

6. La sexta provincia afectada es León, dónde en 1981 hay una diferencia de 605 personas porque no aparece el municipio de Fresnedo (24072) que se agregaría parcialmente a Cubillos de Sil (24064) y a Toreno (24169) posteriormente.

7. En Salamanca también se aprecia una diferencia de 335 habitantes en 1981 porque no se ha tenido en cuenta la población de Castellanos de Villiquera (37093), ni la de Carbajosa de Armuña (37084), que aparecen incorporados al antiguo municipio de La Mata de Armuña (37185) en el censo de 1991, y que a su vez cambia de nombre por el de Castellanos de Villiquera en dicho año.

8. Por último, Zamora muestra una diferencia de 452 habitantes que se debe a la población de Fornillos de Fermoselle (49074) en 1981 en el censo original. Este municipio también pierde su autonomía en 1991 al incorporarse a Villar del Buey (49264).

Estas omisiones provocan una diferencia en 1981 entre el total de población de hecho del censo original y el publicado por el INE en su *Serie histórica municipal de población de hecho* de 3.699 personas de más.

También es interesante constatar algunos problemas de los datos provinciales en la información procedente de las series históricas y que hace que no coincidan con los originales. El total provincial publicado por el INE en su *web* en las series históricas no coincide con la suma de los totales provinciales en 1930, 1940 y 1970. Las diferencias se deben, para los dos primeros años, a que el dato de población de Illes Baleares en 1950 se ha imputado también a 1940 y el de éste año a 1930 y, por lo tanto, el dato original de población de hecho de 1930, 365.512 personas, solo aparece si miramos a los datos agrupados por Comunidades Autónomas, donde sí es correcto. Un error similar es el que se aprecia en 1970, en el que para la provincia de Toledo se ha repetido

[14] En realidad los 204 habitantes de más de Zorita de Canes (19335) y los 204 de menos de Albaltzisketa (20001) se deben a un error en la base de datos. Las poblaciones de estos municipios para todos los años comprendidos entre 1900 y 1981 se encuentran intercambiadas. Ello sólo tiene efectos a nivel provincial en 1981 porque en el resto de años la "*Población en municipios desaparecidos*" actúa de cajón de sastre a efectos de cuadrar con las cifras agregadas provinciales.

la cifra de población de 1960 en dicho año y, por lo tanto, el dato original de población de hecho de Toledo de 1970, 468.925, no aparece en ningún sitio.[15]

Así pues, la utilización de las *Series históricas municipales de población de hecho* puestas a disposición por parte del INE no soluciona adecuadamente los dos problemas mencionados al principio de este epígrafe, ya que la información contiene numerosos errores y además, ni se trata de las poblaciones censales originales,[16] ni de ningún tipo de homogeneización territorial con arreglo a criterios claros y explícitos.

Puesto que los errores y las alteraciones municipales son cuantiosas pero afectan fundamentalmente a municipios pequeños y de escasa entidad de población cabe preguntarse si estas "imperfecciones en los datos" son potencialmente relevantes para el problema que nos ocupa, la distribución del tamaño de las ciudades y la relación rango-tamaño. Un ejemplo bastará para cerciorarnos de que sí.

Como es bien sabido, los dos municipios de mayor importancia en la historia reciente en España son Madrid, capital del estado, y Barcelona, centro de gran tradición industrial (Nadal 2003). Su importancia destacada del resto de municipios españoles se observa bastante antes de principios del siglo XX, si bien Madrid desaparece si nos retrotraemos suficientemente en el tiempo (de Vries 1984, Capítulo 6). A lo largo del siglo XX ambas poblaciones experimentan un crecimiento notable como atestiguan los datos del cuadro 1. De acuerdo con los datos censales, que recogen las dos primeras columnas, la población de ambos municipios estaba notablemente igualada en la primera mitad del siglo XX. Madrid aparece el primero en el ranking, excepto en 1930, año en el que Barcelona presenta un mayor volumen de población. Sin embargo, y dejando al margen el hecho de si estos municipios deben ser considerados como ciudades, o debe considerarse como tal su correspondiente área metropolitana, estamos comparando aglomeraciones de población no homogéneas. En la década de los 40 y 50 Madrid incorporó, fruto de su gran crecimiento, muchos pequeños municipios colindantes (Aravaca, Barajas de Madrid, Canillas, Canillejas, Carabanchel Alto, Carabanchel Bajo, Chamartín de la Rosa, Fuencarral, Hortaleza, El Pardo, Vallecas, Vicálvaro y Villaverde). Por su parte, Barcelona tan sólo ha incorporado dos municipios a principios del siglo XX, Horta y Sarriá. En consecuencia, una comparación histórica de la población de ambos municipios exigiría homogeneizar lo que en la actualidad son

[15] En todos estos casos la causa última de los errores puede rastrearse hasta valores incorrectos en la "*Población en municipios desaparecidos*". Los valores de dicha población son también incorrectos en Málaga en 1910 y en Salamanca en 1960.

[16] Disponibles en la actualidad a través del enlace *Alteraciones de los municipios en los Censos de Población desde 1842*.

sus términos municipales y reconstruir sus poblaciones hacia atrás. Esta reconstrucción son las poblaciones homogéneas en las columnas 3 y 4 del cuadro 1. En este caso los crecimientos de Madrid en la primera mitad del siglo XX se ven amortiguados pero, lo que es más importante desde nuestro punto de vista, Madrid aparece como el municipio más poblado en todos los años del siglo XX, mostrando diferencias mucho más acusadas con Barcelona que las que se observan en las cifras originales. La conclusión es, por tanto, que si queremos realizar un contraste de la relación rango-tamaño, o la ley de Zipf, un gran número de pequeñas alteraciones municipales pueden afectar de forma importante a los resultados y deben ser tenidas en cuenta de alguna forma.[17]

Cuadro 1. Población de hecho de los censos de 1900 a 1991

	Población Censal		Población Homogeneizada	
	Madrid	**Barcelona**	**Madrid**	**Barcelona**
1900	539.835	533.000	576.538	543.930
1910	599.807	587.411	659.775	595.484
1920	750.896	710.335	848.383	721.869
1930	952.832	1.005.565	1.137.943	1.005.565
1940	1.088.647	1.081.175	1.326.674	1.081.175
1950	1.618.435	1.280.179	1.645.215	1.280.179
1960	2.259.931	1.557.863	2.259.931	1.557.863
1970	3.146.071	1.745.142	3.146.071	1.745.142
1981	3.188.297	1.754.900	3.188.297	1.754.900
1991	3.084.673	1.681.132	3.084.673	1.681.132

Fuente: Censos de población. Elaboración propia para la población homogeneizada teniendo en cuenta las alteraciones municipales.

Consciente de este problema García Fernández (1985) emprendió la laboriosa tarea de homogenizar las poblaciones de hecho municipales a partir de los censos de población "*…para eliminar la influencia de las alteraciones territoriales de los municipios, debidas a fusiones, agregaciones parciales, segregaciones…*" (Luis Ruiz-Maya Pérez, director general del INE en su momento, en García Fernández, 1985, Presentación, p. III), en base a la estructura de municipios vigente en la fecha del censo de 1981. En nuestra opinión estos son los datos que el INE debería haber difundido de

[17] Si realizamos la misma comparación que la ofrecida en el cuadro 1 pero para la población de derecho, entonces encontramos dos alteraciones de *ranking* con las poblaciones originales, pero ninguna de nuevo con las poblaciones homogéneas. Madrid en el mayor municipio en este caso en todos los años del siglo XX (Goerlich, Mas, Azagra y Chorén, 2006, Cuadro 2.7. pág.-78).

forma adecuada a través de medios electrónicos, y más modernamente vía *Internet*, en lugar de las *Series históricas de población de hecho municipal* que hemos comentado anteriormente y que contienen numerosos errores e imprecisiones, pero que por su fácil acceso han sido las más utilizadas por los investigadores.

Lamentablemente el trabajo de García Fernández (1985) no fue actualizado al censo de 1991, y adicionalmente el censo de 2001 dejó de investigar la población de hecho para centrarse en la población de derecho o residente. Por ello un trabajo riguroso sobre localización de la población a escala municipal que cubriera todo el siglo XX o un contraste estadístico de la ley de Zipf, como que el ahora nos ocupa, requería de una elaboración *ex-novo* del trabajo de García Fernández (1985) que tomara como referencia la estructura de términos municipales del censo más reciente, 2001, y como variable de estudio la población de derecho.

Por esta razón Goerlich, Mas, Azagra y Chorén (2006) elaboraron poblaciones municipales de derecho homogéneas a partir de *dos principios* básicos:

1. El criterio de asignación de poblaciones es un criterio territorial, los lindes municipales, y

2. El criterio que determina los territorios son los municipios existentes de acuerdo con el censo de 2001.

Los detalles de elaboración de dichas poblaciones homogéneas, que van desde 1900 hasta 2001, está descrito minuciosamente en Goerlich, Mas, Azagra y Chorén (2006), y donde la mencionada homogeneidad se refiere al mantenimiento de los lindes municipales existentes en el censo de 2001, reconstruyéndose hacia atrás las poblaciones de derecho de los 8.108 municipios que aparecen en el último censo efectuado en España.

Ello permite salvar: (*i*) todos los problemas de comparabilidad de poblaciones municipales debidas a las alteraciones, que hubieran sido importantes de utilizar las series originales de los censos; (*ii*) los problemas de inherentes a las *Series históricas de población de hecho municipal para el periodo 1900 – 1991*, disponibles en la *web* del INE, y que hemos comentado anteriormente; y (*iii*) el problema derivado de la necesidad de mezclar poblaciones de hecho con poblaciones de derecho, tal y como hacen, por ejemplo Lanaspa, Perdiguero y Sanz (2004).

Puesto que el censo de 2001 no investiga la población de hecho, los últimos datos sobre este concepto de población disponibles en las estadísticas oficiales son los del censo de 1991. Un análisis histórico que incorpore años más reciente requiere, por tanto, mezclar conceptos de población, o reconstruir ciertas poblaciones de derecho, ya que tradicionalmente la población que más se ha difundido históricamente en las estadísticas demográficas es la población de hecho. Para prologar su análisis más allá de 1991, Lanaspa, Perdiguero y Sanz (2004) utilizan los datos de población de derecho del padrón de 1999 (obviamente sin ningún tipo de homogeneización con datos anteriores). Aunque previsiblemente esta "mezcla" de poblaciones no afecte a los resultados, un análisis somero indica que, incluso a nivel provincial, ambos conceptos de población pueden experimentar divergencias por encima y por debajo del 10% (Goerlich, Mas, Azagra y Chorén 2006, Cuadro 2.5).[18] Por otra parte, algunos ejercicios simples de contabilidad demográfica sugieren que tampoco es conveniente mezclar poblaciones censales con poblaciones padronales (Goerlich 2007).

Por todo ello este trabajo parte, para el análisis de la distribución del tamaño de las ciudades, de las poblaciones de derecho municipales homogéneas procedentes de los censos de 1900 a 2001, donde la homogeneidad debe entenderse según la estructura de municipios existente en el último censo disponible. Nuestra utilización de dichas poblaciones, para definir las ciudades objeto de análisis, se explica brevemente en el apartado siguiente.

III. Sobre el concepto de ciudad

Cualquier análisis sobre el grado de urbanización o la distribución del tamaño de las ciudades requiere ser explícitos respecto a un concepto previo: el de *ciudad* (Goerlich y Mas 2007c). Se trata de un concepto cuya delimitación cuantitativa es difusa, como es bien conocido por los especialistas (de Vries 1990), si bien cualquier análisis referente a la población urbana depende de él.

Las ciudades pueden ser definidas de muchas formas, a partir de núcleos o a partir de umbrales mínimos referentes a poblaciones municipales. En este caso pueden

[18] No es cierto, como argumentan Lanaspa, Perdiguero y Sanz (2004, nota al pie 3, pág. 8), que sólo para 1991 se dispongan de las poblaciones de hecho y derecho. Existen poblaciones municipales de hecho y derecho en todos los censos desde 1877, en los censos de 1857 y 1860 sólo se dispone de las poblaciones de hecho y en el censo de 1842 sólo de la población de derecho. Sin embargo es cierto que, tradicionalmente, la información accesible con cierta facilidad anteriormente a 1991 ha sido solamente la población de hecho.

considerarse los municipios aisladamente o, por el contrario, considerar la posibilidad de que varios términos municipales puedan constituir una sola ciudad, como en el caso de las áreas metropolitanas o grandes áreas urbanas, donde el proceso de concentración de la población ha desbordado los lindes municipales en muchos casos. Por otra parte, en un análisis histórico de largo plazo, podemos analizar siempre las mismas *ciudades* a lo largo del tiempo, o por el contrario podemos establecer umbrales o definiciones (fijos o variables) de forma que el número de ciudades es diferente conforme transcurre el tiempo, unas aparecen y otras (menos frecuentemente) desaparecen. También es posible definir las ciudades a partir de un tamaño mínimo que englobe a un porcentaje dado del total de población. Todos estos criterios alternativos han sido considerados por la literatura (Cheshire 1999).

Estas dificultades a la hora de concretar el concepto de *ciudad* se acentúan en el caso español por la ausencia, en nuestra tradición estadística y administrativa, de una definición de áreas urbanas o metropolitanas (Feria 2004) que sirva de base para un consenso entre los especialistas del tema sobre el que basar un concepto operativo de *ciudad* (Capel 1975).[19]

Los municipios, que es la forma en la que la información nos viene dada, son sólo una aproximación al concepto de asentamiento de población, y resulta obvio que no es una buena aproximación en algunos casos. Las *ciudades* de Madrid, Barcelona, Valencia, Sevilla, Bilbao, o muchas otras, se extienden más allá de los términos municipales con el mismo nombre. Por ello, y aunque reconocemos que es difícil escapar en estos temas de los lindes municipales, nuestro análisis de la distribución del tamaño de las ciudades en España no se basará en los datos de población estrictamente municipal, sino en el concepto de área urbana recogido en el *Atlas estadístico de las áreas urbanas en España* del Ministerio de Fomento (2000), identificando *ciudad* como *área urbana* en dicho atlas.

El trabajo de definición de áreas urbanas del Ministerio de Fomento (2000) parte del Censo de 1991, del Padrón de 1996 y de determinados criterios que combinan umbrales mínimos de población, densidades, dinámicas demográficas, redes de

[19] En general esta tradición está ausente en Europa, frente a la costumbre de países como EEUU (*Office of Management and Budget* 2000) o Canadá (*Statistics Canada* 2002, Mendelson y Lefebvre 2003) en los que dicha tradición está fuertemente arraigada y diversas acepciones de áreas urbanas, metropolitanas y aglomeraciones de población se definen, con carácter cambiante en el tiempo, en los propios censos de población.

transporte y estructuras sectoriales.[20] A partir de esta información se determinan 68 *Grandes Áreas Urbanas* de más de 50.000 habitantes, 31 comprenden más de un término municipal y las restantes (37) solamente uno, englobando un total de 495 municipios. Además, el atlas determina 226 *Pequeñas Áreas Urbanas* entre los 10.000 y los 50.000 habitantes (estas son todas uni-municipales). Por su relevancia describimos brevemente el proceso de determinación de áreas urbanas (ciudades) del Ministerio de Fomento (2000), así como su actualización a nuestra fecha de referencia, la del Censo de 2001.

Para la determinación de las *Grandes Áreas Urbanas* se establece, en principio, un umbral mínimo de población de 50.000 habitantes, si bien la última *Gran Área Urbana* es Huesca, que en el momento de la clasificación (Padrón de 1996) es la única que contaba con una cifra inferior de población. Esto deja fuera de este grupo a algunas capitales de provincia, como Ávila, Cuenca o Soria. Puesto que las Comunidades Autónomas tienen competencias para estructurar su territorio, en seis Comunidades Autónomas (Andalucía, Principado de Asturias, Cataluña, Comunidad Valenciana, Comunidad de Madrid y País Vasco) se partió de trabajos de ordenación del territorio realizados en el seno de dichas comunidades, y que estaban encaminados básicamente a la definición de áreas metropolitanas. Las cinco *Grandes Áreas Urbanas* mayores son, por este orden, Madrid, Barcelona, Valencia, Sevilla y Bilbao. Sólo esta última tenía, en la fecha de referencia (Padrón de 1996), una población inferior al millón de habitantes (concretamente 914.542). Estas podrían ser consideradas como grandes *Áreas Metropolitanas*, con rasgos cuantitativos y cualitativos que les otorgan un elemento diferencial respecto al resto de áreas urbanas (Goerlich y Mas 2007c).

Para las *Pequeñas Áreas Urbanas* se fija en principio un umbral mínimo de población municipal de 10.000 habitantes. El Ministerio de Fomento (2000) distingue, dentro de este grupo, dos sub-estratos.

- Por una lado, los municipios entre 20.000 y 50.000 habitantes no incluidos en las *Grandes Áreas Urbanas*. Este sub-grupo incluye 102 municipios,[21] tres de los

[20] El Ministerio de Fomento (2000) menciona, además, un documento interno específico de marzo de 1995: *Necesidades de Suelo Urbanizado hasta el año 2011 en las Ciudades Españolas*, realizado por la Subdirección General de Urbanismo del entonces Ministerio de Obras Públicas, Transporte y Medio Ambiente.

[21] La publicación del Ministerio de Fomento (2000) indica que el número de municipios de este sub-estrato son 103, pero ello se debe a un error, al computar el municipio de Alfafar (Valencia) dos veces. Una vez dentro del *Gran Área Urbana de Valencia* y otra dentro del grupo de *Pequeñas Áreas Urbanas*. La población de derecho de este municipio en el censo de 1991 es de 20.151 habitantes.

cuales, Puertollano (13071, Ciudad Real), Orihuela (03099, Alicante), y Motril (18140, Granada), tienen en la fecha de referencia más de 50.000 habitantes.

- Por otro lado, los municipios entre 10.000 y 20.000 habitantes, si bien este conjunto por situarse en el límite del umbral es objeto de un tratamiento más detallado, ya que no todos los municipios de más de 10.000 habitantes pueden ser considerados como urbanos, y en consecuencia como ciudades a nuestros efectos.

Por esta razón, el conjunto inicial de los 219 municipios que tienen una población entre los 10.000 y los 20.000 habitantes es sometido a cuatro filtros que tratan de incorporar aspectos no relacionados directamente con el tamaño, pero que entran dentro de la acepción de *ciudad*:

1. Sólo se incluyen, de estos 219 municipios, aquéllos en los que la población del núcleo superase los 10.000 habitantes. Para ello se utilizan los datos del nomenclator de 1991 y se obtiene la población del núcleo como la población municipal total menos la "población en diseminado". Esto reduce los municipios potencialmente urbanos a 174.

2. Se eliminan aquéllos cuya evolución demográfica en el periodo 1960 – 1991 fuese negativa. Ello elimina 37 municipios, dejando un total de 137 municipios potencialmente urbanos.

3. De estos sólo se incluyen aquellos municipios cuyo porcentaje de población activa en el sector servicios (comercio y otros servicios) fuese al menos igual a la media en dicho sector del sub-estrato anterior. Esto reduce los municipios potencialmente urbanos a solamente 47.

4. Finalmente, y para incorporar municipios eminentemente turísticos, se construyó un indicador de "potencial de acogida" que tuviera en cuenta el número de viviendas secundarias.[22] Ello permitió recuperar 77 municipios anteriormente excluidos.

Por tanto, finalmente tenemos 124 municipios entre los 10.000 y los 20.000 habitantes que forman parte de las *Pequeñas Áreas Urbanas*. Si a estos sumamos los

[22] Dicho indicador se construye multiplicando el número de viviendas secundarias del municipio por el índice de ocupación media en España (3,1) y sumándole la población de derecho del municipio que reside en núcleo. La cifra resultante trata de medir la capacidad de acogida de dicho municipio para la población flotante estacional, de forma que todos los municipios que sobrepasen (arbitrariamente) el umbral de 15.000 habitantes son incluidos como área urbana.

102 del primer sub-estrato tenemos que las *Pequeñas Áreas Urbanas* engloban a 226 municipios.

En conjunto, la definición de *Área Urbana* del Ministerio de Fomento (2000) está constituida por 721 municipios englobados en 294 áreas urbanas. El resto, 7.378 municipios de los 8.099 que aparecen en el Padrón de 1996, son definidos simplemente como áreas no urbanas (lo que no significa necesariamente que sean rurales).[23]

El concepto de *ciudad* utilizado en este trabajo coincide con el *Área Urbana* del Ministerio de Fomento (2000), y se mantiene fijo para el periodo de estudio, 1900 – 2001, si bien debe ajustarse por las diferentes fechas de referencia. Entre el Padrón de 1996 y el censo de 2001 aparecieron diez nuevos municipios.[24] Los criterios para la clasificación de dichos municipios fueron los siguientes:

(1) Si la segregación del municipio en cuestión es de un municipio perteneciente a una de las *Grandes Áreas Urbanas* pluri-municipales definidas por el Ministerio, entonces pasa directamente a formar parte de dicha área urbana. Este es el caso de La Palma de Cervelló (08905), segregado del municipio de Cervelló (08068), perteneciente al área urbana de Barcelona, y de San Antonio de Benagéber (46903), segregado del municipio de Paterna (46190), perteneciente al área urbana de Valencia.

(2) En el resto de casos todos los municipios segregados tienen poblaciones inferiores a los 10.000 habitantes, tanto en 1991 como en 2001, y en consecuencia no pasaron a formar parte de ningún área urbana.[25]

[23] La publicación del Ministerio de Fomento (2000) indica que el total de municipios en 1996 es de 8.097, pero el listado ofrecido por el INE en su *web* para el Padrón de 1996 indica un total de 8.099. Dicha publicación contiene otros dos pequeños errores adicionales en el listado de *Pequeñas Áreas Urbanas*, el municipio de Orihuela es asignado a la provincia de Murcia y el municipio de Monóvar es asignado a la provincia de Almería, cuando en realidad ambos pertenecen a la provincia de Alicante.

[24] Dichos municipios son: Pozo Cañada (02901, Albacete), Pueblo Nuevo del Guadiana (06902, Badajoz), La Palma de Cervelló (08905, Barcelona), Arenales de San Gregorio (13903, Ciudad Real), Llanos del Caudillo (13904, Ciudad Real), Marchamalo (19171, Guadalajara), Arroyo del Ojanco (23905, Jaén), Riu de Cerdanya (25913, Lleida), San Cristóbal de Segovia (40906, Segovia) y San Antonio de Benagéber (46190, Valencia). Además el municipio de Darrícal (04039, Almería) que aparece en el Padrón de 1996 fue incorporado en 1997 a Alcolea (04007), de forma que no aparece en el censo de 2001.

[25] Algunos de ellos como Pozo Cañada (02901), segregado de Albacete (02003); Pueblo Nuevo del Guadiana (06902), segregado de Badajoz (06015); o Marchamalo (19171) segregado de Guadalajara (19130); proceden de municipios clasificados como *Grandes Áreas Urbanas*, pero dada su reducida dimensión y alejamiento del núcleo principal no se les consideró dentro del área urbana de estos municipios. En otros casos, como Arenales de San Gregorio (13903), segregado de Campo de Criptana (13028); o Llanos del Caudillo (13904), segregado de Manzanares (13053); ambos en Ciudad Real, proceden de municipios clasificados como *Pequeñas Áreas Urbanas*, pero tampoco se decidió su clasificación como áreas urbanas dentro del municipio de origen. Entre otras razones se prefirió mantener

En definitiva, los 8.108 municipios del censo de 2001 quedan agrupados en las 294 *Áreas Urbanas* definidas por el Ministerio de Fomento (2000), lo que supone un total de 723 municipios urbanos, y 7.385 áreas (o municipios) no urbanos. La población urbana es la asentada en estos 723 municipios urbanos. Las 294 *Áreas Urbanas* a las que pertenecen dichos municipios –de las cuales 31 son ciudades pluri-municipales y 261 uni-municipales- constituyen nuestro concepto de *ciudad* para el análisis de la distribución del tamaño de las ciudades.[26]

El apéndice estadístico ofrece información sobre la población que ha vivido en estas ciudades a lo largo del siglo XX, así como la población que vive en municipios mayores de 10.000 habitantes a efectos comparativos.

Una vez delimitada la fuente de información -las poblaciones de derecho homogéneas de los censos comprendidos entre 1900 y 2001- y nuestra definición de *ciudad* efectuaremos un contraste estadístico de la ley de Zipf (1949) complementario al efectuado por Lanaspa, Perdiguero y Sanz (2004).

IV. Un contraste de la ley de Zipf

Consideremos un conjunto de n ciudades ordenadas de acuerdo al tamaño de su población, x, de forma no creciente

$$x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(r)} \geq \cdots \geq x_{(n-1)} \geq x_{(n)} \tag{1}$$

así $x_{(1)}$ es la ciudad con mayor población y $x_{(n)}$ es la ciudad más pequeña.[27] Por tanto r representa el rango de la ciudad correspondiente.

La conocida como ley de Zipf (1949)[28] postula una relación muy estrecha entre r y $x_{(r)}$. En concreto el producto de ambas debe ser una constante, c, para todo r. Es decir,

---

las *Pequeñas Áreas Urbanas* constituidas por un único municipio. En el resto de casos, ni el municipio segregado, ni el municipio de origen, tenían una dimensión adecuada para su clasificación como áreas urbanas.

[26] El listado completo de municipios urbanos, y su clasificación en grandes o pequeñas áreas, así como los pertenecientes a las grandes áreas pluri-municipales, puede consultarse en la publicación mencionada del Ministerio de Fomento (2000), *Atlas estadístico de las áreas urbanas en España*.

[27] Obsérvese que (1) no son más que los estadísticos de orden del conjunto $\{x_i\}_{i=1}^{n}$, pero que el orden ha sido revertido respecto a la definición estándar para facilitar la formulación. De igual forma esta es la ordenación contraria respecto a la habitual en el análisis de la distribución de la renta (Goerlich y Villar 2007), a pesar de las similitudes analíticas entre ambos campos.

$$rx_{(r)} = c \qquad (2)$$

Obsérvese que para $r = 1$ obtenemos $x_{(1)} = c$, es decir c representa la ciudad de mayor tamaño. Por tanto un gráfico de r frente a $x_{(r)}$ debe producir una hipérbola rectangular. Alternativamente, y tomando logaritmos en (2), el gráfico de logaritmo del rango frente al logaritmo del tamaño debe producir una línea recta con pendiente igual a -1.

Aunque la ley de Zipf (1949) apareció como una regularidad empírica observada (al menos de forma aproximada) para un gran número de países y periodos de tiempo (Carroll 1982; de Vries 1984; Smith 1990), la literatura teórica reciente ha tratado de desarrollar modelos que sean capaces de generar estas regularidades empíricas observadas, descansen sobre mecanismos económicos plausibles y sean consistentes con otras características básicas de los sistemas de ciudades, como las economías de aglomeración y los costes de congestión (Eaton y Eckstein 1997; Brakman, Garretsen, Van Marrewijk y van den Berg 1999; Duranton 2002).

Estos modelos generan situaciones en las que las ciudades crecen de forma aleatoria (independientemente de su tamaño), con una media y varianza común, ya que como ha demostrado Gabaix (1999), si el proceso de crecimiento de las ciudades es homogéneo en este sentido, entonces la distribución límite del tamaño de las ciudades converge a la ley de Zipf (1949). Sin embargo un proceso de crecimiento homogéneo de este estilo (las ciudades creciendo de forma aleatoria a la misma tasa esperada, igual a la tasa de la ciudad promedio, y la misma varianza) es conocido en la literatura como la ley de Gibrat (1931) o ley de crecimiento proporcional (Sutton 1997). De esta forma Gabaix (1999) ha transformado la obscura regularidad de la ley de Zipf (1949), en una regularidad mucho más fácil de entender y explicar, la ley de Gibrat (1931), y desviaciones de la ley de Zipf (1949) pueden ser entendidas como desviaciones de la ley de crecimiento proporcional (Ioannides y Overman 2003).

Puesto que la versión determinista de la ley de Zipf (1949) postulada en (2) es poco probable que se verifique en la práctica, la literatura aplicada (Carroll 1982; Smith 1990; Eaton y Eckstein 1997; Soo 2002, Lanaspa, Perdiguero y Sanz 2004) ha tendido a contrastar la ley de Zipf (1949) a partir de la estimación de la ecuación lineal

$$\log r = \beta_1 + \beta_2 \log x_{(r)} + \varepsilon_r \qquad (3)$$

[28] El argumento al que hace referencia la ley de Zipf (1949) aparece ya en Auerbach (1913).

para a continuación examinar hasta qué punto es o no posible rechazar H_0: $\beta_2 = -1$. Desviaciones de la ley de Zipf (1949) representan, en este contexto, desviaciones de β_2 de -1. En este último caso β_2 puede tomar otro valor fijo, distinto de -1 (hablamos entonces de la *distribución rango-tamaño*, lo que guarda relación con la distribución de Pareto (1896) como veremos a continuación), o bien β_2 puede ser a su vez una función de $x_{(r)}$, y en consecuencia (3) estar incorrectamente especificada (Rosen y Resnick 1980; Fan y Casetti 1994).[29]

El gráfico correspondiente a la ecuación (3) para nuestras ciudades en 2001 se ofrece en el gráfico 1.[30] Las dos observaciones separadas del resto son Barcelona y Madrid. Resulta interesante constatar como estas aglomeraciones urbanas ya aparecen por encima del resto de la distribución en el siglo XIX (de Vries 1984, Capítulo 6), y en consecuencia parece que merecen un tratamiento diferencial.[31]

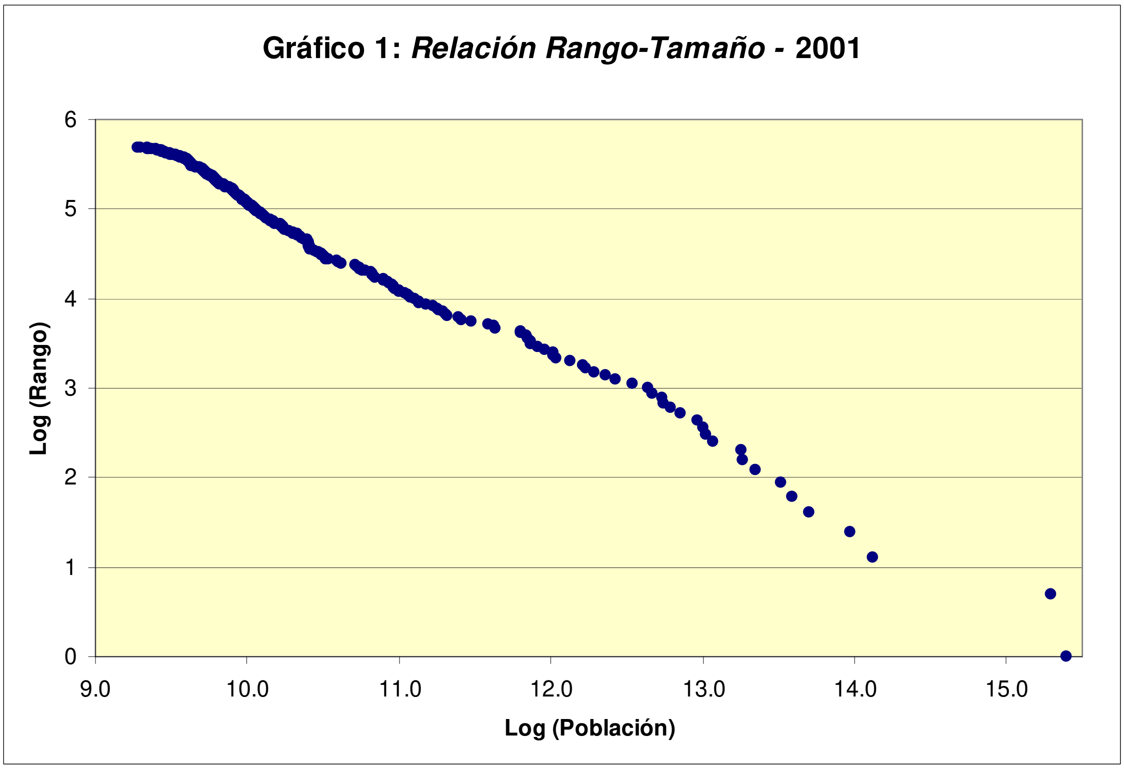
En cualquier caso, la estimación de la ecuación (3) por mínimos cuadrados ordinarios (OLS) para este conjunto de datos produce una estimación de $\hat{\beta}_2 = -0.90$, si bien la eliminación de las observaciones de Barcelona y Madrid acerca dicha estimación a la ley de Zipf (1949), $\hat{\beta}_2 = -0.96$.

---

[29] Existe algo de confusión terminológica en la literatura (Read 1988). En lo que a este trabajo se refiere entenderemos por ley de Zipf (1949) la versión estadística de $rx_{(r)} = c$, o en términos lineales, $\log r = \log c - \log x_{(r)}$. Mientras que hablaremos de la *distribución rango-tamaño* cuando (3) pueda seguir siendo considerado válido, pero para valores de β_2 distintos de -1.

[30] Mientras los historiadores urbanos (de Vries 1984; Smith 1990) tienden a dibujar las observaciones en el espacio (*rango* , *tamaño*); los economistas urbanos (Gabaix 1999; Duranton 2002, Lanaspa, Perdiguero y Sanz 2004) tienden a hacerlo en el espacio (*tamaño* , *rango*). Nosotros seguimos esta última práctica. Desde el punto de vista de la estimación de (3), y el contraste mediante esta ecuación de la ley de Zipf (1949), el procedimiento es irrelevante; aunque hay que prestar cierta atención cuando se habla de la concavidad o convexidad de la relación.

[31] Estas dos ciudades representan, en 2001, algo más del 30% del total de población urbana. La existencia de dos ciudades de tamaño similar, y claramente fuera de la distribución del resto, es una de las desviaciones de la ley de Zipf (1949) señaladas por Smith (1990). En el caso de Madrid, capital del estado, este "excesivo tamaño" tiene una clara justificación política (Ades y Glaeser 1995; Bahamonde y Otero 1999); el caso de Barcelona es, quizá, de más difícil explicación, pero esta hay que buscarla, sin duda, en los orígenes de la revolución industrial (Nadal 2003).



Nuestro objetivo no es, sin embargo, un análisis exhaustivo de la ecuación (3) para el caso español, esto ya ha sido realizado por Lanaspa, Perdiguero y Sanz (2004) salvando las diferencias metodológicas, sino profundizar algo más en las implicaciones estadísticas de (2). Puesto que r es un número entero, no es posible justificar la normalidad de la perturbación ε_r en (3) y presumiblemente podemos aumenta la eficiencia de nuestras estimaciones. Además, los resultados de Monte Carlo en Gabaix y Ioannides (2004) muestran los sesgos del estimador de mínimos cuadrados ordinarios (OLS) en (3) en muestras del tamaño habitual en trabajo aplicado, así como el sesgo a la baja en el estimador del error estándar de la estimación de OLS en (3), lo que afecta negativamente a la inferencia.

Diversos autores (Quandt 1964; Rapoport 1978; Alperovich 1988; Kamecke 1990; Urzúa 2000) han señalado con acierto que es necesario preguntarse por los fundamentos probabilísticos detrás de la relación (2) antes de proceder a estimar directamente la ecuación (3). Dicho de otra forma, es necesario traducir la relación rango-tamaño (2) en una relación frecuencia-tamaño. Consideremos en este sentido, sin pérdida de generalidad, que el tamaño de una ciudad, x, es una variable aleatoria continua y estrictamente positiva, con función de densidad de probabilidad $f(x)$. Dada una muestra aleatoria de tamaño n, $\{x_i\}_{i=1}^{n}$, y suponiendo que todos los tamaños son diferentes podemos escribir,

$$r_x = n.\int_x^{\infty} f(z).dz \qquad (4)$$

donde hemos añadido el subíndice x en r_x para enfatizar la dependencia del rango respecto al tamaño. Podemos ahora derivar el proceso probabilístico, $f(x)$, subyacente a (2). Substituyendo esta última relación en (4),

$$\frac{c}{x} = n.\int_x^{\infty} f(z).dz \qquad (5)$$

y derivando a ambos lados de la igualdad obtenemos,

$$f(x) = \frac{c}{n}.\frac{1}{x^2} \qquad (6)$$

Así pues, cualquier contraste eficiente de la ley de Zipf (1949), interpretada como (2), debe basarse en la ley de potencia recogida por la densidad (6).[32]

Una forma relativamente sencilla de hacer uso de este resultado es observar que la densidad (6) es un caso particular de la función de densidad de Pareto (1896), introducida por este autor para el estudio de la parte superior de la distribución de la renta y cuyas propiedades más notables se recogen en el apéndice. La densidad de Pareto (1896) viene dada por,

$$f_P(x) = \frac{\theta}{\mu}\left(\frac{\mu}{x}\right)^{\theta+1} \qquad x \geq \mu \qquad (7)$$

donde $\mu > 0$ es un parámetro de posición que puede interpretarse como el tamaño mínimo, y $\theta > 0$ es un parámetro de forma indicativo de la dispersión de la distribución. A mayores valores de θ se obtienen densidades más concentradas en las proximidades del mínimo, es decir menos dispersas. Aunque el espacio paramétrico está compuesto por dos parámetros, μ y θ, puesto que la densidad sólo está definida a partir de un valor

[32] Aunque la literatura ha tendido a enfatizar la ley potencial para el extremo superior de la distribución, algunos autores han argumentado que existen razones para esperar un comportamiento similar en la cola inferior (Reed 2001).

mínimo, $\mu > 0$, este parámetro suele fijarse en el mínimo observado en la muestra, $\hat{\mu} = x_{(n)}$, lo que equivale a realizar el análisis condicionado en este valor.[33]

La ley de Zipf se obtiene fijando $\theta = 1$ en (7), en cuyo caso $\mu = \frac{c}{n}$, o de forma equivalente $nx_{(n)} = x_{(1)}$, tal y como requiere (2).[34] Así pues, un contraste sencillo en este contexto de la ley de Zipf (1949) puede derivarse fácilmente bajo la hipótesis mantenida de que se cumple la ley de Pareto (1896). El estimador de máxima verosimilitud (MLE) de θ en (7) puede obtenerse con facilidad (Hill 1975) y un contraste de H_0: $\theta = 1$ puede realizarse por los métodos habituales. Aunque esta aproximación es seguida por algunos autores (Kamecke 1990; Soo 2002), se trata de una opción poco robusta, ya que considera como hipótesis mantenida la distribución de Pareto (1896).

El cuadro 2 muestra los resultados de la estimación MLE de θ en (7), así como la *t*-ratio del contraste de H_0: $\theta = 1$ *versus* H_1: $\theta \neq 1$, y su nivel de significación asociado. Los resultados muestran que el exponente de Pareto se mantiene relativamente estable en la primera mitad del siglo XX (en un entorno próximo al 0,4), para mostrar una decidida tendencia creciente en la segunda mitad del siglo XX. En todos los casos, sin embargo, el coeficiente estimado se mantiene por debajo de la unidad. Incluso cuando dicho coeficiente alcanza su valor máximo, en 2001 con $\hat{\theta} = 0,862$, un contraste formal de la hipótesis de la ley de Zipf nos permite rechazarla con un amplio margen de confianza.

Cuadro 2. Estimación Máximo Verosímil de q en (7) y contraste de q = 1. 1900 - 2001

	1900	**1910**	**1920**	**1930**	**1940**	**1950**	**1960**	**1970**	**1981**	**1991**	**2001**
MLE de θ	0,365	0,374	0,383	0,379	*0,364*	0,411	0,496	0,474	0,473	0,703	**0,862**
Estadístico del contraste, H_0: θ = 1	-29,85	-28,69	-27,68	-28,11	*-29,92*	-24,60	-17,40	-19,04	-19,10	-7,25	**-2,75**
Nivel de significación (2 colas)	0,0000	0,0000	0,0000	0,0000	*0,0000*	0,0000	0,0000	0,0000	0,0000	0,0000	**0,0059**

Nota: En ***cursiva*** el valor **mínimo** del periodo. En **negrita** el valor **máximo** del periodo. El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

[33] El valor mínimo de la muestra es, además, el estimador de máxima verosimilitud (MLE) de μ, aunque dada la dependencia entre el espacio muestral y el espacio paramétrico en este caso no puede obtenerse dicha estimación mediante cálculo, ni tiene las propiedades habituales de estos estimadores.

[34] La dependencia de los parámetros de la distribución respecto a *n* muestra que la ley de Zipf sólo se verificará, en el mejor de los casos, para un determinado tamaño muestral.

Existen, al menos, dos formas adicionales de utilizar el resultado anterior. En primer lugar, sería posible considerar varias distribuciones alternativas a la distribución de Pareto (1896) y una vez ajustadas a los datos buscar aquella que proporcionara el mejor ajuste según un determinado criterio (Quandt 1964, Alperovich y Deutsch 1995). Aunque no seguiremos este procedimiento el cuadro 3 ofrece una comparación de los datos del tamaño de las ciudades frente a una distribución lognormal. Claramente esta no es una aproximación razonable para la descripción del tamaño de las ciudades. El estadístico de Jarque y Bera (1987), basado en la asimetría y el exceso de curtosis en las observaciones,[35] rechaza claramente esta distribución. El rechazo es más fuerte al final del periodo que al principio, lo que indica un progresivo alejamiento de esta distribución. Además, el porcentaje de observaciones por encima de la moda de la distribución lognormal estimada esta comprendido entre un 87,1%, a principios del siglo XX, y el 99,3% a finales, indicando claramente como la práctica totalidad de las observaciones se encuentran en la parte decreciente de la densidad y, en consecuencia, los datos se aproximan más a la ley potencial que a ley lognormal. Adicionalmente, obsérvese como el índice de Gini derivado de la distribución lognormal es sustancialmente menor que el índice muestral calculado directamente a partir de los datos, por lo que esta distribución tiende a mostrar una menor dispersión de la realmente observada en los datos.

En segundo lugar, podemos buscar una densidad que anide la ley de Pareto representada por (7) y, en consecuencia, derivar a partir de ella un contraste paramétrico por los métodos habituales en estadística (Urzúa 2000). Nuestra aplicación final sigue esta última opción.

Considérese la siguiente densidad, que es un miembro particular de la familia de distribuciones de Burr (1942) (Johnson y Kotz 1970, pág. 31),

$$f_B(x) = \frac{\theta}{\sigma}\left(1 + \frac{x-\mu}{\sigma}\right)^{-(\theta+1)} \qquad x \geq \mu \tag{8}$$

[35] Este no es más que el estadístico de normalidad aplicado a los logaritmos de las observaciones.

Cuadro 3. Contraste de Lognormalidad: Jarque y Bera. 1900 - 2001.

	1900	**1910**	**1920**	**1930**	**1940**	**1950**	**1960**	**1970**	**1981**	**1991**	**2001**
Estadístico de Jarque y Bera de lognormalidad	*136,41*	143,67	159,29	180,29	161,62	160,48	226,56	315,90	330,94	**340,65**	337,18
Nivel de significación (Chi-Cuadrado, 2)	**0,000**	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	*0,000*	0,000
Moda de la distribución lognormal estimada	*4.400*	4.973	5.275	5.669	5.934	6.171	7.022	7.773	8.693	9.724	**10.900**
Tamaño de ciudad más pequeño, $x(n)$	*753*	894	1.039	1.118	1.133	1.677	2.886	3.018	3.484	7.556	**10.728**
Observaciones mayor que la Moda	*256*	*256*	*256*	261	262	267	273	281	289	**292**	**292**
% observaciones mayor que la Moda	*87,1%*	*87,1%*	*87,1%*	88,8%	89,1%	90,8%	92,9%	95,6%	98,3%	**99,3%**	**99,3%**
Gini de la distribución lognormal estimada	0,515	*0,511*	0,518	0,524	0,540	0,548	0,547	0,555	**0,561**	0,556	0,551
Gini muestral	0,644	*0,642*	0,659	0,677	0,692	0,706	0,725	0,756	**0,766**	0,760	0,753

Nota: En ***cursiva*** el valor **mínimo** del periodo. En **negrita** el valor **máximo** del periodo. El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

Esta es una distribución más general que la de Pareto, (7), en el sentido de que si σ = μ, entonces $f_P(x) = f_B(x)$. En consecuencia un contraste de la ley de Zipf (1949) en este contexto puede reformularse como un contraste de la hipótesis H_0: σ = μ y θ = 1. Al igual que antes, μ se fija al valor mínimo observado en la muestra, por lo que tenemos sólo dos parámetros a estimar en (8). Lamentablemente la función de verosimilitud asociada a (8) no admite solución explicita por lo que el estimador de máxima verosimilitud (MLE) de σ y θ debe ser obtenido por métodos numéricos. Sin embargo, Urzúa (2000) muestra como un contraste de los multiplicadores de Lagrange (LM) puede ser fácilmente calculado, requiriendo sólo para su cálculo momentos muestrales.

En concreto, algo de algebra (ver Apéndice) muestra que un contraste LM de H_0: σ = μ y θ = 1, puede realizarse a partir del estadístico (Urzúa 2000),

$$LM_Z = 4.n\left[z_1^2 + 6z_1z_2 + 12z_2^2\right] \overset{asy}{\sim} \chi^2(2) \qquad \text{bajo } H_0 \tag{9}$$

donde $z_1 = 1 - \frac{1}{n}\sum_{i=1}^{n}\log\frac{x_i}{x_{(n)}}$ y $z_2 = \frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n}\frac{x_{(n)}}{x_i}$. Bajo H_0: $\sigma = \mu$ y $\theta = 1$, LM_z se distribuye asintóticamente como una χ^2 con dos grados de libertad, constituyendo un contraste conjunto de distribución Paretiana y de exponente de Pareto igual a la unidad. La evidencia mostrada en Urzúa (2000) indica que el contraste tiene buen comportamiento en muestras mucho más pequeñas qua las utilizadas en nuestra aplicación.

El contraste LM_z para nuestro concepto de ciudad se ofrece en el cuadro 4. De nuevo los resultados son concluyentes, la ley de Zipf (1949) se rechaza por un amplio margen. Puesto que este contraste lleva incorporado la distribución de Pareto observamos como esta distribución tampoco parece ser una buena aproximación a los datos, aunque estrictamente hablando no podemos determinar el origen del rechazo.

Como ha mostrado Gabaix (1999) desviaciones de la ley de Zipf (1949) pueden interpretarse como desviaciones de la ley de crecimiento proporcional. En consecuencia podemos concluir que, en promedio, las ciudades no han tendido a crecer a tasas similares a lo largo del siglo XX.

Cuadro 4. Contraste LM de la ley de Zipf. 1900 - 2001.

	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Contraste LM de la ley de Zipf	904,46	827,96	767,81	793,36	**911,16**	606,28	338,30	394,78	397,99	95,19	*13,92*
Nivel de significación (Chi-Cuadrado, 2)	0,0000	0,0000	0,0000	0,0000	*0,0000*	0,0000	0,0000	0,0000	0,0000	0,0000	**0,0010**

Nota: En ***cursiva*** el valor **mínimo** del periodo. En **negrita** el valor **máximo** del periodo. El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

V. Conclusiones

Este trabajo ha examinado el cumplimiento de la ley de Zipf (1949) para el sistema urbano español a lo largo del siglo XX. Los mensajes fundamentales del mismo son tres.

1. Los datos habitualmente utilizados para este tipo de trabajos, y que proceden fundamentalmente de las *Series históricas de población de hecho municipal para el periodo 1900 – 1991* de la *web* del INE, contienen numerosas imprecisiones y errores, y ni se trata de las series originales de los censos (actualmente también disponibles), ni se trata de poblaciones homogéneas bajo ningún criterio. En consecuencia lo mejor que puede hacerse con dichas estadísticas es ignorar su uso, ya que existen hoy día fuentes alternativas de mejor calidad: García (1985), Goerlich, Mas, Azagra y Chorén (2006) o los datos originales de los censos recopilados por el propio INE en su *web* en: *Alteraciones de los municipios en los Censos de Población desde 1842*.

2. El concepto de área urbana, ausente en nuestra tradición estadística y administrativa, debe ser puesto en la lista de prioridades de la investigación en geografía y economía urbana. Sería conveniente disponer de un listado de áreas urbanas y metropolitanas sobre el que hubiera cierto consenso entre la profesión, y que permitiera cierta homogeneidad y comparabilidad en los estudios realizados sobre el tema.[36] En ausencia de dicho listado se ha optado por asimilar el concepto

[36] En este sentido es de destacar la que parece ser la única iniciativa independiente e individual, apoyada en un modelo y que además tienen en cuenta el *conmuting*, en la definición de áreas urbanas en España a partir de criterios claros y totalmente públicos, y a partir de aquí la determinación de las áreas metropolitanas como áreas urbanas de mayores dimensiones y con características peculiares: el proyecto

de ciudad al de la definición de área urbana del Ministerio de Fomento (2000), y se ha mantenido dicho concepto a lo largo del periodo de análisis.

3. Finalmente, se ha realizado un contraste de la ley de Zipf (1949) enmarcado en un contexto probabilístico. Los resultados son bastante contundentes, la ley de Zipf (1949) puede ser rechazada en todos los años analizados con un amplio margen de confianza, tanto cuando utilizamos la distribución de Pareto (1896) cómo hipótesis mantenida, cómo cuando generalizamos dicha distribución.[37] Los datos parecen tener una estructura bastante más compleja de la que sugieren estos modelos de comportamiento, y con dos grandes ciudades, Madrid y Barcelona, sobresaliendo por encima del resto de forma mantenida en todos los años considerados.[38] La conclusión de este rechazo de la ley de Zipf es que las ciudades no han presentado un proceso de crecimiento homogéneo, sino por el contrario comportamientos heterogéneos, tanto por periodos como por grupos de ciudades, especialmente en la segunda mitad del siglo XX.

   Así, aunque el índice de Gini para el tamaño de las ciudades muestra un ligero crecimiento de la dispersión relativa hasta 1970 y un mantenimiento de la misma a partir de entonces, si distinguimos entre los dos grupos del Ministerio de Fomento: las grandes ciudades (*Grandes Áreas Urbanas*) por una parte y las pequeñas ciudades (*Pequeñas Áreas Urbanas*) por otra, observamos que las primeras prácticamente mantienen su dispersión, mostrando un aumento muy ligero de la misma durante los dos primeros tercios del siglo XX, mientras que las segundas muestran una estabilidad notable en la dispersión relativa durante la primera mitad del siglo XX y una tendencia decreciente en dicha dispersión en la segunda mitad.

AUDES5, Áreas Urbanas de España 2005 (http://alarcos.inf-cr.uclm.es/per/fruiz/audes5/), que intenta establecer un listado completo de áreas urbanas de España, con fecha de referencia 2005, indicando, para cada una de ellas, los municipios que la integran.

[37] La distribución lognormal tampoco parece ser una aproximación razonable a los datos.

[38] Aunque estas dos ciudades se destacan por encima del resto en todo el periodo analizado la importancia relativa de ambas si se altera a lo largo del siglo XX, en términos de nuestra definición de ciudades, que no de municipios, Madrid sólo alcanza la supremacía en el último tercio del siglo XX. Por el contrario en términos de poblaciones de derecho municipales Madrid se sitúa por delante de Barcelona en todos los años de nuestras series homogéneas.

Apéndice estadístico

Este apéndice estadístico ofrece información adicional, a nivel provincial, sobre la población que vive en las ciudades tal y como han sido definidas en este trabajo, y acudiendo a un criterio alternativo: los municipios de más de 10.000 habitantes en cada momento censal. Información más exhaustiva puede encontrarse en Goerlich y Mas (2007c).

Los cuadros que se ofrecen son los siguientes:

- Cuadro A1. Efectivos de población en ciudades. 1900 a 2001.
- Cuadro A2. Porcentaje provincial de población en ciudades. 1900 a 2001.
- Cuadro A3. Municipios con más de 10.000 habitantes. 1900 a 2001.
- Cuadro A4. Población en municipios de más de 10.000 habitantes. 1900 a 2001.
- Cuadro A5. Porcentaje provincial de población en municipios de más de 10.000 habitantes. 1900 a 2001.
- Cuadro A6. Índices de Gini según determinados criterios y agrupaciones. 1900 a 2001.
- Cuadro A7. Tamaños medios según determinados criterios y agrupaciones. 1900 a 2001.
- Cuadro A8. Tamaños medianos según determinados criterios y agrupaciones. 1900 a 2001.
- Cuadro A9. Desviación típica según determinados criterios y agrupaciones. 1900 a 2001.

Cuadro A1. Población en ciudades. 1900 - 2001

	Provincia	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	**Álava**	*36.484*	38.291	40.207	44.547	52.340	54.174	77.029	149.448	210.429	226.591	**235.783**
02	**Albacete**	*53.800*	65.103	75.741	94.016	120.668	134.363	134.235	150.016	177.630	194.903	**222.743**
03	**Alicante/Alacant**	*277.597*	304.540	318.770	343.457	393.156	420.159	499.837	695.572	903.815	1.028.504	**1.157.363**
04	**Almería**	*89.550*	90.315	90.458	94.827	120.392	123.316	145.768	184.446	229.089	268.048	**327.420**
05	**Ávila**	*13.457*	13.458	15.368	17.036	20.903	23.598	28.032	31.534	40.173	45.977	**49.712**
06	**Badajoz**	*96.706*	108.814	123.082	130.093	148.594	180.471	216.984	228.796	248.007	272.389	**297.447**
07	**Balears (Illes)**	*160.837*	173.586	183.746	200.312	239.135	253.660	282.216	369.681	482.645	519.665	**616.572**
08	**Barcelona**	*919.139*	993.991	1.186.239	1.558.437	1.760.886	2.038.534	2.636.755	3.699.336	4.393.824	4.418.863	**4.548.716**
09	**Burgos**	*45.731*	46.669	49.831	61.402	78.306	94.553	123.879	170.265	217.420	226.642	**232.054**
10	**Cáceres**	*28.873*	31.696	38.436	45.050	55.702	70.063	84.105	101.580	120.557	136.846	**146.729**
11	**Cádiz**	*340.200*	343.068	403.189	395.237	467.035	560.266	671.486	749.469	869.703	956.168	**993.508**
12	**Castellón/Castelló**	*101.636*	106.045	107.902	113.216	128.611	143.667	170.454	230.955	285.926	307.099	**341.885**
13	**Ciudad Real**	*107.077*	129.747	156.960	181.963	208.815	218.364	235.117	232.501	237.124	249.071	**257.906**
14	**Córdoba**	*173.614*	191.021	225.885	282.084	349.922	382.561	401.222	420.700	459.514	492.771	**506.041**
15	**Coruña (A)**	*234.279*	249.969	273.101	304.209	351.042	406.490	464.109	502.058	595.165	630.119	**654.532**
16	**Cuenca**	*12.630*	13.748	15.186	17.818	23.554	26.091	28.053	34.290	40.007	42.817	**46.341**
17	**Girona**	*73.164*	83.607	85.534	92.760	97.477	100.347	122.989	167.466	220.181	249.539	**272.057**
18	**Granada**	*198.892*	219.230	250.183	287.282	343.276	367.220	365.044	402.654	483.986	531.932	**576.895**
19	**Guadalajara**	*12.662*	13.054	15.459	17.734	20.554	20.619	23.265	32.105	52.313	60.114	**68.248**
20	**Guipúzcoa**	*98.364*	117.050	149.509	175.764	198.356	235.662	307.545	422.232	**480.025**	470.777	473.100
21	**Huelva**	*60.672*	72.162	86.720	99.177	113.923	121.265	138.787	164.364	206.472	229.144	**243.310**
22	**Huesca**	*35.112*	36.033	38.797	39.526	40.099	46.476	54.018	69.738	81.928	84.473	**87.614**
23	**Jaén**	*161.870*	171.531	194.085	216.182	251.854	273.143	274.289	277.670	304.182	318.743	**330.665**
24	**León**	*41.983*	44.715	53.108	65.156	90.457	119.071	151.743	193.326	223.841	**252.798**	243.846
25	**Lleida**	*31.210*	36.029	49.501	51.021	49.496	65.887	77.779	107.748	128.050	133.700	**135.934**
26	**Rioja (La)**	*28.269*	31.894	38.998	44.634	57.264	63.263	73.830	99.212	127.393	141.035	**153.586**
27	**Lugo**	*54.000*	65.000	57.583	61.528	77.908	88.094	92.761	96.668	107.638	118.658	**122.745**
28	**Madrid**	*649.464*	694.748	909.529	1.141.978	1.426.027	1.668.560	2.347.899	3.598.313	4.500.020	4.709.785	**5.014.495**
29	**Málaga**	*283.081*	289.621	313.563	349.028	412.078	466.466	496.292	611.209	814.052	958.204	**1.083.497**
30	**Murcia**	*457.418*	486.819	510.999	522.220	590.969	613.566	657.328	687.576	804.906	887.863	**1.019.232**
31	**Navarra**	*56.034*	57.603	62.710	68.033	85.232	103.501	134.437	216.413	272.758	289.773	**321.899**
32	**Ourense**	*34.619*	38.297	42.541	47.763	56.211	73.916	81.676	93.163	115.436	125.171	**132.948**
33	**Asturias**	*347.065*	399.817	456.297	510.016	559.852	598.865	716.260	815.229	**921.433**	912.473	901.241
34	**Palencia**	*15.610*	17.710	20.016	24.332	32.469	41.122	47.902	56.816	71.716	77.863	**79.797**
35	**Palmas (Las)**	*105.732*	140.050	151.592	182.461	238.586	289.959	361.127	464.131	621.025	666.553	**747.193**
36	**Pontevedra**	*179.053*	207.245	237.015	269.866	306.303	327.792	347.446	415.886	507.203	531.222	**544.886**
37	**Salamanca**	*43.687*	49.111	51.048	64.008	90.633	105.454	122.045	155.502	189.130	**201.712**	197.572
38	**Sta. Cruz de Tenerife**	*111.460*	151.868	154.423	183.925	217.043	271.747	340.592	408.083	504.821	559.743	**618.660**
39	**Cantabria**	*85.321*	96.039	111.611	123.194	144.803	153.264	178.161	230.633	280.541	**297.879**	294.332
40	**Segovia**	*17.088*	17.892	18.347	21.062	24.772	30.043	35.212	43.099	50.759	**54.375**	54.368
41	**Sevilla**	*338.380*	356.882	442.863	499.767	636.993	762.219	900.376	1.026.808	1.189.052	1.316.202	**1.412.707**
42	**Soria**	*7.928*	8.649	8.316	10.788	13.897	16.753	19.799	24.659	30.326	32.360	**35.151**
43	**Tarragona**	103.761	*99.737*	115.509	119.126	124.796	140.383	153.294	221.430	295.311	316.819	**345.510**
44	**Teruel**	*20.444*	24.229	23.741	25.345	25.127	30.149	31.966	32.951	37.586	41.129	**44.589**
45	**Toledo**	*34.364*	35.556	40.974	43.340	49.242	61.599	74.318	90.602	119.175	128.502	**143.751**
46	**Valencia/València**	*483.059*	534.888	572.038	679.007	883.470	956.697	1.049.214	1.376.339	1.658.392	1.705.306	**1.763.538**
47	**Valladolid**	*81.846*	86.053	91.447	108.452	128.602	140.021	171.727	256.654	347.663	**364.825**	361.585
48	**Vizcaya**	*198.805*	233.937	298.531	355.279	387.375	430.020	619.086	895.283	**1.033.920**	1.003.321	969.769
49	**Zamora**	*21.403*	22.642	24.240	27.124	36.191	45.285	53.195	60.713	71.351	78.900	**81.435**
50	**Zaragoza**	*118.902*	131.338	163.159	188.877	235.669	275.249	336.757	505.397	611.036	634.834	**660.868**
51	**Ceuta**	*13.843*	24.249	35.453	50.293	65.982	56.909	64.728	62.607	65.264	67.615	**71.505**
52	**Melilla**	*10.182*	40.929	53.577	69.133	69.684	**76.247**	72.430	60.843	53.593	56.600	66.411
	España	*7.306.357*	**8.036.275**	**9.233.117**	**10.718.885**	**12.701.731**	**14.367.163**	**17.294.598**	**22.394.169**	**27.093.506**	**28.626.415**	**30.309.691**

Nota : En ***cursiva*** el valor **mínimo** de cada provincia en el periodo.
En **negrita** el valor **máximo** de cada provincia en el periodo.

Fuente : INE y elaboración propia.

Cuadro A2. Porcentaje provincial de población en ciudades. 1900 - 2001

	Provincia	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	**Álava**	*37,2%*	38,5%	39,7%	42,1%	46,5%	47,5%	57,6%	74,8%	81,6%	**83,2%**	82,3%
02	**Albacete**	*22,3%*	24,0%	25,5%	28,1%	31,8%	33,5%	35,8%	44,0%	52,3%	56,9%	**61,1%**
03	**Alicante/Alacant**	*58,5%*	59,4%	60,6%	62,0%	64,7%	66,2%	69,6%	75,4%	78,6%	**79,6%**	79,2%
04	**Almería**	24,5%	*22,7%*	23,6%	26,3%	32,2%	34,1%	39,5%	48,8%	55,8%	58,8%	**61,0%**
05	**Ávila**	6,5%	*6,1%*	6,8%	7,2%	8,5%	9,1%	11,1%	14,9%	21,9%	26,4%	**30,4%**
06	**Badajoz**	18,6%	*18,3%*	18,9%	18,5%	19,9%	22,1%	25,5%	32,6%	38,5%	41,9%	**45,4%**
07	**Balears (Illes)**	*50,8%*	51,7%	52,4%	53,4%	58,1%	60,4%	63,9%	69,4%	**73,6%**	73,3%	73,3%
08	**Barcelona**	*87,3%*	87,5%	88,5%	90,2%	91,0%	92,0%	92,9%	94,5%	**95,0%**	94,9%	94,6%
09	**Burgos**	13,3%	*13,3%*	14,4%	17,0%	20,7%	24,3%	31,9%	47,2%	59,8%	64,2%	**66,5%**
10	**Cáceres**	8,1%	*8,0%*	9,3%	10,0%	11,0%	12,8%	15,1%	21,7%	28,6%	33,3%	**36,4%**
11	**Cádiz**	77,9%	*76,7%*	78,4%	77,3%	79,1%	80,8%	82,6%	85,3%	88,0%	88,7%	**89,0%**
12	**Castellón/Castelló**	32,1%	*31,9%*	33,3%	35,8%	40,3%	43,7%	49,9%	59,9%	66,3%	68,8%	**70,6%**
13	**Ciudad Real**	*33,6%*	34,5%	36,7%	37,0%	38,8%	38,1%	39,9%	45,3%	49,9%	52,4%	**53,8%**
14	**Córdoba**	*38,9%*	39,2%	40,7%	42,3%	46,0%	48,4%	49,9%	57,5%	63,7%	65,3%	**66,4%**
15	**Coruña (A)**	*34,4%*	34,4%	35,4%	36,4%	38,5%	41,8%	44,8%	48,7%	54,4%	57,4%	**59,7%**
16	**Cuenca**	*5,0%*	5,1%	5,3%	5,7%	6,9%	7,6%	8,5%	13,6%	18,5%	20,9%	**23,1%**
17	**Girona**	*24,1%*	25,8%	25,9%	28,0%	30,0%	31,1%	35,0%	40,6%	47,1%	**49,0%**	48,1%
18	**Granada**	*40,2%*	41,6%	43,1%	43,8%	45,9%	46,3%	47,0%	54,3%	63,8%	67,3%	**70,2%**
19	**Guadalajara**	6,2%	*6,1%*	7,3%	8,3%	9,7%	9,9%	12,3%	21,4%	36,5%	**41,3%**	39,0%
20	**Guipúzcoa**	*50,1%*	52,9%	57,4%	59,3%	61,0%	63,5%	64,9%	67,4%	69,1%	69,6%	**70,2%**
21	**Huelva**	*23,5%*	24,4%	26,2%	27,9%	30,4%	32,8%	34,3%	40,7%	49,3%	51,7%	**52,6%**
22	**Huesca**	13,8%	*13,6%*	14,6%	15,3%	16,2%	19,6%	23,1%	31,4%	38,1%	40,6%	**42,4%**
23	**Jaén**	34,4%	32,7%	32,9%	*32,1%*	33,1%	35,0%	36,7%	41,6%	47,5%	50,0%	**51,4%**
24	**León**	*10,5%*	10,7%	12,2%	14,1%	17,8%	21,6%	25,7%	34,4%	42,7%	48,1%	**49,9%**
25	**Lleida**	*11,0%*	12,2%	15,2%	16,0%	16,0%	20,4%	23,2%	31,0%	36,3%	**37,8%**	37,5%
26	**Rioja (La)**	*14,8%*	16,7%	19,6%	21,5%	25,4%	27,4%	31,9%	42,3%	50,1%	53,5%	**55,5%**
27	**Lugo**	11,3%	12,7%	*11,1%*	11,7%	14,3%	16,9%	18,9%	22,8%	26,6%	30,9%	**34,3%**
28	**Madrid**	84,0%	*83,6%*	86,7%	88,5%	90,6%	91,5%	93,5%	95,7%	**96,0%**	95,2%	92,5%
29	**Málaga**	*54,4%*	54,7%	55,7%	57,3%	59,9%	61,7%	63,5%	71,6%	79,4%	82,5%	**84,2%**
30	**Murcia**	78,7%	78,4%	*78,1%*	80,1%	80,8%	81,2%	81,9%	82,6%	84,2%	84,9%	**85,1%**
31	**Navarra**	18,1%	*17,8%*	18,5%	19,3%	23,4%	27,0%	33,0%	46,4%	53,6%	55,8%	**57,9%**
32	**Ourense**	*8,2%*	8,6%	9,1%	10,0%	11,4%	15,0%	16,7%	21,1%	26,8%	35,4%	**39,3%**
33	**Asturias**	*54,4%*	56,3%	58,6%	61,1%	64,4%	66,9%	72,0%	77,5%	81,6%	83,4%	**84,8%**
34	**Palencia**	*8,1%*	9,0%	10,2%	11,5%	14,6%	17,4%	20,1%	28,2%	38,0%	42,0%	**45,8%**
35	**Palmas (Las)**	*66,9%*	70,7%	70,5%	71,6%	74,0%	76,3%	78,6%	84,5%	**87,6%**	86,8%	84,2%
36	**Pontevedra**	*36,1%*	37,9%	40,5%	42,8%	44,7%	45,9%	48,6%	53,2%	57,4%	59,2%	**60,3%**
37	**Salamanca**	*13,4%*	13,9%	14,9%	17,9%	22,8%	25,4%	29,3%	40,9%	51,9%	56,4%	**57,2%**
38	**Sta. Cruz de Tenerife**	*54,0%*	58,3%	56,5%	57,9%	59,4%	63,5%	67,2%	70,8%	76,6%	**77,1%**	76,7%
39	**Cantabria**	*30,6%*	31,2%	32,5%	33,3%	35,9%	37,8%	41,2%	49,2%	54,7%	**56,5%**	55,0%
40	**Segovia**	10,5%	*10,3%*	10,5%	11,5%	12,8%	14,8%	17,4%	26,6%	34,0%	**36,9%**	36,8%
41	**Sevilla**	61,3%	*60,4%*	62,9%	63,1%	66,5%	69,2%	72,4%	76,8%	80,4%	81,3%	**81,8%**
42	**Soria**	*5,1%*	5,3%	5,2%	6,6%	8,4%	10,2%	13,0%	21,0%	30,1%	34,2%	**38,7%**
43	**Tarragona**	30,2%	*29,1%*	32,1%	33,9%	36,0%	39,3%	42,2%	51,1%	57,6%	**58,5%**	56,7%
44	**Teruel**	*8,1%*	9,1%	9,0%	9,6%	10,2%	12,4%	14,3%	19,0%	24,5%	28,6%	**32,8%**
45	**Toledo**	9,0%	*8,5%*	9,1%	8,7%	10,0%	11,5%	14,0%	19,0%	25,1%	26,2%	**26,6%**
46	**Valencia/València**	*59,8%*	60,2%	61,2%	65,0%	69,5%	71,1%	72,9%	77,7%	80,3%	**80,5%**	79,6%
47	**Valladolid**	*28,9%*	29,4%	31,1%	34,8%	38,6%	40,2%	46,7%	62,1%	72,2%	**73,8%**	72,6%
48	**Vizcaya**	*64,6%*	66,4%	70,9%	73,6%	75,9%	77,6%	82,4%	86,0%	**86,9%**	86,9%	86,4%
49	**Zamora**	*7,6%*	7,8%	8,3%	9,3%	11,9%	14,3%	17,2%	23,5%	31,3%	36,9%	**40,9%**
50	**Zaragoza**	*28,2%*	29,2%	32,7%	35,6%	41,0%	45,2%	52,5%	66,7%	73,7%	75,8%	**76,7%**
51	**Ceuta**	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*
52	**Melilla**	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*
	España	*38,8%*	**39,5%**	**41,9%**	**44,6%**	**48,1%**	**51,0%**	**56,2%**	**65,8%**	**71,9%**	**73,6%**	**74,2%**

Nota : En ***cursiva*** el valor **mínimo** de cada provincia en el periodo.
En **negita** el valor **máximo** de cada provincia en el periodo.

Fuente : INE y elaboración propia.

Cuadro A3. Municipios con más de 10.000 habitantes. 1900 - 2001

	Provincia	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	Álava	*1*	*1*	*1*	*1*	*1*	*1*	*1*	**2**	**2**	**2**	**2**
02	Albacete	*4*	*4*	5	**7**	**7**	**7**	6	5	5	5	5
03	Alicante/Alacant	*8*	*8*	*8*	9	11	11	14	17	22	27	**30**
04	Almería	8	7	*6*	7	8	8	7	8	8	**9**	**9**
05	Ávila	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*
06	Badajoz	*6*	14	15	17	16	**18**	17	11	10	10	9
07	Balears (Illes)	*4*	*4*	5	6	9	9	9	8	13	15	**17**
08	Barcelona	*9*	*9*	10	11	14	17	25	38	55	59	**68**
09	Burgos	*1*	*1*	*1*	2	**3**	**3**	**3**	**3**	**3**	**3**	**3**
10	Cáceres	*1*	*1*	4	4	**5**	**5**	4	4	4	4	4
11	Cádiz	*10*	13	13	14	16	20	21	**22**	21	21	21
12	Castellón/Castelló	*3*	*3*	*3*	*3*	*3*	5	8	8	9	9	**10**
13	Ciudad Real	*7*	8	10	12	14	**15**	14	12	10	11	12
14	Córdoba	*13*	14	18	19	22	**23**	22	16	14	*13*	*13*
15	Coruña (A)	*9*	14	15	19	25	25	24	**26**	25	25	24
16	Cuenca	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	**2**	**2**
17	Girona	*3*	4	4	*3*	*3*	*3*	5	10	11	**12**	**12**
18	Granada	*6*	7	7	8	12	13	12	11	10	14	**15**
19	Guadalajara	**2**	**2**	**2**	**2**	**2**	*1*	*1*	*1*	*1*	**2**	**2**
20	Guipúzcoa	*1*	2	4	4	5	8	9	19	**20**	19	19
21	Huelva	*2*	4	5	6	7	6	9	8	8	11	**12**
22	Huesca	*1*	*1*	*1*	2	*1*	2	3	**5**	**5**	**5**	**5**
23	Jaén	*8*	13	14	19	22	**23**	20	16	15	14	14
24	León	*2*	*2*	*2*	*2*	3	3	3	5	5	**7**	**7**
25	Lleida	*1*	*1*	*1*	*1*	*1*	*1*	*1*	3	4	4	**5**
26	Rioja (La)	*1*	*1*	2	2	2	2	2	2	**3**	**3**	**3**
27	Lugo	12	13	14	**15**	**15**	13	10	7	6	6	*5*
28	Madrid	*3*	*3*	*3*	*3*	*3*	4	4	15	23	30	**38**
29	Málaga	*6*	7	8	8	9	11	11	13	17	**19**	18
30	Murcia	*14*	16	16	16	18	17	17	19	21	23	**26**
31	Navarra	*1*	*1*	2	3	3	2	2	4	4	6	**7**
32	Ourense	*1*	*1*	2	2	3	2	3	3	**4**	**4**	**4**
33	Asturias	*17*	18	21	23	24	24	24	**25**	22	22	21
34	Palencia	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*
35	Palmas (Las)	*1*	3	3	4	5	6	12	11	12	13	**19**
36	Pontevedra	*11*	12	14	16	20	22	20	19	24	**25**	23
37	Salamanca	*1*	*1*	*1*	*1*	3	3	3	3	3	3	**4**
38	Sta. Cruz de Tenerife	*2*	3	5	5	8	9	10	13	13	14	**17**
39	Cantabria	*2*	*2*	3	3	5	4	5	7	9	8	**10**
40	Segovia	*1*	*1*	*1*	*1*	**2**	*1*	*1*	*1*	*1*	*1*	*1*
41	Sevilla	*8*	9	12	15	20	23	23	26	24	29	**32**
42	Soria	*0*	*0*	*0*	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
43	Tarragona	*4*	*4*	*4*	*4*	*4*	5	5	5	9	10	**13**
44	Teruel	*1*	*1*	*1*	*1*	*1*	*1*	**2**	**2**	**2**	**2**	**2**
45	Toledo	*2*	*2*	*2*	3	3	6	6	6	*2*	3	**7**
46	Valencia/València	*10*	11	13	15	17	17	21	34	39	39	**46**
47	Valladolid	*1*	*1*	2	2	2	2	2	2	2	**3**	**3**
48	Vizcaya	*3*	*3*	6	6	8	7	11	15	17	18	**19**
49	Zamora	*1*	*1*	*1*	*1*	*1*	2	**3**	2	2	2	2
50	Zaragoza	*2*	*2*	*2*	*2*	3	4	4	4	4	4	**5**
51	Ceuta	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*
52	Melilla	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*
	España	*220*	**258**	**297**	**335**	**395**	**420**	**445**	**502**	**549**	**595**	**650**

Nota: En ***cursiva*** el valor **mínimo** de cada provincia en el periodo.
En **negrita** el valor **máximo** de cada provincia en el periodo.

Fuente: INE y elaboración propia.

Cuadro A4. Población en municipios de más de 10.000 habitantes. 1900 - 2001

	Provincia	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	Álava	*34.234*	35.824	37.495	41.754	49.303	50.257	69.849	149.448	210.429	226.591	**235.783**
02	Albacete	*53.800*	65.103	86.319	126.698	155.671	170.593	157.918	161.868	190.162	208.071	**236.702**
03	Alicante/Alacant	*183.994*	205.717	222.697	252.373	314.831	340.930	433.867	636.697	876.046	1.042.569	**1.204.821**
04	Almería	141.962	140.562	124.603	*122.003*	155.242	158.596	163.593	206.116	251.491	303.033	**369.571**
05	Ávila	*13.457*	13.458	15.368	17.036	20.903	23.598	28.032	31.534	40.173	45.977	**49.712**
06	Badajoz	*98.033*	199.444	226.660	255.598	267.785	324.413	**348.416**	274.834	280.749	305.199	320.761
07	Balears (Illes)	*105.400*	113.574	132.210	154.320	220.411	234.823	261.976	331.558	484.906	539.993	**664.688**
08	Barcelona	*678.122*	736.854	913.354	1.231.533	1.433.827	1.728.086	2.324.728	3.366.102	4.129.478	4.120.115	**4.212.285**
09	Burgos	32.641	*31.961*	33.311	52.445	78.306	94.553	123.879	170.265	217.420	226.642	**232.054**
10	Cáceres	*13.617*	14.562	53.720	59.847	83.033	96.521	91.896	101.580	120.557	136.846	**146.729**
11	Cádiz	*280.043*	308.611	366.556	365.006	448.966	583.989	703.491	787.133	894.053	981.346	**1.018.001**
12	Castellón/Castelló	*59.764*	63.587	65.583	68.486	79.317	110.875	165.304	224.743	289.118	308.881	**349.833**
13	Ciudad Real	*95.845*	123.087	168.933	216.883	267.727	290.889	**300.870**	268.091	249.016	270.730	291.379
14	Córdoba	*213.358*	241.255	326.823	415.532	527.310	562.321	**564.084**	491.100	502.954	526.359	540.230
15	Coruña (A)	*196.062*	259.805	293.674	365.408	477.323	535.501	585.032	632.400	716.333	748.315	**765.230**
16	Cuenca	*12.630*	13.748	15.186	17.818	23.554	26.091	28.053	34.290	40.007	53.805	**58.127**
17	Girona	*39.477*	52.514	53.153	51.401	56.598	59.841	93.087	175.649	236.837	272.811	**296.886**
18	Granada	*147.466*	166.443	194.356	234.122	326.398	359.624	339.624	347.968	407.252	473.061	**491.668**
19	Guadalajara	23.243	23.802	25.909	28.555	30.747	*20.619*	23.265	32.105	52.313	71.738	**88.921**
20	Guipúzcoa	*39.641*	59.274	104.516	121.280	143.512	203.987	277.158	480.957	**551.857**	529.448	525.934
21	Huelva	*31.393*	63.338	84.528	107.522	129.972	118.786	168.953	181.089	220.824	274.109	**303.642**
22	Huesca	*13.602*	13.872	15.948	27.505	17.418	31.773	44.193	79.872	93.114	95.964	**99.714**
23	Jaén	*155.347*	218.357	265.632	346.046	429.623	**460.327**	418.308	362.654	374.021	378.352	391.659
24	León	*30.185*	32.722	38.142	46.421	77.618	103.336	134.172	193.326	223.841	**273.588**	264.184
25	Lleida	*21.352*	25.122	38.260	38.423	37.235	51.432	61.822	110.789	140.635	146.737	**159.297**
26	Rioja (La)	*18.866*	22.045	38.998	44.634	57.264	63.263	73.830	99.212	139.124	153.458	**166.975**
27	Lugo	167.079	198.131	202.709	215.765	**236.687**	220.758	188.096	147.864	*146.737*	157.563	150.997
28	Madrid	*599.503*	638.279	849.760	1.070.117	1.359.342	1.600.922	2.244.257	3.522.288	4.469.271	4.714.319	**5.115.759**
29	Málaga	*233.325*	248.111	282.587	319.435	391.992	469.255	493.490	612.522	842.002	996.431	**1.113.248**
30	Murcia	*458.353*	505.490	529.149	519.301	603.632	611.167	653.033	699.289	833.429	937.164	**1.106.898**
31	Navarra	*30.609*	31.271	43.921	59.610	76.012	82.240	109.429	189.508	229.608	261.223	**287.004**
32	Ourense	*21.198*	23.122	36.037	40.631	58.176	65.388	82.527	94.585	125.980	135.666	**145.907**
33	Asturias	*369.511*	422.860	511.362	581.656	637.196	675.092	773.741	874.452	**949.675**	933.437	909.993
34	Palencia	*15.610*	17.710	20.016	24.332	32.469	41.122	47.902	56.816	71.716	77.863	**79.797**
35	Palmas (Las)	*48.357*	94.265	101.231	134.561	184.544	239.048	358.687	449.885	610.933	660.677	**807.587**
36	Pontevedra	*206.825*	243.878	291.620	346.147	425.269	468.548	463.534	521.219	676.600	714.078	**719.873**
37	Salamanca	*25.625*	30.710	32.812	45.040	90.254	104.879	120.887	153.876	186.603	194.895	**197.572**
38	Sta. Cruz de Tenerife	*48.207*	88.544	104.230	125.954	179.298	240.147	306.765	402.469	490.901	550.035	**642.925**
39	Cantabria	*67.414*	74.881	98.189	107.380	149.308	146.133	180.787	251.628	325.845	333.771	**354.265**
40	Segovia	*17.088*	17.892	18.347	21.062	35.064	30.043	35.212	43.099	50.759	**54.375**	54.368
41	Sevilla	*256.201*	276.003	382.819	452.646	630.021	773.172	894.282	1.031.695	1.154.352	1.318.625	**1.426.496**
42	Soria	*0*	*0*	*0*	10.788	13.897	16.753	19.799	24.659	30.326	32.360	**35.151**
43	Tarragona	86.333	*81.658*	93.020	96.903	102.005	126.103	135.824	193.203	293.832	325.134	**387.166**
44	Teruel	*12.745*	15.832	15.145	16.384	16.436	20.337	31.966	32.951	37.586	41.129	**44.589**
45	Toledo	*34.364*	35.556	40.974	54.360	59.664	104.662	116.071	131.439	119.175	138.856	**196.922**
46	Valencia/València	*338.015*	377.953	419.873	523.824	724.025	782.474	872.023	1.280.766	1.592.059	1.631.175	**1.767.720**
47	Valladolid	*71.328*	72.571	87.813	104.040	123.631	135.286	166.103	250.850	339.179	**362.014**	355.500
48	Vizcaya	*114.848*	135.447	207.663	243.444	294.959	325.850	527.142	823.631	**979.756**	951.521	920.638
49	Zamora	*16.646*	17.391	18.419	20.778	28.537	46.464	64.465	60.713	71.351	78.900	**81.435**
50	Zaragoza	*112.249*	123.685	153.123	177.925	234.937	285.450	345.488	512.762	616.356	637.720	**671.448**
51	Ceuta	*13.843*	24.249	35.453	50.293	65.982	56.909	64.728	62.607	65.264	67.615	**71.505**
52	Melilla	*10.182*	40.929	53.577	69.133	69.684	**76.247**	72.430	60.843	53.593	56.600	66.411
	España	*6.108.992*	**7.081.059**	**8.571.783**	**10.240.158**	**12.732.915**	**14.549.473**	**17.354.068**	**22.416.999**	**27.295.598**	**29.076.884**	**31.195.960**

Nota: En ***cursiva*** el valor **mínimo** de cada provincia en el periodo.
En **negrita** el valor **máximo** de cada provincia en el periodo.

Fuente: INE y elaboración propia.

Cuadro A5. Porcentaje de población provincial en municipios de más de 10.000 habitantes. 1900 - 2001

	Provincia	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	Álava	*34,9%*	36,0%	37,0%	39,5%	43,8%	44,0%	52,2%	74,8%	81,6%	**83,2%**	82,3%
02	Albacete	*22,3%*	24,0%	29,1%	37,9%	41,1%	42,6%	42,1%	47,5%	56,0%	60,7%	**64,9%**
03	Alicante/Alacant	*38,8%*	40,1%	42,3%	45,6%	51,8%	53,7%	60,4%	69,1%	76,2%	80,7%	**82,4%**
04	Almería	38,8%	35,3%	*32,5%*	33,9%	41,5%	43,8%	44,3%	54,6%	61,2%	66,5%	**68,9%**
05	Ávila	6,5%	*6,1%*	6,8%	7,2%	8,5%	9,1%	11,1%	14,9%	21,9%	26,4%	**30,4%**
06	Badajoz	*18,9%*	33,6%	34,8%	36,3%	35,9%	39,7%	41,0%	39,2%	43,6%	46,9%	**49,0%**
07	Balears (Illes)	*33,3%*	33,8%	37,7%	41,1%	53,6%	56,0%	59,3%	62,2%	73,9%	76,1%	**79,0%**
08	Barcelona	*64,4%*	64,9%	68,1%	71,2%	74,1%	78,0%	81,9%	86,0%	**89,3%**	88,5%	87,6%
09	Burgos	9,5%	*9,1%*	9,6%	14,6%	20,7%	24,3%	31,9%	47,2%	59,8%	64,2%	**66,5%**
10	Cáceres	3,8%	*3,7%*	13,0%	13,3%	16,4%	17,6%	16,5%	21,7%	28,6%	33,3%	**36,4%**
11	Cádiz	*64,1%*	69,0%	71,3%	71,3%	76,1%	84,2%	86,6%	89,6%	90,5%	91,0%	**91,2%**
12	Castellón/Castelló	*18,9%*	19,1%	20,3%	21,6%	24,8%	33,7%	48,4%	58,3%	67,0%	69,2%	**72,2%**
13	Ciudad Real	*30,0%*	32,8%	39,5%	44,2%	49,8%	50,8%	51,0%	52,3%	52,4%	56,9%	**60,8%**
14	Córdoba	*47,8%*	49,5%	58,9%	62,3%	69,3%	**71,2%**	70,2%	67,2%	69,8%	69,8%	70,9%
15	Coruña (A)	*28,8%*	35,8%	38,0%	43,7%	52,3%	55,1%	56,5%	61,4%	65,5%	68,2%	**69,8%**
16	Cuenca	*5,0%*	5,1%	5,3%	5,7%	6,9%	7,6%	8,5%	13,6%	18,5%	26,2%	**29,0%**
17	Girona	*13,0%*	16,2%	16,1%	15,5%	17,4%	18,6%	26,5%	42,6%	50,7%	**53,5%**	52,5%
18	Granada	*29,8%*	31,6%	33,5%	35,7%	43,7%	45,3%	43,7%	46,9%	53,7%	**59,8%**	59,8%
19	Guadalajara	11,4%	11,2%	12,3%	13,4%	14,5%	*9,9%*	12,3%	21,4%	36,5%	49,3%	**50,8%**
20	Guipúzcoa	*20,2%*	26,8%	40,1%	40,9%	44,2%	55,0%	58,5%	76,8%	**79,4%**	78,3%	78,1%
21	Huelva	*12,2%*	21,4%	25,5%	30,3%	34,6%	32,1%	41,8%	44,9%	52,8%	61,8%	**65,6%**
22	Huesca	5,3%	*5,2%*	6,0%	10,7%	7,0%	13,4%	18,9%	36,0%	43,3%	46,2%	**48,3%**
23	Jaén	*33,1%*	41,6%	45,0%	51,3%	56,5%	58,9%	56,0%	54,3%	58,5%	59,3%	**60,8%**
24	León	*7,5%*	7,8%	8,7%	10,1%	15,3%	18,8%	22,7%	34,4%	42,7%	52,0%	**54,1%**
25	Lleida	*7,5%*	8,5%	11,8%	12,0%	12,1%	15,9%	18,5%	31,9%	39,8%	41,5%	**44,0%**
26	Rioja (La)	*9,9%*	11,5%	19,6%	21,5%	25,4%	27,4%	31,9%	42,3%	54,7%	58,3%	**60,3%**
27	Lugo	35,1%	38,7%	39,1%	41,2%	**43,4%**	42,4%	38,2%	*35,0%*	36,2%	41,0%	42,2%
28	Madrid	77,6%	*76,8%*	81,0%	82,9%	86,4%	87,8%	89,4%	93,6%	**95,4%**	95,3%	94,3%
29	Málaga	*44,8%*	46,9%	50,2%	52,4%	57,0%	62,1%	63,1%	71,8%	82,1%	85,8%	**86,5%**
30	Murcia	*78,8%*	81,4%	80,9%	79,6%	82,6%	80,9%	81,3%	84,0%	87,2%	89,6%	**92,4%**
31	Navarra	9,9%	*9,7%*	12,9%	16,9%	20,8%	21,5%	26,9%	40,6%	45,1%	50,3%	**51,6%**
32	Ourense	*5,1%*	5,2%	7,7%	8,5%	11,8%	13,2%	16,9%	21,4%	29,3%	38,4%	**43,1%**
33	Asturias	*57,9%*	59,6%	65,6%	69,7%	73,3%	75,4%	77,8%	83,1%	84,1%	85,3%	**85,6%**
34	Palencia	*8,1%*	9,0%	10,2%	11,5%	14,6%	17,4%	20,1%	28,2%	38,0%	42,0%	**45,8%**
35	Palmas (Las)	*30,6%*	47,6%	47,1%	52,8%	57,3%	62,9%	78,1%	81,9%	86,2%	86,0%	**91,0%**
36	Pontevedra	*41,7%*	44,6%	49,8%	54,9%	62,0%	65,6%	64,8%	66,7%	76,6%	79,6%	**79,7%**
37	Salamanca	*7,9%*	8,7%	9,6%	12,6%	22,7%	25,3%	29,1%	40,5%	51,2%	54,5%	**57,2%**
38	Sta. Cruz de Tenerife	*23,4%*	34,0%	38,1%	39,7%	49,0%	56,1%	60,5%	69,8%	74,5%	75,8%	**79,7%**
39	Cantabria	*24,2%*	24,3%	28,6%	29,0%	37,1%	36,0%	41,8%	53,6%	63,5%	63,3%	**66,2%**
40	Segovia	10,5%	*10,3%*	10,5%	11,5%	18,1%	14,8%	17,4%	26,6%	34,0%	**36,9%**	36,8%
41	Sevilla	*46,4%*	46,7%	54,4%	57,1%	65,8%	70,2%	71,9%	77,2%	78,1%	81,4%	**82,6%**
42	Soria	*0,0%*	*0,0%*	*0,0%*	6,6%	8,4%	10,2%	13,0%	21,0%	30,1%	34,2%	**38,7%**
43	Tarragona	25,1%	*23,8%*	25,9%	27,6%	29,4%	35,3%	37,4%	44,6%	57,3%	60,0%	**63,5%**
44	Teruel	*5,1%*	6,0%	5,7%	6,2%	6,7%	8,4%	14,3%	19,0%	24,5%	28,6%	**32,8%**
45	Toledo	9,0%	*8,5%*	9,1%	10,9%	12,2%	19,6%	21,8%	27,5%	25,1%	28,4%	**36,4%**
46	Valencia/València	*41,8%*	42,6%	44,9%	50,2%	57,0%	58,2%	60,6%	72,3%	77,0%	77,0%	**79,8%**
47	Valladolid	25,2%	*24,8%*	29,8%	33,4%	37,1%	38,9%	45,2%	60,7%	70,4%	**73,3%**	71,4%
48	Vizcaya	*37,3%*	38,5%	49,3%	50,4%	57,8%	58,8%	70,2%	79,1%	**82,4%**	82,4%	82,0%
49	Zamora	*5,9%*	6,0%	6,3%	7,2%	9,4%	14,7%	20,9%	23,5%	31,3%	36,9%	**40,9%**
50	Zaragoza	*26,7%*	27,5%	30,7%	33,6%	40,9%	46,8%	53,9%	67,7%	74,4%	76,2%	**77,9%**
51	Ceuta	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*
52	Melilla	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*
	España	*32,4%*	**34,8%**	**38,9%**	**42,6%**	**48,3%**	**51,6%**	**56,4%**	**65,9%**	**72,4%**	**74,8%**	**76,4%**

Nota: En ***cursiva*** el valor **mínimo** de cada provincia en el periodo.
En **negrita** el valor **máximo** de cada provincia en el periodo.

Fuente: INE y elaboración propia.

Cuadro A6. Índices de Gini según determinados criterios y agrupaciones. 1900 – 2001

Áreas	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Ciudades (294)	0,644	*0,642*	0,659	0,677	0,692	0,706	0,725	0,756	**0,766**	0,760	0,753
Grandes Ciudades (68)	0,613	*0,605*	0,616	0,632	0,633	0,631	0,645	**0,666**	0,663	0,653	0,649
Pequeñas Ciudades (226)	**0,307**	0,302	0,301	0,302	0,303	0,303	0,278	0,239	0,220	*0,210*	0,217
Áreas no Urbanas (7385)	*0,516*	0,520	0,527	0,535	0,544	0,549	0,564	0,599	0,637	0,656	**0,676**
Todas las Áreas (7679)	*0,665*	0,671	0,688	0,706	0,728	0,745	0,778	0,836	0,874	0,886	**0,893**
Municipìos de más de 10.000 habitantes	0,480	*0,475*	0,497	0,519	0,545	0,571	0,601	0,634	**0,638**	0,629	0,609
Municipìos de menos de 10.000 habitantes	0,511	*0,510*	0,511	0,516	0,517	0,521	0,540	0,580	0,617	0,633	**0,648**
Todos los Municipios (8.108)	*0,637*	0,643	0,660	0,678	0,701	0,719	0,750	0,808	0,846	0,857	**0,862**

Nota : En ***cursiva*** el valor **mínimo** de cada agrupación en el periodo.

En **negrita** el valor **máximo** de cada agrupación en el periodo.

El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente : INE, Ministerio de Fomento (2000) y elaboración propia.

Cuadro A7. Tamaños medios según determinados criterios y agrupaciones. 1900 – 2001

Áreas	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Ciudades (294)	*24.852*	27.334	31.405	36.459	43.203	48.868	58.825	76.171	92.155	97.369	**103.094**
Grandes Ciudades (68)	*75.208*	82.929	97.748	116.216	140.805	162.551	201.673	272.588	335.355	352.992	**370.503**
Pequeñas Ciudades (226)	*9.700*	10.607	11.444	12.461	13.836	14.662	15.844	17.071	18.980	20.456	**22.635**
Áreas no Urbanas (7385)	1.560	1.669	1.730	1.802	1.853	**1.869**	1.826	1.577	1.434	*1.387*	1.427
Todas las Áreas (7679)	*2.452*	2.651	2.867	3.129	3.436	3.669	4.008	4.433	4.907	5.062	**5.319**
Municipìos de más de 10.000 habitantes	27.768	*27.446*	28.861	30.568	32.235	34.642	38.998	44.655	**49.719**	48.869	47.994
Municipìos de menos de 10.000 habitantes	1.613	1.692	1.721	**1.774**	1.770	1.772	1.752	1.528	1.374	1.304	*1.294*
Todos los Municipios (8.108)	*2.322*	2.511	2.715	2.963	3.254	3.475	3.796	4.199	4.648	4.794	**5.038**

Nota: En ***cursiva*** el valor **mínimo** de cada agrupación en el periodo.

En **negrita** el valor **máximo** de cada agrupación en el periodo.

El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

Cuadro A8. Tamaños medianos según determinados criterios y agrupaciones. 1900 – 2001

Áreas	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Ciudades (294)	*11.084*	11.862	12.685	13.845	15.304	15.757	17.594	18.842	20.348	21.134	**23.217**
Grandes Ciudades (68)	*31.344*	33.846	37.878	42.996	53.511	64.115	70.052	89.539	108.336	120.030	**141.473**
Pequeñas Ciudades (226)	*8.544*	9.544	10.401	11.042	12.453	12.500	14.263	15.429	17.293	17.994	**20.179**
Áreas no Urbanas (7385)	910	962	982	**1.008**	1.003	1.005	920	708	563	497	*463*
Todas las Áreas (7679)	954	992	1.031	**1.057**	1.051	1.047	989	764	606	540	*509*
Municipìos de más de 10.000 habitantes	*14.899*	15.031	15.208	15.566	15.416	15.050	15.783	16.940	18.830	18.150	**19.250**
Municipìos de menos de 10.000 habitantes	945	986	1.013	**1.036**	1.022	1.019	953	729	578	506	*468*
Todos los Municipios (8.108)	972	1.017	1.063	**1.094**	**1.094**	1.089	1.028	829	662	604	*568*

Nota: En ***cursiva*** el valor **mínimo** de cada agrupación en el periodo.

En **negrita** el valor **máximo** de cada agrupación en el periodo.

El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

Cuadro A9. Desviación típica según determinados criterios y agrupaciones. 1900 – 2001

Áreas	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Ciudades (294)	*69.459*	75.404	92.262	117.535	139.933	162.424	214.025	310.090	377.749	388.645	**405.172**
Grandes Ciudades (68)	*132.141*	142.996	175.914	226.489	268.472	311.495	413.855	604.446	734.713	753.543	**785.149**
Pequeñas Ciudades (226)	*5.469*	5.855	6.296	6.812	7.557	8.129	8.235	7.573	7.718	8.063	**9.322**
Áreas no Urbanas (7385)	*1.846*	1.992	2.093	2.184	2.294	2.329	2.325	2.155	2.140	2.163	**2.353**
Todas las Áreas (7679)	*14.421*	15.677	19.041	24.036	28.596	33.115	43.343	62.376	75.965	78.273	**81.677**
Municipìos de más de 10.000 habitantes	*56.076*	56.230	66.761	80.956	91.813	104.998	132.645	**168.210**	167.634	155.646	144.962
Municipìos de menos de 10.000 habitantes	*1.761*	1.825	1.849	1.910	1.901	1.915	**1.981**	1.902	1.859	1.834	1.871
Todos los Municipios (8.108)	*10.315*	11.148	13.876	17.525	21.380	25.052	32.270	43.165	**45.316**	43.986	42.996

Nota: En ***cursiva*** el valor **mínimo** de cada agrupación en el periodo.

En **negrita** el valor **máximo** de cada agrupación en el periodo.

El concepto de ciudad se asimila al de área urbana del Ministerio de Fomento (2000). Entre paréntesis el número de observaciones.

Fuente: INE, Ministerio de Fomento (2000) y elaboración propia.

Apéndice 1: Anomalías detectadas en "Alteraciones de los municipios en los Censos de Población desde 1842"

En el transcurso de la elaboración de este trabajo se efectuaron algunas comparaciones entre las cifras municipales de población de la *web* del INE en la sección *Poblaciones de hecho desde 1900 hasta 1991* y las más recientes de la sección *Alteraciones de los municipios en los Censos de Población desde 1842*. Esta última sección respeta rigurosamente la información original de los censos, está notablemente más completa, y permite seguir las alteraciones desde la aparición misma de los municipios en los censos en 1842 (*Censo de la Matrícula Catastral*).

Sin embargo se han detectado, para el periodo 1900 a 1991, algunas incorrecciones, de las que queremos dejar constancia, ya que son fácilmente subsanables.[39]

- Las poblaciones de hecho y derecho para los municipios de Ceuta (51001) y Melilla (52001) en 1991 están intercambiadas.

- En el municipio de Ceuta (51001) existe un error de codificación de la población de hecho en 1960. Se indica que esta población es 73.728, cuando la cifra censal original es 73.182.[40]

- La población de hecho en 1981 de Fornillos de Fermoselle (49074) de la provincia de Zamora aparece como dato inexistente, cuando en realidad es de 452 habitantes.

- En el municipio de Bassagoda (17502) de la provincia de Girona existe un error de codificación de la población de hecho en 1920. Se indica que esta población es 403, cuando la cifra censal original es 463.

- Para la provincia de Madrid existen numerosos municipios en los que ni las poblaciones de hecho, ni las de derecho, en las *Alteraciones de los municipios en*

[39] Obviamente dicho listado no es, en modo algunos, exhaustivo. Las cifras referentes a las consultas *on-line* de este trabajo fueron realizadas en Junio de 2007.

[40] En el caso de Ceuta es de destacar que dicha ciudad se segregó de la provincia de Cádiz en 1930. Aunque la consulta en la *web* deja constancia de este hecho, la población de dicha ciudad se considera siempre excluida de dicha provincia, por lo que los totales provinciales de Cádiz de los censos originales y los de las *Alteraciones de los municipios en los Censos de Población desde 1842* no coinciden con anterioridad a esta fecha. Ceuta es el único municipio para el que se sigue esta práctica. Por ejemplo el municipio de Gátova está correctamente incluido en la provincia de Castellón en el censo de 1991 y anteriores, y correctamente incluido en Valencia en el censo de 2001.

los Censos de Población desde 1842 coinciden para 1991 con las cifras procedentes de la publicación en papel del censo de dicho año (INE 1992), ni tampoco con las poblaciones de derecho para dichos municipios que aparecen en la *web* del INE en el apartado *Censo de Población y Viviendas 1991 (Resultados municipales)*.[41] Estas discrepancias son llamativas al menos por tres razones: (*i*) por encontrarse sólo en la provincia de Madrid y para dicho año, (*ii*) por afectar a un número sustancial de municipios de la provincia (68 en el caso de la población de hecho y 69 en el caso de la población de derecho), y (*iii*) por lo abultado de las discrepancias en algunos casos. Notablemente en el caso de municipio de Madrid (28079). En concreto, para dicho municipio y 1991 la consulta a la sección de *Alteraciones de los municipios en los Censos de Población desde 1842* proporciona para Madrid una población de hecho de 2.984.576 y de derecho de 2.909.792, mientras que en INE (1992) obtenemos unas poblaciones de 3.084.673 y 3.010.492 de hecho y derecho respectivamente. Es decir alrededor de unas 100 mil personas más en ambos casos en la publicación en papel.

[41] Los resultados de esta sección *web* se corresponden con los de la publicación en papel, INE (1992).

Apéndice 2: La distribución de Pareto

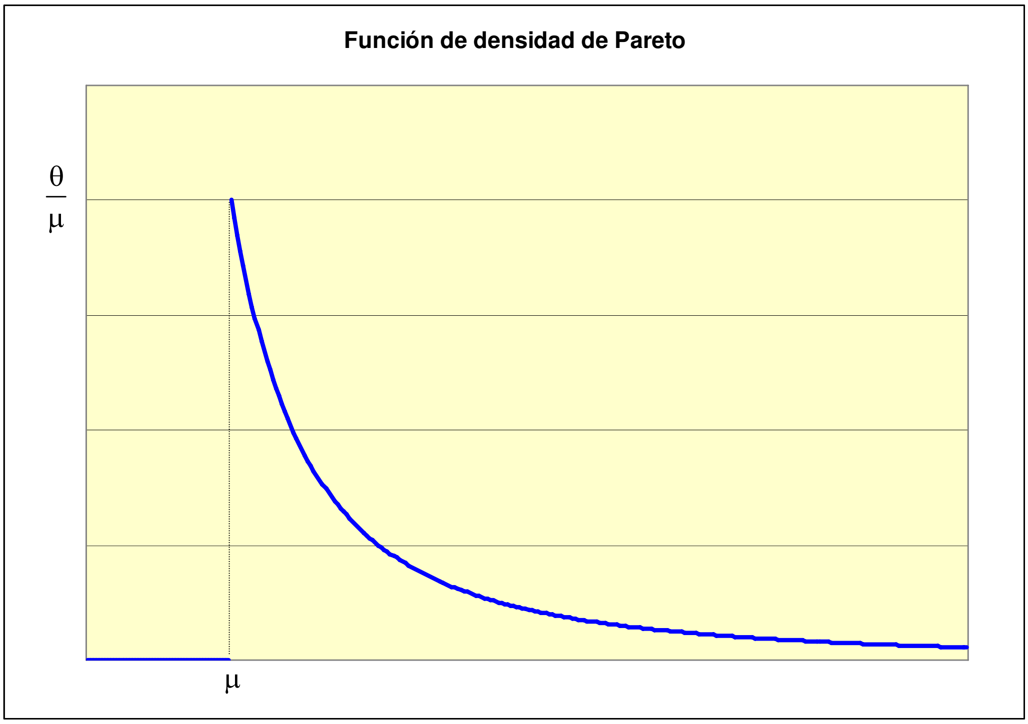
Una variable aleatoria continua y estrictamente positiva, $x \in \mathbb{R}_{++}$, sigue la ley de Pareto (1896) si su función de densidad viene dada por,

$$f_P(x) = \frac{\theta}{\mu}\left(\frac{\mu}{x}\right)^{\theta+1} \qquad x \geq \mu > 0, \quad \theta > 0 \tag{A.1}$$

Para un análisis exhaustivo de esta distribución puede consultarse Johnson y Kotz (1970, Cap. 19).

Junto con la distribución lognormal, la distribución de Pareto (1896) es una de las distribuciones más frecuentemente utilizadas en la modelización de la distribución de la renta, de hecho fue propuesta por Pareto (1896) como modelo para la cola superior de la distribución, es decir para estudiar el comportamiento de las rentas mayores de μ (Cowell 1977).

De la fórmula (A.1) deducimos el gráfico para la función de densidad de $x \in \mathbb{R}_{++}$, $f_P(x) = \dfrac{\theta\mu^{\theta}}{x^{\theta+1}}$.



El parámetro μ es un parámetro de posición, el valor mínimo para el cual la densidad está definida. Por su parte, el parámetro θ es un parámetro de forma que tiene que ver con la altura de la densidad y la dispersión de la misma, en el sentido de que a mayores valores de θ se obtienen densidades de Pareto más concentradas en la proximidades del mínimo, es decir menos dispersas (decaen más rápidamente, aunque desde una altura superior).

A partir de (A.1) podemos obtener, por integración, la función de distribución acumulativa,

$$\begin{aligned} F_P(x) &= \int_{\mu}^{x} \frac{\theta}{\mu}\left(\frac{\mu}{z}\right)^{\theta+1} dz \\ &= \theta\mu^{\theta}\int_{\mu}^{x} z^{-(\theta+1)}dz \\ &= -\frac{\mu^{\theta}}{z^{\theta}}\bigg|_{\mu}^{x} = 1-\left(\frac{\mu}{x}\right)^{\theta} \qquad x \geq \mu > 0, \quad \theta > 0 \end{aligned} \tag{A.2}$$

Por tanto $1 - F_P(x) = \left(\frac{\mu}{x}\right)^{\theta}$ representa la proporción de la población con valores de la variable superiores a x.

Medidas de Posición

Resulta ilustrativo conocer las medidas de dispersión clásicas: Media, Mediana y Moda. La **media** viene dada por,

$$\begin{aligned} E(X) &= \int_{\mu}^{+\infty} x f_P(x)dx = \int_{\mu}^{+\infty} x\frac{\theta}{\mu}\left(\frac{\mu}{x}\right)^{\theta+1} dx \\ &= \theta\mu^{\theta}\int_{\mu}^{+\infty} x^{-\theta}dx = \frac{\theta}{\theta-1}\mu \qquad \theta > 1 \end{aligned} \tag{A.3}$$

Obsérvese que la media sólo está definida para valores de θ > 1 y que conforme θ crece $E(X)$ se aproxima a μ (por la derecha), de forma que la distribución tiende a concentrarse en este valor conforme θ aumenta.

Bajo la ley de Zipf (1949), θ = 1, la media de la distribución no existe.

La **mediana**, $\xi_{0.5}$, viene dada por el valor de x para el cual $F_P(x)$ es igual a 0.5,

$$F_P(\xi_{0.5}) = 1 - \left(\frac{\mu}{\xi_{0.5}}\right)^{\theta} = 0.5 \qquad \Rightarrow \qquad \xi_{0.5} = \mu.2^{\frac{1}{\theta}} \tag{A.4}$$

que de nuevo tiende a μ conforme θ aumenta.

Finalmente se observa fácilmente que la **moda** coincide con μ, puesto que $f_P(x)$ es decreciente a partir de μ.

Medidas de Dispersión

La varianza se obtiene fácilmente observando que los momentos respecto al origen vienen dados por

$$\begin{aligned} E(X^r) &= \int_{\mu}^{+\infty} x^r f_P(x)dx = \int_{\mu}^{+\infty} x^r \frac{\theta}{\mu}\left(\frac{\mu}{x}\right)^{\theta+1} dx \\ &= \theta\mu^{\theta} \int_{\mu}^{+\infty} x^{r-(\theta+1)}dx = \frac{\theta}{\theta - r}\mu^r \qquad \theta > r \end{aligned} \tag{A.5}$$

Por lo que aplicando la fórmula para la varianza, $V(X) = E(X^2) - \left(E(X)\right)^2$, obtenemos

$$V(X) = \frac{\theta}{(\theta-1)^2(\theta-2)}\mu^2 \qquad \theta > 2 \tag{A.6}$$

Obsérvese como la varianza sólo está definida para valores de θ > 2 y que conforme θ crece la varianza tiende a cero, lo que de nuevo justifica la afirmación de que valores mayores de este parámetro se asocian a una menor dispersión en la distribución.

A partir de la media y la varianza podemos obtener el coeficiente de variación,

$$CV(X) = \frac{1}{\sqrt{\theta.(\theta-2)}} \qquad \theta > 2 \tag{A.7}$$

que no depende del parámetro de posición, μ, y de nuevo sólo está definido para valores de θ > 2.

Curva de Lorenz (1905) e Índice de Gini (1912)

Resulta de interés derivar la curva de Lorenz (1905) correspondiente a la distribución de Pareto (1896), así como el índice de Gini (1912) asociado.[42]

Por definición, las ordenadas de la curva de Lorenz (1905) vienen dadas por,

$$\begin{aligned}
\Phi_P(x) &= \frac{1}{E(X)}\int_{\mu}^{x} z dF_P(z) = \frac{1}{E(X)}\int_{\mu}^{x} z f_P(z)dz \\
&= \frac{1}{E(X)}\theta\mu^{\theta}\int_{\mu}^{x} z^{-\theta}dz = \frac{1}{E(X)}\theta\mu^{\theta}\left|\frac{z^{-\theta+1}}{-\theta+1}\right|_{\mu}^{x} \\
&= \frac{1}{E(X)}\theta\mu^{\theta}\left(\frac{\mu^{-\theta+1}-x^{-\theta+1}}{1-\theta}\right) = \frac{1}{E(X)}\underbrace{\frac{\theta}{1-\theta}\mu}_{E(X)}\left(1-\left(\frac{\mu}{x}\right)^{\theta-1}\right) \\
&= 1-\left(\frac{\mu}{x}\right)^{\theta-1}
\end{aligned} \tag{A.8}$$

Por tanto, dado que a partir de (A.2) podemos escribir,

$$\left(\frac{\mu}{x}\right) = \left(1-F_P(x)\right)^{1/\theta} \tag{A.9}$$

la curva de Lorenz (1905), $\Phi_P(x)$ como función de $F_P(x)$, viene dada por,

$$\Phi_P(x) = 1-\left(1-F_P(x)\right)^{\frac{\theta-1}{\theta}} \tag{A.10}$$

El correspondiente índice de Gini (1912) se obtiene como dos veces el área entre la curva de Lorenz (1905) y la línea de igualdad perfecta, $\Phi_P(x)=F_P(x)$. Llamando $p=F_P(x)$, $p\in[0,1]$, obtenemos el índice de Gini (1912) como

$$G(X) = 2\int_0^1\left(p-1+\left(1-p\right)^{\frac{\theta-1}{\theta}}\right)dp \tag{A.11}$$

Y puesto que,

[42] Véase Cowell (1995, Tabla A.2, p.-142) para la relación entre la distribución de Pareto y otros índices de desigualdad habituales en la literatura de la distribución de la renta.

$$\int_0^1 \left(p-1+(1-p)^{\frac{\theta-1}{\theta}} \right) dp = \left.\frac{p^2}{2}\right|_0^1 - \left. p\right|_0^1 - \left.\frac{(1-p)^{\frac{\theta-1}{\theta}+1}}{\frac{\theta-1}{\theta}+1}\right|_0^1 = \frac{1}{2}-1+\frac{\theta}{2.\theta-1} \qquad \text{(A.12)}$$

obtenemos

$$G(X) = 2\left(\frac{1}{2}-1+\frac{\theta}{2.\theta-1}\right) = \frac{1}{2.\theta-1} \qquad \theta \geq 1 \qquad \text{(A.13)}$$

Dado que se trata de un índice relativo el índice de Gini no depende del parámetro de posición, μ, sino solamente del parámetro de forma, θ. Obsérvese que conforme θ crece el índice de Gini disminuye y que dicho índice sólo está definido para θ ≥ 1. En particular conforme θ tiende a 1 (por la derecha) el índice de Gini tiende a su valor máximo, la unidad.

Estimación Máximo Verosímil de θ

Consideremos ahora el problema de estimar los parámetros a partir de una muestra aleatoria de *n* observaciones, $\{x_i\}_{i=1}^n$. Aunque es posible obtener estimadores a partir de diversos criterios (Johnson y Kotz 1970, Cap.- 19) los estimadores de máxima verosimilitud (MLE) son especialmente atractivos si estamos interesados en la inferencia, tanto por su sencillez de cómputo como por sus propiedades (Hill 1975).

Dada una muestra aleatoria de *n* observaciones, $\{x_i\}_{i=1}^n$, el logaritmo de la función de verosimilitud, log *L*, a partir de (A.1) viene dado por

$$\log L = n \log\theta + n\,\theta \log\mu - (\theta+1)\sum_{i=1}^{n} \log x_i \qquad \text{(A.14)}$$

En primer lugar observamos que log *L* no está acotada superiormente con respecto al parámetro μ, y en consecuencia el MLE para este parámetro no puede obtenerse mediante diferenciación. Dado que μ constituye el límite inferior de la variable aleatoria *x* la maximización de (A.14), respecto a este parámetro, debe hacerse sujeta a la restricción de que $\hat{\mu} \leq \min_i \{x\}_{i=1}^n$. Por inspección, el valor de μ que maximiza (A.14) sujeto a esta restricción es precisamente $\hat{\mu} = \min_i \{x\}_{i=1}^n$, es decir $\hat{\mu} = x_{(n)}$ si la muestra está ordenada como en (1). Quandt (1966) ha demostrado la consistencia de este estimador. Sin embargo, desde nuestro punto de vista consideraremos que el parámetro μ es conocido, lo que equivale a realizar el análisis condicionado en él, y finalmente sustituiremos en las fórmulas el parámetro μ por el valor mínimo en la muestra, que como acabamos de ver es el MLE.

Así pues, supuesto que μ es conocido, derivando $\log L$ respecto a θ e igualando a cero, $\frac{n}{\hat{\theta}} + n\log\mu - \sum_{i=1}^{n}\log x_i = 0$, obtenemos el MLE de θ

$$\hat{\theta} = n\left[\sum_{i=1}^{n}\log\frac{x_i}{\mu}\right]^{-1} \tag{A.15}$$

que en la práctica computamos como $\hat{\theta} = n\left[\sum_{i=1}^{n}\log\frac{x_i}{x_{(n)}}\right]^{-1}$, siendo $x_{(n)}$ el mínimo de la muestra.[43]

Puesto que la segunda derivada de $\log L$ respecto a θ es simplemente $-\frac{n}{\theta^2}$, resultados estándares en teoría de la estimación máximo verosímil (Spanos 1986) indican que la distribución asintótica de $\hat{\theta}$ viene dada por

$$\sqrt{n}(\hat{\theta}-\theta) \overset{asy}{\sim} N(0,\theta^2) \tag{A.16}$$

Resultado que puede utilizarse para la realización de contrastes estándar sobre θ. En concreto, el contraste de Wald para la ley de Zipf, H_0: $\theta = 1$, puede realizarse a partir del estadístico[44]

$$W = \sqrt{n}\,\frac{\hat{\theta}-1}{\hat{\theta}} \overset{asy}{\sim} N(0,1) \quad \text{bajo } H_0\text{: } \theta = 1 \tag{A.17}$$

Resulta interesante observar que en este caso el contraste del Wald es numéricamente idéntico al contraste de los multiplicadores de Lagrange (*LM*). Obsérvese que W puede escribirse como

---

[43] Obsérvese que (A.15) no es más que la inversa del logaritmo de la media geométrica de $\left\{\frac{x_i}{\mu}\right\}_{i=1}^{n}$.

[44] Este es el contraste que se ofrece en el texto.

$$W = \sqrt{n}\left(1-\frac{1}{\hat{\theta}}\right) = \sqrt{n}\left(1-\frac{1}{n}\sum_{i=1}^{n}\log\frac{x_i}{\mu}\right) \tag{A.18}$$

Por su parte para derivar el contraste *LM* evaluamos la primera y segunda derivada de log *L* en el estimador restringido, $\tilde{\theta} = 1$,

$$\left.\frac{\partial \log L}{\partial \theta}\right|_{\tilde{\theta}=1} = n - \sum_{i=1}^{n}\log\frac{x_i}{\mu} = n\left(1-\frac{1}{\hat{\theta}}\right) \tag{A.19}$$

y

$$\left.\frac{\partial^2 \log L}{\partial \theta^2}\right|_{\tilde{\theta}=1} = -n \tag{A.20}$$

A partir de aquí el contraste *LM* puede escribirse como (Spanos 1986, pág.- 330, 16.24),

$$LM = \frac{1}{\sqrt{n}}\left(n - \sum_{i=1}^{n}\log\frac{x_i}{\mu}\right) = \sqrt{n}\left(1-\frac{1}{\hat{\theta}}\right) = \sqrt{n}\,\frac{\hat{\theta}-1}{\hat{\theta}} = W \tag{A.21}$$

Generalización de la distribución de Pareto (1896)

Consideremos ahora la generalización de la distribución de Pareto (1896) utilizada en el texto,[45] que como ya hemos indicado es un miembro de la familia de distribuciones de Burr (1942) (Johnson y Kotz 1970, pág.-31),

$$f_B(x) = \frac{\theta}{\sigma}\left(1+\frac{x-\mu}{\sigma}\right)^{-(\theta+1)} \qquad x \geq \mu > 0, \quad \theta > 0, \quad \sigma > 0 \tag{A.22}$$

A partir de (A.22) podemos obtener, por integración, la función de distribución acumulativa,

[45] Esta no es la única generalización posible de la distribución de Pareto, el propio Vilfredo Pareto (1848-1923) propuso otras distribuciones relacionadas con (A.1). Johnson y Kotz (1970, Cap.- 19) y Cowell (1995, Apéndice) ofrecen otras generalizaciones utilizadas en la literatura, muchas de ellas en el contexto del estudio de la distribución de la renta (Champernowne 1952).

$$
\begin{aligned}
F_B(x) &= \int_{\mu}^{x} \frac{\theta}{\sigma}\left(1+\frac{z-\mu}{\sigma}\right)^{-(\theta+1)} dz \\
&= \frac{\theta}{\sigma}\int_{\mu}^{x} \left(1+\frac{z-\mu}{\sigma}\right)^{-(\theta+1)} dz \qquad\qquad (A.23) \\
&= -\left(1+\frac{z-\mu}{\sigma}\right)^{-\theta}\Bigg|_{\mu}^{x} = 1-\left(1+\frac{x-\mu}{\sigma}\right)^{-\theta} \qquad x \geq \mu > 0, \quad \theta > 0, \quad \sigma > 0
\end{aligned}
$$

Por tanto, $1-F_B(x) = \left(1+\frac{x-\mu}{\sigma}\right)^{-\theta}$ representa la proporción de la población con valores de la variable superiores a *x*.

Resulta sencillo observar que para $\sigma = \mu$ la densidad (A.22) se reduce a la densidad de Pareto, (A.1).[46]

Dada una muestra aleatoria de *n* observaciones, $\{x_i\}_{i=1}^{n}$, el logaritmo de la función de verosimilitud, log *L*, a partir de (A.22) viene dado por,

$$
\log L = n\log\theta - n\log\sigma - (\theta+1)\sum_{i=1}^{n}\log\left(1+\frac{x_i-\mu}{\sigma}\right) \qquad (A.24)
$$

Por los mismos motivos que antes supondremos que μ es conocido, lo que equivale a realizar el análisis condicionado en él, y finalmente sustituiremos en las fórmulas el parámetro μ por el valor mínimo en la muestra, que de nuevo coincide con el MLE.

Así pues, supuesto que μ es conocido, derivando $\log L$ respecto a θ y σ, e igualando a cero, obtenemos que el estimador MLE debe resolver:

$$
\begin{aligned}
\frac{n}{\hat{\theta}} - \sum_{i=1}^{n}\log\left(1+\frac{x_i-\mu}{\hat{\sigma}}\right) &= 0 \\
-\frac{n}{\hat{\sigma}} + (\hat{\theta}+1)\sum_{i=1}^{n}\frac{x_i-\mu}{\hat{\sigma}^2\left(1+\frac{x_i-\mu}{\hat{\sigma}}\right)} &= 0
\end{aligned} \qquad (A.25)
$$

Un sistema de ecuaciones no lineales que no admite solución explícita y debe ser resuelto por métodos numéricos. No obstante, una vez obtenida numéricamente la

[46] Consecuentemente la distribución acumulativa (A.23) se reduce a la (A.2).

matriz asintótica de información de Fisher la distribución asintótica del MLE es estándar y puede ser utilizada para construir contrastes de Wald sobre los parámetros. En concreto podríamos construir un contraste de Wald conjunto de la hipótesis de distribución paretiana y la ley de Zipf, H_0: $\sigma = \mu$ y $\theta = 1$.

Sin embargo, en este caso, resulta mucho más sencillo construir un contraste de los Multiplicadores de Lagrange (*LM*) de dicha hipótesis a partir del estadístico $\tilde{q}'\tilde{\Im}_n^{-1}\tilde{q}$, donde q representa el vector de primeras derivadas, $\Im_n$ la matriz de información de Fisher (el negativo de la esperanza del Hessiano de log L), y el símbolo ~ indica que estas cantidades están evaluadas en los estimadores restringidos, $\tilde{\sigma} = \mu$ y $\tilde{\theta} = 1$ (Spanos 1986, pág.- 330, 16.24).

Para derivar el contraste observamos en primer lugar un resultado de interés, y que será de aplicación en la obtención de $\tilde{\Im}_n$. Si x es una variable que verifica la ley de Zipf, es decir su función de densidad viene dada por (A.1) con $\theta = 1$, $f_Z(x) = \dfrac{\mu}{x^2}$, entonces $y_1 = \log\dfrac{x}{\mu}$ sigue una distribución exponencial estándar, es decir con parámetro igual a la unidad, por tanto $E(y_1) = Var(y_1) = 1$.[47] Adicionalmente, $y_2 = \dfrac{\mu}{x}$ sigue una distribución uniforme en el intervalo [0, 1], por tanto $E(y_2) = \dfrac{1}{2}$, $V(y_2) = \dfrac{1}{12}$ y $E(y_2^2) = \dfrac{1}{3}$.[48] Curiosamente una variable que verifica la ley de Zipf, x, no tiene media, mientras que sus transformaciones, y_1 e y_2, si.

---

[47] Este resultado puede derivarse fácilmente a partir de resultados estándares sobre distribuciones de funciones de variables aleatorias (Mood, Graybill y Boes 1974, Cap.- V). En concreto, dado que la función que define y_1 es uno-a-uno y $x = \mu e^{y_1}$, la función de densidad de y_1 vendrá dada por

$$f(y_1) = \left|\frac{dx}{dy_1}\right| f_P(x) = \left|\mu e^{y_1}\right| . \frac{\mu}{\mu^2 e^{2y_1}} = e^{-y_1} \qquad y_1 \geq 0$$

que es la función de densidad de una variable exponencial con parámetro igual a la unidad.

[48] De nuevo, la función que define y_2 es uno-a-uno, está acotada en el intervalo [0, 1] y $x = \dfrac{\mu}{y_2}$, por tanto la función de densidad de y_2 vendrá dada por

$$f(y_2) = \left|\frac{dx}{dy_1}\right| f_P(x) = \left|-\frac{\mu}{y_2^2}\right| . \frac{\mu}{\mu^2 / y_2^2} = \left|-\frac{\mu}{y_2^2}\right| . \frac{y_2^2}{\mu} = 1 \qquad y_2 \in [0, 1]$$

En segundo lugar, evaluamos las primeras derivadas de $\log L$ en los estimadores restringidos, $\tilde{\sigma}=\mu$ y $\tilde{\theta}=1$,

$$\left.\frac{\partial \log L}{\partial \theta}\right|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}} = n - \sum_{i=1}^{n} \log\left(\frac{x_i}{\mu}\right) = n\, z_1$$
$$\left.\frac{\partial \log L}{\partial \sigma}\right|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}} = -\frac{n}{\mu} + \frac{2}{\mu}\sum_{i=1}^{n}\left(1-\frac{\mu}{x_i}\right) = \frac{2n}{\mu}\left(\frac{1}{2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}\right) = \frac{2n}{\mu} z_2 \tag{A.26}$$

definiendo, como en el texto, $z_1 = 1-\frac{1}{n}\sum_{i=1}^{n}\log\frac{x_i}{\mu}$ y $z_2 = \frac{1}{2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}$. Obsérvese que z_1 y z_2 miden simplemente las discrepancias entre las medias poblacionales y muestrales de las variables aleatorias y_1 e y_2. Por tanto, $\tilde{q} = n.\begin{bmatrix} z_1 \\ \frac{2}{\mu} z_2 \end{bmatrix}_{2\times 1}$.

Finalmente, para derivar la matriz de información, $\tilde{\mathfrak{I}}_n$, obtenemos las segundas derivadas de $\log L$,

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{n}{\theta}$$
$$\frac{\partial^2 \log L}{\partial \sigma \partial \theta} = \sum_{i=1}^{n}\frac{x_i-\mu}{\sigma^2\left(1+\frac{x_i-\mu}{\sigma}\right)} \tag{A.27}$$
$$\frac{\partial^2 \log L}{\partial \sigma^2} = \frac{n}{\sigma^2} - (\theta+1)\sum_{i=1}^{n}\frac{(x_i-\mu)(2\sigma+x_i-\mu)}{\sigma^2(\sigma+x_i-\mu)^2}$$

las evaluamos en los estimadores restringidos,

$$\left.\frac{\partial^2 \log L}{\partial \theta^2}\right|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}} = -n$$
$$\left.\frac{\partial^2 \log L}{\partial \sigma \partial \theta}\right|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}} = \sum_{i=1}^{n}\frac{x_i-\mu}{\mu x_i} = \frac{1}{\mu}\sum_{i=1}^{n}\left(1-\frac{\mu}{x_i}\right) = \frac{n}{\mu} - \frac{1}{\mu}\sum_{i=1}^{n}\frac{\mu}{x_i} = \frac{n}{\mu}\left(1-\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}\right) \tag{A.28}$$
$$\left.\frac{\partial^2 \log L}{\partial \sigma^2}\right|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}} = \frac{n}{\mu^2} - 2\sum_{i=1}^{n}\frac{(x_i-\mu)(x_i+\mu)}{\mu^2 x_i^2} = \frac{n}{\mu^2} - \frac{2}{\mu^2}\sum_{i=1}^{n}\left(1-\frac{\mu^2}{x_i^2}\right)$$
$$= -\frac{n}{\mu^2} + \frac{2}{\mu^2}\sum_{i=1}^{n}\frac{\mu^2}{x_i^2} = -\frac{2n}{\mu^2}\left(\frac{1}{2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\mu^2}{x_i^2}\right)$$

---

que es la función de densidad de una variable uniforme.

y les cambiamos el signo y tomamos esperanzas,

$$E\left[-\frac{\partial^2 \log L}{\partial\theta^2}\bigg|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}}\right]=n$$

$$E\left[-\frac{\partial^2 \log L}{\partial\sigma\partial\theta}\bigg|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}}\right]=-\frac{n}{\mu}\left(1-E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\mu}{x_i}\right]\right)=-\frac{n}{\mu}\left(1-E(y_2)\right)=-\frac{n}{2\mu} \tag{A.29}$$

$$E\left[-\frac{\partial^2 \log L}{\partial\sigma^2}\bigg|_{\substack{\tilde{\sigma}=\mu\\ \tilde{\theta}=1}}\right]=\frac{2n}{\mu^2}\left(\frac{1}{2}-E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\mu^2}{x_i^2}\right]\right)=\frac{2n}{\mu^2}\left(\frac{1}{2}-E(y_2^2)\right)=\frac{n}{3\mu^2}$$

Por tanto, $\tilde{\Im}_n=\begin{bmatrix} n & -\dfrac{n}{2\mu} \\ -\dfrac{n}{2\mu} & \dfrac{n}{3\mu^2}\end{bmatrix}_{2\times 2}$.

En consecuencia el estadístico que estamos buscando es,

$$\tilde{q}'\tilde{\Im}_n^{-1}\tilde{q}=n.\begin{bmatrix} z_1 & \dfrac{2}{\mu}z_2\end{bmatrix}.\begin{bmatrix} 1 & -\dfrac{1}{2\mu} \\ -\dfrac{1}{2\mu} & \dfrac{1}{3\mu^2}\end{bmatrix}^{-1}.\begin{bmatrix} z_1 \\ \dfrac{2}{\mu}z_2\end{bmatrix}$$

$$=12\mu^2 n.\begin{bmatrix} z_1 & \dfrac{2}{\mu}z_2\end{bmatrix}.\begin{bmatrix} \dfrac{1}{3\mu^2} & \dfrac{1}{2\mu} \\ \dfrac{1}{2\mu} & 1\end{bmatrix}.\begin{bmatrix} z_1 \\ \dfrac{2}{\mu}z_2\end{bmatrix} \tag{A.29}$$

$$=n.\begin{bmatrix} z_1 & \dfrac{2}{\mu}z_2\end{bmatrix}.\begin{bmatrix} 4 & 6\mu \\ 6\mu & 12\mu^2\end{bmatrix}.\begin{bmatrix} z_1 \\ \dfrac{2}{\mu}z_2\end{bmatrix}$$

$$=4n.\left[z_1^2+6z_1z_2+12z_2^2\right]=LM_z \overset{asy}{\sim} \chi^2(2) \qquad \text{bajo } H_0:\sigma=\mu \text{ y } \theta=1$$

que es el estadístico utilizado en el texto, (9), sustituyendo μ por el valor mínimo en la muestra, $x_{(n)}$.

Referencias

ADES, A. Y GLAESER, E. L. (1995) "Trade and circuses: Explaining Urban Giants". *Quarterly Journal of Economics,* 110, 195-228. Madrid.

AJENJO, M. Y SABATER, A. (2004) "El Impacto de los Movimientos Migratorios sobre la Movilidad Habitual por Trabajo en Cataluña". *Scripta Nova*, VIII, 158, (Febrero). [http://www.ub.es/geocrit/sn/sn-158.htm].

ALPEROVICH, G. A. (1988) "A new testing procedure of the rank size distribution", *Journal of Urban Economics*, 23, 251-259.

ALPEROVICH, G. A. Y DEUTSCH, J. (1995) "The size distribution of urban areas: Testing the appropriateness of the Pareto distribution using a generalized Box-Cox transformation function", *Journal of Regional Science*, 35, 267-276.

ARTÍS, M.; ROMANÍ, J. Y SURIÑACH, J. (1998) "Commuting in Catalonia: Estimates from a place-to-place model". 38th Congress of the European Regional Science Association, August 1998, Vienna. [http://ideas.repec.org/p/wiw/wiwrsa/ersa98p60.html].

AUERBACH, F. (1913) "Das gesetz der bevölkerungskonzentration", *Petermanns Geographische Mitteilungen*, LIX, 73-76.

BAHAMONDE MAGRO, A. Y OTERO CARVAJAL, L. E. (1999) "Madrid, de capital imperial a región metropolitana. Cinco siglos de terciarización", *Papeles de Economía Española*, 18, 18-30. [http://www.ucm.es/info/hcontemp/leoc/madrid%20servicios.htm].

BRAKMAN, S.; GARRETSEN, H.; VAN MERREWIJK, C. Y VAN DER BERG (1999) "The return of Zipf: Towards a further understanding of the rank-size distribution", *Journal of Regional Science*, 39, 1, 183-213.

BOIX DOMÈNECH, R. (2000) "Redes de ciudades y externalidades", *Investigaciones Regionales*, 4, 5-27.

BURR, I. W. (1942) "Cumulative frequency functions", *The Annals of Mathematical Statistics*, 13, 2, (June), 215-232.

CAMPS CURA, E. (1990) "Urbanización y migraciones internas durante la transición al sistema fabril: El caso catalán", *Boletín de la Asociación de Demografía Histórica*, 8, 2, 73-95.

CAPEL SÁEZ, H. (1972) "La validez del modelo *rank-size*", 6, 1, (Enero-Junio), 122-138.

CAPEL SÁEZ, H. (1975) "La definición de lo urbano", *Estudios Geográficos*, 138-139, (Febrero-Mayo), 265-301. (Número especial de "Homenaje al Profesor Manuel de Terán").

CARROLL, G. (1982) "National city-size distribution: What do we know after 67 years of research?", *Progress in Human Geography*, 6, 1-43.

CASTAÑER, M.; VICENTE, J. Y BOIX G. (2000, EDS.) *Áreas urbanas y movilidad laboral en España*, Girona, Universitat de Girona. 17 y 18 de marzo de 2000.

CHAMPERNOWNE, D. G. (1952) "The graduation of income distributions", *Econometrica*, 20, 4, (October), 591-615.

CHESIRE, P. (1999) "Trends in sizes and structures or urban areas". En P. Chesire y E. S. Mills (Eds.), 1339 – 1372.

CHESIRE, P. Y MILLS, E. S. (1999) *Handbook of Regional and Urban Economics*. Volume 3. Amsterdam, Elsevier Science, B. V.

CORREAS, P. (1988) "Poblaciones españolas de más de 5.000 habitantes entre los siglos XVII y XIX", *Boletín de la Asociación de Demografía Histórica*, 6, 1, 5-23.

COWELL, F. A. (1977) *Measuring Inequality*. (Primera edición), Oxford: Phillip Allan. (Segunda edición 1995).

DE COS, O. Y REQUES, P. (2005) "Los cambios en los patrones territoriales de la población española (1900-2001)". *Papeles de Economía española,* 104, 167-192. Madrid.

DE VRIES, J. (1984) *European Urbanization. 1500 - 1800*. Methuen and Co. Ltd. London.

DE VRIES, J. (1990) "Problems in the measurement, description, and analysis of historical urbanization". En Ad van der Woude, Jan de Vries y Akira Hayami (Eds.), 43-60.

DURANTON, G. (2002) "City size distributions as a consequence of the growth process", Centre for Economic Policy Research, Discussion Paper Series Nº 3577, (October). [http://www.cepr.org/pubs/dps/DP3577.asp].

EATON, J. Y ECKSTEIN, Z. (1997) "Cities and growth: Theory and evidence from France and Japan", *Regional Science and Urban Economics*, 27, 443-474.

ESTEVE, D. Y DEVOLDER, D. (2004) "De la ley rango-tamaño (*rank-size*) a la ley log-normal: Los procesos aleatorios en el crecimiento demográfico de las agregados de población". Trabajo presentado en el *VII Congreso de la Asociación de Demografía Histórica*, Sesión 11. Dinámicas espaciales de la población en el largo plazo (siglos XIX y XX). Coordinador: Vicente Pinilla. Granada, 1 al 3 de abril de 2004.

FAN, C. C. Y CASETTI, E. (1994) "The spatial and temporal dynamics of US regional income inequality, 1950 – 1989", *Annals of Regional Science*, 28, 177-196.

FERIA TORIBIO, J. M. (2000) “Pautas Estructurales Diferenciadas de Movilidad en las Áreas Metropolitanas Andaluzas”, en M. Castañer, J. Vicente y G. Boix (Eds.), 121-138. [http://iei.ua.es:9673/commuting/cap/feria2000.pdf].

FERIA TORIBIO, J. M. (2004) “Problemas de definición de las áreas metropolitanas en España”. Boletín de la A.G.E., 38, 85-99.

GABAIX, X. (1999) “Zipf´s law for cities: An explanation”, *The Quarterly Journal of Economics*, 114, 3, (August), 739-767.

GABAIX, X. Y IOANNIDES, Y. M. (2004) “The evolution of city size distributions”, Capítulo 53 en J. V. Henderson y J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, Volume 4, Amsterdam, North-Holland Publishing Company.

GARCÍA FERNÁNDEZ, P. (1985) *La población de los actuales términos municipales 1900-1981. Poblaciones de hecho según los censos.* Madrid. INE.

GIBRAT, R. (1931) *Les inégalités économiques. Applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d’une loi nouvelle, la loi de l’effet proportionnel.* Paris: Libraire du Recueil Sirey.

GINI, C. (1912): “Variabilità e mutabilità, contributo allo studio delle distribuzioni e relazioni statistiche”, *Studi Economico-Giuridici dell’ Universiti di Cagliari*, 3, part 2, 1-158.

GOERLICH, F. J. (2007) “¿Cuántos somos? Una excursión por las estadísticas demográficas del Instituto Nacional de Estadística”. Documento de Trabajo 2007-04. Instituto Valenciano de Investigaciones Económicas. [http://www.ivie.es].

GOERLICH, F. J. Y MAS, M. (2007a) “Pautas de localización de la población a lo largo del siglo XX”, *Investigaciones Regionales*, pendiente de publicación.

GOERLICH, F. J. Y MAS, M. (2007b) “Los motores de la localización: Geografía *versus* historia”, WP FBBVA, en evaluación.

GOERLICH, F. J. Y MAS, M. (2007c) “Localización de la población y urbanización a lo largo del siglo XX”, mimeo.

GOERLICH, F. J.; MAS, M. AZAGRA, J. Y CHORÉN, P. (2006) *La Localización de la Población sobre el Territorio. Un Siglo de Cambios. Un Estudio Basado en Series Homogéneas 1900-2001.* Fundación BBVA. Bilbao.

GOERLICH, F. J. Y VILLAR, A. (2007) *Desigualdad y Bienestar Social: De la Teoría a la Práctica*. Fundación BBVA. Bilbao.

GÓMEZ MENDOZA, A. Y LUNA RODRIGO, G. (1986) "El desarrollo urbano en España, 1860-1930", *Boletín de la Asociación de Demografía Histórica*, 4, 2, 3-22.

HILL, B. M. (1975) "A simple general approach to inference about the tail of a distribution", *The Annals of Statistics*, 3, 5, (September), 1163-1174.

INSTITUTO NACIONAL DE ESTADÍSTICA (1986) *Índice alfabético de los municipios con sus códigos al 1 de abril de 1986.* I.N.E. Madrid.

INSTITUTO NACIONAL DE ESTADÍSTICA (1992) *Poblaciones de derecho y hecho de los municipios españoles. Censo de Población de 1991.* I.N.E. Madrid.

IOANNIDES, Y. M. Y OVERMAN, H. G. (2003) "Zipf´s law for cities: An empirical examination", *Regional Science and Urban Economics*, 33, 127-137.

JARQUE, C. M. Y BERA, A. K. (1987) "A test for normality of observations and regression residuals", *International Statistical Review*, 55, 2, (August), 163-172.

JOHNSON, N. L. Y KOTZ, S. (1970) *Distributions in Statistics: Continuous Univariate Distributions*. Volume 1. Boston: Houghton Mifflin Company.

KAMECKE, U. (1990) "Testing the rank size rule hypothesis with an efficient estimator", *Journal of Urban Economics*, 27, 222-231.

LANASPA, L.; PERDIGUERO, A. M. Y SANZ, F. (2004) "La distribución del tamaño de las ciudades en España, 1900-1999", *Revista de Economía Aplicada*, XII, 34, 5-16.

LASUÉN, J. R.; LORCA, A. Y ORIA, J. (1967) "City-Size distributions and economic growth", *Ekistics*, 24, 221-226.

LORENZ, M. C. (1905) "Methods of measuring the concentration of wealth", *Publications of the American Statistical Association*, 9, 209-219.

LUNA RODRIGO, G. (1988) "La población urbana en España, 1860-1930", *Boletín de la Asociación de Demografía Histórica*, 6, 1, 25-68.

MENDELSON, R. Y LEFEBVRE, J. (2003) "Reviewing Census Metropolitan Areas (CMA) and Census Agglomerations (CA) in Canada According to Metropolitan Functionality". Geography Working Paper Series Nº 2003-001. Statistics Canada.

[http://www.statcan.ca:8096/bsolc/english/bsolc?catno=92F0138M2003001].

MINISTERIO DE FOMENTO (2000) *Atlas estadístico de las áreas urbanas de España*. Centro de Publicaciones. 1ª edición. Ministerio de Fomento. Madrid.

MOOD, A. M.; GRAYBILL, F. A. Y BOES, D. C. (1974) *Introduction to the Theory of Statistics*. Third Edition. International Student Edition. London: McGraw-Hill Book Company.

NADAL, J. (2003, Director) *Atlas de la Industrialización de España, 1750-2000*. Fundación BBVA. Bilbao. Editorial Crítica.

OFFICE OF MANAGEMENT AND BUDGET (2002) "Standards for Defining Metropolitan and Micropolitan Statistical Areas; Notice". Federal Register of USA, 65, 249, (December, 27). [http://www.census.gov/population/www/estimates/00-32997.pdf].

PARETO, V. (1896) *Cours d´Economie Politique*. Geneva, Switzerland. Droz.

PUYOL, R. (1997, Ed.) *Dinámica de la Población en España. Cambios Demográficos en el último cuarto del siglo XX.* Colección: Espacios y Sociedades. Serie Mayor nº 7. Editorial Síntesis. Madrid.

QUANDT, R. E. (1964) "Statistical discrimination among alternative hypotheses and some economic regularities", *Journal of Regional Science*, 5, 1-23.

QUANDT, R. E. (1966) "Old and new methods of estimation and the Pareto distribution", *Metrika*, 10, 55-82.

RAPOPORT, A. (1978) "Rank-size relations", en H. Kruskal y J. M. Tanur (Eds.) *International Encyclopedia of Statistics*, Volume 2, New York: Free Press, 847-854.

READ, C. B. (1988) "Zipf´s law", en S. Kotz, N. L. Johnson y C. B. Read (Eds.) *Encyclopedia of Statistical Sciences*, New York: Wiley.

REED, W. J. (2001) "The Pareto, Zipf and other power laws", *Economics Letters*, 74, 15-19.

REHER, D. S. (1986) "Desarrollo urbano y evolución de la población: España 1787 – 1930", *Revista de Historia Económica*, 4, 1, 39-66.

REHER, D. S. (1994) "Ciudades, procesos de urbanización y sistemas urbanos en la Península Ibérica, 1550 – 1991". En M. Guardia, F. J. Monclús y J. L. Oyón (Dirs.) *Atlas histórico de ciudades europeas*. Barcelona: Centre de Cultura Contemporània de Barcelona.

RICHARDSON, H. (1973) "Theory of the distribution of city sizes: Review and prospects", *Regional Studies*, 7, 239-251.

ROSEN, K. T. Y RESNICK, M. (1980) "The size distribution of cities: An examination of the Pareto law and primacy", *Journal of Urban Economics*, 8, 165-186.

STATISTICS CANADA (2002) "Census Metropolitan Area (CMA) and Census Agglomeration (CA)" [http://www12.statcan.ca/english/census01/Products/Reference/dict/geo009.htm]

SMITH, C. A. (1990) "Types of city-size distributions: A comparative analysis". En Ad van der Woude, Jan de Vries y Akira Hayami (Eds.), 20-42.

SOO, K. T. (2002) "Zipf´s law for cities: A cross country investigation", mimeo. London School of Economics. (December).

SPANOS, A. (1986) *Statistical foundations of econometric modelling*. Cambridge: Cambridge University Press.

SUTTON, J. (1997) "Gibrat´s legacy", *Journal of Economic Literature*, 35, 1, (March), 40-59.

TAFUNELL, X. (2005) "Urbanización y vivienda". En A. Carreras y X. Tafunell (Coords.) *Estadísticas Históricas de España. Siglos XIX – XX*. Volumen I. Segunda Edición Revisada y Ampliada. Bilbao: Fundación BBVA.

TRULLEN, J. Y BOIX DOMÈNECH R. (2003) "Barcelona, Metrópolis policéntrica en red", Working Papers, Universitat Autònoma de Barcelona, Departament d'Economia Aplicada, nº 3.

URZÚA, C. M. (2000) "A simple and efficient test for Zipf´s law", *Economics Letters*, 66, 257-260.

VALERO LOBO, A. (1989) "El sistema urbano español en la segunda mitad del siglo XIX", *Boletín de la Asociación de Demografía Histórica*, 7, 1, 7-29.

VAN DER WOUDE, A.; DE VRIES, J. Y HAYAMI, A. (1990, Eds.) *Urbanization in History. A Process of Dynamic Interactions*. Oxford: Claredon Press.

VINUESA ANGULO, J. (1996) "Dinámica de la población urbana en España (1857 – 1991)", *Ciudad y Territorio*, 28, 107-108, 185-216.

VINUESA ANGULO, J. (1997) "El crecimiento de la población y los desequilibrios en la distribución espacial". En Rafael Puyol (Ed.), 265-311.

ZIPF, G. (1949) *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

ZOIDO, F. Y ARROYO, A. (2004) "La población de España", en A. Arroyo (Coord.) *Tendencias demográficas durante el siglo XX en España*, disponible en la web del INE, http://www.ine.es/prodyser/pubweb/tend_demo_s20/tend_demo_s20.htm, Madrid, Instituto Nacional de Estadística, 17-75.

Nota sobre autores

FRANCISCO J. GOERLICH GISBERT. Licenciado y Doctor en Economía por la Universitat de València, MSc in Economics por la London School of Economics, Universidad de Londres. Actualmente es profesor del Departamento de Análisis Económico de la Universidad de València e Investigador Asociado (creo que soy eso!) del Instituto Valenciano de Investigaciones Económicas (Ivie). Sus campos de especialización son la Econometría Aplicada, la Economía Regional y la Distribución de la Renta. Ha publicado más de cuarenta artículos en revistas especializadas, tanto nacionales como extranjeras y ha colaborado en la elaboración de más de diez libros.

MATILDE MAS IVARS. Licenciada y doctora en Economía por la Universitat de València, profesora titular de Análisis Económico en dicha universidad y profesora investigadora del Ivie desde 1990. Sus campos de especialización son el crecimiento, el cambio técnico, el capital público, la economía regional y las nuevas tecnologías de la información. Ha publicado treinta libros y capítulos de libro, y más de cuarenta artículos en revistas especializadas, nacionales y extranjeras.