

## PANORÁMICA DE LA ESTRUCTURA INDUSTRIAL DE UNA GRAN CIUDAD

M. M. MAYORAL\*

### INTRODUCCIÓN

Cuando se trabaja con poblaciones finitas, la Teoría de Muestras proporciona junto con la estimación de diferentes características de interés, una cuantificación de la precisión de esa estimación, permitiéndonos comparar diferentes métodos de selección de muestras respecto a la precisión de la estimación. Habitualmente, los estudios que se llevan a cabo en Teoría de Muestras hacen referencia a la estimación de características poblacionales, tales como la media o el total de cierta variable, la proporción de individuos que presentan cierta propiedad, etc., así como la determinación de las precisiones de las estimaciones, como puede verse, por ejemplo, en Cochran (1977). En este trabajo nos hemos centrado en otro tipo de característica, la ubicación de los diferentes establecimientos industriales según su mayor o menor dispersión espacial en una determi-

---

Manuscrito recibido en noviembre de 2000, versión final, agosto de 2001.

La autora expresa su agradecimiento a los dos dictaminadores del presente trabajo por sus útiles comentarios y sugerencias.

\* Departamento de Metodología de la Investigación y T. C. Escuela Universitaria de Estadística, Universidad Complutense de Madrid, 28040-Madrid, correo electrónico: mayoralp@eucmax.sim.ucm.es

nada región. Característica que analizaremos utilizando las medidas de entropía procedentes de la teoría de la información estadística.

Estas medidas han sido utilizadas por diferentes autores como medidas de diversidad. En este sentido, Patil y Taille (1972) y Pielou (1975) utilizan índices de entropía en estudios de ecología, Lewontin (1972) considera la entropía de Shannon en estudios de biología, Agresti y Agresti (1978) utilizan la entropía de Havrda-Charvat en estudios sociológicos, Lieberman (1969) en estudios de economía, etcétera.

El concepto de diversidad aparece en numerosos campos de investigación, asociado siempre a la idea de variabilidad de los elementos de una determinada población. Así, por ejemplo, en una comunidad animal se dirá que hay más diversidad ecológica cuanto mayor sea el número de especies y más semejantes sean sus respectivas proporciones. En un sector industrial, se dirá que hay más diversidad (o, paralelamente, menos concentración) cuanto mayor sea el número de empresas que se reparten el mercado y más semejantes sean sus cuotas de mercado.

Siguiendo en la línea de utilizar medidas de la teoría de información estadística, se considera en este trabajo la familia de entropías de orden  $r$  introducida por Renyi (1961) como medidas que permiten cuantificar la incertidumbre o falta de certeza que se tiene sobre una determinada situación proporcionando, en nuestro caso, información de la mayor o menor dispersión de los establecimientos industriales de una determinada zona geográfica.

Esta medida permite cuantificar desde un punto de vista teórico, como puede verse, entre otros, en Mayoral (1995, 1997), el grado de concentración de los establecimientos industriales de una determinada región. Así, se dirá que hay menos concentración cuanto mayor sea el número de establecimientos o empresas que se reparten el mercado y más semejantes sean sus cuotas de mercado.

En la práctica, a veces, ocurre que la entropía poblacional resulta desconocida, bien por referirse a una población excesivamente grande y representar un costo muy elevado la observación de todos los individuos del colectivo o por existir algún tipo de problema que imposibilite su obtención. En estas situaciones, cabe plantearse obtener información

sobre ella a partir de una muestra de individuos seleccionada de la población, sobre la base de la cual se harán, posteriormente, diferentes inferencias. Para ello, en el apartado 2, veremos cómo estimar el índice de concentración basado en la entropía de Renyi tanto en el muestreo aleatorio simple como en el muestreo estratificado, centrándonos también en el estudio del comportamiento asintótico.

Este planteamiento se utiliza en el apartado 3 en el análisis de la distribución de los establecimientos industriales en la Comunidad de Madrid a partir de diferentes muestras aleatorias consideradas en cada una de las subpoblaciones (zonas geográficas, distritos, ...) en que se encuentran repartidos los 18 881 establecimientos industriales existentes en la Comunidad de Madrid en 1992 según la información proporcionada por el directorio industrial de esta comunidad. Obviamente esta metodología puede aplicarse a otros ámbitos.

Como indican Iranzo y otros (1993), la economía ocupa un lugar importante a la hora de desarrollar un modelo territorial. Por ello, señalaremos los principales rasgos distintivos de la industria dentro de la economía madrileña por ser este el ámbito de nuestro estudio.

El papel de la industria ha sido relegado con frecuencia a un segundo plano debido al protagonismo alcanzado por el sector terciario, pero la industria madrileña es un sector fuerte que en 1992 (año de referencia de nuestro estudio) daba empleo a 284 000 personas. El sector industrial va unido al efecto de capitalidad como factor que explica la economía de Madrid y sus funciones al estar dotada de elementos geoestratégicos que la convierten en principal nudo logístico y de comunicaciones (ofrece servicios de la administración, acogiendo la mayoría de las instituciones y organismos administrativos del gobierno central, a la vez que genera necesidades diversas) dominando el sector servicios al mismo tiempo que actúa como principal motor de atracción de instalaciones fabriles. La interacción entre ambos sectores ha llevado consigo la aparición de un entramado productivo, tecnológicamente avanzado, el cual es uno de los rasgos distintivos de la actividad industrial madrileña.

## 1. APROXIMACIÓN AL ÍNDICE POBLACIONAL DE RENYI

Se considera una población finita constituida por  $N$  elementos (en nuestro estudio  $N$  establecimientos) que, de acuerdo a una variable o proceso de clasificación  $X$ , puede dividirse en  $M$  categorías (ramas, subramas, nivel de empleo,...) que denotamos como  $(x_1, \dots, x_M)$  y que aparecen en cada una de las categorías con frecuencias relativas  $(p_1, \dots, p_M)$ , respectivamente. Entonces se denomina entropía de orden  $r$  de Renyi (1961) al valor de la expresión

$$H_M^r(p_1, \dots, p_M) = \frac{1}{1-r} \log \left( \sum_{i=1}^M p_i^r \right), \text{ si } r \neq 1, r > 0$$

donde  $r$  es un parámetro real positivo (es decir, para cada valor de  $r$  tenemos una medida de entropía diferente) y  $\log$  representa la función logaritmo que se puede tomar respecto de cualquier base, aunque habitualmente se utiliza en base 2. Para  $r = 1$  esta expresión tiende a la entropía de Shannon (1948), utilizada entre otros por Mayoral (1992) en el análisis de la población activa española. A partir de ahora a esta expresión la denominaremos *índice poblacional* de orden  $r$ , y puede interpretarse como una valoración del grado de concentración en la distribución de los establecimientos que integran la industria de una determinada región. Es decir, representa una medida de la variabilidad asociada al proceso de clasificación  $X$  como puede verse en Nayak (1985) y Mayoral (1995, 1997).

Con el fin de proceder a la estimación del índice poblacional se supone, en primer lugar, que se extrae una muestra aleatoria simple de tamaño  $n$  y se denota por  $Y_i$  el número de elementos de la muestra que caen en la clase  $x_i$ , ( $i=1, \dots, M$ ).

Considerando ahora  $(Y_1, \dots, Y_M)$  se tiene una variable aleatoria que sigue una distribución multinomial de parámetros  $(n; p_1, \dots, p_M)$  y te-

niendo en cuenta que la frecuencia relativa asociada a la clase  $i$ -ésima  $\hat{p}_i = Y_i / n$  es un estimador consistente de  $p_i$ , se tiene que

$$H_M^r(\hat{p}_1, \dots, \hat{p}_M) = \frac{1}{1-r} \log \left( \sum_{i=1}^M \hat{p}_i^r \right), \text{ si } r \neq 1, r > 0$$

es un estimador consistente de  $H_M^r(p_1, \dots, p_M)$  y recibe el nombre de *índice muestral* de orden  $r$ .

En lo que se refiere al comportamiento asintótico del *índice muestral*, se verifica que (Mayoral, 1995)

$$n^{1/2} (H_M^r(\hat{p}_1, \dots, \hat{p}_M) - H_M^r(p_1, \dots, p_M))$$

sigue aproximadamente una distribución  $N(0, \sigma^2(r))$ , donde

$$\sigma^2(r) = \frac{r^2}{(1-r)^2} \left( \sum_{i=1}^M p_i^r \right)^{-2} \left[ \sum_{i=1}^M p_i^{2r-1} - \left( \sum_{i=1}^M p_i^r \right)^2 \right], \text{ si } r \neq 1, r > 0,$$

siempre que  $\sigma^2(r) > 0$ .

Si  $p_1 = \dots = p_M = 1/M$  el resultado anterior no tiene sentido ya que  $\sigma^2(r) = 0$ . En esta situación se tiene (Mayoral, 1995) que

$$\frac{2n}{r} (\log M - H_M^r(\hat{p}_1, \dots, \hat{p}_M))$$

sigue aproximadamente una distribución  $Ji$ -cuadrada con  $M-1$  grados de libertad.

Si se supone además que la población considerada anteriormente se puede dividir en  $L$  subpoblaciones lo más homogéneas posible respecto a

la variabilidad del proceso de clasificación  $X$ , llamadas estratos, la variabilidad asociada al proceso de clasificación puede cuantificarse por

$${}^{st}H_M^r(p_1, \dots, p_M) = \frac{1}{1-r} \log \sum_{i=1}^M \left( \sum_{k=1}^L p_{ik}^r \right), \text{ si } r \neq 1, r > 0,$$

que en lo sucesivo denominaremos *índice poblacional estratificado* de orden  $r$  y donde  $p_{ik}$  es la frecuencia relativa asociada al estrato  $k$  y a la clase  $x_i$  y  $p_{i.} = \sum_{k=1}^L p_{ik}$  la frecuencia relativa de la clase  $x_i$ .

Sea  $N_k$  el número de individuos en el  $k$ -ésimo estrato, con lo que los tamaños relativos de los estratos vendrán dados por  $W_k = N_k/N$ ,  $k=1, \dots, L$ .

En estas condiciones con el fin de proceder a la estimación del índice poblacional se supone que se extrae una muestra aleatoria estratificada de tamaño  $n$  de la población, donde las submuestras correspondientes a los distintos estratos se tomarán de forma independiente y de acuerdo a una afijación predeterminada  $w_1, \dots, w_L$ .

Sea  $n_k$  el tamaño muestral correspondiente al  $k$ -ésimo estrato según una afijación predeterminada, es decir,  $n_k = w_k n$ ,  $k=1, \dots, L$  con  $\sum_{k=1}^L n_k = n$ ,  $\hat{p}_{ik}$  la frecuencia relativa correspondiente al número de individuos que pertenecen a la clase  $x_i$  en el  $k$ -ésimo estrato obtenida al observar la muestra aleatoria estratificada de tamaño  $n$  y  $\hat{p}_{i.}$  el estimador insesgado de  $p_{i.}$  en muestreo estratificado, es decir  $\hat{p}_{i.} = \sum_{k=1}^L \frac{W_k}{w_k} \hat{p}_{ik}$ ,  $i=1, \dots, M$ .

En estas condiciones, el índice muestral viene dado por:

$${}^{st}H_M^r(\hat{p}_1, \dots, \hat{p}_M) = \frac{1}{1-r} \log \sum_{i=1}^M \left( \sum_{k=1}^L \frac{W_k}{w_k} \hat{p}_{ik} \right)^r, \text{ si } r \neq 1, r > 0,$$

y se denominará *índice muestral estratificado* de orden  $r$ .

En cuanto al comportamiento asintótico del índice muestral en el muestreo estratificado con reposición y afijación  $w_1, \dots, w_L$  se verifica (Mayoral, 1995) que

$$n^{1/2} ({}^{st}H_M^r(\hat{p}_1, \dots, \hat{p}_M) - {}^{st}H_M^r(p_1, \dots, p_M))$$

sigue aproximadamente una distribución  $N(0, {}^{st}\sigma^2(r))$ , donde

$${}^{st}\sigma^2(r) = \frac{r^2}{(1-r)^2} \left( \sum_{i=1}^M p_i^r \right)^{-2} \sum_{k=1}^L \frac{W_k}{w_k} \left[ \sum_{i=1}^M p_i^{2r-1} p_{ik} - \frac{1}{W_k} \left( \sum_{i=1}^M p_i^{r-1} p_{ik} \right)^2 \right], \text{ si } r \neq 1, r > 0,$$

siempre que  ${}^{st}\sigma^2(r) > 0$ .

A partir de estos comportamientos asintóticos se pueden definir procedimientos para realizar algunos contrastes de hipótesis en relación con el índice poblacional de orden  $r$ , así como construir intervalos de confianza.

En relación con los contrastes de hipótesis se pueden realizar, entre otros, los siguientes:

- i) El valor del índice poblacional es uno especificado  $C_0$ , es decir se está interesado en contrastar

$$H_0 : H_M^r(p_1, \dots, p_M) = C_0 \text{ frente a } H_1 : H_M^r(p_1, \dots, p_M) \neq C_0.$$

Ahora bien, si  $H_0$  es cierta el estadístico

$$Z_1^r = \frac{n^{1/2}(H_M^r(\hat{p}_1, \dots, \hat{p}_M) - C_0)}{\hat{\sigma}(r)}$$

sigue aproximadamente una distribución  $N(0,1)$  para  $n$  suficientemente grande, donde  $\hat{\sigma}(r)$  se obtiene tomando la raíz cuadrada en  $\sigma^2(r)$  y sustituyendo  $p_i$  por  $\hat{p}_i$  para todo  $i=1, \dots, M$ . Por tanto, una vez calculado el valor del estadístico  $Z_1^r$  a partir de la muestra de tamaño  $n$  seleccionada según el muestreo aleatorio simple de la población considerada, si éste es menor que  $-z_{\alpha/2}$  o mayor que  $z_{\alpha/2}$ , donde  $z_{\alpha/2}$  es el cuantil de orden  $1 - \alpha/2$  de una distribución normal de media cero y varianza 1, se rechaza la hipótesis nula a un nivel  $\alpha$  y en otro caso no se rechaza.

Supongamos ahora que se extrae una muestra de tamaño  $n_1$  a partir de una población mediante muestreo aleatorio simple, y consideremos otra muestra de tamaño  $n_2$  a partir de una segunda población mediante el mismo tipo de muestreo, siendo las dos muestras independientes entre sí. Consideremos sendos procesos de clasificación sobre las poblaciones anteriores, que en la práctica en muchos casos coincidirán. Entonces si  $H_M^r(\hat{P}_{(1)})$  y  $H_M^r(\hat{P}_{(2)})$  son los estimadores de la concentración para la primera y segunda población respectivamente, se puede realizar el siguiente contraste.

ii) Las concentraciones  $C_1$  y  $C_2$  de dos poblaciones independientes son iguales. Es decir, se está interesado en contrastar

$$H_0 : C_1 = C_2 \text{ frente a } H_1 : C_1 \neq C_2.$$



Ahora bien, bajo la hipótesis nula,  $H_0$ , es inmediato que el estadístico

$$Z_2^r = \frac{(n_1 n_2)^{1/2} (H_M^r(\hat{P}_{(1)}) - H_M^r(\hat{P}_{(2)}))}{(n_2 \hat{\sigma}_{(1)}^2(r) + n_1 \hat{\sigma}_{(2)}^2(r))^{1/2}}$$

sigue aproximadamente una distribución  $N(0,1)$  para  $n_1$  y  $n_2$  suficientemente grandes. En cuanto a la región crítica se tiene que es similar a la del caso anterior.

iii) Las concentraciones  $C_1, \dots, C_s$  de  $s$  poblaciones independientes son iguales. Esto es, se quiere contrastar

$$H_0 : C_1 = \dots = C_s \text{ frente a } H_1 : \exists i, j / C_i \neq C_j, i, j = 1, \dots, s.$$

Si se denota por  $n_i$  el tamaño de la muestra aleatoria de la población  $i$ , aplicando el teorema de Cochran y teniendo en cuenta los resultados anteriores, se prueba que (Nayak, 1983), bajo  $H_0$ , el estadístico

$$T_1^r = \sum_{i=1}^s \frac{n_i (H_M^r(\hat{P}_{(i)}) - \hat{H}^*)^2}{\hat{\sigma}_{(i)}^2(r)}$$

sigue aproximadamente una distribución Ji-cuadrada ( $\chi^2$ ) con  $s-1$  grados de libertad, donde

$$\hat{H}^* = \frac{1}{s} \sum_{i=1}^s H_M^r(\hat{P}_{(i)}).$$

Por tanto, una vez calculado el valor del estadístico  $T_1^r$  para la muestra considerada, se rechaza la hipótesis nula al nivel  $\alpha$  si el valor de  $T_1^r$  es mayor que  $\chi_{s-1, \alpha}^2$ , donde  $\chi_{s-1, \alpha}^2$  es tal que  $P(\chi_{s-1}^2 > \chi_{s-1, \alpha}^2) = \alpha$ .

iv) La población es homogénea, esto es, se está interesado en contrastar

$$H_0 : p_1 = \dots = p_M = 1/M \text{ frente a } H_1 : \exists i, j / p_i \neq p_j.$$

Para lo cual, como bajo la hipótesis nula,  $H_0$ , el estadístico

$$T_2^r = \frac{2n}{r} \left( \log M - H_M^r(\hat{p}_1, \dots, \hat{p}_M) \right)$$

sigue una Ji-cuadrada ( $\chi^2$ ) con  $M-1$  grados de libertad, deberemos rechazar la hipótesis nula si  $T_2^r > \chi_{M-1, \alpha}^2$ .

Además la potencia asintótica de este *test* en  $(p_1, \dots, p_M)$  viene dada por

$$\beta(p_1, \dots, p_M) = F_{N(0,1)} \left( \frac{n^{1/2} \left( \log M - r \frac{\chi_{M-1, \alpha}^2}{2n} - H_M^r(p_1, \dots, p_M) \right)}{\hat{\sigma}(r)} \right)$$

donde  $F_{N(0,1)}$  es la función de distribución de una normal con media 0 y varianza 1, por lo que el cálculo efectivo de la potencia puede efectuarse a través de las tablas de la función de distribución de la  $N(0,1)$ .

Si en lugar de considerar  $H_M^r$  y  $\hat{\sigma}^2(r)$  consideramos  ${}^{st}H_M^r$  y  ${}^{st}\hat{\sigma}^2(r)$  se obtienen los correspondientes estadísticos  ${}^{st}Z_1^r$ ,  ${}^{st}Z_2^r$ ,  ${}^{st}T_1^r$  y  ${}^{st}T_2^r$  con el mismo comportamiento que los anteriores.

Finalmente, los intervalos de confianza al nivel  $1 - \alpha$  para  $H_M^r$  y  ${}^{st}H_M^r$  vienen dados, respectivamente, por

$$\left( H_M^r(\hat{P}) - \frac{\hat{\sigma}(r)z_{\alpha/2}}{n^{1/2}}, H_M^r(\hat{P}) + \frac{\hat{\sigma}(r)z_{\alpha/2}}{n^{1/2}} \right) \text{ y}$$

$$\left( {}^{st}H_M^r(\hat{P}) - \frac{{}^{st}\hat{\sigma}(r)z_{\alpha/2}}{n^{1/2}}, {}^{st}H_M^r(\hat{P}) + \frac{{}^{st}\hat{\sigma}(r)z_{\alpha/2}}{n^{1/2}} \right),$$

donde  $\hat{P} = (\hat{p}_1, \dots, \hat{p}_M)$  y  $\hat{P} = (\hat{p}_1, \dots, \hat{p}_M)$ .

## 2. ESTUDIO DE LA DISTRIBUCIÓN DE LOS DIFERENTES ESTABLECIMIENTOS INDUSTRIALES

En primer lugar, se analiza la distribución de los establecimientos industriales en las diferentes subramas de actividad según subzona geográfica y nivel de empleo. Posteriormente se aborda el estudio de la distribución de los establecimientos en relación con el año de inicio de su actividad por subzona geográfica. La gran diversidad existente junto con el peso de la capital hace necesario establecer una tipología de zonas con profundas y marcadas diferencias que definen situaciones opuestas —de tal forma, que la homogeneidad intragrupo tan sólo es relativa, al no ser que optemos por un número mayor de grupos—. Las cuatro grandes zonas y subzonas consideradas aparecen recogidas en el Directorio Industrial de Establecimientos Industriales de la Comunidad de Madrid. En cuanto a la información utilizada ésta ha sido extraída del listado proporcionado por la Subdirección General de Tratamiento de la Información del Ministerio de Industria y Energía referente al año 1992, a partir de muestras aleatorias simples seleccionadas en cada una de las poblaciones analizadas (subzona geográfica o nivel de empleo), mediante la generación de secuencias de números aleatorios.

En este análisis se utiliza el *índice muestral* de orden 2 dado su buen comportamiento como puede verse en Mayoral (1995, 1997). De forma más precisa, el índice muestral para el muestreo aleatorio simple nos proporcionará estimaciones de la mayor o menor dispersión de los establecimientos en cada una de las subzonas geográficas en el caso de la Comunidad de Madrid o en cada uno de los 21 distritos en el caso de la capital, así como en cada uno de los niveles de empleo. Por otra parte, si tenemos presente que tanto las subzonas como los distritos pueden agruparse en diferentes zonas (zona oeste, zona norte-este, zona sur, Madrid centro-oeste, Madrid sur y Madrid norte-este) o que los distritos se agrupan formando Madrid capital y que las subzonas se agrupan en Madrid provincia podemos analizar qué ocurre en cada una de estas zonas a partir del *índice muestral* para el muestreo estratificado. Esto es posible ya que en cada zona tendremos una muestra estratificada de tamaño muestral igual a la suma de los tamaños muestrales considerados en cada una de las subzonas que componen la zona en estudio y una determinada afijación.

Además, la realización de algunos contrastes de hipótesis nos conducirá a comprobar cuestiones como si la distribución de los establecimientos en dos o más poblaciones diferentes (dos o más subzonas, dos o más distritos, ...) es igual o no, o si la distribución de los establecimientos en una determinada zona es homogénea o no, etcétera.

Para tal fin y en un intento de solucionar el problema que se presenta para seleccionar los tamaños muestrales debido a la diversidad de poblaciones consideradas y sin olvidar la distribución de la población, las características que deseamos estudiar así como los objetivos de este trabajo, se ha optado, por considerar 5% de los establecimientos en cada una de las poblaciones, excepto en aquellos casos en los que este porcentaje es inferior a 10 en los que se ha procedido al aumento del tamaño muestral hasta completar dicho número.

### 2.1. *Por subramas de actividad*

Uno de los rasgos característicos de la industria madrileña en lo que se refiere a la distribución de establecimientos es su densificación en el área metropolitana (Comunidad de Madrid, 1994). No obstante, hay que señalar que el proceso de industrialización surge en los distintos barrios del centro de Madrid de forma aglomerada, pero tal concentración va disminuyendo en un proceso de difusión espacial del centro a la periferia, proceso discontinuo, existiendo diferencias en el comportamiento espacial en función de diferentes aspectos: volumen de empleo, dependencia de otras empresas, etc. (Comunidad de Madrid, 1994; Méndez, 1986). Este proceso, aún no consolidado, está actuando en distintos espacios intersticiales entre ejes industriales ya consolidados próximos a las carreteras radiales. Estos aspectos serán analizados a partir del *índice muestral* de orden 2, teniendo en cuenta las subramas de actividad.

#### 2.1.1. Según zonas geográficas

Observando el cuadro 1, en la que se recogen los valores del *índice muestral* para el muestreo aleatorio simple y el mapa 1 que proporciona una visión acerca de la mayor o menor dispersión de los establecimientos en las distintas zonas geográficas, nos encontramos cuatro subzonas geográficas con los valores más altos, siendo éstos muy similares: Cono Nordeste -3.859823, Eje Burgos -3.750772, Sur Metropolitano -3.729039 y Eje Barcelona -3.712824 y a una cierta distancia el Distrito Suroeste con un valor de 3.554168. (La concentración es mínima cuando las proporciones en todas las categorías sean iguales, es decir, todas iguales a 1/33 ya que el número de subramas es de 33, obteniéndose un valor máximo del índice de 5.044395).

Esto nos lleva a cuestionarnos si los establecimientos están uniformemente repartidos entre las distintas subramas en alguna de estas cuatro zonas. Para ello se contrasta en cada zona la hipótesis nula

$$H_0 : p_1 = \dots = p_{33} = 1/33 \text{ frente a } H_1 : \exists \text{ al menos } k, j / p_k \neq p_j,$$

utilizando el estadístico  $T_2^2$ .

Los resultados de aplicar este *test* a las cuatro zonas así como sus potencias quedan recogidos en la tabla 2, donde se pone de manifiesto que para dos de ellas (Sur Metropolitano y Eje Barcelona) se debe rechazar la homogeneidad, es decir, todas las subramas no tienen la misma presencia en estas zonas.

Continuando con el análisis se observa que el valor más alto se da en el Cono Nordeste, zona de reciente industrialización auspiciada por los traslados de fábricas y por los fenómenos de descentralización productiva del eje industrial del Corredor del Henares. Los establecimientos en esta zona se encuentran repartidos entre las distintas ramas, principalmente en "construcciones metálicas", "artículos metálicos" y "madera".

Sigue a este valor el que se corresponde con Eje Burgos con dos núcleos de indudable atracción industrial sobre los que gravita todo el peso de la zona: Alcobendas y San Sebastián de los Reyes. Aunque encontremos establecimientos en prácticamente todas las ramas, hay que destacar aquellos dedicados a "construcciones metálicas", "material electrónico" y "máquinas de oficina y precisión".

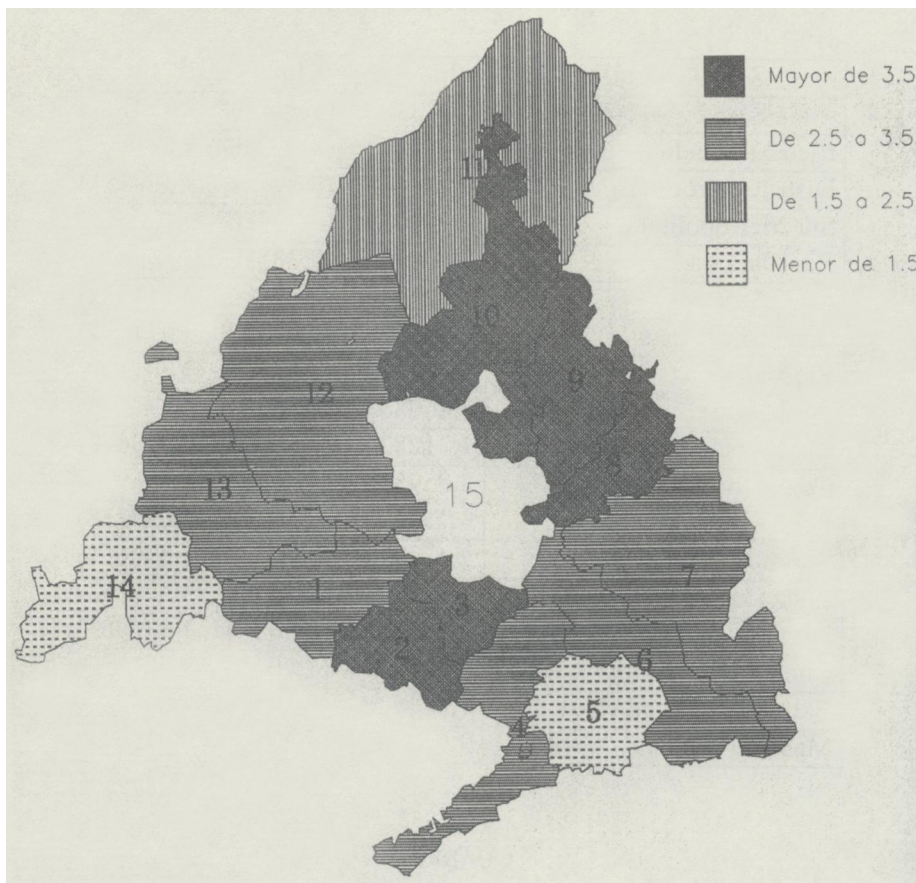
CUADRO 1.  
*Índice muestral de orden 2 en cada zona geográfica  
 según subrama de actividad*

Zona Geográfica	Índice	Intervalo al 95%
Zona Sur	3.783182	(3.614239, 3.952124)
Eje Extremadura	3.152307	(2.719804, 3.584811)
Distrito Suroeste	3.554168	(3.298785, 3.809551)
Sur Metropolitano	3.729039	(3.334267, 4.123811)
Bajo Jarama	2.901992	(2.385536, 3.418449)
Vega Sureste	1.321928	(0.403062, 2.240794)
Zona Norte-Este	3.949417	(3.773270, 4.071811)
Eje Valencia	3.263894	(2.826648, 3.701139)
Cono Este	3.060883	(2.523850, 3.597916)
Eje Barcelona	3.712824	(3.396999, 4.028648)
Cono Nordeste	3.859823	(3.532724, 4.186923)
Eje Burgos	3.750772	(3.403327, 4.098218)
Sierra Norte	2.473932	(1.747205, 3.200658)
Zona Oeste	2.917725	(2.429912, 3.405538)
Eje Coruña-Sierra Central	2.757024	(2.067144, 3.446903)
Cono Oeste	2.643857	(1.784337, 3.503377)
Extremo Suroccidental	1.321928	(0.403062, 2.240794)
Madrid Provincia	3.915902	(3.809144, 4.022658)

CUADRO 2.  
*Contraste de Homogeneidad*

Zona Geográfica	Conclusión al 95%	Potencia
Sur Metropolitano	Rechazar homogeneidad	0.865829
Eje Barcelona	Rechazar homogeneidad	0.999940
Cono Nordeste	No rechazar homogeneidad	0.009726
Eje Burgos	No rechazar homogeneidad	0.451130

MAPA 1.  
*Índice muestral de orden 2*



Subzonas según subrama de actividad (a excepción de la Capital)

- |                      |                     |
|----------------------|---------------------|
| 1. Eje Extremadura   | 6. Eje Valencia     |
| 2. Distrito Suroeste | 11. Sierra Norte    |
| 3. Sur Metropolitano | 7. Cono Este        |
| 4. Bajo Jarama       | 12. Eje Coruña S.C. |
| 5. Vega Sureste      | 8. Eje Barcelona    |
|                      | 13. Cono Oeste      |
|                      | 9. Cono Nordeste    |
|                      | 14. Ext. Surocc.    |
|                      | 10. Eje Burgos      |
|                      | 15. Capital         |



La subzona Sur Metropolitano está formada por tres municipios: Getafe, Leganés y Pinto que actúan como prolongación de los distritos del sur de Madrid en el proceso de industrialización. Los establecimientos obtenidos al seleccionar la muestra pertenecen, principalmente, a las ramas de “construcciones metálicas”, “material de transporte”, “papel e imprentas” y “maquinaria industrial”, aunque en todas las ramas analizadas se observan establecimientos.

El desarrollo del Eje Barcelona también se ve favorecido por la expansión en el proceso industrializador de los distritos de Madrid así como por la importancia de núcleos urbanos de relativa entidad como es el caso de Alcalá de Henares.

El Distrito Suroeste, primera zona industrial si tenemos en cuenta el número de establecimientos, está formado por varios municipios pero el peso recae en Fuenlabrada y Humanes; actualmente en proceso de expansión a municipios próximos. De los establecimientos analizados hay un reparto entre todas las ramas, absorbiendo más establecimientos “construcciones metálicas”, “industria del mueble”, “madera” y “maquinaria industrial”, justificando el valor obtenido 3.554168.

Ante la aparente similitud de los valores obtenidos en el Sur Metropolitano y en el Eje Barcelona podemos realizar un contraste de hipótesis para ver si existen o no diferencias en cuanto a la distribución de los establecimientos entre las distintas subramas de actividad en estas dos subzonas. Para lo cual se contrastará

$$H_0 : \text{Índice (Sur Metropolitano)} = \text{Índice (Eje Barcelona)}$$

frente a

$$H_1 : \text{Índice (Sur Metropolitano)} \neq \text{Índice (Eje Barcelona)}$$

siendo el valor que toma el estadístico a utilizar en este caso  $Z_2^2 = 0.0527444$ .

Luego, con un nivel de confianza de 95% podemos afirmar que no existe diferencia significativa en cuanto al grado de distribución de los establecimientos en estas dos zonas, ya que, en este caso  $z_{0.025} = 1.96$  lo que nos lleva a no rechazar la hipótesis nula.

Con valores muy bajos y similares hay que destacar la Vega Sureste y el Extremo Suroccidental, con nulo protagonismo a efectos industriales a excepción de algunos núcleos como Belmonte del Tajo, San Martín de Valdeiglesias y Villa del Prado. El valor 1.321928 indica un aglutinamiento de establecimientos en una subrama, en concreto la relativa a “pan y molinería”, 6 de los 10 establecimientos se encuentran en esta subrama, tanto en la Vega Sureste como en el Extremo Suroccidental, lo que explica la coincidencia en el resultado.

Si estas 14 subzonas analizadas anteriormente las agrupamos en tres grandes zonas, podemos realizar estimaciones en cada una de estas zonas mediante el índice para el muestreo estratificado, ya que, por ejemplo, en la Zona Sur tenemos cinco estratos que coinciden con cada una de las subzonas que la componen. De esta forma tenemos una muestra estratificada de tamaño 218, donde el tamaño muestral en cada uno de los cinco estratos es 46, 97, 44, 21 y 10, respectivamente. De forma análoga se procede en las restantes zonas así como en el total de la provincia de Madrid donde podemos considerar los 14 estratos correspondientes a las subzonas que la componen obteniéndose las estimaciones que también han quedado incluidas en el cuadro 1. El valor más alto se localiza en la Zona Norte-Este, 3.949417, la cual engloba 6 subzonas con valores todas ellas superiores a tres, indicando un reparto muy equitativo de los establecimientos entre las distintas subramas. Sucede a ésta la Zona Sur con valor 3.783182 valor inferior explicado por la concentración de establecimientos en una única subrama en la Vega Sureste.

Por último y a la vista de los valores del *índice muestral* en estas dos zonas, se analiza si el grado de concentración en estas zonas diferentes con índice semejante es el mismo o no.

Esto nos lleva a contrastar

$$H_0 : \text{Índice (Norte-Este)} = \text{Índice (Sur)}$$

frente a

$$H_1 : \text{Índice (Norte-Este)} \neq \text{Índice (Sur)}$$

siendo el valor del estadístico de contraste  $^{st}Z_2^2 = 1.1200597$ , lo que nos lleva a asegurar que no existen diferencias significativas entre las dos zonas consideradas en cuanto a la distribución de los establecimientos.

En la capital, la industria sigue siendo importante con un total de 9 901 establecimientos industriales en oposición al resto de la provincia con 8 980 establecimientos.

Aplicando el Índice Muestral en el análisis de la distribución de establecimientos industriales en la capital observamos mayor homogeneidad en los valores obtenidos para cada uno de los distritos que la componen y que aparecen recogidos en el cuadro 3. El valor más alto se da en el Distrito de Usera con un valor de 3.542899 frente al distrito de Barajas con un valor de 2.321928.

Según se observa en el mapa 2, los cinco valores más altos se encuentran localizados en Usera -3.542899, Tetuán -3.532496, Puente de Vallecas -3.517100, Arganzuela -3.395302 y Carabanchel -3.375762. El valor 3.532496 de Tetuán indica un reparto bastante equitativo de establecimientos entre las distintas subramas. Sólo en la rama “industria química” no se ha observado ningún establecimiento. Los cuatro distritos restantes, Usera, Arganzuela y Carabanchel que integran la zona Sur y el distrito Puente de Vallecas perteneciente a la zona Norte-Este marcan la pauta del Sur Metropolitano, actuando esta zona como periferia de los distritos señalados.

Esto nos lleva a cuestionarnos si los establecimientos están uniformemente repartidos entre las distintas subramas en alguno de estos

cinco distritos. Para ello se realiza, en cada distrito, el contraste de la hipótesis nula

$$H_0 : p_1 = \dots = p_{33} = 1/33$$

frente a

$$H_1 : \exists \text{ al menos } k, j / p_k \neq p_j,$$

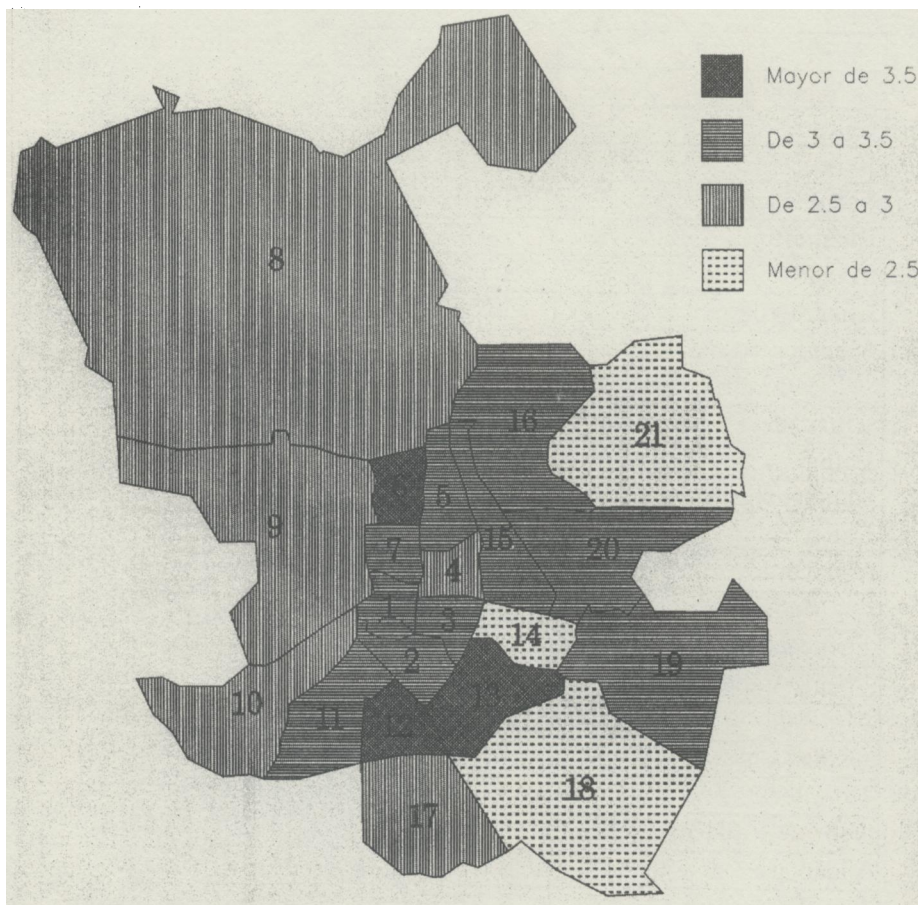
que proporciona los resultados que se recogen en la tabla 4 donde también se incluyen las potencias. Estos valores reflejan como en Tetuán, Carabanchel y Puente de Vallecas se debe rechazar la homogeneidad, luego los establecimientos existentes en estos distritos no están repartidos por igual entre las diferentes subramas.

CUADRO 3.

*Índice muestral de orden 2 en cada distrito según subrama de actividad*

Distrito	Índice	Intervalo al 95%
Madrid Centro-Oeste	3.713905	(3.523067, 3.904743)
Centro	3.157958	(2.730136, 3.585780)
Retiro	3.011973	(2.507280, 3.516666)
Salamanca	2.895784	(2.266351, 3.525216)
Chamartín	3.058894	(2.566604, 3.551185)
Tetuán	3.532496	(3.102295, 3.962697)
Chamberí	3.056715	(2.510386, 3.603044)
Moncloa-Aravaca	2.847998	(2.149552, 3.546443)
Madrid Sur	3.667386	(3.453298, 3.881473)
Arganzuela	3.395302	(3.072085, 3.718520)
Latina	2.810361	(2.191228, 3.429495)
Carabanchel	3.375762	(2.974030, 3.777494)
Usera	3.542899	(3.153845, 3.931953)
Villaverde	2.532496	(1.785228, 3.279764)
Madrid Norte-Este	4.016413	(3.791602, 4.241225)
Fuencarral-El Pardo	2.912538	(2.476765, 3.348310)
Puente Vallecas	3.517100	(3.106127, 3.928073)
Moratalaz	2.473932	(1.747206, 3.200657)
Ciudad Lineal	3.363625	(2.922445, 3.804804)
Hortaleza	3.029748	(2.441349, 3.618147)
Villa de Vallecas	2.473932	(1.747206, 3.200657)
Vicálvaro	3.058894	(2.558727, 3.559062)
San Blas	3.329308	(2.825035, 3.833581)
Barajas	2.321928	(1.740787, 2.903070)
Madrid Capital	3.867921	(3.748035, 3.987808)

MAPA 2.  
*Índice muestral de orden 2*



Districts según subrama de actividad

- |               |                 |                 |
|---------------|-----------------|-----------------|
| 1. Centro     | 8. Fuencarral   | 15. C. Lineal   |
| 2. Arganzuela | 9. Moncloa      | 16. Hortaleza   |
| 3. Retiro     | 10. Latina      | 17. Villaverde  |
| 4. Salamanca  | 11. Carabanchel | 18. V. Vallecás |
| 5. Chamartín  | 12. Usera       | 19. Vicálvaro   |
| 6. Tetuán     | 13. P. Vallecás | 20. San Blas    |
| 7. Chamberí   | 14. Moratalaz   | 21. Barajas     |

CUADRO 4.  
*Contraste de Homogeneidad*

Distritos	Conclusión al 95%	Potencia
Tetuán	Rechazar homogeneidad	0.993954
Arganzuela	No Rechazar homogeneidad	0.011004
Carabanchel	Rechazar homogeneidad	0.999777
Usera	No rechazar homogeneidad	0.122670
Puente Vallecas	Rechazar homogeneidad	0.797082

Además, dada la similitud de los valores del índice obtenidos es interesante realizar un contraste de hipótesis entre los distritos de Tetuán, Arganzuela, Carabanchel, Usera y Puente de Vallecas, con la finalidad de verificar si esas diferencias observadas son o no significativas.

Esto nos lleva a contrastar si es cierta la hipótesis

$$\text{Índ (Tetuán)} = \text{Índ (Arganzuela)} = \text{Índ (Usera)} = \text{Índ (Carabanchel)} = \text{Índ (P. de Vallecas)},$$

para lo cual el estadístico a utilizar es  $T_1^2$  que en este caso toma el valor  $T_1^2 = 0.4197863$ , lo que nos lleva a no poder rechazar dicha hipótesis.

Por tanto, podemos concluir que la presencia de las diferentes subramas en estos distritos es similar en cuanto al número pero, evidentemente, no tiene por qué serlo en cuanto al tipo de subramas representadas en cada una de las zonas consideradas.

Donde se observa una menor equidad en el reparto de las distintas subramas es en el distrito de Barajas -2.321928, Villa de Vallecas -2.473932, Moratalaz -2.473932 y Villaverde -2.532496. Estos valores están explicados por el menor número de establecimientos en los citados distritos así como la concentración en determinadas subramas: “imprentas”, “artículos metálicos”, “maquinaria industrial”, “pan y molinería”, “otras alimenticias” y “estructuras metálicas”.

Si agrupamos los distritos en zonas geográficas, al igual que hemos realizado anteriormente, el máximo valor se da en Madrid Norte-Este (4.016413) zona integrada por distritos con comportamientos muy dife-

rentes como aparece reflejado en el cuadro 3 (Barajas y Ciudad Lineal, por ejemplo), lo que explica el valor alcanzado.

Los distritos que integran las dos zonas restantes tienen un comportamiento similar lo que viene a corroborar los valores obtenidos, no obstante, existen diferencias lo que hará aconsejable la realización del correspondiente contraste de hipótesis para tomar una decisión sobre si las diferencias existentes entre estas dos zonas son o no significativas, lo que nos lleva a evaluar el estadístico  $'Z_2^2$ , que en este caso toma el valor  $'Z_2^2=0.2633898$ , indicándonos que no debemos rechazar la hipótesis nula, es decir no existen diferencias significativas entre ambas zonas, en las subramas de actividad.

#### 2.1.2. Según nivel de empleo

Si consideramos las tablas de resultados, por municipios (cuadro 5) y por distritos (cuadro 6), vemos la similitud en el comportamiento del índice. En los dos casos el valor máximo se encuentra en el estrato de 1-19 empleados con valores 3.848797 por distritos y 3.842307 por municipios lo que indica que en todas las ramas hay establecimientos con este estrato de empleo. Sucede a estos valores los que se relacionan con el estrato de 20-49 trabajadores, observándose diferencias en los valores obtenidos, siendo mayor por municipios 3.707582 frente al valor 3.221624 por distritos. En el estrato de empleo de 100-499 trabajadores los valores son idénticos, hay una coincidencia en la ubicación de las subramas de actividad.



CUADRO 5.

*Índice muestral de orden 2 en cada distrito según nivel de empleo*

Nivel de empleo	Índice	Intervalo al 95%
De 1 a 19 trabajadores	3.848797	(3.722348, 3.975246)
De 20 a 49 trabajadores	3.221624	(2.694305, 3.748943)
De 50 a 99 trabajadores	2.473932	(1.747206, 3.200658)
De 100 a 499 trabajadores	3.058894	(2.558727, 3.559062)
500 y más trabajadores	3.058894	(2.558727, 3.559062)
Madrid Capital	3.863126	(3.740472, 3.985780)

CUADRO 6.

*Índice muestral en cada zona geográfica según nivel de empleo*

Nivel de empleo	Índice	Intervalo al 95%
De 1 a 19 trabajadores	3.842307	(3.714860, 3.969754)
De 20 a 49 trabajadores	3.707582	(3.412186, 4.002979)
De 50 a 99 trabajadores	2.836502	(2.311436, 3.361568)
De 100 a 499 trabajadores	3.058894	(2.558727, 3.559062)
500 y más trabajadores	2.643857	(2.184424, 3.103289)
Madrid Provincia	3.919682	(3.806021, 4.036762)

Por otra parte, las diferencias hay que encontrarlas en el estrato de 50-99 trabajadores y en el estrato de 500 y más trabajadores. En concreto, el valor más bajo por distritos se da en el estrato de 50-99 trabajadores lo que indica una mayor localización de establecimientos en determinadas subramas, en particular ésta se produce en “imprentas” y en “máquinas de oficina”. Por el contrario, por municipios, el valor más bajo se encuentra en el estrato de 500 y más trabajadores, ya que de 10 establecimientos 2 se encuentran en la subrama de “maquinaria industrial”, 2 en “material electrónico” y 2 en “vehículos y sus piezas”.

CUADRO 7.  
*Contraste de Homogeneidad*

Nivel de empleo	Conclusión al 95%	Potencia
Madrid Capital		
De 1 a 19 trabajadores	Rechazar homogeneidad	1.000000
Madrid Provincia		
De 1 a 19 trabajadores	Rechazar homogeneidad	1.000000
De 20 a 49 trabajadores	Rechazar homogeneidad	0.958050

Esto nos lleva a plantearnos si existe homogeneidad en cuanto a la distribución de los establecimientos en el estrato de 1 a 19 trabajadores en el caso de Madrid Capital, así como en este estrato y en el de 20 a 49 trabajadores en el caso de Madrid Provincia, los resultados de estos contrastes se recogen en la tabla 7.

Por tanto, el número de establecimientos dentro de cada uno de estos niveles de empleo no se reparte por igual entre los diferentes distritos, en el primer caso, ni entre las diferentes zonas geográficas en el segundo, existiendo zonas donde la presencia de este tipo de establecimientos se ve acentuado con respecto a los demás.

## *2.2. Por año de inicio de la actividad*

A partir de la información proporcionada por el Ministerio de Industria y Energía es posible realizar un análisis de la antigüedad de los establecimientos industriales. Las categorías que intervienen en esta variable son 14, con un nivel de desagregación muy diferente, ya que una de ellas hace referencia al año de inicio de actividad de los establecimientos antes de 1980 y las 13 restantes hacen referencia a cada uno de los años hasta el año 1992. El valor máximo que indica un reparto similar de los establecimientos entre todos los años de inicio es 3.807556, teniendo en cuenta las 14 categorías señaladas.

El análisis de la antigüedad por municipios refleja como dos subzonas geográficas Eje Extremadura y Distrito Suroeste tienen valores muy similares y próximos al máximo (cuadro 8 y mapa 3), 3.477269 y 3.460046, respectivamente, valor que indica un reparto equitativo de establecimientos en los años analizados. El Distrito Suroeste mantiene un comportamiento muy diferente, ya que hay dos años en los que hay más establecimientos en los que se inicia la actividad: antes de 1980 y el año 1991, también es importante destacar como a partir del año 1986 se produce una recuperación económica ya que de los 97 establecimientos seleccionados y analizados 63 inician su actividad a partir de este año.

Sin embargo, será aconsejable realizar un contraste de hipótesis para ver si realmente existen diferencias significativas entre los dos valores observados y comprobar si es posible aceptar un comportamiento homogéneo en cada una de estas zonas.

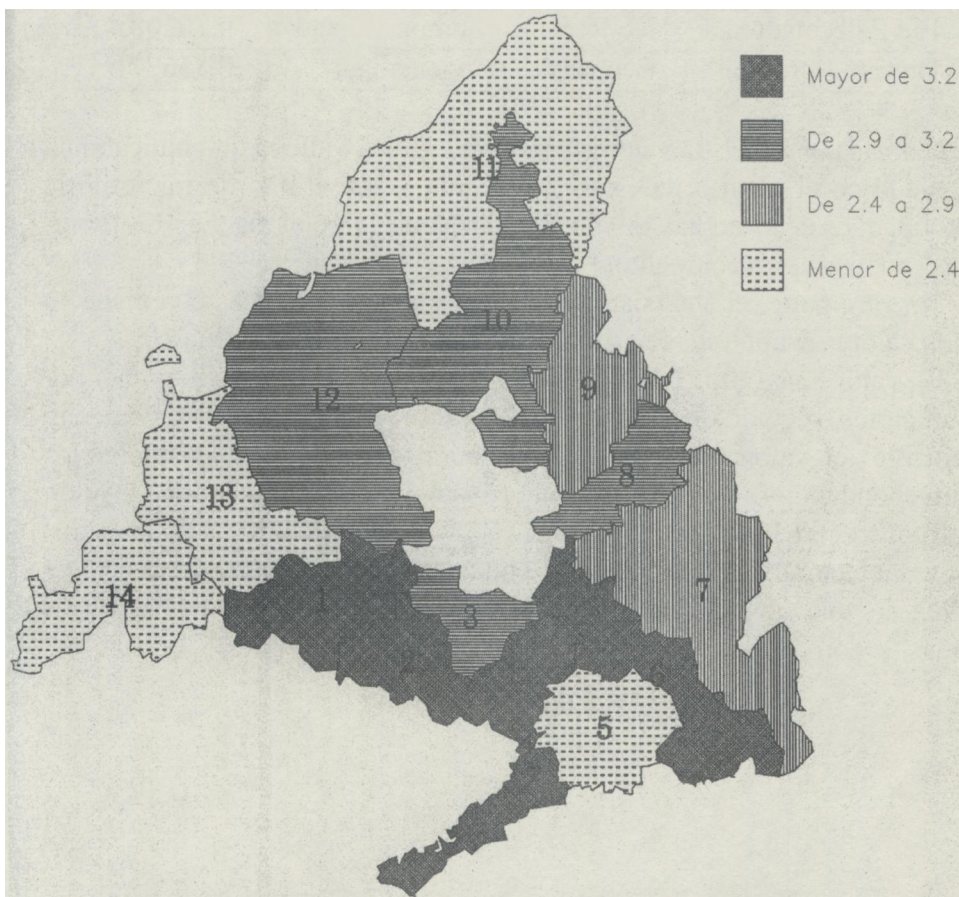
En cuanto a lo que se refiere a la comprobación de si realmente son diferentes es suficiente con evaluar  $Z_2^2$ , que en este caso al ser  $-1.96 = -z_{0.025} < Z_2^2 = 0.104671 < z_{0.025} = 1.96$ , se debe rechazar la existencia de diferencias significativas. En el cuadro 9 se recogen los resultados obtenidos en relación a la homogeneidad.

Por el contrario, el valor más bajo se corresponde con el Cono Oeste 1.836502, valor que indica concentración de establecimientos en unos determinados años en el inicio de la actividad, de los 10 establecimientos analizados cuatro se ubican en la categoría antes de 1980, 3 en el año 1991 y el resto en el año 1985, 1986 y 1991. Con un valor ligeramente superior 1.943417, se encuentra Extremo Suroccidental, también se analizan 10 establecimientos, tres antes de 1980, tres en 1987, dos en 1985 y dos en 1989.

CUADRO 8.  
*Índice muestral en cada zona geográfica  
 según año de inicio de la actividad*

Zonas Geográficas	Índice	Intervalo al 95%
Zona Sur	3.494931	(3.402219, 3.587642)
Eje Extremadura	3.477269	(3.275588, 3.678950)
Distrito Suroeste	3.460046	(3.279645, 3.640447)
Sur Metropolitano	3.098685	(2.600429, 3.596941)
Bajo Jarama	3.230047	(2.883106, 3.576987)
Vega Sureste	2.058894	(1.558727, 2.559062)
Zona Norte-Este	3.255511	(3.097172, 3.413851)
Eje Valencia	3.247499	(2.873264, 3.621734)
Cono Este	2.427011	(1.859620, 2.994401)
Eje Barcelona	3.107281	(2.820728, 3.393833)
Cono Nordeste	2.886790	(2.386884, 3.386696)
Eje Burgos	3.098695	(2.713524, 3.483866)
Sierra Norte	2.058894	(1.558727, 2.559062)
Zona Oeste	3.096970	(2.871990, 3.321949)
Eje Coruña-Sierra Central	3.191427	(2.868362, 3.514492)
Cono Oeste	1.836502	(1.167171, 2.505832)
Extremo Suroccidental	1.943417	(1.660688, 2.226146)
Madrid Provincia	3.406569	(3.312361, 3.500777)

MAPA 3.  
*Índice muestral de orden 2*



Subzona geográfica según año de inicio (a excepción de la Capital)

- |                      |                  |                      |
|----------------------|------------------|----------------------|
| 1. Eje Extremadura   | 6. Eje Valencia  | 11. Sierra Norte     |
| 2. Distrito Suroeste | 7. Cono Este     | 12. Eje Coruña S. C. |
| 3. Sur Metropolitano | 8. Eje Barcelona | 13. Cono Oeste       |
| 4. Bajo Jarama       | 9. Cono Nordeste | 14. Ext. Surocc.     |
| 5. Vega Sureste      | 10. Eje Burgos   | 15. Capital          |

CUADRO 9.  
*Contraste de Homogeneidad*

Nivel de empleo	Conclusión al 95%	Potencia
Eje Extremadura	No Rechazar homogeneidad	0.102082
Distrito Suroeste	Rechazar homogeneidad	0.856771

El análisis por zonas da valores superiores a tres, valores que indican una igualdad en el reparto de establecimientos según el año de inicio de la actividad, es decir no hay un año que destaque sobre el resto en la apertura de nuevos establecimientos.

Por otra parte, el análisis por distritos marca diferencias con respecto a los valores encontrados por municipios, como refleja la tabla 10.

El valor más alto 3.155279, San Blas, es inferior al máximo por municipios, seguido de Ciudad Lineal -3.070179 y Latina -3.062903. Sin embargo, el valor más bajo es inferior al encontrado por municipios correspondiéndose con el distrito de ~~Moncloa-Aravaca~~ Moncloa-Aravaca, 0.77608, seguido de Centro con 1.476203 y Salamanca con 1.851776. En estos distritos, los establecimientos analizados se concentran en una categoría, antes de 1980, lo que nos indica que los establecimientos existentes en estas zonas no son de reciente creación.

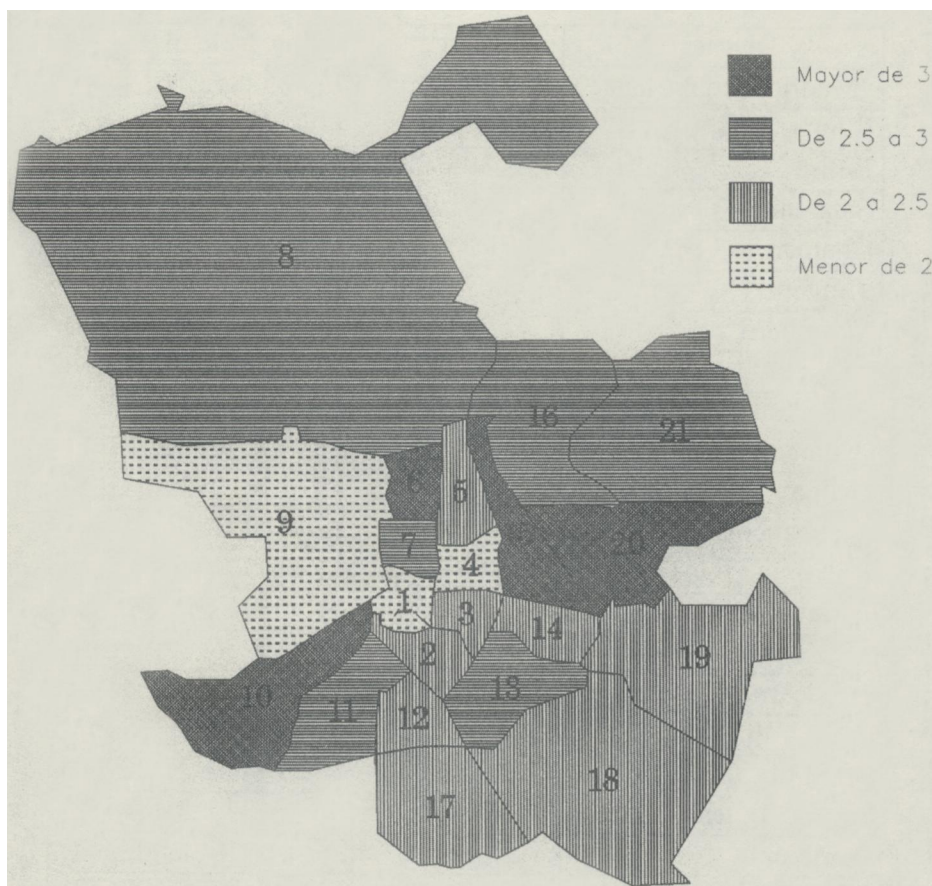
CUADRO 10.

*Índice muestral en cada distrito según año de inicio de la actividad*

Distritos	Índice	Intervalo al 95%
Madrid Centro-Oeste	2.350606	(2.085245, 2.615967)
Centro	1.476203	(1.012793, 1.939613)
Retiro	2.395302	(1.407858, 3.382746)
Salamanca	1.851776	(1.213293, 2.490260)
Chamartín	2.395929	(1.840300, 2.951559)
Tetuán	3.025960	(2.634435, 3.417485)
Chamberí	2.762267	(2.290247, 3.234288)
Moncloa-Aravaca	0.777608	(0.099749, 1.455466)
Madrid Sur	2.839535	(2.549977, 3.929094)
Arganzuela	2.163976	(1.455790, 2.872162)
Latina	3.062903	(2.621610, 3.504196)
Carabanchel	2.726201	(2.358271, 3.094132)
Usera	2.470143	(1.824536, 3.115749)
Villaverde	2.481870	(1.957318, 3.006421)
Zona Norte-Este	3.254935	(3.015256, 3.494614)
Fuencarral-El Pardo	2.752073	(2.169725, 3.334421)
Puente Vallecas	2.790961	(2.242947, 3.338976)
Moratalaz	2.058894	(1.210820, 2.906969)
Ciudad Lineal	3.070179	(2.696340, 3.444018)
Hortaleza	2.614710	(2.226419, 3.003002)
Villa de Vallecas	2.184425	(1.775197, 2.593653)
Vicálvaro	2.643857	(2.184424, 3.103290)
San Blás	3.155279	(2.816701, 3.493858)
Barajas	2.643857	(2.184424, 3.103290)
Madrid Capital	2.822638	(2.653984, 2.991291)

Como puede observarse, el mapa 4, en el que se consideran 4 niveles para el índice, sintetiza este último análisis.

MAPA 4.  
*Índice muestral de orden 2*



Distrito según año de inicio

- |               |                 |                 |
|---------------|-----------------|-----------------|
| 1. Centro     | 8. Fuencarral   | 15. C. Lineal   |
| 2. Arganzuela | 9. Moncloa      | 16. Hortaleza   |
| 3. Retiro     | 10. Latina      | 17. Villaverde  |
| 4. Salamanca  | 11. Carabanchel | 18. V. Vallecás |
| 5. Chamartín  | 12. Usera       | 19. Vicálvaro   |
| 6. Tetuán     | 13. P. Vallecás | 20. San Blas    |
| 7. Chamberí   | 14. Moratalaz   | 21. Barajas     |



## CONCLUSIONES

El objetivo final de este trabajo es disponer de un nuevo instrumento científico así como de una técnica de medición en el análisis de la distribución espacial de los establecimientos industriales en una determinada región. Por este motivo, el trabajo se inclina en una doble vertiente.

Por una parte, en la propuesta de las bases teóricas para la obtención de un nuevo índice dando a conocer diferentes resultados del comportamiento del mismo cuando su valor se estima mediante un muestreo estratificado. Ello nos permite obtener diferentes informaciones de la población bajo estudio que en este caso será la Comunidad de Madrid. Para este primer objetivo nos sumergimos en un análisis de las características o variables básicas que son más utilizadas en los diferentes trabajos existentes: número de establecimientos productivos, personas ocupadas, antigüedad, productividad, valor añadido, etc., como se observa, por ejemplo en el trabajo de Martínez (1998) que utiliza, también, como fuente los Directorios Industriales, fuente de carácter general que ofrecen visiones estáticas de la actividad económica y permite ser útil para investigaciones globales. Es importante señalar que a diferencia de lo que ocurre en estos trabajos, y que es la forma habitual de proceder, nosotros no trabajamos con toda la población.

En relación con el segundo objetivo, el trabajo consiste en presentar, de forma estructurada, las principales conclusiones. Entre ellas, destaca el hecho de que los resultados aquí obtenidos son similares a los proporcionados por otros estudios donde se trabaja, como ya se ha señalado, con valores absolutos y considerando toda la población. En este trabajo se utilizan observaciones procedentes de diferentes muestras con las ventajas asociadas a la utilización de las mismas: coste menor para obtener la información adicional cuando así lo requieran los objetivos de la investigación al ser menor el número de observaciones a realizar o para contrastar la información en los casos en los que se dude de la veracidad de la fuente, etc., muestras que nos permitirán obtener estimaciones de los índices considerados con las que realizamos el estudio.

Para lograr estos objetivos se ha recurrido a la utilización de medidas procedentes de la Teoría de la Información utilizadas en diferentes campos como la ecología, biología, sociología, economía, entre otras, como puede verse en Kapur (1989). En este trabajo se ha optado por la utilización de diversos índices basados en la entropía de orden  $r$  de Renyi, por el buen comportamiento que tiene como medida de diversidad como puede verse en Mayoral (1997). Cabe destacar en favor de esta medida que permite detectar de forma clara las posibles diferencias existentes, debido a que en ella interviene la función logaritmo, función que hace que las diferencias se acentúen. Esta metodología se pone en práctica con el estudio empírico que se realiza sobre una de las comunidades españolas como es la Comunidad de Madrid. Aunque esta metodología es totalmente general y puede ser puesta en práctica en el estudio de cualquier zona geográfica.

En relación con el estudio empírico realizado, los resultados obtenidos reflejan la importante diversidad espacial de la industria en la Comunidad de Madrid:

- a) La actividad industrial muestra una fuerte concentración en torno a la capital. Los distritos de Usera, Tetuán, Puente de Vallecas y Arganzuela aparecen como enclaves en los que se da un reparto equitativo entre las distintas subramas mientras que la industria apenas tiene presencia en la periferia: Barajas, Villa de Vallecas y Moratalaz, quedando representada en unas pocas subramas.
- b) El núcleo central genera unos efectos difusores que potencian el desarrollo de ejes industriales sobre las grandes vías de comunicación que incluso se extiende a provincias colindantes. El cuadro industrial es el resultado no sólo de cierres y aperturas de nuevos establecimientos industriales sino también de la relocalización de establecimientos ya existentes que cambian de ubicación en toda o en parte de su estructura. Los valores altos se observan en Cono Nordeste, Sur Metropolitano, Eje Burgos y Eje Barcelona (estos dos últimos favorecidos por las actuaciones públicas y privadas creándose parques tecnológicos de gran dinamismo).

- c) Los sectores que incorporan servicios avanzados en la producción alcanzando un mayor desarrollo en la comunidad madrileña son edición e imprentas junto con la fabricación de material eléctrico y electrónico. En contraposición, los sectores con un menor protagonismo son metálicas básicas, industria no metálica y textil y calzado. Se observa que todas las ramas están presentes en la región, con mayor o menor peso, permitiendo hablar de diversificación de la industria a la vez que da un carácter equilibrado a la estructura regional.

La industria madrileña no es ajena al comportamiento del conjunto de la industria española, de la que forma parte y por ello sujeta a la propia dinámica del sector. Debe señalarse por tanto, que la industria madrileña seguirá una evolución similar a la del conjunto del país, con unos rasgos estructurales que le hacen tener una situación privilegiada ante las posibles fluctuaciones del ciclo económico, traducido en una doble vertiente: peso de la industria regional dentro de la nacional y, derivada de la anterior, la industria madrileña está en condiciones óptimas de competitividad dentro del marco europeo.

## BIBLIOGRAFÍA

- Agresti, A. y Agresti, B. F., "Statistical analysis of qualitative variation", *Social Methodology*, 1978, pp. 204-237.
- Celada Crespo, F., "Industria y reestructuración territorial en la comunidad de Madrid", *Situación*, Serie Estudios Regionales, Banco Bilbao Vizcaya, 1998, pp. 285-304.
- Cochran, W. G., *Sampling Techniques*, John Wiley, 1977.
- Comunidad de Madrid, *La economía madrileña en sus sectores*, Consejería de Economía, Dirección General de Economía y Planificación, 1988.
- , *Anuario estadístico 1992*, Departamento de Estadística, Consejería de Economía, Departamento de Estadística, 1993.
- , *Directorio industrial de la Comunidad de Madrid 1992*, vol. 1, Explotación Estadística, Departamento de Estadística, Consejería de Economía, 1993.
- , *Directorio industrial de la Comunidad de Madrid 1992*, vol. 2, Guía de Establecimientos, Departamento de Estadística, Consejería de Economía, 1993.
- , *Atlas de la industria en la Comunidad de Madrid*, Departamento de Estadística, Consejería de Economía, 1994.
- , *Estructura industrial de la Comunidad de Madrid*, Departamento de Estadística, Consejería de Economía, 1994.
- Iranzo, J. E., Del Río, C. y Molina, M., "Madrid, entre dos modelos de desarrollo", *Papeles de Economía Española*, Fundación para la Investigación Económica y Social, CECA, 55, 1993, pp. 256-277.
- Kapur, J. N., *Maximun Entropy Models in Science and Engineering*, John Wiley, 1989.
- Lewontin, R. C., "The apportionment of human diversity", *Evolutionary Biol.*, 6, 1972, pp. 381-398.
- Liebersson, S., *Testing Statistical Hyphotesis*, John Wiley, Nueva York, 1969.

- Martínez, R., "Dinámica del sector de fabricación de material electrónico en la Comunidad de Madrid", en *El futuro de la industria en la Comunidad de Madrid*, Asociación Madrileña de Ciencia Regional, Instituto de Estadística, Comunidad de Madrid, 1998.
- Mayoral, M. M., "Estudio de la población activa española a través de la diversidad de Shannon", *Esic-Market*, 75, 1992, pp. 147-168.
- , "Una aproximación a la realidad industrial en la Comunidad de Madrid a través de la entropía de orden  $r$ ", Ph. D. thesis, Universidad Complutense de Madrid, 1995.
- , "Algunas medidas de concentración industrial", *Estudios Geográficos*, LVII (227), 1997, pp. 193-202.
- Méndez, R., "Actividad industrial y estructura territorial en la región de Madrid", Comunidad de Madrid, Consejería de Trabajo, Industria y Comercio, 1986.
- Nayak, T. K., "Applications of entropy functions in measurement and analysis of diversity", Ph. D. thesis, University of Pittsburgh, 1983.
- , "On diversity measures based on entropy functions", *Commun Statist, Theory and Methods*, 14, 1985, pp. 203-215.
- Patil, G. P. y Talle, C., "An overview of diversity", *Ecological Diversity in Theory and Practice*, 1, 1979, pp. 3-28.
- Pielou, E. C., *Ecological Diversity*, Nueva York, John Wiley, 1975.
- Rényi, A., "On the measures of entropy and information", *Proc. 4th. Berkeley Symp. Math, Statist and Prob.*, 1, 1961, pp. 547-561.
- Shannon, C. E., "A mathematical theory of communication", *The Bell System Technical Journal*, 27, 1948, pp. 379-423.