UNA APROXIMACION A LA LOCALIZACION DE PEQUEÑAS AREAS GEOGRAFICAS DE POBREZA

Ernesto J. Veres Ferrer
Departamento de Economía Aplicada
Universidad de Valencia

SUMARIO

En este trabajo se presenta una metodología que pretende estimar la localización de pequeñas áreas geográficas dentro de un municipio, susceptibles de poseer altos porcentajes de población pobre. Estas pequeñas áreas son las secciones estadísticas y la pobreza considerada es, esencialmente, pobreza económica en el sentido definido en el estudio *Las condiciones de vida de la población pobre de la Comunidad Valenciana*, realizado por la Fundación FOESSA (1.995). En cualquier estudio o investigación dirigida a la población pobre, la aplicación de la metodología es inmediata: el método establece una gradación de las secciones censales atendiendo a su probable mayor o menor porcentaje de pobreza, lo que permite optimizar los recursos de la investigación centrándola en las unidades territoriales donde la pobreza esté más extendida.

1. Introducción

En los últimos años - sobre todo, desde la década de los ochenta - se han realizado diversas investigaciones sobre la pobreza, inicialmente dirigidas a su cuantificación global, para más tarde ampliarse a la determinación de las características de la población pobre. Es el caso, por ejemplo, de los informes sobre *Las condiciones de vida de la población pobre,* de la Fundación FOESSA, a los que haremos referencia más adelante. De esos estudios se aprecia la existencia de zonas geográficas donde la pobreza tiene mayor intensidad. Pero dichas zonas - directamente ligadas a la estratificación empleada - son muy amplias. Precisamente, el estudio de Díe (1998) para Cáritas Diocesana de Valencia sobre *La Realidad Social de la Parroquia de Oliva,* utiliza como unidad territorial para el análisis la sección censal. En el presente trabajo se desarrolla una metodología de jerarquización de esas pequeñas unidades territoriales, atendiendo a su porcentaje esperado de pobreza.

Independientemente de la consideración de la pobreza como realidad pluridimensional - y, por tanto, en la que intervienen variados factores no exclusivamente económicos, sino afectivos, culturales, sanitarios, étnicos, etc. - aquí entenderemos como pobreza la que está medida en términos sólo económicos. Por tanto, si el objetivo de cierta investigación es el estudio de la pobreza de cierto territorio, puede resultar interesante a priori excluir de la misma aquellas zonas cuya población, de entrada, tiene ya características socioeconómicas que la alejan de esa pobreza económica. De esta formas la investigación puede centrarse en aquellas otras zonas cuyas características invitan a suponer la existencia de bolsas de pobreza. Esta forma de proceder pretende, claramente, conseguir una mayor eficacia a la hora de utilizar los recursos económicos y humanos, siempre escasos, con los que suele trabajarse en cualquier estudio o investigación.

También el estudio territorial de cualquier variable, el conocer cómo se distribuye espacialmente, resulta sugerente a la hora de establecer los motivos y las causas de su evolución, y proporciona pistas para estimar el comportamiento futuro si aceptamos un modelo de repulsión-atracción de la población, de o hacia zonas con o sin presencia, respectivamente, de esa posible característica no deseada. Consecuentemente, el establecimiento de un mapa de riqueza-pobreza, esto es, su localización gráfica, podría ser otra de las utilidades de este trabajo, y serviría para facilitar la decisión del responsable político o del planificador social a la hora de determinar prioridades de actuación sobre el territorio.

2. Definición de pobreza

La definición de pobreza adoptada en este trabajo es la utilizada en el informe ya citado de la Fundación FOESSA. Por lo tanto, consideraremos como listón en el que ubicar una situación de pobreza el de unos ingresos de la unidad familiar inferiores al 50% de la renta media de las familias del país (criterio UE). Esta pobreza es definida como de *precariedad social o pobreza relativa*, frente a las pobrezas moderada, grave y extrema que suponen bajar el listón a porcentajes inferiores (35%, 20% y 15%, respectivamente). Utilizando este criterio, y para datos recabados en 1.994 para el conjunto de la Comunidad Valenciana, el 20'60% de los hogares son pobres, afectando al 24'6% de la población valenciana.

Se define el *indicador básico o tasa de pobreza (H)* como la proporción de familias cuyos ingresos son inferiores a la línea de pobreza considerada. Esta linea de pobreza mide el nivel de ingresos que deben percibirse por una familia para que, en un determinado contexto social, no sea considerada pobre. Dicha línea, como hemos indicado, se ha establecido en el 50% de los ingresos medios percibidos por el conjunto de familias, por lo que H mide, en definitiva, el porcentaje de familias que no alcanzan a percibir unos ingresos que socialmente se valoran como suficientes. El informe FOESSA realiza la estimación del indicador de pobreza H a partir de la Encuesta de Presupuestos Familiares 1990-91, utilizando la distribución del ingreso familiar.

Los valores del indicador básico H para España y para las provincias valencianas se recoge en la tabla siguiente:

Tabla 1

Indicador básico o tasa de pobreza (H). Año 1.994

Ambito territorial	H
España	19'4
Comunidad Valenciana	20'6
Alicante	20'2
Castellón	23'2
Valencia	20'3

3. Aproximación a la localización territorial de la pobreza: el caso de un único municipio

La finalidad prevista con la metodología a desarrollar es el establecimiento de una ordenación de más a menos entre las secciones estadísticas que forman parte de un municipio, señalando, pues, aquellas en las que probablemente puedan encontrarse mayores porcentajes de pobreza. En Veres (1991) se justificaba el uso de la sección estadística o sección censal en las investigaciones de cáracter estadístico, en cuanto que la sección censal se constituye en el elemento último de referencia de la información estadística que va a ser tratada. En efecto, es unidad territorial más pequeña a la que están referidos los datos del Censo de Población de 1.991, últimos que, con carácter general y exhaustivo, están disponibles. Concretamente, para cada sección censal son conocidas ciertas características socioeconómicas básicas de su población, características que han sido utilizadas en los informes FOESSA a la hora de establecer correlaciones con el indicador básico o tasa de pobreza H.

La sección censal se define, salvo excepciones, como un territorio continuo habitado por un mínimo de 500 habitantes y un máximo de 2.500 personas, y perfectamente delimitado por accidentes geográficos y/o urbanísticos. Por ello, y desde un punto de vista social, su pequeña amplitud asegura una constatable homogeneidad para sus habitantes y familias - sobre todo en términos de su posición relativa comparada con la de las demás secciones -, por lo que pueden aceptarse válidas para cada unidad individual que la compone, las características medias de todos los demás individuos de la sección.

3.1 Descripción de la metodología propuesta

El indicador básico o tasa de pobreza H definido en el informe FOESSA está correlacionado con ciertas características socioeconómicas básicas de la población. Concretamente, en dicho informe se publican las correlaciones - medidas a través del respectivo coeficiente de correlación simple - entre el coeficiente H y diez características socieconómicas fundamentales, tal como se recoge en la tabla siguiente:

Tabla 2

Correlaciones entre el indicador básico (H)

y ciertas variables socioeconómicas

Variables	**Coeficientes**
X_1 : Tasa de analfabetismo	0.614453
X_2 : Tasa de población mayor de 65 años	0.231722
X_3 : Tasa de población mayor de 16 años	-0.216540

X_4 : Tasa de población activa	-0.462340
X_5 : Porcentaje de parados	0.275401
X_6 : Tasa de población ocupada	-0.275400
X_7 : Tasa de empleo agrícola	0.494691
X_8 : Tasa de empleo industrial	-0.665550
X_9 : Tasa de empleo en la construcción	0.569038
X_{10} : Tasa de empleo en el sector servicios	-0.120450

Las anteriores correlaciones están calculadas atendiendo a las características socioeconómicas de la población de todas las provincias de España, y valores cercanos a ±1 indican la existencia de una fuerte relación entre la pobreza y la variable considerada. Se aprecia, pues, que las correlaciones más altas son las expresadas al poner en relación al indicador básico de pobreza H y la tasa de analfabetismo (0'6145) y la tasa de empleo industrial (-0'6656). El distinto signo de ambas correlaciones indican una relación inversa entre ambas: la mayor presencia de empleo industrial proporciona estabilidad económica y, destierra, consecuentemente, la posibilidad de pobreza. Respecto el analfabetismo - sin entrar en consideraciones sobre si constituye causa o es efecto de la pobreza - su signo positivo lo confirma como claro indicador de posible marginación económica.

El conocimiento de las correlaciones, junto con el conocimiento de la información de las características socioeconómicas de la tabla 2, permiten definir un coeficiente para cada una de las secciones censales, expresado a través de una combinación lineal. Este coficiente o indicador tiene un sentido de globalidad, e intenta acercarse al conocimiento del nivel de pobreza a partir de la información indirecta proporcionada por cierta información socioecónomica disponible para la población. Así pues, sean conocidos los siguientes elementos:

- un conjunto $\{X_i\}_{i=1}^{I}$ de variables que, a priori, pueden tomar cualquier valor;

- las variables X_i anteriores están correlacionadas con otra variable, H, a través de sendos coeficientes de correlación $\{-1 \leq r_{X_i H} \leq 1\}_{i=1}^{I}$. Para simplificar la notación, estableceremos $r_{X_i H} = r_i$, $\forall$i ;

Las variables X_i toman sus valores sobre ámbitos territoriales concretos, es decir, tienen una intrínseca referencia territorial. Y así, teniendo en cuenta que la estructura de las correlaciones de la Tabla 2 se refiere a todo el país, supondremos que dicha estructura obedece a la correlación existente al considerar todos los posibles ámbitos territoriales más pequeños (comunidad autónoma, provincias, municipios, distritos y secciones censales, y todas las combinaciones posibles entre esos ámbitos), al correlacionar el nivel de pobreza con los valores que las variables X_i toman en todos esos subámbitos. Con X_{ij} denotaremos, pues, el valor que la variable X_i toma sobre la unidad territorial j-ésima.

Estableceremos como hipótesis - perfectamente aceptable en la práctica - la de la no uniformidad en los valores de cada variable, esto es:

$$\forall i \quad \exists\, j_{1i} \text{ y } j_{2i} \text{ tales que } X_{j1i} \neq X_{j2i}$$

asegurándose, pues, la existencia de un valor máximo y un valor mínimo para cada X_i. En lo que sigue, y una vez hayan sido determinados, consideraremos siempre que esos dos valores tienen la condición de constantes.

Definimos la siguiente transformación de las variables anteriores:

$$Y_{ij} = \frac{X_{ij} - \min_j(X_{ij})}{\max_j(X_{ij}) - \min_j(X_{ij})} \qquad \{1\}$$

transformación empleada, por ejemplo, en ciertos índices utilizados por las Naciones Unidas para establecer una jerarquía de países atendiendo a su desarrollo (los Indices de Desarrollo Humano y de Desarrollo de Género, incluídos en el Informe sobre Desarrollo Humano 1.998 del Programa de las Naciones Unidas para el Desarrollo-PNUD). La transformación {1} permite ubicar el valor de cada variable y para cada uno de los ámbitos territoriales, atendiendo a su posición relativa respecto al rango de variación posible. El efecto conseguido es normalizar las variables con las que se va a trabajar, al darles un rango de variación común a todas ellas, con la consiguiente comparabilidad del orden de magnitud de las varianzas y covarianzas que aparecerán en las expresiones desarrolladas adelante. Consecuentemente:

$$0 < X_{ij} < 1 \quad \forall i,j$$

$$\forall i \ \exists\, j_{Mi} \text{ y } j_{mi} \text{ tales que } X_{jMi} = 1 \text{ y } X_{jmi} = 0$$

Dado el carácter lineal de las transformaciones anteriores, se verifica:

$$r_{X_i H} = r_{Y_i H} = r_i \quad \forall i \qquad \{2\}$$

Pretendemos definir un *indicador global* sobre toda unidad territorial, en particular sobre cada sección censal como unidad territorial mínima de la que se dispone de información estadística. Trabajando con las variables transformadas, se define el siguiente indicador Ψ_j asociado al ámbito territorial j-ésimo:

$$\Psi_j = \sum_{i=1}^{I} sig(r_i) \cdot r_i^2 \cdot Y_{ij} \qquad \{3\}$$

donde $sig(\rho_i)$ es la función "signo".

La expresión {3} se interpreta como que Ψ es una combinación lineal de las variables transformadas Y_i - variables que en cada unidad territorial j-ésima están relacionadas con el porcentaje de pobreza en dicha unidad expresado por H_j - , ponderadas por sus respectivos coeficientes de determinación. Estos coeficientes pueden concebirse como la contribución de aquéllas sobre el indicador básico H. Se respeta el signo del coefiente de correlación respectivo para que en Ψ se mantenga el sentido de la contribución en H de cada variable socioeconómica considerada.

Se acepta, pues, como hipótesis fundamental de trabajo que las correlaciones deducidas con la información provincial y que establecen las relaciones a nivel estatal, son válidas para su aplicación a ámbitos infraprovinciales y, en especial, municipales. Sin embargo, explotaciones específicas de la información del estudio FOESSA permitiría calcular la estructura de las ponderaciones propia de una única Comunidad Autónoma o de una única provincia, y que debería ser la estructura a utilizar en {3} para aquellas aplicaciones que traten de ámbitos territoriales subsumidos en éstas.

La correlación entre el indicador básico de pobreza H y el coeficiente global Ψ se configura como elemento esencial para valorar su bondad. Sean, al respecto, los siguientes resultados:

Proposición 1ª

El coeficiente de correlación lineal entre Y y H es:

$$r_{\Psi H} = \frac{1}{s_{\Psi}} \cdot \sum_{1=1}^{I} r_i^3 \cdot sig(r_i) \cdot s_i \geq 0 \qquad \text{\{4\}}$$

siendo:

$$s_{\Psi}^2 = \sum_{i=1}^{I} r_i^4 \cdot s_i^2 + 2 \cdot \sum_{i} \sum_{k>i}^{I} r_i^2 \cdot r_k^2 \cdot sig(r_i) \cdot sig(r_k) \cdot s_{ik}$$

y en donde s_i e s_{ik} indican, respectivamente, la desviación standard y la covarianza de las variables transformadas Y_i e Y_k.

En efecto: la obtención de s_{Ψ}^2 es inmediata, como varianza de una combinación lineal de variables no necesariamente independientes. De ahí que:

$$s_{\Psi H} = \sum_{i=1}^{I} r_i^2 \cdot sig(r_i) \cdot s_{iH}$$

$$r_{\Psi H} = \frac{s_{\Psi H}}{s_{\Psi} s_H} = \frac{1}{s_{\Psi)}} \cdot \sum_{1=1}^{I} r_i^2 \cdot sig(r_i) \cdot \left\{ \frac{s_{iH}}{s_H} \right\} =$$

$$= \frac{1}{s_{\Psi}} \cdot \sum_{1=1}^{I} r_i^2 \cdot sig(r_i) \cdot \left\{ \frac{s_{iH}}{s_H \cdot s_i} \right\} \cdot s_i =$$

$$\frac{1}{s_{\Psi}} \cdot \sum_{1=1}^{I} r_i^3 \cdot sig(r_i) \cdot s_i \geq 0$$

toda vez que se verifica {2}, y siendo $r_i^3 \cdot sig(r_i) \geq 0 \quad \forall$i , *c.q.d.*

Dado el valor positivo de la correlación anterior, se deduce, pues, el sentido positivo de la relación entre el indicador Ψ y H , por lo que a mayor porcentaje de pobreza en la sección j-ésima es de esperar un mayor valor para su respectivo coeficiente Ψ_j . Además se verifica:

Corolario

$$\sum_{i=1}^{I} \frac{r_i \cdot sig(r_i)}{1 + \frac{1}{r_i^2 \cdot s_i} \sum_{k \neq i}^{I} r_k^2 \cdot s_k} = \sum_{i=1}^{I} \frac{r_i^3 \cdot sig(r_i) \cdot s_i}{\sum_{i=1}^{I} r_i^2 \cdot s_i} \leq r_{\Psi H}$$

$$r_{\Psi H} \leq \sum_{i=1}^{I} \frac{r_i \cdot sig(r_i)}{\sqrt{\left(1 + \frac{1}{r_i^4 \cdot s_i^2} \sum_{k \neq i} r_k^4 \cdot s_k^2\right)}} \qquad \{5\}$$

En efecto:

a) Basta considerar que $0 \leq sig(\rho_i).sig(\rho_k).\sigma_{ik} \leq \sigma_i.\sigma_k \quad \forall i,k$:

$$r_{\Psi H} \geq \frac{\sum_i r_i^3 \cdot sig(r_i) \cdot s_i}{\sqrt{\sum_i r_i^4 \cdot s_i^2 + 2 \cdot \sum_i \sum_{k>i} r_i^2 \cdot r_k^2 \cdot s_i s_k}} =$$

$$= \frac{\sum_i r_i^3 \cdot sig(r_i) \cdot s_i}{\sqrt{\left(\sum_i r_i^2 \cdot s_i\right)^2}} = \frac{\sum_i r_i^3 \cdot sig(r_i) \cdot s_i}{\sum_i r_i^2 \cdot s_i} =$$

$$= \sum_i \frac{r_i^3 \cdot sig(r_i) \cdot s_i}{\sum_i r_i^2 \cdot s_i} = \sum_i \frac{r_i^3 \cdot sig(r_i) \cdot s_i}{r_i^2 \cdot s_i \left(1 + \frac{1}{r_i^2 \cdot s_i} \sum_{k \neq i} r_k^2 \cdot s_k\right)} =$$

$$= \sum_i \frac{r_i \cdot sig(r_i)}{\left(1 + \frac{1}{r_i^2 \cdot s_i} \sum_{k \neq i} r_k^2 \cdot s_k\right)} \qquad\qquad c.q.d.$$

b) Basta considerar que sig(ρ_i).sig(ρ_k).$\sigma_{ik} \geq 0 \;\; \forall$i,k :

$$r_{\Psi H} \leq \frac{\sum_i r_i^3 \cdot sig(r_i) \cdot s_i}{\sqrt{\sum_i r_i^4 \cdot s_i^2}} = \sum_i \frac{r_i^3 \cdot sig(r_i) \cdot s_i}{\sqrt{\sum_i r_i^4 \cdot s_i^2}} =$$

$$= \sum_i \frac{r_i^3 \cdot sig(r_i) \cdot s_i}{\sqrt{r_i^4 \cdot s_i^2 \left(1 + \frac{1}{r_i^4 \cdot s_i^2} \sum_{k \neq i} r_k^4 \cdot s_k^2\right)}} = \sum_i \frac{r_i \cdot sig(r_i)}{\sqrt{\left(1 + \frac{1}{r_i^4 \cdot s_i^2} \sum_{k \neq i} r_k^4 \cdot s_k^2\right)}} \qquad c.q.d.$$

La desigualdad del Corolario de la Proposición 1ª anterior sugiere, con gran generalidad, una correlación entre Ψ y H no inferior a la de las variables componentes. Concretamente, para la estructura de ponderaciones de la Tabla 2 y con los datos de todas las secciones censales de la Comunidad Valenciana, la cota inferior anterior toma el valor de 0'5440331, por lo que la correlación entre el indicador Ψ y H es, presumiblemente, superior.

En efecto, la aplicación del indicador Ψ tiene pleno sentido siempre que su correlación con H sea superior a la existente entre éste y las variables Y_i que definen a aquél. Al respecto, se demuestra fácilmente que al menos existe siempre una de las variables X_i cuya correlación con H es inferior a la que esta tasa de pobreza tiene con el indicador Ψ. Pero, además, la proposición siguiente presenta una condición necesaria y suficiente para que la correlación entre Ψ y H sea superior a la existente entre ésta y todas las variables componentes que definen a Ψ:

Proposición 2ª

Es condición necesaria y suficiente para que el indicador Y gane correlación respecto la de todas sus variables componentes Y, que se verifique:

$$\sum_{i>1}^{I} r_i^4 \cdot s_i^2 \cdot (r_i^2 - r_1^2) + 2 \cdot \sum_i \sum_{k>i} r_i^2 \cdot r_k^2 \cdot \left\{ |r_i| \cdot |r_k| \cdot s_i s_k - r_1^2 \cdot |s_{ik}| \right\} \geq 0$$

donde Y_1 es la variable tal que $|r_1| \geq |r_i| \;\; \forall i \neq 1$.

En efecto. Siendo Y_1 la variable de mayor correlación absoluta:

$$r_{\Psi H} \geq |r_1| \;\; \leftrightarrow \;\; r_{\Psi H}^2 \geq r_1^2 \;\; \leftrightarrow$$

$$\Leftrightarrow \quad \frac{\sum_{i=1}^{I} r_i^6 \cdot s_i^2 + 2\cdot \sum_{i>k}\sum_{k} r_i^3 \cdot r_k^3 \cdot sig(r_i)\cdot sig(r_k)\cdot s_i s_k}{\sum_{i=1}^{I} r_i^4 \cdot s_i^2 + 2\cdot \sum_{i>k}\sum_{k} r_i^2 \cdot r_k^2 \cdot sig(r_i)\cdot sig(r_k)\cdot s_{ik}} \geq r_1^2 \quad \leftrightarrow$$

$$\leftrightarrow \quad \sum_{i>1} r_i^6 \cdot s_i^2 + 2\cdot \sum_{i>k}\sum_{k} r_i^3 \cdot r_k^3 \cdot sig(r_i)\cdot sig(r_k)\cdot s_i s_k \geq$$

$$\geq \sum_{i>1} r_i^4 \cdot r_1^2 \cdot s_i^2 + 2\cdot r_1^2 \cdot \sum_{i>k}\sum_{k} r_i^2 \cdot r_k^2 \cdot sig(r_i)\cdot sig(r_k)\cdot s_{ik} \quad \leftrightarrow$$

$$\leftrightarrow \quad \sum_{i>1} r_i^4 \cdot s_i^2 \cdot (r_i^2 - r_1^2) +$$

$$+2\cdot \sum_{i>k}\sum_{k} r_i^2 r_k^2 \left\{ r_i \cdot sig(r_i)\cdot r_k \cdot sig(r_k)\cdot s_i s_k - r_1^2 \cdot sig(r_i)\cdot sig(r_k)\cdot s_{ik} \right\} \geq 0 \quad \leftrightarrow$$

$$\leftrightarrow \quad \sum_{i>1}^{I} r_i^4 \cdot s_i^2 \cdot (r_i^2 - r_1^2) + 2\cdot \sum_{i>k}\sum_{k} r_i^2 \cdot r_k^2 \cdot \left\{ |r_i|\cdot|r_k|\cdot s_i s_k - r_1^2 \cdot |s_{ik}| \right\} \geq 0 \qquad c.q.d.$$

La aplicación de la expresión de la *Proposición 2ª* sobre todas las secciones censales de la Comunidad Valenciana confirma su signo positivo (+0'01595003), lo que implica para esta Comunidad una ganancia en correlación entre el indicador Ψ y H , respecto la existente entre éste y cada variable componente Y de Ψ, suponiendo válida para la Comunidad Valenciana la estructura de correlaciones recogida en la Tabla 2.

Pero aún más, ya que el uso del indicador global Ψ tiene pleno sentido siempre que suponga ganancia de correlación con H respecto la existente entre ésta y las variables por separado, la exigencia de alcanzar una mayor correlación obliga a preguntarse por el número N de variables que deben formar parte del indicador. Al respecto hay que notar que de las diez variables estudiadas en la Tabla 2, algunas de ellas presentan correlaciones muy bajas (y, con mayor motivo, coeficientes de determinación). Por ello, su inclusión en la expresión {3} podría hacer disminuir la correlación final entre H y Ψ.

La elección del número de variables que deben formar parte de Ψ se realiza teniendo en cuenta el siguiente resultado, previa ordenación de las mismas según su correlación absoluta con H:

Proposición 3ª

Suponiendo una ordenación de las variables Y_i, atendiendo a la fuerza de su relación lineal con H , según $|r_i| \geq |r_{i+1}|$ "i , y definiendo

$$\Psi(N) = \sum_{i=1}^{N} sig(r_i)\cdot r_i^2 \cdot Y_i \qquad \{6\}$$

se verifica la siguiente condición necesaria y suficiente:

$$r_{\Psi(N+1)H} \leq r_{\Psi(N)H} \quad \text{«} \quad 1+\frac{r_{N+1}^{3}\cdot sig(r_{N+1})\cdot s_{N+1}}{\sum_{i=1}^{N} r_i^3 \cdot sig(r_i)\cdot s_i} \leq$$

$$\sqrt{1+\frac{r_{N+1}^{4}\cdot s_{N+1}^{2}+2\cdot\sum_{i=1}^{N} r_i^2 sig(r_i)\cdot r_{N+1}^{2} sig(r_{N+1})\cdot s_{i(N+1)}}{\sum_{i=1}^{N} r_i^4\cdot s_i^2+2\cdot\sum_{i=1}^{N}\sum_{k>i}^{N} r_i^2 sig(r_i)\cdot r_k^2 sig(r_k)\cdot s_{ik}}}$$

En efecto, siendo:

$$\Psi(N)=\sum_{i=1}^{N} sig(r_i)\cdot r_i^2\cdot Y_i$$

$$\Psi(N+1)=\sum_{i=1}^{N+1} sig(r_i)\cdot r_i^2\cdot Y_i=\Psi(N)+sig(r_{N+1})\cdot r_{N+1}^{2}\cdot Y_{N+1}$$

$$s_{\Psi(N+1)}^{2}=s_{\Psi(N)}^{2}+r_{N+1}^{4}s_{N+1}^{2}+2\cdot\sum_{i=1}^{N} r_i^2 sig(r_i)\cdot r_{N+1}^{2} sig(r_{N+1})\cdot s_{i(N+1)}=$$

$$=s_{\Psi(N)}^{2}+d$$

$$r_{\Psi(N+1)H}=\frac{1}{s_{\Psi(N+1)}}\cdot\sum_{1=1}^{N+1} r_i^3\cdot sig(r_i)\cdot s_i=$$

$$=\frac{\sum_{1=1}^{N} r_i^3\cdot sig(r_i)\cdot s_i+r_{N+1}^{3} sig(r_{N+1})\cdot s_{N+1}}{\sqrt{s_{\Psi(N)}^{2}+d}}=\frac{a+c}{\sqrt{s_{\Psi(N)}^{2}+d}}$$

De donde:

$$r_{\Psi(N+1)H} \leq r_{\Psi(N)H} \quad \text{«} \quad \frac{a+c}{\sqrt{s_{\Psi(N)}^{2}+d}} \leq \frac{a}{\sqrt{s_{\Psi(N)}^{2}}} \quad \text{«}$$

$$\text{«} \quad (a+c)\cdot\sqrt{s_{\Psi(N)}^{2}} \leq a\cdot\sqrt{s_{\Psi(N)}^{2}+d} \quad \text{«}$$

$$\text{«} \quad \left(1+\frac{c}{a}\right) \leq \sqrt{1+\frac{d}{s_{\Psi(N)}^{2}}}$$

toda vez que *a,c,d > 0 , c.q.d.*

La proposición anterior proporciona un criterio para la elección del número de variables que deben formar parte en la definición de {3} : dada la ordenación entre las variables Y_i atendiendo a su correlación absoluta con el nivel de pobreza H , en la expresión del indicador Ψ irán incorporándose variables hasta encontrar la primera de ellas cuya contribución de lugar a una correlación entre Ψ y H inferior a la existente sin considerarla. Esta variable no se considera en la definición de Ψ , y se repite el estudio de correlaciones entre las ya aceptadas y las siguientes variables, hasta que, definitivamente, son rechazadas las que no aumentan la correlación final entre Ψ y H .

La aplicación que va desarrollarse posteriormente se realiza sobre ciertos municipios de la Comunidad Valenciana. Por ello, la Proposición anterior, aplicada a la información del Censo de Población de 1.991 para el conjunto de secciones censales de la Comunidad Valenciana (un total de 2.938 secciones) proporciona el resultado de la Tabla 3, en la que las columnas I y II expresan los miembros primero y segundo, respectivamente, de la desigualdad de la Proposición 3ª:

Tabla 3

Comparación entre	**I**	**II**
$\rho_{\Psi(1)}$ y $\rho_{\Psi(2)}$	1’5489	1’2240
$\rho_{\Psi(2)}$ y $\rho_{\Psi(3)}$	1’2845	1’1927
$\rho_{\Psi(3)}$ y $\rho_{\Psi(4)}$	1’1797	1’1654
$\rho_{\Psi(4)}$ y $\rho_{\Psi(5)}$	1’0899	1’0647
$\rho_{\Psi(5)}$ y $\rho_{\Psi(6)}$	1’0197	1’0310
$\rho_{\Psi(6)}$ y $\rho_{\Psi(7)}$	1’0175	1’0202
$\rho_{\Psi(7)}$ y $\rho_{\Psi(8)}$	1’0109	1’0214
$\rho_{\Psi(8)}$ y $\rho_{\Psi(9)}$	1’0092	1’0198
$\rho_{\Psi(9)}$ y $\rho_{\Psi(10)}$	1’0023	1’0112

Se aprecia, pues, que a partir de la sexta variable (porcentaje de parados) no se incrementa la correlación entre Ψ y H . Por ello, en la definición de Ψ a desarrollar en la aplicación posterior, sólo intervendrán las cinco variables siguientes: tasa de empleo industrial, tasa de analfabetismo, tasa de empleo en la construcción, tasa de empleo agrícola y tasa de población activa.

Se dispone, por tanto, de una valoración sobre el porcentaje de pobreza en cada una de las secciones censales de los municipios. Sin embargo, esa valoración no puede entenderse en términos absolutos, sino sólo relativos. Y ello porque, en primer lugar, faltan otras variables que también pueden influyen sobre el porcentaje de pobreza; en segundo lugar, las mismas variables X_i pueden ser en parte redundantes, presentando autocorrelaciones entre ellas; y, finalmente, la influencia de cada X_i sobre H no viene dada necesariamente por el respectivo coeficiente de determinación de la regresión lineal, ρ_i^2 , sino que esta correlación es sólo aproximación a su influencia real. Todas estas consideraciones suponen la no significatividad absoluta de la aditividad definida en {3} .

En definitiva, el indicador Ψ_j no puede interpretarse como medida directa y absoluta del nivel de pobreza de esa sección. Por ello, por ejemplo, un valor de Ψ_j para una sección que sea doble que para otra no indica un porcentaje doble de pobreza. Pero sí tiene pleno sentido utilizarlo en términos relativos, pudiéndose establecer una jerarquía probabilística sobre las secciones censales susceptibles de tener altos porcentajes de pobreza. Y así, este indicador admite una ordenación, que supone la ordenación equivalente según el porcentaje de pobreza esperado en cada sección.

De ahí que la aplicabilidad del método expuesto exija las siguientes tres consideraciones:

1ª No tiene sentido realizar a través Ψ_j comparaciones absolutas entre secciones. Por ello, carece de sentido comparar los resultados obtenidos al aplicar el procedimiento por separado a dos municipios o territorios distintos. Los distintos valores máximo y mínimo correspondientes a cada variable X_{ij} en cada municipio impide la estricta comparabilidad en aplicaciones diferentes de la metodología a diferentes municipios.

2ª La metodología es de plena aplicación para la ubicación de aquellas secciones censales de un municipio concreto que es de preveer posean altos porcentajes de pobreza. Esto es, permite el establecimiento de una gradación de las secciones de un municipio, atendiendo a su porcentaje de pobreza probable.

3ª No obstante lo indicado en el punto anterior, la aplicabilidad final del indicador Ψ tendrá pleno sentido siempre y cuando haya ganancia en correlación (condición expresada en la *Proposición 2ª*), o cuando $\Psi(N)$ se haya construído previa selección de ciertas variables en línea al resultado de la *Proposición 3ª*.

Finalmente, el método descrito podría aplicarse, teóricamente, a unidades territoriales mayores que la sección censal (distritos, municipios, comarcas, etc.), sin más que considerar los valores que sobre ellas toman las variables socioeconómicas de interés. Sin embargo, conforme van englobándose secciones, los valores de las variables sobre las nuevas unidades agrupadas van adoptando la condición de medias, por lo que las diferencias entre unidades se suavizan, dificultando el posterior análisis basado en el estudio de las diferencias. De ahí que, para estos ámbitos mayores, la metodología vaya perdiendo utilidad.

3.2 Aplicación al municipio de Gandía

La aplicación de la metodología anterior a las secciones censales del municipio de Gandía (Valencia) da como resultado la Tabla 4, que resume las diferentes etapas de su aplicación. Concretamente, la Tabla 4 incluye la información transformada de la censal según la expresión {1}, sobre la que se calculan los indicadores $\Psi(5)$ definidos según {6} y que se recogen en su última columna. Las secciones aparecen ya ordenadas, atendiendo al valor descendente del indicador global. La bondad del método se manifiesta, por otra parte, por la aproximación a {4} deducida a partir de la información de las 37 secciones censales de Gandía, en cuanto que ofrece una correlación entre $\Psi(5)$ y H igual a 0'894784, superando la máxima correlación (absoluta) entre H y la variable "tasa de empleo industrial". La ordenación de las secciones confirma, pues, a la sección 004 del Distrito 02 como la de potencial mayor porcentaje de pobreza. Por el contrario, la de esperado menor porcentaje de pobreza es la que tiene un coeficiente Ψ menor, que corresponde a la sección 012 del Distrito 02.

Tabla 4

Información transformada e indicadores Yj. Secciones censales de Gandía

Sección	Tasa industria	Tasa Analfab	Tasa construc	Tasa agricola	Tasa activos	Ψ_j.100
Gandía	0.438	0.202	0.350	0.139	0.472	-7.154
02 004	0.039	1.000	1.000	0.250	0.523	63.334
07 002	0.360	0.958	0.491	1.000	0.233	55.635
04 002	0.222	0.508	0.389	0.619	0.045	36.124
05 002	0.547	0.569	0.618	0.127	0.256	14.910
06 001	0.498	0.349	0.643	0.198	0.210	12.325
01 003	0.000	0.088	0.216	0.032	0.006	10.969
07 001	0.419	0.437	0.403	0.317	0.392	10.365
03 003	0.256	0.442	0.198	0.155	0.341	8.232
04 004	0.424	0.109	0.339	0.615	0.239	6.298
06 002	0.532	0.378	0.449	0.119	0.341	0.880
05 009	0.429	0.171	0.657	0.147	0.557	0.435
04 003	0.064	0.076	0.283	0.060	0.563	-1.393
03 004	0.621	0.303	0.657	0.087	0.432	-1.875
01 001	0.158	0.140	0.000	0.036	0.051	-1.931
01 002	0.286	0.081	0.138	0.083	0.000	-3.106
04 001	0.443	0.127	0.325	0.464	0.483	-3.279
05 001	0.256	0.000	0.194	0.067	0.045	-4.374
05 007	0.271	0.153	0.364	0.008	0.483	-4.566
02 006	0.384	0.231	0.435	0.107	0.625	-4.952
05 003	0.153	0.072	0.166	0.071	0.545	-8.567
02 003	0.562	0.320	0.300	0.143	0.460	-9.425
02 001	0.310	0.067	0.226	0.040	0.352	-10.443
02 008	0.635	0.185	0.548	0.123	0.483	-10.741
05 004	0.488	0.009	0.509	0.198	0.540	-11.459
05 005	0.384	0.033	0.163	0.111	0.193	-11.929
02 011	0.261	0.004	0.512	0.024	0.926	-14.030
03 001	0.507	0.096	0.145	0.000	0.045	-15.136

05 008	0.635	0.176	0.329	0.044	0.250	-15.147
02 010	0.463	0.151	0.495	0.083	0.949	-17.020
02 009	0.478	0.230	0.279	0.052	0.716	-17.467
02 002	0.498	0.022	0.350	0.107	0.551	-19.045
03 002	0.734	0.163	0.357	0.099	0.341	-19.656
05 006	0.399	0.006	0.113	0.056	0.597	-25.184
06 003	0.453	0.089	0.293	0.028	0.983	-27.544
02 007	0.626	0.113	0.208	0.032	0.938	-35.969
02 005	1.000	0.138	0.318	0.163	0.682	-39.383
02 012	0.675	0.084	0.265	0.000	1.000	-39.513

4. Comparación de los niveles de pobreza de varios municipios: la aplicación sucesiva del método

Dependiendo del objetivo perseguido por la investigación concreta a desarrollar, la consideración conjunta de dos o más municipios presenta el inconveniente de la posible heterogeneidad de sus situaciones respectivas frente a la pobreza. Una posible actuación posterior sobre todos los municipios invalida el tratamiento conjunto a sus secciones, salvo que pudiera establecerse como razonable la hipótesis de que las características socioeconómicas de la población de todos ellos son semejantes, con la consiguiente igualdad de los valores máximo y mínimo para cada una de las variables Xij en los municipios considerados. Si la hipótesis no fuera cierta, el cálculo a partir del conjunto de todas las secciones de los coeficientes Ψ_j - coeficientes que no tienen significación absoluta - , estaría afectado por un factor de media, enmascarando las posibles diferencias entre las secciones.

En este sentido, se aprecia fácilmente que los valores de Ψ_j tienden a concentrase conforme aumenta el número de secciones consideradas, esto es, su varianza tiende a hacerse más pequeña conforme aumenta el número de secciones, si bien su distribución es claramente asimétrica negativa y no admite un buen ajuste normal, como también puede comprobarse a partir del correspondiente contraste de bondad de ajuste efectuado sobre el conjunto de secciones de la Comunidad Valenciana. De ahí que ese efecto tamaño invalide la comparación de los coeficientes deducidos por la aplicación por separado a las secciones de distintos municipios.

Por ello, se propone la aplicación de la anterior metodología en dos fases diferentes.

4.1 Descripción del método para varios municipios

4.1.1 Fase 1ª, por separado, ordenación de las secciones

La primera fase implica sendos cálculos de los coefientes Ψ_j , realizados por separado a cada municipio. De esta manera, se dispondría de sendas ordenaciones de secciones, atendiendo al valor del indicador Ψ_j respectivo, y con validez exclusiva dentro del municipio al que pertenezcan las secciones.

Dado el carácter relativo de este coeficiente, las ordenaciones anteriores no son comparables. Por ello, hay que utilizar un criterio único de selección de secciones censales para una nueva y única ordenación.

4.1.2 Elección del criterio de selección de las secciones probablemente con alto nivel de pobreza

La elección prevista supone definir un criterio único para todos los municipios considerados, que determine cuáles son las secciones susceptibles de atención por su nivel de pobreza esperado. Esta elección es, claramente, arbitraria y, de entrada, podría formularse en términos absolutos o relativos.

Sin embargo, se propone seleccionar, para un tratamiento conjunto posterior, aquellas secciones de un municipio cuyo coeficiente Ψ_j cumpla una condición del tipo siguiente:

$$\Psi_j \geq \Psi_{municipio} + k \cdot S_{\forall j} \qquad \{7\}$$

donde $\Psi_{municipio}$ es el valor que toma el coeficiente sobre las respectivos valores transformados de las variables X_i referidas a todo el municipio, y $S_{\forall j}$ es la desviación típica de los coeficientes Ψ_j de las secciones del municipio. El valor de la constante k es arbitrario, y dependerá de la extensión final de la investigación que sobre la pobreza desee realizarse, una vez concretadas las secciones con un alto porcentaje potencial de pobreza. Esta extensión deberá tener en cuenta, evidentemente, la existencia de recursos disponibles, tanto económicos como humanos. En la aplicación que desarrollamos después se toma un valor de $k=1$.

El anterior criterio de selección se aplica por separado a las secciones de cada municipio. Por tanto, todas las secciones seleccionadas finalmente verifican la propiedad común de que "*su coeficiente Y_j dista por encima del valor que dicho coefiente toma sobre los valores correspondientes a todo el municipio, más de una vez la desviación típica de los coeficientes correspondientes al conjunto de secciones de ese municipio*". Evidentemente que esta selección es sensible al número total de secciones por municipio: a más secciones, también aumenta el número de secciones seleccionadas.

4.1.3 Fase 2ª, conjunta con las secciones seleccionadas, jerarquización de la pobreza probable

Una vez seleccionadas las secciones de los dos municipios considerados, se aplica nuevamente el método a todas ellas, de forma conjunta. Se obtienen una vez más los coefientes Ψ_j , cuya ordenación tiene ahora pleno sentido, en cuanto que, todas ellas, satisfacen la misma propiedad definida a través de la condición $\{7\}$ anterior.

La ordenación puede ser muy ilustrativa. Si, por ejemplo, todas las secciones de un municipio, o mayoritariamente, se colocan en los primeros puestos de la ordenación, se está manifestando un posible porcentaje de pobreza superior en las secciones pobres de ese municipio respecto las de los otros. Si, por el contrario, la ordenación dispone las distintas secciones de los municipios entrelazadas, los niveles de porcentaje de pobreza en las secciones de ambos municipios no tienen una tendencia mayoritaria en uno respecto al otro. Es necesario, pues, dar significatividad estadística a la apreciación anterior.

Se propone, pues, el test de Kruskal-Wallis para contrastar las hipótesis:

H_0: No hay diferencia en la distribución de los porcentajes de pobreza
en las secciones seleccionadas de los municipios considerados
H_1: Si hay diferencias

completado con el método de Dunn para determinar sobre qué municipios se producen las diferencias significativas, en el caso de que el contraste anterior nos invite a rechazar H_0 .

4.2 Aplicación a los municipios entre 50.000 y 100.000 habitantes de la Comunidad Valenciana

Se aplica la metodología propuesta a los municipios de la Comunidad Valenciana cuya población está comprendida entre 50.000 y 100.000 habitantes. Son: Alcoy, Elda y Orihuela, en la provincia de Alicante; y Sagunto, Torrente y Gandía, en la provincia de Valencia.

La aplicación, por separado, a cada uno de los seis municipios anteriores proporciona los siguientes resultados que ahora comentamos. El número de secciones seleccionadas aplicando el criterio expresado por $\{7\}$, utilizando como coeficiente $k=1$, con expresión de los límites inferiores para cada $\Psi(5)$ respectivo a partir del cual se efectúa la selección, el número total de secciones censales de cada municipio y la aproximación a $\{4\}$ deducida a partir de la información de las dichas secciones censales, quedan recogida en la tabla 5 :

Tabla 5

Situación comparativa tras la aplicación del método a los municipios de la Comunidad Valenciana cuya población está comprendida entre 50.000 y 100.000 habitantes

Municipio	(1)	(2)	(3)	(4)
Alcoy	51	10	11.725	0.8995
Elda	39	5	12.768	0.9193
Orihuela	35	6	30.7505	0.8680
Gandía	37	4	14.3043	0.8948
Sagunto	42	8	18.1528	0.8561
Torrente	39	7	17.5185	0.8144

Columna (1): Nº total de secciones del municipio
Columna (2): Nº de secciones seleccionadas según el criterio {7}
Columna (3): Límite inferior de $\Psi(5)$ a partir del cual se seleccionan las secciones de la columna (2)
Columna (4): Aproximación a la correlación entre $\Psi(5)$ y H expresada según {4} y calculada con las secciones del municipio

La última columna, que expresa la aproximación a {4} deducida a partir de la información de las correspondiente secciones censales de cada municipio, ofrece correlaciones entre $\Psi(5)$ y H en todos ellos muy superiores a la máxima correlación (absoluta) entre H y la variable "tasa de empleo industrial".

Finalmente, y tras la aplicación por separado del método a cada uno de los seis municipios, la Tabla 6 recoge la ordenación de las secciones seleccionadas, una vez aplicada por segunda vez la metodología sobre ellas, atendiendo al valor mayor o menor de su coeficiente $\Psi(5)$:

Tabla 6

Ordenación según el indicador Y de las Secciones censales seleccionadas de los municipios de la Comunidad Valenciana cuya población está comprendida entre 50.000 y 100.000 habitantes

Posición	Sección	Ψ*100	Posición	Sección	Ψ*100
1	08 002 Orihuela	56.142	21	01 006 Sagunt	4.69
2	06 004 Orihuela	55.124	22	02 007 Torrent	1.989
3	02 004 Gandía	52.571	23	02 001 Torrent	0.446
4	06 007 Orihuela	47.731	24	01 005 Alcoi	-1.78
5	08 003 Orihuela	46.397	25	02 003 Torrent	-2.925
6	08 005 Orihuela	42.354	26	04 002 Torrent	-4.231
7	07 002 Gandía	39.078	27	01 001 Elda	-7.154
8	07 005 Orihuela	33.004	28	07 006 Alcoi	-7.348
9	03 003 Sagunt	24.121	29	01 004 Alcoi	-8.392
10	02 001 Sagunt	22.058	30	02 006 Alcoi	-9.78
11	04 002 Gandía	21.609	31	03 005 Alcoi	-10.889
12	01 004 Sagunt	20.203	32	01 006 Torrent	-10.988

13	01 008 Torrent	18.15	33	02 005 Alcoi	-12.892
14	05 002 Gandía	16.382	34	01 006 Alcoi	-17.403
15	04 018 Sagunt	11.663	35	03 001 Elda	-18.474
16	01 003 Sagunt	8.269	36	02 004 Alcoi	-18.573
17	01 001 Alcoi	7.051	37	08 003 Alcoi	-22.819
18	02 005 Torrent	5.827	38	01 002 Elda	-28.411
19	04 002 Sagunt	5.799	39	01 005 Elda	-32.689
20	03 006 Sagunt	4.806	40	02 001 Elda	-34.587

La aplicación de la teoría estadística establece la significatividad de la tabla 6. En efecto, la aplicación del test de Kruskal-Wallis a sus datos invita a rechazar la hipótesis de igualdad de comportamiento ante la pobreza de las secciones seleccionadas de esos municipios, afirmación con validez estadística para prácticamente cualquier nivel de significación.

Para determinar dónde se producen las diferencias significativas, la aplicación del criterio de Dunn proporciona la siguiente Tabla 7 de resultados:

Tabla 7

Significatividad de los comportamientos ante la pobreza de los municipios de la Comunidad Valenciana cuya población está comprendida entre 50.000 y 100.000 habitantes

Municipio	Elda	Orihuela	Gandía	Sagunt	Torrent
Alcoi	No	Si	Si	No	No
Elda		Si	Si	Si	No
Orihuela			No	No	No
Gandía				No	No
Sagunt					No

Se aprecia, pues, comportamientos diferenciales ante la pobreza en las secciones censales seleccionadas de Alcoi respecto Orihuela y Gandía; y de Elda respecto Orihuela, Gandía y Sagunt. Los demás emparejamientos entre municipios no ofrecen significatividad estadística a la hora de valorar las diferencias de los niveles de pobreza entre sus secciones seleccionadas.

5. Bibliografía

CASAS SANCHEZ, J.M. (1.996): *Inferencia estadística para Economía y Administración de empresas*. Ed. Centro de Estudios Ramón Areces, S.A. Madrid.

DIE OLMOS, L. (1998): *Estudio de la Realidad Social de la Parroquia de San Roque. Oliva.* Informe realizado para Cáritas Diocesana de Valencia y Cáritas Parroquial de San Roque (Oliva).

FUNDACION FOESSA (1995): *Las condiciones de vida de la población pobre de la Comunidad Valenciana.* Fundación Foessa, Madrid.

FUNDACION FOESSA (1.998): *Las condiciones de vida de la población pobre en España.* Fundación Foessa, Madrid.

INSTITUTO VALENCIANO DE ESTADÍSTICA (1.994): *Cens de Població 1.991.* Generalitat Valenciana.

ONU (1998): Informe sobre Desarrollo Humano 1.998. *Programa de las Naciones Unidas para el Desarrollo. PNUD-1998*. Ed. Mundi-Prensa Libros, S.A. Madrid.

SIEGEL, S. (1.970): *Estadística no paramétrica, aplicada a las ciencias de la conducta.* Ed. Trillas, México.

VERES FERRER, E. (1991): “Tipología de secciones estadísticas”, en *Actas de las I Jornadas Internacionales sobre Demografía Urbana y Regional*. Instituto de Demografía del Consejo Superior de Investigaciones Científicas. Madrid.