

ANEXO 15

UNIFORMIDAD Y EQUILIBRIO DEL TERRITORIO

1. DEFINICIÓN Y CONCEPTOS PREVIOS

Es frecuente oír hablar, en los últimos tiempos, tanto a los políticos como a los sociólogos, economistas, urbanistas, geógrafos, ingenieros, juristas, etc., acerca de la conveniencia de que el "territorio se halle equilibrado", o bien de que deba alcanzarse un cierto grado de "equilibrio territorial", o de que no debe trastocarse el ya existente, sin que, a continuación, se intente explicar o clarificar el significado real que tales palabras y expresiones encierran (FRANQUET, 1990/91).

Desde un primer punto de vista, **el equilibrio territorial se logrará cuando las masas socioeconómicas de población y de renta** (que vienen siendo objeto de estudio en anteriores capítulos de nuestro trabajo) **se hallen distribuidas por el territorio del modo más uniforme y homogéneo posible**, sin discontinuidades, pero también sin grandes concentraciones desequilibradoras. En términos matemáticos, y también en los que hemos venido utilizando en nuestro Análisis Territorial, ello equivaldría a procurar la asimilación del territorio que se planifica hacia un espacio de tipo continuo y no discreto.

En otro orden de ideas, el equilibrio territorial puede asimilarse a un estado del sistema físico en el cual todos sus componentes o elementos disfruten de un similar grado de accesibilidad a los servicios públicos y privados (educación, sanidad, cultura, transporte, comercio, ocio, ...).

El caso de Cataluña, que por cierto nos servirá para analizar paramétricamente el equilibrio de un territorio específico, es bastante típico por lo que se refiere al desequilibrio territorial que aquí tratamos de estudiar. Desde luego, en el fenómeno de la capital macrocefálica (en este caso, Barcelona), la compresión de los desequilibrios en la distribución espacial de las masas de población y de renta debe entenderse ligada a la concentración tradicional de las inversiones en infraestructuras y servicios diversos (abastecimientos de agua, depuradoras, hipermercados, servicios culturales y recreativos, etc...) en el área de la gran capital y en detrimento del resto del territorio de su jurisdicción, es decir, en este caso, de las 4 provincias, 41 comarcas, 944 municipios y 54 entidades locales menores (o entidades municipales descentralizadas) que conforman, actualmente, el conjunto catalán.

Es digno de remarcar, no obstante, que si las inversiones mencionadas se han venido concentrando y limitando al área macrocefálica es, esencialmente, la naturaleza del tipo de infraestructura escogido la que provoca que siempre se localicen dichas actuaciones -por razones de rentabilidad interna- en las proximidades de la capital congestionada del país. No juzgamos necesario entrar aquí en la discusión pormenorizada acerca de los mecanismos financieros y/o especulativos que, con gran frecuencia, acompañan a la promoción y construcción de los grandes elementos de infraestructura, aunque sí creemos conveniente denunciar el funcionamiento absurdamente autónomo con que estas operaciones acostumbran a plantearse; de este modo, prácticamente los únicos estímulos que explican la implantación de las mismas serán el consumo generado por la propia infraestructura y los beneficios esperables de su misma construcción.

Aceptando esta lógica constructiva con la que aquí discreparemos (autopistas, grandes hipermercados, etc.) el punto de destino beneficiado por la inversión acabará estando siempre en la capital del territorio o en sus aledaños. Y si, por alguna razón fortuita, dicha inversión notable se localiza a mayor distancia, lo será con un grado de concentración o peligrosidad tan fuerte (centrales nucleares para la producción de energía eléctrica, plantas de tratamiento de residuos industriales, ...) que ello constituye una forma subrepticia de colonizar una porción más del territorio para la capital.

Por todo ello, juzgamos necesario crear nuevos tipos o nuevas condiciones de los elementos de las infraestructuras, que permitan romper la lógica rutina de la concentración en la capital o para la capital y que, además, refuercen y hagan posible la acción de fomento armónico del crecimiento en el conjunto del territorio. Dicha política no es técnicamente imposible, ni debe suponer, siquiera, la aparición de deseconomías de gestión. Y ello, precisamente, porque la gestión concentrada resulta compatible con una cierta dispersión espacial, y porque es posible rechazar la naturaleza tan "determinísticamente" definida de los prototipos de infraestructura que venimos criticando.

Todo ello implicará conceder un papel fundamental a las comarcas y a los municipios (en el orden político, económico y cultural), potenciando la ligazón y los intercambios entre los diferentes enclaves territoriales; de este modo, se producirá una interrelación saludable de las actividades económicas y culturales, de tal suerte que los sub-territorios no se centren en una sola especialidad. Debe tenderse, en fin, a que la estructura económica de las unidades territoriales de ámbito local presente una diversificación notoria entre los diferentes sectores económicos, mediante la aplicación de conceptos tales como el de la "industrialización difusa", el aprovechamiento de los recursos turísticos y otros.

El desarrollo coherente de estos principios de actuación, concretándolos en actuaciones específicas, debería configurar, ciertamente, el núcleo de una política territorial eficiente y moderna. Pero el hecho es que, en numerosas ocasiones, las actuaciones que los gobiernos emprenden están impulsadas por la necesidad urgente de cubrir déficits de todo tipo. Y es que, debe distinguirse entre la perentoriedad de las acciones puntuales (con frecuencia ineludibles) y las actuaciones programadas para el conjunto del territorio con miras al logro de su reequilibrio, lo cual exige tomar decisiones para influir positivamente en la orientación de las grandes tendencias. *Para conseguir todo ello, juzgamos deseable, en nuestros administradores y políticos, un buen bagaje de ideas claras junto al sentido de la perseverancia y de la continuidad.*

Ahora bien, cabe hacerse diversas preguntas, como por ejemplo: ¿cuáles son y habrán de ser las grandes tendencias de distribución de la población y de las actividades socioeconómicas en el territorio? ¿Qué papel deberá desempeñar la capital más o menos macrocefálica y el resto del territorio? ¿Es razonable que los ritmos de crecimiento cuantitativo de dicha capital se mantengan en el futuro o, por el contrario, resulta más lógico que se modere o incluso anule el ritmo de crecimiento de la región metropolitana en beneficio del resto del territorio de su jurisdicción? De conseguirse esto último, se podría llegar a una doble situación: a) que en la región metropolitana se pondría más énfasis en las mejoras cualitativas, y b) que, en el resto del territorio, un crecimiento más acelerado permitiría obtener, también, un mayor crecimiento cualitativo, puesto que tanto en tecnología como en cultura, sanidad, enseñanza, etc., las dimensiones demográficas y económicas posibilitan la aparición de iniciativas mayormente ambiciosas. Trátase, en definitiva, de que la región metropolitana encauce su futuro en base, fundamentalmente, a mejoras cualitativas, mientras que el resto del territorio debería realizarlo simultáneamente con mejoras cualitativas y cuantitativas, sin perder de vista, no obstante, que las tendencias de la distribución de las masas socioeconómicas por el territorio no se pueden manipular o "pilotar" con facilidad, en la mayor parte de los casos.

Sí podemos intuir, sin embargo, que la oferta de puestos de trabajo por parte de los centros urbanos e industriales no tendrá, en el futuro, la misma intensidad que en pasado. La diversificación sectorial de las actividades y la disminución de las distancias físicas o de los tiempos de desplazamiento hacen que la creatividad económica se halle cada vez más generalizada a lo largo y ancho del territorio, sin tener que concentrarse, necesariamente, en poblaciones de rancia tradición industrial, o bien poseedoras de otras ventajas diferenciales. *En cualquier caso, con independencia de tendencias espontáneas, la acción de gobierno deberá enfocar al desarrollo equilibrado y armónico del conjunto territorial, basándose en su diversificación sectorial, en el*

aumento general del nivel de conocimientos, en la abolición de la frontera existente entre el mundo urbano y el rural, en la facilidad de las comunicaciones, en la industrialización difusa y en el apoyo, generoso y decidido, a las zonas relativamente deprimidas.

La transformación, en los términos expuestos, de los modelos macrocefalos actuales en la articulación del territorio, exigirá, en fin, unas bases de gestión urbana diferentes a las que se han venido empleando mayoritariamente hasta la fecha, y que tomen como alternativa el fomento decidido de la capacidad de todos y cada uno de los enclaves del territorio en cuestión, conjugando eficazmente la unidad con la diversidad (FRANQUET, 1990/91).

2. EQUILIBRIO TERRITORIAL: SÍMIL ESTÁTICO-MECÁNICO

2.1. CONDICIONES DE EQUILIBRIO TERRITORIAL

El concepto de "fuerza económica" y "momento de una fuerza económica", que ya han sido expuestos por el doctorando en algunas publicaciones¹, servirán de base a la exposición que sigue. Consideremos, ahora, un territorio sobre el que actúan varias fuerzas económicas \vec{F}_i , aplicadas en lugares geográficos diversos del mismo, así como un cierto punto 0 relevante de aquel (podiera ser su centro de masas, o una capitalidad importante o un nudo de comunicaciones, etc.) y para componer dichas fuerzas las trasladamos a 0 agregando sus correspondientes momentos \vec{M}_i aplicados en 0, con la dirección y sentido que les corresponden. Sumando vectorialmente las fuerzas por un lado y los momentos por otro, obtendremos una sola fuerza económica aplicada en 0, llamada **resultante general**, y un solo vector momento o **momento resultante**; es decir, que al componer un sistema de fuerzas económicas cualesquiera, aplicadas a un territorio, se obtendrá una **resultante general** y un par o **momento resultante**.

Pues bien, utilizando el símil estático, podríamos afirmar que para que un territorio -sobre el que actúan diversas fuerzas económicas- "esté equilibrado", deberá suceder que la "resultante general" y el "momento resultante" de aquéllas sean simultáneamente nulos. Esto es, deberá cumplirse que:

¹ Vide al respecto, el capítulo 10 ("Una interpretación mecanicista de los desplazamientos de las masas económicas") del libro del mismo doctorando titulado: *Análisis Territorial. División, Organización y Gestión del Territorio*, en CADUP (Estudios), Centro Asociado de la UNED. Tortosa, 1990/91. Citado en la bibliografía.

$$\begin{cases} \vec{F} = \sum \vec{F}_i = 0 \\ \vec{M} = \sum \vec{M}_i = 0 \end{cases}$$

, ecuaciones vectoriales que, para un territorio bidimensional, equivaldrán a las 4 ecuaciones escalares siguientes:

$$\begin{cases} F_x = F_y = 0 \\ M_x = M_y = 0 \end{cases}$$

2.2. CLASES DE EQUILIBRIO TERRITORIAL

Comparativamente con la Estática, podemos considerar dos clases de equilibrio en un territorio (FRANQUET, 1990/91):

a) *Estable*, en la que al aplicar una determinada fuerza económica, la posición geográfica del centro de masas no varía substancialmente. Entonces, la energía del centro de masas no varía substancialmente. Entonces, la energía económica potencial del territorio es mínima².

b) *Inestable*, sería un estado de equilibrio territorial en que, al contrario que en el caso anterior, la aplicación de un determinada fuerza económica o social implicara un desplazamiento especial significativo del centro territorial de masas. La cuestión exigible, a continuación, estribaría en determinar la cuantía necesaria de dicho desplazamiento para poderlo calificar de "significativo" y, en su consecuencia, considerar al territorio en estudio como en estado de "equilibrio inestable o desequilibrio".

Desde luego, en este último caso, la energía económica potencial es máxima, y debido a la tendencia "natural" de transformarse la energía potencial económica en energía cinética económica, en el equilibrio inestable cualquier pequeña acción o fuerza económica podría dar lugar a una desestabilización notoria, pues no es necesaria energía económica exterior al sistema para que pueda tener lugar.

Podríamos generalizar la explicación mecanicista anterior de equilibrio de un territorio diciendo que, bajo la acción de cualquier fuerza económico-social (son multitud las fuerzas las que actúan, en cada momento), dicho territorio se hallará en equilibrio estable o inestable según que al variar

² *Vide* al respecto, el capítulo 10 ("Una interpretación mecanicista de los desplazamientos de las masas económicas") del libro del mismo doctorando titulado: *Análisis Territorial. División, Organización y Gestión del Territorio*, en *CADUP (Estudios)*, Centro Asociado de la UNED. Tortosa, 1990/91. Citado en la bibliografía.

ligeramente la posición de su centro de masas, aumente o disminuya su energía potencial económica. Por el contrario, si dicha energía permaneciera constante después de la expresada perturbación, se podría hablar, por último, de un territorio en régimen de equilibrio "indiferente".

3. EQUILIBRIO TERRITORIAL: UNIFORMIDAD

3.1. INTRODUCCIÓN

De la definición de equilibrio territorial que hemos visto en el primer apartado del presente anexo, se infiere su relación biunívoca, concomitancia o biyectividad con los conceptos de uniformidad y homogeneidad en la distribución espacial de las masas socioeconómicas por el territorio (FRANQUET, 1990/91).

En el siguiente epígrafe 3.5. ("Aplicación a Cataluña"), hemos calculado el coeficiente de variación de Pearson (que, como es sabido, tratase de una medida abstracta de dispersión relativa de los valores de la variable aleatoria estadística, profusamente utilizada) para cada una de las nuevas comarcas y nuevas regiones en que proponemos la división del territorio catalán, y para diferentes variables explicativas: población y superficies municipales, densidades de población, altitud topográfica, etc. Es obvio que, desde los respectivos puntos de vista, el territorio en cuestión se hallará tanto más equilibrado cuanto menores sean los valores de su CV referido a la variable aleatoria estadística que toma valores para cada una de las partes en que se considera espacialmente dividido dicho territorio. Destaca, del coeficiente elegido como medida de la variabilidad, su adimensionalidad, es decir, su independencia de las unidades de medida, permitiendo la comparación entre grupos diferentes de datos, lo que no resulta posible establecer mediante el exclusivo empleo de la varianza o de la desviación típica de la correspondiente distribución de frecuencias.

Al respecto, y como medida de la uniformidad en la distribución de las masas socioeconómicas por un territorio cualquiera, pueden utilizarse los diversos coeficientes que propondremos a continuación (expresados en %), de sentido contrario a la variabilidad antedicha.

El primero de ellos podría ser el siguiente:

$$CU_1 = 100(1 - CV), \text{ de gran sencillez y aplicabilidad, siendo: } CV = \sigma/\bar{X}$$

, en que \bar{X} es la media aritmética de los valores de la variable analizada (población, renta, etc.) y σ es su desviación típica o "standard" (desviación cuadrática media).

3.2. ESTUDIO PORMENORIZADO DE LA DISTRIBUCIÓN NORMAL

3.2.1. La distribución teórica de probabilidad

El significado físico del CV se deduce claramente si aceptamos que todos los valores de la variable elegida en el territorio en estudio, se distribuyen de acuerdo con la curva campaniforme de una distribución normal y, por lo tanto, se tendrá lo siguiente:

a) Prácticamente, todos los valores observados se hallarán comprendidos en el entorno: $(1 \pm 3 \cdot CV)\bar{X}$.

b) Aproximadamente, el 95% de las observaciones se encuentran comprendidas en el entorno: $(1 \pm 2 \cdot CV)\bar{X}$.

c) Si se toman las $n/4$ observaciones de valores más bajos del total de los n valores medidos de la variable en cuestión (cuyo valor superior será el primer cuartil Q_1 de la distribución de frecuencias), su media aritmética será igual a: $q_{25} = (1 - 1'27 \cdot CV)\bar{X}$.

d) El 68'27% de las observaciones realizadas estarán en el intervalo: $(1 \pm CV)\bar{X}$.

Por todo ello, y según los valores que adopte dicho coeficiente CV (o el CU_1), podríamos establecer una clasificación de los territorios según su grado o índice de uniformidad en relación a la correspondiente característica o variable territorial digna de evaluación.

Otros coeficiente de uniformidad territorial podrían definirse a partir de las siguientes expresiones:

$$CU_2 = (Q_1 / \bar{X}) \times 100 \text{ (de menor aplicabilidad) y } CU_3 = (q_{25} / \bar{X}) \times 100,$$

siendo q_{25} , como ya se ha visto, el valor medio del cuarto inferior de los valores de la variable analizada.

Desde luego, la ecuación matemática de la función de la distribución normal viene dada, como es sabido, por la expresión:

$$y = 1 / \sigma \sqrt{2\pi} \cdot e^{-1/2(x-\alpha/\sigma)^2}$$

, en la que se ha tomado, como es usual, $\bar{X} = \alpha$, y que coincide con la expresión matemática de la célebre "ley de los errores", debida a Gauss, siendo $y = f(x)$ la denominada "función de densidad normal". Así mismo, se tendrá:

$$F(x_j) = P(x \leq x_j) = P(-\infty < x \leq x_j) = \int_{-\infty}^{x_j} f(x) dx$$

, (función de distribución normal), que es la probabilidad de que la variable aleatoria territorial toma un valor $\leq x_j$.

Las áreas bajo la curva normal representan probabilidades; en estas condiciones, la probabilidad de que $x \in]x_1, x_2]$ será:

$$\begin{aligned} P(x_1 < x \leq x_2) &= F(x_2) - F(x_1) = P(x \leq x_2) - P(x \leq x_1) = \\ &= \int_{-\infty}^{x_2} f(x) \cdot dx - \int_{-\infty}^{x_1} f(x) \cdot dx = \int_{x_1}^{x_2} f(x) \cdot dx. \end{aligned}$$

Cuando x viene expresada en unidades de desviación: $Z = (x-\alpha)/\sigma$, se tiene la **distribución normal tipificada**, así:

$$y = 1 / \sqrt{2\pi} \cdot e^{-z^2/2}$$

, y decimos que la variable Z se distribuye normalmente con media cero ($\alpha=0$) y varianza uno ($\sigma^2 = 1$).

Vamos a proceder, seguidamente, al estudio más pormenorizado de ambas curvas.

1) **Como acabamos de ver, la ecuación de la curva normal tipificada es:**

$$y = 1 / \sqrt{2\pi} \cdot e^{-z^2/2} ; \alpha = 0 ; \sigma^2 = 1 ,$$

que es la función de densidad normal reducida.

- *Extremos relativos y puntos de inflexión:*

Se tiene:

$$\begin{aligned} y' &= 1/\sqrt{2\pi} \cdot e^{-z^2/2} \cdot (-z) = 0 ; z = 0 \text{ (origen)} ; \\ y'' &= 1/\sqrt{2\pi} \cdot e^{-z^2/2} \cdot (z^2 - 1) ; \forall z = 0 \rightarrow y'' = 1/\sqrt{2\pi} < 0 \end{aligned}$$

luego existe un **máxima relativo o local** en el punto $(0, 1 / \sqrt{2\pi})$, de intersección con el eje de ordenadas OY, y no habrá intersección con el eje de abscisas OZ a distancia finita, esto es:

$$\forall y = 0 : z = +\infty \quad \text{y} \quad z = -\infty .$$

De $y'' = 0$, o sea: $z^2 - 1 = 0$, saldrán los **puntos de inflexión**, que son: $z = \pm 1$; o sea, los puntos de coordenadas: $(1, 1 / \sqrt{2\pi \cdot e})$ y $(-1, 1 / \sqrt{2\pi \cdot e})$.

- *Crecimientos y decrecimientos:*

Veamos que:

$$\forall z < 0 \rightarrow y' > 0 \rightarrow \text{CRECIENTE}$$

$$\forall z > 0 \rightarrow y' < 0 \rightarrow \text{DECRECIENTE}$$

- *Asíntotas:*

$$\lim_{z \rightarrow +\infty} 1/\sqrt{2\pi} \cdot e^{-z^2/2} = 1/\sqrt{2\pi} \cdot e^{-\infty} = 1/\sqrt{2\pi} \cdot 1/e^{\infty} = 1/\infty = 0$$

$$\lim_{z \rightarrow -\infty} 1/\sqrt{2\pi} \cdot e^{-z^2/2} = 1/\sqrt{2\pi} \cdot e^{-\infty} = 1/\sqrt{2\pi} \cdot 1/e^{\infty} = 1/\infty = 0$$

Luego tiene por asíntota horizontal el eje OZ.

- *Simetrías:*

- Respecto al eje OY, pues al cambiar (z) por (-z), no sufre variación (se trata de una función par).

- *Concavidades y convexidades:*

De: $y'' = 1/\sqrt{2\pi} \cdot e^{-z^2/2} \cdot (z^2 - 1)$, se deduce que:

$\forall z \in]-\infty, -1[\rightarrow y'' > 0 \rightarrow \text{CÓNCAVA}$ $\forall z \in]-1, -1[\rightarrow y'' < 0 \rightarrow \text{CONVEXA}$ $\forall z \in]1, +\infty[\rightarrow y'' > 0 \rightarrow \text{CÓNCAVA}$
--

Su representación gráfica, en definitiva, será la siguiente:

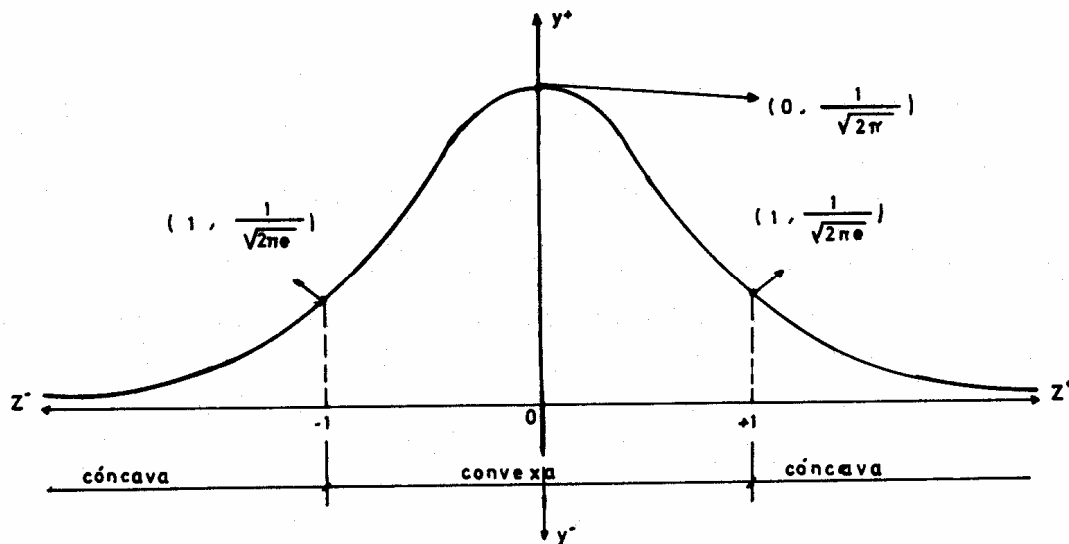


Fig. A-15.1. Curva normal tipificada.

Veamos que en esta distribución teórica de probabilidad, como en todas las simétricas, se cumple, como puede observarse fácilmente, la igualdad entre la media aritmética, la mediana y la moda, valores que, en este caso, coinciden con cero.

2) La ecuación de la curva normal sin tipificar es:

$$y = 1/\sigma \sqrt{2\pi} \cdot e^{-(x-\alpha)^2/2\sigma^2} \text{ (función de densidad normal)}$$

•Asíntotas:

$$\lim_{x \rightarrow +\infty} y = 1/\sigma \sqrt{2\pi} \cdot e^{-(x-\alpha)^2/2\sigma^2} = 1/\sigma \sqrt{2\pi} \cdot e^{-\infty} = 0 ;$$

$$\lim_{x \rightarrow -\infty} y = 1/\sigma \sqrt{2\pi} \cdot e^{-(x-\alpha)^2/2\sigma^2} = 1/\sigma \sqrt{2\pi} \cdot e^{-\infty} = 0$$

luego tiene por asíntota horizontal el eje OX.

•Cortes con los ejes:

Cortará al eje OY en el punto de coordenadas $(0, 1 / \sqrt{2\pi} \cdot e^{-\alpha^2/2\sigma^2})$ y no cortará al eje OX a distancia finita.

•Extremos relativos y puntos de inflexión:

$$\begin{aligned}
 y' &= 1/\sigma\sqrt{2\pi} \cdot e^{-(x-\alpha)^2/2\sigma^2} \cdot [(\alpha-x)/\sigma^2] = y \cdot (\alpha-x)/\sigma^2 = 0 ; \rightarrow x = \alpha \\
 y'' &= y' \cdot (\alpha-x)/\sigma^2 - y \cdot 1/\sigma^2 = y \cdot [(\alpha-x)/\sigma^2]^2 - y \cdot 1/\sigma^2 = \\
 &= y \cdot (\alpha-x)^2/\sigma^4 - y \cdot \sigma^2/\sigma^4 = y \cdot [(\alpha-x)^2 - \sigma^2]/\sigma^4 ; \\
 \forall x = \alpha &\rightarrow y'' = y \cdot (-\sigma^2)/\sigma^4 = -y/\sigma^2 < 0
 \end{aligned}$$

luego existe un máximo relativo o local en el punto $(\alpha, 1/\sigma\sqrt{2\pi})$.

Por otra parte, de $y''=0$, o sea: $(\alpha-x)^2 = \sigma^2$, saldrán los **puntos de inflexión**, que son: $\pm(\alpha-x)^2 = \pm\sigma^2$; o sea: $x = \alpha + \sigma$, o bien: $x = \alpha - \sigma$, esto es, los puntos de las coordenadas cartesianas rectangulares:

$$\left(\alpha + \sigma, 1/\sigma\sqrt{2\pi e}\right) \quad \text{y} \quad \left(\alpha - \sigma, 1/\sigma\sqrt{2\pi e}\right)$$

, siendo convexa la curva entre dichos puntos, y cóncava en el resto del intervalo de existencia, como puede comprobarse del estudio de la segunda derivada y'' . Así pues, es cóncava hacia la región positiva de OY, para: $-\infty < x < \alpha - \sigma$, y para $\alpha + \sigma < x < +\infty$, y cóncava hacia la región negativa del eje de ordenadas (convexa hacia las $y+$) en el intervalo o dominio de definición: $\alpha - \sigma < x < \alpha + \sigma$.

•*Simetrías:*

La curva es simétrica con respecto a la ordenada correspondiente al punto α , por ser una función par con respecto a la diferencia: $(x-\alpha)$.

•*Crecimientos y decrecimientos:*

$$\text{Veamos que } \begin{cases} \forall x < \alpha \rightarrow y' > 0 \rightarrow \text{CRECIENTE} \\ \forall x > \alpha \rightarrow y' < 0 \rightarrow \text{DECRECIENTE} \end{cases}$$

Su representación gráfica será:

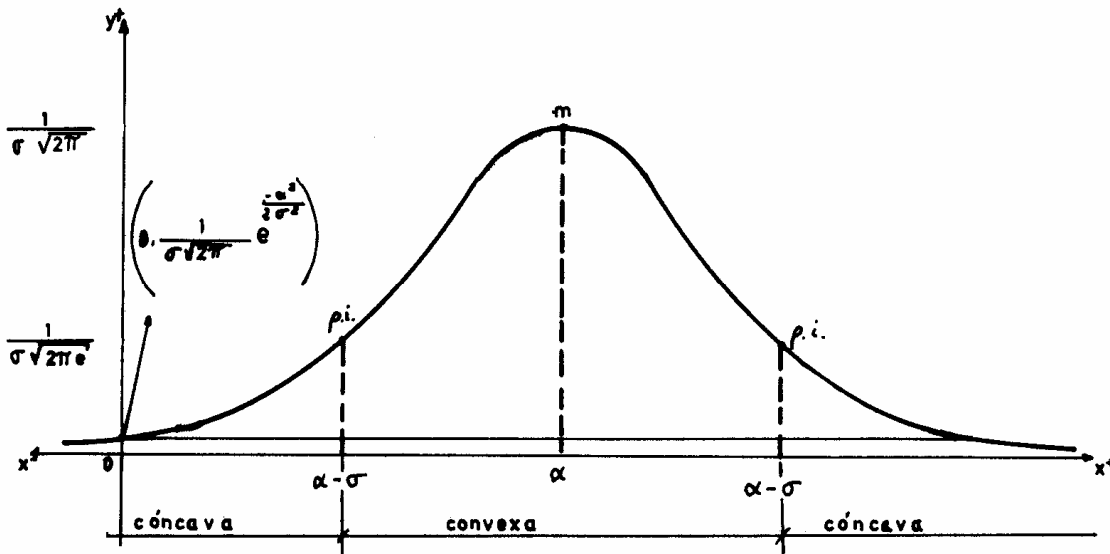


Fig. A-15.2. Curva normal sin tipificar.

3.2.2. Las áreas bajo la curva normal

El estudio detenido que acabamos de realizar, desde el punto de vista del análisis matemático, de las distribuciones normales tipificadas y sin tipificar, nos permitirá aprovechar los conocimientos que la ciencia estadística proporciona acerca de dicha distribución teórica de frecuencias para obtener ciertas conclusiones de tipo cuantitativo, de gran aplicación en el Análisis Territorial.

Y así, se tendrá lo siguiente:

% de casos	INTERVALOS
50'00	$[(1 - 0'68CV) \bar{X}, (1 + 0'68 CV)\bar{X}]$
64'24	$[(1 - 0'92CV) \bar{X}, (1 + 0'92 CV)\bar{X}]$
68'27	$[(1 - CV) \bar{X}, (1 + CV)\bar{X}]$
95'00	$[(1 - 1'96CV) \bar{X}, (1 + 1'96 CV)\bar{X}]$
95'45	$[(1 - 2 CV) \bar{X}, (1 + 2 CV)\bar{X}]$
99'00	$[(1 - 2'58 CV) \bar{X}, (1 + 2'58 CV)\bar{X}]$
99'73	$[(1 - 3 CV) \bar{X}, (1 + 3 CV)\bar{X}]$

, que podría representarse, gráficamente, del siguiente modo:

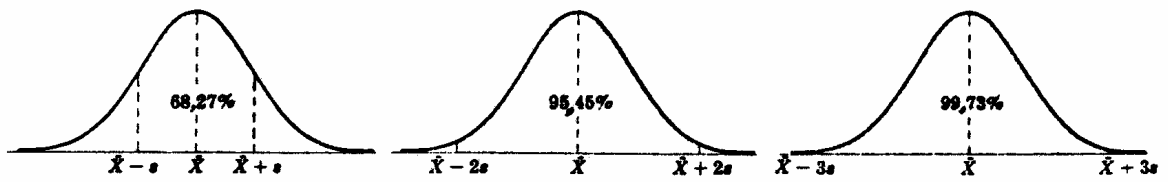


Fig. A-15.3. Áreas bajo la curva normal (I).

Fig. A-15.4. Áreas bajo la curva normal (II).

Fig. A-15.5. Áreas bajo la curva normal (III).

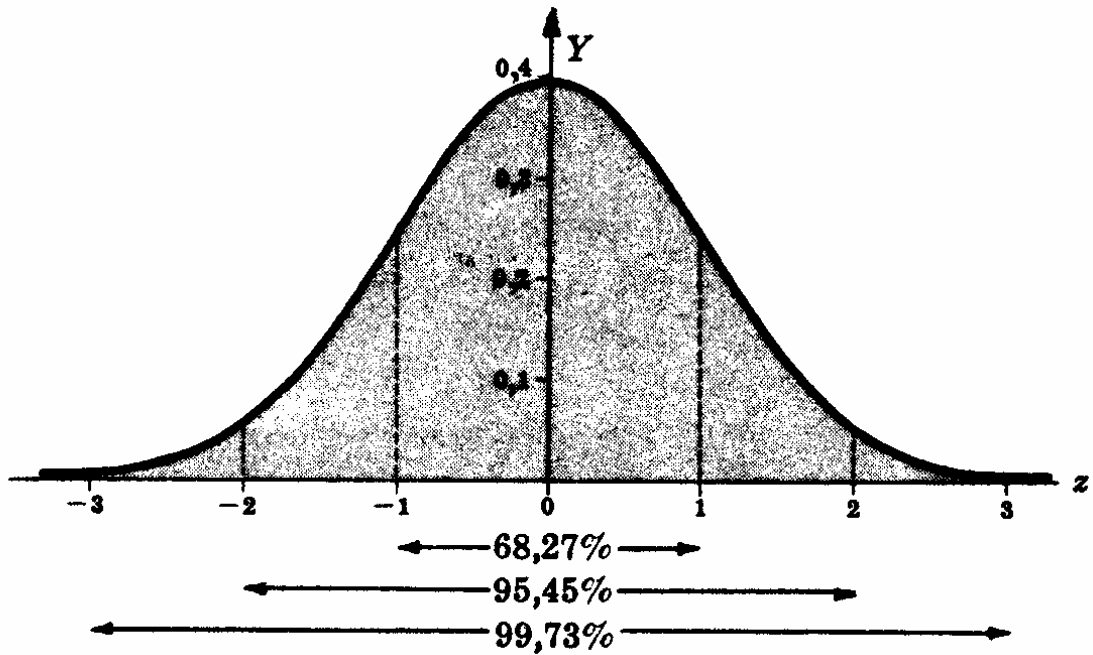


Fig. A-15.6. Áreas bajo la curva normal (IV).

En la siguiente tabla se presentan las áreas: $\int_0^x f(x) dx$ (multiplicadas por 1.000) bajo la curva normal. A saber:

x	1.000 A	x	1.000 A	x	1.000 A	x	1.000 A
0,0	000	1,0	341	2,0	477	3,0	498,6
0,1	040	1,1	364	2,1	482	3,1	499,0
0,2	079	1,2	385	2,2	486	3,2	499,3
0,3	118	1,3	403	2,3	489	3,3	499,5
0,4	156	1,4	419	2,4	492	3,4	499,7
0,5	191	1,5	433	2,5	493,8	3,5	499,77
0,6	226	1,6	445	2,6	495,3	3,6	499,84
0,7	258	1,7	455	2,7	496,5	3,7	499,89
0,8	288	1,8	464	2,8	497,4	3,8	499,93
0,9	316	1,9	471	2,9	498,1	3,9	499,95

Tabla A-15.1. Áreas bajo la curva normal.

De aquí, pueden resolverse las siguientes cuestiones (RODRÍGUEZ-ARENALES, 1988):

a) **Área total bajo la curva normal y probabilidad de que la variable territorial tome un valor cualquiera de su recorrido o campo de variación (de $-\infty$ a $+\infty$).**

La simple observación de la tabla A-15.1 nos dice que el área bajo la curva normal, desde 0 a 3'9, toma el valor:

$$499'95 / 1.000 = 0'49995 \approx 0'5$$

Por la simetría de la curva de Gauss, ésta es la mitad del área total, que vale la unidad (probabilidad total).

b) **Área bajo la curva determinada por las ordenadas en los extremos de los intervalos (1, 2) y (-1, -2). ¿Cuál es el valor de la probabilidad de que la variable territorial x tome un valor comprendido entre 1 y 2? ¿Y entre -2 y -1?**

Según puede verse en la tabla anterior, las áreas bajo la curva comprendidas entre el eje de ordenadas ($x=0$) y las ordenadas $x=2$ y $x=1$, son, respectivamente:

$$477 / 1.000 = 0'477 \quad \text{y} \quad 341 / 1.000 = 0'341 \quad ;$$

entonces, el área pedida será la diferencia:

$$\int_0^2 f(x) dx - \int_0^1 f(x) dx = 0'477 - 0'341 = 0'136 = \int_1^2 f(x) dx = 13'6\%$$

que es también la probabilidad de que la variable territorial x tome un valor comprendido entre 1 y 2.

El área comprendida entre las ordenadas $x = -2$ y $x = -1$ es la misma anterior y la probabilidad de que x tome un valor del intervalo $(-2, -1)$ es también igual, en virtud de la simetría de la figura, a:

$$\int_{-2}^{-1} f(x) dx = 0'136 \quad \text{y} \quad P(-2 < x < -1) = 0'136 \quad .$$

c) **Intervalo $(-a, a)$ cuyas ordenadas extremas delimiten el 50 por 100 del área total normal y su expresión probabilística.**

Hemos de encontrar un valor $x = a$, tal que delimite hasta el eje de ordenadas el 25 por 100 del área total (por simetría, el intervalo $(-a, 0)$ delimitará el otro 25 por 100).

Según la tabla, este valor comprendido entre $x = 0'6$ y $x = 0'7$, y las áreas respectivas, a saber, $0'226$ y $0'258$, incluyen la de valor $0'250$ pedido.

De la proporción:

$$\frac{0'258 - 0'226}{0'7 - 0'6} = \frac{0'250 - 0'226}{a - 0'6}$$

obtendremos $a = 0'68$, con lo que:

$$\int_{-0'68}^{0'68} f(x) dx = 0'50 \quad \text{y} \quad P(-0'68 < x < 0'68) = 0'50 .$$

d) Valor de a tal que las colas (áreas a la izquierda de $-a$ y a la derecha de $+a$) que existen bajo la curva normal sumen el 5 por 100 del área total.

Por razón de simetría, el área de cada cola debe medir el 2'5 por 100 del área total; entonces el valor de a ha de satisfacer la condición:

$$\int_0^a f(x) dx = 0'500 - 0'025 = 0'475$$

Según la tabla, este valor de a está comprendido entre $1'9$ y $2'0$ y se puede estimar según la proporción:

$$\frac{477 - 471}{2 - 1'9} = \frac{475 - 471}{a - 1'9} \quad ; \quad a = 1'9667 \approx 2.$$

En la práctica, se suelen tomar los valores de -2 y 2 para definir la cola del 5 por 100, o lo que es igual:

$$P(-2 < x < 2) = 0'95 .$$

3.3. EFECTO DEL COEFICIENTE DE VARIACIÓN EN LA UNIFORMIDAD DEL TERRITORIO

En relación a la uniformidad territorial a la que nos venimos refiriendo en el presente capítulo, veamos que la propiedad más interesante de la distribución normal de los valores de la variable territorial analizada (población o su densidad, renta, etc.) es que si se toma el 25% de los valores más bajos, su valor medio, es decir, lo que hemos denominado q_{25} , valdrá (según hemos visto en el punto c) del epígrafe 3.2.1. anterior):

$$q_{25} = (1 - 1'27 \cdot CV) \cdot \bar{X} \quad ,$$

con lo que el coeficiente de uniformidad CU_3 anteriormente definido, tomará el valor:

$$CU_3 = 100 (1 - 1'27 \cdot CV) < CU_1$$

Si suponemos, v.gr., un territorio en el que analizando la distribución espacial de sus rentas municipales obtendremos un $CV = 0'32$, veamos que:

$$CU_1 = 100 (1 - 0'32) = 68'00\%$$

$$CU_3 = 100 (1 - 1'27 \times 0'32) = 59'36\% ,$$

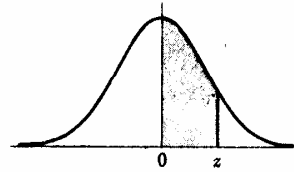
aunque dependería de las circunstancias el escoger uno u otro índice para la medida de la uniformidad territorial que se pretende. De hecho, el CU_3 siempre ofrecerá, expresado en %, por su propia configuración analítica, valores absolutos más bajos que el correspondiente CU_1 , tanto si se trata de valores positivos como negativos (véase, al respecto, el gráfico de la figura A-15.8).

Por otra parte, consideremos, ahora, una región cualquiera estructurada en comarcas que, a su vez, se dividen en municipios para los que conocemos los valores de las diferentes variables territoriales analizadas (población, renta, etc.). Llamando "e" al número de municipios que integran una comarca determinada, resultará que cuanto mayor sea "e" menor será la probabilidad de que todos los municipios de esa comarca pertenezcan al 25% más bajo de la región en estudio. En este caso, la expresión del coeficiente de uniformidad propuesto se transformaría en:

$$CU_3 = (1 - 1'27 \cdot CV / \sqrt{e})$$

Habida cuenta de su interés para la realización de este tipo de cálculos, a continuación se presenta una tabla que ofrece las áreas existentes bajo la curva normal tipificada, limitadas por la ordenada $z = 0$ y cualquier valor positivo de z . A partir de esta misma tabla, se pueden encontrar las áreas comprendidas entre dos ordenadas cualesquiera, utilizando la simetría de la curva de Gauss en relación al eje de ordenadas $z = 0$. Por último, se incluye también una tabla con los valores de las ordenadas (y) de la curva normal tipificada para los diferentes valores de z .

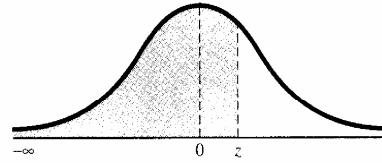
Veámoslas a continuación:



z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

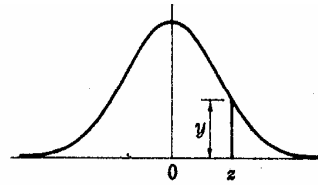
Tabla A-15.2. Áreas bajo la curva normal tipificada de 0 a z .

$$F(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{z^2}{2}} dz$$



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

Tabla A-15.3. Áreas bajo la curva normal tipificada de $-\infty$ a z.



z	0	1	2	3	4	5	6	7	8	9
0.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
0.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
0.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
0.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
0.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
0.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
0.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
0.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
0.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
0.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0022	.0022	.0021	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001

Tabla A-15.4. Ordenadas (y) de la curva normal tipificada en z.

Por otra parte, según acabamos de ver en el epígrafe anterior, se cumplirá que: $Q_1 = (1 - 0.68 CV) \cdot \bar{X}$, que será el intervalo correspondiente al

50% de los casos o "rango intercuartílico" ($Q_3 - Q_1$) de la distribución de probabilidad, con lo que también:

$$CU_2 = 100 (1 - 0'68 CV) ,$$

que, lógicamente, será el mayor de los cuatro coeficientes de uniformidad territorial aquí definidos (ver figura A-15.7).

Así pues, y en base al valor de dichos coeficientes, resulta un \overline{CU} (medio) de : $Z = -0'9375$ (media aritmética), o bien $Z = -0'9117$ (media geométrica), por lo que podríamos considerar, como medida "standard" de la uniformidad de un territorio cualquiera, un $\overline{CU} = 100 (1 - 0'92 CV)$, cuyo intervalo, bajo la hipótesis de normalidad en la distribución espacial de los valores de la variable territorial, abarcaría un 64'24% de los casos, según puede comprobarse mediante las tablas más completas de áreas y ordenadas bajo la función normal, que adjuntamos en las páginas precedentes.

En el ejemplo anteriormente propuesto, se tendrá:

$$\begin{cases} CU_2 = 100 (1 - 0'68 \times 0'32) = 78'24\% \\ \overline{CU} = 100 (1 - 0'92 \times 0'32) = 70'56\% \end{cases}$$

, pudiendo, en la práctica, escoger cualquiera de ellos como medida de la uniformidad territorial que deseamos realizar.

3.4. OTROS COEFICIENTES DE UNIFORMIDAD TERRITORIAL

3.4.1. Basados en la desviación media absoluta

La "desviación media" es la media aritmética de las desviaciones absolutas de los n valores de la variable socioeconómica analizada respecto a un promedio cualquiera (FRANQUET, 1990/91). Si tomamos como dicho promedio la media aritmética $\overline{X} = \alpha$, su expresión será, en el caso de una distribución de frecuencias unitarias:

$$DM = \frac{|x_1 - \overline{X}| + |x_2 - \overline{X}| + \dots + |x_n - \overline{X}|}{n} = \frac{\sum_{i=1}^n |x_i - \overline{X}|}{n}$$

(este valor resultaría mínimo si en vez de considerar la \overline{X} hubiéramos tomado la mediana $M_e = Q_2$ o valor central de la correspondiente distribución de frecuencias).

Por otra parte, en el caso de operar con frecuencias agrupadas o conjuntas, lo que sucederá cuando se opte por agrupar los valores de la variable territorial analizada por intervalos de clase, se tendrá que:

$$n_1 + n_2 + \dots + n_n = \sum_{i=1}^h n_i = n$$

$$DM = \frac{|x_1 - \bar{X}| \cdot n_1 + |x_2 - \bar{X}| \cdot n_2 + \dots + |x_h - \bar{X}| \cdot n_h}{n} =$$

$$= \frac{\sum_{i=1}^h |x_i - \bar{X}| \cdot n_i}{n} = \sum_{i=1}^h |x_i - \bar{X}| \cdot f_i$$

Pues bien, en base a ella, podríamos definir el siguiente nuevo coeficiente de uniformidad:

$$CU_4 = 100 (1 - DM / \bar{X}) ,$$

que, en realidad, resulta similar al CU_1 , habiendo substituido la desviación típica o "standard" por la desviación media absoluta, como medida absoluta de dispersión espacial por el territorio de los valores de la variable aleatoria estadística. Normalmente, para un mismo territorio, se cumplirá que:

$$CU_3 < CU_1 < \bar{CU} < CU_4 < CU_2 < CU_5$$

estando los valores de todos estos coeficientes de uniformidad limitados o acotados superiormente en el 100%, según podrá comprobarse de modo gráfico en la figura A-15.7.

3.4.2. Basados en índices diversos

Veamos, en fin, que también resulta posible la conceptualización y posterior medición o cuantificación de la uniformidad territorial, mediante el empleo del índice de GINI (variable entre 0 y 1) y la correspondiente curva de LORENZ (línea poligonal a un lado de la bisectriz del primer cuadrante). Para ello, habrá que calcular los porcentajes acumulados de los valores de las variable territoriales en estudio que, normalmente, serán:

- a) Superficie (p_i) / Población (q_i)
- b) Superficie (p_i) / Renta (q_i)
- c) Superficie (p_i) / Depósitos bancarios (q_i)

Obviamente, la distribución espacial de las masas socioeconómicas (q_i) por el territorio será tanto más equitativa cuanto más bajo sea el valor de dicho índice, de expresión general (según la fórmula dada por Pulido):

$$G = \frac{\sum_{i=1}^{n-1} (p_i - q_i)}{\sum_{i=1}^{n-1} p_i}$$

Así, por ejemplo, a un 50% de la superficie del territorio deberá corresponder, aproximadamente, un porcentaje similar de la variable territorial, y a la inversa, para que se puede afirmar con propiedad que el territorio hállese "equilibrado".

Teóricamente, la distribución perfecta de la variable territorial analizada tendrá lugar cuando todos sus valores sean iguales, lo que constituye un *desideratum* ideal pero, en cualquier caso, presenta una medida de la uniformidad en su distribución territorial. En este caso, al representar los porcentajes acumulados de la variable frente a los porcentajes acumulados del territorio, se obtendrá la recta de ecuación: $q_i = p_i$, coincidente con la bisectriz del primer cuadrante, y el índice de GINI valdrá 0. Obviamente, este índice se encuentra más próximo a 1 cuanto peor está distribuida, por el territorio, la variable que estamos evaluando (FRANQUET, 1990/91).

En los libros de A. PULIDO SAN ROMÁN³ y de A. ALCAIDE INCHAUSTI⁴, podemos encontrar presentaciones diferentes de la medida que hemos empleado para parametrizar la concentración de la variable. Para interpretar correctamente su significado, resulta suficiente con observar que G varía entre los valores extremos 0 y 1, tomando el valor mínimo o nulo cuando cada p_i es igual a su correspondiente q_i , lo que provoca la anulación del numerador de su expresión definitoria.

Este concepto puede precisarse mejor representando, en un diagrama, la función: $p_i = f(q_i)$, o bien la inversa: $q_i = \varphi(p_i)$, que nos permitirán la obtención de una línea poligonal construida sobre (o por debajo) la diagonal de un cuadrado que tiene un extremo en el centro u origen de coordenadas cartesianas rectangulares (0,0) y el otro extremo en el punto de coordenadas (100,100). Esta figura constituye la mencionada curva de LORENZ, frecuentemente usada en el Análisis estructural económico, y la distribución por el territorio de la variable en estudio q_i será tanto más equitativa cuanto la superficie del área comprendida entre dicha línea poligonal (que tenderá a convertirse en una curva al aumentar el número de puntos en estudio) y la diagonal de equidistribución sea menor, y recíprocamente.

³ *Estadística y Técnicas de Investigación Social*. Ed. ANAYA. Madrid, 1971, pág. 111.

⁴ *Estadística Económica*. Ed. SAETA. Madrid, 1973, pág. 294.

En el mismo orden de ideas, también es recomendable la utilización, a estos mismos efectos, del denominado "índice de Williamson", que nos dará idea acerca del nivel de agrupación de los valores de la variable territorial (q_i) respecto al valor central o promedio de la correspondiente distribución de frecuencias.

Las fórmulas pertinentes, en relación a la superficie territorial, serán, respectivamente, las siguientes (FRANQUET, 1995):

$$\begin{aligned} \text{a) } W_{P,S} &= \frac{\sqrt{\sum_{i=1}^n [(P_i/S_i) - (P/S)^2] \cdot S_i/S}}{P/S} \\ \text{b) } W_{Y,S} &= \frac{\sqrt{\sum_{i=1}^n [(Y_i/S_i) - (Y/S)^2] \cdot S_i/S}}{Y/S} \\ \text{c) } W_{D,S} &= \frac{\sqrt{\sum_{i=1}^n [(D_i/S_i) - (D/S)^2] \cdot S_i/S}}{D/S} \end{aligned}$$

$\forall i \in (1, 2, \dots, n)$, siendo n el número de subunidades territoriales a las que se refieren los valores de la variable territorial en estudio (municipios, comarcas, ...), y también:

Y_i = renta de cada subunidad territorial i .
 D_i = depósitos bancarios de cada subunidad territorial i .
 S_i = superficie de cada subunidad territorial i .
 P_i = población de cada subunidad territorial i .
 Y = renta total del territorio.
 D = depósitos bancarios totales del territorio.
 S = superficie total del territorio.
 P = población total del territorio.

En otros trabajos de este doctorando, puede verse la aplicación de los índices de Williamson aquí definidos, por ejemplo, a la estimación del grado de concentración/dispersión de la propiedad de la tierra en el territorio que pudiéramos denominar "región catalana del Ebro", constituida por las cuatro comarcas de la veguería de les *Terres de l'Ebre* (Baix Ebre, Montsià, Ribera

d'Ebre i Terra Alta), por lo que nos remitiremos simplemente a ellos para el logro de mayores especificaciones y detalles⁵.

A mayor abundamiento, desarrollaremos el cálculo de un nuevo índice denominado "Índice de concentración de Lorenz" desde el mismo diagrama o curva que hemos propuesto anteriormente. Tal como se ha venido considerando, se obtendrán siempre curvas cóncavas hacia las *y* positivas, y que se hallan situadas por debajo de la diagonal del cuadrado que pasa por el origen de coordenadas y por el punto (100,100). El cálculo de ambos índices (Gini y Lorenz) se ha venido realizando profusamente en el anterior Anexo nº: 12 de nuestro trabajo ("Análisis estadístico y clasificación comarcal"), como puede comprobarse.

Así pues, tendremos (FRANQUET, 1995):

$$L = \frac{(a - q_1) + (2a - q_2) + \dots + [(n-1)a - q_{n-1}]}{a + 2a + \dots + (n-1)a} \quad (1),$$

donde *a* es la media aritmética de los porcentajes de la variable correspondientes a cada intervalo de clase, o sea:

$$X_i = \frac{x_i \cdot n_i}{\sum_{i=1}^n x_i \cdot n_i} \times 100 \quad ; \quad a = \frac{\sum_{i=1}^n X_i}{n} = \frac{q_n}{n} \quad ;$$

De esta manera, se cumplirá también que:

$$q_1 = X_1$$

$$q_2 = X_1 + X_2$$

$$q_3 = X_1 + X_2 + X_3$$

.....

$$q_n = X_1 + X_2 + \dots + X_n$$

⁵ Vide FRANQUET BERNIS, José M^o: *Estructura de la Propietat Agrària. Aplicació a la Regió Catalana de l'Ebre*. Tesis doctoral. Departamento de Econometría, Estadística y Economía española. Facultad de Ciencias Económicas y Empresariales. Universidad de Barcelona. Barcelona, 1995. Citada en la bibliografía.

$$\text{o sea: } q_i = \sum_{j=1}^i X_j$$

que es justamente el criterio que hemos seguido para la elaboración de la tabla correspondiente. **Debe tenerse bien presente que, en este caso, la ordenación de los valores de las X_j es preciso realizarla de menor a mayor** (FRANQUET, 1990/91).

Desarrollando la expresión anterior (1), obtendremos:

$$\begin{aligned} L &= \frac{a + 2a + \dots + (n-1)a - (q_1 + q_2 + \dots + q_{n-1})}{a + 2a + \dots + (n-1)a} = \\ &= 1 - \frac{q_1 + q_2 + \dots + q_{n-1}}{a + 2a + \dots + (n-1)a} = 1 - \frac{\sum_{i=1}^{n-1} q_i}{\frac{n(n-1)}{2} a} = \\ &= 1 - \frac{2}{n-1} \times \frac{\sum_{i=1}^{n-1} q_i}{q_n} \quad , \end{aligned}$$

$$\text{ya que: } 1 + 2 + \dots + (n-1) = n \cdot (n-1) / 2 \quad ,$$

dado que se trata de la adición de los (n-1) primeros términos consecutivos de una progresión aritmética de razón igual a la unidad (demostrable por inducción completa), y además: $n \cdot a = q_n$, por la propia definición que hemos considerado de la media aritmética **a**.

Veamos, entonces, los valores que adopta este nuevo índice en los casos extremos posibles. Efectivamente, **si la concentración de la variable territorial es máxima**, tendremos que:

$$X_1 = X_2 = \dots = X_{n-1} = 0, \quad \text{y también: } q_n = \sum_{i=1}^n X_i$$

$$L = 1 - \frac{2}{n-1} \times \frac{\sum_{i=1}^{n-1} q_i}{q_n} = 1 - \frac{2}{n-1} \times \frac{0}{q_n} = 1 \quad ,$$

$$\text{dado que: } \sum_{i=1}^{n-1} q_i = 0 \quad .$$

Sin embargo, **si la concentración de la variable territorial es mínima**, o sea, la distribución de la misma variable territorial es teóricamente perfecta desde el punto de vista estadístico, se tendrá lo siguiente:

$$X_1 = X_2 = \dots = X_n = a,$$

$$\sum_{i=1}^{n-1} q_i = \frac{n(n-1)}{2} a$$

en cuyo caso, el índice de concentración de Lorenz será:

$$L = 1 - \frac{2}{n-1} \times \frac{\sum_{i=1}^{n-1} q_i}{q_n} = 1 - \frac{2}{n-1} \times \frac{n \cdot (n-1) \cdot a}{2 \cdot n \cdot a} = 1 - 1 = 0$$

De hecho, estos valores extremos del índice analizado se corresponden con similares valores del índice de Gini anteriormente estudiado. Podemos ver que, cuando: $L = 0$ ($X_1 = X_2 = \dots = X_n = a$), sucede justamente que: $q_n = n \cdot a$, razón por la cual la curva pertinente es el segmento recto coincidente con la diagonal del cuadrado al que nos hemos referido con anterioridad. En el caso de la concentración máxima, resulta: $L = 1$ ($X_1 = X_2 = \dots = X_{n-1} = 0$), y la curva poligonal de Lorenz, que constituye un triángulo rectángulo, viene dada por los dos lados normales o perpendiculares del cuadrado construido al objeto de trazar el diagrama en cuestión. Obviamente, cuanto más se aproxime la curva a la diagonal relacionada, más perfecta será -al menos desde el punto de vista estadístico- la distribución de la variable territorial en estudio. Incluso podemos dar una interpretación geométrica del índice de Lorenz de esta manera: el numerador de la fórmula (1) se puede considerar como la adición de las áreas de $(n-1)$ rectángulos de base igual a la unidad y altura: $(h \cdot a - q_h)$, $\forall h \in [1, 2, \dots, (n-1)]$. El denominador, en este caso, es la suma de las áreas de $(n-1)$ rectángulos de base unidad y altura: $h \cdot a$, $\forall h \in [1, 2, \dots, (n-1)]$. Si observamos lo que representa la suma de estos rectángulos, deduciremos que el numerador de la expresión (1) es el área comprendida entre la curva poligonal de Lorenz y la diagonal del cuadrado, mientras que el denominador es precisamente el área de la mitad de dicho cuadrado⁶.

Este índice es equivalente al anteriormente estudiado de Gini y obliga a la realización del cálculo de la superficie rayada de la figura, comprendida entre la diagonal y la correspondiente curva o poligonal de Lorenz. Un valor aproximado es el que se obtiene mediante la aplicación de la fórmula basada en los porcentajes acumulados, que resulta muy empleada en los trabajos prácticos (FRANQUET, 1995).

⁶Así pues, también el índice de concentración de Lorenz será tanto más pequeño cuanto menor sea el valor del área limitada por la diagonal del primer cuadrante y la misma curva poligonal.

3.5. APLICACIÓN A CATALUÑA

En el Capítulo 7 de nuestro estudio, se procede a la regionalización del territorio de Cataluña, esto es, a su división territorial en 7 regiones o veguerías, utilizando como elementos componentes de las mismas las comarcas que pudiéramos denominar "clásicas" (ya definidas hace unos 70 años por la Generalitat republicana) o bien las "nuevas" (resultantes de la aplicación de los modelos objetivos de división territorial aquí propugnados a la realidad socioeconómica del país imperante en los últimos años), que, a su vez, han sido establecidas en el Capítulo 6.

Los resultados de dicha división territorial, que exponemos en este estudio, pueden ser analizados desde el punto de vista del equilibrio territorial de las regiones resultantes, mediante el cálculo y posterior análisis de los correspondientes coeficientes de uniformidad anteriormente establecidos. Habida cuenta de la definición que hemos hecho de los mismos, es posible la obtención de CU (coeficientes de uniformidad territorial) negativos cuando los coeficientes de variación -o, en definitiva, el grado de dispersión de las masas socioeconómicas por el territorio- sean muy elevados, lo que sucederá en territorios altamente desequilibrados.

Utilizaremos, como variable territorial cuya distribución espacial se cuantifica, la "densidad de población" (expresada en habitantes/km², en el censo del año 1986, justamente un año antes de la aparición de las cuatro leyes de organización territorial a las que nos hemos referido con anterioridad), puesto que subsume, a la vez, los conceptos de población y de superficie; en caso de ser conocida, por las mismas razones, resultaría también interesante emplear la "densidad de renta" (expresada en €/km²).

Es obvio, por otra parte, que amén de las restantes consideraciones que ya han sido expuestas en nuestro Análisis Territorial, el "diseño" de nuevas unidades territoriales deberá tener bien presente el concepto aquí estudiado del equilibrio territorial. Y así, por ejemplo, deberá procurarse que los nuevos territorios resultantes -bien sea de la partición o de la división geográfica de una unidad territorial mayor, o bien de la reunión o agrupación de otras- tengan coeficientes de uniformidad positivos (refiriéndonos, generalmente, al CU₁) y, obviamente, que éstos sean los mayores posibles.

En el caso del ejemplo catalán al que nos venimos refiriendo, sucederá lo siguiente:

Capitales de Regiones o Veguerías		COEFICIENTES DE UNIFORMIDAD					
		Regionalización con comarcas clásicas			Regionalización con comarcas nuevas		
		CU1	CU3	CU4	CU1	CU3	CU4
01	Barcelona	- 94%	- 146%	- 35%	- 63%	- 107%	- 13%
02	Girona	53%	40%	64%	53%	40%	55%
03	Tarragona	± 0%	- 27%	22%	± 0%	- 27%	23%
04	Lleida	18%	- 4%	34%	36%	19%	46%
05	Tortosa	48%	34%	50%	37%	20%	48%
06	Manresa	26%	6%	31%	30%	11%	40%
07	Seu d'Urgell	46%	31%	46%	36%	19%	47%
TOTAL	CATALUÑA	- 82%	- 131%	- 27%	- 82%	- 131%	- 27%

Tabla A-15.5. Coeficientes de uniformidad de la distribución espacial de la variable territorial "densidad de población" de Cataluña, en el año 1986.

Resulta curioso comprobar cómo la regionalización efectuada con las comarcas "nuevas" aporta el mismo grado de equilibrio territorial que el efectuado con las comarcas clásicas. De las 7 regiones que componen el conjunto catalán, 3 de ellas aumentan su índice de uniformidad o equilibrio (Barcelona, Ponent y Catalunya central), 2 de ellas permanecen igual (Girona y Camp de Tarragona) y las dos restantes pierden uniformidad (Terres de l'Ebre y Alt Pirineu). Ello podría conducir a determinaciones correctoras de dichas regionalizaciones; y así, en el caso de estas dos últimas regiones o veguerías, podría pensarse en adoptar la regionalización clásica, puesto que las dota de un mayor equilibrio territorial, mientras que podría mantenerse la regionalización con las comarcas "nuevas" en todos los demás casos.

Es digno de resaltar, así mismo, cómo el fenómeno distorsionante de la macrocefalia interviene, en el caso de Cataluña, desequilibrando el conjunto del territorio hasta niveles decididamente preocupantes ($CU_1 = -82\%$).

4. OTRAS CARACTERÍSTICAS INTERESANTES DE LA DISTRIBUCIÓN ESPACIAL DE LAS VARIABLES TERRITORIALES

4.1. ECUACIONES DE LIGADURA ENTRE LOS COEFICIENTES DE UNIFORMIDAD

Evidentemente, existen en la metodología estadística otras medidas del grado de concentración y/o dispersión de las variables territoriales que pueden emplearse eficazmente en la medida de la uniformidad o del equilibrio territorial

(el recorrido "semi-intercuartílico", el "coeficiente de apertura", el "recorrido relativo", etc.), debiéndose tener en cuenta que, para distribuciones moderadamente asimétricas, se pueden aplicar, con buena aproximación, las fórmulas empíricas siguientes (donde Q_1 i Q_3 son, respectivamente, el primer y tercer cuartil de la correspondiente distribución de frecuencias):

$$DM \approx (4/5) \cdot \sigma \quad ; \quad (Q_3 - Q_1)/2 \approx (2/3) \cdot \sigma$$

, que no son más que consecuencias directas del hecho de que, para distribuciones normales, se tiene que la desviación media absoluta DM y el "rango semi-intercuartílico" son, respectivamente, iguales a 0,7979 y 0,6745 veces la desviación típica o "standard" σ .

Desde esta perspectiva, y para distribuciones territoriales aproximadamente normales con suficiente número de valores de la variable territorial en estudio ($n \geq 30$), los coeficientes de uniformidad anteriormente definidos pueden representarse, geoméricamente, por rectas o funciones lineales cuya variable independiente o explicativa sea el coeficiente de variación de Pearson CV (FRANQUET, 1990/91).

En concreto, se tendrá que:

$$CU_4 = 100 (1 - 0,7979 \cdot \sigma/\bar{X}) \approx 100 \cdot (1 - 0,80 \cdot CV)$$

La representación gráfica será la siguiente:

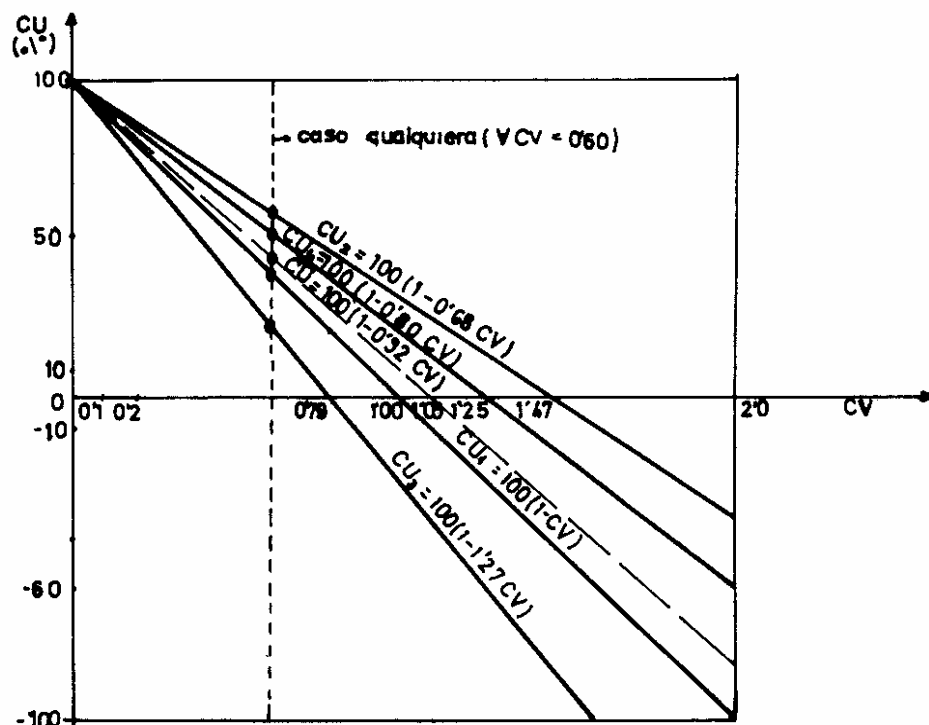


Fig. A-15.7. Coeficientes de uniformidad en función del coeficiente de variación de Pearson.

A su vez, las relaciones que ligan entre sí los diferentes coeficientes de uniformidad territorial aquí definidos, pueden deducirse del siguiente modo:

$$\begin{aligned} CU_1 &= 100 (1 - CV) = 100 - 100 \cdot CV \\ CU_2 &= 100 (1 - 0,68 \cdot CV) = 100 - 68 \cdot CV \\ CU_3 &= 100 (1 - 1,27 \cdot CV) = 100 - 127 \cdot CV \\ CU_4 &= 100 (1 - 0,80 \cdot CV) = 100 - 80 \cdot CV \\ \overline{CU} &= 100(1 - 0,92 \cdot CV) = 100 - 92 \cdot CV \end{aligned}$$

De donde se obtiene:

$$CU_1 - CU_3 = 100 - 100CV - 100 + 127 CV = 27 CV$$

$$CU_3 - CU_4 = 100 - 127 CV - 100 + 80 CV = -47 CV$$

$$CU_1 - CU_4 = \dots\dots(27 CV - 47 CV) \dots\dots = -20 CV$$

Se tendría que:

$$CU_1 / CU_3 = (1 - CV) / (1 - 1,27 \cdot CV) \quad ;$$

$$CU_1 - 1,27 \cdot CV \cdot CU_1 = CU_3 - CV \cdot CU_3 \quad ;$$

$$CU_1 - CU_3 = 27 CV = 1,27 CV \cdot CU_1 - CV \cdot CU_3 \quad ;$$

$$27 = 1,27 \cdot CU_1 - CU_3 \quad ; \quad CU_3 + 27 = 1,27 \cdot CU_1 \quad ; y:$$

$$\mathbf{CU_1 = (CU_3 + 27) / 1,27}$$

Así mismo:

$$CU_1 / CU_4 = (1 - CV) / (1 - 0,8 \cdot CV) \quad ;$$

$$CU_1 - 0,8 \cdot CV \cdot CU_1 = CU_4 - CV \cdot CU_4 \quad ;$$

$$CU_1 - CU_4 = -20 \cdot CV = 0,8 \cdot CV \cdot CU_1 - CV \cdot CU_4 \quad ;$$

$$-20 = 0,8 \cdot CU_1 - CU_4 \quad ; \quad CU_4 - 20 = 0,8 \cdot CU_1 \quad ; y:$$

$$\mathbf{CU_1 = (CU_4 - 20) / 0,8}$$

Si observamos la representación gráfica adjunta A-15.8, la convergencia de ambas rectas tendrá lugar para los valores:

$$(CU_3 + 27) / 1,27 = (CU_4 - 20) / 0,8 \quad y \quad CU_3 = CU_4$$

, lo que implica que, en dicho punto, tendrá lugar la máxima uniformidad territorial posible, con:

$$\mathbf{CU_1 = CU_3 = CU_4 = 100\% = CU_2 = \overline{CU}}$$

También se cumplirá que:

$$CU_3 / CU_4 = (1 - 1,27 \cdot CV) / (1 - 0,8 \cdot CV) \quad ;$$

$$CU_3 - 0,8 \cdot CV \cdot CU_3 = CU_4 - 1,27 \cdot CV \cdot CU_4 \quad ;$$

$$CU_3 - CU_4 = - 47 \cdot CV = 0,8 \cdot CV \cdot CU_3 - 1,27 \cdot CV \cdot CU_4 \quad ;$$

$$1,27 \cdot CU_4 - 47 = 0,8 \cdot CU_3 \quad ; y$$

$$\mathbf{CU_3 = (1,27 \cdot CU_4 - 47) / 0,8}$$

Las relaciones existentes entre los diferentes coeficientes de uniformidad territorial anteriormente definidos, y que ya han sido expresadas analíticamente, pueden verse gráficamente a continuación:

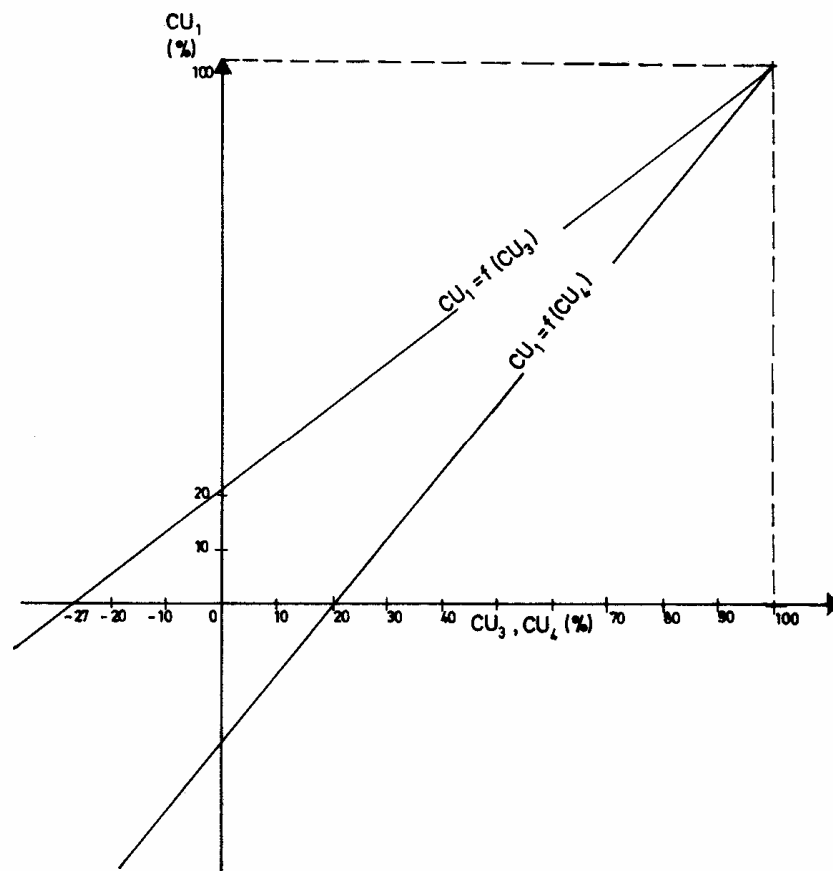


FIG. A-15.8. Relaciones entre los diferentes coeficientes de uniformidad, para distribuciones territoriales aproximadamente normales (I).

El gráfico anterior puede complementarse, por su elevado interés práctico, con el siguiente, que relaciona el coeficiente de uniformidad CU_3 con el CU_4 del siguiente modo:

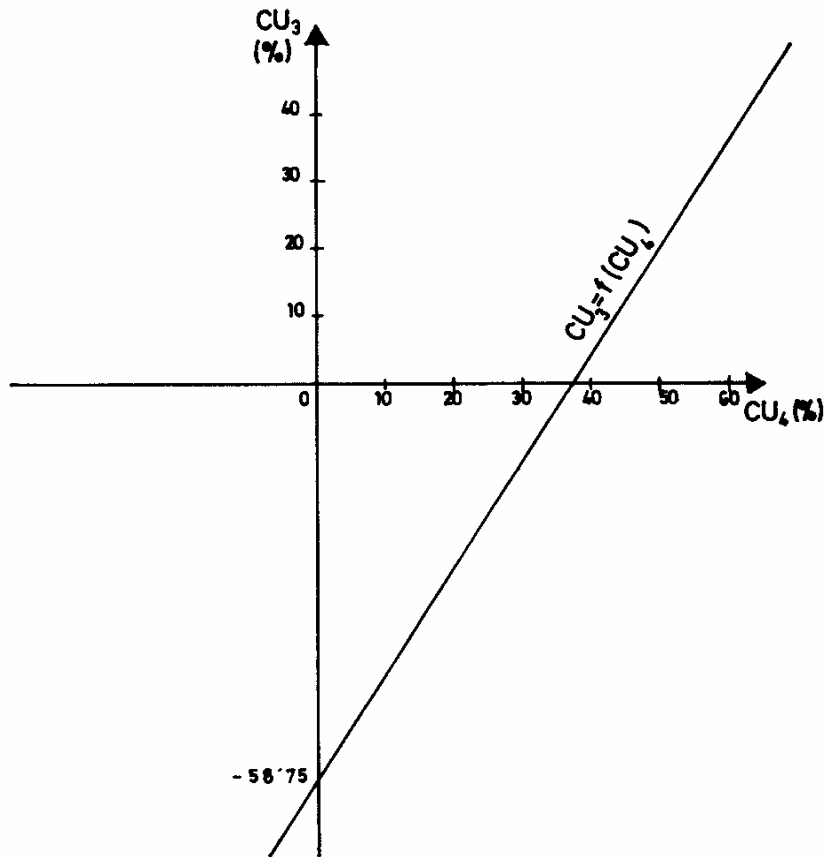


FIG. A-15.9. Relaciones entre los diferentes coeficientes de uniformidad, para distribuciones territoriales aproximadamente normales (II).

De las expresiones anteriores, se deducen las tres siguientes, hasta completar las seis relaciones posibles entre los índices territoriales de tal suerte definidos:

$$CU_3 = 1,27 \cdot CU_1 - 27 \quad ; \quad CU_4 = 0,8 \cdot CU_1 + 20 \quad ;$$

$$CU_4 = (0,8 \cdot CU_3 + 47) / 1,27$$

Idénticas consideraciones podríamos realizar respecto a CU_2 y a \overline{CU} en relación con los tres restantes coeficientes de uniformidad territorial anteriormente definidos.

4.2. AGRUPAMIENTO EN "CLASES" Y OTRAS CARACTERÍSTICAS DE LAS DISTRIBUCIONES TERRITORIALES

4.2.1. Los intervalos de clase

Así mismo, cuando el número n de los valores de la variable territorial analizada sea grande, lo que tendrá lugar en aquellos territorios compuestos por un número elevado de subunidades territoriales menores, resultarán poco manejables las tablas estadísticas que recojan todos los valores con sus correspondientes frecuencias (FRANQUET, 1990/91). En tales casos, se agruparán los valores de la variable en "clases", que podrán ser de la misma o diferente amplitud; una norma práctica genérica pudiera ser el establecer una misma amplitud equivalente al 10% de la observación mayor, con lo que el número de clases oscilará alrededor de la decena. Cuando esto acontezca, el cálculo de la desviación típica necesaria para el hallazgo de los CV y de los pertinentes coeficientes de uniformidad registrará algo de error, debido, precisamente, al "error de agrupamiento" en clases. Para ajustarnos mejor a la realidad, se utilizará entonces la varianza corregida, ofrecida por la denominada "corrección Sheppard", a saber (SPIEGEL, 1981):

$$\sigma_c^2 = \sigma^2 - c^2/12$$

siendo c la amplitud del intervalo de clase escogido y σ^2 la varianza de los datos agrupados, y ello tendrá lugar en distribuciones continuas donde las "colas" van gradualmente convergiendo a 0 en ambas direcciones.

En líneas generales, veamos que un número excesivo de "clases" reduce las ventajas de la agrupación, pero un número escaso puede anular la significación de los datos. Respecto a la amplitud de las "clases" establecidas, conviene observar que, en general, es conveniente sea la misma para todas; sin embargo, esto dependerá mucho de los propios datos y del objetivo final de la distribución territorial de la variable en estudio. En principio, si la distribución es más o menos uniforme, todas las "clases" serán de igual amplitud, y si, por el contrario, presenta grandes oscilaciones, puede ser interesante considerar intervalos de amplitud diferente.

De hecho, la construcción de una distribución numérica -como la mayoría de las que elaboraremos aquí- consta de tres etapas fundamentales: 1) determinar las "clases" con sus intervalos más procedentes, tal como ya hemos expresado antes, en las que se han de agrupar los datos de la variable territorial en estudio, 2) clasificar (o distribuir) los datos en las clases apropiadas, y 3) contar el número de casos de cada clase. Como sea que las dos últimas etapas son puramente mecánicas, así como el establecimiento de la correspondiente "marca de clase" (obtenida normalmente, a falta de más datos, como la semisuma de los valores extremos del intervalo de clase), nos

fijaremos sólo en la primera. Por esto, hace falta determinar el número de clases así como la amplitud del intervalo de los valores de la variable aleatoria estadística con la que trabajamos (población, superficie, depósitos bancarios, ...). Con ello, en términos generales, se pueden observar al efecto las siguientes normas:

- a) Pocas veces emplearemos menos de 6 ó más de 15 clases; el número exacto de las mismas dependerá de la naturaleza, cuantía e intervalo que cubren los datos.
- b) Siempre escogeremos las clases de manera que todos los datos queden comprendidos.
- c) Se procurará, siempre que sea posible, que todos los intervalos de clase tengan la misma amplitud, lo que obviará la determinación de las “densidades de frecuencia” -que determinan la altura de los rectángulos yuxtapuestos del histograma- para el cálculo de algunas medidas centrales de la correspondiente distribución de frecuencias (como la “moda”) o la representación gráfica de los histogramas.

Veamos, por último, que mediante el razonamiento que sirve para definir la “desviación típica o standard” como una medida de dispersión absoluta de los valores de la variable territorial, se puede afirmar que si este estadístico resulta pequeño, los valores se encuentran concentrados en el entorno de la media aritmética y, además, si la desviación típica es grande, los valores están mucho más esparcidos o dispersos en relación a los centrales. Para comprender este razonamiento sobre una base algo menos intuitiva, nos referiremos brevemente al conocido *Teorema de Tchebyshev*, que expresa que **para cualquier clase de datos (poblaciones o muestras), al menos el 75% de los datos se encuentran sobre el intervalo que se extiende a cada lado de la media aritmética en dos veces el valor de la desviación típica ($\pm 2\sigma$)**. Según este teorema, también se puede afirmar que por lo menos el 88'8% de los datos se encuentran dentro del intervalo de tres veces ($\pm 3\sigma$) la desviación típica (a ambos lados de la media aritmética) y que al menos el 96% de los mismos se hallan comprendidos dentro del intervalo de amplitud de cinco veces la desviación típica ($\pm 5\sigma$).

Precisamente, una característica importante del expresado teorema de Tchèbyshev es que resulta válido para cualquier tipo de datos. No obstante, si se dispone de alguna información adicional en relación a la forma global de la distribución que estamos trabajando, también se pueden realizar afirmaciones mucho más estrictas. Por ejemplo, si una distribución es *campaniforme* o gaussiana, se puede esperar que aproximadamente el 95% de los datos (en lugar de al menos el 75%) se encuentren dentro del intervalo $\pm 2\sigma$ y el 99% de

los datos (en lugar de al menos el 88'8%) se encuentran dentro del intervalo $\pm 3\sigma$. Estos porcentajes, en definitiva, corresponden a la llamada *distribución normal*, que es objeto de estudio en diversas partes de nuestro trabajo.

4.2.2. Forma de la distribución de frecuencias

También existen otras características de menor interés práctico, que tratan de precisar la *forma de la distribución* de la variable que se estudia en relación con una distribución normal. Así, la curva de Gauss sabemos que es simétrica respecto de la ordenada $x = \alpha$ (el parámetro α , como ya se ha visto anteriormente, es la media aritmética o esperanza matemática de una distribución normal de frecuencias) y la distribución observada puede ser *asimétrica* respecto a la ordenada correspondiente e, incluso, dicha asimetría puede representar una mayor área bajo la curva, a la derecha o a la izquierda de dicha ordenada; por otra parte, a la distribución observada puede corresponderle un área bajo la curva más achatada (platicúrtica) o más alargada (leptocúrtica o apuntada) que la correspondiente área de una distribución normal (mesocúrtica); a tal característica la denominaremos *kurtosis* o *medida de apuntamiento*.

Todas estas características, bien conocidas por otra parte, de una distribución de frecuencias originan medidas exactas si están utilizando todos los posibles valores de la variable territorial, es decir, si corresponden a la población o universo de datos; pero un objetivo esencial de la Inferencia Estadística es el de estimar dichas características poblacionales a partir de una muestra o subconjunto poblacional, lo que en el Análisis Territorial sucederá sólo en casos muy concretos de territorios para los que se manejen grandes cantidades de datos de la variable territorial. En estos casos, las características de la *distribución muestral* suelen ser, en general, los mejores *estimadores* de las características de la distribución de la población, pero las *estimaciones* que originan han de presentarse, como sabemos, acompañadas de los errores de muestreo y de otras medidas de naturaleza probabilística, que permiten apreciar el grado de confianza o fiabilidad de la correspondiente estimación de la característica poblacional, y que no procede analizar aquí con mayor profundidad por razones obvias (FRANQUET, 1990/91).

4.2.3. Otros coeficientes de uniformidad territorial

Veamos, por último, que en base a los mismos o parecidos conceptos, sería posible la definición de otros coeficientes de uniformidad territorial. Y así, valga como ejemplo el que tendría en cuenta el valor del primer y tercer cuartil de la distribución de frecuencias de la variable territorial contemplada, a saber:

$$CU_5 = 100 \times \sqrt{\frac{Q_1}{Q_3}}$$

que ofrecerá, en el caso de una distribución moderadamente asimétrica (aproximadamente normal), un valor en función de Q3 y de σ equivalente a:

$$Q_3 - Q_1 \approx 4\sigma/3 ; \text{ esto es:}$$

$$(Q_3 - Q_1)/Q_3 \approx 4\sigma/3Q_3 \approx 1 - (Q_1/Q_3); \text{ de donde:}$$

$$\sqrt{\frac{Q_1}{Q_3}} = \sqrt{1 - \left(\frac{4\sigma}{3Q_3}\right)}, \text{ con lo que:}$$

$$CU_5 = 100 \cdot \sqrt{\frac{Q_1}{Q_3}} = \sqrt{10.000} \cdot \sqrt{1 - \frac{4\sigma}{3Q_3}} = \sqrt{10.000 - \frac{40.000 \cdot \sigma}{3Q_3}}$$

5. EJEMPLO PRÁCTICO

5.1. DATOS Y ENFOQUE DEL PROBLEMA

Sea una nación geográficamente determinada, constituida por 1.000 pequeñas subunidades territoriales o municipios, de los que conocemos la correspondiente densidad de renta según datos del año 2000, expresada en 10⁶ ptas./km². Los valores de esta variable territorial se hallan agrupadas en ocho "clases" de la misma amplitud, concretamente: 5 x 10⁶ ptas./km², constituyéndose la primera y última clases de la tabla en "intervalos abiertos", ya que no se han fijado el extremo inferior de la primera y el extremo inferior de la última, que corresponderían, respectivamente, a las densidades de renta más baja y más alta del territorio en cuestión.

La tabla en cuestión es la siguiente:

DENSIDADES MUNICIPALES DE RENTA						
L _i 10 ⁶ ptas./km ²	f. simple n _j	f. acumul. ascen. N _j ↑	f. acumul. descen. N _j ↓	f. simple f _i	f. acumul. ascen. F _j ↑	f. acumul. descen. F _j ↓
Menos de 150	3	3	997	0'003	0'003	0'997
De 150 a 155	15	18	982	0'015	0'018	0'982
De 155 a 160	50	68	932	0'050	0'068	0'932
De 160 a 165	240	308	692	0'240	0'308	0'692
De 165 a 170	312	620	380	0'312	0'620	0'380
De 170 a 175	235	855	145	0'235	0'855	0'145
De 175 a 180	108	963	37	0'108	0'963	0'037
De 180 y más	37	1.000	0	0'037	1'000	0'000
TOTAL	1.000	///	///	1'000	///	///

Tabla A-15.6. Tabla de frecuencias municipales de la variable "densidad de renta".

Para interpretar correctamente dicha tabla, deben formularse las siguientes observaciones:

a) La columna de las frecuencias simples, encabezada por n_i , $\forall i \in (1, 2, \dots, 8)$, recoge la *frecuencia absoluta* de cada uno de los ocho *intervalos de clase*; así, 240 es el número de municipios cuya densidad de renta está comprendida entre 160 (inclusive) y 165×10^6 ptas./km². La *frecuencia total* toma el valor $n = 1.000$.

b) Las *frecuencias acumuladas ascendentes* las hemos designado por $N_i \uparrow$ ($i=1, 2, \dots, 8$) y a cada una de ellas corresponde el número de municipios con densidad de renta inferior al extremo superior de la clase; así, $N_4 \uparrow = 308$, significa que se han hallado 308 municipios con una densidad de renta inferior a 165×10^6 ptas./km².

c) Las *frecuencias relativas ordinarias y acumuladas ascendentes* encabezadas, respectivamente, por f_i , $F_i \uparrow$ y $F_i \downarrow$ vienen determinadas por el cociente de dividir la correspondiente frecuencia absoluta o acumulada (ascendente o descendente) por la frecuencia total. Así:

$$f_4 = 240 / 1.000 = 0'240; \quad F_4 \uparrow = 308 / 1.000 = 0'308; \quad F_4 \downarrow = 692 / 1.000 = 0'692$$

La suma de las frecuencias relativas ordinarias y la última frecuencia relativa acumulada ascendente han de ser siempre iguales a la unidad, puesto que representan la probabilidad total.

d) La representación gráfica de una tabla de frecuencias ordinarias (absolutas o relativas), cuando la variable es continua (como es el caso de la mayoría de las variables territoriales, que pueden tomar valores entre dos consecutivos) está distribuida en clases -como las de la tabla número A-15.6-, debe realizarse mediante un *histograma de rectángulos yuxtapuestos* (fig. A-15.10), que se construye tomando como abscisas los extremos de los intervalos y levantando sobre cada intervalo, tomado como base, un rectángulo cuya área sea directamente proporcional a la correspondiente frecuencia (absoluta o relativa). Si todos los intervalos de clase son de la misma amplitud (caso de la tabla núm. A-15.6), en este caso 5×10^6 ptas./km², la altura de cada rectángulo es proporcional a la correspondiente frecuencia. En nuestra figura siguiente se ha supuesto que el extremo inferior de la variable es 145 y el superior 185×10^6 ptas./km², aunque, de hecho, los extremos inferior y superior, respectivamente, del campo de existencia de la misma, son de 120 y 208×10^6 ptas./km², correspondientes a los municipios de densidades municipales de renta mínima y máxima.

Veamos que a las alturas de cada rectángulo del histograma se le denomina “densidad de frecuencia” del intervalo considerado. En este caso, al ser todos los intervalos de clase de la misma amplitud, las alturas de los rectángulos yuxtapuestos coinciden con las correspondientes frecuencias simples sin acumular n_i .

Se tendría, pues, la siguiente representación gráfica:

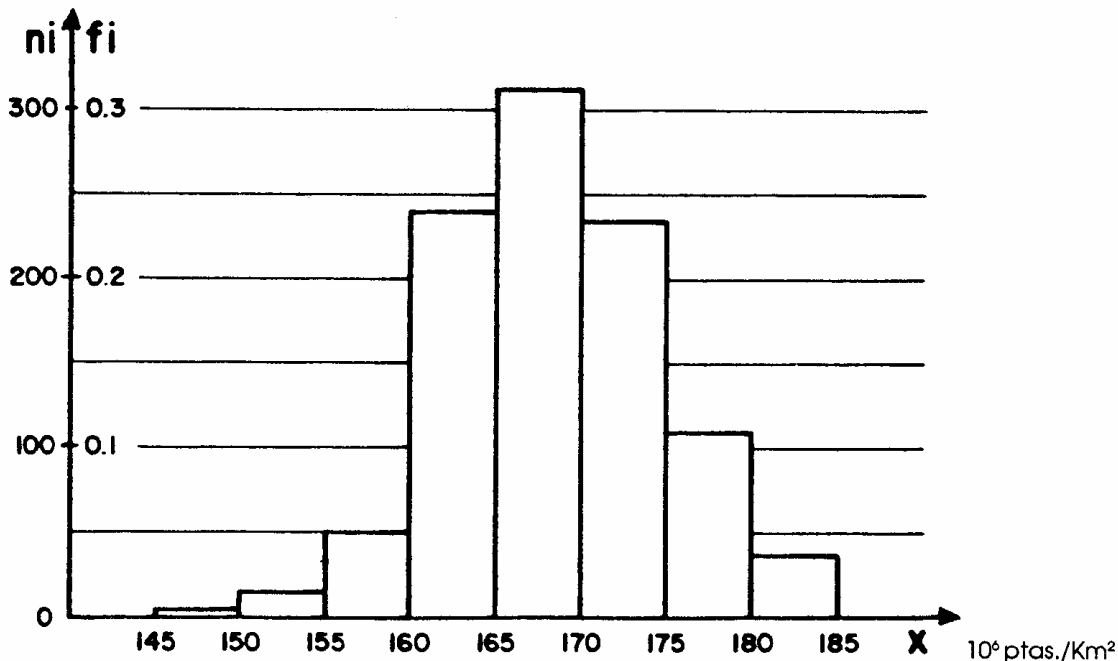


Fig. A-15.10. Histograma de las frecuencias municipales.

e) La representación gráfica de una tabla de frecuencias acumuladas ascendentes, cuando la variable se ha distribuido en intervalos de clase (figura A-15.11), se obtiene mediante un **polígono de frecuencias o diagrama acumulativo ascendente**, que consiste en unir mediante una línea poligonal los puntos cuyas abscisas son los extremos superiores de cada intervalo de clase (150, 155, 160, etc., en nuestro ejemplo) y cuyas ordenadas son las correspondientes frecuencias acumuladas. En la figura A-15.11, para un valor cualquiera de la variable territorial x , la ordenada correspondiente representa el número (absoluto o relativo) de municipios que tienen su densidad de renta inferior o dicho valor de x . El diagrama suele partir de un punto de ordenada cero cuya abscisa es el extremo inferior de la tabla (en nuestro ejemplo hemos tomado para dicho extremo la abscisa $x = 145 > 120$); a partir del extremo superior de la tabla ($185 < 208$, en nuestro caso) la línea es una recta paralela al eje de abscisas cuya altura constante es la frecuencia total ($n = 1.000$, en nuestro caso).

Veamos que, como regla general para formar los intervalos de las distribuciones de frecuencia de las variables territoriales, deberíase:

1) Determinar el mayor y el menor entre los datos registrados y así encontrar el recorrido o rango entre ellos (por diferencia entre ambos valores). 2) Dividir el rango entre un número conveniente de intervalos de clase del mismo tamaño; si ello no fuera posible, utilizar intervalos de clase de diferente tamaño o bien abiertos, eligiéndose de tal forma que las marcas de clase o puntos medios coincidan con los datos realmente observados, lo que tiende a aminorar el denominado "error de agrupamiento" que se comete en los tratamientos matemáticos posteriores. Sin embargo, los límites de clase podrán no coincidir con dichos datos reales.

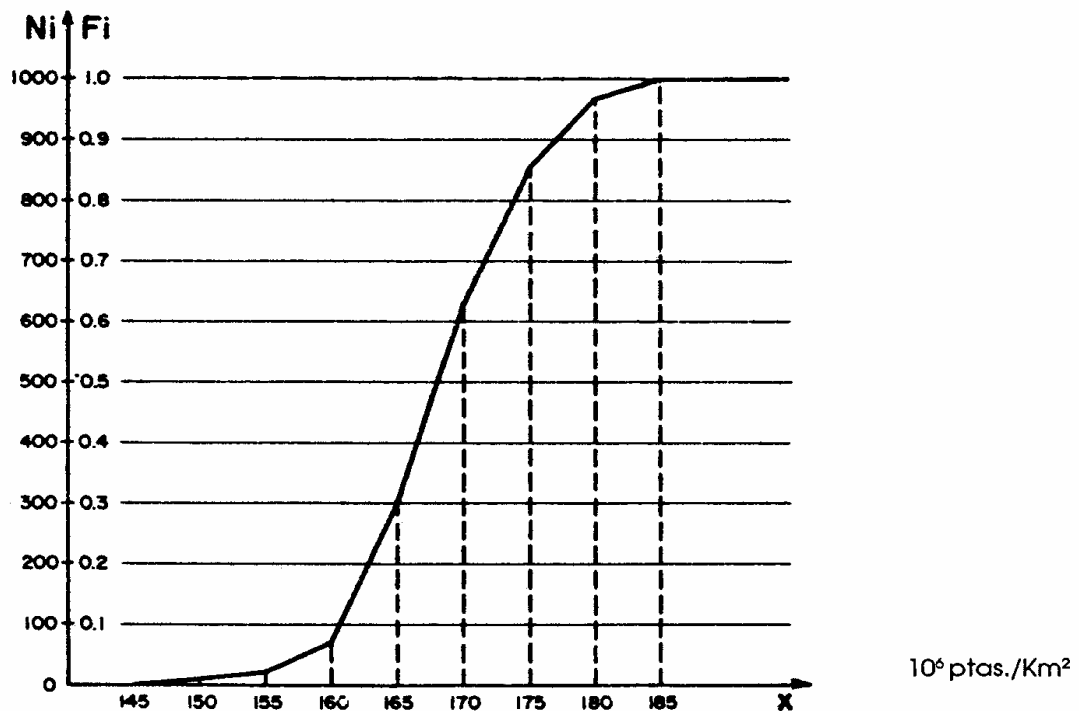


Fig. A-15.11. Diagrama acumulativo ascendente de las densidades municipales de renta.

5.2. "NORMALIZACIÓN" DEL PROBLEMA

Una variable aleatoria que se distribuya con la función de densidad que ya hemos relacionado en otras partes de este mismo trabajo, recibe el nombre de *normal* debido a que la distribución binomial, considerada en su caso límite, es la que corresponde corrientemente o habitualmente a la mayor parte de las variables empíricas, entre ellas las territoriales.

Desde luego, entre las muchas distribuciones continuas que se utilizan en Estadística y que pueden tener provechosas aplicaciones en el estudio de las variables territoriales, tal como vemos en este mismo capítulo, la *curva normal* o *distribución normal* es, con mucho, la más importante de ellas. Su estudio data de investigaciones sobre la naturaleza de los errores experimentales, llevadas a cabo en el siglo XVIII. Se observaba entonces que las discrepancias entre las medidas repetidas de la misma cantidad física

mostraban un sorprendente grado de regularidad; sus aspectos (distribución), según se encontró, podían aproximarse muy bien mediante un cierto tipo de curva de distribución continua, denominada “curva normal de errores” y atribuida a las leyes del azar. Las propiedades matemáticas de este tipo de curva de distribución continua y su base teórica fueron investigadas, por primera vez, por Abraham de Moivre (1667-1745), Pierre Laplace (1749-1827) y Karl Gauss (1777-1855). Este último fue un gran matemático alemán que, con sus brillantes aportaciones sobre las geometrías no euclidianas, hizo posible la aparición de las ciencias formales.

La importancia de la distribución normal radica, en primer lugar, en que son muy numerosas las variables aleatorias que la siguen, resultando adecuada para describir la distribución de muchos conjuntos de datos. En efecto, numerosas medidas físicas, datos meteorológicos, características biológicas, variables económicas y sociales, etc., siguen la ley normal, así como también aparece en muchas investigaciones teóricas. En segundo lugar, como consecuencia del teorema del límite central, que establece que la suma de un número elevado de variables aleatorias converge a una distribución normal, sea cual sea la distribución de estas variables. Además, en ciertas condiciones, las distribuciones discretas pueden ser substituidas por distribuciones normales, lo que simplifica notablemente los cálculos correspondientes. Hay que tener mucho cuidado, en fin, al suponer que un determinado conjunto de observaciones se puede aproximar por una distribución normal, pues será necesario realizar una comprobación previa.

Como se sabe, la curva normal tiene forma de campana extendida indefinidamente en ambas direcciones, positiva y negativa, siendo asintótica en relación al eje de abscisas. Rara vez es necesario extender las colas de la curva normal muy lejos de la media, porque el área comprendida bajo la curva y el eje horizontal que queda a más de cuatro o cinco desviaciones típicas de la media aritmética o esperanza matemática resulta insignificante para la mayoría de los fines prácticos y, entre ellos, los propios del Análisis Territorial. Debe tenerse en cuenta que no todas las distribuciones acampanadas simétricas en relación al eje de ordenadas son distribuciones normales, y las palabras *distribución normal* refiérense al hecho de que el área bajo la curva se distribuye de una manera determinada.

Una importante propiedad de la curva normal es que está completamente determinada por su media y su desviación típica. Es decir, la ecuación matemática de dicha curva es tal que se puede determinar el área existente bajo la curva entre dos puntos cualesquiera del eje horizontal si se conoce el valor que adoptan ambos parámetros, aunque en la práctica dichas áreas se obtienen valiéndose de tablas especiales elaboradas al efecto. Por otra parte, la probabilidad o frecuencia relativa con que una variable territorial

tomará valores entre dos puntos es el área bajo la curva comprendida entre los dos puntos del eje horizontal.

Si se representa gráficamente esta función (Fig. A-15.12.), encontraremos que está definida para todos los valores de x (desde $-\infty$ hasta $+\infty$) y con las siguientes propiedades:

- a) Solamente existe la curva para valores positivos de las ordenadas.
- b) El eje de abscisas OX es una asíntota de la curva.
- c) Existe un máximo para el punto $x = \alpha$.
- d) Es creciente hasta el máximo y después es decreciente.
- e) Existen dos puntos de inflexión: para $x = \alpha - \sigma$ y para $x = \alpha + \sigma$.
- f) Es cóncava hacia la región positiva del eje OY , para $-\infty < x < \alpha - \sigma$ y para $\alpha + \sigma < x < +\infty$ y cóncava hacia la región negativa del eje de ordenadas en el intervalo: $\alpha - \sigma < x < \alpha + \sigma$.

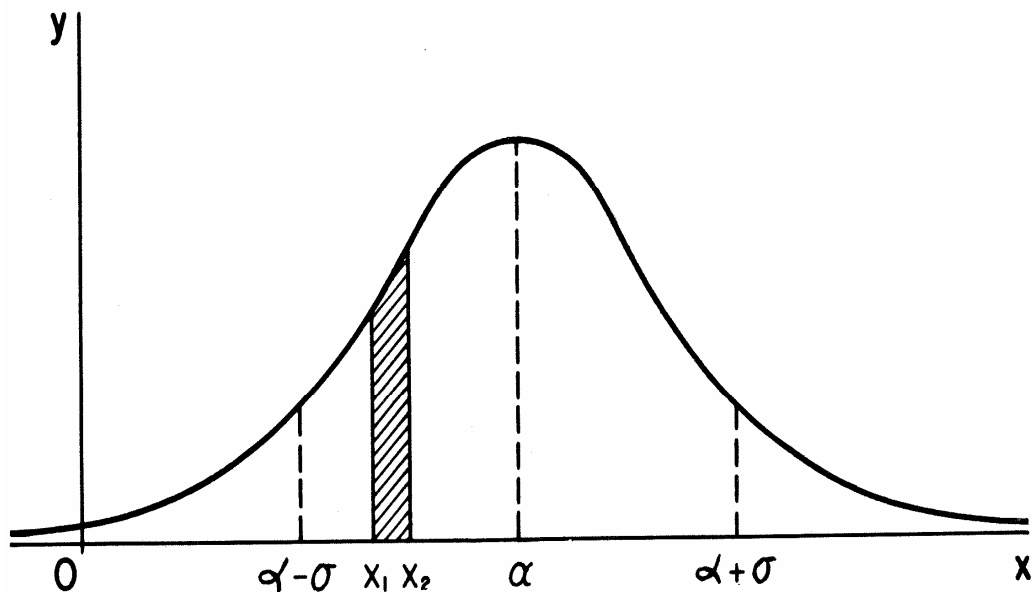


Fig. A-15.12. Área bajo la curva normal entre los puntos x_1 y x_2

Dentro de la ley normal los valores más próximos a α son más frecuentes que los más alejados de α . La utilización e interpretación de esta función para Gauss puede verse a través de un ejemplo como el siguiente: supongamos que se efectúan 1.000 estimaciones de la densidad de renta de un municipio determinado, cuyo valor verdadero es de 168×10^6 ptas./km²; en este caso se dice que $\alpha = 168$ ptas./km². Las medidas efectuadas se

acumularán en mayor proporción alrededor de α cuanto más próximas estén a dicho valor y serán menos frecuentes cuanto más alejadas se encuentren de α .

El parámetro σ está relacionado con la precisión de las mediciones realizadas al medir el parámetro α . Si se hace igual a la unidad el área comprendida entre la curva y el eje de abscisas y se supone que las densidades municipales de renta estimadas que han estado comprendidos entre los valores x_1 y x_2 representan el diez por ciento de las 1.000 estimaciones realizadas, el área rayada de la Fig. A-15.15. debe ser igual a 0'10. También sabemos hoy que el área comprendida entre la curva, el eje de abscisas y las ordenadas $x = \alpha - \sigma$ y $x = \alpha + \sigma$ corresponde, aproximadamente, al 64 por ciento del área total. Por lo tanto, ya tenemos un significado del hasta ahora desconocido parámetro σ : cuanto menor sea el valor de σ (que se denomina *desviación típica o standard* en la terminología estadística) serán de menor cuantía los errores que se cometen al estimar la densidad de renta α ; es decir, si la estimación de la variable territorial es más perfecta o si es más hábil la persona encargada de medirla, σ será un número menor que en el caso contrario.

La función de densidad normal o la función de distribución normal son dos modelos matemáticos inspirados por la conocida "ley de los errores" o ley de Gauss. Pues bien, la Estadística ha incorporado a su metodología –como un modelo probabilístico esencial– esta ley de Gauss con el nombre de *función de densidad normal*. Si $y = f(x)$ es dicha función de densidad, se puede definir a partir de ella la denominada *función de distribución normal* $F(x)$, tal que:

$$F(x) = \int_{-\infty}^x f(x) \cdot dx \quad ,$$

y dicha integral impropia de primera especie significa que a cada valor de x corresponde un número $F(x)$ determinado por la probabilidad de que la variable territorial tome un valor menor o igual a x .

Las áreas comprendidas bajo la curva normal y el eje de abscisas representan probabilidades. En estas condiciones, la probabilidad de que la variable x tome un valor comprendido entre x_1 y x_2 , vendría dada de la siguiente forma, aplicando la propiedad de la aditividad del intervalo de integración:

$$\begin{aligned}
 P_r(x_1 < x \leq x_2) &= F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(x) \cdot dx - \int_{-\infty}^{x_1} f(x) \cdot dx = \int_{x_1}^{x_2} f(x) \cdot dx = \\
 &= \int_{x_1}^{x_2} \frac{1}{\sigma \sqrt{2\pi}} \times e^{-\frac{(x-\alpha)^2}{2\sigma^2}} \times dx = \frac{1}{6'36 \times \sqrt{2\pi}} \times \int_{x_1}^{x_2} e^{-\frac{(x-168'325)^2}{2 \times 6'36^2}} \times dx = \\
 &= 0'063 \times \int_{x_1}^{x_2} e^{-\frac{(x-168'325)^2}{80'9}} \times dx
 \end{aligned}$$

es decir, como la última integral permite conocer el área rayada de la Fig. A-15.12., dicha área representa la probabilidad de que la variable $x = Q$ (que en el Cálculo de Probabilidades es una *variable aleatoria* o *estocástica*, en lugar de una *variable estadística*) tome un valor comprendido entre $x_1 = Q_1$ y $x_2 = Q_2$.

En el caso de que la Tabla anterior correspondiera a 1.000 mediciones distintas de una misma densidad de renta (en lugar de corresponder a las densidades de renta medidas en 1.000 municipios distintos), la distribución de frecuencias determinada por las columnas encabezadas por "10⁶ ptas./km²" y por "fi" presentaría una *imagen empírica* de una función de densidad normal. Si los resultados se hubieran obtenido a partir del modelo matemático, en lugar de constituir una observación de la realidad, para: $x_1 = 165$ y $x_2 = 170$, se tendría que:

$$\int_{165}^{170} f(x) dx = 0'312$$

Conocidas la media y la desviación típica se pueden tabular las áreas bajo la curva normal.

En realidad, para $a = 168$ (media aritmética de aquella distribución de frecuencias) y para $s = 6'36$ (desviación típica) se tiene que:

$$\int_{165}^{170} f(x) dx = 0'303 \quad ,$$

cuya diferencia: $0'312 - 0'303 = 0'009$ es una medida de la discrepancia existente entre la realidad y el modelo teórico que ha sido considerado y para la clase particular "de 165 a 170".

Las áreas comprendidas entre menos infinito y x son los valores de la función de distribución para cada valor de la variable territorial x .

De la misma manera, la distribución de frecuencias determinada por la primera columna de la Tabla A-15.6. y la encabezada por $F_i \uparrow$, constituyen una *imagen empírica de la función de distribución normal* que venimos considerando. En este caso, los resultados empírico y teórico, para $x = 170$, serían, respectivamente:

$$\int_{-\infty}^{170} f(x) dx = 0'620, \text{ y tambi3n: } \int_{-\infty}^{170} f(x) dx = 0'622$$

que implicarían una discrepancia absoluta de 0'002 y relativa del 0'32% entre el valor teórico y el correspondiente valor observado, que resulta ser francamente baja.

5.3. CARACTERÍSTICAS DE LA DISTRIBUCIÓN TERRITORIAL

5.3.1. Medidas centrales o promedios

Comenzaremos, como es habitual, por el cálculo de los diferentes promedios o valores de la distribución de la variable territorial "densidad municipal de renta". Para ello, y a partir de la tabla inicial, elaboraremos la siguiente:

TABLA DE CÁLCULOS DE LA DISTRIBUCIÓN DE DENSIDADES MUNICIPALES DE RENTA								
$n_i \cdot \log x_i$	$\log x_i$	(10° ptas./Km²)	x_i	n_i	$x_i \cdot n_i$	$N_i \uparrow$	$x_i^2 \cdot n_i$	n_i / x_i
6'506	2'1687920	Menos de 150	147'5	3	442'5	3	65.268'75	0'020
32'749	2'1832698	De 150 a 155	152'5	15	2.287'5	18	348.843'75	0'098
109'864	2'1972806	De 155 a 160	157'5	50	7.875'0	68	1.240.312'50	0'318
530'605	2'2108534	De 160 a 165	162'5	240	39.000'0	308	6.337.500'00	1'477
693'893	2'2240148	De 165 a 170	167'5	312	52.260'0	620	8.753.550'00	1'863
525'645	2'2367891	De 170 a 175	172'5	235	40.537'5	855	6.992.718'75	1'362
242'913	2'2491984	De 175 a 180	177'5	108	19.170'0	963	3.402.675'00	0'608
83'667	2'2612629	De 180 y más	182'5	37	6.752'5	1.000	1.232.331'25	0'203
2.225'842	////////	TOTAL	///	1000	168.325'0	///	28.373.200'00	5'949

Tabla A-15.7. Cálculos auxiliares de la distribución de densidades municipales de renta.

* **Media Aritmética:** a partir de los correspondientes resultados se tiene que (distribución conjunta):

$$\bar{X} = \sum_{i=1}^h x_i \cdot n_i / n = \alpha = 168.325 / 1.000 = 168.325.000 \text{ ptas./km}^2$$

que constituye la media aritmética o esperanza matemática de la distribución de frecuencias de la variable territorial analizada.

Por otra parte, y por lo que se refiere a la determinación de otros promedios o medidas centrales de esta distribución de frecuencias, veamos que, a partir de la primera columna y de la encabezada por $N_i \uparrow$ (ver las tablas A-15.6. y A-15.7.) se obtienen los siguientes resultados promedios (como $N_i \uparrow =$

500 está en el intervalo del 165 a 170, en él se encuentra la mediana o segundo cuartil):

$$* \text{ Mediana: } M_e = 165 + (500-308) / 312 \times 5 = 168 \times 10^6 \text{ ptas./km}^2$$

que se halla situada en el punto de intersección de los diagramas acumulativos ascendente y descendente, por definición.

$$* \text{ 1.}^{\text{er}} \text{ cuartil: } Q_1 = 160 + (250-68) / 240 \times 5 = 164 \times 10^6 \text{ ptas./km}^2$$

$$* \text{ 3.}^{\text{er}} \text{ cuartil: } Q_3 = 170 + (750-620) / 325 \times 5 = 173 \times 10^6 \text{ ptas./km}^2$$

Es decir, entre los municipios integrantes de la nación en estudio, la mitad de ellos tienen una densidad de renta inferior a 168×10^6 ptas./km², la cuarta parte, una densidad de renta inferior a 164, otra cuarta parte, más de 173 y entre las densidades 164 y 173 se encuentran el 50% de aquellos municipios.

Otros promedios interesantes serían:

$$* \text{ Moda: } M_o = 165 + (235 / 240+235) \times 5 = 167'47 \times 10^6 \text{ ptas./km}^2$$

* **Media cuadrática:**

$$C = \sqrt{\sum_{i=1}^8 x_i^2 \cdot n_i / n} = \sqrt{28.373.200 / 1.000} = 168'44 \times 10^6 \text{ ptas./km}^2$$

* **Media geométrica:**

$$G = \sqrt[n]{\prod_{i=1}^8 x_i^{n_i}} = \text{antilog. } \sum_{i=1}^8 n_i \cdot \log x_i / n = \\ = \text{antilog } 2.225'842 / 1.000 = 168'2062 \times 10^6 \text{ ptas./km}^2,$$

* **Media armónica:**

$$H = n / \sum_{i=1}^8 n_i / x_i = 1.000 / 5'949 = 168'0955 \times 10^6 \text{ ptas./km}^2,$$

Desde luego, las cuatro medias aquí estudiadas quedan ordenadas con arreglo a su magnitud del modo siguiente:

armónica < geométrica < aritmética < cuadrática

$$(H = 168'0955) < (G = 168'2062) < (\bar{X} = 168'3250) < (C=168'4435) \quad (*)$$

De hecho, la relación matemática existente entre las medias aritmética, geométrica, cuadrática y armónica, para una misma distribución de frecuencias con todos sus datos positivos, debe cumplir la monotonía ascendente:

$$H \leq G \leq \bar{X} \leq C$$

En efecto. Consideremos el caso más sencillo de una distribución con dos valores de la variable con frecuencias unitarias y que con dichos valores pueden calcularse los cuatro promedios antedichos:

$$H = \frac{2}{\frac{1}{x_1} + \frac{1}{x_2}} = \frac{2}{\frac{x_1 + x_2}{x_1 x_2}} = \frac{2x_1 x_2}{x_1 + x_2}$$

$$G = \sqrt{x_1 x_2}$$

$$\bar{X} = \frac{x_1 + x_2}{2}; C = \sqrt{\frac{x_1^2 + x_2^2}{2}}$$

Vamos a demostrar en primer lugar que $H \leq G$, o sea:

$$\frac{2x_1 x_2}{x_1 + x_2} \leq \sqrt{x_1 x_2}$$

Elevando al cuadrado los miembros de la anterior desigualdad y operando:

$$4x_1^2 x_2^2 \leq x_1 x_2 (x_1 + x_2)^2 \quad ; \quad 4x_1 x_2 \leq (x_1 + x_2)^2$$

$$4x_1 x_2 \leq x_1^2 + x_2^2 + 2x_1 x_2 \quad ; \quad 0 \leq x_1^2 + x_2^2 - 2x_1 x_2$$

$$0 \leq (x_1 - x_2)^2$$

Con lo que queda demostrado que $H \leq G$. Por otro lado $G \leq \bar{X}$ ya que:

$$\sqrt{x_1 x_2} \leq \frac{x_1 + x_2}{2} \quad ; \quad 4x_1 x_2 \leq (x_1 + x_2)^2$$

Con lo que se tendrá que:

$$0 \leq (x_1 - x_2)^2$$

Por tanto, queda demostrado que $H \leq G \leq \bar{X}$. Por último, teniendo en cuenta la relación que liga la media cuadrática con la aritmética y la varianza, esto es: $C^2 = \bar{X}^2 + \sigma^2$, se deduce que: $\bar{X} \leq C$, como se quería demostrar.

Esta demostración puede generalizarse para cualquier número de valores de la variable territorial en estudio. En cualquier caso, la demostración

de la última desigualdad también puede realizarse analíticamente a partir de la definición de ambas medias:

$$\bar{X} = \frac{x_1 + x_2}{2} ; C = \sqrt{\frac{x_1^2 + x_2^2}{2}}$$

$$\bar{X}^2 = \frac{x_1^2 + x_2^2 + 2x_1x_2}{4} ; C^2 = \frac{x_1^2 + x_2^2}{2} = \frac{x_1^2 + x_2^2}{4} + \frac{x_1^2 + x_2^2}{4}$$

$$\bar{X}^2 = \frac{x_1^2 + x_2^2}{4} + \frac{x_1x_2}{2} ; C^2 - \bar{X}^2 = \frac{x_1^2 + x_2^2}{4} - \frac{x_1x_2}{2} ;$$

Obsérvese que esta diferencia de fracciones debe ser necesariamente positiva o nula. En efecto, se trata de comparar las expresiones:

$$\frac{x_1^2 + x_2^2}{4} \text{ y } \frac{2x_1x_2}{4}, \text{ o más concretamente sus numeradores, o sea:}$$

$x_1^2 + x_2^2$ y $2x_1x_2$. Como se tiene que el cuadrado de una diferencia ofrece:

$$(x_1 - x_2)^2 = x_1^2 + x_2^2 - 2x_1x_2 \geq 0 \Rightarrow x_1^2 + x_2^2 \geq 2x_1x_2,$$

con lo que también: $C^2 \geq \bar{X}^2$, y además: $\bar{X} \leq C$, c.s.q.d.

5.3.2. Medidas de dispersión o concentración

Por lo que se refiere a la desviación típica o "standard", como medida de la dispersión absoluta de la distribución por el territorio de nuestra variable territorial, veamos que su valor vendrá dado por:

$$\sigma = \sqrt{C^2 - \bar{X}^2} = \sqrt{168'44^2 - 168'33^2} = 6'36 \times 10^6 \text{ ptas./km}^2,$$

y un coeficiente de variación de Pearson (medida de dispersión relativa) de :

$$CV = \sigma / \bar{X} = 6'35 \times 10^6 / 168'32 \times 10^6 \approx 0'038 = 3'8\% .$$

Recorrido intercuartílico:

$$Q_3 - Q_1 = 173 - 164 = 9 \times 10^6 \text{ ptas./km}^2,$$

Recorrido semi-intercuartílico:

$$(Q_3 - Q_1) / (Q_3 + Q_1) = 9 / (173 + 164) = 0'02671,$$

es decir, que en ambos casos, la medida de dispersión empleada viene a representar sólo alrededor del 3% del correspondiente promedio. **Desde luego,**

ello indicaría que el grado de uniformidad y equilibrio del territorio que nos ocupa resulta francamente elevado. En efecto, los coeficientes de uniformidad anteriormente definidos, tomarán, en este caso, los siguientes valores:

$$\begin{cases} CU_1 = 100 \cdot (1 - CV) = 100 \cdot (1 - 0'038) = 96'2\% \\ CU_3 = 100 \cdot (1 - 1'27 \cdot CV) = 100 \cdot (1 - 1'27 \cdot 0'038) = 95'2\% \\ CU_4 = 100 \cdot (1 - 0'80 \cdot CV) = 100 \cdot (1 - 0'80 \cdot 0'038) = 97'0\% \end{cases}$$

debiendo considerarse también que:

$$CU_2 = (Q_1 / \bar{X}) \cdot 100 = (16379 / 168'32) \cdot 100 = 97'3\%$$

, si bien otra determinación del mismo coeficiente de uniformidad territorial conduciría al valor:

$$CU_2 = 100 (1 - 0'68 \cdot CV) = 100 (1 - 0'68 \times 0'038) = 97'4\% ,$$

cuya pequeña discrepancia (+0'1%) con el resultado anterior débese al propio ajuste de normalidad, o bien al proceso de cálculo decimal.

Veamos, por último, que el "coeficiente de uniformidad territorial medio", ofrecerá un valor de:

$$\overline{CU} = 100 (1 - 0'92 \cdot CV) = 100 (1 - 0'92 \times 0'038) = 96'5\% ,$$

mientras que también:

$$CU_5 = 100 \cdot \sqrt{\frac{Q_1}{Q_3}} = 100 \cdot \sqrt{\frac{16379}{17277}} = 97'4\%$$

La representación gráfica de los valores de los diferentes coeficientes de uniformidad hallados, en definitiva, es la siguiente:

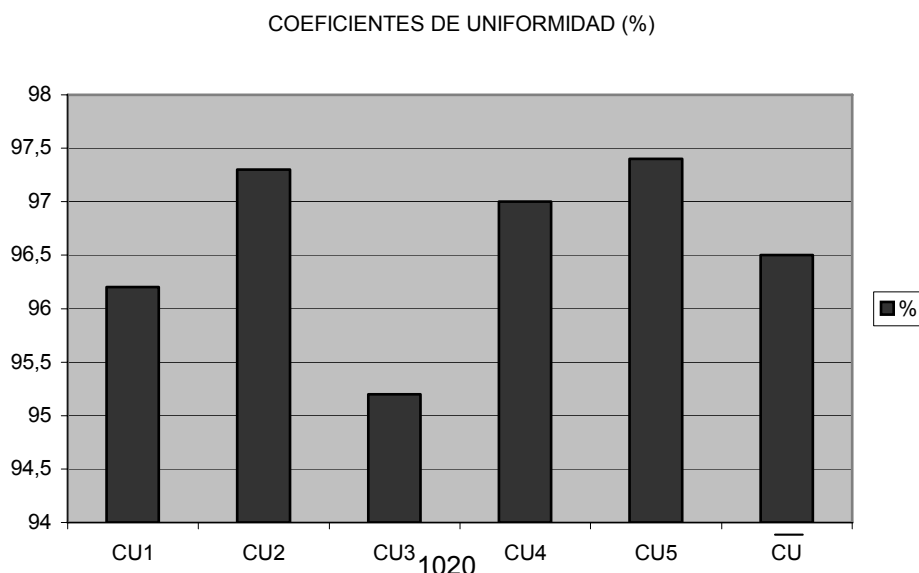


Fig. A-15.13. Valores de los diferentes coeficientes de uniformidad.

5.3.3. Otras características de la distribución espacial

Por lo que se refiere a las restantes características de la distribución espacial de la variable territorial "densidad municipal de renta", veamos que una de la asimetría o sesgo la constituye el denominado "primer coeficiente de asimetría de Pearson":

$$P_1 = \frac{\bar{X} - M_0}{\sigma} = \frac{168'32 - 167'47}{6'36} = 0'134 \approx 0 ,$$

luego se trata de una distribución prácticamente simétrica.

También podríamos calcular el "segundo coeficiente de asimetría de Pearson", a saber:

$$P_2 = \frac{3 \cdot (\bar{X} - M_e)}{\sigma} = \frac{3 \cdot (168'32 - 168'08)}{6'36} = 0'113 ,$$

o bien por el "coeficiente de sesgo cuartílico", de valor:

$$\frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{172'77 - 2 \cdot 168'08 + 163'79}{172'77 - 163'79} = 0'045 ,$$

que conducen, en todos los casos, a conclusiones similares. Cabe observar, no obstante, una ligera asimetría o sesgo hacia la derecha, puesto que: $P_1 > 0$, $P_2 > 0$ y también:

$$\bar{X} = 168'32 > M_e = 168'00 > M_0 = 167'47$$

Así mismo, se elaborará la siguiente tabla:

n_i	$x_i - 168$	$(x_i - 168)^3 n_i$	$(x_i - 168)^4 n_i$	$ x_i - 168 n_i$
3	- 20'5	- 25.844'4	529.810'2	61'5
15	- 15'5	- 55.851'2	865.693'6	232'5
50	- 10'5	- 57.881'3	607.753'6	525'0
240	- 5'5	- 39.930'0	219.615'0	1.320'0
312	- 0'5	- 39'0	19'5	156'0
235	4'5	21.414'6	96.365'7	1.057'5
108	9'5	92.596'5	879.666'7	1.026'0
37	14'5	112.800'0	1.635.600'0	536'5
1.000	////	47.265'2	4.834.524'3	4.915'0

$$\sum_{i=1}^8$$

Tabla A-15.8. Tabla auxiliar de cálculo.

De los resultados de la tabla anterior, se deduce que el momento central (respecto a la media aritmética) de tercer orden es:

$$m_3 = 47'2652 \text{ ,}$$

por lo que se tendrá un "coeficiente directo de asimetría" o "coeficiente de sesgo de Fisher" de:

$$g_1 = m_3 / \sigma^3 = 47'2652 / 6'36^3 = 0'18 \text{ ,}$$

que confirma la existencia de una distribución aceptablemente simétrica (en curvas simétricas como la normal, se cumple que: $g_1 = g_1^2 = 0$).

Por otra parte, $m_4 = 4.834'5243$ (momento central o respecto al origen de 4º orden) y, por tanto, se tendrá un "coeficiente de curtosis o apuntamiento de Fisher" de:

$$g_2 = (m_4 / \sigma^4) - 3 = (4.834'5243 / 6'36^4) - 3 = 2'95 - 3 = - 0'05 \approx 0 \text{ ,}$$

lo que permite asegurar que la distribución es **mesocúrtica** y aproximadamente normal (la curva normal tiene: $g_2 = 0$).

Veamos, así mismo, que la anterior tabla auxiliar de cálculo A-15.8, en su última columna, nos permitirá el cálculo de otra medida de dispersión absoluta a la que ya nos hemos referido con anterioridad: la desviación media respecto a la media aritmética (que será mínima con respecto a la mediana), a saber:

$$\begin{aligned} DM &= \frac{\sum_{i=1}^8 |x_i - \bar{X}| \cdot n_i}{n} = 4.915 / 1.000 = 4'915 \approx 4\sigma / 5 = \\ &= 4 \cdot 6'36 / 5 = 5'088 \end{aligned}$$

Este valor de la DM ($5'088 \times 10^6$ ptas./km²) en la presente distribución agrupada de frecuencias, conducirá a una determinación más ajustada y directa del valor del coeficiente de uniformidad territorial CU₄, a saber:

$$CU_4 = 100 \cdot (1 - 4'915 / 168'325) = 97'08 \% \text{ ,}$$

prácticamente coincidente con el anteriormente calculado mediante procedimientos indirectos.

5.3.4. Índice de Gini y curva de Lorenz

Para la determinación de dichos ítems, es necesaria la elaboración de la siguiente tabla de cálculos auxiliares:

L_i	x_i	n_i	$Q_i = x_i \cdot n_i$	$\frac{n_i}{n} \times 100$	$\frac{x_i \cdot n_i}{\sum x_i \cdot n_i} \times 100$	p_i	q_i	$p_i - q_i$
<150	135'0	3	405'0	0'3	0'24	0'3	0'24	0'06
150-155	152'5	15	2.287'5	1'5	1'36	1'8	1'60	0'20
155-160	157'5	50	7.875'0	5'0	4'67	6'8	6'27	0'53
160-165	162'5	240	39.000'0	24'0	23'12	30'8	29'39	1'41
165-170	167'5	312	52.260'0	31'2	30'97	62'0	60'36	1'64
170-175	172'5	235	40.537'5	23'5	24'03	85'5	84'39	1'11
175-180	177'5	108	19.170'0	10'8	11'36	96'3	95'75	0'55
>180	194'0	37	7.178'0	3'7	4'25	100	100	0'00
$\sum_{i=1}^8$		$n = 1.000$	$Q = 168.713'0$	100%	100%	383'5	378'0	5'50

Tabla A-15.9. Cálculos auxiliares para la determinación del índice G.

En este caso, al objeto de precisar mejor los cálculos, los valores de la primera y última marca de clase: $x_1 = 135'0 \times 10^6$ ptas./km² y $x_8 = 194'0 \times 10^6$ ptas./km² se han estimado teniendo en cuenta los valores extremos: $Q_0 = 120'0 \times 10^6$ ptas./km² y $Q_4 = 208'1 \times 10^6$ ptas./km²

Así pues, según la fórmula dada por Pulido, el valor del índice de Gini, en este caso, será de:

$$G = \frac{\sum_{i=1}^{K-1} (p_i - q_i)}{\sum_{i=1}^{K-1} p_i} = \frac{5'50}{283'5} = 0'02 = 2\%$$

Habida cuenta del bajísimo valor del índice obtenido, podemos afirmar que la distribución de las densidades municipales de renta por el territorio estudiado es casi perfecta desde el punto de vista estadístico.

La correspondiente curva poligonal de Lorenz apenas se separa de la diagonal o bisectriz del primer cuadrante, dado el bajísimo valor del índice de Gini ya obtenido, por lo que obviaremos su representación gráfica, a efectos prácticos.

5.3.5. Índice de Williamson

En nuestro caso, como ya es sabido, la variable territorial en estudio es la densidad de renta municipal. Por ello, la fórmula pertinente, en relación al número de municipios, vendrá dada por la expresión (FRANQUET, 1990/91):

$$W_{Q,n} = \sqrt{\frac{\sum_{i=1}^8 \left(\frac{Q_i}{n_i} - \frac{Q}{n} \right)^2 \times \frac{n_i}{n}}{Q/n}}, \forall i \in (1,2,\dots,8)$$

Esto es:

$$\begin{aligned} W &= \sqrt{\frac{\sum_{i=1}^8 \left(x_i - \frac{168.713}{1.000} \right)^2 \times f_i}{168'713}} = \\ &= \sqrt{\frac{(135'0-168'713)^2 \times 0'003 + (152'5-168'713)^2 \times 0'015 + (157'5-168'713)^2 \times 0'05 +}{168'713}} \\ &\sqrt{\frac{+ (162'5-168'713)^2 \times 0'24 + (167'5-168'713)^2 \times 0'312 + (172'5-168'713)^2 \times 0'235 +}{168'713}} \\ &\sqrt{\frac{+ (177'5-168'713)^2 \times 0'108 + (194'0-168'713)^2 \times 0'037}{168'713}} = \sqrt{\frac{58'73}{168'713}} = 0'59 \end{aligned}$$

que, como era de esperar, resulta ser muy bajo.

5.3.6. Índice de concentración de Lorenz

En el caso que nos ocupa, veamos que un valor aproximado de este índice es el que se obtiene mediante la aplicación de la fórmula basada en los porcentajes acumulados, que se emplea comúnmente en los trabajos prácticos. Aquí, la fórmula correspondiente explicada en epígrafes anteriores de este mismo capítulo de nuestro trabajo, tomará la configuración simplificada (con $n = 8$ y $q_n = 100$):

$$L = 1 - \frac{2}{7} \times \frac{\sum_{i=1}^7 q_i}{100} = 1 - \frac{\sum_{i=1}^7 q_i}{350}$$

El resultado que ofrece la aplicación de la fórmula anterior es el siguiente, teniendo en cuenta, como ya se ha dicho, que resulta necesario ordenar los valores de la variable territorial en estudio de menor a mayor, para la aplicación correcta de la fórmula, así:

	X_i	q_i
	0'24	0'24
	1'36	1'60
	4'25	5'85
	4'67	10'52
	11'36	21'88
	23'12	45'00
	24'03	69'03
	30'97	100'00
$\sum_{i=1}^8$	100'00	254'12

Tabla A-15.10. Ordenación creciente de los porcentajes acumulados de las densidades municipales de renta.

A esta distribución de probabilidad le corresponde, pues, el índice de Lorenz:

$$L = 1 - \frac{154'12}{350} = 0'56$$

5.4. AJUSTE A UNA DISTRIBUCIÓN NORMAL

5.4.1. La hipótesis de normalidad y el estadígrafo χ^2

La distribución que venimos estudiando de la variable territorial "densidad municipal de renta", en la nación considerada, se ajusta muy aceptablemente a una distribución normal. De este modo, la distribución teórica de las densidades municipales de renta vendrá dada, como ya se ha visto, por la función de densidad normal:

$$y = (1/6'36 \cdot \sqrt{2\pi}) \cdot e^{-\frac{(x-168'325)^2}{80'8992}}$$

Pese a que existen muchas de ellas, la distribución normal o de Gauss es, sin duda, la más importante y la de más provechosas aplicaciones territoriales de entre todas las distribuciones teóricas de probabilidad de tipo continuo. No obstante, hay que tener mucho cuidado al suponer que un determinado conjunto de observaciones de variables territoriales se puede aproximar con suficiente exactitud por una distribución normal, pues será necesario llevar a cabo una comprobación previa.

Mediante los métodos adecuados de la Estadística Matemática, sería posible justificar la bondad de este ajuste contrastando la **hipótesis de normalidad** mediante el empleo del estadígrafo χ^2 de Pearson, tal como ya se realiza en algún otro apartado de esta misma tesis doctoral. En la fig. A-15.14 se ha representado gráficamente la función anterior. En realidad, tal representación significa el **ajuste de una función de densidad normal** a la distribución de las densidades municipales de renta observadas. Utilizando una tabla de la distribución normal, podemos obtener las áreas de las superficies que están limitadas por la curva normal, el eje de abscisas y las ordenadas en los extremos o límites de los intervalos de clase. Dado que, como es bien sabido, dicha área total es igual a la unidad, resultará que cada una de ellas expresa la probabilidad p_i de que la variable territorial "densidad de renta" tome un valor perteneciente al correspondiente intervalo de clase.

Para proceder al ajuste es necesario transformar los valores observados de la variable territorial (en este caso representados por el límite superior L_i de cada intervalo de clase) en *valores tipificados* o "estandarizados", lo que se consigue restando de cada L_i la media 168 (columna 2ª de la Tabla A-15.11) y dividiendo dicha diferencia por la desviación típica $6'36 \times 10^6$ ptas./km² de la distribución observada (3ª columna), al objeto de poder emplear las tablas de la distribución normal, obtenidas para una distribución de media aritmética o esperanza matemática $\alpha = 0$ y desviación típica $\sigma = 1$.

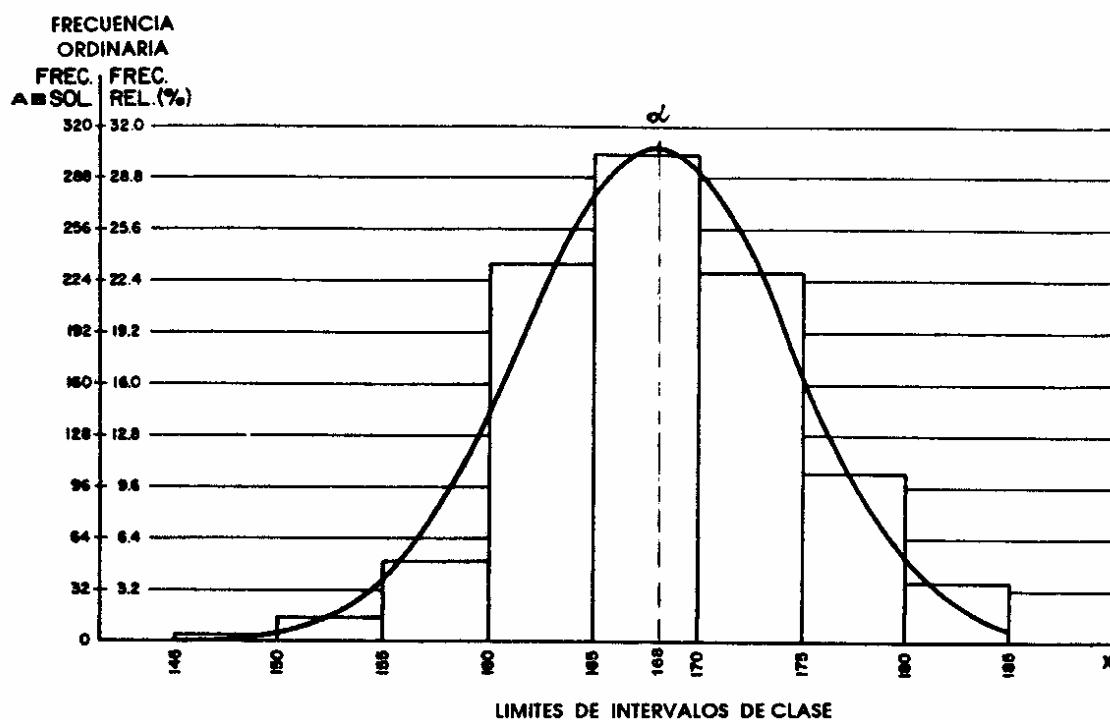


Fig. A-15.14. Distribución campaniforme o gaussiana de la variable territorial.

AJUSTE DE UNA DISTRIBUCIÓN NORMAL A LA DE "DENSIDADES DE RENTA"						
L_i	L_i-168	$t = \frac{L_i-168}{6,36}$	P	P_i	T_i	n_i
150	-18	-2,817	0,002	0,002	2	3
155	-13	-2,034	0,021	0,019	19	15
160	-8	-1,252	0,105	0,084	84	50
165	-3	-0,469	0,319	0,214	214	240
170	2	0,313	0,622	0,303	303	312
175	7	1,095	0,864	0,242	242	235
180	12	1,878	0,971	0,107	107	108
"	"	"	1,000	0,029	29	37
				1,000	1,000	1,000

Tabla A-15.11. Ajuste de una distribución normal a la de "densidades municipales de renta".

Las áreas bajo la curva normal se obtienen como diferencia entre los valores correspondientes de la función de distribución. A partir de estos valores tipificados se puede emplear la Tabla de la Distribución Normal, calculando mediante las oportunas interpolaciones, en su caso, los valores P, tales que:

$$P_1 = \int_{-\infty}^{-2,817} f(t) dt = 0'002 \quad ; \quad P_2 = \int_{-\infty}^{-2,034} f(t) dt = 0'021 \quad ; \text{etc.},$$

siendo:

$$t = Z = (x - \alpha) / \sigma$$

La diferencia entre dos P_i consecutivas determina las probabilidades p_i de que la variable territorial tome un valor comprendido entre L_{i-1} y L_i , de tal manera que:

$$p_1 = P_1 = 0'002 \quad ;$$

$$p_2 = P_2 - P_1 = 0'021 - 0'002 = 0'019 \quad ; \text{etc.}$$

Las frecuencias teóricas son proporcionales a las correspondientes áreas bajo la curva normal. Distribuyendo la frecuencia total, $n = 1.000$ en nuestro caso, proporcionalmente respecto a las probabilidades p_i , se obtienen *frecuencias teóricas* T_i , que corresponderían a las densidades municipales de renta observadas si dicha variable territorial se ajustara *exactamente* al modelo teórico de la distribución normal. En nuestro ejemplo, las discrepancias entre los valores observados n_i y los teóricos T_i son lo suficientemente pequeñas como para que pueda aceptarse la hipótesis de que la *variable territorial sigue una distribución normal*, a pesar de la determinación que realizaremos a continuación.

En cualquier caso puede realizarse, como ya se ha apuntado, un contraste de la bondad del ajuste utilizando el estadígrafo χ^2 de Pearson, con (n-1) grados de libertad, a saber:

$$\chi^2 = \sum_{i=1}^n \frac{(n_i - T_i)^2}{T_i}$$

puesto que $\forall T_i > 5$, a excepción del correspondiente al primer intervalo de clase L_1 , ya que $T_1 = 2$.

De este modo, se tendrá un valor del estadígrafo:

$$\begin{aligned} \chi^2 = & \frac{(3-2)^2}{2} + \frac{(15-19)^2}{19} + \frac{(50-84)^2}{84} + \frac{(240-214)^2}{214} + \frac{(312-303)^2}{303} + \\ & + \frac{(235-242)^2}{242} + \frac{(108-107)^2}{107} + \frac{(37-29)^2}{29} = 20'95 \end{aligned}$$

El valor de χ^2 con: 8-1 = 7 g.l., para una probabilidad del 5% es 14'067 < 20'95, luego se rechaza la hipótesis nula y, en su consecuencia, el ajuste o asimilación a una distribución normal será sólo orientativo, pero no aceptable estrictamente.

5.4.2. Determinación y fiabilidad del coeficiente de correlación no lineal

Podemos, no obstante, calcular el coeficiente de determinación o crítico del ajuste anterior, a partir de la siguiente fórmula de la varianza residual:

$$\begin{aligned} S_e^2 &= \frac{\sum_{i=1}^n e_i^2}{n} = \frac{2.044}{8} = 255'5 \quad ; \quad S_{n_i}^2 = 12.617 \quad ; \\ R = r^2 &= 1 - \frac{S_e^2}{S_{n_i}^2} = 1 - \frac{255'5}{12.617} = 0'980 \quad , \end{aligned}$$

lo que implica un coeficiente de correlación de:

$$r = \pm \sqrt{0'980} = \pm \mathbf{0'990}$$

que resulta suficientemente aceptable, razón por la que aceptaremos como válido el presente ajuste normalizado.

Por otra parte, y por lo que se refiere a la fiabilidad del coeficiente de correlación hallado r , definimos la variable aleatoria:

$$z = 1/2 \cdot \ln [(1+r) / (1-r)] \quad ; \quad e^{2z} = [(1+r) / (1-r)] \quad ;$$

con $\rho = 0'990$ y $n = 8$.

Se trata de determinar un intervalo de valores entre los que puede razonablemente esperarse (con una probabilidad del 95%) que se encuentre r , con media:

$$\mu_z = \frac{\ln \frac{1+\rho}{1-\rho}}{2} = \frac{\ln \frac{1'99}{0'01}}{2} = 2'647$$

y desviación típica:

$$\sigma_z = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{5}} = 0'447$$

El intervalo será:

$$Z = m_z \pm 2 \cdot s_z = 2'647 \pm 0'894 = 3'541 \text{ y } 1'753 ,$$

valores éstos que corresponden, respectivamente, a: $r_1 = 0'998$ y $r_2 = 0'942$.

Así pues, puede afirmarse que la probabilidad de que se cumpla la desigualdad: $0'942 < r < 0'998$, es del 95%.

Aunque la relación precedente simplifique, de modo notable, el problema de determinar la exactitud de r como estimador de ρ , tiene la desventaja de no ser fiable si las dos variables analizadas no poseen una distribución normal conjunta. En su consecuencia, a menos que se esté bastante seguro de que estas variables tengan tal distribución -por lo menos con buena aproximación- no debe confiarse grandemente en los resultados obtenidos.

5.4.3. Distribución exponencial

También se le suele llamar *distribución exponencial negativa* (FRANQUET, 2003).

Diremos que una variable aleatoria X , de tipo continuo, sigue una **distribución exponencial** de parámetro a , siendo $a \in \mathfrak{R}$ y $a > 0$, si su *función de densidad* es:

$$f(x) = \begin{cases} a \cdot e^{-ax} & , \forall x > 0 \\ 0 & , \forall x \leq 0 \end{cases}$$

Abreviadamente lo indicaremos por:

$$X \rightarrow \text{Exp}(a)$$

Constituye un caso particular de la distribución teórica de probabilidad "gamma" $\Gamma(\alpha, a)$ que se presenta cuando: $\alpha = 1$, circunstancia que da lugar a

esta “distribución exponencial”. O sea, se tiene la siguiente función de densidad:

$$f(x) = \begin{cases} \frac{e^{-x/\beta}}{\beta}, & \forall x > 0 \\ 0 & \forall x \leq 0 \end{cases}$$

En nuestro caso, asimilamos: $\beta = \bar{X} = 168'713 \times 10^6$ ptas./km² (media aritmética de la distribución o esperanza matemática de la misma), con lo que:

$$f(x) = \frac{1}{168'713} \cdot e^{-\frac{x}{168'713}}$$

Así pues, si se trata, por ejemplo, de saber la probabilidad de encontrar, en esta nación, un municipio de densidad de renta inferior a 160×10^6 ptas./km², el problema estriba en calcular la integral definida:

$$F(x < 160) = \int_{135}^{160} \frac{e^{-\frac{x}{168'713}}}{168'713} \cdot dx = \left\{ \begin{array}{l} t = x/168'713 \\ dx = 168'713 \cdot dt \end{array} \right\} =$$

(haciendo el correspondiente cambio de variable)

$$= \int_{0'80}^{0'95} e^{-t} \cdot dt = [-e^{-t}]_{0'80}^{0'95} = -e^{-0'95} + e^{-0'80} = 0'45 - 0'39 = 0'06 = 6\%$$

resultado éste que diverge algo del que se deduce de la tabla anterior, donde correspondería el 6'8%, como se puede comprobar. De hecho, la determinación del grado de exactitud del ajuste de la distribución anterior a la distribución teórica exponencial se habría de contrastar mediante un *test* de hipótesis χ^2 con (k-1) grados de libertad, siendo **k** el número de pares de clase comparados. En nuestro caso, tendríamos que:

$$e_i = \frac{e^{-\frac{x_i}{168'713}}}{168'713} \cdot n = \frac{5'93}{e^{x_i/168'713}};$$

Obviamente, el valor del estadígrafo es:

$$\chi^2 = \sum_{i=1}^8 \frac{(n_i - e_i)^2}{e_i}$$

que resulta ser elevado e inapropiado, habida cuenta del carácter descendente de esta distribución teórica de probabilidad, que no se corresponde con la realidad del caso práctico cuya resolución nos ocupa, con lo que, la

discrepancia entre los valores teóricos (propios de la distribución exponencial) y los reales es bastante grande, y haría falta buscar una distribución mejor, circunstancia que excede las pretensiones de nuestro problema, ya que $\chi^2_{0.95}$ (5 g.l.) = 11'07, en el caso de escoger una región crítica del 5%.

5.4.4. Corrección por agrupamiento en "clases"

Veamos, por último, que por haberse realizado, en este ejemplo práctico, un agrupamiento en clases o intervalos de amplitud: $c = 5 \times 10^6$ ptas./km² de los valores de la variable territorial analizada (la "densidad municipal de renta"), procede aplicar la corrección de Sheppard para la determinación de la desviación típica más ajustada de los datos del problema (SPIEGEL, 1981). Esto es:

$$\sigma_c^2 = \sigma^2 - \frac{c^2}{12} = 168'44^2 - 168'32 - \frac{5^2}{12} = 38'33 \quad (\text{varianza corregida})$$

$$\sigma_c = \sqrt{38'33 \times 10^{12}} = 6'19 \times 10^6 \text{ ptas./km}^2 \quad (\text{desviación típica corregida}),$$

que lógicamente, resultan ser de alguna menor cuantía que en la primera determinación efectuada. Ello podría obligar a una ligera revisión de los cálculos posteriores en los que se haya hecho intervenir a la expresada medida de la dispersión espacial de los valores de la variable territorial estudiada.



(*) Por lo que se refiere a las caracterización del valor central, veamos que G. Udny Yule, estadístico inglés, en su "Introducción a la Teoría de la Estadística", ha precisado las condiciones que debe cumplir una buena caracterización del valor central de una serie. En resumen son las siguientes:

a) *La característica del valor central debe ser definida objetivamente a partir de los datos de la serie, sin que haya lugar a intervenir ninguna apreciación del estadístico.*

b) *Debe depender de todas las observaciones de la serie, a ser posible. Señalemos que, no obstante, hay veces que se plantea el problema de decidir si debe tenerse en cuenta una observación que es notablemente distinta de todas las demás de su conjunto o si puede ser rechazada por considerar que tal observación tiene carácter excepcional debido a algún factor extraño a la serie como, por ejemplo, un error de observación.*

c) *Debe tener, en la medida de lo posible, una significación concreta, sencilla y fácil de comprender. Si se tiene en cuenta que muchos de los valores centrales de las series han de ser utilizados por personas poco familiarizadas con la Estadística, se comprende la preferencia*

que en la realidad se ha dado a la media aritmética como característica del valor central de que goza esta propiedad, de una interpretación sencilla.

d) *Debe ser de cálculo fácil y rápido.*

e) *Debe ser poco sensible a las fluctuaciones del muestreo.* Frecuentemente las observaciones se efectúan, no sobre el conjunto completo de elementos a estudiar, sino sobre una parte de éstos que reciben el nombre de muestra. Las observaciones hechas sobre los elementos componentes de la muestra constituyen la serie estadística de la cual se determina el valor central. Es evidente que "a priori" no puede asegurarse que el valor central correspondiente a la muestra adoptada coincida con el valor central que se obtendría si se hiciese una serie estadística que abarcase todo el conjunto completo de elementos a estudiar, ni que coincidan siquiera con los correspondientes a distintas muestras que se eligiesen al azar. Ahora bien, dado que en la práctica se procede casi siempre por muestreo, conviene que la característica elegida del valor central sea de tal naturaleza que dicho valor central sea sensiblemente el mismo para las distintas muestras. (Conviene hacer notar que esta elección del valor central sólo será posible cuando se conozca la ley de distribución del fenómeno en estudio; la variación del valor central y de otros estadísticos en las distintas muestras entra de lleno en la parte de la Teoría Estadística conocida por la denominación de Teoría de las Muestras).

f) *Debe ser adecuada a los cálculos algebraicos posteriores.* Se comprende fácilmente la importancia de tal condición con sólo pensar en el caso muy frecuente de tratar de determinar el valor central que corresponde a una serie global resultado de reunir varias series estadísticas parciales.

De entre las cuatro medias expresadas, se ve inmediatamente que la aritmética es la que mejor reúne las anteriores condiciones de Yule, si bien ni ella ni las otras tres proporcionan indicación alguna acerca de la repartición de los datos de la serie o de sus posiciones respectivas ni sobre las desviaciones de unos respecto a otros. Se limitarán a condensar todos los datos de la serie en uno solo, **la media**, como síntesis de todos ellos.

En particular, las medias aritméticas (\bar{X}) y cuadrática (C) dan mucho relieve a los elementos grandes de la serie y, desde luego, la segunda todavía más que la primera. Por el contrario, las medias geométrica y armónica destacan la influencia de los valores pequeños y reducen la influencia de los valores grandes, lo que habrá que tener bien presente en los estudios de Análisis Territorial.

Recordemos, por último, que las medias deben calcularse a partir de datos homogéneos y numerosos, condiciones ambas inherentes a toda buena estadística en materias territoriales (FRANQUET, 1990/91).

