

Notas y comentarios

La región en el pensamiento geográfico anglosajón: notas metodológicas

Boris Graizbord*

El interés por la región se ha acentuado –casi podría decirse renacido– en los últimos años debido, paradójicamente, al proceso globalizador de la economía y a la conformación de bloques regionales pero, no menos, al deseo de hacer efectivo el federalismo o, bien, por el surgimiento de movimientos separatistas o regionalismos de carácter étnico o económico.¹

La región –que representa el tema central en el pensamiento geográfico– es, a pesar de la crítica conceptual y epistemológica al interior de la geografía, una forma lógica y satisfactoria para organizar información geográfica y tomar decisiones respecto a la localización espacial de las actividades humanas y la asignación de los recursos.

La clasificación de la superficie terrestre por regiones, con base en diferencias observadas, representa una descripción de la naturaleza de la vida social y la forma en que ésta se organiza y organiza el espacio geográfico para su funcionamiento y reproducción.

Los geógrafos han considerado tradicionalmente que existen categorías de regiones dependiendo del número de criterios utilizados para reconocerlas.² Así, por ejemplo, podríamos reconocer una región con base en un solo atributo natural o humano como el tipo de suelo, la vegetación o la principal actividad económica, la densidad de ocupación del suelo, el tipo de cultivo predominante, etcétera. En este caso nuestra óptica es estrecha y específica: una sola característica observable permite identificarla. Por el contrario, si analizamos la calidad de vida o el nivel de desarrollo, entraremos necesariamente en una gama más amplia de criterios que res-

* Profesor-Investigador del Centro de Estudios Demográficos y de Desarrollo Urbano de El Colegio de México.

¹ Un enfoque económico reciente puede verse en Krugman y Obsfield (1994); uno político en Ohmae (1993).

² Para algunos geógrafos, no necesariamente de óptica regionalista, la existencia objetiva de regiones es un hecho independiente de los propósitos y criterios subjetivos del analista u observador. Para efectos de esta presentación se acepta la *subjetividad* implícita en los métodos de regionalización presentados. Sobre la existencia independiente y objetiva de regiones en México, véase Bassols (1964). Para una regionalización de corte subjetivo utilizando, por cierto, una de las técnicas presentadas en este trabajo, véase Stern (1973).

ponden a situaciones más generales y complejas. Éste es generalmente el caso al que nos enfrentaremos como analistas y planificadores.

Las unidades geográficas que se caracterizan por sus múltiples atributos —y no se trata de la visión holística de los regionalistas— pueden, para efectos analíticos, clasificarse en:

1) *Regiones formales o uniformes*, en las cuales las áreas más pequeñas que las forman presentan una mínima diferencia entre ellas respecto del factor o conjunto de factores considerados. Esta similitud permite clasificarlas en *regiones homogéneas*; el criterio de pertenencia, por lo tanto, está dado por la “distancia” (varianza) mínima entre las unidades que forman la región y máxima con las unidades que no pertenecen a ésta.

2) *Regiones nodales o funcionales*. Son aquellas unidades geográficas que se organizan a partir de la relación que se establece entre elementos constitutivos de un área y un centro o núcleo dominante. Esta relación o interacción espacial que se manifiesta entre éste y aquéllos a partir de flujos de algún tipo, los constituye en regiones nodales o funcionales cuya base de pertenencia está dada por la intensidad de estas relaciones y el grado de organización funcional interna que el núcleo central mantiene con el área y los elementos circundantes.

En ambos casos, ya sea para diferenciar el espacio geográfico en áreas o para delimitar esferas de influencia de centros o nodos, se requiere de decisiones subjetivas que con la ayuda de técnicas cuantitativas es posible racionalizar y formalizar.

Existe otra clase de regiones que no siendo específicamente geográficas forman parte importante de la investigación del geógrafo. Se trata de las *regiones administrativas* que pueden subdividirse a su vez en: *i)* regiones político-administrativas de derecho como, por ejemplo, un país, estado, municipio, distrito electoral, fundo legal de una ciudad, etcétera, y *ii)* regiones administrativas *ad-hoc*, las cuales se constituyen con objeto de facilitar el cumplimiento de algún propósito, generalmente temporal como, por ejemplo, la planificación, la distribución de algún bien o servicio público, el levantamiento de información censal, etcétera.

Las tres categorías de regiones no son necesariamente excluyentes, pudiendo darse el caso de que una región nodal sea considerada como área homogénea para el levantamiento de algún tipo de información. Muchas veces, sin embargo, el constituirse como unidad organizadora de la información geográfica le confiere a un espacio o área un carácter homogéneo artificial, ya que se pierde información sobre las diferencias al interior de ella y se desconoce el tamaño de cada una de las unidades que la forman. Para el

analista y planificador esto representa un problema metodológico y práctico que debe tomar en consideración, pues es difícil casi siempre que coincidan las unidades geográficas de registro de la información con un problema geográfico y con las áreas relevantes para la implementación de las políticas. Esto se debe, entre otras cosas, a que la fragmentación del territorio en unidades político-administrativas no corresponde a la realidad de los fenómenos; o bien, porque no siempre las unidades espaciales de las que se tiene información delimitan con precisión el área del problema o donde éste incide o es más grave. Es probable que la población que requiere atención esté distribuida en dos o más áreas, lo que haría que la gravedad del problema se diluyera u ocultara ya sea porque la información se obtiene por separado o porque ésta se maneja de forma agregada, dando lugar a la denominada “falacia ecológica” y, en el caso de la planificación y la formulación de las políticas, a errores de diagnóstico o de estrategia (Graizbord, 1990).

Otros problemas que se deben enfrentar en el análisis regional se refieren a la calidad de la información con que se cuenta y al grado en que las variables o indicadores utilizados reflejan realmente las condiciones o situaciones que se quieren conocer. Al regionalizar o adoptar una regionalización será necesario, por lo tanto, aclarar con precisión el propósito del estudio y tomar en cuenta el carácter de la información, es decir, atender el *qué, cómo y para qué*.

Conceptualmente, podemos pensar que los indicadores a tomarse en cuenta en una regionalización socioeconómica pueden clasificarse de distintas maneras, independientemente de los aspectos o variables que se quieren representar y conocer, como serían vivienda, salud, educación, empleo, ingreso, estructura demográfica, participación política, etcétera. Así, por su naturaleza, podemos referirnos a indicadores *absolutos*, *relativos* y *autónomos*. Los primeros exigen un conocimiento científico –aceptado– de los valores que deben alcanzar los distintos aspectos socioeconómicos de la población. Por ejemplo: consumo mínimo de calorías por día y *per cápita*, metros cuadrados por habitante, proteínas según la edad, etcétera. Los indicadores relativos son aquellos que miden las condiciones de un grupo en un área respecto a otro en esa misma o en otra área. Esto implica, desde luego, aceptar la relatividad de los niveles socioeconómicos en un determinado momento y lugar. Implica, además, reconocer la necesidad inevitable de introducir juicios de valor acerca de situaciones sociales particulares. Por último, los indicadores autónomos son aquellos que pueden utilizarse para conocer condiciones específicas de la población de un área que sirvan a un propósito particular del análisis.

De acuerdo con el carácter, intención o alcances del estudio para el cual van a utilizarse, podemos clasificar los indicadores en:

i) Informativos: son aquellos esencialmente descriptivos de condiciones particulares en un momento dado. Requieren de una decisión *a priori* acerca de cuáles son los aspectos (socioeconómicos, físicos, biológicos, etcétera) que se desea o que es necesario medir.

ii) Predictivos: son aquellos vinculados a un modelo o teoría del cambio social que permiten inferir, a través de variaciones en ellos, los cambios en los procesos social y (o) espacial que interesan.

iii) Específicos: constituyen el equivalente cuantitativo de problemas particulares (v.gr. vivienda deficitaria, acceso a la educación secundaria, estado nutricional de un grupo etario, etcétera) y la base de las decisiones políticas, puesto que identifican los problemas sociales discretamente en el tiempo y el espacio.

iv) Evaluativos: son aquellos que permiten el seguimiento, progreso y efectividad de programas y políticas al representar cuantitativamente las condiciones existentes en relación con las metas y la situación inicial.

Conviene aclarar que estos indicadores no se excluyen entre sí y que, más bien, permiten construir convenientemente un sistema de información, de monitoreo y evaluación de condiciones (socioeconómicas) y cambios logrados a través de la aplicación de políticas sociales, económicas y espaciales.

En los cuadros 1 y 2 se reproducen, a guisa de ejemplo, dos listados de indicadores socioeconómicos utilizados para propósitos semejantes por una misma fuente (véanse los cuadros 1 y 2).

En el caso de las regiones nodales o funcionales, las variables e indicadores para delimitarlas se refieren a aspectos de circulación e intercambio de personas, dinero, mercancías, información o mensajes e ideas. En ocasiones, como en el caso del análisis de la difusión de innovaciones o enfermedades, por ejemplo, se estudia la evolución temporal del fenómeno en el espacio geográfico para reconocer el área de influencia y la fuente (nodal o zonal) del fenómeno. En otras ocasiones, se identifica el lugar de residencia de los consumidores, usuarios o suscriptores de algún bien o servicio localizado en un punto para conocer su alcance y su área tributaria. Graizbord y Garrocho, entre otros (véanse referencias bibliográficas), han empleado los datos de Teléfonos de México sobre el origen y destino de las llamadas telefónicas para delimitar regiones funcionales y reconocer o identificar la estructura nodal y organización jerárquica de sistemas urbanos (nacional, regionales y estatales).

CUADRO 1
Índice general de nivel de vida (Naciones Unidas)

<i>Tipo de necesidad</i>	<i>Componente</i>	<i>Indicador</i>
A) Física	1. Nutrición	a. Calorías <i>per cápita</i> b. Proteínas <i>per cápita</i> c. Proporción de calorías derivadas de cereales, raíces, azúcares, etcétera
	2. Habitación	a. Calidad de vivienda b. Densidad de ocupación c. Viviendas unifamiliares
	3. Salud	a. Acceso a servicios de salud b. Mortalidad por causas c. Tasa de mortalidad
B) Cultural	4. Educación	a. Índice de inscripción b. Eficiencia terminal c. Relación maestro-alumno
	5. Esparcimiento y recreación	a. Tiempo promedio de ocio b. Circulación de periódicos c. Número de radios o televisores
	6. Seguridad	a. Incidencia de mujeres violadas b. Porcentaje de población con seguro de desempleo c. Porcentaje de población pensionada
C) Conspicua	7. Ingreso excedente	a. Excedente del ingreso familiar una vez cubiertas las necesidades físicas y culturales básicas

Fuente: Drewnowski and Scott (1968), "The Level of living Index", *Ekistics*, núm. 25, pp. 266-275, reproducido en Knox (1975).

En resumen, como problema puramente taxonómico la clasificación de regiones representa en su aspecto geográfico-espacial lo que una clasificación significaría en cualquiera de las ciencias naturales. De tal forma que una "región de características únicas" representa una "categoría"; un "límite regional" equivale a un "intervalo entre clases" y una "región homogénea" a una "clase con mínima varianza interna". Sin embargo, esta lógica deja fuera un aspecto sin resolver: la localización única de cada región. En otras palabras, el hecho de que dos unidades regionales pudieran pertenecer a una misma clase y estar localizadas en latitudes apartadas como, por ejemplo, una en el sur del país y otra en el norte, lo que, si bien las hace diferentes en muchos aspectos, las constituye en elementos de un sistema migratorio en el que ambas forman

CUADRO 2
Índice sistemático de nivel de vida (Naciones Unidas)

<i>Variable</i>	<i>Aspecto básico</i>	<i>Indicador</i>
A. Vivienda	Características de la habitación	1. Densidad ocupacional
		2. Hacinamiento
		3. Vivienda compartida
		4. Vivienda vacía
		5. Hogares pequeños
		6. Hogares grandes
		7. Vivienda propia
		8. Vivienda rentada
		9. Vivienda pública
		10. Valor catastral
	Calidad de la vivienda	11. Agua dentro de la vivienda
		12. Agua caliente
		13. Acceso a W.C.
		14. Baño dentro de la vivienda
		15. Total de nuevas viviendas
Construcción de vivienda	16. Viviendas públicas construidas	
	17. Mortalidad infantil	
B. Salud	Enfermedad	18. Mortalidad por bronquitis
		19. Mortalidad por tuberculosis
	Acceso a servicios médicos	20. Número de médicos generales
C. Educación	Amplitud de la educación	21. Dentista escolar
		22. Adultos que abandonaron la escuela antes de los 15 años
	Ambiente de la enseñanza	23. Alumnos registrados
		24. Maestros de educación media superior por alumno
		25. Maestro de educación media por alumno
D. Estatus social	Estatus social	26. Trabajadores profesionales
E. Empleo	Uso de la mano de obra	27. Trabajadores manuales
		28. Desempleados
F. Riqueza	Ingreso monetario	29. Población económicamente activa
		30. Tasa de participación femenina
	Consumo	31. Ingreso personal mínimo
		32. Ingreso personal elevado
G. Recreación	Acceso a servicios	33. Comercio al menudeo
		34. Personas propietarios de autos
		35. Bibliotecas
		36. Cines
		37. Restaurantes
H. Seguridad social	Seguridad	38. Hoteles
		39. Servicios de seguridad pública
	Bienestar	40. Servicios asistenciales
		41. Servicios para infantes
I. Estabilidad social	Crimen	42. Delitos perseguibles
	Estabilidad familiar	43. Divorcios
J. Estructura demográfica	Edad y ciclo de vida	44. Parejas en unión libre
		45. Grupo de edad 0-14
		46. Grupo de edad 15-44
K. Medio ambiente físico	Cambios poblacionales y movilidad	47. Grupo de edad 60 y más
		48. Tasa de crecimiento natural
		49. Saldo migratorio
		50. Cambio de residencia
L. Participación política	Urbanismo	51. Personas por hectárea
		52. Curules de oposición
	Participación democrática	53. Tamaño del padrón electoral

Fuente: P. L. Knox (1974), "Spatial variations in levels of living in England in Wales in 1961", *твс*, núm. 62, 1-24. Repr. en Knox (1975).

parte del mismo fenómeno.

Esta última reflexión –que no deja de ser un tema polémico entre geógrafos– no elimina, sin embargo, la necesidad de clasificar, por una parte, unidades espaciales uniformes u homogéneas que requieren alguna forma de trato igual o selectivo y discriminatorio (Graizbord, 1990 basado en Papageorgiou, 1984); y, por otra, identificar una región con base en criterios funcionales, utilizando variables como migración, viajes al trabajo o de compras, circulación de periódicos, frecuencia de corridas, etcétera. En este caso enfrentamos el problema de una demarcación o frontera distintas para cada criterio o variable que se usa, lo cual genera espacios difusos y límites no resueltos, e introduce un elemento de incertidumbre que implica asumir que cada indicador tiene el mismo peso como determinante de la región o bien, en su caso, aceptar sólo una de las delimitaciones de la región tomando en cuenta la base metodológica o sistemática del procedimiento, no menos que el propósito específico del estudio.

A continuación se presentan algunos métodos o técnicas utilizadas en la regionalización socioeconómica que responden a estas consideraciones y, a su vez, representan las dos clases de región mencionadas anteriormente: región homogénea y región nodal. En otras palabras, distinguimos metodológicamente entre: *i) la asignación y desagregación de áreas o unidades de información geográfica a unidades, sistemas o conjuntos regionales homogéneos, y ii) la identificación o delimitación de regiones funcionales.*

En el primer caso (*i*) se trata de asignar áreas a una clase de acuerdo con una o varias características, si bien al tomar en cuenta la contigüidad se trataría de una regionalización propiamente, mientras que si no lo hacemos estaríamos hablando sólo de una clasificación o tipología de unidades geográficas. Para el segundo grupo (*ii*) se hace una distinción entre región nodal y región funcional, propiamente. Cabe decir que éstas no agotan las técnicas posibles para la identificación o construcción estadística y funcional de regiones, pero se espera que sirvan para alentar su uso.

Regiones homogéneas

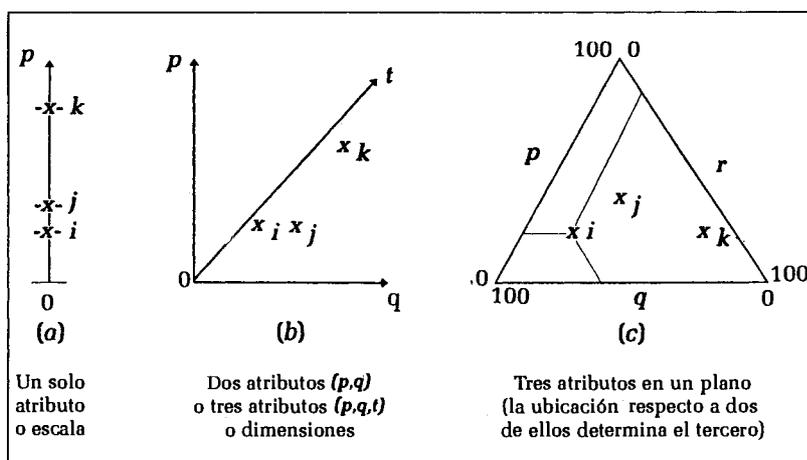
1) La clasificación de áreas basada en un solo criterio no presenta ningún problema: requiere solamente de decisiones acerca del número de clases y de los valores o rangos que van a utilizarse para determinar los límites de cada clase. Estos pueden resultar obvios al sólo revisar la distribución de la variable representativa según el criterio utilizado; o bien derivarse de un “*corte*” estadístico

(v.gr. los valores menores y los mayores a la media, una desviación estándar, etcétera) o simplemente pueden ser arbitrarios. Con el resultado se puede construir un mapa de coropletas que, de hecho, representa una diferenciación geográfica a partir de valores semejantes o diferentes respecto a una variable, independientemente de que se incorpore el criterio de contigüidad de las unidades espaciales para entonces sí hablar de regionalización y no sólo de clasificación.

2) Si lo que se desea es una clasificación basada en dos o tres aspectos o atributos, el uso de gráficos con la escala apropiada para los valores absolutos o relativos de las variables representativas es suficiente (figura 1).

FIGURA 1

Posición de tres observaciones (i, j, k) en una dos y tres dimensiones o escalas



Fuente: adaptado de Haggett (1965).

Al ubicar nuestras observaciones en la escala (a) o en las (b) o (c) según se trate de uno, dos o tres atributos o variables utilizados, podremos identificar cuáles unidades se asemejan o están más cercanas unas de otras. Esto no siempre es posible, especialmente cuando son muchas las unidades u observaciones, o cuando hay una gran semejanza en los valores de todas o algunas de ellas. El ejemplo muestra que la *cercanía* o proximidad entre los

valores de las unidades observadas que se ubican en el plano (la escala lineal, ortogonal o triangular) de las gráficas sirve como criterio relevante para delimitar grupos o clases. En las tres gráficas, las observaciones o áreas i y j constituyen de entrada una clase, dada la proximidad o cercanía mínima entre ellas comparada con la distancia respecto de k .

3) Al introducir uno o varios atributos o aspectos adicionales, el método convencional para una o dos dimensiones deja de ser útil. Sin embargo, con base en el mismo criterio de proximidad, o sea, en el análisis de la *distancia* entre observaciones, se cuenta con un método general para analizar semejanzas y diferencias en un espacio n -dimensional, formalizado en la siguiente expresión:

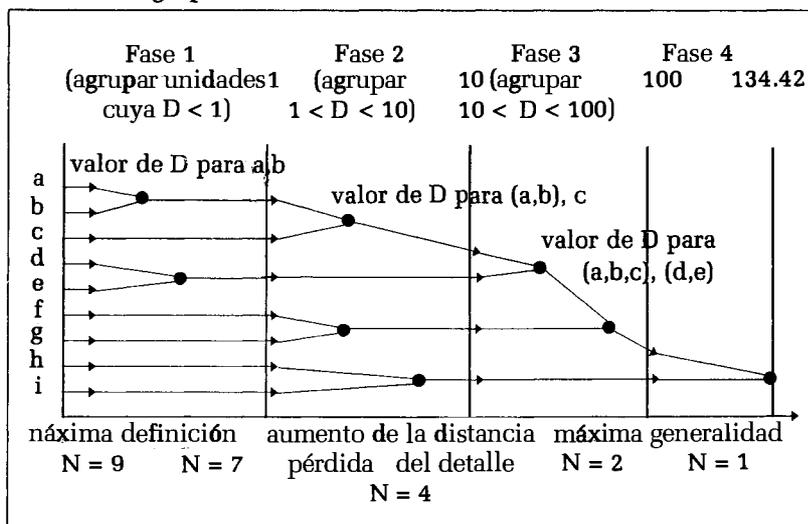
$$D (\text{distancia}) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

en donde la mínima distancia representa similitud entre cada par de observaciones x, y ; mientras que x_i, y_i son los valores respectivos para la característica i . El método tiene por objeto agrupar en un espacio de n dimensiones a las unidades espaciales de observación en función de su cercanía entre ellas y su lejanía con las demás unidades, conjuntos o grupos que van formándose simultánea o consecutivamente. El proceso permite ir ganando en generalidad a la vez que se pierde en definición. El número de pasos dependerá de los propósitos que se tengan para llegar a un número eficiente de grupos (o regiones). Este proceso por fases se representa en la figura 2.

4) Cuando se *construyen* regiones a partir de unidades o áreas pequeñas existen dos fuentes de observación: una variación entre y una variación al interior de cada una de ellas.

Para medir estas variaciones intra e inter-regiones se ha utilizado el *análisis de varianza*. La varianza representa el total de variación de un conjunto de observaciones y es igual a la suma de cuadrados de las diferencias de los valores respecto a la media aritmética de todas las observaciones. La varianza *entre* regiones nos da la variación de las regiones alrededor de la media total, mientras que la *intra* nos indica la variación de las unidades alrededor de la media de la región a la que pertenecen. Para comprobar la validez de una regionalización a base de un agrupamiento de unidades semejantes, se utiliza el factor que resulta de dividir la varianza mayor entre la menor. Este método se ha empleado,

FIGURA 2

Proceso de agrupación de unidades u observaciones

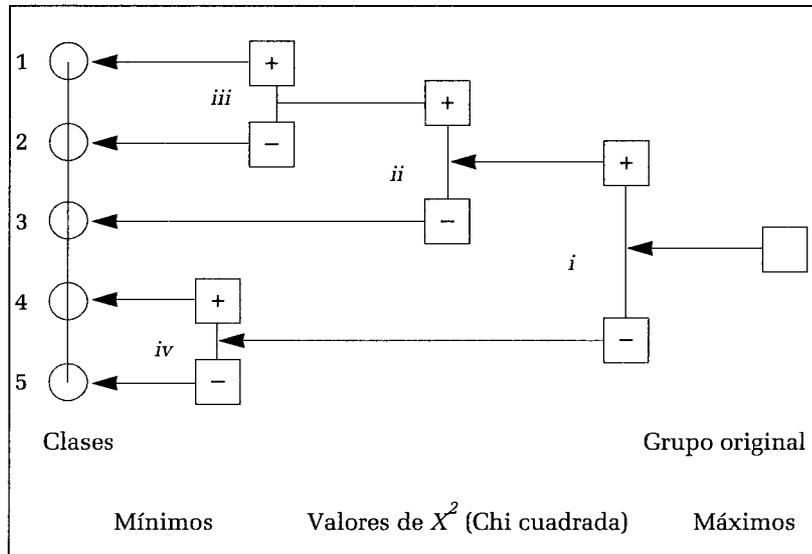
Fuente: adaptado de Haggett (1965).

asimismo, para resolver el problema que se presenta al tener que asignar unidades a una de tantas regiones, comparando la *distancia* de la unidad a cada una de las regiones a partir del valor de la variación al interior de cada una de éstas, es decir, de comprobar que la distancia entre la unidad y la región (o conjunto) a la que se asigna es menor o similar a la distancia que permitió agrupar a las últimas unidades incluidas en ella.

5) Por el contrario, desagregar o subdividir un espacio o área en varias categorías o subregiones resulta un caso distinto al de asignar y agrupar áreas para formar una región. Se trata en este caso de separar unidades con base en un indicador de asociación regional. El empleo de la Chi cuadrada (X^2) permite precisamente subdividir la unidad o unidades subsecuentes de manera sucesiva.

Como se indica en la figura 3, gráficamente se tendría la siguiente secuencia en el proceso de subdivisión hipotética, de acuerdo con la *presencia* o *ausencia* de un indicador. Como puede apreciarse, a partir de una región única se van haciendo cortes sucesivos según los valores mayores de Chi cuadrada (X^2) que en cada fase van disminuyendo. El proceso se detiene en el momento en que el número de subdivisiones sea aceptable o cumpla con

FIGURA 3
Desagregación por etapas de acuerdo con presencia-ausencia de atributos (i a iv)



Fuente: adaptado de Haggett (1965).

criterios establecidos previamente.

6) Por último, lo que no necesariamente agota todas las técnicas de clasificación que puedan ser utilizadas en una regionalización homogénea, se presenta un método estadístico más complejo: el análisis de factores o de componentes principales.

Este procedimiento se ha utilizado en sus distintas modalidades desde los años sesenta para la identificación de sistemas regionales en donde se estima necesario considerar un número relativamente elevado de variables o criterios para un gran número de observaciones. Se trata, como en casos anteriores, de una forma de medir correlaciones, es decir, asociaciones entre valores de las variables representativas, lo cual permite agrupar en *regiones* lo más homogéneas posibles a las unidades de observación geográfica.

Este enfoque multifactorial incluye dos fases principales: la primera, que permite reducir la información a proporciones manejables, extrayendo de ella un número pequeño de factores (o componentes) significativos que representan variables compuestas que toman en cuenta la amplia variación de la información

original. Esto es posible ya que seguramente algunas variables pueden decir lo mismo que todo el conjunto de ellas acerca de las diferencias entre las unidades espaciales consideradas. En la segunda fase, una vez comprimida la información, y al darse a cada unidad espacial una calificación respecto a los factores, se procede a agruparlas en clases o regiones a partir de un análisis de "conglomerados (*cluster analysis*). Aquí cabría preguntarse: ¿cuántos criterios y variables son relevantes?, ¿cuántos casos o unidades espaciales deben incluirse? Ambas preguntas tienen respuesta condicionada que dependerá, entre otras cosas, del propósito de la regionalización, de la escala, la forma y el nivel en que se presenta y obtiene la información, el grado de variación de cada variable, etcétera.

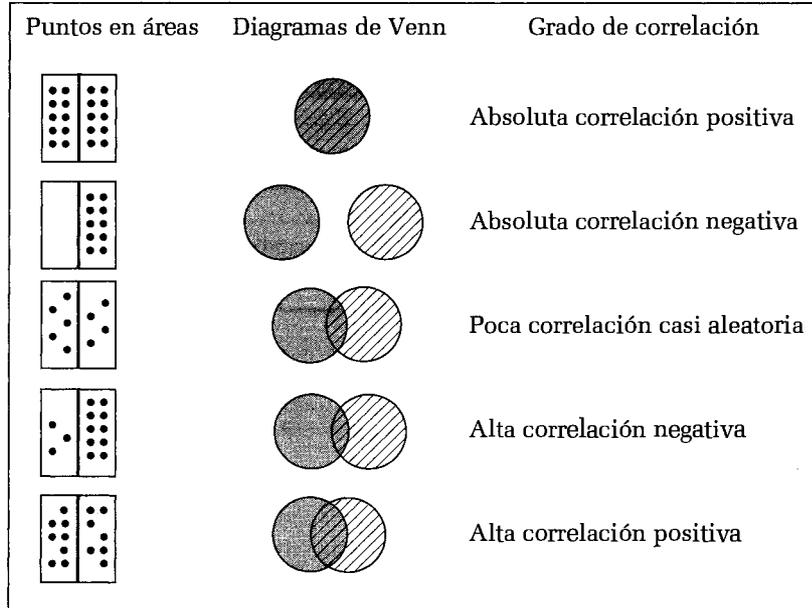
Al tener acceso a un programa de computadora, nuestra capacidad para incluir un número elevado de variables o atributos aumenta extraordinariamente. Sin embargo, ¿cuáles son las relevantes o cuáles tienen sentido en un determinado contexto? En términos generales se puede decir que nos enfrentamos al problema de rechazar o aceptar una hipótesis y, en este caso, se trata de utilizar estos instrumentos analíticos para hacer inferencias de carácter causal, o bien, no teniendo hipótesis establecidas, tratar abiertamente de identificar relaciones entre variables o entre unidades espaciales o entre ambas. En efecto, se podría pensar en relaciones entre: *a)* aspectos físicos y aspectos físicos; *b)* aspectos físicos y aspectos humanos; *c)* aspectos humanos y aspectos físicos y, finalmente, entre *d)* aspectos humanos y aspectos humanos.

Esta idea de relacionar objetos, atributos, eventos, individuos o conjuntos de ellos es una experiencia cotidiana. Ha sido parte fundamental del trabajo del geógrafo. Para éste se trata más bien de encontrar relaciones potenciales entre atributos que inciden en un lugar o se encuentran próximos en un espacio geográfico. La afirmación cuantitativa acerca de la co-ocurrencia y co-variación de valores asignados a pares de objetos o atributos, que como datos numéricos se denominan variables, es implícita al hablar de correlación.

Por la complejidad de ciertas situaciones geográficas es necesario considerar más de dos variables o conjuntos de variables. La relación entre pares de variables o de conjuntos de variables se representa numéricamente como el *coeficiente de correlación*, al cual se le puede dar un carácter descriptivo o bien explicativo e inferencial.

Véamos gráficamente (figura 4) distintos grados de correlación (positiva y negativa) entre dos atributos representados por puntos y líneas en áreas en un plano (dos dimensiones), utilizando diagramas Venn.

FIGURA 4
Correlación



Es necesario señalar que un índice de correlación, que indica una débil o nula relación entre áreas y atributos, es tan importante como uno muy elevado puesto que el primero podría interpretarse como una independencia absoluta en la relación entre dos variables o áreas, y llevar, incluso, al abandono de una hipótesis o vía para analizar un problema.

A continuación, de manera sucinta, se describe el método de componentes principales o análisis multifactorial que se ha utilizado como método de regionalización en aquellos casos en que, como ya se dijo, la complejidad de la situación requiere de la consideración de múltiples atributos, aspectos, o variables para un número considerable de observaciones.

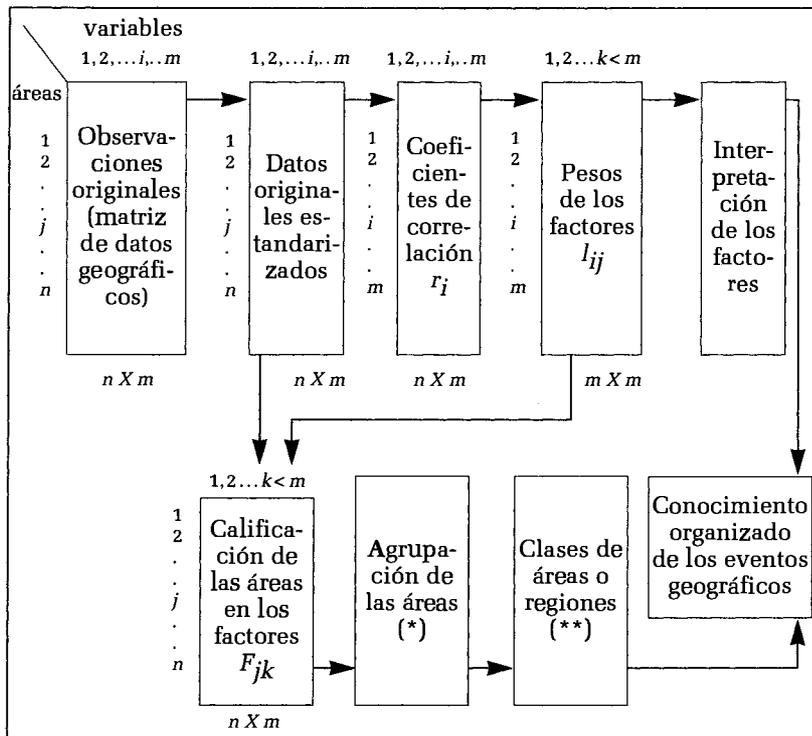
El análisis de componentes principales se diferencia del análisis factorial en que el primero reemplaza a las variables originales por un nuevo conjunto de igual número de componentes no correlacionados entre sí; los principales de ellos contienen o explican una parte relativamente grande de la varianza original. El análisis factorial hace básicamente lo mismo, excepto que extrae del número original de variables un número reducido de factores, dejando parte de la varianza en un término llamado error. Toman-

do en cuenta que cada variable se distribuye en el conjunto de áreas consideradas y que cada una de ellas representa una dimensión, una matriz de tamaño m variables por n observaciones (o áreas) representaría un espacio de m dimensiones. En el análisis, cada variable contribuye a cargar a cada factor o componente. Éstos, a su vez, permitirán agrupar a las áreas en función de la calificación que cada una de ellas obtenga respecto al o los factores.

Los pasos a seguir en este proceso de regionalización utilizando el análisis de factores (incluido el método de agrupación o análisis de conglomerado), se pueden ver más claro en la figura 5.

El análisis multifactorial ofrece básicamente dos posibilidades:

FIGURA 5
Análisis factorial



Notas:

(*) Los procedimientos de agrupación pueden ser el de mínimas distancias en una, dos, tres o n dimensiones. Este último caso lo resuelve el programa *cluster analysis*.

(**) Dependiendo de si se consideró o no la contigüidad de las áreas para agruparlas. Véase Spence y Taylor (1970).

así como podemos registrar nuestra información en una matriz $n \times m$ en la que n = número de áreas y m = número de variables, procediendo a encontrar los coeficientes de correlación entre las variables en una tabla $m \times m$, podemos también encontrar coeficientes de correlación entre lugares en una tabla $n \times n$, derivando de ellas factores. El análisis (referido como modalidad R o Q) producirá dimensiones o factores que corresponden al patrón que caracteriza a las áreas en un espacio multidimensional de m variables. En este caso, la correlación mide el grado de correspondencia, no entre variables respecto a cada área, sino entre lugares en términos de cada variable, lo que permite identificar conjuntos de lugares en un espacio de variables, en vez de grupos de variables en un espacio de n lugares.

Las dos modalidades (R y Q) han sido utilizadas para definir regiones homogéneas. Ahora bien, cabe reconocer que el análisis es estático en el sentido de que el patrón resultante no hace referencia a los procesos que lo originan –si construimos una matriz de flujos entre áreas nos estaríamos refiriendo al análisis de un proceso: la interacción espacial. La representación matricial de esta interacción daría por resultado una tabla cuadrada de orígenes y destinos (O-D) en el que cada cruce representaría el volumen del flujo considerado que tiene origen en i y destino en j . Trasformar esta tabla en una matriz de correlación entre el O y el D permitiría obtener valores de los coeficientes que indicarían una similitud del origen de los flujos para cada par de destinos. En estos términos, el análisis factorial a partir de una matriz O-D producirá conjuntos de áreas semejantes en función del origen del flujo que reciben. Los factores (la carga en los factores) indican la correlación de las áreas con ese factor y la calificación de cada una con el factor las identifica respecto al origen más importante del flujo que llega a ellas. Así, podríamos identificar regiones productoras de flujos y regiones receptoras de éstos, con implicaciones económicas y funcionales en términos de exportación-importación o especialización-complementariedad, aspectos fundamentales del Análisis Espacial y la Economía Regional.

Regiones nodales o funcionales

Antes de presentar algunas de las técnicas empleadas en la identificación de estas regiones formadas por entidades espaciales o locacionales (puntos o nodos) que presentan una intensa interacción o conexión entre ellas, convendría hacer la distinción entre región nodal y región funcional.

La diferenciación del espacio geográfico a partir de la actividad humana resulta en un sistema de regiones o áreas que se distinguen por una particular característica que las hace homogéneas o les da cohesión funcional, es decir, las hace funcionalmente complementarias. Podríamos imaginar en este último caso un paisaje estructurado a partir de una constelación de nodos, redes y áreas tributarias.

Partimos para la identificación de este tipo de regiones del hecho de que las sociedades organizan y llevan a cabo sus actividades en localizaciones específicas o puntuales. La producción, el lugar de residencia, los asentamientos, etcétera, forman entidades locacionales puntuales o nodales que organizan la vida y las actividades de la sociedad alrededor de ellas. Un nodo es, pues, un centro de actividad de un grupo o colectividad que ocupa un área geográfica. De tal forma que una región nodal es aquella que se identifica por el grado de cohesión que la población de los alrededores establece con el centro. Una región nodal es un caso especial de región funcional en la que sólo hay un punto central o focal que introduce la noción de orden y dominación. Si la identificación se basa en la interacción entre entidades locacionales y la relación jerarquizada entre ellas, identificándose un centro que *domina* a todas las entidades que interactúan entre sí, se tiene una región nodal o polarizada, como en ocasiones se les llama. Por el contrario, una región funcional es aquella en la que las entidades que interactúan lo hacen con mayor intensidad al interior del área que con otras áreas, independientemente del rol y jerarquía de los nodos que resultara de ese patrón de interacción. Es importante en esta distinción que se hace entre región nodal y región funcional el hecho de que las relaciones interlocacionales no son simétricas y que el patrón de interacción es independiente de la importancia relativa de un centro respecto a los demás.

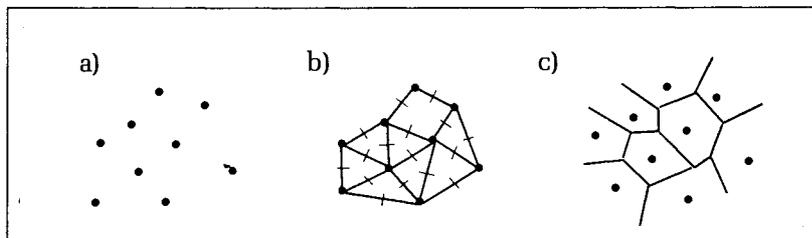
Debe quedar claro que una región funcional puede estar formada por un conjunto de regiones nodales y, a su vez, una región nodal o polarizada puede contener subgrupos de entidades locacionales que forman regiones funcionales y pueden interactuar con el centro dominante a través del intercambio de flujos sean mensajes, personas, dinero, mercancía, etcétera.

Como se mencionó antes, la actividad que se genera a partir de un centro o polo y su área tributaria, permite delimitar una región nodal. Cabe señalar que de cada flujo específico que se considere se obtendrá una delimitación o frontera que superpuesta a otras obtenidas a partir de los demás indicadores generaría una zona incierta o difusa de líneas no coincidentes. Esta incertidumbre respecto a la línea demarcatoria de una región nodal se

convertiría en problema si de lo que se trata es de racionalizar la asignación de recursos y de conocer precisamente el mercado existente en la región circundante o la localización puntual de la oferta de un bien o servicio producido y distribuido pública o privadamente. No cabe duda, pues, de la importancia de definir claramente la línea demarcatoria del área de influencia de centros o nodos contiguos; para tal efecto existe un gran número de técnicas no explotadas aún suficientemente por los geógrafos. A continuación se presentan algunas de ellas.

1) *Polígonos Thiessen*. Esta técnica geométrica permite trazar fronteras entre puntos o centros distribuidos en un plano de dos dimensiones. La validez del método se fundamenta en dos supuestos: primero —y ésta es una simple propiedad geométrica de los polígonos—, que cualquier punto en el área que queda delimitada al interior de las intersecciones lineales del polígono es más cercano al centro del mismo que a cualquier otro centro de otro polígono y, segundo, que no existe una jerarquía de los centros distribuidos espacialmente, es decir, que cada centro “pesa” igual que los demás, que el área circundante es homogénea y que, en términos agregados, el espacio se distribuye equitativamente respecto a la posición original de los centros en el plano. Los polígonos Thiessen producen un conjunto teóricamente óptimo de centros y sus respectivas áreas tributarias. Es relevante señalar que de una distribución triangular regular de centros se obtendrían polígonos hexagonales *a la Christaller*. Los pasos para formar los polígonos se muestran en la figura 6.

FIGURA 6

Polígonos Thiessen

Fuente: adaptado de Haggett (1965).

Dibujar polígonos en la práctica, aunque es muy simple, consume mucho tiempo, especialmente si es elevado el número de centros. Además, no es muy precisa la delimitación de las diagonales en el trazo, por lo que éste puede hacerse con arcos de círculo

desde puntos adyacentes que permiten encontrar en su intersección el punto medio por el que deben pasar los lados del polígono. Un algoritmo basado en los triángulos pitagóricos lo resolveríamos rápidamente si contamos con una PC. En fin, en el trazo no entra en consideración la importancia de cada punto o nodo y así, no puede distinguirse su influencia en el conjunto.

2) *Minimización de la distancia.* Con los conceptos de optimización de la programación lineal se ha desarrollado un método de delimitación de regiones que permite minimizar el costo-distancia a recorrer desde los puntos dentro de la línea demarcatoria al centro o nodo correspondiente. De manera formal, se tiene una función sujeta a un número de restricciones que puede expresarse algebraicamente como:

$$\sum_{i=1}^n \sum_{j=1}^m d_{ij} X_{ij} = \text{mínimo}$$

en donde d_{ij} es la distancia de la sección o punto i al centro j y x_{ij} es el costo del movimiento desde esa sección o punto i al centro j .

Hay varios ejemplos en que se aplica este método para optimizar áreas tributarias de un servicio localizado puntualmente. Por ejemplo, cuando se trata de minimizar el recorrido agregado de los niños hacia varias escuelas distribuidas en una zona de la ciudad, la función objetivo sería optimizar el área tributaria de cada una de las escuelas, de tal forma que *i)* la distancia agregada de recorrido por todos los niños se minimizara, y *ii)* cada escuela maximizara su captación de niños.³

3) *La identificación de regiones nodales a partir del análisis de la interacción espacial* puede hacerse con base en dos tipos de información y dos métodos.

a) La estructura nodal derivada de una tabla de flujos. Una estructura regional puede definirse a partir de *i)* un conjunto de centros en un área, y *ii)* una medida de asociación entre ellos. Esta asociación puede medirse a partir del flujo o interacción espacial entre pares de nodos (o bien con base en el hecho de que estén o no conectados entre sí). En la tabulación es posible identificar las relaciones sobresalientes entre pares de centros lo que a su vez permite la identificación de relaciones funcionales dependientes entre los mismos. La estructura jerarquizada a partir

³ Los detalles de la aplicación del método pueden verse en Yeates (1963) y Maxfield (1972), entre otros.

de esas relaciones indicará las regiones que se forman alrededor de los centros dominantes. El grado de dependencia relativa de los centros al centro dominante estará dado por el volumen del flujo (o el número efectivo de conexiones), lo cual permite conocer la fuerza relativa de los vínculos regionales (la multiplicidad y complejidad de la red), siendo este conocimiento útil, sin duda, para toda gestión administrativa de carácter territorial (Kansky, 1963).

b) *Los modelos gravitacionales.* En 1931, Reilly ya había identificado dos reglas básicas acerca del comportamiento espacial del comercio al menudeo: *i)* mientras más grande fuese un centro más comercio atraería, y *ii)* un centro capta mayor volumen del flujo que le llega de los lugares más cercanos que de los más alejados. Ambas reglas observadas empíricamente dieron por resultado la llamada ley de la gravitación del mercado al menudeo, como analogía a la ley gravitacional del sistema newtoniano. En el lenguaje algebraico el modelo puede escribirse como:

$$\frac{b_{ix}}{b_{iy}} = \frac{P_x}{P_y} \left(\frac{d_{iy}}{d_{ix}} \right)^2$$

siendo b_{ix} y b_{iy} el volumen del comercio que los centros x y y captan de un sitio o área i ; P_x y P_y representan el peso o la población de x y y ; d_{ix} y d_{iy} son las distancias respectivas a i .

Esta formulación simple y lógicamente plausible nos dice que "dos centros atraen comercio de un tercero intermedio en proporción directa a sus poblaciones e inversamente al cuadrado de sus distancias respectivas con aquél". De tal forma que, por ejemplo: dos centros x , y de población igual y a una distancia también igual a un tercero i recibirán el mismo volumen de comercio (interacción) que sale de este último. Así, por ejemplo, si P_x es dos veces más grande que P_y , pero las distancias a i son iguales, x recibirá el doble del comercio que y .

Debe quedar claro que es posible, aplicando una ligera variante de este modelo, encontrar los límites del área tributaria de cada centro; ésta se definiría como el punto en la línea que une dos centros en el que ambos logran atraer la misma proporción de la interacción o flujo que ahí se genera.

Es decir que $\frac{b_{ix}}{b_{iy}} = 1$ y por tanto, en la primera expresión:

$$\frac{P_x}{P_y} \left(\frac{d_{iy}}{d_{ix}} \right)^2 = 1$$

si $d_{ix} = d_{xy} - d_{iy}$, representando d_{xy} la distancia entre x y y tenemos:

$$\frac{P_x}{P_y} \left(\frac{d_{iy}}{d_{xy} - d_{iy}} \right)^2 = 1$$

$$\left(\frac{d_{iy}}{d_{xy} - d_{iy}} \right) = 1 + \sqrt{\frac{P_y}{P_x}} \text{ o } d_{iy} = \frac{d_{xy}}{1 + \sqrt{\frac{P_x}{P_y}}}$$

Es decir, que la distancia de y a i es el punto de inflexión en el que la interacción con x y con y es la misma [o en el que la probabilidad de generación de flujo desde i es la misma (0.5) hacia ambos centros] y, por lo tanto, es el límite de las áreas tributarias de x y de y .

Un ejemplo numérico permitirá ver claramente el método:

$$P_x = 50\,000 \text{ habitantes}$$

$$P_y = 200\,000 \text{ habitantes}$$

$$d_{xy} = 20 \text{ km}$$

$$d_{iy} = \frac{20}{1 + \sqrt{\frac{50000}{200000}}} = \frac{20}{1.5} = 13.3 \text{ km}$$

Es decir que el límite del área tributaria de y respecto a x se encuentra a 13.3 km desde y , y a 6.7 km desde x . Este punto de inflexión puede encontrarse para un centro x respecto a varios que le rodean. Veamos el siguiente ejemplo numérico:

nodos	habitantes	distancia x y en km
x	50 000	
y_2	50 000	18
y_3	100 000	10
y_4	200 000	40
y_5	30 000	10
y_1	200 000	20 (conocida en el ejemplo anterior)

Para cada par xy existe un punto de inflexión i (figura 7). Éste permite conocer las distancias d_{iy} que equivalen a una interacción entre x y con probabilidad 50% las cuales *delimitan* el área de influencia o área de mercado tributario de x respecto a los nodos alternos y .

$$\text{Así, } d_{iy_2} = \frac{18}{1 + \sqrt{\frac{50000}{50000}}} = \frac{18}{2} = 9 \text{ km}$$

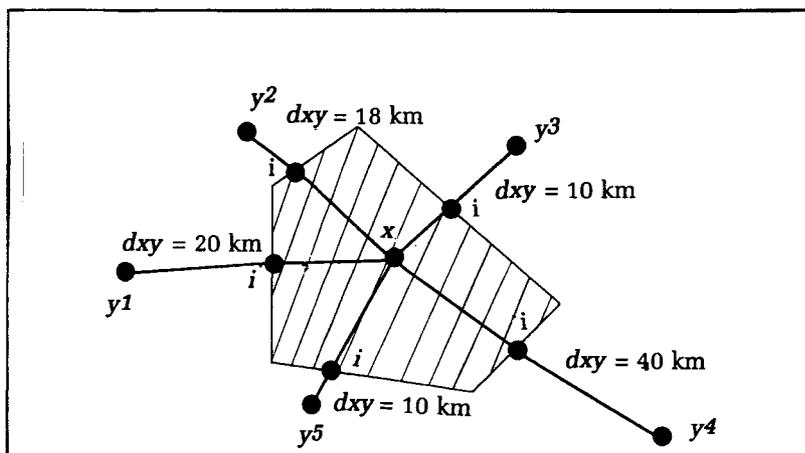
$$d_{iy_3} = \frac{10}{1 + \sqrt{\frac{50000}{100000}}} = \frac{10}{1.7} = 5.9 \text{ km}$$

$$d_{i_4} = 26.7 \text{ km}$$

$$d_{i_5} = 4.4 \text{ km}$$

$$d_{i_1} = 13.3 \text{ km conocido del ejemplo anterior.}$$

FIGURA 7
Límite del área tributaria de x



La identificación de regiones nodales o polarizadas ofrece una importante información en la toma de decisiones para diseñar políticas orientadas a la asignación racional de los recursos en el espacio geográfico como serían la localización puntual óptima de un servicio público, el punto de oferta de un bien o servicio orientado a un mercado determinado, la distribución jerárquica eficiente de equipamiento social, etcétera.

Ahora bien ¿qué tan exacto es este método de delimitación de áreas tributarias? De hecho, el modelo ha sido aplicado ininidad de veces tanto en su carácter descriptivo como predictivo. Su aplicación es rápida, simple y nada costosa. Se puede decir que en función de la predicción certera de la población futura de una área o región, la aplicación del modelo permitirá predecir con relativa precisión tanto el patrón de áreas tributarias como el flujo que podría generarse entre centros en una región.⁴

En fin, debe considerarse que un cambio en el número de los demandantes o en su nivel de ingreso así como en el número o calidad de las unidades oferentes de un servicio generará un cambio en el volumen del flujo y, por tanto, en la influencia de los centros donde unos y otros se localicen. Asimismo, que existe una interdependencia mutua, económica y social, tanto entre el centro de oferta y su área de mercado o tributaria como entre los centros de diferente jerarquía. Además, que el área de influencia es una expresión territorial de toda actividad humana localizada y que tal localización y su área de influencia representan una unidad significativa para la planificación en general (Friedmann, 1956) y para la oferta y la demanda de servicios, en particular.

En síntesis, cualquier decisión de localizar puntualmente una inversión (infraestructura física, equipamiento social o actividad productiva directa) dará lugar a una región nodal y el conjunto de estas formarán una región funcional o, lo que es lo mismo, un sistema regional jerarquizado. De esta suerte, si el problema es de eficiencia, es decir, de dar servicio al mayor número de personas dado un monto de inversión determinado (recursos escasos o previamente asignados), la regionalización o el conocimiento de la estructura regional, a partir de los métodos descritos en esta última parte del texto, *contribuirá* a mejores y menos costosas decisiones sin que sea necesario el uso de ambiciosas metodologías que pretenden considerar "todo". Y, si bien, estos modelos tratan

⁴ Aguilar y Graizbord (1993), Garrocho (1988a y 1988b), Graizbord *et al.* (1987 y 1994), Juárez (1993), han utilizado esta metodología para resolver diversos problemas de política pública y ventajas locacionales en diferentes contextos y escalas: nacional, regional, metropolitano, estatal y subregional.

sólo un tipo de interacción cada vez, esto no debe preocupar si se reconoce que pueden utilizarse indicadores de “masa” y “distancia” relevantes y que una interacción representada por algún tipo de flujos (como podrían ser llamadas telefónicas, carga, pasajeros, dinero, mensajes por cualquier medio, etcétera) es, de hecho, una síntesis de las complejas relaciones de tipo económico y político que resultan de la agregación de las múltiples decisiones individuales de los diversos agentes económicos y sociales en un tiempo y en un espacio geográfico determinados.

Bibliografía

- Aguilar, G. y B. Graizbord (1993), “Escenarios del crecimiento urbano para el ordenamiento territorial del país”, Sedesol (mimeo.).
- Bassols Batalla, A. (1964), “La división económica regional de México”, en *Investigación Económica XXIV*, 95.
- Berry D. y A. Wrobel (1968), “Economic Regionalization and Numerical Methods”, en *Geographia Polonica*, núm. 15.
- Boudeville, J. (1970), *Les Espaces Economiques*, PUF, Colección Que Sais-je?, núm. 950.
- Brown L. y J. Holmes (1971), “The delimitation of functional regions, nodal regions and hierarchies by functional distance approaches”, en *Journal of Regional Science*, vol. II, núm. 1, pp. 57-72.
- Bunge, W. (1966), *Theoretical Geography*, Lund Studies C.
- Friedmann, J. (1956), “The Concept of a Planning Region”, en *Land Economics*, vol. 32, núm. 1, pp. 1-13.
- Garrocho, C. (1988a), “Análisis del sistema de ciudades de San Luis Potosí”, tesis de Maestría en Desarrollo Urbano, El Colegio de México.
- (1988b), *Estructura funcional del sistema de asentamientos del Estado de México*, Cuaderno de Trabajo núm. 6, El Colegio Mexiquense.
- George, P. (1970), “La región, tema geográfico de la organización del espacio”, capítulo IV en P. George, *La acción del hombre y el medio geográfico*, Península.
- Golledge, R.G. y D. H. Amadeo (1966), “Some introductory notes on regional division and set theory”, en *Professional Geographer*, núm. 18, pp. 14-19.
- Gould, P. y T. Leinbach (1966), “An approach in the geographic assignment of hospital services”, en *Tidjschrift voor Economische en Sociale Geographie*, núm. 57, pp. 203-206.
- Graizbord, B. (1990), “Programa Nacional de Desarrollo Urbano, 1990-1994: aspectos cualitativos y cuantitativos para una evaluación ex ante”, en *Estudios Demográficos y Urbanos*, vol. 5, núm. 3, pp. 755-763.
- et al. (1987), “MESUREC: modelo exploratorio del sistema de ciudades de la región Centro del país”, Comisión de Conurbación del Centro del País (mimeo.).

- *et al.* (1994), "Análisis del impacto de la inversión pública y el crecimiento económico y demográfico en el estado de Guanajuato", El Colegio de México y Gobierno del estado de Guanajuato (mimeo.).
- Grigg, D. (1965), "The Logic of Regional Systems", en *Annals, Association of American Geographers*, núm. 55, pp. 465-491.
- Haggett, P. (1965), *Locational Analysis in Human Geography*, Arnold.
- Huff, D. (1963), "A Probabilistic Analysis of Shopping Centre Trade Areas", en *Land Economics*, núm. 27, pp. 81-90.
- Jones, E. (1956), "Cause and Effect in Human Geography", en *Annals, Association of American Geographers*, núm. 46.
- Juárez, E. (1993), "Interacción de mercados laborales municipales en el estado de Tabasco: una aproximación a través del uso de modelos gravitatorios", en *Estudios Demográficos y Urbanos*, vol. 8, núm. 1, pp. 157-189.
- Kansky, K. J. (1963), *Structure of Transportation Networks. Relationships between Network Geometry and Regional Characteristics*, Ph. dissertation, University of Chicago.
- Knox, P.L. (1975), *Social Weil-Being: A Spatial Perspective*, Oxford University Press.
- Krugman, P. y M. Obsfield (1994), *International Economics-Theory and Policy*, Harper and Collins.
- Little, A. y C. Mabey (1972), "An Index for Designation of Educational Priority Areas", en A. Shonfield y S. Shaw (comps.), *Social Indicators and Social Policy*, Heineman.
- Maxfield, D.W. (1972), "Spatial Planning of School Districts", en *Annals, Association of American Geographers*, núm. 62, pp. 582-590.
- Nystuen, J. y M. Dacey (1961), "A Graph Theory Interpretation of Nodal Regions", en *Papers, Regional Science Association*, núm. 7, pp. 29-42.
- Ohmae, K. (1993), "The Rise of the Region State", en *Foreign Affairs*, núm. 72, pp. 78-87.
- Papageorgiou, G. J. (1980), "Social values and Social Justice", en *Economic Geography*, vol. 56, núm. 2.
- Smith, D.M. (1974), "Who Gets What, Where, and How", en *Geography*, núm. 59, pp. 289-297.
- (1975), *Patterns in Human Geography*, Penguin, Harmondsworth.
- Spence, N. y P. Taylor (1970), "Quantitative Methods in Regional Taxonomy", en *Progress in Geography*, núm. 2, pp. 1-63.
- Stern, C. (1973), "Las regiones de México y sus niveles de desarrollo socioeconómico", en *Jornadas* núm. 72, México, El Colegio de México.
- Taylor, J.P. (1969), "The Location Variable in Taxonomy", en *Geographical Analysis*, núm. 1, pp. 181-195.
- (1977), *Quantitative Methods in Geography*, Houghton Mifflin.
- Yates, M. (1963), "Hinterland Delimitation: A Distance Minimizing Approach", en *Professional Geographer*, núm. 15 (6).