

ÍNDICE

<i>Presentación</i>	13
Tema 1. CONOCIMIENTO PREVIO DE LA NATURALEZA DE LA INFORMACIÓN GEOGRÁFICA	17
1. La naturaleza de la información geográfica.....	17
2. La medición de datos en Geografía. Las unidades espaciales de recogida de la información.....	19
3. La componente temática de la información.....	21
3.1. Las escalas de medida	21
3.2. Tipos de variables o atributos.....	24
3.3. La componente temporal de la información geográfica.....	26
4. La matriz de datos geográfica	27
5. La medición topológica de los objetos geográficos.....	30
6. La problemática particular de la información espacial	32
6.1. La autocorrelación espacial	33
6.2. El problema de la unidad espacial modificable (PUEM).....	34
7. La obtención de la información geográfica	35
7.1. La infraestructura de datos espaciales.....	35
7.1.1. El origen de la información geográfica de base espacial.....	36
7.2. Las principales fuentes de la información temática.....	39
7.3. Hacia una normalización global de la información geográfica.....	40
7.4. La recogida de la información geográfica por medio del muestreo	42
7.4.1. El muestreo de base espacial.....	42
7.4.2. El muestreo como técnica de investigación social.....	44
8. La calidad de los datos	47
8.1. Componentes de la calidad de la información geográfica	48
8.2. Los errores inherentes a la información geográfica. Tipos de errores	50

Tema 2. ANÁLISIS DE LOS DATOS. LA COMPONENTE TEMÁTICA DE LA INFORMACIÓN GEOGRÁFICA (1). ANÁLISIS UNIVARIADO Y MULTIVARIADO ENTRE VARIABLES	53
1. Algunos comentarios sobre el significado de la estadística clásica y su relación con la ciencia geográfica	53
2. El análisis exploratorio de datos	54
3. El análisis de las distribuciones de frecuencias	57
3.1. Distribuciones de frecuencia para datos nominales (variables cualitativas).....	58
3.2. Distribuciones de frecuencia para datos ordinales (variables cuasicuantitativas).....	59
3.3. Distribuciones de frecuencia para datos de variables cuantitativas	60
3.3.1. Las medidas de tendencia central.....	62
3.3.2. Las medidas de posición no centrales: los cuantiles	70
3.3.3. Las medidas de dispersión. La varianza. La desviación típica	73
3.3.4. Las formas de distribución: asimetría y curtosis.....	76
4. Correlación y regresión bivariada	80
4.1. El concepto de correlación	80
4.1.1. Relación entre variables cualitativas. La relación de variables nominales. La tabla de contingencia. El coeficiente de contingencia	80
4.1.2. Relación entre variables cualitativas. La correlación de variables ordinales. El coeficiente de Spearman. El coeficiente de Kendall	82
4.1.3. Relación de variables cuantitativas. El coeficiente de Pearson	85
4.2. El modelo de regresión lineal simple	87
4.2.1. El concepto de residual. La interpretación de los resultados del análisis de regresión a partir de los residuales ...	91
4.2.2. Inferencia y análisis de regresión	94
4.2.3. Regresión no lineal	98
5. El análisis multivariado	102
5.1. El análisis factorial	102
5.1.1. El modelo del análisis factorial. Las relaciones fundamentales de forma intuitiva	104
5.1.2. Fundamento matemático del análisis factorial.....	114

5.1.3. Fases del análisis.....	117
5.1.4. La aplicación del modelo factorial en el campo geográfico.....	119
5.2. El análisis de regresión múltiple.....	122
5.3. La regresión logística.....	126
5.3.1. Introducción.....	126
5.3.2. El modelo de regresión logística.....	127
5.3.3. El concepto de odds ratio.....	128
5.3.4. Interpretación del modelo con una variable independiente dicotómica.....	130
5.3.5. Interpretación del modelo con una variable independiente continua.....	132
5.3.6. Las variables cualitativas, de diversas categorías, en el modelo logístico.....	133
5.3.7. Análisis multivariable.....	134
5.3.8. Presentación habitual de los resultados de una regresión logística.....	135
5.3.9. Bondad del ajuste.....	136
5.3.10. Un ejemplo sencillo que facilite la interpretación.....	136
Tema 3. ANÁLISIS DE LOS DATOS. LA COMPONENTE TEMÁTICA DE LA INFORMACIÓN GEOGRÁFICA (2). MÉTODOS DE CLASIFICACIÓN.....	141
1. El objeto de la clasificación de los datos geográficos.....	141
2. La metodología clasificatoria: fases y procedimientos de ejecución.....	142
3. Métodos monotéticos, basados en la optimización de un único criterio.....	144
3.1. Método de agrupamiento con la ayuda de parámetros.....	144
3.1.1. El índice de especialización de Nelson.....	145
3.2. Método de partición por discontinuidades.....	147
3.3. La clasificación monotética, en función de un índice o indicador geográfico sintético.....	150
3.4. Delimitación de regiones funcionales.....	151
3.5. La clasificación por el gráfico triangular.....	152
4. Métodos basados en la consideración de diversos criterios o variables.....	155
4.1. Procedimientos de representación mediante dos ejes cartesianos.....	156
4.2. Métodos jerárquicos.....	157

4.2.1. Los procedimientos jerárquicos deductivos	157
4.2.2. Los procedimientos jerárquicos aglomerativos.....	159
4.3. Los métodos nucleares o con transferencia	164
Tema 4. ANÁLISIS DE LOS DATOS. LA COMPONENTE TEMÁTICA DE LA INFORMACIÓN GEOGRÁFICA (3). TÉCNICAS DE ESTADÍSTICA INFERENCIAL	167
1. La estadística inferencial. El concepto de muestra y población o universo. El concepto de población finita e infinita	167
2. El concepto de curva de probabilidad. De la distribución de frecuencias muestral a la distribución teórica	168
2.1. Algunas de las distribuciones teóricas más frecuentes	169
2.1.1. La distribución binomial.....	170
2.1.2. La distribución de Poisson	173
2.1.3. La distribución normal o de Gauss.....	178
2.1.4. La distribución de la χ^2 (ji cuadrado).....	184
2.1.5. La distribución t de Student	185
2.1.6. La distribución F de Snedecor.....	186
3. La distribución muestral de algunos estadísticos.....	187
4. Tests de significación estadística	190
4.1. Tests para una sola muestra.....	192
4.1.1. El test de la ji cuadrado χ^2	192
4.1.2. El test de Kolmogorov-Smirnov	194
4.1.3. El test z.....	195
4.2. Tests para la comparación de dos muestras.....	196
4.2.1. El test de la ji cuadrado χ^2 para comparación de dos muestras independientes.....	196
4.2.2. El test de Kolmogorov-Smirnov para comparación de dos muestras independientes	198
4.2.3. El test t de Student	200
4.3. El análisis de la varianza.....	202
4.3.1. Aplicación del análisis de la varianza en un ejemplo didáctico concreto	205
Tema 5. ESTUDIO DE LOS FENÓMENOS PUNTUALES	215
1. Fenómenos puntuales: concepto	215
2. Visualización de fenómenos puntuales: los mapas de puntos	216
3. Análisis espacial de fenómenos puntuales: conceptos básicos.....	219

3.1. Patrones espaciales puntuales	219
3.2. Procesos espaciales puntuales	219
4. Análisis espacial de patrones puntuales simples	222
4.1. Descripción estadística sumaria de los patrones puntuales simples.....	223
4.2. Estadística avanzada para los patrones puntuales simples. Modelo de aleatoriedad espacial completa	225
4.3. Modelos alternativos al modelo de aleatoriedad espacial com- pleta	227
4.4. Expresión matemática de los efectos de primer y segundo orden	230
4.4.1. Función de intensidad de primer orden	230
4.4.2. Función de intensidad de segundo orden.....	232
4.4.3. Momento reducido de segundo orden	233
4.5. Caracterización analítica de patrones puntuales simples	234
4.5.1. Análisis mediante cuarteles.....	235
4.5.2. Análisis del vecino más próximo	243
4.5.3. Análisis refinado del vecino más próximo	246
4.5.4. Análisis mediante la función «K» (momento reducido de 2º orden).....	252
Tema 6. ESTIMACIÓN DE DENSIDADES	267
1. Introducción	267
2. Métodos no paramétricos de estimación de densidades	269
3. Estimadores zonales	271
3.1. Estimación de densidades mediante cuarteles.....	271
3.2. Estimación de densidades mediante coropletas.....	275
3.3. Problemas de la estimación zonal de densidades	276
3.3.1. Arbitrariedades en las estimaciones.....	276
3.3.2. Discontinuidad de las estimaciones	278
4. Estimadores focales	279
4.1. Estimación de densidades mediante ventanas móviles.....	280
4.1.1. Aportaciones de la estimación de densidades mediante ventanas móviles	283
4.2. Estimadores kernel	284
4.2.1. Fundamento de la estimación kernel	285
4.2.2. Características de las funciones kernel	288
4.2.3. Procedimiento de la estimación kernel.....	289

4.2.4. Factores que afectan a la estimación kernel.....	291
4.2.5. Elección del tipo de kernel	293
4.2.6. Elección del tamaño de la ventana	296
4.2.7. Extensión de los estimadores kernel a los patrones marcados con cargas.....	305
4.2.8. Aportaciones y aplicaciones de los estimadores kernel.....	309
Tema 7. ANÁLISIS GEOESTADÍSTICO DE DATOS CONTINUOS	313
1. Los principios básicos del análisis geoestadístico	313
2. La autocorrelación espacial	314
2.1. Expresión general de la autocorrelación espacial.....	315
2.1.1. El índice de autocorrelación espacial de Morán	316
2.1.2. El Índice de autocorrelación espacial de Geary	319
2.1.3. Test de la χ^2 , mediante el método de las frecuencias en una rejilla cuadrada superpuesta (análisis mediante cuarteles).....	319
3. La teoría de las variables aleatorias regionalizadas	322
4. El análisis estructural. La construcción del variograma	324
4.1. Conceptos básicos.....	324
4.2. Construcción del variograma	328
4.3. Problemas más comunes encontrados en el cálculo del semi- variograma.....	336
4.4. Tipos de variogramas	336
4.5. El ajuste del variograma experimental a un modelo teórico ..	341
4.6. Diferencias del variograma según la anisotropía del territorio..	342
4.7. La importancia de la estacionariedad de los datos	348
4.8. Los conceptos de covarianza y correlograma	349
5. Modelización de la variación espacial	352
5.1. Los principales procedimientos de interpolación existentes ..	352
5.1.1. Métodos de poligonación	354
5.1.2. Métodos de triangulación.....	355
5.1.3. La interpolación polinómica o de superficie de tenden- cia.....	358
5.1.4. El método de interpolación de las medias móviles con ponderación de la distancia	362
5.1.5. El modelo del krigeado	365
<i>Bibliografía</i>	387

TEMA 1
CONOCIMIENTO PREVIO DE LA NATURALEZA
DE LA INFORMACIÓN GEOGRÁFICA

1. LA NATURALEZA DE LA INFORMACIÓN GEOGRÁFICA

La Geografía puede concebirse como la ciencia cuyo objetivo fundamental es la explicación, comprensión y predicción de la localización y dinámica de las diferentes actividades y realizaciones naturales y humanas que tienen lugar sobre la superficie terrestre. Desde esta perspectiva, cualquier investigación geográfica requiere la disponibilidad de determinada información, estructurada y georreferenciada espacialmente, lo que confiere a esta ciencia de unas características y una especificidad muy concretas.

Sin embargo, no cabe asimilar la «información geográfica» a la mera localización de sucesos. Los datos geográficos poseen un contenido más complejo, pudiendo ser considerados como relativos a entidades espacio temporales que cuantifican la distribución, el estado y los vínculos de los distintos fenómenos u objetos naturales o sociales en el espacio. Su naturaleza particular se deriva de la conjunción, dentro de la misma, de una serie de propiedades de carácter múltiple, que podríamos resumir en:

1. **Locacional**, para referir la posición de un objeto sobre la superficie terrestre, mediante un sistema de coordenadas terrestre. Esta situación puede ser establecida, igualmente, de manera relativa, frente a otros sistemas definidos de manera más o menos arbitraria.
2. **Espacial**, para describir que, en cuanto objeto material, ocupa un lugar en el espacio, y posee una forma y un volumen geométrico representativo.
3. **Temática**, para expresar los atributos o propiedades de dicho objeto, de manera que se establezcan las características de cada elemento o fenómeno.
4. **Temporal**, para relatar los cambios ocurridos con el transcurso del tiempo.

Estas y otras consideraciones hacen que los datos geográficos posean propiedades especiales que justifican el desarrollo de sistemas de información específicos para su tratamiento. Entre las más relevantes, por su influencia en el modelado de datos geográficos, se puede destacar la autocorrelación espacial. La *primera ley de la Geografía*, enunciada por W. Tobler (1970), establece que todos los objetos de la superficie terrestre están relacionados entre sí, pero las relaciones son más intensas con los objetos más próximos que con los más alejados. Además, como tendremos ocasión de exponer más adelante, los datos geográficos están afectados por la agregación de *unidades espaciales arbitrarias* (PUEM), e, incluso, también por la autocorrelación temporal. Estos principios fundamentan los estudios geográficos de base geoestadística, pero, a la vez, condicionan el análisis de los datos por los métodos tradicionales.

Por todos los motivos señalados, los métodos utilizados para el análisis espacial son muy diversos. El empleo de la estadística clásica conduce a olvidar, en un primer momento, la localización de los objetos, centrándose el análisis, exclusivamente, en la componente temática y temporal de la información geográfica. Desde esta perspectiva, resultan de interés determinadas técnicas denominadas como *análisis de datos*. Por el contrario, determinadas variables, analizadas, no sólo por la componente temática, sino, además, por su situación en el espacio, muestran con frecuencia estructuras espaciales no aleatorias que pueden condicionar los resultados de las técnicas de análisis de la estadística convencional. Ello obliga a la utilización de métodos geoestadísticos y de la morfología matemática, agrupados en la denominada *estadística espacial*, que analizan, directamente, las informaciones geocodificadas.

Además, la posición relativa de los objetos geográficos en el espacio introduce nuevos términos a los que referirse, más allá de la concepción individualizada de un único componente, que tienen que ver con el concepto de estructura espacial, como relación relativa e interrelacionada de un conjunto de elementos. Entre estas nuevas concepciones a considerar, estarían:

1. La **accesibilidad** de un lugar, definida como la mayor o menor facilidad con la cual un lugar puede ser alcanzado desde uno o otro lugares.
2. La **distancia** entre lugares, como manera de medir su separación relativa. Tan importante es el concepto que la definición matemática de la distancia define, por sí sola, un espacio métrico.

3. La **interacción espacial**, como fenómeno de relación mutua entre los elementos de un sistema, que se manifiesta por los diferentes flujos existentes entre cada dos de ellos, generalmente decrecientes con la distancia.
4. La **polarización** o atracción que ejerce un lugar sobre un espacio más o menos extenso y heterogéneo, que se encuentra en una situación de dependencia con respecto a un lugar central.
5. La **escala geográfica** considerada, que analiza los fenómenos geográficos a distintos niveles de aproximación y que modifica el nivel de organización de los sistemas terrestres.

La consideración de la estructura de los elementos geográficos conlleva otro tipo de análisis, relacionado con la *modelística espacial*. El análisis espacial pone en evidencia las formas de organización de las actividades en el territorio, permitiendo distinguir, mediante la reproducción y simulación de la realidad por medio de modelos, las propiedades que definen sus estructuras, a partir de la información contenida en sus elementos e interrelaciones mutuas. Durante los últimos años, numerosos estudios han puesto de relieve que muchos patrones o comportamientos de la realidad geográfica no pueden ser explicados, únicamente, mediante los factores, de forma aislada, sino que requieren otro tipo de técnicas de análisis. En este sentido, los modelos, como concepciones simplificadas del mundo real, constituyen, actualmente, interesantes herramientas capaces de plantear las complejas problemáticas geográficas existentes y su posible solución de manera formalizada.

2. LA MEDICIÓN DE DATOS EN GEOGRAFÍA. LAS UNIDADES ESPACIALES DE RECOGIDA DE LA INFORMACIÓN

La recopilación y posterior tratamiento de la información, en cualquiera de los campos científicos existentes, implica la definición previa de las unidades de análisis que sirven de soporte a la recogida de los datos, así como la concreción del tipo de variables o atributos temáticos de interés que de cada elemento se desea conocer.

La particularidad de los estudios geográficos viene determinada por el carácter espacial de los objetos a analizar. Esta peculiaridad supone que los

componentes de un sistema geográfico se hallan definidos, no sólo por su forma y extensión, sino por la situación que ocupan en el espacio, establecida por la localización relativa respecto a algún procedimiento de referencia exterior (en general, las proyecciones sobre el elipsoide terrestre y sus coordenadas de longitud y latitud geográficas). La reducción geométrica de los elementos geográficos permite referirnos a diversos objetos geográficos según las dimensiones consideradas. Así, podríamos diferenciar *objetos puntuales* (sin dimensiones), *objetos lineales* (una dimensión), *objetos superficiales planos o polígonos* (dos dimensiones) y *objetos superficiales o volúmenes* (tres dimensiones).

Estas unidades de observación geográfica pueden ser naturales o artificiales. Las unidades naturales son intrínsecas al propio fenómeno observado. Así, los usos del suelo constituyen un buen ejemplo de este tipo de elementos, en donde los límites espaciales entre las diversas categorías representadas vienen determinados por las características del fenómeno en sí. Por su parte, las unidades artificiales no se hallan relacionadas con ningún hecho geográfico, sino que han sido definidas de forma arbitraria, de acuerdo a algún criterio artificioso. En este sentido, las unidades administrativas constituyen la muestra de cómo subdividir el espacio sin una relación geográfica directa y definida.

Otros autores diferencian los conceptos de *entidad* o ente (derivado del inglés «entity») y *fenómeno, recubrimiento o campo* (derivado, igualmente, de «field»). El término «entidad» se refiere a un objeto discreto e individualizable, representación de un ente del mundo real, cuyos límites y extensión espacial se hallan relativamente bien definidos. Entre las entidades, se incluyen aquellas realidades que puedan situarse en el espacio y en el tiempo: edificios, ciudades, oleoductos, carreteras, etc. Por su parte, el concepto «recubrimiento» se caracteriza por la variación cuantitativa de un fenómeno, de modo continuo en el espacio, como los campos escalares (campos de presión o temperatura) o modelos digitales del terreno (variación de la topografía de un territorio) o bien pueden ir asociados a variables nominales (v.g., los usos del suelo) que se manifiesten de manera discontinua en el espacio. Las entidades vienen representadas por elementos geométricos discretos (puntos, líneas y polígonos), mientras que los campos o fenómenos requieren la imagen de todos los puntos de un espacio concreto, lo que puede lograrse por un mapa de isolíneas.

3. LA COMPONENTE TEMÁTICA DE LA INFORMACIÓN

Las variables temáticas de interés geográfico son muy variadas, como corresponde a la naturaleza de nuestra disciplina. Información relativa al medio físico (litología, relieve, clima, vegetación, suelo, hidrografía, etc.) o humano (características sociales, económicas, políticas, culturales, etc.) tienen cabida en un análisis geográfico, cuyo foco de interés se centra en el estudio de las variaciones espaciales de un fenómeno determinado.

El proceso de medida de cualquier fenómeno geográfico, cuya manifestación espacial deseamos evaluar, necesita de algún procedimiento que permita establecer la correspondencia entre las unidades espaciales, objeto de estudio y de recogida de la información, y las variables temáticas a valorar. Esta operación concluye con la asignación de símbolos, normalmente numéricos, a cada uno de los individuos geográficos que integran la estructura espacial de referencia. Este procedimiento de medida de las cualidades espaciales de cada elemento del universo espacial considerado se halla condicionado por la naturaleza del hecho a medir, lo que incide en las relaciones matemáticas que se pueden establecer entre el andamiaje metodológico y el mundo real.

3.1. Las escalas de medida

La medición implica la asignación de números o nombres a las características o dimensiones de los objetos-unidades de estudio (Sánchez Carrión, 1996). La medición de las diferentes variables que integran el inventario de un determinado trabajo de investigación territorial exige que nos detengamos, previamente, en analizar su variada naturaleza, que se halla en relación con las diversas escalas de medida.

Las variables *nominales* representan el nivel más bajo de medición en cuanto a su significado. La verificación mediante esta escala de medida se realiza por la asignación de los individuos de un colectivo a una categoría o clase de entre un número limitado existente de las mismas, con un nombre representativo de su significado. Los nombres de las clases, en que se subdivide cada una de las variables nominales, sólo sirven para clasificar a los individuos u objetos de estudio, determinando si son iguales o diferentes en la característica que estamos analizando.

Gran parte de las variables ambientales que manejamos pertenecen a este grupo de variables. Por ejemplo, la definición en el territorio de la vegetación, dividida en tipos como pinar, encinar, matorral, etc., exige que cada unidad espacial considerada se haga corresponder con una de las categorías definidas, no pudiendo pertenecer, en general, a dos de ellas a la vez. La forma de expresar la medida es la asignación de un código a cada clase (pueden ser números correlativos 1, 2, 3, etc.), no teniendo entre sí ninguna correspondencia de orden o valor.

El escalón inmediatamente superior, en la escala de medida, es el representado por las variables *ordinales*, llamadas también de atributos jerarquizados. Se trata de valores que tienen la condición de estar ordenados según un criterio determinado, de manera que los códigos asignados: 1, 2, 3, 4, etc., sí representan, en este caso, una ordenación de los valores, de acuerdo a un orden, aunque no signifique que el número 2 sea el doble del número 1. En este caso, los individuos objeto de estudio no sólo quedan clasificados sino ordenados. Imaginemos que deseamos medir la productividad del suelo de forma categórica en tres clases: alta, media y baja producción. El tratamiento de la variable «producción agraria» podría realizarse, bien mediante variable nominal, o también mediante variable ordinal, al existir un orden ascendente desde la baja a la alta producción.

Tanto las variables nominales como las ordinales se denominan, asimismo, como variables categóricas o cualitativas. Las categorías definidas deben cumplir una serie de propiedades: ser exhaustivas, mutuamente excluyentes y basadas en un único principio de clasificación. Las categorías son *exhaustivas* cuando permiten clasificar a todas las unidades de análisis. La existencia de individuos difíciles de clasificar supone la necesidad de crear una clase que los integre y a la cual se le dé la etiqueta de «otros» o «resto». El principio de *mutua exclusión* supone que no exista ambigüedad, de manera que cada individuo sea asignado a una categoría y sólo a ella. Esta circunstancia implica que, cuando existen variables categóricas delimitadas por intervalos, se debe definir con exactitud a qué categoría pertenece el valor frontera. Por ejemplo, si definimos la altitud como variable categórica, mediante las clases: 100-200 metros, 200-300 metros y 300-400 metros, deberíamos señalar a qué categoría pertenecen los valores 200 y 300. Finalmente, las clases definidas en cada variable deben proceder de *un único principio clasificatorio*.

Por su parte, las variables *cuantitativas*, propiamente dichas, son las de intervalo y las de razón, porque aplican una unidad de medida a todos los objetos o individuos, de manera que es posible diferenciarlos por valores concretos y precisos.

La diferencia entre ambos se debe a que la *escala de intervalo* asigna un valor cero como origen de medición, que es elegido arbitrariamente, y que no significa que el objeto tenga ausencia de la característica que se mide. Un ejemplo característico de esta escala de medida es la temperatura. La temperatura 0 °C (centígrados) no significa que no exista temperatura alguna, sino que ese punto es el origen de la medición. Por tanto, entre dos temperaturas, de 20 °C y 40 °C, no sería correcto decir que una es el doble de la otra.

Muchas variables geográficas pertenecen, igualmente, a esta escala de medida. Por ejemplo, la renta per cápita de las provincias españolas, expresada con relación a una renta base igual a 100, sería una variable medida en escala de intervalo. La diferencia entre dos provincias con valores, de 80 y 90, sería la misma que otras dos, entre 110 y 120, sin que ocurra lo mismo entre los valores absolutos de las mismas.

Finalmente, las variables medidas en escala de razón se caracterizan por tener un cero absoluto, que representa la ausencia de la característica medida; por ejemplo, si se considera la altura media de la cubierta vegetal, dicha variable tendría valor cero en las zonas totalmente despejadas de vegetación; por esta causa, el cociente entre dos valores cualesquiera de una variable de este tipo es significativo en relación con la característica medida. Siguiendo con el mismo ejemplo, si se consideran dos zonas, cuya altura media de la cubierta vegetal sea, respectivamente, de 2 m y de 1 m, tiene sentido decir que la altura media de dicha cubierta es doble en la primera zona que en la segunda.

La selección de un tipo u otro de escala viene, en general, condicionado por las características de la variable a medir, aunque, a veces, una misma variable pueda medirse en una escala u otra. Pensemos en la vegetación de un área concreta, que únicamente reflejemos el tipo de planta existente (variable nominal), o el número de especies vegetales por unidad de superficie (variable cuantitativa). En general, la transformación de variables de carácter cuantitativo a ordinal o nominal supone una disminución de tiempo y esfuerzo en la realización del inventario de las mismas, aunque limite su tratamiento a métricas determinadas (no paramétricas).

Además, una variable puede ser modificada o transformada, con objeto de obtener una mejor representación de los datos, sin que ello represente nada más que una simple operación de cálculo. Solo las variables de mayor escala de medida pueden ser transformadas en otras de escala inferior o de la misma escala. Así, a partir de variables cuantitativas, es posible obtener escalas cualitativas, por categorización de la variable. Por ejemplo, la variable representativa de la pendiente del terreno, de tipo de razón, podría ser transformada en una variable cualitativa, al definir varios intervalos: 0-10%, 10-20%, 20-30% y 30-40%. La nueva variable sería una variable ordinal de cuatro clases, ordenadas de mayor a menor. Igualmente, una variable ordinal podría ser transformada en nominal o en una variable nominal podrían agruparse las categorías, reduciendo su número.

3.2. Tipos de variables o atributos

La naturaleza específica de la información geográfica implica la existencia de un variado conjunto de variables o atributos, relativos, bien a cada objeto discreto de análisis bien a cada una de las celdas que integran un espacio raster. Esta amplia variedad de características procede, tanto de la presencia de seres en el interior de cada una de las unidades descritas (personas, edificios, especies animales o vegetales, etc.), como de la medición de los variados fenómenos que se manifiestan en el espacio (variables climáticas, tipos de suelo o roquedo, etc.), la medida respecto a una posición de referencia geográfica (coordenadas geográficas, altitud) o la magnitud de los flujos o relaciones mutuas existentes entre las unidades espaciales del análisis (Duncan, Cuzzort y Duncan, 1961). En un intento de mostrar esta amplia gama de atributos geográficos, así como de sus unidades de medida, podríamos referirnos a:

1. Si el territorio de estudio es analizado como una *colección de «items»* o *«entes»* contables, se pueden considerar una serie de variables resultantes de medir el número de unidades presentes en la unidad espacial. Así, tendríamos el volumen de población, medido por el número de personas censadas en un espacio geográfico concreto, o el número de familias, establecimientos industriales o de otro tipo, el número de cabezas de ganado o el de especies vegetales determinadas. El número de variables geográficas de este tipo puede crecer en amplia progresión, si consideramos las sub pobla-

ciones que se derivan de cada grupo de ítems o entes medibles, como el relativo a la población de sexo masculino o aquélla que posee una determinada edad o nivel de actividad. Este tipo de variables pertenece a la escala de razón, ya que los resultados son fácilmente comparables con una unidad de medida (la persona, el establecimiento, el edificio, la especie animal o vegetal, etc.), y existe un cero absoluto en cada una de ellas.

2. Otro tipo de variables surge de la relación entre variables extensivas, del tipo de las anteriores, lo que se denominan *índices*. Este tipo de variable se expresa como relación entre las poblaciones de dos estratos de la población total (v. g. el índice de masculinidad como relación entre el número de hombres y el de mujeres en tanto por ciento) o la relación existente entre la población perteneciente a un estrato de población respecto a la población total (v. g. el índice de juventud, medido por el porcentaje de personas de edad menor de 15 años respecto al total de la población).

3. Si la unidad espacial de recogida espacial es al menos bidimensional, las variables anteriores pueden venir expresadas en forma de *densidad*, al medir la relación existente entre el número de ítems y la superficie (en el caso de polígonos) o longitud (en el caso de líneas). Es muy frecuente referirse al volumen de población por unidad de superficie o al número de estaciones de servicios de una carretera determinada, expresado por unidad de longitud. Este tipo de variables, como en el caso anterior, son de escala de razón.

4. La medida de determinados *fenómenos geográficos*, como las características climáticas del aire atmosférico, presentes en el objeto geográfico de estudio, vendrían expresadas en las unidades de medida específicas a este tipo de variables. Así, nos referiríamos a la temperatura media anual o la velocidad de viento máxima en un observatorio. Este tipo de variables puede pertenecer a la escala cuantitativa (como la velocidad en metros por segundo) o a la de intervalo (como la temperatura).

5. Otro tipo de variables geográficas se deriva de la *medición de un segmento del espacio* (bien un trozo de superficie o de línea) respecto a la superficie total. La medida de variables de carácter nominal, como la existencia de determinadas categorías de usos del suelo o de litología, puede ser evaluada por la presencia o ausencia de este fenómeno, en cuyo caso la escala de medida sería de tipo cualitativo, pero también por la relación de la superficie ocupada respecto al total de superficie de la unidad de análi-

sis, en cuyo caso nos hallaríamos en presencia de una variable de razón. Este sería el caso de medir la superficie destinada al regadío de un municipio concreto, bien en hectáreas o en porcentaje respecto al total de la superficie del mismo.

6. El cálculo de las relaciones o *flujos existentes entre las unidades espaciales* del análisis, medido en número de personas, bienes o capitales, es una variable de uso frecuente en los estudios geográficos. Así, la investigación acerca del mercado laboral de una gran ciudad requiere del conocimiento de los desplazamientos diarios entre los lugares de residencia y trabajo, medidos entre los municipios o distritos que la integran. Las variables derivadas estarían referidas a los flujos o movimientos de personas existentes entre cada unidad espacial objeto de estudio y las demás.

7. Finalmente, otro tipo de variable geográfica resulta de relacionar el valor de una variable concreta, de las expuestas con anterioridad, con respecto al espacio total. Así, podríamos referirnos al volumen de población de una provincia, respecto al total español.

3.3. La componente temporal de la información geográfica

El tiempo juega un papel de primera magnitud en la interpretación de los fenómenos en el espacio. El mundo real sólo puede ser explicado a partir de la evolución de las estructuras espaciales y de los procesos temporales que intervienen en su modificación. Así, un estudio de los usos del suelo de una determinada región requiere del conocimiento de los mismos en una etapa anterior, para comprender la modificación de la realidad según un patrón o modelo que ofrezca las claves de la transformación acontecida.

El cambio puede afectar, exclusivamente, a la componente temática y no a la espacial. En este caso, el estudio de la evolución temporal de un fenómeno se ve facilitada. Ello supondría que las unidades espaciales continuarían siendo las mismas. Sería el caso de que los municipios de una región mantuvieran idénticos perfiles espaciales. Es frecuente, en las estadísticas ofrecidas por las administraciones públicas, una organización diferente de la información, de un año censal a otro, por el cambio de categorías en la distribución de la población de un colectivo, aunque no afecte a las unidades administrativas de referencia. Este hecho hace difícil la com-

paración de valores a lo largo del tiempo y la interpretación de la evolución temporal de un determinado fenómeno.

Si los cambios suponen una modificación de los objetos de análisis, el estudio temporal complica la interpretación del fenómeno a analizar. Si los elementos geográficos son el resultado de la división o agrupación de los ya existentes, la información temática podría ser reconstruida, haciendo intervenir la totalidad de los mismos en las diversas etapas del estudio.

El estudio temporal, cuando se dispone de información suficiente, es más fácil de realizar en el modelo raster. En este modelo, la atención se centra, exclusivamente, en el cambio temático, ya que las unidades espaciales, para el mismo nivel de resolución (tamaño de celda), permanecen idénticas.

4. LA MATRIZ DE DATOS GEOGRÁFICA

La organización de la información geográfica, de acuerdo a las tres dimensiones a que nos hemos referido en apartados anteriores: unidades espaciales, atributos y momentos en el tiempo, ha supuesto la ordenación de los datos en una matriz geográfica (Berry, 1964), con vistas a la realización de cualquier análisis regional.

	Lugares (columnas)							
atributos (filas)								

Figura 1.1. Matriz de datos geográfica reducida (componentes espacial y temática).

Esta primitiva manera de preparar la información geográfica estaba orientada al tratamiento cuantitativo de las características geográficas de una serie de lugares, concebidos éstos como objetos geográficos, sin consideración de su situación espacial. La aportación de Berry suponía la disposición de los lugares y atributos en una matriz rectangular, donde los primeros eran situados en columnas y los segundos en filas. Los lugares eran cualquier elemento geográfico definido: país, región, municipio, sección censal, cuadrícula del terreno, puntos relativos a observatorios, etc. (figura 1.1). La intersección de una fila con una columna determina una celda, cuyo valor cuantitativo expresaría la medición del atributo considerado en la unidad espacial correspondiente. Los elementos de una fila reflejarían la variación espacial de una variable, mientras que los de una columna constituirían el inventario de valores temáticos correspondientes a un lugar determinado.

Elaborada la matriz geográfica, podríamos abordar el análisis regional desde diversos puntos de vista (Estébanez y Bradshaw, 1979):

1. Examinar la disposición de las celdas, siguiendo una fila completa o una parte de la misma. Supondría el estudio de la variación espacial de una única variable y podría elaborarse un mapa temático de su distribución espacial.
2. Examinar la disposición de las celdas dentro de cada columna, lo que reflejaría el inventario locacional o estructura local de un lugar determinado.
3. Comparar dos filas o todo el conjunto de filas. Se estudiaría, así, la covariación o asociación espacial de dos o más hechos geográficos.
4. Comparar un par de columnas o toda la serie de columnas de la matriz.
5. Estudiar una submatriz de la matriz general, matizando las variaciones espaciales en una subregión, referidas a unas determinadas variables geográficas.
6. Comparar una fila o parte de ella en dos momentos de tiempo diferentes.
7. Comparar una columna o parte de ella en dos momentos de tiempo diferentes.