

Discretización

La información geográfica puede revestir varias formas. Entre éstas, la representación cartográfica presenta la ventaja de la lectura instantánea de una imagen, una lectura que debe ir de lo general a lo particular. La ventaja que constituye esta lectura sólo puede existir al precio de una inversión previa hecha durante la concepción y la realización de la «carta». La inversión involucra tanto la selección pertinente de la información como el tratamiento de dicha información. El tratamiento previo de la información depende del tipo de carácter estadístico que se desea cartografiar. En función de dicho carácter, las elecciones de los métodos de discretización son más o menos numerosas y los resultados cartográficos pueden dar imágenes muy variadas. La discretización es la operación que permite separar en clases una serie de «variables cualitativas» o de «variables cuantitativas». Esta operación simplifica la información agrupando los objetos geográficos que presentan las mismas características en distintas clases. Una discretización es satisfactoria cuando permite la creación de clases homogéneas y distintas entre sí: los objetos geográficos de una misma clase deben parecerse más entre sí que con los objetos de las otras clases. Discretizar una serie estadística constituye a menudo la última etapa de la reducción, de la organización y la jerarquización de la información antes de construir una carta que da cuenta de la repartición geográfica de esa misma serie estadística.

La operación de discretización debe satisfacer a la vez las exigencias de la representación cartográfica y las de los principios estadísticos. Debe conservar las características esenciales que presentan los datos, perder el mínimo posible de información, pero también respetar las reglas de la percepción visual con el fin de transmitir una información geográfica eficaz y de calidad. El número óptimo de clases a obtener en una partición es siempre función del número de individuos observados (unidades espaciales). Existe un índice que permite conocer el número ideal de clases para una distribución; es necesario considerarlo únicamente como una ayuda indicativa. Se trata del índice de Huntsberger:

$$N(c) = 1 + 3,3 \log_{10}(N)$$

N = número de observaciones

$N(c)$ = número de clases

Los principios de la discretización difieren según la naturaleza de la información:

Si la información es cualitativa:

ordinal: se busca conservar la jerarquía de las informaciones;

nominal: se pretende definir un criterio común de agrupamiento para lograr la construcción de una tipología.

No existe un método específico para discretizar la información cualitativa: cada agrupamiento de objetos geográficos es propio de los objetivos de la simplificación de los datos elegida.

Si la información es cuantitativa: de stock o de relación: se pretende conservar la forma de la distribución o resaltar valores particulares.

La elección de un método de discretización depende de las propiedades de la distribución y los objetivos que se han fijado respecto de la información a comunicar.

Se deben preservar tres informaciones sobre las características de la distribución:

el orden de amplitud, medido a través de los valores centrales de la distribución: modo, media y mediana. Estos valores pueden elegirse como límites de clases y hacer aparecer los valores excepcionales que informan sobre las particularidades de la distribución;

la dispersión, que es una medida de la desigualdad de los valores, caracterizada por el desvío tipo, el intervalo intercuartiles o el coeficiente de variación; esos valores pueden elegirse como amplitudes de clases. Dicha medida toma en cuenta la variancia: ésta permite minimizar las diferencias entre los objetos de una misma clase (varianza intra-clase) y de maximizar las diferencias entre las clases (varianza inter-clase);

la forma, que puede ser de tipo normal, simétrica o disimétrica. Si la serie estadística observada es pequeña, la realización del diagrama de distribución permite la evaluación de la forma de la distribución. Si la serie es grande, la comparación de los valores centrales posibilita determinarla.

Los principios de la discretización difieren según los objetivos a alcanzar:

si se trata de construir una cartografía, el mejor método es el que da mejor cuenta de la estructura interna de los datos: los «umbrales naturales» observados sobre el diagrama de distribución u otro método que respete la forma de la distribución;

si se trata de establecer comparaciones sobre el mismo conjunto de objetos geográficos (por ejemplo por medio de una serie de cartas), es necesario utilizar obligatoriamente métodos de discretización basados en el cálculo de umbrales significativos estadísticamente. Allí todo depende aún del objetivo de la comparación:

-Para comparar las posiciones de los mismos lugares según varios caracteres: las unidades de medida son diferentes, entonces es

necesario emplear un método que permita utilizar los mismos parámetros estadísticos de cada serie (por ejemplo la media y la desviación estándar);

-Para comparar las dispersiones de varias series estadísticas con los mismos «lugares»: esta comparación sólo es válida si los datos son de la misma naturaleza con las unidades de medida idénticas, y se empleará el método de las amplitudes o los efectivos iguales jugando sobre la gama de los grises o los colores.

-Para comparar una misma serie con fechas diferentes: hay que establecer la discretización sobre el conjunto de los valores observados de la serie estadística para el conjunto de los datos.

Los métodos de discretización:

Existe un gran número de métodos de discretización y de variantes. La elección de un método depende a la vez de las propiedades de la distribución, de las posibilidades de la representación cartográfica y de los objetivos que uno se fija (mensaje, público, soporte).

Los cuatro métodos más fáciles de implementar y los más corrientemente utilizados en geografía son los siguientes: el método de los umbrales naturales, el de las amplitudes iguales, el de los efectivos iguales y el de la progresión geométrica.

Christine Zanin

Bibliographie

Bibliografía:

-Cauvin (C), Raymond (H), Serradj (A), Discrétisation et représentation cartographique, Montpellier, GIP RECLUS (Col. Reclus modes d'emploi), 1987.

-Palsky G., Des chiffres et des cartes. Naissance et développement de la cartographie quantitative au XIXe siècle, Paris, CTHS, 1996.

-M. Beguin y D. Pumain, La représentation des données géographiques, coll. Cursus, Armand Colin, 2ª edición, 2000.

-E. Blin y J.P. Bord, Initiation géo-graphique ou comment visualiser son information, Sedes, 2ª edición, 1995.

-C. Zanin y M.L. Trémolo, Savoir faire une carte, coll. Sup Géographie Belin, 2003.