

# CELADE

## CENTRO LATINOAMERICANO DE DEMOGRAFIA

Distribución interna  
Roberto Bachi

5383

BIBLIOTECA "GREGORIO MARTÍNEZ"  
CENTRO LATINOAMERICANO  
DE DEMOGRAFIA

Serie D, nº 60.  
Agosto, 1970.  
300.

REPRESENTACION GRAFICA Y ANALISIS DE  
INFORMACION ESTADISTICO-GEOGRAFICA  
(Traducción del artículo aparecido en  
el Boletín del Instituto Internacional  
de Estadística, vol. XLI, tomo I,  
Actas de la 35ª Sesión, Belgrado, 1965,  
pp. 219-230)

Las opiniones y datos que figuran en este trabajo son responsabilidad del autor, sin que el Centro Latinoamericano de Demografía (CELADE) sea necesariamente partícipe de ellos.

## INDICE

	<u>Página</u>
1. LA REPRESENTACION GRAFICA DE LAS DISTRIBUCIONES GEOGRAFICAS .....	7
(a) Sistemas de símbolos y patrones gráficos y (b) normas para su selección .....	7
(c) Adaptación de la escala de símbolos gráficos de los datos .....	4
2. VALORES RESUMENES PARA LA DESCRIPCION DE UNA DISTRIBUCION ESTADISTICO-GEOGRAFICA .....	6
(a) Ubicación .....	7
(b) Dispersión .....	7
(c) La principal dirección de la distribución, y la representación gráfica de los valores resúmenes .....	9
(d) Asimetría .....	10
(e) Concentración .....	10
3. EL ESTUDIO DE LA DISTANCIA ENTRE DOS DISTRIBUCIONES GEOGRAFICAS .....	10
APENDICE .....	12

The  
 Department of the Interior  
 Bureau of Land Management  
 Office of Realty and Operations  
 1601 North Lincoln Street  
 Denver, Colorado 80202

Este trabajo se propone: 1) discutir algunos problemas de la representación gráfica de las distribuciones geográficas; y describir algunos trabajos de investigación realizados por el autor y colaboradores 2) en la construcción de valores resúmenes para la descripción de las distribuciones geográficas, y 3) en la medición de la "distancia" entre dos distribuciones geográficas.

## 1. LA REPRESENTACION GRAFICA DE LAS DISTRIBUCIONES GEOGRAFICAS

A pesar del papel vital desempeñado por los gráficos en la presentación de las series geográficas, y a pesar de los grandes avances alcanzados en la representación gráfica, tanto por los estadísticos como por los geógrafos, todavía queda mucho por hacer para que los mapas estadísticos sean exactos y racionales: la intuición y el enfoque artístico todavía juegan un papel demasiado importante en este campo. Los principales problemas que debemos enfrentar están relacionados con: (a) la disponibilidad de sistemas apropiados de símbolos y patrones gráficos; (b) la disponibilidad de pautas para elegir el sistema apropiado de símbolos o patrones para cada serie específica de datos a representarse; (c) la disponibilidad de pautas para adaptar las "gradaciones" de la escala de símbolos gráficos al tipo de datos a representarse. Destacaremos para su discusión algunos puntos escogidos en relación a cada uno de los temas precedentes.

### (a) Sistemas de símbolos y patrones gráficos y (b) normas para su selección

No es fácil obtener símbolos o patrones que puedan diferenciarse entre sí, que sean proporcionales a los datos a representarse y que brinden una impresión visual correcta al lector. Inter alia, se crean dificultades por la falibilidad del ojo humano: (i) al estimar magnitudes representadas por símbolos (tales como cuadrados, círculos, esferas, etc.) especialmente si éstos se hallan alejados entre sí en el mapa, y (ii) a veces incluso para reconocer el orden correcto de estos símbolos (véase por ejemplo [1], págs. 85-91)<sup>1/</sup> (iii) en la apreciación de la proporción exacta del negro sobre el blanco, que presentan ciertos tipos de patrones gráficos (véase por ejemplo [2]). A pesar de esto, parece haber lugar para el perfeccionamiento de algunos de los sistemas de símbolos o patrones actualmente en uso. Se presentaron algunas proposiciones en este campo a las reuniones de este Instituto realizadas en Roma y París (véase [3] y [4]). Se ha tomado en consideración la necesidad de sencillez en los nuevos proyectos gráficos, condición que se destacó en la discusión de [4] (véase [5]), de modo que ahora podemos presentar una nueva proposición para una escala simplificada de Patrones Racionales Gráficos individuales, los que se conocerán como PRG. Los PRG que presentamos aquí aún se hallan en estado provisional.

Como se muestra en el Gráfico 1, la idea fundamental de esta escala de PRG es la de representar cualquier<sup>2/</sup> entero  $n=10t + u$  por un patrón formado por  $u$  marcas unitarias cuadradas del área  $a$  (que indica las

1/ Los números en paréntesis cuadrados son referencias bibliográficas a la lista del Apéndice.

2/ Para mayor sencillez, el Gráfico 1 muestra principalmente una escala para  $0 \leq n \leq 100$ . Sin embargo, se muestran algunos métodos de alternativa para representar valores de  $n > 100$  al pie de este Gráfico, y, resulta inútil decirlo, la escala principal puede adaptarse a otros valores  $0 \leq n \leq 100k$ , atribuyendo cualquier valor arbitrario  $k$ , a la unidad elemental cuadrado del área  $a$ , y multiplicando por  $k$  los valores de todos los patrones.

unidades) y por  $t$  marcas cuadradas del área  $10a$  (que indica las decenas). La correspondencia entre  $n$  y el patrón que lo representa se asegura de dos modos: 1) la negrura de este patrón es  $n$  veces la del patrón que representa a la unidad (hecho que puede atraer la vista, a pesar de su falibilidad); 2) las marcas que forman cada patrón se agrupan de modo tal que permiten al lector captar rápida y fácilmente el valor relativo del patrón y le permiten "leer" su valor numérico, siempre que esto sea necesario.

Los patrones están dibujados en forma tal que requieren un espacio comparativamente pequeño, y para mayor comodidad, pueden encerrarse dentro de un pequeño marco cuadrado. Esto facilita la aplicación de PRG individuales para representar gráficamente las tablas de correlación, las tablas de contingencia y cualquier serie de datos clasificados en forma cruzada, para los cuales una de las características está dada por las hileras y la otra por las columnas del cuadro. Como ejemplo de este tipo de aplicación a los datos geográficos, que no están presentados sobre un mapa, el Gráfico 2 muestra algunos indicadores de características socio-económicas para cada una de las 9 divisiones principales de Estados Unidos de acuerdo al Censo de Población de 1960. Se pegaron todos los indicadores para cada división en franjas móviles, en el gráfico original, de modo de poder variar el orden de las divisiones. En el gráfico presentado, las divisiones están dispuestas en orden descendente de acuerdo al porcentaje de familias con ingresos de US\$ 10 000 y más.

Los PRG pueden utilizarse en tres formas distintas para representar las distribuciones estadístico-geográficas sobre mapas:

- 1) como patrones individuales, que representan un número  $n$  en un punto dado (escala del Gráfico 1);
- 2) como patrones repetidos, en los cuales el símbolo de  $n$  se repite muchas veces para demostrar que un valor  $n$  ha de asignarse a una zona dada. El Gráfico 3 muestra varios ejemplos de este método. Para los mapas menos detallados, puede bastar una escala en gradas de diez: se muestra un ejemplo de esta escala en el Gráfico 3A. Puede obtenerse fácilmente una escala intermedia de más de 10 valores alternando cuadrados rellenos con espacios vacíos. La Figura 3B da un ejemplo de este método, en que se emplea una escala construida para mostrar el movimiento en las carreteras u otros datos del tránsito.<sup>3/</sup> La Figura 3C muestra un patrón repetido de  $n = 63$  tomado de una escala detallada de 101 valores de 0 a 100;
- 3) representando los patrones aislados o repetidos en colores o en negro sobre un fondo de color, para que los patrones muestren los valores cuantitativos y los colores los atributos cualitativos.<sup>4/</sup>

<sup>3/</sup> No mostramos aquí una aplicación de este método, puesto que este trabajo no trata de las series geográficas que indican movimiento. La Oficina Central de Estadística de Israel ha hecho muchas aplicaciones prácticas de una escala semejante para representar el tráfico de carreteras y la frecuencia de los accidentes de carretera en las carreteras inter-urbanas en todo el país, y un perfil de la intensidad cambiante del tráfico en el transcurso de 24 horas en cada segmento de la carretera Tel Aviv-Haifa.

<sup>4/</sup> No se dan aquí ejemplos del 3) por motivos técnicos. Se espera que pronto podrá disponerse de material ya hecho que permita la preparación rápida de PRG (bajo limitaciones de patente y de propiedad literaria).

Describiremos ahora algunas aplicaciones de PGR Únicos:

A) Se presenta en el Gráfico 4 una aplicación a una frecuencia geográfica, que muestra la población de cada Estado de Estados Unidos para el Censo de 1950. Este tipo de representación se usará cuando se desea comunicar separadamente información sobre el tamaño total de la población de cada región. <sup>5/</sup>

B) Las distribuciones que se relacionan al  $\bar{X}_i$  promedio de una variable (que puede ser un índice, una razón, una tasa, un porcentaje, etc.) para la población  $P_i$  de cada región  $i$  (de área  $A_i$ ) pueden representarse por PRG individuales, según el método que se muestra en el Gráfico 5. En este mapa, los patrones indican el porcentaje  $\bar{X}_i$  de la población rural de cada uno de los Estados de Estados Unidos. Se aplica un patrón a cada unidad de 2 000 000 habitantes (que se indica por un marco cuadrado). De este modo el gráfico comunica la siguiente información: (i) Los promedios estatales  $\bar{X}_i$  (porcentaje rural); (ii) el tamaño de la población de cada Estado ( $P_i$ ); (iii) una impresión general de la densidad de la población en la superficie del Estado  $\frac{P_i}{A_i}$ ; (iv) la magnitud absoluta del fenómeno estudiado (número de la población rural)

$P_i \bar{X}_i$ ; y (v) una impresión general de la densidad del fenómeno  $\frac{P_i \bar{X}_i}{A_i}$ . La principal ventaja de este tipo

de representación es que da visualmente a  $\bar{X}_i$  su ponderación apropiada  $P_i$ . En el método de alternativa, que se emplea corrientemente para representar a  $\bar{X}_i$ , al cubrir toda el área  $A_i$  con un patrón aproximadamente proporcional a  $\bar{X}_i$ , el lector puede cometer inadvertidamente el error de ponderar visualmente  $\bar{X}_i$  según  $A_i$ , (en lugar de hacerlo según  $P_i$ ). Esto puede llevar a conclusiones sumamente erróneas: en nuestro ejemplo, el lector podría impresionarse más por los valores de  $\bar{X}_i$  en los Estados que tienen poca población y una superficie extensa, tales como los Estados Montañosos, que por los valores en los Estados Centrales del Atlántico con poblaciones muy grandes en superficies comparativamente pequeñas.

C) Estrictamente hablando, el método corriente de cubrir toda la superficie de cada región con un patrón proporcional a  $\bar{X}_i$  sólo debe usarse cuando  $\bar{X}_i$  es una razón de densidad (tal como  $\frac{P_i}{A_i}$ ). En este caso, puede recomendarse el uso de patrones repetidos (descritos en la forma "2" más arriba).

D) El Gráfico 6 muestra los porcentajes de población rural en cada una de las divisiones de Estados Unidos en 1940, 1960 y 1960. Este es un ejemplo de representación de distintos valores  $\bar{X}_i$  con respecto a cada región  $i$ . Estos valores pueden relacionarse con diferentes períodos, con diferentes clases cuantitativas o atributos cualitativos. El uso de un color diferente para el patrón que se relaciona con cada período, clase o atributo, facilita la lectura del mapa.

<sup>5/</sup> Cuando se desea una representación refinada que también muestre la distribución de la población dentro de cada región, debe preferirse el mapa punteado acostumbrado. Cuando el énfasis se coloca en la densidad de la población (por km. cuadrado) en cada región, se recomienda el uso de patrones repetidos proporcionales a la densidad y que cubren la superficie de cada región.

(c) Adaptación de la escala de símbolos gráficos a los datos

Algunos autores, que han tratado de resolver este problema en forma racional, (véase por ejemplo [6], págs. 175-180), toman como punto de partida la disponibilidad de un número muy limitado  $s$  de patrones discernibles (digamos:  $s=7$ ). Las regiones se ordenan de acuerdo a valores crecientes de  $\bar{X}_i$ , y se subdividen en  $s$  clases. Si se aprovecha el hecho que la escala de PRG forma prácticamente un continuo (como el número de patrones disponibles es muy grande), este procedimiento puede invertirse. Cualquier número deseado  $s$  de clases de regiones puede formarse con los intervalos de clase más convenientes, según el tipo de datos y la finalidad de la representación gráfica; y luego  $s$  patrones gráficos se seleccionan de modo conveniente, para representar los promedios  $\bar{X}_i$ , para cada clase de región. En forma más específica: - 1) A veces nos interesa un estudio detallado de las diferencias de  $\bar{X}_i$ , entre las regiones, aun cuando estas diferencias sean pequeñas. En este caso, habría que emplear muchas clases, con pequeños intervalos de clase, o incluso puede resultar conveniente usar un patrón característico para cada región. Por ejemplo, en el Gráfico 4 se emplea una escala con 201 patrones posibles, y la población de cada Estado de Estados Unidos se ilustra con una exactitud de  $\pm 50\ 000$ ; 2) En otros casos se desea una visión global de toda la distribución: entonces es preferible usar unas pocas clases con grandes intervalos de clase (véase, por ejemplo, el Gráfico 7).

Otro punto que debe tomarse en consideración es si el mapa se construye principalmente: (i) para el estudio de la distribución de cierto fenómeno en la superficie; o (ii) para la comparación de la distribución de muchos fenómenos X, Y, Z..., a representarse en el mismo mapa, o en mapas semejantes. En el caso (i), se puede utilizar escalas simples, por ejemplo la del Gráfico 4. En el caso (ii) puede asegurarse la comparabilidad utilizando escalas más sofisticadas. Por ejemplo:

- (A) si  $\bar{X}_i$  son porcentajes (o resultados numéricos, etc.) capaces de variar entre 0 y 100, cada porcentaje (expresado como un entero) puede representarse convenientemente por medio del patrón correspondiente (véanse Gráficos 5 y 6).
- (B) Otros datos  $\bar{X}_i$  pueden representarse convenientemente por el patrón más cercano al índice  $k \frac{\bar{X}_i}{\bar{X}}$ , calculado sobre la base de  $\bar{X}$  -el promedio ponderado para todo el territorio. Aquí  $k$  es una constante, tal como  $k = 1$  o  $k = 100$ . (Véase [4], figura 12).
- (C) El Gráfico 2 da un ejemplo de la transformación de  $X_i$  en un resultado numérico normalizado

$50 \left( 1 + h \frac{\bar{X}_i - \bar{X}}{\sigma_x} \right)$ , en que  $h$  es una constante y  $\sigma_x$  es la desviación normal ponderada de la  $\bar{X}_i$  regional con respecto a su promedio  $\bar{X}$ . Aquí se pone énfasis en el estudio del status relativo de la región  $i$  respecto a varios fenómenos, X, Y, Z..., después de hacerse un ajuste para los diferentes tipos de magnitud y variabilidad de estos fenómenos. Las desviaciones relativas de  $X_i$  respecto a  $\bar{X}$  pueden amplificarse, usando un valor grande para  $50h$  (en el Gráfico 2:  $h = \frac{1}{2 \cdot 5}$  y  $50 \cdot h = 20$ ), para poner énfasis en las comparaciones gráficas entre los "resultados numéricos" de X, Y, Z.



(D) El Gráfico 7 muestra una escala ideada para el estudio gráfico de la concentración geográfica (datos obtenidos de [7], para un proyecto de investigación realizado junto con el Profesor G. Tagliacarne). En nuestro ejemplo, se ha representado los ingresos medios per capita en las provincias de Italia en 1960-62: (i) para averiguar si existen provincias cuya proporción de los ingresos totales de Italia sea considerablemente en exceso de su proporción de la población total de Italia, y (ii) si estas provincias se hallan cercanas o distantes entre sí. Las provincias han sido ordenadas según los ingresos per capita y luego divididas en cinco clases, cada una de las cuales incluye alrededor de un quinto de los ingresos totales. Para cada clase  $c$  el índice  $k \frac{\bar{X}_c}{\bar{X}}$  (con  $k = 1$ ) ha sido calculado como en (B), comparando el ingreso medio per capita  $\bar{X}_c$  de la clase con el promedio nacional  $\bar{X}$ . El patrón que muestra este índice se presenta en el gráfico para todas las provincias que pertenecen a la misma clase  $c$ . Se da un patrón para cada 1 por ciento de la población total de Italia (o alrededor de 500 000 habitantes). La siguiente distribución de ingresos, que muestra una concentración considerable, llega a su expresión cabal en el gráfico.

Clase	Porcentaje del ingreso nacional en la clase	Porcentaje de la población de la clase real	Porcentaje de la población de la clase indicada por el número de patrones	Ingreso medio per capita anual de la clase (en miles de libras)	Índice del ingreso medio basado en el promedio nacional = 1,00
1	20,1	9,9	10	627	2,04
2	18,2	12,9	13	430	1,40
3	20,8	13,0	13	356	1,16
4	20,4	23,7	24	264	0,86
5	20,5	35,5	35	170	0,58
Total	100,0	100,0	100	307	1,00

Los centros de gravedad de los ingresos en cada clase se muestran en el Gráfico 7. Puede verse que las 5 clases se diferencian en forma marcada en cuanto a sus características geográficas. Las provincias con ingresos promedios altos se hallan muy concentradas en el Norte y el Oeste; aquellas que tienen ingresos promedios bajos se hallan dispersas en el Sur, Este y en las Islas. La "distancia normal" (standard distance) (medida de dispersión) de los ingresos, que se explica más adelante en las secciones 2 (b) y (c), aumenta de la clase 1 a la 5 (clase 1: 33 kms; 2: 167; 3: 169; 4: 247; 5: 259).

## 2. VALORES RESUMENES PARA LA DESCRIPCIÓN DE UNA DISTRIBUCIÓN ESTADÍSTICO-GEOGRÁFICA

La distribución de una población en un territorio puede demostrarse:

(i) indicando la ubicación exacta de cada uno de los  $n$  individuos que componen la población, o (ii) por agrupamiento, i.e., indicando el número de individuos enumerados ( $P_i$ ) en cada una de las regiones (de área  $A_i$ ) en que se divide el territorio. Entonces se puede calcular, entre otras cosas, las razones de densidad de la población  $P_i/A_i$ . Sólo se puede analizar detalladamente toda la información proporcionada por la condición (i) o incluso por la condición (ii) si manejamos una o pocas poblaciones, y también si el número de individuos que se incluyen en (i), o el número de regiones que se incluyen en (ii) no es demasiado grande. Sin embargo, el estudiante a menudo se ve enfrentado a distribuciones extremadamente detalladas, y con el problema de tener que comparar muchas distribuciones (tales como la distribución de una población dada en diferentes períodos, o el análisis de muchas sub-poblaciones, etc.). Entonces resulta conveniente condensar cada distribución estadístico-geográfica en expresiones o valores resúmenes que se puedan interpretar y comparar con facilidad, aun cuando esto ocasione una pérdida considerable de la información en detalle. Este método es semejante al del cálculo de valores resúmenes para las distribuciones cuantitativas y las series cronológicas.

Se han sugerido diversos métodos para lograr este objetivo, por ejemplo:

(i) En lugar de las razones de densidad  $\left(\frac{P_i}{A_i}\right)$  demasiado detalladas y a menudo ambiguas (véase [11], Capítulo 2) pueden usarse funciones, que miden en cierto modo la "influencia" que ejerce la población sobre cada punto del territorio. Stewart y colaboradores (véase [11] y [12]) han desarrollado y aplicado el concepto de "potencial de población" en el punto  $j$ , definido como 
$$\sum_{i=1}^n \frac{P_i}{d_{ij}}$$
 (en que  $d_{ij}$  es la distancia entre dos puntos  $i$  y  $j$  cualquiera).<sup>6/</sup>

(ii) Los  $n$  puntos de la distribución original se sustituyen por un pequeño número de puntos que la resumen. Sviatlovsky ha ensayado este método para calcular valores de partición para las distribuciones geográficas (véase por ejemplo [13]): desgraciadamente estos valores varían con los cambios en el sistema de coordenadas. Otro método en el que estamos trabajando actualmente es el de agrupar los  $n$  puntos en un número limitado  $s$  de sub-series, calculando los  $s$  centros de gravedad de los puntos incluidos dentro de cada sub-serie. Se realiza la partición tratando de maximizar la "distancia normal" (standard

<sup>6/</sup> Estamos tratando ahora de desarrollar (siguiendo líneas análogas) una fórmula algo diferente, que tiene la ventaja de evitar las dificultades que surgen del hecho que para  $i=j$ ,  $d_{ij} = 0$  y  $\frac{P_i}{d_{ij}}$  no es definible. Además, hay una cierta conveniencia en dividir los potenciales reales por los potenciales teóricos que se habrían encontrado si la población se hubiese distribuido en forma pareja por el territorio. Esto elimina la influencia del tamaño y la forma del territorio y los efectos de las fronteras, y recalca la influencia de la distribución de la población.

distance) (como se la define en las secciones (b) y (c) más adelante) entre los centros de las sub-series (para un valor fijo de s), minimizando así la distancia normal dentro de las sub-series. El Gráfico 4 muestra, por ejemplo, un intento de dividir Estados Unidos en tres regiones, y calcular los centros de gravedad de la población incluida en cada región (1950). La distancia normal (standard distance) de la población dentro de cada región es, como promedio, de 500 kms., en comparación con la distancia normal de 1 247 kms. de toda la población respecto a su centro. De este modo se da cuenta de alrededor del 70 por ciento del cuadrado de la distancia global respecto al centro nacional al describir la distribución por medio de los tres centros regionales.

(iii) Mediante el cálculo de indicadores que expresan las características principales de la distribución, como son las siguientes: a) ubicación; b) dispersiones; c) dirección principal; d) asimetría; e) concentración. Las siguientes secciones describen algunos estudios recientes respecto a cada una de estas características.

#### (a) Ubicación

Son conocidos los métodos para indicar la "ubicación" general de una distribución estadístico-geográfica por medio de un "centro". Gini y Galvani han discutido ampliamente los problemas relacionados con el "centro de gravedad" y el "centro mediano" (véase [14], [15]). Estos centros se definen como los puntos del plano desde los cuales la suma de los cuadrados de las distancias a los puntos de la distribución, o la suma de las distancias, respectivamente, es mínima. H. Wernitz y D. Neft han estudiado el centro que minimiza la media armónica de las distancias, i.e. maximiza el potencial medio (véase [17]). Entre estos centros, el más usado es el centro de gravedad, que presenta la ventaja de ser fácil de calcular. Sin embargo, la principal característica del centro mediano, a saber, la de minimizar la suma de las distancias, podría ser útil tanto para propósitos descriptivos como también para solucionar problemas prácticos de ubicación de fábricas, tiendas, oficinas, etc. Por otra parte, la comparación entre el centro de gravedad y el centro mediano parece ser de gran interés, y las mediciones de la dispersión alrededor del centro mediano son valiosas (véase (b), más abajo). El centro mediano por lo tanto merece mayor atención y aplicación que la que ha tenido hasta aquí. Puesto que los motivos de este descuido probablemente se relacionan con las dificultades de su cálculo (véase, por ejemplo [16]), se han realizado esfuerzos para hallar una solución a este problema. B. Schaerf (Israel) ha desarrollado un método para el cálculo rápido del centro mediano por medio de una computadora electrónica [18] [40]. Se ha aplicado este cálculo a poblaciones humanas, distribuciones agrícolas y a otras distribuciones sociales y económicas en diversos países (véase [18]).

#### (b) Dispersión

Pueden utilizarse diversos métodos para medir la dispersión o diseminación de una distribución por su territorio, tales como : (i) calcular el promedio de la distancia entre los puntos; (ii) calcular el

promedio de la distancia entre cada punto y el centro de gravedad; (iii) calcular el promedio de la distancia entre cada punto y el centro mediano; (iv) medir los radios de los círculos cuyos centros se hallan en la mediana e incluir proporciones dadas de la distribución. El autor y colaboradores han hecho muchas aplicaciones de los métodos (ii), (iii) y (iv), tanto para medir la dispersión de los fenómenos en áreas comparativamente pequeñas consideradas como superficies planas (véase [10] y [18]) y para superficies mayores en las cuales no se puede dejar de tomar en cuenta la esfericidad de la tierra (véase [19]). En particular se ha consagrado mucha atención a la "distancia normal" (standard distance) o cuadrado del promedio de las distancias entre los puntos y el centro de gravedad. Se han examinado diferentes definiciones y propiedades para esta medición; se ha calculado su valor para diversos modelos teóricos, y se han propuesto métodos para medir separadamente los principales factores que la afectan. Se han realizado aplicaciones de la distancia normal (standard distance) a la medición de los cambios en la dispersión de poblaciones enteras, de grupos étnicos y socio-económicos, y de funciones económicas por zonas urbanas, o países, y también al estudio de la distribución geográfica de cultivos agrícolas y de diferentes tipos de industrias, etc. Además, se ha empleado esta medición para evaluar las influencias características del crecimiento natural, migraciones externas e internas, sobre el cambio en la dispersión de la población (véase [10] y [18]).

Se ha elaborado un programa (para una computadora electrónica IBM 7401) para calcular el centro de gravedad, la distancia normal (standard distance) y algunas de las mediciones que se describen en la letra (c) más abajo. Este programa se ha usado para calcular estos valores para muchas distribuciones geográficas de las características demográficas, económicas y sociales de la población de Israel, de acuerdo al Censo de Población de 1961. Se han descubierto interesantes características de estas distribuciones, que no hubiesen emergido en forma clara de las incómodas distribuciones, detalladas de acuerdo a 370 lugares habitados ó 35 regiones naturales, (por ejemplo: la disminución de la dispersión geográfica de los grupos de población de acuerdo al tamaño creciente del lugar habitado, mayor permanencia en el país, mayor edad, y nivel educacional más elevado). Se han encontrado regularidades semejantes con respecto a la dispersión de los grupos de población clasificados según edad, período de inmigración, y densidad por pieza en las zonas urbanas de Tel Aviv - Yafo y Jerusalén.

Para un análisis más detallado de las distribuciones geográficas de ciertos fenómenos, el método (iv) (radios de los círculos) ha resultado particularmente apropiado. Este método se ha aplicado al estudio

de la distribución de ciertas funciones económicas en los países (tales como Italia) y en las ciudades (tales como Tel Aviv-Yafo). Como un ejemplo, citaremos del estudio de Tel Aviv (elaborado por A. Shahar) la longitud (en metros) de los radios de los círculos que tienen su centro en el centro mediano del fenómeno estudiado, que incluyen respectivamente 1/4, 1/2 y 3/4 de la población apropiada: -las compañías de seguros: 20, 340, 390; los hoteles: 30, 640, 1 290; las oficinas de los ingenieros y arquitectos: 600, 990, 1 360; los restaurantes: 790, 1 390, 2 130; las farmacias: 730, 1 430, 2 210; las barberías: 1 030, 1 570, 2 550; los cines: 640, 1 370, 2 460; la zona municipal: 2 510, 3 770, 4 650; la población: 1 430, 2 190, 3 410.

(c) La principal dirección de la distribución, y la representación gráfica de los valores resúmenes

Consideremos nuevamente una serie de puntos  $i$  (o individuos de una población) en que las coordenadas  $(X_i, Y_i)$  se refieren a un sistema ortogonal de ejes. Supongamos además que los ejes tienen su origen en el centro de gravedad. Rotemos ahora los ejes de las coordenadas alrededor del centro de gravedad en sentido contrario al de las manecillas del reloj y señalemos las nuevas coordenadas del punto  $i$  por medio de  $(X'_i, Y'_i)$ . Que  $\sigma_x, \sigma_y$  indiquen las desviaciones normales (standard deviations) de  $X$  e  $Y$  y  $\sigma'_x, \sigma'_y$ , las desviaciones normales (standard deviations) de  $X'$  e  $Y'$ . La distancia normal (standard distance)

$d = \sqrt{\sigma_x^2 + \sigma_y^2}$  es constante con respecto a la rotación de los ejes de las coordenadas, pero  $\sigma'_x$  y  $\sigma'_y$  no lo son. Es interesante calcular el ángulo de rotación (en el radio de acción  $-45^\circ \leq \alpha' \leq 45^\circ$ ) que hace que uno de estos valores sea mínimo ( $\min \sigma$ ), y que el otro sea entonces máximo ( $\max \sigma$ ). Linders (véase [20]) ha demostrado que  $\min \sigma$  es útil para describir una distribución geográfica: el "eje principal", con respecto al cual el cuadrado del promedio de la distancia de los puntos de distribución es  $\min \sigma$ , puede considerarse que demuestra la "dirección principal" de la distribución. Podemos considerar a  $\min \sigma$  y  $\max \sigma$  como las dos principales componentes de la distancia normal (standard distance)  $d$ , puesto que indican la dispersión de los puntos en las dos direcciones en que la dispersión es respectivamente mínima y máxima.

Los Gráficos 8 y 9 muestran la presentación de un cuadro resumen de estas características principales de una distribución geográfica por medio de una cruz, cuyo centro está en el centro de gravedad, cuyos cuatro brazos descansan sobre los dos ejes principales y son de longitud  $\sigma'_x$  y  $\sigma'_y$ , respectivamente. Los dos ejes están situados en un ángulo  $\alpha'$  respecto a los ejes de coordenadas originales. Este tipo de representación es útil, por cuanto permite la presentación simultánea en el mismo mapa del resumen de muchas distribuciones. El Gráfico 8 presenta, como ejemplo, la distribución de las zonas dedicadas a diversos cultivos agrícolas en Francia; se han agregado puntos que representan una información detallada de la distribución de la remolacha ("betterave industrielle") - (tomado de [21]), para permitirle al lector comparar el cuadro detallado con el resumen que proporciona la cruz.

En el ejemplo del Gráfico 9, se presentan algunas distribuciones de funciones económicas en la zona Tel Aviv - Yafa; se observa que las tiendas de comestibles siguen de cerca la distribución geográfica de la población; los abogados constituyen un grupo muy concentrado, con sus centros en la "Ciudad"; los hoteles también se hallan concentrados, pero menos que los abogados, y su centro se aleja de la "Ciudad" en dirección al mar. La población está mucho menos diseminada que en la hipótesis de la equi-distribución por toda la zona municipal, y también se haya más densamente distribuida en el Sur.

(d) Asimetría

Se ha ideado una medición de esta característica, que varía desde el 0 por ciento en el caso de una simetría perfecta hasta un máximo de 100 por ciento. Como ejemplo, el índice de asimetría para la distribución de la población de Estados Unidos en 1960 (véase Gráfico 4) sería del 27,0 por ciento, en comparación con un índice de asimetría del 3,9 por ciento en el caso de una equi-distribución de la población en la superficie de Estados Unidos (sin incluir Alaska, Hawai, ni Puerto Rico).

(e) Concentración

Los índices de concentración y los índices de segregación de Gini y Lorentz han sido ampliamente aplicados al estudio de las distribuciones estadísticas geográficas (véase [II]). Desgraciadamente estas mediciones son variables con respecto a los cambios en el sistema de subdivisión del territorio (por ejemplo, por Estados o por condados en Estados Unidos); por otra parte, no son sensibles a la concentración "geográfica", por cuanto proporcionan valores iguales sea que los distritos en que el fenómeno se concentra en su mayor parte se hallan cercanos o distantes entre sí (ceteris paribus). A fin de superar estas dificultades, se han empleado dos líneas diferentes de ataque en la investigación que se lleva a cabo en el momento que escribo: (i) el cálculo de un índice basado en el cuadrado del promedio de las distancias que deben cubrirse para diseminar la población estudiada en forma pareja por el territorio; (ii) El uso de la función potencial modificada sugerida anteriormente (sección 2, i).

3. EL ESTUDIO DE LA DISTANCIA ENTRE DOS DISTRIBUCIONES GEOGRAFICAS

Al comparar la distribución geográfica de dos poblaciones, se ha usado el "cuadrado de la distancia media" (<sup>2</sup>D) entre ellas. Este se define como la raíz cuadrada del "viaje" medio por unidad de población que se requiere para transportar una población a otra, siendo medido el viaje de cada unidad por el cuadrado de la distancia que tiene que cubrir. Este transporte puede lograrse de diversas maneras, por ejemplo:

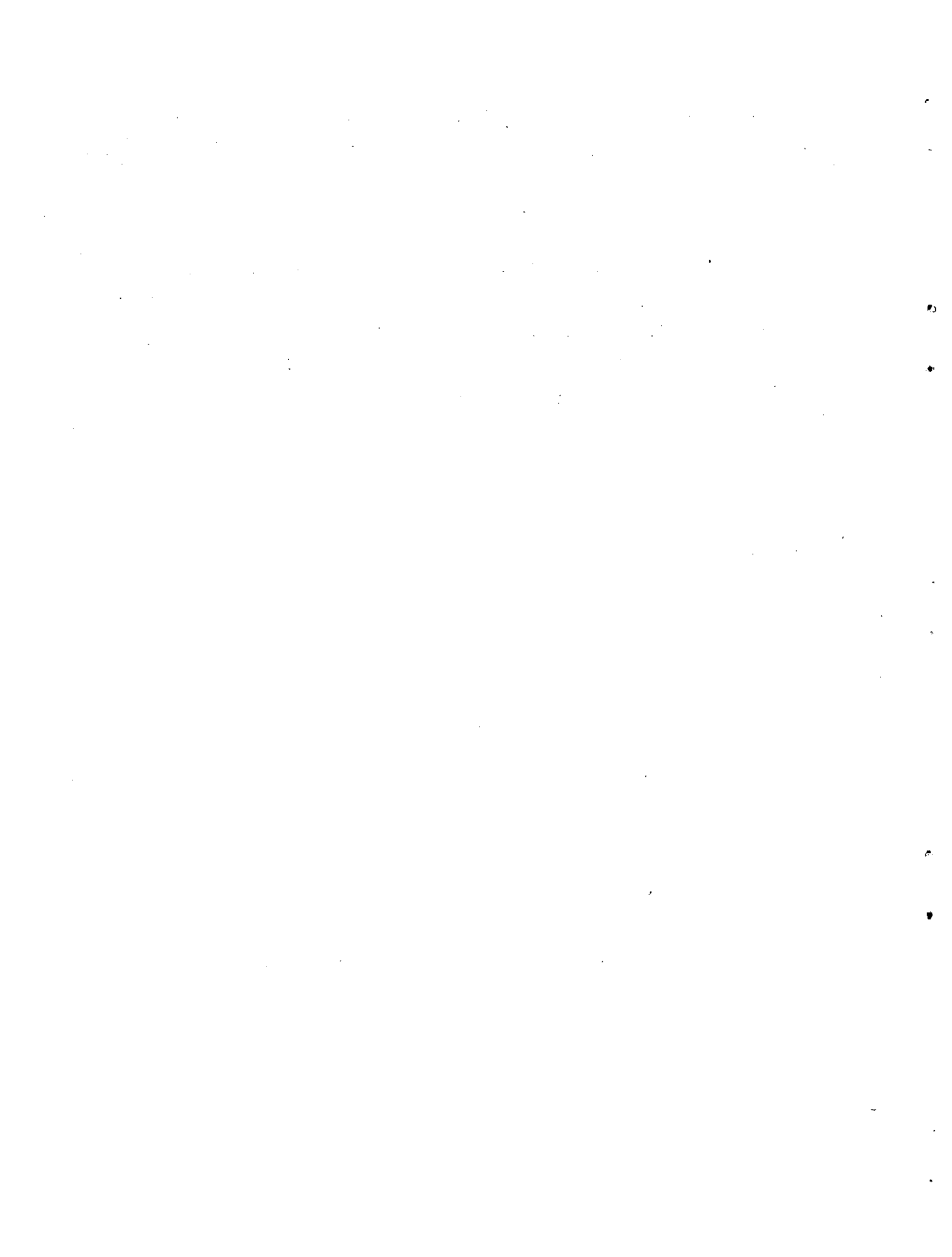
- a) En la migración interna los migrantes viajan de un punto a otro. Calculamos por lo tanto el cuadrado del promedio de las distancias (a vuelo de pájaro) para los movimientos de los migrantes;
- b) En otros casos, podemos definir este transporte como aquél que minimiza la suma del cuadrado de las distancias desde los puntos de una población a otra. Si bien se dispone de una solución sencilla para

calcular la distancia entre series unidimensionales [22], no parece disponerse de una solución sencilla para calcular tales distancias para las series bi-dimensionales, y por lo tanto se ha adaptado un algoritmo de transporte de programación lineal para la computación de este valor (en una computadora IBM 1401 y 1520). Se realizan investigaciones respecto a la posibilidad de calcular en forma más sencilla aproximaciones de  $^2D$ .

Este valor ha sido utilizado: a) para medir las distancias entre distintos grupos étnicos que viven en el mismo territorio; b) para comparar la distribución real en un territorio con una equi-distribución (principalmente para estudiar si la población tiende a diseminarse más o a concentrarse más con el tiempo); c) para analizar la relación entre la distribución geográfica de dos variables; d) para analizar movimientos tales como migraciones internas (véase [3], [9], [10], [13] <sup>7/</sup>).

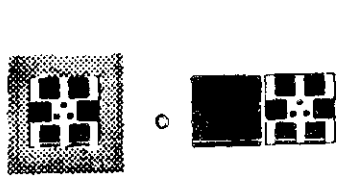
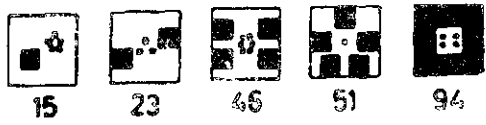
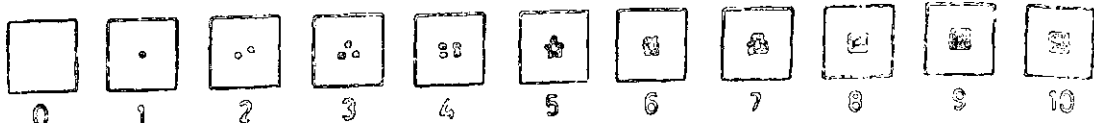
---

<sup>7/</sup> Parte de la investigación resumida en las secciones 2 y 3 fue apoyada por donaciones de la Fundación Ford.

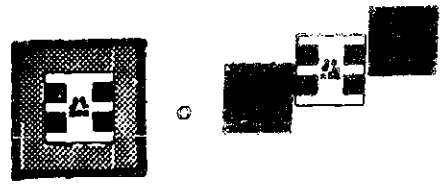




# I. PATRONES UNICOS



163



247

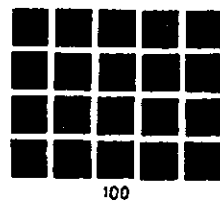
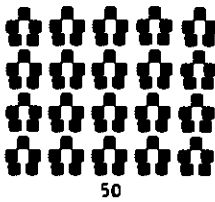
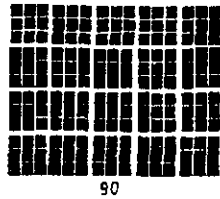
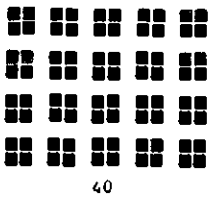
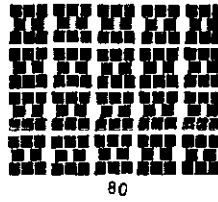
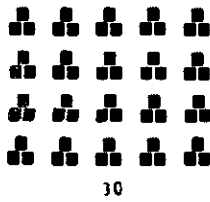
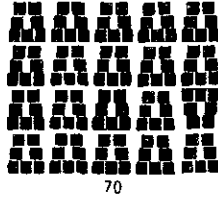
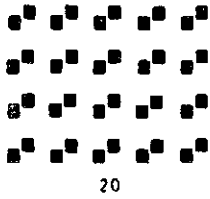
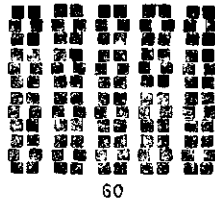
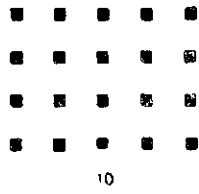
2. INDICADORES<sup>a/</sup> DE CARACTERISTICAS SOCIALES Y ECONOMICAS  
EN LAS DIVISIONES DE ESTADOS UNIDOS, 1960

División	(a) Ingresos de 10 000 dólares y más	(b) Ocupacio- nes no manuales	(c) Años de escuela comple- tados	(d) 4 años ó más de escuela media	(e) Ingresos de 3 000 dólares ó menos	(f) Empleo agrícola	(g) Población rural
Pacífico							
Atlántico Medio							
Este del Norte central							
Nueva Inglaterra							
Montaña							
Oeste del Norte central							
Atlántico Sur							
Oeste del Sur central							
Este del Sur central							

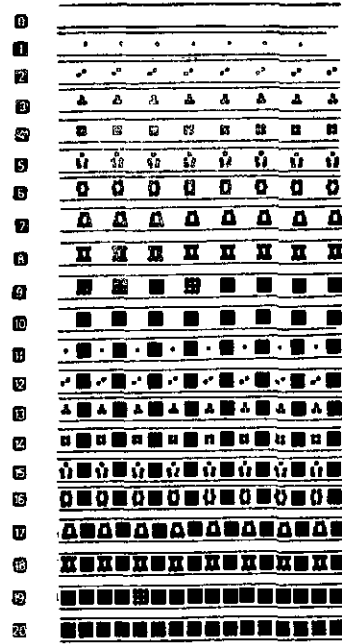
a/ Los valores originales  $X_i$  se transforman en resultados numéricos

$$S_i = 50 + 20 \frac{X_i - \bar{X}}{\sigma_x}$$

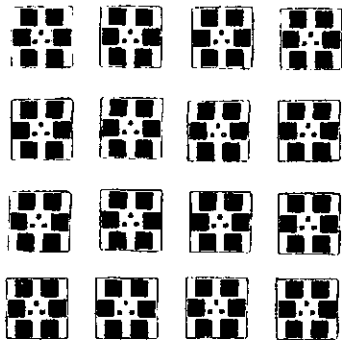
### 3. PATRONES REPETIDOS



A

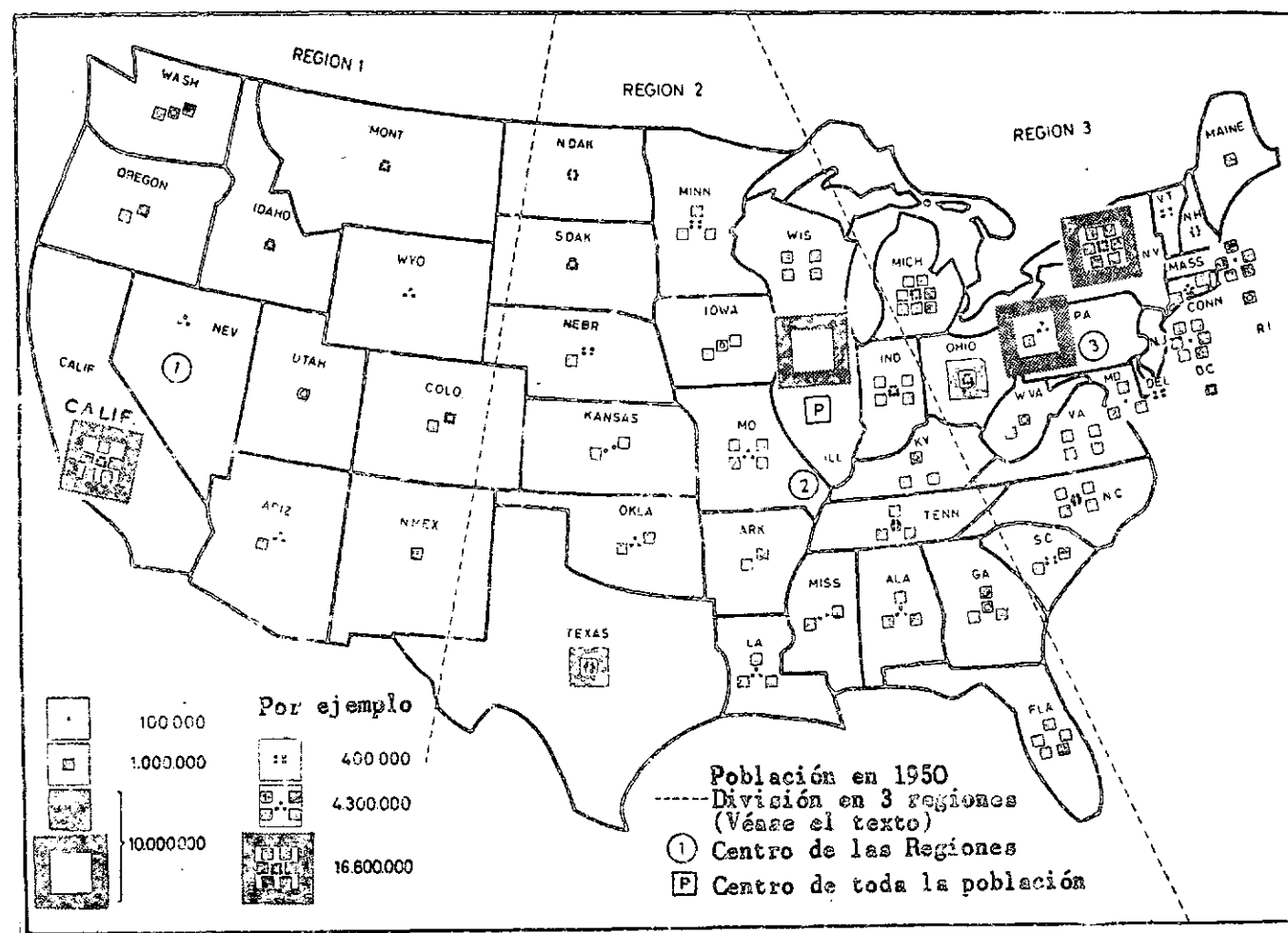


B

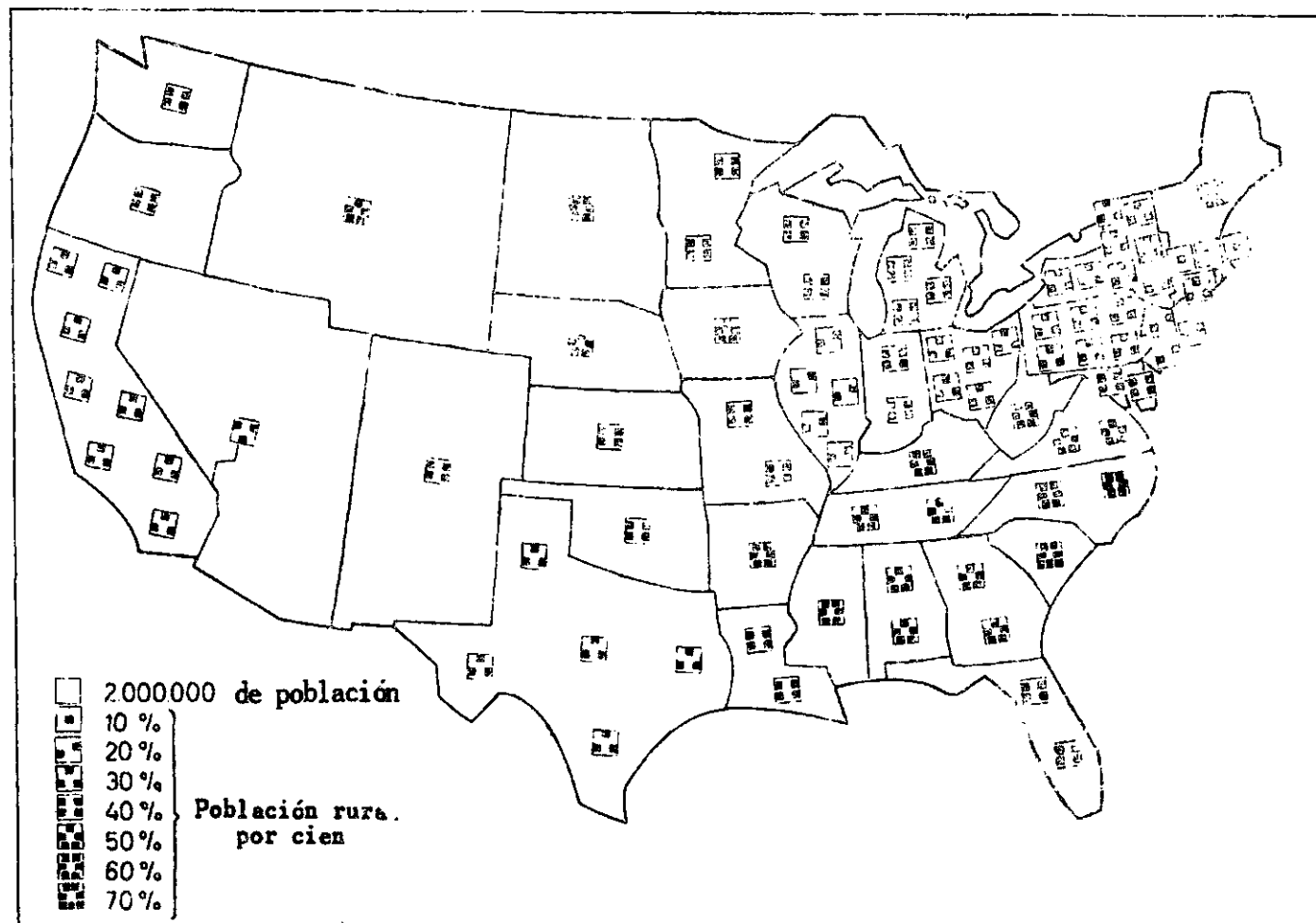


C

4. ESTADOS UNIDOS: POBLACION, 1960

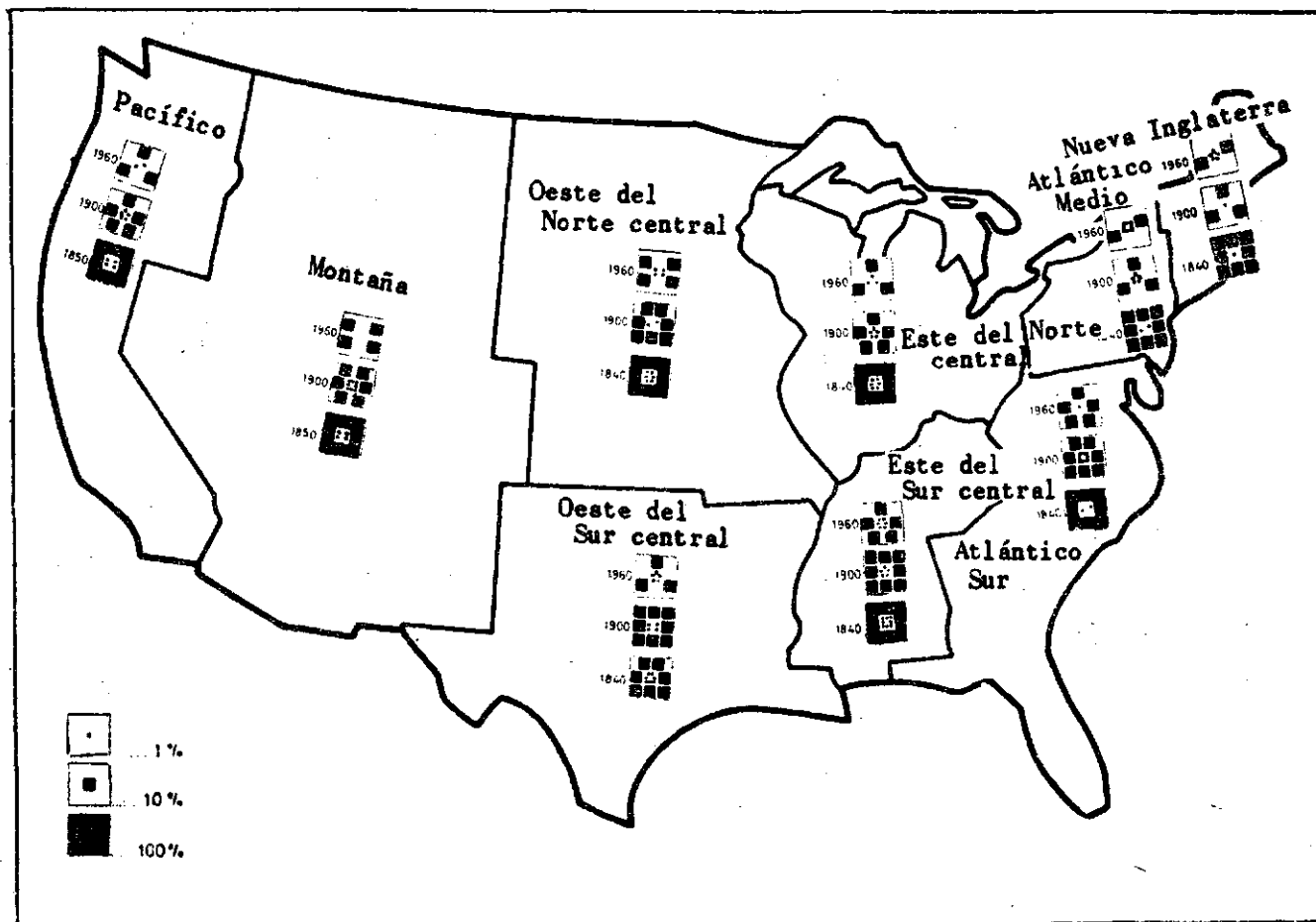


5. ESTADOS UNIDOS: PORCENTAJE DE POBLACION RURAL<sup>a/</sup>, 1960

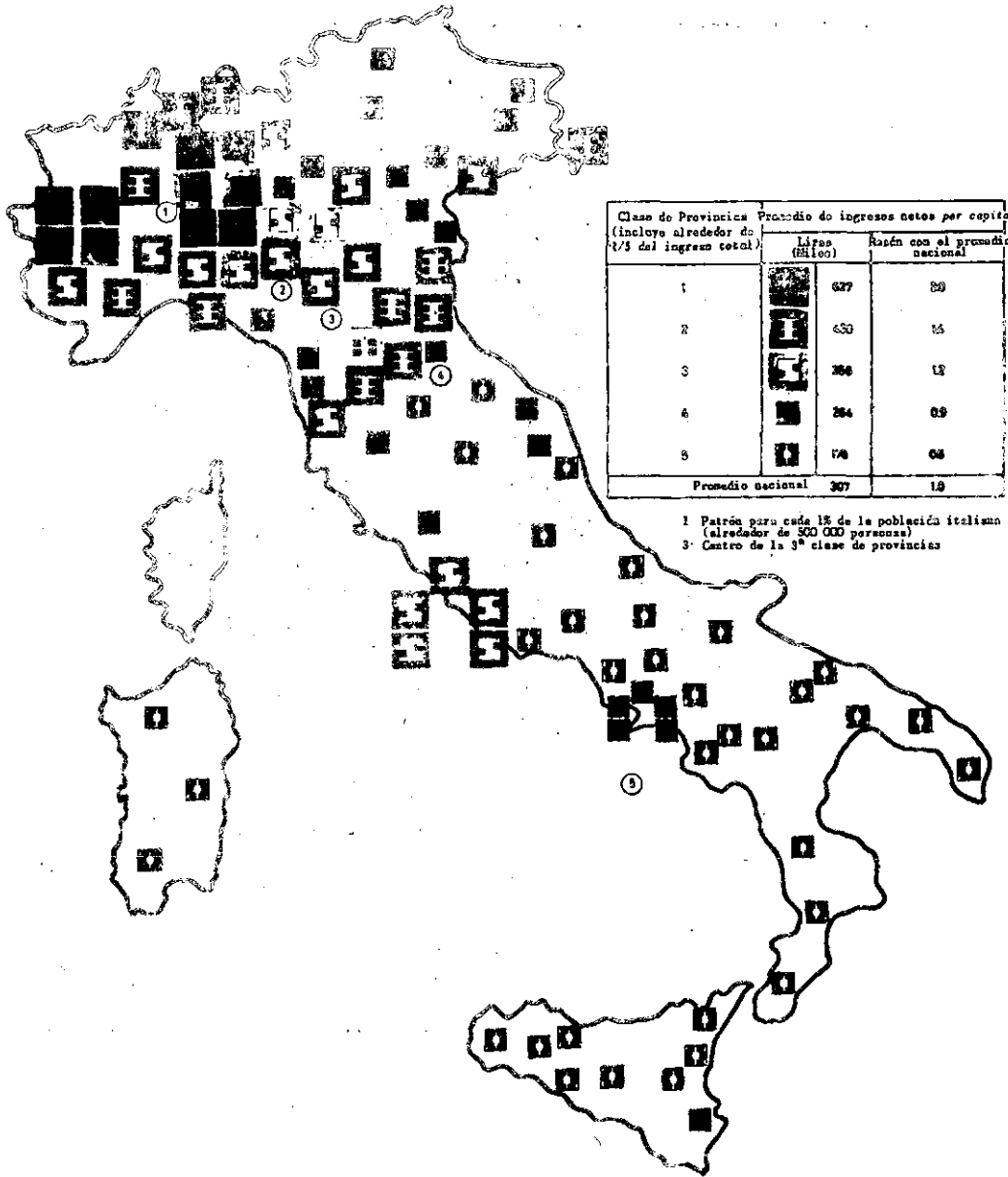


a/ Antigua definición.

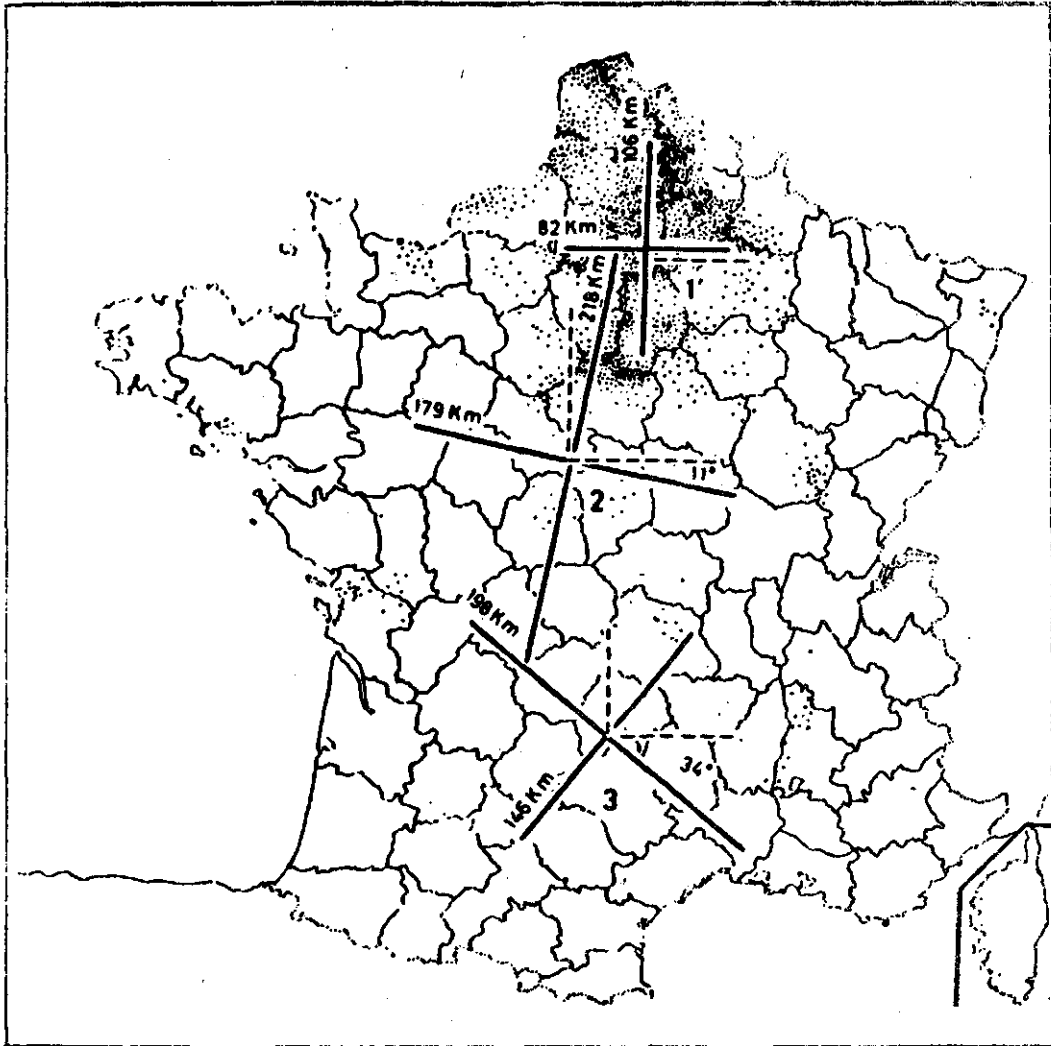
6. PORCENTAJE DE POBLACION RURAL EN CADA DIVISION DE ESTADOS UNIDOS, 1840-1960



7. ITALIA, 1960-62



## 8. REMOLACHA, TRIGO Y VIÑAS



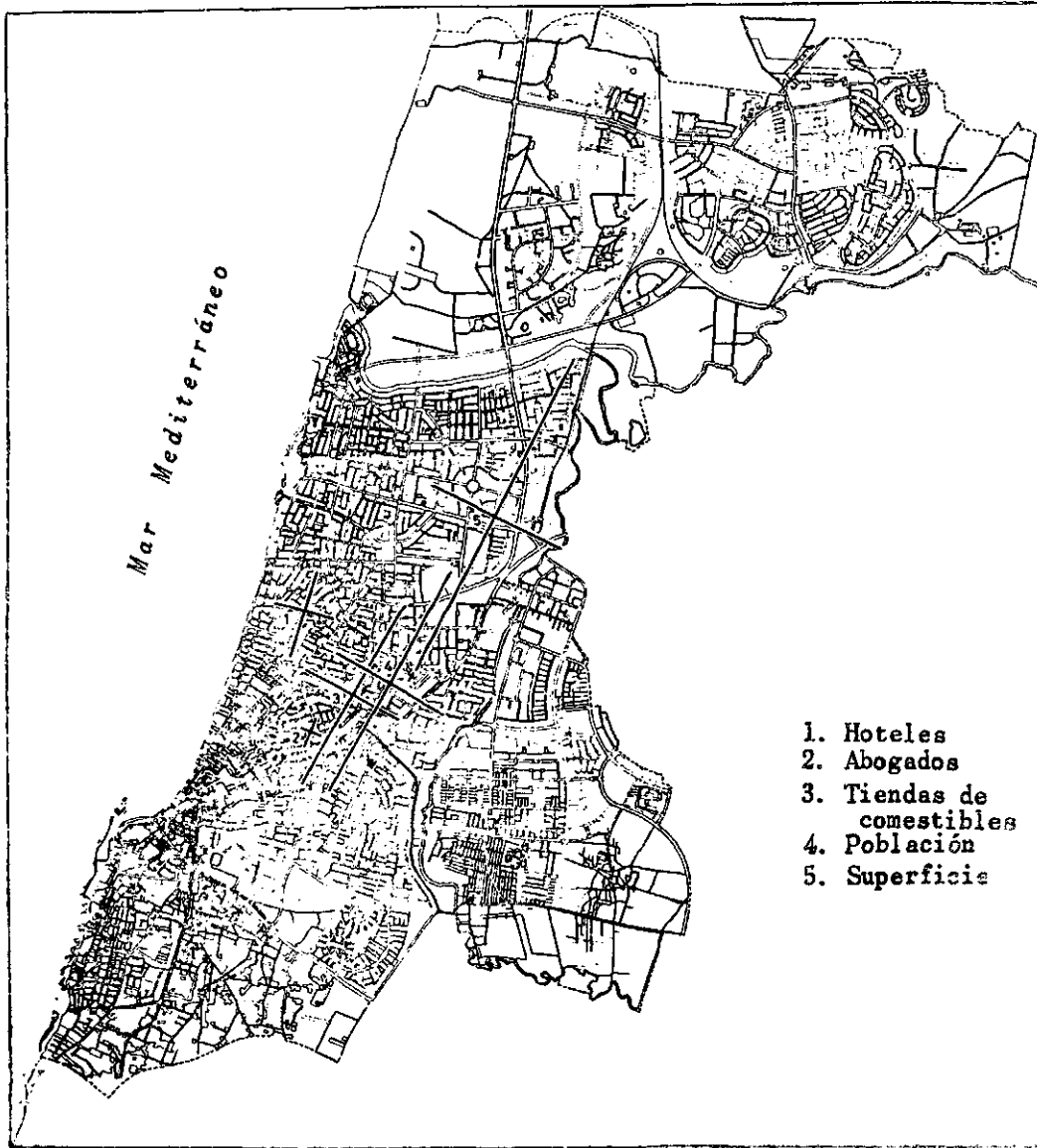
1. Remolacha

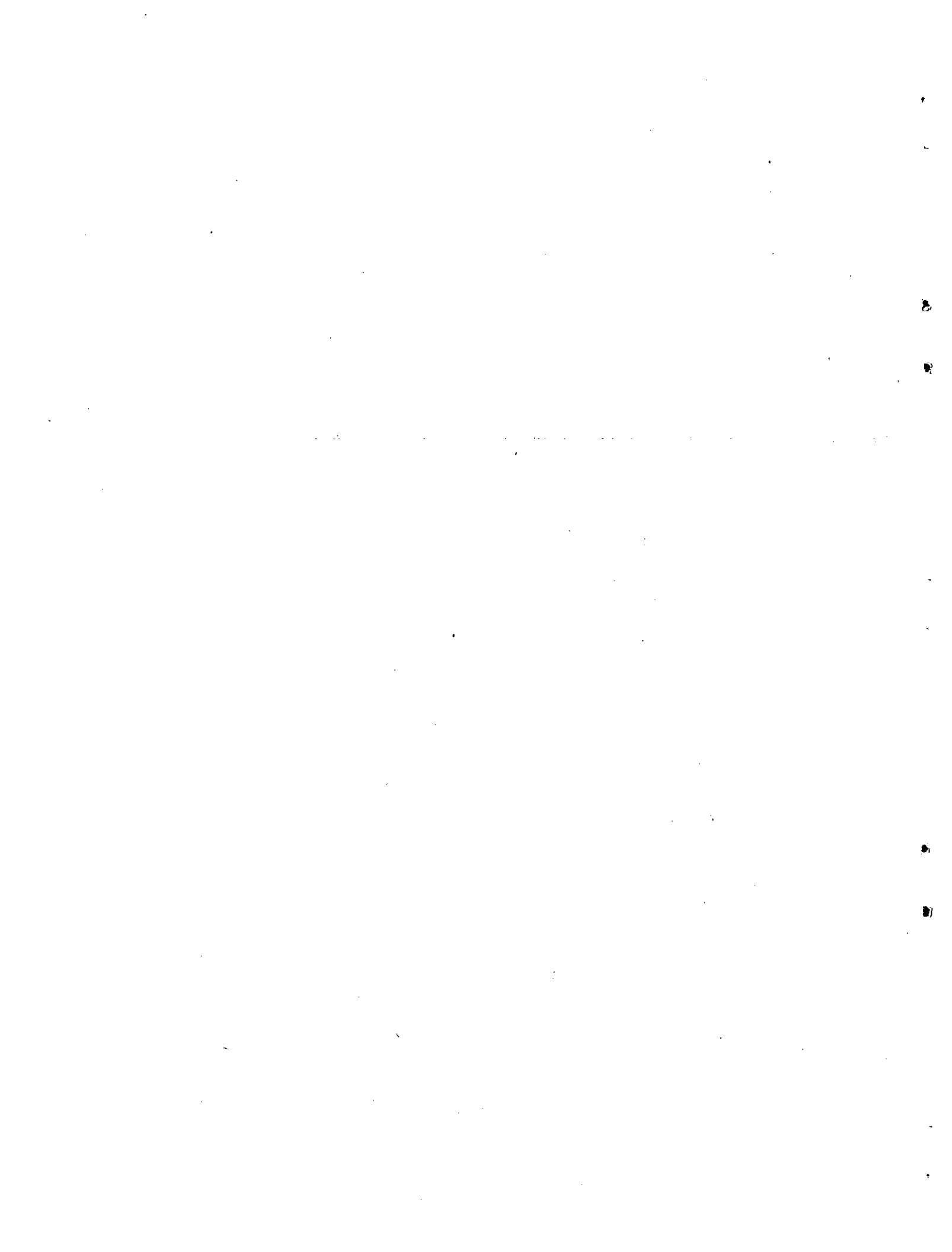
2. Trigo

3. Viñas



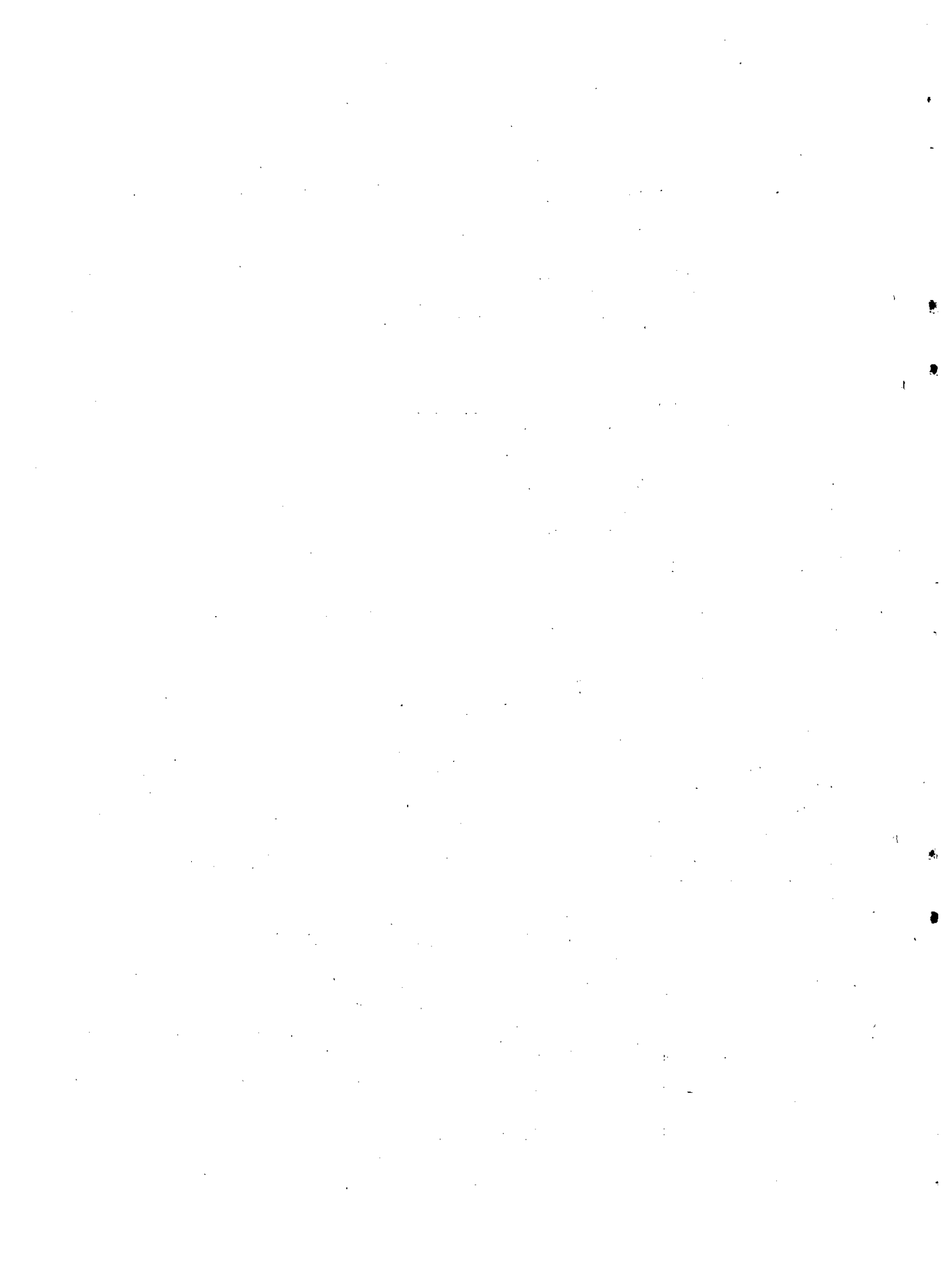
9. TEL AVIV — YAFO

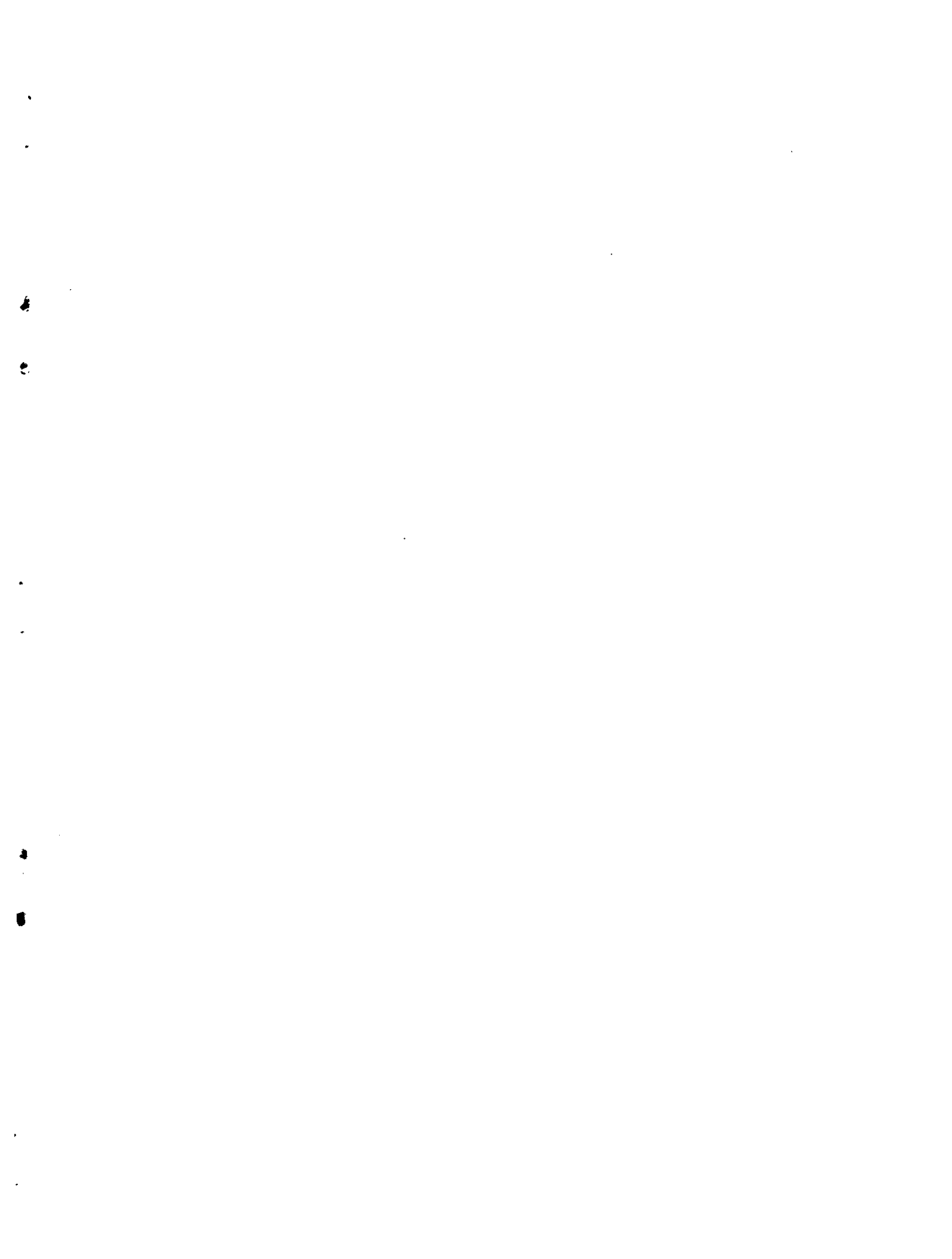


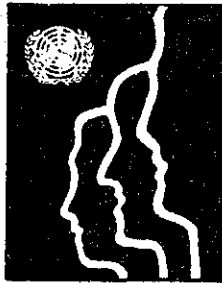


A P E N D I C E  
REFERENCIAS BIBLIOGRAFICAS

1. Dickinson G.C., Statistical mapping and the presentation of statistics, Londres, Arnold, 1963;
2. a) Williams R.L., Map symbols: equal-appearing intervals for printed screens. Annals, Assoc. of American Geogr., Vol. 48 (1958), p. 132-139;  
b) Robinson A.H., The curve of the grey spectrum : a review. Annals, Assoc. of American Geogr., Vol. 49 (1959), p. 457-460;  
c) Williams R.L., Map symbols: the curve of the grey spectrum - an answer. Annals, Assoc. of American Geogr., Vol. 50 (1960), p. 487-491;  
d) Jenks F. & Knos D.S., The use of shading patterns in graded series. Annals, Assoc. of American Geogr., Vol. 51 (1961), p. 316-334;
3. Bachi, R., The graphical representation of geographical series. Bulletin of the International Statistical Institute. Tomo XXXIV - 4<sup>a</sup> Parte. Roma, 1955;
4. Bachi R., Some methods for the graphical representation of statistical data. Bulletin of the International Statistical Institute - Tomo XXXIX, 2<sup>a</sup> Parte. París, 1962;
5. Discussion of paper [4]. Tomo XXXIX, 1<sup>a</sup> Parte, pp. 111-113;
6. Pépe P., Présentation des statistiques. París, Dunod, 1959;
7. Tagliacarne G., Calcolo del reddito prodotto dal settore privato e dalla pubblica amministrazione nelle provincie e regioni d'Italia nel 1962 e confronto col 1961. "Moneta e credito", septiembre 1963;
8. Bachi R., Statistical analysis of geographical series. Bulletin of the International Statistical Institute. Vol. 36, 2<sup>a</sup> Parte, Estocolmo, 1958;
9. Bachi R., Some methods for the study of geographical distributions of internal migrations. Trabajo presentado a la International Population Union Conference, Nueva York, 1961;
10. Bachi R., Standard distance measures and related methods for spatial analysis. Regional Science Association; Papers, X, Congreso de Zurich, 1962;
11. Duncan O.D., Cuzzort R. P., Duncan B., Statistical geography, The Free Press of Glencoe, Illinois, 1961;
12. Isard W., Methods of regional analysis, Wiley J., Nueva York, 1960;
13. Sviatlovsky E. E., y Ellis W.C., The Centrophical method and regional analysis. Geographical Review. Vol. 27, Nº 2, 1937;
14. Galvani L., Sulla determinazione del centro di gravità e del centro mediano di una popolazione. Metron. Vol. XI, Nº 1, 1933;
15. Gini C., Boldrini M., Galvani L., Venere A., Sui centri delle popolazioni e sulle loro applicazioni. Metron. Vol. XI, Nº 2, 1933;
16. Austin T.L., An approximation to point of minimum aggregate travel. Metron. Vol. XIX, Nº 3-4, 1959;
17. Warntz W., Neft D., Contributions to a statistical methodology for areal distributions. Journal of Regional Science. Vol. 2, Nº 1, primavera 1960;
18. Bachi R., Methodi di sintesi delle serie geografiche, con applicazioni a statistiche italiane. Scritti di economia e statistica in memoria di A. Molinari, Milán, Giuffrè, 1963;
19. Shimony Z., Theoretical research in the field of statistical geographical series. Tesis de Ph. D., Universidad Hebrea, Jerusalén, 1964. (En hebreo con un resumen en inglés);
20. Linders F., Ueber die Berechnung des Schwerpunktes und der Tragheitsellipse einer Bevölkerung. Metron. XI, 1, 1933;
21. Klatzmann J., La localisation des cultures et des productions animales en France. París, Imprimerie Nationale, 1953;
22. Samuel E. y Bachi R., Measures of distances of distribution functions and some applications. Metron. XXIII, 1-4, 1964.







**CENTRO LATINOAMERICANO DE DEMOGRAFIA  
CELADE**

*Sede:* J.M. Infante 9. Casilla 91. Teléfono 257806  
Santiago (Chile)

*Subsede:* Ciudad Universitaria Rodrigo Facio  
Apartado Postal 5249  
San José (Costa Rica)