

CONCENTRACIÓN DEMOGRÁFICA Y ESTRUCTURA URBANA: UN ENFOQUE *VÍA* TEORÍA DE LA INFORMACIÓN

PEDRO URIBE JR.
El Colegio de México

1. INTRODUCCIÓN

Recientemente, H. Theil, J. Hooper, M. Scholes y el autor (ver referencias [1], [6] y [7] al final de este artículo) han propuesto y aplicado a problemas económicos, medidas de concentración basadas en conceptos de la teoría de la información. En este trabajo se aplican medidas de este tipo a fenómenos demográficos; en particular, se propone un método para la medición y análisis de la concentración geográfica de la población y de su concentración en núcleos urbanos.¹

Las medidas tradicionales de concentración pueden suponer o no una forma particular de la distribución de frecuencias del fenómeno al que se aplican; se acostumbra llamar paramétricas a las que dependen de la distribución y no paramétricas o libres de distribución a las demás. Por ejemplo, el coeficiente de Pareto, comúnmente usado en la medición de la desigualdad del ingreso, es el parámetro de la llamada distribución de Pareto; el coeficiente de Gini, también de uso frecuente, es una medida no paramétrica de concentración.

Una de las características interesantes del enfoque paramétrico es su relación con proposiciones o hipótesis sobre la dinámica del fenómeno investigado. Por ejemplo, se sabe que ciertos procesos de Markov dan origen a una distribución estable del tipo Pareto; el proceso de "crecimiento proporcional" de Galton, Kapteyn y Gibrat da origen a una distribución estable normal en el logaritmo de la variable en cuestión. El método propuesto en este trabajo es un método *no paramétrico*, y no se intenta un análisis dinámico, sino de estática comparada. Sin embargo, se conjetura que el enfoque presentado en este artículo puede llevar a resultados interesantes en el análisis dinámico.²

Limitando la discusión al análisis estático de la concentración, la

¹ El autor agradece la colaboración de Adalberto García Rocha y Raúl de la Peña, del Centro de Estudios Económicos y Demográficos de El Colegio de México, y del Dr. Rómulo H. González Zubieta, de CEIR de México, S. A., en la programación y el cómputo del material empírico de este trabajo en la computadora GE415 de CEIR.

² Esta conjetura se basa en que el *cambio* en la información indirecta (medida de concentración usada en este trabajo) es, aproximadamente, otra información indirecta. El punto se trata más adelante.

siguiente propiedad es de considerable importancia para fines analíticos. Supóngase que se desea medir la concentración de la población en localidades urbanas en n estados que forman el país. La medida de concentración para el país es, digamos, c y las de los estados son c_1, c_2, \dots, c_n . ¿Qué relación existe entre c y c_1, c_2, \dots, c_n ? Para las medidas paramétricas, la respuesta depende de que la distribución de la variable en cuestión en los estados y en el agregado sean de la misma forma y de que la relación entre los parámetros de los estados y los del agregado sea suficientemente simple. En el caso de los enfoques paramétricos más comúnmente usados, Pareto y lognormal, la primera condición no se cumple: una media aritmética (ponderada)³ de variables Pareto no es Pareto y la de variables lognormales no es lognormal. Por el lado de las medidas no paramétricas, resulta claro que el coeficiente de Gini, construido sobre la base de diferencias absolutas, es inadecuado para descomposición y agregación. La varianza de la variable estudiada, aunque poco usada como medida no paramétrica de concentración, es susceptible de análisis; se verá más adelante que las medidas de información simplifican el problema dimensional del análisis de varianza.

Aunque originalmente fraseada en el lenguaje de la ingeniería de comunicaciones, la teoría de la información⁴ puede verse como una rama de la teoría de la probabilidad. Parte de la idea de que un mensaje cuyo contenido es *a priori* poco probable, contiene más información que otro cuyo contenido es, también *a priori*, más probable. La información esperada de un experimento es, entonces, función del grado de incertidumbre asociado, antes de llevarlo a cabo, a sus posibles resultados. Resulta así que, si uno de estos resultados es cierto *a priori*, la información esperada es cero; si todos los resultados tienen *a priori* la misma probabilidad de ocurrir, la incertidumbre —y por tanto la información— es máxima. Es posible demostrar⁵ que, dadas ciertas condiciones muy generales, la medida de información esperada de un experimento cuyos posibles resultados tienen probabilidades *a priori* p_1, \dots, p_n está totalmente determinada hasta unidades de escala y está dada por:

$$H = \lambda \sum_{j=1}^n p_j \log \frac{1}{p_j}$$

en que λ es una constante, que define las unidades y que depende de la

³ El tamaño de las localidades en el país se distribuye como la variable

$$\frac{k_1 X_1 + k_2 X_2 + \dots + k_n X_n}{K}$$

en que k_j es el número de localidades en el estado j , $K = k_1 + k_2 + \dots + k_n$, y la variable aleatoria X_j es el tamaño de las localidades en el estado j .

⁴ Para una introducción simple y lúcida a los conceptos de la teoría de la información, puede consultarse [3].

⁵ Ver [2], pp. 9-13.

base escogida para los logaritmos. H se llama *entropía*, por analogía con el concepto similar de la termodinámica.⁶

En general, en las ciencias sociales hay un número considerable de fenómenos que se manifiestan en forma de proporciones o de frecuencias y que se pueden manejar como probabilidades: transacciones interregionales e interindustriales (reducidas a fracciones de las transacciones totales), participación de los factores en el ingreso nacional, proporciones regionales de población, participaciones de satisfactores en el presupuesto del consumidor, etc. Parece natural intentar examinar estos fenómenos a la luz de la información asociada con experimentos que descubran o modifiquen tales probabilidades.

La entropía de las participaciones de los productores en el mercado fue sugerida por Theil, Scholes y el autor [7] como medida de concentración industrial. La diferencia de la entropía con su máximo:

$$R = \log n - H$$

se llama *redundancia* y es un caso especial de la llamada información indirecta. Fue sugerida recientemente por Theil [6] para medir la desigualdad de la distribución del ingreso.

El plan de este trabajo es el siguiente: en las secciones 2 y 3 se propone una medida de concentración geográfica de la población, analizándose sus propiedades de descomposición y agregación, a base de participaciones de estados y regiones (agregados de estados) en la población total. Estos conceptos se ilustran, en la sección 4, con datos de los censos de población de México, de 1930 a 1960. La sección 5 analiza el uso de la redundancia como medida de concentración de la población por tamaños de localidades habitadas, a la que se llama aquí "concentración urbana"; se utiliza material empírico de los censos de población de México de 1930 a 1960. Las secciones 6, 7 y 8 se destinan a analizar el siguiente problema: dada la distribución conjunta de la población en estados y en diversos tamaños de localidad habitada, obtener expresiones para las concentraciones marginales (esto es, geográfica y urbana) que reflejen la interrelación que existe entre ambas distribuciones, en términos de la estructura urbana de los estados y regiones. Los conceptos sugeridos se analizan numéricamente con base en material de los censos de población de México de 1930 a 1960. Por último, el apéndice I indica las fuentes de la información estadística utilizada; el apéndice II sugiere una medida de concentración para la distribución conjunta de la población, geográfica y urbana.

2. UNA MEDIDA DE LA CONCENTRACIÓN GEOGRÁFICA DE LA POBLACIÓN

En lo que sigue, los estados serán las unidades de observación y las regiones serán agregados de estados. (La exposición podría refra-

⁶ Si, en lugar de una "distribución de masa" $\{p_1, \dots, p_n\}$ se tiene una densidad $p(x)$, H se define como la integral $-\int_{-\infty}^{\infty} p(x) \log [p(x)] dx$. La entropía para distribuciones continuas presenta algunas dificultades (por ejemplo, puede ser negativa). En este trabajo se trata solamente con distribuciones discretas.

searse en términos de otras unidades: municipios, localidades, países, etc.) El punto de partida del análisis son las participaciones de n estados en el total de la población. Se tiene así una colección de n números no negativos p_1, \dots, p_n , cuya suma es la unidad. La medida de concentración de la población en los n estados será una función de las p_j ($j = 1, 2, \dots, n$).

La más simple de las medidas de concentración que se pueden sugerir en términos de la teoría de la información es la *entropía* de las participaciones p_j :

$$(2.1) \quad H = \sum_{j=1}^n p_j \log \frac{1}{p_j}$$

H es una medida *inversa* de concentración. En efecto, en el caso de máxima concentración existe un estado, digamos k , que absorbe a la totalidad de la población; se tiene entonces que $p_k = 1$, $p_j = 0$, si $j \neq k$. Conviene en definir, como es usual en la teoría de la información:

$$0 \log \frac{1}{0} = \lim_{x \rightarrow 0} x \log \frac{1}{x} = 0$$

se tiene que $H = 0$ en el caso descrito. Supóngase ahora que todos los estados tienen igual parte ($1/n$) en la población total; la concentración es la mínima y el valor de H es el logaritmo de n , que es el máximo de (2.1) respecto a las p_j , bajo la condición de que su suma sea la unidad. El rango de variación de la entropía es, entonces, de cero al logaritmo del número de estados. Este último es una consecuencia de fenómenos geográficos, históricos, políticos, etc., y en una mínima parte de fenómenos demográficos. El uso de la entropía como medida de concentración geográfica de la población llevaría, pues, a la siguiente consecuencia: considérense dos países, uno con n y otro con m estados; en ambos, la población se distribuye uniformemente, o sea que la distribución en el primer país es $\{1/n\}$ y en el segundo $\{1/m\}$. La concentración geográfica de la población en el primero sería $\log n$ y en el segundo $\log m$, de modo que el país con menos estados sería el más concentrado.

Cuando la entropía se usa como medida de concentración industrial, la diferencia de concentración entre las industrias en que los competidores participan en proporciones iguales en su mercado, está justificada si se considera que, en estas condiciones iguales de distribución, hay menor concentración mientras mayor sea el número de competidores (ver [7]). Esto no puede sostenerse usando el número de estados.

Al enfrentarse a un problema semejante en el análisis de la distribución del ingreso (en condiciones iguales de distribución, el grupo o el país más poblado sería el de menor desigualdad), Theil [6] sugirió el uso de la redundancia, que se define en la teoría de la información como la diferencia entre la entropía y su máximo:

$$(2.2) \quad R = \log n - H = \sum_{j=1}^n p_j \log np_j$$

Tampoco la redundancia resuelve el problema que se plantea en este trabajo. Para demostrarlo, conviene ver a la redundancia como un caso especial del concepto de información indirecta. Supóngase que, antes de un experimento (o, como es usual en la investigación demográfica, antes de un censo), el investigador tiene evidencia (*a priori*) consistente en participaciones q_1, \dots, q_n de los estados en el total de la población. Después del experimento, su evidencia consiste en las participaciones "verdaderas" (*a posteriori*) p_1, \dots, p_n . La media ponderada de cambios de información:

$$(2.3) \quad I(p:q) = \sum_{j=1}^n p_j \log \frac{p_j}{q_j}$$

es una medida de la información ganada en el experimento. Se le conoce en la teoría de la información como información indirecta con distribución *a priori* $\{q_j\}$ y distribución *a posteriori* $\{p_j\}$. La redundancia es una información indirecta cuya distribución *a priori* es uniforme.

En el caso de la distribución del ingreso, tiene sentido hablar de concentración o desigualdad en tanto las participaciones de los individuos en el ingreso de su comunidad difieran entre sí, o sea, en tanto su distribución se desvíe de la uniforme. Tiene sentido entonces partir de una distribución *a priori* uniforme y medir la desigualdad a través de una información indirecta, que en este caso resulta ser la redundancia. En el caso de la concentración de la población, no tiene mucho sentido asignar iguales probabilidades *a priori* a estados grandes y pequeños. En cierto sentido, la forma tradicional de medir la concentración geográfica en términos de densidades de población equivale al criterio de información, con distribución *a priori* igual a la distribución territorial. En efecto, considérese la media geométrica de densidades de población:

$$(2.4) \quad G = \prod_{j=1}^n d_j^{p_j} = \prod_{j=1}^n (h_j/t_j)^{p_j} = \prod_{j=1}^n (hp_j/tq_j)^{p_j} = D \prod_{j=1}^n (p_j/q_j)^{p_j}$$

en que d_j es la densidad de población del estado j ; h_j su número de habitantes; t_j su territorio; $\{p_j\}$, $\{q_j\}$ son las distribuciones de población y territorio; y D , h , t son la densidad, habitantes y territorio del país. Tomando logaritmos en (2.4) se advierte que:

$$(2.5) \quad \log \frac{G}{D} = I(p:q)$$

En este trabajo se propone la información indirecta $I(p:q)$ con distribución *a priori* igual a la distribución territorial y distribución *a posteriori* igual a la de población, como medida de la concentración geográfica de la población.

Conviene ahora hacer explícitas algunas de las propiedades de $I(p:q)$, en especial las que motivan esta proposición. En primer

lugar, $I(p:q) \geq 0$; en particular, $I(p:q) = 0$ si (y sólo si) $p_j = q_j$ para $j = 1, 2, \dots, n$. La demostración es muy simple: basta con notar que $-I(p:q)$ es el logaritmo de la media geométrica (con pesos p_j) de los cocientes q_j/p_j la cual, por la desigualdad de las medias, es menor o igual a la correspondiente media aritmética, que es la unidad, con igualdad sólo cuando $p_j = q_j$ para todas las j . Nótese que $I(p:q) = 0$ cuando las densidades de población de todos los estados son iguales.

Supóngase ahora que, en un momento dado, se inicia un movimiento migratorio del estado 1 al estado 2, sin afectar la participación conjunta de los dos estados ($p_1 + p_2$) en la población total. Suponiendo que el estado 1 tiene mayor densidad de población que el estado 2, ¿cuál es el efecto del movimiento migratorio sobre el valor de la concentración? Para obtenerlo, escríbase $I(p:q)$ como sigue:

$$\begin{aligned}
 (2.6) \quad I(p:q) &= p_1 \log \frac{p_1}{q_1} + p_2 \log \frac{p_2}{q_2} + \sum_{j=3}^n p_j \log \frac{p_j}{q_j} \\
 &= (p_1 + p_2) \left[\frac{p_1}{p_1 + p_2} \log \frac{p_1/p_1 + p_2}{q_1/q_1 + q_2} \right. \\
 &\quad \left. + \frac{p_2}{p_1 + p_2} \log \frac{p_2/p_1 + p_2}{q_2/q_1 + q_2} + \log \frac{p_1 + p_2}{q_1 + q_2} \right] \\
 &\quad + \sum_{j=3}^n p_j \log \frac{p_j}{q_j}
 \end{aligned}$$

Dado el supuesto de que ($p_1 + p_2$) no se altera, cambian por la migración sólo los dos primeros términos entre corchetes en la última igualdad:

$$\frac{p_1}{p_1 + p_2} \log \frac{p_1/p_1 + p_2}{q_1/q_1 + q_2} + \frac{p_2}{p_1 + p_2} \log \frac{p_2/p_1 + p_2}{q_2/q_1 + q_2}$$

Pero ésta es una información indirecta restringida a los estados 1 y 2. Se sabe que esta información es mínima ($= 0$) si su distribución *a priori* es igual a su distribución *a posteriori*, o sea, si:

$$\begin{aligned}
 p_1/p_1 + p_2 &= q_1/q_1 + q_2 \\
 p_2/p_1 + p_2 &= q_2/q_1 + q_2
 \end{aligned}$$

relaciones que pueden escribirse como sigue:

$$\begin{aligned}
 \frac{h_1}{t_1} &= d_1 = \frac{h_1 + h_2}{t_1 + t_2} \\
 \frac{h_2}{t_2} &= d_2 = \frac{h_1 + h_2}{t_1 + t_2}
 \end{aligned}$$

Se concluye que $I(p:q)$ disminuye hasta el momento en que las densidades de población de los dos estados sean iguales. Si la migración continúa, $I(p:q)$ crece consecuentemente.

La propiedad de $I(p:q)$ que motiva en buena parte el enfoque de este trabajo está de hecho implícita en la ecuación (2.6). Allí se ve que la contribución a $I(p:q)$ de los estados 1 y 2 es la siguiente:

$$(2.7) \quad P_1 \left[\frac{p_1}{P_1} \log \frac{p_1/P_1}{q_1/Q_1} + \frac{p_2}{P_2} \log \frac{p_2/P_1}{q_2/Q_1} \right] + P_1 \log \frac{P_1}{Q_1}$$

en que $P_1 = p_1 + p_2$, $Q_1 = q_1 + q_2$ son las participaciones en población y territorio del agregado constituido por los estados 1 y 2. El primer término de (2.7) es P_1 multiplicando a una información indirecta con distribución *a priori* $\{q_j/Q_1\}$ y distribución *a posteriori* $\{p_j/P_1\}$ ($j = 1, 2$) y mide la concentración dentro del agregado. El segundo es uno de los sumandos de una información indirecta; su probabilidad *a priori* es Q_1 y su probabilidad *a posteriori* es P_1 . Esto sugiere el siguiente enfoque del problema de agregación: considérense G agregados de estados, S_1, \dots, S_G , a los que conviene llamar regiones. Defínanse las participaciones regionales en población y territorio como sigue:

$$P_g = \sum_{j \in S_g} p_j \quad g = 1, 2, \dots, G$$

$$Q_g = \sum_{j \in S_g} q_j \quad g = 1, 2, \dots, G$$

Se tiene entonces:

$$(2.8) \quad I(p:q) = \sum_{g=1}^G P_g \log \frac{P_g}{Q_g} + \sum_{g=1}^G P_g \sum_{j \in S_g} \frac{p_j}{P_g} \log \frac{p_j/P_g}{q_j/Q_g} = \\ = I_0 + \sum_{g=1}^G P_g I_g,$$

digamos. El primer término de (2.8), I_0 , es una información indirecta con distribución *a priori* $\{Q_g\}$ y distribución *a posteriori* $\{P_g\}$. Mide la concentración geográfica de la población al nivel de regiones; será llamado aquí concentración interregional de población. Cuando la información de que se dispone se reduce a datos regionales, I_0 estima la concentración $I(p:q)$.

Como puede verse, el siguiente término de la descomposición (2.8) es no negativo, por lo que I_0 subestima la concentración $I(p:q)$.

El segundo término de la descomposición (2.8) es una media ponderada de G términos del tipo siguiente:

$$(2.9) \quad I_g = \sum_{j \in S_g} \frac{p_j}{P_g} \log \frac{p_j/P_g}{q_j/Q_g}$$

Cada I_g es una información indirecta; su distribución *a priori* es la distribución condicional $\{q_j/Q_g\}$, que no es sino la distribución interna de territorio en la región S_g . La distribución *a posteriori* en (2.9) es la distribución condicional $\{p_j/P_g\}$, esto es, la distribución interna de población en la región S_g . I_g se llamará la concentración interna de la región S_g ; la media, $\Sigma P_g I_g$, será la concentración intrarregional de población.

3. OTRAS PROPIEDADES DE $I(p:q)$

La ecuación (2.5) muestra que $I(p:q)$ es el logaritmo de un cociente de densidades; la media geométrica de las densidades de los estados es el numerador; el denominador es la densidad de población del país. I_0 e I_g se pueden expresar en la misma forma:

$$(3.1) \quad I_0 = \log \prod_{g=1}^G (P_g/Q_g)^{P_g} = \log \prod_{g=1}^G \left(\frac{H_g/h}{T_g/t} \right)^{P_g} = \\ = \log t/h \prod_{g=1}^G (H_g/T_g)^{P_g} = \log \prod D_g^{P_g} / D = \log G_0/D$$

digamos, en que D_g , H_g y T_g son la densidad, población y territorio de la región S_g . A su vez:

$$(3.2) \quad I_g = \log \prod_{j \in S_g} \left(\frac{p_j/P_g}{q_j/Q_g} \right)^{p_j/P_g} = \log Q_g/P_g \prod_{j \in S_g} (p_j/q_j)^{p_j/P_g} = \\ = \log \frac{T_g/t}{H_g/h} \prod_{j \in S_g} \left(\frac{h_j/h}{t_j/t} \right)^{p_j/P_g} = \\ = \log \frac{T_g}{H_g} \prod_{j \in S_g} (h_j/t_j)^{p_j/P_g} = \\ = \log \frac{\prod d_j^{p_j/P_g}}{D_g} = \log \frac{G_g}{D_g}$$

digamos. Se concluye que, en cada región S_g , I_g es cero si $G_g = D_g$, o sea, si todos los estados tienen la misma densidad. En este sentido, I_g mide la heterogeneidad de la región S_g en cuanto a las densidades de población de los estados que la componen. Por otra parte, I_0 es cero si $G_0 = D$, o sea, si todas las regiones tienen la misma densidad de población. Un esquema de regionalización define las regiones "más homogéneas" si I_0 es un máximo (de donde $\Sigma P_g I_g$ es un mínimo, dada una concentración total I).

Por otra parte, la concentración media en términos de información se presta a una interpretación interesante: supóngase que la totalidad de la población se concentra en una fracción q_0 del territorio, y

el resto queda despoblado. En esta fracción q_0 la población se distribuye uniformemente y su densidad es:

$$(3.3) \quad d_0 = \frac{h}{q_0 t} = \frac{d}{q_0}$$

Entonces la concentración geográfica está dada por:

$$(3.4) \quad I(p:q) = 1 \times \log \frac{1}{q_0} + 0 \log \frac{0}{q_1} + \dots + 0 \log \frac{0}{q_n}$$

de donde $q_0 = 2^{-I(p:q)}$ si los logaritmos son de base 2, como es común en la teoría de la información. Usando este resultado se concluye que $d_0 = 2^I D$; ver (3.3). Esta interpretación será útil en el análisis de los datos empíricos. Por ejemplo, una concentración de 2 *bits* (unidades binarias de información, que resultan de usar logaritmos de base 2) equivale a la concentración de la población en una cuarta parte del territorio del país, con una densidad de 4 veces la densidad del país; es obvio que el mismo razonamiento se aplica al nivel regional.

Interesa ahora explorar las posibilidades de análisis dinámico *vía* información indirecta. Sean $\{p_j^1\}$ y $\{p_j^2\}$ las distribuciones de población en dos períodos sucesivos. Sean I^1, I^2 las informaciones indirectas en ambos períodos y defínanse:

$$\Delta I = I^2 - I^1, \Delta \log p_j = \log p_j^2 - \log p_j^1.$$

Se tiene entonces:

$$(3.4) \quad \begin{aligned} \Delta I &= \sum_{j=1}^n p_j^2 \log \frac{p_j^2}{q_j} - \sum_{j=1}^n p_j^1 \log \frac{p_j^1}{q_j} = \\ &= \sum_{j=1}^n p_j^2 \log \frac{p_j^2}{p_j^1} + \sum_{j=1}^n (p_j^2 - p_j^1) \log \frac{p_j^1}{q_j} = \\ &= \sum_{j=1}^n p_j^2 \Delta \log p_j + \sum_{j=1}^n w_j \log \frac{p_j^1}{q_j} \end{aligned}$$

Se nota de inmediato que $\sum w_j = 0$ y que $\sum p_j^2 \Delta \log p_j$ es una información indirecta, de modo que ΔI puede escribirse:

$$(3.5) \quad \Delta I = I(p^2: p^1) + R$$

en que R puede tratarse como residuo, ya que se trata de una suma ponderada de $\log(p_j^1/q_j)$ con pesos que suman cero. Es probable que, si los períodos no son muy lejanos, este residuo sea suficientemente

pequeño y se pueda aproximar ΔI por $I(p^2: p^1)$. Esto lleva a conjeturar que, siendo el cambio de información aproximadamente una información indirecta, con todas las propiedades de ésta, el enfoque de la teoría de la información puede llevar a resultados interesantes en el análisis dinámico.

4. CONCENTRACIÓN GEOGRÁFICA DE LA POBLACIÓN EN MÉXICO, 1930-1960

Se ha computado la ecuación (2.8), midiendo la información en *bits*, respecto a la distribución de la población de las 32 entidades (estados, territorios y Distrito Federal, en adelante "estados") de la República, divididas en nueve regiones, con base en datos de los censos de 1930, 1940, 1950 y 1960. (Véase el Apéndice I.)

Las regiones fueron sugeridas en una ponencia presentada por la Dirección de Estudios Hacendarios de la Secretaría de Hacienda y Crédito Público al V Congreso Internacional de Planificación [5]. Las regiones y estados que las componen se enumeran a continuación; las cifras entre paréntesis son las participaciones porcentuales medias (sobre el período 1930-1960) de cada estado y región en la población total y en el territorio del país, en ese orden.

Región 1. Pacífico Norte (6.49, 21.07): Baja California Norte (0.79, 3.56), Baja California Sur (0.25, 3.75), Nayarit (1.09, 1.40), Sinaloa (2.45, 2.95) y Sonora (1.91, 9.40).

Región 2. Norte (5.80, 20.27): Chihuahua (3.17, 12.56) y Coahuila (2.63, 7.70).

Región 3. Golfo Norte (5.22, 7.34): Nuevo León (2.80, 3.28) y Tamaulipas (2.42, 4.06).

Región 4. Centro Norte (9.46, 13.38): Aguascalientes (0.73, 0.28), Durango (2.49, 6.08), San Luis Potosí (3.30, 3.19) y Zacatecas (2.64, 3.81).

Región 5. Occidente (13.08, 7.39): Colima (0.41, 0.28), Jalisco (7.02, 4.07) y Michoacán (5.65, 3.04).

Región 6. Centro. (34.81, 6.56): Distrito Federal (10.50, 0.08), Guanajuato (5.34, 1.55), Hidalgo (3.51, 1.07), México (5.85, 1.09), Morelos (0.97, 0.25), Puebla (6.34, 1.72), Querétaro (1.19, 0.60) y Tlaxcala (1.11, 0.20).

Región 7. Golfo (10.11, 4.96): Tabasco (2.04, 1.25) y Veracruz (8.07, 3.70).

Región 8. Pacífico-Istmo (12.03, 11.85): Chiapas (3.30, 3.76), Guerrero (3.63, 3.24) y Oaxaca (5.74, 4.85).

Región 9. Península de Yucatán (2.63, 7.19): Campeche (0.49, 2.85), Quintana Roo (0.10, 2.14) y Yucatán (2.05, 2.21).

El cuadro 1 contiene los componentes inter e intrarregional de la concentración $I(p: q)$, de 1930 a 1960. El crecimiento de la concentración total de 1.07 a 1.36 *bits* aproximadamente equivale a una relación entre la media geométrica de las densidades de los estados y la del país, creciente de $2^{1.07} \cong 2.021$ a $2^{1.36} = 2.701$. Acudiendo a la interpretación de $I(p: q)$ mencionada en la sección 3, se puede ver que los valores observados equivalen a la concentración que se obtendría de

distribuir a la población con densidad uniforme en un 49.48 % del territorio en 1930, proporción que ha disminuido a 37.02 % en 1960.

Cuadro 1

MÉXICO: COMPONENTES INTER E INTRARREGIONAL DE LA CONCENTRACIÓN GEOGRÁFICA DE LA POBLACIÓN, 1930-1960

(Cifras absolutas en *bits*; cifras entre paréntesis en porcentos de $I(p:q)$)

| Año del censo | Concentración Total (I) | Concentración Interregional (I_0) | Concentración intrarregional ($\sum P_g I_g$) |
|--------------------|-----------------------------|---------------------------------------|---|
| 1930 | 1.0711 (100.00) | 0.7436 (69.42) | 0.3275 (30.58) |
| 1940 | 1.1133 (100.00) | 0.7414 (66.59) | 0.3719 (33.41) |
| 1950 | 1.2456 (100.00) | 0.7251 (58.21) | 0.5205 (41.79) |
| 1960 | 1.3630 (100.00) | 0.7355 (53.96) | 0.6275 (46.04) |
| Promedio 1930-1960 | 1.1982 (100.00) | 0.7364 (61.46) | 0.4616 (38.54) |

Aunque el valor numérico de la concentración interregional es bastante estable, su contribución tiende a ser cada vez menor; casi todo el crecimiento de la concentración se explica en términos del incremento de la concentración intrarregional. Se puede pues pensar que las "disparidades" entre regiones van cediendo lugar, en términos relativos, a fenómenos intrarregionales de concentración. Conviene, pues, explorar la estructura del término intrarregional, para precisar las regiones que desempeñan el papel más importante en los cambios de concentración. El cuadro 2 contiene las concentraciones internas y sus contribuciones porcentuales a la concentración intrarregional, $P_g I_g / \sum P_g I_g$ (cifras entre paréntesis).

Resulta así evidente que a la región Centro puede atribuirse la mayor parte de los cambios de concentración y, aún más, que contrarresta la tendencia decreciente de la mayoría de las demás concentraciones internas. Esto es reflejo, claro está, del crecimiento del Distrito Federal. Sólo las regiones Centro Norte y Golfo presentan valores aproximadamente estables en sus concentraciones internas; sin embargo, dado que su participación en la población del país decrece en el período, su contribución relativa es también decreciente.

Si en 1930 seis de las nueve regiones contribuían con más del 1 % de la concentración intrarregional, y una estaba cerca del 1 % (Golfo Norte), para 1960 sólo cuatro tenían tal contribución; la región más

Cuadro 2

MÉXICO: CONCENTRACIÓN INTERNA (EN "BITS") Y CONTRIBUCIÓN (EN PORCIENTOS) DE CADA REGIÓN A LA CONCENTRACIÓN INTRARREGIONAL, 1930-1960

| Región | 1930 | 1940 | 1950 | 1960 | Promedio 1930 - 1960 |
|----------------------------|-------------------|-------------------|-------------------|-------------------|-------------------------|
| 1. Pacífico Norte | 0.5837 (10.14) | 0.5450 (8.95) | 0.4340 (5.58) | 0.3529 (4.20) | 0.4789 (7.22) |
| 2. Norte | 0.0252 (0.44) | 0.0174 (0.25) | 0.0190 (0.22) | 0.0061 (0.06) | 0.0169 (0.24) |
| 3. Golfo Norte | 0.0695 (0.90) | 0.0265 (0.35) | 0.0106 (0.11) | 0.0126 (0.12) | 0.0298 (0.37) |
| 4. Centro Norte | 0.1939 (5.72) | 0.1242 (3.23) | 0.1749 (3.05) | 0.1863 (2.44) | 0.1698 (3.61) |
| 5. Occidente | 0.0040 (0.17) | 0.0032 (0.11) | 0.0015 (0.04) | 0.0000 (0.00) | 0.0022 (0.08) |
| 6. Centro | 0.6906 (72.45) | 0.8893 (80.44) | 1.2768 (86.52) | 1.5744 (90.25) | 1.1078 (82.42) |
| 7. Golfo | 0.0554 (1.66) | 0.0211 (0.68) | 0.0444 (0.79) | 0.0417 (0.61) | 0.0406 (0.94) |
| 8. Pacífico-Istmo | 0.0257 (1.66) | 0.0225 (0.75) | 0.0050 (0.12) | 0.0020 (0.04) | 0.0138 (0.50) |
| 9. Península de Yucatán | 0.8244 (7.42) | 0.7425 (5.23) | 0.7200 (3.57) | 0.5994 (2.27) | 0.7216 (4.62) |

importante (en contribución) después de la región Centro, aporta el 4.20 % a la concentración intrarregional.

Es interesante preguntarse si los datos o, en última instancia, los criterios de información, pueden revelar la existencia de "polos de crecimiento" de la población. Desde luego, una región cuya concentración interna es creciente contiene al menos un "polo de crecimiento" de influencia regional. Pero de las concentraciones internas decrecientes no puede deducirse la ausencia de tales polos. En efecto, en vista de la ecuación (2.6) y del razonamiento subsecuente, el decrecimiento puede provenir de la migración a una zona originalmente de baja densidad. Esta zona puede contener un polo regional de crecimiento, cuya influencia no es aún suficientemente fuerte para convertirla en zona de alta densidad. De continuar la migración hacia el o los polos, llegará un momento en que las posiciones relativas de los estados se inviertan y la concentración inicie un movimiento ascendente. Por otra parte, está claro que los criterios de información, aplicados al nivel

Cuadro 3
OBTENCIÓN DE I_1 A PARTIR DE DENSIDADES

| Año del censo | Densidad $(d_j)_a$, relación con la densidad de la región (d_j/D_1) y población relativa $(100p_j/P_1)$ | | | | Media geométrica b/ de densidades de los estados (G_1) | $\frac{G_1}{D_1}$ | $I_1 = \frac{G_1}{\log \frac{1}{D_1}}$ |
|---------------|--|-----------------------|---------------------|---------|--|-------------------|--|
| | Estado | Baja California Norte | Baja California Sur | Nayarit | | | |
| 1930 | d_j | 0.69 | 0.64 | 6.09 | 6.81 | 1.46 | .5837 |
| | d_j/D_1 | .3082 | .2827 | 2.7200 | 3.0444 | 0.6504 | |
| | $100p_j/P_1$ | 5.21 | 5.07 | 18.07 | 42.63 | 29.02 | |
| 1940 | d_j | 1.49 | 0.62 | 8.03 | 8.89 | 1.96 | .5450 |
| | d_j/D_1 | .4931 | .2051 | 2.6657 | 2.9505 | 0.6492 | |
| | $100p_j/P_1$ | 8.36 | 3.65 | 17.71 | 41.31 | 28.96 | |
| 1950 | d_j | 3.08 | 0.83 | 10.54 | 10.96 | 2.76 | .4340 |
| | d_j/D_1 | .7396 | .1983 | 2.5323 | 2.6332 | .6638 | |
| | $100p_j/P_1$ | 13.16 | 3.53 | 16.83 | 36.87 | 29.61 | |
| 1960 | d_j | 7.42 | 1.11 | 14.15 | 14.44 | 4.23 | .3529 |
| | d_j/D_1 | 1.1780 | .1741 | 2.2455 | 2.2913 | .6719 | |
| | $100p_j/P_1$ | 19.90 | 3.12 | 14.92 | 32.08 | 29.97 | |
| Promedio | d_j | 3.17 | 0.80 | 9.70 | 10.28 | 2.60 | .4789 |
| 1930-1960 | d_j/D_1 | .6797 | .2151 | 2.5409 | 2.7299 | 0.6588 | |
| $100p_j/P_1$ | 11.66 | 3.84 | 16.88 | 38.22 | 29.39 | | |

a/ Habitantes por km² b/ Con pesos p_j/P_1 . c/ Logaritmos de base 2.

Cuadro 4
OBTENCIÓN DE I_6 A PARTIR DE DENSIDADES

| Año del censo | Densidades y participación de población de los estados de la región 6 | | | | | Densidad de la región \bar{e} (D_6) | Media geométrica de densidades de los estados (G_6) ^{b/} | $I_6 = \log \frac{G_6}{D_6}$ | | | | | |
|--------------------|---|------------|---------|--------|-----------------------------------|---|---|------------------------------|--------|-------|--------|--------|--------|
| | Distrito Federal | Guanajuato | Hidalgo | México | Morelos Puebla Querétaro Tlaxcala | | | | | | | | |
| 1930 | d_j | 780.58 | 32.33 | 32.20 | 46.18 | 26.85 | 34.00 | 19.83 | 52.37 | 43.42 | 70.08 | 1.6139 | 0.6906 |
| | d_j/D_6 | 17.9769 | .7446 | .7416 | 1.0634 | .6185 | .7830 | .4567 | 1.2062 | | | | |
| | 100 p_j/P_6 | 21.92 | 17.59 | 12.09 | 17.67 | 2.36 | 20.53 | 4.18 | 3.68 | | | | |
| 1940 | d_j | 1 127.04 | 34.82 | 36.31 | 61.78 | 37.38 | 38.38 | 20.91 | 56.43 | 53.27 | 98.67 | 1.8523 | 0.8893 |
| | d_j/D_6 | 21.1576 | .6537 | .6816 | 1.1597 | .7017 | .7206 | .3926 | 1.0594 | | | | |
| | 100 p_j/P_6 | 25.80 | 15.44 | 11.12 | 19.27 | 2.67 | 18.89 | 3.59 | 3.23 | | | | |
| 1950 | d_j | 1 938.34 | 43.58 | 40.40 | 64.95 | 55.49 | 48.05 | 24.25 | 72.32 | 70.43 | 170.66 | 2.4230 | 1.2768 |
| | d_j/D_6 | 27.5198 | .6187 | .5736 | .9221 | .7878 | .6822 | .3443 | 1.0268 | | | | |
| | 100 p_j/P_6 | 33.55 | 14.61 | 9.35 | 15.32 | 3.00 | 17.88 | 3.15 | 3.13 | | | | |
| 1960 | d_j | 3 095.09 | 56.92 | 47.25 | 88.51 | 78.54 | 58.34 | 30.08 | 88.12 | 97.31 | 289.80 | 2.9781 | 1.5744 |
| | d_j/D_6 | 31.8061 | .5849 | .4856 | .9096 | .8071 | .5995 | .3091 | .9056 | | | | |
| | 100 p_j/P_6 | 38.77 | 13.82 | 7.92 | 15.11 | 3.08 | 15.71 | 2.83 | 2.76 | | | | |
| Promedio 1930-1960 | d_j | 1 735.26 | 41.91 | 39.04 | 65.35 | 49.57 | 44.69 | 23.77 | 67.31 | 66.13 | 142.52 | 2.1552 | 1.1078 |
| | d_j/D_6 | 24.5161 | .6504 | .6206 | 1.0137 | .7287 | .6963 | .3756 | 1.0495 | | | | |
| | 100 p_j/P_6 | 30.00 | 15.36 | 10.12 | 16.84 | 2.77 | 18.25 | 3.43 | 3.20 | | | | |

\bar{e} / Habitantes por km². ^{b/} Con pesos p_j/P_6 ^{c/} Logaritmos de base 2.

Cuadro 5
OBTENCIÓN DE I_0 A PARTIR DE DENSIDADES

| Año del censo | Densidad (D_g) ^{a/} y población relativa (100 P_g) de las regiones | | | | | | | | | Media geométrica de densidades de las regiones país (G_0) (D) | Densidad del país G_0/D | $I_0 = \log_2 G_0/D$ | | |
|--------------------|--|------|------|-------|-------|-------|-------|-------|-------|---|---------------------------|----------------------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | | | |
| 1930 | D_g | 2.24 | 2.32 | 4.80 | 5.99 | 15.45 | 43.44 | 16.41 | 9.68 | 3.40 | 13.88 | 8.29 | 1.6743 | 0.7436 |
| | 100 . P_g | 5.69 | 5.67 | 4.25 | 9.66 | 13.75 | 34.37 | 9.82 | 13.84 | 2.95 | | | | |
| 1940 | D_g | 3.01 | 2.75 | 7.02 | 7.51 | 18.39 | 53.29 | 25.28 | 10.91 | 3.79 | 17.37 | 10.39 | 1.6718 | 0.7414 |
| | 100 . P_g | 6.11 | 5.36 | 4.96 | 9.67 | 13.08 | 33.65 | 12.07 | 12.44 | 2.62 | | | | |
| 1950 | D_g | 4.16 | 3.93 | 10.11 | 8.89 | 22.57 | 70.46 | 24.62 | 13.93 | 4.70 | 21.66 | 13.10 | 1.6530 | 0.7251 |
| | P_g | 6.69 | 6.08 | 5.66 | 9.08 | 12.73 | 35.27 | 9.32 | 12.60 | 2.58 | | | | |
| 1960 | D_g | 6.30 | 5.35 | 14.56 | 10.91 | 30.68 | 97.34 | 33.04 | 17.69 | 5.88 | 29.56 | 17.75 | 1.6650 | 0.7355 |
| | P_g | 7.48 | 6.11 | 6.02 | 8.22 | 12.77 | 35.97 | 9.23 | 11.81 | 2.38 | | | | |
| Promedio 1930-1960 | D_g | 3.93 | 3.59 | 9.12 | 8.32 | 21.77 | 66.13 | 24.84 | 13.05 | 4.44 | | | | |
| | 100 . P_g | 6.49 | 5.80 | 5.32 | 9.15 | 13.08 | 34.81 | 10.11 | 12.67 | 2.63 | 20.62 | 12.38 | 1.6660 | 0.7364 |

^{a/} Habitantes por km².

de agregación usado en este artículo, no pueden revelar la existencia de polos de crecimiento cuya influencia se restrinja a un estado.

Para descubrir la existencia de polos del segundo tipo —los que provocan la desconcentración de la población dentro de su región— es necesario ver los cambios en la contribución de los estados a la concentración interna de su región. Esto se puede ver también a partir de los cambios en las densidades de población de los estados. A continuación se ejemplifica la obtención de la concentración interna de una región, la Pacífico Norte, a partir de las densidades de los estados, usando la ecuación (2.5). Esto servirá para precisar la existencia de polos del segundo tipo, como podría haber quizá en la zona fronteriza de Baja California.

Se advierte en el cuadro 3 que, efectivamente, la densidad y la contribución del estado de Baja California Norte han ido en constante aumento, mientras que la contribución de los demás, salvo el caso de Sonora, que ha permanecido bastante estable, es decreciente. En el momento en que el incremento de densidad de Baja California Norte a expensas de los demás sea suficientemente grande, la concentración interna de la región 1 volverá a aumentar.

Otro ejemplo interesante es el de la región Centro (cuadro 4). El Distrito Federal llega a absorber casi el 39 % de la población de la región en 1960 y su densidad crece de 18 a 32 veces la densidad de la región; la posición relativa de los demás estados decrece, salvo en el caso de Morelos, aunque el descenso de los estados de México y Tlaxcala no es muy pronunciado.

La región Centro aporta un ejemplo de polarización con crecimiento en la concentración. Es interesante ver que el crecimiento de Morelos, en particular de la ciudad de Cuernavaca, constituida en un "suburbio turístico" de la ciudad de México y en un pequeño polo de concentración de población, se puede advertir aun a este nivel de agregación.

Por último, el cuadro 5 ilustra la obtención de I_0 , la concentración intrarregional, a partir de las densidades de las regiones.

5. CONCEPTO Y MEDICIÓN DE LA CONCENTRACIÓN URBANA

En condiciones ideales de información, la medición de la concentración urbana parte de las participaciones f_1, \dots, f_K de las K localidades que componen el país, en la población total. Se dirá que no existe concentración urbana si todas las localidades tienen el mismo número de habitantes, es decir, si $f_1 = f_2 = \dots = f_K = 1/K$.

Se concluye que la media de concentración urbana es una información indirecta con distribución *a priori* uniforme, que de inmediato se identifica con la redundancia de la distribución *a posteriori* $\{f_k\}$ ($k = 1, \dots, K$); ver ecuación (2.2). Se tiene entonces:

$$(5.1) \quad I(f: 1/K) = \sum_{k=1}^K f_k \log \frac{f_k}{1/K} =$$

$$\begin{aligned}
&= \sum_{k=1}^K f_k \log (K f_k) = \log K - \sum_{k=1}^K f_k \log \frac{1}{f_k} = \\
&= \log K - H = R
\end{aligned}$$

El valor máximo de R es el logaritmo del número de localidades (que se daría si *una* localidad concentra a toda la población del país). Fuera de este caso, bastante poco interesante, el valor de R depende de proporciones de población y no del número de localidades. En efecto, considérese el caso en que una proporción p de la población se reparte uniformemente en una proporción θ de las localidades; el restante $(1-p)$ de la población se reparte, también uniformemente, en el $(1-\theta)$ de las localidades. La participación en la población de una localidad arbitraria del primer grupo es $p/K\theta$ y la de una localidad del segundo es $(1-p)/K(1-\theta)$. Habiendo $K\theta$ localidades en el primer grupo y $K(1-\theta)$ en el segundo, se tiene:

$$\begin{aligned}
(5.2) \quad R &= K\theta \frac{p}{K\theta} \log K (p/K\theta) + K(1-\theta) \frac{1-p}{K(1-\theta)} \log K [(1-p)/K(1-\theta)] \\
&= p \log \frac{p}{\theta} + (1-p) \log \frac{(1-p)}{(1-\theta)} = I(p; \theta)
\end{aligned}$$

Es obvio que (5.2) puede generalizarse a m grupos de localidades; el grupo i absorbe una proporción p_i de la población y una proporción θ_i de las localidades. Éste es, de hecho, el camino natural para computar R cuando la información está dada en la forma de "datos agrupados", como en los censos de población de México. En tal caso, se conocen la proporción p_i de la población que vive en localidades entre x_{i-1} y x_i habitantes ($i = 1, \dots, m$), y el número de éstas, k_i ($\sum p_i = 1$, $\sum k_i = K$). La concentración urbana se puede aproximar en la forma siguiente: suponiendo que la distribución *interna* de cada intervalo sea uniforme, $f_k = p_i/k_i$ si la localidad k está entre x_{i-1} y x_i habitantes. Haciendo $k_i/K = \theta_i$ se tiene:

$$(5.3) \quad R = \sum_{i=1}^m k_i \frac{p_i}{k_i} \log K \frac{p_i}{k_i} = \sum_{i=1}^m p_i \log \frac{p_i}{\theta_i} = I(p; \theta)$$

que es la forma general de (5.2).

El tamaño supuesto para cada localidad en $[x_{i-1}, x_i)$ es el mismo en todo el intervalo:

$$\bar{x}_i = h \frac{p_i}{k_i} = \bar{x} \frac{p_i}{\theta_i}$$

en que $\bar{x} = h/K$ es el tamaño medio de las localidades del país (h es la población total del país, al igual que en las secciones 2 y 3). R se puede escribir, pues, en otra forma:

$$(5.4) \quad R = \frac{\sum_{i=1}^m p_i \log \frac{\bar{x}_i}{\bar{x}}}{\sum_{i=1}^m p_i}$$

La ecuación (5.3) expresa a R como el logaritmo del cociente de la media geométrica de las participaciones en localidades de los intervalos $[x_{i-1}, x_i)$ sobre la media geométrica de sus participaciones en la población:

$$R = \frac{\prod \theta_i^{p_i}}{\prod p_i} = \log \frac{G_\theta}{G_p}$$

mientras que, de acuerdo con (5.4), R es el logaritmo del cociente de la media geométrica de las "marcas de clase" \bar{x}_i sobre el tamaño medio de las localidades del país:

$$R = \frac{\prod \bar{x}_i^{p_i}}{\bar{x}} = \log \frac{G_x}{\bar{x}}$$

Una interpretación interesante de la concentración urbana es la siguiente: si a un gobernante se le ocurriera repartir a la población uniformemente en una parte de las localidades del país, dejando deshabitado el resto, ¿cuántas localidades necesitaría para que la concentración urbana fuera *igual* a la concentración observada antes

Cuadro 6

MÉXICO: CONCENTRACIÓN URBANA, 1930-1960
(En bits)

| Año del censo | Concentración urbana (R) | Proporción de localidades habitada en el caso de distribución uniforme (2^{-R}) |
|--------------------|-----------------------------|--|
| 1930 | 2.7440 | 0.1493 |
| 1940 | 3.0208 | 0.1232 |
| 1950 | 3.5565 | 0.0850 |
| 1960 | 3.6828 | 0.0779 |
| Promedio 1930-1960 | 3.2510 | 0.1050 |

de hacer ese cambio? Haciendo $p = 1$ en (5.2), se tiene $R = \log 1/\theta$, de donde, si R está dada en *bits*, $\theta = 2^{-R}$, y serían suficientes $K2^{-R}$ localidades.

Es evidente que, como cualquier información indirecta, R posee las propiedades de descomposición y agregación analizadas en las secciones anteriores; los agregados "naturales" serían aquí los grupos de intervalos. Para el propósito de este trabajo, esto no es particularmente interesante. En cambio, la posibilidad de ver a R como una medida asociada a una distribución *marginal*, proveniente de una distribución de la población por estados y tamaños de localidades, sí presenta interés, en tanto se puedan estudiar los componentes estatales y regionales de la concentración urbana. Esto se hará en la sección 8.

En el caso particular de México en el período 1930-1960, la concentración urbana es bastante elevada: en 30 años crece de 2.75 a 3.68 *bits* aproximadamente; en otras palabras, la relación de la media geométrica de tamaños de localidad por intervalos, G_m , a la localidad media del país va de $2^{2.744} \cong 6.69$ en 1930 a $2^{3.688} \cong 12.81$ en 1960. (Véase el cuadro 6.)

Acudiendo a la interpretación, mencionada más arriba, de la concentración urbana en términos de la proporción de localidades en que se reparte uniformemente la población dejando el resto deshabitadas, se advierte que esta proporción ha descendido de un valor cercano a 1/7 en 1930 a un valor cercano a 1/13 en 1960.

6. ESTRUCTURA URBANA Y CONCENTRACIÓN GEOGRÁFICA

Hasta el momento, la concentración de la población en el territorio y su concentración en las localidades del país se han considerado como dos fenómenos aparte. El paso natural al intentar avanzar en el análisis consiste en considerar la distribución bidimensional cuyas distribuciones marginales sean las distribuciones *a posteriori* estudiadas en las secciones anteriores. Sea p_{ij} la proporción de la población que vive en localidades entre x_{i-1} y x_i habitantes (por brevedad, localidades de tamaño i) en el estado j . En nuestro caso, $i = 1, \dots, 15$; $j = 1, \dots, 32$. Conviene introducir la siguiente notación:

$$p_{i\cdot} = \sum_{j=1}^{32} p_{ij} \quad (i = 1, \dots, 15)$$

$$p_{\cdot j} = \sum_{i=1}^{15} p_{ij} \quad (j = 1, \dots, 32)$$

$$c_{ig} = \sum_{j \in S_g} p_{ij} \quad (i = 1, \dots, 15; g = 1, \dots, 9)$$

$$P_{\cdot g} = \sum_{j \in S_g} p_{\cdot j} = \sum_{i=1}^{15} c_{ig} \quad (g = 1, \dots, 9)$$

$\{p_{i\cdot}\}$ y $\{p_{\cdot j}\}$ son las distribuciones marginales de la población por

tamaños de localidades y por estados, $\{P_{\cdot g}\}$ es la distribución marginal por regiones y $\{c_{ig}\}$ es la distribución conjunta por tamaños de localidades y por regiones. Se usarán con frecuencia las distribuciones condicionales:

$$\{p_{ij}/p_{i\cdot}\}; \{p_{ij}/p_{\cdot j}\}; \{c_{ig}/P_{\cdot g}\}; \{p_{ij}/c_{ig}\}$$

cuyo significado es evidente.

En esta sección se considerarán las concentraciones geográficas de la población para cada tamaño de localidad, por ejemplo, las informaciones indirectas con distribución *a priori* $\{q_j\}$ y *a posteriori* $\{p_{ij}/p_{i\cdot}\}$ ($i = 1, \dots, 15$):

$$(6.1) \quad I_{i\cdot} = \sum_{j=1}^{32} \frac{p_{ij}}{p_{i\cdot}} \log \frac{p_{ij}/p_{i\cdot}}{q_j} \quad (i = 1, \dots, 15)$$

Es inmediato obtener componentes inter e intrarregional de $I_{i\cdot}$:

$$(6.2) \quad I_{i\cdot} = \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \log \frac{c_{ig}/p_{i\cdot}}{Q_g} + \sum_{g=1}^9 \sum_{j \in S_g} \frac{p_{ij}}{p_{i\cdot}} \dots$$

$$\dots \log \frac{p_{ij}/c_{ig}}{q_j/Q_g} = I_{i0} + \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} I_{ig}$$

($I_{ig} = \sum_{j \in S_g} p_{ij}/c_{ig} \log \frac{p_{ij}/c_{ig}}{q_j/Q_g}$). I_{i0} es la concentración interregional de las localidades de tamaño i ; I_{ig} es su concentración interna en la región S_g . A su vez, cada uno de estos componentes puede descomponerse en dos:

$$(6.3) \quad I_{i0} = \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \log \frac{P_{\cdot g}}{Q_g} + \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \log \frac{c_{ig}}{p_{i\cdot} P_{\cdot g}}$$

$$(6.4) \quad I_{ig} = \sum_{j \in S_g} \frac{p_{ij}}{c_{ig}} \log \frac{p_{\cdot j}/P_{\cdot g}}{q_j/Q_g} + \sum_{j \in S_g} \frac{p_{ij}}{c_{ig}} \log \frac{p_{ij}/c_{ig}}{p_{\cdot j}/P_{\cdot g}}$$

Los componentes de I_{i0} son, el primero, un término de la concentración geográfica (marginal) interregional, analizada en las secciones 2 y 3. En efecto, promediando este término sobre i se tiene:

$$\sum_{i=1}^{15} p_{i\cdot} \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \log \frac{P_{\cdot g}}{Q_g} = \sum_{g=1}^9 P_{\cdot g} \log \frac{P_{\cdot g}}{Q_g} = I_{\cdot 0}$$

Se puede considerar a $\sum c_{ig} \log P_{\cdot g}/Q_g$ como la contribución de las localidades de tamaño i a la concentración marginal interregional.

El segundo término a la derecha de la igualdad en (6.3) es una información indirecta cuya distribución *a priori* es la distribución marginal $\{P_{\cdot g}\}$ y cuya distribución *a posteriori* es la distribución con-

dicional de población por regiones, restringida a localidades del tamaño i , $\{c_{ig}/p_{i\cdot}\}$. Obviamente, este término es cero si $c_{ig}/p_{i\cdot} = P_{\cdot g}$, o sea si la distribución de población por regiones es la misma para todos los tamaños de localidad. Por otra parte, la condición $c_{ig}/p_{i\cdot} = P_{\cdot g}$ puede escribirse también $c_{ig}/P_{\cdot g} = p_{i\cdot}$, o sea, la proporción de la población de la región S_g que habita en localidades de tamaño i es independiente de g , o sea, es la misma para todas las regiones.

Por otra parte, la condición puede verse como una condición de independencia estocástica de la distribución $\{c_{ig}\}$: $c_{ig} = p_{i\cdot} P_{\cdot g}$. Promediando sobre los tamaños de localidad se tiene:

$$\sum_{i=1}^{15} p_{i\cdot} \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \log \frac{c_{ig}}{p_{i\cdot} P_{\cdot g}} = \sum_{i=1}^{15} \sum_{g=1}^9 c_{ig} \log \frac{c_{ig}}{p_{i\cdot} P_{\cdot g}} = L_0$$

que es una información indirecta en dos dimensiones; su distribución *a priori* es estocásticamente independiente ($\sum_i c_{ig} = P_{\cdot g}$; $\sum_g c_{ig} = p_{i\cdot}$).

L_0 mide la desviación de la distribución $\{c_{ig}\}$ del patrón de independencia; se le ha llamado "contenido de información" o "transinformación" de la distribución $\{c_{ig}\}$; ver [4], Cap. VI.

En vista de la interpretación esbozada en el párrafo anterior, se le llamará aquí la heterogeneidad de estructura urbana (distribución de población por tamaños de localidad) entre las regiones o simplemente heterogeneidad urbana interregional.

Se tiene entonces que:

$$(6.5) \quad \sum_{i=1}^{15} p_{i\cdot} I_{i0} = I_{\cdot 0} + L_0$$

en que $I_{\cdot 0}$ es la concentración geográfica marginal interregional, y L_0 es la heterogeneidad urbana interregional.

Por otra parte, considérense los componentes I_{ig} , ecuación (6.4). El primero es un término de la concentración marginal interna de la región S_g ; en efecto, promediando sobre tamaños de localidad *dentro de la región S_g* se tiene:

$$\sum_{i=1}^{15} \frac{c_{ig}}{P_{\cdot g}} \sum_{j \in S_g} \frac{p_{ij}}{c_{ig}} \log \frac{p_{\cdot j}/P_{\cdot g}}{q_{j\cdot}/Q_g} = \sum_{j \in S_g} \frac{p_{\cdot j}}{P_{\cdot g}} \log \frac{p_{\cdot j}/P_{\cdot g}}{q_{j\cdot}/Q_g} = I_{\cdot g}$$

El segundo es una información indirecta, referida a la distribución de la población que habita en localidades de tamaño i , en los estados que forman la región S_g . La distribución *a priori* es independiente de i y es simplemente la distribución marginal condicional $\{p_{\cdot j}/P_{\cdot g}\}$; la distribución *a posteriori* es la distribución de la población en los estados de la región S_g , restringida a localidades de tamaño i . Este término es cero si:

$$\frac{p_{ij}}{c_{ig}} = \frac{p_{\cdot j}}{P_{\cdot g}}$$

o sea, si la distribución por regiones es la misma para todos los tamaños de localidad. La condición enunciada arriba se puede escribir también:

$$\frac{p_{ij}}{p_{\cdot j}} = \frac{c_{ig}}{P_{\cdot g}}$$

o sea, la condición de que la proporción de la población que habita en localidades de tamaño i sea la misma para todos los estados que constituyen la región S_g .

El segundo componente de $I_{i\cdot}$ en la ecuación (6.2) es la media sobre las regiones de I_{ig} . Usando estos resultados se tiene:

$$\begin{aligned} \sum_{i=1}^{15} p_{i\cdot} \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} I_{ig} &= \sum_{g=1}^9 P_{\cdot g} I_{\cdot g} + \sum_{i=1}^{15} \sum_{g=1}^9 \frac{c_{ig}}{p_{i\cdot}} \dots \\ \dots \sum_{j \in S_g} \frac{p_{ij}}{c_{ig}} \log \frac{p_{ij}/c_{ig}}{p_{\cdot j}/P_{\cdot g}} &= \sum_{g=1}^9 P_{\cdot g} I_{\cdot g} + \sum_{g=1}^9 P_{\cdot g} \dots \\ \sum_{j \in S_g} \frac{p_{\cdot j}}{P_{\cdot g}} \sum_{i=1}^{15} \frac{p_{ij}}{p_{i\cdot}} \log \frac{p_{ij}/p_{\cdot j}}{c_{ig}/P_{\cdot g}} &= \sum_{g=1}^9 P_{\cdot g} I_{\cdot g} + \\ + \sum_{g=1}^9 P_{\cdot g} \sum_{j \in S_g} \frac{p_{\cdot j}}{P_{\cdot g}} L_{\cdot j} \end{aligned}$$

en que $L_{\cdot j}$ es una información indirecta con distribución *a priori* $\{c_{ig}/P_{\cdot g}\}$ y *a posteriori* $\{p_{ij}/p_{\cdot j}\}$. Evidentemente, $L_{\cdot j}$ mide la desviación de la "estructura urbana" del estado respecto a la de su región. La media sobre los estados de S_g de $L_{\cdot j}$:

$$L_{\cdot g} = \sum_{j \in S_g} \frac{p_{\cdot j}}{P_{\cdot g}} L_{\cdot j}$$

se llamará *heterogeneidad urbana interna* de la región S_g ; la media sobre regiones, $\sum P_{\cdot g} L_{\cdot g}$, será la *heterogeneidad urbana intrarregional*. Se tiene así que:

$$(6.6) \quad \sum_{i=1}^{15} \sum_{g=1}^9 c_{ig} I_{ig} = \sum_{g=1}^9 P_{\cdot g} I_{\cdot g} + \sum_{g=1}^9 P_{\cdot g} L_{\cdot g}$$

en que $I_{\cdot g}$ es la concentración geográfica marginal interna de la región S_g y $L_{\cdot g}$ su heterogeneidad urbana interna.

Sumando los componentes de $I_{i\cdot}$ y promediándolos sobre i se tiene:

$$(6.7) \quad \sum_{i=1}^{15} p_{i\cdot} I_{i\cdot} = I_{\cdot 0} + L_{\cdot 0} \sum_{g=1}^9 P_{\cdot g} I_{\cdot g} + \sum_{g=1}^9 P_{\cdot g} L_{\cdot g} =$$

$$\begin{aligned}
&= \sum_{i=1}^{15} \sum_{j=1}^{32} p_{ij} \log \frac{p_{ij}/p_{i\cdot}}{q_j} = \\
&= \sum_{j=1}^{32} p_{\cdot j} \log \frac{p_{\cdot j}}{q_j} + \sum_{i=1}^{15} \sum_{j=1}^{32} \log \frac{p_{ij}}{p_{i\cdot} p_{\cdot j}} = \\
&= I(p:q) + L
\end{aligned}$$

en que $L = L_{\cdot 0} + \sum P_{\cdot g} L_{\cdot g}$ es la *heterogeneidad urbana total* (igual al "contenido de información" de la distribución $\{p_{ij}\}$). Resumiendo se tiene:

$$\begin{aligned}
&\text{Concentración geográfica marginal interregional} \\
&+ \text{Heterogeneidad urbana interregional} \\
&\hline
&= \text{Concentración media interregional}
\end{aligned}$$

$$\begin{aligned}
&\text{Concentración geográfica marginal intrarregional} \\
&+ \text{Heterogeneidad urbana intrarregional} \\
&\hline
&= \text{Concentración media intrarregional}
\end{aligned}$$

$$\begin{aligned}
&\text{Concentración geográfica marginal interregional} \\
&+ \text{Concentración geográfica marginal intrarregional} \\
&\hline
&= \text{Concentración geográfica marginal total}
\end{aligned}$$

$$\begin{aligned}
&\text{Heterogeneidad urbana interregional} \\
&+ \text{Heterogeneidad urbana intrarregional} \\
&\hline
&= \text{Heterogeneidad urbana total}
\end{aligned}$$

$$\begin{aligned}
&\text{Concentración media interregional} \\
&+ \text{Concentración media intrarregional} \\
&\hline
&= \text{Concentración media total}
\end{aligned}$$

7. CONCENTRACIÓN GEOGRÁFICA Y ESTRUCTURA URBANA REGIONAL DE MÉXICO, 1930-1960

Es fácil visualizar las interrelaciones presentadas en la sección anterior, al examinarlas a la luz de los datos censales de México, de 1930 a 1960. El cuadro 7 presenta la descomposición de la media $\sum p_{i\cdot} I_{i\cdot}$, de acuerdo con las ecuaciones (6.2) a (6.6); la media $\sum p_{i\cdot} I_{i\cdot}$ está dada en el último renglón del cuadro (las cifras entre paréntesis son porcentajes de esta media). La descomposición (6.2) corresponde a los últimos renglones de las dos secciones del cuadro:

$$\sum_{i=1}^{15} p_{i\cdot} I_{i\cdot} = \sum_{i=1}^{15} p_{i\cdot} I_{i0} + \sum_{i=1}^{15} \sum_{g=1}^9 c_{ig} I_{ig}$$

Cuadro 7

MÉXICO: CONCENTRACIÓN GEOGRÁFICA DE LA POBLACIÓN Y ESTRUCTURA
URBANA REGIONAL, 1930-1960
(Componentes en *bits*)

| | 1930 | 1940 | 1950 | 1960 | Promedio 1930-1960 |
|---------------------------|----------|----------|----------|----------|-----------------------|
| Concentración geográfica | 0.3754 | 0.3290 | 0.3400 | 0.3321 | 0.3441 |
| Marginal interregional | (18.40) | (15.40) | (14.56) | (13.94) | (15.48) |
| Heterogeneidad urbana | 0.2707 | 0.3209 | 0.3500 | 0.2862 | 0.3070 |
| Interregional | (13.27) | (15.02) | (14.99) | (12.62) | (13.81) |
| Concentración media | 0.6461 | 0.6499 | 0.6900 | 0.6183 | 0.6511 |
| Interregional | (31.67) | (30.41) | (29.55) | (25.96) | (29.28) |
| Concentración geográfica | 1.0095 | 1.0472 | 1.2103 | 1.3097 | 1.1442 |
| Marginal intrarregional | (49.48) | (49.01) | (51.83) | (54.99) | (51.46) |
| Heterogeneidad urbana | 0.3849 | 0.4398 | 0.4349 | 0.4535 | 0.4283 |
| Intrarregional | (18.86) | (20.59) | (18.62) | (19.04) | (19.26) |
| Concentración media | 1.3943 | 1.4870 | 1.6452 | 1.7632 | 1.5724 |
| Intrarregional | (68.33) | (69.60) | (70.45) | (74.04) | (70.72) |
| Concentración geográfica | 1.3849 | 1.3762 | 1.5503 | 1.6418 | 1.4883 |
| Marginal total | (67.87) | (64.40) | (66.39) | (68.94) | (66.94) |
| Heterogeneidad urbana | 0.6556 | 0.7607 | 0.7849 | 0.7397 | 0.7352 |
| Total | (32.13) | (35.60) | (33.61) | (31.06) | (33.06) |
| Concentración media total | 2.0404 | 2.1369 | 2.3352 | 2.3815 | 2.2235 |
| | (100.00) | (100.00) | (100.00) | (100.00) | (100.00) |

La descomposición (6.5) corresponde a la primera sección del cuadro; la (6.6) a la segunda; la última ecuación de (6.7) a la última sección del cuadro. Se advierte que la heterogeneidad urbana total —bastante estable como fracción de la concentración media— no es despreciable, de modo que si se considera que se incurre en un sesgo al estimar la concentración geográfica de la población *via* su distribución marginal en lugar de la media de concentraciones condicionales, este sesgo resulta, en promedio para el período 1930-1960, del orden del 33 % de la concentración media total. Al nivel interregional, heterogeneidad y concentración marginal son casi iguales.

Desde luego, la forma en que la heterogeneidad urbana se reparte en componentes inter e intrarregional depende de la definición de regiones. Si se considera deseable que, al menos respecto a su estructura urbana, las regiones definidas resulten lo más homogéneas posi-

ble, deberían escogerse aquellas regiones para las que el componente intrarregional de la heterogeneidad urbana sea mínimo (y el intrarregional sea máximo, ya que la heterogeneidad total está dada). Como se puede ver, el esquema de las regiones que se escogió para el ejemplo numérico de este trabajo no es particularmente feliz en este sentido; la heterogeneidad intrarregional es siempre algo superior a la interregional. Dada una heterogeneidad intrarregional, $\sum P \cdot_g L \cdot_g$, es siempre posible encontrar qué regiones y qué estados contribuyen más a la heterogeneidad; en una región en la que el valor de $P \cdot_g L \cdot_g$ sea elevado, el o los estados de mayor contribución (medida por $\frac{P \cdot_j}{P \cdot_g} L \cdot_j$) son candidatos a formar parte de otra región.

8. ANÁLISIS DE LA CONCENTRACIÓN URBANA

En forma análoga a la empleada en la sección 6, se pueden definir las concentraciones urbanas de las regiones y de los estados:

$$(8.1) \quad R \cdot_g^0 = \sum_{i=1}^{15} \frac{c_{ig}}{P \cdot_g} \log \frac{c_{ig}/P \cdot_g}{\theta_i}$$

es la concentración urbana interna de la región S_g ;

$$(2.8) \quad R \cdot_j = \sum_{i=1}^{15} \frac{p_{ij}}{P \cdot_j} \log \frac{p_{ij}/P \cdot_j}{\theta_i}$$

es la concentración urbana del estado j . Se encuentra de inmediato la siguiente expresión para la concentración urbana media de las regiones:

$$(8.3) \quad \sum_{g=1}^9 P \cdot_g R \cdot_g^0 = \sum_{g=1}^9 \sum_{i=1}^{15} c_{ig} [\log c_{ig}/p_i \cdot P \cdot_g + \log p_i \cdot / \theta_i] = L_0 + R$$

esto es, la heterogeneidad urbana interregional más la concentración urbana marginal. Por otra parte, promediando $R \cdot_j$ para los estados de una misma región, se tiene:

$$(8.4) \quad \sum_{j \in S_g} \frac{P \cdot_j}{P \cdot_g} R \cdot_j = \sum_{j \in S_g} \frac{P \cdot_j}{P \cdot_g} \sum_{i=1}^{15} \frac{p_{ij}}{P \cdot_j} [\log \frac{p_{ij}/c_{ig}}{P \cdot_j/P \cdot_g} + \log \frac{c_{ig}/P \cdot_g}{\theta_i}] = \sum_{j \in S_g} \frac{P \cdot_j}{P \cdot_g} L \cdot_j + R \cdot_g^0 = L \cdot_g + R \cdot_g^0$$

Finalmente, tomando la media de $R \cdot_j$ sobre todos los estados del país se tiene:

$$(8.5) \quad \sum_{j=1}^{32} p \cdot_j R \cdot_j = \sum_{g=1}^9 P \cdot_g (L \cdot_g + R \cdot_g^0) =$$

$$= R + L_0 + \sum_{g=1}^9 P \cdot_g L \cdot_g$$

La suma de los términos primero y tercero de (8.5) es la concentración urbana media de las regiones. Se tiene así que:

$$\begin{aligned} & \text{Concentración urbana marginal} \\ & + \text{Heterogeneidad urbana interregional} \\ & \hline & = \text{Concentración urbana media de las regiones} \\ & + \text{Concentración urbana intrarregional} \\ & \hline & = \text{Concentración urbana media de los estados} \end{aligned}$$

Si de nuevo la heterogeneidad urbana se considera como un sesgo —de agregación al tomar medias regionales en lugar de medias estatales, y de “marginalización” al tomar concentración marginal en lugar de la media de concentraciones (condicionales) “internas” de estados y regiones— se puede advertir que este sesgo puede ser importante; en el caso de México, en el período 1930-1960 significa alrededor de un 20 % de la concentración media estatal; de este 20 %, 8 se “recupera” al pasar de la distribución marginal a la media regional; el restante 12 se recupera al desagregar las regiones y promediar sobre los estados. (Véase el cuadro 8.)

Cuadro 8

MÉXICO: COMPONENTES DE LA CONCENTRACIÓN URBANA, 1930-1960
(En *bits*)

| | 1930 | 1940 | 1950 | 1960 | Promedio 1930- 1960 |
|-------------------------------|----------|----------|----------|----------|---------------------------|
| Concentración urbana | 2.7440 | 3.0208 | 3.5565 | 3.6828 | 3.2510 |
| Marginal | (80.72) | (79.88) | (81.92) | (83.28) | (81.56) |
| Heterogeneidad | 0.2707 | 0.3209 | 0.3500 | 0.2862 | 0.3070 |
| Urbana interregional | (7.96) | (8.49) | (8.06) | (6.47) | (7.70) |
| Concentración urbana media | 3.0147 | 3.3417 | 3.9065 | 3.9690 | 3.5580 |
| Regional | (88.68) | (88.37) | (89.98) | (89.75) | (89.26) |
| Heterogeneidad urbana | 0.3849 | 0.4398 | 0.4349 | 0.4535 | 0.4283 |
| Intrarregional | (11.32) | (11.63) | (10.02) | (10.25) | (10.74) |
| Concentración urbana | 3.3996 | 3.7815 | 4.3414 | 4.4225 | 3.9863 |
| Media estatal | (100.00) | (100.00) | (100.00) | (100.00) | (100.00) |

REFERENCIAS

- [1] HOOPER, J. y H. THEIL, "The Information Approach to the Measurement of Income Inequality". Documento mimeografiado 6501, Instituto de Econometría, Rotterdam, 1965.
- [2] KHINCHIN, A. I., "The Entropy Concept in Probability Theory" (traducción al inglés del original ruso en *Uspiehi Matematicheskii Nauk*, 8 (1953), pp. 3-20), en *Mathematical Foundations of Information Theory*, by A. I. Khinchin, Nueva York, Dover, 1957, pp. 1-28.
- [3] KOOPMAN, B. O. y G. KIMBALL, "Information Theory", en *Notes on Operations Research*. Operations Research Center, Cambridge, Mass., M. I. T., 1959.
- [4] REZA, F. F., *Introduction to Information Theory*, Nueva York, McGraw Hill, 1961.
- [5] SECRETARÍA DE HACIENDA Y CRÉDITO PÚBLICO, Ponencia presentada por el Depto. de Programación Económica y Fiscal de la Dirección General de Estudios Hacendarios, al V Congreso Internacional de Planificación, México, 1965.
- [6] THEIL, H., *Economic Applications of Information Theory*. Notas mimeografiadas. Universidad de Chicago, 1965, Caps. IV y VIII.
- [7] THEIL, H., M. SCHOLLES y P. URIBE, "An Informational Approach to the Measurement of Industrial Concentration", *The Journal of Political Economy*, por aparecer.

APÉNDICE I

Sobre los datos estadísticos básicos

La distribución de territorio utilizada en el artículo es la dada por el *Anuario Estadístico de los Estados Unidos Mexicanos*, Dirección General de Estadística, 1963. Las distribuciones de población y localidades $\{h_{ij}\}$ y $\{k_{ij}\}$, han sido obtenidas de los V, VI, VII y VIII Censos Generales de Población. La distribución $\{k_{ij}\}$ es la marginal de localidades, obtenida directamente de los censos, que dan el número de localidades por intervalos en cada estado. La distribución $\{h_{ij}\}$ es la dada por los censos en 1950 y 1960; el censo de 1940 no proporciona distribución de población. Ésta se ha estimado como sigue; para $i \leq 5$:

$$h_{ij} (1940) = \bar{x}_{ij} (1950) k_{ij} (1940)$$

en que \bar{x}_{ij} es el tamaño medio de una localidad entre x_{i-1} y x_i habitantes en el estado j y k_{ij} es el número de estas localidades en el estado j . La misma relación sirvió para estimar la distribución h_{ij} para el Distrito Federal en 1930. Para $i > 5$, se usaron datos estimados (inéditos) del Centro de Estudios Económicos y Demográficos de El Colegio de México. Para el resto de los estados y el año 1930, se usaron datos del Censo.

Ciertas afinaciones y correcciones han sido necesarias en casi todos los cuadros (por ejemplo, h_{ij}/k_{ij} salía del intervalo $[x_{i-1}, x_i)$ al usarse el dato censal exacto; en tales casos se recalculó $h_{ij} = \bar{x}_{ir} k_{ij}$, en que r es un estado vecino o de condiciones geográficas y económicas semejantes a las de j , etc.).

Por estas razones, puede haber algunas discrepancias entre el material utilizado en este trabajo y las cifras de los censos.

Por último, los intervalos utilizados fueron aquellos que permitieron comparar los cuatro censos, ya que en cada uno se presentan intervalos diferentes.

Las tabulaciones de este material fueron tiradas en mimeógrafo y están a la disposición de quien las solicite al Centro de Estudios Económicos y Demográficos de El Colegio de México.

APÉNDICE II

Una medida de concentración para la distribución conjunta (p_{ij})

Considérese la matriz (p_{ij}), en que p_{ij} es la proporción de población que vive en localidades entre x_{i-1} y x_i habitantes (por brevedad, de tamaño i), en el estado j . Las medidas simples como la entropía y la redundancia poseen extensiones inmediatas; la entropía de (p_{ij}) está dada por:

$$(A2.1) \quad H' = \sum_{i=1}^m \sum_{j=1}^n p_{ij} \log \frac{1}{p_{ij}}$$

cuyo máximo, bajo la sola restricción de que

$$\sum_{i=1}^m \sum_{j=1}^n p_{ij} = 1,$$

está dado por $\log m n$. La matriz (p_{ij}^0) de entropía máxima es entonces la definida por $p_{ij}^0 = 1/m n$.

La redundancia de (p_{ij}) queda así definida como:

$$(A2.2) \quad R' = \log m n - H' = \sum_{i=1}^m \sum_{j=1}^n p_{ij} \log \frac{p_{ij}}{1/m n}$$

que es una información indirecta con probabilidades *a priori* iguales a $1/m n$. La matriz (p_{ij}^0) de máxima entropía tiene entonces distribuciones marginales:

$$p_{i\cdot}^0 = \sum_{j=1}^n p_{ij}^0 = 1/m \quad (i = 1, \dots, m)$$

$$p_{\cdot j}^0 = \sum_{i=1}^m p_{ij}^0 = 1/n \quad (j = 1, \dots, n)$$

por lo cual las probabilidades p_{ij} son no sólo uniformes, sino estocásticamente independientes ($p_{ij}^0 = p_{i\cdot}^0 \cdot p_{\cdot j}^0$).

Las medidas marginales de concentración usadas en las secciones 6, 7 y 8 implican probabilidades *a priori* no uniformes: para q_i se ha usado θ_i , la proporción de localidades en el intervalo i respecto a la totalidad de localidades del país y para q_j se ha usado q_j , la proporción del estado j en el territorio del país. Sin embargo, no se ha implicado ningún supuesto respecto a la estructura de una distribución conjunta *a priori* $\{q_{ij}\}$.

Conviene suponer que, además de la información *a priori* $\{\theta_i\}$ y la información *a priori* $\{q_j\}$, el experimentador no posee *ninguna información previa* a la realización del experimento. En consecuencia, la información indirecta

$$\sum_{i=1}^m \sum_{j=1}^n q_{ij} \log \frac{q_{ij}}{\theta_i q_j}$$

es cero idénticamente. Esto define por completo a las q_{ij} , y se tiene:

$$(A2.3) \quad q_{ij} = \theta_i q_j$$

Se conserva, entonces, el supuesto de independencia estocástica, pero no el de uniformidad de la distribución *a priori*.

La medida de concentración resultante de (A2.3) es entonces una información indirecta con probabilidades *a priori* $\theta_i q_j$ y probabilidades *a posteriori* p_{ij} .

Considérese el logaritmo:

$$(A2.4) \quad \log \frac{p_{ij}}{\theta_i q_j} = \log \frac{p_{i\cdot}}{\theta_i} + \log \frac{P_{\cdot g}}{Q_g} + \log \frac{p_{\cdot j}/P_{\cdot g}}{q_j/Q_g} + \\ + \log \frac{c_{ig}/P_{\cdot g}}{p_{i\cdot}} + \log \frac{p_{ij}/p_{\cdot j}}{c_{ig}/P_{\cdot g}}$$

Se identifican inmediatamente los términos logarítmicos siguientes: el primero corresponde a la concentración urbana marginal R , el segundo a la concentración geográfica marginal interregional I_0 , el tercero a la intrarregional, el cuarto a la heterogeneidad urbana interregional L_0 , y el quinto a la intrarregional.

De acuerdo con esto, se puede escribir:

$$(A2.5) \quad C = \sum_{i=1}^m \sum_{j=1}^n p_{ij} \log \frac{p_{ij}}{\theta_i q_j} = \sum_{i=1}^m p_{i\cdot} \log \frac{p_{i\cdot}}{\theta_i} + \\ + \sum_{g=1}^G P_{\cdot g} \log \frac{P_{\cdot g}}{Q_g} + \sum_{g=1}^G \sum_{j \in S_g} p_{\cdot j} \log \frac{p_{\cdot j}/P_{\cdot g}}{q_j/Q_g} + \\ + \sum_{i=1}^m \sum_{g=1}^G c_{ig} \log \frac{c_{ig}/P_{\cdot g}}{p_{i\cdot}} + \sum_{i=1}^m \sum_{g=1}^G \sum_{j \in S_g} \dots \\ \dots p_{ij} \log \frac{p_{ij}/p_{\cdot j}}{c_{ig}/P_{\cdot g}} =$$

$$\begin{aligned}
 &= R + I_0 + \sum_{g=1}^G P_{\cdot g} I_{\cdot g} + L_0 + \sum_{g=1}^G P_{\cdot g} L_{\cdot g} = \\
 &= R + I + L,
 \end{aligned}$$

en que L es el *contenido de información* (ver [4], Cap. VI) de la distribución (p_{ij}) :

$$L = \sum_{i=1}^m \sum_{j=1}^n p_{ij} \log \frac{p_{ij}}{p_{i\cdot} p_{\cdot j}}$$