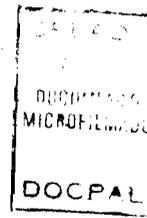


RESERVADO

September de 1983



INFORME DE LA MISION DE ASESORAMIENTO REALIZADA EN LA  
REPUBLICA DE VENEZUELA DESDE EL 27 DE JULIO AL 25 DE AGOSTO DE 1983

Carlos Cavallini

Asesor en Estadísticas en misión para  
las Naciones Unidas

Las opiniones expresadas en este informe son de responsabilidad del autor.

83-9- 1533

MISION EN LA REPUBLICA DE VENEZUELA

1. Organismo solicitante. Oficina Central de Estadística e Informática, OCEI.
2. Organismo auspiciador. Comisión Económica para América Latina, CEPAL.
3. Términos de la misión. Asesorar en la Encuesta de Evaluación Censal y en el Marco Muestral Computarizado.
4. Lugar físico de trabajo. OCEI, en Caracas.
5. Tiempo de duración de la misión. Desde el 27 de julio hasta el 25 de agosto de 1983, lo cual abarcó un total de 30 días.
6. Objetivos principales llevados a cabo Párrafo
  - i) Evaluación del XI Censo de Población y Vivienda de Venezuela de 1981 9
  - ii) Tasas de omisión de personas, por estrato, sexo y edad, del Censo de Población de 1981 10
  - iii) Marco Muestral 17
    - a) Tamaños muestrales 19.1
    - b) Programa logicial para la selección de muestras 20
7. Funcionarios. Para definir los alcances de la misión y para participar en las tareas de desarrollo de la misma, se mantuvo una estrecha colaboración con los siguientes funcionarios.

Supelano, Ofelia R. de

Jefe de la OCEI

Quintero, Carmen M. de

Directora General Sectorial de Estadística

CELADE - SISTEMA DOCPAL  
DOCUMENTACION  
SOBRE POBLACION EN  
AMERICA LATINA

Grisanti, Mireya de	Directora de Estadísticas de Población
Cáceres, Josefina	Directora de Estadísticas Sociales
Rodríguez, Rahiza de	Encuesta de Evaluación
Fervenza, María José	Encuesta de Evaluación
Brando, José	Encuesta de Evaluación
Escalona, Nelson	Asesor Dirección Estadísticas de Población
Saint Sourin, Jacques	Asesor Dirección Estadísticas de Población
Urbáez, Luis Domingo	Asesor Censal
Bonnin, Cristóbal	Director de Computación
Bellucci, Juan Carlos	Computación
González, Roger	Encuesta de Hogares
Alvarez, Daysi	Operaciones de campo
Ríos, Francisco	Encuesta de Hogares
Ricci, Elizabeth de	Cartografía

8. Las tareas estadísticas llevadas a cabo, fueron desarrolladas bajo la condición de que las mismas pudieran realizarse a través de programas paramétricos computarizados. Estos programas genéricos permiten procesar la información, en función de la variación de los parámetros introducidos, tantas veces como sea necesario.

La construcción del modelo estadístico para evaluar la cobertura del censo de población, consideró, entre otros, como características paramétricas, al estrato geográfico y al grupo de edad por sexo, así como a la ponderación de los estratos. En los modelos para las estimaciones de los tamaños muestrales y para la selección de las unidades muestrales, son parámetros los supuestos propios del diseño, las unidades de selección en las distintas etapas y el número de etapas a considerar. Si bien, en este escrito se presenta el flujograma de selección de unidades con la limitación de un máximo de 3 etapas, su paso a un número mayor de etapas es inmediato, dado que la lógica de selección para cada etapa es similar.

## 9. EVALUACION DEL XI CENSO GENERAL DE POBLACION Y VIVIENDA DE VENEZUELA, 1981

### 9.1. Introducción.

El Censo de Población y Vivienda es, sin dudas, el operativo estadístico de mayor envergadura al cual se deben enfrentar, periódicamente, las Oficinas Nacionales de Estadística de los distintos países del mundo. En él interviene toda la población de un país, a la cual, en un momento dado, se trata de medir, persona por persona, sobre distintas características demográficas, económicas y sociales, que son fundamentales conocer para poder planificar y establecer las diferentes políticas de desarrollo, tanto regionales como nacionales. Por otro lado, decenas de miles de agentes en el terreno y en las oficinas, previamente entrenados para este trabajo, son los encargados, a lo largo de todo el país, de entrevistar a estas personas, registrar la información requerida, transportarla a los lugares de centralización, analizarla y procesarla. En todas estas etapas intervienen personas, unas para dar la información y otras para recibirla. En todas estas etapas se pueden cometer errores introducidos consciente o inconscientemente, ya sea por el que da o por el que recibe la información, o también por accidente, o por pérdida de algún registro, o por fallas mecánicas del proceso, etc. Es por todo ello, que un censo requiere de una planificación rigurosa y de procedimientos que permitan conocer el grado de eficiencia con el cual se ha trabajado en cada una de estas etapas, de la confiabilidad de la información recogida, y de la acuracidad de los resultados que le serán otorgados a los usuarios.

La práctica ha demostrado que un censo exacto es imposible, los errores existirán inevitablemente. El público en general, y los usuarios de la información censal en particular, deben estar en situación de comprender que la estadística no trabaja con valores exactos sino con valores confiables, y para que esta estadística sea de utilidad, esa confiabilidad debe conocerse. Todos los valores que se conocen del universo físico, regido por leyes naturales, como del universo social, regido por la voluntad del hombre, son valores estimados, todos con error, obtenidos a través de la aplicación de adecuadas técnicas estadísticas. Lo importante es tratar de descubrir estos errores y tratar de limitarlos.

### 9.2. Evaluación Censal.

Evaluar, en el uso coloquial, significa calcular el valor de una cosa, o también, el de otorgarle algún valor a algo. Desde el punto de vista estadístico, se denomina evaluación al conocimiento del valor de una investigación dentro de determinados límites de confiabilidad. En este caso, "valor" tiende a ser equivalente a "calidad".

La evaluación de un censo no es sólo conveniente sino también necesaria. Una obligación que debe prevalecer en toda investigación censal es dar suficiente información sobre la calidad de los resultados que se presenten. Con ello se logrará, por un lado, que el usuario de los resultados conozca las limitaciones de los mismos, y por otro lado, mejorar el diseño de la investigación para sucesivas ocasiones. Además, la evaluación es la única protección real contra ataques injustificados a la calidad del censo y a la competencia e integridad de sus productores. Las Naciones Unidas<sup>1/</sup>, la Organización de Estados Americanos, el Instituto Internacional de Estadística, y los numerosos centros de investigación estadística, recomiendan la evaluación de los censos de población y la inclusión, en la publicación de los resultados del censo, de una estimación del error de cobertura, es decir, del grado de omisión, así como de la bondad de la información recogida. Por ello, instan a los países a considerar dentro de sus cronogramas censales, la evaluación estadística de los mismos.

### 9.3. Objetivos de la evaluación censal.

El objetivo más importante que persigue la evaluación de un censo es medir su acuracidad, es decir, estimar el grado de desvío existente entre los valores reales de la población y los valores arrojados por el censo. En base a este conocimiento se podrán ajustar los valores censales y se podrán obtener los valores proyectados de la población para un momento dado. Una categoría de error, que se intenta medir a través de un programa de evaluación, es el error de cobertura, que comprende tanto a la sobreenumeración como a la subenumeración de personas. Pero, dado que el problema crónico de los censos es la omisión de personas, el error de cobertura se refiere a esta deficiencia. La otra categoría de error se refiere a la bondad de la información recogida. Este error puede resultar de una información errónea o inconsistente de las características que se recogen, o del fracaso para obtener parte de la información solicitada debido a descuido o mala comprensión del empadronador. También puede resultar de ignorancia del informante o mala interpretación o conceptos equivocados del dato requerido, a los errores de oficina o mecánicos en el procesamiento mismo de los datos, a los métodos de revisión e imputación usados para datos inaceptables u omitidos, etc.

La mayoría de los países necesitan conocer la omisión censal y la bondad de las características medidas para distintas áreas geográficas y de acuerdo con ciertas categorías de la población. Así, por ejemplo, se querrán estimaciones separadas de la tasa de omisión por sexo y edad para distintas entidades, para población urbana y rural, para grandes conglomerados, etc. Se querrán conocer las distintas tasas de confiabilidad que permitan estimar

<sup>1/</sup> Naciones Unidas, "Principios y Recomendaciones relativos a los Censos de Población" Series M, No. 44, 1970.

la bondad de las variables "número de cuartos por vivienda", "edad", "lugar de nacimiento", "situación de trabajo", etc.

#### 9.4. Distintos métodos de evaluación.

Para evaluar la cobertura y la calidad de un censo existen varios métodos. Todos ellos tienen sus propias virtudes y la elección de uno u otro dependerá de la infraestructura de la Oficina de Estadística, del costo, de la precisión, de los niveles de estimación, de la coordinación con otros trabajos, de los subproductos que dicha investigación puede aportar, etc.

Entre los métodos para evaluar la cobertura y la calidad censal podemos citar:

- i) Métodos analíticos. Basados, generalmente, en proyecciones demográficas, que a su vez toman como base un censo anterior. Utilizando estos métodos se puede estimar la población esperada de un censo y su composición por sexo y edad.
- ii) Métodos de Registro. Este método utiliza la organización de registros que posee un país. Básicamente consiste en la siguiente ecuación compensatoria: Censo anterior + nacimientos en el período + entradas al país - defunciones - salidos del país = censo actual. Es un método eficiente cuando existen buenos registros administrativos, pero costoso. Por ejemplo, si se selecciona de un registro de nacimiento un niño nacido antes del censo, para controlar si fue o no censado, el seguimiento de este niño puede resultar oneroso. Otro problema que presenta este método con respecto a su aplicación en los países latinoamericanos, es que éstos, generalmente, carecen de registros confiables.
- iii) Métodos directos o de re-entrevista. Son métodos que utilizan la reentrevista censal. Se llevan a cabo, generalmente, inmediatamente después del levantamiento censal. Estos métodos permiten estimar, entre otros, la omisión censal de personas y de viviendas, la distribución geográfica de las mismas, las causas de la omisión, la duplicación de personas y de viviendas, las causas de la duplicación y, además, la calidad de la información censal para distintas variables.
- iv) Método de Chandra-Sekar y Deming. Una variación del método directo descrito, es el ingenioso método denominado, en honor de sus autores, de Chandra-Sekar y Deming. El mismo permite estimar, además, la omisión de personas omitidas por el censo y la encuesta.
- v) Métodos directos coordinados con una Encuesta Continua de Hogares. Para reducir los costos de los métodos directos, generalmente se coordinan éstos para que la muestra seleccionada sirva para el comienzo de una encuesta actualizada continua de hogares. De esta manera, los costos atribuibles a la evaluación son prácticamente bajos, dado que la mayor parte de los costos son absorbidos por la Encuesta de Hogares.

- vi) Métodos de autoevaluación. Se basa en el supuesto de que la población de hecho de un país es igual a la población de derecho. Por ejemplo, si un censo es de hecho, se pueden registrar también las personas de derecho de la vivienda y cotejar si éstas han sido censadas. En el caso de que el censo sea de derecho se registra, además, las personas que están de hecho para el posterior cotejo.
- vii) Encuestas en funcionamiento. En algunos casos, para evaluar la cobertura y la bondad de la información de un censo, se utiliza una encuesta que ya está en funcionamiento y que ha sido programada para otros propósitos. Si bien esto es mejor que nada, generalmente estas encuestas pueden estar basadas en marcos ya anacrónicos, lo cual podría afectar la representatividad de la misma.
- viii) Método de las Submuestras interpenetrantes. Consiste en dividir aleatoriamente a las unidades censales del censo en subcensos, o submuestras, denominadas interpenetrantes, dado que cada una de ellas tiene el mismo grado de representatividad. Estos subcensos pueden ser analizados independientemente y en base a los resultados de cada uno de ellos, se puede estimar la variación o confiabilidad del censo.

#### 9.5. Elección y justificación del método.

La muestra fue programada para evaluar el XI Censo de Población y Vivienda de 1981 en su cobertura y en su contenido. El método que se utilizó fue el desarrollado por Chandra-Sekar y Deming<sup>2/</sup>. Este método se basa en el procedimiento de encuestar a una muestra de la población de estudio y luego cotejar la información de la encuesta con la información censal. La decisión de usar este método directo de evaluación, surgió luego de analizar los distintos métodos enunciados previamente, tanto en función de la confiabilidad de los resultados que se obtendrían de los mismos, como del costo y del uso que se les podría dar en futuras investigaciones.

A los métodos analíticos, o basados en registros, no se les veía mucha aplicación; por lo general, debido a la pobre información existente en los sistemas de registro de los hechos vitales. Ello motivó que se decidiera utilizar un método directo de re-entrevista sobre una muestra de los hogares censados. Se eligió el método de Chandra-Sekar y Deming debido a las siguientes consideraciones.

- i) Permite estimar la omisión de personas que han sido omitidas tanto por el censo como por la encuesta, lo cual reduce el sesgo de la información.
- ii) Permite estimar la omisión por pérdida de material censal.

<sup>2/</sup> Chandra-Sekar y W.E. Deming. "On a Method of Estimating Death Rates and the Extent of Registration", JASA. Vo. 44, marzo de 1949.

- iii) Permite ser estructurado, de tal manera, que la muestra seleccionada para la evaluación, pueda luego, ser utilizada como una muestra actualizada de hogares para satisfacer otras necesidades, lo cual acarrea una reducción significativa del costo por unidad de información de ambas investigaciones.
- iv) Su formulación matemática es simple y flexible, en el sentido de que los resultados pueden ser adicionados o desagregados en forma fácil, según las distribuciones que se requieran.
- v) Su diseño se basa en la teoría de las muestras probabilísticas, lo cual permite calcular la precisión de sus estimaciones.
- vi) En América ha sido usado con éxito en numerosas investigaciones, lo cual aumenta la experiencia que del mismo se tiene. En las evaluaciones censales se le ha aplicado entre otros, en la evaluación de los censos de los Estados Unidos de Norte América de 1960, del Ecuador de 1974 y de Bolivia de 1976.

#### 9.6. Descripción del método.

El método directo de Chandra-Sekar y Deming, permite evaluar la sobreenumeración y subenumeración de personas y viviendas, la distribución geográfica de las mismas, las causas de omisión y duplicación y la bondad de las preguntas censales.

##### 9.6.1. El método comprende los siguientes pasos:

- a. Selección aleatoria de una muestra de áreas
- b. Enumeración de las viviendas de las áreas seleccionadas
- c. Selección aleatoria de una muestra de viviendas de las áreas seleccionadas.
- d. Entrevista a las viviendas seleccionadas
- e. Cotejo de la encuesta versus el censo
- f. Conciliación de la información en caso de discrepancia
- g. Estimación de resultados

9.6.2. Cobertura geográfica.

Abarcó el ámbito geográfico del país, tal cual fue definido para el XI Censo de Población y Vivienda.

9.6.3. Cobertura de estudio.

Abarcó los hogares particulares de acuerdo con la definición dada por el Censo, y además, aquellos hogares colectivos con un máximo de 30 personas.

9.6.4. Niveles geográficos de estimación.

La muestra se diseñó para proveer estimaciones confiables de la evaluación del Censo, en los siguientes niveles geográficos de estimación: nacional, regional, estatal, grandes conglomerados (localidades con más de 100.000 habitantes), resto urbano y área rural.

9.6.5. Principales características de estudio.

La principal característica de estudio para la evaluación de la cobertura del censo, será la omisión de personas. Es decir, aquellas personas que debiendo haber sido censadas no lo fueron, o que habiendo sido censadas no ingresaron al proceso censal. Por otra parte, se darán estimaciones de la omisión de viviendas particulares, es decir, viviendas que no fueron censadas o que habiéndolo sido no ingresaron al proceso censal. Respecto del estudio para estimar la bondad del Censo, fueron fijadas las siguientes características.

- a) Relación de parentesco con el Jefe del Hogar
- b) Edad
- c) Sexo
- d) Situación conyugal
- e) Lugar de nacimiento
- f) Nacionalidad actual
- g) Situación de Trabajo

9.6.6. Dominios de estudio.

Las tasas de omisión de personas y bondad de las variables, serán dadas por sexo y grupos de edades, dentro de los niveles geográficos de estimación establecidos.

9.6.7. Marco muestral.

Se utilizó el marco elaborado por la Unidad de Computación. El mismo está basado en el marco del Censo, el cual comprende los grandes conglomerados, las localidades con más de 10.000 habitantes, las localidades de 1.000 a 9,999 habitantes y los centros poblados con menos de 1.000 habitantes.

9.6.8. Unidades de selección:

Primera etapa: Manzanas (área urbana)

Segmento censal (área urbana-zona de ranchos)

Sector (área rural)

Segunda etapa: La unidad de selección fue la vivienda particular de acuerdo con la definición dada por el Censo, las cuales fueron seleccionadas dentro de las unidades de selección de primera etapa que integran la muestra.

9.6.9. Unidad muestral de medición.

La unidad última de selección fue la vivienda, pero la unidad de medición fue el hogar o los hogares que integran dicha vivienda en el momento de la investigación.

9.6.10. Método de selección de las unidades muestrales.

Las unidades muestrales de selección de primera etapa (USPE) se seleccionaron con probabilidad proporcional aproximada al tamaño de viviendas, y dentro de las USPE seleccionada, la selección de viviendas en segunda etapa se realizó con igual probabilidad. Es decir, todas las viviendas del país tuvieron la misma probabilidad de ser seleccionadas.

9.6.11. Estratificación geográfica para niveles de estimación.

La muestra proveerá estimaciones para los siguientes estratos geográficos:

- a. Regiones (9) Capital, Central, Centro Occidental, Andes, Nor-Oriental, Zuliana, Guayana, Insular y de los Llanos.
- b. Entidades Federales (23) Distrito Federal, Apure, Anzoátegui, Aragua, Bolívar, Barinas, Carabobo, Cojedes, Guárico, Falcón, Lara, Mérida, Miranda, Monagas, Portuguesa, Táchira, Trujillo, Yaracuy, Nueva Esparta, Sucre, Zulia, Territorios Amazonas y Delta Amacuro.
- c. Grandes Conglomerados (25) Area Metropolitana de Caracas, Valencia. Barquisimeto, Maracay, Ciudad Guayana, San Cristóbal, Ciudad Bolívar, Pto. La Cruz, Cumaná, Cabimas, Maturín, Barcelona, Ciudad Ojeda, Punto Fijo (Cardón-Judibana), Mérida, Valera, Guarenas, Pto. Cabello, Barinas, Los Teques, Coro, Acarigua, Turmero y Porlamar.
- d. Area urbana y rural, ambas por Entidad Federal.(46)

9.6.12. Tamaño muestral de viviendas para todo el país (Ver Cuadro No. 1)

En base a un estudio realizado por la División de Encuesta de Hogares sobre el tamaño estimado de la muestra de hogares a nivel regional y estatal, se configuraron los tamaños muestrales de viviendas para los distintos niveles de estimación. Para evaluar la cobertura de viviendas, se estimó suficiente seleccionar una submuestra de unidades muestrales de selección de primera etapa (USPE) del 10%.

9.6.13. Recolección de la información.

Se comenzó el levantamiento de la encuesta una vez terminado el levantamiento del censo, el 19.11.81., culminando la misma el 30.12.81. El trabajo operativo se inició en los grandes conglomerados, luego en el resto urbano y finalmente en el área rural.

#### 9.6.14. Procesamiento de la Encuesta

Esta etapa comprende, i) el cotejo de la información versus la información del censo, ii) la conciliación de divergencias entre el censo y la encuesta y iii) el procesamiento de la información para obtener resultados de la evaluación. Es de hacer notar que en aquellos casos en los cuales se detectaban diferencias entre la información censal y la información recogida por la encuesta, se volvía a medir la vivienda, para establecer si la divergencia era error del censo o de la encuesta. De esta manera, los resultados que se introducían en el proceso iterativo de los cálculos obedecían a una información real.

Como la etapa de cotejo debió coordinarse en función de la organización que tenía el censo para su programa de análisis, codificación y lectura, la misma se inició en febrero de 1983, después que la Unidad de Lectura Óptica devolvió las primeras cajas con material censal para su archivo. (Ver cuadro 2)

#### 9.7. Estimación de resultados.

El proceso completo de esta evaluación será terminado cuando se hayan procesado totalmente los datos censales.

Los resultados que hasta el momento se han presentado, son aquellos correspondientes a las tasas de omisión por Entidad Federal. El procedimiento utilizado para calcular la estimación de las tasas de omisión de personas del Censo, por Entidad fue, en forma general, el siguiente:

- a) Dentro de cada Entidad se estimaron las tasas de omisión, en forma separada, i) para las ciudades de más de 100.000 habitantes, ii) para el resto de área urbana y iii) para el área rural.

Los estimadores utilizados fueron,

A personas idénticamente clasificadas por el Censo y la Encuesta.

B personas identificadas por el Censo pero no por la Encuesta.

C personas identificadas por la Encuesta pero no por el Censo.

$$X = \frac{BC}{A}$$

estimación de personas no identificadas por el Censo ni por la Encuesta, bajo el supuesto de independencia operativa entre el Censo y la Encuesta.

$$N = A+B+C+X$$

estimación de personas incluidas en la investigación muestral.

$$q = \frac{C+X}{N}$$

estimación de la tasa de omisión.

- b) Las tasas de omisión para el área urbana se construyeron ponderando las tasas obtenidas para las ciudades y las tasas del resto del área urbana, por las respectivas poblaciones que dió el Censo.
- c) Las tasas de omisión por Entidad se obtuvieron ponderando las tasas del área urbana y las tasas del área rural, por las respectivas poblaciones que dió el Censo.
- d) La tasa para el país se construyó, ponderando por la población censada, las tasas de las Entidades Federales.
- e) Se hace notar que " q " es la tasa de omisión con respecto a la población arrojada por el Censo. La tasa de omisión con respecto a la población teórica,  $q^{\circ}$ , se obtiene con

$$q^{\circ} = \frac{q^{\circ}}{1 + q}$$

Oportunamente, a medida que continúe el proceso censal, estas tasas se presentarán desagregadas por sexo y grupos de edades, así como los resultados que estimen la bondad de las principales características medidas por el censo.

TAMAÑO MUESTRAL DE VIVIENDAS DISTRIBUIDO SEGUN LOS DISTINTOS NIVELES GEOGRAFICOS DE ESTIMACION URBANO

REGION	ESTADO	GRAN CONCENTRADO	CENTROS POBLADOS DE 100,000 Y MAS HABITANTES	RESTO	RURAL	TOTAL
VENEZUELA	CAPITAL		14,280	9,720	4,980	28,980
			3,600	360	240	4,200
		Área Metropolitana Miranda (2)	3,000	-	-	3,000
		Guarema Los Teques	600 300	360	240	1,200
CENTRAL		2,040	1,600	960	4,600	
	Aragua	840	420	240	1,500	
	Mareny Turmero	600 240	-	-	-	
	Carabobo	1,200	540	240	1,980	
CENTRO OCCIDENTAL		-	240	-	-	
	Cojedes Guárico	-	240 420	240 240	480 660	
	Falcón	1,740	2,160	1,020	4,920	
	Pto. Fijo, Cardón, Judibona Coro	560 300 240	240	240	1,020	
ANDES		960	780	240	1,980	
	Lara	960	250	240	720	
	Barquisimeto	960	900	300	1,200	
	Portuguesa	240	-	-	-	
	Acarigua	240	900	300	1,200	
	Trujillo	1,140	2,040	1,320	4,500	
	Apure	240	240	240	720	
	Mérida	300	600	300	1,200	
	Táchira	300	600	300	1,200	
	San Cristóbal	300	360	240	900	
MOR-ORIENTAL		300	240	240	780	
	Valera	-	240	240	480	
	Anzoátegui	1,440 600	2,100 720	960 240	4,500 1,560	
	Puerto la Cruz Barcelona	300 300	-	-	-	
	Monagas	300	660	240	1,200	
	Nueva Esparta	300	360	240	840	
	Sucre	240 300	360	240	900	
	Cumaná	300	360	240	900	
	Zulia	2,520	240	240	3,000	
		1,920	240	240	-	
ZULIANA		300	-	-	-	
	Maracaibo Caimán Ciudad Ojeda	300	-	-	-	
	Cumaná	300	360	240	900	
		2,520	240	240	3,000	
GUAYANA(4)		1,800	1,200	240	3,240	
	Bolívar	1,800	720	240	2,760	
	Upata	540	240	-	-	
	Caicara	-	240	-	-	
Delta Amacuro Amazonas		1,260	240	-	-	
	Ciudad Guayana Curi	-	240	-	-	
	Tucupica	-	240	-	-	
	Puerto Ayacucho	-	240	-	-	

1) Incluye Distrito Federal y el Distrito Sucre y el Sector Panamericano del Estado Miranda

2) Excluye el Distrito Sucre y el Sector Panamericano

3) Considerado como gran conglomerado, aunque tiene menos de 100,000 habitantes.

4) De acuerdo con lo solicitado por la Corporación Venezolana de Guayana, C.V.G.

NOTA: Es de hacer notar que en el caso de solicitarse una desagregación distinta e contemplado en el diseño muestral es posible realizar una post-estratificación.

CUADRO No. 2

VENEZUELA: POBLACION EMPADRONADA EN EL XI CENSO GENERAL DE POBLACION Y VIVIENDAS  
Y CORREGIDA MEDIANTE LA ENCUESTA DE EVALUACION DE LA COBERTURA

PAIS ENTIDADES FEDERALES	Población Empadronada (1)	Población Omitida	Población Corregida (2)	Tasa Omision res- pecto a la pobl. técnica
VENEZUELA	14.516.735	1.110.220	15.626.955	7,10
Distrito Federal	2.070.742	212.458	2.283.200	9,31
Estados:				
Anzoátegui	683.717	34.137	717.854	4,76
Apure	188.187	20.155	208.342	9,67
Aragua	891.623	69.691	961.314	7,25
Barinas	326.166	19.612	345.778	5,67
Bolívar	668.340	33.091	701.431	4,72
Carabobo	1.062.268	95.042	1.157.310	8,21
Cojedes	133.991	11.375	145.366	7,83
Falcón	503.896	29.247	533.143	5,49
Guárico	393.467	17.473	410.940	4,25
Lara	945.064	54.767	999.831	5,48
Mérida	459.351	45.679	505.040	9,04
Miranda	1.421.442	123.022	1.544.464	7,97
Monagas	388.536	20.056	408.592	4,91
Nueva Esparta	197.198	10.448	207.646	5,03
Portuguesa	424.984	36.364	461.348	7,88
Sucre	585.698	29.729	615.427	4,83
Táchira	660.234	37.596	697.830	5,39
Trujillo	433.735	39.779	473.514	8,40
Yaracuy	300.597	13.682	314.279	4,35
	1.674.252	150.964	1.825.216	8,27

TASAS DE OMISION DE PERSONAS POR ESTRATO, SEXO Y EDAD DEL CENSO DE POBLACION  
DE 1981

10. Niveles geográficos de estimación:

Se consideraron los siguientes niveles:

Nivel Geográfico de Estimación	Cantidad de Cuadros
(*) .Grandes Conglomerados y Ciudades de más de 100.000 habitantes.....	26
(*) .Resto de Area urbana por Entidad.....	23
. Area urbana por Entidad.....	23
(*) .Area rural por Entidad.....	23
. Entidades.....	23
. País.....	1
. Total General de los grandes conglomerados y ciudades de más de 100.000 habitantes del país.....	1
. Resto Area urbana del país.....	1
. Area urbana del país.....	1
. Area rural del país.....	1
	123

(\*) Estos cuadros no son ponderados. Los demás, son ponderados por la cantidad de habitantes según censo.

11. Dominios de Estudio

Serán por sexo y según grupos de edades. En una reunión en la cual participaron los Demógrafos de la OCEI, Luis Urbáez, Nelson Escalona y Jaques Saint Sourin, los Estadísticos Mireya de Grisanti y Rahiza de Rodriguez se acordó, que los grupos de edades en años, de las distintas distribuciones serán 8, con los siguientes intervalos: menos de 1, 1-4, 5-14, 15-24, 25-44, 45-64 65 y más y "no declarado". Por tanto, los cuadros adoptarán la siguiente presentación, según los distintos niveles geográficos.

Grupos de edades en años	Total	Varones	Mujeres
Menos de 1			
1 - 4			
5 - 14			
15 - 24			
25 - 44			
45 - 64			
65 y más			
No Declarado			
Total			

12. En este punto conviene aclarar que la distribución de las tasas de omisión de personas, por grupos de edades y niveles geográficos, no obedece a una estructura fija. Esto significa que los usuarios estarán en posición de solicitar otras agrupaciones, ya sea de edad o de nivel geográfico, dado que la programación mecánica de esta tarea se está llevando a cabo a través de un programa paramétrico de computación que permite dicha flexibilidad. La información unitaria está archivada en cinta magnética. Hasta los 5 años está agrupada por año y a partir de los 5 años por intervalos de 5 años, hasta 80 y más años. Esto indica que los datos no pueden ser desagregados dentro del intervalo sino por agrupación de intervalos.

Por ejemplo, el programa paramétrico no estará en condiciones de dar la información de "17 y más años" pero sí podrá dar de "15 y más años" o de "20 y más años".

13. A continuación se presentan las fórmulas básicas, que deberán ser programadas en forma genérica, y que serán utilizadas en las estimaciones de los distintos valores, que hacen a la investigación censal, en relación a la cobertura de personas.

### 13.1 Simbología

$A_{hji}$  Personas idénticamente clasificadas por el Censo y la Encuesta del grupo de edad  $i$ , sexo  $j$ , nivel geográfico  $h$ , donde.

$$i = \overline{1,8}$$

$$j = \overline{1,2}$$

$B_{hji}$  Personas identificadas por el Censo pero no por la Encuesta, en  $hji$ -celda.

$C_{hji}$  Personas identificadas por la Encuesta pero no por el censo, en la  $hji$ -celda.

$X_{hji} = \frac{B_{hji} C_{hji}}{A_{hji}}$  Personas no identificadas por el Censo ni por la Encuesta, en la  $hji$ -celda.

$N_{hji} = (A+B+C+X)_{hji}$  Total de personas medidas en la  $hji$ -celda.

$q_1_{hji} = \frac{(C+X)_{hji}}{N_{hji}}$  Tasa de omisión de personas con respecto a la población censada en la  $hji$ -celda.

$P_1_{hji} = \frac{(A+B)_{hji}}{N_{hji}}$   
 $= 1 - q_1_{hji}$  Tasa de personas censadas en la  $hji$ -celda.

$q_2_{hji} = \frac{(B+X)_{hji}}{N_{hji}}$  Tasa de personas no identificadas por la encuesta en la  $hji$ -celda

$P_2_{hji} = \frac{(A+C)_{hji}}{N_{hji}}$

$$= 1 - q_2 \text{ hji}$$

Tasa de personas identificadas por la encuesta en la hji-celda.

$$v (X_{\text{hji}}) \approx \left\{ \frac{N q_1 q_2 (1 - P_1 P_2)^2}{P_1 P_2} \right\}_{\text{hji}}$$

Varianza aproximada de  $X_{\text{hji}}$

$$D (X_{\text{hji}}) = \sqrt{v (X_{\text{hji}})}$$

Desviación estándar de  $X_{\text{hji}}$

$$X_{\text{Ihji}} = X_{\text{hji}} - t D (X_{\text{hji}})$$

Límite de confianza inferior de  $X_{\text{hji}}$ . Los valores de "t" que para estos casos, generalmente, se utilizan son,

<u>t</u>	<u>Confianza Aproximada</u>
1	90 %
2	95 %
3	99 %

$$X_{\text{Shji}} = X_{\text{hji}} + t D (X_{\text{hji}})$$

Límite de confianza superior de  $X_{\text{hji}}$ .

$$N_{\text{I hji}} = (A+B+C+X_{\text{I}})_{\text{hji}}$$

Límite inferior de  $N_{\text{hji}}$

$$N_{\text{S hji}} = (A+B+C+X_{\text{S}})_{\text{hji}}$$

Límite superior de  $N_{\text{hji}}$

$$q_{\text{I hji}} = \frac{(C+X_{\text{I}})_{\text{hji}}}{N_{\text{I hji}}}$$

Límite inferior de  $q_{\text{I hji}}$

$$q_{\text{S hji}} = \frac{(C+X_{\text{S}})_{\text{hji}}}{N_{\text{S hji}}}$$

Límite superior de  $q_{\text{I hji}}$

Se debe tener en cuenta que la siguiente expresión debe prevalecer

$$q_{\text{I I hji}} \leq q_{\text{I hji}} \leq q_{\text{I S hji}}$$

13.2 Así mismo, para totalizar por sexo "j", se tiene

$$A_{h.i} = \sum_j A_{hji} \quad \text{Total de la hi-celda}$$

En forma similar

$$B_{h.i} = \sum_j B_{hji}$$

$$C_{h.i} = \sum_j C_{hji}$$

$$X_{h.i} = \sum_j X_{hji}$$

$$X_{Ih.i} = \sum_j X_{Ihji}$$

$$I_{Sh.i} = \sum_j X_{Shji}$$

$$N_{h.i} = \sum_j N_{hji}$$

$$N_{Ih.i} = \sum_j N_{Ihji}$$

$$N_{Sh.i} = \sum_j N_{Shji}$$

Luego

$$q_{1h.i} = \frac{(C+X)_{h.i}}{N_{h.i}}$$

$$q_{I1h.i} = \frac{(C+XI)_{h.i}}{N_{I1h.i}}$$

$$q_{S1h.i} = \frac{(C+X_S)_{h.i}}{N_{S1h.i}}$$

donde, se debe verificar que

$$q_{I1h.i} \leq q_{1h.i} \leq q_{S1h.i}$$

13.3 En forma análoga, para sumar por grupos de edad "i", se tiene

$$A_{hj.} = \sum_i A_{hji} \quad \text{total de la } hj\text{-celda}$$

En igual forma

$$B_{hj.} = \sum_i B_{hji}$$

$$C_{hj.} = \sum_i C_{hji}$$

$$X_{hj.} = \sum_i X_{hji}$$

$$X_{Ihj.} = \sum_i X_{I hji}$$

$$X_{Shj.} = \sum_i X_{S hji}$$

$$N_{hj.} = \sum_i N_{hji}$$

$$N_{Ihj.} = \sum_i N_{I hji}$$

$$N_{Shj.} = \sum_i N_{S hji}$$

Luego

$$q_{1hj.} = \frac{(C+X)_{hj.}}{N_{hj.}}$$

$$q_{I1hj.} = \frac{(C+X_I)_{hj.}}{N_I \text{ } hj.}$$

$$q_{S1hj.} = \frac{(C+X_S)_{hj.}}{N_S \text{ } hj.}$$

donde debe ser

$$q_{Ilhj} \leq q_{lhj} \leq q_{Slhj}$$

13.4 Con un razonamiento similar, sumando sobre las "i" y sobre las "j" se obtendrán los valores de  $A_h \dots$ ,  $B_h \dots$ ,  $C_h \dots$ ,  $X_h \dots$ ,  $X_{Ih} \dots$ ,  $X_{Sh} \dots$ ,  $N_h \dots$ ,  $N_{Ih} \dots$ ,

$$N_{Sh} \dots, q_{lh} \dots, q_{Ilh} \dots \text{ y } q_{Slh} \dots$$

donde

$$q_{Ilh} \dots \leq q_{lh} \dots \leq q_{Slh} \dots$$

13.5 Una vez obtenidos los tabulados básicos sin ponderar, que corresponden a aquellos cuadros marcados con (\*) en el punto 1, se procederá a la construcción del resto de los tabulados. Estos se obtendrán ponderando los tabulados básicos en función de la población de personas, según Censo de 1981.

Llamando con

$M_h$  Total de personas, según Censo de población de 1981, en el nivel geográfico "h"

$M = \sum_h M_h$  Total de personas en el nivel geográfico de estimación

$W_h = \frac{M_h}{M}$  ponderación en el nivel geográfico "h"

las tasas de omisión serán, respectivamente, para las celdas sexo-edad "ji",

$$q_{1..ji} = \sum_h q_{1hji} W_h$$

$$q_{I1..ji} = \sum_h q_{I1hji} W_h$$

$$q_{S1..ji} = \sum_h q_{S1hji} W_h$$

Para los totales por sexo "j"

$$q_{1..j.} = \sum_h q_{1hj.} W_h$$

$$q_{I1..j.} = \sum_h q_{I1hj.} W_h$$

$$q_{S1..j.} = \sum_h q_{S1hj.} W_h$$

Para los totales por edad "i"

$$q_{1..i} = \sum_h q_{1h.i} W_h$$

$$q_{I1..i} = \sum_h q_{I1h.i} W_h$$

$$q_{S1..i} = \sum_h q_{S1h.i} W_h$$

En todos ellos, como medida de control de cálculo, se debe verificar que

$$q_{I1..} \leq q_{1..} \leq q_{S1..}$$

13.6 Ejemplo ficticio para  $h = \overline{1;2}$  ;  $i = \overline{1;3}$ ;  $j = \overline{1;2}$

Edad	$h=1; j=1$ (Hombres)										
$i$	A	B	C	X	N	$q_1$	$P_1$	$q_2$	$P_2$	V(x)	T(x)
1	100	4	12	.4800	116.4800	.1071	.8930	.0385	.9615	.0112	.1057
2	60	2	10	.3333	72.3333	.1429	.8571	.0323	.9677	.0117	.1082
3	40	3	8	.6000	51.6000	.1667	.8333	.0698	.9302	.0392	.1979
$i=.$	200	9	30	1.4133	240.4133	.1307	.8693	.0433	.9567		

$i$	$X_I$	$X_S$	$N_I$	$N_S$	$q_I$	$q_S$
	$X - D(x)$	$X + D(x)$				
1	.3743	.5857	116.3743	116.5857	.1063	.1080
2	.2251	.4415	72.2251	72.4415	.1416	.1441
3	.4021	.7979	51.4021	51.7979	.1635	.1699
$i=.$	1.0015	1.8251	240.0015	240.8251	.1292	.1322

13.7 h = 1; j = 2 (Mujeres)

i	A	B	C	X	N	q <sub>1</sub>	P <sub>1</sub>	q <sub>2</sub>	P <sub>2</sub>	V(x)	T(x)
1	80	2	8	.2000	90.2000	.0990	.9090	.0244	.9756	.0029	.0540
2	40	1	5	.1250	46.1250	.1111	.8888	.0244	.9756	.0025	.0500
3	90	3	5	.1667	98.1667	.0526	.9474	.0323	.9677	.0013	.0355
i=.	210	6	18	.4917	234.4917	.0789	.9211	.0277	.9723		

i	<u>XI</u>	<u>XS</u>	NI	NS	q <sub>1 I</sub>	q <sub>1 S</sub>
	X - D(x)	X + D(x)				
1	.1460	.2540	90.1460	90.2540	.0904	.0915
2	.0750	.1750	46.0750	46.1750	.1101	.1121
3	.1312	.2022	98.1312	98.2022	.0523	.0530
i=.	.3522	.6312	234.3522	234.6312	.0783	.0794

13.8  $h = 1$ ;  $j = \cdot$  (Hombres + Mujeres)

$i$	A	B	C	X	N	$q_1$
1	180	6	20	.6800	206.6800	.1001
2	100	3	15	.4583	118.4583	.1305
3	130	6	13	.7667	149.7667	.0919
$i = \cdot$	410	15	48	1.9050	474.9050	.1051

26

$i$	$x_I$	$x_S$	$N_I$	$N_S$	$q_{iI}$	$q_{iS}$
1	.5203	.8397	206.5203	206.8397	.0994	.1008
2	.3001	.6165	118.3031	118.6165	.1293	.1317
3	.5333	1.0001	149.5333	150.0001	.0905	.0933
$i = \cdot$	1.3537	2.4563	474.3537	475.4563	.1040	.1061

13.9 Resumen para el Usuario para el nivel geográfico, sin ponderar,  $h = 1$

i	<u>j = 1 Hombres</u>			<u>j=2 Mujeres</u>			<u>j= . Total</u>		
	$q_i$	Li	LS	$q_i$	Li	LS	$q_i$	Li	LS
1	.1071	.1063	.1080	.0909	.0904	.0915	.1001	.0994	.1008
2	.1429	.1416	.1441	.1111	.1101	.1121	.1305	.1293	.1317
3	.1667	.1635	.1699	.0526	.0523	.0530	.0919	.0905	.0933
i=.	.1307	.1292	.1322	.0789	.0783	.0794	.1051	.1040	.1061

Supongamos que simbolizamos con  $M_h$  la población censal de  $h$ . En forma ficticia asignamos a este nivel 100,000 personas. Luego  $M_1 = 100.000$  será la ponderación correspondiente a este tabulado.

13.10 Para  $h = 2$  aceptemos el siguiente cuadro.

i	j=1 Hombres			j=2 Mujeres			J= Total		
	$q_{1j}$	$L_i$	$L_s$	$q_{1j}$	$L_i$	$L_s$	$q_{1j}$	$L_i$	$L_s$
1	.1248	.1041	.1400	.1200	.1100	.1300	.1220	.1050	.1350
2	.0821	.0700	.0980	.0800	.0700	.0900	.0810	.0700	.0940
3	.0723	.0621	.0820	.0700	.0600	.0800	.0710	.0610	.0810
i=.	.0900	.0800	.1000	.1000	.0900	.1100	.09500	.08500	.1050

Suponiendo  $M_2 = 20.000$ , en el cuadro de total del nivel 1 más el nivel 2 será

$$q_{1.ji} = \sum_h q_{1hji} w_h$$

donde  $w_h = \frac{M_h}{M}$  con  $M = \sum_h M_h$

13.11 Para el total  $h = .$  , será

i	<u>j = 1 Hombres</u>		
	$q_1$	$L_i$	$L_a$
1	.1101	.1059	.1133
2	.1328	.1297	.1364
3	.1510	.1466	.1552
$i = .$	.1239	.1210	.1268

donde

$$W_1 = \frac{100\ 000}{120\ 000} = .8333$$

$$W_2 = \frac{20\ 000}{120\ 000} = .1667$$

<u>j = 2    Mujeres</u>			<u>j = .    Total</u>		
$q_1$	$L_i$	$L_s$	$q_1$	$L_i$	$L_s$
.0958	.0937	.0979	.1038	.1003	.1065
.1059	.1034	.1084	.1222	.1194	.1254
.0555	.0536	.0575	.0884	.0856	.0912
.0824	.0803	.0845	.1034	.1008	.1059

14. Tasa de omisión de personas con respecto a la población teórica .

La población teórica está definida como la población censada más la omisión de esa población , llamando con

N	Población absoluta de personas censadas
q	Tasa de omisión de personas con respecto a la población censada
qN	Valor absoluto de la omisión de personas
N+ qN	Población teórica

La tasa de omisión de personas con respecto a la población teórica,  $q'$ , viene dada por

$$q' = \frac{qN}{N+qN}$$
$$= \frac{q}{1 + q}$$

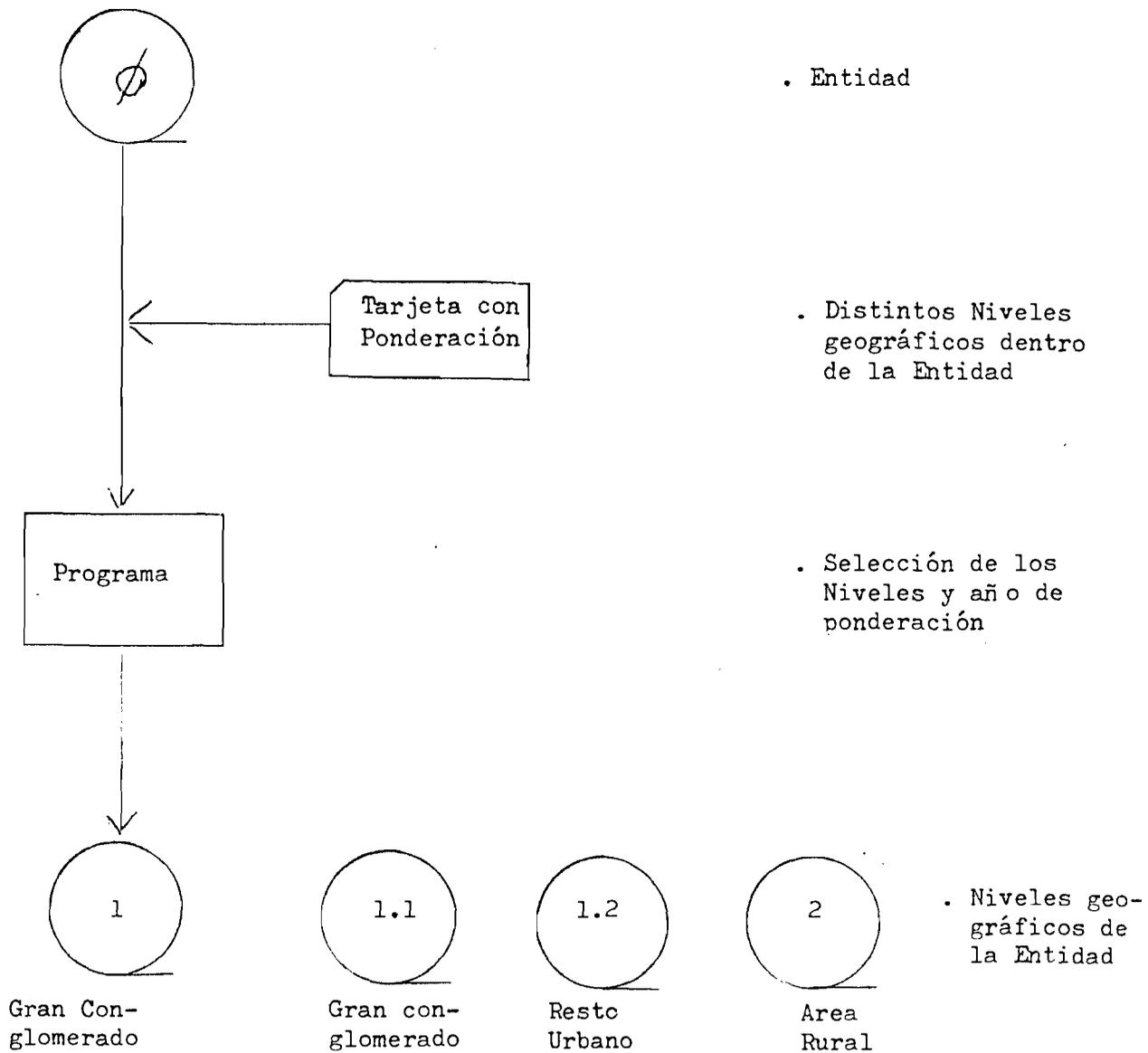
Por tanto, si se desean los cuadros en función de  $q'$  se deberán dividir todas las celdas que tienen  $q$  por  $(1+q)$ .

15. Diseño del Registro en Cinta Ø por Entidad

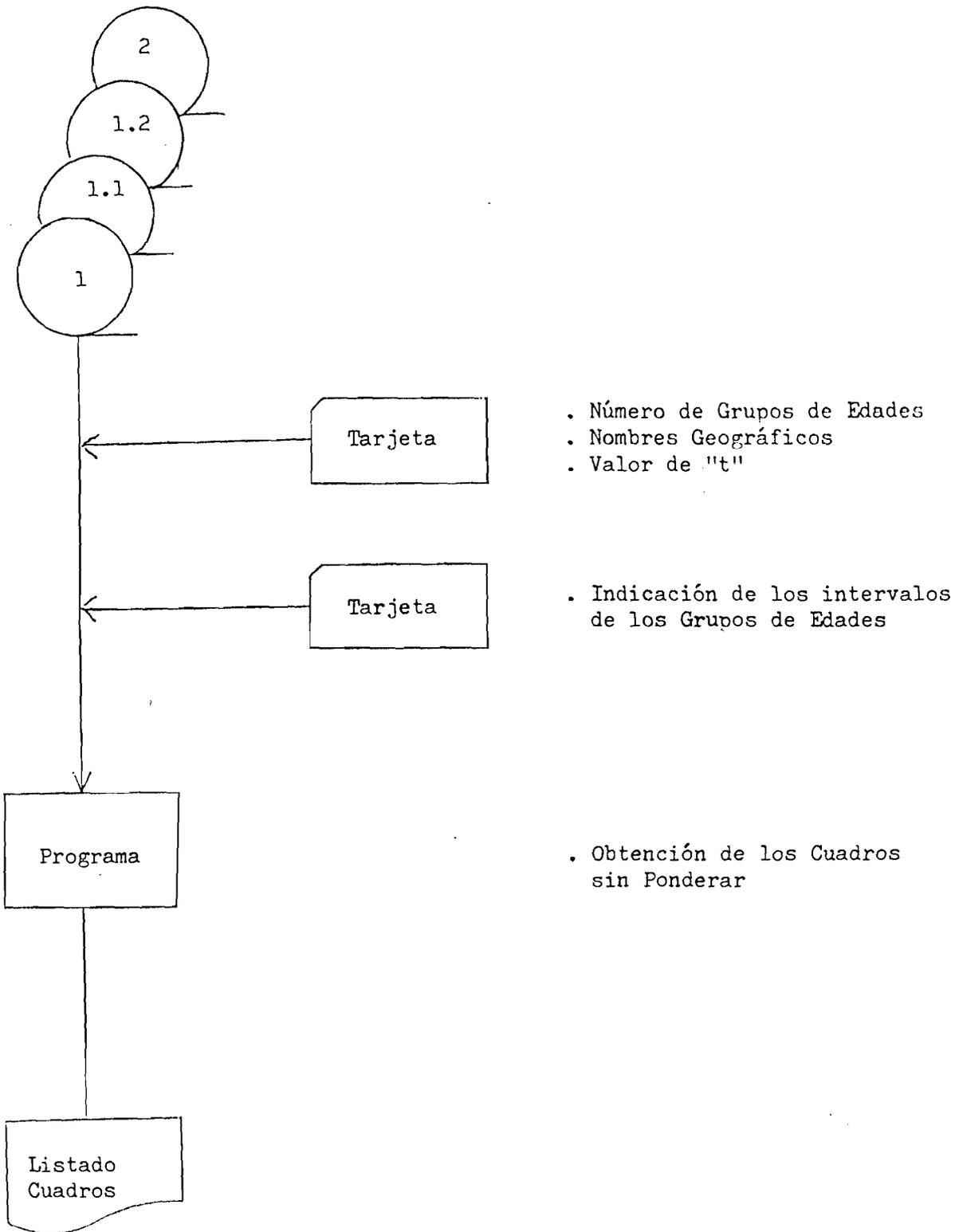
Nº Dato	De A	Tamaño	Nombre del dato
*	1 - 11	2	Identificación
1	1 - 2	2	Región
2	3 - 4	2	Area
3	5 - 5	1	Area
4	6 - 6	1	Ciudad
5	7 - 9	3	Carpeta
6	10 - 11	2	Código edad
*	12 - 23		TOTALES
7	12 - 15	4	A censo Si; muestra Si
8	16 - 19	4	B censo Si; muestra No
9	20 - 23	4	C censo No; muestra Si
*	24 - 35		HOMBRES
10	24 - 27	4	A censo Si; muestra Si
11	28 - 31	4	B censo Si; muestra No
12	32 - 35	4	C censo No; muestra Si
*	36 - 47		MUJERES
13	36 - 39	4	A censo Si; muestra Si
14	40 - 43	4	B censo Si; muestra No
15	44 - 47	4	C censo No; muestra Si
16	48 - 51	4	Control
17	52 - 54	3	Control
18	55 - 57	3	Control
19	58 - 60	3	Fila-Entidad (Para cuadros)

16. Flujograma del Sistema de Evaluación de Cobertura del Censo de Población 1981 - EVACE 81 - Programa Paramétrico

16.1 Selección de Cintas de Niveles Geográficos por Entidad Federal -  
EVACE 81 - 4

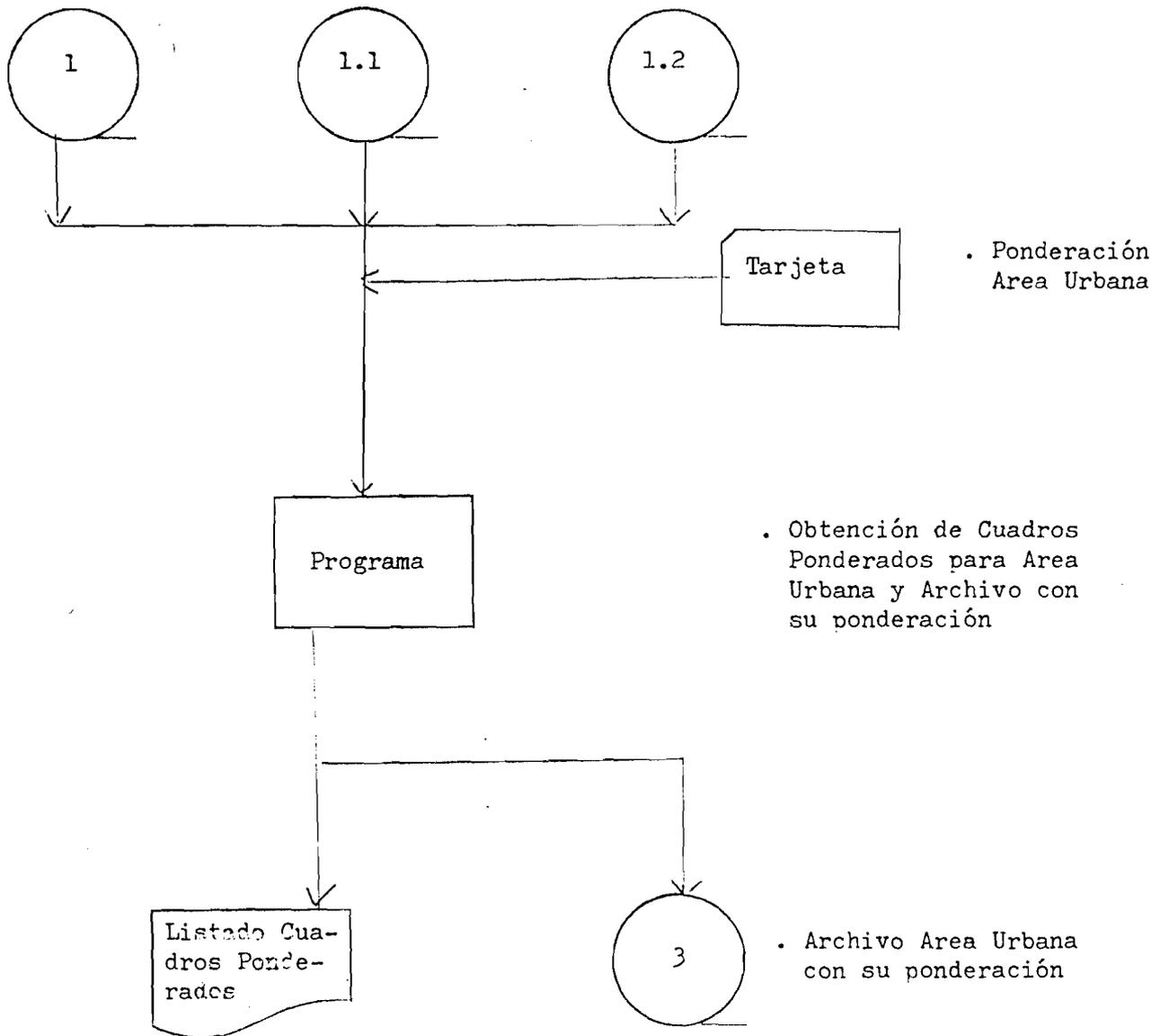


16.2 Proceso de Cuadros sin Ponderar - EVACF 81-5 -

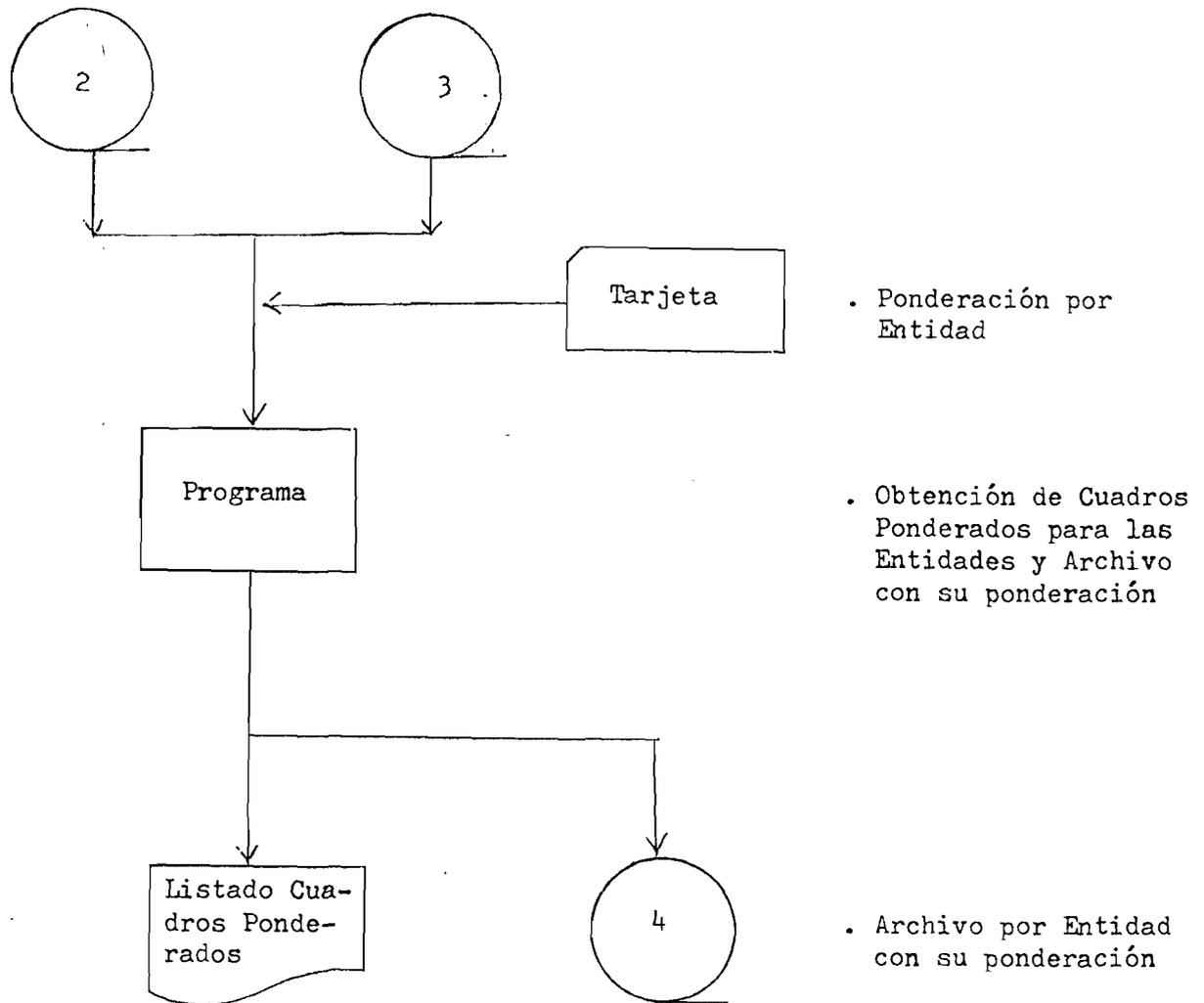


CELADE - SISTEMA GENERAL  
DOCUMENTOS  
SOBRE POBLACION  
AMERICA LATINA

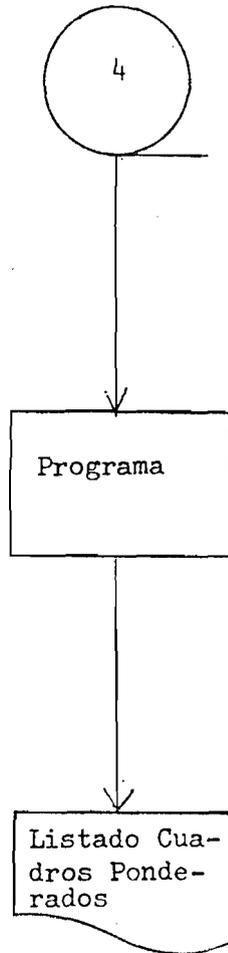
16.3 Proceso de Cuadros Ponderados para Areas Urbanas por Entidad -  
EVACE 81-6 -



16.4 Proceso de Cuadros Ponderados para las Entidades



### 16.5 Proceso de Cuadros Ponderados para el País



• Obtención de Cuadros Ponderados para el País

17. Marco Muestral. El Censo de Población y Vivienda de 1981 se encuentra registrado bajo 2 condiciones. Una, contiene todos los datos levantados en el operativo censal, denominado LO1, la otra, contiene solamente los datos considerados básicos y que son necesarios conocer anticipadamente, denominados MO1. La condición MO1 otorga los datos de la ubicación geográfica del hogar y los siguientes datos principales de las personas: edad, sexo, relación de parentesco con el jefe del hogar, estado civil y nacionalidad. El Marco Muestral, por tanto, constará de 2 etapas de construcción. La primera, que ya ha sido elaborada, abarca la información provista por la condición MO1, (ver informe de junio de 1983, párrafos 10 y 11). La segunda, que se estima poder finalizar a lo largo del año 1984, contendrá todos los datos medidos por el Censo.

18. La diferencia substancial entre ambos marcos radica en la mayor cantidad de estratos que podrán construirse con el Marco LO1, en especial, utilizando las variables de tipo socioeconómico que provee el Censo, lo cual, sin lugar a dudas, reducirá significativamente los costos por nivel de confiabilidad en las encuestas de hogares.

19. Los métodos que se presentan a continuación para estimar los tamaños muestrales y para la selección de unidades muestrales, están basados en el Marco MO1.

19.1 Tamaños muestrales. Se utilizaron funciones paramétricas para calcular el tamaño muestral, para un estrato geográfico "h", según distintas unidades de selección y de medición.

Parámetros dados:

$p_h$                       valor esperado o conjeturado de "p", siendo "p" la  
variable principal de estudio

- $d_h$  error absoluto de aceptación de "p"
- $t_h$  confiabilidad de que los límites  $p \pm d$  encierren al valor real de "p"
- $a_h$  tasa estimada de la PEA, población económicamente activa
- $b_h$  promedio de personas por unidad de medición de última etapa generalmente el hogar particular
- $c_h$  promedio de hogares particulares por unidad de selección generalmente la vivienda particular
- $g_h$  tasa de viviendas que se espera no podrán medirse, no-respuesta, debido a causas que no obedecen a factores inherentes a la unidad de medición.
- $l_h$  carga de trabajo diario de viviendas del entrevistador por unidad de selección de primera etapa, generalmente segmentos.
- $r_h$  coeficiente de correlación intraclase de la variable principal de estudio

$m_h$

carga de trabajo por Sector, o sea,  
Segmentos a seleccionar por Sector

19.2 La programación seguirá el orden siguiente:

Fórmula para el tamaño muestral del estrato "h". Por comodidad se han suprimido los subcriptos de los parámetros	Unidad de selección o de medición en la muestra
$n_{h1} = \frac{t^2}{j^2}$	PEA
$n_{h2} = \frac{n_{h1}}{a}$	Personas
$n_{h3} = \frac{n_{h2}}{b}$	Hogares
$n_{h4} = \frac{n_{h3}}{c}$	Viviendas
$n_{h5} = n_{h4} (1 + g)$	Viviendas aumentadas debido al factor de no-respuesta
$n_{h6} = n_{h5} (1+r)$	Viviendas aumentadas debido al hecho de usar conglomerados de "1" viviendas por unidad de selección de primera etapa

$$n_{h7} = \frac{n_{h6}}{1}$$

Segmentos o equivalentes. Redondeado a entero por exceso

$$n_{h8} = \frac{n_{h7}}{m}$$

Sectores o equivalentes. Redondeado a entero por exceso.

En el caso de no especificarse, los siguientes parámetros tomarán los valores,  $c = 1$ ,  $g = 0$ ;  $r = 0$ .

19.3 El error absoluto de "p" viene dado por "d". El error relativo de "p" simbolizado con "e", es

$$e = \frac{d}{p}$$

y el coeficiente de variación de "p", es

$$cv(p) = \frac{d}{tp}$$

Se observa que si se hace  $p = .10$ ,  $d = .01$  y  $t = 2$ , es  $e = .10$  y  $cv(p) = .05$ .

20. Programa logicial paramétrico, para la selección de muestras.

20.1 Este programa permitirá, por estrato geográfico, seleccionar a las unidades muestrales de selección, utilizando el Marco Muestral basado en el Censo de Población y Vivienda de 1981. El programa está concebido para ser utilizado en las investigaciones que se lleven a cabo a través de una encuesta de hogares. La selección podrá realizarse en 3 etapas o en 2 etapas, según

corresponda al estrato que se investigue. La probabilidad de selección de la unidad será en base a la proporción aproximada al tamaño de unidades de última etapa, ppat, que contenga la unidad.

20.2 Para fijar ideas, utilizaremos las siguientes denominaciones, para el h-estrato.

Etapa	Unidad de selección	Método de Selección
1	Sector	Probabilidad proporcional aproximado al tamaño de viviendas, ppatv
2	Segmento	ppatv
3	Vivienda	Probabilidad igual, previamente a la actualización del Segmento

En rigor, el Sector es la Unidad Primaria, el Segmento la Unidad Segunda y la Vivienda la Unidad Tercera. La probabilidad de selección puede también hacerse en función de los hogares, o de las personas, o de otra característica que observe el Marco Muestral, en vez de las viviendas.

20.3 Clasificación del Marco Muestral. El Marco entrará al Programa por estrato, clasificado por Sector, en orden creciente de Viviendas, y dentro del Sector, clasificado por Segmento en orden creciente de Viviendas.

20.4 Parámetros

$n_{gh}$

Sectores a seleccionar en h-estrato

$m_h$

Segmentos a seleccionar por Sector

$V_h$

Viviendas en el h-estrato

20.5 Calcular

$$is_h = \frac{V_h}{n_h}$$

Intervalo de selección del h-estrato para la selección de Sectores, con 6 decimales

20.6 Generar

$aa_h$

Arranque aleatorio para seleccionar los Sectores, con 6 decimales. Condición

$$1 \leq aa_h \leq is_h$$

20.7 Imprimir

" $is_h$ " y " $aa_h$ "

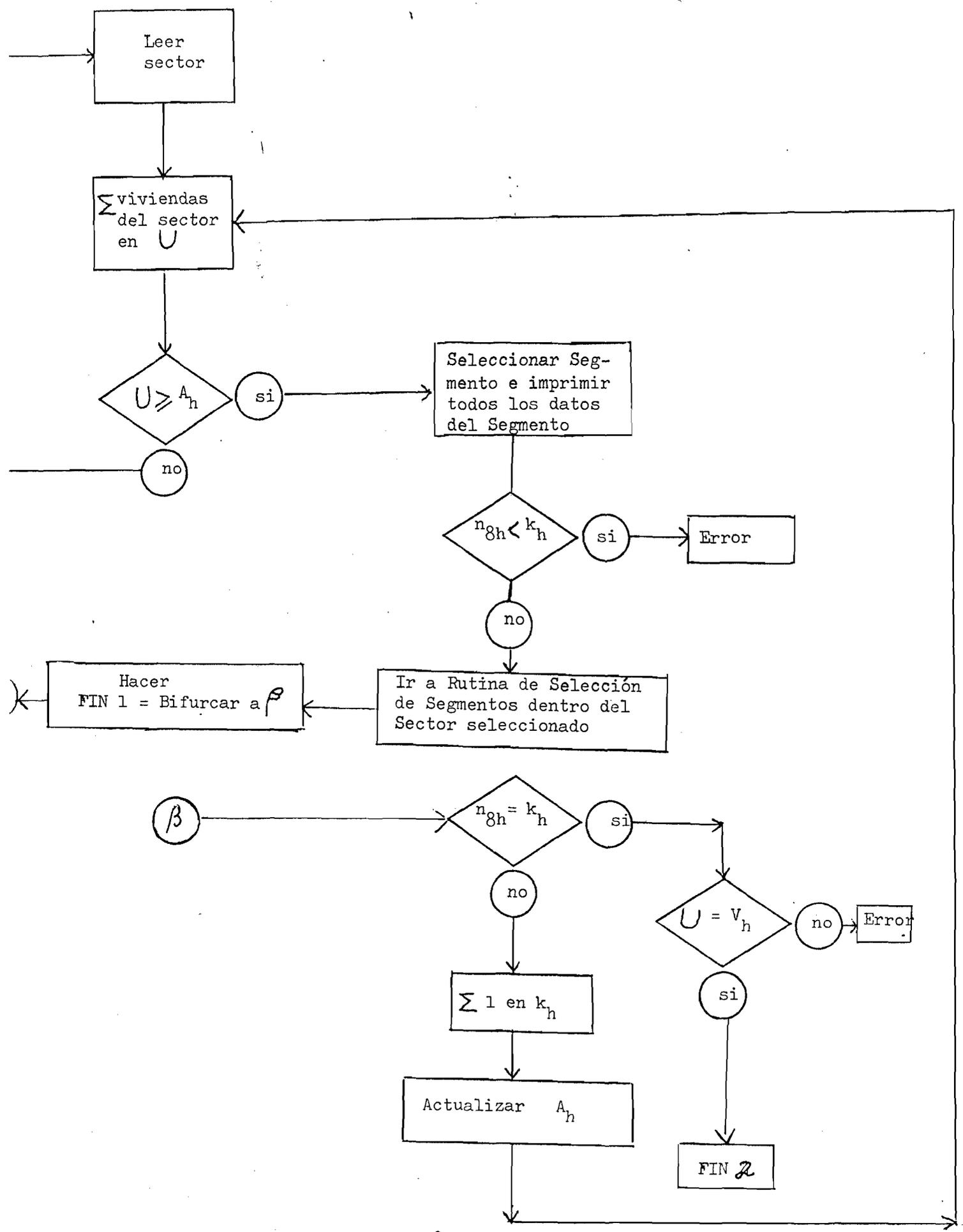
20.8 Definir

$$U = 0 \quad \text{y} \quad k_h = 1$$

20.9 Calcular

$$A_h = aa_h + (k_h - 1) is_h$$

20.10 • Flujograma para la selección de sectores



20.11

α Registrar

$V_{hj}$

Viviendas en el hj-Sector seleccionado

20.12 Calcular

$$is_{hj} = \frac{V_{hj}}{m_h}$$

Intervalo de selección del hj-Sector para seleccionar los Segmentos

20.13 Generar

$aa_{hj}$

Arranque aleatorio para seleccionar los Segmentos dentro del hj-Sector, con 6 decimales. Condición

$$1 \leq aa_{hj} \leq is_{hj}$$

20.14 Imprimir

" $is_{hj}$ " y " $aa_{hj}$ "

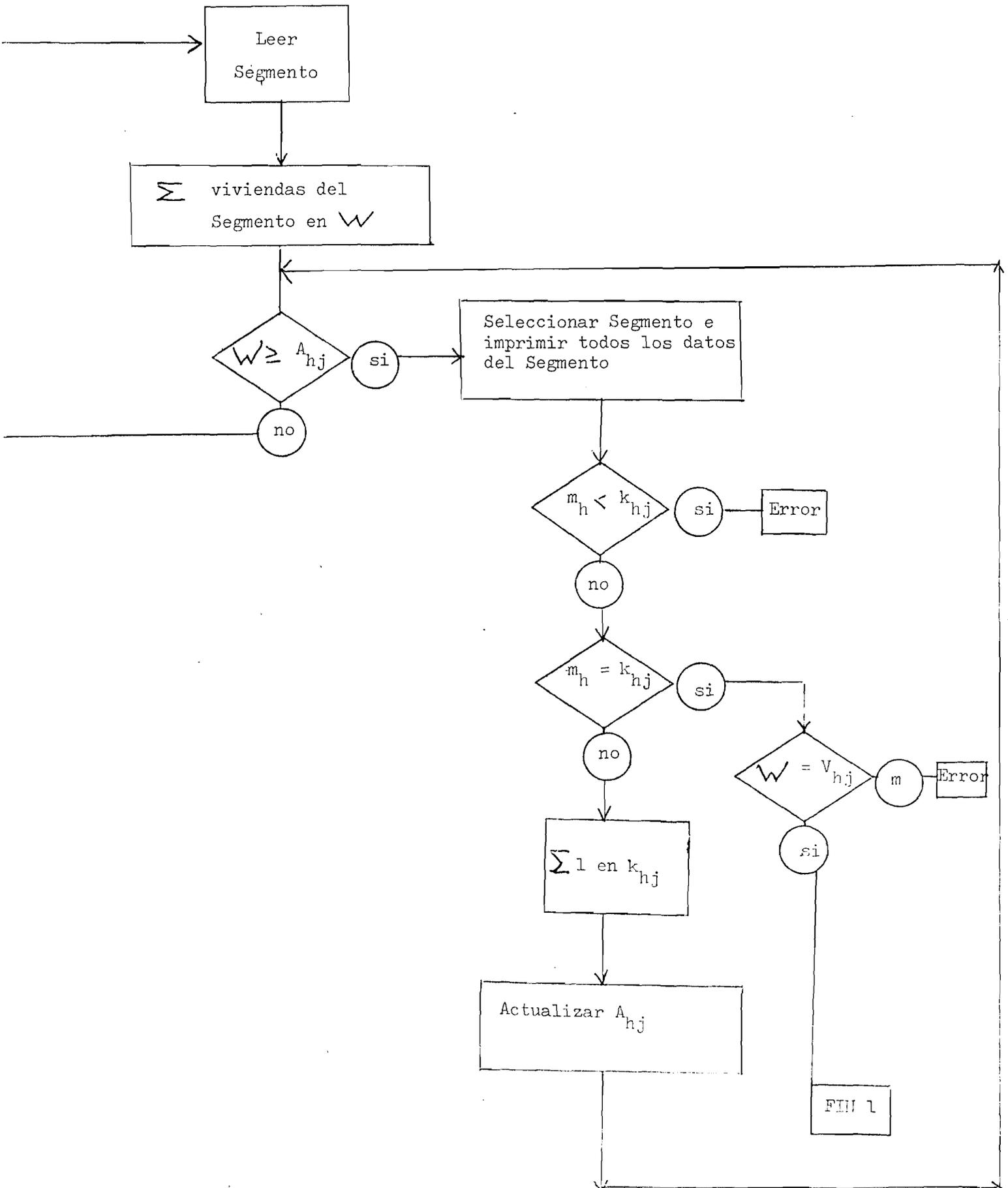
20.15 Definir

$$W = 0, \text{ y } k_{hj} = 1$$

20.16 Calcular

$$A_{hj} = aa_{hj} + (k_{hj} - 1) is_{hj}$$

20.17 Flujograma para la selección de segmentos



20.18 En el caso que se desee seleccionar la muestra en 2 etapas, como podrá ser para las áreas urbanas donde el Segmento es la Unidad Primera de selección, se deberá hacer,

$$n_{8h} = n_{7h}$$

Segmentos a seleccionar en el h-estrato

$$\alpha = \beta$$

Con esta condición el programa no bifurca a y el Flujograma de Sectores pasa a ser el Flojograma de Segmentos

21. Se observa que ambos Flujogramas tienen una lógica similar. Por tanto, debe ser un objetivo de los Analistas de Sistemas, establecer un programa paramétrico, bajo determinadas condiciones, que permita la selección de unidades en un número variable de etapas.

22. La selección de las Unidades Ultimas, en nuestro caso particular, de Viviendas, se hará una vez actualizadas las Unidades de Penúltima etapa, o sea el Segmento. De esta manera, todas las Viviendas tendrán una probabilidad conocida de pertenecer a la muestra.

23. Se hace notar que la clasificación de las unidades de selección en orden creciente de viviendas no es una condición necesaria, pero sí puede ser beneficiosa, dado que la misma conforma una estratificación de las unidades por tamaño, lo cual, utilizando la selección sistemática aumenta la representatividad del estrato.