



1960

UNIVERSIDAD DE CHILE



UNIVERSIDAD
DE CHILE

D.1/1

(DISTRIBUCION RESTRINGIDA)

A 13

RESUMEN DEL DISEÑO DE LA MUESTRA
DE MYSORE

(ESTE TRABAJO SE DISTRIBUYE PARA
USO EXCLUSIVO DE BECARIOS DEL
CENTRO LATINOAMERICANO DE DEMOGRAFIA)

2210

(resumen realizado por el
Profesor Albino Bocaz.)

MUESTRA DE MYSORE

1. Objetivos de la muestra:

- a) Determinar tasas vitales (natalidad, mortalidad, nupcialidad, fecundidad) de las diferentes regiones del Estado y su relación con el nivel socio-económico, dado a que la calidad de las cifras oficiales no lo permiten, o bien, influencia sobre las tasas vitales de los programas de salud pública de la educación, industrialización y otras medidas del status social.
- b) Tendencias futuras del crecimiento demográfico.
- c) Determinación de las necesidades de empleo, inversión en bienes de producción, de viviendas, de escuelas.....
- d) Política a seguir para mantener, dentro de límites controlables, el crecimiento demográfico.

A pesar de que una encuesta a base de muestras permite obtener información muy detallada en diversos rubros, solamente se incluyeron en la encuesta aquellas características más fundamentales para programar la política a seguir que era imposible deducir de las cifras censales.

En India no se conocía -antes de la Muestra Nacional- ninguna medida exacta de la tasa de natalidad, de la de mortalidad y la tasa de incremento natural.

La muestra tuvo además por objetivo, demostrar que era posible obtener cifras de esta naturaleza con un grupo de confianza cuantificable mediante encuestas dirigidas a las viviendas. De esta manera se investigó dentro de la vivienda el número de miembros que la integraban, el número de nacidos y fallecidos durante un período de 12 meses y se confrontaron estos datos con los registrados por los organismos oficiales. (ver en parte I, capítulo 4 y 6 y en parte III, capítulo)

2. Áreas cubiertas por la muestra

Debía buscarse áreas que presentaran diferentes grados y formas de desarrollo económico y social.

Esto tenía el propósito de ver como se afectan las características demográficas si el área es una ciudad industrial o una comunidad poco urbanizada o una área esencialmente rural. Además, como se modifican las características demográficas en áreas rurales relativamente prósperas en comparación con las de otras que se van quedando más atrás en el desarrollo económico y social.

Podría haberse solucionado el problema realizando una muestra a nivel

nacional, ya que la India en sus diversos Estados presenta toda una gama de variación de los niveles socio-económicos, pero resultando esto muy complicado por dificultades de orden administrativo entre otras, se decidió limitar la encuesta al Estado de Mysore, que se encuentra situado al SSO en la India. El Estado de Mysore tenía al Censo de 1951 una población de 9.075.000 habitantes siendo su capital la ciudad de Bangalore con cerca de 800.000 habitantes.

En el Estado podía encontrarse una suerte conveniente de áreas con diferentes niveles de desarrollo, ya sea porque en ciertas áreas se ha desarrollado la potencia hidroeléctrica lo que conduce al crecimiento de las minas de la producción de acero y otras actividades industriales; ya porque presenta áreas con desarrollos en la explotación agrícola con importantes obras de riego; ya porque también dentro del Estado se realizan programas de salud. Esto unido a que Mysore además es un Estado importante dentro de la India y a que presenta una estabilidad política y un importante progreso económico desde 1881 bajo la dirección de gobernantes capaces ("Dewan", como es su nombre)

3. Zonificación de Mysore para fines de muestreo

Para la zonificación del Estado se tuvo que tomar en cuenta la relativa homogeneidad de los individuos que viven en el área, a base de la información estadística disponible. Los criterios que se usaron fueron:

- la urbanización
- las condiciones de la agricultura
- el desarrollo de los programas de salud

Con referencia a la urbanización, las áreas identificadas en el Censo de 1951 se clasificaron en 3 grupos principales de acuerdo con el tamaño de la población en:

- ciudades: todas aquellas unidades de población que tenían 50.000 o más habitantes. Para el caso de Mysore siendo 4 estas unidades: Bangalore Corporation, Mysore City, Kolar Gold Fields y Davangere.
- pueblos: comunidades con 5.000 a 50.000 habitantes subdivididas en 3 subgrupos: 5.000-10.000; 10.000-25.000 y 25.000-50.000
- aldeas: todas aquellas comunidades con menos de 5.000. Para el Estado de Mysore, las aldeas abarcan 7 de los 9 millones de la población del Estado. Estas aldeas se agruparon en 6 subgrupos tomando en consideración: la naturaleza del suelo, la cantidad de lluvia caída, tipo de irrigación, las cosechas principales y la importancia de acuerdo al programa anti-maléfico.

Agrupando los "taluks" (unidades administrativas de población aproximadamente igual a 100.000 habitantes y mínimas en cuanto a la disposición de información estadística) se formaron 4 zonas geográficas:

- la región montañosa o Melnad, ubicada en la parte Este del Estado.
- la región "irrigada por tanques de agua", que corresponde a las llanuras central y oriental.
- la región "irrigada por canales" de las llanuras sur-central
- las áreas de secano de las llanuras norte y occidentales.

Las zonas rurales montañosas y las "irrigadas por canal" se subdividieron en dos partes respectivamente según que se estuviese desarrollando o no un programa antimalárico en gran escala. Todo lo cual nos lleva al agrupamiento de las aldeas en 6 zonas rurales.

Tabla 1.4 Población de zonas del Estado de Mysore, 1901 al 1951

Zona	Población (en miles)						Porcentaje de Incremento 1901-1951
	1901	1911	1921	1931	1941	1951	
Total, Estado de Mysore	5,546	5,814	5,988	6,566	7,338	9,075	64
Total, ciudades y pueblos	650	666	798	962	1,252	2,053	216
Ciudades de 50,000 y más en 1951	311	358	428	524	723	1,238	298
Pueblos de 25,000 a 50,000 en 1951	39	36	55	75	103	177	354
Pueblos de 10,000 a 25,000 en 1951	133	123	150	175	217	343	158
Pueblos de 5,000 a 10,000 en 1951	167	149	165	188	209	295	77
Total aldeas	4,896	5,148	5,190	5,604	6,086	7,022	43
Rural montañosa, (zonas I y II)	1,031	985	938	925	932	1,055	2
Llanura rural, áreas irrigadas por tanque (Zona III)	2,199	2,390	2,427	2,673	2,967	3,523	60
Llanura rural, áreas irrigadas por canal (Zonas IV y V)	457	482	518	556	599	648	42
Llanura rural, áreas de secano (Zona VI)	1,207	1,291	1,307	1,450	1,588	1,796	49

Tabla 1.5 Características de las zonas rurales en el Estado de Mysore

Características	Rural montañosa (Zonas I y II)	Llanuras rurales, áreas irri- gadas por tanques (Zona III)	Llanuras rurales, áreas de irri- gación por ca- nal (Zonas IV y V)	Llanuras rurales, áreas de sacano (Zona VI)
Población, 1951	1,054,655	3,522,215	648,001	1,796,210
Area (millas cuadra- das) a/	7,887	11,540	1,762	7,801
Población por milla cuadrada a/	134	305	368	230
Número de taluks	20	36	7	19
Número de aldeas	3,536	8,851	1,138	2,799
Población por aldea	298	398	569	642
Natural de la región	Cerros	Llanuras	Llanuras	Llanuras
Caídas de agua (pulgadas por año)	35-148	21-34	26-29	18-26
Tipo de irrigación	Lluvia	Tanque	Canal	Ninguna
Principal cosecha de productos alimenticios	Arroz	Ragi	Arroz	Jola
Principal cosecha de otros productos	Araca, café	Coco, mani	Caña de azúcar	Algodón, maní

a/ Las cantidades son aproximadas, ya que las áreas de las ciudades (y de Ciudad de Davangere, pero no Bangalore, Mysore, o Kolar Gold Fields Cities), localizadas en cada zona no están incluidas en los totales de área de las zonas. Las cantidades en áreas están basadas en las estimaciones del Departamento de Investig. de Mysore las que difieren muy poco de las estimaciones de los Estudios de India.

4.- Selección de las zonas que cubriría la muestra

No era posible debido a las limitaciones de recursos económicos y por tratarse de un "proyecto piloto" extender el área bajo estudio a todo el Estado de Mysore. Se decidió limitar el estudio a la ciudad de Bangalore, las ciudades de 10.000 a 25.000 habitantes y las aldeas de las zonas rurales montañosas y las de las áreas "irrigadas por tanques", ésta última conteniendo la mitad de la población de las aldeas.

5. Estratificación en la ciudad de Bangalore

Se dividió la ciudad en 5 estratos tomando en consideración la composición religiosa y el alfabetismo de la población. El marco lo constituyó la información sobre alfabetismo recogida en el Censo de 1951, que se encontraba disponible en el momento de diseñar la muestra y la información sobre religión obtenida del Censo de 1941 ya que hasta ese momento no se disponía de la del Censo de 1951. Los 5 estratos fueron los siguientes:

- Estrato 1: más de 35% de musulmanes en 1941
- " 2: más de 35% de cristianos en 1941
- " 3: más de 35% de Caste Hindú en 1941
- " 4: otras áreas indués, con alfabetismo en hombres de 60% o más
- " 5: otras áreas indués, con alfabetismos en hombres menores de 60%

De las 50 divisiones municipales en que se encuentra dividida la ciudad, 5 quedaron incluidas en el primer estrato; 6 en 2°, 3° y 4° y las 27 restantes en el 5° estrato.

6. Tipo de encuestas que se realizaron

El trabajo en el terreno se dividió en el siguiente tipo de encuestas:

I La encuesta en las aldeas: hecha en base de una muestra de 194 aldeas tomada de las 3 zonas seleccionadas para el estudio y estratificadas según el tamaño de la aldea. La información sobre las características de estas aldeas se obtuvo entrevistando una o dos personas en cada aldea que estuvieran bien informadas aparte de consultar los registros oficiales. Las características incluidas fueron:

- a) facilidades de transportes y comunicaciones (caminos y ferrocarriles, servicios postales y telegráficos).
- b) condiciones agrícolas, incluyendo áreas cultivadas, cultivos en crecimiento, magnitud de la irrigación, tenencia de la tierra, tipo de explotación, técnicas agrícolas y venta de productos agrícolas.
- c) tipo de empleos no agrícolas (cercanía de los establecimientos industriales).
- d) servicios sociales y de entronamiento, incluyendo bibliotecas, radio, cine escuelas, facilidades médicas y de salud y sanidad.

Este tipo de información junto a la obtenida de la encuesta de las viviendas permitió obtener el grado de desarrollo socio-económico y la "modernización" de las comunidades rurales en 3 zonas, y estudiar su relación con las tendencias demográficas.

II La encuesta en las viviendas: hecha en base de una muestra de 10.744 viviendas tomadas al azar de las aldeas elegidas en las 3 zonas rurales, y de una selección de bloques censales en cada pueblo de 10.000 a 25.000 habitantes y en cada estrato de la ciudad de Bangalore. La información en este caso se obtuvo por entrevistas a miembros de la vivienda. Los principales objetivos de este trabajo fueron:

- a) nacimientos y muertes ocurridas en los 12 meses que precedieron a la encuesta.
- b) hijos nacidos vivos, tenidos por las madres durante su vía matrimonial y factores relacionados con esto como ser edad al casarse, incidencia de viudez y frecuencia de nuevas nupcias.
- c) características demográficas básicas de los miembros de la vivienda (sexo, edad, estado civil).
- d) características socio-económicas de la vivienda y sus miembros, incluyendo religión, casta, alfabetismo y nivel educacional, ocupación en la industria, situación ocupacional, desempleo, tipo de vivienda, alumbrado doméstico y para personas que cultivan predios agrícolas, tenencia y tamaño de las tierras cultivadas.

Los datos de esta encuesta sobre características sociales y económicas junto con las obtenidas en la encuesta de aldeas se usó para identificar grupos afectados diferentemente por los procesos de evolución socio-económica.

III La encuesta de Fecundidad y Actitudes: realizada en matrimonios jóvenes que se encontraban en las viviendas incluidas en la encuesta de viviendas en la zona rural montañosa sin los programas anti-maláricos intensivos y en 4 de los estratos de la ciudad de Bangalore. Los propósitos de esta encuesta incluíen:

- a) información detallada sobre el desarrollo del embarazo de las mujeres casadas.
- b) opiniones y actitudes con respecto al tamaño de la familia, edad al casarse, deseo de tener o no más hijos, razones para tal deseo y asuntos relacionados.
- c) conocimiento y práctica de métodos de limitación de la familia.

En la encuesta de Fecundidad y Actitudes, se obtuvo información por medio de entrevistas personales con un grupo de encuestadoras bien entrenadas que entrevistaron a las mujeres casadas, y un grupo de encuestadores bien entrenados que entrevistaron a los maridos. Las entrevistas de esta encuesta fueron las que consumieron más tiempo, por lo cual se justificó que fuese más restringida (en área) que las otras. La información se obtuvo a base de una muestra de 1403 mu-

jerres casadas y 1804 casados.

Antes de que las encuestas se hicieran por completo, se probó la efectividad de los procedimientos, primeramente en un aldea del Estado de Delhi en Marzo de 1951, y luego en algunas áreas urbanas y rurales del Estado de Mysore en Julio de 1951. La recolección de cifras comenzó en Diciembre de 1951 y se completó en Septiembre de 1952, habiéndose realizado primero las "Encuestas de Aldeas" y de "Viviendas" y posteriormente la de "Fecundidad y Actitudes".

También se realizó un estudio sobre inmigración a la ciudad de Bangalore por el Profesor Gist de la Universidad de Missouri, coordinado con el estudio de Mysore.

7. La planificación de la "Muestra de Viviendas"

En una "Muestra de Viviendas" se debe conocer:

- número de viviendas de la población en estudio
- número de viviendas a incluirse en la muestra.

La eficiencia del diseño puede mejorarse si estas decisiones se hacen a la luz de ciertos antecedentes tales como magnitud de la variación de los rubros que se considerarán, entre los diferentes grupos de población.

Para el caso de Mysore, se disponía de escasa información y por lo tanto la muestra no ha podido aprovechar estas ventajas.

El número de viviendas a entrevistar se determinó separadamente para las áreas en estudio, esto es, para cada una de las zonas rurales, para los pueblos de 10,000 a 25,000 habitantes para cada estrato de la ciudad de Bangalore.

El criterio esencial al fijar el número de viviendas fué que la encuesta permitiera estimar la tasa de natalidad en cada área con un error típico cuya magnitud estuviera comprendido entre 6 a 8% de la tasa. Con errores de muestreo de esta magnitud, se esperaba que cualquier diferencia importante entre las tasas de natalidad de 2 áreas distintas fuera estadísticamente significativa? Podría haberse usado el error tipo de cualquier otro rubro importante en la investigación, pero se esperó que la muestra sería adecuada para todos sus propósitos principales si se satisfacía el criterio establecido anteriormente para las tasas de natalidad.

El método adoptado para la selección de las viviendas fué un "muestreo en 2 etapas", según el cual se elegían primeramente ciertas localidades (aldeas en el área rural, manzanas en las áreas urbanas) dentro del área bajo estudio, para seleccionarse posteriormente una cierta proporción de viviendas dentro de cada una de

esas localidades.

El esquema puede delinarse de la siguiente manera:

- 1) Para cada una de las 3 zonas rurales:
 - a) selección al azar de una proporción constante (1 en a) de aldeas de cada uno de los 6 estratos establecidos agrupando las aldeas de acuerdo a su población, esto es, bajo 100; 100-299; 300-499; 500-999; 1000-1999; 2000-4999 habitantes.
 - b) selección al azar de una proporción constante (1 en b) de viviendas en cada aldea de la muestra.
- 2) Para los pueblos de 10.000 a 25.000 habitantes tomados como grupo:
 - a) selección al azar de una proporción constante (1 en a) de manzanas censales, en cada una de los 15 pueblos que se incluían dentro del área de muestreo.
 - b) selección al azar de una proporción constante (1 en b) de viviendas en cada manzana seleccionada.
- 3) Para cada uno de los estratos de la ciudad de Bangalore
 - a) selección al azar de una proporción (1 en a) de las manzanas censales, entre aquellas pertenecientes a un estrato dado.
 - b) selección al azar de una proporción constante (1 en b) de viviendas en cada manzana seleccionada.

El número deseado de viviendas podría haberse obtenido tomando un número bastante grande de aldeas y una fracción de submuestreo pequeña, o bien, restringiendo el número de aldeas a un cierto número y entrevistando más viviendas por aldea de la muestra.

La guía para la determinación de las fracciones de muestreo de primera y segunda etapa se hizo de manera que se obtuviera la máxima precisión para el costo dado. Es verdad que tomando más aldeas se habría obtenido una mayor precisión en las estimaciones (reducción del error de muestreo) pero habrían subido los costos, en especial el de movilización entre aldeas, el de alojamiento del equipo de terreno en cada aldea, contactos con las oficinas locales y obtención de su cooperación, y trabajo preliminar de terreno tal como prelistado de viviendas.

Tabla 2.1 Valor de las constantes de muestreo, número esperado de localidades y viviendas a encontrarse, y errores tipos de las estimaciones de las tasas de natalidad para zonas y estratos.

Zona y Estratos	Constante a	Constante b	Número de aldeas o bloques a ser encues- tados ^{a/}	Número de vivien- das esperado	Error típico expresado en porcentaje de la tasa de na- talidad espe- rada
(1)	(2)	(3)	(4)	(5)	(6)
Zona I (rural montañosa*)	32	3	46	991	8
Zona II (rural montaña- sa **)	32	3	65	1,306	7
Zona III (llanura rural)	118	3	75	1,734	6
Pueblos (10.000 a 25.000 habitantes)	6	6	66	1,499	6
Ciudad de Bangalore:					
Estrato 1					
(porcentaje de Musulmanes sobre 35)	2.3	6	37	834	7
Estrato 2					
(porcentaje de Cristia- nos sobre 35)	2.0	6	36	811	7
Estrato 3					
(porcentaje de Caste Re- conocida sobre 35)	3.4	6	39	893	7
Estrato 4					
(Otra Hindú, alfabetis- mo masculino sobre un 60%)	3.0	6	39	883	7
Estrato 5					
(Otra Hindú, alfabetis- mo masculino bajo un 60%)	12.8	6	41	942	7

* Con programas anti-maláricos

** Sin programas anti-maláricos

a/ Pequeñas diferencias entre el número de aldeas a ser encuestadas y el verdadero número encuestado (Tabla 2.2) surgieron de la aplicación de la constante "b" separadamente para cada uno de los seis estratos formados dentro de las zonas de acuerdo al tamaño de población. En la Ciudad de Bangalore la diferencia se produjo por el arreglo en el campo de trabajo de acuerdo al diseño de cuadrado latino.

8.- Selección de las aldeas de la muestra

Lo primero fué establecer el universo de aldeas, esto es, una lista de los lugares con una población de menos de 5.000 habitantes en 1951. Como la lista de Municipalidades del Censo de 1951 aún no estaba terminada cuando se diseñó la muestra, se tuvo que usar la del Censo de 1941. Esta lista se suplementó con una publicación de lista de municipalidades que tenían una población de 5.000 o más al momento del Censo de 1951, lo que permitió poner, parte de la lista de 1941, al día. Todas las municipalidades con una población igual o superior a 5.000 se removió de la lista de 1941, y las aldeas restantes se usaron como un universo provisional para con el muestreo. Este universo difería sin embargo del real. La lista contenía algunas aldeas extrañas como por ejemplo, aldeas que estaban pobladas en 1941 pero que en 1951 estaban deshabitadas y aldeas que no existían en 1941 pero que se habían formado en el período intercensal. Las primeras diferencias se removieron durante el curso de la encuesta cuando se tuvo información acerca de las aldeas deshabitadas, y en cuanto a la magnitud de la última diferencia, que permaneció sin corregirse, se discutirá posteriormente.

La lista de aldeas de 1941, modificada en la forma descrita anteriormente permitió clasificarlas en 6 estratos de acuerdo al tamaño de la población en 1941.

A fin de separar aldeas de las zonas I (rurales montañosas con programas anti-maláricos) y II (rurales montañosas, sin programas anti-maláricos) de la lista censal, se le pidió al Director de Salud Pública una lista de aldeas (o "hamlets" que podían considerarse como satélites de la aldea principal) en que el trabajo anti-malárico se estuviese realizando bajo programas que hubiesen comenzado en 1946. Un extenso número de lugares en la lista dada por el Director de Salud Pública podía identificarse fácilmente en la lista de aldeas. Una vez que se tuvo el nombre de los "hamlets" fué necesario hacer referencia al "amildar" del "taluk" para suplementar el nombre de la aldea principal a la que pertenecía. En algunos casos, el "amildar" no fué capaz de dar la información, en tales casos no podían entonces incluirse en la zona I. De este modo, en la zona I se incluyeron todas las aldeas de las áreas rurales montañosas para las cuales se tenía información de que la aldea en su totalidad o en parte (aldea principal o hamlet) había sido incluida en el programa anti-malárico. El resto de aldeas que figuraban en la lista se incluyeron en la zona II.

Se da a continuación una tabla con el número de aldeas incluidas en la muestra en los diferentes estratos de las zonas I, II y III, junto con el número de aldeas incluidas en la muestra y el número medio de personas en la muestra por aldea muestreada.

Cuando el número de aldeas a seleccionarse era menor que 2 se aumentó este número a 2 a fin de poder calcular la variancia de las estimaciones. En estos casos la proporción de viviendas incluidas se redujo también.

Tabla 2.2 Número de aldeas en las tres zonas rurales y número escogido para la "Muestra de Aldeas", según su población.

Item y zona	Tamaño de la población de la aldea						
	Todo tamaño	Bajo 100	100- 299	300- 499	500- 999	1,000- 1,999	2,000- 4,999
<u>Número total de aldeas</u>							
Zona I (rural montañosa*)	1,481	442	605	242	138	45	9
Zona II (rural montañosa**)	2,072	689	842	301	185	48	7
Zona III (llanura rural)	8,847	1,539	3,833	1,850	1,246	316	63
<u>Número de aldeas escogidas para la muestra</u>							
Zona I (rural montañosa*)	48	13	20	7	4	2	2
Zona II (rural montañosa**)	67	21 ^{a/}	26	10	6	2	2
Zona III (llanura rural)	76	13	32 ^{b/}	16	10	3	2
<u>Promedio de personas por aldea muestreada</u>							
Zona I (rural montañosa*)	106	22	70	115	245	275	546
Zona II (rural montañosa**)	106	23	81	148	240	204	584
Zona III (llanura rural)	148	31	76	166	298	558	473

* Con programas anti-maléricos

** Sin programas anti-maléricos

a/ Excluyendo dos aldeas que se deshabitaron en 1951

b/ Excluyendo una aldea que se deshabitó en 1951.

La selección de las aldeas a encuestarse dentro de cada estrato se hizo al azar excepto por una restricción. Algunas aldeas en las zonas I, II y III se habían incluido en la Muestra Nacional. Se consideró conveniente incluir todas esas aldeas, a fin de comparar los resultados de los dos estudios. El número de aldeas de este tipo se da en la siguiente tabla.

Tabla 2.3 Número de aldeas de la "Muestra Nacional" incluidas en las tres zonas rurales

Tamaño de población	Zona I (rural montañosa*)	Zona II (rural montañosa**)	Zona III (llanura rural)
Todo tamaño	3	3	16
Bajo 100	2	1	4
100-299	-	-	2
300-499	1	-	-
500-999	-	1	6
1,000-1,999	-	-	3
2,000-4,999	-	1	1

* Con programas anti-maláricos

** Sin programas anti-maláricos

Como se indicó más atrás algunas de las aldeas seleccionadas en base de los datos de Censo de 1941 se encontraban -al momento de la encuesta- deshabitadas. Este tipo de aldeas se excluyeron de cualquier consideración posterior. La lista de aldeas de 1951 se usó también para encontrar aquellas que se habían formado a partir de 1941 y la población que tenían. En las zonas I y II, en conjunto, se encontró que un total de 8.120 personas estaban viviendo en aldeas que se habían formado posteriormente a 1941. En la zona III este número fué de 9.153, de modo que un 8% de personas en las zonas I y II y un 3% en las zonas III quedaron excluidas del universo de muestreo.

Excluyendo las aldeas deshabitadas, el número de aldeas seleccionadas para la muestra fué de 48 en la zona I; 67 en la zona II y 76 en la zona III, o sea

un total de 191. En la encuesta de aldeas, sin embargo, no fué posible, debido a dificultades administrativas, obtener la información deseada para unas cuantas de estas aldeas.

De esa manera en la "Encuesta de Aldeas" se incluyó un total de 45 aldeas en la zona I, 66 en la zona II, y 75 en la zona III.

9. Selección de los bloques (o manzanas) en las áreas urbanas

La Oficina del Censo disponía de una lista de los bloques censales con particular identificación para cada uno de los 15 pueblos de 10.000 a 25.000 habitantes. En este caso, el problema básico fué la determinación del número de bloques censales a seleccionarse en un pueblo que no contenía un número múltiplo de "a". Por ejemplo, en el pueblo de Ramanagaram con 20 bloques censales, donde la constante "a" era de 6, el problema era elegir 3 o 4 bloques. Esta decisión en sí misma fué dejada al azar. El procedimiento adoptado fué elegir de una tabla de números aleatorios uno entre 1 y 6. Si el número que aparecía era 1 ó 2 se elegían 4 bloques. Si el número aleatorio era 3, 4, 5 o 6, solamente 3 bloques. Una vez que se había decidido el número de bloques a tomar dentro del pueblo, se los elegía al azar hasta completar el número pre-determinado.

Una lista de los bloques censales en cada uno de los 5 estratos en la ciudad de Bangalore permitió incluir los 40 bloques que se necesitaban de cada estrato, con el fin de poder diseñar posteriormente un esquema de "cuadrado latino" para medir la calidad del trabajo de los encuestadores. (Parte III, Cap.21)

Se da a continuación el universo de bloques censales para la ciudad de Bangalore con indicación del número incluido en la muestra y la población media por bloque que se incluiría en la muestra.

revisiones hechas por los operadores. En algunos casos, tuvieron que cambiar la clasificación original de "viviendas colectivas" a "viviendas especiales" u "otra". En otros casos, la unidad de vivienda permaneció desocupada después del 15 de Septiembre de 1950. En estos casos, las entradas fueron apuntadas como no pertenecientes al universo de muestreo. Ulteriores indagaciones en el terreno mostraron que 2 o más entradas se referían a la misma vivienda. Estos casos se presentaron cuando parientes - como por ejemplo 2 hermanos- vivían en la misma vivienda. Ya que el Censo había listado erróneamente cada hermano como un jefe de familia, o 2 hermanos que vivían separadamente al momento del Censo después de él, estaban en la misma vivienda. En tales casos, las 2 viviendas se consideraron como una, siempre que hubieran salido en la muestra las dos. En casos excepcionales, los operadores no pudieron encontrar una o más unidades de vivienda que aparecían en la lista; estas unidades se sacaron del universo de muestreo. Esto se hizo, porque de otra manera se habría introducido un sesgo. (bías)

El número de viviendas colectivas listadas en las aldeas o en los bloques censales seleccionadas para cada zona y estrato y el número seleccionado para la encuesta se da en la tabla siguiente:

(ver tabla pág. siguiente)

Tabla 2.5 Número de viviendas en las aldeas y bloques seleccionados de las diferentes zonas y estratos, números seleccionados para el estudio y número efectivamente muestreado

Zona y estrato	Viviendas colectivas				Viviendas especiales	
	Número listado	Número seleccionado	Número muestreado	Porcentaje de los seleccionados	Número listado	Número muestreado
<u>Zona I (rural montañosa*)</u>						
Todas las aldeas	4.393	1.163	1.131	99	10	6
Población de las aldeas						
Bajo 100	212	68	68	100	1	1
100- 299	995	328	326	99	-	-
300- 499	582	198	195	98	-	-
500- 999	594	196	196	100	-	-
1000-1999	759	131	131	100	6	4
2000-4999	1.251	215	215	100	3	1
<u>Zona II (rural montañosa**)</u>						
Todas las aldeas	5.615	1.567	1.561	99	13	10
Población de las aldeas						
Bajo 100	333	107	107	100	-	-
100- 299	1.565	530	528	99	2	1
300- 499	932	320	319	99	3	3
500- 999	891	295	295	100	-	-
1000-1999	591	98	96	98	2	1
2000-4999	1.303	217	216	99	6	5
<u>Zona III (llanura rural)</u>						
Todas las aldeas	7.173	2.220	2.198	99	6	4
Población de las aldeas						
Bajo 100	233	82	82	100	-	-
100- 299	1.518	514	505	98	1	-
300- 499	1.544	519	515	99	-	-
500- 999	1.718	583	576	99	3	3
1000-1999	944	318	316	99	2	1
2000-4999	1.216	204	204	100	-	-
<u>Pueblos</u>	7.578	1.273	1.257	99	21	17
<u>Ciudad de Bangalore</u>						
Todos los estratos	27.711	4.548	4.436	98	51	43
Estrato 1 (sobre un 35% de Musulmanes)	5.407	885	857	97	9	9
Estrato 2 (sobre un 35% de Cristianos)	5.596	915	891	97	19	15
Estrato 3 (sobre un 35% de Casta Reconocida)	6.299	1.030	1.004	97	3	2
Estrato 4 (otra Hindú, alfabetismo masc. sobre un 60%)	5.136	849	832	98	13	10
Estrato 5 (otra Hindú, alfabetismo masc. bajo un 60%)	5.273	869	852	98	7	7

* Con programas anti-maláricos

** Sin programas anti-maláricos

Esta tabla da también el número de viviendas seleccionadas y el que fué encuestado. El número que no podía encuestarse por una razón u otra fué en general muy pequeño y alcanzó un máximo de 3% del total de número de viviendas seleccionadas para la encuesta en cierto estrato de la ciudad de Bangalore. En esta tabla se indica también el número de viviendas especiales listadas en las aldeas seleccionadas o en los bloques y el número de los que no podían encuestarse.

12. Estimaciones de números absolutos

Tal como se han descrito anteriormente la "Muestra de Viviendas" y la "Muestra de Aldeas" consistieron esencialmente en dividir las áreas componentes de cada zona (aldeas o bloques censales) en diferentes estratos, seleccionando al azar una cierta fracción ($1/a$) de áreas en cada estrato, y luego seleccionando igualmente al azar ($1/b$) de viviendas dentro de las áreas seleccionadas en la primera etapa.

Si hubiera sido posible mantener los valores de las fracciones de 1^a y 2^a etapa estrictamente constantes dentro de las áreas, se podrían haber usado estimaciones no sesgadas a través de fórmulas muy sencillas. Así por ejemplo, el número de personas que tenían cierta característica, o el número de nacimientos y defunciones en cada área podían haberse estimado multiplicando el valor dado por la muestra por el producto de "a" y "b". De la misma manera, el número de aldeas que pertenecían a cierta categoría podía haberse estimado de la Encuesta de Aldeas multiplicando los resultados de la muestra por "a".

Ya que no fué posible mantener fracciones de muestreo ($1/a$ y $1/b$) constantes, las fórmulas que nos permiten estimar números absolutos son más complejas.

Si denotamos por

L = número de estratos en la zona

M_h = " " aldeas del estrato

m_h = " " " de la muestra incluídas en la muestra

N_{hi} = " " viviendas de la i -ésima aldea del estrato h

n_{hi} = " " " " " " " " " " h incluídas en la muestra

x_{hij} = número de personas que tienen determinada característica en la j -ésima vivienda del i -ésima aldea del estrato h .

Entonces una estimación del total X_I del número de personas que tienen cierta característica en la zona I está dada por

$$x_I = \frac{L}{I} \left[\frac{M_h}{m_h} \frac{N_{hi}}{n_{hi}} \left(\frac{n_{hi}}{I} x_{hij} \right) \right] \quad (1)$$

cuya variancia está dada por

$$\text{Var } x_I = \sum_{h=1}^L \left(\frac{M_h}{m_h}\right)^2 \left[(1-f_{1h}) m_h S_{h(bx)}^2 + \frac{m_h}{M_h} \sum_{i=1}^{M_h} N_{hi}^2 (1-f_{2hi}) \frac{S_{hix}^2}{n_{hi}} \right] \quad (2)$$

siendo $f_{1h} = \frac{m_h}{M_h}$; $f_{2hi} = \frac{n_{hi}}{N_{hi}}$

$$S_{h(bx)}^2 = \frac{1}{M_h - 1} \sum_{i=1}^{M_h} \left(\frac{N_{hi}}{M_h} x_{hij} - \frac{\sum_{i=1}^{M_h} \sum_{j=1}^{n_{hi}} x_{hij}}{M_h} \right)^2$$

$$S_{hix}^2 = \frac{1}{N_{hi} - 1} \sum_{j=1}^{N_{hi}} \left(x_{hij} - \frac{\sum_{j=1}^{N_{hi}} x_{hij}}{N_{hi}} \right)^2 \quad (3)$$

y que puede ser determinada a base de la muestra.

La fórmula anterior para la variancia puede reemplazarse por una más sencilla de cálculo, pero que no obstante introduce un sesgo de sobreestimación. Si las fracciones de muestreo de 1 etapa son pequeñas entonces la fórmula de sustitución es

$$\text{Var } x_I = \sum_{h=1}^L \left(\frac{M_h}{m_h}\right)^2 \frac{m_h}{m_h - 1} E \sum_{i=1}^{m_h} (x'_{hi} - \bar{x}'_h)^2 \quad (4)$$

siendo $x'_{hi} = \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} x_{hij}$; $\bar{x}'_h = 1/m_h \sum_{i=1}^{m_h} x'_{hi}$ (5)

Puede notarse que esta fórmula es más rápida en el sentido que hace uso de la población de todas las viviendas de la i-ésima aldea del estrato h y no de cada vivienda como aparece en la fórmula general.

El total X_I también se podría haber estimado mediante la relación

$$x'_I = \sum_{h=1}^L \frac{M_h}{m_h} f_h \sum_{i=1}^{m_h} \frac{n_{hi}}{N_{hi}} x_{hij} \quad (6)$$

siendo $f_h = \frac{m_h N_{hi}}{\sum_{h=1}^L m_h n_{hi}}$

debido al hecho que fué posible mantener los valores de la fracción 1/b casi uniforme. En base de la muestra se estimaron el total de población, el número de nacimientos y el de defunciones de acuerdo con 3 fórmulas: la primera tomando en consideración las fracciones "a" y "b" diferentes en cada etapa de muestreo; la segunda si las fracciones "a" y "b" se hubieran mantenido constantes y la tercera, que supone que las fracciones de submuestreo fueron iguales. Estos resultados se dan en la siguiente tabla:

Tabla 6.1 Ejemplos ilustrativos de la población residente y del número de nacimientos y muertes en viviendas corrientes usando diferentes fórmulas para zonas y estratos

Zona y Estrato		Población			Nacimientos			Muertes		
		Fórmula X_1	Fórmula X_1''	Fórmula X_1'	Fórmula X_1	Fórmula X_1''	Fórmula X_1'	Fórmula X_1	Fórmula X_1''	Fórmula X_1'
Zona I (Rural montañoso *)	Número	443,948	490,080	442,694	17,955	20,160	17,861	5,938	6,528	5,799
	% de X_1	100.0	110.3	99.7	100.0	112.3	99.5	100.0	109.9	97.7
Zona II (Rural montañoso**)	Número	598,419	681,408	596,577	24,916	27,360	24,707	9,423	9,600	9,065
	% de X_1	100.0	113.9	99.7	100.0	109.8	99.2	100.0	101.9	96.2
Zona III llamu- ra montañosa)	Número	3,740,997	3,933,648	3,746,109	140,365	149,034	140,719	65,666	69,384	65,799
	% de X_1	100.0	105.1	100.1	100.0	106.2	100.3	100.0	105.7	100.2
<u>Pueblos</u>	Número	223,282	223,534	223,320	8,090	8,100	8,073	2,762	2,808	2,794
	% de X_1	100.0	100.1	100.0	100.0	100.1	99.8	100.0	101.7	101.2
<u>Ciudad Bangalore</u>	Número	689,957	665,949	690,286	20,665	19,938	20,665	5,599	5,423	5,615
Total a/	% de X_1	100.0	96.5	100.0	100.0	96.5	100.0	100.0	96.9	100.3
<u>Estrato 1</u> (predominantemen- te musulmán)	Número	67,354	64,147	67,452	1,829	1,739	1,828	343	328	344
	% de X_1	100.0	95.2	100.1	100.0	95.1	99.9	100.0	95.6	100.3
<u>Estrato 2</u> (predominantemen- te cristiano)	Número	53,356	51,035	53,422	1,468	1,406	1,472	477	458	479
	% de X_1	100.0	95.6	100.1	100.0	95.8	100.3	100.0	96.0	100.4
<u>Estrato 3</u> (predominantemente Casta Reconocida)	Número	111,803	106,992	111,876	3,687	3,531	3,692	820	798	834
	% de X_1	100.0	95.7	100.1	100.0	95.8	100.1	100.0	97.3	101.7
<u>Estrato 4</u> (otra Hindú, alfabe- tismo masc. sobre 60 %)	Número	84,550	81,989	84,354	2,412	2,332	2,399	616	592	609
	% de X_1	100.0	97.0	99.8	100.0	96.7	99.5	100.0	96.1	98.9
<u>Estrato 5</u> (otra Hindú, alfabe- tismo masc. bajo 60 %)	Número	372,894	361,786	373,182	11,269	10,930	11,274	3,343	3,247	3,349
	% de X_1	100.0	97.0	100.1	100.0	97.0	100.0	100.0	97.1	100.2

* Con programas anti-maláricos.

** Sin programas anti-maláricos.

a/ Estimaciones correctas de tasas de natalidad y de mortalidad, tomando en cuenta las omisiones de nacimientos y muertes.

13. Estimación de números relativos

Estimación de razones, porcentajes y cuocientes en general, basados en totales no ponderados sufren sesgos menores, debido a que estas fracciones diferentes afectan casi de la misma manera el numerador y el denominador. Siendo las fórmulas de uso más sencillo se encontró que su aplicación en general no se desvió en más de 2% que al usar la razón de medias ponderadas, llegándose excepcionalmente en un caso, a un sesgo de 10%.

Debe hacerse notar que las estimaciones de nacimientos y defunciones que aquí se han incluido se han hecho únicamente con el propósito de ilustrar como usar los métodos estadísticos de estimación. En el capítulo 8, parte 3, del Informe sobre la Encuesta de Mysore se indican estimaciones que toman en consideración las omisiones al informar los nacimientos y las defunciones en las Cédulas de la Vivienda.

Para la mayoría de los propósitos analíticos del estudio de Mysore no se necesitan estimaciones de números absolutos. La mayoría de los rubros que interesan son en general razones y promedios tales como tasa de natalidad y de mortalidad, número medio de hijos por mujer, porcentaje de personas (o aldeas) que tienen ciertas determinadas características. De hecho, aún la magnitud de estas razones y promedios son en general menos importantes, para los propósitos del estudio, que sus variaciones entre diversas categorías de la población y entre las diferentes áreas seleccionadas. En la parte II del informe se analizan las cifras encontradas, y allí puede verse que la presentación más corriente es en la forma de tasas, porcentajes y promedios calculados directamente de las cifras tabuladas. A pesar de que estas medidas están sujetas en algunos casos a sesgos considerables, tal como pudo verse para la estimación "no ponderada" de la tasa de mortalidad en la zona II (tabla 6.2), la presentación de los resultados en esta forma se justifica por el ahorro de trabajo al hacer estimaciones más correctas.

Tabla 6.2 Ejemplo de estimaciones ^{a/} de tasas de natalidad y mortalidad por 1.000, según las fórmulas (1) y $X''_1 = ab \sum \sum x_{ijk}$ para zonas y estratos

Zona y Estrato	Nacimientos			Muertes		
	Número por 1.000		X''_1 como porcentaje de edad de	Número por 1.000		X''_1 como porcentaje de edad de
	X_1	X''_1	X_1	X_1	X''_1	X_1
Zona I (rural montañosa*)	40.4	41.1	101.7	13.3	13.3	100.0
Zona II(rural montañosa**)	41.6	40.2	96.6	15.7	14.1	89.8
Zona III(llanura rural)	37.5	37.9	101.1	17.6	17.6	100.0
Pueblos	36.2	36.2	100.0	12.4	12.6	101.6
Ciudad de Bangalore, total	30.0	29.9	99.7	8.1	8.1	100.0
Estrato 1						
(predominio de Musulmanes)	27.2	27.1	99.6	5.1	5.1	100.0
Estrato 2						
(predominio de Cristianos)	27.5	27.5	100.0	8.9	9.0	101.1
Estrato 3						
(predominio de Casta Reconocida)	33.0	33.0	100.0	7.3	7.5	102.7
Estrato 4						
(otra Hindú, alfabetismo masculino sobre un 60%)	28.5	28.4	99.6	7.3	7.2	98.6
Estrato 5						
(otra Hindú, alfabetismo masculino bajo un 60%)	30.2	30.2	100.0	9.0	9.0	100.0

* Con programas anti-maláricos

** Sin programas anti-maláricos

^{a/} Estimaciones corregidas tomando en consideración las omisiones en la información sobre nacimientos y defunciones.

14. Estimación para combinación de zonas y estratos

Para diversos propósitos es deseable obtener de las tabulaciones de la muestra, estimaciones para combinaciones de áreas, tales como el total para los 5 estratos de la ciudad de Bangalore y el total para las 3 zonas rurales.

Como se usaron fracciones de muestreo diferentes en las diferentes zonas y estratos a fin de evitar "sesgos" en las estimaciones de conjunto deben "ponderarse" apropiadamente las estimaciones de las áreas componentes.

En la tabla siguiente se dan los pesos relativos o "ponderaciones" para el cálculo de estimaciones para diversa combinación de áreas.

Tabla 6.3. Ponderaciones para el cálculo de estimaciones para diferentes combinaciones de áreas, según datos de las encuestas de viviendas, aldeas, fecundidad y actitudes.

Zona y estrato	Ponderaciones para datos de encuesta de vivienda			Ponderaciones para datos de encuesta de viviendas a/	Ponderaciones para datos de encuesta de fecundidad y actitudes
	Combinando todas las zonas	Combinando las zonas rurales	Combinando los estratos de la ciudad de Bangalore		
Total, todas las zonas	1.0000	1.0000	...	1.0000	...
Zona I (Rural montañosa *)	.0779	.09281179	...
Zona II (Rural montañosa **)	.1051	.12511660	...
Zona III (Llanura rural)	.6567	.78217161	...
Pueblos	.0392
Ciudad de Bangalore	1.0000	...	1.0000 b/
Total	.1211
Estrato 109760977
Estrato 20773
Estrato 316211535
Estrato 412251349
Estrato 554056140

* Con programa anti-malárico

** Sin programa anti-malárico

a/ Las ponderaciones fueron calculadas después de considerar que una cierta proporción de aldeas no pudo ser cubierta por la encuesta de aldeas.

b/ Se refiere sólo al total de los cuatro estratos cubiertos por las encuestas de fecundidad y actitudes.

Tabla 6.5. Errores típicos de estimaciones de población, nacimientos y muertes.

Zona y estrato	Error típico					Error típico expresado en % de estimación				
	Pobla- ción	Nacimien- tos	Muertes	Tasa de na- talidad	Tasa de mor- talidad	Pobla- ción	Nacimien- tos	Muertes	Tasa de na- talidad	Tasa de mor- talidad
Zona I (rural mon- tañosa *)	19293	1422.5	827.7	2.72	1.87	4.35	7.92	13.94	6.73	14.00
Zona II (rural mon- tañosa **)	21648	2204.0	1085.2	2.55	1.84	3.62	8.85	11.52	6.13	11.71
Zona III (llanura rural)	146882	9515.5	4915.7	2.40	1.29	3.93	6.78	7.49	6.41	7.35
Pueblos	10250	552.4	328.2	2.84	1.52	4.59	6.83	11.88	7.82	12.29
Ciudad de Bangalore										
Total	23754	1116.2	502.6	1.27	0.75	3.44	5.40	8.98	4.25	9.28
Estrato 1 (predominantemente musulmán)	3280.3	169.15	63.99	2.49	0.88	4.87	9.25	18.65	9.6	17.27
Estrato 2 (predominantemente cristiano)	3342.5	160.63	88.50	2.28	1.52	6.26	10.94	18.55	8.29	16.98
Estrato 3 (predominantemente de casta reconocida)	4738.8	349.11	141.1	2.54	1.22	4.24	9.47	17.20	7.72	16.66
Estrato 4 (otra hindú, alfabe- tismo masc. sobre un 60 %)	4214.9	256.17	99.55	2.86	1.12	4.99	10.62	16.16	10.02	15.45
Estrato 5 (otra hindú, alfabe- tismo masc. bajo un 60 por ciento)	22408	1002.1	459.2	2.07	1.31	6.01	8.89	13.73	6.84	14.61

a/ Los errores típicos se refieren a las estimaciones de población, nacimientos y muertes, y a las tasas de natalidad y de mortalidad presentadas en las Tablas 6.1 y 6.2.

* Con programa anti-maléfico.

** Sin programa anti-maléfico.

En la práctica las variancias estimadas con las fórmulas anteriores resultaron mayores para la información tabulada para la "Encuesta de Fecundidad y Actitudes" y la "Encuesta de aldeas" que las correspondientes a la "Encuesta de Viviendas". Los datos de la "Encuesta de Fecundidad y Actitudes", en particular, pueden considerarse como indicadores de orden general de sus magnitudes verdaderas y como estimaciones precisas.

Nota: El uso de la fórmula "abreviada" (4) para la variancia produce, en general, un sesgo menor que $\frac{m_h}{M_h - m_h}$ veces la variancia obtenida por la fórmula más exacta (2). En los 5 estratos^h de la ciudad de Bangalore, el sesgo producido por la fórmula abreviada es igual a

$$\frac{L}{l} \frac{m_h^2}{M_h - 1} \sigma^2 \quad (13)$$

y usando los valores

$$M_h = 71 \text{ (número de bloques censales en ese estrato)}$$
$$\sigma_h^2 = 68.101 \text{ (población en esos bloques según el Censo de 1951)}$$

para el estrato 2, se obtiene como sesgo el valor 4.904.274.

La variancia de la población estimada en el estrato 2, usando la fórmula abreviada (4) nos da 11.172.017 y la fórmula exacta (2) en cambio, 6.267.743, lo que nos indica que la estimación (4) exagera la variancia en 78 % y el error de muestreo en 33 %.

VII-69/50
VII-61/30
XI-62/30

