

*Reprinted from Estadística, Journal of the Inter American Statistical Institute, Junio 1957, N.º 55.*

**SOBRE ALGUNOS ASPECTOS DE LA ENCUESTA  
NACIONAL DE LA INDIA POR MUESTRAS \***

# SOBRE ALGUNOS ASPECTOS DE LA ENCUESTA NACIONAL DE LA INDIA POR MUESTRAS\*

P. C. MAHALANOBIS y S. B. SEN, India\*\*

## SUMMARY

In the following article, mention is made of the large-scale surveys taken previous to the establishment of the National Sample Survey (NSS) in October 1950.

The NSS is intended as a series of continuing surveys similar in nature but varying in content. The field organization consists of about 350 full-time investigators and a supervisory staff of about 100. The main body of data is collected from sample households. The preparation of the sample designs, the questionnaires, and the analysis of the data, are the responsibilities of the Indian Statistical Institute. The questionnaires cover a wide range of subjects.

The sampling for the 1952 survey was based on 1951 Population Census data and was accomplished in 3 stages. The rural areas were grouped into 240 strata, according to geographic location and population density. From these, 480 tenshils (administrative units) were selected. The second stage consisted of 960 villages in these tenshils. In the final stage, households were selected for investigation.

In the urban areas, the towns were stratified into 4 strata from which 50 towns were selected. The cities of Calcutta, Bombay, Delhi and Madras were also included. Census enumeration blocks composed the second stage of the sample from which the households were selected for investigation.

The article indicates research that has been done by the Institute relating to the question of the reference period, and the evaluation of the validity of the data. The article concludes with criticisms of the survey, and an indication of the increasing interest in the survey and in its development.

Si bien en las últimas dos décadas se organizaron, con diferentes propósitos, muchas encuestas de muestras al azar en determinadas localidades o en Estados de la India, la organización de encuestas en una escala nacional comenzó recién a partir de la segunda guerra mundial. El primer ejemplo de tal clase de investigación fué la Encuesta sobre Trabajo Agropecuario, llevada a cabo entre diciembre de 1949 y diciembre de 1950, que cubrió la totalidad de las áreas rurales del país. Fué ésta una investigación *ad hoc organizada* por el Ministerio de Trabajo del Gobierno de la India y su alcance estuvo restringido a las condiciones de los trabajadores agropecuarios. La Encuesta Nacional por Muestras (National Sample Survey—NSS) se inició en octubre de 1950 como una investigación continua a ser llevada a cabo en fases sucesivas, para reco-

\* Presentado a la XXIX Sesión del ISI en Petrópolis, Brasil, en junio de 1955; se publicará también en el *Boletín* del ISI No. 35. Traducción hecha por la Secretaría del IASI.

\*\* Director y Miembro, respectivamente, Instituto de Estadística de la India, Calcuta, India.

lectar información socio-económica esencial acerca de todas las clases de la población. En el campo de estadísticas de los cultivos, ha estado en operación desde 1952 un esquema coordinado de encuestas por muestras, cubriendo prácticamente el total de la India, para obtener las tasas de rendimiento de los principales cultivos alimenticios.

#### ALGUNAS ENCUESTAS POR MUESTRAS, EN GRAN ESCALA, REALIZADAS EN EL PASADO

Algunas de las encuestas realizadas hasta ahora han sido, aunque no en una escala nacional, bastante amplias y merecen citarse. En 1936, el Instituto de Estadística de la India (Indian Statistical Institute) realizó una encuesta en Bengala, entre los tejedores manuales de telar, para el Gobierno del Estado. Se investigaron 5,372 familias, pertenecientes a 500 aldeas muestreadas. A partir de 1939, se iniciaron en Bengala encuestas anuales por muestras, sobre la base de la extensión completa del Estado, para determinar la superficie en acres y los rendimientos de los principales cultivos. El Instituto de Estadística estuvo organizando estas encuestas hasta 1951, en que las mismas quedaron a cargo de la Oficina de Estadística del Gobierno de Bengala Occidental.

El Consejo de Investigación Agropecuaria de la India (Indian Council of Agricultural Research) inició a partir de 1943, encuestas por muestras para estimar los rendimientos de los cultivos, al nivel de la extensión completa del Estado. El Consejo suministró el plan de la muestra y la guía técnica y el trabajo de campo estuvo a cargo de funcionarios locales, pertenecientes a los departamentos de ingresos o de agricultura, de los diferentes gobiernos de los Estados. Las encuestas sobre el rendimiento de trigo, para la totalidad del Estado, se iniciaron en 1943-44 en Punjab y en Uttar Pradesh y en el mismo período se comenzó en Orissa la encuesta sobre el rendimiento de arroz de invierno. En 1944-45 se inició en la Madhya Pradesh una encuesta para estimar el rendimiento de algodón, al nivel de la extensión completa del Estado. Estas encuestas se extendieron a otros Estados en los años subsiguientes y en 1947, las encuestas por muestras para estimar los rendimientos de los principales cultivos, cubrían todos los Estados principales, exceptuando Bengala Occidental y tres de los pequeños Estados con administración centralizada. Otros de los Estados con administración centralizada y los Estados bajo la autoridad de los Príncipes, fueron excluidos. Por los años 1952-53, estaba en marcha un plan coordinado para estimar los rendimientos de los principales cultivos alimenticios en toda la Unión de la India. La naturaleza general de la organización de estas encuestas permaneció sin alteraciones, excepto una importante modificación: el personal técnico que guiaba estas operaciones fué transferido del Consejo a la NSS, en 1952.

Durante el período 1943-46, la Oficina de Trabajo del Ministerio de Trabajo realizó 28 encuestas diferentes, bajo el Esquema de Costo de Vida del Gobierno de la India, relativas a presupuesto familiar y esparcidas sobre diferentes centros industriales y mineros y sobre áreas de

plantaciones. El tamaño de las muestras para las diferentes encuestas varió entre 123 y 5.000 familias censales.

El Instituto de Estadística llevó a cabo una encuesta en julio de 1944—febrero de 1945, para determinar los efectos de la gran escasez de alimentos de 1943 en Bengala. Se investigaron 15.640 familias, diseminadas en diferentes partes del Estado. En 1946-47 el Instituto emprendió una investigación para el Gobierno del Estado, sobre las deudas rurales y las condiciones económicas de los trabajadores agropecuarios de Bengala. Se investigaron cerca de 16.000 familias.

Además de estas investigaciones, las Oficinas de Estadística de los Estados de Bombay, Uttar Pradesh y Bengala Occidental han realizado, en los últimos años, algunas encuestas por muestras para recollectar varios tipos de datos socio-económicos.

### ENCUESTA NACIONAL POR MUESTRAS

El buen resultado de la organización de estas encuestas creó una base favorable para el planeo de una investigación del tamaño y alcance de la Encuesta Nacional por Muestras. La encuesta se inició en octubre de 1950 bajo la dirección del Ministerio de Finanzas del Gobierno de la India. La organización de campo, integrada por alrededor de 350 investigadores de tiempo completo y cerca de 100 empleados supervisores, está controlada directamente por el Ministerio y la guía técnica es proporcionada por el ISI. La preparación de los planes de muestreo, de los cuestionarios y el análisis de los datos, también están bajo la responsabilidad del Instituto.

#### *Alcance de la encuesta*

El alcance de la encuesta es amplio. El cuerpo principal de datos es recolectado por medio de muestras de familias censales e incluye características tales como edad, sexo, situación económica y ocupación de los miembros individuales de la familia censal; cantidades y valores de los diferentes artículos de consumo y gastos de la familia censal; tamaño indole de la tenencia de la tierra poseída, renta e impuestos pagados; procedencia de otros bienes tales como edificios, herramientas e implementos, costo de mantenimiento y reparaciones y valor de mejoramiento y construcción; detalles sobre empresas agropecuarias y de otra naturaleza, dirigidas por la familia censal, incluyendo información sobre el valor de la producción, cantidades de los principales renglones de producción, costos bajo varios encabezamientos, cantidades o principales renglones de materias primas, valor de los bienes de capital y número de empleados, además de los trabajadores que son miembros de la familia censal.

Para las aldeas de la muestra se recolectan características que indiquen la naturaleza de las comunicaciones, el número y naturaleza de las empresas, las épocas de las cosechas, el número de acres cultivados por observación directa de una muestra de campos y las particularidades de los pesos y medidas locales. Además, semanalmente se recolectan en las

aldeas y en las manzanas urbanas de la muestra, cotizaciones de precios al por mayor y al por menor, de 32 mercaderías seleccionadas.

Además de estas características, algunas veces se usan cuestionarios *ad hoc*, para recolectar información especialmente requerida por un Ministerio o por una Agencia de Gobierno. También puede mencionarse que las encuestas anuales a base a muestras, de establecimientos manufactureros que producen en gran escala, fueron llevadas a cabo en los últimos dos años, por un grupo especial de investigadores de la NSS.

### *Diseño*

La NSS ha sido planeada como una serie de investigaciones continuas que se realizarán en fases sucesivas. Hasta ahora se han completado cinco fases de la investigación y se espera completar la sexta en septiembre de 1953. La naturaleza de las diferentes fases será similar, aunque la cantidad de características a ser recolectadas, puede variar algo entre las fases. Debe notarse, sin embargo, que entre las primeras fases, que pueden considerarse como el período formativo de la investigación, se han verificado cambios significativos. No había sido posible cubrir la totalidad del país en un solo paso y de ahí que las primeras dos fases de la encuesta estuvieran confinadas solamente a las áreas rurales. La cobertura se extendió a las áreas urbanas a partir de la tercera.

El diseño de la encuesta también cambió entre las fases. El material básico para la preparación del plan de muestreo en las tres primeras, se refirió al Censo de Población de 1941. No se disponía de registros completos y el país experimentó cambios territoriales considerables a partir de ese año. Fué necesaria, por tanto, una cierta cantidad de improvisación para construir el marco de la muestra. En las primeras tres fases, el muestreo se hizo en dos etapas. Las 560.000 aldeas de la India fueron agrupadas en 256 estratos, de los cuales se seleccionó una muestra de aldeas. En la segunda etapa, a partir de estas aldeas se seleccionaron muestras de familias censales, para una investigación detallada.

Hacia fines de 1951, fué posible compilar una parte del material del Censo de Población de 1951 y se preparó un nuevo plan, sobre la base de los nuevos datos, para la cuarta fase de la encuesta, que se inició en abril de 1952. Para las áreas rurales, el muestreo fué ejecutado en tres etapas. Los 2,500 tehsils (áreas administrativas de aproximadamente 500 millas cuadradas) del país, fueron agrupados en 240 estratos, de acuerdo con la ubicación geográfica y con la densidad de población y de los mismos se seleccionaron 480. En la segunda etapa, se eligieron 960 aldeas de estos tehsils y en la etapa final, dentro de estas aldeas se seleccionó una cierta cantidad de familias censales, para investigación. El planeo de las dos fases siguientes, es decir, de la quinta y la sexta, también fué igual al de la cuarta fase.

Para las áreas urbanas, las ciudades se agruparon en cuatro estratos de acuerdo con sus tamaños y se seleccionaron 50 ciudades para la muestra. Además se incluyeron en la muestra las cuatro ciudades más grandes, es decir, Calcuta, Bombay, Delhi y Madras. En la segunda etapa se se-

leccionó dentro de ellas, una cierta cantidad de manzanas de enumeración del censo y en la etapa final se eligió para investigación, una muestra de familias censales dentro de esas manzanas. El número de aldeas y de manzanas urbanas, así como también el de familias censales dentro de aquéllas, varió en cierto grado entre las distintas fases.

#### EL PROBLEMA DEL PERÍODO DE REFERENCIA

Más significativos, desde el punto de vista de las dimensiones de las estimaciones, han sido los cambios en el período de referencia. En la primera fase, la información relativa a las varias partidas de los cuestionarios se refirió a un período de un año. Esto permitía, en la medida de lo posible, tomar información acerca de un período pasado; pero se pensó que no era deseable mantener el año como período de referencia para todas las partidas de información, ya que se podía también obtener datos sobre los cambios estacionales y sobre otros cambios a corto plazo, a través de un sistema repetido de encuestas. Por lo tanto, en la segunda fase, el período de referencia para partidas de consumo, principalmente alimentos, los cuales se compran muy frecuentemente, fué cambiado a una semana: una semana anterior a la fecha de la entrevista. Para un grupo de otros artículos que se compran con menos frecuencia, el período de referencia fué cambiado a un mes y para ropas y artículos de naturaleza comparativamente durable, el período de referencia siguió siendo un año.

Surgió el problema de si era más adecuado un período de referencia de una semana o de un mes, para determinar el volumen correcto del consumo de alimentos. Se puede argüir tanto en favor de un mes como de una semana; pero en lugar de depender de nociones *a priori*, decidimos probar los méritos por experimentación real. Con referencia a los alimentos y a otros pocos artículos, se introdujeron en la cuarta fase dos cuestionarios con períodos de referencia diferentes: una semana y un mes; pero idénticos en todo otro sentido. Las muestras asignadas a un investigador fueron divididas en dos sub-muestras, destinándose a una los cuestionarios semanales y a la otra los mensuales. En vista de la importancia de este problema, se enviaron dos investigadores a cada aldea, para examinar una muestra diferente de familias censales y obtener así datos replicados. En consecuencia, para cada aldea muestreada, obtuvimos del primer investigador cifras para un período semanal y para un período mensual y también dos clases de cifras similares, del segundo investigador.

Todos los formularios recolectados en la cuarta fase, no han sido todavía analizados, pero más abajo se dan las cifras de 5.521 familias censales muestreadas, para mostrar los resultados comparativos obtenidos con la semana y con el mes como períodos de referencia. Puede notarse que las cifras dadas en el siguiente cuadro, se han obtenido por tabulación directa, sin usar las ponderaciones apropiadas para lograr las estimaciones rigurosas, que se obtendrán en el debido momento. Como las ponderaciones de las familias censales muestreadas varían dentro de límites estrechos en las diferentes aldeas, la comparabilidad de las cifras dadas abajo sigue siendo correcta. Las col. 2 y 4 muestran los valores de consumo

por mes y por persona, con un período de referencia de una semana y de un mes respectivamente, tales como fueron obtenidos por el primer conjunto de investigadores y las col. 3 y 5 presentan las cifras similares, recogidas por el segundo conjunto de investigadores.

Se verá que para ambos conjuntos de investigadores, los valores de consumo con la semana como período de referencia, para los artículos 1 a 4, son más elevados que los valores con el mes como período de referencia. En tanto que en los valores para los artículos 5 a 7, que estaban referidos a períodos mensuales o anuales en ambos conjuntos de cuestionarios, no es perceptible una diferencia consistente. Las cifras para las cuatro columnas son del mismo orden.

CUADRO 1. VALORES COMPARATIVOS DE CONSUMO (POR MES Y POR PERSONA) EN RUPIAS, CON DOS PERIODOS DE REFERENCIA: UNA SEMANA Y UN MES. CUARTA FASE: ABRIL-OCTUBRE DE 1952

	Período de referencia: semana		Período de referencia: mes	
	936	936	936	935
Aldeas muestreadas .....	936	936	936	935
Familias censales muestreadas .....	1383	1381	1381	1376
ARTÍCULOS	Submuestra 1 s Rs.	Submuestra 2 s Rs.	Submuestra 1 M Rs.	Submuestra 1 M Rs.
(1)	(2)	(3)	(4)	(5)
1. Cereales .....	9,56	9,47	8,87	8,66
2. Otros artículos alimenticios .....	5,82	6,21	5,06	5,02
(i) Alimentos total .....	15,38	15,68	13,93	13,68
3. Hojas de betel, tabaco y bebidas alcohólicas .....	0,82	0,82	0,68	0,65
4. Combustible y luz .....	1,54	1,59	1,22	1,17
5. Artículos mensuales .....	1,97	2,34	2,34	2,21
6. Vestuario (anual) .....	1,62	1,73	1,79	1,65
7. Otros artículos anuales .....	2,39	2,42	2,42	2,19
GASTO TOTAL .....	23,72	24,58	22,38	21,55

Así, parece existir una clara indicación de que las cifras obtenidas con la semana como período de referencia, fueron más elevadas que las obtenidas con el mes como período de referencia. El problema que ahora surge es qué cifras de los dos períodos de referencia deberán ser consideradas correctas.

#### UNA INVESTIGACIÓN ESPECIAL SOBRE EL PERÍODO DE REFERENCIA

En marzo-abril de 1952, se llevó a cabo una investigación intensiva en 76 aldeas de Bengala Occidental, para suministrar alguna contestación a esta pregunta. Simultáneamente se seleccionó una muestra de 1.254

familias censales de estas aldeas y las mismas fueron divididas en dos grupos. Del primero, las características de consumo de ciertos artículos alimenticios se obtuvieron por interrogación, mientras que del segundo grupo, las características de consumo se obtuvieron por pesadas directas de los artículos, realizadas por los investigadores. Los artículos en estudio fueron: arroz limpio, legumbres, azúcar y sal. La medición fué hecha de la siguiente manera: dos enumeradores solicitaron a los jefes de familia que colocaran en algunos recipientes adecuados, las provisiones para aproximadamente diez días y les pidieron que sacaran de estas existencias, para el consumo de los siete días siguientes. Los investigadores pesaban las provisiones iniciales y regresaban una semana después para pesar el remanente. La diferencia daba la cantidad de consumo. En un grupo de familias censales, la pesada se hizo con un intervalo de 7 días, mientras que para otro grupo, la pesada se realizó dos veces en 7 días, una vez después de 3 días y nuevamente después de 4 días. Estas operaciones se repitieron durante 3 semanas. Los días iniciales de estos períodos de 7 días, fueron esparcidos entre los distintos días de la semana, para las diferentes familias.

En la muestra de familias censales en que los datos se obtuvieron por interrogación, se recolectaron, para la mitad de las familias aproximadamente, características con referencia a un período de una semana y para la otra mitad se recogieron datos sobre un mes. El cuadro que sigue, muestra las cantidades consumidas (por persona y por día) de los cuatro artículos especificados, en onzas. Los cuatro conjuntos de cifras de consumo obtenidas en la encuesta, se presentan en el cuadro siguiente. La col. 1 enumera los artículos para los cuales se obtuvieron los datos. La col. 2 presenta las cantidades de consumo (por persona y por día) en onzas, obtenidas por interrogación con una semana como período de referencia y la col. 3 muestra las cifras que tienen el mes como período de referencia. La col. 4 da las cifras obtenidas por pesadas efectuadas una vez a la semana y la col. 5 las cifras por pesadas verificadas dos veces a la semana.

CUADRO 2. CIFRAS COMPARATIVAS DE CONSUMO (POR PERSONA Y POR DIA) EN ONZAS, POR DOS METODOS DE INVESTIGACION Y POR DOS PERIODOS DE REFERENCIA, UNA SEMANA Y UN MES\*

Artículos	Por interrogación		Por medida	
	Período de referencia semana	Período de referencia mes	Una vez por semana	Dos veces por semana
(1)	(2)	(3)	(4)	(5)
1. Arroz (limpio).....	16,21	15,14	15,61	14,87
2. Legumbres.....	1,56	1,19	1,42	1,19
3. Azúcar.....	0,23	0,16	0,16	0,14
4. Sal.....	1,07	0,76	0,86	0,82

\* Los datos del cuadro son estimaciones efectuadas con base en los siguientes número de familias: 390 (col. 2), 386 (col. 3), 287 (col. 4) y 386 (col. 5).

Se verá que las cifras obtenidas por medio de pesadas una vez a la semana, son ligeramente superiores a las obtenidas por dos pesadas semanales. Cualquiera que sea la razón de esas diferencias, azar u otro motivo, el punto que debe notarse es que las cifras obtenidas por interrogación, con el mes como período de referencia, concuerdan estrechamente con las obtenidas por medio de pesadas, en tanto que las cifras de la col. 2 con una semana como período de referencia, son más elevadas. Los resultados de este experimento indican, por lo tanto, que los datos recolectados con el mes como período de referencia son más satisfactorios.

### RECUENTO DE LOS DATOS RECOLECTADOS

#### *Verificaciones de rutina*

Se está siguiendo la práctica usual de determinar el orden de precisión de las estimaciones calculando sus errores estándares o bien evaluando la confiabilidad de las cifras por comparaciones de las mismas con datos disponibles de otras fuentes y no es necesario discutir dicha práctica en este breve artículo. En cambio, puede hacerse mención al sistema de recuento usado en la NSS. En una investigación en gran escala, de la magnitud de la presente, donde los trabajadores de campo están esparcidos por todas partes de este inmenso país, o aún en el laboratorio, donde debe manejarse un enorme volumen de datos, no puede evitarse el empleo de un gran número de trabajadores primarios. Por tanto, es esencial un sistema de verificaciones y recuentos, para controlar la calidad de los datos.

En el campo, donde hay un inspector para aproximadamente cuatro investigadores, los inspectores verifican los datos independientemente, por medio de una muestra de familias censales y los funcionarios supervisores también hacen alguna verificación independiente. En el laboratorio es posible un sistema más riguroso de recuento y actualmente, el control se hace en cinco etapas, a saber: (a) a medida que los cuestionarios completados comienzan a llegar del campo, son confrontados para verificar la consistencia de los datos y las cifras fuera de lo común son señaladas y devueltas al campo para referencia. (b) Posteriormente, se verifican los diversos cálculos de los cuestionarios, tales como sumas, proporciones, etc., así como también los códigos. (c) Después que las características son perforadas en las tarjetas para la tabulación mecánica, las mismas son verificadas para detectar perforaciones incorrectas. (d) Cuando los cuadros salen de las tabuladoras, una unidad de empleados vinculados a la sección de máquinas tabuladoras, examinan las cifras. (e) Las tablas son revisadas nuevamente por los empleados a cargo de la preparación de los cuadros finales y se extrae una muestra de estimaciones y de cifras fuera de lo común que luego se confrontan con los cuestionarios originales.

#### *Submuestras interpenetrantes*

Sin embargo, nosotros no tomamos todo este examen como suficiente e introdujimos submuestreo interpenetrante en el plan de la encuesta,

primeramente para verificar la calidad del trabajo de campo y en segundo lugar, para comprobar la consistencia interna de los datos comparando los resultados de las submuestras. A partir de la segunda fase en adelante, las aldeas de la muestra en cada estrato fueron divididas en dos grupos, para hacer dos muestras independientes que constituyeran cortes cruzados de la población de la India: dos grupos de investigadores trabajaron en estos dos conjuntos de aldeas.

CUADRO 3. DISTRIBUCION DE ALDEAS O DE MANZANAS URBANAS DE LA MUESTRA Y DE FAMILIAS CENSALES DE LA MUESTRA, POR SUBMUESTRAS 1 Y 2, EN LA SEGUNDA Y TERCERA FASES DE LA NSS

Fases y áreas (1)	Muestra total		Submuestra 1		Submuestra 2	
	Aldeas o manzanas (2)	Familias censales (3)	Aldeas o manzanas (4)	Familias censales (5)	Aldeas o manzanas (6)	Familias censales (7)
<i>Segunda fase</i>						
1. Áreas rurales .....	1142	10870	570	5395	572	5475
<i>Tercera fase</i>						
2. Áreas rurales .....	905	5098	452	2527	453	2571
3. Pueblos.....	316	1180	163	609	153	571
4. Ciudades.....	174	510	87	255	87	255

CUADRO 4. TOTAL DE LOS GASTOS DE CONSUMO EN RUPIAS (POR FAMILIA CENSAL Y POR MES) PARA LAS SUBMUESTRAS 1 Y 2 OBTENIDAS EN LA SEGUNDA Y TERCERA FASES DE LA NSS

Fases y áreas (1)	Gastos totales en rupias			Diferencia en rupias (5)	Diferencia como porcentaje de col. 2 (6)
	Muestra total (2)	Submuestra 1 (3)	Submuestra 2 (4)		
<i>Segunda fase</i>					
1. Áreas rurales.....	130,3	130,2	120,9	0,3	0,23
<i>Tercera fase</i>					
2. Áreas rurales.....	124,4	122,9	121,3	- 1,4	1,13
3. Pueblos.....	149,6	153,3	143,8	9,5	6,39
4. Ciudades.....	243,4	247,9	238,0	9,8	4,02

A título de ilustración, más arriba se dan dos tablas, una de las cuales presenta la distribución de las muestras entre las dos submuestras en la segunda y tercera fases y la otra ofrece una comparación del gasto total por familia censal y por mes, para las submuestras 1 y 2, en esas dos fases.

Las diferencias entre las cifras de las dos submuestras, se presentan en la col. 5 y en la col. 6, las mismas aparecen expresadas como

porcentajes de las cifras de las muestras totales. Se presentan aquí las cifras para mostrar el grado de similitud de las cantidades de las dos submuestras. En el curso de nuestro examen de las cifras relativas a los diversos artículos; siempre que se observaban grandes diferencias entre las estimaciones de las dos submuestras, se verificaban las hojas de trabajo pertinentes y se detectaban errores con mayor frecuencia que cifras correctas. Los investigadores también se alertaban cuando el personal de campo notificaba amplias divergencias.

El diseño del submuestreo interpenetrante ha sido algunas veces criticado como ineficiente desde el punto de vista del mejoramiento, en base a que, si el sesgo en la respuesta es uniformemente sistemático, entonces ni siquiera una concordancia entre las cifras de las submuestras, asegurará la confiabilidad. Alternativamente, si el sesgo en la respuesta es al azar, entonces los desvíos se cancelarán de todos modos y el sistema de interpenetración no proveerá beneficio adicional. Además, se argumenta que el tamaño de las sub-muestras no puede hacerse suficientemente grande, con el resultado de que los márgenes de error serán tan grandes, que aún las diferencias más amplias no aparecerán como estadísticamente significativas. No hay ningún adelanto en decir que un sistema de muestreo interpenetrante por si solo no puede asegurar confiabilidad, pero nos gustaría destacar la cuestión de que tal sistema suministra un excelente control operacional sobre la calidad de los datos. Más aún, salvo las situaciones extremas, la concordancia entre los resultados de las sub-muestras, suministra una evidencia adicional de confiabilidad. En investigaciones en gran escala, el control operacional es un asunto importante y el costo adicional que implica este diseño, se encuentra recompensado por los provechos obtenidos del mismo.

#### OBSERVACIONES FINALES

##### *interés creciente*

Los resultados obtenidos a través de la NSS, ya están creando mucho interés. Se han recibido solicitudes de datos de parte de los distintos Ministerios del Gobierno de la India o de agencias gubernamentales. Hasta ahora, además del trabajo regular, la NSS condujo para el Ministerio de Rehabilitación, una Encuesta de Refugiados en Bengala Occidental e inició otra en Bombay. Ha recolectado información en un nivel nacional para la Comisión de Prensa (Ministerio de Radiodifusión e Información) y ha iniciado, para la Comisión de Planificación, investigaciones sobre desempleo en Calcuta y en varias ciudades principales de la India. Se ha hecho algún trabajo preliminar con relación a la Encuesta sobre Unidades de Explotación Agropecuaria, para el Ministerio de Alimentación y Agricultura y un Censo de Tierras para la Comisión de Planificación. Se tiene planeado un trabajo de análisis bastante extenso para la Comisión de Investigación de Impuestos (Ministerio de Finanzas) y se están preparando los planes para una Encuesta de Edificios Industriales a pedido del Ministerio de Obras, Energía y Abastecimientos.

Además, la NSS ayudó activamente en el trabajo de campo y también tomó íntegramente a su cargo la labor estadística de un Estudio de Población en Mysore, que recientemente se ha llevado a cabo como un proyecto cooperativo de las Naciones Unidas y del Ministerio de Salud del Gobierno de la India. Desde el año antepasado, la NSS también está conduciendo encuestas a base de muestreo sobre Establecimientos Manufactureros que producen en gran escala.

### *Críticas*

Algunas de las cifras obtenidas a través de la encuesta difieren con las cifras existentes publicadas por el Gobierno y por tanto, están dando lugar a controversias. Se está discutiendo no solamente la confiabilidad de algunas de las cifras de la NSS, sino también el sistema de recolección de datos. El método de muestreo al azar, sin embargo, no ha sido puesto en tela de juicio.

Ha surgido un problema con respecto a si los datos recolectados por medio de entrevistas son o no dignos de confianza. En realidad la cuestión se reduce al sesgo en la respuesta e indudablemente deberían aumentarse las precauciones para reducirlo al mínimo. Pero si se rechaza la entrevista como método, entonces la única alternativa, en la práctica, será la de recolectar apenas algunos datos, porque el método de la observación física no se presta para encuestas relativas a datos socio-económicos, cuando se toma en consideración el costo. Algunos tipos de datos, tales como estimación de la superficie en acres y rendimiento de las cosechas, pueden obtenerse por medio de la observación física y se están considerando los planes para realizar tales encuestas en un nivel nacional.

Otra objeción ha sido la de que el alcance de la encuesta es muy amplio y que cuando se hacen tantas preguntas a los jefes de familia, se agota su paciencia y por tanto se resiente la calidad de los datos. Debe destacarse que aunque se usan muchos cuestionarios en la NSS, no todos ellos son completados por la totalidad de las familias de la muestra. Por ejemplo, las familias que suministran información sobre gastos de consumo, no lo hacen sobre empresas. Para estas últimas, se entrevista una muestra distinta de familias. Así, el cuestionario para cada familia no es indebidamente largo. Es cierto, sin embargo, que los cuestionarios de la NSS están pormenorizados en gran detalle y esto se hace con vistas a reducir a un mínimo las omisiones, así como también, para ayudar a recordar a los informantes, paso a paso, acerca de todos los detalles.

La objeción puede ser vista desde otro ángulo, o sea, que el alcance de la investigación es muy amplio con relación a las habilidades de los investigadores. Probablemente, esto hubiera sido así, si la encuesta abarcara investigaciones que requieren observaciones expertas, como sucede en las investigaciones de las ciencias físicas o médicas. Pero estos campos están excluidos de la NSS. Los datos recolectados son de carácter socio-económico y todos ellos se obtienen entrevistando a los jefes de familia.

Luego está la cuestión muy importante del costo. Por ejemplo, el presupuesto anual de la NSS es del orden de 4 millones de rupias y si

solamente un Ministerio del Gobierno hubiera estado interesado en las cifras de la NSS, la información hubiera resultado cara. Pero cuando varios Ministerios y Agencias del Gobierno están obteniendo información a través de una agencia de recolección, la porción individual de los costos totales no es muy onerosa.

En conclusión, puede decirse que tomará algún tiempo más, poner esta investigación en gran escala en un plan de ejecución uniforme. La experiencia de estos pocos primeros años, mostrará qué tipo de datos pueden obtenerse con confianza y qué es lo que no puede hacerse. La NSS ha creado en la India un interés general. Es probable que los resultados de la investigación estimularán en el país varios estudios intensivos, regionales y típicos, para verificar y complementar los resultados de la NNS, ya que frecuentemente no es posible hacer estudios detallados en una escala nacional. En este proceso, se espera que pronto se verán llenados los diversos vacíos en las estadísticas esenciales del país.

\* \* \*