

COMISION ECONOMICA PARA AMERICA LATINA

SEMINARIO SOBRE LA PREPARACION Y USO DE TABULACIONES
DE CENSOS DE POBLACION Y HABITACION

Organizado por las Naciones Unidas, a través de la Comisión Económica para América Latina, la Oficina de Estadística de las Naciones Unidas el Fondo de las Naciones Unidas para Actividades en materia de Población con la colaboración del Centro Latinoamericano de Demografía

Santiago de Chile, 14 al 19 de agosto de 1972

Documento de referencia Nº 1

ALGUNAS CONSIDERACIONES SOBRE EL MUESTREO Y SU APLICACION
EN LOS CENSOS DE POBLACION Y VIVIENDA

por

Carlos Cavallini
Asesor Regional en Muestreo
para Estadísticas Demográficas
adscrito a la CEPAL

Agradecimientos

Agradezco al Dr. Gustaaf F. Loeb, ex Director de la División de Estadística de la CEPAL, el haberme sugerido escribir el presente documento con motivo del seminario sobre preparación y uso de tabulaciones de censos de población y habitación que se llevará a cabo próximamente, así como por sus consejos en la preparación del mismo.

Agradezco al Dr. Francisco Azorín Poch, actual Director de la División de Estadística de la CEPAL, la lectura del manuscrito y los comentarios que sobre él me brindó.

Carlos Cavallini
Asesor Regional en Muestreo para
Estadísticas Demográficas,
adscrito a la CEPAL

INDICE

	<u>Página</u>
ALGUNAS CONSIDERACIONES SOBRE EL MUESTREO Y SU APLICACION EN LOS CENSOS DE POBLACION Y VIVIENDA	1
A. EL MUESTREO	1
1. ¿Qué es el muestreo?	1
2., Errores muestrales	1
3. Errores no atribuibles al muestreo	3
B. EL MUESTREO Y LOS CENSOS DE POBLACION Y VIVIENDA	7
1. Sistema integrado	7
2. Aplicación del muestreo estadístico antes del empadronamiento censal	8
a) Chile	10
b) Guatemala	14
c) República Dominicana	14
3. Aplicación del muestreo estadístico durante el empadronamiento censal (muestreo simultáneo)	16
a) República Dominicana	17
4. Aplicación del muestreo con posterioridad al empadronamiento censal	18
a) Anticipación de resultados	18
b) Evaluación del censo en sus errores de cobertura y en sus errores de conte- nido	25
c) Control de calidad en la elaboración de los datos	26
d) Comparación del recorrido o campo de error medio de resultados censales entre algunos países latinoamericanos	28
e) Aplicación en América Latina: Argentina	31

	<u>Página</u>
5. Aplicación del muestreo entre períodos censales...	32
a) El censo como marco muestral	33
b) Experiencia en América Latina	34
6. Censos por muestreo	35
a) Ventajas de los censos por muestras	36
b) Desventajas de los censos por muestras	37
c) Experiencias con censos por muestreo en América Latina	37
C. VENTAJAS Y DESVENTAJAS DEL MUESTREO EN LOS CENSOS DE POBLACION	39
1. Ventajas del muestreo	39
a) Reducción de los gastos	39
b) Reducción del tiempo	40
c) Mejor calidad del trabajo	40
d) Evaluación de la eficiencia	40
2. Desventajas del muestreo	41
a) Alcance limitado de la información	41
b) Errores de muestreo	41
c) Posibilidad de demora en la obtención de los resultados definitivos de los censos	41
d) Confianza pública	42
D. CONCLUSIONES	42
Bibliografía	47

ALGUNAS CONSIDERACIONES SOBRE EL MUESTREO Y SU APLICACION
EN LOS CENSOS DE POBLACION Y VIVIENDA

A. EL MUESTREO

1. ¿Qué es el muestreo?

El muestreo estadístico es una disciplina derivada de la matemática aplicada, que tiene relación preponderante (aunque no exclusiva) con los fenómenos estocásticos. Basándose en el cálculo de probabilidades y en la teoría de los errores, emplea el método inductivo para hacer inferencias rigurosas.

Dentro del campo social ayuda mucho a explicar los fenómenos estocásticos y a formular soluciones para los problemas que en este campo se plantean.

En los censos de población y habitación el muestreo cumple una función primordial en la planificación, ejecución, análisis y evaluación; además los suplementa en los períodos intercensales y, en ocasiones, los sustituye.

2. Errores muestrales

La muestra es un subconjunto de unidades o elementos unitarios que se selecciona en forma probabilística del conjunto o población, con la esperanza de que sea representativo.

Los resultados que se obtengan de la muestra estarán influidos en parte por la variación que exista de una a otra muestra y en parte por la desviación que exista entre el parámetro que se quiere estimar y el resultado que se obtenga promediando todas las muestras posibles.

El primer error, o error de muestreo, dependerá del proceso de muestreo que se adopte, y en particular del tamaño de la muestra.

El segundo error o sesgo es difícil de evaluar. Puede resultar del método de selección de la muestra o de los estimadores que se utilicen en la computación de los resultados. En ciertos casos surgirá de errores no atribuibles al muestreo.

El sesgo que se introduce en los métodos de selección de la muestra es un componente constante de error que generalmente no decrece al extender la muestra. Entre sus causas principales cabe mencionar

a) La selección deliberada de ciertas unidades muestrales, como por ejemplo, de las que parezcan responder a la definición de "unidades típicas de población" y por lo tanto, representar mejor a dicha población;

b) La selección de unidades que dependan de ciertas características asociadas con las variables de estudio, como puede ocurrir en ocasiones con el procedimiento de selección sistemática;

c) La sustitución de ciertas unidades seleccionadas por otras, por ser de más fácil acceso o por no poder medirse la variable en estudio. Como ejemplo cabe señalar lo sucedido en Chile cuando se dieron resultados adelantados del censo de población y vivienda de 1952. La muestra para efectuar estas estimaciones se diseñó tomando en cuenta únicamente el costo y los errores muestrales y terminó con un sesgo en cuya corrección debió trabajar alrededor de dos años un experto en muestreo de la Dirección del Censo de los Estados Unidos. Los encargados de seleccionar la muestra (una selección sistemática del 2% de los hogares de todo el país) tendieron a evitar hogares numerosos y los sustituyeron por la cédula censal anterior o posterior. Este sesgo, difícil de descubrir sin medidas de control estadístico, en verdad sólo se detectó una vez seleccionada toda la muestra, hechas las perforaciones y comparados los resultados totales preliminares con los totales que previamente habían sido computados a mano.

/d) La

d) La no cobertura de todas las unidades seleccionadas. Cualquiera sea la causa por la cual no se puede obtener la información deseada sobre una unidad muestral seleccionada hace que esta unidad tenga cero probabilidades de pertenecer a la muestra.

El sesgo que se introduce en la estimación de los resultados, como en el caso de los estimadores por cociente, puede evitarse utilizando otros métodos de estimación. A veces este sesgo puede no ser importante con respecto al error de muestreo y en tales casos puede aceptarse.

El error total muestral es función del error de muestreo y del sesgo que se introduce en la selección y en la estimación. Estos errores muestrales sirven para medir el grado en que los resultados de una muestra se aproximan a los resultados de un censo de enumeración completa efectuada con métodos iguales o equivalentes.

La magnitud del error aleatorio muestral se puede calcular por los datos obtenidos en la muestra, lo cual hace del muestreo un método de precisión.

La eliminación del sesgo que se introduce en la selección de las unidades muestrales dependerá de la técnica de selección que se aplique, que a su vez dependerá de la investigación que se quiera realizar y del lugar donde se ha de aplicar. El avance teórico y experimental de estas técnicas ha aumentado considerablemente su eficiencia, y su aplicación exige conocimiento y experiencia en los métodos de muestreo y en sus aplicaciones.

3. Errores no atribuibles al muestreo

Entre los errores no atribuibles al muestreo se consideran principalmente aquéllos que resultan del error de medición de la unidad muestral y los errores que surgen a través del procesamiento de los datos. Los resultados de un censo de enumeración completa también se ven afectados por errores de esta índole.

/La identificación

La identificación de las fuentes o causas de estos errores preocupa a muchos investigadores, dado que ellos afectan a los resultados finales de las investigaciones censales o muestrales más de lo que comúnmente se piensa. Pero no es fácil detectarlas. La información que proporciona algún encuestado puede ser falsa, y a veces filtrarse a través de todos los controles hasta penetrar en la computación de los resultados finales. Se procura reducir estos errores mediante diseños de experimentos, reentrevistas, controles de calidad en el procesamiento de los datos, uso de muestras interpenetrantes, medidas especiales de supervisión, etc.

En ocasiones la necesidad urgente de contar con resultados hace que se omitan ciertos controles estadísticos. Otras veces estos controles parecen tan triviales que no se aplican, confiando en el buen juicio de los funcionarios. Por ejemplo, en ciertas oportunidades es necesario utilizar un sistema de enumeración consecutiva de unidades para luego efectuar el muestreo. Corresponde al control estadístico verificar que el último número asignado concuerde con un recuento manual independiente de las unidades. En caso de no haber concordancia, deberá establecer normas como las siguientes: "... buscar números omitidos y duplicados, asignar los números omitidos a las últimas unidades cancelando los previamente asignados a estas últimas unidades ...", etc. Estas instrucciones, que pueden parecer tontas, son necesarias, como lo ha demostrado la experiencia. El censo de población y vivienda de 1960 en la Argentina tropezó con este tipo de dificultades al no establecer controles eficientes en los distintos distritos censales encargados de compilar y remitir las cédulas censales a la oficina central de estadística. Esto hizo que mucho tiempo después de efectuado el levantamiento censal se recibieran comunicaciones inquiriendo sobre lo que debería hacerse con las cédulas censales que permanecían en los archivos. La creciente demanda de datos estadísticos abundantes y de calidad de los países latinoamericanos y lo limitado de los recursos para

/satisfacerla obliga

satisfacerla obliga a emplear cada vez más las modernas técnicas estadísticas de muestreo para suministrar datos confiables y para evaluar los datos censales.

Los errores de medición se agrupan generalmente en errores de no respuesta o errores que surgen por no haberse podido obtener información de toda o parte de la unidad muestral y en errores de respuesta o errores que ocurren cuando la información que se obtiene de la unidad muestral es imperfecta.

Los errores de no respuesta se deben generalmente a que la persona seleccionada i) ya no pertenece al ámbito geográfico de la investigación, ii) no se puede entrevistar por encontrarse muy raramente en la unidad, iii) no tiene capacidad para responder o iv) se niega a cooperar por falta de tiempo, desgano, problemas familiares, enfermedad, porque no obtendrá beneficios personales al hacerlo, etc.

En el estudio de la no respuesta se puede dividir a la población en dos estratos: uno compuesto por las unidades de las cuales se podrá obtener información que se supone buena si quedan dentro de la muestra, y otro compuesto por todas aquellas unidades de las cuales no se podrá obtener información si son seleccionadas. La muestra no proporcionará ninguna información sobre el segundo estrato, lo que no importaría mucho si las características del segundo estrato fuesen iguales a las del primero. Pero generalmente no es así. A veces la diferencia entre las unidades del primer y del segundo estrato es apreciable, por lo cual el sesgo que se puede introducir al hacer estimaciones para toda la población puede llegar a ser grande.

Entre las causas de errores de respuesta se pueden citar principalmente i) la falta de información por parte del que responde; ii) la falta de motivación para responder con precisión; iii) los problemas de memoria (por ejemplo, sobre la duración de algunos sucesos) y la pérdida de

memoria frecuente en personas de edad avanzada; iv) la información deliberadamente errónea, por suspicacia o por razones de prestigio; v) la redacción de la pregunta en el cuestionario, que puede dar lugar a diversas interpretaciones y vi) el influjo del entrevistador (su buen o mal entendimiento y su participación en preguntas con problemas de definición e interpretación puede tener importantes efectos en los resultados).

Los errores de respuesta pueden tomar las formas siguientes:

- a) Sesgo de la respuesta, o sea, aquella parte del error de respuesta que no tiende a promediarse entre los distintos entrevistadores que se podrían asignar a un área en particular. Este sesgo puede diferir considerablemente entre distintos tipos de área o entre una y otra investigación pero generalmente tiende a una misma dirección, y a veces puede ser muy grande.
- b) Varianza de la respuesta, es decir, aquella parte del error de respuesta que tiende a compensarse en un gran número de observaciones pero que en un conjunto limitado puede afectar a la exactitud de los resultados. La selección y preparación de buenos entrevistadores y supervisores y la confianza y cooperación que se obtenga de las unidades muestrales seleccionadas ayudan a reducir este tipo de error de medición. Asimismo, mediante ciertos controles de verificación y el análisis de la consistencia de los datos es posible determinar su confiabilidad.

El procesamiento de los datos se inicia con los procedimientos para la revisión completa de los formularios; el primer control se realiza sobre el terreno y luego se hace una revisión de oficina donde se corrigen manualmente las inconsistencias y se llenan las respuestas faltantes. Vienen luego las fases de codificación, perforación y verificación de las fichas, análisis de consistencia mecánica, etc. Los errores que surgen a

/través de

través de este proceso pueden influir en los resultados finales, por lo cual es conveniente efectuar controles de verificación en cada una de estas fases. Por ejemplo, en ciertos casos las fichas perforadas se verifican parcialmente, de modo que es preciso establecer controles de calidad por perforadora. La calidad del trabajo podría apreciarse, por ejemplo, introduciendo deliberados errores en el material que deberá procesarse. Una vez terminado el procesamiento, el informe sobre los errores corregidos debería mostrar los introducidos deliberadamente. Así, la inclusión de fichas con errores antes del análisis de consistencia mecánica puede dar una medida de la eficiencia del programa.

B. EL MUESTREO Y LOS CENSOS DE POBLACION Y VIVIENDA

1. Sistema integrado

Para efectuar una planificación científica se precisan datos acerca de todo el país y también de sus diversas regiones. Esos datos provienen de los censos y del muestreo, que no se excluyen mutuamente, sino que se complementan con ventaja.

Los censos de población y vivienda constituyen por su magnitud una de las tareas estadísticas de mayor envergadura y por ello requieren una gran aplicación de recursos y personal, que se justifica por sus resultados. Los datos que proporcionan son utilizados por los diversos países en la elaboración de los planes de desarrollo económico y social, como base analítica para muchos tipos de investigaciones y como marco muestral para numerosas investigaciones estadísticas basadas en el muestreo.

La aplicación de los métodos de muestreo en conexión con las operaciones censales ha abierto considerablemente el camino de la estadística moderna. El uso apropiado de los métodos muestrales aumenta las

/posibilidades de

posibilidades de los censos en lo que se refiere a sus alcances, la exactitud de sus resultados, la rapidez del trabajo, el uso racional de sus recursos, etc.

La coordinación entre los censos de población y viviendas y los métodos de muestreo estadístico desemboca en un sistema integrado de reunión de datos. La eficacia con que se establezca y se mantenga esta labor coordinada en función de los recursos disponibles redundará en una mayor eficiencia de todo el proceso censal-muestral.

Para aplicar en forma eficaz los métodos de muestreo a los censos de población y habitación es preciso tener un amplio conocimiento de la teoría del muestreo y cierta práctica en la aplicación de los mismos. Sin la teoría el resultado práctico será malo o limitado a la aplicación de unos cuantos principios básicos que suelen no permitir la elaboración de planes eficaces. A su vez, la experiencia práctica es necesaria porque la teoría no basta para tener en cuenta todos los detalles que deben considerarse para resolver adecuadamente los problemas que se plantean.

El muestreo estadístico tiene aplicación en los censos de población y vivienda a) antes del empadronamiento (ensayo censal); b) durante el empadronamiento; c) después del empadronamiento; d) entre períodos censales, y en ciertos casos e) como sustituto del censo.

2. Aplicación del muestreo estadístico antes del empadronamiento censal

Al planificar los censos surgen muchos problemas prácticos que varían según el país y su experiencia censal anterior. El muestreo permite experimentar con distintas opciones en una pequeña parte de la población antes de decidir su aplicación en toda ella, lo que da al uso del muestreo una enorme utilidad antes del empadronamiento censal.

La etapa precensal generalmente incluye lo siguiente:

/i) Comprobación

- i) Comprobación de la validez y facilidad de aplicación de los conceptos y definiciones vertidos;
- ii) Redacción y determinación del contenido de las preguntas;
- iii) Apreciación del valor relativo de los distintos sistemas de enumeración (entrevistas, correspondencia, autoempadronamiento, combinación de algunos sistemas, etc.);
- iv) Costo de las distintas fases de las operaciones censales;
- v) Decisiones sobre los diversos tipos de cuestionario censal;
- vi) Actuación de los enumeradores y determinación de la intensidad de su capacitación;
- vii) Ensayo de los distintos planes de revisión y codificación;
- viii) Propiedad de los planes de tabulación;
- ix) Cronología de las operaciones censales, y
- x) Procesamiento mecánico (adecuada dimensión del equipo mecánico de registro unitario o de programa almacenado; entrada de datos en tarjetas perforadas, formularios, etc.; programa del análisis de coherencia de los datos, que aunque puede hacerse con datos ficticios o antiguos resulta mejor combinado con datos precensales).

Los procedimientos que se empleen deben simular las condiciones del censo y pueden abarcar desde la publicidad que se piensa realizar, pasando por todas las etapas de recolección, control y procesamiento, hasta la publicación de los resultados obtenidos. La manera más efectiva de realizar este censo experimental es concentrando las operaciones en una ciudad, comuna, distrito, etc. o en zonas pequeñas en diferentes puntos del país. La actividad muestral en este caso, se traduce en ayuda para seleccionar los lugares donde se realizará la investigación.

Como las investigaciones precensales pueden resultar caras, se utilizan distintos métodos para limitar sus costos. En algunos casos

/se llevan

se llevan a cabo conjuntamente con censos especiales solicitados por algunas ciudades que corren con la mayor parte de los gastos. (Por ejemplo, en el censo de prueba efectuado en Santiago de Chile en 1950 se incluyó la recolección y compilación de datos de actualidad sobre población y vivienda en la comuna de Quinta Normal, que interesaban vivamente a algunos organismos del Gobierno). En otros casos, cuando solamente se quiere medir la calidad de las respuestas a través de preguntas nuevas o modificadas para probar la interpretación, reacción y capacidad del interrogado para contestar correctamente, se puede utilizar una muestra aleatoria simple de unidades en una determinada población.

Estos ensayos censales se consideran esenciales para la eficiente planificación de un censo general de población y vivienda. Cuando no es preciso basar la decisión final de un procedimiento en medidas cuantitativas de los datos recogidos por medio del ensayo precensal, no siempre es necesario el muestreo probabilístico. Pero cuando la decisión depende de medidas cuantitativas para comparar la eficiencia de distintos procedimientos (por ejemplo, errores de respuesta derivados de diferentes sistemas de empadronamiento), el uso del muestreo probabilístico es útil.

Las técnicas de muestreo permiten adoptar los planes censales más apropiados ante los problemas y condiciones prevalecientes en un determinado país. Y cuando existen varias opciones, se puede hallar la solución óptima con bastante precisión si los ensayos se hacen sobre la base de una muestra.

A continuación se reseñan algunas experiencias precensales en ciertos países latinoamericanos.

a) Chile

El Instituto Nacional de Estadística de Chile realizó en conexión con el Censo Nacional de Población y Vivienda de 1970 dos censos experimentales de población y vivienda, uno en la comuna de Valdivia y otro en el departamento de Arica.

i) Censo de prueba y encuesta demográfica de la comuna de Valdivia.

Se llevó a cabo el 28 de abril de 1967 con la colaboración del Servicio Nacional de Salud y del Ministerio de la Vivienda y Urbanismo y con el asesoramiento de la CEPAL y del CELADE. Se eligió la comuna de Valdivia porque había sufrido cambios estructurales en su población luego de los sismos de 1960, porque la Intendencia de Valdivia mostró vivo interés en el proyecto y por la posibilidad de contar con la colaboración de la Universidad Austral y de las oficinas públicas de esa región. La comuna de Valdivia tiene una superficie de 1 046 kilómetros cuadrados y su población, según este censo de prueba, alcanzaba aproximadamente a las 87 000 personas.

Los objetivos principales del censo de prueba fueron: adiestrar al personal en las labores censales de campo; aplicar métodos, procedimientos y formas de organización y tomar en consideración las normas internacionales para los censos de 1970 y examinar su posible aplicación en el futuro Censo Nacional. Además, para investigar la cobertura del recuento censal se elaboró una muestra que tuvo los siguientes propósitos: valorar la calidad de la información recogida; adelantar cifras y ser utilizada en una encuesta demográfica con la finalidad de evaluar la cobertura del registro de hechos vitales (nacimientos y defunciones) y la migración de la comuna de Valdivia.

En el diseño de la muestra se consideró suficiente seleccionar en forma sistemática el 8% de las zonas de empadronamiento del conjunto total de zonas de empadronamiento usadas en el censo. Es decir, se seleccionaron 85 zonas de las 1 065 zonas en que se dividió la comuna para el censo. En dichas zonas vivían unas 7 000 personas de las 87 000 que arrojó aproximadamente dicho censo como población total de facto.

Estas 85 zonas seleccionadas fueron interrogadas el 5 de mayo de 1967 y las cédulas de recenso así obtenidas se compararon con las cédulas

censales; en casos de discrepancia se realizó una tercera visita a la familia para determinar la veracidad de los datos. Esta comparación directa de las cédulas censales con las cédulas muestrales reveló que la omisión del censo del 28 de abril de 1967 se aproximó al 3%, que las respuestas correctas del censo alcanzaron al 77% y que el 23% restante de las respuestas del censo cayeron dentro del calificativo de omitidas (12.9%), incorrectas (3.0%), innecesarias (5.5%), incompletas (1.2%) y mal ubicadas (0.4%).

La muestra se utilizó también como prueba de una metodología que permitiera medir omisiones en los registros de nacimientos y defunciones de Valdivia, y como prueba para medir la migración interna. Para ello se realizaron tres ciclos de entrevistas en tres períodos distintos del año en los cuales se midieron características relativas a nacimientos y defunciones, que luego se cotejaron con los registros de hechos vitales correspondientes, y características sobre la migración.

La organización administrativa de este censo de prueba estuvo en manos de una Oficina Central, un Jefe del Censo y un Comité Censal que tuvo como principal objetivo el de asesorar en materias técnicas y administrativas sobre el censo programado.

ii) Censo de prueba realizado en el departamento de Arica. Dado el interés del Gobierno por conocer el número de extranjeros radicados en el departamento de Arica y el objetivo técnico inmediato por parte del Servicio de Estadística de adquirir experiencia respecto a los problemas regionales que podrían plantearse en la organización de un censo de carácter nacional, como el de 1970, en el departamento de Arica se realizó un ensayo censal de población y vivienda, en el que se sometió a prueba una cédula de empadronamiento que consultaba nuevas preguntas en una región con características socioeconómicas diferentes de las del resto del país.

El departamento de Arica forma parte de la provincia de Tarapacá, ubicada dentro de la zona tropical, y consta de cinco comunas: Arica, General Lagos, Putre, Belén y Codpa. Tiene una superficie de 16 457 kilómetros cuadrados y en junio de 1968 su población se estimaba en alrededor de 76 000 habitantes.

La organización de este ensayo censal estuvo a cargo del Instituto Nacional de Estadística, asesorado por técnicos nacionales e internacionales.

El empadronamiento se realizó el 15 de mayo de 1969. En algunas zonas, especialmente rurales, el levantamiento censal demandó varios días (en algunas zonas de difícil acceso demoró 10 días).

Simultáneamente se hizo uso de una muestra para medir ciertas características de la población y de las viviendas. El diseño de la muestra previó una selección sistemática del 25% de la población, tanto de hogares privados como de hogares colectivos. Sólo uno de cada cuatro hogares privados y alrededor de una de cada cuatro personas de los hogares colectivos quedaron dentro de la muestra.

Para las viviendas seleccionadas en la muestra se incluían seis preguntas adicionales sobre vivienda y para las personas seleccionadas en la muestra había 14 preguntas más sobre características demográficas y de empleo. La cédula censal de la muestra incluía en total diez preguntas para la vivienda y 29 preguntas para cada persona.

Los datos muestrales obtenidos se perforaron en tarjetas para su procesamiento mecánico, pero los resultados no han sido tabulados.

Este ensayo censal mostró que si bien la coordinación entre el censo y la muestra es factible y reduce significativamente el costo de medir ciertas características, las instrucciones impartidas no estuvieron al alcance de todos los empadronadores, especialmente en las zonas rurales.

b) Guatemala

Con motivo del Censo de Población y Vivienda de 1964, el Consejo Nacional de Estadística de Guatemala elaboró y llevó a cabo, en la etapa precensal, varias pruebas en el terreno con el objetivo principal de verificar la bondad de las boletas censales y conocer el rendimiento medio por entrevistador. Además de la participación de los técnicos nacionales en la programación de las pruebas, se contó con la asistencia de técnicos destacados por la Agencia Internacional de Desarrollo (AID) de los Estados Unidos.

Las pruebas en el terreno se realizaron especialmente en los departamentos de Guatemala y Santa Rosa y en la región occidental del país. Estas experiencias permitieron conocer los problemas prácticos que se presentan en el empadronamiento, y fueron de gran utilidad en la programación final del censo.

Como resultados positivos cabe señalar, primero que los formularios, pese a algunas dificultades en la interpretación y respuesta de ciertas preguntas, demostraron ser funcionales y segundo, que se pudo determinar el tiempo medio que emplearon empadronadores adiestrados en llenar un formulario de población y vivienda (aproximadamente 30 minutos), dato fundamental para calcular el personal necesario y la duración de la etapa de empadronamiento.

c) República Dominicana

Como ensayo del Censo Nacional de Población y Habitación de 1970, entre el 17 y el 18 de julio de 1968 se realizó un censo experimental en el distrito municipal de Sabana Grande de Palenque, con el propósito principal de capacitar personal, ensayar tipos de organización y procedimientos y probar el cuestionario censal y otros documentos. Como para aprovechar al máximo los recursos disponibles, el Gobierno estaba estudiando la posibilidad de levantar el censo agropecuario conjuntamente

con el censo de población y habitación, este censo de prueba serviría también para aquilatar la conveniencia de esta asociación.

El distrito municipal de Sabana Grande de Palenque tenía al 1º de julio de 1967 una población estimada de 6 000 habitantes, número que se consideró suficiente para capacitar al núcleo de funcionarios de la División de Censos y para cumplir con los objetivos deseados. Además, se seleccionó este distrito para el censo de prueba porque su cercanía de la capital (alrededor de 45 kilómetros) hizo posible que los funcionarios de la Oficina Nacional participaran en la organización y levantamiento sin cambiar de domicilio, y porque tiene sectores urbanos y rurales.

Se decidió que el censo cubriese todo el distrito. El proceso de recenso o verificación censal se realizó el 23 y el 24 de julio de 1968, es decir, cuatro días después del empadronamiento. Este plazo de cuatro días era lo suficientemente corto como para que el informante no olvidase la situación del hogar en el momento del censo, y además permitía organizar el proceso en el terreno. La muestra consistió en un 10% del total de las zonas de empadronamiento en que se dividió el distrito. Las informaciones obtenidas en los cuestionarios del recenso de la muestra se cotejarían con las del censo.

Esta experiencia muestral sirvió fundamentalmente para probar procedimientos. La preparación del censo de prueba de Palenque estuvo a cargo del Director General de Estadística, auxiliado por una Comisión central de la Oficina Nacional y por una comisión local del Distrito Municipal. También se contó con la asistencia técnica de las Naciones Unidas.

3. Aplicación del muestreo estadístico durante el empadronamiento censal (muestreo simultáneo)

Simultáneamente con el empadronamiento censal el muestreo puede usarse con eficacia principalmente en dos casos: para comprobar información adicional necesaria en función del tiempo, la precisión y el costo, y para sustituir al censo en ciertas zonas o estructuras de la población en las cuales es difícil o no se justifica un empadronamiento total.

En el primer caso, las modernas técnicas del muestreo estadístico confirman que no es necesario reunir información completa para obtener datos adicionales de aceptable exactitud. El problema reside en determinar cuál es la zona mínima para la cual se desean estimaciones, dado que el método muestral generalmente no permite hacer cálculos suficientemente exactos sobre zonas administrativas pequeñas. Para obtener esta información adicional se pueden emplear los mismos empadronadores del censo o se pueden contratar empadronadores especialmente preparados para obtener datos por muestreo; la decisión depende de si el factor preponderante es el costo de las operaciones o la exactitud de los resultados. En los casos en que la técnica corriente de empadronamiento no permite obtener respuestas exactas a las preguntas complementarias, será útil el empleo de empadronadores especializados.

Como referencia ilustrativa cabe mencionar que en el censo de población de 1970 los Estados Unidos formuló sólo nueve preguntas al 100% de la población, diez preguntas adicionales al 20% de ella, ocho preguntas adicionales más al 15% y siete preguntas adicionales más al 5% de la población. Por su parte, la Unión Soviética formuló once preguntas al 100% de la población y siete preguntas adicionales al 25% de la población.

El segundo caso se plantea cuando en ciertas zonas es difícil levantar un censo o éste no se justifica ante ciertas estructuras de la población, como tribus indígenas, dado lo cuantioso de los recursos que

/obligaría a

obligaría a movilizar. Entre las dificultades para levantar un censo podemos mencionar las siguientes condiciones sociales desfavorables:

- i) Población diseminada por colinas
- ii) Ausencia de medios de transporte
- iii) Gran movilidad de población
- iv) Escasez de enumeradores calificados
- v) Superstición
- vi) Miedo a operaciones censales
- vii) Analfabetismo e ignorancia
- viii) Falta de interés
- ix) Insuficiencia de ayuda por parte de los servicios administrativos
- x) Problemas políticos o laborales de la población.

Los métodos de muestreo ayudan a superar estos obstáculos porque reducen y simplifican las operaciones.

A continuación se enumeran algunas experiencias latinoamericanas en esta aplicación del muestreo estadístico.

a) República Dominicana

En el censo nacional de población y habitación llevado a cabo en enero de 1970 en todo el país, se utilizaron para el empadronamiento cédulas censales grandes y cédulas censales pequeñas. La cédula grande contenía todas las preguntas de la pequeña más algunas adicionales. La información solicitada en la cédula pequeña equivalía casi a la mitad de la solicitada en la cédula grande. La cédula pequeña contenía 23 preguntas, 10 sobre vivienda y 13 para población. La cédula grande contenía 42 preguntas, 19 sobre vivienda y 23 sobre población.

El plan de muestreo para el empadronamiento consistió en una muestra sistemática con arranque aleatorio cuya unidad muestral fue la "zona de empadronamiento". La "zona de empadronamiento" se definió como la porción

/de territorio

de territorio o área de trabajo asignada a cada empadronador. Cada una de estas zonas abarcaba unas 20 viviendas y aproximadamente 100 personas. La fracción muestral del total de "zonas de empadronamiento" fue del 50% para el área urbana y del 25% para el área rural. Las cédulas grandes se usaron exclusivamente para empadronar las "zonas de la muestra" y las cédulas pequeñas para empadronar las "zonas fuera de la muestra". Cada empadronador llenaría así un solo tipo de cédula. Para verificar si la muestra había sido seleccionada equitativamente para distintos tamaños de hogares se pensó en la conveniencia de utilizar por municipio la prueba estadística del chi-cuadrado.

Una de las razones que impulsaron a utilizar el muestreo fue la falta de funcionarios con adecuada preparación para el trabajo en el terreno. El número de empadronadores y jefes de grupo utilizado en este censo, que se realizó en dos días, se aproximó a 35 000.

4. Aplicación del muestreo con posterioridad al empadronamiento censal

El uso de métodos de muestreo después del empadronamiento censal sirve principalmente para adelantar resultados, evaluar errores de cobertura y de contenido y controlar la calidad en la elaboración de los datos.

a) Anticipación de resultados

Como la tabulación completa de todos los conceptos del programa censal lleva mucho tiempo, aunque se disponga de un equipo normal de tabulación mecánica, las necesidades urgentes de información estadística quedan insatisfechas y se menoscaba seriamente la utilidad práctica del censo. Para reducir las largas demoras provenientes sobre todo del recargo de trabajo que significa la edición, codificación, perforación y tabulación de los datos, por medio del muestreo se realiza un programa de informes preliminares, para proseguir luego con los informes finales basados en el censo total. Muestras del 25 al 10% de la población

/eliminan del

eliminan del 75 al 90% del trabajo de producción. La utilización de una muestra grande (por ejemplo del 25%) reduce significativamente los errores de muestreo; además, las estimaciones para zonas pequeñas quedan sujetas solamente a errores de muestreo moderados.

i) Muestreo simultáneo. Cuando el muestreo se emplea como parte integral del empadronamiento completo para ampliar el alcance del censo, se proporciona también una muestra para las tabulaciones preliminares del censo propiamente dicho. La crítica y codificación de los cuestionarios incluidos en la muestra y la perforación de los mismos en tarjetas puede servir tanto para las tabulaciones censales completas como para la estimaciones preliminares. Para ello bastará duplicar las tarjetas perforadas de la muestra y guardar los originales para incluirlos en las tabulaciones normales del censo. De este modo, las estimaciones preliminares sólo costarán lo que cueste preparar las tabulaciones.

ii) Muestra de cédulas de empadronamiento. Dentro de este esquema se pueden seguir distintos planes: Para la muestra sistemática se seleccionan en forma sistemática los hogares que se incluirán en la muestra de acuerdo con la fracción muestral establecida. Esta selección se facilita si los empadronadores o las oficinas locales del censo han preparado listas de todos los hogares con el número de personas por hogar en alguna etapa del trabajo censal. Luego de seleccionar la muestra se ubican los cuestionarios correspondientes a ella y se someten a las tareas de crítica, codificación y perforación. Se preparan dos juegos iguales de tarjetas y uno de ellos se intercala en el momento oportuno en el lote censal. Con el otro juego se preparan las tabulaciones preliminares. Este proceso es muy flexible en sus detalles; sin embargo, deberá establecerse con precisión el procedimiento que habrá de usarse, y ejercer luego una supervisión adecuada.

La muestra por conglomerados en una sola etapa tiene como unidad muestral el distrito de empadronamiento, conforme al cual se elaboran

/todos los

todos los datos censales correspondientes a cada distrito seleccionado. No es necesario repetir la crítica, codificación y perforación de las cédulas de empadronamiento de estos distritos en el análisis completo del censo. Una variante es la de extraer primero una muestra pequeña (por ejemplo, del 1%) y elaborar los primeros cálculos preliminares, y luego ampliar la muestra (por ejemplo al 10%) y dar los segundos cálculos preliminares.

La muestra por conglomerados en dos etapas suele utilizarse cuando la extensión de los distritos de empadronamiento varía mucho y resulta imposible estratificarlos satisfactoriamente. El distrito de empadronamiento es la unidad muestral de la primera etapa y el hogar censal la unidad muestral de la segunda etapa. En este caso la muestra incluye un número mayor de distritos que en el caso anterior, pero dentro de cada distrito sólo incluye una muestra de hogares censales. Resulta más eficiente si la selección de las unidades primarias se efectúa con probabilidad proporcional al tamaño de las mismas.

iii) Muestra de tarjetas perforadas. Cuando se tienen los datos censales perforados se pueden adelantar resultados seleccionando mecánicamente una fracción de estas tarjetas. Es fácil calcular el ahorro de tiempo que deriva del uso de tales muestras, ya que es proporcional a la fracción de muestreo que se use.

iv) Algunos resultados en América Latina. En el cuadro 1 se muestran nueve países latinoamericanos que adelantaron resultados tomados de una muestra del censo. Entre estos países hubo variaciones en cuanto al tamaño de la muestra, pero la meta común fue apresurar las publicaciones.

Cuadro 1

País	Fecha del censo	Fecha de publicación	
		De los datos preliminares basados en la muestra	De los datos finales completos
Argentina	1960	1963	1965
Costa Rica	1963	1964	1966
Chile	1960	1962	1964
Ecuador	1962	1964	1964
El Salvador	1961	1962	1965
Honduras	1961	1962	1964
Nicaragua	1962	1964	1964
Perú	1961	1963	1966
Venezuela	1961	1963	1966

Fuente: Tabulación inédita elaborada por el Instituto Interamericano de Estadística. Informe ST/ECLA/Conf.32/L.22, Consejo Económico y Social, Naciones Unidas, 16 de marzo de 1968.

Para estos nueve países el tiempo medio (sin considerar los meses) que transcurrió entre la publicación de los resultados preliminares y la publicación de los del censo fue aproximadamente de dos años.

Los cuadros 2, 3 y 4 ofrecen una comparación entre algunos datos preliminares publicados y los datos finales.

Cuadro 2

VENEZUELA: CENSO DE 1961

	Muestra preliminar a/	Informe final
Total hombres	3 824 000	3 822 000
Total mujeres	3 700 000	3 702 000
Hombres menores de 1 año	148 000	156 000
Mujeres menores de 1 año	153 000	151 000
Hombres 25 a 29 años	286 000	283 000
Mujeres 25 a 29 años	268 000	268 000
Hombres de 85 años y más	4 675	5 808
Mujeres de 85 años y más	13 850	11 494
Hombres jefes de familia	1 017 000	1 009 000
Mujeres jefes de familia	337 000	335 000
Hombres que residen en habitación especial	208 000	224 000
Mujeres que residen en habitación especial	90 000	94 000
Hombres que no se han casado nunca, 15 años y más	964 000	944 000
Mujeres en unión consensual	413 000	412 000
Hombres económicamente activos	1 957 000	1 929 000
Hombres sin empleo	287 000	274 000
Hombres no incluidos entre los obreros	600 000	567 000
Mujeres no incluidas entre los obreros	2 044 000	2 007 000

a/ Los datos preliminares se obtuvieron en una muestra de hogares que abarcó 4% de los hogares de Caracas y 1% de los del resto del país.

Cuadro 3

COSTA RICA: CENSO DE 1963

	Muestra preliminar <u>a/</u> (1964)	Informe final (1966)
Hombres de menos de 1 año	25 300	26 000
Mujeres de menos de 1 año	26 100	25 200
Hombres de 25 a 29 años	41 900	41 300
Mujeres de 25 a 29 años	43 000	43 500
Hombres de 75 años y más	7 478	7 383
Mujeres de 75 años y más	7 553	7 822
Jefes de hogar	231 300	231 200
Esposas de jefes de hogar	171 200	170 700
Hombres alfabetos de 10 años y más	369 800	371 300
Mujeres alfabetas de 10 años y más	371 900	374 200
Hombres analfabetos de 10 años y más	61 200	60 900
Mujeres analfabetas de 10 años y más	65 500	63 600
Empleados	371 500	367 800
Sin empleo	28 400	27 500

a/ Los datos preliminares se basan en la muestra de 5% de los hogares.

Cuadro 4

EL SALVADOR: CENSO DE 1961

	Muestra preliminar <u>a/</u> (1962)	Informe final (1965)
Hombres menores de 1 año	50 800	49 100
Mujeres menores de 1 año	48 900	48 100
Hombres de 25 a 29 años	81 800	80 900
Mujeres de 25 a 29 años	93 800	91 600
Hombres de 75 años y más	12 900	13 000
Mujeres de 75 años y más	17 100	16 200
Hombres nunca casados, 14 años y más	308 900	310 600
Mujeres viudas	47 000	45 100
Empleados	763 400	766 000
Sin empleo	43 200	41 100

a/ Los datos preliminares se basan en la muestra del 5% de los hogares.

/b) Evaluación

b) Evaluación del censo en sus errores de cobertura y en sus errores de contenido

En el levantamiento de un censo que abarca a todas las unidades censales de la población es muy difícil evitar enteramente los errores en el empadronamiento, de modo que los datos censales tendrán validez siempre que los errores no excedan de límites razonables. La medida de estos errores debe ser conocida, para que los usuarios puedan apreciar tanto las limitaciones como el grado de precisión de los datos. La comprobación total de un censo, por su costo prohibitivo y por las molestias que impondría al público, es prácticamente irrealizable. La manera más económica y eficiente de evaluar un censo es efectuar enseguida de él y lo antes posible, comprobaciones basadas en muestras.

Los dos factores adversos que más influyen en la calidad de los datos censales son los errores de cobertura y los errores de contenido.

Los errores de cobertura emanan de la omisión o duplicación de unidades censales. Para verificarlos se utilizan principalmente esquemas muestrales en virtud de los cuales se toma contacto con personas seleccionadas de listas independientes del censo original o de archivos administrativos para comprobar si han sido censadas. (Por ejemplo, se puede verificar si un niño cuyo nombre se ha obtenido de un registro de nacimientos ha sido enumerado en el censo.) El problema de este método reside generalmente en obtener la dirección actualizada de las personas seleccionadas. Para seleccionar este tipo de muestra se acostumbra a utilizar listas combinadas de personas que tiendan a representar a todo el país y a todos los grupos de población. Por ejemplo, listas combinadas de personas enumeradas en el censo anterior, de nacidos en el período intercensal y de extranjeros registrados en los servicios de inmigración. De cada una de estas tres listas se seleccionan personas por el método del muestreo probabilístico y se procura averiguar si fueron enumeradas en el censo.

El coeficiente de penetración muestral en cada lista generalmente se hará en función de las dificultades de enumeración de las personas de esa lista. Así, por ejemplo, aunque los registros de nacimientos en el período intercensal pueden abarcar sólo una pequeña fracción de la población total por mostrar, la muestra puede comprender una fracción mayor, dado que se sabe por experiencias que generalmente la tasa de omisión de niños en los censos es elevada.

Otro método muestral para medir este tipo de error de cobertura es el de la reenumeración. Es decir, se seleccionan ciertas zonas censales y se vuelven a enumerar repitiéndose el proceso censal. Este método no elimina completamente las fallas inherentes a los métodos de enumeración, de modo que en la reenumeración se suele omitir el mismo tipo de personas (por ejemplo, niños) que fue omitido la primera vez. El uso de enumeradores especializados permite reducir este tipo de faltas.

Los dos métodos descritos permiten detectar omisiones censales. Si bien el primero generalmente resulta más caro, es más eficiente porque la selección de las unidades muestrales no es equivalente a la del censo.

Los errores de contenido o de observación comprenden los errores que se cometen tanto por parte del entrevistado como del entrevistador y los motivados por imperfecciones de diseño del cuestionario. Para detectarlos generalmente se seleccionan unidades censales que se vuelven a empadronar y los datos así obtenidos se comprueban con los datos censales. En caso de haber diferencia se realiza una tercera visita a la unidad.

Para que estas investigaciones sean eficientes es preciso utilizar entrevistadores especializados.

c) Control de calidad en la elaboración de los datos

El muestreo puede usarse con eficiencia para medir y controlar la calidad de las distintas operaciones de un proceso censal. Tanto la bondad de los cuestionarios como la codificación, la perfoverificación de la

información y la tabulación de los resultados pueden verificarse por muestreo. Los métodos de control de calidad permiten comprobar eficaz y económicamente la calidad de la elaboración de los datos censales, con lo cual se hace innecesario y poco práctico descubrir y rectificar todos los errores de elaboración repitiendo cada operación para comprobarla. Además, no vale la pena aumentar los costos de verificación del procesamiento de datos una vez logrado cierta mayor exactitud en los resultados censales, dado que éstos siempre contendrán errores por defectos de enumeración y otros errores ajenos al procesamiento. El control estadístico de la calidad de la elaboración de datos tiene por objetivo principal verificar que los errores de las operaciones sean tan poco frecuentes que no exijan rectificación y, en seguida, evitar que se incluyan en los resultados lotes de datos cuya calidad sea excesivamente defectuosa. Con muestras pequeñas es posible localizar las unidades de trabajo muy defectuosas para separarlas y someterlas a una verificación completa. El control debe ejercerse principalmente sobre la labor de los empleados, pues aquéllos cuyo trabajo contenga errores que sobrepasen los límites de aceptación pueden ser desplazados de su labor. La calidad del trabajo de los empleados debe controlarse periódicamente, dado que varía con el tiempo. A su vez es preciso controlar la labor de los supervisores, ya que la experiencia ha mostrado que en muchos casos la supervisión sólo detecta algunos de los errores cometidos por los empleados.

Una inspección del $X\%$ no asegura un ahorro del $(100-X)\%$. Cabe tener en cuenta que el costo de inspección abarca el 100% de la inspección durante el período de prueba, costos fijos de manipuleo de lotes, 100% de la inspección de lotes rechazados, costos de diseño y selección de la muestra y la mantención de registros.

d) Comparación del recorrido o campo de error medio de resultados censales entre algunos países latinoamericanos

Una gran variedad de indicadores muestra que la calidad de los datos estadísticos sobre población varía considerablemente de uno a otro país. La necesidad de conocer estos datos y su grado de precisión ha sido reconocida por todos los países, de modo que las verificaciones postcensales van en aumento y hoy son componentes esenciales de un proceso censal.

Con el objeto de comparar el grado de error envuelto en los resultados censales de distintos países se ha ideado una fórmula (véase Naciones Unidas, Demographic Yearbook, 1960) que indica aproximadamente el recorrido o campo de error que puede atribuirse a las estimaciones de población de cada país para 1967. Esta fórmula considera el número de años desde que se recolectaron los datos, por lo tanto, dichos valores son comparables solamente con respecto al año 1967. Estos recorridos no son márgenes de error que pueden ser aplicados hacia arriba o hacia abajo del valor estimado, sino intervalos dentro de los cuales todos los errores tienen posibilidades de darse (pero como el problema crónico de los censos de población es la omisión censal, lo más probable es que el error sea función de esa deficiencia).

Cuadro 5

RECORRIDO APROXIMADO DEL ERROR QUE PUEDE ATRIBUIRSE A LA
ESTIMACION DE LA POBLACION PARA 1967

(Porcentajes)

Países latinoamericanos

Caribe

Cuba	3.8
Jamaica	2.4
Puerto Rico	2.4
Rep. Dominicana	6.6
Trinidad y Tabago	2.4

Centroamérica

Costa Rica	2.6
El Salvador	5.8
Guatemala	2.2
Honduras	5.8
México	5.2
Nicaragua	4.2
Panamá	5.2

Sudamérica tropical

Bolivia	20.0
Brasil	5.2
Colombia	3.4
Ecuador	5.0
Guyana	6.6
Perú	8.3
Venezuela	4.2

Sudamérica templada

Argentina	6.6
Chile	5.2
Paraguay	7.0
Uruguay	5.2

/Otros países

Otros países

Argelia	1.6
Australia	1.2
Canadá	1.2
Dinamarca	1.2
Estados Unidos	2.4
Francia	2.0
Holanda	1.7
Israel	1.4
Italia	2.2
Japón	1.4
Portugal	2.4
Reino Unido	2.2
República Árabe Unida	1.8
República Federal de Alemania	2.2
Tanzania	1.0
Tunisia	1.6
Turquía	2.2

En las diferentes regiones del mundo el recorrido del error ha sido el siguiente:

Norteamérica y Europa	2.65%
Otras regiones	8.08%
América Latina	6.30%
Asia y Oceanía	7.97%
Asia Menor y África	9.67%

Fuente: The American Statistician, Vol. 25, Nº 1, febrero de 1971.

e) Aplicación en América Latina: Argentina

i) Resultados preliminares del censo de población, familias y vivienda de 1970. Para adelantar resultados sobre el censo de 1970 el Instituto Nacional de Estadística y Censos de la Argentina programó una muestra de tamaño del 2%, de selección sistemática con arranque aleatorio, cuya unidad muestral fue la cédula censal. Cada cédula censal fue identificada con un código numérico secuencial; una vez seleccionados estos códigos numéricos se ubicaron las cédulas censales que pasaron a constituir la muestra. Estas cédulas pasaron luego por todas las etapas del proceso censal: crítica, codificación, traspaso a los formularios de lectura óptica, análisis de consistencia por programa almacenado y grabación de la información en cinta magnética. Todo este proceso se halla aún en etapa de elaboración. Una vez que la información muestral esté registrada en cinta magnética se podrán anotar resultados para el ámbito nacional, para la mayoría de las provincias y para los departamentos más importantes.

Estos registros muestrales se intercalarán luego en las cintas magnéticas que hayan sido grabadas con el resto de los registros censales, para obtener así la información final.

ii) Evaluación del censo de población, familias y viviendas de 1970. Para evaluar las omisiones del censo de 1970 el Instituto Nacional de Estadística y Censos de la Argentina decidió utilizar el método muestral y las proyecciones de los censos de 1947 y de 1960. Se aplicó el método de la encuesta utilizando el procedimiento conocido como verificación de registros y la encuesta continua de empleo y desempleo.

La verificación de registros consiste en el estudio de muestras de personas extraídas de registros administrativos independientes y la comprobación de que esas personas no fueron omitidas en el empadronamiento censal. La persona omitida en la cobertura censal es la que debería haber sido contada en el censo pero que no lo fue. Se hicieron encuestas para controlar la cobertura en tres grupos de población: niños que el día del

/empadronamiento censal

empadronamiento censal tenían hasta dos años de edad; jóvenes cuyo empadronamiento cívico se efectuó entre el 1° de enero y el 30 de junio de 1970, es decir, que el día del levantamiento censal tenían alrededor de 18 años de edad, y jóvenes nacidos en 1950. Se consideraron estos grupos que habitualmente son los que más inciden en la omisión censal con el fin de obtener así una tasa máxima de omisión. El marco muestral que se usó para seleccionar la muestra de niños y la muestra de jóvenes empadronados fue el Registro Nacional de las personas. Los jóvenes de la clase de 1950 fueron muestrados cuando se les llamó a examen médico para cumplir con el servicio militar obligatorio. La muestra incluyó a todos los jóvenes cuyo número de libreta de enrolamiento terminaba en 6.

En el cuestionario muestral, además de preguntárseles si habían sido censados, se les pedía información sobre la cantidad de personas censadas en el hogar donde se encontraban el día del censo.

Para medir la omisión censal se usó también la encuesta de empleo y desempleo que el instituto efectúa tres veces al año. En esta investigación la unidad muestral es el hogar familiar. Ambas encuestas se hallan actualmente en elaboración.

5. Aplicación del muestreo entre períodos censales

La información proporcionada por los censos de población y habitación debe actualizarse permanentemente con el objeto de medir la evolución de las diversas características y de conocer la realidad socioeconómica actual de la población. De allí que un censo de enumeración completa es sólo una de las dos poderosas herramientas de recolección de nuevos datos.

La otra es la investigación muestral entre los períodos censales, lo que al diseñar un sistema estadístico se ganará en eficiencia si se planifica conjuntamente el papel de los censos y de las muestras intercensales. Son muchos ya los países latinoamericanos donde se ha establecido por decreto que para mantener un sistema permanente de

/información, las

información, las oficinas estadísticas nacionales están obligadas a organizar y dirigir encuestas de hogares con propósitos múltiples, como labor de continuidad censal.

En estas encuestas, además de recolectarse información propia de los censos, se miden otras características que no se incluyen en los programas censales, ya sea porque es difícil medirlas en profundidad o porque se consideran perecederas (es decir, que pueden sufrir grandes cambios en breve plazo). Cabe citar entre otras las variables propias de los estudios sobre desocupación, consumo, ingresos, problemas de salud, hábitos alimenticios, actitud de la población frente a políticas socioeconómicas, etc.

Estas encuestas se utilizan para obtener estimaciones en el ámbito nacional y por grandes regiones o áreas. En ellas generalmente se emplean entrevistadores a los que se ha seleccionado, capacitado y supervisado mejor que a los utilizados en el levantamiento censal. Asimismo, en la elaboración de los datos el programa muestral permite mayor control y supervisión tanto de los empleados como del procesamiento mismo. De allí que estas investigaciones muestrales de nivel nacional generalmente suministran, mejores datos que los censos.

a) El censo como marco muestral

Uno de los objetivos de los censos es el de servir de marco muestral para el desarrollo de programas de muestreo. Por lo tanto, antes de seleccionar una muestra, es preciso analizar dicho marco para descubrir si adolece de inexactitudes o duplicaciones, si es incompleto o si está obsoleto. La deficiencia más grave de un marco de este tipo es el ser incompleto, primero porque no es posible descubrir este defecto durante el desarrollo de la encuesta, y segundo, porque es probable que en un marco incompleto se hayan omitido solamente las unidades que poseen ciertas características especiales, lo que puede introducir desviaciones en los resultados de la encuesta.

Los censos se realizan con poca frecuencia, de manera que pueden resultar anticuados como marco si se los usa mucho tiempo después con este fin. Por otra parte, en las primeras etapas de los censos es posible que la información no sea fácilmente accesible. Estas y otras desventajas pueden eliminarse, por lo menos parcialmente, construyendo una muestra maestra adecuada en la época en que se levanta el censo, para extraer posteriormente de ella las submuestras que se necesiten en el futuro. En estos casos conviene que la unidad de muestreo de la muestra maestra sea la vivienda y no la persona o el hogar censal que la ocupe en el momento del censo, para que la muestra conserve actualidad.

b) Experiencia en América Latina

i) Argentina. El Instituto Nacional de Estadística y Censos de la Argentina iniciará una encuesta continua de hogares de propósitos múltiples a fines de 1972, primordialmente para posibilitar la actualización permanente de la información censal y poder estudiar a fondo la realidad socio-económica argentina. La encuesta permitirá realizar investigaciones sobre mano de obra; condiciones de vivienda; estructura del hogar (desde el punto de vista demográfico, educativo y de su comodidad); ingresos y gastos familiares; migración interna de la población; actitud de la población hacia aspectos sociales, económicos y políticos y otros estudios de importancia que se realicen en una o varias ocasiones. La encuesta comenzará en la zona metropolitana (Capital Federal y Gran Buenos Aires), para extenderse luego hasta a estimaciones de carácter nacional y por regiones de desarrollo, provincias, áreas y grandes conglomerados. Su marco muestral es el Censo Nacional de Población, Familias y Viviendas de 1970.

La encuesta se llevará a cabo tres veces al año y abarcará aproximadamente 20 000 hogares en todo el país, distribuidos en 4 000 de la zona metropolitana, 5 000 de los grandes conglomerados, 6 100 de la zona urbana del resto del país y 4 900 de la zona rural.

ii) Chile.

ii) Chile. El Instituto Nacional de Estadística de Chile proyecta iniciar a mediados de 1972 una encuesta continua de hogares de propósitos múltiples, en el ámbito nacional. Se darán estimaciones separadas urbanas y rurales para toda la nación; urbanas y rurales para cada una de las 12 zonas definidas del país; para cada uno de los tres grandes conglomerados del país y para cada una de las 17 ciudades o agrupaciones de ciudades contiguas que tienen más de 40 000 habitantes. La muestra abarcará alrededor de 33 000 viviendas. La encuesta tendrá por objeto principal medir características relativas a empleo y desempleo, vivienda, hogar y otros aspectos económicos y sociales de la población.

El marco muestral utilizado para seleccionar la muestra de viviendas es el Censo Nacional de 1970. Este marco está formado por las hojas de control de empadronamiento que fueron confeccionadas durante el levantamiento censal. Cada hoja de control de empadronamiento consta de 34 líneas. La línea corresponde a una unidad de vivienda, de la cual se especifica la dirección y el número de personas empadronadas en ella; además, su ubicación geográfica y en un recuadro aparece dibujado un croquis del sector al que corresponde la hoja. La preparación de la muestra se hará seleccionando primero las hojas de control de empadronamiento y luego las viviendas.

6. Censos por muestreo

Cuando no se dispone de los elementos necesarios para un censo de enumeración completa, el actual desarrollo de la teoría y la práctica de las encuestas por muestra ofrece una buena solución, la de los "censos por muestras".

Para los censos por muestras es preciso disponer del mismo tipo de medios que para los censos de enumeración completa pero en una medida más limitada.

/a) Ventajas

a) Ventajas de los censos por muestras

Para levantar un censo por muestreo los datos se recogen de una muestra de unidades que representa al universo. Como en relación con este universo la muestra es pequeña, existe análoga diferencia en la cantidad de trabajo y recursos necesarios.

Entre las condiciones que se imponen a un censo por muestreo está en primer lugar la precisión exigida o la magnitud de los errores de muestreo tolerados. Si se desea una gran precisión se debe trabajar con muestras muy grandes, y en estos casos se reducen las ventajas de este tipo de censos. Si se acepta un grado de precisión moderado, los censos por muestras permiten una mayor flexibilidad en el trabajo que los censos de enumeración completa.

Para conseguir estimaciones de una precisión determinada se necesita una muestra compuesta por un número de unidades casi invariable, sea cual sea el tamaño del universo en cuestión: para dar estimaciones de carácter nacional el tamaño de la muestra será igual que para dar estimaciones respecto a una provincia determinada. Así, si un país tiene 10 provincias cuyas medidas y varianzas son las mismas que las del universo, el tamaño de la muestra para cada una de esas 10 provincias, dada una cierta precisión, será igual al tamaño de la muestra para el universo con igual precisión. En otras palabras, para tener estimaciones respecto a las 10 provincias por separado se requiere una muestra 10 veces mayor; por lo tanto, si se desea obtener estimaciones relativas a un número elevado de unidades administrativas son menguadas las esperanzas de reducir las dificultades recurriendo al empleo de censos por muestras.

Los censos por muestras pueden adaptarse a una gran variedad de circunstancias. Por ejemplo la recolección de datos puede hacerse en varias encuestas que se efectúan en distintos períodos del año y se dedican a diversas partes del programa, para obtener luego la información censal compilando los datos de todas ellas.

/b) Desventajas

b) Desventajas de los censos por muestras

Los censos por muestras no pueden tabularse por unidades administrativas demasiado pequeñas (pueblos, ayuntamientos, etc.) dado que para obtener estimaciones en ese nivel la muestra tendría que ser demasiado grande y, por lo tanto, dejaría de ser ventajosa. Además, los censos por muestras son poco eficientes para obtener informaciones sobre cuestiones poco generalizadas, como por ejemplo, el número de personas de edad avanzada y su distribución por sexo y lugar de residencia.

Los censos por muestra, de otro lado, tienen poca utilidad como marco muestral.

c) Experiencias con censos por muestreo en América Latina

i) Haití. Uno de los factores principales que impulsaron a la República de Haití a realizar un censo por muestreo fue el alto costo de los censos de enumeración completa. Se decidió hacer un censo de enumeración completa sólo en la zona urbana y aplicar el muestreo en la zona rural.

La zona rural se dividió en secciones rurales y estas secciones rurales se dividieron en secciones de enumeración. Se seleccionó el 50% de las secciones rurales de manera que cada una de ellas tuviese igual posibilidad de resultar seleccionada. Dentro de las secciones rurales se eligieron las secciones de enumeración en forma sistemática, con arranque aleatorio. La fracción de muestreo fue del 10% de las secciones de enumeración.

Las secciones de enumeración fueron luego censadas en su totalidad.

El levantamiento de datos en las zonas urbana y rural tomó 6 semanas. Para ello se contó con el siguiente personal:

	<u>Personas</u>
Zonas urbanas	
Entrevistadores	727
Supervisores de barrio	107
Supervisores de zona	51
Supervisores generales	3
Asistentes de supervisores	4
Total zona urbana	892
Zonas rurales	
Entrevistadores	514
Supervisores	257
Total zona rural	771
Otros	
Supervisores de región	112
Supervisores de distrito	27
Supervisores de departamento	13
Inspectores generales	5
Total otros	157
<u>Total</u>	<u>1 820</u>

C. VENTAJAS Y DESVENTAJAS DEL MUESTREO EN LOS CENSOS DE POBLACION

El censo clásico supone la enumeración de todas las unidades constitutivas del universo, es decir, la universalidad de la enumeración. La realización de encuestas por muestreo, por otro lado, es una práctica que supone una desviación respecto del principio de universalidad. Esta es la principal diferencia entre el censo y la muestra.

Sin embargo ambos se complementan, y en el desarrollo científico de las labores estadísticas de los países ambos desempeñan un papel fundamental. De allí que las ventajas y desventajas del muestreo en los censos por lo general dependan principalmente de las necesidades, de los recursos y de la forma en que ambas técnicas se coordinen.

1. Ventajas del muestreo

Entre las principales ventajas podemos citar las siguientes:

a) Reducción de los gastos

Interrogar a una parte de las unidades componentes de un universo generalmente es más económico que interrogarlas a todas. Las distintas operaciones que determinan el costo de una encuesta (entrevistas, supervisión, transporte, elaboración manual y mecánica de datos, etc.) tienen sólo una fracción del costo total de realizar esas mismas operaciones abarcando a todas las unidades.

Aun cuando el costo por unidad sea superior en una muestra que en una investigación del 100% de la población, dado los métodos más precisos que generalmente se utilizan para obtener y elaborar los datos en las encuestas, la aplicación del muestreo permite ahorrar dinero. Este ahorro depende principalmente del plan muestral, es decir, del tamaño de la muestra, del grado de complejidad del método para seleccionar unidades, de los objetivos que se quieran alcanzar y del nivel y la eficiencia deseados en las estimaciones.

b) Reducción

b) Reducción del tiempo

La elaboración y tabulación de los datos de la muestra toma mucho menos tiempo que la tabulación final de un censo, de modo que los resultados obtenidos a través de muestras tienen más actualidad. Así, encuestas realizadas en distintas ocasiones permiten medir la dinámica de una población.

c) Mejor calidad del trabajo

En una investigación por muestreo el personal que trabaja en el terreno y el de oficina puede seleccionarse más cuidadosamente que en una investigación censal, lo que generalmente redundará en una mejor calidad de los datos obtenidos. Además, dado el menor volumen de información que se debe elaborar, es más fácil verificar las inconsistencias detectadas por los controles estadísticos que se hayan establecido.

d) Evaluación de la eficiencia

Una técnica es más eficiente que otra cuando los resultados estimados que ella arroja tienen una menor dispersión con respecto al valor esperado. Este concepto de dispersión es inverso al de precisión. La característica más importante del muestreo estadístico es que permite prever los márgenes de error dentro de los cuales se desea obtener el valor estimado, con una probabilidad dada. Es decir, al planear una encuesta se puede determinar el grado de precisión que habrá de tener, según el destino de los datos. La encuesta es ante todo función de la precisión, del costo, del tiempo, de la técnica que se utilice, de la cartografía disponible y de la idoneidad del personal que la realice. Además, su éxito depende en gran parte de la colaboración del entrevistado, de la posibilidad de medir las unidades seleccionadas y de la calidad de esta medición. La exactitud de una encuesta adecuadamente planeada y realizada puede determinarse mediante un examen estadístico de los resultados.

2. Desventajas del muestreo

Entre ellas cabe mencionar las siguientes:

a) Alcance limitado de la información

Un caso extremo al cual no puede aplicarse el muestreo se presenta cuando se necesita información sobre cada uno de los individuos que constituyen la población. Por lo general, el muestreo no permite hacer cálculos relativos a zonas administrativas pequeñas. Cuando la muestra sobrepasa cierta proporción de la población de la zona para la cual debe arrojar resultados, tal vez sea más económico y eficiente el empadronamiento completo.

La precisión de un estimador suele depender del tamaño de la muestra y no de la fracción de muestreo, lo que puede considerarse hasta cierto punto como una limitación. En ciertos casos tal vez se necesite una muestra casi del mismo tamaño para obtener datos de una provincia que para obtenerlos de todo el país.

b) Errores de muestreo

Además de estar sujeto a los errores que pueden resultar de la medición de la unidad muestral, comunes también al censo, el muestreo lo está también a la variación existente entre las distintas muestras que es posible seleccionar.

c) Posibilidad de demora en la obtención de los resultados definitivos de los censos

Los resultados preliminares que proporcionan las muestras se refunden posteriormente con los demás datos para obtener los resultados finales. Este proceso de intercalación de datos puede contribuir a que la duración y el costo totales de la tabulación sean superiores a lo que serían si no se tabularan previamente los resultados preliminares.

/d) Confianza

d) Confianza pública

La diferencia que existe entre los resultados de una muestra y los resultados de un censo no debe interpretarse como la diferencia entre los resultados de una muestra y las verdaderas características de la población. Bajo la impresión errónea de que los censos de enumeración completa ofrecen necesariamente una garantía de exactitud, mucha gente desconfía aún de los resultados de las encuestas basadas en muestras probabilísticas.

D. CONCLUSIONES

Un censo de enumeración completa no es sino un caso especial de una investigación por muestreo; por lo tanto, es posible concebir un modelo general. Este modelo deberá considerar ciertas condiciones esenciales, como el diseño del cuestionario, los procedimientos de investigación, el tipo de personal calificado para realizar distintas actividades, su adiestramiento, el control de las operaciones, etc.

Consideramos que existe un parámetro desconocido que deseamos estimar, ya sea utilizando el censo o la muestra. Designemos con U al parámetro y con X a la estimación del parámetro obtenido mediante el censo o la muestra. El error medio cuadrático de esta estimación puede analizarse así:

$$\begin{aligned} E (X - U)^2 &= \text{Varianza muestral} \\ &+ \text{Varianza de los errores no muestrales.} \\ &+ \text{Covarianza entre los errores muestrales y no} \\ &\quad \text{muestrales} \\ &+ (\text{sesgo})^2 \end{aligned}$$

La varianza muestral depende del diseño de la muestra. Los términos de la varianza muestral y de la covarianza valen cero en el caso de un censo de enumeración completa.

/Cualquiera que

Cualquiera que sea el diseño de trabajo que se realice (censo o muestra), éste se verá afectado por el problema de la variación. Por ello, siendo el objetivo principal de toda investigación el de conseguir resultados fidedignos dentro de normas presupuestarias y de plazos establecidos, la metodología estadística actual hace uso de ambas técnicas, complementándolas entre sí en la forma más racional y adecuada.

Los censos de población, que generalmente se levantan cada 10 años, constituyen un buen marco muestral cuando cada uno de los elementos de la población que lo componen están claramente identificados. Además, los valores censales desempeñan un papel importante en el ajuste de estimaciones muestrales.

A su vez, las muestras ayudan al censo al permitir una programación más efectiva del mismo, complementándolo en el período de levantamiento, evaluándolo y continuándolo en los períodos intercensales.

El avance continuado de la metodología estadística y de los complejos sistemas de computación digital, así como la evolución del comportamiento de las sociedades humanas, hacen que el muestreo y los censos que se basan en ellos estén en continuo proceso de cambio, que a veces encuentra resistencia. Al diseñarse procedimientos nuevos que exigen el uso de tecnologías más modernas y que tienden a mejorar la calidad y la rapidez de los resultados, surgen resistencias nacidas de razones intuitivas.

La investigación generalmente ha demostrado que no basta con la intuición y la experiencia. Las decisiones de adoptar métodos más eficientes deben ser guiadas por comprobaciones científicas y experiencias rigurosas, y no por inferencias o deducciones no comprobadas.

Los modelos futuros que se están programando para recolectar y producir información exigen una complementación cada vez más estrecha entre los censos y el muestreo. Se estima que muchos censos futuros se realizarán en parte por correo. Los cuestionarios censales se dirigirán a las

direcciones de las viviendas y no a las personas. Los componentes de esos hogares llenarán los cuestionarios y los remitirán de vuelta por correo. Parte de los hogares recibirán cuestionarios con preguntas censales, y otra parte recibirá cuestionarios que, además de las preguntas censales, contendrán preguntas muestrales. Un control del registro de direcciones permitirá que los entrevistadores visiten a los que no respondan para obtener la información. Estos entrevistadores también tomarán contacto personalmente o por teléfono con quienes hayan devuelto cuestionarios incompletos.

La oficina de correos desempeñará un papel importante. Los carteros generalmente conocen bien la zona en que trabajan. Por lo tanto, ellos procurarán identificar y añadir cualquier dirección omitida en la lista de direcciones preparada por la oficina del censo. La experiencia ha demostrado que en la enumeración censal suelen omitirse muchas direcciones, generalmente en grandes ciudades y en zonas rurales. Con la participación de la oficina de correos se espera reducir considerablemente estas omisiones. En sectores de la población donde no se utilice el correo y donde se cuente con entrevistadores especializados se podrá recoger la información en formularios especiales que luego serán leídos por una unidad de lectura óptica integrada a un sistema de sistematización de datos; los datos entrarán así directamente al proceso de análisis y registro mecánico. Estos formularios pueden diseñarse tomando en cuenta las preguntas que habrán de codificarse con posterioridad, además de las que no necesitan codificación. Los formularios que sean rechazados por la unidad lectora, ya sea por ilegibles o por desgaste del formulario, deberán ser reproducidos. Este método evita la cédula censal y el proceso de perfoverificación.

Los países latinoamericanos, mucho de los cuales tienen gran superficie y en su mayoría poblaciones escasas y heterogéneas, lo que puede

elevar el costo por unidad censada, deben aprovechar al máximo todos los adelantos de la técnica, adaptarlos a sus necesidades estadísticas y censales y conformarlos a las estructuras socioeconómicas propias de cada país. Un país latinoamericano con variados climas y geografía, con una o dos grandes ciudades que aglutinan a la mayoría de su población, con distintos niveles sociales que van desde la clase analfabeta hasta la sociedad cerrada pasando por todos los niveles intermedios, con grandes poblaciones marginadas, con poblaciones autóctonas que aún viven en sus antaños lugares manteniendo costumbres, lenguaje y una economía doméstica propia, con no pocas zonas rurales de difícil acceso, con planes regionales de desarrollo, con problemas políticos y sociales, debe estudiar y programar cuidadosamente las investigaciones censales-muestrales, y aplicar a cada sector los métodos que optimicen los resultados con los recursos disponibles.

Por todo lo expuesto en este trabajo, se puede llegar a las siguientes conclusiones:

1. El muestreo es una disciplina científica cuya aplicación es indispensable para ayudar a planificar, ejecutar, analizar, evaluar los censos y a darles continuidad.
2. A su vez, los datos censales son de suma utilidad en los procedimientos muestrales, primero porque permiten construir marcos muestrales, y segundo porque los parámetros pueden usarse como valores de ajuste en las estimaciones muestrales.
3. Los países latinoamericanos, al formular los correspondientes planes censales nacionales, deben prever el uso de métodos muestrales en las distintas etapas de programación del censo, con el fin de optimizar métodos, recursos y resultados.
4. El avance de las técnicas muestrales ha aumentado considerablemente su precisión, de modo que para obtener una máxima eficiencia en los

diseños muestrales éstos deben ser ejecutados y llevados a cabo por personas con conocimiento y experiencia en los métodos de muestreo y en sus aplicaciones.

5. Los países latinoamericanos deben aprovechar al máximo los recursos de que disponen para desarrollar sus sistemas estadísticos de captación, elaboración y publicación de datos, coordinando sobre bases científicas el método censo-muestreo.

Bibliografía

Muchos de los conceptos y ejemplos vertidos en este trabajo se extractaron de la siguiente bibliografía:

1. S.S. Zarkovich, Quality of Statistical Data, FAO, Roma, 1966.
2. S.S. Zarkovich, Los métodos de muestreo y los censos, FAO, Roma, 1967.
3. E.S. Marks, "Techniques of Sampling and Statistical Evaluation for Census Work", Proceedings of the American Statistical Association, Washington D.C., 1961, pág. 49.
4. Dirección del Censo de los Estados Unidos, Procedural Report on the 1960 Censuses of Population and Housing, Working Paper N° 16. Washington, D.C.
5. Dirección del Censo de los Estados Unidos, Evaluation and Research Program of the U.S. Censuses of Population and Housing, Serie ER 60, N° 7. Washington, D.C., 1960.
6. The Future of Statistics, Documentos de una conferencia realizada en la Universidad de Wisconsin en junio de 1967, Academic Press, Inc., Nueva York.
7. Des Raj, Sampling Theory, Mc Graw-Hill, Nueva York, 1968.
8. F. Yates, Sampling Methods for Censuses and Surveys, Hafner Publishing Company, Nueva York, 1960.
9. Naciones Unidas, Manual de Métodos de Censos de Población, Estudios de Métodos, Serie F N° 5, Rev. 1, Nueva York, 1958.
10. Naciones Unidas, Principles and Recommendations for the 1970 Housing Censuses, Statistical Papers, Serie M N° 45, Nueva York, 1967.
11. Naciones Unidas, Use of Sampling in Population and Housing Censuses (ST/ECLA/Conf.32/L.12), 26 de marzo de 1968.
12. Naciones Unidas, Censo experimental de población de la República Dominicana (ST/ECLA/Conf. 32/L.16).
13. Naciones Unidas, Censo de prueba de la comuna de Valdivia, Chile (ST/ECLA/Conf.32/L.14).
14. H.H. Hansen y J. Waksberg, "Research on Non-Sampling Errors in Censuses and Surveys", Review of the International Statistical Institute, Volumen 38, N° 3, pág. 317, La Haya, 1970.

