


# **Departamento Administrativo Nacional de Estadística**



**Dirección de Censos y Demografía - DCD**

**Lineamientos para la Selección de Muestra  
Cocensal en DMC  
Censo General 2005 - CGRAL**

**Mayo 2005**

|   |   |   |
|---|---|---|
|  | <p align="center"><b>LINEAMIENTOS PARA LA SELECCIÓN DE<br/>MUESTRA COCENSAL EN DMC<br/>CENSO GENERAL 2005 - CGRAL</b></p> | <p>CÓDIGO: CM-CGRAL-DIM-01<br/>VERSIÓN : 02<br/>PÁGINA: 2<br/>FECHA: 15-05-05</p> |
| ELABORÓ: Equipo CM  | REVISÓ: Coordinador CM  | APROBÓ : Director DCD   |

## TABLA DE CONTENIDO

|   |   |
|---|---|
| INTRODUCCIÓN.....   | 3 |
| OBJETIVO .....  | 5 |
| PROCEDIMIENTO PARA SELECCIÓN DE LA MUESTRA COCENSAL ..... | 5 |
| 1. Temario básico y temario ampliado.....                 | 5 |
| 2. Implementación del diseño de muestreo Bernoulli .....  | 8 |
| 3. Consideraciones finales .....                          | 9 |

## INTRODUCCIÓN

El operativo censal que se llevará a cabo durante el presente año, tiene varias características innovadoras, tanto desde el punto de vista metodológico como operativo y tecnológico. Desde el punto de vista operativo, pueden destacarse:

- **La inmovilización parcial**, llevando a cabo el operativo en un periodo de tiempo más prolongado (alrededor de 4 meses), con lo que se consigue un mayor control del operativo, la posibilidad de efectuar correcciones en campo y se aprovecha mejor el trabajo de capacitación de los entrevistadores que, usualmente, no alcanzan el tope de su capacidad sino hasta el segundo o tercer día de trabajo.
- **La utilización de cartografía IGAC**, en algunas zonas en que ésta cuenta con un nivel mayor de precisión y actualización que la cartografía DANE.

En lo metodológico se distinguen, entre otras, las siguientes innovaciones:

- **Su carácter general**, abarcando no solamente el tema de población y vivienda usual, sino también el de unidades económicas y unidades agropecuarias.
- **La indagación de un temario ampliado de población y vivienda**, mediante la aplicación de un cuestionario ampliado a una muestra probabilística de hogares en cada uno de los municipios.
- **La utilización de modelos demográficos para la generación de cifras**, considerando que ya no se cuenta con una inmovilización total sino parcial de la población a censar.

En lo tecnológico, cabe subrayar las siguientes características:

- **La utilización de Dispositivos Móviles de Captura – DMC para la recolección de información**, optimizando la inversión de recursos,

mejorando la calidad de la información recolectada, por los controles y validaciones que es posible efectuar directamente en campo, y disminuyendo el tiempo usualmente requerido para la consolidación de información y generación de cifras censales.

- **La creación de mecanismos especializados de monitoreo y control del operativo**, haciendo uso de la tecnología de punta en cuanto a bases de datos y permitiendo con esto, mantener bajo control el operativo de campo y minimizar el tiempo de respuesta frente a los problemas que puedan presentarse.

Frente a los retos planteados por la implantación de estas novedades, cada uno de los equipos que conforman la estructura organizacional del presente censo ha tenido que generar soluciones y alternativas que satisfagan las necesidades del proyecto.

En cuanto al tema particular que es de interés para este documento, esto es, la indagación de un temario ampliado para las temática de población y vivienda, fue necesario llevar a cabo un estudio concienzudo de las necesidades de información de cada uno de los municipios, considerando las capacidades tecnológicas y operativas de este censo, así como la teoría estadística que soporta el diseño de muestras probabilísticas.

Como resultado de este estudio, por la simplicidad de su implementación y por su robustez frente a errores operativos, se eligió el diseño de muestreo denominado Bernoulli para la selección de los hogares, sujeto de entrevista utilizando el cuestionario ampliado. Este diseño, proveniente de la rama estadística de Control de Calidad, es muy simple en su implementación, lo que disminuye las posibilidades de errores no controlables por efecto de la selección en campo.

## **OBJETIVO**

Proporcionar lineamientos para efectuar la selección de muestra de hogares para la aplicación del cuestionario ampliado, utilizando el dispositivo móvil de captura.

## **PROCEDIMIENTO PARA SELECCIÓN DE LA MUESTRA COCENSAL**

### **1. Temario básico y temario ampliado**

El cuestionario del Censo General 2005, en lo que corresponde a Población y Vivienda, está compuesto por un temario básico y un temario ampliado, con cerca de 35 y 39 preguntas respectivamente. Conforme a lo estipulado en la metodología, el temario básico será indagado en todas las viviendas, hogares y personas residentes en el país, en el momento del censo en el respectivo municipio; adicionalmente, el temario ampliado, será indagado en todos los hogares seleccionados como muestra, en todas las personas pertenecientes a estos hogares y en las viviendas en donde, al menos uno de los hogares residentes, haya sido seleccionado como muestra.

La elección del hogar como la unidad de muestreo en lugar de la vivienda o la persona, obedece a criterios tanto estadísticos como de costos:

- Una muestra con la vivienda como unidad de muestreo, será útil para indagar las características de vivienda del temario ampliado pero seguramente, será insuficiente en términos de tamaño de muestra final, para obtener estimaciones con buena precisión para variables de hogar o persona. Esto, debido a que si bien el tamaño de muestra final en hogares o personas es más grande que el de viviendas, se reduce por efecto la correlación que suele presentarse para una importante cantidad de variables, entre personas y hogares que habitan en la misma vivienda.

- Una muestra con la persona como unidad de muestreo puede ser mucho más costosa que una de hogares puesto que, si bien el tamaño de muestra para personas es más grande, puede verse ostensiblemente reducido en términos de las variables de vivienda o de hogar, como resultado de la repetición de información de estos dos niveles, por la selección de varias personas que habitan en una misma vivienda o que son miembro de un mismo hogar.

Dentro del temario ampliado, se incluyen variables tanto de hogar, como de vivienda y persona. Por esto es importante que, como parte del módulo de vivienda, se haga una identificación de los hogares residentes y se lleve a cabo la selección de muestra para cada uno de ellos.

El procedimiento elegido para la selección de muestra, denominado Diseño de muestreo Bernoulli, es muy usado en Control de Calidad por su simplicidad; puede explicarse fácilmente, a partir de un esquema de producción en cadena, en donde, sobre una banda transportadora van pasando uno a uno, los productos: en la medida en que los productos pasan por un punto determinado, se efectúa un procedimiento aleatorio para decidir si el producto en cuestión será o no parte de la muestra que será sometida a control de calidad.

Este caso es asimilable al que se tiene en el caso del Censo General 2005, en donde, un entrevistador pasa por cada uno de los hogares y decide, al encontrar a cada hogar, si hará o no parte de la muestra cocensal y con ello, si se aplicará el cuestionario ampliado o el básico.

El proceso aleatorio usado para elegir a los hogares que pertenecen a la muestra cocensal debe ser llevado a cabo para todos y cada uno de los hogares del municipio, en forma independiente para cada uno de los hogares, esto es, sin

tener en cuenta si dos o más hogares residen en la misma vivienda, edificación o AG y sin considerar si otros hogares han sido o no seleccionados.

En el caso particular en que en una vivienda solo resida un hogar, debe aplicarse a éste el proceso aleatorio que determinará si hace o no parte de la muestra. Si dos o más hogares residen en una misma vivienda, el proceso debe ser efectuado para cada uno de ellos en forma tal que, es posible que uno, los dos, o ninguno de los hogares sea seleccionado.

Cuando un hogar es seleccionado, tanto él como todas las personas que lo componen se consideran como sujeto de aplicación del cuestionario ampliado, y deben indagarse también las características de vivienda del temario ampliado, para la vivienda que habitan.

Cuando en una vivienda reside más de un hogar, debe indagarse el temario ampliado de vivienda si al menos uno de los hogares residentes es seleccionado. Si son seleccionados dos o más de los hogares residentes, se indagarán las características de vivienda ampliadas, solo la primera vez.

Para llevar a cabo la selección para cada uno de los hogares residentes, se han incluido dentro del módulo de vivienda las preguntas CTL1 y CTL2, que indagan sobre el número de hogares en la vivienda y hacen la identificación de cada uno de ellos. Una vez generada por el DMC una respuesta para CTL2 (número de orden del hogar en la vivienda), debe efectuarse el procedimiento de selección de muestra sobre el hogar identificado. Una vez terminado el proceso de diligenciamiento de la información de hogar y persona para un hogar en particular, y luego de recolectar la información de unidades económicas o agropecuarias asociadas, debe volverse a la pregunta CTL2 del módulo de vivienda para

identificar el siguiente hogar, si existe, y efectuar el proceso de selección de muestra nuevamente.

Es importante enfatizar que, es necesario llevar a cabo la selección de muestra para todos y cada uno de los hogares residentes del municipio, independientemente de si habitan o no la misma vivienda, edificación o manzana, e incluso, de si otros han sido o no seleccionados.

## **2. Implementación del diseño de muestreo Bernoulli**

Para la implementación del diseño de muestreo Bernoulli, solamente se requieren un valor de comparación y un mecanismo para generar números aleatorios con distribución uniforme entre 0 y 1, tal como la función aleatorio() de Excel. El procedimiento de selección consiste en, para cada uno de los hogares, generar un número aleatorio como el descrito arriba, comparar este número con el valor de comparación preestablecido y, si el aleatorio es menor o igual que el valor preestablecido, se considera que el elemento ha sido seleccionado. En caso contrario, no se incluye este hogar dentro de la muestra y se aplica para él, el cuestionario básico.

El valor de comparación es un valor fijo, preestablecido desde el diseño de la muestra, que cambia de AG en AG. Este valor, se entregará como complemento a las tablas con la información de AG.

Para ejemplificar la selección de muestra, suponga que se desea llevar a cabo el proceso de selección para el hogar 01 de la vivienda 05 de la edificación 22 de la AG 897456 del municipio de Túquerres, que tiene como valor de comparación 0.7. Se procede como sigue:

- Genere un número aleatorio con distribución uniforme entre 0 y 1



Respuesta: 0.478975

- Lleve a cabo la comparación entre el valor y el aleatorio generado  
¿ $0.478975 <= 0.7$ ? Respuesta: Sí.
- El hogar 01 es seleccionado y se aplica para él, las personas que lo componen y la vivienda donde reside, el cuestionario ampliado.

Suponga que en la misma vivienda, residen otros dos hogares. Pueden presentarse tres casos:

- **Que ninguno de los dos hogares restantes sea seleccionado:** A pesar que todas las personas residan en la misma vivienda, se aplicará el cuestionario ampliado a aquellas que pertenezcan al hogar seleccionado y el básico a las demás. Se indagará una sola vez la información de vivienda ampliada.
- **Que uno de los dos hogares restantes sea seleccionado:** En este caso, se aplicará el cuestionario ampliado a las personas pertenecientes a los dos hogares seleccionados y el básico a las demás. En cuanto a las características de vivienda ampliada, se indagarán solo la primera vez.
- **Que los dos hogares restantes sean seleccionados:** Se indagará el temario ampliado a todos los hogares seleccionados y las personas que los componen. El temario de vivienda ampliada se indagará solo la primera vez.

### 3. Consideraciones finales

En cuanto al manejo de la información a nivel de base de datos, tanto el aleatorio generado, como el valor de comparación utilizado deben ser capturados, guardados y formar parte de la información de cada hogar. En cuanto a la



**LINEAMIENTOS PARA LA SELECCIÓN DE  
MUESTRA COCENSAL EN DMC  
CENSO GENERAL 2005 - CGRAL**

CÓDIGO: CM-CGRAL-DIM-01  
VERSIÓN : 02  
PÁGINA: 10  
FECHA: 15-05-05

vivienda, no es necesario capturar ninguna información adicional, ni siquiera el, o los números del, o los hogares por causa de los que se aplicó el formulario ampliado.

Para el procesamiento de información a partir de la muestra seleccionada se requieren, la información de aleatorio, valor de comparación (para hogares) y conteos de hogares, viviendas y personas que pueden efectuarse sobre la base censal consolidada.

En el caso de resguardos indígenas o territorios colectivos de comunidades negras, se ha decidido que se aplicará siempre el formulario ampliado. Por esto, siempre que el entrevistador se encuentre en AG clase 3 y marque en Territorialidad uno de estos dos, deberá asignarse para este hogar el valor de comparación 1 y así, se aplicará el formulario ampliado en todos los casos.

Finalmente, cabe resaltar que el diseño elegido exige, para su correcto funcionamiento, que el procedimiento aleatorio mencionado sea efectuado para TODOS y cada uno de los hogares del municipio, ausentes o presentes; así como que el número aleatorio generado para la selección de muestra sea efectivamente aleatorio y con distribución uniforme en el intervalo 0, 1. Este diseño tiene tamaño de muestra aleatorio, y la estabilidad del tamaño de muestra depende de la aleatoriedad del número generado para cada hogar.