

METODOLOGÍA DEL DISEÑO DE LA MUESTRA ENCUESTA DE INGRESOS Y GASTOS DE HOGARES URBANOS Y RURALES - ENIGHUR 2010

1. Antecedentes

La Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales - ENIGHUR 2010- tiene como objetivo general proveer los insumos indispensables para la formulación del Nuevo Índice de Precios al Consumidor (IPC) del INEC.

En este contexto, el nuevo IPC deberá tener como peculiaridad la comparabilidad entre series históricas estadísticas, por lo tanto se deberá considerarse la totalidad o parte de las ciudades donde actualmente investiga el INEC precios al consumidor final.

Cabe mencionar que en el diseño muestral de la ENIGHUR 2010 se ha considerado las siete Zonas Regionales de Planificación propuesta por la Secretaría Nacional de Planificación y Desarrollo (SENPLADES) como dominios co-representados para las futuras estimaciones, para que de esta manera el nuevo IPC mida la inflación que ocurra en el Área Urbana de dichas zonas.

2. Objetivo General

La Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos - Rurales ENIGHUR 2010 - 2011, tiene como objeto proporcionar información sobre la distribución, monto y estructura del ingreso y gasto de los hogares, a través de la caracterización demográfica y socioeconómica de los mismos.

3. Objetivo Especifico

-Proveer el marco referencial para la estructuración de un nuevo Índice de Precios al Consumidor (IPC).

-Proporcionar la información estadística adecuada para realizar el Cambio de Base del Índice de Precios al Consumidor Urbano (IPC), es decir, la información requerida para actualizar la “Canasta Básica Uniforme de Bienes y Servicios de Consumo de los Hogares”.

-Proporcionar al IPC el cálculo de Índices Complementarios Estratificados, según ingresos, disponibilidad de bienes de los hogares segmentaciones urbanas etc., los mismos que servirán para el análisis del impacto socioeconómico del embate inflacionario de la población.

-Identificar las fuentes de aprovisionamiento de los hogares de los diferentes tipos de bienes y servicios de los hogares urbanos – rurales.

-Identificar nuevos patrones de consumo de los hogares urbanos y rurales.

- Viabilizar el conocimiento detallado de la estructura del presupuesto de los hogares urbanos – rurales, de acuerdo a los distintos orígenes y montos de los ingresos y la forma del gasto, así como los hábitos de consumo de los diferentes bienes y servicios.
- Traducir la relación anterior en la medición del ahorro o endeudamiento de los hogares urbanos – rurales.
- Posibilitar la medición de variaciones periódicas (trimestrales) de aspectos socio-económicos a nivel nacional, regionales naturales, zonas regionales de planificación y por ciudades principales.
- Posibilitar información que sirva de insumo para medir la pobreza a través de cualquier método, esto es: según acceso a los servicios básicos, pobreza por consumo, flujos monetarios, estructurales y pobreza subjetiva de los hogares urbanos – rurales.
- Generar información para la estimación global del consumo final de los hogares para las Cuentas Nacionales.

4. Unidades de Análisis y de Observación

Las unidades de análisis son los hogares tanto de las áreas urbanas como rurales y las personas que habitan dentro del hogar.

Las unidades de observación son las viviendas de las áreas urbanas y rurales, los hogares que habitan dentro de ellas y las personas. Definiéndose como hogar el conjunto de personas, emparentadas o no entre sí que residen habitualmente en una vivienda y se constituye o agrupan para atender sus necesidades vitales u otras necesidades.

5. Marco Muestral y Unidades Muestrales

El marco muestral está conformado por el listado actualizado de todas las viviendas ocupadas del territorio nacional, urbano y rural, incluyendo las Islas Galápagos, que proviene de la actualización cartográfica del censo 2010.

Las unidades primarias de muestreo (UPM) son los sectores censales, las unidades secundarias son las viviendas. Las UPM's son seleccionadas con probabilidad proporcional al tamaño y las viviendas dentro de la UPM son seleccionadas mediante un muestreo aleatorio simple considerando que todas las viviendas dentro del sector tienen igual probabilidad de ser seleccionada.

6. Delimitación del Área Urbana

Se define como área urbana a los centros poblados con una población de 2.000 y más habitantes conforme a las recomendaciones de la Comunidad Andina de Naciones, sin importar si son cabeceras administrativas o no.

Debido a que se trabajará con el marco muestral del 2010, la definición de área urbana de 2.000 habitantes y más se construirá a partir de los resultados del precenso o actualización cartográfica.

7. Dominios de estudio

Los dominios de estudio también denominados dominios de estimación los podemos observar en el siguiente cuadro:

Tabla 1.- Dominios autorepresentados para la ENIGHUR 2010

N	Provincia		Dominios Autorepresentados	
		Ciudad	Provincial Urbano	Total provincial
1	Azuay	Cuenca	Azuay Urbano	Todo Azuay
2	Bolívar		Bolívar Urbano	Todo Bolívar
3	Cañar		Cañar Urbano	Todo Cañar
4	Carchi		Carchi Urbano	Todo Carchi
5	Cotopaxi		Cotopaxi Urbano	Todo Cotopaxi
6	Chimborazo		Chimborazo Urbano	Todo Chimborazo
7	El Oro	Machala	El Oro Urbano	Todo El Oro
8	Esmeraldas	Esmeraldas	Esmeraldas Urbano	Todo Esmeraldas
9	Guayas	Guayaquil	Guayas Urbano	Todo Guayas
10	Imbabura		Imbabura Urbano	Todo Imbabura
11	Loja	Loja	Loja Urbano	Todo Loja
12	Los Ríos		Los Ríos Urbano	Todo Los Ríos
13	Manabí	Manta	Manabí Urbano	Todo Manabí
14	Morona Santiago		Morona Santiago Urbano	Todo Morona Santiago
15	Napo		Napo Urbano	Todo Napo
16	Pastaza		Pastaza Urbano	Todo Pastaza
17	Pichincha	Quito	Pichincha Urbano	Todo Pichincha
18	Tungurahua	Ambato	Tungurahua Urbano	Todo Tungurahua
19	Zamora Chinchipe		Zamora Chinchipe Urbano	Todo Zamora Chinchipe
20	Galápagos		Galápagos Urbano	Todo Galápagos
21	Sucumbíos		Sucumbíos Urbano	Todo Sucumbíos
22	Orellana		Orellana Urbano	Todo Orellana
23	Santa Elena		Santa Elena Urbano	Todo Santa Elena
24	Santo Domingo de los Tsáchilas	Santo Domingo	Santo Domingo de los Tsáchilas Urbano	Todo Santo Domingo de los Tsáchilas

En total son 9 ciudades autorepresentadas, y 48 dominios provinciales.

Se los llama dominios auto-representado, debido a que la estimación de indicadores resultantes de ésta encuesta es estadísticamente representativa en cada uno de estos dominios.

Además de los dominios mencionados también tenemos los estratos socioeconómicos en las 9 ciudades autorepresentadas y a nivel Nacional Urbano, lo que aumenta a 27 dominios en las ciudades, así por ejemplo tenemos Quito Estrato Alto, Quito Estrato Medio y Quito estrato Bajo, igualmente para las 9 ciudades.

Los dominios se pueden resumir de la siguiente manera:

- 24 capitales provinciales estratificadas (3 estratos): 72 dominios
- Provincial Urbano y total: 48 dominios
- 1 dominio rural con asignación provincial
- En total hay 121 dominios autorepresentados

Productos calculables por agregación de los dominios

- 3 estratos socioeconómicos a nivel nacional
- Restos urbanos estratificados
- Regiones geográficas
- Regiones administrativas de planificación estratificadas
- Resto país Urbano y rural
- Nacional urbano rural y total

8. Estratificación

Tenemos un marco 2001 estratificado con un criterio socioeconómico, pero debido a que se trabajará con el marco actualizado se requiere la cartografía estratificada del censo 2010, para poder tener una buena estimación de estratos en el marco actualizado necesariamente tenemos que pasar por el censo 2010, pero debido a que la ENIGHUR 2010 empezará antes de que se tenga los resultados definitivos se tomará la estratificación propuesta por la consultoría de David Meguill para el Marco 2001.

En este marco, definir los estratos en el marco muestral 2.010, esto es casi imposible, la estrategia es utilizar los estratos del marco 2.001, como éstos están conectados mediante un código con la nueva cartografía, podemos ver la correspondencia de sectores y sus estratos de la cartografía anterior con la nueva, de esta forma hacemos la selección de sectores en la nueva cartografía de acuerdo al diseño muestral definido en el primer paso. Por último para llegar a una validación de los estratos se realizará un proceso de verificación, que consiste en visitar todos los sectores seleccionados (con un 35% de sectores adicionales), con la finalidad de hacer correcciones en caso de haber cambios de estratos en los sectores, para esta etapa de verificación, se diseñará un formulario con algunas preguntas claves que serán llenadas por observación por el verificador y que finalmente a partir de un modelo matemático se tendría la clasificación definitiva de sectores.

Esta estratificación ayudará a tener un diseño muestral estratificado, y es útil además para dispersar y optimizar la muestra, posteriormente luego del levantamiento de la información es necesario correr un proceso de validación o post-estratificación con la finalidad de tener los estratos definitivos para la generación de canastas por estratos.

9. Tamaño de la Muestra

El tamaño de muestra queda determinado tomando en consideración varios factores. Los 3 factores más importantes son: (1) el grado de precisión (confiabilidad) deseado en las estimaciones; (2) el costo y las limitaciones operacionales y (3) la eficiencia del diseño.

Para poder calcular los tamaños de muestra por estimación de medias, se generó los gastos medio por hogar para los 12 rubros de gastos del IPC, que se presentan a continuación:

Nº	VARIABLE	NOMBRE DE LA VARIABLE
1	galimone	01 ALIMENTOS - MONETARIOS
2	gbebmone	02 BEBIDAS ALCOHOLICAS, TABACO Y ESTUP. - MONETARIO
3	gvesmone	03 PRENDAS DE VESTIR Y CALZADO - MONETARIO
4	galomone	04 ALOJAMIENTO, AGUA, ELECTRIC, GAS. - MONETARIO
5	gmuemone	05 MUEBLES Y ENSERES - MONETARIO
6	gsalmone	06 SALUD - MONETARIO
7	gramone	07 TRANSPORTE - MONETARIO
8	gcommone	08 COMUNICACIONES - MONETARIO
9	grecmone	09 RECREACION Y CULTURA - MONETARIO
10	gedumone	10 EDUCACION - MONETARIO
11	ghotmone	11 HOTELES Y RESTAURANTES - MONETARIO
12	gbiemone	12 BIENES Y SERVICIOS DIVERSOS - MONETARIO

De todas las variables anteriores, se realizó un análisis descriptivo teniendo el gasto en alimento monetario la varianza más grande en las ciudades, por lo tanto se la seleccionó como variable de diseño para el cálculo del tamaño de la muestra para poder cubrirnos en la cantidad de la muestra.

Determinación del tamaño de la muestra

Sea X la variable de interés, N el tamaño de la población conocida y con el $(1-\alpha)$ % se fija el error del muestreo que es igual a $e = |\bar{x} - \mu|$, nos interesa conocer dicho error según las diferentes características poblacionales a estimar.

Para la determinación del tamaño de la muestra se utilizó la fórmula siguiente:

$$n = \frac{Z_{\alpha}^2 \sigma^2}{e^2} * deff^{1/2}$$

Z_{α}^2 = valor "z" para un nivel de confiabilidad deseada en la distribución normal de

Probabilidades igual a 1,96

σ^2 = la varianza del estimador (gasto en alimento)

e = error relativo máximo esperado

n = Tamaño muestral

$$deff^{1/2} = \sqrt{\text{efecto de diseño}} = \text{factor de diseño}$$

Al aplicar las ecuaciones obtuvimos que:

Este es el resultado después de trabajar con un error de diseño del 5%, con un 95% de confianza, El tamaño de muestra es de 40.932 viviendas que corresponden a 3.411 sectores en los que se investiga 12 viviendas por sector.

La Distribución Geográfica de la Muestra queda de la siguiente manera:

Ciudades/provincias	Total de Sectores	Sectores Urbanos	Sectores Rurales	Total de Viviendas
Cuenca	143			1.716
Resto Azuay	81	18	63	972
Guayaquil	199			2.388
Resto Guayas	150	66	84	1.800
Quito	178			2.136
Resto Pichincha	118	34	84	1.416
Machala	131			1.572
Resto El Oro	63	34	29	756
Esmeraldas	150			1.800
Resto Esmeraldas	80	34	46	960
Loja	146			1.752
Resto Loja	65	18	47	780
Santo Domingo	132			1.584
Resto Santo Domingo	13		13	156
Manta	139			1.668
Resto Manabi	150	66	84	1.800
Ambato	152			1.824
Resto Tungurahua	72	18	54	864
Bolívar	81	58	23	972
Cañar	92	66	26	1.104
Carchi	83	66	17	996
Cotopaxi	107	58	49	1.284
Chimborazo	114	66	48	1.368
Imbabura	97	66	31	1.164
Los Ríos	120	66	54	1.440
Santa Elena	65	50	15	780
Napo	71	58	13	852
Pastaza	71	58	13	852
Zamora Chinchipe	71	58	13	852
Sucumbíos	78	58	20	936
Orellana	72	58	14	864
Morona Santiago	75	58	17	900
Galápagos	52	32	20	624
Total	3.411	1.164	877	40.932

10. Diseño y selección de la muestra

El tipo de muestreo utilizado es: probabilístico bietápico para el caso de las ciudades auto-representadas, donde la unidad primaria de selección es el sector censal (UPM) y la unidad secundaria es la vivienda (USM); mientras que es probabilístico trietápico para el resto de dominios, la primera unidad de selección es la ciudad (UPM), la segunda es el sector censal (USM) y la tercera es la vivienda (UTM).

