

Actualización del Diseño Muestral ENIGHUR



DINEM

Diciembre, 2024

1. Dominios de Estudio

El dominio de diseño de la encuesta consiste en una subpoblación que puede ser identificada en el marco de muestreo y, por lo tanto, puede ser manejada independientemente en el tamaño de muestra, procedimientos de muestreo y demás.

El levantamiento de información de la ENIGHUR está planificada para 13 periodos, cada periodo está conformado de 4 semanas, por lo que, el levantamiento está previsto para un periodo de 52 semanas y la muestra está distribuida espacial y temporalmente respetando los periodos indicados.

La ENIGHUR para cumplir con su objetivo considera dos grupos de dominios, los cuales son:

- **Grupo 1:** considera las 24 provincias del Ecuador, las cuales constituyen el primer nivel de desagregación geográfico según el “Clasificador Geográfico Estadístico 2010 - Esquema de codificación de la División Político Administrativa del país”.
- **Grupo 2:** considera 9 cantones autorepresentadas que son de interés debido a su comportamiento económico.

Tabla 1. Dominios de estudio

N	Provincia	Cantón
1	Resto Azuay	Cuenca
2	Bolívar	
3	Cañar	
4	Carchi	
5	Cotopaxi	
6	Chimborazo	
7	Resto El Oro	Machala
8	Resto Esmeraldas	Esmeraldas
9	Resto Guayas	Guayaquil
10	Imbabura	
11	Resto Loja	Loja
12	Los Ríos	
13	Resto Manabí	Manta
14	Morona Santiago	
15	Napo	
16	Pastaza	
17	Resto Pichincha	Quito (DMQ)
18	Resto Tungurahua	Ambato
19	Zamora Chinchipe	
20	Galápagos	
21	Sucumbíos	

22	Orellana	
23	Santa Elena	
24	Resto Santo Domingo de los Tsáchilas	Santo Domingo

Fuente: Gestión de diseño muestral – DINEM

2. Marco de Muestreo

El Marco de muestreo a utilizar para la ENIGHUR es el Marco Maestro de Muestreo para encuestas de hogares del INEC (MMM), el cual se presenta como una lista organizada en forma de base de datos que contiene las viviendas registradas en el precenso del VIII Censo de Población y VII de Vivienda del Ecuador (CPV - 2022).

3. Diseño Muestral

3.1. Tipos y etapas del diseño muestral

El diseño muestral implementado en la ENIGHUR es un muestreo probabilístico bietápico estratificado de elementos. En la primera etapa, se selecciona una muestra estratificada de UPM con probabilidad proporcional al tamaño (PPT), donde la medida de tamaño de cada UPM está dada por el total de viviendas particulares ocupadas. Luego, se enlista la totalidad de viviendas que conforman cada UPM para, en una segunda etapa, seleccionar aleatoriamente un total fijo de 12 viviendas en cada UPM seleccionada.

3.2. Tamaño de la Muestra

Para poder calcular los tamaños de muestra por estimación de medias, se generó el gasto promedio por hogar para el rubro de gastos en ALIMENTOS - MONETARIOS, con el fin de obtener las estimaciones a partir de la información de la encuesta ENIGHUR 2012. Dichas estimaciones serán el insumo para el cálculo del tamaño de muestra.

Para calcular un tamaño de muestra que responda a todas los requerimientos planteados se sigue el siguiente esquema:

1. Se define como variable de diseño los gastos ALIMENTOS - MONETARIOS
2. Se calcula un tamaño de muestra independiente para cada dominio de diseño.
3. El tamaño de muestra anterior se transforma a tamaño de muestra de UPM.

4. El tamaño de muestra por UPM se distribuye para todos los estratos presentes en el marco. De ser necesario, se aumenta el tamaño de muestra de UPM por estrato a 2, para asegurar el cálculo de los estimadores de varianza resultantes de la encuesta.

3.2.1. Tamaño de muestra de viviendas

El tamaño de muestra de la ENIGHUR se obtuvo mediante la siguiente ecuación.

$$n_k = \frac{n_{0,k}}{1 + \frac{n_{0,k}}{N_k}} * \frac{1}{1 - tnr}$$

Donde:

$$n_0 = \frac{Z_{1-\frac{\alpha}{2}}^2 S^2}{\varepsilon^2 \mu^2}$$

$$S^2 = \sigma^2 DEFF$$

- n_k : El tamaño de la muestra de viviendas para el dominio de estudio k .
- N_k : El tamaño de la población en el dominio k .
- μ_k : La media poblacional de la variable de interés en el dominio k .
- σ_k^2 : La varianza poblacional de la variable de interés en el dominio k .
- $DEFF_k$: El efecto del diseño muestral en el dominio k .
- α : Nivel de confianza $\alpha = 0,95$.
- tnr : Tasa de no respuesta.

3.2.2. Efecto de diseño (DEFF)

Con el objetivo de poder tener un acercamiento a los parámetros reales relacionados a los efectos de diseño para cada uno de los dominios de estudio, se elaboró una reconstrucción de los estratos en la base de cobertura ENIGHUR-2012. Dichos estratos se elaboraron considerando la parte geográfica subdivida en áreas urbanas (estrato 1) y rurales (estrato 2). Mientras que para las provincias que cuentan con cantones auto-representadas se consideró el cantón (estrato 3), el resto del área urbana (estrato 1) y el resto del área rural (estrato 2).

3.2.3. Nivel de confianza

Con una puntuación estándar Z_{α}^2 = valor "z" para un nivel de confiabilidad deseada en la distribución normal con probabilidad igual, se fija una puntuación de 1,96.

3.2.4. Margen de error

Para el cálculo del margen de error relativo (MER) en cada uno de los dominios de estudio se tomó como referencia el margen de error obtenido en las estimaciones de la ENIGHUR-2012, con lo que se definió un MER entre el 3% y 4% para las cantones auto-representadas, para el caso de los dominios ubicados en la amazonia se definió un MER entre 3% y 10% y un valor de 14% para Morona Santiago, siendo este el máximo MER, mientras que para el resto de provincias el MER varía entre el 2,5% y 10%.

3.2.5. Tasa de no respondientes (TNR)

Esta tasa permite conocer la proporción de viviendas elegibles no efectivas para el total de viviendas visitadas. En este sentido, la tasa de no respuesta para la encuesta fue calculada a partir de la base de cobertura de la ENIGHUR-2012 y finalmente definida como el promedio de las TNR's de cada dominio de estudio, cuyo resultado se aproxima al 20%, resultado aproximado de la TNR promedio por provincia.

$$TNR = 1 - \frac{N^{\circ} \text{ viviendas efectivas}}{N^{\circ} \text{ de viviendas seleccionadas}}$$

3.2.6. Tamaño de muestra de UPM

Es necesario calcular el número de UPM que deben ser seleccionadas en la primera etapa de muestreo a partir de la relación:

$$n_{UPMk} = \frac{\text{Numero de viviendas calculadas}}{\text{Carga tecnica operativa}}$$

Para la presente operación estadística se definió operativamente que el número de viviendas a investigarse por UPM será doce (12).

3.2.7. Asignación de la muestra

La muestra es distribuida proporcionalmente por cada estrato dentro de cada dominio de estudio.

$$n_{ki} = n_{UPMk} * \frac{N_{ki}}{N_k}$$

Donde:

- n_{ki} : Tamaño de la muestra para el estrato i del dominio de estudio k .
- n_{UPMk} : Tamaño de la muestra para el dominio de estudio k .
- N_{ki} : Total de viviendas ocupadas en el estrato i del dominio de estudio k .
- N_k : Total de viviendas ocupadas en el dominio de estudio k .

En los estratos, donde $n_{ki} = 1$ se aproxima a mínimo muestral $n_{ki} = 2$.

Es importante mencionar que antes de la distribución proporcionalmente por cada estrato en cada dominio de estudio, a los tamaños calculados se realizan un ajuste. Este tipo de ajuste responde estrictamente a solventar requerimientos referentes a la parte operativa de la encuesta, por lo que, se busca mantener un tamaño muestral por zonal para garantizar que el número de encuestadores y supervisores estén acordes a lo planificado y presupuestado por el equipo encargado, sin que afecte la calidad del diseño muestral.

3.2.8. Distribución espacial y temporal de la muestra

Con el objetivo de definir claramente el levantamiento de información de cada UPM en cada dominio de estudio, se realiza una distribución en espacio y tiempo para definir los periodos y semanas en la que se debe visitar cada UPM.

Se considera levantar la misma cantidad de UPM en cada periodo. La asignación de la muestra queda definida con un total de 264 UPM en cada periodo y en cada dominio de estudio se busca mantener una cantidad similar de UPM a levantar entre cada periodo. Esto garantiza que la encuesta no se concentre en determinados periodos y esté distribuida equitativamente a lo largo de toda la línea de tiempo que dura la operación estadística.

3.2.9. Resultados del tamaño muestral

Una vez realizado el proceso de asignación y ajuste muestral mencionados en el aparatado anterior, los resultados del tamaño muestral para cada dominio son los siguientes:

Tabla 2. Tamaño muestral por dominio

Dominio	Nombre del dominio	Tamaño (UPM)
01	Azuay	69
02	Bolívar	95
03	Cañar	92
04	Carchi	83
05	Cotopaxi	97
06	Chimborazo	104
07	El Oro	65
08	Esmeraldas	77
09	Guayas	126
10	Imbabura	95
11	Loja	74
12	Los Rios	111
13	Manabí	150
14	Morona Santiago	98
15	Napo	66

16	Pastaza	83
17	Pichincha	144
18	Tungurahua	81
19	Zamora Chinchipe	94
20	Galápagos	104
21	Sucumbíos	76
22	Orellana	67
23	Santo Domingo de los Tsachilas	27
24	Santa Elena	60
33	Quito	164
42	Guayaquil	210
30	Cuenca	131
40	Machala	124
34	Ambato	135
41	Esmeraldas Ciudad	141
44	Santo Domingo	120
43	Manta	132
32	Loja Ciudad	137

Fuente: Gestión de diseño muestral – DINEM

3.3. Observaciones

- Es importante señalar que el actual documento recoge información de los procesos implementando hasta el momento, mas no de los resultados finales; por lo que, existirán cambios a medida que la operación estadística avance y con ello se podrá elaborar un documento final en el que se detalle toda la metodología utilizada.

FIRMAS DE RESPONSABILIDAD		
Nombre y Apellido	Dirección	Firmas
Elaborado por: OMAR LLAMBO	DINEM	
Revisado y aprobado por: CHRISTIAN GARCÉS	DINEM	

