



**ENCUESTA
NACIONAL SOBRE
RELACIONES
FAMILIARES Y
VIOLENCIA DE
GÉNERO CONTRA
LAS MUJERES
ENVIGMU**

— Noviembre - 2019 —



Buenas cifras,
mejores vidas

Dirección de Infraestructura Estadística y Muestreo

Metodología del Diseño Muestral de la Encuesta Nacional sobre Relaciones Familiares de las Mujeres

Año 2019

Quito - Ecuador, 2019

Dirección

Dirección de Infraestructura Estadística y Muestreo

Unidad

Gestión de Diseño Muestral

Elaborado por:

Ángel Gaibor
Salomé Borja
Pilar Carrillo

Revisado por:

Javier Núñez

Aprobado por:

Christian Garcés

Tabla de contenido

INTRODUCCIÓN	6
ANTECEDENTES	7
1 PLANIFICACIÓN	9
1.1 OBJETIVOS Y DELIMITACIÓN DE LA OPERACIÓN ESTADÍSTICA	9
1.1.1 OBJETIVOS	9
<i>Objetivo General</i>	9
<i>Objetivos Específicos</i>	9
1.1.2 DELIMITACIÓN	9
<i>Tipo de operación estadística</i>	9
<i>Universo de estudio</i>	9
<i>Población objetivo</i>	9
<i>Unidad de observación y análisis</i>	9
<i>Cobertura geográfica</i>	9
2 DISEÑO Y CONSTRUCCIÓN	10
2.1 DISEÑAR EL MARCO Y LA MUESTRA	10
2.1.1 TIPO Y ETAPAS DE DISEÑO MUESTRAL	10
2.1.2 DOMINIOS DE ESTUDIO	10
2.1.3 MARCO MUESTRAL	11
2.1.4 ESTRATIFICACIÓN DEL MARCO DE MUESTREO	12
2.1.5 TAMAÑO MUESTRAL	12
2.1.6 SELECCIÓN DE LA MUESTRA	14
2.1.7 CONTROL DE COBERTURA DE CAMPO Y MUESTRAL	14
2.1.7.1 <i>Cobertura muestral por viviendas</i>	14
2.1.7.2 <i>Cobertura muestral por mujeres esperadas</i>	15
2.1.8 FACTORES DE EXPANSIÓN	16
2.1.8.1 <i>Ponderación de selección de la primera etapa (UPM)</i>	16
2.1.8.2 <i>Tasa de no respuesta de la primera etapa de selección</i>	17
2.1.8.3 <i>Ponderación de selección de la segunda etapa (viviendas)</i>	17
2.1.8.4 <i>Ajuste de no respuesta de viviendas</i>	18
2.1.8.5 <i>Ajuste de no elegibilidad de viviendas</i>	18
<i>Dado que la tasa de respuesta a nivel de vivienda se calcula a nivel de estrato, todas las UPM del mismo estrato comparten la misma corrección de cobertura</i>	18
2.1.8.6 <i>Ponderación de tercera etapa (mujer de 15 años y más)</i>	18
2.1.8.7 <i>Ajuste por no respuesta a nivel de mujer de 15 años y más</i>	19
2.1.8.7 <i>Factores de expansión ENVIGMU 2019</i>	19

2.1.9. POSTESTRATIFICACIÓN	20
2.2.0 ESTIMACIÓN DE CARACTERÍSTICAS	20
2.2.1 ERROR DE MUESTREO.....	21
2.2.2. MÉTODOS DE ESTIMACIÓN DE ERRORES PARA DISEÑOS MUESTRALES COMPLEJOS	21
BIBLIOGRAFÍA	23

Lista de tablas

Tabla 1. Tabla de KISH para la selección de una mujer dentro del hogar	10
Tabla 2. Variables de diseño.....	12
Tabla 3. Distribución de la muestra por Provincia	13
Tabla 4. Seguimiento de cobertura muestral por viviendas	15
Tabla 5. Seguimiento de cobertura muestral según mujeres esperadas	16
Tabla 6. Variables requeridas para la declaración del diseño muestral.....	22

INTRODUCCIÓN

El Instituto Nacional de Estadística y Censos – INEC, como ente rector de la estadística nacional es el encargado de llevar a cabo una serie de encuestas que alimenten a la producción estadística del país, en esta ocasión se desarrolla la “Encuesta Nacional sobre Relaciones Familiares de las Mujeres, “ENVIGMU 2019”.

La Planificación nacional considerando como prioritario realizar esta investigación, y el INEC como ente rector de la estadística nacional ponen en marcha el desarrollo de la misma, donde el Diseño Muestral es parte del Modelo de Producción Estadística. En este sentido, la Dirección de Infraestructura Estadística y Muestreo en la creación del diseño muestral para la ENREFAM 2019, tiene como objetivo realizar un diseño muestral de calidad para actualizar la información producida en el año 2011, con el fin de aportar información para la toma de decisiones en política pública.

El proceso que se realiza para la generación del diseño muestral consta de tres etapas:

En la primera etapa se toma en cuenta:

- Determinación del marco de muestreo
- El universo de la investigación
- Las unidades de muestreo y análisis
- Cobertura geográfica

En la segunda etapa:

- Metodología para el diseño muestral
 - Determinación de los dominios de estudio
- Determinación del tamaño de la muestra
- Selección de la muestra
- Validación de la cobertura obtenida en campo

Por último, se realiza:

- El cálculo de los factores de expansión
- Estimación de características
- Error de muestreo
- Declaración del plan de muestreo

ANTECEDENTES

En América Latina y el Caribe, hasta los años noventa, la violencia contra la mujer acaecida en el ámbito familiar era considerada un asunto privado en el cual el Estado no debía intervenir. Esta concepción se expresaba en la invisibilidad del problema tanto en los ámbitos legislativos, judicial y ejecutivo, como en otros sectores de la sociedad. El Ecuador no era una excepción, de manera que ninguna mujer que sufría violencia tenía posibilidad de denunciarla, pues una disposición del Código de Procedimiento Penal prohibía la denuncia entre cónyuges o entre ascendientes o descendientes; y, además, esta violencia no se encontraba especificada en las leyes del país.

En 1994 la Comisión Interamericana de Mujeres de la Organización de Estados Americanos, junto a varias organizaciones y redes latinoamericanas, consiguió la promulgación de la "Convención Interamericana para prevenir, sancionar y erradicar la violencia contra la mujer" (Belém do Pará), la misma que ha sido ratificada por casi todos los países de la región, entre ellos, el Ecuador. Dicho marco legal obliga a los gobiernos a implementar y dar seguimiento a leyes y políticas para prevenir, atender y sancionar la violencia contra las mujeres.

Organizaciones de la sociedad civil como el Centro de Planificación y Estudios Sociales (CEPLAES), Comité Ecuatoriano de Cooperación con la Comisión Interamericana de Mujeres (CECIM), Fundación María Guare, Centro de Estudios e Investigaciones Multidisciplinarias del Ecuador (CEIME) mediante investigaciones locales o con determinados grupos poblacionales han producido información que sirvió para mostrar la gravedad y la complejidad del problema de la violencia y sustentar la necesidad de la acción estatal.

En el país también se utilizaron los registros de los servicios, particularmente los de las Comisarías de la Mujer y la Familia (CMF), para a partir del número de denuncias hechas por las mujeres en estas instancias, mostrar la magnitud del problema. Sin embargo, durante muchos años sólo se contó con datos parciales o sobre pequeños períodos de tiempo, recogidos o sistematizados por ONG de mujeres interesadas en el tema. A pesar de que hubo varias iniciativas desde el Consejo Nacional de las Mujeres (CONAMU) y de la Dirección Nacional de Comisarías de la Mujer, posteriormente constituida como Dirección Nacional de Género (DINAGE), por muchos años no se logró poner en marcha un sistema de registro e información, mediante el cual las CMF provean datos estadísticos que permitan caracterizar mejor el problema, evaluar la respuesta estatal o definir nuevas iniciativas o políticas para erradicar esta práctica.

Entre 1999 y 2004, la DINAGE únicamente cuenta con información sobre el número de denuncias atendidas en cada una de las CMF; y a partir de 2005 dispone de información sobre el número de denuncias presentadas por mujeres y hombres, como también sobre la respuesta institucional, entre las que se destacan: número de citaciones y de audiencias realizadas, número de las distintas medidas de amparo emitidas, de indemnizaciones o pensiones de subsistencia dispuestas, de sentencias o resoluciones absolutorias dictadas. Contar con esta información ha sido sumamente importante y ha contribuido a dar seguimiento a la actuación de las CMF y a la toma de medidas para mejorar su desempeño.

La única información estadística sobre la prevalencia de la violencia en contra de las mujeres que se ha producido a nivel nacional, es la proporcionada por la Encuesta Demográfica y de Salud Materna e Infantil (ENDEMAIN), realizada en 1999 por una entidad privada Centro de Estudios de Población y Desarrollo Social (CEPAR) que, desde un enfoque de salud, incluyó algunas preguntas sobre la violencia sufrida por las mujeres. En 2004 éstas fueron ampliadas y la encuesta incluyó un módulo sobre el tema y produjo un documento de análisis de sus resultados en 2006. Si bien es innegable el aporte de la información arrojada por la ENDEMAIN, la información producida tiene algunas limitaciones.

En forma concomitante a la visibilidad y reconocimiento social que ha alcanzado la violencia contra la mujer en el Ecuador, la legislación nacional se ha ido modificando de forma progresiva, de manera que en la actualidad, se cuenta con un marco normativo que garantiza el ejercicio a una vida libre de violencia para las mujeres, pero también con una ley específica sobre la violencia que acaece al interior de la familia (Ley 103)

Y es así que en el año 2007, se promulga el Decreto Ejecutivo N° 620, que declara como política de Estado la erradicación de la violencia de género hacia la niñez, adolescencia y mujeres. Para ejecutarlo se formula ese mismo año el "Plan nacional de erradicación de la violencia de género", en el marco del cual se realiza la presente encuesta.

Posteriormente, en el año 2011 el INEC ejecutó la Encuesta Nacional sobre Violencia de Género contra las Mujeres con una desagregación nacional urbano, rural y por regiones naturales. La encuesta abarcó por un lado la contextualización de las mujeres encuestadas y del hogar en que viven (Aspectos socioeconómicos y demográficos que se realizó de acuerdo con la metodología, los criterios y formatos definidos por el INEC); y por otro, la prevalencia e incidencia de las distintas formas de violencia (psicológica, física, sexual y patrimonial) en diversos ámbitos y sus perpetradores con el fin de que esta encuesta constituya un instrumento para la ejecución del Plan Nacional de Erradicación de la Violencia de Género hacia la Niñez, Adolescencia y Mujeres que se ejecuta en el Ecuador.

En cuanto a la planificación del diseño muestral, en el año 2016 se inició con las actividades previas de la fase de planificación estimando el tamaño de la muestra con propósitos presupuestales, sin embargo, por razones ajenas a diseño muestral se retomó las siguientes fases de la encuesta en el año 2018 , concluyendo el proceso en el año 2019.

En este contexto, se desarrolla el diseño muestral de la (ENVIGMU) 2019, con la finalidad de obtener datos acerca de la evolución de los tipos de violencia de género que sufren o que han sufrido las mujeres, tanto en el ámbito público como en el privado, y finalmente, contribuir a la formulación de las políticas públicas.

1 PLANIFICACIÓN

1.1 Objetivos y delimitación de la operación estadística

1.1.1 Objetivos

Objetivo General

Elaborar el Diseño Muestral de la “Encuesta Nacional sobre Relaciones Familiares de las Mujeres, Julio/Agosto 2019”, utilizando técnicas de muestreo probabilístico que permitan emitir resultados confiables de las principales variables que se investigarán en dicha encuesta.

Objetivos Específicos

- Garantizar la representatividad de la encuesta tanto a nivel geográfico como de los dominios de estudio.
- Calcular factores de expansión y errores asociados acordes al diseño muestral elaborado.

1.1.2 Delimitación

Tipo de operación estadística

Encuesta por muestreo

Universo de estudio

El total de mujeres de 15 años y más en el territorio ecuatoriano, residentes en los hogares del Ecuador, exceptuando la población temporal en viviendas colectivas, viviendas flotantes y población indigente (sin techo). Este universo es ajustado de acuerdo a las proyecciones poblacionales a nivel de sexo y grupo de edad para cada dominio de estudio.

Población objetivo

Mujeres de 15 años y más

Unidad de observación y análisis

Las unidades de observación son las viviendas ocupadas de las áreas urbanas y rurales, con los hogares que habitan dentro de ellas y las personas.

Las unidades de análisis son todas las mujeres miembro de hogar de 15 años y más, tanto de las áreas urbanas como rurales, a nivel nacional.

Cobertura geográfica

La cobertura geográfica está definida por las viviendas ocupadas que se encuentren ubicadas dentro del territorio ecuatoriano incluyendo la región Insular.

2 DISEÑO Y CONSTRUCCIÓN

2.1 Diseñar el marco y la muestra

2.1.1 Tipo y etapas de diseño muestral

El diseño muestral para esta investigación corresponde a un muestreo probabilístico con tres etapas de selección:

- **Primera etapa:** selección aleatoria de Unidades Primarias de Muestreo (UPM) por estrato.
- **Segunda etapa:** selección de 8 viviendas ocupadas dentro de cada uno de las UPM seleccionadas en la primera etapa.
- **Tercera etapa:** selección de una mujer de 15 años y más dentro del hogar.

Para la selección de la mujer de 15 años y más, se utilizó la tabla de Kish, donde las filas vienen dadas por la letra que se asigna de manera aleatoria para cada vivienda y en las columnas el número de mujeres de 15 años residentes habituales del hogar seleccionado.

Tabla 1. Tabla de KISH para la selección de una mujer dentro del hogar

LETRA DE LA VIVIENDA	NÚMERO DE MUJERES DE 15 AÑOS Y MÁS MIEMBROS DEL HOGAR											
	1	2	3	4	5	6	7	8	9	10	11	12
A	1	1	1	1	1	1	1	1	1	1	1	1
B	1	1	1	1	1	1	1	1	2	2	2	2
C	1	1	1	1	2	2	2	2	2	3	3	3
D	1	1	1	2	2	2	2	3	3	4	4	4
E	1	1	2	2	3	3	3	3	4	5	5	5
F	1	1	2	2	3	3	4	4	4	5	6	6
G	1	2	2	3	3	4	5	5	5	6	7	7
H	1	2	2	3	4	4	5	6	6	7	8	8
I	1	2	3	3	4	5	6	6	7	7	9	9
J	1	2	3	4	4	5	6	7	8	8	9	10
K	1	2	3	4	5	6	7	8	9	9	10	11
L	1	2	3	4	5	6	7	8	9	10	11	12

2.1.2 Dominios de estudio

Los dominios de estudio de la encuesta están definidos geográficamente a nivel de Provincias:

01. Azuay
02. Bolívar
03. Cañar
04. Carchi
05. Cotopaxi
06. Chimborazo
07. El Oro
08. Esmeraldas
09. Guayas
10. Imbabura
11. Loja
12. Los Ríos
13. Manabí

14. Morona Santiago
15. Napo
16. Pastaza
17. Pichincha
18. Tungurahua
19. Zamora Chinchipe
20. Galápagos
21. Sucumbíos
22. Orellana
23. Santo Domingo
24. Santa Elena

La encuesta es representativa a nivel nacional urbano y nacional rural, por este motivo se tiene muestra en las zonas no delimitadas con el fin de completar la representatividad nacional.

2.1.3 Marco muestral

El Marco Maestro de Muestreo (MMM) para la “Encuesta Nacional sobre Relaciones Familiares de las Mujeres, 2019”, parte del Marco Maestro de Muestreo diseñado para el sistema integrado de Encuesta a Hogares -SIEH; el cual presenta como una lista organizada en forma de base de datos que contiene a las viviendas registradas en el CPV 2010, que participarán en cada una de las fases de diseño, distribución y selección de la muestra de una encuesta dirigida a hogares.

El MMM principalmente se emplea para identificar y seleccionar las unidades de muestreo y como base para realizar estimaciones basadas en los datos de la muestra; esto implica que la población a ser seleccionada para la muestra debe estar representada de forma física, es decir, el MMM también está formado por todos los mapas y planos a diferentes escalas que permiten identificar en forma precisa y clara los límites físicos que tienen las diferentes unidades de selección, considerándose como parte principal de éste los registros y listados en los que se detalla las referencias que faciliten identificar en forma exacta las viviendas seleccionadas.

El MMM contiene información sobre la división político-administrativa y geográfica del país-subdivisiones políticas o zonificación estadística definida para efectuar la enumeración del censo de población-, como también sobre los volúmenes de viviendas y de población total, por grupos de edad y sexo, entre otras variables necesarias para clasificar a los hogares de acuerdo a determinadas características según los objetivos específicos de la encuesta. Todos y cada uno de los elementos de los que está compuesto el MMM tienen una probabilidad conocida y diferente de cero de ser seleccionados de alguna de las muestras que se puedan extraer del mismo.

En resumen, el marco de la muestra se basa en los resultados definitivos y la cartografía del VII Censo de Población y VI de Vivienda del 2010 (CPV-2010), complementado con la información parcialmente actualizada por distintos procesos cartográficos que ha llevado el INEC en el período 2014-2017. Para más detalle dirigirse al documento de Marco Maestro de Muestreo para Encuestas a Hogares¹.

¹https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2018/Disenio_Muestral_2018/SIEH%20-MMM.pdf

2.1.4 Estratificación del Marco de Muestreo

La estratificación se refiere a la subdivisión de una población determinada en subconjuntos con características propias. Esta acción se lleva a cabo como una etapa previa a la selección de la muestra y la(s) variable(s) que se utiliza(n) para ello debe(n) contener información acerca de todas las unidades de la población.

Los objetivos del proceso de estratificación son los siguientes:

- Agrupar a las unidades de marco en grupos de acuerdo a un conjunto de características socioeconómicas.
- Formar grupos de UPM homogéneos en su interior y heterogéneos entre ellos.
- Mejorar el diseño muestral incrementando la eficiencia del mismo y controlando la varianza de los estimadores.

2.1.5 Tamaño muestral

Durante el desarrollo de la fase de Planificación de la Operación Estadística (2016), se determinó el tamaño de muestra de la encuesta considerando tres aspectos fundamentales: el efecto de diseño, el nivel de confianza y el error relativo, realizando ajustes al diseño a través del coeficiente de variación de las variables de diseño. Todos estos estimadores fueron calculados a partir de la Encuesta de Violencia de Género del 2011. En la Tabla 2 se detallan las variables de diseño utilizadas.

Tabla 2. Variables de diseño

Siglas	Variable de diseño
Vsi	Violencia psicológica
Vfi	Violencia física
Vse	Violencia sexual
Vpa	Violencia patrimonial
Vtl	Violencia total

Fuente: Encuesta de Violencia de Género, 2011

También se consideró una tasa de no respuesta del 5% y un nivel de confianza del 95%. Para el cálculo del tamaño de la muestra se utilizó la fórmula siguiente:

$$n = \frac{z^2 q Deff}{r^2 p(1 - tnr)} \cdot p(\geq 15m)$$

Dónde:

n	=	Tamaño de la muestra
p	=	Estimación de la proporción de interés
q	=	$1 - p$
r	=	Error relativo máximo esperado
z	=	Coficiente que representa el nivel de seguridad o confianza
$Deff$	=	Efecto de diseño
tnr	=	Tasa de no respuesta
$p(\geq 15m)$	=	probabilidad de encontrar una mujer de 15 años y más.

Es importante destacar que éste cálculo se realizó utilizando como insumo el marco de muestreo para encuestas de hogares vigente en el 2016, el cual estaba constituido por 40.628 unidades primarias de muestreo (UPM). Como resultado se obtuvo un tamaño

muestral de 20.228 viviendas distribuidas en 1.268 UPM en los que se planificó investigar 16 viviendas por UPM, en una primera propuesta.

En la tabla Tabla 3 se presenta el tamaño y la distribución de la muestra de ENREFAM 2018.

Posteriormente, para el año 2018 se retoma las actividades de la operación estadística, sin embargo para ese año se introdujo mejoras técnicas en cuanto a la constitución de UPM en el MMM. Cabe mencionar que en la actualidad el MMM consta de 107.014 UPM.

Debido a ésta reestructura, se revisó el tamaño de muestra y la distribución de la ENREFAM 2019 y se estableció:

- Número de UPM: 2.606.
- Número de viviendas seleccionadas por UPM: 8.
- Total de número de viviendas: 20.848.

Tabla 3. Distribución de la muestra por Provincia

Provincia	No. De UPM	No. De viviendas por UPM
Azuay	106	848
Bolívar	102	816
Cañar	98	784
Carchi	86	688
Cotopaxi	84	672
Chimborazo	80	640
El Oro	102	816
Esmeraldas	98	784
Guayas	184	1.472
Imbabura	114	912
Loja	104	832
Los Ríos	106	848
Manabí	98	784
Morona Santiago	82	656
Napo	82	656
Pastaza	98	784
Pichincha	172	1.376
Tungurahua	106	848
Zamora Chinchipe	84	672
Galápagos	74	592
Sucumbíos	88	704
Orellana	100	800
Santo Domingo	94	752
Santa Elena	256	2.048
Zonas no delimitadas	8	64
Total	2.606	20.848

Fuente: Elaboración propia, INEC 2018

En la planificación del levantamiento de información constan 8 viviendas originales y 3 viviendas de reemplazo, es decir, un total de 11 viviendas por UPM, lo que es igual a 28.666 viviendas entre originales y reemplazos. La prioridad es que sean efectivas las viviendas originales y no utilizar las viviendas de reemplazo, sin embargo, existen ciertas condiciones donde es necesario aplicar los reemplazos:

Cuando la vivienda está habitada (no respuesta):

1. Rechazo de la encuesta
2. Nadie en casa

Cuando la vivienda está deshabitada (no elegibilidad):

3. Vivienda Temporal
4. Vivienda desocupada
5. Vivienda en construcción
6. Vivienda inhabitable o destruida
7. Vivienda convertida en negocio

Estas últimas están dadas por la falta de actualización del MMM, lo que repercute en que las viviendas originales tengan una alta probabilidad de no ser efectivas, siendo necesario utilizar reemplazos de viviendas e incluso de UPM.

En la ENREFAM, se estableció un escenario mínimo de 6 viviendas efectivas siempre y cuando se hayan visitado el total de las viviendas entre originales y reemplazos (11). En el caso de existir menos de 6 viviendas efectivas, se realizaba una solicitud para reemplazar la UPM, con la condición de que el reemplazo generado debía garantizar las mismas condiciones de la UPM original.

Cabe mencionar que por inconvenientes cartográficos se realizaron 37 reemplazos de UPM. De igual manera, en la provincia de Santa Elena no se realizó un reemplazo de UPM, debido a que ya no se contaba con tiempo ni recursos logísticos para realizarlo, sin embargo, se verificó que no realizarlo tenía una afectación mínima para las estimaciones de la provincia, quedando un total de 2605 UPM.

2.1.6 Selección de la muestra

La selección de las UPM que formaron parte de la muestra se realizó de manera independiente en cada uno de los dominios de forma aleatoria. De la misma forma, la selección de viviendas es sistemática dentro de cada UPM seleccionada; y en la última unidad de muestreo, la mujer de 15 años y más se selecciona según la letra Kish de la vivienda a la que pertenece y el número de mujeres elegibles por hogar, de esta manera se garantiza igual probabilidad de ser seleccionada en cada una de las etapas.

2.1.7 Control de cobertura de campo y muestral

La cobertura hace referencia a la información obtenida en campo y proporcionada por los informantes, se divide en dos partes: viviendas y mujeres de 15 años y más.

Posteriormente, se realizan seguimientos de cobertura tanto para viviendas como para mujeres de 15 años y más, con el fin de verificar el total de viviendas efectivas y mujeres esperadas, ya que de esto dependerán las posteriores estimaciones en los indicadores.

2.1.7.1 Cobertura muestral por viviendas

Como se puede observar en la Tabla 4, en la ENREFAM 2019 se visitaron un total de 25.427 viviendas, entre originales y reemplazos, de las cuales 20.014 viviendas fueron efectivas, obteniendo una cobertura del 78,71% a nivel nacional. En cuanto al rechazo, nadie en casa y las viviendas no elegibles suman el 21,29% del total de las viviendas visitadas siendo, a nivel nacional, el 5% de nadie en casa y rechazo, y el 16,22% de viviendas no elegibles.

Las provincias que mayor porcentaje de rechazo o que no hubo nadie en casa son Pichincha (209), Azuay (140) y Guayas (132), las cuales suman el 37,37% del total de viviendas rechazadas. Por otro lado, se puede observar que las viviendas no elegibles están

distribuidas proporcionalmente entre todas las provincias, sin embargo, en Santa Elena (329), Pichincha (257), Guayas (234) y Orellana (212) se acumula el 25% del total de viviendas no elegibles, esto es debido a la desactualización del marco muestral lo que genera gastos de recursos humanos y económicos en campo.

Por último, cuando se habla de la cobertura final, se hace referencia a las efectivas sobre el total de visitadas, lo que permite conocer la cobertura por dominio de estudio.

Tabla 4. Seguimiento por cobertura muestral por viviendas

Provincias	Total de viviendas visitadas	Efectivas	Rechazo y nadie en casa	No elegibles	% Cobertura final
Azuay	1.094	780	140	174	71,30%
Bolívar	982	787	46	149	80,14%
Cañar	982	727	55	200	74,03%
Carchi	791	683	18	90	86,35%
Cotopaxi	832	638	32	162	76,68%
Chimborazo	795	609	18	168	76,60%
El Oro	976	782	47	147	80,12%
Esmeraldas	941	763	21	157	81,08%
Guayas	1.795	1.429	132	234	79,61%
Imbabura	1.087	903	40	144	83,07%
Loja	1.043	774	80	189	74,21%
Los Ríos	1.004	842	40	122	83,86%
Manabí	956	749	40	167	78,35%
Morona Santiago	838	603	37	198	71,96%
Napo	769	640	16	113	83,22%
Pastaza	960	754	43	163	78,54%
Pichincha	1.762	1.296	209	257	73,55%
Tungurahua	1.019	823	41	155	80,77%
Zamora Chinchipe	842	630	34	178	74,82%
Galápagos	718	577	23	118	80,36%
Sucumbíos	862	680	21	161	78,89%
Orellana	990	763	15	212	77,07%
Santo Domingo	930	718	83	129	77,20%
Santa Elena	2.383	2.000	54	329	83,93%
Zonas no delimitadas	76	64	2	10	84,21%
Totales	25.427	20.014	1.287	4.126	78,71%

Fuente: Elaboración propia, INEC 2019

2.1.7.2 Cobertura muestral por mujeres esperadas

En la tabla 5 se puede observar que existen 27.842 mujeres que son elegibles (15 años y más) dentro de los hogares a nivel nacional, sin embargo, por la naturaleza de la encuesta solo se seleccionó a una por hogar dando como resultado 19.161 mujeres seleccionadas, de las cuales 17.211 fueron efectivas, es decir, se obtuvo una efectividad del 89,82%.

Cabe mencionar, que la cobertura final hace referencia al total de mujeres efectivas sobre el total de mujeres seleccionadas, lo que permite conocer la cobertura por dominio de estudio.

Tabla 5. Seguimiento de cobertura muestral según mujeres esperadas

Provincias	Mujeres esperadas	Total de mujeres elegibles	Total de mujeres seleccionadas (y)	Total de mujeres efectivas (x)	% Cobertura final (x/y)
Azuay	786	1.171	787	663	84,24%
Bolívar	728	1.072	751	673	89,61%
Cañar	718	1.093	734	634	86,38%
Carchi	635	964	662	605	91,39%
Cotopaxi	620	926	625	568	90,88%
Chimborazo	579	783	567	514	90,65%
El Oro	722	1.049	742	636	85,71%
Esmeraldas	690	1.072	737	654	88,74%
Guayas	1.309	2.093	1.384	1.235	89,23%
Imbabura	841	1.271	870	782	89,89%
Loja	755	1.104	763	657	86,11%
Los Ríos	745	1.141	776	720	92,78%
Manabí	718	1.067	715	659	92,17%
Morona Santiago	582	803	580	493	85,00%
Napo	593	863	600	556	92,67%
Pastaza	701	1.016	728	640	87,91%
Pichincha	1.251	1.775	1.254	1.106	88,20%
Tungurahua	777	1.165	797	725	90,97%
Zamora Chinchipe	585	763	575	503	87,48%
Galápagos	484	630	490	460	93,88%
Sucumbíos	606	862	636	557	87,58%
Orellana	692	958	719	665	92,49%
Santo Domingo	678	1.011	684	640	93,57%
Santa Elena	1.849	3.098	1.919	1.805	94,06%
Zonas no delimitadas	57	92	66	61	92,42%
Totales	18.702	27.842	19.161	17.211	89,82%

Fuente: Elaboración propia, INEC 2019

2.1.8 Factores de expansión

El objetivo de una encuesta por muestreo es hacer inferencias acerca de la población a partir de la información contenida en la muestra. En este sentido, se utilizan factores de expansión para llevar la información de la muestra a niveles poblacionales.

El procedimiento de ponderación general para la ENREFAM requiere de tres etapas:

- La primera etapa consiste en calcular una ponderación de UPM. De ser necesario, se aplica un ajuste por no respuesta a nivel de UPM.
- La segunda etapa consiste en calcular una ponderación de viviendas dentro de la UPM. De ser necesario, se aplica un ajuste por no respuesta a nivel de vivienda.
- La tercera etapa consiste en calcular una ponderación de mujer. De ser necesario se aplica un ajuste por no respuesta a nivel de mujer a la ponderación de mujer. La ponderación de mujer se calcula de manera independiente para cada hogar.

2.1.8.1 Ponderación de selección de la primera etapa (UPM)

La ponderación de primera etapa representa el inverso de la probabilidad de selección de la primera etapa asignada a cada UPM.

Una vez realizado el equilibrio en el número de viviendas que constituyen las UPM, es posible obtener la probabilidad de selección de cada UPM de una manera aleatoria simple dentro de cada estrato. La ponderación de primera etapa (PPE) está definida por:

$$PPE_{hi} = \frac{m_h}{M_h}$$

Dónde:

- PPE_{hi} = Probabilidad de selección de primera etapa para la UPM i del estrato h .
 m_h = Número de UPM de la muestra en el estrato h .
 M_h = Número de UPM en el marco muestral h .

2.1.8.2. Tasa de no respuesta de la primera etapa de selección

Para realizar este ajuste, es necesario definir la tasa de no respuesta a nivel de UPM, la cual mide la proporción de UPM seleccionadas que efectivamente participaron en la encuesta.

El ajuste por no respuesta a nivel de UPM se calcula como sigue:

$$TNR_h = \frac{m_{h_ef}}{m_h}$$

Dónde:

- TNR_h = Tasa de no respuesta de la primera etapa de selección en el estrato h .
 m_{h_ef} = Número de UPM efectivas de la muestra en el estrato h .
 m_h = Número de UPM de la muestra en el estrato h .

Las ponderaciones de primera etapa se calculan para las UPM originalmente muestreadas y las UPM de reemplazo que participaron. Cualquier UPM que ya no era elegible fue removida del cálculo del ajuste por no respuesta y queda definido por:

Dónde:

- PPE_{Fhi} = Probabilidad de selección final de primera etapa para la UPM i del estrato h .

2.1.8.3. Ponderación de selección de la segunda etapa (viviendas)

La ponderación de segunda etapa representa el inverso de la segunda probabilidad de selección asignada a cada vivienda. Las viviendas fueron seleccionadas con igual probabilidad de selección, por lo que la ponderación básica de segunda etapa está dada por:

$$PSE_{nij} = \frac{n_i}{N_i}$$

Dónde:

- PSE_{nij} = Probabilidad de selección de segunda etapa para la vivienda j en la UPM i del estrato h .
 n_i = Número de viviendas visitadas en la UPM i del estrato.
 N_i = Total de viviendas de la UPM i .

Para cada UPM, n_i toma un valor de 8 a 11 dependiendo del número de viviendas de reemplazo investigadas.

Adicionalmente es necesario ajustar las probabilidades de inclusión mediante la proporción de viviendas seleccionadas que efectivamente participaron en la encuesta (efectivas), las viviendas rechazo y nadie en casa (no respuesta) y por cambio de ocupación (no elegibilidad). En la ENREFAM 2019, se realiza este ajuste a nivel de estrato.

2.1.8.4. Ajuste de no respuesta de viviendas

El cálculo del ajuste por no respuesta se realizó por estrato y está dada por la siguiente expresión:

$$Aj_{nr,h} = \frac{n_h}{n_h - n_{h_{nr}} - n_{h_{nc}} - n_{h_{ex}}}$$

Dónde:

- $Aj_{nr,h}$ = Ajuste por no respuesta para las viviendas del estrato h .
- n_h = Número de viviendas visitadas en el estrato h .
- $n_{h_{nr}}$ = Número de viviendas que no respondieron en el estrato h .
- $n_{h_{nc}}$ = Número de viviendas con nadie en casa en el estrato h .
- $n_{h_{ex}}$ = Número de viviendas no elegibles en el estrato h .

2.1.8.5. Ajuste de no elegibilidad de viviendas

Al igual que el ajuste de no respuesta, el ajuste de no elegibilidad se realizó por estrato y está dado por la siguiente expresión:

$$Aj_{ne,h} = \frac{n_h - n_{h_{ex}}}{n_h}$$

Dónde:

- $Aj_{ne,h}$ = Ajuste por no elegibilidad de viviendas en el estrato h .
- n_h = Número de viviendas visitadas en el estrato h .
- $n_{h_{ex}}$ = Número de viviendas no elegibles en el estrato h .

Dado que la tasa de respuesta a nivel de vivienda se calcula a nivel de estrato, todas las UPM del mismo estrato comparten la misma corrección de cobertura.

La ponderación de segunda etapa se calcula para las viviendas seleccionadas dentro de las UPM incluidos en la muestra. En el caso de que existan viviendas excluidas de la investigación, se debe recalcular su probabilidad de selección. Por lo tanto, la ponderación de segunda etapa para la j -ésima vivienda en la UPM i del estrato h queda definida por:

$$PSE_{F_{hij}} = \frac{PSE_{hij}}{Aj_{nr,h} \cdot Aj_{ne,h}}$$

Dónde:

- $PSE_{F_{hi}}$ = Probabilidad de selección final de segunda etapa para la UPM i del estrato h .

2.1.8.6. Ponderación de tercera etapa (mujer de 15 años y más)

La ponderación de tercera etapa representa el inverso de la probabilidad de selección de la tercera etapa asignada a la mujer de 15 años y más. Puesto que dentro de cada hogar de la vivienda j de la UPM i en el estrato h se selecciona únicamente a una mujer, se

simplifica la notación con subíndice k. Por lo tanto, la ponderación básica de tercera etapa para la mujer k es:

$$PTE_k = \frac{1}{M_k}$$

Dónde:

PTE_k = Probabilidad de selección de tercera etapa para la mujer k de la vivienda j, UPM i y estrato h.
 M_k = Número de mujeres elegibles dentro del hogar k de la vivienda i, UPM j y estrato h.

2.1.8.7. Ajuste por no respuesta a nivel de mujer de 15 años y más

Las ponderaciones de tercera etapa se calculan para las mujeres seleccionadas dentro del hogar, sin embargo, se debe considerar que existieron hogares con mujeres mayores de 15 años que fueron seleccionadas y no facilitaron información, por lo que se debe recalculan su probabilidad de selección. El ajuste se realizó por estrato y queda definido por:

$$A_{j_{hk}} = \frac{me_h}{\sum_h 1/PTE_k}$$

Dónde:

$A_{j_{hk}}$ = Ajuste por no respuesta a la mujer k del estrato h.
 me_h = Total de mujeres elegibles en el estrato h.
 $\sum_h 1/PTE_k$ = Total de mujeres elegibles en los hogares con una encuesta efectiva en el estrato h.

Por último, la ponderación de tercera etapa se calcula para las mujeres seleccionadas dentro de los hogares incluidos en la muestra. Por lo tanto, la ponderación de tercera etapa para la mujer k queda definida por:

$$PTE_{F_k} = \frac{PTE_k}{A_{j_{hk}}}$$

2.1.8.7. Factores de expansión ENVIGMU 2019

ENREFAM 2019, tiene tres unidades de análisis: vivienda, hogar y mujeres mayores de 15 años. Debido a lo antes mencionado, es necesario generar factores de expansión diferenciados.

Para viviendas y hogares el factor de expansión es:

$$fexp = \frac{1}{PPE_{F_{hi}}} \cdot \frac{1}{PSE_{F_{hi}}}$$

Mientras que para mujeres el factor de expansión es:

$$fexp_M = \frac{1}{PPE_{F_{hi}}} \cdot \frac{1}{PSE_{F_{hi}}} \cdot \frac{1}{PTE_{F_k}}$$

2.1.9. Postestratificación

Considerando que la muestra fue seleccionada a partir del Marco de Muestra para encuestas dirigidas a hogares del INEC, el cual está constituido en su mayoría por información que data del 2010 y que el periodo de levantamiento de información es 2019, existe una diferencia sustancial en la constitución de los dominios de estudio. Con el fin de solventar este inconveniente se utiliza un método denominado postestratificación.

El método de postestratificación, permite realizar un ajuste a los factores de expansión mediante la incorporación de información auxiliar. En este caso, la información auxiliar utilizada proviene de las proyecciones poblacionales del INEC, dándoles la posibilidad de ajustar los totales poblacionales investigados a los resultantes en las proyecciones poblacionales. Por ende, se decide calibrar por:

- Provincia,
- Sexo,
- grupo de edad (menos de 15 años y de 15 años y más)

los cuales coinciden con los dominios de estudio de la muestra. Luego, la fórmula de cálculo del factor de expansión por persona es:

$$fexp_{kp} = a_k \times \frac{N_p}{\hat{N}_p}$$

Dónde:

$fexp_{kp}$ = Factor de expansión de la persona k perteneciente al postestrato p .

a_k = Factor de expansión de la persona k (P_T^j).

\hat{N}_p = Tamaño estimado de la población del postestrato p .

N_p = Tamaño de la población del postestrato p .

En el caso del factor de mujer la calibración considera los mismos grupos de calibración a excepción de sexo.

2.2.0 Estimación de características

Una vez obtenidos los factores de expansión ajustados, se calcula los estimadores de ENVIGMU 2019 mediante el estimador de Horvitz-Thompson, el cual sirve para estimar el valor total de una característica determinada. Está dado por (Carl-Erik Särndal, 1992).

$$\hat{Y}_{mHT} = \sum_i \sum_j fexp_{kp} \cdot y_k$$

Dónde:

- $\hat{Y}_{m_{HT}}$ = estimador insesgado para el total Y_k de una variable y en el dominio de estudio m
 i = número de UPM investigadas en el dominio de estudio m .
 j = número de viviendas investigadas en la UPM i del dominio de estudio m
 y_k = valor de la variable y para la unidad de análisis k .

2.2.1 Error de muestreo

Luego de realizar la estimación respectiva para la variable de interés a nivel de dominio de estudio, el error de muestreo podrá ser calculado a partir de la estimación de la varianza del estimador del total $\hat{Y}_{m_{HT}}$. Para calcular adecuadamente los errores de muestreo de cada estimador, se debe tomar en cuenta los diferentes aspectos del diseño muestral, es decir, las dos etapas de muestreo, la estratificación presente en los dominios de estudio y los procesos de selección en cada una de las etapas.

Con todos estos elementos, el coeficiente de variación para el estimador $\hat{Y}_{k_{HT}}$ viene dado por la siguiente expresión:

$$CV(\hat{Y}_{k_{HT}}) = \frac{\sqrt{\hat{V}_{2st}(\hat{Y}_{k_{HT}})}}{\hat{Y}_{k_{HT}}}$$

Un estimador insesgado para la varianza está dado por:

$$\hat{V}_{2st}(\hat{t}_{\pi}) = \sum \sum_{s_{Ij}} \check{\Delta}_{Iij} \frac{\hat{t}_{i\pi}}{\pi_{Ii}} \frac{\hat{t}_{j\pi}}{\pi_{Ij}} + \sum_{s_{Ij}} \frac{\hat{V}_i}{\pi_{Ii}}$$

En el cual el \hat{V}_i apropiado es

$$\hat{V}_i = \sum \sum_{s_{IIi}} \check{\Delta}_{IIqr|i} \frac{\hat{t}_{iq\pi}}{\pi_{IIq|i}} \frac{\hat{t}_{ir\pi}}{\pi_{IIr|i}}$$

Dónde:

- i, j = Índice que recorre los UPM i, j en el dominio de estudio m .
 q, r = Índices que recorren las viviendas de la UPM i en el dominio de estudio m .
 π_{Ii} = Probabilidad de selección de la Etapa I para la i – ésima UPM en el dominio de estudio m .
 $\pi_{IIq|i}$ = Probabilidad de selección de la Etapa II para la q – ésima vivienda, dada la i – ésima UPM.
 $\check{\Delta}_{Iij}$ = Cantidad Δ expandida asociada a las UPM i, j .
 $\check{\Delta}_{IIqr|i}$ = Cantidad Δ expandida asociada a las viviendas q, r dada la selección de la i – ésima UPM.

2.2.2. Métodos de estimación de errores para diseños muestrales complejos

Aunque la selección del diseño de muestreo y el estimador sean de libre elección para los investigadores, no lo es el cálculo de las medidas de confiabilidad y precisión. Dado que la base científica sobre la cual descansa el muestreo es la inferencia estadística, se deben respetar las normas básicas para la asignación y posterior cálculo del margen de error, que constituye una medida unificada del error total de muestreo el cual cuantifica la incertidumbre acerca de las estimaciones en una encuesta.

Los métodos de estimación de los errores muestrales pueden clasificarse en cuatro categorías:

- a) Métodos exactos
- b) Métodos de la última UPM
- c) Aproximaciones por linealización
- d) Técnicas de replicación

Para la descripción de los métodos se ha tomado como referencia los textos de Kish y Frankel (1974), Wolter (1985) y Lehtonen y Pahkinen (1995). A continuación se describe las principales características de cada uno de los métodos de estimación de errores para el muestreo complejo:

- Los métodos exactos pueden ser utilizadas para estimar totales, medias, tamaños y proporciones.
- La linealización de Taylor debe ser utilizada para estimar parámetros no lineales como razones, medias dentro de dominios, cuartiles o funciones de distribución.
- La técnica de la última UPM junto con la linealización de Taylor puede ser utilizada para estimar la varianza de los indicadores de interés de las encuestas dirigidas a hogares que tengan diseños muestrales complejos. Esta es la técnica que por defecto utiliza el software SPSS.
- Las técnicas de replicación pueden ser usadas para estimar eficientemente todos los parámetros de interés, sin importar su forma funcional.
- La comparación general entre los métodos de linealización y replicación, es que no generan resultados idénticos del error de muestreo, pero hay que señalar que existen estudios (Kish y Frankel, 1974) que concluyen que las diferencias presentadas no son significativas cuando se trata de grandes muestras.

El INEC utiliza para la estimación de los parámetros de interés y sus correspondientes errores de muestreo diversos programas estadísticos tales como SPSS, Stata y R. En virtud de las características de cada uno de los métodos, es la técnica de la última UPM en combinación con la linealización de Taylor, la cual induce a una muy buena aproximación del error muestral sobre los indicadores más importantes de las encuestas dirigidas a hogares, además de su facilidad de cálculo y replica. En este sentido, será esta la técnica la utilizada para la estimación de los errores muestrales en la ENVIGMU 2019.

Las variables requeridas para declarar el diseño muestral en los programas estadísticos (SPSS, Stata y R) y ejecutar el cálculo de los errores de muestreo son presentadas en la Tabla 6, donde se describe las etiquetas de las variables identificadoras de las UPM, estratos y ponderación.

Tabla 6. Variables requeridas para la declaración del diseño muestral

Característica	Variable	Descripción
Unidad primaria de muestreo	upm	Agrupación de viviendas ocupadas en un número entre 30 a 60, próximas entre sí y con límites definidos.
Estratos	Estrato	Identificación de estrato muestral (aproximación clasificación socio-económica).
Ponderación	Fexp	Factor de expansión calibrado.

Fuente: Elaboración propia, INEC 2019

BIBLIOGRAFÍA

Hidiroglu, M., Särndal, C., & Binder, D. (1995). Weighting and Estimation in Business Surveys. En Cox, Binder, Chinnappa, Christianson, Colledge, & Kott, *Business Survey Methods* (págs. 477-502). Jhon Wiley & Sons.

Kish, L. (1972). *Muestreo de Encuestas*. México: Trillas.

Särndal, C., Swensson, B., & Wretman, J. (2003). *Model Assisted Survey Sampling*, Nueva York, USA. New York: Springer.

**CADA
HECHO
DE TU
VIDA**
Cuenta

 @ecuadorencifras

 INEC/Ecuador

 @InecEcuador

 INECEcuador

 t.me/ecuadorencifras

 INEC Ecuador