



República
del Ecuador

Diseño muestral

ESPAC 2021

Abril, 2022



Buenas cifras,
mejores vidas



Documento de diseño muestral de la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC 2021)

Quito - Ecuador, 2022

Dirección

Dirección de Infraestructura Estadística y Muestreo

Unidad

Gestión de Diseño Muestral

Elaborado por:

Andrea Conde

Revisado por:

Javier Núñez

Francisco Céspedes

Andrés Albán

Aprobado por:

Julio Muñoz



Contenido

Introducción	6
Objetivo	7
Determinación del marco de muestreo	7
Marco de áreas	7
Estratificación del marco de áreas	9
Segmentación y definición de tamaño por estrato	10
Marco de lista	12
Diseño muestral	14
Universo de investigación	14
Población objetivo	14
Unidades de muestreo y análisis	14
Cobertura geográfica	15
Tipo de muestreo	15
Tamaño de muestra	15
Selección de la muestra	16
Cobertura muestral	18
Cálculo de los factores de expansión	20
Probabilidad de selección	20
Probabilidad de selección de primera etapa	20
Probabilidad de selección de segunda etapa	20
Ajustes por rechazo	21
Ajuste por subdivisión	21
Ajuste por fuera del territorio	21
Ajuste por intersección	22
Estimación de características	23
Error de muestreo	24
Bibliografía	26
Anexos	27



Lista de tablas

Tabla 1. Estratos del marco de muestreo para investigaciones agropecuarias	9
Tabla 2. Tamaño de las Unidades Primarias de Muestreo (UPM)	11
Tabla 3. Distribución del número de segmentos del marco de áreas por provincia y estrato	11
Tabla 4. Número de segmentos y superficie por estrato del marco de áreas.....	12
Tabla 5. Fuentes de información consideradas para el marco de lista.....	13
Tabla 6. Código de estrato por producto para el marco de lista	14
Tabla 7. Tasas de muestreo por estrato	15
Tabla 8. Tamaños de muestra por provincia y estrato	16
Tabla 9. Cobertura muestral	18
Tabla 10. Distribución de los segmentos por provincia y tipo de rechazo.....	19
Tabla 11. Distribución del número de terrenos por provincia y estrato.....	19

Lista de figuras

Figura 1. Generación de las Unidades Mínimas de Estratificación (UME)	8
Figura 2. Límites provinciales del marco de muestreo de áreas para investigaciones agropecuarias. Un ejemplo para las provincias de Manabí y Esmeraldas	9
Figura 3. Áreas para estratificación	10
Figura 4. Distribución del número de segmentos por estrato	12
Figura 5. Zonas de selección del estrato 02 de la provincia de Santo Domingo de los Tsáchilas.....	17
Figura 6. Selección de los segmentos muestrales de estrato 04	17
Figura 7. Segmentos que se encuentran fuera del territorio continental ecuatoriano	22
Figura 8: superficie de los terrenos declarados en campo	23

Anexos

Anexo 1. Distribución por provincia del número de productos del marco de lista .	27
---	----



Introducción

El Instituto Nacional de Estadística y Censos (INEC) es la institución pública que coordina, norma y evalúa la producción de la información estadística oficial proveniente del Sistema Estadístico Nacional con el propósito de entregar a la sociedad y al Estado información de calidad, pertinente, veraz y oportuna, que contribuya al desarrollo nacional.

La necesidad de generar herramientas de decisión para la planificación y formulación de política agropecuaria en el país es permanente, por lo que es prioritario contar con estadísticas agropecuarias con regularidad, tales como la producción de cultivos y de ganado. Esta información se recoge generalmente mediante registros administrativos y/o encuestas por muestreo.

Con este antecedente y con la finalidad de estimar los principales indicadores agropecuarios y ambientales, el INEC realiza anualmente la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC), la cual es el pilar fundamental del actual Sistema de Estadísticas Agropecuarias Nacionales de Ecuador.

El proceso para la generación del diseño muestral consta de las siguientes etapas:

Primera etapa:

- Definición del marco de muestreo
- Las unidades de muestreo y análisis
- Cobertura geográfica
- Cobertura temática

Segunda etapa:

- Metodología para el diseño muestral
- Determinación de los dominios de estudio
- Determinación del tamaño de la muestra
- Selección de la muestra

Tercera etapa:

- Validación de la cobertura obtenida en campo
- Cálculo de los factores de expansión
- Estimación de características
- Error de muestreo
- Declaración del plan de muestreo

Objetivo

Elaborar el diseño muestral de la "Encuesta de Superficie y Producción Agropecuaria Continua 2021", utilizando técnicas de muestreo probabilístico que permitan emitir resultados confiables de las principales variables que se investigarán en esta encuesta.

Determinación del marco de muestreo

(Kish, 1972) manifiesta que el marco de muestreo es la piedra angular en torno a la cual los procesos de selección deben ser diseñados. Por tal motivo, es indispensable que el marco se encuentre completo, validado y depurado antes de iniciar el cálculo de la muestra.

El marco de muestreo de la ESPAC concierne a un **marco de muestreo múltiple**, el cual se define como una combinación de un marco de áreas y un marco de lista (marco dual). Las estimaciones basadas en marcos de áreas tienden a ser sensibles a valores extremos, por lo que para conseguir estabilidad se incluye una lista de unidades de observación más extrema, denominado marco de lista.

Un muestreo de marcos múltiples es más eficiente que un marco de áreas, puesto que proporciona estimaciones más exactas y el costo adicional de elaborar el marco de lista es muy bajo. En encuestas agropecuarias se usa frecuentemente un marco dual, el cual es completo, en el sentido de que cualquier unidad de observación de la población objetivo está en al menos uno de sus dos componentes.

Marco de áreas

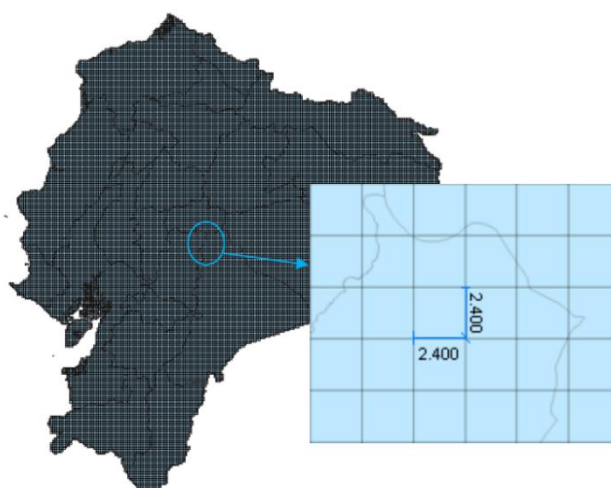
Un marco de áreas se construye a partir de la división del área a ser investigada en pequeños pedazos, generalmente conocidos como segmentos.

El conjunto de segmentos conforma el marco de áreas, en cuya construcción se utiliza material cartográfico (mapas, imágenes de satélite y fotografías aéreas) para delimitar las unidades de muestreo y medir su superficie.

Para la construcción del marco de muestreo de áreas para investigaciones agropecuarias, se empieza definiendo una malla cuadrada homogénea sobre el territorio continental ecuatoriano, cuyos elementos tienen un área de 576 hectáreas (ha), de tal forma que esta forme un recubrimiento finito del mismo (Ver **Figura 1**). A los elementos de la malla cuadrada homogénea se las denomina Unidad Mínima de Estratificación (UME)¹.

¹ Reciben este nombre porque las UME serán estratificadas en función a la intensidad de uso de suelo, procedimiento que se explicará más adelante.

Figura 1. Generación de las Unidades Mínimas de Estratificación (UME)



Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Para la construcción del marco de áreas se debe excluir las zonas del territorio ecuatoriano en las que la producción agropecuaria es ínfima y no afecta significativamente a los objetivos de la operación estadística. Entre estas zonas se encuentran:

- Segmentos completamente contenidos en superficies mayores a los 3.000 metros sobre el nivel del mar (msnm).
- Segmentos completamente contenidos en superficies sin uso agropecuario, tales como reservas naturales, selva amazónica, entre otros.
- Segmentos completamente contenidos en superficies urbanas.
- Segmentos con intersección parcial o total con sectores amanzanados, pertenecientes al marco de muestreo generado para el Sistema Integrado de Encuestas de Hogares (SIEH)².

Se puede garantizar que el marco de muestreo es completo puesto que el conjunto de todas las UME forman un recubrimiento finito del territorio continental ecuatoriano. Además, puesto que cada UME ocupa una superficie única del mismo, no existe intersección entre cualquier par de UME, esto implica que dentro del marco de muestreo podemos identificar de manera única cada UME, evitando así la duplicidad de elementos del mismo. Cabe destacar que esta propiedad se mantiene en el caso de dividir cualquier UME.

Las UME ubicadas sobre las zonas fronterizas y costeras del territorio ecuatoriano abarcan superficie de otros países y del Océano Pacífico, respectivamente. Para evitar que población ajena o sin información agropecuaria (en el caso de las UME costeras) formen parte del marco de muestreo, todo subconjunto territorial de cada UME que no forme parte de la superficie continental ecuatoriana no será investigado.

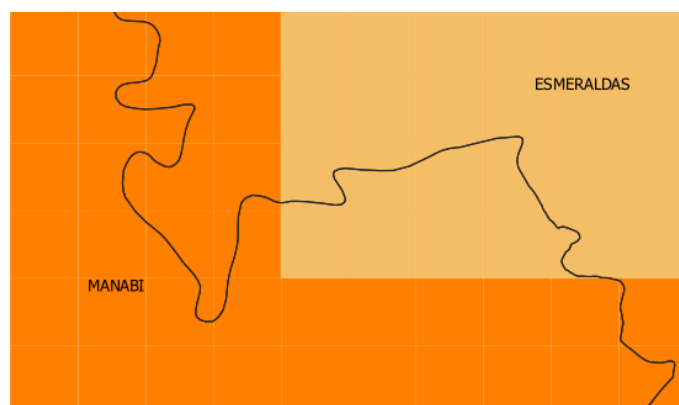
Por otro lado, puesto que el requerimiento de información contempla la estimación de información a nivel provincial, los límites provinciales se han ajustado al perfil de las UME, es decir, dentro del marco de muestreo de áreas para investigaciones

² El Sistema Integrado de Encuestas de Hogares (SIEH) es una estrategia de producción de información sobre las características demográficas y socioeconómicas de los hogares y las personas, basado en el diseño y ejecución coordinada de un conjunto de encuestas que comparten un cierto marco conceptual y metodológico.



agropecuarias, cada provincia tiene límites poligonales definidos en el porcentaje de cada UME que pertenece a cada provincia (Ver **Figura 2**)

Figura 2. Límites provinciales del marco de muestreo de áreas para investigaciones agropecuarias. Un ejemplo para las provincias de Manabí y Esmeraldas



Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Estratificación del marco de áreas

La estratificación del marco de áreas inicia con la estratificación del territorio nacional, del cual se desprenden dos estratos primarios, de donde el interés se centra en las zonas con uso agropecuario y con baja densidad poblacional:

- (i) Zonas urbanas densamente pobladas.
- (ii) Resto del territorio, incluyendo áreas agropecuarias, forestales, agua y otras áreas rurales con baja densidad de población.

Del estrato (ii) se separan únicamente los cuerpos de agua, la superficie restante se estratifica en función de los siguientes criterios:

- **Distribución Político Administrativa del Ecuador actualizada a junio de 2018 hasta el nivel provincial:** debido a que el marco de muestreo es nacional y que el requerimiento de información contempla estimaciones a nivel provincial, cada provincia es tratada como un dominio de estudio independiente.
- **Intensidad de uso de suelo:** para la selección de la muestra, se tomó la variable "superficie de suelo con uso agropecuario" como variable de diseño de muestra, estratificándose el marco de muestreo en función al porcentaje de superficie con uso agropecuario que tiene cada UME de la siguiente forma:

Tabla 1. Estratos del marco de muestreo para investigaciones agropecuarias

Estrato	% de la UME con uso agropecuario	Superficie de la UME con uso agropecuario en hectáreas (ha)
Estrato 1	60% - 100%	345,6 – 576
Estrato 2	20% - 60%	115,2 – 345,6
Estrato 3	0% - 20%	0 – 115,2

Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Los límites entre estratos son los estándares recomendados en [FAO (1996, 1998)], citados en (Instituto Nacional de Estadística y Censos, 2015):

- Áreas con una proporción mayor al 60% de uso de suelo agropecuario se consideran UME de estrato 1.



- Áreas con una proporción de superficie cultivada entre el 20% y 60% se consideran UME de estrato 2.
- Áreas con una proporción de superficie cultivada inferior al 20% se consideran UME de estrato 3.

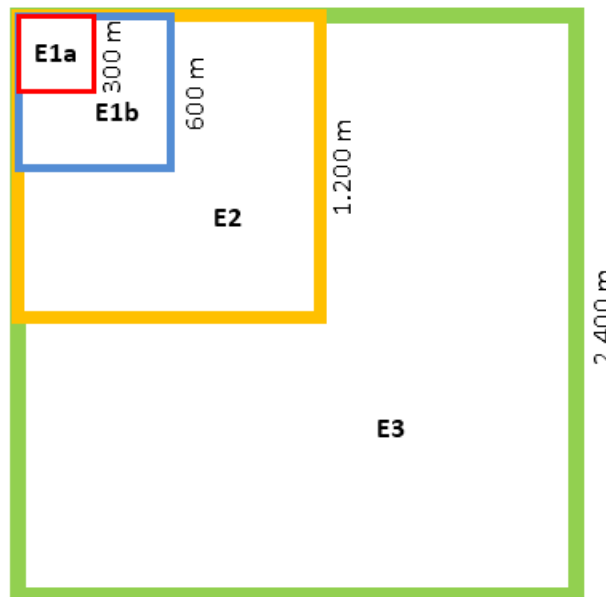
Segmentación y definición de tamaño por estrato

Teniendo en cuenta que las características de producción de la región natural Sierra (producción intensiva) son distintas a las características de las regiones naturales Costa y Amazonía (producción extensiva), se decidió diferenciar las UME del estrato 1 en dos estratos. Estrato 1a para las regiones naturales de la Sierra y estrato 1b para las regiones naturales de la Costa y Amazonía. Esta diferenciación fue establecida por la Dirección de Cartografía Estadística del INEC, de acuerdo a las regiones biogeográficas del Ecuador.

Además, gracias a la experiencia obtenida al recolectar la información en campo, se pudo demostrar que investigar una superficie de 576 hectáreas (ha) resulta sumamente dificultoso por varios factores, entre ellos el alto número de entrevistas a realizar, por lo que para construir Unidades Primarias de Muestreo (UPM) más acordes a las características de producción se dividió sucesivamente cada UME en función al estrato al que pertenece, de la siguiente forma (Ver **Figura 3**):

- UME de estrato 2 se dividieron en 4 partes iguales,
- UME de estrato 1a y 1b se dividieron en 64 y 16 partes iguales, respectivamente,
- UME de estrato 3 no fueron divididas.

Figura 3. Áreas para estratificación



Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Bajo estas consideraciones, los tamaños de las UPM quedan definidos de la siguiente manera:



Tabla 2. *Tamaño de las Unidades Primarias de Muestreo (UPM)*

Estrato ³	Tamaño UPM en hectáreas (ha ²)
01	9
02	36
03	144
04	576

Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

De ahora en adelante, las UPM se denominarán segmentos.

Es importante mencionar que para la ESPAC 2021 no se realizaron actualizaciones del marco de muestreo de áreas, por lo que tiene la misma conformación del marco de muestreo del año 2019. Únicamente se reclasificaron 3 segmentos de estrato 4 de la provincia de Los Ríos de la siguiente forma: 2 de ellos en estrato 3 (8 segmentos) y el restante en estrato 2 (6 segmentos); así se cuenta con 357.620 segmentos.

La **Tabla 3** muestra la distribución por provincia y estrato de los segmentos que conforman el marco de áreas, recordando que para este ejercicio se consideró únicamente la superficie continental del Ecuador.

Cabe indicar que la determinación de los límites provinciales es competencia de la Secretaría Técnica del Comité Nacional de Límites Internos, ante esto y de acuerdo a las resoluciones emitidas hasta diciembre de 2018 en la División Político Administrativa del territorio, las zonas no delimitadas que se presentaban en años anteriores, se anexaron a las provincias constituidas, donde el territorio conocido como Manga del Cura pasa a ser parte de la provincia de Manabí, Las Golondrinas a la provincia de Imbabura y El Piedrero a las provincias de Guayas y Cañar.

Tabla 3. *Distribución del número de segmentos del marco de áreas por provincia y estrato*

Provincia	01	02	03	04	Total
Azuay	29.963	190	1.008	189	31.350
Bolívar	17.038	252	601	51	17.942
Cañar	9.176	691	503	94	10.464
Carchi	10.154	0	405	90	10.649
Cotopaxi	23.425	412	660	123	24.620
Chimborazo	17.749	60	816	155	18.780
El Oro	12.375	3.354	657	83	16.469
Esmeraldas	1.905	16.830	1.713	296	20.744
Guayas	275	15.880	1.620	302	18.077
Imbabura	14.455	98	661	105	15.319
Loja	19.674	0	1.271	270	21.215
Los Ríos	2.144	14.900	455	3	17.502
Manabí	0	26.210	2.886	366	29.462
Morona Santiago	9.790	2.898	2.235	521	15.444
Napo	6.835	127	1.485	291	8.738
Pastaza	4.586	367	2.454	726	8.133
Pichincha	26.430	373	914	196	27.913
Tungurahua	9.522	0	263	52	9.837
Zamora Chinchipe	3.556	669	754	167	5.146
Sucumbíos	1.994	4.874	2.404	468	9.740
Orellana	324	2.311	1.959	422	5.016
Santo Domingo de los Tsáchilas	12.514	1.815	286	0	14.615
Santa Elena	0	181	193	71	445
Total	233.884	92.492	26.203	5.041	357.620

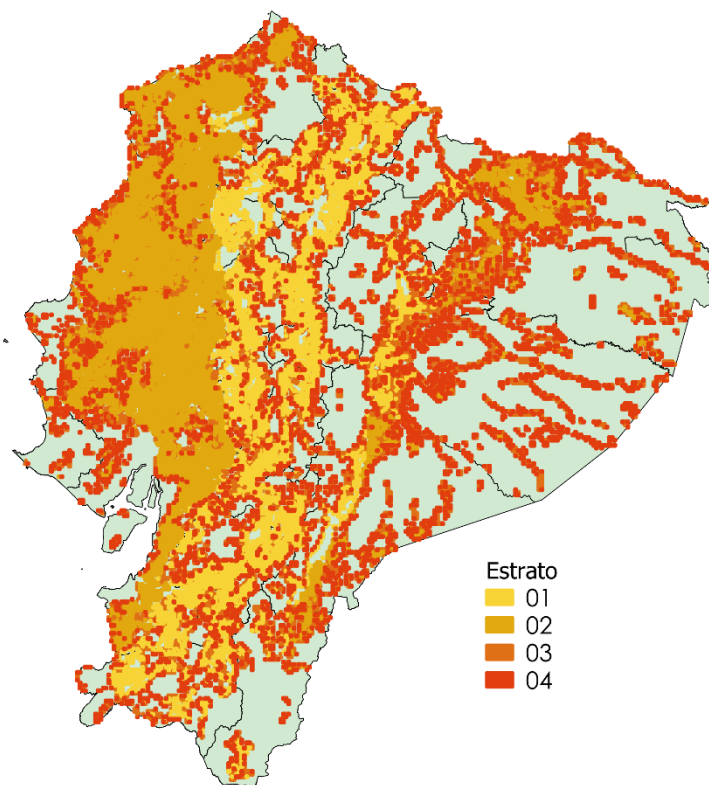
Fuente: Marco de áreas ESPAC 2021

³ Con el objetivo de llevar una lógica en la codificación, se decidió cambiar las denominaciones iniciales de los estratos de la siguiente forma: estrato 1a por estrato 01, estrato 1b por estrato 02, estrato 2 por estrato 03 y estrato 3 por estrato 04.



En la **Figura 4** se puede observar la distribución del número de segmentos muestrales por estrato, en el territorio continental ecuatoriano:

Figura 4. Distribución del número de segmentos por estrato



Fuente: Marco de áreas ESPAC 2021

Por otra parte, en la **Tabla 4** se muestra el número de segmentos por estrato y su respectiva superficie:

Tabla 4. Número de segmentos y superficie por estrato del marco de áreas

Estrato	Total	Superficie (ha ²)	Total (ha ²)	Superficie (%)
01	233.884	9	2.104.956	17,4%
02	92.492	36	3.329.712	27,5%
03	26.203	144	3.773.232	31,2%
04	5.041	576	2.903.616	24,0%
Total	357.620		12.111.516	100,0%

Fuente: Marco de áreas ESPAC 2021

Marco de lista

Con el objetivo de mejorar la calidad de las estimaciones de las variables de superficie y producción de algunos cultivos, se elaboró un directorio de investigación con los principales productores de ciertos productos.

Los productos considerados para el marco de lista son aquellos que actualmente son de importancia macroeconómica y determinados como sensibles o estratégicos para instituciones como el Banco Central del Ecuador, Ministerio de Agricultura, Secretaría Nacional de Planificación y Vicepresidencia de la República:



- Arroz
- Banano
- Café
- Cacao
- Caña de azúcar para azúcar
- Maíz duro seco
- Mango
- Maracuyá
- Palma africana
- Ganado vacuno
- Ganado porcino
- Flores
- Aves
- Quinoa

Dada la complejidad que representa generar un marco de lista de fuente homogénea para todos estos productos, se elaboró el directorio de investigación utilizando varias fuentes de información que cumplieren las siguientes características:

- Directorio actualizado.
- Cobertura superior al 85% de superficie y/o producción a nivel nacional, según las referencias entregadas por MAG.
- Directorio validado (superficie, producción, existencia de animales).

Una vez analizada la consistencia de la información disponible proveniente de varias instituciones, se elaboró un directorio para cada producto. En la **Tabla 5** se detalla las fuentes de información consideradas para el directorio del marco de lista.

Además, se estableció un parámetro de aporte dentro de la actividad agrícola o pecuaria. La determinación del parámetro de aporte para cada producto dentro de su actividad es de vital importancia y debe ser construido con cautela debido a que el objetivo es mejorar las estimaciones, alcanzando niveles aceptables que reflejen la realidad de las actividades agropecuarias, es decir, no sobreestimar las variables de interés.

Tabla 5. Fuentes de información consideradas para el marco de lista

Producto	Fuente del directorio	Elaboración
Arroz	ESPAC 2013	INEC
Banano	Catastro bananero 2012-2013	MAG
Brócoli	ESPAC 2013	INEC
Café	ESPAC 2013	INEC
Cacao	ESPAC 2013	INEC
Caña de azúcar para azúcar	ESPAC 2013	INEC
Maíz duro seco	ESPAC 2013	INEC
Mango	Directorio de productores, empaques y exportadores de mango 2014	Fundación Mango Ecuador
Maracuyá	ESPAC 2013	INEC
Palma africana	Censo palmicultor 2005	MAG
Ganado vacuno (cabezas)	Registro de productores (vacunación fiebre aftosa) 2013	MAG – AGROCALIDAD
Ganado porcino (cabezas)	Censo porcicultor 2010	MAG
Flores	Catastro florícola 2010	MAG
Planteles avícolas	Censo avícola 2006	MAG - CONAVE
Quinoa	ESPAC 2014, ESPAC 2015	INEC

Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Los directorios se consideraron de forma individual, y una vez esquematizada la información en función de su tamaño se realizó una selección por producto. Para mantener un control y seguimiento de los directorios, a cada producto se le asignó un código de estrato:



Tabla 6. Código de estrato por producto para el marco de lista

Estrato	Producto
51	Arroz
52	Banano
53	Brócoli
54	Cacao
55	Café
56	Caña de azúcar para azúcar
57	Maíz duro seco
58	Mango
59	Maracuyá
60	Palma africana
61	Ganado vacuno
62	Ganado porcino
63	Flores
64	Aves
65	Quinua

Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

La distribución por provincia de los productos considerados en el marco de lista se puede verse en el **Anexo 1**.

Diseño muestral

Universo de investigación

El universo de investigación para la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC) está conformada por el conjunto de todos los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano.

Se entiende por terreno a una extensión de tierra continua que se encuentra con un tipo de cultivo (solo o asociado) o uso de suelo determinado, a cargo de una persona responsable o productora; de existir otro cultivo diferente en un área continua, este se considerará otro terreno que será identificado con otro número secuencial y la clave del cultivo o uso del suelo allí existente.

Por ejemplo, en un espacio físico continuo de una hectárea sin divisiones de cercas o accidentes físicos se encuentran un cultivo de maíz solo, uno con fréjol, otro con plátano, una casa, otro con pasto cultivado y un último espacio con bosque, así ésta extensión de tierra tendrá 6 terrenos.

Población objetivo

La población objetivo comprende todos los terrenos con superficie agropecuaria en el territorio continental ecuatoriano.

Unidades de muestreo y análisis

La unidad elemental de muestreo es el segmento, también denominado como Unidad Primaria de Muestreo (UPM). Por otro lado, las unidades de análisis son todos los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano.

Un segmento es el conjunto de terrenos, cuya superficie varía de acuerdo al estrato al que pertenece.



Cobertura geográfica

La cobertura geográfica de la ESPAC 2021 está definida por los terrenos que presentan superficie con producción agropecuaria en el territorio continental ecuatoriano, por lo que se excluye la región Insular.

El nivel de estimación es uniforme para cada provincia del Ecuador, manteniendo así la representatividad provincial, por lo que cada provincia es considerada como un dominio de diseño.

Tipo de muestreo

El tipo de muestreo corresponde a un muestreo probabilístico estratificado por conglomerados y por réplicas, con selección aleatoria.

Tamaño de muestra

En cada dominio (provincia) se calcula un número determinado de segmentos en relación a una tasa de muestreo específica por estrato. Con el propósito de mejorar la eficiencia de las estimaciones, estas tasas de muestreo se establecieron en función al número de segmentos del marco de áreas, es decir, si el número de UPM del marco de áreas es mayor entonces la tasa de muestreo será menor, y viceversa.

Tabla 7. Tasas de muestreo por estrato

Estrato	Tasa de muestreo
01	1,50%
02	1,50%
03	1,20%
04	4,00%

Fuente: INEC (Instituto Nacional de Estadística y Censos, 2015)

Cabe señalar que, en casos donde la muestra sea menor a 10, se debe ajustar por los mínimos muestrales⁴, de manera que se puede calcular los errores de muestreo asociados a las estimaciones. En la **Tabla 8** se muestra los tamaños de muestra por provincia y estrato para la ESPAC 2021:

⁴ El número mínimo de segmentos por provincia y estrato en la muestra es 10, cuando el número de segmentos del marco de áreas sea menor a 10 se incluirán en la muestra todos aquellos segmentos.



Tabla 8. Tamaños de muestra por provincia y estrato

Provincia	01	02	03	04	Total
Azuay	450	10	15	10	485
Bolívar	255	10	10	10	285
Cañar	140	10	10	10	170
Carchi	155	0	10	10	175
Cotopaxi	355	15	15	10	395
Chimborazo	270	10	15	10	305
El Oro	185	50	10	10	255
Esmeraldas	30	255	30	15	330
Guayas	10	240	20	15	285
Imbabura	220	10	10	10	250
Loja	295	0	20	15	330
Los Ríos	35	215	10	3	263
Manabí	0	395	45	20	460
Morona Santiago	150	45	30	20	245
Napo	105	10	20	15	150
Pastaza	70	10	30	30	140
Pichincha	400	10	20	10	440
Tungurahua	145	0	10	10	165
Zamora Chinchipe	55	10	10	10	85
Sucumbíos	30	75	30	20	155
Orellana	10	35	25	20	90
Santo Domingo de los Tsáchilas	190	40	10	0	240
Santa Elena	0	10	10	10	30
Total	3.555	1.465	415	293	5.728

Fuente: Muestra ESPAC 2021

Selección de la muestra

Con el objetivo de seleccionar de manera uniforme y eficiente los segmentos que conformarán la muestra, de acuerdo al tamaño establecido, se construyen zonas de selección, las cuales son conjuntos de segmentos muestrales construidas en cada estrato de cada provincia.

Dado que el muestreo de la ESPAC es replicado, se escogen de forma independiente en cada zona de selección⁵ (estrato implícito) 5 segmentos, los cuales corresponden a las réplicas del año 2021⁶.

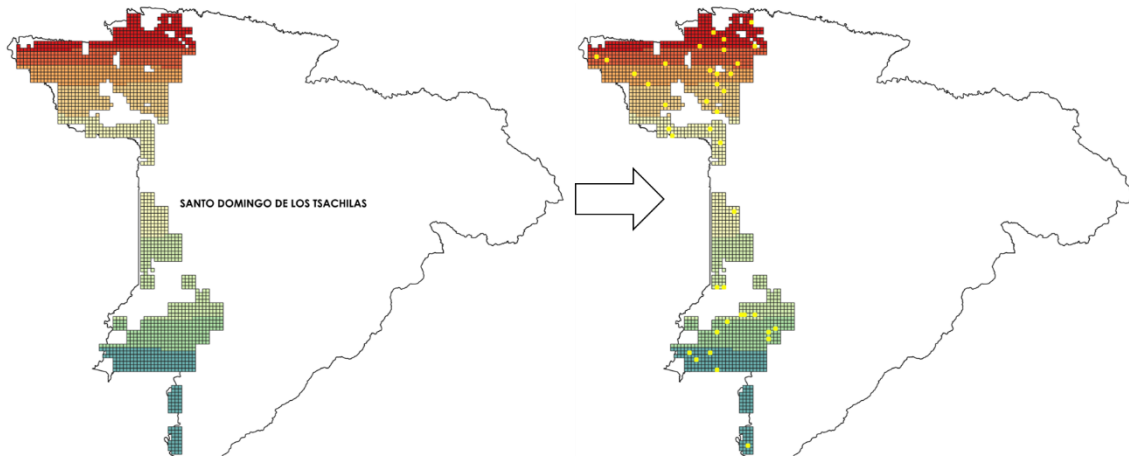
En la **Figura 5** se muestra las zonas de selección del estrato 02 de la provincia de Santo Domingo de los Tsáchilas junto con los segmentos que fueron seleccionados en la muestra.

⁵ Para el cálculo de los errores de muestreo debe existir mínimo dos zonas de selección por provincia y estrato, por tal motivo los mínimos muestrales son 10 segmentos, 5 por cada zona de selección.

⁶ En el año 2014 se seleccionaron 5 segmentos correspondientes a las réplicas 1,2,3,4 y 5, en el año 2015 las réplicas 2,3,4,5 y 6, en el año 2016 las réplicas 3,4,5,6 y 7, en el año 2017 las réplicas 4,5,6,7 y 8, en el año 2018 5,6,7,8 y 9, en el año 2019 las réplicas 6,7,8,9 y 10, en el año 2020 las réplicas 7,8,9,10 y 11, finalmente para el año 2021 se tomaron las réplicas 8,9,10,11 y 12.



Figura 5. Zonas de selección del estrato 02 de la provincia de Santo Domingo de los Tsáchilas

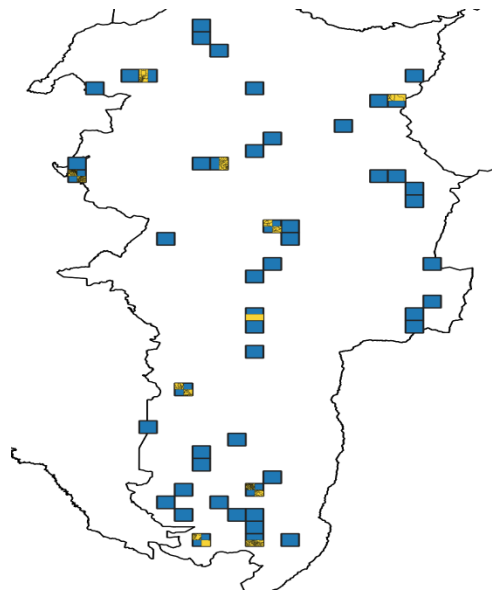


Fuente: Marco de áreas y muestra ESPAC 2021

De acuerdo a las asistencias técnicas proporcionadas por la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), para optimizar el tamaño de la muestra y por tanto mejorar la eficiencia de las estimaciones, la selección de la muestra de los segmentos del estrato 04 se realiza en dos etapas.

En la primera etapa se seleccionan los segmentos de estrato 04 en función al tamaño de muestra establecido, para luego en la segunda etapa dividir estos segmentos en cuatro subsegmentos de igual superficie (144 hectáreas) y escoger únicamente dos de ellos (Ver **Figura 6**) para investigar.

Figura 6. Selección de los segmentos muestrales de estrato 04



Fuente: marco de áreas y muestra ESPAC 2021

Es importante señalar que únicamente en la provincia de Los Ríos se investigaron todos los segmentos de estrato 04, es decir, 3 segmentos. (Ver **Tabla 8**)



Cobertura muestral

La información de cobertura de una encuesta nos permite evaluar su rendimiento muestral, un seguimiento oportuno y adecuado de la misma permite generar estrategias para reducir los errores no muestrales⁷ presentes en la etapa de recolección de datos.

Esta información se convierte en el insumo principal al momento de ajustar los factores de expansión de una encuesta en función a la relación entre segmentos efectivos y segmentos investigados, para cada uno de los estratos de la población objetivo.

Una vez seleccionada la muestra, los encuestadores se encargan de levantar la información correspondiente a los terrenos; sin embargo, no siempre pueden obtener la información completa del segmento, pues la o las personas dueñas de los terrenos contenidos en el segmento a investigar se rehúsan a dar información de sus terrenos o desconocen ciertas características de sus tierras que resultan ser importantes para la ESPAC.

Así se tienen los siguientes rechazos:

- **Rechazo total:** si de todos los terrenos del segmento no se tiene información.
- **Rechazo parcial:** si al menos se obtuvo la información de un terreno.

Por otro lado, la gran extensión de usos del suelo de los segmentos de estrato 03 y 04, hizo que se considerara realizar una nueva división en dichos segmentos y escoger únicamente un subsegmento para levantar la información (36 y 72 hectáreas respectivamente).

Finalmente, considerando la construcción del marco de áreas, las UME forman un recubrimiento finito del territorio continental ecuatoriano por lo que algunas zonas de selección contienen segmentos que se encuentran por fuera del territorio continental y que fueron seleccionados en la muestra (Ver **Figura 7**). Su identificación es importante para calcular los factores de expansión, lo cual se explica más adelante.

Es así que, una vez levantada la información de la ESPAC 2021 se identificó que de los 5.728 segmentos muestrales a ser investigados únicamente se obtuvieron 5.693 segmentos efectivos, es decir, segmentos con al menos un terreno con información efectiva.

La **Tabla 9**, muestra el número de segmentos por cobertura muestral.

Tabla 9. Cobertura muestral

Cuestionario	Descripción	Nº segmentos
5555	Subdividido	21
6666	Fuera del territorio	16
7777	Rechazo total	35
8888	Rechazo parcial	99
Total		171

Fuente: Muestra ESPAC 2021

La **Tabla 10** muestra la distribución de los segmentos por provincia y cobertura muestral, siendo Bolívar la provincia que registra un mayor número de segmentos con rechazo total y las provincias de Azuay y Pichincha las que presentan mayor número de segmentos con rechazo parcial.

⁷ Este error no muestral se conoce como error debido a la no respuesta.



Tabla 10. Distribución de los segmentos por provincia y tipo de rechazo

Provincia	Subdividido	Fuera del territorio	Rechazo total	Rechazo parcial	Total
Azuay	0	0	2	19	21
Bolívar	0	0	11	2	13
Cañar	0	0	2	4	6
Carchi	0	2	0	7	9
Cotopaxi	1	0	3	5	9
Chimborazo	5	0	2	4	11
El Oro	0	1	0	4	5
Esmeraldas	0	3	0	6	9
Guayas	2	0	3	3	8
Imbabura	2	0	1	5	8
Loja	1	1	1	4	7
Los Ríos	0	0	1	5	6
Manabí	2	2	2	2	8
Morona Santiago	0	0	2	2	4
Pastaza	0	0	0	1	1
Pichincha	2	0	3	14	19
Tungurahua	6	0	2	2	10
Zamora Chinchipe	0	1	0	0	1
Sucumbíos	0	6	0	4	10
Orellana	0	0	0	2	2
Santo Domingo de los Tsáchilas	0	0	0	3	3
Santa Elena	0	0	0	1	1
Total	21	16	35	99	171

Fuente: Muestra ESPAC 2021

Finalmente, 88.293 terrenos conforman la muestra de la ESPAC 2021. En la **Tabla 11** se muestra la distribución de los mismos por provincia y estrato.

Tabla 11. Distribución del número de terrenos por provincia y estrato

Provincia	E1	E2	E3	E4	Total
Azuay	7.201	169	576	416	8.362
Bolívar	2.797	285	391	566	4.039
Cañar	2.303	136	284	242	2.965
Carchi	1.364	0	527	94	1.985
Cotopaxi	7.380	323	381	706	8.790
Chimborazo	6.063	138	999	294	7.494
El Oro	1.083	508	181	207	1.979
Esmeraldas	103	2.646	632	268	3.649
Guayas	33	4.230	802	425	5.490
Imbabura	2.608	173	292	148	3.221
Loja	2.747	0	955	406	4.108
Los Ríos	173	4.317	402	64	4.956
Manabí	0	6.277	1.664	602	8.543
Morona Santiago	775	453	468	137	1.833
Napo	675	279	380	131	1.465
Pastaza	359	45	260	190	854
Pichincha	4.424	131	380	344	5.279
Tungurahua	5.087	0	795	340	6.222
Zamora Chinchipe	316	159	240	272	987
Sucumbíos	105	1.345	527	294	2.271
Orellana	24	359	464	307	1.154
Santo Domingo de los Tsáchilas	1.209	533	517	0	2.259
Santa Elena	0	45	124	219	388
Total	46.829	22.551	12.241	6.672	88.293

Fuente: Muestra ESPAC 2021



Cálculo de los factores de expansión

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

Es importante resaltar que el cálculo de factores de expansión se realiza exclusivamente para el marco de áreas, pues al ser el marco de lista un "censo" se tiene la información de la población.

Probabilidad de selección

Probabilidad de selección de primera etapa

La probabilidad de selección de primera etapa para el i -ésimo segmento muestreado dentro de la zona de selección h queda definido por:

$$PPE = \frac{n_h}{N_h}$$

Donde:

n_h Número de segmentos en la muestra en la zona de selección h ⁸

N_h Número de segmentos en el marco de áreas en la zona de selección h

Probabilidad de selección de segunda etapa

Recordando que los segmentos de estrato 04 fueron divididos en cuatro subsegmentos de igual superficie, exceptuando los segmentos de la provincia de Los Ríos que se levantaron por completo, y se escogió únicamente dos subsegmentos de 144 hectáreas (ha) para investigar, la probabilidad de inclusión de segunda etapa se define como:

$$PSE = \begin{cases} \frac{\text{número de subsegmentos del estrato 04 seleccionados}}{\text{número de subsegmentos del segmento de estrato 04}}, & \text{si es un segmento de estrato 04, excepto Los Ríos.} \\ 1, & \text{caso contrario} \end{cases}$$

Así, para obtener la superficie total del marco de áreas se calcula el factor de expansión teórico, el cual se define como el inverso de la probabilidad de selección de primera etapa por el inverso de la probabilidad de selección de segunda etapa.

$$fexp_{teo} = \frac{1}{PPE} \cdot \frac{1}{PSE} = \begin{cases} \frac{N_h}{n_h} \cdot \frac{4}{2} = 2 \cdot \frac{N_h}{n_h}, & \text{si es un segmento de estrato 04, excepto Los Ríos} \\ \frac{N_h}{n_h} \cdot 1 = \frac{N_h}{n_h}, & \text{caso contrario} \end{cases}$$

⁸ Por cada zona de selección se tomaron aleatoriamente 5 segmentos del total



Ajustes por rechazo

Para los segmentos que poseen rechazo parcial o total es necesario realizar un ajuste, con el fin de obtener la superficie total del marco de áreas. En la ESPAC se realizan estos ajustes a nivel de estrato y zona de selección, por lo que los segmentos pertenecientes al mismo estrato y zona de selección comparten el mismo ajuste.

- **Ajuste por rechazo parcial:**

$$aju_{rp} = \frac{\text{superficie total a investigar del segmento}}{\text{superficie investigada del segmento}}$$

- **Ajuste por rechazo total:**

$$aju_{rt} = \frac{\text{número de segmentos en la muestra en la zona de selección } h}{\text{número de segmentos efectivos (completos) en la zona de selección } h}$$

Ajuste por subdivisión

Se realiza un ajuste por subdivisión para recuperar la superficie que no se envió a investigar.

$$aju_{sub} = \begin{cases} 4, & \text{si el segmento fue subdividido una vez} \\ 8, & \text{si el segmento fue subdividido dos veces} \end{cases}$$

Así, el factor de expansión ajustado por cobertura se define como:

$$fexp_{aju} = fexp_{teo} \cdot aju_{rp} \cdot aju_{rt} \cdot aju_{sub}$$

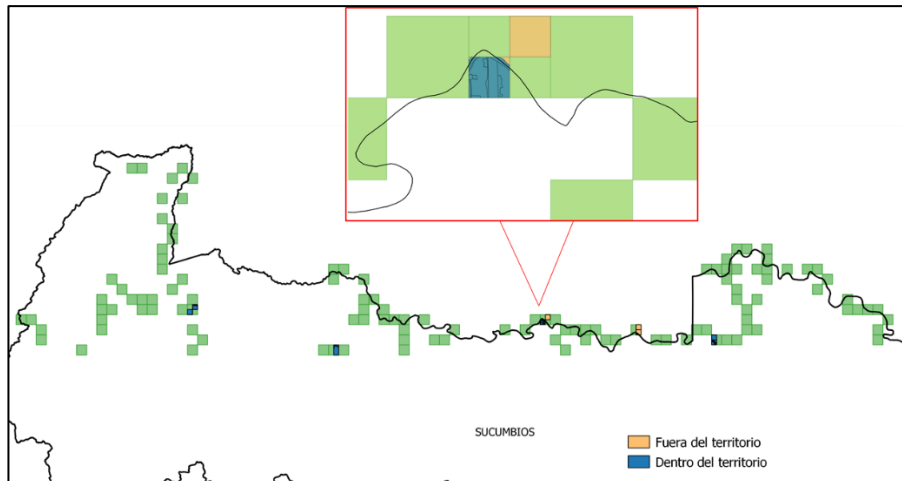
Ajuste por fuera del territorio

Para los segmentos seleccionados en la muestra que tienen superficie fuera del territorio continental ecuatoriano se realiza una corrección de la superficie por fuera del continente ecuatoriano, denotado por δ :

$$\delta = \frac{\text{superficie no cuadrada (determinada por los límites reales de las provincias)}}{\text{superficie expandida de la encuesta}}$$



Figura 7. Segmentos que se encuentran fuera del territorio continental ecuatoriano



Fuente: Marco de áreas y muestra ESPAC 2021

Ajuste por intersección

Existen segmentos cuya información fue levantada en el marco de áreas y en el marco de lista (intersecciones), por lo que la información proveniente del marco de áreas es innecesaria ya que la información obtenida en el marco de lista (censo) es completa. En la ESPAC 2021 existen 896 intersecciones.

$$\text{intersección} = \begin{cases} 1, & \text{si el segmento pertenece al marco de áreas y al marco de lista} \\ 0, & \text{si el segmento pertenece al marco de áreas o al marco de lista} \end{cases}$$

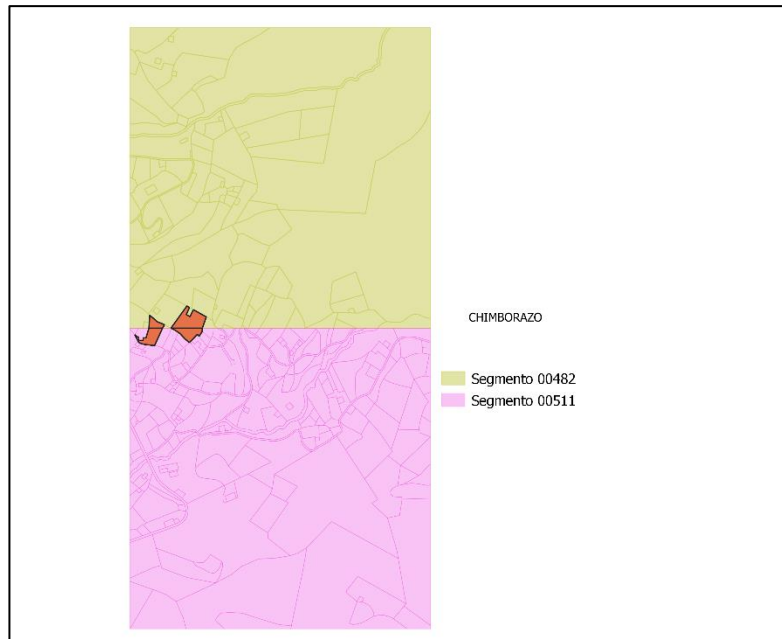
El ajuste por intersección junto con el ajuste de corrección de superficie permite obtener el factor de expansión calibrado:

$$fexp_{\text{calibrado}} = fexp_{\text{ajustado}} \cdot \delta \cdot \text{intersección}$$

El objetivo de la ESPAC es expandir la superficie contenida dentro de los segmentos seleccionados en la muestra; sin embargo, en campo se levanta la información del terreno completo y no de la parte del terreno que está contenida en el segmento (Ver **Figura 8**).



Figura 8: superficie de los terrenos declarados en campo



Fuente: Marco de áreas y muestra ESPAC 2021

Para esto es necesario realizar una corrección entre la superficie declarada en campo y la superficie dentro del segmento (denominada como superficie de cartografía). Esta corrección se denota por θ :

$$\theta = \frac{\text{sup_carto}}{\text{sup_campo}}$$

Donde:

sup_carto superficie de cartografía del terreno
sup_campo superficie declarada en campo del terreno

Así, el factor de expansión final es el resultado de multiplicar el factor de expansión calibrado por el ponderador, de forma que este factor permita expandir la superficie de campo de los terrenos a la superficie del marco de áreas.

$$fexp_{final} = fexp_{calibrado} \cdot \theta$$

Estimación de características

De acuerdo con el Manual sobre marcos maestros de muestreo para estadísticas agropecuarias (Strategy, 2015), la estimación del total poblacional para determinada característica o variable Y , se obtiene a partir del estimador de marco dual de Hartley de acuerdo a la siguiente fórmula:

$$\hat{Y}_H = \hat{Y}_a + \rho \hat{Y}_{ab} + (1 - \rho) \hat{Y}_l$$



Donde:

- \hat{Y}_H estimador de Hartley para el total poblacional de una determinada característica Y .
- \hat{Y}_a estimador total poblacional de la característica Y , correspondiente al marco de áreas.
- \hat{Y}_{ab} estimador del total poblacional de la característica Y , correspondiente a la intersección entre el marco de áreas y el marco de lista.
- \hat{Y}_l estimador del total poblacional de la característica Y correspondiente al marco de lista.
- ρ coeficiente de la combinación lineal entre el marco de áreas, marco de lista y su intersección.

Así mismo en (Strategy, 2015), se muestra que “se puede utilizar cualquier valor de ρ tal que $0 \leq \rho \leq 1$. Uno de ellos es una opción óptima ($\hat{\rho}_0$), en el sentido de que reduce al mínimo la varianza $Var(\hat{Y}_H)$. En la práctica, el valor de $\hat{\rho}_0$ puede ser muy cercano a cero, en estos casos es conveniente y ventajoso adoptar el estimador de examen preliminar (elegir $\rho = 0$).”

Por tanto, la expresión del estimador de examen preliminar resultante es:

$$\hat{Y}_S = \hat{Y}_a + \hat{Y}_l$$

$$\hat{Y}_S = \sum_{i \in S_a}^{N_a} fexp_{final_a} \cdot y_{i_a} + \sum_{i \in S_l}^{N_l} fexp_{final_l} \cdot y_{i_l}$$

Donde:

- $fexp_{final_a}$ es el factor de expansión final del terreno i del marco de áreas.
- $fexp_{final_l}$ es el factor de expansión final del terreno i del marco de lista.
- y_{i_a} característica o variable que se está investigando del terreno i del marco de áreas.
- y_{i_l} característica o variable que se está investigando del terreno i del marco de lista.

La fórmula anterior puede simplificarse como:

$$\hat{Y}_S = \sum_{i \in S}^N fexp_{final} \cdot y_i$$

Donde:

- $fexp_{final}$ factor de expansión final del terreno i , tanto para el marco de áreas como para el marco de lista.
- y_i característica o variable del terreno i que se está investigando, tanto para el marco de áreas como para el marco de lista.

Error de muestreo

Según (Strategy, 2015), el error de muestreo asociado a la estimación de una característica se calcula a través del estimador de marco dual de examen preliminar de Hartley:

$$\hat{Y}_S = \hat{Y}_a + \hat{Y}_l$$

Dado que los dominios son mutuamente excluyentes, y que los estimadores utilizan información de diferentes marcos, la varianza del estimador de examen preliminar es:

$$Var(\hat{Y}_S) = Var(\hat{Y}_a) + Var(\hat{Y}_l)$$



Donde:

$Var(\hat{Y}_a)$ es el estimador de la varianza del estimador del total poblacional de la variable Y del marco de áreas.

$Var(\hat{Y}_l)$ es el estimador de la varianza del estimador del total poblacional de la variable Y del marco de lista.

La varianza del estimador del total poblacional del marco de lista $Var(\hat{Y}_l)$ es igual a cero debido a que se levanta la información de todos los productores, es decir, el tamaño de muestra es igual al de la población lo que anula la varianza.

Así, la varianza del estimador de examen preliminar será:

$$Var(\hat{Y}_S) = Var(\hat{Y}_a)$$

El estimador de la varianza del estimador del total poblacional de la variable Y se calcula a partir de un muestreo probabilístico aleatorio simple de conglomerados en una etapa de acuerdo a la siguiente fórmula:

$$Var(\hat{Y}_S) = Var(\hat{Y}_a) = \sum_{h=1}^H M_h^2 (1 - f_h) \frac{\hat{S}_h^2}{m_h}$$

Donde:

M_h número de segmentos en el marco de áreas en la zona de selección h .

m_h número de segmentos en la muestra en la zona de selección h .

f_h fracción global de muestreo.

$$f_h = \frac{m_h}{M_h}$$

\hat{S}_h^2 varianza estimada de la característica Y en los segmentos de la zona de selección h

$$\hat{S}_h^2 = \frac{1}{m_h - 1} \sum_{i=1}^m (y_i - \bar{y}_h)^2$$

Así, el coeficiente de variación para la estimación de una determinada característica se define como la división entre la raíz de la varianza del estimador del total poblacional (error estándar) y la suma de los totales poblacionales estimados del marco de áreas y del marco de lista:

$$CV(\hat{Y}_S) = \frac{\sqrt{Var(\hat{Y}_S)}}{\hat{Y}_S}$$

$$CV(\hat{Y}_S) = \frac{\sqrt{Var(\hat{Y}_S)}}{\hat{Y}_a + \hat{Y}_l}$$



Bibliografía

- Ambrosio, L. (2017). *Curso de muestreo con marcos múltiples: Estimación con marcos sencillos. Muestreo Aleatorio Simple sin información auxiliar*. Madrid: España: Escuela Técnica Superior de Ingenieros Agrónomos.
- Instituto Nacional de Estadística y Censos. (2015). *Metodología de la Encuesta de Superficie y Producción Agropecuaria Continua ESPAC 2014*. Quito.
- Kish, L. (1972). *Muestreo de encuestas*. México: Trillas.
- Strategy, G. (2015). *Manual sobre Marcos Maestros de Muestreo para las Estadísticas Agropecuarias: elaboración del marco, diseño de la muestra y estimación*.



Anexos

Anexo 1. Distribución por provincia del número de productos del marco de lista

Provincia	Arroz	Aves	Banano	Brócoli	Cacao	Café	Caña de azúcar para azúcar	Flores	Ganado porcino	Ganado vacuno	Maíz duro seco	Mango	Maracuyá	Palma africana	Quinua
Azuay	0	104	0	0	9	14	0	24	9	22	0	0	0	0	0
Bolívar	0	16	0	0	18	2	0	0	17	12	0	0	8	0	0
Cañar	6	40	12	0	14	0	29	5	13	8	0	0	0	0	0
Carchi	0	52	0	0	0	23	0	35	10	75	0	0	0	0	0
Cotopaxi	0	56	10	43	4	0	0	221	9	130	0	0	0	5	0
Chimborazo	0	109	0	0	0	0	0	8	21	43	0	0	0	0	0
El Oro	0	367	35	0	69	9	0	0	125	146	0	0	13	0	0
Esmeraldas	0	12	0	0	23	3	0	0	32	556	0	0	0	286	0
Guayas	115	146	134	0	169	36	90	7	49	359	28	45	11	17	0
Imbabura	0	107	0	0	0	22	16	75	6	142	0	0	0	12	26
Loja	0	59	0	0	0	17	4	0	8	10	6	0	0	0	0
Los Ríos	78	52	172	0	29	5	6	0	19	103	48	0	15	76	0
Manabí	0	304	3	0	52	40	0	0	50	796	9	0	18	1	0
Morona Santiago	0	31	0	0	0	2	0	0	12	7	0	0	0	0	0
Napo	0	23	0	0	3	0	0	0	8	13	0	0	0	0	0
Pastaza	0	76	0	0	0	4	0	0	4	1	0	0	0	0	0
Pichincha	0	350	0	0	0	35	0	652	40	545	0	0	0	25	0
Tungurahua	0	384	0	0	0	0	0	4	6	16	1	0	0	0	0
Zamora Chinchipe	0	32	0	0	0	0	0	0	8	7	0	0	0	0	0
Sucumbíos	0	8	0	0	0	33	4	0	18	2	0	0	0	54	0
Orellana	0	64	0	0	3	21	0	0	0	19	0	0	0	54	0
Santo Domingo de los Tsáchilas	0	191	4	0	1	0	0	0	64	260	0	0	67	66	0
Santa Elena	0	10	3	0	0	8	0	1	8	0	15	0	0	0	0
Total	199	2.593	373	43	394	274	149	1.032	536	3.272	107	45	132	596	26

Fuente: Marco de lista ESPAC 2021



@ecuadorencifras



@ecuadorencifras



@InecEcuador



t.me/euadorencifras



INEC/Ecuador



INECEcuador



INEC Ecuador