



# METODOLOGÍA ESPAC

2016



**Metodología de la Encuesta de Superficie y Producción Agropecuaria Continua 2016**

**Dirección responsable de la información estadística y contenidos:**

**DIRECCIÓN DE ESTADÍSTICAS AGROPECUARIAS Y AMBIENTALES**

**Realizadores:**

**Unidad de Estadísticas Agropecuarias**

**Corrección de textos:**

**DIRECCIÓN DE COMUNICACIÓN SOCIAL**

## 1. INTRODUCCIÓN

La información agropecuaria del Ecuador se obtiene a través de la Encuesta de Superficie y Producción Agropecuaria Continua (ESPAC), este proyecto viene ejecutándose desde el año 2002 hasta la actualidad; desde su primer diseño y construcción de los marcos muestrales, no ha existido una actualización de los mismos.

La desactualización de los marcos muestrales, ha provocado que la estimación de los resultados pierda representatividad de áreas geográficas y productos, además no permitía obtener información desagregada por provincia para la región oriental.

Bajo éstos antecedentes, el INEC decidió realizar la actualización del marco muestral, lo que permite mejorar los niveles de estimación de los resultados agropecuarios provistos por la ESPAC, con la finalidad principal de obtener y producir datos que midan de manera permanente la dinámica del sector agropecuario de forma científica, moderna, eficiente y con innovación tecnológica.

Para la ESPAC 2014, el INEC realizó una actualización del marco maestro de muestreo utilizado para la encuesta. Esta actualización implicó la utilización de información cartográfica digital, la cual sirvió de insumo para generar una nueva estratificación.

En el año 2016 se realizó la investigación de 5.638 segmentos. Los resultados tienen representatividad a nivel provincial.

## 2. MARCO MAESTRO DE MUESTRO Y DISEÑO DE LA MUESTRA

Se trata de un Marco Múltiple de Muestreo (MMM), que combina marco de áreas (MM) y marco de lista (ML).

### 2.1 Marco de áreas

El material cartográfico de base utilizado para la construcción del marco es el Mapa de Usos del Suelo elaborado por el MAGAP y actualizado durante el período 2012-2014. Este mapa cubre el 85% del territorio nacional, aunque a diferentes escalas: 1:25000, 1:50000 y 1:100000. El 15% restante se cubre con imágenes de satélite.

### 2.2 Estratificación

El territorio nacional se estratifica en dos estratos primarios: (i) las zonas urbanas densamente pobladas y (ii) el resto del territorio, incluyendo áreas agropecuarias, forestales, agua y otras áreas rurales con baja densidad de población. Del estrato (ii) se separan solo los cuerpos de agua. El resto (áreas agropecuarias y forestales, reservas naturales y otras áreas no agropecuarias) se estratifican en cuatro estratos, según la proporción de superficie cultivada.

Los límites entre estratos son los estándares recomendados en FAO (1996,1998)]: áreas con una proporción mayor del 60% (Estrato1 y 2); áreas con una proporción de superficie cultivada entre 20% y 60% (Estrato 3) y áreas con una proporción de superficie cultivada inferior al 20% (Estrato 4). El Estrato 1, se divide en dos estratos, atendiendo al tamaño de los terrenos (parcelas o campos): terrenos pequeños (Estrato 1) y terrenos grandes (Estrato 2).

En primer lugar se segmenta todo el territorio en una retícula de celdillas cuadradas de 576 hectáreas. Esta segmentación se realiza sobre el mapa de usos y se calcula el porcentaje de tierras cultivadas (incluyendo las praderas cultivadas) en cada celdilla. Las celdillas se clasifican en los estratos 1 al 4, en función de ese porcentaje. Las celdillas del Estrato 1 se clasifican a su vez en dos subestratos: Sierra (Estrato 01) y el resto (Estrato 02).

### 2.3 Tamaño del segmento

Se consideran segmentos de límites geométricos. El criterio usual para elegir el tamaño del segmento consiste en fijar el número idóneo de unidades de observación dentro de sus límites y ese número idóneo se ha fijado en 10 terrenos.

## 2.4 La muestra de segmentos y la muestra de agricultores

A partir del marco de áreas se selecciona una muestra aleatoria de segmentos. Sobre los segmentos de la muestra se observa parte de la información requerida (la superficie de los cultivos y demás usos del suelo y algunas variables ambientales, tales como tipos de riego o técnicas de laboreo). Mediante entrevistas a los agricultores que cultivan las tierras dentro de los límites del segmento, se recoge toda la información de las variables que contiene el cuestionario ESPAC.

## 2.5 Tamaño de la muestra de segmentos

Las tasas de muestreo consideradas son, respectivamente, del 3%, 2%, 1% y 0,5% para los estratos 1 (Superficie cultivada >60% y campos pequeños (Sierra)), 2 (Superficie cultivada >60% y campos grandes), 3 (20% < Superficie cultivada <60%) y 4 (Superficie cultivada <20%).

Se espera que esos tamaños de muestra sean suficientes para que las precisiones de las estimaciones de los cultivos mayoritarios estén dentro de los límites de tolerancia estándares (15% de coeficiente de variación). Sin embargo, es de temer que la precisión de las estimaciones de los cultivos muy localizados en el espacio (como los hortícolas, y algunos cultivos industriales) y la ganadería, supere los límites de tolerancia: para mejorar esas precisiones se usarán marcos múltiples, complementando la muestra de áreas con una lista de los mayores productores.

## 2.6 Selección de la muestra de segmentos.

### 2.6.1. Réplicas. Rotación de la muestra

El muestreo sistemático es eficiente, pero el muestreo por zonas es más simple y será el que utilicemos. La muestra se selecciona por zonas en 5 réplicas, cada una de tamaño igual a una quinta parte de la muestra del estrato.

El número de zonas en un estrato es  $m_h = n_h / r_h$ , donde  $n_h$  es el tamaño de la muestra y  $r_h = 5$  es el número de réplicas.

La primera réplica se forma seleccionando un segmento de cada zona, con probabilidades iguales. La segunda réplica se forma seleccionando un nuevo segmento de cada zona, sin reposición y con probabilidades iguales, y así sucesivamente hasta completar las 5 réplicas.

Cuando se prevé repetir una encuesta periódicamente, es necesario considerar en el momento del diseño un plan de rotación de la muestra, de modo que en cada período se encuentren en la muestra, junto a elementos seleccionados en períodos anteriores, otros

nuevos incluidos en el período en cuestión. La rotación de la muestra reduce el sesgo debido a la "fatiga del encuestado" y permite una mejor estimación de los cambios y de las tendencias de las características en estudio.

La rotación se lleva a cabo haciendo salir de la muestra en cada período una réplica e incluyendo otra nueva. Las réplicas que salen de la muestra después del primero y del segundo año de la vida del marco, pueden ser repetidas seis años después. Sin embargo, las que salen tras el tercero y cuarto año no vuelven a entrar. Las que entran durante los cuatro últimos años de la vida del marco están menos de cinco años.

Para la ESPAC 2014, se seleccionó un total de 5.638 segmentos en el marco de áreas y para el marco de lista se seleccionaron 3.969 unidades productivas, distribuidas en 14 estratos.

Estrato	Descripción
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	Porcinos
63	Flores
64	Aves
65	Quinua

Las Zonas No Delimitadas<sup>1</sup> El Piedrero y Las Golondrinas forman un único dominio de estudio independiente denominado Zonas No Delimitadas; para el año 2016 el área conocida como La Manga del Cura considerada hasta el 2015 como una zona no delimitada fue incluida en la muestra dentro de la provincia de Manabí, ya que mediante consulta popular que se realizó el 27 de septiembre de 2015 esta zona pertenece a la provincia en mención.

## 2.7 Recolección de datos

El soporte de la información a recoger en campo es la ortofoto y los cuestionarios, lo que permite recoger datos directamente observables sobre el terreno, como las superficies de los cultivos, tipo de riego y tipo de laboreo.

<sup>1</sup> Superficie de territorio ecuatoriano que no está asignado a ninguna de sus provincias.

Se dispone de ortofotos para cubrir alrededor del 85% del territorio. Para el 15% restante no se dispone de más soporte gráfico que el mapa de usos del suelo. Una opción es dibujar un croquis de los segmentos de la muestra para los que no se dispone de ortofoto, incluyendo el croquis de los terrenos existentes dentro del segmento y tomar en campo con el GPS las coordenadas del mínimo número de puntos requeridos para poder calcular en gabinete la superficie de cada terreno.

## 2.7.2. Levantamiento de información

### a) Material gráfico: fotografías, mapas e imágenes de satélite.

Los límites de los segmentos de la muestra se imprimen en la ortofoto a escala 1:5.000 (o superior) o directamente sobre la ortofoto.

En la cartulina figura impreso el número de la hoja del Mapa de Usos del Suelo en la que se encuentra el segmento. También contiene una representación reducida de esta última, con el segmento seleccionado. Esta información es útil para la localización del segmento sobre el terreno.

### b) Identificación de los límites del segmento.

Una vez situado sobre el segmento, el encuestador empieza por asegurarse de que los límites del segmento son identificables al 100 %. Los segmentos con límites no identificables serán devueltos al supervisor.

### c) Delimitación de terrenos.

Sobre la ortofoto se delimitan los terrenos, rotulando los límites entre éstos. Se entiende por terreno o "unidad de uso del suelo" *una superficie continua ocupada por el mismo cultivo o uso del suelo*, aunque esté integrada por varios terrenos catastrales pertenecientes a distintos propietarios. Cualquier cambio de cultivo o uso del suelo será considerado como límite de terreno.

Si los límites del segmento son accidentes naturales o artificiales del terreno, todas las parcelas delimitadas dentro de un segmento serán enteras, no habrá parcelas partidas. Si los límites son geométricos las parcelas quedan partidas por los límites. En todo caso hay que asegurarse de que todos los recintos o terrenos quedan cerrados.

Puede haberse producido cambios en los límites de parcela que se observan en la fotografía, respecto de los realmente observados sobre el terreno. Es necesario asegurarse de que los límites entre terrenos trazados sobre la ortofoto son los actuales. Es conveniente el disponer de los límites de terrenos en soporte informático trasladar los de un año sobre el material a utilizar el año siguiente.

#### **d) Numeración de terrenos**

Los terrenos se numeran de forma correlativa, siguiendo el mismo criterio ya utilizado para numerar los segmentos. Por ejemplo, la numeración de los terrenos puede realizarse, empezando por la situada más al noroeste y siguiendo de forma serpenteante continua, de oeste a este y posteriormente de este a oeste, hasta completar todo el segmento.

#### **e) Identificación de cultivos o usos del suelo.**

El cultivo o uso del suelo en cada uno de los terrenos del segmento se identifican y se registran en el cuestionario, asignándole el código correspondiente, según figure en la codificación de cultivos y usos del suelo que se proporciona al encuestador.

#### **f) El Cuestionario.**

El cuestionario para la recogida de datos sobre superficies consta en su cabecera de un espacio para anotar el código completo del segmento, el nombre del encuestador y la fecha de realización de la encuesta.

El resto del cuestionario lo constituyen varios capítulos que contienen varias filas y columnas. Los datos se registran por filas. En cada fila se registra la información relativa al cultivo o uso del suelo en un terreno.

- **CAPITULO 1.-** Características generales de los terrenos con y sin actividad agropecuaria.
- **CAPITULO 2.-** Superficie de los terrenos.
- **CAPITULO 3.-** Características de uso del suelo.
- **CAPITULO 4.-** Cultivos permanentes.
- **CAPITULO 5.-** Cultivos transitorios.
- **CAPITULO 6.-** Árboles o plantas permanentes dispersos.
- **CAPITULO 7.-** Floricultura.
- **CAPITULO 8.-** Ganado vacuno (bovino).
- **CAPITULO 9.-** Ganado porcino.
- **CAPITULO 10.-** Ganado ovino.
- **CAPITULO 11.-** Otras especies de ganado.
- **CAPITULO 12.-** Aves de campo y de planteles avícolas.
- **CAPITULO 13.-** Empleo en los terrenos.
- **MÓDULO PRÁCTICAS AMBIENTALES.-** Prácticas en la agricultura.
- **CAPITULO 14.-** Datos adicionales del informante.
- **DATOS** del personal participante.



### **g) Identificación de tramos**

Los terrenos pertenecientes a una misma Unidad de producción deben ser agrupados en un mismo tramo y el conjunto de tramos en un segmento debe ser identificado.

## **2.8 Control de calidad de los datos recogidos.**

Se realiza un control del trabajo realizado en campo y otro en gabinete del material que sirve de soporte a la información.

### **a) Control de campo.**

Consiste en una inspección por muestreo del trabajo de campo realizado por cada uno de los encuestadores. Del conjunto de segmentos investigados por cada encuestador se selecciona una muestra. La delimitación de terrenos en cada uno de los segmentos de esa muestra y la asignación de usos a cada una de los terrenos, es inspeccionado por un supervisor. También la identificación de los límites.

### **b) Control en gabinete.**

En gabinete se procede a un control, primero del material que sirve de soporte a la información recogida en campo (las ortofoto y los cuestionarios), y después de la propia información, previa traslación de ésta a un programa informático.

El control del material se realiza manualmente y consiste en verificar la exacta correspondencia entre el código de identificación de la ortofoto y el cuestionario, así como la coincidencia de ambos con el código del segmento al que corresponden.

Un primer control de la información gráfica y alfanumérica, también realizado de forma manual, consiste en verificar (i) que el número de terrenos delimitados sobre la ortofoto coincide con el del cuestionario, (ii) que a todos los terrenos se les ha asignado un número y que éste es único tanto en ortofoto como en el cuestionario, (iii) que no existen recintos abiertos de entre los delimitados en el segmento y (iv) que en los cuestionarios, los códigos de cultivo coinciden con la denominación del cultivo.

Los errores detectados que es posible corregir en gabinete se corrigen, previa consulta a los encuestadores. El material cuya información errónea no ha sido posible corregir en gabinete será devuelto a los encuestadores para su revisión.

Un segundo y definitivo control de la información gráfica y alfanumérica será realizado por procedimientos informáticos.

## 2.9 Estimación

Las características directamente observadas sobre el terreno, tales como la superficie de los cultivos y demás usos del suelo, la superficie regada con cada tipo de riego y la superficie labrada con cada técnica de laboreo, se estiman a partir de los datos recogidos en la muestra de segmentos.

Las variables relativas a la economía de las unidades de producción agropecuaria, incluyendo la ganadería, y las variables relativas al hogar, han de ser estimadas a partir de los datos recogidos mediante entrevistas a los agricultores y sus hogares.

### 2.9.1 Estimaciones basadas en la muestra de segmentos.

El total  $Y$  de la variable en estudio se expresa como suma de los totales,  $Y_h$ , de cada estrato:

$$Y = \sum_{h=1}^L Y_h$$

#### 2.9.1 Estimadores de dominio

La ESPAC es una encuesta nacional: se ha diseñado para obtener estimaciones de la producción agropecuaria a nivel nacional, con un grado de precisión aceptable (coeficiente de variación inferior al 15%). No obstante, es posible obtener también estimaciones (llamadas de dominio) para ámbitos administrativos más reducidos (cantones y provincias) y/o para ciertos subgrupos de la población de UPAs, tal como la agricultura familiar (cuántos son, cuál es su renta, cuáles son sus condiciones de vida...) a partir de la muestra maestra diseñada para la ESPAC.

Las estimaciones de dominio se basan solo en una parte de la muestra (la parte de la muestra total que cae dentro de los límites del dominio) y de ahí que su precisión sea inferior a la de las estimaciones obtenidas a nivel nacional. Si se requiere mejorar las estimaciones de dominio, es necesario complementar la muestra original con una muestra adicional seleccionada dentro del dominio en cuestión.

##### 2.9.1.1 Estimadores de dominio basados en la muestra original

Se define la variable auxiliar:

$$y'_{hij} = \begin{cases} y_{hij} & \text{si } i \in \text{al dominio } D \\ 0, & \text{en otro caso} \end{cases}$$

$$[1]: \hat{Y}_D = \sum_{h=1}^L \hat{Y}_{Dh}, \text{ donde } \hat{Y}_{Dh} = \sum_{j=1}^{m_h} \hat{Y}_{Dhj}$$

Y se usa el estimador definido en  $\hat{Y}_{Dh}$  es el estimador

Del total de la variable de muestreo en el estrato h dentro del dominio  $\{Y_{Dh}, e \hat{Y}_{Dhj}\}$

Es el estimador del total de la variable en estudio en la j-ésima zona del estrato h-ésimo, dentro del dominio

$$\hat{Y}_{Dhj} = \frac{N_{hj}}{r_h} \sum_{i=1}^{r_h} y'_{hij}$$

### Varianza.

La varianza del estimador del total de dominio viene dada por la siguiente expresión [2],

$$V(\hat{Y}_D) = \sum_{h=1}^L V(\hat{Y}_{Dh}), \text{ donde } V(\hat{Y}_{Dh}) = \sum_{j=1}^{m_h} V(\hat{Y}_{Dhj}) \text{ y}$$

$$V(\hat{Y}_{Dhj}) = N_{hj}^2 \left(1 - \frac{r_h}{N_{hj}}\right) \frac{1}{r_h} \frac{1}{N_{hj} - 1} \sum_{i=1}^{r_h} (y'_{hij} - \bar{y}'_{hj})^2, \text{ donde } \bar{y}'_{hj} = \frac{1}{N_{hj}} \sum_{i=1}^{r_h} y'_{hij}$$

$$\text{y se estima por } \hat{V}(\hat{Y}_D) = \sum_{h=1}^L \hat{V}(\hat{Y}_{Dh}), \text{ donde } \hat{V}(\hat{Y}_{Dh}) = \sum_{j=1}^{m_h} \hat{V}(\hat{Y}_{Dhj}),$$

$$\text{y, } \hat{V}(\hat{Y}_{Dhj}) = N_{hj}^2 \left(1 - \frac{r_h}{N_{hj}}\right) \frac{1}{r_h} \frac{1}{r_h - 1} \sum_{i=1}^{r_h} (y'_{hij} - \bar{y}'_{hj})^2$$

$$\text{donde } \bar{y}'_{hj} = \frac{1}{r_h} \sum_{i=1}^{r_h} y'_{hij}$$

La varianza del estimador de la media es

$$V(\hat{Y}_D) = \frac{1}{N_D^2} V(\hat{Y}_D)$$

$$\hat{V}(\hat{Y}_D) = \frac{1}{N_n^2} \hat{V}(\hat{Y}_D)$$

Y se estima por Los coeficientes de variación del total y de la media coinciden.

## 2.10 Estimadores de dominio basados en muestras múltiples

Si la precisión de las estimaciones obtenidas usando los estimadores presentados en el epígrafe 7.1 no alcanza el nivel requerido, entonces es necesario complementar la muestra original con una muestra adicional seleccionada dentro del dominio en cuestión.

En este epígrafe se considera el problema de estimar el total de dominio de una variable de muestreo,  $Y_D$ , a partir de dos muestras: la inicial de la ESPAC,  $A$ , y una complementaria,  $B$ . Las muestras se seleccionan sucesivamente de sendos marcos  $\{A, B\}$ : primero se muestrea  $A$  y después  $B$ . Los tamaños de los estratos en ambos marcos son conocidos a priori.

$$\{N_h^{(A)}; h=1, 2, \dots, L\} \text{ con } \sum_{h=1}^L N_h^{(A)} = N_A \text{ y } \{N_h^{(B)}; h=1, 2, \dots, L\} \text{ con } \sum_{h=1}^L N_h^{(B)} = N_B$$

$N_A$  es el número de unidades de muestreo en el marco  $A$  y  $N_B$  es el número de unidades del marco  $B$  (es el tamaño del dominio  $D$ ). Generalmente, el marco  $B$  es la parte del marco  $A$  dentro de los límites del dominio  $D$ : por ejemplo, si el dominio  $D$  es un cantón o una provincia,  $N_B$  es el número de segmentos del marco  $A$  que están dentro del dominio  $D$ .

$$; \{h^{(A)}; h=1, 2, \dots, L\} \text{ y } \{h^{(B)}; h=1, 2, \dots, L\}$$

Pero los criterios de estratificación de  $B$  pueden ser distintos de los de  $A$  y, en ese caso, los estratos coincidirán:

Por ejemplo, si el dominio  $D$  es el conjunto de UPAs que integran la agricultura familiar y el marco  $B$  es un registro de UPAs familiares,  $N_B$  es el número de UPAs incluidas en el marco  $B$  y los criterios de estratificación de  $B$  pueden ser distintos de los de  $A$ .

Se asume que  $B \subset A$ , por lo que seleccionar una muestra aleatoria estratificada a partir del marco  $B$  es equivalente a seleccionar una muestra aleatoria estratificada de  $A$ , de modo que uno de los estratos contiene los elementos de la población no incluidos en el marco  $B$  y este estrato se muestrea a tasa nula [Bankier (1986)]. Sin pérdida de generalidad, supondremos que el estrato que contiene los elementos de la población no incluidos en el marco  $B$  es  $(L+1)^{(B)}$

Si los criterios de estratificación de  $B$  son los mismos que los de  $A$ , entonces.

$$N_{h^{(A)}h^{(B)}} = 0; \forall \left\{ \left( h^{(A)}, h^{(B)} \right) \middle| h^{(A)} \neq h^{(B)} \right\} \text{ y } N_{h^{(B)}} = N_{h^{(A)}h^{(B)}}; \forall h^{(A)} = h^{(B)} \text{ y}$$

$$N_{h^{(A)}h^{(B)}} + N_{h^{(A)},(L+1)^{(B)}} = N_{h^{(A)}}; \forall \{h=1, 2, \dots, L\}$$

### El estimador del total

Sea  $y_{h^{(A)}h^{(B)}i}$  el valor de la variable en estudio en el  $i$ ésimo – elemento de la población en el estrato  $H^{(A)}H^{(B)}$

El total de dominio

$$Y_D = \sum_{h^{(A)}=1}^{L^{(A)}} \sum_{h^{(B)}=1}^{L^{(B)}} Y_{h^{(A)}h^{(B)}} \text{ ($$

Donde:

$$Y_{h^{(A)}h^{(B)}} = \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} y_{h^{(A)}h^{(B)}i} \text{ es el total del dominio en el estrato } h^{(A)}h^{(B)}.$$

### Varianza del estimador

La varianza de  $\hat{Y}_B$  viene dada por la siguiente expresión:

$$V\hat{Y}_B = \sum_{h^{(A)}=1}^{L^{(A)}} \sum_{h^{(B)}=1}^{L^{(B)}} V\hat{Y}_{h^{(A)}h^{(B)}}$$

donde,

$$V\hat{Y}_{h^{(A)}h^{(B)}} = \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} \frac{1 - \varpi_{h^{(A)}h^{(B)}i}}{\varpi_{h^{(A)}h^{(B)}i}} y_{h^{(A)}h^{(B)}i}^2 + \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} \sum_{i'(\neq i)=1}^{N_{h^{(A)}h^{(B)}}} \frac{\varpi_{h^{(A)}h^{(B)}i'} - \varpi_{h^{(A)}h^{(B)}i} \varpi_{h^{(A)}h^{(B)}i'}}{\varpi_{h^{(A)}h^{(B)}i} \varpi_{h^{(A)}h^{(B)}i'}} y_{h^{(A)}h^{(B)}i} y_{h^{(A)}h^{(B)}i'}$$

o bien,

$$V\hat{Y}_{h^{(A)}h^{(B)}} = \frac{1}{2} \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} \sum_{i'(\neq i)=1}^{N_{h^{(A)}h^{(B)}}} \left( \varpi_{h^{(A)}h^{(B)}i'} - \varpi_{h^{(A)}h^{(B)}i} \varpi_{h^{(A)}h^{(B)}i'} \right) \left( \frac{y_{h^{(A)}h^{(B)}i}}{\varpi_{h^{(A)}h^{(B)}i}} - \frac{y_{h^{(A)}h^{(B)}i'}}{\varpi_{h^{(A)}h^{(B)}i'}} \right)^2$$

### Estimador de la varianza del estimador

Siempre que  $\varpi_{h^{(A)}h^{(B)}i'} > 0, \forall i, i'$ , un estimador insesgado de  $V\hat{Y}_{h^{(A)}h^{(B)}}$  es:

$$\hat{V}\hat{Y}_{h^{(A)}h^{(B)}} = \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} \frac{1 - \varpi_{h^{(A)}h^{(B)}i}}{\varpi_{h^{(A)}h^{(B)}i}} \frac{y_{h^{(A)}h^{(B)}i}^2}{\varpi_{h^{(A)}h^{(B)}i}} + \sum_{i=1}^{N_{h^{(A)}h^{(B)}}} \sum_{i'(\neq i)=1}^{N_{h^{(A)}h^{(B)}}} \frac{\varpi_{h^{(A)}h^{(B)}i'} - \varpi_{h^{(A)}h^{(B)}i} \varpi_{h^{(A)}h^{(B)}i'}}{\varpi_{h^{(A)}h^{(B)}i} \varpi_{h^{(A)}h^{(B)}i'}} \frac{y_{h^{(A)}h^{(B)}i} y_{h^{(A)}h^{(B)}i'}}{\varpi_{h^{(A)}h^{(B)}i}}$$

En ocasiones, esa varianza resulta negativa. En ese caso se toma como estimación de la varianza el valor cero.

Con muestras simples,  $\varpi_{h^{(A)h^{(B)}}i} = 1 - (1 - f_h^{(A)})(1 - f_h^{(B)}) = \varpi_{h^{(A)h^{(B)}}i}; \forall i$  y

$\varpi_{h^{(A)h^{(B)}}i'} = \varpi_{h^{(A)h^{(B)}}i} \times \varpi_{h^{(A)h^{(B)}}i'}$  donde  $\varpi_{h^{(A)h^{(B)}}i'} = 1 - (1 - \tilde{f}_h^{(A)})(1 - \tilde{f}_h^{(B)}) = \tilde{\varpi}_{h^{(A)h^{(B)}}i'}; \forall i, i', y$

$\tilde{f}_h^{(A)} = \frac{n_h^{(A)} - 1}{N_h^{(A)} - 1}$  y lo mismo para  $\tilde{f}_h^{(B)}$ .

Esto es,

$$\hat{V}_{h^{(A)h^{(B)}}} = \frac{1 - \varpi_{h^{(A)h^{(B)}}}}{\varpi_{h^{(A)h^{(B)}}}} \frac{1}{\varpi_{h^{(A)h^{(B)}}}} \sum_{i=1}^{n_{h^{(A)h^{(B)}}}} y_{h^{(A)h^{(B)}}i}^2 + \frac{\varpi_{h^{(A)h^{(B)}}} \tilde{\varpi}_{h^{(A)h^{(B)}}} - \varpi_{h^{(A)h^{(B)}}} \varpi_{h^{(A)h^{(B)}}}}{\varpi_{h^{(A)h^{(B)}}} \varpi_{h^{(A)h^{(B)}}}} \frac{1}{\varpi_{h^{(A)h^{(B)}}} \tilde{\varpi}_{h^{(A)h^{(B)}}}} \sum_{i=1}^{n_{h^{(A)h^{(B)}}}} \sum_{i'(i) \neq i}^{n_{h^{(A)h^{(B)}}}} y_{h^{(A)h^{(B)}}i} y_{h^{(A)h^{(B)}}i'}$$

## 2.11 Post-estratificación

El ajuste o calibración de los factores de expansión consiste en la creación de un componente extra ( $\delta$ ) en la fórmula de cálculo, para poder cuadrar la población estimada por medio de la encuesta con un dato o un parámetro conocido

Este componente “delta” permitirá el ajuste de los factores de expansión a fin de que las estimaciones de las superficies muestrales presenten una mejor aproximación de la superficie real de las provincias.

Los límites de los dominios de estudio se ajustan al perfil de las UME, mas estos no corresponden a los límites reales de cada uno de ellos. Para tener concordancia en la superficie de los dominios de estudio, se multiplican los factores de expansión a nivel de dominio por un valor  $\delta \neq 0$ , donde

$$\delta = \frac{\text{Superficie provincia límites reales}}{\text{Superficie provincia marco ESPAC}}$$

Además, algunos segmentos presentaban superficie fuera del universo de estudio, perteneciente a Colombia, Perú o el Océano Pacífico. La utilización de la post-estratificación concilia esta particularidad con el objetivo de la encuesta.

**Gráfico 1.-** Segmento de investigación de ESPAC con superficie en otra provincia y fuera del universo de estudio.



### 3. SISTEMA DE ALMACENAMIENTO Y MANEJO DE DATOS

El Sistema de Almacenamiento y Manejo de datos deberá cumplir funciones básicas:

- Difundir las estadísticas oficiales, haciéndolas accesibles tan pronto como estén elaboradas, e identificándolas claramente según su procedencia y, en su caso, dando detalles del diseño de la encuesta.
- Proporcionar el marco idóneo para el almacenamiento de los resultados (agregados) de las encuestas y las observaciones muestrales (desagregadas), junto con su georreferenciación y, en su caso, los datos administrativos utilizados como información auxiliar.



[www.ecuadorencifras.gob.ec](http://www.ecuadorencifras.gob.ec)



@ecuadorencifras



INEC/Ecuador



Inec



INECEcuador



INEC Ecuador