

Muestreo de Encuestas de Hogares Con R Día 5

-
- INEC – Diseño Muestral
-

Diciembre - 2025

www.ecuadorencifras.gob.ec



• • • •
Enlace día 4



CONTENIDO

01

Factores de expansión

- Qué son y por qué se necesitan?
- Cálculo
- Evaluación

02

Estimación de características

- Ejercicio con ENDI

Factores de expansión

Factores de expansión

Qué son y por qué se necesitan

Por definición, son el **inverso de la probabilidad de inclusión** para cada una de las viviendas en la muestra, siendo un componente clave al producir estimadores poblacionales desde una muestra.

Sin embargo, puesto que la muestra capta viviendas categorizadas como rechazo, no elegibilidad y elegibilidad desconocida esta información puede utilizarse para ajustar las probabilidades de inclusión de la muestra resultante.

Sin el uso de ponderadores, las estimaciones pueden reflejar sólo matices de una muestra particular y pueden contener niveles significativos de sesgo.

Weighting Survey Data

How to do it right?



Factores de expansión

Cálculo

El diseño muestral de la ENDI es **bietápico estratificado por conglomerados**, por lo que para el cálculo de los factores de expansión se consideran las probabilidades de inclusión en cada una de las etapas, así como los respectivos ajustes por cobertura.



Factores de expansión

Probabilidad de primera etapa

En la primera etapa, se seleccionan UPM con una probabilidad proporcional al tamaño (número de viviendas ocupadas). Por lo que la probabilidad de inclusión de primera etapa de la UPM i en el estrato h es:

$$ppe_{i,h} := m_h \frac{N_i}{N_h}$$

donde:

- $ppe_{i,h}$: es la probabilidad de inclusión de primera etapa de la UPM i en el estrato h .
- m_h : es el número de UPM seleccionadas en el estrato h .
- N_i : el número de viviendas de la UPM i según el marco de muestreo.
- N_h : el número de viviendas en el estrato h según el marco de muestreo.

Factores de expansión

Ajuste de cobertura de primera etapa

Debido a diferentes motivos, se puede tomar la decisión de no investigar una UPM. De ser el caso, se debe realizar un ajuste en el que se tomen en cuenta únicamente las UPM que fueron investigadas, siendo:

$$a_{1h} := \frac{m_h}{m'_h}$$

donde m'_h es el número de UPM investigadas en el estrato h .

Factores de expansión

Probabilidad de segunda etapa

El enlistamiento permitió identificar las viviendas con niños menores a 5 años, la selección de segunda etapa se realiza sobre este grupo, siendo:

$$pse_i := \frac{k_i}{K_i}$$

donde:

- pse_i : probabilidad de segunda etapa de las viviendas en la UPM i .
- k_i : número de viviendas con niños menores a cinco años seleccionadas en la UPM i .
- K_i : número total de viviendas con niños menores a cinco años en la UPM i .

Factores de expansión

Factor de expansión teórico

El factor de expansión teórico es el inverso del producto de las probabilidades de selección:

$$d_{0i,h} := \frac{1}{ppe_{i,h} \cdot pse_i}$$

El factor de expansión ajustado por primera etapa es:

$$d_{1i,h} := a_{1h} \cdot d_{0i,h}$$

Factores de expansión

Ajuste por cobertura de viviendas

Cada vivienda muestreada se clasifica en uno de los cuatro conjuntos anteriormente descritos (*RE*, *NR*, *NE* y *ED*).

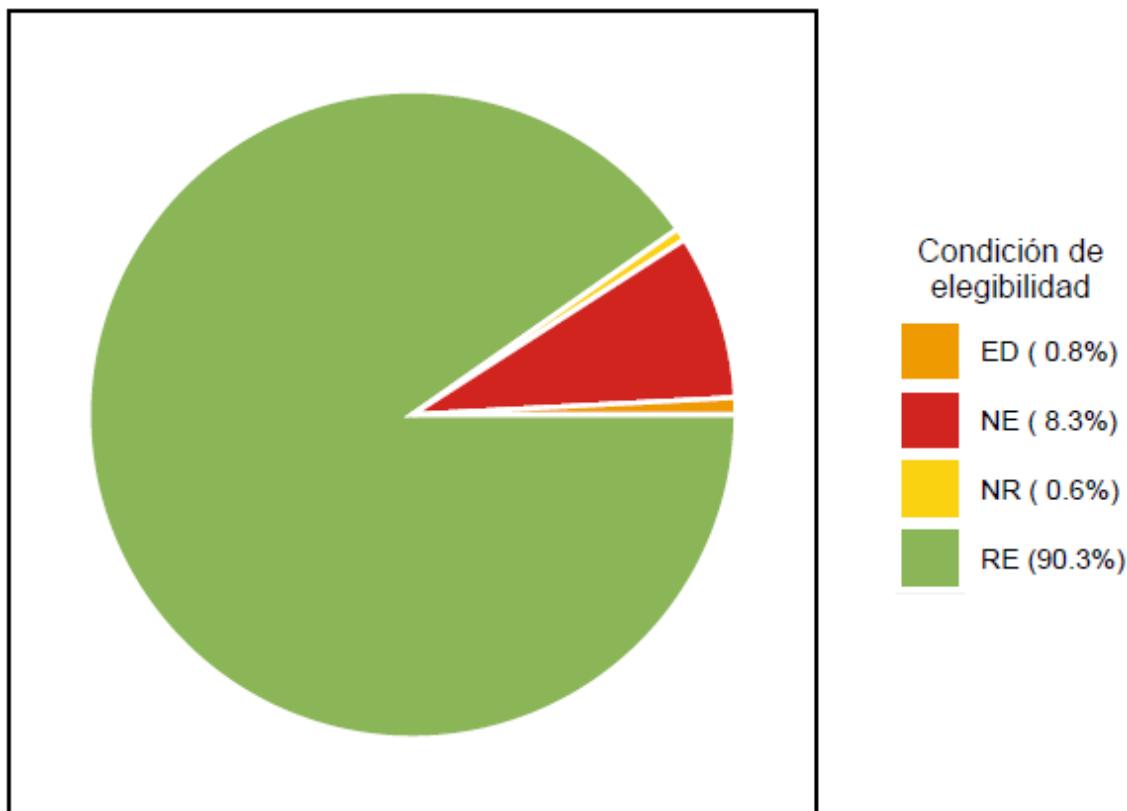
- Si existen unidades que son no elegibles en la muestra, esto es evidencia de que hay unidades no elegibles tanto sobre *ED* como sobre lo no muestreado. Diferentes decisiones se pueden realizar sobre cómo manejar las unidades en *ED*.
- Además, tanto los marcos como las muestras seleccionadas de ellos pueden contener unidades cuya elegibilidad no puede ser determinada.

Por lo tanto, se realizan ajustes por **elegibilidad desconocida** y **no respuesta** sobre las unidades elegibles para que estas se equiparen al total de la población elegible.

Factores de expansión

Ajuste por cobertura de viviendas

Para tal efecto, es necesario considerar la distribución inicial estimada de todas las viviendas con niños menores de 5 años por condición de elegibilidad considerando el factor ajustado por cobertura de primera etapa:



Factores de expansión

Ajuste por elegibilidad desconocida

Una forma de manejar las unidades en *ED* es distribuir el peso total de la muestra entre aquellas unidades cuyo estado de elegibilidad se conoce (es decir, sobre las unidades en *RE*, *NR* y *NE*).

Por lo general, la elegibilidad desconocida se trata mediante métodos simples en parte porque se sabe poco sobre los casos con elegibilidad desconocida y en parte porque la falta de respuesta se considera un problema más serio que debe recibir más atención.

La distribución de su peso utiliza un enfoque basado en clases, es decir, el ajuste se puede realizar a nivel de UPM, estrato o cualquier otro conjunto definido para tal efecto.

Para la ENDI se realizará el ajuste a nivel de estrato, basándonos en el análisis de la cobertura a nivel de UPM y viviendas, así como por la naturaleza de la población objetivo de la misma.

Factores de expansión

Ajuste por elegibilidad desconocida

Por lo tanto, el ajuste por elegibilidad desconocida se define como:

$$a_{2h} := \frac{\sum_{j \in s_h} d_{1j}}{\sum_{j \in s_h} d_{1j} - \sum_{j \in s_h \cap ED} d_{1j}}$$

donde:

- a_{2h} : ajuste por elegibilidad desconocida en el estrato h .
- s_h : muestra total de viviendas en el estrato h .
- d_{1j} : factor de expansión de la vivienda j .

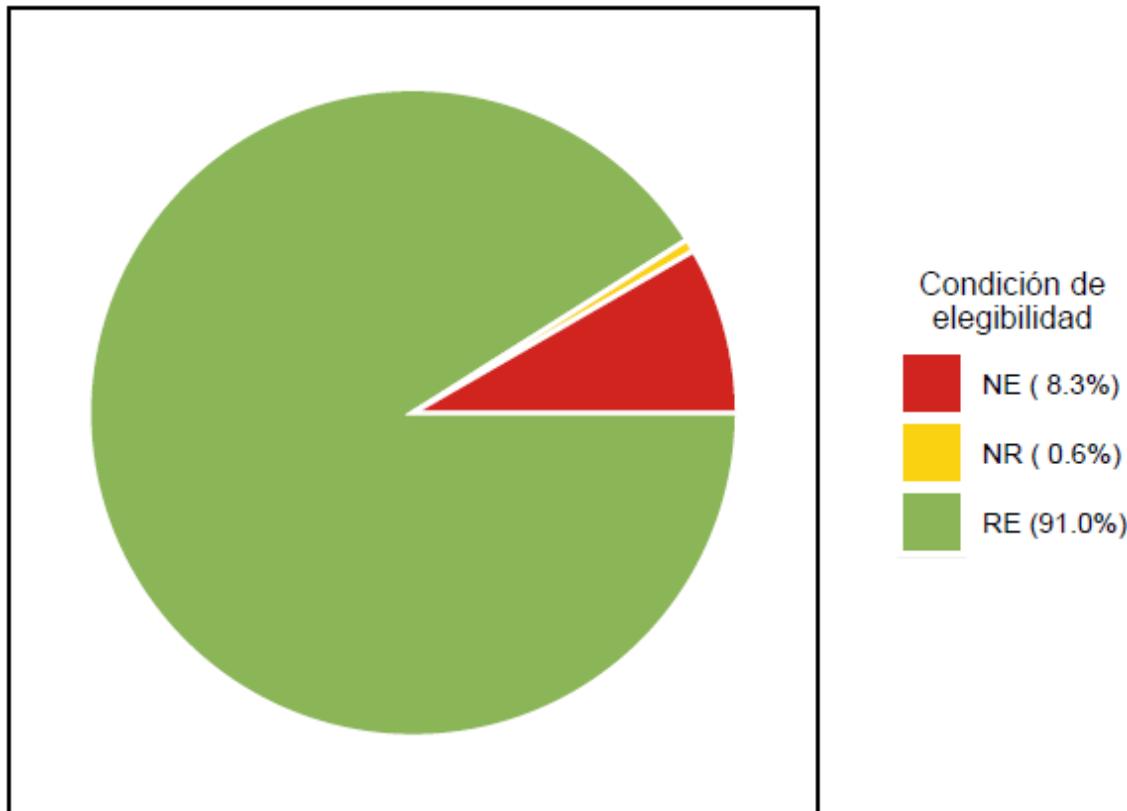
Por lo tanto el factor de expansión ajustado por elegibilidad desconocida (d_{2j}) queda definido de la siguiente manera:

$$d_{2j} := \begin{cases} a_{2h} \cdot d_{1j} & j \in s_h \cap RE, \\ a_{2h} \cdot d_{1j} & j \in s_h \cap NR, \\ a_{2h} \cdot d_{1j} & j \in s_h \cap NE, \\ 0 & j \in s_h \cap ED \end{cases}$$

Factores de expansión

Ajuste por cobertura de viviendas

Luego de aplicar este ajuste, las unidades pertenecientes a ED se distribuyen de manera proporcional entre las otras categorías, dando como resultado la distribución presentada a continuación:



Factores de expansión

Ajuste por no respuesta

El ajuste por no respuesta para la vivienda j del estrato h se calcula utilizando los factores de expansión ajustados por elegibilidad desconocida, quedando determinado por:

$$a_{3h} := \frac{\sum_{j \in s_{h,R}} d_{2j}}{\sum_{s_h \cap RE} d_{2j}}$$

donde:

- a_{3h} : ajuste por no respuesta en el estrato h .
- s_h, R : $s_h \cap (RE \cup NR)$.
- d_{2j} : factor de expansión ajustado por elegibilidad desconocida de la vivienda j .

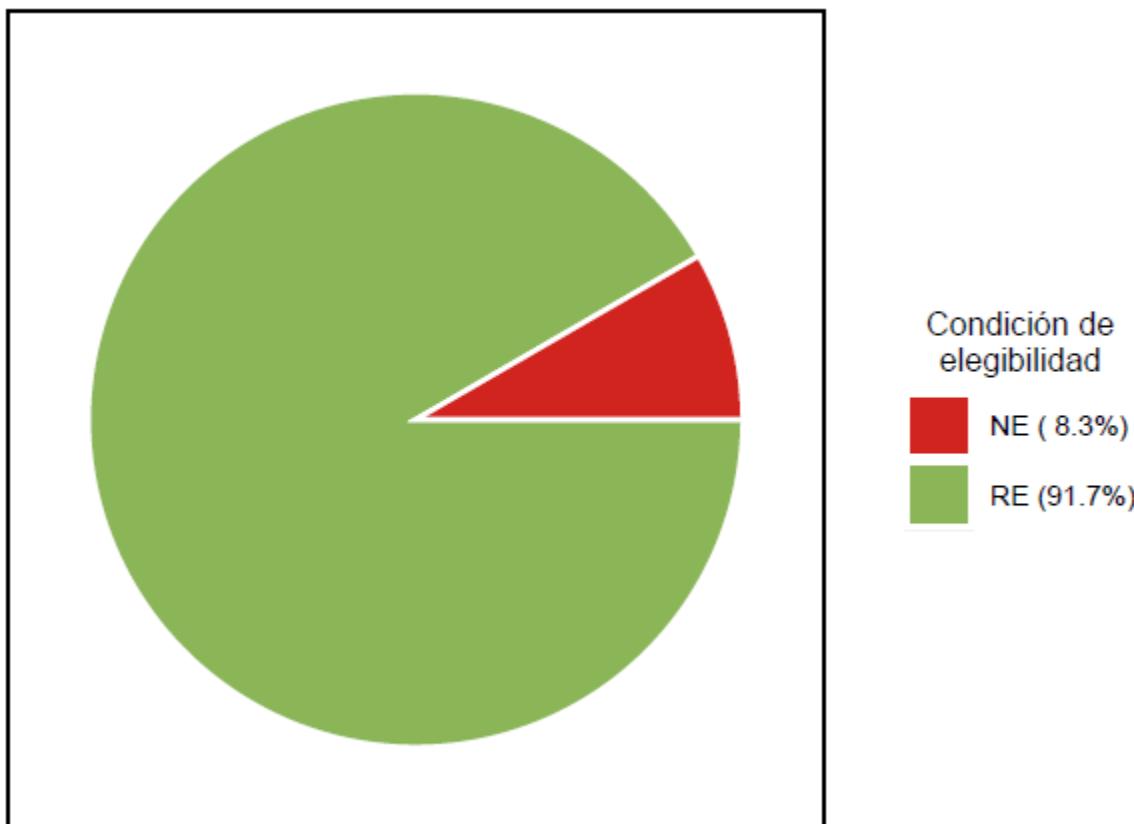
Por lo tanto el factor de expansión ajustado por no respuesta (d_{3j}) queda definido de la siguiente manera:

$$d_{3j} := \begin{cases} a_{3h} \cdot d_{2j} & j \in s_h \cap RE, \\ d_{2j} & j \in s_h \cap NE, \\ 0 & j \in s_h \cap NR \end{cases}$$

Factores de expansión

Ajuste por cobertura de viviendas

El ajuste por no respuesta considera que las viviendas no respondientes tienen características similares a las respondientes. Por lo tanto, la distribución final, considerando los ajustes por cobertura es:

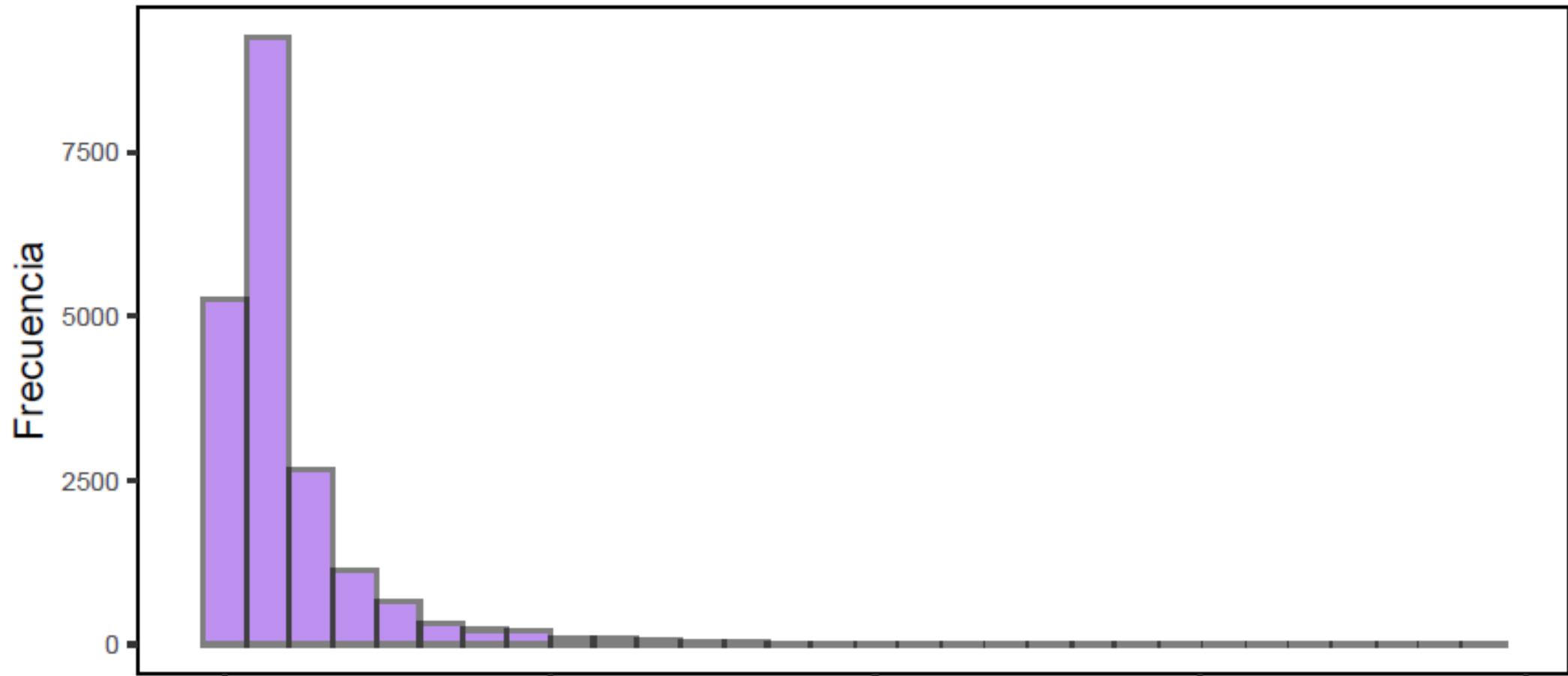


Factores de expansión



Buenas cifras,
mejores vidas

Distribución de los factores de expansión



Ejercicio

Ejercicio 7

Estimación de características

Para terminar el taller, calcularemos los estimadores de desnutrición crónica para niños menores de 5 años desde la ENDI Ronda 2. Para esto:

- 1) Descargaremos la base desde el página web del INEC.
- 2) Fijaremos el diseño muestral de la encuesta.
- 3) Calcularemos los estimadores de desnutrición crónica para diferentes desagregaciones.



Buenas cifras,
mejores vidas

