

Cálculo del tamaño y selección de muestra

Unidades Educativas



DINEM

JULIO· 2025

Su

1. **Introducción**

El Índice de Precios al Consumidor (IPC) es un índice mensual y nacional, que permite medir los cambios en el tiempo del nivel general de los precios de consumo final de bienes y servicios de los hogares residentes en el área urbana del país. Es un índice de gran importancia para la vida económica del país dado que su objetivo principal es calcular la inflación, a través de una recopilación mensual de precios al consumidor final en el área urbana, con base en una canasta de bienes y servicios representativas del gasto monetario de consumo final que realizan los hogares del Ecuador, para satisfacer sus necesidades.

Para mantener la representatividad del índice con respecto al consumo de los hogares, el INEC con el Proyecto Cambio de Base del Sistema de Indicadores de Precios al Consumidor en el Ecuador (CAB-SIPCE) 2024-2026 está realizando el séptimo cambio de año base en los más de 70 años de historia estadística de este indicador. El CAB-SIPCE busca actualizar de manera integral la metodología del Índice de Precios al Consumidor base 2014=100. Una de las mejoras a proponer es contar con la implementación de un diseño de muestreo probabilístico, debido a que el IPC necesita basarse en metodologías sólidas que se amparen en las recomendaciones internacionales, detalladas en el Manual de Teoría y Práctica de Índice de Precios al Consumidor del FMI. En las cuales se indica que deben preferirse las técnicas de muestreo probabilístico cuando existan los marcos de muestreo adecuados y los recursos necesarios para su ejecución.

En atención a la Acción de Mejora y en lo que compete a la Dirección de Infraestructura Estadística y Muestreo (DINEM), este informe presenta el ejercicio referente al cálculo del tamaño y selección de la muestra a partir de un marco de muestreo que está conformado por unidades educativas ubicadas en las ciudades de interés.

1. **Planificación**

Objetivo General

Cálculo del tamaño de muestra y selección aleatoria de instituciones educativas.

Tipo de operación estadística

La operación estadística es una encuesta dirigida a unidades educativas.

Cobertura temática

En referencia a los niveles de educación, las unidades educativas a investigar pertenecen a aquellas particulares que dispongan de al menos tres niveles:

* Nivel de educación inicial
* Nivel de educación básica
* Nivel de educación bachillerato

Universo de estudio

Se consideran aquellas unidades educativas que cumplen con los aspectos definidos en torno a la cobertura temática.

Unidad de observación y de análisis

Unidades educativas.

Cobertura geográfica

Las unidades educativas ubicadas en el área urbana de los siguientes cantones:

* Manta
* Machala
* Santo Domingo
* Esmeraldas
* Ambato
* Cuenca
* Loja

En el presente ejercicio no se consideran las ciudades de Quito y Guayaquil debido a las limitaciones operativas. Esto genera un sesgo geográfico relevante, ya que estas ciudades concentran una parte considerable de las unidades educativas urbanas y del gasto en pensiones/matrículas.

1. **Metodología del marco muestral y el diseño de la muestra**

Tipos y etapas del diseño muestral

Muestreo probabilístico con selección aleatoria de elementos.

Dominios de estudio

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

Debido a las limitaciones operativas del proyecto, el presente ejercicio considera un solo dominio de estudio para el cálculo del tamaño muestral, posterior a ello se procede con la distribución en cada uno de los cantones en los que se pretende llevar la encuesta.

## Marco muestral

Un marco muestral es una lista de los elementos que forman el universo que será objeto de estudio, dichos elementos son llamados unidades muestrales que tienen definidas características que permitan identificarlos. Con el marco muestral definido correctamente, podremos extraer la población que se va a investigar, es decir, la muestra que se analizará.

Para realizar un muestreo probabilístico es sumamente necesario construir un marco muestral ya que permite identificar claramente la unidad de estudio. El marco muestral debe ser preciso, libre de omisiones y duplicaciones. Los datos y las unidades deben cubrir a toda la población que sea objeto de estudio e identificarlas correctamente.

El marco de muestreo que se considera está construido a partir del listado de unidades educativas que fue elaborado por el equipo CAB-SIPCE que usó como insumo la información de precios de pensiones y matrículas de Unidades Educativas (registros administrativos) del Ministerio de Educación (MINEDUC). Se complementó este insumo de precios con el Archivo Maestro de Instituciones Educativas 2023-2024 del MINEDUC para obtener la información general de todas las instituciones particulares urbanas de 7 ciudades investigadas por el IPC junto con los precios de matrícula y pensión de cada nivel educativo que oferte la institución.

Para la construcción del marco muestral se seleccionaron aquellas unidades educativas que cumplan las siguientes condiciones:

* Las instituciones particulares con al menos 3 niveles educativos.
* Las instituciones que se encuentren ubicadas en el área urbana de Manta, Machala, Santo Domingo, Esmeraldas, Ambato, Cuenca y Loja.
* Se excluyen unidades educativas que no disponen de información en las variables costo de matrícula y/o costo de pensión.

El marco de muestreo queda constituido por 186 unidades educativas

Tabla 1. Marco de muestreo

|  |  |
| --- | --- |
| **Etiquetas de fila** | **Cuenta de cantón** |
| AMBATO | 36 |
| CUENCA | 35 |
| ESMERALDAS | 11 |
| LOJA | 13 |
| MACHALA | 29 |
| MANTA | 35 |
| SANTO DOMINGO | 27 |
| **Total general** | **186** |

## Variables de diseño

Para el análisis del tamaño muestral se ha considerado como variable de diseño el *costo medio de matrícula* que es un dato actualizado en cada año lectivo de las instituciones educativas.

## Tamaño muestral

Para la determinación del tamaño de la muestra se requiere establecer la característica o características a estimar, el nivel de confianza y la precisión requerida de tal manera que los resultados obtenidos no sean demasiado costosos y/o imprecisos.

Bajo este criterio, para determinar el tamaño de la muestra se usa como variable de diseño el *costo medio de matrícula*, dicha variable está disponible en el marco muestral. La fórmula para dicho cálculo se presenta a continuación:

* = Tamaño de la muestra.
* N = Tamaño del dominio i.
* = Cuasivarianza de la variable de diseño para el dominio i.
* e = Error relativo máximo admisible.
* z = Coeficiente que representa el nivel de seguridad o confianza.
* = Variable de diseño (costo medio de matricula) en el dominio i.

Debido a las limitaciones operativas y a las características del mercado, se definieron instituciones educativas que deben investigarse de manera obligatoria, por tanto, se cuenta con un listado de instituciones que fueron definidas por el equipo CAB-SIPCE como de inclusión forzosa (probabilidad 1).

Tabla 1. Inclusión forzosa

|  |  |
| --- | --- |
| **Etiquetas de fila** | **Cuenta de cantón** |
| AMBATO | 3 |
| CUENCA | 4 |
| ESMERALDAS | 1 |
| LOJA | 2 |
| MACHALA | 3 |
| MANTA | 6 |
| SANTO DOMINGO | 3 |
| **Total general** | **22** |

*Fuente: Gestión de Diseño Muestral – DINEM*

El tamaño muestral final comprende los resultados obtenidos al calcular la muestra para las ciudades de estudio y las unidades educativas de inclusión forzosa.

Posterior a la distribución de la muestra se presentan los resultados del tamaño muestral definitivo (Tabla 2).

## Distribución de la muestra

Se definen como estratos las ciudades en las que se lleva a cabo la encuesta: Manta, Machala, Santo Domingo, Esmeraldas, Ambato, Cuenca y Loja. 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. Con la estratificación definida se realiza una distribución proporcional al tamaño (PPT) definido por la cantidad de unidades educativas y considerando un mínimo muestral en cada estrato. Con esto se garantiza la participación de cada uno de las ciudades que son de interés para la encuesta.

Para la distribución de la muestra en cada uno de los estratos se procede de la siguiente manera:

Donde:

* = Tamaño m𝑢𝑒𝑠𝑡𝑟𝑎l 𝑝𝑎𝑟𝑎 𝑒𝑙 𝑒𝑠𝑡𝑟𝑎𝑡𝑜 h.
* = 𝑇𝑎𝑚𝑎ñ𝑜 𝑑𝑒 𝑚𝑢𝑒𝑠𝑡𝑟𝑎.
* = 𝑇𝑎𝑚𝑎ñ𝑜 𝑑𝑒𝑙 𝑒𝑠𝑡𝑟𝑎𝑡𝑜 h.
* = 𝑁ú𝑚𝑒𝑟𝑜 𝑑𝑒 𝑒𝑠𝑡𝑟𝑎𝑡𝑜𝑠.

Tabla 2. Tamaño muestral

|  |  |  |  |
| --- | --- | --- | --- |
| **Cantón** | **Calculo Tamaño** | **Inclusión Forzosa** | **Tamaño Final** |
| AMBATO | 12 | 3 | 15 |
| CUENCA | 12 | 4 | 16 |
| ESMERALDAS | 5 | 1 | 6 |
| LOJA | 5 | 2 | 7 |
| MACHALA | 10 | 3 | 13 |
| MANTA | 11 | 6 | 17 |
| SANTO DOMINGO | 9 | 3 | 12 |
| **Total general** | **64** | **22** | **86** |

*Fuente: Gestión de Diseño Muestral – DINEM*

## Selección de la muestra

La selección de las unidades educativas que forman parte de la muestra, de acuerdo con el tamaño establecido, se realiza independientemente en cada uno de los estratos de manera aleatoria, asignando a cada unidad educativa igual probabilidad de participar en la selección.

