

## INFORME DE LOS AVANCES REALIZADOS CON RESPECTO A LA METODOLOGIA PARA EL DISEÑO MUESTRAL DEL NUEVO IPP

### Introducción

Actualmente el Índice de Precios al Productor de Disponibilidad Nacional (IPP-DN) se desarrolla mediante un muestreo dirigido o intencional en su totalidad. En el marco del proceso de certificación del IPP-DN, ejecutado en el año 2017, se emitieron diversas observaciones relacionadas con cada una de las fases de producción de esta operación estadística. Particularmente, en las fases de metodología sólida y diseño, se evidenció la necesidad de evaluar la implementación de un diseño de muestreo probabilístico a partir de un marco de muestreo en donde se incluyan a todas las empresas productoras, así como evaluar el uso de material cartográfico. En tal sentido, el presente documento tiene como objetivo exponer los alcances realizados que se han implementado para obtener un diseño muestral probabilístico en los dominios de estudio que así lo permitan.

### Objetivo General

- Elaborar un diseño muestral probabilístico para la encuesta que mide los Índices de Precios al Productor de Disponibilidad Nacional (IPP-DN).

### 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 nos 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 encontrar la unidad que va a ser el objeto de estudio. El marco muestral debe ser preciso, libre de omisiones y duplicaciones. Los datos y las unidades deben cubrir a toda la población e identificarlas correctamente.

El marco de muestreo que se considera está definido por las empresas que constan en el Directorio de Empresas y establecimientos Económicos 2021 (INEC, 2023).

Para la construcción del marco muestral se consideraron las actividades económicas mediante el Código CIIU Rev 4 - 6 dígitos, dichos códigos consideran una correspondencia a nivel de los productos definidos en la canasta del índice del Precio al Productor (IPP) que fue elaborada por el equipo a cargo y que se encuentra vigente a la fecha DECON/CAB-SIPP, además, con los Código CIIU Rev 4 - 6 dígitos, se construye una estratificación dentro de cada dominio.

Dichas actividades económicas están agrupadas por las actividades principales que se enlistan a continuación:

- A Agricultura, ganadería, silvicultura y pesca
- B Explotación minas y canteras

- C Industrias manufactureras
- E Distribución agua; alcantarillado, desechos y saneamiento
- G Comercio, reparación automotores y motocicletas
- H Transporte y almacenamiento
- I Actividades de alojamiento y de servicio de comidas
- J Información y comunicación
- K Actividades financieras y de seguros
- L Actividades inmobiliarias
- M Actividades profesionales, científicas y técnicas
- N Actividades de servicios administrativos y de apoyo
- P Enseñanza
- R Artes, entretenimiento y recreación

Una vez revisada la estructura de DICE-2021 que está conformado por 858.101 empresas, para la construcción del marco muestral, por pedido del equipo DECON/CAB-SIPP, se seleccionaron aquellas que cumplan las siguientes condiciones:

Empresas de tamaño Pequeña empresa (2), Mediana empresa “A” (3), Mediana empresa “B” (4) y Grande empresa (5). Se excluyen las empresas que constan como “empresas no ubicadas” en el DICE-2021.

El marco de muestreo queda constituido de la siguiente forma:

Código Sección	TAMAÑO DE EMPRESA				Total
	Pequeña empresa	Mediana “A”	Mediana “B”	Grande Empresa	
A	116	39	43	44	242
B	213	35	27	41	316
C	2.281	461	341	602	3.685
E	31	3	3	2	39
G	12	1	-	-	13
H	113	22	12	16	163
I	77	4	3	-	84
J	87	10	8	7	112
K	21	9	9	63	102
L	225	16	5	3	249
M	523	35	20	5	583
N	219	31	20	17	287
P	45	13	7	2	67
R	22	1	-	-	23

Fuente: DICE-2021

## 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.

Los dominios de estudio para el análisis muestral están definidos por los sectores económicos (Código CIIU4 - Sección - 1 dígito) agrupados por el tamaño de la empresa a nivel nacional. Dentro de cada uno de los dominios de estudio se considera un grupo de inclusión forzosa, es decir, un subconjunto de empresas que formarán parte de la muestra con probabilidad uno (1), este grupo corresponde a aquellas empresas catalogadas como “Grande Empresa” (tamaño 5).

## Variables de diseño

Para el análisis del tamaño muestral se ha considerado como variable de diseño las “Ventas Totales 2021” obtenida del DICE-2021 para cada empresa.

Cabe señalar que mediante reuniones mantenidas se ha tratado de discutir el hecho de incluir otras variables de diseño como los gastos energéticos, sin embargo, con la información que se dispone al respecto no se podría ser enlazada con la base de datos DICE-2021, este cruce de información no es viable debido a que no existe una variable de identificación, código único o número de RUC que permita dicho proceso. De la misma manera, se ha realizado la consulta con respecto a la disponibilidad de obtener información con respecto a las ventas totales o precios para cada uno de los productos que genere la empresa o la cantidad de empleados destinados estrictamente a la elaboración de cada uno de los productos que genere la empresa, sin embargo; la institución no recolecta dicha información.

## 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 requeridas 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 “Ventas Totales”, dicha variable está disponible en el DICE-2021. La fórmula para dicho cálculo se presenta a continuación:

$$n_i \geq \frac{(N_i * S_i)^2}{\left(\frac{N_i - 1}{N_i}\right) * \left(\frac{e * v}{z}\right)^2 + N_i * S_i^2} * \frac{1}{(1 - Tnr_i)}$$

- $n_i$  = Tamaño de la muestra por dominio.
- $N_i$  = Tamaño del dominio  $i$ .
- $S_i$  = Cuasivarianza del dominio  $i$ .
- $e$  = Error relativo máximo admisible.
- $z$  = Coeficiente que representa el nivel de seguridad o confianza.

- $v_i$  = ventas totales en el dominio  $i$ .
- $tnri$  = Tasa de no respuesta del dominio  $i$ .

Para definir la tasa de no respuesta (TNR) en cada uno de los dominios, se usó como referencia las TNR's resultantes de la Encuesta Estructural Empresarial (ENESEM-2021). Para los dominios en los que no se especifica una TNR se realizó un promedio acorde a los tamaños de las empresas y actividad principal.

## Selección de la muestra

Con la estratificación realizada dentro de cada uno de los dominios a nivel de Código CIIU Rev 4 - 6 dígitos, se realiza una distribución proporcional al tamaño (PPT). Con esto se busca obtener la representatividad de cada uno de los Códigos CIIU Rev 4 – 6 dígitos en la muestra. 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, utilizando variables que aportan información para todas las unidades de la población. El objetivo de la estratificación nos permite mejorar la eficiencia del diseño maestral mediante la disminución de la varianza de los estimadores. La distribución por estratos está construida de la siguiente manera:

$$n_{i_h} = n_i * \frac{N_{i_h}}{\sum_{h=1}^{H_i} N_{i_h}}$$

donde:

- $n_{i_h}$  = Muestra para el estrato  $h$  del dominio de diseño  $i$ .
- $n_i$  = Tamaño de muestra del dominio de diseño  $i$ .
- $N_{i_h}$  = Tamaño del estrato  $h$  del dominio de diseño  $i$ .
- $H_i$  = Número de estratos en el dominio de diseño  $i$ .

Para el caso de las empresas cuya actividad principal corresponde a Industrias manufactureras se está elaborando un análisis particular debido a la cantidad de empresas y Códigos CIIU Rev 4 - 6 dígitos para los que se desea obtener una adecuada representatividad, por tanto, se está planteando alternativas para el cálculo del tamaño en los dominios que pertenecen a este conjunto.

## Conclusiones y recomendaciones

- Para elaborar un acertado plan de diseño se debería disponer de una actualización del marco muestral completo y preciso con variables que permitan una correcta selección de empresas que son el objeto de estudio, dicha actualización debería contemplar variables como el precio de cada uno de los productos (variable de diseño ideal) que oferten cada una de las empresas presentes en el universo de estudio.
- El marco de muestreo disponible presenta ciertas limitaciones que no permiten una aplicabilidad directa de un muestreo probabilístico más robusto, por ejemplo,

para el caso de las microempresas no se dispone de información en el campo de ventas, que es la variable que fue considerada como de diseño. Es importante señalar que la variable de diseño utilizada (Ventas Totales) no es la ideal pero es la única disponible para los cálculos del tamaño muestral en el presente ejercicio.

### Observaciones

- Es importante señalar que el actual documento recoge información de los procesos implementando hasta el momento, mas no de los resultados finales; por lo que, existirán cambios a medida que se discutan y se tomen decisiones definitivas y con ello se podrá elaborar un documento final en el que se detalle la metodología utilizada.

### Bibliografía

INEC. (2023). *ENCUESTA ESTRUCTURAL EMPRESARIAL (ENESEM) 2021*. QUITO.

FIRMAS DE RESPONSABILIDAD		
Nombre y Apellido	Dirección	Firmas
OMAR LLAMBO	DINEM	
CHRISTIAN GARCÉS	DINEM	