



**METODOLOGÍA MUESTRAL
ENCUESTA LONGITUDINAL DE EMPRESAS**

AÑO CONTABLE 2009

INSTITUTO NACIONAL DE ESTADÍSTICAS
Noviembre / 2011

DEPARTAMENTO DE INVESTIGACIÓN Y DESARROLLO

Metodología Muestral Encuesta Longitudinal de Empresas

Instituto Nacional de Estadísticas.

Noviembre / 2011.

Jefe Departamento: Charles Durán Artigas.

Coordinadora Sección Estadísticas Económicas: Carolina Morales López

ÍNDICE

Diseño Muestral	1
1. Población Objetivo.....	1
2. Marco Muestral.....	1
3. Estratificación	1
4. Tamaño de la Muestra	2
5. Método de Selección	3
6. Precisión Muestral.....	4
7. Factores de Expansión.....	8
7.1. Factor de Expansión Ventas.....	9
7.2. Factor de Expansión por número de empresas	10
8. Niveles de Estimación.....	11

Diseño Muestral

1. Población Objetivo

La población objetivo está conformada por empresas formales que desarrollan actividad productiva, comercial o de servicio, dentro de los límites territoriales del país y que su nivel de venta es superior a 800,1 UF durante el año 2009.

Para efectos de éste estudio, se entiende como empresa formal a las personas naturales o jurídicas, que cuentan con inicio de actividades en el Servicio de Impuestos Internos (SII) y que tributan en primera categoría. Se excluyen las actividades económicas realizadas por el Gobierno.

2. Marco Muestral

El marco muestral, se construye a partir del Directorio Empresas 2009, del Instituto Nacional de Estadísticas, el cual, está conformado principalmente por registros administrativos del Servicio de Impuestos Internos, año 2009 (Declaración Anual de Impuestos a la Renta, formulario 22 y Declaración Mensual del IVA, formulario 29) y directorios internos del INE.

3. Estratificación

Para Facilitar la comparabilidad de la información se realiza la primera estratificación a nivel de clasificación económica, utilizando la Clasificación Industrial Uniforme de todas las actividades Económicas, CIIU Rev. 3, a nivel de categoría de tabulación.

Cuadro 1. Clasificaciones económicas, según CIIU rev.3, consideradas para el estudio.

Categoría de Tabulación	Descripción
A	Agricultura, ganadería, caza y silvicultura.
B	Pesca
C	Explotación de minas y canteras.
D	Industrias manufactureras.
E	Suministro de electricidad, gas y agua.
F	Construcción.
G	Comercio al por mayor y por menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos.
H	Hoteles y restaurantes.
I	Transporte, almacenamiento y comunicaciones.
J	Intermediación financiera.
K	Actividades inmobiliarias, empresariales y de alquiler.
O	Otras actividades de servicios comunitarios, sociales y personales.

La segunda estratificación se realiza a nivel de tamaño, considerando el nivel de ventas. Esta segmentación por tamaño, incluye los siguientes estratos:

Cuadro 2. Descripción de los intervalos de corte de ventas, según tamaño.

Tamaño de ventas	Ventas anuales (UF)	
	Limite Inferior	Limite Superior
Microempresa	800,1	2.400
Pequeñas empresa 1	2.400,1	5.000
Pequeñas empresa 2	5.000,1	25.000
Mediana empresa	25.000,1	100.000
Empresa grande	100.000,1	Más

Como una forma de obtener mejor representación y disminuir el error de muestreo que proviene de la dispersión observada en las ventas anuales, se considera, una vez ordenados los estratos por la variable de interés, integrar las unidades más grandes como Tramo a Censar (TC), y el resto como Tramo a Muestrear (TM).

El punto de corte para estos tramos, se determina según el aporte de las ventas de las empresas a censar y el error de muestreo que resulta de restar al marco muestral las empresas que afectan notoriamente la precisión de las estimaciones. Estas empresas adquieren carácter censal o de Inclusión Forzosa (IF).

4. Tamaño de la Muestra

La estimación del tamaño muestral a nivel nacional, se obtuvo a partir de un muestreo probabilístico estratificado, según categoría de tabulación CIIU Rev.3, con un nivel de confianza del 95%, alcanzando a 7.062 empresas, que corresponde al 2,83% del total de las unidades del marco muestral. Para la determinación del tamaño muestral de cada estrato, se distribuye el total de unidades muestrales en forma proporcional al número de empresas de cada uno de ellos. Aplicándose los siguientes algoritmos:

Tamaño Muestral

$$n = \frac{\left(\sum_{h=1}^H N_h \cdot S_h \right)^2 * Z^2}{e_r * \sum_{h=1}^H y_h + \left(\sum_{h=1}^H N_h \cdot S_h^2 \right) * Z^2}$$

Donde:

$$N = \sum_h N_h$$

n : Tamaño de la muestra.

Z : Estadístico normal estándar correspondiente a un nivel de confianza 95%.

e_r : Error relativo.

S_h : Desviación estándar en el estrato h .

N_h : Tamaño de la población en el estrato h .

Distribución del tamaño muestral en forma proporcional de acuerdo al nivel de ventas de cada estrato.

$$n_h = n * \frac{\sum_{i=1}^k y_{hi}}{\sum_{h=1}^H \sum_{i=1}^k y_{hi}}$$

Donde:

n_h : Tamaño de la muestra en el estrato h .

n : Tamaño de la muestra.

$\sum_{i=1}^k y_{hi}$: Sumatoria de las ventas de las empresas en el tramo aleatorio del estrato h .

$\sum_{h=1}^H \sum_{i=1}^k y_{hi}$: Sumatoria de las ventas de las empresas del tramo aleatorio de la clasificación económica.

El tamaño definitivo de las agrupaciones, está sujeto a los costos versus precisión, por lo cual se debe aumentar o disminuir el número de empresas de algunos estratos para fortalecer su representación y precisión necesaria para el estudio.

5. Método de Selección

La selección de unidades a encuestar se realiza de dos formas. En el caso de inclusiones forzosas, las empresas son incluidas en su totalidad en el levantamiento.

La selección que corresponde al tramo a muestrear, se realiza en forma independiente para estrato (sector, tamaño), aplicando una selección sistemática y empleando como variable auxiliar la ubicación geográfica.

Se selecciona aproximadamente 30% de unidades de reemplazo. Para cubrir eventuales pérdidas de unidades muestrales, debido a quiebras, paralizaciones, cierre, cambio de giro, etc. En caso de las unidades de reemplazo seleccionadas no logre cubrir los requerimientos de tamaño total de la muestra, el INE selecciona unidades adicionales a entrevistar. Todo esto, en los casos en que el Marco Muestral de empresas contenga estas unidades adicionales.

6. Precisión Muestral

Para determinar la precisión muestral respectiva, se utilizan los siguientes algoritmos:

Coefficiente de variación del estrato h .

$$CV_h = \frac{\sqrt{V(\hat{Y})_h}}{\sum_{i=1}^k y_{hi}}$$

Donde:

CV_h : Coeficiente de variación del estrato h .

$V(\hat{Y})_h$: Varianza muestral estimada de las ventas del estrato h .

$\sum_{i=1}^k y_{hi}$: Sumatoria de las ventas del estrato h .

Varianza muestral de las ventas de las empresas

$$V(\hat{y}) = \sum_h^H N_h * (N_h - n_h) * \frac{S_h^2}{n_h}$$

Donde:

$V(\hat{y})$: Varianza muestral estimada de las ventas.

S_h^2 : Cuasi-varianza de las ventas de las empresas del estrato h .

N_h : Número de empresas en el marco muestral en el estrato h .

n_h : Tamaño muestral del estrato h .

Cuasi-varianza de las ventas de las empresas del estrato h .

$$S_h^2 = \frac{N_h}{(N_h - 1)} * \sigma_h^2$$

Donde:

S_h^2 : Cuasi-varianza que mide la dispersión de las ventas de las empresas de la muestra en el estrato h .

N_h : Número de empresas en el estrato h .

σ_h^2 : Varianza de las ventas del estrato h .

Varianza de las ventas del estrato h .

$$\sigma_h^2 = \frac{\sum_{i=1}^k \left(y_{hi} - \bar{y}_h \right)^2}{N_h}$$

Donde:

σ_h^2 : Varianza de las ventas del estrato h .

y_{hi} : Venta de la i -ésima empresa del estrato h .

\bar{y}_h : Media aritmética de las ventas en el estrato h .

N_h : Número de empresas en el estrato h .

Cuadro 3. Marco Muestral y Muestra Efectiva, según categoría de tabulación CIU Rev. 3.

Categoría	Descripción	Marco Muestral	Muestra Efectiva		CV	
			Total	Inclusión Forzosa		Tramo Aleatorio
	Total nacional	249.793	7.062	326	6.736	1,18
A	Agricultura, ganadería, caza y silvicultura.	28.751	449	15	434	1,52
B	Pesca	1.611	270	9	261	3,60
C	Explotación de minas y canteras.	1.508	314	24	290	1,80
D	Industrias manufactureras.	26.861	872	31	841	3,69
E	Suministro de electricidad, gas y agua.	147	106	31	75	1,03
F	Construcción.	21.807	788	11	777	2,93
G	Comercio al por mayor y por menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos.	88.151	1.401	39	1.362	2,95
H	Hoteles y restaurantes.	11.835	636	17	619	1,22
I	Transporte, almacenamiento y comunicaciones.	25.220	748	28	720	2,47
J	Intermediación financiera.	740	288	30	258	2,82
K	Actividades inmobiliarias, empresariales y de alquiler.	34.291	809	64	745	1,86
O	Otras actividades de servicios comunitarios, sociales y personales.	8.871	381	27	354	1,20

Departamento de Investigación y Desarrollo, INE
Sección Estadísticas Económicas

Cuadro 4. Marco Muestral y Muestra efectiva, según tamaño de ventas.

Tamaño	Marco Muestral	Muestra Efectiva		CV	
		Total	Inclusión Forzosa		Tramo Aleatorio
Total	249.793	7.062	326	6.736	1,18
100.000,1 y más	7.580	2.252	326	1.926	1,43
25.000,1 hasta 100.000	15.772	1.441		1.441	1,31
5.000,1 hasta 25.000	57.205	1.680		1.680	1,37
2.400,1 hasta 5.000	52.396	725		725	0,89
800,1 hasta 2.400	116.840	964		964	1,15

Departamento de Investigación y Desarrollo, INE
Sección Estadísticas Económicas

Cuadro 5. Marco Muestral y Muestral Efectiva, según Categoría de tabulación CIU Rev. 3 y Tamaño.

Categoría	Tamaño	Marco Muestral	Muestra Efectiva			CV
			Total	Inclusión Forzosa	Tramo Aleatorio	
	Total	249.793	7.062	326	6.736	1,18
A	Total	28.751	449	15	434	1,52
	100.000,1 y más	301	90	15	75	2,11
	25.000,1 hasta 100.000	1.147	128		128	3,50
	5.000,1 hasta 25.000	7.271	127		127	4,06
	2.400,1 hasta 5.000	6.483	54		54	2,87
	800,1 hasta 2.400	13.549	50		50	4,48
B	Total	1.611	270	9	261	3,60
	100.000,1 y más	50	28	9	19	4,59
	25.000,1 hasta 100.000	103	69		69	2,96
	5.000,1 hasta 25.000	521	97		97	4,24
	2.400,1 hasta 5.000	338	31		31	3,61
	800,1 hasta 2.400	599	45		45	4,47
C	Total	1.508	314	24	290	1,80
	100.000,1 y más	135	60	24	36	1,82
	25.000,1 hasta 100.000	120	70		70	2,99
	5.000,1 hasta 25.000	433	101		101	4,16
	2.400,1 hasta 5.000	296	29		29	3,68
	800,1 hasta 2.400	524	54		54	3,96
D	Total	26.861	872	31	841	3,69
	100.000,1 y más	1.526	352	31	321	4,20
	25.000,1 hasta 100.000	2.329	154		154	3,18
	5.000,1 hasta 25.000	7.270	198		198	3,35
	2.400,1 hasta 5.000	5.227	69		69	2,54
	800,1 hasta 2.400	10.509	99		99	3,13
E	Total	147	106	31	75	1,03
	100.000,1 y más	114	81	31	50	1,03
	25.000,1 hasta 100.000	33	25		25	4,01
F	Total	21.807	788	11	777	2,93
	100.000,1 y más	748	302	11	291	4,12
	25.000,1 hasta 100.000	1.624	153		153	3,14
	5.000,1 hasta 25.000	5.755	169		169	3,61
	2.400,1 hasta 5.000	4.671	78		78	2,42
	800,1 hasta 2.400	9.009	86		86	3,38
G	Total	88.151	1.401	39	1.362	2,95
	100.000,1 y más	2.592	552	39	513	3,79
	25.000,1 hasta 100.000	5.533	193		193	2,88
	5.000,1 hasta 25.000	18.789	271		271	2,91
	2.400,1 hasta 5.000	17.624	145		145	1,78
	800,1 hasta 2.400	43.613	240		240	2,06

Departamento de Investigación y Desarrollo, INE
 Sección Estadísticas Económicas

Cuadro 5. Marco Muestral y Muestral Efectiva, según Categoría de tabulación CIU Rev. 3 y Tamaño.

Categoría	Tamaño	Marco Muestral	Muestra Efectiva			CV
			Total	Inclusión Forzosa	Tramo Aleatorio	
	Total	249.793	7.062	326	6.736	1,18
H	Total	11.835	636	17	619	1,22
	100.000,1 y más	130	87	17	70	1,76
	25.000,1 hasta 100.000	488	132		132	3,01
	5.000,1 hasta 25.000	2.367	210		210	3,21
	2.400,1 hasta 5.000	2.570	103		103	2,06
	800,1 hasta 2.400	6.280	104		104	3,13
I	Total	25.220	748	28	720	2,47
	100.000,1 y más	609	211	28	183	3,00
	25.000,1 hasta 100.000	1.370	133		133	3,35
	5.000,1 hasta 25.000	4.995	199		199	3,42
	2.400,1 hasta 5.000	5.829	78		78	2,43
	800,1 hasta 2.400	12.417	127		127	2,81
J	Total	740	288	30	258	2,82
	100.000,1 y más	347	178	30	148	2,89
	25.000,1 hasta 100.000	393	110		110	3,26
K	Total	34.291	809	64	745	1,86
	100.000,1 y más	918	239	64	175	2,81
	25.000,1 hasta 100.000	2.387	183		183	2,89
	5.000,1 hasta 25.000	8.447	192		192	3,44
	2.400,1 hasta 5.000	7.632	92		92	2,21
	800,1 hasta 2.400	14.907	103		103	3,12
O	Total	8.871	381	27	354	1,20
	100.000,1 y más	110	72	27	45	1,20
	25.000,1 hasta 100.000	245	91		91	3,47
	5.000,1 hasta 25.000	1.357	116		116	4,33
	2.400,1 hasta 5.000	1.726	46		46	3,10
	800,1 hasta 2.400	5.433	56		56	4,34

Departamento de Investigación y Desarrollo, INE
Sección Estadísticas Económicas

7. Factores de Expansión

La información recolectada se hace a través de una muestra de empresas, para obtener una visión completa de las actividades que componen los sectores se realiza una expansión de la información levantada.

Para realizar la mencionada expansión, se considera en primer lugar las empresas de inclusión forzosa que son unidades muestrales autorepresentadas, ya que tienen carácter censal, en segundo las empresas seleccionadas en forma aleatoria.

7.1. Factor de Expansión Ventas

El factor de expansión calculado a partir de las ventas anuales, se emplea para obtener una mejor estimación de las variables cuantitativas del estudio. Para las unidades de inclusión forzosa en casos de existir no respuesta, se aplica el siguiente algoritmo:

Factor de expansión de las empresas de inclusión forzosa (IF).

$$FE(IF)_h = \frac{\sum_{i=1}^K y(IF)_{h,i}}{\sum_{i=1}^K y(IFe)_{h,i}}$$

Donde:

$FEV(IF)_h$: Factor de expansión de ventas de las empresas IF en el estrato h .

$\sum_{i=1}^K y(IF)_{h,i}$: Sumatoria de las ventas del marco muestral correspondiente a las empresas IF del estrato h .

$\sum_{i=1}^k y(IFe)_{h,i}$: Sumatoria de las ventas del marco muestral que corresponden a las empresas IF efectivamente logradas, del estrato h .

Recuérdese que para homogenizar la variable en estudio se realizó una segunda estratificación en el tramo aleatorio, de acuerdo al tamaño según el nivel de ventas que declararon las empresas. Esto cobra relevancia en el momento de realizar la expansión, debido a que debe efectuarse con la misma desagregación, así se obtendrá una mejor estimación de las ventas totales, pues las empresas de distintos tamaños tendrán su propio tratamiento.

La expresión algebraica del factor de expansión de los establecimientos pertenecientes al Tramo Aleatorio (TA), tiene la siguiente forma:

Factor de expansión de las empresas de inclusión aleatoria (TA).

$$FEV(TA)_h = \frac{\sum_{i=1}^K y(TA)_{h,i}}{\sum_{i=1}^K y(TAe)_{h,i}}$$

Donde:

$FEV(TA)_h$: Factor de expansión de ventas del tramo aleatorio del estrato h .

$\sum_{i=1}^K y(TA)_{h,i}$: Sumatoria de las ventas del marco muestral correspondiente a las empresas del TA del estrato n .

$\sum_{i=1}^k y(TAe)_{h,i}$: Sumatoria de las ventas del marco muestral que corresponden a las empresas efectivamente logradas en el TA del estrato h .

7.2. Factor de Expansión por número de empresas

En esta oportunidad para obtener una mejor estimación de las variables cualitativas del estudio, se obtiene un factor de expansión a partir del número de empresas, los que son calculados empleando la cantidad de ellas en el marco muestral y muestra, a nivel de estratos por tipo de inclusión, clasificación económica y tamaño. Para las unidades de inclusión forzosa en caso de que exista no respuesta, se aplica el siguiente algoritmo:

Factor de expansión de las empresas de inclusión forzosa (IF).

$$FEN(IF)_h = \frac{\sum_{i=1}^K y(IF)_{h,i}}{\sum_{i=1}^K y(IFe)_{h,i}}$$

Donde:

$FEN(IF)_h$: Factor de expansión por número de empresas de las unidades IF en el estrato h .

$\sum_{i=1}^K y(IF)_{h,i}$: Número de empresas del marco muestral, correspondiente a las empresas IF, del estrato h .

$\sum_{i=1}^k y(IFe)_{h,i}$: Número de empresas del marco muestral, corresponden a las empresas IF efectivamente logradas, del estrato h .

La expresión algebraica del factor de expansión de las empresas pertenecientes al Tramo Aleatorio (TA), tiene la siguiente forma:

Factor de expansión de las empresas de inclusión aleatoria (TA).

$$FEN(TA)_h = \frac{\sum_{i=1}^K y(TA)_{h,i}}{\sum_{i=1}^K y(TAe)_{h,i}}$$

Donde:

$FEN(TA)_h$: Factor de expansión de número de empresas del tramo aleatorio del estrato h .

$\sum_{i=1}^K y(TA)_{h,i}$: Número de empresas del marco muestral correspondiente a las empresas del TA del estrato h .

$\sum_{i=1}^k y(TAe)_{h,i}$: Número de empresas efectivamente logradas en el TA del estrato h .

8. Niveles de Estimación

De acuerdo a los objetivos del estudio, la validez estadística de las variables a estimar tiene representatividad para los siguientes niveles de estimación:

- Nacional según actividad económica.
- Nacional según tamaño
- Nacional según actividad económica y tamaño.