



# **ENCUESTA NACIONAL DE EMPLEO DISEÑO MUESTRAL**

**SUB DIRECCIÓN TÉCNICA  
DEPARTAMENTO DE INVESTIGACIÓN Y DESARROLLO**

SEPTIEMBRE 2006

## ÍNDICE

	Páginas
<b>Introducción</b>	3
1 Objetivos de la Encuesta	3
2 Población Objetivo	3
3 Marco Muestral y Estratificación	4
4 Niveles de Estimación	6
5 Unidades de Muestreo y Análisis	6
6 Tamaño de la Muestra	7
7 Tamaño Medio de Viviendas por Sección	10
8 Selección de las Unidades de Muestreo	11
9 Distribución de La Muestra Ponderaciones	12
10 Rotación de Viviendas de la muestra	12
11 Método de Estimación: Factores de Expansión de la muestra	12
11.1 Construcción del Factor de Expansión	13
12 El Estimador estándar del Total	15
13 Estimador de Razón Separado del Total	15
14 Estimadores de varianzas	16
14.1 Estimador de la varianza para un Total	16
14.2 Estimador de la varianza para la Tasa de Desocupación	17
15 Estimador de la varianzas del estimador estándar (sin Ajuste de Población)	17
16 Error de Muestreo	17
16.1 Error Absoluto de un estimador al 95% de confianza	17
16.2 Error Relativo de un estimador al 95% de confianza	17
17 Intervalo de Confianza al 95%	18
 <b>ANEXOS</b>	
Anexo 1: Muestra por Áreas de Estimación y Error Relativo para los Desocupados (Base: Oct Dic 2003)	19
Anexo 2: Procedimiento de Selección de las Unidades de primera y segunda Etapa	22
Anexo 3: Áreas de descuento ADA	23
Anexo 4: Comparación de Niveles de Estimación entre Los Niveles Actuales y la nueva propuesta	24
Anexo 6: Grandes Ciudades	28
a. Gran Santiago	
b. Gran Valparaíso	
c. Gran Concepción	
d. Gran Temuco	
e. Gran Serena	
Anexo 7: Estimadores de Totales, Proporciones, Tasas y Varianzas	29
Procedimientos de Cálculo de los Estimadores de la Encuesta Nacional del Empleo	29
 <b>BIBLIOGRAFÍA</b>	33

## **Introducción.**

En líneas generales el diseño para la selección de la muestra de la ENE 2006 es similar al utilizado en el año 1996 <sup>(1)</sup>, esto es, el método de selección es probabilístico en dos etapas con estratificación geográfica por región y área urbana-rural. Los estimadores asociados al diseño no son auto-ponderados y se ajustan por una proyección exógena de población calculada según métodos demográficos de acuerdo a INE-CELADE. Como se aprecia, en lo sustancial, el nuevo diseño es idéntico al vigente. En ese sentido, los cambios van por el lado de permitir la medición de más fenómenos del mercado laboral que lo que permitía el diseño vigente al 2005 como se presenta más adelante.

Dentro de los cambios relevantes destaca la actualización del marco muestral SIEH (Sistema Integrado de Encuestas de Hogares) con los antecedentes del último Censo de Población y Vivienda del año 2002.

### **1 Objetivos de la Encuesta.**

Caracterizar a la población mayor de 15 años del país, haciendo particular énfasis en la Fuerza de Trabajo. Esta caracterización contempla diversos perfiles entre los que destacan: sexo, edad, grupo de ocupación, rama de actividad, categoría de ocupación, y nivel de instrucción.

La representatividad muestral es nacional (urbana-rural); regional (urbana-rural); grandes centros urbanos (Ciudades de mas de 40.000 habitantes); y, resto de área urbana.

El diseño muestral contempla la distinción entre fenómenos coyunturales y estructurales en el mercado de trabajo. Asimismo, se inscribe en la conformación del nuevo marco maestro para el SIDEH, que tendrá como una de sus mejoras más importantes la actualización permanente del mismo.

### **2 Población Objetivo.**

La población objetivo de la muestra para la Encuesta Nacional de Empleo, ENE, está constituida por todas las personas del país residentes en viviendas particulares ocupadas. Esta definición excluye a las personas que habitan en viviendas colectivas, como hospitales, cárceles, conventos, cuarteles y otros, pero incluye a las personas que residen en viviendas particulares dentro de dichos centros, como son los porteros, conserjes y otros. Como así también, en la cobertura geográfica, se excluye la población residente en las áreas de difícil acceso denominadas ADA's.

---

<sup>(1)</sup> Instituto Nacional de Estadísticas, (1996), METODOLOGÍA DE LA ENCUESTA NACIONAL DEL EMPLEO.

### 3 Marco Muestral y Estratificación.

#### a) Marco Muestral:

Dadas las características del marco de áreas, la construcción del marco de muestreo se basó en el material cartográfico y de base del Precenso 2001 y el de Censo de Población y Viviendas 2002, compuesto por los mapas regionales, provinciales, comunales y distritales que contienen la delimitación de las áreas urbanas y rurales, y la demarcación de las áreas de difícil acceso (ADA).

Todos y cada uno de los elementos con que esta compuesto el Marco tienen una probabilidad conocida y diferente de cero de ser seleccionados.

**Conformación de los estratos del Marco Muestral:** De acuerdo a las características geográficas del marco muestral de áreas, los estratos a considerar según la División Político-Administrativa (DPA), corresponden a las comunas y dentro de éstas las áreas urbanas y Rurales, con demarcación de las áreas de difícil acceso no consideradas en el marco.

**Áreas de descuento (ADA):** Las áreas de difícil acceso (ADA) son aquellas que por razones de orden climático, topográfico o por ausencias de vías y medios de comunicación expeditos permanecen aisladas parcial o totalmente en el año (ver anexo 4).

**Partición del marco Muestral en Conglomerados o Secciones:** Dadas las características mencionadas del marco de áreas, la construcción del nuevo marco muestral se basó en el material cartográfico nacional, compuestos por los mapas regionales y comunales, conteniendo la delimitación de las áreas urbanas y rurales, y la demarcación de las ADA's.

Con estos antecedentes, se procedió a particionar el territorio en distintos estratos y en áreas geográficas o conglomerados denominados secciones.

Tales secciones se conformaron considerando la cantidad de población y el número de viviendas. El tamaño de cada sección varía entre los distintos estratos que conforman las provincias o regiones del país.

**Tabla N° 1: Tamaño Medio de Secciones por Estrato.**

<b>Estrato</b>	<b>Tamaño medio viviendas por sección</b>
Ciudad	100 - 200 viviendas
Resto Área Urbana	80 - 150 viviendas
Rural	60 - 80 viviendas

En la construcción de las secciones se consideró además de su estabilidad en el tiempo, su fácil identificación en terreno (respetando los límites comunales tanto en las secciones urbanas como rurales), su homogeneidad <sup>1</sup> y su compacidad<sup>2</sup>.

<sup>1</sup> La clasificación de la sección dentro del estrato respecto a los Desocupados, según su error  $\rho$ .

**Cobertura.** La cobertura geográfica del Marco Muestral para el SIEH 2003, es nacional, considerando todo el territorio continental de Chile, con exclusión de las áreas de difícil acceso (ADA). Estas últimas, determinadas en una fase previa a la elaboración del marco muestral, y que representan un 0.53% de viviendas particulares ocupadas del territorio nacional.

**b) Estratificación del Marco Muestral Maestro ajustado a la Encuesta de Empleo:**

El Marco Muestral Maestro se divide en comunas, sin embargo, para efectos de comparación se llevan a los de la actual ENE, incorporando aquellos resultantes de creación o división de comunas, etc., llegando a conformarse 158 estratos definidos según la condición geográfica (División Política Administrativa) y el número de viviendas y población que contenían en el censo de población y vivienda de 2002, según las definiciones siguientes:

**Ciudades o grandes centros urbanos (CD):** Constituidos por ciudades o conjuntos de ciudades adyacentes con 40.000 o más habitantes.

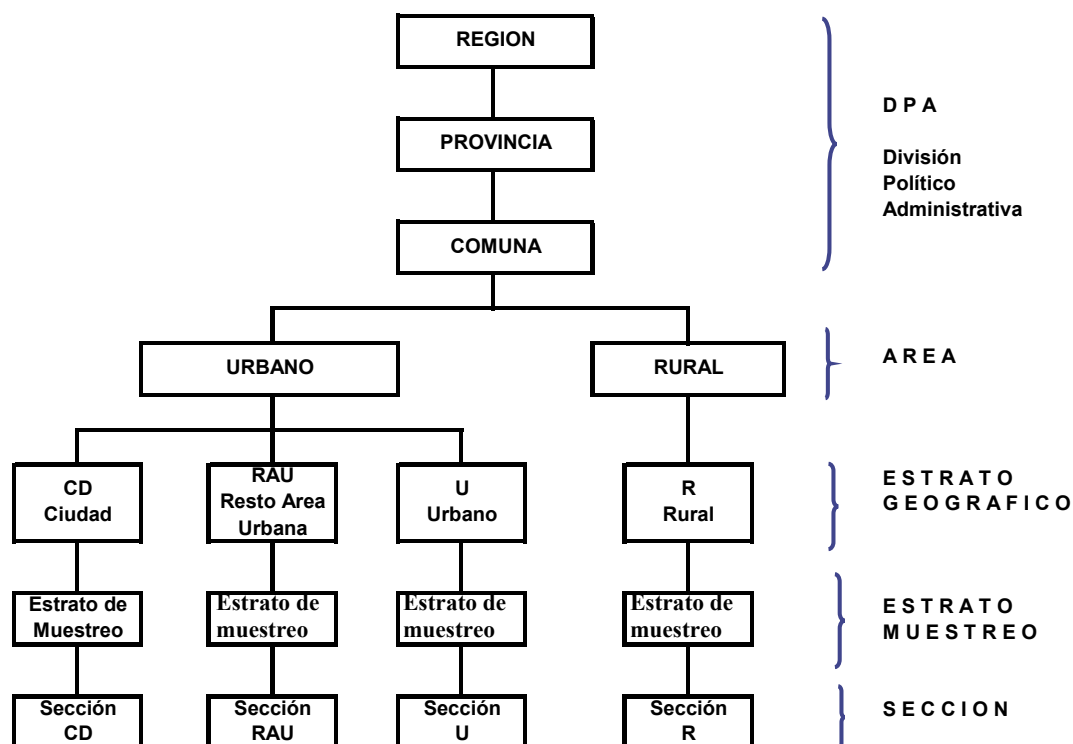
**Resto de Área Urbana (RAU):** Conjunto de viviendas concentradas con menos de 40.000 habitantes y más de 2.000 habitantes, o entre 1.001 y 2.000 con el 50 % o más de su población económicamente activa, dedicada a actividades secundarias y/o terciarias.

**Áreas Rurales (R):** Conjunto de viviendas concentradas o dispersas con 1.000 habitantes o menos, o entre 1.001 y 2.000 con menos del 50 % de su población económicamente activa, dedicada a actividades secundarias y/o terciarias.

---

<sup>2</sup> Habla de la similitud de la secciones dentro del estrato en relación a las estimaciones y errores.

**Figura\_1 : Estratificación del Marco de Selección**



#### 4 Niveles de Estimación.

Los Niveles de estimación en todo el país se desagrega de la siguiente forma:

NACIONAL: Urbano, Grandes Ciudades y Resto de Área Urbana; Rural.

REGIONAL: Urbano y Rural.

PROVINCIAL Y GRANDES CIUDADES (40.000 habitantes y más).

#### 5 Unidades de Muestreo y Análisis.

Las Unidades de Primera Etapa (UPE) son las **Secciones** que corresponden a conglomerados geográficos o secciones bien delimitados que contienen aproximadamente entre 60 y 200 viviendas cada uno, según el estrato de selección.

Las Unidades de Segunda Etapa (USE) son las **viviendas** particulares que se encuentran en cada sección (Unidades de Primera Etapa). A su vez, las Unidades de análisis son todas las **personas** que conforman los distintos hogares de las viviendas particulares ocupadas seleccionadas.

## 6 Tamaño de la Muestra.

El tamaño de muestra se fijó en términos del total de viviendas, de manera que resultase compatible con la obtención de un error de muestreo de baja magnitud en la estimación del total de desocupados del orden del 2% como error relativo, al 95% de confiabilidad.

La fórmula utilizada para la determinación del tamaño, en número de secciones, corresponde al Muestreo Aleatorio Simple (MAS) ajustado con el efecto de diseño (DEFF).

Número de secciones en la muestra en el estrato h:

$$n_h = \left( \frac{z_{0,95}}{e_a} \right)^2 \frac{S_h^2 \cdot M_h^2}{\bar{m}_h} \cdot Deff_h$$

Obteniéndose un tamaño total de:  $n = \sum_{h=1}^{158} n_h = 3.558 \text{ secciones}$

Siendo:

$n_h$  : El número de secciones por estrato

$e_{a_h}$  : El error absoluto del total de Desocupados por estrato h.

$S_h^2$  : El valor de la cuasi varianza estimada por estrato.

$M_h$  : correspondiente al número de viviendas que contiene el estrato h al censo 2002.

$Deff_h$  : Es el "efecto del diseño" en el estrato h, que se calculó como el cuociente entre la varianza de un estimador estándar dado el diseño en etapas y el mismo estimador dado un muestreo aleatorio simple de viviendas<sup>3</sup>.

El error absoluto fue calculado con la información obtenida y procesada de la Encuesta de Empleo del periodo Oct- Dic 2003, del total de desocupados en cada estrato. Este valor se consideró como fijo para posteriormente establecer el tamaño muestral resultante, así manteniendo los errores de muestreo de origen para garantizar la precisión de un coeficiente de variación del 2% a nivel Nacional.

<sup>3</sup> El detalle de este Cálculo se encuentra en el anexo 8 o en el informe del Departamento de Metodología del INE: "Cálculo del efecto del diseño en la muestra nacional del empleo en el trimestre Octubre- Noviembre –Diciembre 2003".

Para la afijación de la muestra en los distintos estratos se tuvo en cuenta además de la estructura piramidal<sup>4</sup> de los errores de muestreo, otros factores relativos de cada estrato tales como:

1. La tasa de desocupación por estrato
2. Coeficiente de variación
3. Costos de actualización y entrevistas
4. La varianza poblacional con el fin de definir y analizar los márgenes de error para los distintos estratos y así poder optar por un tamaño adecuado de muestra.

De acuerdo con las etapas del diseño el tamaño de la muestra se desagrega de la siguiente manera:

Unidades de primera etapa: 3.558 secciones

Unidades de segunda etapa: 34.511 viviendas particulares

En la tabla 2 siguiente se presentan los tamaños muestrales a nivel nacional y regional por área urbano rural.

Es necesario destacar que en algunas áreas de estimación el error relativo es muy alto, debido a una baja tasa de Desocupación (por ejemplo, área Rural de la región IX). Para reducir el error en estos casos, con tasas muy pequeñas, sería necesario aumentar la muestra a niveles prohibitivos<sup>5</sup>.

---

<sup>4</sup> Es decir, el error es mayor a medida que se va desagregando de la forma: Nacional, Urbano, Rural; Regional, Urbano, Rural; Comunal y Grandes Centros, y estratos.

<sup>5</sup> En los niveles de estimación con tasas de desocupación bajas, no se justifica conocerlas con precisión. En el caso anterior, basta el orden de magnitud sugerido por la estimación y su intervalo de confianza para concluir que el desempleo en esos niveles no es un problema importante comparado con otros donde se hace urgente saber si la tasa baja de uno o dos puntos con políticas públicas apropiadas.



**Tabla 2: Tamaño Muestral y Error Relativo por Área de Estimación para Desocupados**

NOMBRE NIVEL Y ESTRATO	Desocupados	Tasa de Desocupación	Muestra Secciones	Muestra Viviendas	Error Relativo	Calificación
NACIONAL	461.072	0,08	3.551	34.455	3,6	B
URBANO	432.988	0,09	3.018	26.460	3,7	B
GRANDES CENTROS	367.021	0,09	2.439	19.512	4,2	B
RAU	65.967	0,07	579	6.948	7,9	B
RURAL	28.084	0,04	533	7.995	10,7	R
I REGION	19.444	0,09	158	1.398	10,8	R
II REGION	13.634	0,08	144	1.317	16,6	R
III REGION	10.032	0,09	145	1.371	11,3	R
IV REGION	15.208	0,07	211	2.185	15,8	R
URBANO	14.401	0,09	164	1.480	16,4	R
RURAL	807	0,02	47	705	56,2	M
V REGION	63.176	0,11	400	3.822	9,4	B
URBANO	61.076	0,11	358	3.192	9,7	B
RURAL	2.100	0,04	42	630	35,3	A
VI REGION	7.440	0,03	218	2.372	22,2	R
URBANO	5.414	0,04	158	1.472	27,8	R
RURAL	2.026	0,02	60	900	33,6	A
VII REGION	33.300	0,09	278	3.025	10,5	R
URBANO	26.114	0,11	195	1.780	12,1	R
RURAL	7.187	0,06	83	1.245	21,3	R
VIII REGION	50.945	0,08	496	4.820	8,8	B
URBANO	44.843	0,08	424	3.740	9,4	B
RURAL	6.102	0,05	72	1.080	25,0	R
IX REGION	14.406	0,05	212	2.252	15,2	R
URBANO	13.202	0,07	160	1.472	15,1	R
RURAL	1.205	0,01	52	780	74,2	M
X REGION	21.822	0,06	295	3.196	11,8	R
URBANO	17.618	0,07	211	1.936	12,6	R
RURAL	4.204	0,03	84	1.260	30,6	A
XI REGION	1.926	0,05	106	954	24,1	R
XII REGION	4.289	0,07	94	860	19,3	R
REGION METROPOLITANA	205.449	0,09	794	6.883	6,3	B
URBANO	203.479	0,09	764	6.356	6,4	B
RURAL	2.950	0,04	41	615	25,8	R

R A N G O % de Error Relativo al 95 %					coeficiente de variación	
Buena	B	Menor o igual que:	-	10	-	5,1
Regular	R	Entre:	11	30	5,6	15,3
Aceptable	A	Entre:	31	50	15,8	25,5
Mala	M	Mayor o igual que:	51		26,0	-

## 7 Tamaño Medio de Viviendas por Sección.

El Tamaño medio de viviendas por sección en cada estrato se determinó considerando el problema de minimización de la varianza del total de Desocupados, considerando una función de costos en donde los gastos de viaje entre unidades son de poca importancia. Suponiendo fijo y dado el número de secciones por estratos, se obtiene:

$$m_{h\_optimo} = \frac{S_w}{\sqrt{S_b^2 - S_w^2 / M_h}} \sqrt{C_1 / C_2}$$

Donde:

$$C = C_1 \cdot n + C_2 \cdot n \cdot m.$$

C: función de costo por el levantamiento de la encuesta en el estrato h.

$C_1 \cdot n$  : es proporcional al número de Unidades Primarias de la Muestra.

$C_2 \cdot n \cdot m$  : es proporcional al número Total de Unidades de Segunda Etapa.

$$S_b^2 = \frac{\sum_{i=1}^{n_h} m_{hi} (\bar{Y}_{hi} - \bar{\bar{Y}}_h)^2}{n_h - 1} : \text{Varianza entre secciones (unidades primarias).}$$

$$S_w^2 = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} (y_{hij} - \bar{Y}_{hi})^2}{\sum_{i=1}^{n_h} m_{hi} - n_h} : \text{Varianza entre viviendas dentro de las secciones (unidades secundarias).}$$

$M_h$  : Número de Viviendas poblacionales dentro del estrato h.

$n_h$  : Número de secciones en la muestra dentro del estrato h.

$m_h$  : Número de Viviendas en la muestra dentro del estrato h.

$m_{hi}$  : Número de viviendas en el estrato h, sección i.

Así el número de viviendas a tomar en cada sección de acuerdo al estrato está dado según la tabla siguiente:

**Tabla N° 3: Número Medio de viviendas a encuestar por sección según estrato.**

Estrato	Tamaño medio de sección	Número viviendas por Sección
Ciudad	100 - 200 viviendas	8 (con 1 más de reemplazo)
Resto Área Urbana	80 - 150 viviendas	12
Rural	60 - 80 viviendas	15

## 8 Selección de las Unidades de Muestreo.

La selección de unidades de primera etapa se efectuó, en cada estrato geográfico, con probabilidad proporcional al tamaño de las secciones, es decir según el número de viviendas (ver anexo 3).

La probabilidad de inclusión de la i-ésima unidad primaria (sección), que es proporcional al tamaño de la sección, es igual a:

$$\pi_{hi} = n_h \cdot \frac{M_{hi}}{M_h},$$

Donde:

"h" representa el índice del estrato CD, RAU/U, R, " $M_{hi}$ ", el número total de viviendas en la sección "i" del estrato "h", al Censo 2002; y " $M_h$ " es el total de viviendas en el estrato "h" al mismo Censo.

La selección de las unidades de segunda etapa (viviendas dentro de cada estrato) se realizó con probabilidad igual para las viviendas de cada sección, mediante una selección sistemática con probabilidad  $\pi_{hij} = \frac{m_{hi}}{M'_{hi}}$  donde  $M'_{hi}$  es el número de viviendas actualizado, en la sección i del estrato h, y  $m_{hi}$  es el número de viviendas a encuestar en la sección i del estrato h.

En cada vivienda todas las personas son encuestadas de modo que no hay una tercera etapa de muestreo (ver anexo2).

## 9 Distribución de La Muestra.

La muestra total (3.558 secciones) se distribuye en un período trimestral, dividiéndose en tres muestras de un tamaño aproximadamente similar, cada una de las cuales se asigna a uno de los tres meses. De esta manera, las consultas de cada una de estas muestras más pequeñas duran un mes, las cuales cada una por sí sola no es representativa de los niveles de estimación.

Así, resulta que todas las viviendas encuestadas en el mes "t", vuelven a serlo en el mes "t+3" y cada vivienda de la muestra es encuestada sólo una vez por trimestre<sup>6</sup>.

Las estimaciones para el período (mes) "t" se calculan utilizando la información de los meses "t", "t-1", "t-2".

## 10 Rotación de Viviendas de la muestra.

La rotación de secciones es un procedimiento que tiene como finalidad mantener actualizada la muestra y evitar el cansancio de los informantes.

Lo anterior consiste en actualizar la enumeración de la sección luego de un cierto período de tiempo, para registrar cambios y posibles crecimientos o decrecimientos que se hayan producido, para posteriormente efectuar una nueva selección de viviendas, con el objeto de incorporar los cambios en la estimación de las cifras que proporciona la muestra y al mismo tiempo, evitar el "cansancio" de los informantes.

Para efectuar esta rotación, las secciones de la muestra se dividen en "Turnos de Rotación" (TR), los que se estructuran considerando 1/6 aproximadamente del total de secciones urbanas (CD Y RAU) y 1/12 aproximadamente de las secciones rurales de cada región mensualmente. Esto permite rotar la totalidad de la muestra en un período de 18 meses en el caso de las secciones urbanas y en 36 meses para las secciones rurales, con lo que se logra el objetivo de mantenerla permanentemente actualizada y vigente por la incorporación de los cambios que la afectan.

## 11 Método de Estimación: Factores de Expansión de la muestra.

Los parámetros poblacionales básicos que se estiman en la ENE son los totales de personas en la fuerza de trabajo y personas desempleadas en la fuerza de trabajo. Así, el parámetro más importante del estudio: la tasa de desocupación, que es el cociente de los totales anteriores, se estima dividiendo las estimaciones correspondientes.

Por ejemplo, para estimar el total de personas desocupadas en la fuerza de trabajo se utiliza un estimador que consiste en "expandir" el valor de la variable en cada persona de la muestra al universo según la siguiente fórmula:

---

<sup>6</sup> La información de cada mes se entrega acumulada en la muestra completa.

$$\hat{Y}_{Sep} = \sum_{hij} F_{hi}^{(2)} \cdot y_j$$

Donde:

- $y_j$  número de personas en la vivienda j.
- El término  $F_{hi}^{(2)}$  se denomina factor de expansión de la sección i en el estrato h.
- Los factores anteriores, se construyen de modo que el estimador tenga propiedades estadísticas convenientes y dependen del diseño muestral, es decir, del método de selección de la muestra y de un ajuste por una proyección de población.
- El factor de expansión se puede interpretar como la cantidad de personas en la población que representa una persona en la muestra.
- Para estimar totales en otros dominios de estudio, por ejemplo, personas en la fuerza de trabajo en el tramo de edad 15 y 24 años, sólo es necesario definir una nueva variable de interés dummy con la condición requerida.

### 11.1 Construcción del Factor de Expansión

El factor de expansión esta definido como:

$$F_{hi}^{(2)} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}} \cdot \frac{P_h^{SE}}{\hat{P}_h^{SE}}$$

Donde:

$M_h$  : Número de viviendas en el estrato "h" según el censo de población y viviendas de 2002.

$M_{hi}$  : Número de viviendas en la sección "i" según el censo 2002.

$n_h$  : Número de secciones en la muestra del estrato "h".

$M'_{hi}$  : Número de viviendas actualizadas en la sección "i".

$m_{hi}$  : Número de viviendas en la muestra de la sección "i" del estrato "h".

$P_h^{SE}$  : Proyección del número de personas del sexo "S" y tramo de edad "E" según el Censo de población y Vivienda 2002, en el estrato h.

$\hat{P}_h^{SE}$  : estimación de número de personas del sexo "S" y tramo de edad "E"<sup>7</sup>, en el estrato h, según un **estimador estándar** que será explicado más adelante.

Para explicar la lógica estadística del factor anterior, conviene visualizarlo como el producto de dos componentes:

$$F_{hi}^{(2)} = F_{hi}^{(1)} \cdot \frac{P_h^{SE}}{\hat{P}_h^{SE}}$$

donde: 
$$F_{hi}^{(1)} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}}$$

El factor de expansión  $F_{hi}^{(1)}$  denominado comúnmente como factor teórico o estándar, depende solo del diseño de muestreo y se puede interpretar como el inverso de la probabilidad de inclusión en la muestra de la persona k.

La probabilidad de inclusión anterior resulta del producto de las probabilidades de inclusión de la primera etapa (secciones) por la probabilidad de inclusión de la segunda etapa (viviendas) por el total de secciones en el estrato:

$$\pi_{hij} = \pi_{hi} \cdot \pi_{j/hi} \quad : \text{Probabilidad de inclusión de la vivienda } j, \text{ en la sección } i, \text{ del estrato } h^8.$$

$$\pi_{hi} = \frac{n_h \cdot M_{hi}}{M_h} \quad : \text{Probabilidad de inclusión de la sección } i\text{-ésima en el estrato } h.$$

$$\pi_{j/hi} = \frac{m_{hi}}{M'_{hi}} \quad : \text{Probabilidad condicional de inclusión de la vivienda } j \text{ en la sección } i \text{ del estrato } h.$$

Donde:

$M_h$  : Número de viviendas en el estrato "h" según el censo de población y viviendas de 2002.

$M_{hi}$  : Número de viviendas en la sección "i" según el censo 2002.

$n_h$  : Número de secciones en la muestra en el estrato "h".

$M'_{hi}$  : Número de viviendas actualizadas en la sección "i" según el censo de población y viviendas de 2002.

<sup>7</sup> Tramos de edad E: 1 = menor de 15 años y 2 = de 15 o más años.

<sup>8</sup> Nótese que la probabilidad de inclusión cambia según el estrato y sección, pero es constante para todas las personas de cada vivienda de una sección.

$m_{hi}$  : Número de viviendas en la muestra de la sección "i" del estrato "h".

## 12 El Estimador estándar del Total.

La componente  $F_{hi}^{(1)}$  define por si sola un factor de expansión, que da lugar al **estimador estándar**<sup>9</sup> del total:

$$\hat{Y}_{hi} = \sum_{j, i \in h} F_{hi}^{(1)} \cdot y_{hi}$$

,con  $y_{hi}$  número total de personas de la característica de interés.

El estimador anterior resulta insesgado y consistente respecto del total de la variable de interés, sin embargo, su eficiencia estadística (error cuadrático medio) puede mejorarse ajustándolo por la razón  $P_h^{SE} / \hat{P}_h^{SE}$ , obteniéndose el estimador final utilizado en la ENE para estimar totales.

Con la definición del estimador estándar se completa la descripción del factor de ajuste  $P_h^{SE} / \hat{P}_h^{SE}$ , puesto que  $\hat{P}_h^{SE} = \sum_{j, i \in h} F_{hi}^{(1)} \cdot P_{hi}^{SE}$  es la estimación estándar del

número de personas en el estrato h, de sexo S y edad E, donde  $P_{hi}^{SE}$  es el número de personas en la sección i del estrato h, según el sexo S y la edad E.

## 13 Estimador de Razón Separado del Total

El estimador del total  $\hat{y}_h = \sum_{i \in h} F_{hi}^{(2)} \cdot y_{hi}$  puede anotarse en una forma más

compacta como, la suma de estimadores estándar por estrato ajustados por la proyección de población del estrato correspondiente:

$$\hat{y}_h = \sum_{i \in h} \frac{P_h^{SE}}{\hat{P}_h^{SE}} \cdot \hat{Y}_{hi}$$

El estimador anterior, llamado de Razón Separado, tiene las siguientes propiedades:

1. Consistente.
2. Tiene un **error cuadrático medio** menor que el estimador estándar.

<sup>9</sup> Este estimador se denomina de Horvitz-Thompson (1952). Para más detalles acerca de sus propiedades ver Carl-Erik Sarndal, Bengt Swensson, Jan Wretman (1992), Model Assisted Survey Sampling, página 43.

3. Es sesgado, pero el sesgo es despreciable en comparación con la ganancia que se obtiene de la estimación.
4. Permite atenuar el sesgo de no respuesta, dado que tiene la propiedad que al estimar el total de personas por género mayores de 15 años en cada estrato  $h$ , resultan exactamente las proyecciones de población en el estrato  $h$ .

La última propiedad implica que al sumar los factores  $F_{hi}^{(2)}$  para todas las personas mayores de 15 años en la muestra, se obtiene la proyección correspondiente del estrato  $h$ . En cambio, si se suman los factores  $F_{hi}^{(1)}$  para todas las viviendas de la muestra en el dominio (1 factor por vivienda), se obtiene una estimación del número de viviendas en el dominio.

## 14 Estimadores de varianzas.

La varianza o el error estándar de un estimador permite estudiar la calidad de una **inferencia** de la muestra a la población, ya que al conocer los rangos de variación de los estimadores se puede decidir con seguridad si un fenómeno observado puede o no ser debido al azar.

Las fórmulas siguientes se deducen considerando el diseño muestral bietápico estratificado. El supuesto anterior, permite simplificar las fórmulas obteniéndose estimadores conservadores, es decir, sus estimaciones sobrestiman las varianzas verdaderas.

### 14.1 Estimador de la varianza para un Total.

$$\hat{Var}(\hat{Y}_h) = \frac{n_h}{n_h - 1} \cdot \sum_{i \in h} \left[ F_{hi}^{(1)} \cdot y_{hi} - \hat{R}_h \cdot F_{hi}^{(1)} \cdot p_{hi} \right]^2$$

Donde:

$Y_{hi} = \sum_{Vivienda\ j} y_{hij}$  : Número de personas, por ejemplo, desocupados en la muestra de la sección "i".

$p_{hi} = \sum_{Vivienda\ j} p_{hij}$  : Número de personas mayores de 15 años en la muestra de la sección "i".

$\hat{R}_h = \hat{Y}_h / \hat{P}_h$  : Razón entre el número de personas de la variable de interés y el total de personas en el estrato, con  $\hat{y}_h = \sum_{i \in h} F_{hi}^{(1)} \cdot y_{hi}$  y  $\hat{p}_h = \sum_{i \in h} F_{hi}^{(1)} \cdot p_{hi}$ .



#### 14.2 Estimador de la varianza para la Tasa de Desocupación.

$$\hat{TD}_h = \frac{\text{Estimador del Total de Desocupados}}{\text{Estimador del Total de Fuerza de Trabajo}} = \frac{\hat{Y}_h}{\hat{X}_h}$$

$$\hat{Var}(\hat{TD}) = \sum_h \frac{n_h}{(n_h - 1)} \sum_i \left[ (F_{hi}^{(1)} Y_{hi} - \frac{1}{n_h} \sum_i F_{hi}^{(1)} Y_{hi}) - \hat{TD} \cdot (F_{hi}^{(1)} X_{hi} - \frac{1}{n_h} \sum_i F_{hi}^{(1)} X_{hi}) \right]^2$$

#### 15 Estimador de la varianzas del estimador estándar (sin Ajuste de Población).

$$\hat{Var}(\hat{Y}_{Estan}) = \sum_h \frac{n_h}{n_h - 1} \cdot \sum_i \left( F_{hi}^{(1)} Y_{hi} - \frac{1}{n_h} \sum_i F_{hi}^{(1)} \cdot Y_{hi} \right)^2$$

Donde:

$$Y_{hi} = \sum_{\text{Vivienda } j} y_{hij} \quad \text{:Total de personas, por ejemplo, desocupadas, en la muestra de la sección i.}$$

#### 16 Error de Muestreo.

Los errores de Muestreo de un estimador  $\hat{\theta}$  de un parámetro cualquiera  $\theta$ , por ejemplo, de un total o tasa, que se definen a continuación son el Error Absoluto y el Error Relativo.

##### 16.1 Error Absoluto de un estimador al 95% de confianza.

$$e_{a(0,95)} = \text{Error Absoluto} = 1,96 \cdot \sqrt{\hat{V}(\hat{\theta})}.$$

##### 16.2 Error Relativo de un estimador al 95% de confianza.

$$\text{Error relativo} = \frac{\text{Error Absoluto}}{\hat{\theta}} = \frac{e_{a(0,95)}}{\hat{\theta}}$$

El error relativo con una confianza del 68%, se conoce como "Coeficiente de Variación" de un estimador y resulta del cuociente entre la desviación estándar de la estimación (raíz cuadrada de la varianza) y el valor de la estimación:

$$CV(\hat{\theta}) = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}.$$

### 17 Intervalo de Confianza al 95%.

Estimada la varianza de un estimador  $\hat{\theta}$ , es posible obtener el intervalo de confianza para un parámetro  $\theta$  del total de una variable con una confianza del 95 %.

$$\left( \hat{\theta} - e_{a(0,95)} ; \hat{\theta} + e_{a(0,95)} \right)$$

## **ANEXOS**

### **ANEXO 1: Muestra por Áreas de Estimación y Error Relativo para los Desocupados (Base: Oct Dic 2003).**

NOMBRE NIVEL Y ESTRATO	NUMERO SECCIONES MARCO 2002	NUMERO VIVIENDAS MARCO 2002	Desocupados	Tasa de Desocupación	Muestra Secciones	Muestra Viviendas	Coefficiente de Variación	Error Absoluto	Error Relativo
NACIONAL	32.964	4.000.403	461.072	0,08	3.551	34.455	1,8	16.399	3,6
URBANO	26.279	3.476.817	432.988	0,09	3.018	26.460	1,9	16.120	3,7
GRANDES CENTROS	18.759	2.808.345	367.021	0,09	2.439	19.512	2,1	15.256	4,2
RAU	7.520	668.472	65.967	0,07	579	6.948	4,0	5.209	7,9
RURAL	6.685	523.586	28.084	0,04	533	7.995	5,5	3.011	10,7
I REGION	1.132	108.142	19.444	0,09	158	1.398	5,5	2.093	10,8
ARICA CD	446	45.409	8.369	0,11	65	520	8,4	1.376	16,4
IQUIQUE CD	405	41.573	5.904	0,08	60	480	9,8	1.130	19,1
ALTO HOSPICIO CD	117	12.388	4.815	0,09	10	80	11,2	1.061	22,0
II REGION	1.012	111.756	13.634	0,08	144	1.317	8,5	2.269	16,6
CALAMA CD	274	28.063	2.351	0,05	60	480	16,7	768	32,7
ANTOFAGASTA CD	538	68.579	9.806	0,10	48	384	11,0	2.114	21,6
III REGION	778	62.427	10.032	0,09	145	1.371	5,8	1.138	11,3
COPIAPO CD	338	28.959	3.276	0,08	60	480	10,7	687	21,0
VALLENAR CD	133	11.481	3.075	0,13	42	336	10,4	627	20,4
IV REGION	1.701	164.615	15.208	0,07	211	2.185	8,1	2.408	15,8
URBANO	1.219	126.752	14.401	0,09	164	1.480	8,4	2.365	16,4
COQUIMBO CD	314	39.951	5.533	0,10	38	304	15,0	1.625	29,4
LA SERENA CD	302	39.249	5.575	0,11	37	296	13,7	1.494	26,8
OVALLE CD	217	17.742	1.362	0,08	47	376	17,4	464	34,0
PR ELQUI	852	98.380	11.563	0,09	104	1.011	9,8	2.223	19,2
PR LIMARI	535	43.527	1.871	0,04	72	724	16,2	595	31,8
PR CHOAPA	314	22.708	1.773	0,06	35	450	20,4	710	40,0
RURAL	482	37.863	807	0,02	47	705	28,7	454	56,2
V REGION	4.065	449.301	63.176	0,11	400	3.822	4,8	5.944	9,4
URBANO	3.590	415.291	61.076	0,11	358	3.192	4,9	5.897	9,7
RURAL	475	34.010	2.100	0,04	42	630	18,0	742	35,3
PR PETORCA	261	18.867	3.147	0,10	31	402	12,1	747	23,8
PR LOS ANDES	313	24.831	1.590	0,05	43	424	18,7	583	36,6
LOS ANDES CD	185	15.246	1.058	0,05	29	232	22,5	467	44,1
PR SAN FELIPE	460	36.312	3.611	0,07	40	504	13,3	943	26,1
PR QUILLOTA	536	47.233	2.962	0,06	46	482	18,3	1.063	35,9
PR VALPARAISO	1.795	258.227	43.676	0,13	160	1.325	6,4	5.490	12,6
VALPARAISO CD	488	76.355	15.884	0,15	54	432	9,1	2.841	17,9
VIÑA DEL MAR CD	663	102.724	12.973	0,19	66	528	11,4	2.899	22,3
QUILPUE CD	248	37.495	7.497	0,16	22	176	14,8	2.172	29,0
VILLA ALEMANA CD	186	28.965	5.578	0,17	16	128	21,9	2.389	42,8
PR SAN ANTONIO	497	39.891	6.054	0,12	51	453	10,0	1.184	19,6

NOMBRE NIVEL Y ESTRATO	NUMERO SECCIONES MARCO 2002	NUMERO VIVIENDAS MARCO 2002	Desocupados	Tasa de Desocupación	Muestra Secciones	Muestra Viviendas	Coefficiente de Variación	Error Absoluto	Error Relativo
GRAN VALPARAISO	1.531	236.936	40.850	0,13	151	1.208	6,4	5.133	12,6
SAN ANTONIO CD	285	22.620	4.627	0,14	42	336	9,6	866	18,7
QUILL-CAL-CR CD	322	33.253	2.289	0,18	47	376	18,3	821	35,9
VI REGION	2.323	207.603	7.440	0,03	218	2.372	11,3	1.654	22,2
URBANO	1.542	148.272	5.414	0,04	158	1.472	14,2	1.507	27,8
RURAL	781	59.331	2.026	0,02	60	900	17,2	681	33,6
RANCAGUA CD	458	57.026	1.619	0,05	65	520	21,2	673	41,6
SAN FERNANDO CD	159	13.592	852	0,05	41	328	18,3	306	35,9
PR CACHAPOAL	1.496	145.051	5.617	0,04	109	1.120	14,2	1.567	27,9
PR COLCHAGUA	618	51.386	1.591	0,02	78	832	16,3	507	31,9
PR CARDENAL CARO	209	11.166	233	0,02	31	420	32,3	147	63,2
VII REGION	2.815	248.150	33.300	0,09	278	3.025	5,4	3.510	10,5
URBANO	1.813	167.355	26.114	0,11	195	1.780	6,2	3.157	12,1
RURAL	1.002	80.795	7.187	0,06	83	1.245	10,9	1.532	21,3
CURICO CD	311	25.472	5.586	0,14	40	320	10,0	1.091	19,5
TALCA CD	426	53.504	8.386	0,11	57	456	10,2	1.678	20,0
LINARES CD	209	17.621	1.762	0,08	43	344	15,4	531	30,2
PR TALCA	958	97.024	13.028	0,10	85	849	9,3	2.380	18,3
PR LINARES	829	68.812	7.207	0,08	87	959	12,2	1.719	23,9
PR CAUQUENES	232	16.498	762	0,04	35	468	22,1	329	43,2
PR CURICO	796	65.816	12.304	0,12	71	749	7,9	1.894	15,4
VIII REGION	4.573	483.336	50.945	0,08	496	4.820	4,5	4.465	8,8
URBANO	3.452	397.465	44.843	0,08	424	3.740	4,8	4.196	9,4
RURAL	1.121	85.871	6.102	0,05	72	1.080	12,8	1.525	25,0
CHILLAN CD	357	44.017	4.925	0,17	62	496	13,2	1.271	25,8
LOTA CD	107	12.539	2.276	0,13	44	352	15,1	672	29,5
CORONEL CD	191	23.550	3.347	0,12	44	352	14,0	919	27,5
LOS ANGELES CD	2.402	311.251	37.716	0,28	305	2.672	5,6	4.151	11,0
PR ÑUBLE	1.221	112.371	8.270	0,06	98	1.071	11,3	1.833	22,2
PR CONCEPCION	1.797	237.308	30.097	0,09	247	2.092	6,1	3.616	12,0
CONCEPCION CD	655	100.237	11.184	0,24	66	528	11,8	2.576	23,0
TALCAHUANO CD	409	61.730	8.945	0,20	44	352	11,2	1.964	22,0
PR ARAUCO	485	37.365	4.014	0,07	45	570	11,0	862	21,5
PR BIO-BIO	1.349	130.499	12.391	0,07	140	1.359	8,4	2.050	16,5
GRAN CONCEPCION	1.139	173.380	22.196	0,09	119	952	7,6	3.325	15,0
IX REGION	2.543	234.996	14.406	0,05	212	2.252	7,7	2.186	15,2
URBANO	1.609	159.098	13.202	0,07	160	1.472	7,7	1.994	15,1
RURAL	934	75.898	1.205	0,01	52	780	37,9	894	74,2
ANGOL CD	141	11.703	1.089	0,08	40	320	17,7	378	34,8
TEMUCO CD	558	70.462	6.887	0,15	72	576	11,5	1.554	22,6
PR MALLECO	1.008	104.609	8.681	0,05	100	1.018	9,7	1.644	18,9
PR CAUTIN	1.460	121.075	4.815	0,05	99	1.130	14,4	1.359	28,2

NOMBRE NIVEL Y ESTRATO	NUMERO SECCIONES MARCO 2002	NUMERO VIVIENDAS MARCO 2002	Desocupados	Tasa de Desocupación	Muestra Secciones	Muestra Viviendas	Coefficiente de Variación	Error Absoluto	Error Relativo
X REGION	3.084	276.360	21.822	0,06	295	3.196	6,0	2.573	11,8
URBANO	2.038	191.644	17.618	0,07	211	1.936	6,5	2.228	12,6
RURAL	1.046	84.716	4.204	0,03	84	1.260	15,6	1.288	30,6
VALDIVIA CD	316	33.400	2.973	0,07	54	432	14,3	836	28,1
OSORNO CD	359	36.638	2.873	0,07	48	384	17,6	993	34,6
PUERTO MONTT CD	369	38.892	3.194	0,07	47	376	15,2	952	29,8
PR VALDIVIA	1.064	94.405	7.321	0,06	99	1.044	11,1	1.596	21,8
PR OSORNO	667	61.953	3.978	0,06	72	723	14,5	1.134	28,5
PR LLANQUIHUE	866	79.676	4.379	0,04	83	874	12,8	1.099	25,1
PR CHIOLE-PALENA	487	40.326	6.144	0,09	41	555	10,4	1.258	20,5
XI REGION	384	24.886	1.926	0,05	106	954	12,3	465	24,1
COIHAIQUE CD	168	12.473	1.221	0,08	52	416	17,2	411	33,7
PUERTO AISEN CD	68	4.338	214	0,03	35	280	33,1	139	64,8
XII REGION	546	43.190	4.289	0,07	94	860	9,9	828	19,3
PUNTA ARENAS CD	408	34.710	3.628	0,07	73	584	11,2	798	22,0
REGION METROPOLITANA	8.008	1.585.641	205.449	0,09	794	6.883	3,2	13.027	6,3
URBANO	7.618	1.549.965	203.479	0,09	764	6.356	3,3	13.016	6,4
RURAL	432	44.303	2.950	0,04	41	615	13,2	761	25,8
GRAN SANTIAGO	6.993	1.425.287	185.820	0,09	630	5.040	3,5	12.726	6,8
PUENTE ALTO CD	629	127.753	15.283	0,10	63	504	12,0	3.591	23,5
SAN BERNARDO CD	292	60.770	11.549	0,10	56	448	8,9	2.022	17,5
MELIPILLA CD	71	14.640	1.560	0,06	38	304	24,0	735	47,1
COLINA CD	72	14.717	2.908	0,12	24	192	14,4	818	28,1
PROV CHACABUCO	200	32.910	5.572	0,05	36	354	12,5	1.364	24,5
PR CORDILLERA	692	135.481	16.852	0,10	72	627	10,9	3.599	21,4
PR MAIPO	515	95.990	15.933	0,09	77	730	8,4	2.628	16,5
PR MELIPILLA	265	37.754	2.908	0,05	54	529	20,5	1.169	40,2
PR TALAGANTE	306	55.369	6.178	0,07	55	643	11,5	1.393	22,5

## ANEXO 2: Procedimiento de Selección de las Unidades de primera y segunda Etapa.

Las unidades de primera etapa (secciones) serán seleccionadas con probabilidad proporcional a su tamaño, según el número de viviendas, según el procedimiento sistemático siguiente.

Se construyen "N" ( = número de secciones en el estrato) intervalos, de la siguiente manera:

Sección	Nº de Viviendas	Intervalos
1	$M_1$	1 y $M_1$
2	$M_2$	$M_1 + 1$ y $M_1 + M_2$
3	$M_3$	$M_1 + M_2 + 1$ y $M_1 + M_2 + M_3$
$\vdots$	$\vdots$	$\vdots$
N	$M_N$	$M_1 + \dots + M_{N-1} + 1$ y $M_1 + \dots + M_N = M$

A continuación, se genera un número aleatorio "A" entre 1 y  $k = M/n$ , entonces, las secciones seleccionadas están determinadas por el intervalo a cual pertenecen las cantidades:

$$A, A + k, A + 2k, \dots, A + (n - 1)k.$$

El procedimiento anterior, no permite repeticiones y se puede demostrar que la probabilidad de inclusión de una unidad con "Mi" viviendas es "n Mi/M".

### ANEXO 3: Áreas de descuento ADA.

Región	Localidad	Entidad	Situación
I	Valle de Lluta (P)	Sora	Difícil acceso, camino mal estado arenoso, dunas y quebrada peligrosas. Ent. Sora (Pc.), difícil acceso, camino que cruza río de gran caudal y torrencioso por quebrada, sin puente.
II			Toda la Comuna de Ollagüe.
III	Ciénaga Redonda	Campo Marte	Distrito N° 19 de Copiapó completo es un ADA pues es un sector muy alejado y Cordillerano.
IV	Caldera y Damas	La Sombría	Difícil acceso
V	Alta Montaña	Los Libertadores	Solo en invierno
VI	Chancón (P)	Anita	Acceso sólo meses enero a marzo. DC12, se ratifica conformación de ADA
VII	El Baúl	El Baúl	Se mantiene ADA, sector de muy difícil acceso, requiere de vehículo con tracción en las 4 ruedas y de buen tiempo. Muy alejado de los centros urbanos.
VIII	Puerto Sur	Ex Colonia Penal	Área de alto Costo todo el DC 10, corresponde a Isla Santa María
IX	Chilpaco (P)	Chilpaco (P)	Se mantiene ADA. Existe entidad Chilpaco (P).
X	Hueyusca (P)	El Mirador	DC7 Localidad Hueyusca, Entidad El Mirador (Cs) es área de difícil acceso.
XI	Las Bandurrias (P)	Las Bandurrias	AREA DE EXCLUSIÓN
XII	Entre Vientos	Monte Bello	AE (ZONA MILITAR)
XIII	El Ingenio (P)	El Ingenio	Difícil acceso

Se mencionan las regiones que presentan comunas completas:

II Región (se excluye la comuna de Ollagüe)

V Región (se excluye la comuna de Isla de Pascua y Juan Fernández)

X Región (se excluye las comunas de Cochamó, Futaleufú, Hualaihué y Palena)

XI Región (se excluye las comunas de Guaitecas, O'Higgins y Tortel)

XII Región (se excluye las comunas de Cabo de Hornos, Antártica)

#### **ANEXO 4: Comparación de Niveles de Estimación entre Los Niveles Actuales y la nueva propuesta.**

Se presenta los Niveles de Estimación de la Muestra de la Encuesta Nacional de Empleo 1996 en forma paralela a los Niveles de Estimación de la Muestra Propuesta para la ENE 2006, basada en el Censo 2002. Se destacarán los Niveles Adicionales a considerar en la propuesta ENE 2006 y los Niveles de la Muestra Actual ENE '96 que no se estarían considerando.

El Cuadro Siguiente Muestra en forma paralela los niveles de ambas Muestra (la actual y la propuesta):

<b>NIVELES ACTUALES</b>	<b>NIVELES NUEVOS</b>
TOTAL NACIONAL	NACIONAL
URBANO	URBANO
GRANDES CENTROS URBANOS	GRANDES CENTROS
R.A.U NACIONAL	RAU
RURAL	RURAL
I REGION DE TARAPACA	I REGION
CIUDAD DE ARICA	ARICA CD
CIUDAD DE IQUIQUE	IQUIQUE CD
	<b>ALTO HOSPICIO CD</b>
II REGION DE ANTOFAGASTA	II REGION
CIUDAD CALAMA	CALAMA CD
<b>CIUDAD CHUQUICAMATA</b>	
CIUDAD ANTOFAGASTA	ANTOFAGASTA CD
III REGION DE ATACAMA	III REGION
CIUDAD COPIAPO	COPIAPO CD
CIUDAD VALLENAR	VALLENAR CD
IV REGION DE COQUIMBO	IV REGION
URBANO	URBANO
RURAL	RURAL
CIUDAD COQUIMBO	COQUIMBO CD
CIUDAD LA SERENA	LA SERENA CD
CIUDAD OVALLE	OVALLE CD
PROVINCIA ELQUI	PR ELQUI
PROVINCIA LIMARI	PR LIMARI
PROVINCIA CHOAPA	PR CHOAPA



NIVELES ACTUALES	NIVELES NUEVOS
V REGION VALPARAISO	V REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA PETORCA	PR PETORCA
PROVINCIA LOS ANDES	PR LOS ANDES
	<b>LOS ANDES CD</b>
PROVINCIA SAN FELIPE	PR SAN FELIPE
PROVINCIA QUILLOTA	PR QUILLOTA
PROVINCIA VALPARAISO	PR VALPARAISO
GRAN VALPARAISO	GRAN VALPARAISO
CIUDAD VALPARAISO	VALPARAISO CD
CIUDAD VIÑA DEL MAR	VIÑA DEL MAR CD
	<b>QUILPUE CD</b>
	<b>VILLA ALEMANA CD</b>
PROVINCIA SAN ANTONIO	PR SAN ANTONIO
CIUDAD SAN ANTONIO	SAN ANTONIO CD
CONJUNTO : QUILLOTA CALERA - LA CRUZ	QUILL-CAL-CR CD
LIBERTADOR GENERAL BERNARDO O'HIGGINS	VI REGION
URBANO	URBANO
RURAL	RURAL
CIUDAD RANCAGUA	RANCAGUA CD
CIUDAD SAN FERNANDO	SAN FERNANDO CD
PROVINCIA DE CACHAPOAL	PR CACHAPOAL
PROVINCIA COLCHAGUA	PR COLCHAGUA
PROVINCIA CARDENAL CARO	PR CARDENAL CARO
VII REGION DEL MAULE	VII REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA CURICO	PR CURICO
CIUDAD CURICO	CURICO CD
PROVINCIA TALCA	PR TALCA
CIUDAD TALCA	TALCA CD
PROVINCIA LINARES	PR LINARES
CIUDAD LINARES	LINARES CD
PROVINCIA CAUQUENES	PR CAUQUENES

NIVELES ACTUALES	NIVELES NUEVOS
VIII REGION DEL BIOBIO	VIII REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA ÑUBLE	PR ÑUBLE
CIUDAD CHILLAN	CHILLAN CD
PROVINCIA CONCEPCION	PR CONCEPCION
GRAN CONCEPCION	GRAN CONCEPCION
CIUDAD CONCEPCION	CONCEPCION CD
CIUDAD TALCAHUANO	TALCAHUANO CD
CIUDAD LOTA	LOTA CD
CIUDAD CORONEL	CORONEL CD
PROVINCIA ARAUCO	PR ARAUCO
PROVINCIA BIO-BIO	PR BIO-BIO
CIUDAD LOS ANGELES	LOS ANGELES CD
IX REGION DE LA ARAUCANIA	IX REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA MALLECO	PR MALLECO
CIUDAD ANGOL	ANGOL CD
CIUDAD TEMUCO	TEMUCO CD
PR CAUTIN	PR CAUTIN
X REGION DE LOS LAGOS	X REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA VALDIVIA	PR VALDIVIA
CIUDAD VALDIVIA	VALDIVIA CD
PROVINCIA OSORNO	PR OSORNO
CIUDAD OSORNO	OSORNO CD
PROVINCIA LLANQUIHUE	PR LLANQUIHUE
CIUDAD PUERTO MONTT	PUERTO MONTT CD
PROVINCIA CHILOE-PALENA	PR CHILOE-PALENA
XI REGION AISEN DEL GENERAL CARLOS IBANEZ DEL CAMPO	XI REGION
CIUDAD COHAIQUE	COIHAIQUE CD
CIUDAD PUERTO AISEN	PUERTO AISEN CD
XII REGION DE MAGALLANES Y DE LA ANTARTICA CHILENA	XII REGION
CIUDAD PUNTA ARENAS	PUNTA ARENAS CD

NIVELES ACTUALES	NIVELES NUEVOS
REGION METROPOLITANA DE SANTIAGO	RM REGION
URBANO	URBANO
RURAL	RURAL
PROVINCIA SANTIAGO	PROVINCIA SANTIAGO
GRAN SANTIAGO	GRAN SANTIAGO
CIUDAD PUENTE ALTO	PUENTE ALTO CD
CIUDAD SAN BERNARDO	SAN BERNARDO CD
CIUDAD MELIPILLA	MELIPILLA CD
	<b>COLINA CD</b>
PROVINCIA CHACABUCO	PROV CHACABUCO
PROVINCIA CORDILLERA	PR CORDILLERA
PROVINCIA MAIPO	PR MAIPO
PROVINCIA MELIPILLA	PR MELIPILLA
PROVINCIA TALAGANTE	PR TALAGANTE

## **ANEXO 6: Grandes Ciudades.**

Como se detalla en la clasificación de Centros Urbanos, se denomina Metrópoli, a la mayor representación urbana que tiene un país; esta concentra a más de 1.000.000 de habitantes y corresponde a un elevado porcentaje de la población total nacional, está conformada por el área urbana de un conjunto de comunas que se han unido producto de la expansión territorial y procesos de "conurbación".

### **a. Gran Santiago.**

Gran Santiago, en el mismo esquema detallado más arriba, del nivel de Grandes Áreas Urbanas, definido para el Censo 1992, pasó por su gran envergadura poblacional a conformar el nivel Metrópoli de Santiago", que tiene un tratamiento operativo de especial atención.

<b>Metrópolis de Santiago (2002)</b>	<b>Gran Santiago (1992)</b>
32 comunas de la provincia de Santiago Comuna de Puente Alto Comuna de San Bernardo Comuna de Padre Hurtado Cd de Pirque (Comuna de Pirque) *  La Obra-Las Vertientes (Comuna de San José de Maipo) *	32 comunas de la provincia de Santiago Comuna de Puente Alto Comuna de San Bernardo Comuna de Padre Hurtado
* Tanto la Ciudad de Pirque con 4.855 hbtes., como la Ciudad La Obra-Las Vertientes con (2.477) hbtes., en la nueva Muestra ENE pertenecen al estrato Provincia de Cordillera RAU.	

### **b. Gran Valparaíso.**

Gran Valparaíso esta compuesto por las Ciudades de Valparaíso, Quilpue, Villa Alemana y Viña del Mar.

### **c. Gran Concepción.**

Gran Concepción esta compuesto por las Ciudades de Concepción, Chiguayante, Penco, San Pedro de la Paz y Talcahuano.

### **d. Gran Temuco.**

Gran Temuco esta conformado por las Ciudades de Temuco y Padre Las Casas.

### **e. Gran Serena**

Gran Serena esta compuesto por las Ciudades de Coquimbo y la Serena.

Los Niveles de Estimación Adicionales que se considerarían en la propuesta para la Nueva Encuesta Nacional del Empleo 2006 serían:

- Ciudad de Alto Hospicio
- Ciudad de Los Andes
- Ciudad de Quilpue
- Ciudad de Villa Alemana
- Ciudad de Colina.

Los Niveles de Estimación, que se encuentran en la Actualidad, de la ENE 96 y no serían considerados en la Nueva ENE 2006 propuesta sería la Ciudad de Chuquicamata.

## **ANEXO 7: Estimadores de Totales, Proporciones, Tasas y Varianzas.**

### **Procedimientos de Cálculo de los Estimadores de la Encuesta Nacional del Empleo.**

La base muestral debe contener los datos de la encuesta y los parámetros utilizados para la definición del Factor de Expansión y las formulas requeridas para Promedios, Tasas y Proporciones.

Esta base que contiene la información de la encuesta de empleo para ser proyectada a la población requiere de procedimientos asociados al diseño de la muestra.

Para ello se adjuntan, en forma técnica, los algoritmos utilizados para expandir al universo la muestra seleccionada. Lo que será aplicado en programas informáticos.

### **Factores de Expansión de la muestra.**

El Factor de Expansión se expresa como las veces que representa una unidad seleccionada con respecto al universo.

Los factores utilizados para la expansión de la Encuesta Nacional del Empleo de acuerdo al diseño se expresa como sigue:

### **Factor de Expansión Teórico en la vivienda $j$ de la sección del estrato $h$ .**

$$FE_{hi}^{(1)} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}}$$

$M_h$  = Número de viviendas en el estrato  $h$  según el censo de población y viviendas de 2002.

$M_{hi}$  = Número de viviendas, en la sección  $i$  del estrato  $h$ , según el censo 2002.

$n_h$  = Número de secciones en la muestra del estrato  $h$ .

$M'_{hi}$  = Número de viviendas actualizadas, en la sección  $i$  del estrato  $h$ .  
 $m_{hi}$  = Número de viviendas en la muestra, en la sección  $i$  del estrato  $h$ .

**Factor de Expansión Ajustado en la vivienda  $i$ , sección  $j$  y estrato  $h$ .**

$$FE_{hi}^{(2)} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}} \cdot \frac{P_h^{SE}}{\hat{P}_h^{SE}},$$

Con:

$P_h^{SE}$  = Proyección de población de personas del sexo "S" y tramo de edad "E" (de 15 años ó más, y menor a 15 años) según el Censo población y Vivienda 2002 en el estrato  $h$ .

$\hat{P}_h^{SE}$  = Número estimado de personas del sexo "S" y tramo de edad "E" en el estrato  $h$ .

$\hat{P}_h^{SE} = \sum_j FE_{hi}^{(1)} \cdot P_{hij}^{SE}$  : Estimación estándar de personas del estrato  $h$ , del sexo "S" y tramo de edad "E" (mayor o igual a 15 años y menor a 15 años).

$P_{hij}^{SE}$  = número de personas de la vivienda  $j$ , sección  $i$ , estrato  $h$ , del sexo "S" y tramo de edad "E" (mayor o igual a 15 años y menor a 15 años) que se encuentra en la fuerza de trabajo.

**Estimadores.**

Los estimadores asociados a este diseño son no-autoponderados debido al número fijo de viviendas seleccionadas por sección.

Las estimaciones se efectúan considerando la muestra completa y sobre la base de períodos móviles trimestrales, en los que se incluye el mes en proceso y los dos meses anteriores. La información exógena sobre proyecciones de población que se utiliza en los cálculos, se refiere al mes central.

La estimación de un total dado para una variable, se obtiene, primero, multiplicando el valor de la variable en cada persona por su Factor de Expansión y luego sumando sobre toda las personas de la muestra.

La estimación de población por estrato, sexo y tramo de edad, se obtiene al ponderar cada persona por el Factor de Expansión Teórico y luego sumar todas las personas de la muestra dentro del estrato, por sexo y tramo de edad.

**Estimador Total de personas con la condición c.**

$$\hat{P}_c = \sum_h \sum_i \sum_j FE_{hi}^{(1)} \cdot [P_{hijc}]$$

$P_{hijc}$  = Total de personas en la vivienda j, de la sección i del estrato h, con la condición c.

**Tasa (Promedio) estimada de la condición b, por persona con la condición c.**

$$\hat{T}_{bc} = \frac{\hat{P}_b}{\hat{P}_c} \quad ; \quad \hat{P}_b = \text{Total estimado de personas con la condición b.}$$

**Varianza del Total Estimado**

$$\hat{Var}(\hat{P}_c) = \frac{n_h}{n_h - 1} \cdot \sum_{i \in h} [F_i^{(1)} \cdot p_{ic} - \hat{R}_c \cdot F_i^{(1)} \cdot p_{ic}]^2$$

**Error de Muestreo.**

Los errores de Muestreo que se definen a continuación son el Coeficiente de Variación, el Error Absoluto y el Error Relativo.

El coeficiente de Variación de una estimación se obtiene del cuociente entre la desviación estándar de la estimación (raíz cuadrada de la varianza) y el valor de la estimación.

$$\text{Coeficiente de Variación:} \quad CV(\hat{P}_c) = \frac{\sqrt{V(\hat{P}_c)}}{\hat{P}_c}.$$

$$\text{Error Absoluto:} \quad Error\ Absoluto = 1,96 \cdot \sqrt{V(\hat{P}_c)}.$$

$$\text{Error Relativo:} \quad Error\ Relativo = \frac{Error\ Absoluto}{\hat{P}_c}$$

**Intervalo de Confianza Estadística del Estimador entre 2 periodos (Trimestres Móviles).**

Conocida la Varianza del Estimador, es posible obtener el intervalo de confianza del total de personas, en la población, con la condición "c" con una confianza del 95 %.

**Intervalo de Confianza al 95%.**

$$\left( \hat{P}_c - 1,96 \times \sqrt{V(P_c)} \quad ; \quad \hat{P}_c + 1,96 \times \sqrt{V(P_c)} \right)$$



## **BIBLIOGRAFÍA**

- Des Raj (1968), "Sampling Theory".
- Cochran W. G. (1998), "Técnicas de Muestreo".
- Carl-Erik Sarndal, Bengt Swensson, Jan Wretman (1992), "Model Assisted Survey Sampling".
- C. J. Skinner, D. Holt y T. M. Smith (1989), "Analysis of Complex Surveys".