INFORME METODOLOGÍA DISEÑO MUESTRAL EFECTIVA
ENCUESTA DE COSECHA DE CULTIVOS ANUALES 2015

ESTADISTICA AGRICOLA CONTINUA INTERCENSAL

INSTITUTO NACIONAL DE ESTADÍSTICAS

Junio / 2015

METODOLOGÍA DISEÑO MUESTRAL TEÓRICO
ENCUESTA DE COSECHA DE CULTIVOS ESENCIALES 2015

ESTADISTICA AGRICOLA CONTINUA INTERCENSAL
Instituto Nacional de Estadísticas.

Junio / 2015

Jefe Depto. Investigación y Desarrollo: Charles Duran Artigas.

Coordinador Sección Estadísticas Económicas: Leticia Morales Maluenda.

Analista Estadístico: Claudio Olivares Concha.

ÍNDICE

1. Objetivos 1
1.1. Objetivo General 1
1.2. Objetivos Específicos 1
2. Población Objetivo 1
3. Unidad de muestreo 1
4. Cobertura 1
4.1. Cobertura temática 1
4.2. Cobertura geográfica 2
5. Niveles de estimación 2
6. Marco Muestral 2
6.1. Estratificación 2
7. Método de muestreo 3
7.1. Tamaño muestral teórico 3
7.2. Afijación de la muestra 4
7.3. Selección de la muestra 5
8. Factores de expansión 5
9. Estimadores 6
9.1. Cálculo del estimador de Rendimiento 6
9.2. Cálculo de la estimación de varianza del estimador del estimador de rendimiento. 6
9.3. Cálculo de la estimación del coeficiente de variación del estimador de rendimiento. 7
9.4. Cálculo de la estimación de cosecha de un cultivo “k” para la región. 8

1. Objetivos

1.1. Objetivo General

Elaborar una muestra que permita obtener estimaciones consistentes para la cosecha de cultivos esenciales en el año 2014-2015.

1.2. Objetivos Específicos

- Elaborar un diseño muestral de acuerdo a las características de interés de la población en estudio.

- De acuerdo al diseño especificado obtener un tamaño de muestra que para un error dado en la principal variable de interés minimice el costo de muestreo.

- De acuerdo al diseño y tamaño de muestra obtenido, seleccionar las unidades que conformaran el directorio de la muestra.

2. Población Objetivo

Explotaciones que hayan informado superficie sembrada con cultivos anuales en el Censo Agropecuario Y Forestal 2007.

3. Unidad de muestreo

La unidad de muestreo está constituida por una explotación que haya informado siembra en cultivos anuales al momento de la Encuesta de Cultivos Esenciales 2014-2015.

4. Cobertura

4.1. Cobertura temática

Explotaciones que hayan informado siembra en cultivos anuales al momento de la Encuesta de Cultivos Esenciales 2014-2015.

4.2. Cobertura geográfica

Regiones de Coquimbo, Valparaíso, Libertador General Bernardo O´Higgins, Del Maule, Biobío, La Araucanía, Los Lagos, Metropolitana y Los Ríos.

5. Niveles de estimación

Regional y nacional considerando un error de estimación teórico relativo de un 5% para las variables de estratificación con un nivel de confianza del 95%.

6. Marco Muestral

El marco muestral se construye a partir de la Encuesta de Cultivos Anuales 2014 – 2015.

6.1. Estratificación

La estratificación del marco muestral se realiza a nivel de región y tamaño de la explotación en hectáreas, de acuerdo al criterio establecido para la Encuesta de Cultivos Esenciales 2014 – 2015, es decir considerando los tamaños informados a momento del Censo Agropecuario y Forestal año 2007, el cual es actualizado parcialmente para muestra de cultivos realizada durante el año 2014.

Cuadro 1. Explotaciones y Superficies Agrícolas según Tamaño de la explotación en hectáreas. Año agrícola 2014/2015

	Explotaciones		Superficie			
Estrato Número de Hectareas	Número	Porcentaje	Hectareas	Porcentaje	Promedio por explotación	Desviación Estándar
Total	49.124	100,0	585.916,0	100,0	11,9	52,2
[1 2]	21.618	44,0	30.735,2	5,2	1,4	0,4
]2 5]	13.615	27,7	46.617,4	8,0	3,4	0,9
]5 10]	6.449	13,1	48.179,5	8,2	7,5	1,5
]10 20]	3.469	7,1	50.520,3	8,6	14,6	2,9
]20 50]	2.022	4,1	66.527,3	11,4	32,9	8,7
]50 100]	924	1,9	66.939,6	11,4	72,4	14,5
]100 200]	583	1,2	84.534,2	14,4	145,0	28,8
]200 500]	342	0,7	104.622,2	17,9	305,9	78,7
]Más de 500]	102	0,2	87.240,4	14,9	855,3	327,3

7. Método de muestreo

El método de muestreo empleado es probabilístico y estratificado por región tamaño de la explotación.

7.1. Tamaño muestral teórico

Para obtener el tamaño de muestra se considera un error relativo de 5% respecto al total de cultivos anuales y un nivel de confianza de 95% para un nivel regional.

La fórmula a utilizada para obtener el tamaño de muestra será la siguiente:

$$n_r^* = \frac{\sum_{i=1}^{9}(n_i\, S_i)^2}{V + \sum_{i=1}^{9} n_i \; \sum_{i=1}^{9} n_i\, S_i^2}$$

$$S_i^2 = \frac{\sum_{i=1}^{9}\left(y_{ij} - \overline{y}_i\right)^2}{n-1}$$

Donde:

n_r^* : Tamaño de muestra a obtener en la región.

n : Cantidad explotaciones en el marco muestral elaborado a partir de la muestra de siembra de cultivos esenciales 2014-2015 en estrato "i" de la región k, i=1,2,…,9.

S_i^2: Estimación de la varianza de la variable de interés "y" en el estrato "i" de la región "k" en la encuesta de cultivos esenciales 2013-2014, i=1,2,…,9.

$\overline{y}_i$: Promedio de la muestra en el estrato de tamaño "i"

La suma de los tamaños de muestra obtenidos en cada región dará el total de la muestra.

Cuadro 2. Tamaños de muestra efectiva, CV a nivel Nacional y región para el Total de Cultivos Anuales de la Encuesta de Cosecha de Cultivos Anuales 2014 – 2015.

Regíon	Tamaño			
	Inclusión Forzosa	Selección Aleatoria	Total	C.V
País	**56**	**1.464**	**1.520**	**1,9**
De Coquimbo	5	82	87	2,3
De Valparaíso	3	96	99	2,2
Metropolitana	5	145	150	1,5
Del Lib. Gral. Bdo. O´Higgins	10	171	181	5,0
Del Maule	9	178	187	5,7
Del Biobío	7	216	223	4,2
De La Araucanía	8	277	285	3,0
De Los Ríos	3	145	148	3,2
De Los Lagos	6	154	160	3,5

7.2. Afijación de la muestra

Obtenidos los tamaños muestrales para cada región se hizo la afijación de la muestra en los estratos tamaño usando la fórmula:

$$n_i^* = n^* \frac{n_i \; S_i}{\sum_{i=1}^{9} n_i \, S_i}$$

n_i^* : Tamaño de muestra a elegir para el estrato "i" de la región.

n_i: Tamaño de la muestra de la Encuesta de Cultivos Anuales 2014- 2015 en el estrato "i" de la región.

Cuadro 3. Tamaños muestral y error de Estimación muestra Efectiva
Encuesta de Cosecha 2014-2015

Regíon	Tamaño			
	Inclusión Forzosa	Selección Aleatoria	Total	Error de Estimación
País	**56**	**1.464**	**1.520**	**3,7**
De Coquimbo	5	82	87	4,4
De Valparaíso	3	96	99	4,3
Metropolitana	5	145	150	2,9
Del Lib. Gral. Bdo. O´Higgins	10	171	181	9,9
Del Maule	9	178	187	11,2
Del Biobío	7	216	223	8,3
De La Araucanía	8	277	285	5,8
De Los Ríos	3	145	148	6,3
De Los Lagos	6	154	160	6,8

7.3. Selección de la muestra

La selección de la muestra es aleatoria dentro de cada estrato, los tamaños regionales y nacionales se obtienen por agregación.

8. Factores de expansión

Los factores de expansión muestran a cuantas explotaciones de la población se expande cada una de las explotaciones seleccionadas en la muestra, se obtienen aplicando la siguiente formula.

$$FE_i = \frac{1}{\frac{n_i}{N_i}} = \frac{N_i}{n_i}$$

Dónde:

N_i : Número de explotaciones que en el marco muestral están contenidas en el estrato "i".

n_i : Número de explotaciones que en la muestra seleccionada están contenidas en el estrato "i".

9. Estimadores

9.1. Cálculo del estimador de Rendimiento

El estimador de rendimiento es un estadístico que se obtiene a partir del cociente entre la expansión simple de la cosecha en quintales y la expansión simple de las superficies sembradas de una especie en particular al momento de levantar la encuesta de cosecha medida en quintales por hectárea, la fórmula que permite calcular este estimador es:

$$\hat{R}_k = \frac{\sum_{i=1}^{9} FE_i \sum_{j=1}^{n_i} Q_{kij}}{\sum_{i=1}^{9} FE_i \sum_{j=1}^{n_i} T_{kij}}$$

Dónde:

$\hat{R}_k$: Estimador de rendimiento a nivel regional expresada en quintales/hectárea obtenida para la especie "k".

Q_{kij} : Cosecha obtenida de la especie "*k*" en la explotación "j" del estrato "i" en la región.

T_{kij} : Superficie sembrada informada para la especie "*k*" en la explotación "j" del estrato "i" en la *región.*

FE_i : Factor de expansión para el estrato "i" de la región.

9.2. Cálculo de la estimación de varianza del estimador del estimador de rendimiento.

Para calcular la varianza del estimador de rendimiento para una especie "k" se aplica la formula.

$$\hat{V}(\hat{R}_k) = \frac{1}{\hat{T}_k^2} \sum_{i=1}^{9} \left(\frac{N_i}{n_i}\right)\left(\frac{N_i - n_i}{n_i - 1}\right) \sum_{j=1}^{n_i} d_{kij}^2$$

$$d_{kij} = \left(\left(Q_{kij} - \hat{R}_k T_{kij}\right) - \left(\bar{Q}_{ki} - \hat{R}_k \bar{T}_{ki}\right)\right)$$

Donde:

$\hat{V}(\hat{R}_k)$: Estimación de la varianza del estimador de rendimiento de cosecha para una especie "k" a nivel regional.

$\hat{R}_k$: Estimador de rendimiento nivel regional expresado en quintales/hectárea u otra medida de rendimiento obtenida para la especie "k".

Q_{kij} : Cosecha obtenida de la especie "k" en la explotación "j" del estrato "i" en la región.

T_{kij} : Superficie sembrada informada para la especie "k" en la explotación "j" del estrato "i" en la región.

$\overline{Q}_{ki}$: Producción promedio por explotación para especie "k" en el estrato "i" de la región.

$\overline{T}_{ki}$: Superficie cultivada promedio por explotación de la especie "k" en el estrato "i" de la región.

n_i : Cantidad de explotaciones contenidas en el estrato tamaño "i" al momento de la encuesta de cosecha.

N_i : Cantidad de explotaciones contenidas en el estrato tamaño "i" al momento en el marco usado para la cosecha de cultivos anuales.

FE_i : Factor de expansión para el estrato "i" de la región.

9.3. Cálculo de la estimación del coeficiente de variación del estimador de rendimiento.

Para calcular el coeficiente de variación del estimador de rendimiento se aplic*a* la formula.

$$C.V.(\hat{R}_k) = \frac{\sqrt{\hat{V}(\hat{R}_k)}}{\hat{R}_k}$$

$C.V.(\hat{R}_k)$: Coeficiente de variación del estimador de rendimiento se aplic*a* la formula.

$\hat{V}(\hat{R}_k)$: Estimación de la varianza del estimador de rendimiento de cosecha para una especie "k" a nivel regional.

$\hat{R}_k$: Estimador de rendimiento a nivel regional expresado en quintales/hectárea u otra medida de rendimiento obtenida para la especie "k".

9.4. Cálculo de la estimación de cosecha de un cultivo "k" para la región.

Para estimar la cosecha para un cultivo "k" se aplica la siguiente formula.

$$\widehat{Q}_k = \widehat{R}_k \widehat{Y}_k$$

Donde:

$\widehat{Q}_k$: Estimación de cosecha a nivel regional expresada en quintales/hectárea u otra medida de rendimiento obtenida para la especie "k".

$\widehat{R}_k$: Estimación del rendimiento de cosecha a nivel regional para una especie "k"

$\widehat{Y}_k$: Superficie estimada obtenida para especie "k" en la encuesta de cultivos anuales 2014 – 2015.