

MEDICIÓN DEL ERROR DE MUESTREO UTILIZANDO TÉCNICAS DE CONGLOMERADOS Y GRUPOS ALEATORIOS EN UNIVERSOS AGROPECUARIOS

Lic. Pinto Ahjuacho, Jaime Tito

✉ titojaime_pinto@yahoo.com

RESUMEN

La estimación de parámetros del sector agropecuario, demanda el planteamiento de metodologías de muestreo que plantean estimar la varianza de los estimadores que se construyan, es importante medir la precisión, medir el error de muestreo, para lo cual acudimos a la técnica del muestreo por conglomerados y grupos aleatorios, realizando una aplicación vemos el procedimiento de determinar el error de muestreo para las variables de estudio.

PALABRAS CLAVE

Muestreo por conglomerados, Grupos Aleatorios, Error de muestreo.

ABSTRACT

The estimation of parameters of the agricultural sector, demands the approach of sampling methodologies that propose to estimate the variance of the estimators that are constructed, it is important to measure the precision, measure the sampling error, for which we turn to the technique of cluster sampling and randomized groups, making an application we see the procedure of determining the sampling error for the study variables.

KEYWORDS

Cluster sampling, Random Groups, Sampling error.

1. INTRODUCCIÓN

La medición del error de muestreo, se puede indicar que es la incertidumbre que se comete al estimar un parámetro poblacional del universo de estudio mediante el valor obtenido a partir de una muestra de ese universo, utilizándose estadígrafos, es una medida de la variabilidad que se observaría entre todas las muestras posibles si fueran seleccionadas usando el mismo diseño de muestral.

Se debe aceptar que el error muestral, puede ser debido a muchos factores, uno puede ser a causa del diseño muestral, que plantea el conocer el universo, la determinación del tamaño de la muestra, la selección de las

unidades de la muestra y el procedimiento de la estimación de la variables de estudio; por lo cual se analiza la varianza del estimador, concretamente la desviación estándar del estimador, al cual se le llama error de muestreo.

El estadígrafo para medir el error de muestreo dependerá de la muestra diseñada, de la técnica de muestreo utilizada, siendo estos cálculos complejos dependiendo del diseño muestral.

Mientras más pequeños sean los márgenes de error, los resultados de las encuestas serán más exactos, por ello se debe trabajar para que las muestras sean más eficientes, para que no haya errores al publicar los resultados del

margen de error de encuestas por muestreo.

Existen técnicas de muestreo que aportan en minimizar la varianza de los estimadores en estudio de población por encuestas por muestreo, tales como el muestreo por conglomerados, métodos de estimaciones de varianza de estimadores.

En el sector agropecuario, se puede plantear algunas técnicas que muestren la estimaciones de los parámetros de estudio y la medición de los errores de muestreo, una de las técnicas es la de conglomerados que muestra que plantea la minimización de la varianza del estimador, pero buscando que sea mas mínima, se puede plantear usar la varianza del error utilizando el método de Grupos Aleatorios.

Hay razones para aplicación del muestreo por conglomerados, se ha encontrado que para muchas encuestas no se tiene una lista confiable de los elementos de la población y además sería demasiado costoso formular dicha lista, no existen listas completas y actualizadas de la gente, las viviendas, o las granjas en grandes regiones geográficas, sin embargo, a partir de los mapas de la región, se puede dividirla en unidades de área, como serian las manzanas de una ciudad y los terrenos de área, como suelen elegirse estos conglomerados porque resuelven al problema de la construcción de una lista de unidades de muestreo(Cochran, 1998).

El Muestreo por Conglomerados, sugiere y desarrolla que las unidades de muestreo se pueden agrupar en subconjuntos, denominados conglomerados, de forma tal que haya heterogeneidad entre las unidades de un mismo conglomerado y homogeneidad entre conglomerados.

El concepto de homogeneidad entre conglomerados se refiere a que las medidas que se pueden calcular para cada conglomerado difieren poco de conglomerado en conglomerado. Al existir un patrón de conglomerados de las unidades muestrales, se obtiene una estimación más precisa si se muestrean aleatoriamente un número determinado de conglomerados.

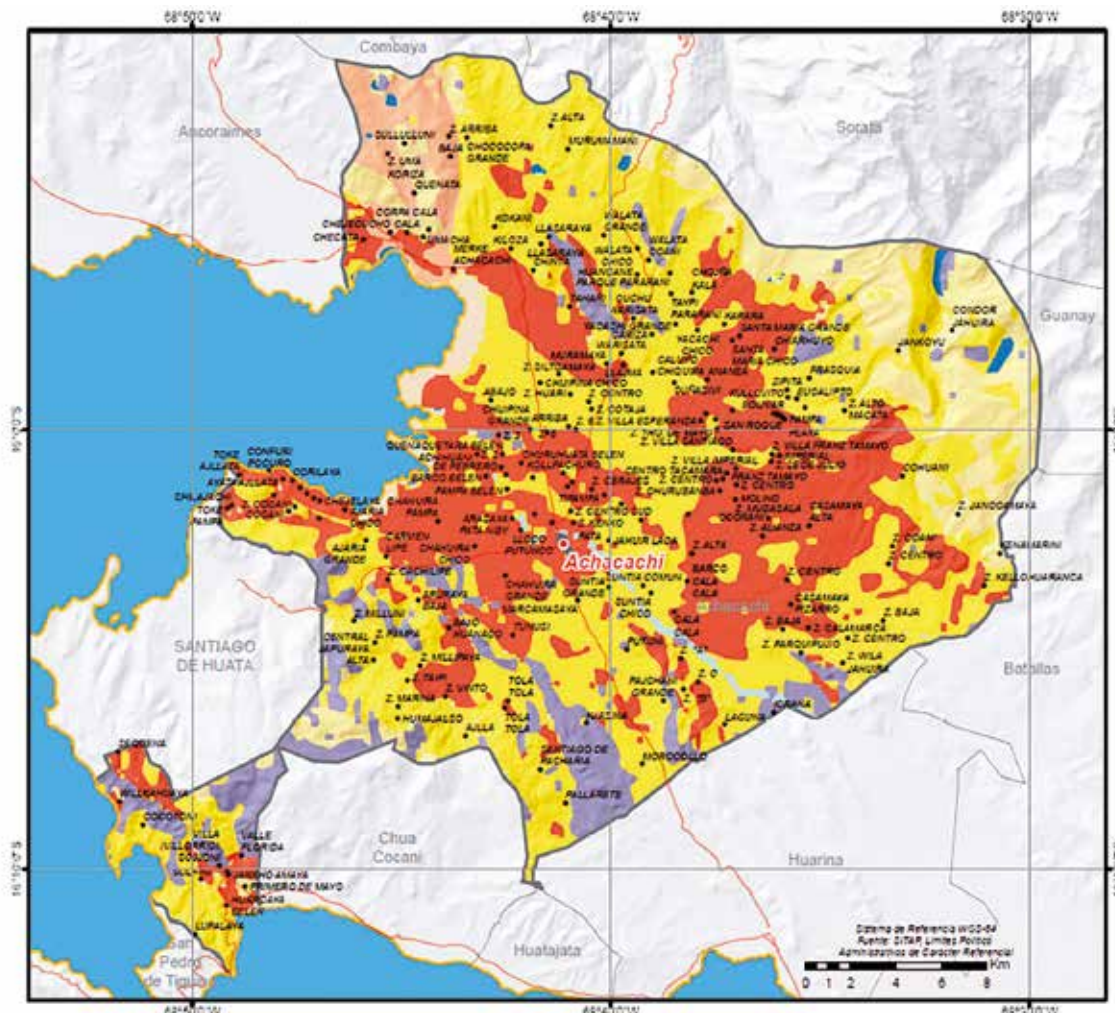
Para la estimación de la varianza de los estimadores, la técnica de los grupos aleatorios sugiere de una población estudio, trabajar formando grupos seleccionados en forma aleatoria y mediante una notación y procedimiento aleatorio realizar la medición de las varianzas (Miras Julio, 1976).

El objetivo general consiste en diseñar una propuesta de actividades didácticas que permita promover un uso de algunos tipos de muestreos, se plantea una aplicación de utilizar los métodos mencionados para mostrar el procedimiento de minimizar la varianza de estimadores en un universo de estudio y explicar la medición de error de muestreo.

Para conocer algunos parámetros poblacionales del sector agropecuario, se diseño una encuesta por muestreo en el municipio de Achacachi del departamento de La Paz, determinándose un tamaño de muestra y aplicar el procedimiento de estimación de la variable “Superficie cultivada de papa”. En el diseño muestral, la técnica plateada fue el muestreo por conglomerados, las unidades de primera etapa fueron los segmentos censales (Conglomerados) y unidades de segunda etapa las Unidades de Producción Agropecuaria (UPAs). El procedimiento y desarrollo de lo indicado se lo presenta a continuación.

Medición del error de muestreo utilizando técnicas de conglomerados y grupos aleatorios en universos agropecuarios

Mapa N° 1
DEPARTAMENTO: La Paz -PROVINCIA: Omasuyos MUNICIPIO: Achacachi



Fuente: Mapa de Referencia estadística, armada en base a Atlas de Potenciales Productivas ., SITAP –UDAPRO –Bolivia

2. MÉTODO DE MUESTREO POR CONGLOMERADOS.-

$N= 498$ Segmentos en el municipio de Achacachi.

$n = 21$ Segmentos muestra.

$M=19.317$ Unidades de Producción Agropecuarias (UPAs) municipio Achacachi.

$VivM$: Numero de viviendas Marco Muestral Agropecuario.

Mi : Número de Unidades de Producciones Agropecuarias (Listado).

$UPAs MU$: Número de Unidades de Producciones Agropecuarias (muestra planificada).

mi : Numero de Unidades de Producciones Agropecuarias (muestra ejecutada).

Cuadro N° 1
Estimación de la Media de la variable “Superficie cultivada de papa” a nivel de unidades de producción agropecuarias

<i>Segto i</i>	<i>VivM</i>	<i>Mi</i>	<i>UPAs MU</i>	m_i	$y_i = \sum_{j=1}^{m_i} y_{ij}$	$\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij}$	$M_i \bar{y}_i$	$(M_i \bar{y}_i - \bar{M} \bar{y})^2$
125	35	25	i	4	0,930	0,2325	5,8125	1,9176
126	14	45	8	4	0,280	0,0725	3,2625	1,3577
127	36	38	8	11	2,770	0,2518	9,5690	26,4329
128	20	32	8	3	0,570	0,1900	6,0800	2,7300
129	39	28	8	7	0,597	0,8520	2,3879	4,1607
130	52	51	8	10	3,267	0,3267	16,6617	149,6707
131	41	36	8	7	1,000	0,1428	5,1428	0,5113
132	72	72	8	6	0,147	0,0245	1,7640	7,0953
134	27	26	8	7	0,100	0,01428	0,3714	16,4535
135	75	73	8	10	0,209	0,0209	1,5257	8,4216
136	60	57	8	4	0,065	0,01625	0,92625	12,2601
137	57	57	8	7	0,209	0,02585	1,7018	7,4305
138	110	105	8	5	0,105	0,0210	2,2050	4,9403
141	70	68	8	6	0,448	0,0746	5,0773	0,4219
142	40	38	8	16	0,205	0,0128	0,4868	15,5306
143	45	45	8	6	0,136	0,0226	1,0199	11,6131
144	26	26	8	13	0,403	0,0310	0,8060	13,1167
145	50	50	8	8	0,360	0,0450	2,2500	4,7423
146	31	31	8	9	0,705	0,0783	2,4283	3,9976
147	35	33	8	9	0,955	0,1061	3,5013	0,8582
148	48	46	8	6	2,610	0,4350	20,0100	242,8080
Total							92,99015	536,4706

Fuente: Elaboración Propia

$\hat{\bar{Y}} = \bar{y} = \left(\frac{N}{Mn} \right) \sum_{i=1}^n M_i \bar{y}_i = \left(\frac{498}{19.317(21)} \right) 92,99015 = 0,11415 =$ Estimación de la media de la variable “Superficie cultivada de papa” a nivel de Unidades Agropecuarias Productivas.

$$\bar{M} \bar{y} = 38,789(0,11415) = 4,4277$$

$$s_1^2 = \frac{1}{n-1} \sum_{i=1}^n (M_i \bar{y}_i - \bar{M} \bar{y})^2 = \frac{536,4706}{21-1} = 26,8235 = \text{Varianza entre segmentos.}$$

$$1 - f_1 = 1 - \left(\frac{n}{N} \right) = 1 - \left(\frac{21}{498} \right) = 0,9578 = \text{Fracción de muestreo entre segmentos.}$$

$$1 - f_2 = 1 - \left(\frac{m_i}{M_i} \right) = \text{Fracción de muestreo dentro del segmento i-ésimo.}$$

Medición del error de muestreo utilizando técnicas de conglomerados y grupos aleatorios en universos agropecuarios

$$s_{2i}^2 = \frac{1}{m_i - 1} \left(\sum_{j=1}^{m_i} y_j^2 - \frac{\left(\sum_{j=1}^{m_i} y_j \right)^2}{m_i} \right) = \text{Varianza dentro del segmento i-ésimo.}$$

Cuadro N° 2
Cálculos del estimador y su varianza

Segto i	Mi	mi			$M_i^2(1 - f_{2i}^2)$	s_{2i}^2	$M_i^2(1 - f_{2i}^2) \frac{s_{2i}^2}{m_i}$
125	25	4	0,1600	0,8400	525,00	0,032225	4,2295
126	45	4	0,0890	0,9110	1.844,70	0,007691	3,5470
127	38	11	0,2894	0,7106	1.026,10	0,025030	2,3348
128	32	3	0,0930	0,9070	928,76	0,072300	22,3833
129	28	7	0,2500	0,7500	588,00	0,005038	0,4232
130	51	10	0,1960	0,8040	2.091,20	0,216170	45,2000
131	36	7	0,1940	0,8060	1.044,57	0,006240	0,9311
132	72	6	0,0830	0,9170	4.753,70	0,0000775	0,0614
134	26	7	0,2690	0,7310	494,16	0,00002857	0,002016
135	73	10	0,1369	0,8631	4.599,40	0,0002018	0,0928
136	57	4	0,070	0,9300	3.021,57	0,0001729	0,1306
137	57	7	0,1228	0,8772	2.850,02	0,001233	0,5020
138	105	5	0,047	0,9530	10.506,80	0,0003425	0,7197
141	68	6	0,088	0,9120	42.170,08	0,010740	7,5485
142	38	16	0,4210	0,5790	836,07	0,00005976	0,00312
143	45	6	0,1330	0,8670	1.755,67	0,00003674	0,01075
144	26	13	0,5000	0,5000	338,00	0,001462	0,03801
145	50	8	0,1600	0,8400	2.100,00	0,007078	1,8579
146	31	9	0,2900	0,7100	682,31	0,005075	0,3847
147	33	9	0,2727	0,7273	792,02	0,015980	1,4062
148	46	6	0,1304	0,8696	1.840,07	0,223830	68,6400
Total							160,4465

Fuente: Elaboración Propia

$$\hat{V}(\hat{\mu}) = \hat{V}(\bar{y}) = (1 - f_1) \left(\frac{1}{n \bar{M}^2} \right) s_1^2 + \frac{1}{n N \bar{M}^2} \sum_{i=1}^n M_i^2 (1 - f_{2i}) \left(\frac{s_{2i}^2}{m_i} \right)$$

$$\hat{V}(\bar{y}) = (0,9578) \left(\frac{1}{21(38,7891)^2} \right) 26,8235 + \frac{1}{21(498)(38,7891)^2} 160,4465 =$$

$$\hat{V}(\bar{y}) = 0,0008131 + 0,00001019 = 0,0008232$$

$$\sqrt{\hat{V}(\bar{y})} = \sqrt{0,0008232} = 0,02869 = \text{Error de Muestreo de variable agropecuaria a nivel de UPAs.}$$

$$\text{Límite para el error de estimación} = 1,96 \sqrt{\hat{V}(\bar{y})}$$

$$\bar{y} - z\sqrt{\hat{V}(\bar{y})} = 0,11415 - 1,96(0,02869) = 0,0579 = \text{Límite Inferior del estimador de la media de la variable "Superficie cultivada de papa" a nivel de UPAs en el municipio}$$

$$\bar{y} + z\sqrt{\hat{V}(\bar{y})} = 0,11415 + 1,96(0,02869) = 0,17038 = \text{Límite Superior del estimador de la media de la variable "Superficie cultivada de papa" a nivel de UPAs en el municipio.}$$

$$CV(\bar{y}) = \frac{\sqrt{\hat{V}(\bar{y})}}{\bar{y}} * 100 = \frac{0,02869}{0,11415} * 100 = 25,13 \% = \text{Coeficiente de variación del estimador de la media.}$$

ESTIMACIÓN DEL TOTAL DE LA VARIABLE "SUPERFICIE CULTIVADA DE PAPA" A NIVEL DE UNIDADES DE PRODUCCIÓN AGROPECUARIAS

$$\hat{Y} = M \bar{y} = 19.317(0,11415) = 2205,03 = \text{Estimador del Total de la variable a nivel UPAs en el municipio.}$$

$$V(\hat{Y}) = V(M \bar{y}) = M^2 V(\bar{y}) = (19.317)^2 (0,0008232) = 307.174,18 = \text{Varianza del estimador del Total.}$$

$$\sqrt{\hat{V}(\hat{Y})} = \sqrt{307.174,18} = 554,23 = \text{Error de muestreo del estimador del Total.}$$

$$\hat{Y} - z\sqrt{\hat{V}(\hat{Y})} = 2.205,03 - 1,96(554,23) = 1.118,7 = \text{Límite Inferior del estimador del Total de la variable a nivel de UPAs en el municipio.}$$

$$\hat{Y} + z\sqrt{\hat{V}(\hat{Y})} = 2.205,03 + 1,96(554,23) = 3.291,3 = \text{Límite Superior del estimador del Total de la variable a nivel de UPAs en el municipio.}$$

$$CV(\hat{Y}) = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}} * 100 = \frac{554,23}{2.205,03} * 100 = 25,13 \% = \text{Coeficiente de variación del estimador del Total.}$$

3. ESTIMACIÓN DE LA VARIANZA POR LA TÉCNICA DE GRUPOS ALEATORIOS

ESTIMACIÓN DE LA VARIANZA DEL ESTIMADOR DE LA MEDIA Y EL TOTAL, MEDIANTE LA TÉCNICA DE GRUPOS ALEATORIOS

Para la estimación de la varianza se utilizó la siguiente notación:

N = Número de Unidades de Producción Agropecuarias en el Municipio.

n = Número de Unidades de Producción Agropecuarias en la muestra.

y_i = Valor observado en la i -ésima UPA

k = Número de grupos aleatorios.

Medición del error de muestreo utilizando técnicas de conglomerados y grupos aleatorios en universos agropecuarios

$m = \frac{n}{k}$ = Número de UPAs seleccionadas aleatoriamente.

$\hat{\bar{Y}} = \frac{\sum_{i=1}^n y_i}{n}$ = Promedio del valor observado a nivel de la muestra de tamaño n

$\hat{\bar{Y}}_r = \frac{\sum_{i=1}^m y_i}{m}$ = Promedio del valor observado aleatoriamente en la submuestra m (para valores de $j=1 \dots m$).

$\hat{V}_{G.A.}(\hat{\bar{Y}}) = \frac{\sum_{r=1}^K (\hat{\bar{Y}}_r - \hat{\bar{Y}})^2}{k(k-1)}$ = Varianza estimada del estimador de la media ($r=1, 2, \dots, k$).

$\sqrt{\hat{V}_{G.A.}(\hat{\bar{Y}})} = \frac{\sum_{r=1}^K (\hat{\bar{Y}}_r - \hat{\bar{Y}})^2}{k(k-1)}$ = Error de muestreo del estimador de la media por el método de Grupos Aleatorizados.

$M = \sum_{i=1}^N M_i$ = Número de unidades de Producción Agropecuaria en el municipio de estudio.

$\hat{V}_{G.A.}(\hat{Y}) = \hat{V}(M \hat{\bar{Y}}) = M^2 \hat{V}(\hat{\bar{Y}}) = M^2 \frac{\sum_{r=1}^K (\hat{\bar{Y}}_r - \hat{\bar{Y}})^2}{k(k-1)}$ = Varianza estimada del estimador del Total por el método de Grupos Aleatorizados. ($r=1, 2, \dots, k$).

$\sqrt{\hat{V}_{G.A.}(\hat{Y})}$ = Error de muestreo del estimador del Total por el método de Grupos Aleatorizados.

Estimación de la Varianza del Estimador de la Media de la variable a nivel de unidades de producción agropecuarias.

n = 158 Unidades de Producción Agropecuarias en la muestra.

M = 19.317 Unidades de Producción Agropecuarias (Marco Muestral).

k = 3 grupos aleatorios.

$m = \frac{n}{k} = \frac{158}{3} = 52,66 \approx 53$ = Número de UPAs seleccionadas aleatoriamente

$\bar{y} - z \sqrt{\hat{V}_{G.A.}(\hat{\bar{Y}})} = 0,11415 - 1,96(0,008906) = 0,09669$ = Límite Inferior del estimador de la media de la variable a nivel de UPAs en el municipio.

$\bar{y} + z \sqrt{\hat{V}_{G.A.}(\hat{\bar{Y}})} = 0,11415 + 1,96(0,008906) = 0,131605$ = Límite Superior del estimador de la media de la variable a nivel de UPAs en el municipio.

$CV(\bar{y}) = \frac{\sqrt{\hat{V}_{G.A.}(\hat{\bar{Y}})}}{\bar{y}} * 100 = \frac{0,008906}{0,11415} * 100 = 7,80 \%$ = Coeficiente de variación del estimador de la media.

Cuadro N° 3
Estimación de la varianza por grupos aleatorios

		GRUPOS ALEATORIOS			RESULTADOS
		1	2	3	
	$\sum_{i=1}^m y_i$	5,44	4,244	5,189	
	m	53	53	52	158
	$\hat{\bar{Y}}_r$	0,102641	0,08007	0,09978	
	$\hat{\bar{Y}}$				0,101778
	$(\hat{\bar{Y}}_r - \hat{\bar{Y}})^2$	0,000000744	0,0004712	0,000003992	0,0004759
	$K(k-1)$				3(3-1)
Media	$\hat{V}_{G.A.}(\hat{\bar{Y}})$				0,00007932
	$\sqrt{\hat{V}_{G.A.}(\hat{\bar{Y}})}$				0,008906
Total	M				19.317
	$\hat{V}_{G.A.}(\hat{Y})$				29.597,97
	$\sqrt{\hat{V}_{G.A.}(\hat{Y})}$				172,04

Fuente: Elaboración Propia

Estimación de la Varianza del Estimador del Total de la variable a nivel de unidades de producción agropecuarias.

$\hat{Y} = M \bar{y} = 19.317(0,11415) = 2205,03$ = Estimador del Total de la variable a nivel UPAs en el municipio.

$\hat{V}_{G.A.}(\hat{Y}) = \hat{V}(M \bar{y}) = M^2 \hat{V}(\bar{y}) = (19.317)^2 (0,00007932) = 29.597,97$ = Varianza del estimador del Total.

$\sqrt{\hat{V}_{G.A.}(\hat{Y})} = \sqrt{29.597,97} = 172,04$ = Error de muestreo del estimador del Total.

$\hat{Y} - z \sqrt{\hat{V}_{G.A.}(\hat{Y})} = 2.205,03 - 1,96(172,04) = 1.867,83$ = Límite Inferior del estimador del Total de la variable a nivel de UPAs en el municipio.

$\hat{Y} + z \sqrt{\hat{V}_{G.A.}(\hat{Y})} = 2.205,03 + 1,96(172,04) = 2.542,22$ = Límite Superior del estimador del Total de la variable a nivel de UPAs en el municipio.

$CV(\hat{Y}) = \frac{\sqrt{\hat{V}_{G.A.}(\hat{Y})}}{\hat{Y}} * 100 = \frac{172,04}{2.205,03} * 100 = 7,80\%$ = Coeficiente de variación del estimador del Total.

Medición del error de muestreo utilizando técnicas de conglomerados y grupos aleatorios en universos agropecuarios

La precisión del estimador de ambos métodos se puede ver en el siguiente cuadro:

Cuadro N° 4
Error de muestreo de la media

	Muestreo por Conglomerados	Técnica de Grupos Aleatorios
Estimador	0,11415	
Error de muestreo	0,02869	0,008906
Límite Inferior (N.C. 95%)	0,0579	0,09669
Límite Superior (N.C. 95%)	0,17038	0,131605
Coefficiente de Variación (%)	25,13	7,80

Fuente: Elaboración Propia

Cuadro N° 5
Error de muestreo del Total

	Muestreo por Conglomerados	Técnica de Grupos Aleatorios
Estimador	2.205	
Error de muestreo	554,23	172,04
Límite Inferior (N.C. 95%)	1.118,7	1.867,83
Límite Superior (N.C. 95%)	3.291,3	2.542,22
Coefficiente de Variación (%)	25,13	7,80

Fuente: Elaboración Propia

4. CONCLUSIONES.-

El muestreo por conglomerados, propone un tratamiento de la información que se puede acomodar al sector agropecuario, donde se trabaja para validar la homogeneidad entre conglomerados y permite la comparación de varianzas entre conglomerados.

Esta técnica de muestreo, facilita al investigador que pueda asignar sus recursos limitados a los pocos conglomerados o áreas seleccionadas aleatoriamente cuando se usan muestras por conglomerados.

El muestreo por conglomerados, muestra un proceso controlable que permite conocer los estimadores de la media y el total como también sus varianzas.

En el proceso de cálculo se puede ver la heterogeneidad dentro del grupo o conglomerado que es fundamental para un buen diseño del muestreo por conglomerados, que muestra que los elementos dentro de cada grupo debe ser tan heterogéneos como la población objetivo.

La técnica de Grupos Aleatorios que estima la varianza de estimadores, muestra una minimización de las varianzas, siendo un buen medidor del error de muestreo.

La utilización de diferentes técnicas de muestreo que minimizan la varianza de los estimadores son buenas herramientas para resolver problemas de estimación en el sector agropecuario.

BIBLIOGRAFÍA

Cochran William G.,(1998), “Técnicas de muestreo”, Decima Cuarta Edición;, JOHN WILEY & SONS, México.

Lohr Sharon L.,(1999), “Muestreo: Diseño y Análisis”, Duxbury Press, USA.

Woodruff, R. S., (1971) “A simple method for approximating the variance of a complicated estimate”. Journal of the American Statistical Association 66:411-414.

Binder D.A., (1983) “On the variances of asymptotically normal estimators from complex surveys International Statistical” Review 51:279-292.

Miras Amor Julio, (1976), “Estimación de errores de muestreo”, INE España.