

Confiabilidad de las estimaciones

El cálculo de los indicadores de calidad de los resultados de la muestra de la Encuesta de Arroz, Maíz y Frijol de Bejuco de abril de 2018 (año agrícola 2017/2018) fueron obtenidos, mediante la inferencia estadística que permite obtener una estimación del verdadero valor de la población.

La muestra tiene una participación significativa tanto en la superficie sembrada como en la cosecha de los rubros investigados, destacándose la contribución de los estratos intencionales de productores (as) grandes. Estas consideraciones incidieron positivamente en los resultados de la encuesta y por ende, en los niveles de tolerancia que contienen las estimaciones.

Las técnicas de muestreo se aplican a las fincas pequeñas y a las fincas grandes contempladas en la investigación. Y en adición, las empresas y organizaciones comunales son investigadas al 100%.

Los errores de muestreo se construyen a partir del uso intensivo del paquete estadístico SPSS que ha sido diseñado para el procesamiento de censos y encuestas, al tiempo que ofrece una amplia gama de opciones para la generación de estadísticas descriptivas y generación de indicadores de confiabilidad, se presentan a nivel nacional, considerando las variables básicas siguientes:

ESTIMACIÓN DE ARROZ, DESVIACIÓN ESTÁNDAR, COEFICIENTE DE VARIACIÓN E INTERVALO DE CONFIANZA EN LA REPÚBLICA, SEGÚN VARIABLE DE ESTUDIO: ENCUESTA AGRÍCOLA DE ABRIL DE 2018

Variable de estudio	Estimación de arroz	Desviación estándar	Coeficiente de variación %	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	93,204.38	4,568.97	5.0	75,652.23	110,756.53
Superficie perdida (ha)	6,266.77	386.47	6.0	4,782.10	7,751.44
Superficie cosechada (ha)	86,937.61	4,459.53	5.0	69,805.86	104,069.36
Cosecha (qq)	6,975,470.90	471,559.80	7.0	5,163,926.77	8,787,015.03

Las estimaciones no incluyen los datos de las comarcas Emberá y Kuna Yala.

ESTIMACIÓN DE MAÍZ, DESVIACIÓN ESTÁNDAR, COEFICIENTE DE VARIACIÓN E INTERVALO DE CONFIANZA EN LA REPÚBLICA, SEGÚN VARIABLE DE ESTUDIO: ENCUESTA AGRÍCOLA DE ABRIL DE 2018

Variable de estudio	Estimación de maíz	Desviación estándar	Coeficiente de variación %	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	53,229.18	2,998.75	6.0	41,709.17	64,749.19
Superficie perdida (ha)	5,069.56	349.1	7.0	3,728.46	6,410.66
Superficie cosechada (ha)	48,159.62	2,948.05	6.0	36,834.40	59,484.84
Maíz nuevo (qq)	77,848,787.00	9,591,776.94	12.0	41,001,016.72	114,696,557.28
Maíz seco (qq)	2,405,011.00	294,643.10	12.0	1,273,110.05	3,536,911.95
Cosecha (qq)	2,560,708.60	295,615.33	12.0	1,425,072.76	3,696,344.44

Las estimaciones no incluyen los datos de las comarcas Emberá y Kuna Yala.

**ESTIMACIÓN DE FRIJOL DE BEJUCO, DESVIACIÓN ESTÁNDAR, COEFICIENTE DE VARIACIÓN E INTERVALO DE CONFIANZA EN LA REPÚBLICA, SEGÚN VARIABLE DE ESTUDIO:
ENCUESTA AGRÍCOLA DE ABRIL DE 2018**

Variable de estudio	Estimación de frijol	Desviación estándar	Coeficiente de variación %	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	12,143.18	1,214.13	10.0	7,478.97	16,807.39
Superficie perdida (ha)	1,646.24	174.97	11.0	974.09	2,318.39
Superficie cosechada (ha)	10,496.94	1,195.98	11.0	5902.46	15,091.42
Cosecha (qq)	130,666.80	17,737.50	14.0	62,526.42	198,807.18

Las estimaciones no incluyen los datos de las comarcas Emberá y Kuna Yala.

A. Estimación puntual:

Es poco probable que las estadísticas que se obtienen mediante una muestra científica de una población en estudio (sean estos totales, promedios, porcentajes, etc.) difieran de las estadísticas que se obtendrían a través de la investigación de todos los elementos de la población, bajo las mismas condiciones y circunstancias, es decir, utilizando iguales cuestionarios, idéntico método de empadronamiento, similar calidad de enumeradores, análogos instructivos, etc. A estas últimas estadísticas se le denominan valores poblacionales. Interesa conocer entonces, cuán cerca está el valor que se obtiene, a través de una muestra (estimación muestral) del valor poblacional correspondiente. La estimación de un parámetro puede tener por resultado un solo punto (estimación puntual).

La diferencia entre una estimación muestral y el promedio para todas las posibles muestras, se conoce como error de muestreo, se expresa en medidas denominadas “error estándar” y “coeficiente de variación”. La primera es el error en términos absolutos y la segunda es el error en términos relativos, es decir, en porcentaje.

El error estándar y el coeficiente de variación de una estimación sirven para medir la variación entre las estimaciones del conjunto de muestras posibles y, por tanto, es una medida de la precisión con la cual una estimación de la muestra investigada, se aproxima al resultado promedio (valor verdadero) de todas las muestras posibles.

B. Estimación por intervalo:

La estimación puntual y la estimación del error estándar permiten construir estimaciones por intervalo, con un nivel de confianza especificado, medido en términos de probabilidad de que el intervalo incluya el resultado promedio de todas las muestras posibles, es decir, el valor verdadero que se desea estimar. Las probabilidades de obtener el valor verdadero entre el límite inferior y el superior del intervalo, considerando un total estimado, se obtiene mediante el siguiente modelo probabilístico:

<u>Intervalo</u>	<u>Nivel de confianza</u>
$\bar{X}' \pm 1 \sigma_{\bar{X}'}$	68.27%
$\bar{X}' \pm 1.96 \sigma_{\bar{X}'}$	95.45%
$\bar{X}' \pm 2.58 \sigma_{\bar{X}'}$	99.73%

Un intervalo dentro del cual exista cierta probabilidad de encontrarlo (estimación por intervalos).

C. Resultados:

Los resultados obtenidos se presentan en las tablas adjuntas a la publicación, conteniendo los indicadores siguientes:

- Estimación de la variable considerada
- Error estándar
- Coeficiente de variación
- Estimación por intervalo al 95% de confianza

Las estimaciones proporcionadas en las tablas permiten evaluar el nivel de fiabilidad de las mismas; se debe tener presente, que a mayor desagregación de las cifras, mayor error de muestreo resultará y por tanto, menor será la confianza que merezca el dato publicado de la encuesta.

Por otra parte, cuando la frecuencia de la variable en el universo es pequeña, su representatividad en la muestra disminuye, resultando con errores de muestreo altos, siendo la única posibilidad de reducción, la obtención de una muestra más grande en función de la característica, lo que incidiría en los costos de investigación.

Ejemplo: El cuadro de la estimación de arroz a nivel de República, nos proporciona la información siguiente:

1. Estimación de la superficie sembrada de arroz es de 93,204.38 ha
2. La desviación estándar es de 4,568.97 ha
3. El coeficiente de variación es 5.0%
4. El intervalo de confianza al 95% (75,652.93, 110,756.53) ha

En términos de probabilidad, se puede asegurar en la estimación por intervalo, que existe un 95% de confianza y que el verdadero valor poblacional de la superficie sembrada de arroz se encuentra entre 75,652.93 ha y 110,756.53 ha.

La estimación del valor verdadero contiene un nivel de error por arriba o por abajo, en términos de error estándar de 4,568.97 ha; siendo su error relativo de tan solo 5.0 %.

Es evidente que corresponde al usuario determinar si una estimación con cierto nivel de error de muestreo le es útil o no, para la toma de decisiones, de acuerdo con el grado de fiabilidad que precisa para ello.