

Confiabilidad de las estimaciones

El cálculo de los indicadores de calidad de los resultados de la muestra aplicada a la Encuesta agrícola de arroz, maíz y frijol de bejuco de febrero 2016 (año agrícola 2015/2016), fueron obtenidos mediante el paquete estadístico especializado SPSS19.0, que utiliza principalmente métodos de linearización, para la estimación de la varianza del total; correspondiente al muestreo aleatorio estratificado con afinación óptima.

Los errores de muestreo 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 FEBRERO DE 2016

Variable de estudio	Estimación de arroz	Desviación estándar	Coeficiente de variación (%)	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	92,380.2	4,730.6	5.1	83,103.9	101,654.9
Superficie pérdida (ha)	11,462.6	900.9	7.9	9,695.7	13,228.8
Superficie cosechada (ha)	80,917.7	4,530.5	5.6	72,033.9	89,800.2
Cosecha (qq)	6,216,485.5	471,017.7	7.6	5,292,938.2	7,140,021.6

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 FEBRERO DE 2016

Variable de estudio	Estimación de maíz	Desviación estándar	Coeficiente de variación (%)	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	65,913.6	4,372.0	6.6	57,340.7	74,485.4
Superficie pérdida (ha)	14,167.0	1,336.1	9.4	11,546.9	16,786.4
Superficie cosechada (ha)	51,746.6	3,977.7	7.7	43,947.2	59,545.6
Maíz nuevo (mazorca)	62,975,560.0	10,572,968.3	16.8	42,244,711.2	83,706,326.5
Maíz seco (qq)	2,135,549.0	324,789.2	15.2	1,498,715.4	2,772,367.5
Cosecha (qq)	2,261,499.8	325,539.1	14.4	1,623,196.0	2,899,788.9

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

ESTIMACIÓN DE FRIJOL, DESVIACIÓN ESTÁNDAR, COEFICIENTE DE VARIACIÓN E INTERVALO DE CONFIANZA EN LA REPÚBLICA, SEGÚN VARIABLE DE ESTUDIO:
ENCUESTA AGRÍCOLA DE FEBRERO DE 2016

Variable de Estudio	Estimación de frijol	Desviación estándar	Coeficiente de variación (%)	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Superficie sembrada (ha)	12,508.1	1,624.1	12.9	9,323.8	15,692.5
Superficie pérdida (ha)	2,775.2	261.8	9.4	2,261.9	3,288.5
Superficie cosechada (ha)	9,732.9	1,586.2	16.3	6,622.8	12,843.1
Cosecha (qq)	110,374.4	19,804.7	17.9	71,539.1	149,202.7

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

A. Estimación puntual:

Una estimación puntual del valor de un parámetro poblacional desconocido (como puede ser la media μ , o la desviación estándar σ), es un número que se utiliza para aproximar el verdadero valor de dicho parámetro poblacional. A fin de realizar tal estimación, tomaremos una muestra de la población y calcularemos el parámetro de la muestra asociado (x para la media, s para la desviación estándar, etc.). El valor de este parámetro de la muestra será la estimación puntual del parámetro poblacional.

Consiste en la estimación del valor del parámetro mediante un solo valor, obtenido de una fórmula determinada. Por ejemplo, si se pretende estimar la talla media de un determinado grupo de individuos, puede extraerse una muestra y ofrecer como estimación puntual la talla media de los individuos.

Interesa conocer entonces, cuán cerca está el valor que se obtiene mediante una muestra (estimación de muestreo) del valor poblacional correspondiente.

La teoría del muestreo provee los procedimientos adecuados para medir en qué magnitud la estimación de muestreo se aproxima al valor poblacional. La muestra utilizada en la encuesta representa una del conjunto de muestras posibles del mismo tamaño, que podrían haberse seleccionado utilizando el mismo diseño de muestreo. Las estimaciones derivadas de las distintas muestras serían diferentes las unas de las otras. La diferencia entre una estimación de muestreo 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 da el error en términos absolutos y la segunda 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:

Intervalo	Nivel de confianza
$X' \pm 1 \sigma_{x'}$	68.3%
$X' \pm 1.96 \sigma_{x'}$	95.5%
$X' \pm 2.58 \sigma_{x'}$	99.7%

C. Resultados:

Los resultados obtenidos se presentan en las tablas adjuntas a la publicación, conteniendo las mismas, 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 92,380.2 ha.
2. La desviación estándar es de 4,730.6 ha.
3. El coeficiente de variación es 5.1%.
4. El intervalo de confianza al 95% (83,103.9, 101,654.9) 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 83,103.9 ha. y 101,654.9 ha.

La estimación del valor verdadero contiene un nivel de error por arriba o por abajo, en términos de error estándar de 4,730.6 ha.; siendo su error relativo de tan solo 5.1 por ciento.

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