

**CONFIABILIDAD DE LAS ESTIMACIONES  
ENCUESTA DE CAFÉ**

ESTIMACIÓN DE CAFÉ, DESVIACIÓN ESTÁNDAR, COEFICIENTE DE VARIACIÓN E INTERVALO DE CONFIANZA EN LA REPÚBLICA, SEGÚN VARIABLE DE ESTUDIO, ENCUESTA DE CAFÉ: ABRIL 2008					
Variable de estudio	Estimación de café	Desviación estándar	Coeficiente de variación (%)	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Árboles de café.....	40,010,889.0	1,106,445.479	3.0	37,842,256	42,179,522
Plantas que producen.....	31,135,377.0	965,320.2276	3.1	29,243,349	33,027,405
Plantas que no producen.....	8,875,512.0	444,102.8447	5.0	8,005,070	9,745,954
Cosecha (quintales pilados).....	204,753.6	6,569.802477	3.2	191,877	217,630

El cálculo de los indicadores de calidad de los resultados de la muestra de la Encuesta Café de abril de 2008, fueron obtenidos mediante el modelo matemático de la varianza de la estimación total del ítem deseado; correspondiente al muestreo estratificado con afijación óptima.

Los errores de muestreo se presentan a nivel nacional, considerando las variables básicas siguientes:

- Total de árboles
- Plantas que producen
- Plantas que no producen
- Cosecha

**A. Estimación Puntual**

Es muy 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 mediante la investigación de todos los elementos de la población bajo las mismas condiciones y circunstancias; es decir, utilizando los mismos cuestionarios, el mismo método de empadronamiento, la misma calidad de enumeradores, los mismos instructivos, entre otros. A estas últimas estadísticas se le denominan valores poblacionales. Interesa conocer entonces, cuán cerca está el valor que se obtiene mediante una muestra (estimación muestral) del valor poblacional correspondiente.

La teoría del muestreo provee los procedimientos adecuados, para medir en qué magnitud la estimación muestral 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 muestral. Las estimaciones derivadas de las distintas muestras serían diferentes las unas de las otras.

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 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 límite superior del intervalo, considerando un total estimado, se obtiene mediante el siguiente modelo:

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

## 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 por ciento de confianza

Las estimaciones proporcionadas en las tablas permiten evaluar el nivel de fiabilidad de las mismas; lo cual 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 árboles de café a nivel de República nos proporciona la información siguiente:

1. Estimación de árboles de café es de 40,010.889 árboles.
2. La desviación estándar es de 1,106.445.48 árboles.
3. El coeficiente de variación es de 2.8 por ciento
4. El intervalo de confianza al 95 por ciento de la superficie sembrada, oscila entre 37,842,255.86 y 42,179,522.14 árboles de café.

En términos de probabilidad, se puede asegurar en la estimación por intervalo, que existe un 95 por ciento de confianza y que el verdadero valor poblacional de los árboles de café se encuentra entre 37,842,255.86 y 42,179,522.14 árboles.

La estimación del valor verdadero contiene un nivel de error por arriba o por abajo, en términos de error estándar de 1,106.445.48 árboles; siendo su error relativo de tan sólo 2.8 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 la toma de decisiones, de acuerdo con el grado de fiabilidad que precisa para ello.