

Departamento Administrativo Nacional de Estadística



**Dirección de Regulación, Planeación,
Estandarización y Normalización
-DIRPEN-**

**Metodología del Diseño Muestral
Encuesta de Calidad de la Gestión
Estatad para el Desarrollo Empresarial
-ECDE-**

Junio 2008



TABLA DE CONTENIDO

1. ESTIMACIÓN DE LOS PARÁMETROS LA VARIANZA DEL ESTIMADOR PARA DOMINIOS DE ESTUDIO	2
1.1 ESTIMACIÓN POR DOMINIOS DE ESTUDIO.....	2
1.2 EL ESTIMADOR DEL TOTAL Y LA VARIANZA PARA EL TOTAL EN UN DISEÑO ESTMAS	2
1.3 EL ESTIMADOR DE LA RAZÓN VARIANZA PARA UNA RAZÓN EN UN DISEÑO ESTAS	3
1.4 LA VARIANZA DEL ESTIMADOR DE RAZÓN	4
2. COEFICIENTE DE VARIACIÓN	4

1. ESTIMACIÓN DE LOS PARÁMETROS LA VARIANZA DEL ESTIMADOR PARA DOMINIOS DE ESTUDIO

1.1 Estimación por Dominios de Estudio

Un Dominio de Estudio es una subpoblación para la cual se requieren estimaciones puntuales separadas con buena precisión y con intervalos de confianza útiles.

Sea la variable Z_{dk} definida como

$$Z_{dk} = \begin{cases} 1 & \text{si } k \in U_d \\ 0 & \text{si } k \notin U_d \end{cases}$$

donde

k = La empresa dependiendo de lo que se este estimando

U_d = Dominio d

Luego

$$\sum_U z_{dk} = N_d$$

N_d = Cantidad de elementos en el universo que pertenecen al dominio d , la cual es desconocida.

Ahora sea la variable

$$y_{dk} = x_{dk} * z_{dk}$$

x_k = Es la variable cuantitativa de interés y observada en la muestra

1.2 El Estimador Del Total Y La Varianza Para El Total En Un Diseño ESTMAS

t_d = total de la variable x en el dominio d

Parámetro

$$t_d = \sum_U x_{dk}$$

Estimador

$$\begin{aligned}\hat{t}_d &= \sum_{h=1}^H N_h \bar{y}_{dh} \\ &= \sum_{h=1}^H \sum_{m_h} F_h y_{hk}\end{aligned}$$

Con

$$\bar{y}_{dh} = \sum_{m_h} y_{hk} / n_h$$

Nota: Los subíndices de las sumas se refieren a:

U = Universo muestral

m_h = Las unidades seleccionadas en la muestra que pertenecen al estrato h

y_{hdk} = Al valor de la variable y para el elemento k en el estrato h en el dominio d

\bar{y}_{dm_h} = Al promedio de la variable y en el estrato h para el dominio d

n_{dh} = Al tamaño de la muestra para el estrato h en el dominio d

F_h = El factor de expansión para los elementos del estrato h

donde

$$\hat{V}_{ESTMAS}(\hat{t}_\pi) = \sum_h \frac{N_I(N_I - n_I)}{n_I} S_{\hat{t}_{S_I}}^2$$

y

$$S_{\hat{t}_{S_I}}^2 = \frac{1}{n_I - 1} \sum_{S_I} (\hat{t}_{i\pi} - \bar{\hat{t}}_{i\pi})^2$$

1.3 El estimador de la razón varianza para una razón en un diseño ESTAS es

Un estimador de razón para un dominio es de la forma

$$R = \frac{t_{dy}}{t_{dx}}$$

donde

t_{dy} Es el total de la variable y en el dominio d

t_{dx} Es el total de la variable x en el dominio d

Y el estimador es

$$\hat{R} = \frac{\hat{t}_{dy}}{\hat{t}_{dx}}$$

\hat{t}_{dy} Es el total estimado de la variable y en el dominio d

\hat{t}_{dx} Es el total estimado de la variable x en el dominio d

1.4 La Varianza Del Estimador De Razón

Dado que los estimadores de razón son funciones no lineales de totales, su estimador es sesgado aunque con un sesgo relativo no muy grande, la varianza aproximada del estimador de razón, se calcula con la misma fórmula que para el total solamente que en lugar de utilizar la variable y se genera una nueva variable U_k para cada uno de los elementos en la muestra de la siguiente forma

$$U_k = \frac{(y_k - R \cdot z_k)^2}{\hat{t}_{hz}^2}$$

donde

y_k es el valor de la variable y (numerador) k (registro del archivo)

z_k es el valor de la variable z (denominador) k (registro del archivo)

R es la razón estimada como se calculó en el numeral anterior

\hat{t}_{hz}^2 es el total estimado de la variable z en el estrato h (denominador) elevada al cuadrado

2. COEFICIENTE DE VARIACIÓN

Cuando la magnitud de la variabilidad es muy grande los parámetros estimados pierden utilidad pues el valor verdadero del parámetro, en el universo, puede estar en un intervalo muy amplio lo cual no proporciona información útil. Por lo tanto esta medida de variabilidad es un criterio muy importante para determinar la calidad de una estimación. Pero el valor de esta magnitud esta directamente relacionada con la unidad de medida de la variable de interés, por ejemplo si se esta estimando el total de los ingresos para una actividad económica dada la unidad de medida en miles de pesos, y como además la varianza está en unidades al cuadrado, se estaría hablando de miles de pesos al cuadrado. La raíz cuadrada de esta varianza es la que se denomina *Desviación estándar de la distribución o error estándar*. Esta medida de dispersión tiene la ventaja de que la unidad de la medida de dispersión corresponde a la unidad de la variable de interés, así por ejemplo, una variabilidad de cien en un millón en la estimación del total es pequeña lo cual indicaría que la estimación tiene muy buena precisión, pero si el parámetro a estimar es el promedio de personas ocupadas por empresas, tener una variabilidad de cinco

personas es alta, y aunque es un valor muchísimo menor puede significar que no es tan preciso.

Un criterio más uniforme para determinar la precisión de un estimador es el *coeficiente de variación* el cual se define como la variación porcentual del error estándar a la estimación central, es decir es el cociente entre el error estándar del estimador y el estimador multiplicado por 100.

$$CV = \frac{\sqrt{V(\hat{\theta})}}{E(\hat{\theta})} * 100$$

Y el coeficiente de variación estimado esta dado por

$$c\hat{v} = \frac{\sqrt{\hat{V}(\hat{\theta})}}{(\hat{\theta})} * 100$$

Aunque la varianza, el error estándar y el coeficiente de variación, miden la magnitud de la variabilidad de la distribución muestral del estimador, es decir, lo que algunos denominan el *error de muestreo*, el coeficiente de variación tiene la ventaja de dar esta medida en términos porcentuales, por lo cual es una medida común para estimaciones.

Se suele considerar que el resultado de una estimación es bueno si su coeficiente de variación es menor del 5 %; aceptablemente práctico, entre el 5 % y el 10%; de baja precisión si es mayor del 10 %. Y menor del 15% y no útil si es mayor del 15%.