MÉTODO DE ESTIMACIÓN CLÁSICO

Se describe el método clásico de estimación de forma general, presentando las especificidades del método usadas por dos de los equipos de trabajo, el de la Mtra. Patricia Romero Mares y el de la Dra. Guillermina Eslava Gómez.

El diseño utilizado en el conteo rápido 2012 es un *Muestreo estratificado simple de casillas con 483 estratos*. El estimador puntual corresponde al conocido como *estimador de razón combinado*. Las expresiones en su forma general aparecen en Särndal et al. (1992). A continuación presentamos las expresiones específicas del estimador y la de su varianza estimada.

Sea N_h el número de casillas en el estrato h y n_h el número de casillas en muestra, $h=1,\ldots,H$. Donde H denota el número total de estratos.

Sea Y_{hi} el número de votos emitidos a favor de un partido en la casilla i del estrato h, además, sea X_{hi} el número de votos totales emitidos en la casilla i del estrato h, $i=1,\ldots n_h,;\ h=1,\ldots,H$. El estimador de la proporción de votos nacional para ese partido se estima como sigue.

$$\hat{P} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_h^H \hat{Y}_h}{\sum_h^H \hat{X}_h} = \frac{\sum_h^H \frac{N_h}{n_h} \sum_i^{n_h} Y_{hi}}{\sum_h^H \frac{N_h}{n_h} \sum_i^{n_h} X_{hi}}$$

Para determinar los errores de estimación que se pueden cometer, es necesario calcular la varianza estimada del estimador anterior. Ésta se obtiene con las expresiones siguientes.

Primero se calcula la varianza estimada de una nueva variable al interior de cada estrato. Esta es:

$$\hat{V}_h\,(G_{hi}) = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left[\frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}} - \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}} \right]^2$$
$$= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(G_{hi} - \overline{G}_h \right)^2,$$

donde $G_{hi} = \frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}}$; con ésta, el estimador de la varianza total del estimador de la proporción es la siguiente.

$$\hat{V}(\hat{P}) = \sum_{h=1}^{H} N_h^2 \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left[\hat{V}_h(G_{hi}) \right]$$

La precisión estimada, con una confianza del (1-α)100% es:

$$delta = z_{1-\alpha/2} \sqrt{\sum_{h=1}^{H} N_h^2 \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left[\hat{V}_h(G_{hi}) \right]}$$

donde $z_{1-\alpha/2}$ es el cuantíl de la función de densidad Gaussiana estándar.

Por ejemplo, la precisión estimada, con una confianza del 99.9% es:

$$delta = 3 \sqrt{\sum_{h=1}^{H} N_h^2 \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left[\hat{V}_h(G_{hi}) \right]}$$

El intervalo de confianza es el estimador del la proporción menos y más la delta.

A medida que los resultados de las casillas en muestra van arribando, las estimaciones se realizan colapsando estratos, para cubrir aquellos que aún no tienen muestra, y a medida que la muestra recibida va aumentando, los estratos colapsados se van separando hasta llegar a su máximo número original.

El cálculo de las varianzas se hace de forma adicional, usando el método de remuestreo de Jackknife, ver e.g. Wolter (2007) o Fuller (2009).

El cálculo de las estimaciones se realizó usando principalmente el programa R.

Referencias

Fuller, W.A. (2009). *Sampling Statistics.* John Wiley & Sons.

Särndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer-Verlag, New York.

Wolter, K.M. (2007). *Introduction to Variance Estimation, 2nd ed*. Springer-Verlag, New York.