6. Muestreo por Conglomerados

6.1 Necesidad y ventajas del muestreo por conglomerados.

6.2 Formación de los conglomerados. Conglomerados y estratos. Notación

$N =$ conglomerados en la población.

$n =$ conglomerados en la muestra.

$m_i =$ elementos en el conglomerado i

$y_i =$ suma de las observaciones en el conglomerado i

$M = \sum_{i=1}^{N} m_i =$ elementos en la población

$m = \sum_{i=1}^{n} m_i =$ elementos en la muestra

$\overline{M} = \frac{1}{N}\sum_{i=1}^{N} m_i =$ tamaño medio de los conglomerados de la población.

$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m_i =$ tamaño medio de los conglomerados de la muestra.

6. Muestreo por Conglomerados

6.3 Estimación de la media, proporción y total poblacionales.

$$\hat{\mu} = \overline{y} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i} \qquad \hat{V}(\overline{y}) = \frac{1}{\overline{M}^2}\frac{N-n}{N}\frac{S_c^2}{n} \qquad S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{y}m_i\right)^2$$

$$\hat{\tau} = M\overline{y} \qquad \hat{V}(\hat{\tau}) = M^2\hat{V}(\overline{y}) = N(N-n)\frac{S_c^2}{n}$$

$$\hat{\tau}_t = N\overline{y}_t \qquad \hat{V}(\hat{\tau}_t) = N^2\hat{V}(\overline{y}_t) = N(N-n)\frac{S_t^2}{n}$$

$$\overline{y}_t = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad \hat{V}(\overline{y}_t) = \frac{N-n}{N}\frac{S_t^2}{n} \qquad S_t^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{y}_t\right)^2$$

$$Si \quad m_1 = m_2 = \ldots = m_N \quad \Rightarrow \quad M\overline{y} = N\overline{y}_t$$

6. Muestreo por Conglomerados

6.4 Determinación del tamaño muestral.

$$\hat{\tau} = M\bar{y} \qquad \hat{\mu} = \bar{y} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i}$$

$$n = \frac{N\sigma_c^2}{ND + \sigma_c^2} \qquad \hat{\sigma}_c^2 = S_c^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}m_i\right)^2 \qquad D = \frac{B^2\overline{M}^2}{4} \quad (media) \qquad D = \frac{B^2}{4N^2} \quad (total)$$

$$\hat{\tau}_t = N\bar{y}_t \qquad \bar{y}_t = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$n = \frac{N\sigma_t^2}{ND + \sigma_t^2} \qquad \hat{\sigma}_t^2 = S_t^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}_t\right)^2 \qquad D = \frac{B^2}{4N^2}.$$

6. Muestreo por Conglomerados

6.5 Muestreo por Conglomerados con Probabilidades Proporcionales al tamaño

$$\pi_i = \frac{m_i}{M}$$

$$\boxed{\begin{aligned} \hat{\tau}_{cppt} &= \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\pi_i} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\left(m_i/M\right)} = M\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{m_i} = M\frac{1}{n}\sum_{i=1}^{n}\bar{y}_i \\ \hat{V}\left(\hat{\tau}_{cppt}\right) &= M^2\hat{V}\left(\hat{\mu}_{cppt}\right) = \frac{M^2}{n}\frac{\sum_{i=1}^{n}\left(\bar{y}_i - \hat{\mu}_{cppt}\right)^2}{n-1} = \frac{M^2 S_{cppt}^2}{n} \end{aligned}}$$

$$\boxed{\hat{\mu}_{cppt} = \frac{1}{M}\hat{\tau}_{cppt} = \frac{1}{n}\sum_{i=1}^{n}\bar{y}_i \qquad \hat{V}\left(\hat{\mu}_{cppt}\right) = \frac{1}{n}\frac{\sum_{i=1}^{n}\left(\bar{y}_i - \hat{\mu}_{cppt}\right)^2}{n-1} = \frac{S_{cppt}^2}{n}}$$