

4.2 Teoría del muestreo aleatorio estratificado.

Sea Ω una población finita de N elementos y $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ una partición de Ω . A cada $\Omega_i = \{y_{i1}, y_{i2}, \dots, y_{iN_i}\}$ de tamaño N_i , con $i = 1, \dots, m$, se le llama **estrato de la población**.

Luego, para cada $i = 1, \dots, m$, sea S_i una muestra aleatoria obtenida del estrato Ω_i de tamaño n_i . Los estadísticos de la población son los siguientes:

$$t_i = \sum_{j=1}^{N_i} y_{ij} \quad \text{es el total de la población en el estrato } \Omega_i \text{ para cada } i = 1, \dots, m$$

$$t = \sum_{i=1}^m t_i \quad \text{es el total de la población total}$$

$$\overline{y_{Ui}} = \frac{t_i}{N_i} = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij} \quad \text{es la media de la población en el estrato } \Omega_i \text{ para cada } i = 1, \dots, m$$

$$\overline{y_U} = \frac{t}{N} = \frac{1}{N} \sum_{i=1}^m t_i = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{N_i} y_{ij} \quad \text{es la media de la población total}$$

$$S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (y_{ij} - \overline{y_{Ui}})^2 \quad \text{es la varianza de la población en el estrato } \Omega_i \text{ para cada } i = 1, \dots, m$$

Como en cada estrato se hace un muestreo aleatorio simple, se tiene que

1. $\overline{y_i} = \frac{1}{n_i} \sum_{j \in S_i} y_{ij}$ es estimador insesgado de $\overline{y_{Ui}}$ para cada $i = 1, \dots, m$.
2. $\hat{t}_i = N_i \overline{y_i} = \sum_{j \in S_i} y_{ij}$ es estimador insesgado de t_i para cada $i = 1, \dots, m$.
3. $\widehat{S_i^2} = \frac{1}{n_i - 1} \sum_{j \in S_i} (y_{ij} - \overline{y_i})^2$ es estimador insesgado de S_i^2 para cada $i = 1, \dots, m$.

Proposición.

1. $\hat{t} = \sum_{i=1}^m \hat{t}_i = \sum_{i=1}^m N_i \overline{y_i}$ es estimador insesgado de t .
2. $\overline{y} = \frac{\hat{t}}{N} = \frac{1}{N} \sum_{i=1}^m t_i$ es estimador insesgado de $\overline{y_U}$.
3. $V(\overline{y}) = \sum_{i=1}^m \left(1 - \frac{n_i}{N_i}\right) \left(\frac{N_i}{N}\right)^2 \frac{\widehat{S_i^2}}{n_i}$.

Demostración.

1. Se tiene que

$$E [\widehat{t}] = E \left[\sum_{i=1}^m \widehat{t}_i \right] = \sum_{i=1}^m E [\widehat{t}_i] = \sum_{i=1}^m t_i = t$$

Por lo tanto, \widehat{t} es un estimador insesgado de t .

2. Por el inciso anterior, se obtiene que

$$E [\bar{y}] = E \left[\frac{\widehat{t}}{N} \right] = \frac{1}{N} E [\widehat{t}] = \frac{t}{N} = \bar{y}_U$$

Por lo tanto, \bar{y} es un estimador insesgado de \bar{y}_U .

3. Se da que

$$\begin{aligned} V(\bar{y}) &= V \left(\frac{1}{N} \sum_{i=1}^m t_i \right) \\ &= \frac{1}{N^2} \sum_{i=1}^m V(t_i) \\ &= \frac{1}{N^2} \sum_{i=1}^m N_i^2 \left(1 - \frac{n_i}{N_i} \right) \frac{\widehat{S}_i^2}{n} \end{aligned}$$