Diseño BE

Daniel Czarnievicz

2017

Estrategia de selección

En el diseño BE primero se elije π tal que $0 \leq \pi \leq 1$. Para cada elemento del marco muestral se sortea ε_k de una distribución $Unif(0;1)$. Si $\varepsilon_k < \pi$, el k-ésimo elemento es incluido en la muestra. Por lo tanto, el diseño muestral es el siguiente:

$$p(s) = P(S = s) = \underbrace{\pi \dots \pi}_{n_S \text{ veces}} \underbrace{(1-\pi)\dots(1-\pi)}_{N-n_S \text{ veces}} = \pi^{n_S}(1-\pi)^{N-n_S}$$

Probabilidades de inclusión

Dada la estrategia de selección, $I_k \sim Ber(\pi) \;\; \forall k \in U$. Por lo tanto:

$$\star \;\; \pi_k = P(k \in s) = \pi \;\; \forall k \in U$$

Dado que los ε_k se sortean de forma independiente, los eventos $\{k \text{ es seleccionado}\}$ y $\{l \text{ es seleccionado}\}$ son independientes, por lo tanto:

$$\star \;\; \pi_{kl} = P(k; l \in s) = E_{BE}(I_k I_l) = \pi^2 \;\; \forall k \neq l \in U$$

$$\star \;\; \Delta_{kl} = COV_{BE}(I_k; I_l) = \begin{cases} \pi(1-\pi) & \text{si} \quad k = l \\ 0 & \text{si} \quad k \neq l \end{cases}$$

El estimador $\hat{t}_\pi$

$$\star \;\; \hat{t}_\pi = \textstyle\sum_s y_k^{\checkmark} = \sum_s \dfrac{y_k}{\pi} \Rightarrow \boxed{\hat{t}_\pi = \dfrac{1}{\pi} \textstyle\sum_s y_k}$$

$$\star \;\; E_{BE}(\hat{t}_\pi) = E_{BE}\left(\frac{1}{\pi}\textstyle\sum_s y_k\right) = \dfrac{1}{\pi} E_{BE}\left(\textstyle\sum_s y_k\right) = \dfrac{1}{\pi}\textstyle\sum_U E_{BE}(I_k) y_k = \sum_U y_k = t_y$$

$$\star \;\; V_{BE}\left(\hat{t}_\pi\right) = \textstyle\sum\sum_U \Delta_{kl}\, y_k^{\checkmark}\, y_l^{\checkmark} = \sum\sum_U \pi(1-\pi)\left(\dfrac{y_k}{\pi}\right)^2 = \dfrac{1-\pi}{\pi}\textstyle\sum_U y_k^2 \Rightarrow$$

$$\Rightarrow \boxed{V_{BE}\left(\hat{t}_\pi\right) = \frac{1-\pi}{\pi}\left[(N-1)S^2_{y_U} + N\,\bar{y}^2_U\right]}$$

Un estimador insesgado para la varianza viene dado por:

$$\star \;\; \hat{V}_{BE}\left(\hat{t}_\pi\right) = \frac{1-\pi}{\pi^2}\left[(n-1)S^2_{y_s} + n\bar{y}^2_s\right]$$

Estimador $\hat{t}_{alt}$

Supongamos que: $E_{BE}\big(n_S\big) = N\,\pi = n$, entonces:

$$\star\ \hat{t}_\pi = \frac{N}{n}\sum\nolimits_s y_k$$

$$\star\ V_{BE}\big(\hat{t}_\pi\big) = N^2\left(\frac{1}{n}-\frac{1}{N}\right)S^2_{y_U}\left[1-\frac{1}{N}+CV^2_{y_U}\right]$$

El estimador $\hat{t}_{alt}$ estará dado por:

$$\star\ \hat{t}_{alt} = N\sum\nolimits_s \frac{y_k}{n_S} = N\bar{y}_s = \frac{n}{n_S}\sum\nolimits_s y_k = \frac{N\,\pi}{n_S}\,\frac{1}{\pi}\sum\nolimits_s y_k = \frac{n}{n_S}\hat{t}_\pi = \left(\frac{N}{{}^{n_S}/_{\pi}}\right)\hat{t}_\pi = \frac{N}{\hat{N}}\,\hat{t}_\pi$$

$$\text{donde } \hat{N} = \sum\nolimits_s \frac{1}{\pi} = \frac{N}{n}\sum\nolimits_s 1 = N\left(\frac{n_S}{n}\right)$$

$$\star\ V_{BE}(\hat{t}_{alt}) \doteq \frac{N}{\pi}(1-\pi)S^2_{y_U} = \frac{N^2}{n}(1-f)S^2_{y_U} \quad \text{si } \pi = {}^{n}/_{N}$$

$$\star\ \frac{V_{BE}(\hat{t}_\pi)}{V_{BE}(\hat{t}_{alt})} \doteq 1+\frac{1}{CV^2_{y_U}}$$

Efecto diseño

Para calcular el Deff se debe considerar un muestreo tal que $n_S = E(n_S) = N\pi$ para asegurar una comparación justa. Luego también es conveniente escribir la varianza del estimador como:

$$\star\ V_{BE}\Big(\hat{t}_\pi\Big) = \frac{1-\pi}{\pi}NS^2_{y_U}\left[1-\frac{1}{N}+\frac{1}{CV^2_{y_U}}\right] \text{ donde } CV_{y_U} = \frac{S_{y_U}}{\bar{y}_U}$$

Por lo tanto, teniendo en cuenta que $\pi \overset{\text{BE}}{=} {}^{n}/_{N} \overset{\text{SI}}{=} f$,

$$Deff(BE;\hat{t}_\pi) = \frac{V_{BE}(\hat{t}_\pi)}{V_{SI}(\hat{t}_\pi)} = \frac{\frac{1-\pi}{\pi}N\,S^2_{y_U}\left[1-\frac{1}{N}+\frac{1}{CV^2_{y_U}}\right]}{\frac{N^2}{n}(1-f)S^2_{y_U}} = 1-\frac{1}{N}+\frac{1}{CV^2_{y_U}} \doteq 1+\frac{1}{CV^2_{y_U}} > 1$$

$$\star\ \frac{V_{BE}(\hat{t}_{alt})}{V_{SI}(\hat{t}_\pi)} \doteq 1$$

Tamaño muestral

$$n_S \sim Binomial(N;\pi) \Rightarrow E_{BE}(n_S) = N\pi;\;\; V_{BE}(n_S) = N\pi(1-\pi)$$

$$IC^{n_S}_{95\,\%} = \left[N\pi \pm z_{(1-{}^{\alpha}/_{2})}\sqrt{N\pi(1-\pi)}\right]$$

Para determinar n, llamemos $n = E(n_S)$ dado que el tamaño es aleatorio. Tomemos $\pi = {}^{n}/_{N}$, luego entonces:

$$\star\ \varepsilon^2 \doteq z^2_{1-{}^{\alpha}/_{2}}V_{BE}(\hat{t}_y) = z^2_{1-{}^{\alpha}/_{2}}\left(\frac{1}{\pi}-1\right)\sum\nolimits_U y_k^2 \Rightarrow \boxed{n = \frac{z^2_{1-{}^{\alpha}/_{2}}N\sum_U y_k^2}{\varepsilon^2 + z^2_{1-{}^{\alpha}/_{2}}\sum_U y_k^2}}$$

Teniendo en cuenta que: $\sum_U y_k^2 = (N-1)S_{y_U}^2 + N\bar{y}_U^2 = \left(1 - \dfrac{1}{N} + \dfrac{1}{cv^2}\right) N S_{y_U}^2$, podemos escribir:

$$\boxed{n = \frac{z_{1-\alpha/2}^2 N^2 k S_{y_U}^2}{\varepsilon^2 + z_{1-\alpha/2}^2 N k S_{y_U}^2} \ \text{ donde } \ k = 1 - \frac{1}{N} + \frac{1}{cv^2}}$$