

Muestreo Estadístico: π -Estimador^{*}

García Prado, Sergio
sergio@garciparedes.me

25 de septiembre de 2017

1. Demostración de insesgadez

Para una población finita U formada por los individuos $\{1, \dots, i, \dots, N\}$ cuya variable de interés se denota por y , siendo y_k el valor que toma dicha variable en el individuo k . Se define el total poblacional como sigue:

$$t = \sum_{k \in U} y_k \quad (1)$$

Sea s una muestra extraída de la población U (por tanto $s \subset U$), se define la variable aleatoria I_k como se indica a continuación:

$$I_k = \begin{cases} 1 & \text{si } k \in s \\ 0 & \text{si } k \notin s \end{cases} \quad (2)$$

De esta manera, se puede utilizar el valor π_k para modelizar la probabilidad de que I_k tome el valor 1, es decir, de que el elemento k -ésimo esté en la muestra s . Por tanto, se puede afirmar que $I_k \sim B(\pi_k)$, una distribución de *Bernoulli* de parámetro π_k . A partir de dicha distribución se tiene que $E[I_k] = \pi_k$. Además, $\pi_k > 0 \forall k \in U$ para que se cumpla la restricción de muestreo probabilístico.

Se define el π -estimador como:

$$\hat{t}_\pi = \sum_{k \in s} \frac{y_k}{\pi_k} = \sum_{k \in U} I_k \frac{y_k}{\pi_k} \quad (3)$$

La demostración acerca de la insesgadez de este estimador deriva de la esperanza de la variable I_k , que toma el valor π_k , tal y como se indicaba anteriormente, lo cual produce que se anule dicho valor π_k con el π_k del divisor del estimador. Además, la restricción de muestreo probabilístico $\pi_k > 0 \forall k \in U$ hace que el valor \hat{t}_π exista siempre (no hay división entre 0). Por estas razones se puede decir que el estimador es insesgado. De manera matemática esto se describe a continuación:

$$E[t] = \sum_{k \in U} y_k \quad (4)$$

$$E[\hat{t}_\pi] = \sum_{k \in U} E[I_k] \frac{y_k}{\pi_k} = \sum_{k \in U} \pi_k \frac{y_k}{\pi_k} = \sum_{k \in U} y_k \quad (5)$$

$$\text{bias}(t, \hat{t}_\pi) = |E[t] - E[\hat{t}_\pi]| = \left| \sum_{k \in U} y_k - \sum_{k \in U} y_k \right| = 0 \quad (6)$$

^{*}URL: <https://github.com/garciparedes/statistical-sampling-pi-estimator>

Nótese por tanto, que el funcionamiento del π -estimador se basa en la “expansión” del conjunto de individuos seleccionados en la muestra s , de tal manera que se siga cumpliendo de la forma más precisa posible la distribución de probabilidades de la población U .

2. Demostración del intervalo de confianza para la varianza

[TODO]

Referencias

- [SSW03] Carl-Erik Särndal, Bengt Swensson, and Jan Wretman. *Model assisted survey sampling*. Springer Science & Business Media, 2003.
- [TG18] Jesús Alberto Tapia García. Muestreo Estadístico 1, 2017/18.