

Título:

Selección y estimación en Ranked Set Sampling utilizando R

Autores:

Beatriz Cobo, Antonio Arcos, María del Mar Rueda

Email:

beacr@ugr.es, arcos@ugr.es, mrueda@ugr.es

Organización:

Departamento de Estadística e Investigación Operativa. Facultad de Ciencias. Universidad de Granada

Resumen:

Un paso importante para poder realizar cualquier inferencia estadística es que los datos se obtengan a través de un mecanismo adecuado, de este modo podremos sacar conclusiones válidas sobre la variable de interés.

Uno de los mecanismos más comunes para obtener datos es el muestreo aleatorio simple. Además de este, los datos pueden ser recogidos mediante muestreo estratificado, muestreo por conglomerados,..., con cualquiera de estos enfoques, una vez que se han elegido los elementos de muestra, se recogen las medidas deseadas de cada uno de los elementos seleccionados.

En este trabajo nos vamos a centrar en Ranked Set Sampling (RSS), conocido en español como muestreo por conjuntos ordenados. Este muestreo fue propuesto por primera vez por McIntyre (1952) para situaciones en las que es difícil tomar las mediciones reales para las observaciones de la muestra, por ejemplo, por ser costosas, destructivas, que consumen mucho tiempo,..., y se lleva a cabo de la siguiente forma:

En el caso de un RSS equilibrado, se selecciona una muestra de m^2 unidades de una población U mediante muestreo aleatorio simple con reemplazamiento (SRSWR). Se asignan estas m^2 unidades al azar en m conjuntos cada uno de tamaño m . Se ordenan todas las unidades en cada uno de los conjuntos con respecto a los valores de la variable de interés, mediante un método de clasificación, como la inspección ocular,... y se selecciona la primera unidad del conjunto 1, la segunda unidad del conjunto 2,..., y la última unidad del conjunto m . Finalmente se miden con precisión usando un instrumento adecuado, para obtener el valor real de la variable de interés. Esto completa un ciclo del muestreo. Este proceso se repite para r ciclos, para obtener el tamaño de muestra deseado $n = m * r$.

Para los procedimientos de estimación hay muchos programas para trabajar con encuestas complejas, pero hay pocos que hayan implementado módulos para trabajar con RSS.

Existe un paquete, denominado NSM3, que calcula el RSS, pero solo considera la opción de RSS equilibrado, sin contemplar el caso de RSS desequilibrado, es decir, cuando n no es múltiplo de m . Actualmente se está desarrollando por parte de los autores un paquete de R que incluye cómo obtener muestras mediante muestreo equilibrado y desequilibrado, así como el cálculo de sus estimaciones y varianzas.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research* **2** 385-390.

Palabras clave:

Ranked set sampling, Muestreo por conjuntos ordenados, Muestreo aleatorio simple, Muestreo complejo