



Estadística española : revista del Instituto Nacional de Estadística, ISSN 0014-1151
Número 193 (mayo-agosto 2017)

On Comparison Of Horvitz-Thompson And Murthy's Sampling Strategies For Estimating Sensitive Finite Population Totals Under Scrambled Randomized Response Plans

P. 5-27

Samindranath Sengupta

Abstract

We consider the problem of unbiased estimation of a finite population total related to a sensitive quantitative variable under two scrambled randomized response plans and compare the relative efficiency of the unequal probability sampling strategies due to Horvitz-Thompson (1952) and Murthy (1957) under a super-population model depending on a parameter g . It is shown that for the linear plan the model expected variance is smaller for Murthy's (1957) strategy if $g \leq 1$, while for the multiplicative plan the model expected variance is smaller for the Horvitz-Thompson (1952) strategy if $g \geq 2$. We also address the problem of unbiased estimation of the variances of these two sampling strategies under the two randomized response plans and study the non-negative property of the variance estimators.

Resumen

Consideramos que el problema de la estimación no sesgada de un total finito de población está relacionado a una variable cuantitativa sensible bajo dos planes aleatorizados de respuesta aleatoria y comparar la eficiencia relativa de las estrategias desiguales de muestreo probabilístico debido a Horvitz-Thompson (1952) y Murthy (1957) bajo una superpoblación modelo dependiendo de un parámetro g . Se muestra que para el plan lineal, el modelo la varianza esperada es menor para la estrategia de Murthy (1957) si $g \leq 1$, mientras que para el plan multiplicativo, la varianza esperada del modelo es menor para los Horvitz-Estrategia de Thompson (1952) si $g \geq 2$. También abordamos el problema de la estimación de las varianzas de estas dos estrategias de muestreo bajo los dos planes de respuesta aleatoria y estudiar la propiedad no negativa de la varianza estimadores.

Posdem: frame and reliability

P. 29-35

Gonzalo Sánchez-Crespo Benítez

Abstract

Try to reduce the sampling error for a specific survey population frame with the stratification and the sample size done is possible at any time. The sampling methods are tested over the frame with the purpose to make the sample more representative. POSDEM allows testing 28 different sampling methods over any population frame. We seek increased sample heterogeneity. Software is free and more sampling procedures can be added. It can be used in surveys, censuses, quality control or to explore frames into big data. Sampling methods have a strong relationship with the survey population frame. Partly by this, there are many sampling methods that are not frequently applied. On the other hand we know that several alternatives of systematic sampling and unequal probabilities sampling are frequently used. However, both may be disturbed by the structure of the investigated population. Trends, cycles and inappropriate relationship between variables are present more often that it seems. In both cases, to find the best sampling method, it is necessary to study the relationship between the sampling plans and the population frame before applying these methods. For different sample sizes, software POSDEM tests this relationship between frames and sampling methods. That test might be done just over the real population or over a super population model inspired on this population. We

find interesting results regarding stability properties procedures.

Resumen

Siempre es posible reducir el error de muestreo para un determinado marco poblacional sin variar el tamaño de muestra o la estratificación utilizada. Para conseguir que la muestra sea más representativa podemos estudiar los diferentes métodos de muestreo disponibles en relación con el marco utilizado. POSDEM permite contrastar 28 métodos sobre cualquier marco poblacional. Buscamos aumentar la heterogeneidad en la muestra. El programa está disponible en la web y puede ser utilizado en encuestas, censos, estudios de control de calidad o para explorar por muestreo marcos dentro de grandes poblaciones. Los métodos de muestreo presentan una relación estrecha con su marco poblacional. Para diferentes tamaños de muestra, el software POSDEM, contrasta estas relaciones entre marcos y métodos de muestreo. Este contraste puede realizarse sobre el marco real o sobre el marco suprapoblacional basado en un modelo de dicho marco. De esta forma se ponen de manifiesto resultados relevantes acerca de la estabilidad de los procedimientos utilizados. Observamos que se utilizan con frecuencia distintas alternativas de muestreo sistemático y muestreo con probabilidades desiguales. Sin embargo ambos pueden ser distorsionados por la estructura de la población investigada. Tendencias, ciclos o inapropiadas relaciones entre variables se presentan con mayor frecuencia de lo aconsejable. En todos los casos para elegir el mejor método de muestreo es necesario estudiar la relación entre el plan de muestreo y el marco poblacional antes de aplicar estos métodos.

Estudio para la estimación de la varianza del cambio neto anual en la Encuesta de Población Activa

P. 37-44

Montserrat Herrador Cansado, Juana Porras Puga, Carlos Pérez Arriero, Juan Vicente Jiménez Llorente

Resumen

La varianza del estimador del cambio neto anual se calcula aplicando métodos indirectos como las semimuestras reiteradas, el jackknife y el bootstrap. Cada uno de ellos está adaptado al diseño muestral de la Encuesta de Población Activa y tiene en consideración el solapamiento de las muestras entre los trimestres de dos años consecutivos. Además y en el caso del método bootstrap, se han usado diferentes metodologías que van desde la más sencilla, que no tiene en cuenta ni el solapamiento ni la calibración, hasta la más compleja con ambos aspectos incorporados. Finalmente, se comparan los errores de muestreo obtenidos por las diferentes técnicas utilizadas.

Abstract

The variance of the annual net change estimators is estimated using indirect methods such as repeated half-samples, jackknife and bootstrap. Each of them is adapted to the sample design of the Spanish Labour Force Survey and takes into account the overlap of the samples between the quarters of two consecutive years. In addition, in the case of the bootstrap method, different methodologies have been used, ranging from the simplest, without overlapping or calibration, to the most complex with both aspects incorporated. Finally, we compare the sampling errors obtained by the different techniques.

Tail weight measures for distributions. A new tail weight coefficient

P. 45-57

Juan Fco. Ortega

Abstract

Specific information on the structure of a distribution is contained in its extreme values, which is measured using the concept of the tail weight. In this paper we analyze the concept of tail of a distribution and propose a new definition for constructing the tail set. Using this set, we define a measure to quantify the tail weight in different distributions, which depends on the distribution and on a cut-off value. Fixing the cut-off value, we define a new tail weight coefficient, where this depends only on the distribution used. From all these elements a sample version is proposed, its main properties are studied and its values are calculated in certain distributions.

Resumen

Los valores extremos de una distribución contienen información importante sobre ella, la cual es cuantificada mediante las medidas de peso de colas. En este trabajo se analiza el concepto de cola de una distribución y se propone cómo determinar sus elementos. Usando este conjunto, se propone una medida para cuantificar su peso en diferentes distribuciones, la cual dependería de una cierta cota. Finalmente, fijando dicha cota se propone un nuevo coeficiente de peso de colas que depende sólo de la distribución en estudio. De todos estos elementos se propone una versión para muestras, se estudian sus principales propiedades y se calculan sus valores en ciertas distribuciones.
