

Probabilistic & nonprobabilistic sampling using auxiliary information

A seminar by Ismael Ramón Sanchèz Borrego

University of Granada – Spain

Friday 1 July 2022 | 12:30 p.m.
Room Benveuti and live Zoom
Department of Statistical Sciences

Auxiliary information is often used in a probabilistic sampling setting to increase the efficiency of estimators of population parameters of interest. There is a list of estimation methods that have been used in the past contemplating such information, like regression, calibration, etc. Nonparametric regression methods have been proven very useful in this area of research outperforming other available estimators.

There is a growing interest in surveying populations that are hidden or hard-to-reach, such as injection-drug users, VIH-infected people, LGBTI community, etc. They usually represent a small size of the overall population, making it difficult to access them and typically obtaining lower response rates. Therefore, they are difficult to reach and using traditional sampling to survey them can become utterly expensive.

There are some non-probabilistic methods available that account for small groups within a population and address some of these issues, like snowball sampling or Respondent-driven sampling (RDS). RDS is a snowball-type sampling method used to survey hidden population, or those lacking a sampling frame. RDS is useful for overcoming limitations of traditional sampling, as it reduces privacy concerns among the participants.

We consider the problem of regression modeling and association for continuous RDS data by proposing a new sample weight method for estimating non-linear parameters such as the covariance and the correlation coefficient.

We also consider the problem of estimating poverty measures for RDS data by proposing estimators of the distribution function and the quantiles.