ADVANCED METHODS IN SAMPLING AND ESTIMATION IN FINITE POPULATIONS

30 August - 1 September 2017, Neuchâtel, Switzerland

About the IASC-ERS summer school

The IASC-ERS Summer school is intended to provide training in special areas of statistics for PhD students, junior researchers and lecturers at universities. Professionals working in industry who are interested in the application of new statistical methods are also welcome to participate. Participants are expected to have good background in statistics at the M.Sc. level although not necessarily related to the subject of the course.

Aims of the summer school

Probabilistic sampling methods from finite populations are used in many fields, from demographics, economics, public health to official statistics, but also in environmetrics or forestry. In this summer course we will start by introducing basic ideas of sampling from finite populations, as well as usual estimation methods. Advanced sampling methods will also be covered, showing recently proposed methods, such as balanced sampling, balanced spatial sampling, etc., as well as advanced estimation methods based on bootstrap in finite populations. The summer course will bring together the theoretical and practical aspects of sampling and estimating in finite populations. Exercises will be provided using the R software. Knowledge of R is not required (an introduction to R will be given), but knowledge of a statistical software is a must. The course is given at the M.Sc. level, and does not require preliminary knowledge of sampling and estimation in finite populations.

Speakers

Guillaume Chauvet, ENSAI - National School for Statistics and Information Analysis, Rennes, France,

Anton Grafström, Swedish University of Agricultural Studies, Umea, Sweden,

Alina Matei, University of Neuchâtel, Switzerland,

Yves Tillé, University of Neuchâtel, Switzerland.

Organizing committee

Alina Matei, University of Neuchâtel, Switzerland

Yves Tillé, University of Neuchâtel, Switzerland

Adminstrative coordinator, Corine Diacon, University of Neuchâtel, Switzerland

Location

Faculty of Science, University of Neuchâtel, UniMail Building, Rue Emile-Argand 11, 2000, Neuchâtel

The summer scholl includes three lectures (in the morning) and three lab sessions using R (in the afternoon).

Lecture 1
Yves Tillé - Survey sampling, main definitions and tools

This first course introduces the main notion, tools and definitions of
the theory of survey sampling. After defining a sampling design, we study the different ways of conducting inference in surveys in the design-based and model-based frameworks. The mains sampling designs are presented including performing algorithms for unequal probability sampling and balanced sampling. The different techniques of estimation based on the theory of calibration are presented in details including the question of variance estimation. The introduction concludes by the establishment of general rules to conceive a sampling strategy composed of a sampling design and an estimator.

Lecture 2
Guillaume Chauvet - Bootstrap for complex sampling designs

Survey data are obtained by means of complex sampling designs, which may vary depending on the populations under study. For example, business surveys are typically performed by means of stratified simple random sampling on a register of businesses, while household surveys and epidemiologic surveys are usually carried through multistage sampling designs. The resulting estimators account for the way the samples are selected, and the measures of accuracy must be defined accordingly. Among the possible ways for computing measures of accuracy, bootstrap enables to produce variance estimators or confidence intervals for parameters via resampling in the original sampled data. The resampling scheme is supposed to mimic the way the data were originally sampled, which is challenging in the context of complex survey designs. In this lecture, we will present some bootstrap methods proposed for survey sampling. We will illustrate them on practical examples and/or simulated data.

Lecture 3
Anton Grafström - Spatial sampling

In spatial sampling a goal has been to select probability samples that are well spread over the population. One of the reasons for that is that for populations with spatial trends in the variables of interest, the estimation can be much improved by selecting samples that are well spread over the population. The concept of spatial balance of a sample is introduced as a measure of spread of a sample, and as a distance between the sample and the sampling distribution. The connection between balanced samples, spatially balanced samples and representative samples will be discussed. Novel sampling methods to select probability samples that are well spread in any auxiliary space will be presented. In particular the local pivotal method, spatially correlated Poisson sampling and the local cube method will be covered. Some examples end exercises will be done using the R package Balanced Sampling.

Alina Matei - Some applications in R (lab sessions)

The software environment R is recognized today as an important statistical computing tool, providing a convenient framework to implement functions related to random sampling and estimation in finite populations. Therefore, applications to some of the topics presented during the

lectures (such as unequal probability sampling designs, balanced sampling, calibration, generalized calibration, etc.) will be illustrated using the R package 'sampling'. Knowledge in R is not required; a short introduction to R will be given. It is however expected from the participants to have an important interest in statistical computing.