

Q2016 Quality Conference course

Theme of the Course/Workshop: Multisource Statistics: Quality and Statistical Methods

Facilitator: Li-Chun Zhang, University of Southampton & Statistics Norway

Programme

Time slot	Topic
09:30 – 09:45	Introduction: Basu's elephants
Morning session	
09:45 – 11:00	Two-phase life-cycle model of integrated statistical data: a total error framework
11:00 – 11:30	Coffee-Break
11:30 – 13:00	Quality indicators and measurement for multisource statistics
13:00 – 14:00	Lunch Break
Afternoon session	
14:00 – 15:30	Categorical data fusion and proxy variables: Techniques and uncertainty
15:30 – 16:00	Coffee-Break
16:00 – 17:00	Optimal adjustment for inconsistency in imputed data

Outline

The production of Official Statistics is undergoing a paradigm shift from census/survey sampling to combination of data from multiple sources, including administrative data and Big Data. Reduction of survey cost and response burden, gains of long-term production efficiency, and potentials for new and detailed spatial-demographic and longitudinal statistics are some of the chief benefits.

The first half of this one-day course provides an introduction to some quality concepts relevant to multisource statistics, divided in two sessions:

- A coherent total error framework for integrated statistical data based on the two-phase life-cycle model that covers the transformation, linkage and integration of multiple datasets
- An overview of quality indicators and measures with an emphasis on statistics produced using administrative data sources

The second half consists of sessions discussing two selected topics of statistical methods:

- The techniques and uncertainties of statistical matching, which aims to bring together separate datasets that do not share common units and cannot be linked at the record level
- Optimal adjustment of either incompatible or implausible measurements in an imputed or linked dataset, in order to produce a numerically consistent and clean dataset

Participant's profile

The presentation of the material is largely non-technical, especially in the first half. However, some experience or familiarity with various administrative (or Big) data and register-based statistics is helpful for a better appreciation of the quality concepts to be introduced. Basic statistical knowledge or understanding is needed to successfully absorb the methods discussed in the second half.

Instructor's profile

Li-Chun Zhang is Professor of Social Statistics at University of Southampton and Senior Researcher at Statistics Norway. He received his Dr. Scient. in Statistics in 1996 at the University of Tromsø, Norway. He joined Statistics Norway in 1997, and University of Southampton in 2012. Li-Chun Zhang has worked and published on a number of subjects in official statistics. These include sampling design and coordination, sample survey estimation, non-response, measurement errors, small area estimation, index number calculations, editing and imputation, register-based statistics, statistical matching, *etc.* He participated in the EU framework projects EURAREA, DACSEIS, RISQ and BLUE-ETS, and the ESSnet projects Small Area Estimation and Data Integration.