Survey Methods Section: Analysis of Complex Sample Survey Data

Steven Heeringa, Institute for Social Research, University of Michigan

In order to extract maximum information at minimum cost, survey designs are typically more complex than simple random samples. Clustered and stratified sample designs are common. But how do you analyze the resulting data - in particular, how do you determine margins of error and make inferences that take into account the complex sample design features? This workshop discusses methods for the analysis of complex sample survey data, including estimation of descriptive parameters, analysis of categorical data, and linear and logistic regression modeling. The workshop is intended for anyone analyzing survey data collected from complex samples, and assumes a background in applied statistical analysis. The course is largely based on selected chapters from the book *Applied Survey Data Analysis* (Chapman and Hall, April 2010) by Steve Heeringa, Brady West, and Pat Berglund, published by Chapman and Hall. Statistical procedures will be illustrated through example analyses that focus on physical and mental health data from major surveys such as the U.S. National Health and Nutrition Examination Survey (NHANES), the National Comorbidity Survey-Replication (NCS-R) and the Health and Retirement Survey (HRS). The workshop will be lecture-based, but participants may bring their own laptop computers with software for the analysis of survey data installed to follow along with the examples.

Part 1: An Overview of Complex Sample Designs and Methods for Inference

-Complex Samples and Design Effects and Weighting in Survey Estimation and Inference
-Sampling Error Calculation Models and Methods
-Inference for Descriptive Statistics: Means, Proportions, Ratios, Percentiles

Part 2: Analytic Methods and Software for Analysis of Survey Data

-Analysis of Categorical Data from Complex Sample Survey Data
-Linear Regression Analysis of Complex Sample Survey Data
-Logistic Regression Analysis of Complex Sample Survey Data

About the Instructor:

Steven G. Heeringa is a Research Scientist in the Survey Methodology Program, the Director of the Statistical and Research Design Group in the Survey Research Center at the Institute for Social Research. He is a member of the faculty of the Michigan Program in Survey Methods and the Joint Program in Survey Methodology at the University of Maryland. Steve is a Fellow of the American Statistical Association and an elected member of the International Statistical Institute. He has over 35 years of statistical sampling experience directing the development of the SRC National Sample design, as well as sample designs for SRC's major longitudinal and cross-sectional survey programs. During this period he has been actively involved in research and publication on sample design methods and procedures such as weighting, variance estimation, and the imputation of missing data that are required in the analysis of sample survey data. He has been a teacher of survey sampling methods to U.S. and international students and has served as a sample design consultant to a wide variety of international research programs based in countries such as Russia, the Ukraine, Uzbekistan, Kazakhstan, India, Nepal, China, Egypt, Iran, and South Africa and Chile. Steve is the lead author of Applied Survey Data Analysis (Chapman and Hall, 2010).

Who Should Attend:

Participants should ideally have a working knowledge of applied statistical analysis, including hypothesis testing, descriptive estimation, confidence interval construction, and linear and generalized linear regression modeling. Background in applied survey sampling is recommended but not required. Example exercises will be illustrated in Stata but no knowledge of specific software is required.