



JOINT PROGRAM IN SURVEY METHODOLOGY

PRACTICAL TOOLS FOR DESIGNING AND WEIGHTING SURVEY SAMPLES

A short course sponsored by the Joint Program in Survey Methodology

Feb 10 – 21, 2025

COURSE OBJECTIVES

The main goal of this short course is to give students tools they can directly use in designing single-stage and multi-stage samples in the real world. This includes deciding on a sample size given a specified budget, creating strata, allocating the sample to the strata given a set of constraints or requirements for detectable differences between group estimates, estimating variance components, and determining what sample sizes to use at different stages in a multi-stage sample.

An important component of the course will be the use of specialized software for calculating sample sizes, computing survey weights, and estimating variances. We emphasize R but will also include some examples in SAS and Stata. Sample size calculations can be done using the R `PracTools` package written by the instructors or with Microsoft Excel, which will be used for mathematical programming. Survey weights can be computed with the R `survey` package for many designs and estimators. The R `survey` package will also compute variances using linearization and replication methods.

Homework problems will be given that cover many of the topics in the course. These will be short but will allow students to test how well they have understood the ideas presented.

WHO SHOULD ATTEND

Professional statisticians, social scientists, and economists who have some familiarity with applied sampling and survey design and want more in-depth knowledge of sample size determination in complex samples, survey weighting, and variance estimation methods.

THE INSTRUCTORS

Jill Dever is a Senior Director at RTI International in Washington, D.C. and received her Ph.D. from the Joint Program in Survey Methodology at the University of Maryland in 2008. She has worked as a survey researcher for more than 30 years in areas such as health, health care, education, and the U.S. military.



JOINT PROGRAM IN SURVEY METHODOLOGY

Her experience includes sampling from address-based lists, creating software for optimizing complex sample designs; constructing linearization and replicate analysis weights using calibration techniques; and analyzing data from complex surveys. Dr. Dever's research foci include variance estimation for calibration estimators with estimated control totals under a complex sample design, and techniques to evaluate the utility of nonprobability sampling. She has co-authored the books, *Practical Tools for Designing and Weighting Survey Samples* with Frauke Kreuter and Richard Valliant and *Survey Weights: A Step-by-Step Guide to Calculation* with Richard Valliant. She is an elected Fellow of the American Statistical Association and a past Associate Editor of the journals, *Journal of Official Statistics* and *Survey Methodology*.

Richard Valliant is a Research Professor Emeritus at the Universities of Michigan and Maryland and an elected Fellow of the American Statistical Association. He received his Ph.D. in Biostatistics at Johns Hopkins University and has over 45 years of experience in sample design and estimation using data from complex surveys. Prior to joining the faculties of the Michigan Program in Survey Methodology and the Joint Program in Survey Methodology at Maryland, he was an Associate Director at Westat and a mathematical statistician with the Bureau of Labor Statistics. He is a past Associate Editor of the *Journal of Official Statistics*, the *Applications and Case Studies* and *Theory and Methods* sections of the *Journal of the American Statistical Association*, and *Survey Methodology*. He delivered the 2022 Morris Hansen Memorial Lecture and was the winner of the 2024 Joseph Waksberg Award.

PREPARATION FOR THE COURSE

Prior to the course, students should install the latest version of R on their computers along with the `PracTools` and `survey` packages. Both are available free at <http://www.r-project.org/>. We will also use RStudio, which can be downloaded at <https://posit.co/download/rstudio-desktop/>. Much of the material will be taken from this text, which students can purchase separately from various booksellers:

Valliant, R., Dever, J.A., & Kreuter, F. (2018). *Practical Tools for Designing and Weighting Survey Samples*, 2nd edition. New York: Springer.



JOINT PROGRAM IN SURVEY METHODOLOGY

Course Topics

1. R and its Packages

Use of R and RStudio; downloading, data structures, calculations

Sample Size Calculations for Single-Stage Designs

2. Single-stage Designs: simple random sampling and stratified simple random sampling

Sample sizes based on target coefficients of variation (CV); means and proportions; power for detecting differences; allocations in stratified sampling

3. Single-stage Designs: probability proportional to size sampling and power calculations

4. Mathematical Programming

Multicriteria optimization; setting up a problem in Excel; allocations in stratified sampling; cost-constrained and CV constrained allocations

Sample Size Calculations for Multi-Stage Designs

5. Sample Size and Allocation for Two-stage Designs

Variances of estimates; variance components; with-replacement sampling of PSUs as an approximation to simplify variance formulae

Cost functions; optimal numbers of 1st and 2nd stage units; measures of homogeneity (*moh*); effects of cluster sizes on *moh*'s; estimating variance components; anticipated variances

Weighting and Estimation

6. Basic Steps in Weighting

Base weights, adjustments for unknown eligibility, nonresponse; cell adjustments; response propensities

7. Weight Calibration

Poststratification, raking, general regression estimators (GREG); weight bounding

8. Variance Estimation

Linearization, jackknife, balanced repeated replication, bootstrap estimators



JOINT PROGRAM IN SURVEY METHODOLOGY

Special Topics

9. Nonprobability sampling Potential problems with nonprobability samples. Approaches to estimation (quasi-randomization, superpopulation modeling, doubly robust). Combining nonprobability and probability samples.
10. Process Control and Quality Measures
Flowcharting; specification writing; documentation

Course Schedule

Live sessions will be Friday Feb 14 and Feb 21, 2025 from 10:00 AM to 11:30 AM EST.

The slides and talk are pre-recorded, spreading it out over 2 weeks. Week 1, Monday – Thursday: The entire Week 1 recordings would be made available Online and participants could watch them at their own pace. Week 1, Friday: A 1.5-hour live Online discussion, Q A with the instructor. Week 2, following the same schedule as Week 1, (recorded lectures M-Th, live Online meeting Friday).