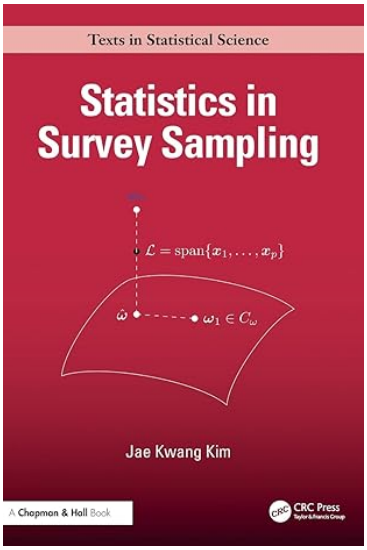
Book Review

Statistics in Survey Sampling

Changbao Wu

University of Waterloo, Canada, cbwu@uwaterloo.ca

Kim, J. K. (2025). *Statistics in Survey Sampling* (1st ed.). Chapman and Hall/CRC, ISBN 9781032997766, DOI: 10.1201/9781003606147, 284 pages.



The title of a book often leads readers to infer its content. This book conveys a clear message through its title, *Statistics in Survey Sampling*, indicating that the focus is on statistical theory and methods used in survey sampling. The author, Dr. Jae Kwang Kim, is a Professor in the Department of Statistics at Iowa State University, a powerhouse in survey sampling research and education with a tradition that spans many decades. Professor Kim is a leading researcher in the field of survey sampling and an innovative theoretical statistician. The content of this book reflects the materials he has used for teaching and some of the research topics he has pursued over the past 25 years. The book is suitable as a textbook or reference for an advanced graduate level course in survey sampling and as a reference for researchers working in the field or on related topics.

The book can be roughly divided into three parts. The first part consists of Chapters 1–7 and covers basic topics in survey design and sampling methods. These include simple and systematic sampling designs described in Chapter 3, stratified sampling in Chapter 4, sampling with unequal probabilities in Chapter 5, single-stage and two-stage cluster sampling in Chapters 6 and 7. The most unique feature of the presentation of this part is the early introduction of the Horvitz-Thompson estimator in Chapter 2.

Chapter 2 starts with the mathematical definition of a probability sample and the probability sampling design through the so-called sampling distribution $P(\cdot)$, which is the probability measure for the sampling design. Design-based expectation, variance, and mean squared error of an estimator, along with first-order and second-order inclusion probabilities, are defined using $P(\cdot)$. Theoretical properties of the Horvitz-Thompson estimator for the population total are then presented for a general sampling design characterized by the first- and second-order inclusion probabilities. Estimation of more general parameters, including those defined through census estimating equations, is briefly discussed. This provides a more rigorous approach to the introduction of survey sampling and is aimed at an advanced level for graduate students in a statistics program.

Chapters 3–7 contain material that goes beyond the basic presentations one often sees in traditional textbooks and may be of interest to students with a strong mathematical and statistical background. Examples include the random sorting method and the reservoir sampling method for implementing simple random sampling (SRS), stratum boundary determination, entropy for sampling designs and

maximum entropy sampling, and balanced sampling.

The second part of the book consists of Chapters 8, 9 and 10 and covers topics on point and variance estimation for the population total and mean. Chapter 8 presents standard material on ratio and regression estimators in the presence of population auxiliary information. Chapter 9 focuses mainly on calibration methods for estimation with survey data and describes several calibration techniques with theoretical details, including the conventional calibration methods, the so-called model-assisted calibration, generalized entropy calibration, and soft calibration. Chapter 10 discusses methods and techniques for variance estimation, including linearization methods and replication methods such as random grouping, the jackknife, and the bootstrap.

The third part of the book covers more specialized topics in Chapters 11–16. Chapter 11 discusses several aspects of two-phase sampling and how it is used to provide auxiliary information for estimation. Chapters 12 and 13 deal with missing survey data, focusing on how to adjust for unit nonresponse and how to impute item nonresponse. Chapter 14 addresses inference for analytic model parameters using survey data, including Bayesian inference. Chapter 15 deals with so-called predictive inference under multi-level models and over time. Chapter 16 includes material on current research topics in the analysis of non-probability survey samples.

Dr. Jae Kwang Kim has worked on many of the topics presented in the second and third parts of the book. He has the insight and technical expertise to introduce the research problems, describe the methodological approaches, and develop theoretical results. These materials are a rich resource for graduate students in statistics and young researchers who are interested in pursuing research on related topics. Each chapter of the book is also followed by a set of practice problems to facilitate teaching and learning, making it easy to use as a textbook.

The book *Statistics in Survey Sampling* contains many references of historical importance to various research topics, some of which remain active today. It is a valuable addition to the existing textbooks and research monographs used for teaching and research in survey sampling, including Cochran (1977), Fuller (2009), Lohr (2022), and Wu and Thompson (2020).

References

Cochran, W. G. (1977) *Sampling Techniques*, 3rd ed., Wiley, New York.

Fuller, W. A. (2009) *Sampling Statistics*. Wiley, Hoboken, NJ.

Lohr, S. L. (2022) *Sampling: Design and Analysis*, 3rd ed., Chapman & Hall/CRC, Boca Raton.

Wu, C., and Thompson, M. E. (2020) *Sampling Theory and Practice*. Springer, Cham.