Book and Software Review

Book review: "Sampling: Design and Analysis, Third Edition" by Sharon L. Lohr

---

Camelia Goga[1]

[1]Laboratoire de Mathématiques de Besançon, Université de Franche-Comté
Besançon, France, camelia.goga@univ-fcomte.fr

Abstract

The book *"Sampling: Design and Analysis, Third Edition"* by Sharon L. Lohr is an outstanding work in the field of survey sampling. This third edition is accompanied by two additional books for readers who wish to implement the survey sampling techniques presented in the book using either R or SAS software. Like the previous two editions, this book is primarily aimed at students and instructors.

Keywords: categorical data analysis, linearization techniques, nonprobability samples, R and SAS companion books, sampling designs.

The book *"Sampling: Design and Analysis, Third Edition"* by Sharon L. Lohr was published by Chapman and Hall/CRS Press in 2022 (Lohr, 2022a). This new edition is structured in 16 chapters, for more than 600 pages, and covers a broad spectrum of survey theory concepts, the whole supported by numerous examples from social sciences, public opinion research, public health, business, agriculture, and ecology. Chapters 1 to 6 are dedicated to the most commonly used probability sampling designs (simple random sampling without replacement, stratified sampling, unequal probability sampling designs, cluster sampling design) to estimate totals and means by using the Horvitz-Thompson, ratio and regression estimators. These chapters also contain sections dedicated to the model-based versus design-based estimation points of view, with discussions of recent research in this field. Chapters 7 to 16 are dedicated to advanced topics in survey sampling theory usually related to complex surveys, such as nonresponse, linearization techniques, categorical data analysis, two-phases sampling designs, rare populations and small area estimation. A chapter on nonprobability samples is included as well in this new edition.

Throughout this book, focus is given on the sampling phase, which is considered by the author to be the most important step in the survey process. The reader is guided in this process by means of many examples on real survey data, inevitably involving graphical analysis of survey data, which can be a real challenge in practice.

As the book is primarily aimed at students and instructors, the book contains more than 550 exercises from which more than 150 exercises are new for this third edition.

As in the previous editions, these exercises are structured in four types intended to meet the objectives of the book: *Introductory*, *Working with survey data*, *Working with theory* and *Project and Activities*. As a teacher of statistics and survey theory myself, I can testify of their utility and great interest.

The novelty of this new edition is the addition of two free downloadable books: *"R Companion for Sampling: Design and Analysis, Third Edition"* by Yan Lu and Sharon L. Lohr (Yan & Lohr, 2022) and *"SAS Software Companion for Sampling: Design and Analysis, Third Edition"* by Sharon L. Lohr (Lohr, 2022b). These two additional books are intended to be a guide for a novice reader in the field of surveys but also in the use of R and SAS software. The companion books have exactly the same structure. They start with an introductory chapter which gives basic instructions to get started with R and SAS (e. g. for R, installing the software and specific packages, reading and saving statistical datasets as well as conducting first statistical analyses and drawing plots). The books then provide respectively R and SAS implementation of the exercises proposed in the book *"Sampling: Design and Analysis, Third Edition"* (Chapters 1 to 11). For the R companion, the R packages `survey` (Lumley, 2020) for computing statistics from complex surveys, `sampling` (Tillé & Matei, 2021) for drawing complex samples and `SDARessources` (Lu & Lohr, 2021) for datasets from the *"Sampling: Design and Analysis, Third Edition"* book are mainly used; some functions of these packages are described in detail. For the SAS part, the focus is on the procedures: `surveyselect`, `surveymeans`, `surveyfreq`, `surveyreg` and `surveylogistic`. The implemented examples are provided with code, annotated output, and helpful tips. Both companions end with a chapter entitled *"Additional Topics for Survey Data Analysis"* containing implementations for some advanced methodology contained in Chapters 12 to 16 from the *"Sampling: Design and Analysis, Third Edition"* book as well as a *"Data Set Descriptions"* in Appendix. The SAS companion contains a specific Appendix, *"Jackknife Macros"*, dedicated to jackknife methods with survey data.

Concluding, the Lohr's book *"Sampling: Design and Analysis, Third Edition"* will be again a reference book in the field of survey sampling as the first two editions. The two companion books for carrying out surveys and estimates with R and SAS software fill a gap in the specialist literature. Students, instructors and anyone wishing to train in survey techniques will find in this book the necessary methodology, as well as many examples of sample surveys on real data and how to implement them using R and SAS software in its two companion books.

References

Lohr, S. L. (2022a). *Sampling: Design and Analysis, Third Edition*. New York: Chapman and Hall/CRC Press.

Lohr, S. L. (2022b). *SAS Companion for Sampling: Design and Analysis, Third Edition*. New York: Chapman and Hall/CRC Press.

Lu, Y. & Lohr, S. L. (2021). SDAResources: Datasets and Functions for 'Sampling: Design and Analysis, 3rd Edition'. R package version 0.1.1, `https://CRAN.R-project.org/package= SDAResources`.

Lumley, T. (2020). *survey: analysis of complex survey samples*. R package version 4.0, `https://CRAN.R-project.org/package=survey`.

Tillé, Y. & Matei, A. (2021). *sampling: Survey Sampling*. R package version 2.9, `https://CRAN.R-project.org/package=sampling`.

Yan, L. & Lohr, S. L. (2022). *R Companion for Sampling: Design and Analysis, Third Edition*. New York: Chapman and Hall/CRC Press.