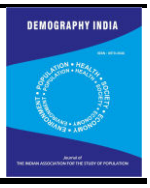
Book Review

Demography India, Vol. 54, Issue 3 (September-December, 2025) ISSN 0970-454X

Demography India
A Journal of Indian Association of Study of Population
Journal Homepage: https://demographyindia.iasp.ac.in/



Laishram Ladusingh (2025): *Statistics for Health Science with R* Himalaya Publishing House (xii+399)

Gudakesh Yadav[1]* and Koustuv Ghosh[2]

Abstract

This review examines *Statistics for Health Science with R* by Prof. Laishram Ladusingh, a comprehensive and application-oriented text designed to strengthen statistical reasoning in health research. The book integrates foundational concepts, ranging from measurement, descriptive statistics, probability, and sampling to advanced analytical techniques including hypothesis testing, ANOVA, categorical data analysis, regression modelling, limited dependent variable models, multilevel analysis, and introductory machine learning. A distinguishing feature is the systematic embedding of R programming throughout the chapters, enabling learners to translate theory into reproducible empirical practice. The review highlights the book's pedagogical strength in balancing conceptual clarity with computational application, particularly within contexts where statistical training has been predominantly formula-driven. While acknowledging its robust coverage and relevance for students and researchers in public health, biostatistics, epidemiology, demography, and clinical research, the review also suggests enhancements for future editions, such as introductory R programming, expanded advanced methods, and greater emphasis on reproducible research workflows. Overall, the book emerges as a timely, practically oriented, and methodologically grounded resource for data-driven health science education.

Keywords

Analytical techniques, Biostatistics, Epidemiology, Health Science Statistics, R Programming.

* *Corresponding Author*
[1] *Assistant Professor, Population Research Centre, The Maharaj Sayajiro University of Baroda, Gujarat. Email-Id: gudakesh@yahoo.com*
[2] *Research Investigator, Population Research Centre, The Maharaj Sayajiro University of Baroda, Gujarat. Email-Id: koustav2020@gmail.com*

The increasing digitisation of health systems, expansion of epidemiological surveillance, and growing emphasis on evidence-based policymaking have fundamentally transformed the landscape of health research. Across clinical practice, public health programmes, and demographic analysis, statistical reasoning is no longer optional but central to informed decision-making. In this rapidly evolving analytical environment, there is a pressing need for pedagogically sound resources that not only explain statistical concepts but also enable learners to apply them effectively using contemporary computational tools. It is within this context that Statistics for Health Science with R by Prof. Laishram Ladusingh, published by Himalaya Publishing House, makes a timely and significant contribution, and the book is priced at 625 rupees and is available on Amazon.

Prof. Ladusingh is a highly esteemed Indian scholar, renowned for his significant contributions to the fields of demography, statistics, public health, and population studies. He has held prominent leadership positions in various academic and research institutions across India. He has served as the Vice-Chancellor of Bodoland University. Previously, he was the Director as well as Professor and Head of the Department of Mathematical Demography and Statistics at the International Institute for Population Sciences (IIPS), Mumbai. Prof. Ladusingh has played a pivotal role in the design and implementation of major national demographic and health surveys in India.

In addition to Statistics for Health Science with R, Prof. Ladusingh has contributed significantly to statistical literature through his book Survey Sampling Methods (January 2018). This work focuses on theoretical and practical aspects of sampling techniques, including stratified, cluster, and multistage sampling designs. The emphasis on applied illustrations reflects the author's consistent pedagogical approach. Compared to Survey Sampling Methods, the present volume adopts a broader scope by extending beyond sampling to comprehensive statistical modelling in health sciences. While the earlier work concentrated primarily on sampling design and estimation, the current book integrates regression analysis, multilevel modelling, epidemiological research design, and machine learning within a unified framework using R. This progression demonstrates the author's evolution from specialised sampling methodology to a more integrated and computationally oriented approach to health statistics, aligning with contemporary demands in data-driven health research.

Statistics for Health Science with R provides a comprehensive foundation in statistical methods for health sciences, covering concepts from basic measurement of variables to advanced modelling techniques for assessing cause-and-effect relationships. What distinguishes the book is its balanced integration of theory and application, presented with minimal mathematical abstraction and supported by practical illustrations using the open-source statistical software R. The rationale of the book lies in addressing a longstanding pedagogical gap in health science education, particularly in India and other developing contexts, where statistical instruction has often remained formula-driven and detached from computational practice. By embedding R programming within statistical explanation, the author not only enhances conceptual clarity but also promotes reproducibility, transparency, and analytical rigor in health research. In doing so, the book aligns itself with contemporary methodological

standards in epidemiology, biostatistics, and public health training.

A particularly noteworthy feature of the book is its systematic integration of statistical theory with practical implementation using the open-source software R. Rather than treating computation as an afterthought, the author embeds R-based illustrations throughout the text, allowing readers to immediately apply theoretical concepts to real or simulated health data. This approach significantly enhances conceptual clarity, as learners can visualise distributions, test hypotheses, and build regression models in a hands-on manner. The use of R not only promotes reproducibility and transparency but also aligns the book with contemporary methodological standards in public health and biostatistics training. Importantly, the theoretical explanations are presented with minimal mathematical abstraction, making complex statistical ideas accessible without oversimplification. By maintaining a balanced proportion between theory and application, the book succeeds in transforming statistical learning from a formula-driven exercise into an analytical and computational practice. This pedagogical design strengthens both interpretative skills and technical competence, which are essential in modern health research. In addition to authoring several books, Prof. Ladusingh has published a large number of research papers, making significant scholarly contributions to the fields of demography and population studies.

The volume comprises thirteen chapters that move systematically from data representation and descriptive statistics to advanced inferential techniques such as hypothesis testing, analysis of variance and covariance, categorical data analysis, and linear regression. Emphasis is placed on limited dependent variable models, logistic, probit, and multinomial regression which are central to analysing binary and categorical health outcomes. The treatment of multilevel regression is especially noteworthy, as it addresses the hierarchical nature of biomedical and community-level data, a common feature in public health research. The book also devotes substantial attention to research designs in epidemiology, providing conceptual orientation and applied illustrations relevant to postgraduate training in public health and biostatistics. The consistent use of graphical and visual tools enhances interpretability, while the inclusion of supervised and unsupervised machine learning algorithms connects traditional biostatistics with contemporary advancements in health data analytics. In this sense, the volume responds to a long-standing need for a practical and methodologically grounded user manual in health statistics.

The book is particularly well-suited for postgraduate students and early-career researchers in public health, biostatistics, epidemiology, community medicine, and demography. Its clear explanations and integrated R-based demonstrations also make it valuable for practitioners involved in health-system research, programme evaluation, and survey-based studies. Faculty members designing courses in health statistics will find the book an excellent teaching companion, given its balanced blend of conceptual clarity and computational application. By bridging theoretical foundations with reproducible R workflows, the book caters to learners who seek both analytical depth and hands-on practical skills.

The book begins by laying a strong conceptual foundation through a discussion on the nature of health data, types of variables, and scales of measurement. By carefully distinguishing between nominal, ordinal, interval, and ratio scales, the author clarifies how measurement frameworks influence statistical choice and interpretation. This methodological grounding is essential in health research, where improper classification or measurement can lead to flawed inference and misguided policy conclusions.

It then moves to data representation and descriptive statistics, emphasising both graphical and numerical summarisation techniques. Measures of central tendency and dispersion are discussed in relation to health indicators such as morbidity and mortality rates. The consistent use of visual representation enhances interpretability and strengthens the ability of students to communicate empirical findings effectively.

Probability theory is introduced within a health research framework rather than as abstract mathematics. Concepts such as conditional probability and Bayes' theorem are linked to diagnostic testing and risk assessment. By situating probability within clinical and epidemiological contexts, the author makes theoretical ideas more accessible and practically meaningful.

The discussion of probability distributions, including binomial, Poisson, and normal distributions, connects statistical models to real-world health events such as disease incidence, rare conditions, and population variability. This linkage between theory and empirical scenarios strengthens inferential reasoning and prepares readers for applied modelling.

Sampling designs and estimation procedures are subsequently examined, with attention to sampling distributions and the logic of inference. The emphasis on representativeness and variability is particularly relevant for survey-based public health research. By explaining the foundations of estimation, the author reinforces the importance of drawing valid population-level conclusions.

The treatment of hypothesis testing introduces both parametric and non-parametric approaches within structured decision-making frameworks. Step-by-step illustrations clarify the logic of statistical significance and error probabilities. This section is especially useful for clinical trials, intervention studies, and community-based health research.

Attention then shifts to categorical data analysis, focusing on contingency tables, measures of association, and interpretation of risk. Given that many health outcomes are binary or categorical in nature, this section provides essential tools for epidemiological investigation and programme evaluation.

The application of analysis of variance and covariance demonstrates how comparisons across multiple groups can be conducted rigorously. By integrating experimental and quasi-experimental contexts, the discussion strengthens understanding of treatment effects and controlled comparisons in health studies.

Linear regression modelling is explored as a method for examining relationships between continuous health variables. The author explains model assumptions, interpretation of coefficients, and diagnostic considerations, ensuring that readers understand both the analytical potential and limitations of regression analysis.

The analytical scope deepens with a detailed treatment of limited dependent variable models, including logistic, probit, and multinomial regression. As binary and categorical outcomes dominate health research, this section represents one of the strongest analytical contributions of the book. The emphasis on interpretation of odds ratios and predicted probabilities enhances practical relevance.

The discussion further expands into multilevel regression modelling, acknowledging the hierarchical and clustered structure of biomedical and community-level data. By addressing contextual and group-level effects, the author moves beyond single-level analysis and strengthens methodological sophistication.

Research design in epidemiology is integrated with statistical reasoning, highlighting issues of bias, validity, confounding, and causal inference. This section strengthens the coherence between statistical tools and study design principles, aligning well with postgraduate curricula in public health and biostatistics.

The volume concludes with an introduction to machine learning techniques, including supervised and unsupervised algorithms. Although primarily introductory, this section effectively situates traditional biostatistics within the broader and rapidly evolving domain of health data science. By incorporating contemporary analytical approaches, the book connects foundational methods with emerging computational advancements.

Overall, Statistics for Health Science with R emerges as a comprehensive and pedagogically robust contribution to health statistics and biostatistical education. The book successfully bridges foundational statistical concepts with applied analytical practice, ensuring that readers move beyond formula-based learning toward conceptual clarity and computational competence. A defining strength of the volume is the systematic integration of R throughout the chapters. Each methodological component ranging from descriptive statistics and probability distributions to hypothesis testing, ANOVA, regression modelling, and multilevel analysis is accompanied by R-based illustrations, enabling readers to translate theory directly into empirical application.

The author adopts a hands-on computational approach, demonstrating how statistical procedures are implemented, interpreted, and visualised within R. For example, regression diagnostics, estimation of odds ratios in logistic models, handling of categorical outcomes, and modelling of hierarchical data structures are presented with executable R commands. This not only enhances interpretability but also promotes reproducibility and transparency qualities central to contemporary health research. Even advanced sections, such as machine learning, incorporate R-based algorithmic demonstrations, connecting classical statistical reasoning with modern data science tools.

The structured progression from measurement and descriptive analysis to advanced modelling techniques reflects the analytical trajectory required in epidemiological and public health research. By combining minimal mathematical abstraction with applied R demonstrations and graphical outputs, the book ensures accessibility without sacrificing analytical depth. While certain advanced themes could be further expanded in future editions, the

volume effectively addresses a longstanding need for a practically oriented, computationally integrated guide tailored to health sciences. In sum, it stands as a valuable resource for students, researchers, and practitioners seeking to strengthen quantitative skills in an era defined by data-driven health research and evidence-based policymaking.

While Statistics for Health Science with R is comprehensive and methodologically sound, certain additions could further strengthen future editions. First, the inclusion of a brief introductory section on basic R programming, data management, and package handling would enhance accessibility for beginners. Second, expanded coverage of advanced analytical areas such as survival analysis, time-to-event models, causal inference techniques, and longitudinal data analysis would increase its relevance for clinical and cohort-based research. Third, greater emphasis on reproducible research practices such as the use of R Markdown, workflow documentation, and real-world national health datasets would align the book more closely with contemporary research standards. Finally, a more detailed discussion of machine learning model validation and performance evaluation would deepen analytical rigor. Incorporating these elements would further consolidate the book's position as a forward-looking and practically oriented guide in health statistics.

A notable strength of the book lies in its smooth narrative flow and highly accessible writing style. In domains such as biostatistics and health analytics, where authors often struggle to maintain readability without compromising technical depth, Prof. Ladusingh succeeds remarkably well. The clarity of exposition, logical sequencing of concepts, and balanced integration of R examples make the text genuinely enjoyable to read. This ability to sustain readability in a technically demanding subject reflects the author's seasoned pedagogical skill and commitment to learner-centric writing.

In conclusion, the book covers a wide range of analytical techniques essential for contemporary health research, including descriptive statistics, probability modelling, hypothesis testing, analysis of variance and covariance, categorical data analysis, linear and logistic regression, multinomial and probit models, multilevel modelling, and introductory machine learning methods. These analytical tools make the volume particularly useful for postgraduate students in public health, biostatistics, epidemiology, community medicine, and demography, as well as for researchers conducting survey-based and clinical studies. The integration of R across these methods enhances its practical value for thesis work, programme evaluation, and research projects requiring reproducible statistical analysis. From a policy perspective, the book equips users to analyse health indicators, assess intervention effectiveness, model risk factors, and interpret hierarchical and contextual data skills essential for evidence-based policymaking and health system planning. By linking statistical techniques with epidemiological research design and real-world applications, the volume not only strengthens methodological learning but also supports data-driven decision-making in public health programmes and institutional research settings.

What sets this book apart is its rare combination of methodological rigor, accessible writing style, and seamless

integration of R for real-world analysis. At a time when statistical education in health sciences often remains formula-driven and disconnected from practical computation, the book offers a refreshing and much-needed model of applied, reproducible learning. By aligning classical biostatistical methods with modern data-science tools, it stands out as a uniquely valuable resource for strengthening evidence-based research and analytical capacity in the health sector.