Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/GC03.20250026
Published: February 2, 2026

Small Area Estimation of Class-Specific Proportions in Ordinal Categories Using Mixed Effect Cumulative Link Mixed Models

Muhammad Hamza
Department of Bio & Medical Big Data
Gyeongsang National University
Jinju-si, South Korea
hamza640@gnu.ac.kr

Shakeel Ahmed
Department of Molecular and Translational Medicine in the Paul L. Foster School of Medicine at Texas Tech University Health Sciences Center El Paso
United States
Department of Mathematics, School of Natural Sciences, National University of Science and Technology
Islamabad, Pakistan
shakeel.ahmed@ttuhsc.edu

Cho Sungbin
Department of Information and Statistics
Gyeongsang National University
Jinju-si, South Korea
sungbin@gnu.ac.kr

Youngsoon Kim
Department of Information & Statistics and Department of Bio & Medical Big Data
Gyeongsang National University
Jinju-si, South Korea
youngsoonkim@gnu.ac.kr

Abstract

Most of the studies on small area estimation (SAE) focused on the quantitative, count, and categorical variables. However, in practical policy making problems we also require information on class-specific proportions of an ordinal response variable in small areas. This study proposed a model-based approach for estimating class-specific proportions in small areas using unit level cumulative linked mixed models (CLMMs). The CLMMs based SAE method captures the ordering structure of the categories as well as the nested structure of the data to capture between area variability. The suggested CLMMs based method was compared with the direct method using a simulation as well as study on immunization coverage estimation in different districts of Pakistan using Pakistan Demographic and Health Survey (PDHS) 2017-18 data. The CLMM-based SAE methods outperform the direct method in terms of all performance measures used in this study.

Keywords
Small Area Estimation, Cumulative Link Mixed Models, Ordinal Response Variable, Model-Based Estimation, Demographic and Health Survey.

Acknowledgements
This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT challenge and Advanced Network of HRD) program (IITP-2025-RS-2022-00156409) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), and by the "Gyeongsangnam-do Regional Innovation System & Education (RISE)" project, supported by the Ministry of Education and Gyeongsangnam-do. Further, we would like to thank the Demographic and Health Surveys (DHS) for providing the comprehensive dataset of Pakistan 2017-18. that served as the foundation of our analysis.

Biographies
Muhammad Hamza is a Ph.D. candidate in Bio & Medical Big Data under the BK21 Four Program at the Graduate School of Gyeongsang National University, South Korea, where he works under the supervision of Professor Youngsoon Kim and in collaboration with Dr. Shakeel Ahmed. He received the B.S. degree in Statistics from the University of Haripur, Pakistan, in 2018, and the M.S. degree in Statistics from the National University of Science and Technology, Islamabad, Pakistan, in 2022, completing a thesis on a double-goal CART approach for domain-specific health outcome estimation in Pakistan under supervision of Dr. Shakeel Ahmed. From January to August 2023, he served as Biostatistician at Skytech Consultants Inc. Canada, performing and interpreting statistical analysis of healthcare datasets. Earlier, he was a lecturer at Department of Mathematics at Army Public School and College Kohat Cantonment, Pakistan (2019-2020). His research interest includes Biostatistics, Big Data, small-area estimation, model-based and model-assisted survey sampling, machine learning, and the study of income and health inequality with measures such as concentration indices and norm-based indices.

Shakeel Ahmed received the M.S. and Ph.D. degrees in Statistics from Quaid-e-Azam University, Islamabad, Pakistan. He is currently a Post-Doctoral Researcher in the Department of Molecular and Translational Medicine, Paul L. Foster School of Medicine, Texas Tech University Health Sciences Center El Paso, USA, and an Assistant Professor in the Department of Mathematics, School of Natural Sciences, National University of Science and Technology (NUST), Islamabad, Pakistan. Prior to these roles, he served as a Biostatistician at Wah Medical College and as an Assistant Professor of Statistics at the Institute of Business Management (IOBM), Karachi. He has published extensively on model-based parameter estimation, small-area estimation, and the application of machine learning and Bayesian methods to public health data. His current research focuses on model-assisted and model-based survey sampling methodologies, high-throughput biomedical data analysis, and statistical methods for health demography.

Cho Sungbin received the B.S. degree in Information and Statistics from Gyeongsang National University, Jinju, South Korea, in 2025. She is currently pursuing the M.S. degree in Information and Statistics at Gyeongsang National University under the supervision of Professor Youngsoon Kim. Her research interests include machine learning, big data analytics, information security, and measurement of income inequality.

Youngsoon Kim received the B.S., M.S., and Ph.D. degrees in Statistics from Gyeongsang National University, Jinju, South Korea, in 1994, 1998, and 2005, respectively. From 2009 to 2020, she served as a senior researcher at the Gyeongnam Institute, South Korea. Between 2017 to 2019, she was a postdoctoral researcher and part-time Assistant Professor at Kennesaw State University, USA. She is currently an Associate Professor in the Department of Information and Statistics and the Department of Bio and Medical Big Data at Gyeongsang National University, South Korea. Her research interests include bioinformatics, machine learning, data mining, and big data analytics.