National Center of Excellence and Advance Research on Diets
Lady Irwin College, University of Delhi

Report
Capsular Course on Nutrition Epidemiology,
Module -2 Data Management Tools and Advanced Statistics (Using STATA)

Background

Applications of software packages for analysing of data of various types, including nutrition related data, has been gaining popularity among the public health researchers with the development of various software packages and publication of large data sets by the Government and private organisations. Applications of these packages help to analyse the data in a rigorous manner that in turn help to draw robust conclusions about the phenomena being studied.

Among various statistical software packages, STATA is considered to be most user-friendly and effective in rigorously analysing data sets of various types (cross-sectional, time-series and panel), applying the tools of statistics. The knowledge in applications of STATA would provide a definite boost in research of the nutrition/social scientists, as for others.

National Centre of Excellence and Advanced Research on Diets (NCEARD), has been established at Department of Food and Nutrition, Lady Irwin College, New Delhi to support the Maternal Health Division, Ministry of Health and Family Welfare, Government of India in convening policy discourse, developing and testing maternal malnutrition guidelines, enhancing knowledge of nutrition epidemiology under PHN training for evidence based nutrition programming in India. UNICEF India has generously provided the financial support to the centre for initial phase of its establishment.

NCEARD, Lady Irwin College, University of Delhi, signed a Memorandum of Understanding (MoU) with the International Institute for Population Sciences (IIPS), Mumbai on 25th October, 2018 to conduct capsular course on nutrition epidemiology at the centre, with technical and financial assistance from UNICEF India.

This course has been proposed given the fact that principles of nutritional epidemiology are not comprehensively covered in any existing course curriculum in India, which co-exists

with vast observational research being conducted in the country. The certificate course, covered through six modules on nutrition epidemiology, aims at enhancing the understanding of relationships between dietary intake, nutritional status and health outcomes among nutrition programmers, health care professionals and researchers.

The first module of the course – Module-1: Survey Methods – was conducted at Lady Irwin College during 26-29 October 2018. This module aimed at enhancing the knowledge on the history of nutrition programmes in India and emerging issues, the survey research techniques and various indicators of nutritional assessment, increasing application of research methodology in different research and programmatic outcomes; understand sources, estimation and implications of measurement errors in nutritional epidemiological research; understand, apply and critique the use of sampling and study designs to be able to get statistically meaningful data; and collaborate efficiently with epidemiologists, clinicians, nutritionists and statisticians on designing, planning, conducting and interpreting epidemiological research focused on the relationship of diet and disease.

The second module of the course focused on Data Management Tools and Advanced Statistics using STATA for handling, investigating and displaying Nutrition and public health data. The themes covered during the course were introduction of STATA and its functionalities; importing and exporting data; data management in STATA; Data manipulation; Statistical analysis; Plotting graphs; applying survey weights, survey sets and regression analysis. Additionally, commands for aforementioned analysis were discussed in details for repetitive analysis.

The Participants

A total of 24 participants included the researchers and program managers from the organizations such as Population Council, Piramal Foundation, Nutrition International, World Vision, IEG, ROSHNI, Patna medical college and hospital, Maulana Azad Medical College, Kalpana Chwala Govt. Medical College etc. completed the course and awarded the "certificate of attendance".

The Facilitators

Two professors from International Institute for Population Sciences, Mumbai and one professor from Institute of Economic Growth, New Delhi were the key facilitators to deliver the course contents to participants. These were Dr. Abhishek Singh (IIPS), Dr. William Joe (IEG) and Dr Kaushalendra Kumar (IIPS).

The Course

The training course of 32 hours, on "Data Management Tools and Advanced Statistics using STATA" was organized during 1-4 May 2019, at the Lady Irwin College. The purpose of the course was to make participants familiar with the STATA interface, its functionalities, managing and analysing data. Practical hands-on training interspersed with discussions and lectures was provided throughout the workshop to build capacity of participants on use of STATA.

The day-wise agenda covered the course contents in the following manner –

Day1: First day of course started with the registration of participants, a welcome note, and overview of course and details of module by professors of Lady Irwin College. The sessions on day-one included –

- Introduction of STATA software
- Operating STATA and its functionalities
- Entering data in STATA

- Data management in STATA

Day-2: The second day of training started with a recap of day-1 sessions and answering the questions and clarifications. The following sessions of day-2 were-
- Data manipulation (I)-Computing of new variables, recoding variables and apply logics
- Data manipulation (II)- Management of data file, deleting variables and sorting data
- Data manipulation (III)- Restructuring of data set
- Data manipulation (IV)- Merging files and aggregating files

Day-3: The day three sessions began with review of the sessions conducted on day-2 while addressing questions and clarifications of participants. The day-3 sessions covered –

- Basic statistical analysis (I)- Understanding nature and distribution of the variables, applying correlations and identifying associations
- Basic statistical analysis (II)- Applying statistical hypothesis tests in STATA
- Graph (I)-Plotting and interpreting graphs in STATA
- Graph (II)- Plotting and interpreting scatter plots and regression graphs in STATA

Day-4: The fourth day of training started with an overall recap of the sessions covered in past three days. The sessions conducted on day-4 included –

- Survey weights and survey set
- Regression analysis-Linear and Logistic

Valedictory: The valedictory session was conducted toward the end of 4-day of training. The session started with welcome note followed by overview of the course. The facilitators expressed their closing remarks and reiterated the importance of the software in their routine research activities. In addition, some of the participants also shared their feedback regarding the training. Overall, the course was conducted successfully and all the participants took active part in all the four days of the course. The participants found the training content very useful for their day-to-day operations.

Feedback from Participants

At the end of sessions on each day, a feedback form was given to the participants to provide feedback on the sessions conducted during the day. Participant feedback was solicited on three key aspects of session, which include – contents covered in the session, presentation/communication skills of facilitator, and interactions between facilitator and participants. In addition, the participants were also asked for their rating on quality of food and snacks served during training.

Overall, the participants rated most of the sessions of "high quality". At the end of day-1, the overall score given by participants was 4.5 points on a five points scale, starting from "1=bad/not useful" to "5=excellent/extremely useful". The qualitative feedback for day-1 showed participants were satisfied with the session's content and guidance received from the speakers. They further appreciated the inclusion of practical exercises in the sessions.

Few of the participants mentioned the need of more interactions.

Learning from day-1, the day-2 sessions were redesigned to introduce greater interaction between facilitators and participants. These efforts from facilitators received positive outcome in terms of the participants giving an average overall rating as 4.6 out of 5 for day-2 sessions.

Throughout the 4-days training the facilitators and course coordinator made efforts to learn and improve upon to meet participants' expectations; some of which included uploading the course material on Google drive, giving homework, providing links to suggested reading, etc. These efforts showed up in positive outcome in terms of overall rating for day-3 and day-4 increasing to 4.8 out of 5 points (both days).

A session-wise analysis of feedback from participants show that sessions on statistical analysis, survey weight and survey set were most appreciated; the ratings for these sessions was 4.9 out of 5 points (Figure 1).

Figure 1: Participant's feedback on Module-2 sessions
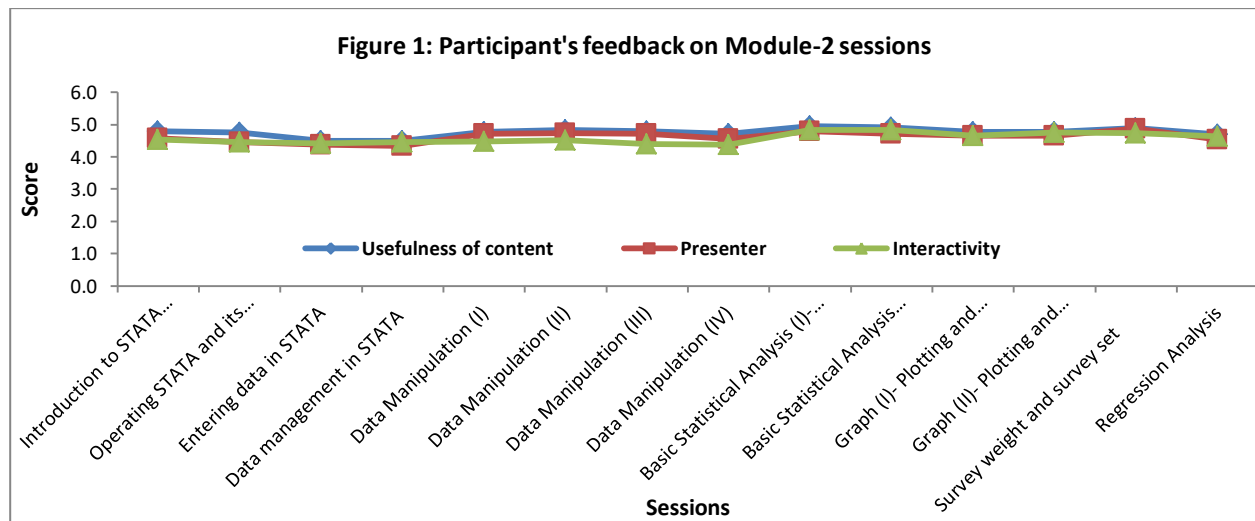
Suggestions received

The following suggestions were received from the participants of module-2.

- The course should be extended for a week long and more emphasis should be given on advanced statistical analysis.

- The technical courses should be divided into two modules-basic and advanced.
- The course should be according to the level of participants. A homogenous group should be selected such type of technical course.
- There should be enough time to prepare for the training, arranging for background material, dates of

application and shortlisting of applicants, etc. it is envisaged that after finalizing the agenda and facilitators, at least 4-weeks' time is needed to have a sizeable number of participants with basic minimum requirements for the module training.