

SOFTWARE FOR SURVEY DATA ANALYSIS



SSDA Version 1.0

DIVISION OF COMPUTER APPLICATION



INDIAN AGRICULTURAL STATISTICS RESEARCH INSTITUTE

LIBRARY AVENUE, PUSA, NEW DELHI- 110 012



SSDA At a Glance

- SSDA 1.0 is a software package with a user friendly interface for the analysis of survey data. The software is completely menu driven and guides users step-by-step through data analysis process.
- SSDA 1.0 has been developed in C# language in the .NET platform. A class library has been developed which contains the methods for the sampling schemes available in the software for analysis.
- Using SSDA 1.0, you can quickly and efficiently manage your data, get estimate of parameters based on the sampling design and design efficiency of the sampling design in comparison to the simple random sampling without replacement.
- It also has the facility to impute missing data, if any, using commonly used imputation methods.
- This package is an aid in teaching the subject of analysis of sample survey data to the post-graduate students and is also helpful to the researchers in statistics with interest in sample surveys.

DATA HANDLING:

The screenshot shows the 'Software for Survey Data Analysis' interface. The main window displays a spreadsheet with columns labeled 'var_1' through 'var_6'. A context menu is open over the spreadsheet, showing options like 'Row', 'Column', 'Sort', and 'Filter'. The 'Row' menu is expanded, showing 'Delete', 'Insert', 'Below', 'Above', and 'At the end' options. The 'Insert' menu is also expanded, showing 'Delete', 'Below', 'Above', and 'At the end' options. The spreadsheet data includes numerical values and some null entries.

var_1	var_2	var_3	var_4	var_5	var_6
1	1	0.00020040	0.010073079	102	
2	1		0.004740272	48	(null)
3	1		0.011159391	113	(null)
4	1	1	55.6		
5	1	1	58.1		
6	1	1	40.8		
7	1	2	57.4		
8	1	2	16.8		
9	1	2	20.3		
10	1	2	70.1		
11	1	2	34.6		
12	1	2	22.6		
13	1	2	44.9		(null)
14	1	2	28.4		(null)
15	1	2	17.5		(null)
16	1	2	33.7		(null)

Features

- Spreadsheet like data editor
- Creating a new data file
- Saving a data file
- Open an existing data file
- Importing input data from MS excel, text and MS access file
- Renaming columns
- Jumping to a given row
- Filtering the data column wise
- Filtering the data row wise
- Insert or delete columns or rows
- Imputation of missing data using mean, average of the preceding and succeeding observations, zero substitution
- Printing a data file

Imputation

The screenshot shows the 'Imputation' dialog box. It has two sections: 'Specify Stratum and Cluster' and 'Imputation'. In the first section, there are two checked checkboxes: 'Whether Stratum present in the data?' with a dropdown menu set to 'var_1', and 'Whether Cluster present in the data?' with a dropdown menu set to 'var_2'. The second section, 'Imputation', has a subtitle 'Columns with missing values (Hold Ctrl to select Multiple Columns)'. It contains a list box with 'var_3', 'var_4', 'var_5', and 'var_6', where 'var_3' is selected. To the right of the list box are three radio buttons for 'Imputation Methods': 'Mean Method', 'Mean of Neighboring Units Method' (which is selected), and 'Zero Method'. At the bottom of the dialog are 'OK' and 'Cancel' buttons.

ESTIMATES

This software provides the estimates of population mean, variance and design efficiency of the sampling scheme in comparison to the simple random sampling without replacement. This software analyzes the data collected through the following sampling designs:

S.No.	Sampling Scheme	Replacement Scheme	Probability Scheme	Estimation Method	S.No.	Sampling Scheme	Replacement Scheme	Probability Scheme	Estimation Method
1.	Random	WR	Equal	TM & RM	4.	Two-stage	WR	Equal	TM
		WOR	Equal	TM & RM			WOR	Equal	TM
		WR	Unequal	TM			WR (first stage)	Unequal (first stage)	TM
2.	Stratified	WR	Equal	TM & RM (S & C)	5.	Stratified Two stage	WOR	Equal	TM
		WOR	Equal	TM & RM (S & C)			WR (first stage)	Unequal (first stage)	TM
		WR	Unequal	TM			6.	Systematic	-
3.	Cluster	WR	Equal	TM	WOR : Without Replacement, WR : With Replacement TM: Traditional Method by using the data of study variable(s). RM: Ratio Method by using the data of study & auxiliary variable(s). S: Separate Type, C: Combined Type				
		WOR	Equal	TM					
		WR	Unequal	TM					

ANALYSIS. RESULTS & HELP

Analysis of survey data collected using the specified sampling schemes. Also provides the basic statistics of the data without considering the design.

Select Sampling Design

Sampling Scheme

- Simple Random Sampling
- Systematic Random Sampling
- Probability Proportional to Size
- Stratified Sampling
- Cluster Sampling
- Two Stage Sampling
- Stratified Two Stage Sampling

Probability

- With Equal Probability
- With Unequal Probability

Replacement

- Without Replacement
- With Replacement

Next Cancel

Results are displayed in the form of reports which can be easily exported to MS-word, MS-excel and PDF formats. Facilities like print, find and zoom are also available

SSDA 1.0 Report

Sampling Scheme: Stratified Two Stage Sampling With Replacement and Unequal Probability (n First Stage)

User Supplied Sampling Parameters for Analysis

Status Variable: var_1
 Probability Variable: var_4
 Study Variable(s): var_2
 Primary Stage: var_2
 Population Size for Secondary Stage: var_3
 Population Size(s) for each Stratum: var_5

Population Sizes of PSU's

Stratum Number	Population Size
1	1,100
2	1,100

Estimate of Parameters

Study Variable	Status	Mean	Variance	RSE (%)	Variance (SE)	Design Efficiency
var_3	1	1.0412224	0.0031901	5.4244550	3%	3%
var_3	2	1.2718889	0.0046429	5.3172853	3%	3%
var_3	Pooled	1.1510047	0.0039278	3.8140638	2.9418454	1,849.547038

Descriptive Statistics

Study Variable	Mean	Variance	StdDev	Skewness	Kurtosis	Coefficient of Variation
var_3	43.6269738	257.1032947	41.8000000	0.4834911	2.8172523	36.7540082

SSDA - Stratified Two Stage

Select Column For Stratum: var_1 No. of Stratum: 2

Select Column For Primary Stage Unit: var_2

Select Column For Probability: var_4

Enter Pop. Sizes For SSU

- Select Column for SSU Pop Sizes: var_5
- Enter Manually

Enter Population Size (N) Where i=1 to L

Help: Enter Sizes by giving spaces()

- Select Column For Pop Sizes: var_5

var_4 var_3

Proceed Cancel

This is an example showing the parameters to be supplied by the user for stratified two stage sampling.

SSDA Help - JAGRU New Delhi

Contents Index Search Favorites

- ABOUT SSDA
- Sampling
 - Definitions
 - Sampling Schemes
 - Simple Random Sampling
 - Stratified Random Sampling
 - Systematic Sampling
 - Cluster Random Sampling
 - Two-Stage Sampling
 - Introduction
 - Notations
 - Estimation of Parameters
 - Use of SSDA

Estimation of Parameters

Estimator of population mean, when M_0 is known:

$$\bar{y}_{st} = \frac{N}{nM_0} \sum_{i=1}^L M_i \bar{y}_i$$

Variance of the estimator \bar{y}_{st} :

$$V(\bar{y}_{st}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{i=1}^L \frac{M_i^2}{M_0^2} \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$$

Estimator of variance $V(\bar{y}_{st})$:

$$V(\bar{y}_{st}) = \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{i=1}^L \frac{M_i^2}{nN \sum_{i=1}^L M_i^2} \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_i^2$$

Estimator of population mean which does not depend on M_0 :

$$\bar{y}_{st} = \frac{1}{n} \sum_{i=1}^L \bar{y}_i$$

Bias of the estimator \bar{y}_{st} :

A detailed manual to use the software and the mathematical formulas related to the methods used in the package are provided. A HTML-based help features content, index and search capabilities.

System Requirements

1. Operating System Windows 2000/XP/2003/Windows Vista (Server or Professional)
2. Hardware Intel Pentium IV or above
3. .Net Framework .Net Framework 1.1/2.0 or Higher
(Available with the Installation Disk)
4. Setup Disk The setup.exe available automatically checks whether .NET Framework is installed. If not present, the setup installs it.
5. Other Requirements Microsoft Data Access Component, Crystal Report Viewer (If not installed, the setup installs these too).

Contact Us

Dr. V. K. Mahajan vkm@iasri.res.in
Sh. S. B. Lal sblall@iasri.res.in
Mrs. Anu Sharma anu@iasri.res.in
Division of Computer Applications

Head
Division of Computer Applications
Email: head_ca@iasri.res.in

Director
Email: director@iasri.res.in

**Indian Agricultural Statistics Research Institute, ICAR,
Library Avenue, New Delhi – 110012, INDIA
Phone: 91-11-25841479, Fax: 91-11-25841564**