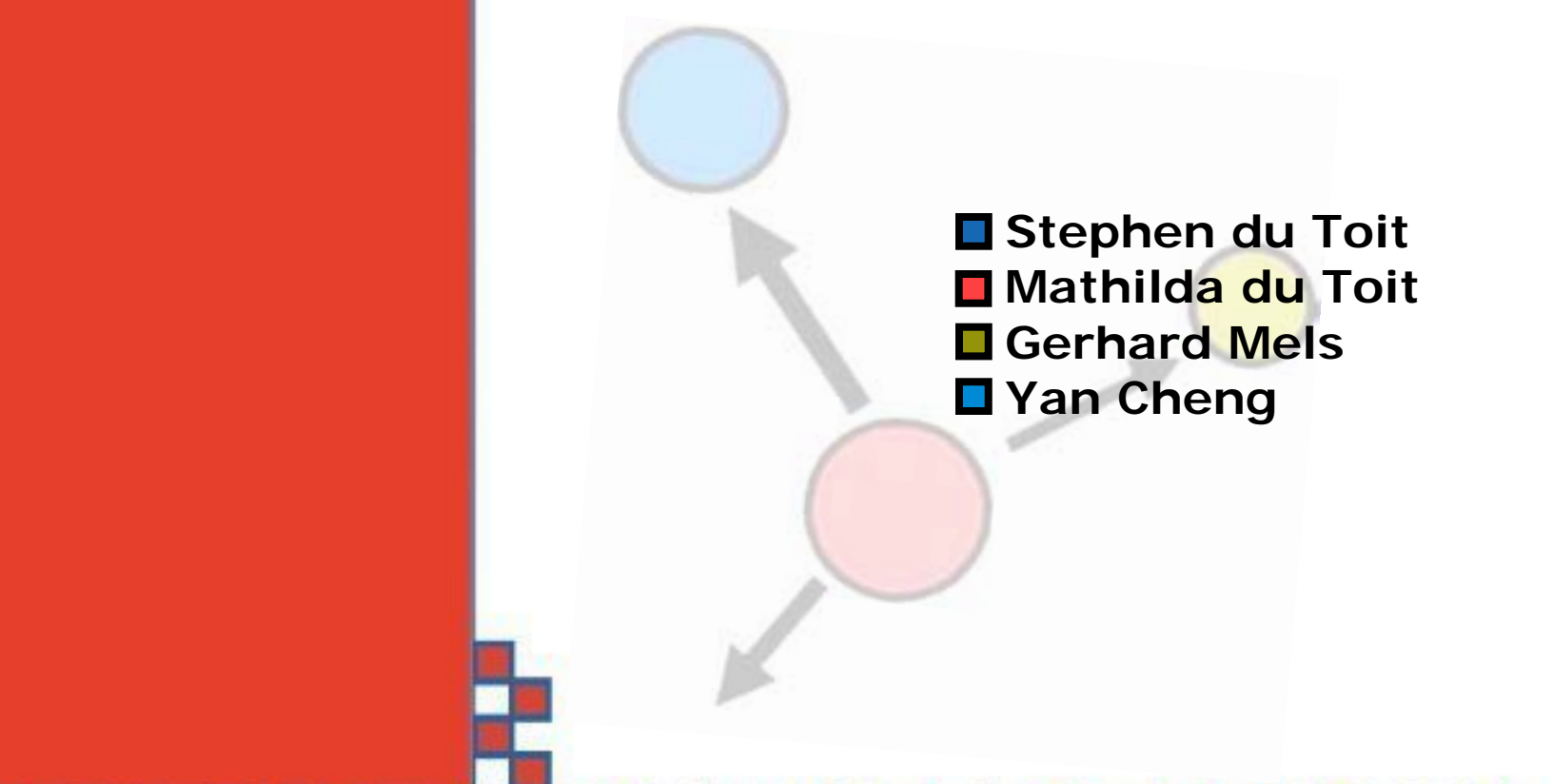


- 
- Stephen du Toit
 - Mathilda du Toit
 - Gerhard Mels
 - Yan Cheng



LISREL for Windows: PRELIS User's Guide



SSI SCIENTIFIC
SOFTWARE
INTERNATIONAL

Table of contents

INTRODUCTION.....	1
GRAPHICAL USER INTERFACE	2
The Data menu	2
The Define Variables dialog box	2
The Add Variables dialog box.....	4
The Variable Types for... dialog box.....	4
The Category Labels for... dialog box.....	4
The Missing Values for... dialog box	4
The Select Data dialog box	5
The Weight Cases dialog box	7
The Survey Design dialog box.....	8
The Transformation menu	9
The Recode Variables dialog box	9
The Compute dialog box.....	10
The Statistics menu.....	13
The Impute Missing Values dialog box	13
The Multiple Imputation dialog box	14
The Equal Threshold Test dialog box	16
The Fix Thresholds dialog box	16
The Homogeneity Test dialog box.....	17
The Normal Scores dialog box	18
The Factor Analysis dialog box	19
The Censored Regression dialog box	21
The Logistic Regression dialog box	21
The Probit Regressions dialog box	22
The Regression dialog box	24
The Two-Stage Lease-Squares dialog box.....	25
The Bootstrapping dialog box.....	26
The Output dialog box	27
PRELIS SYNTAX.....	30
The structure of the PRELIS syntax file.....	30
CA command.....	31
CB command.....	32
CE command	32
CL command	33
CO command.....	33
DA command.....	34
ET command	36
FT command.....	37
HT command.....	37
IM command	38
LA command.....	39
LA paragraph.....	39
LO command	40
LR command	40
MI command	41
MT command	41
NE command.....	42
OR command.....	42

OU command.....	43
PO command.....	49
PR command.....	50
RA command.....	50
RA paragraph.....	51
RE command.....	52
RG command.....	52
SC command.....	53
SD command.....	53
SY command.....	54
TI paragraph.....	54
WE command.....	55
EXAMPLES.....	56
Processing continuous variables.....	56
Data preparation example using fitness data.....	56
The data.....	56
Importing the Excel data file.....	57
Defining the variable types.....	59
Dealing with the missing values.....	60
Defining a global missing value and listwise deletion.....	60
Multiple imputation.....	62
Inserting a new variable.....	65
Computing values for a new variable.....	66
Selecting cases and creating a subset of the data.....	68
Exporting the data to Excel.....	70
Data screening.....	72
Normal scores.....	74
Bootstrapping.....	76
Computing matrices.....	79
Multiple linear regression example using drug abuse data.....	81
The data.....	81
The model.....	82
The analysis.....	82
Two-stage least square example using US economy data.....	84
The data.....	84
The model.....	85
The analysis.....	85
Exploratory factor analysis example using psychological data.....	88
The data.....	88
The mathematical model.....	89
The specific model.....	89
The analysis.....	90
Processing ordinal variables.....	92
Data preparation example using political action survey data.....	92
The data.....	92
Importing the SPSS data file.....	94
Defining the variable types.....	95
Assigning labels to the categories.....	96
Dealing with the missing values.....	97
Defining the missing values.....	97
Listwise deletion.....	98
Imputation by matching.....	100
Data screening.....	103

Exporting the PSF to SPSS	104
Thresholds	106
Equal thresholds test and compute polychoric correlations	106
Fixed thresholds and compute polychoric correlations	109
Bootstrapping.....	111
Computing matrices	113
Homogeneity test.....	115
Logistic regression example using political action survey data.....	117
The model	117
The analysis	118
Multivariate probit regression example using political action survey data.....	119
The model	119
The analysis	120
REFERENCES	123

Introduction

LISREL for Windows (Jöreskog & Sörbom 2005) is a Windows application for structural equation modeling, multilevel structural equation modeling, multilevel linear and nonlinear modeling, generalized linear modeling and formal inference-based recursive modeling.

However, it can also be used to import data, prepare data, manipulate data and basic statistical methods such as multiple linear regression, logistic regression, probit regression, censored regression and exploratory factor analysis. LISREL for Windows consists of a 32-bit Windows application LISWIN32 that interfaces with the 32-bit applications LISREL, PRELIS, MULTILEV, SURVEYGLIM, CATFIRM and CONFIRM.

PRELIS is a 32-bit application for manipulating data, transforming data, generating data, computing moment matrices, computing asymptotic covariance matrices, performing multiple linear, censored, logistic and probit regression analyses, performing exploratory factor analyses, etc.

The 32-bit application LISREL is intended for standard and multilevel structural equation modeling. The Full Information Maximum Likelihood (FIML) method for missing data is also available for both standard and multilevel structural equation modeling. In the case of continuous data, these methods are also available for complex survey data. MULTILEV fits multilevel linear and nonlinear models to raw data while CATFIRM and CONFIRM allow formal inference-based recursive modeling for raw categorical and continuous data respectively. In the case of continuous response variables, MULTILEV allows for design weights at each level of the hierarchy. SUVEYGLIM fits generalized linear models to simple random sample and complex survey data.

LISREL for Windows uses the PSF (PRELIS System File) window to display the contents of a PSF, which contains the raw data to be processed. A PSF is typically obtained by importing external data files such as SPSS, SAS, STATA, Statistica, etc. data files. PRELIS, MULTILEV and SURVEYGLIM can be accessed interactively by means of the PSF window.

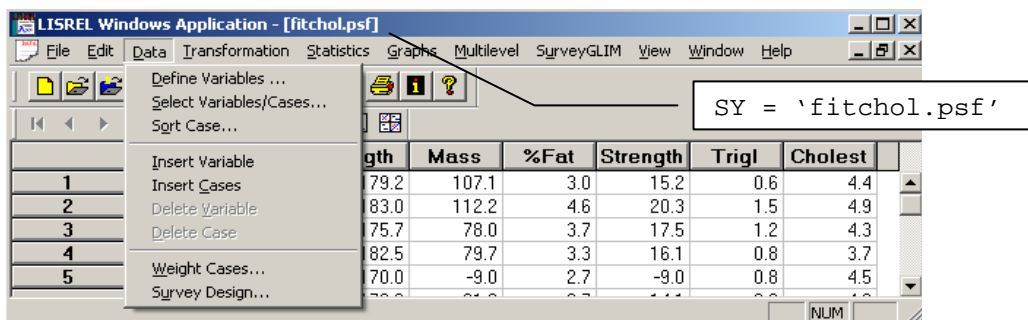
This document is intended as an illustrative user's guide for using the application PRELIS interactively and for preparing graphical displays. It supersedes Chapter 3 and many sections of Chapter 2 of Du Toit & Du Toit (2001). After we overview the graphical user interface of PRELIS, we give an overview of the PRELIS syntax file. Finally, we demonstrate how to use PRELIS to process data on continuous variables as well as data on ordinal variables. For both these variable types, we illustrate the preparation of the data and the application of basic statistical methods.

Graphical User Interface

The graphical user interface (GUI) of the PRELIS module consists of the options and dialog boxes of the Data, Transformation and Statistics menus on the PSF (PRELIS System File) window of LISREL for Windows. This GUI allows the user to interactively generate PRELIS syntax and output files. These three menus and the corresponding options and dialog boxes are reviewed in this section.

The Data menu

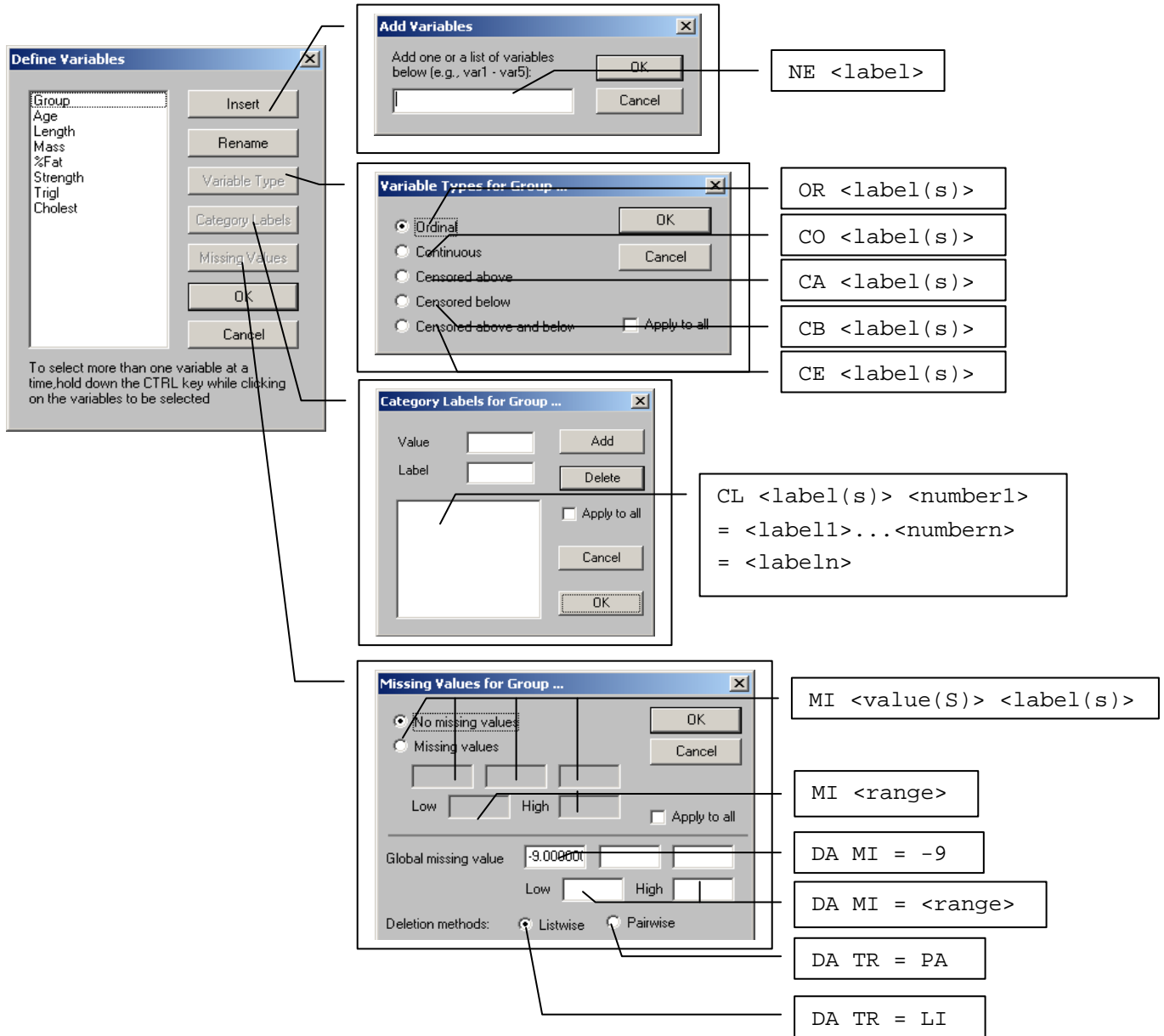
The **Data** menu is located on the PSF window of LISREL for Windows which is used to display, manipulate and process raw data. In other words, one needs to create a PSF and open it in a PSF window before using any option of the **Data** menu interactively. To illustrate this, the PSF window for the file **fitchol.psf** in the **TUTORIAL** subfolder with the **Data** menu expanded is shown below.



The **Data** menu allows the user to specify variables types, manipulate data, specify a case weight variable and complex survey design variables. Only the **Define variables**, **Select Variables/Cases**, **Weight Cases** and **Survey Design** options and the associated dialog boxes form part of the PRELIS GUI. They are discussed separately in the subsequent sections. The **Sort Case**, **Insert Variable**, **Insert Cases**, **Delete Variable** and **Delete Case** options and the corresponding dialog boxes are not part of the PRELIS GUI. Hence, they are not discussed here.

The Define Variables dialog box

The **Define Variables** dialog box allows the user to insert variables, rename variables, define variable types, assign category labels and define missing values. It is accessed by selecting the **Define Variables...** option on the **Data** menu. This selection loads the following **Define Variables** dialog box shown in the image below.



As shown above, the five variable buttons on the **Define Variables** dialog box are the **Insert**, **Rename**, **Variable Type**, **Category Labels** and **Missing Values** buttons. All these buttons become active as soon as a variable(s) is selected from the variable list box.

The **Rename** button enables the user to rename a variable by first selecting the label from the variable list box and then by clicking on the **Rename** button to change the variable label in the variable list box directly.

The other four variable buttons provide the user access to four dialog boxes. Each of these dialog boxes is discussed next.

The Add Variables dialog box

The **Add Variables** dialog box is accessed by clicking on the **Insert** button on the **Define Variables** dialog box. It corresponds with the NE command as indicated on its image.

The user can add a variable by first entering the variable label in the string box and then by clicking on the **OK** button to return to the **Define Variables** dialog box.

A list of variables can be added as well. However, this requires that the variable labels must end with sequential positive integers. For example, if we enter var1 – var3 in the string box, var1, var2 and var3 are added to the in the variable list once we click on the **OK** button.

Note that only the first 8 characters of variable labels are recognized and that the variable labels are **not** case sensitive.

The Variable Types for... dialog box

The **Variable Types for...** dialog box is accessed by clicking on the **Variable Type** button on the **Define Variables** dialog box. It corresponds with the OR, CO, CA, CB and CE commands as indicated on its image.

The user can choose the desired variable type by activating the **Ordinal**, **Continuous**, **Censored above**, **Censored below** or **Censored above and below** radio button. If the variable type applies to all the variables, the user needs to check the **Apply to all** check box. Finally, click on the **OK** button to return to the **Define Variables** dialog box.

The Category Labels for... dialog box

The **Category Labels for...** dialog box is accessed by clicking on the **Category Labels** button on the **Define Variables** dialog box. It corresponds with the CL command as indicated on its image.

To define the category labels, enter the number and the corresponding label in the **Value** and **Label** string boxes respectively. Then, click on the **Add** button to add a command line in the CL grid box. Check the **Apply to all** check box if the category labels apply to all the variables. Finally, click on the **OK** button to return to the **Define Variables** dialog box.

The Missing Values for... dialog box

The **Missing Values for...** dialog box is accessed by clicking on the **Missing Values** button on the **Define Variables** dialog box. It corresponds with the MI and DA commands as indicated on its

image.

The **Missing Values for...** dialog box consists of two sections. The upper section defines the missing values for the selected variable(s) only while the lower section defines the global missing values.

To specify the missing values of a selected variable(s), activate the **Missing values** radio button and enter the missing values in the corresponding string boxes. Three different missing values can be defined at a time. The user can also define a range of missing values by entering the lower and upper bounds of the range in the **Low** and **High** string boxes respectively. The **Apply to all** check box is checked if the missing values defined can be applied to all the variables. In this case, it is the same as specifying a global missing value(s).

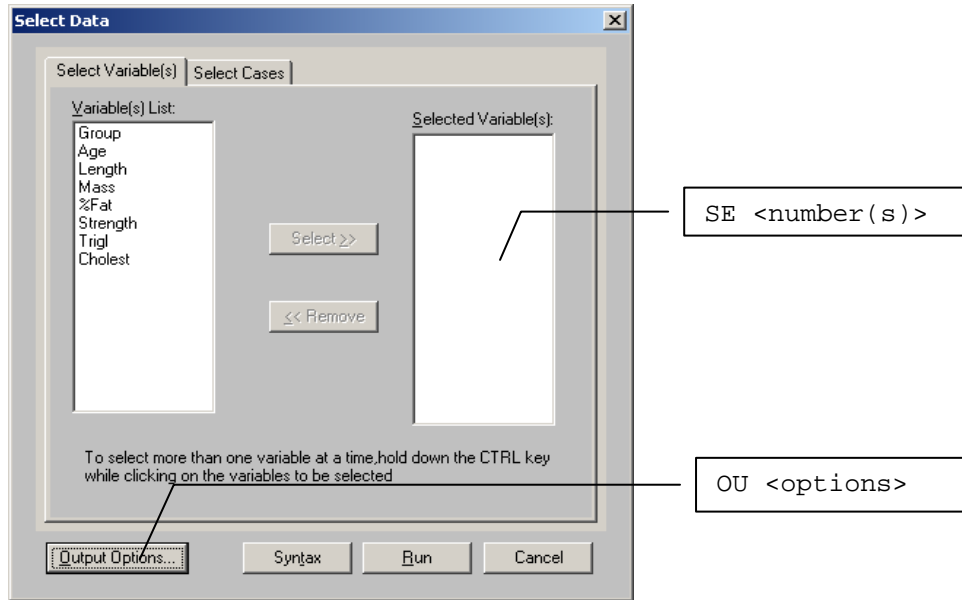
To specify global missing values, enter the missing value(s) in one or all **Global missing value** string box(es). Three different missing values can be defined at a time. A range of the missing values can be specified by entering the lower and upper bounds of the range in the **Low** and **High** string boxes respectively.

Two deletion methods, namely listwise deletion and pairwise deletion, are available in PRELIS. Either of the two deletion methods can be specified by activating one of the **Listwise** or **Pairwise** radio buttons. Click on the **OK** button to return to the **Define Variables** dialog box when done.

Once the user is done with all dialog boxes, click on the **OK** button on the **Define Variables** dialog box to return to the PSF window. Save the changes by selecting the **Save** option on the **File** menu.

The Select Data dialog box

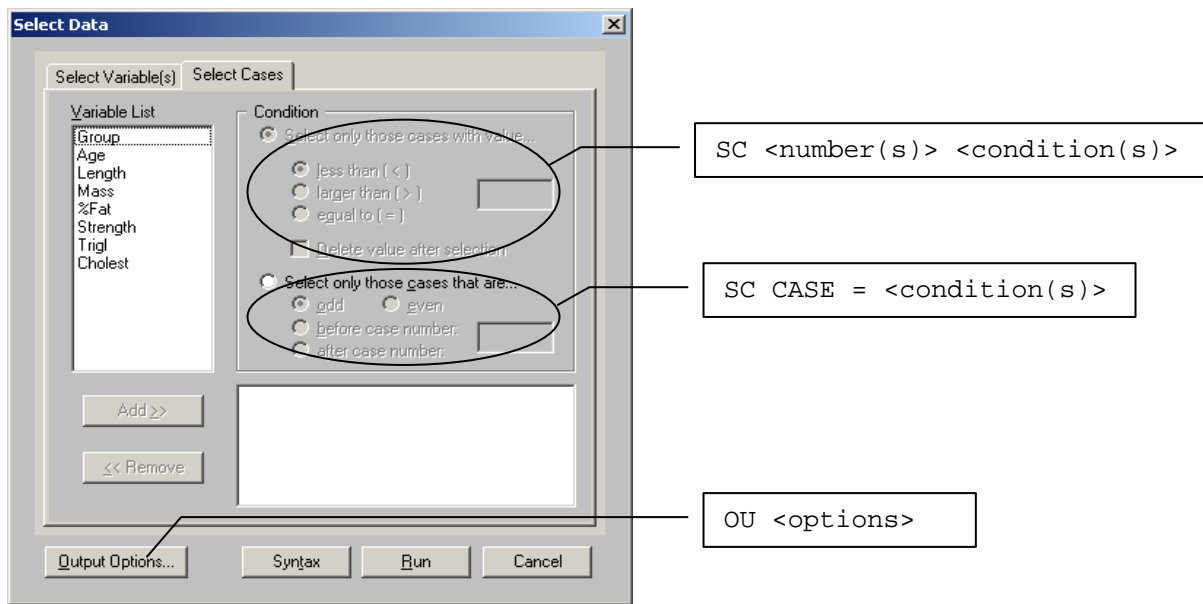
The **Select data** dialog box contains two tabs, namely the **Select Variable(s)** and the **Select Cases** tabs. The **Select Variable(s)** tab allows the user to select variables while the **Select Cases** tab allows the user to select cases. The **Select Variable(s)** tab is accessed by selecting the **Select Variables/Cases...** option on the **Data** menu. This selection loads the following **Select Variable(s)** tab of the **Select Data** dialog box.



Note that the **Select Variable(s)** tab of the **Select Data** dialog box corresponds with the SE and OU commands as indicated on the image above.

The user selects the variable(s) by first selecting the label(s) from the **Variable(s)** list box and then by clicking on the upper **Select** button to add the variables to the **Selected Variable(s)** list box.

A click on the **Select Cases** tab loads the following dialog box.



Note that the **Select Cases** tab of the **Select Data** dialog box corresponds with the SC and OU commands as indicated on the image above.

The user selects the cases by first selecting the desired label from the **Variable list** box to activate **Select only those cases with value** radio button. Then, choose one of the desired conditions of less than, larger than or equal to by activating the corresponding radio button. Enter the desired value in the corresponding string field. By default, the variable(s) selected is not excluded by PRELIS. The user can request PRELIS to ignore the selected variable by checking the **Delete values** check box.

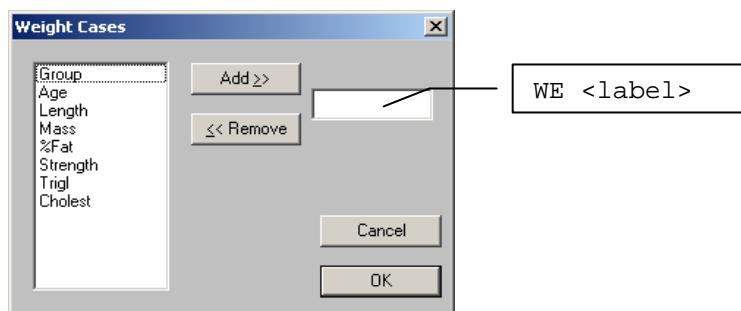
PRELIS offers some other options for selecting cases, which include selecting odd/even numbered cases or selecting the cases before/after a case number. To use these options, first activate the **Select only those cases that are** radio button and then activate the desired condition radio button. If the cases before/after a case number are wanted, enter the case number in the corresponding string box.

The **Output Options** button provides access to the **Output** dialog box, which can be used for example to save the selected variables/cases in a new PSF (see the **Output** dialog box section).

Once the user is done with the **Select Data** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Weight Cases dialog box

The **Weight Cases** dialog box allows the user to define the variable to be used to weight cases. It is accessed by selecting the **Weight Cases...** option on the **Data** menu. This selection loads the following **Weight Cases** dialog box.



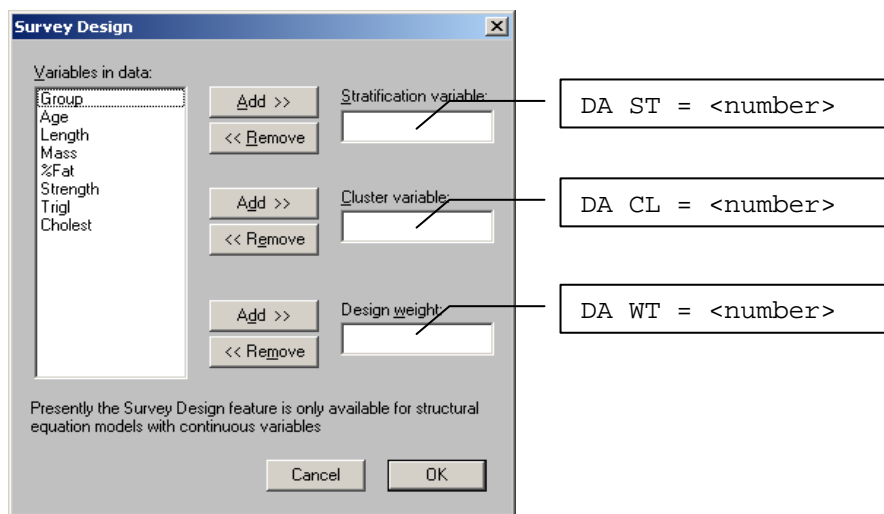
Note that the **Weight Cases** dialog box corresponds with the WE command as indicated on the image above.

The user specifies the weight variable by first selecting the label from the variable list box and then by clicking on the **Add** button to add the variable to the WE grid box.

Once the user is done with the **Weight Cases** dialog box, click on the **OK** button to return to the PSF window. Save the changes by selecting the **Save** option on the **File** menu.

The Survey Design dialog box

The **Survey Design** dialog box allows the user to define the stratification, cluster and design weight variables of a complex survey design. It is accessed by selecting the **Survey Design...** option on the **Data** menu. This selection loads the following **Survey Design** dialog box.



Note that the **Survey Design** dialog box corresponds with the DA command as indicated on the image above.

The user specifies the stratification variable by first selecting the corresponding label from the **Variables in data** list box and then by clicking on the upper **Add** button to add the variable to the **Stratification variable** grid box.

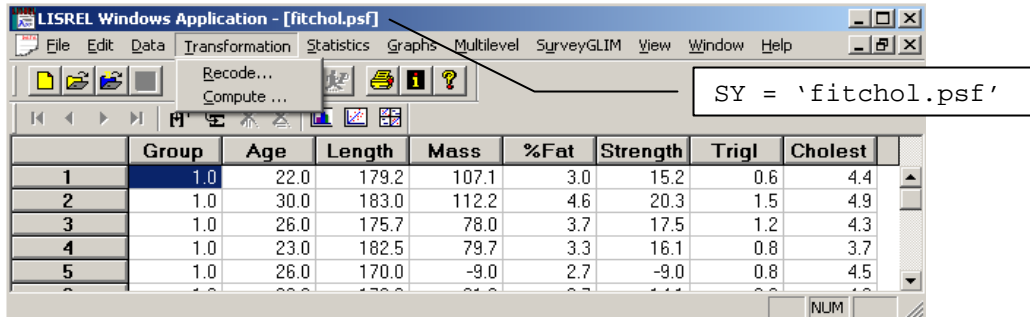
The cluster variable is specified by first selecting the desired label from the **Variables in data** list box and then by clicking on the middle **Add** button to add the variable to the **Cluster variable** grid box.

The weight variable is specified by first selecting the appropriate label from the **Variables in data** list box and then by clicking on the bottom **Add** button to add the variables to the **Design weight** grid box.

Once the user is done with the **Survey Design** dialog box, click on the **OK** button to return to the PSF window. Save the changes by selecting the **Save** option on the **File** menu.

The Transformation menu

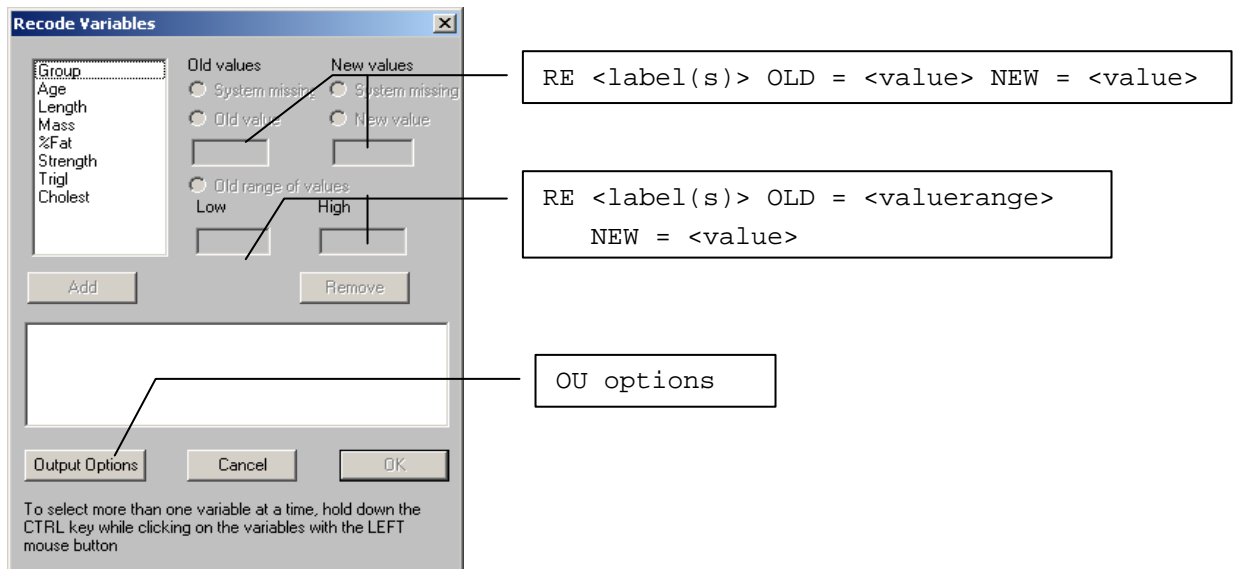
The **Transformation** menu is located on the PSF window of LISREL for Windows which is used to display, manipulate and process raw data. In other words, one needs to create a PSF and open it in a PSF window before using any option of the **Transformation** menu interactively. To illustrate this, the PSF window for the file **fitchol.psf** in the **TUTORIAL** subfolder with the **Transformation** menu expanded is shown below.



The **Transformation** menu allows the user to perform basic data manipulations. The **Recode** option gives the user access to the **Recode Variables** dialog box while the **Compute** option provides the user access to the **Compute** dialog box. Each of these dialog boxes is discussed separately in the next two sections.

The Recode Variables dialog box

The **Recode Variables** dialog box allows the user to recode the values or the range of values of one or more variables. It is accessed by selecting the **Recode...** option on the **Transformation** menu. This selection loads the following **Recode Variables** dialog box.



Note that the **Recode Variables** dialog box corresponds with the RE and OU commands as indicated on the image above.

The user can recode a system missing value to a new value by first selecting the desired label from the variable list box to activate all the radio buttons. Then, activate the **System missing** radio button in the **Old values** section. Once done, the string field of old values is de-activated and the **New value** radio button is activated automatically. Next, enter the desired value in the new value string box. Finally, click on the **Add** button to add a recode command line to the RE grid box. Similarly, the user can recode an old value to a new value or an old value to a system missing value.

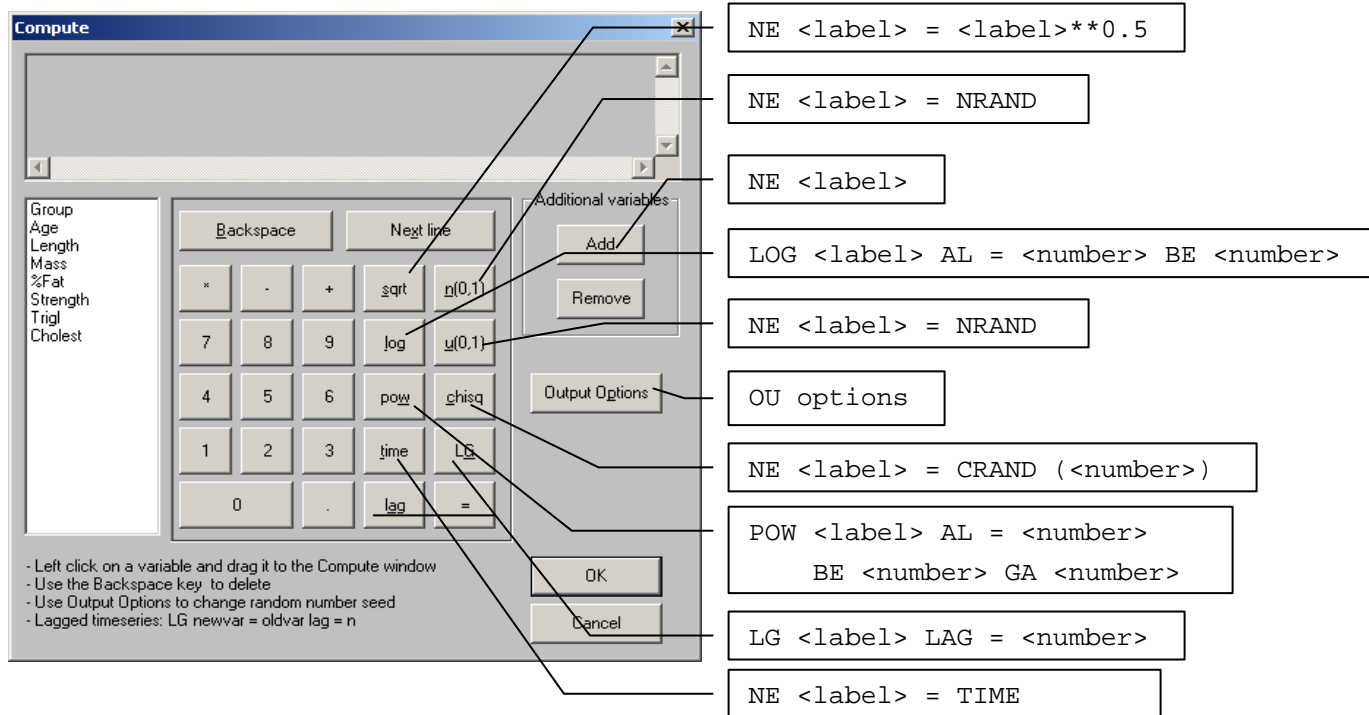
PRELIS also enables the user to recode a range of values of a variable to a new value. This is achieved by first activating the **Old range of values** radio button. Specify the old range of the values by entering the lower bound in the **Low** string box and the upper bound in the **High** string box. Next, activate the **New value** radio button and enter the desired value in the corresponding string box. Finally click on the **Add** button to add a recode command line to the RE grid box.

The **Output Options** button provides the user access to the **Output** dialog box, which is used to specify the desired results to be written to the output file. For example, it may be used to save the recoded data to a new PSF (see the **Output** dialog box section).

Once the user is done with the **Recode Variables** dialog box, click on the **OK** button to run PRELIS.

The Compute dialog box

The **Compute** dialog box allows the user to compute the values of new variables and to transform variables. The **Compute** dialog box is accessed by selecting the **Compute...** option on the **Transformation** menu. This selection loads the following **Compute** dialog box.



Note that the **Compute** dialog box corresponds with the NE, LOG, POW, LG and OU commands as indicated on the image above.

The functions of the functional buttons on the **Compute** dialog box are summarized in the following table.

Button	Function
sqrt	Computes the square root of a positive number.
log	Computes the natural logarithm of a positive number.
pow	Computes the power value of a linear combination.
time	Generates the time points of the observations.
LG & lag	Computes the values of a variable lagged for a given number of points in time.
n(0,1)	Generates a random value from a standard Normal distribution.
u(0,1)	Generates a random value from a Uniform distribution on the interval (0,1).
chisq	Generates a random value from a Chi-square distribution with a given degrees of freedom.

The user can click on any computation function button of the **Compute** dialog box and/or select and drag the labels from the variable list box into the top grid box to specify the desired computation or transformation. Note that the keyboard can not be used for the **Compute** dialog box.

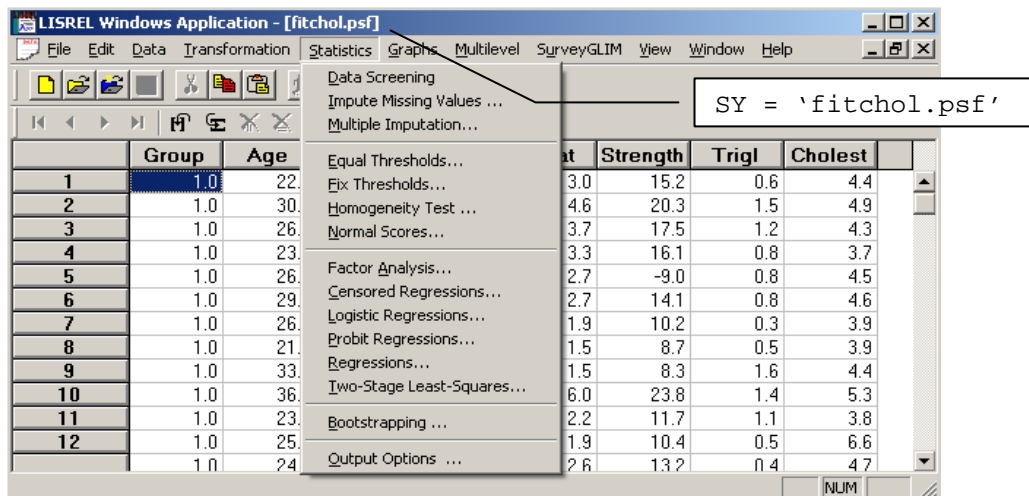
The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file. For example, it can be used to save the transformed

data to a new PSF (see the **Output** dialog box section).

Once the user is done with the **Recode Variables** dialog box, click on the **OK** button to run PRELIS.

The Statistics menu

The **Statistics** menu is located on the PSF window of LISREL for Windows which is used to display, manipulate and process raw data. In other words, one needs to create a PSF and open it in a PSF window before using any option of the **Statistics** menu interactively. To illustrate this, the PSF window for the file **fitchol.psf** in the **TUTORIAL** subfolder with the **Statistics** menu expanded is shown below.



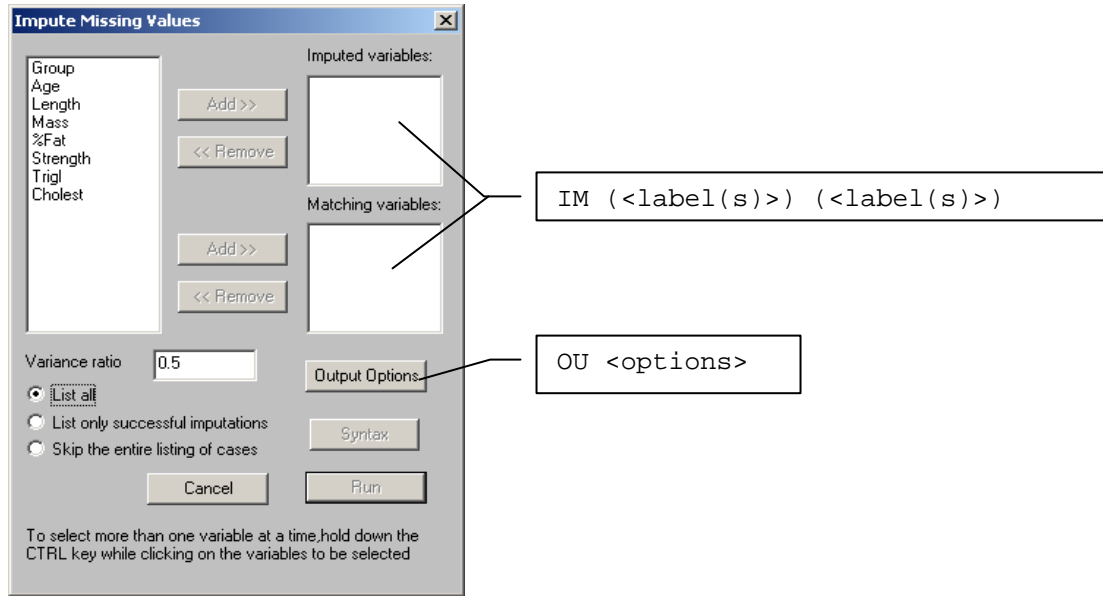
The **Statistics** menu provides the user access to fifteen data analysis options. Except for the **Data Screening** option, each of these options gives the user access to a dialog box. Each of these dialog boxes is discussed separately in the subsequent sections.

Note

The **Data Screening** option is not associated with a dialog box. Instead, it directly generates a PRELIS output file containing descriptive statistics for the ordinal and continuous variables in the opened PSF.

The Impute Missing Values dialog box

The **Impute Missing Values** dialog box allows the user to impute the missing data values by matching. It is accessed by selecting the **Impute Missing Values...** option on the **Statistics** menu. This selection loads the following **Impute Missing Values** dialog box.



Note that the **Impute Missing Values** dialog box corresponds with the IM and OU commands as indicated on the image above.

The user can specify the variable(s) to be imputed by first selecting the label(s) from the variable list box, then clicking on the upper **Add** button to add the variables to the **Imputed variables** list box. Similarly, the matching variables can be selected by using the lower **Add** button.

The default **Variance ratio** is 0.5. The user can specify any number between 0 and 0.5 if 0.5 is not desired.

In the output file, all the imputations are listed by default. The user can request a list of only the successful imputations by activating the **List only successful imputations** radio button. If no imputations are to be listed, the user can activate the **Skip entire listing of cases** radio button.

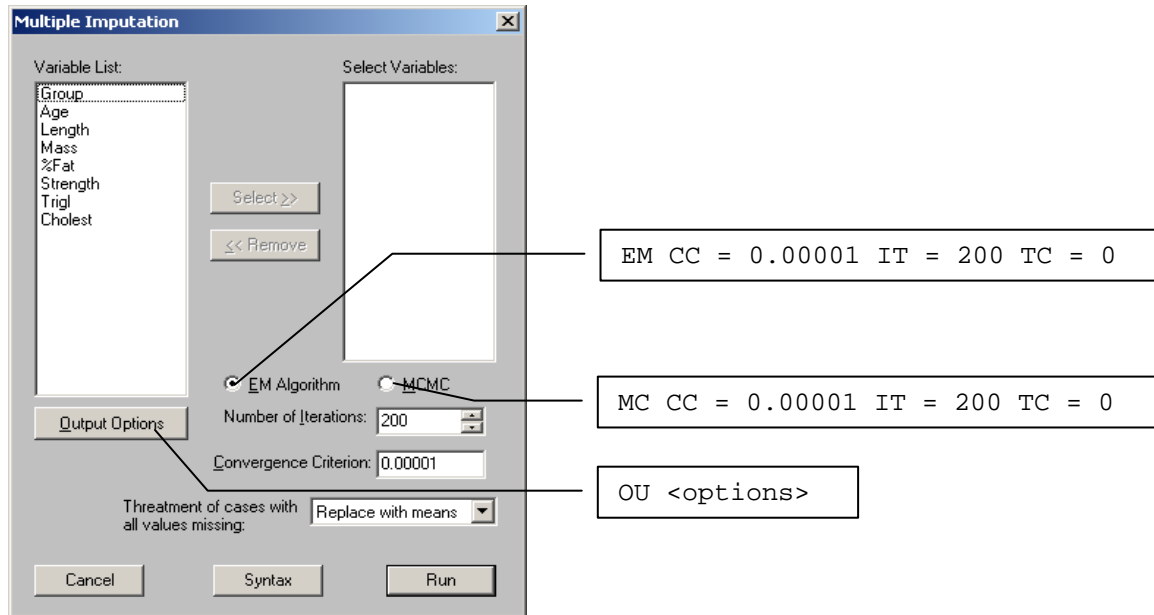
The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Impute Missing Values** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Multiple Imputation dialog box

The **Multiple Imputation** dialog box allows the user to impute the missing values by using the Expected Maximization (EM) or the Monte Carlo Markov Chain (MCMC) method. These methods are described in Du Toit & Du Toit (2001).

The **Multiple Imputation** dialog box is accessed by selecting the **Multiple Imputation...** option on the **Statistics** menu. This selection loads the following **Multiple Imputation** dialog box.



Note that the **Multiple Imputation** dialog box corresponds with the EM, MC and OU commands as indicated on the image above.

The default method for multiple imputation is the EM algorithm. If this method is not desirable, the user can specify the MCMC method by activating the **MCMC** radio button.

Both the EM and the MCMC methods use an iterative algorithm to impute the missing values. The user can enter the maximum number of iterations in the **Number of Iterations** number field if the default of 200 is not appropriate. Enter the appropriate convergence criterion in the **Convergence Criterion** number field if the default value of 0.0001 is not to be used.

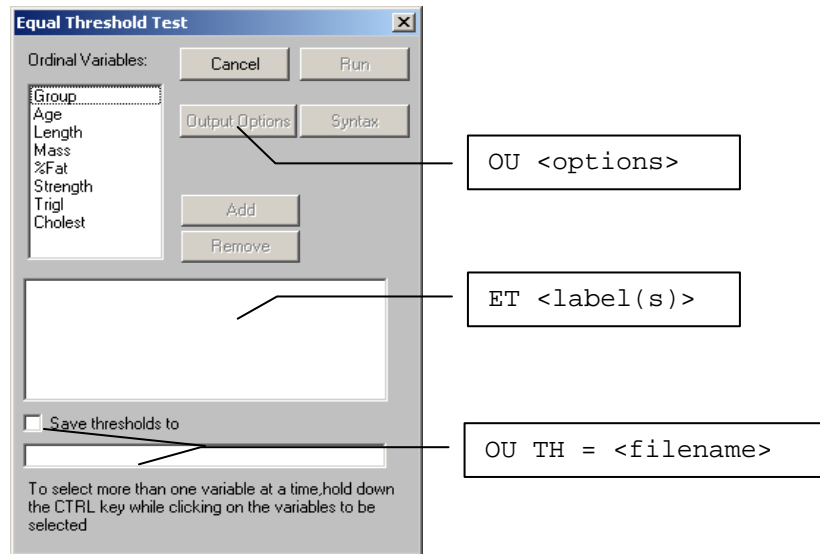
By default, the completely missing cases are replaced with the mean case for the variables. Alternatively, the user may either keep these cases as missing or delete them by selecting the desired option from the **Treatment of cases with all values missing** drop-down list box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Multiple Imputation** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Equal Threshold Test dialog box

The **Equal Threshold Test** dialog box allows the user to perform the equal threshold test for ordinal variables (Jöreskog & Sörbom 1999). It is accessed by selecting the **Equal Thresholds...** option on the **Statistics** menu. This selection loads the following **Equal Threshold Test** dialog box.



Note that the **Equal Threshold Test** dialog box corresponds with the ET and OU commands as indicated on the image above.

At least two ordinal variables should be selected to perform the equal threshold test. Once the desired variables are selected, click on the **Add** button to add an ET command line to the ET grid box.

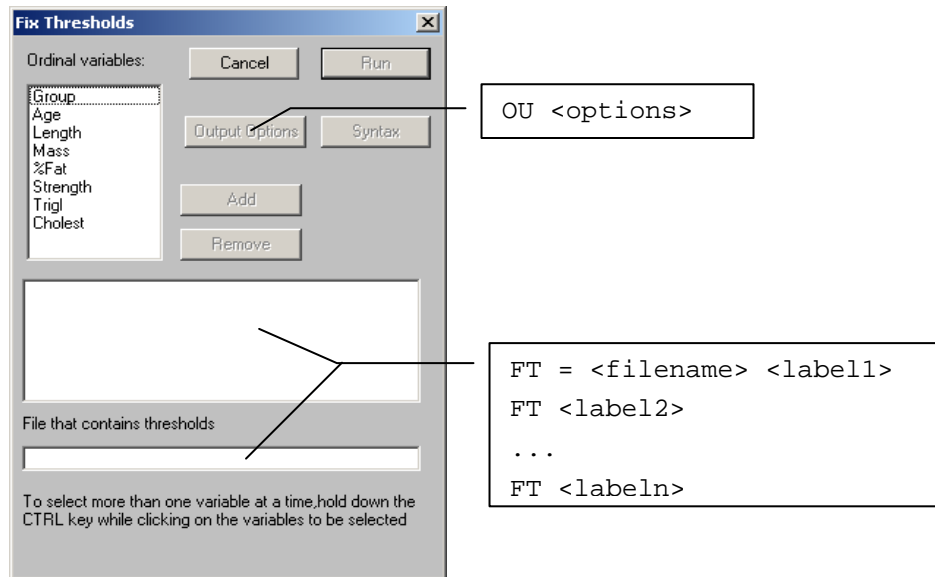
The user can save the thresholds to a separate text file by checking the **Save thresholds to** check box and entering a desired file name in the corresponding string box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Equal Threshold Test** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Fix Thresholds dialog box

The **Fix Thresholds** dialog box allows the user to fix the thresholds for ordinal variables (Jöreskog & Sörbom 1999). It is accessed by selecting the **Fix Thresholds...** option on the **Statistics** menu. This selection loads the following **Fix Thresholds** dialog box.



Note that the **Fix Thresholds** dialog box corresponds with the FT and OU commands as indicated on the image above.

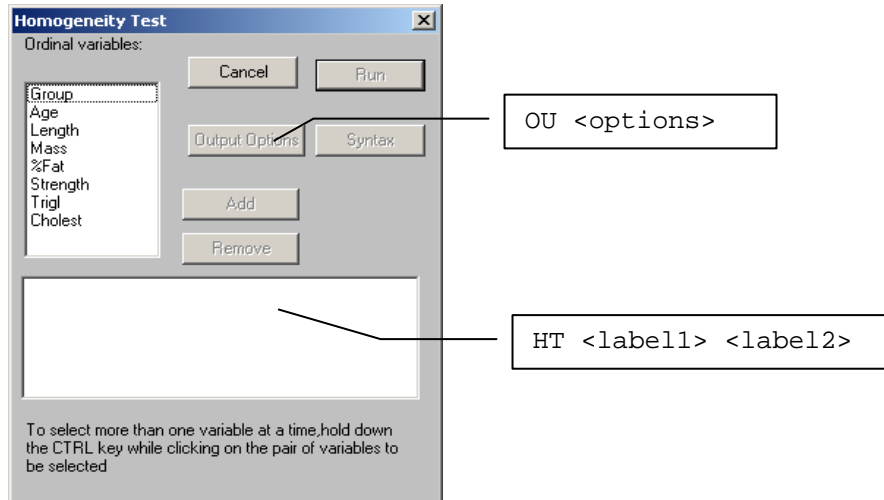
Select the desired variables, and click on the **Add** button to write the corresponding commands to the FT grid box. Before the thresholds can be fixed, a separate text file that contains all the threshold values needs to be prepared. Specify this text file by entering the file name in the **File that contains thresholds** string box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Fix Thresholds** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Homogeneity Test dialog box

The **Homogeneity Test** dialog box allows the user to perform the homogeneity test, which is a test of the hypothesis that the marginal distributions of two categorical variables with the same number of categories are the same (Jöreskog & Sörbom 1999). It is accessed by selecting the **Homogeneity Test...** option on the **Statistics** menu. This selection loads the following **Homogeneity Test** dialog box.



Note that the **Homogeneity Test** dialog box corresponds with the HT and OU commands as indicated on the image above.

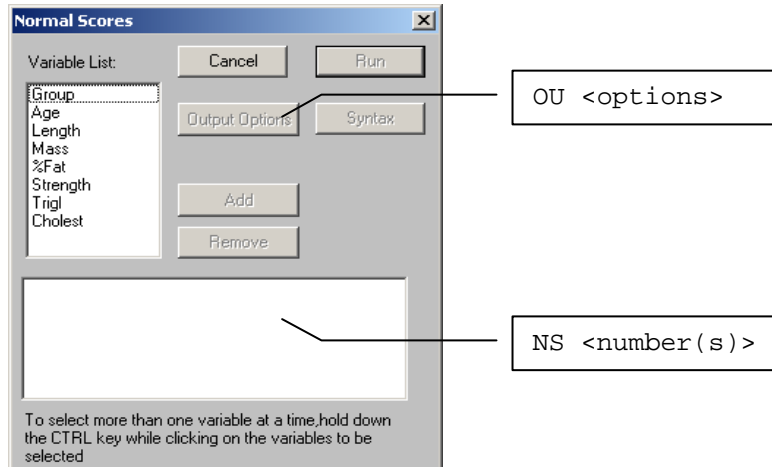
A pair of variables must be selected at a time to activate the **Add** button. Once the desired variables are selected, click on the **Add** button to write a corresponding command line to the HT grid box. Repeat these actions for all the homogeneity tests to be performed.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Homogeneity Test** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Normal Scores dialog box

Normal scores offer an effective way of normalizing a variable. They may be computed for ordinal and continuous variables (Jöreskog et al. 2001) by using the **Normal Scores** dialog box. The **Normal Scores** dialog box is accessed by selecting the **Normal Scores...** option on the **Statistics** menu. This selection loads the following **Normal Scores** dialog box.



Note that the **Normal Scores** dialog box corresponds with the NS and OU commands as indicated on the image above.

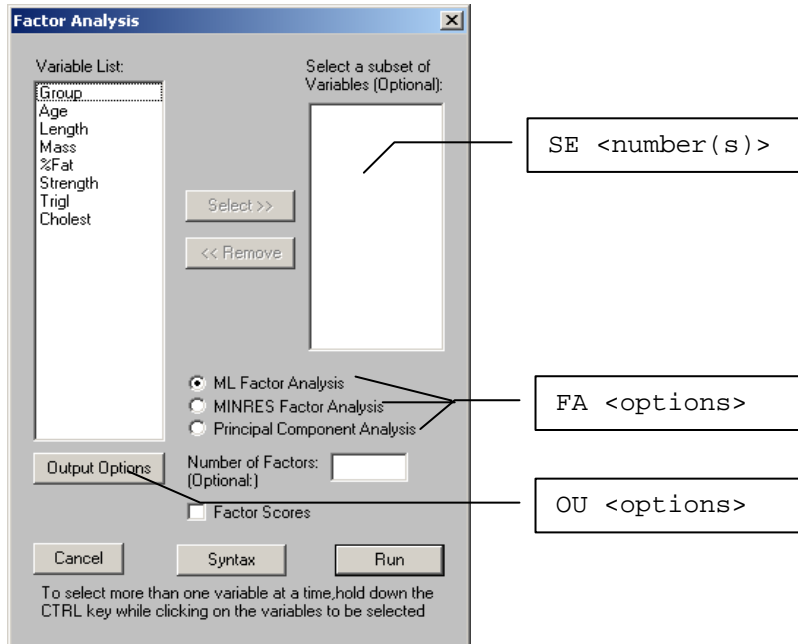
Once the desired variables are selected, click on the **Add** button to write a corresponding command line to the NS grid box. Variables can be added separately or as groups of variables.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Normal Scores** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Factor Analysis dialog box

The **Factor Analysis** dialog box allows the user to perform an exploratory factor analysis (Jöreskog et al. 2001). It is accessed by selecting the on the **Factor Analysis...** option on the **Statistics** menu. This selection loads the following **Factor Analysis** dialog box.



Note that the **Factor Analysis** dialog box corresponds with the SE, FA and OU commands as indicated on the image above.

The user can specify the variables to be used for the factor analysis by first selecting the labels from the **Variable List** list box and then by clicking on the **Select** button to add the variables to the **Select a subset of Variables** list box. This step is optional. If no variable is selected, all the variables in the data set are used for the analysis.

There are three types of exploratory analyses available, namely ML Factor Analysis, MINRES Factor Analysis and Principal Component Analysis. The default analysis is ML Factor Analysis. The user can specify the other methods by activating either the **MINRES Factor Analysis** or the **Principal Component Analysis** radio button.

By default, PRELIS chooses the number of factors. However, the user can specify the desired number of factors by entering a number in the **Number of factors** string box.

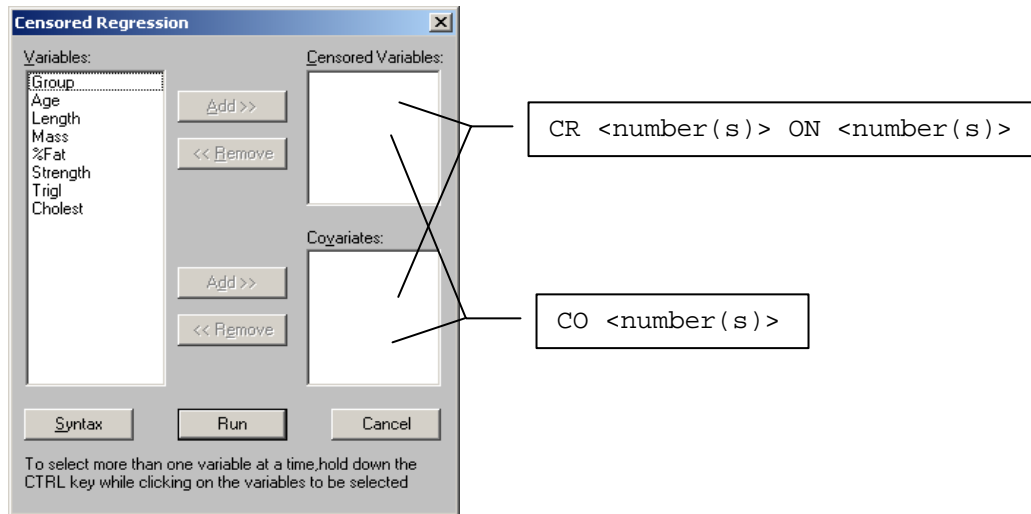
The factor scores are not printed in the output file by default. The user can request the factor scores to be written to the output file by checking the **Factor Scores** check box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Factor Analysis** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Censored Regression dialog box

The **Censored Regression** dialog box allows the user to perform a censored regression analysis (Jöreskog 2002b, 2004). It is accessed by selecting the **Censored Regressions...** option on the **Statistics** menu. This selection loads the following **Censored Regression** dialog box.



Note that the **Censored Regression** dialog box corresponds with the CR and CO commands as indicated on the image above.

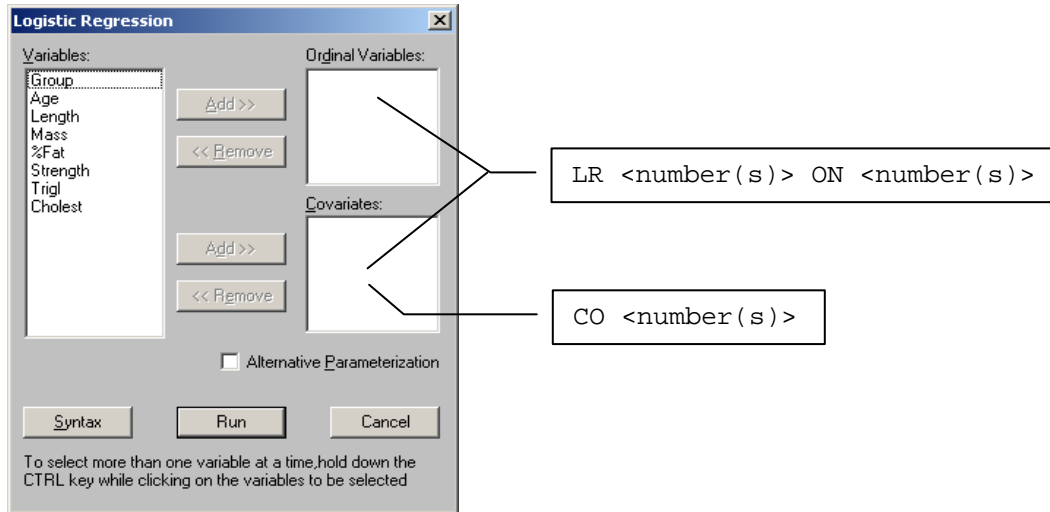
The user can specify the dependent censored variables by first selecting the labels from the **Variables** list box and then by clicking on the upper **Add** button to add the variables to the **Censored Variables** list box.

The independent variables are specified by first selecting the labels from the **Variables** list box and then by clicking on the lower **Add** button to add the variables to the **Covariates** list box.

Once the user is done with the **Censored Regression** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Logistic Regression dialog box

The **Logistic Regression** dialog box allows the user to perform a logistic regression analysis (Jöreskog 2002a). It is accessed by selecting the **Logistic Regressions...** option on the **Statistics** menu. This selection loads the following **Logistic Regression** dialog box.



Note that the **Logistic Regression** dialog box corresponds with the LR and CO commands as indicated on the image above.

Before performing a logistic regression analysis, the type of the dependent variables must be correctly defined as ordinal (see the **Define Variables** dialog box section).

The user can specify the ordinal dependent variables by first selecting the labels from the **Variables** list box and then by clicking on the upper **Add** button to add the variables to the **Ordinal Variables** list box.

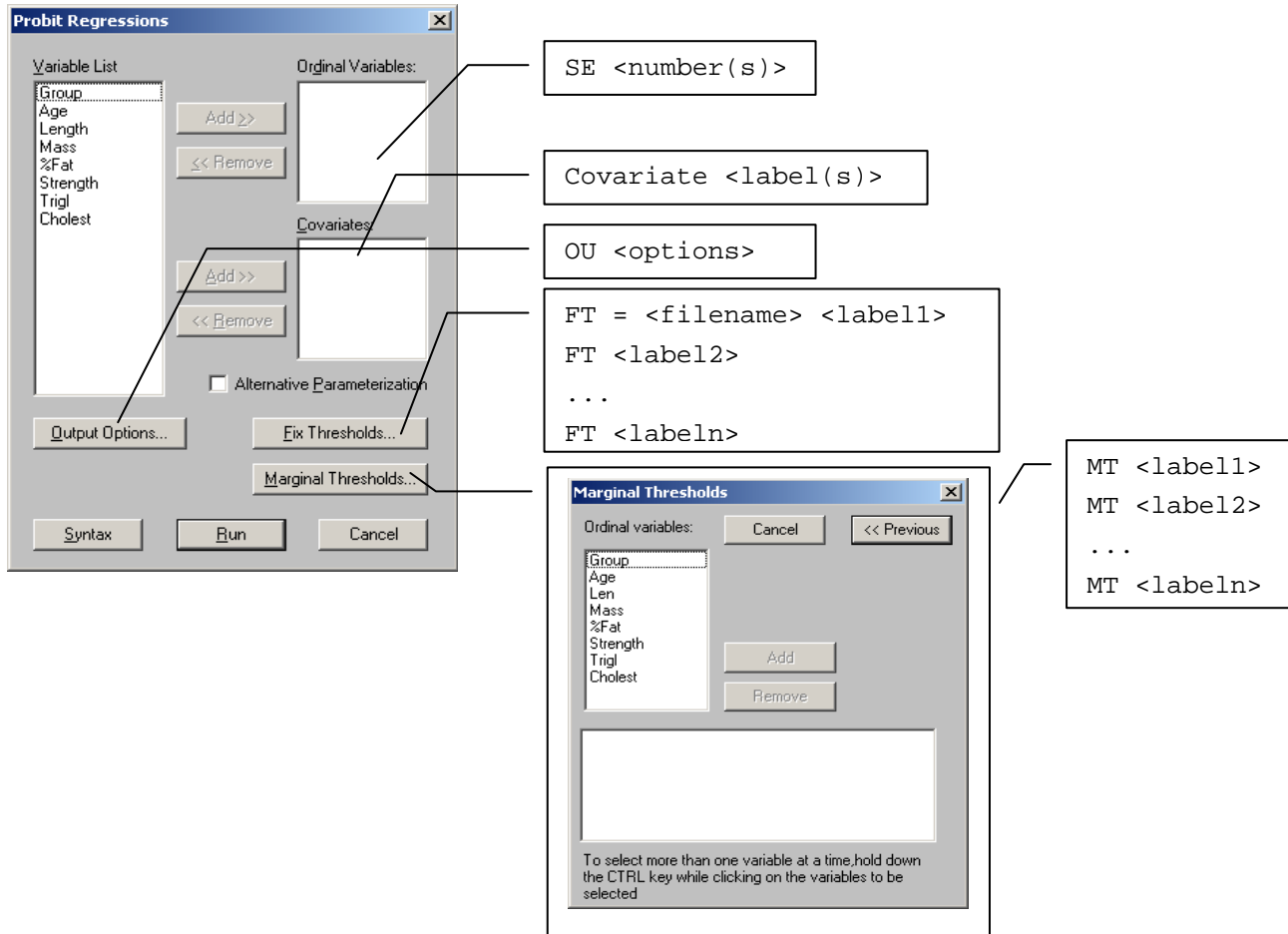
The independent variables are specified by first selecting the labels from the **Variables** list box and then by clicking on the lower **Add** button to add the variables into the **Covariates** list box.

The user may request the alternative parameterization method (Jöreskog 2002a) by checking the **Alternative Parameterization** check box.

Once the user is done with the **Logistic Regression** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Probit Regressions dialog box

The **Probit Regressions** dialog box allows the user to perform a probit regression analysis (Jöreskog 2002a). It is accessed by selecting the **Probit Regressions...** option on the **Statistics** menu. This selection loads the following **Probit Regressions** dialog box.



Note that the **Probit Regressions** dialog box corresponds with the SE, Covariate, OU and FT commands as indicated on the image above.

Before performing a probit regression analysis, the type of the dependent variables must be correctly defined as ordinal (see the **Define Variables** dialog box section).

The user can specify the ordinal dependent variables by first selecting the labels from the **Variable List** list box and then by clicking on the upper **Add** button to add the variables to the **Ordinal Variables** list box.

The independent variables are specified by first selecting the labels from the **Variable List** list box and then by clicking on the lower **Add** button to add the variables to the **Covariates** list box.

The user may request the alternative parameterization method (Jöreskog 2002a) by checking the **Alternative Parameterization** check box.

The **Fix Thresholds** button allows the user to access the **Fix Thresholds** dialog box (see the **Fix**

Thresholds dialog box section).

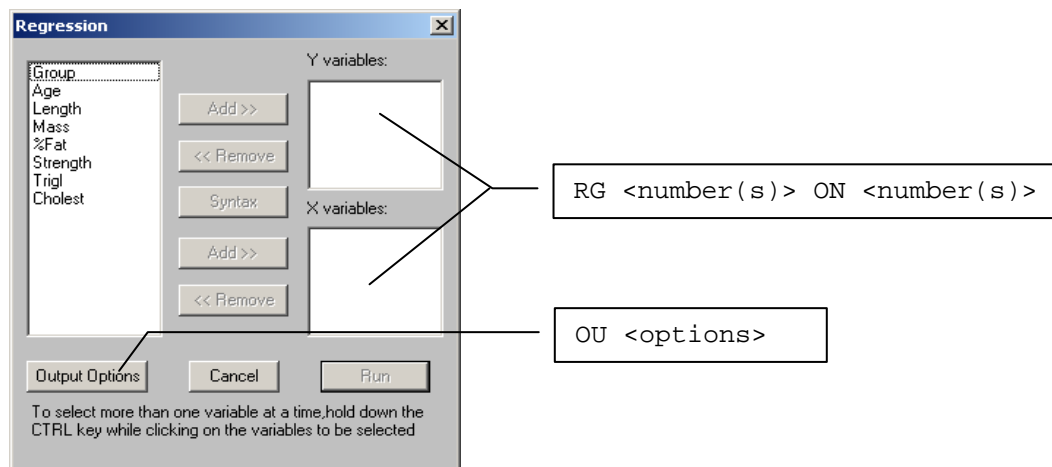
A click on the **Marginal Thresholds** button loads the **Marginal Thresholds** dialog box. Note that the **Marginal Thresholds** dialog box corresponds with the MT command as shown on its image. It allows the user to request marginal thresholds for the desired ordinal variables by first selecting the labels from the **Ordinal Variables** list box and then by clicking on the **Add** button to add the corresponding commands to the MT grid box. Once done, click on the **Previous** button to return to the **Probit Regressions** dialog box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Probit Regression** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Regression dialog box

The **Regression** dialog box allows the user to perform a multiple linear regression analysis. It is accessed by selecting the **Regressions...** option on the **Statistics** menu. This selection loads the following **Regression** dialog box.



Note that the **Regression** dialog box corresponds with the RG and OU commands as indicated on the image above.

The user can specify the dependent variables by first selecting the labels from the variable list box and then by clicking on the upper **Add** button to add the variables to the **Y variables** list box.

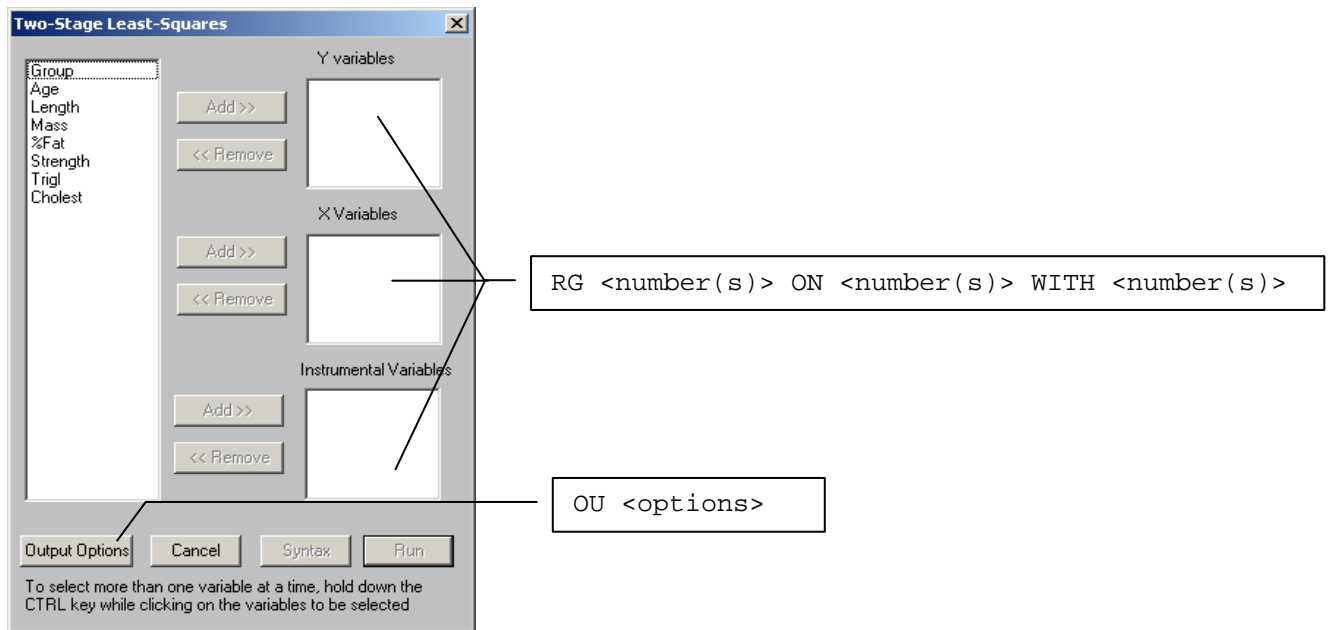
The independent variables are specified by first selecting the labels from the variable list box and then by clicking on the lower **Add** button to add the variables to the **X variables** list box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Regression** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Two-Stage Least-Squares dialog box

The **Two-Stage Least-Squares** dialog box is used to apply the two-stage least-square method (Jöreskog et al. 2001). It is accessed by selecting the **Two-Stage Least-Squares...** option on the **Statistics** menu. This selection loads the following **Two-Stage Least-Squares** dialog box.



Note that the **Two-Stage Least-Squares** dialog box corresponds with the RG and OU commands as indicated on the image above.

The user can specify the dependent variables by first selecting the labels from the variable list box and then by clicking on the upper **Add** button to add the variables to the **Y Variables** list box.

The independent variables are specified by first selecting the labels from the variable list box and then by clicking on the middle **Add** button to add the variables to the **X Variables** list box.

The instrumental variables are specified by first selecting the labels from the variable list box and then by clicking on the bottom **Add** button to add the variables to the **Instrumental Variables** list

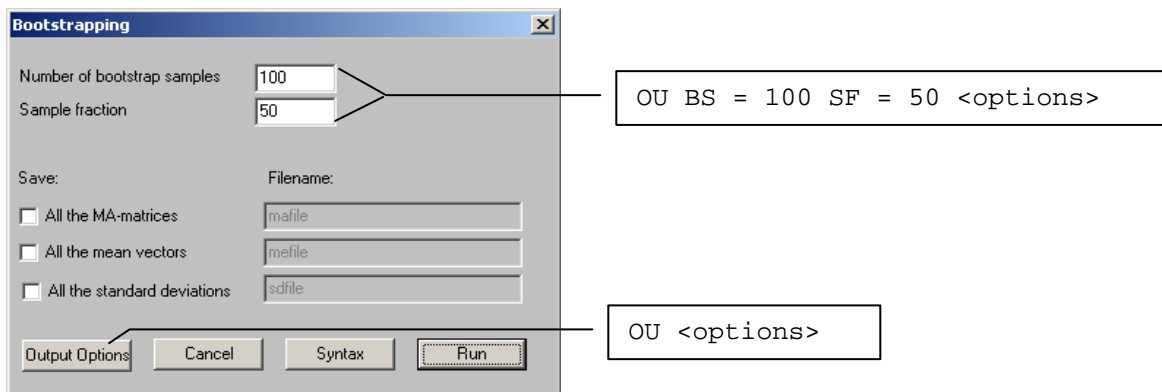
box.

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Two-Stage Least-Squares** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Bootstrapping dialog box

The **Bootstrapping** dialog box allows the user to draw random samples with replacement from an original sample (Jöreskog & Sörbom 1999). It is accessed by selecting the **Bootstrapping...** option on the **Statistics** menu. This selection loads the following **Bootstrapping** dialog box.



Note that the **Bootstrapping** dialog box corresponds with the OU command as indicated on the image above.

The default number of bootstrap samples is 100. The user can specify another number by entering the desired number in the **Number of bootstrap samples** string box. By default, the sample fraction is 50 percent. It can be changed by entering another fraction in the **Sample fraction** string box.

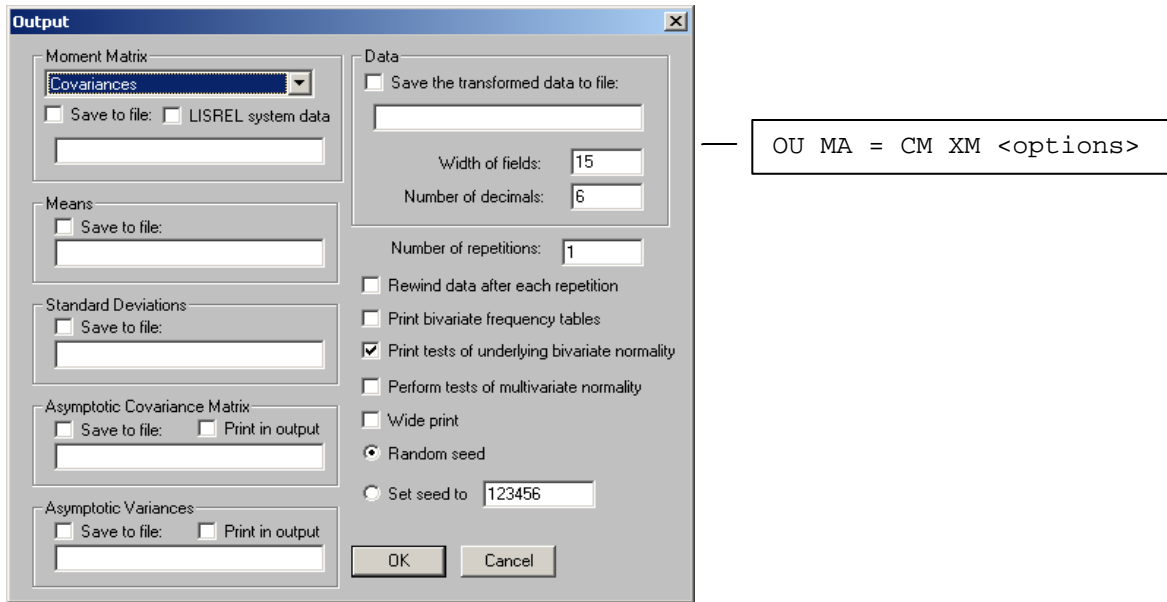
PRELIS provides the option to save the matrices specified in the MA keyword on the OU command line, the mean vectors and the standard deviation vectors to a separate text file(s). This is obtained by checking the **All the MA-matrices**, **All the mean vectors** and/or **All the standard deviations** check box(es), and then entering the desired file name(s) in the **Filename** string box(es).

The **Output Options** button provides access to the **Output** dialog box, which is used to specify the desired results to be written to the output file (see the **Output** dialog box section).

Once the user is done with the **Bootstrapping** dialog box, one may either request the PRELIS syntax file by clicking on the **Syntax** button or run PRELIS directly by clicking on the **Run** button.

The Output dialog box

The **Output** dialog box allows the user to compute sample means, standard deviations, covariance (correlation) matrices and estimated asymptotic covariance matrices of sample variances and covariances (correlations). It also allows the user to save these results and/or the transformed data to file and to request specific results to be written to the output file. It can be accessed by clicking on the **Output Options** button on most of the other dialog boxes associated with the options on the **Statistics** menu or by selecting the **Output Options...** option on the **Statistics** menu. These selections load the following **Output** dialog box.



Note that the **Output** dialog box corresponds with the OU command as indicated on the image above.

The **Moment Matrix** section of the **Output** dialog box enables the user to specify the desired moment matrix to be computed. This is accomplished by selecting the desired moment matrix from the drop-down list box. If it is to be saved to file, either check the **Save to file** check box if a separate text file is wanted or check the **LISREL system data** check box if a data system file (DSF) is desired. Then, enter a name in the corresponding string box.

By default, the means of the variables in the data file are written to the output file, but not saved to a separate text file. The **Means** section of the **Output** dialog box enables the user to save the sample means of the variables to a separate text file by first checking the **Save to file** check box and then by entering a name in the corresponding string box.

By default, the standard deviations of the variables in the data file are printed in the output file, but not saved in a separate text file. The **Standard Deviations** section of the **Output** dialog box allows the user to save the sample standard deviations of the variables to a separate text file by first checking the **Save to file** check box and then by entering a name in the corresponding string box.

The estimated asymptotic covariance matrix of the selected sample moments is not computed by default. The **Asymptotic Covariance Matrix** section of the **Output** dialog box enables the user to compute the estimated asymptotic covariance matrix of the selected sample moments, save it to a binary file and/or print it in the output file. The user checks the **Save to file** check box if a separate binary file is desired and then enter a name in the corresponding string box. If the matrix is to be written to the output file, the user needs to check the **Print in output** check box.

The estimated asymptotic variances of the selected sample moments are not computed by default. The **Asymptotic Variances** section of **Output** dialog box allows the user to compute the estimated asymptotic variances of the sample moments, save it to a binary file and/or write it to the output file. If a separate binary file is desired, check the **Save to file** check box and then enter a name in the corresponding string box. These estimated variances are written to the output file if the **Print in output** check box is checked.

The **Data** section allows the user to save the transformed data to file. To save the transformed data to file, first check the **Save the transformed data to file** check box and then enter a name in the corresponding string box. If the extension of the file name is PSF, a PSF is created. Otherwise, a text file is created. The default width of the column fields is 15 and the default number of decimals is 6. Other numbers can be used by simply enter the desired numbers in the **Width of field** and **Number of decimals** string boxes respectively.

The default number of repetitions is 1. If there are more repetitions, the number of repetitions is specified by entering the appropriate number in the **Number of repetitions** string box.

The rewinding of the data to the first observation after each repetition is not done by default. However, the user may request it by checking the **Rewind data after each repetition** check box.

The bivariate frequency tables are not written to the output file by default. They are requested by checking the **Print bivariate frequency tables** check box.

By default, the underlying bivariate normality test results are written to the output file. If this is not desired, the user may un-check the **Print tests of underlying bivariate normality** check box.

One may request the output file to be formatted to a width of 132 columns instead of 80 columns by checking the **Wide print** check box.

By default, the random seed 123456 is used for the bootstrapping and Monte Carlo methods for simulation. The user can set the seed to a fixed number by activating the **Set seed to** radio button and then by entering an integer in the corresponding string box.

Once the user is done with the **Output** dialog box, click on the **OK** button to either return to the previous dialog box (if applicable) or to run PRELIS.

PRELIS Syntax

The structure of the PRELIS syntax file

The PRELIS syntax file, which is generated by the graphical user interface of the PRELIS module, can also be prepared by the user by using the LISREL for Windows text editor or any other text editor such as Notepad and Wordpad. The general structure of the PRELIS syntax file depends on the data to be processed. If the raw data file to be processed is a PSF, the PRELIS syntax file has the following structure.

```
TI
<string>
SY= <psfname>
<commands>
OU <options>
```

where <string> denotes a character string, <psfname> denotes the complete name (including the drive and folder names) of the PSF and <commands> denotes a set of optional PRELIS commands to specify data manipulations or basic statistical analyses and <options> denotes a list of options for the analysis each of which either has the syntax:

```
<keyword> = <selection>
```

where <keyword> is one of AC, AM, BM, BS, CM, IX, KM, MA, ME, ND, NP, OM, PM, RA, RM, SA, SD, SF, SM, SR, TH, TM or WI and <selection> denotes a number, a value or a name (see the OU command section) or the syntax:

```
<option>
```

where <option> is one of PA, PV, WP, XB, XM or XT (see the OU command section).

The PRELIS syntax file has the following structure if the data file to be processed is in the form of a text file.

```
TI
<string>
DA <data-specifications>
LA
<labels>
```

RA= <filename>
<commands>
OU <options>

where <data-specifications> has the syntax

<keyword> = <selection>

where <keyword> is one of CL, MI, NI, NO, RP, ST, TR or WT and <selection> denotes a number, a value or a name (see the DA command section), <filename> denotes the complete name (including the drive and folder names) of a text data file and <labels> denotes the labels of the variables in the text data file.

The SY command is a required command only if a PSF is used. If the data to be analyzed are not provided as a PSF, the DA command, the LA paragraph (or LA command) and the RA command are required. In this case, the DA command should be the first command after the titles and comments. The OU command is a required command and should be the final command in the PRELIS syntax file. The remaining PRELIS commands are all optional and can be entered in any order.

In the following sections, the PRELIS commands and paragraphs are discussed separately in alphabetical order.

CA command

The CA command is used to specify the variable type of a variable(s) as censored above. It is an optional command.

Syntax

CA <labels>
or
CA <numbers>

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <numbers> is a list of variable positions in the raw data file in free or abbreviated format.

Examples

CA REPAIRS

CB command

The purpose of the CB command is to specify the variable type of a variable(s) as censored below. It is an optional command.

Syntax

CB <labels>

or

CB <numbers>

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <numbers> is a list of variable positions in the raw data file in free or abbreviated format.

Examples

CB AFFAIRS

CB 11 27

CE command

Variables that are censored above and below are specified by using the CE command. It is an optional command.

Syntax

CE <labels>

or

CE <numbers>

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <numbers> is a list of variable positions in the raw data file in free or abbreviated format.

Examples

CE WAGES

CE 1 9

CL command

The purpose of the CL command is to assign category labels to the categories of ordinal variables. It is an optional command.

Syntax

CL <labels> <assignments>

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <assignments> is a list of assignment statements for the category labels each of which has the syntax

<number>=<label>

where <number> denotes a value of the ordinal variable(s) and <label> denotes the corresponding category label.

Example

CL NOSAY – INTEREST 1=AS 2=A 3=D 4=DS 8=DK 9=NA

CO command

The CO command is used to specify the variable type of a variable(s) as continuous. It is an optional command.

Syntax

CO <labels>

or

CO <numbers>

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <numbers> is a list of variable positions in the raw data file in free or abbreviated format.

Examples

CO JOBSEC AGE WEIGHT

CO 3 4 12 29

DA command

The DA command is used to specify the number of variables, the number of observations, the missing values and their treatment, the number of repetitions, the clustering variable, the stratification variable and the design weight variable for the data to be processed. If the data to be processed is in the form of a text data file, the DA command is required.

Syntax

DA <data-specifications>

where <data-specifications> has the syntax

<keyword> = <selection>

where <keyword> is one of CL, MI, NI, NO, RP, ST, TR or WT and <selection> denotes a number, a value or a name.

Examples

DA NI=12 NO=579 CL=3 ST=4 WT=11 MI=-9

DA NI=37

DA NI=76 TR=PA

Note

Except for the NI keyword, the keywords of the DA command are optional.

CL keyword

The purpose of the CL keyword is to specify the clustering variable of a complex survey data set.

Syntax

CL=<number>

where <number> denotes the column position of the clustering variable in the raw data file.

MI keyword

The MI keyword is used to specify the global missing value for the data file.

Syntax

MI=<number>

where <number> denotes the global missing value for the data file.

NI keyword

The purpose of the NI keyword is to specify the number of variables to be read from the raw data file.

Syntax

NI=<number>

where <number> denotes the number of input variables in the raw data file.

NO keyword

The number of observations to be read from the raw data file is specified by using the NO keyword.

Syntax

NO=<number>

where <number> denotes the number of observations in the raw data file.

Default

NO=0

RP keyword

The purpose the RP keyword is used to specify the number of subsets of data in the raw data file to be processed.

Syntax

RP=<number>

where <number> denotes the number of data subsets in the raw data file.

Default

RP=1

ST keyword

The stratification variable of a complex survey data set is specified by using the ST keyword.

Syntax

ST=<number>

where <number> denotes the column position of the stratification variable in the data file.

TR keyword

The purpose of the TR keyword is to specify how the missing values in the data should be treated.

Syntax

TR=<treatment>

where <treatment> in one of LI for list-wise deletion and PA for pair-wise deletion.

Default

TR=LI

WT keyword

The purpose of the WT keyword is to specify the variable that contains the design weights for the observations of a complex survey data set.

Syntax

WT=<number>

where <number> denotes the column position of the design weight variable in the data file.

ET command

The purpose of the ET command is to specify an equal thresholds test for ordinal variables. It is an optional command.

Syntax

ET <labels>

OR

ET <numbers>

where <labels> denotes a list of ordinal variable labels in free or abbreviated format and <numbers> denotes a list of the variable positions of ordinal variables in free or abbreviated format.

Example

```
ET VOTING COMPLEX NOCARE
```

FT command

The FT command is used to specify fixed thresholds for ordinal variables. It is an optional command.

Syntax

```
FT=<filename> <labels>
```

```
FT <labels>
```

or

```
FT=<filename> <numbers>
```

```
FT <numbers>
```

where <filename> denotes the name (including drive and folder names) of the text file that contains the fixed threshold values, <labels> denotes a list of ordinal variable labels in free or abbreviated format and <numbers> denotes a list of the variable positions of ordinal variables in free or abbreviated format.

Example

```
FT=USA.TH NOSAY
```

```
FT NOCARE VOTING COMPLEX
```

HT command

The HT command allows the user to specify a homogeneity test for the marginal distributions of a pair of ordinal variables. It is an optional command.

Syntax

HT <labels>

or

HT <numbers>

where <labels> denotes a list of two ordinal variable labels in free format and <numbers> denotes a list of the variable positions of two ordinal variables in free format.

Example

HT NOSAY INTEREST

IM command

The purpose of the IM command is to impute missing values by means of imputation by matching. It is an optional command.

Syntax

IM <ilabels> <mlabels> <options>

or

IM <inumbers> <mnumbers> <options>

where <ilabels> denotes a list of the labels of the variables to be imputed in free or abbreviated format, <mlabels> denotes a list of the labels of the matching variables in free or abbreviated format, <inumbers> denotes the variable positions of the variables to be imputed in free or abbreviated format, <mnumbers> denotes the variable positions of the matching variables in free or abbreviated format and <options> is one of XN if only the successful imputations are to be listed in the output file, XL if no imputations are to be listed in the output file or

VR=<number>

where <number> denotes the cutoff criterion for the variance ratio for imputation.

Examples

IM (NOSAY – INTEREST) (NOSAY – INTEREST)

IM (1 – 9) (10 – 18) VR=0.6

LA command

If the labels of the observed variables are in the form of a text file, the LA command is used to specify descriptive names for the observed variables in the raw data file. If a PSF or an LA command is not used and an LA paragraph is not specified, it is a required command. The labels of the observed variables can also be specified as part of the PRELIS syntax file. In this regard, the LA paragraph rather than the LA command is used (see the LA paragraph section).

Syntax

```
LA=<filename>
```

where <filename> denotes the name of the text file containing the descriptive names of the observed variables in the raw data file.

Example

Suppose that the name of the text file containing the variable labels is **ABUSE.LAB**, which is located in the folder **Projects\ABUSE** on the **C** drive. In this case, the corresponding LA command is given by

```
LA='C:\Projects\ABUSE\ ABUSE.LAB'
```

LA paragraph

The LA paragraph is used to provide descriptive names to the variables in the raw data file. It is a required command, unless a PSF or an LA command is used. If the labels of the observed variables are in the form of a text file, the LA command instead of the LA paragraph is used (see the LA command section).

Syntax

```
LA  
<labels>
```

where <labels> denotes the descriptive names of the observed variables in the raw data file. These names are provided in free or abbreviated format and only the first 8 characters of each name are utilized.

Examples

```
LA  
Age Gender MSCORE SSCORE ESCORE
```

```
LA  
JS1 – JS6 OC1 – OC10
```

LO command

The LO command is used to compute the natural logarithm of a variable or a linear combination of a variable. It is an optional command.

Syntax

```
LO <labels> AL=<number> BE=<number>
```

or

```
LO <numbers> AL=<number> BE=<number>
```

where <labels> denotes a list of variable labels in free or abbreviated format, <number> is a numerical value and <numbers> denotes a list of variable positions in free or abbreviated format.

Example

```
LO INCOME AL=0 BE=1
```

LR command

The user can specify a logistic regression analysis by using the LR command. It is an optional command.

Syntax

```
LR <labels> ON <labels>
```

or

```
LR <numbers> ON <numbers>
```

where <labels> denotes a list of variable labels in free or abbreviated format and <numbers>

denotes a list of variable positions in free or abbreviated format.

Example

```
LR Y1 Y2 ON X1 - X10
```

MI command

If the raw data to be processed include missing values, the MI command is used to specify the missing values. In this case, it is a required command, unless a PSF, which contains all the missing value information, is to be processed.

Syntax

```
MI <values> <labels>
```

where <values> denotes a list of values in free or abbreviated format and <labels> is a list of variable labels in free or abbreviated format.

Example

Suppose that the missing values in a text data file are listed as -100, -99 and -98 for all the variables. For this example, the MI command is

```
MI -100 -99 -98 AGE JOBSAT ITEM1 – ITEM8
```

MT command

The MT command allows the user to compute the marginal thresholds for ordinal variables from their marginal distributions. It is an optional command.

Syntax

```
MT <labels>
```

or

```
MT <numbers>
```

where <labels> denotes a list of ordinal variable labels in free or abbreviated format and <numbers> denotes a list of the variable positions of ordinal variables in free or abbreviated format.

Example

```
MT NOSAY NOCARE VOTING COMPLEX INTEREST
```

NE command

The purpose of the NE command is to create new variables as functions of the variables in the raw data file and to transform the variables in the raw data file. It is an optional command.

Syntax

```
NE <label> =<function>
```

where <label> denotes the label of the new variable and <function> is one of NRAND, URAND, CHISQ(<k>) or a linear combination of the variables where <k> denotes a positive integer.

Examples

```
NE V1=NRAND
```

```
NE X6 = 0.27*V1+0.25*V2+0.38*V8
```

OR command

Ordinal variables are specified by using the OR command. It is an optional command.

Syntax

```
OR <labels>
```

or

```
OR <numbers>
```

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables and <numbers> is a list of variable positions in the raw data file in free or abbreviated

format.

Examples

OR NOSAY – INTEREST

OR ALL

OU command

The OU command is used to request the results to be printed in the PRELIS output file, to specify external files for specific results to be written to and to specify the format of the results. It is a required command.

Syntax

OU <options>

<options> denotes a list of options for the analysis each of which either has the syntax:

<keyword> = <selection>

where <keyword> is one of AC, AM, BM, BS, CM, IX, KM, MA, ME, ND, NP, OM, PM, RA, RM, SA, SD, SF, SM, SR, TH, TM or WI and <selection> denotes a number, a value or a name or the syntax:

<option>

where <option> is one of PA, PV, WP, XB, XM or XT.

Examples

OU RA=TREATMENT.PSF WP=8 ND=3 KM=TREATMENT.KM

OU PM=NIH.PM AC=NIH.ACM

OU NP=6 XM

AC keyword

The purpose of the AC keyword is to specify the name of the binary file for the estimated

asymptotic covariance matrix of the selected sample moments.

Syntax

AC=<filename>

where <filename> denotes the name of the binary file for the estimated asymptotic covariance matrix.

AM keyword

The name of the text file for the augmented moment matrix is specified by using the AM keyword.

Syntax

AM=<filename>

where <filename> denotes the name of the text file for the augmented moment matrix.

BM keyword

The purpose of the BM keyword is to specify the name of the text file for the moment matrices of the bootstrap samples.

Syntax

BM=<filename>

where <filename> denotes the name of the text file for the moment matrices of the bootstrap samples.

BS keyword

The BS keyword is used to specify the number of bootstrap samples to be generated.

Syntax

BS =<number>

where <number> denotes the number of bootstrap samples to be drawn.

CM keyword

The name of the text file for the sample covariance matrix is specified by using the CM keyword.

Syntax

CM=<filename>

where <filename> denotes the name of the text file for the sample covariance matrix.

IX keyword

The IX keyword is used to specify the starting value for the random number generator (random seed).

Syntax

IX =<number>

where <number> denotes the random seed.

KM keyword

The name of the text file for the sample correlation matrix is specified by using the KM keyword.

Syntax

KM=<filename>

where <filename> denotes the name of the text file for the sample correlation matrix.

MA keyword

The MA keyword is used to specify the type of moment matrix to be computed.

Syntax

MA=<matrix>

where <matrix> is one of AM for the augmented moment matrix, CM for the covariance matrix, KM for the correlation matrix, MM for the matrix of sample moments about zero, OM for the matrix of sample correlations based of optimal scores, PM for the matrix of product-moment, poly-choric and poly-serial sample correlations, RM for the Spearman rank correlation matrix and TM for the matrix of Kendall's tau-c sample correlations.

ME keyword

The name of the text file for the sample means is specified by using the ME keyword.

Syntax

ME=<filename>

where <filename> denotes the name of the text file for the sample means.

ND keyword

The ND keyword is used to specify the number of decimal places to be used for writing the transformed data to the external file specified in the RA keyword.

Syntax

ND =<number>

where <number> denotes the number of decimal places.

NP keyword

The NP keyword is used to specify the number of decimal places to be used for writing the results

of a basic statistical analysis to the output file.

Syntax

NP =<number>

where <number> denotes the number of decimal places.

Default

NP=3

OM keyword

The name of the text file for the sample correlation matrix based on optimal scores is specified by using the OM keyword.

Syntax

OM=<filename>

where <filename> denotes the name of the text file for the sample correlation matrix based on optimal scores.

PA option

The estimated asymptotic covariance matrix of the sample moments is listed in the output file if the PA option is specified.

PM keyword

The purpose of the PM keyword is to specify the name of the text file for the matrix of product-moment, poly-choric and poly-serial sample correlations.

Syntax

PM=<filename>

where <filename> denotes the name of the text file for the matrix of product-moment, poly-choric and poly-serial sample correlations.

PV option

The purpose of the PV option is to request that the estimated asymptotic variances of the sample moments are written to the output file.

RA keyword

The RA keyword is used to specify the name of the binary or text file for the transformed raw data

set.

Syntax

RA=<filename>

where <filename> denotes the name of the binary or text file for the transformed raw data. If the extension PSF is used, a binary file is created. Otherwise, a text file is produced.

RM keyword

The name of the text file for the Spearman rank sample correlation matrix is specified by using the RM keyword.

Syntax

RM=<filename>

where <filename> denotes the name of the text file for the Spearman rank sample correlation matrix.

SA keyword

The purpose of the SA keyword is to specify the name of the binary file for the estimated asymptotic covariance matrix of the selected sample moments.

Syntax

SA =<filename>

where <filename> denotes the name of the binary file for the estimated asymptotic covariance matrix.

SD keyword

The name of the text file for the sample standard deviations is specified by using the SD keyword.

Syntax

SD=<filename>

where <filename> denotes the name of the text file for the sample standard deviations.

SF keyword

The SF keyword is used to specify the fraction of the sample size to be used for the sample size of the bootstrap samples to be generated.

Syntax

SF =<number>

where <number> denotes the fraction of the sample size to number of bootstrap samples to be drawn.

SM keyword

The name of the text file for the selected sample moment matrix is specified by using the SM keyword.

Syntax

SM=<filename>

where <filename> denotes the name of the text file for the selected sample moment matrix.

SR keyword

The SR keyword is used to specify the name of the binary or text file for the transformed raw data set.

Syntax

SR=<filename>

where <filename> denotes the name of the binary or text file for the transformed raw data. If the extension PSF is used, a binary file is created. Otherwise, a text file is produced.

TH keyword

The name of the text file for the estimated thresholds for the ordinal variables is specified by using the TH keyword.

Syntax

TH=<filename>

where <filename> denotes the name of the text file for the estimated thresholds for the ordinal variables.

TM keyword

The purpose of the TM keyword is to specify the name of the text file for the Kendall tau-c sample correlation matrix.

Syntax

TM=<filename>

where <filename> denotes the name of the text file for the Kendall tau-c sample correlation matrix.

WI keyword

The WI keyword is used to specify the column width to be used for writing the transformed data to the external file specified in the RA keyword.

Syntax

WI =<number>

where <number> denotes the column width.

WP option

An output file with a column width of 132 columns is produced if the WP option is specified.

XB option

The XB option is used to suppress the printing of the bivariate frequency tables for the ordinal variables in the output file.

XM option

The purpose of the XM option is to suppress the computing and reporting of the tests for multivariate normality.

XT option

The printing of test statistic value (if applicable) is suppressed by specifying the XT option.

PO command

The PO command allows the user to transform a variable by raising it or a linear combination of it to a power. It is an optional command.

Syntax

PO <labels> AL=<number> BE=<number> GA=<number>

or

PO <numbers> AL=<number> BE=<number> GA=<number>

where <labels> denotes a list of variable labels in free or abbreviated format, <number> is a numerical value and <numbers> denotes a list of variable positions in free or abbreviated format.

Example

PO AGE AL=0 BE=1 GA=2

PR command

The user can specify a Probit regression analysis by using the PR command. It is an optional command.

Syntax

```
PR <labels> ON <labels>
```

or

```
PR <numbers> ON <numbers>
```

where <labels> denotes a list of variable labels in free or abbreviated format and <numbers> denotes a list of variable positions in free or abbreviated format.

Example

```
PR NOSAY ON AGE GENDER RACE
```

RA command

The RA command is used to specify the name of the text file containing the raw data. It is a required command only if a text data file is to be processed. The raw data matrix can also be specified as part of the PRELIS syntax file. In this regard, the RA paragraph rather than the RA command is used (see the RA paragraph section).

Syntax

```
RA=<filename>
```

where <filename> denotes the name of the text data file.

Example

Suppose that the name of the text file containing the raw data is **SELECT.DAT** and that it is located in the folder **Projects\SELECT** on the **E** drive. In this case, the corresponding RA command is given by

```
RA='E:\Projects\ SELECT \ SELECT.DAT'
```

RA paragraph

The RA paragraph is used to provide the raw data to be analyzed as part of the PRELIS syntax file. It is a required command only if the raw data are provided as part of the PRELIS syntax file. If the raw data are in the form of a text file or a PSF, the RA command instead of the RA paragraph is used (see the RA command section).

Syntax

```
RA
<values>
```

where <values> denotes the rows of the raw data matrix. These rows can be provided in free or fixed formats. However, in the case of a fixed format, the first line of the RA paragraph should be a Fortran format statement.

Examples

```
RA
1 7 5
2 1 7
5 5 5
2 2 4
3 3 1
5 6 6
7 7 7
1 1 1
1 2 1
```

```
RA
(2F6.3)
12.34514.417
16.24519.205
10.33411.276
15.11416.267
13.24715.589
```

RE command

The RE command allows the user to recode values or a range of values of a variable to new values. It is an optional command.

Syntax

```
RE <labels> OLD=<valueranges> NEW=<values>
```

or

```
RE <numbers> OLD=<valueranges> NEW=<values>
```

where <labels> denotes a list of variable labels in free or abbreviated format or ALL for all the variables, <valueranges> denotes a list of value ranges (a value range can also be a single value) separated by commas, <values> is a list of values separated by commas and <numbers> denotes a list of variable positions in free or abbreviated format.

Examples

```
RE NOSAY – INTEREST OLD=1,2,3,4 NEW=4,3,2,1
```

```
RE 11-14 OLD=1-3,4-5 NEW=0,1
```

RG command

The RG command is used to specify a univariate or a multivariate multiple linear regression analysis or a two stage least squares regression with instrumental variables. It is an optional command.

Syntax

```
RG <labels> ON <labels> [WITH <labels>] [RES=<label>]
```

or

```
RG <numbers> ON <numbers> [WITH <numbers>] [RES=<label>]
```

where <labels> denotes a list of variable labels in free or abbreviated format, <numbers> denotes a list of variable positions in free or abbreviated format and <label> denotes the label to be used for the regression residuals.

Examples

RG Y1 ON Y2 WITH X1 X2 X3

RG Y1 ON Y2 X2 X3 WITH X1 X2 X3 RES=Y1RES

RG 2 ON 1 4 7 with 3 5 6 8 9

SC command

The purpose of the SC command is to select specific cases to be processed. It is an optional command.

Syntax

SC <labels> <conditions>

or

SC <numbers> <conditions>

or

SC CASE= <condition>

where <labels> denotes a list of variable labels in free or abbreviated format, <numbers> denotes a list of variable positions in free or abbreviated format, <conditions> is a list of conditions in free or abbreviated format and <condition> is one of ODD, EVEN, ><case> or <<case> where <case> denotes a case number.

Examples

SC 2 7-11 > 2 < 5

SC GENDER=2

SD command

The SD command is used to obtain a subset of the raw data and exclude variables from the analysis. It is an optional command.

Syntax

SD <labels> <conditions>

or

SD <numbers> <conditions>

where <labels> denotes a list of variable labels in free or abbreviated format, <numbers> denotes a list of variable positions in free or abbreviated format and <conditions> is a list of conditions in free or abbreviated format.

Examples

SD GENDER=1

SD 15-30

SY command

The SY command is used to specify the PSF to be processed. It is a required command only if a PSF is to be processed.

Syntax

SY=<filename>

where <filename> denotes the complete name (include drive and folder names) of the PSF.

Example

Suppose that the PSF, **DEPRESSION.PSF**, which is located in the folder **Projects\DEPRESSION** on the **F** drive, is to be processed. In this case, the corresponding SY command is given by

SY='F:\Projects\ DEPRESSION \ DEPRESSION.DSF '

Tl paragraph

The Tl paragraph is used to specify a descriptive heading for the analysis. It is an optional command. If the Tl paragraph is used, avoid using any other PRELIS commands or paragraphs in the title string.

Syntax

```
TI  
<string>
```

where <string> denotes a character string.

Example

```
TI  
A PRELIS syntax file for Example 6
```

WE command

The purpose of the WE command is to specify the variable to be used to weight the cases in the raw data file. It is an optional command

Syntax

```
WE <label>  
or  
WE <number>
```

where <label> denotes the descriptive name and <number> denotes the variable position of the case weight variable.

Example

```
WE PAWEIGHT
```

Examples

PRELIS is a versatile tool that, among other uses, can be used interactively for data screening, graphical displays and regression analyses. It includes methods for data on continuous and ordinal variables. In this section, we will illustrate these interactive features by using data on continuous variables obtained from a study of the relationship between cholesterol and fitness and political data on ordinal variables.

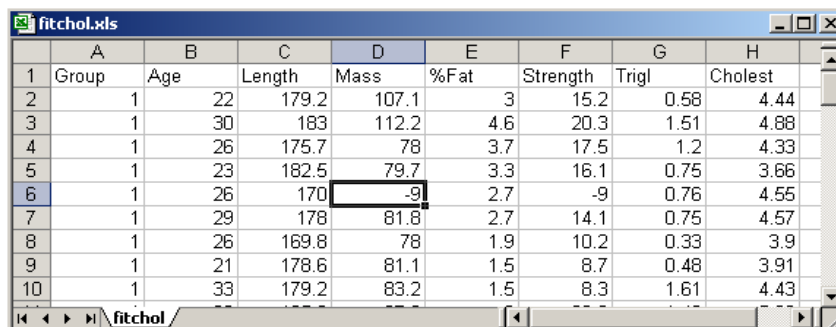
Processing continuous variables

PRELIS can be used to process data on continuous variables interactively. Next, we illustrate these interactive features for continuous variables by using data obtained from a study of the relationship between cholesterol and fitness. The specific data set is described next.

Data preparation example using fitness data

The data

Two of the many factors that are known to have some influence or relevance on the condition of the human heart are physical fitness and blood cholesterol level. In a related research project, three different homogeneous groups of adult males were considered. A number of plasma lipid parameters were measured on each of the 60 individuals and fitness parameters were also measured on three groups. The data file is saved as an Excel file, **fitchol.xls**, in the **TUTORIAL** subfolder. The first 10 lines of the data set are shown below.



	A	B	C	D	E	F	G	H
1	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
2	1	22	179.2	107.1	3	15.2	0.58	4.44
3	1	30	183	112.2	4.6	20.3	1.51	4.88
4	1	26	175.7	78	3.7	17.5	1.2	4.33
5	1	23	182.5	79.7	3.3	16.1	0.75	3.66
6	1	26	170	-9	2.7	-9	0.76	4.55
7	1	29	178	81.8	2.7	14.1	0.75	4.57
8	1	26	169.8	78	1.9	10.2	0.33	3.9
9	1	21	178.6	81.1	1.5	8.7	0.48	3.91
10	1	33	179.2	83.2	1.5	8.3	1.61	4.43

Note

The -9 entries represent missing values in this data set. There are 3 observations with missing values.

The variables in the data set are:

- Group is the group ID (1 for a weightlifter, 2 for a student and 3 for a marathon athlete.)
- Age is the age in years.
- Length is the height in cm.
- Mass is the weight in kg.
- %Fat is the percentage fat.
- Strength represents the breast-strength in lb.
- Trigl is the triglycerides.
- Cholest represents the total cholesterol.

More details about the data are available in Du Toit, Steyn and Stumpf (1986).

We now use the data in **fitchol.xls** to illustrate how to accomplish the following tasks.

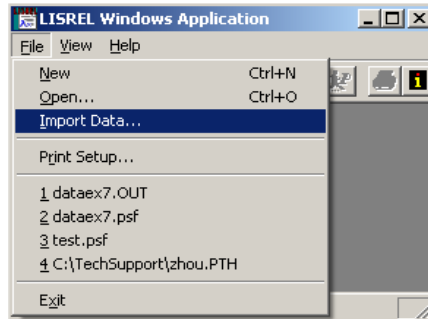
- importing the Excel data file
- defining the variable types
- dealing with the missing values

We then use the complete data after listwise deletion to illustrate the following tasks.

- inserting a new variable
- computing values for a new variable
- selecting cases and create a subset of the data
- exporting the data to Excel
- performing data screening
- computing normal scores
- computing matrices

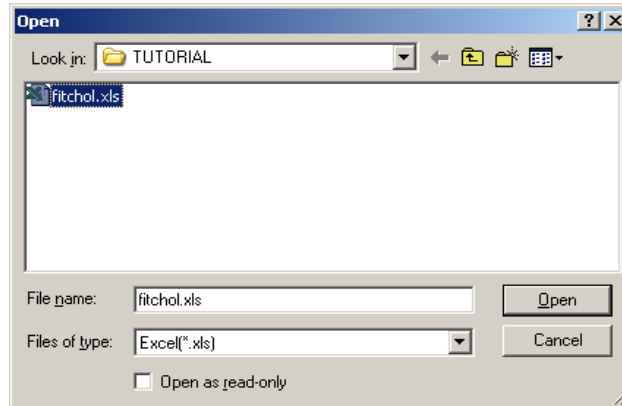
Importing the Excel data file

- Open the main window of LISREL for Windows by using the LISREL shortcut on the LISREL program menu.
- On the **File** menu of the main window, select the **Import Data** option as shown below

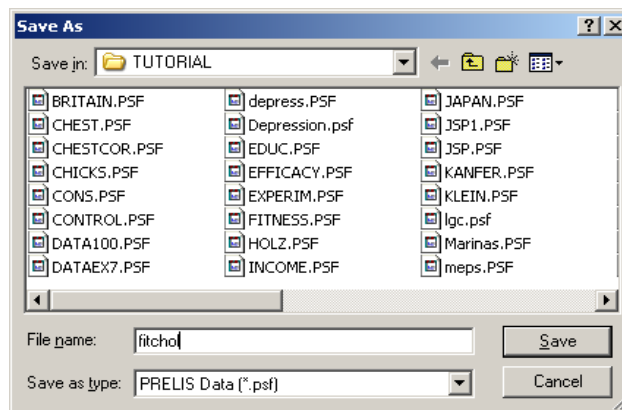


to load the **Open** dialog box.

- Select the **Excel (*.xls)** option from the **Files of type** drop-down list box.
- Browse for the file **fitchol.xls** in the **TUTORIAL** subfolder and select it to produce the following dialog box.



- Click on the **Open** button to load the **Save As** dialog box.
- Enter the name **fitchol** in the **File name** string field to create the following dialog box.



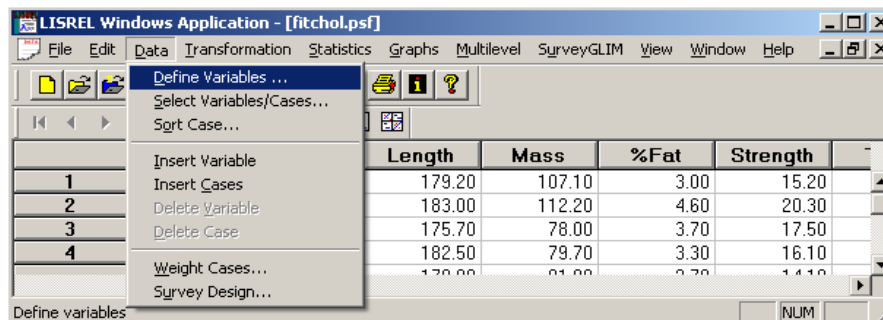
- Click on the **Save** button to open the following PSF window.

	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
5	1.00	26.00	170.00	-9.00	2.70	-9.00	0.76	4.55
6	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57
7	1.00	26.00	169.80	78.00	1.90	10.20	0.33	3.90
8	1.00	21.00	178.60	81.10	1.50	8.70	0.48	3.91
9	1.00	33.00	179.20	83.20	1.50	8.30	1.61	4.43
10	1.00	36.00	185.20	87.80	6.00	23.80	1.42	5.33

Defining the variable types

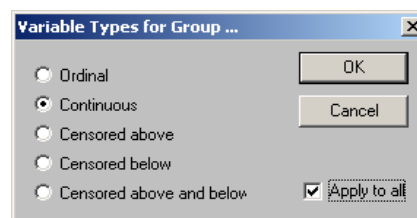
Our next step is to specify the variable types for the variables in **fictchol.psf**. All the variables except for Group are continuous. If data are imported, the default variable type is ordinal. To modify this for the continuous variables in **fictchol.psf**, we proceed as follows.

- o Select the **Define Variables** option on the **Data** menu of the PSF window as shown below



to activate the **Define Variables** dialog box.

- o Select the variable Group from the variable list to activate all the buttons on the **Define Variables** dialog box.
- o Click on the **Variable Type** button to load the **Variable Types for Group...** dialog box.
- o Activate the **Continuous** radio button and check the **Apply to all** check box to obtain the following dialog box.



- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **fictchol.psf**.
- Select the **Define Variables** option on the **Data** menu to activate the **Define Variables** dialog box.
- Select the variable Group from the variable list.
- Click on the **Variable Type** button to load the **Variable Types for Group...** dialog box.
- Activate the **Ordinal** radio button.
- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **fictchol.psf**.

Note

PRELIS doesn't make a distinction between nominal and ordinal variables. Both of them are defined as ordinal.

Dealing with the missing values

Dealing with the missing values is an important step in data preparation. One may choose either to delete or to impute the missing values. There are two methods in PRELIS to delete the missing values, namely listwise and pairwise deletion. There are also two methods to impute the missing values, namely imputation by matching and multiple imputation. In this section, we illustrate the following tasks interactively.

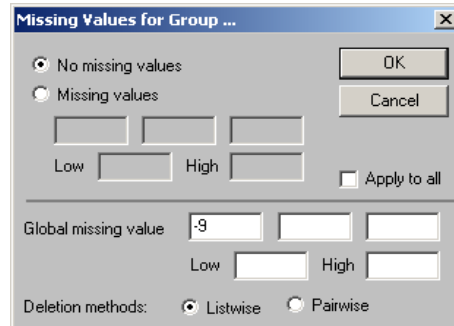
- defining a global missing value and listwise deletion
- multiple imputation

Defining a global missing value and listwise deletion

The value -9 is the global missing value in **fitchol.psf**. We specify this global missing value as follows.

- Select the **Define Variables** option on the **Data** menu of the PSF window to activate the **Define Variables** dialog box.
- Select the variable Group from the variable list to activate the buttons on the **Define Variables** dialog box.
- Click on the **Missing Values** button to load the **Missing Values for Group...** dialog box.

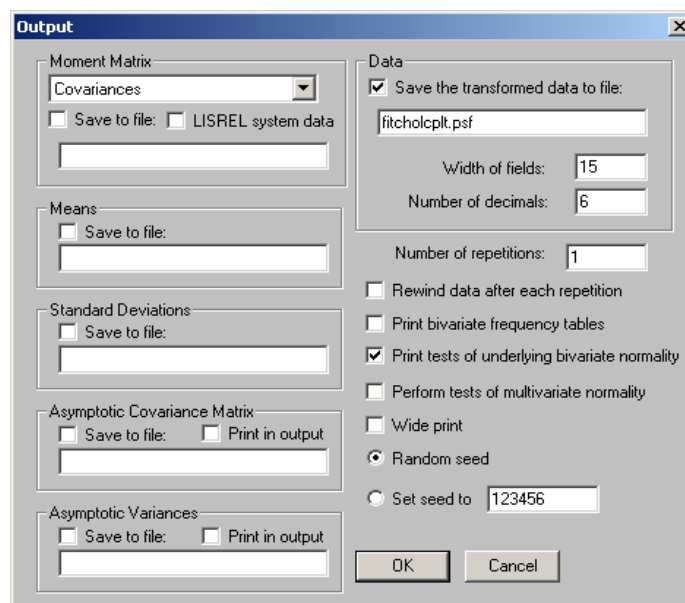
- Enter -9 in the first string field in the **Global missing value** section.
- Activate the **Listwise** radio button in the **Deletion methods** section to produce the following dialog box.



- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **fictchol.psf**.

Next, we illustrate listwise deletion and creating a PSF which only contains the complete cases. This is accomplished as follows.

- Select the **Output Options** option on the **Statistics** menu to obtain the **Output** dialog box.
- Check the **Save the transformed data to file** check box in the **Data** section.
- Enter the name **fitcholcplt.psf** in the corresponding string field to create the following dialog box.



- Click on the **OK** button to produce the text editor window for **fitchol.out**. A portion of this output file is shown below.

fitchol.OUT

```

Number of Missing Values per Variable

  Group   Age   Length   Mass   %Fat   Strength   Trigl   Cholest
-----
         0     0       1       1       1         2       0         1

Distribution of Missing Values

Total Sample Size =      60

Number of Missing Values   0   1   2
Number of Cases           57  0   3
Listwise Deletion

```

The file **fitcholcplt.psf**, which contains 57 observations, is saved in the same subfolder as that of **fitchol.psf**. We access the file as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitcholcplt.psf** and select it.
- Click on the **Open** button to open the PSF window for **fitcholcplt.psf**.

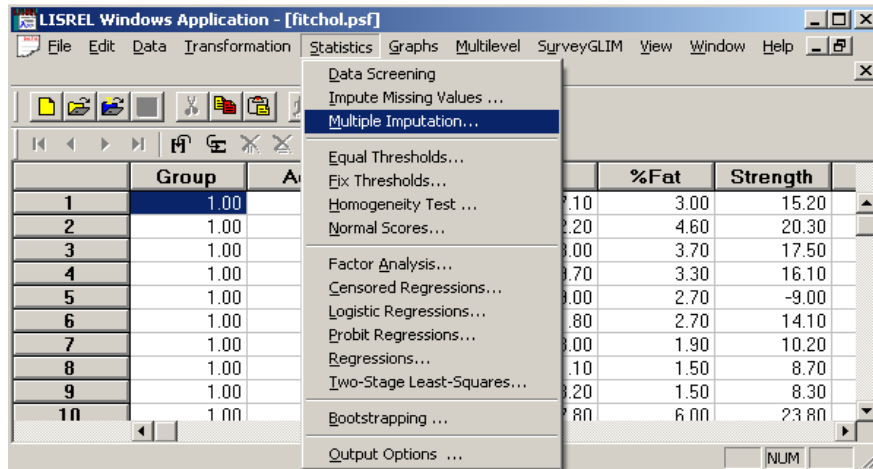
	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
5	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57
6	1.00	26.00	169.80	78.00	1.90	10.20	0.33	3.90
7	1.00	21.00	178.60	81.10	1.50	8.70	0.48	3.91
8	1.00	33.00	179.20	83.20	1.50	8.30	1.61	4.43

Multiple imputation

Missing values can be imputed using either the Expected Maximization (EM) or the Monte Carlo Markov Chain (MCMC) method for multiple imputation. The EM method for multiple imputation is applied interactively to the data in **fitchol.psf** as follows.

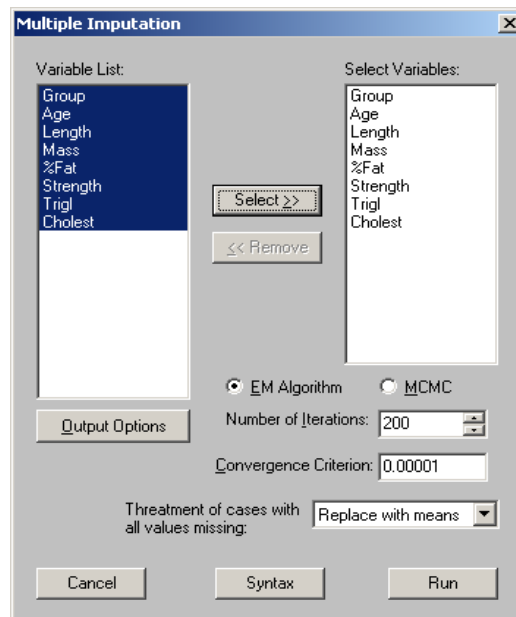
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitchol.psf** and select it.

- Click on the **Open** button to open the PSF window for **fitchol.psf**.
- Select the **Multiple Imputation** option on the **Statistics** menu of the PSF window as shown below

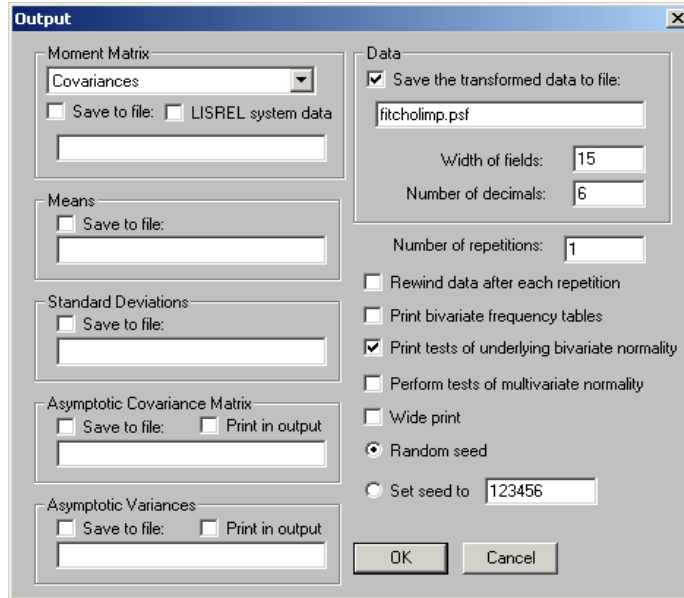


to activate the **Multiple Imputation** dialog box.

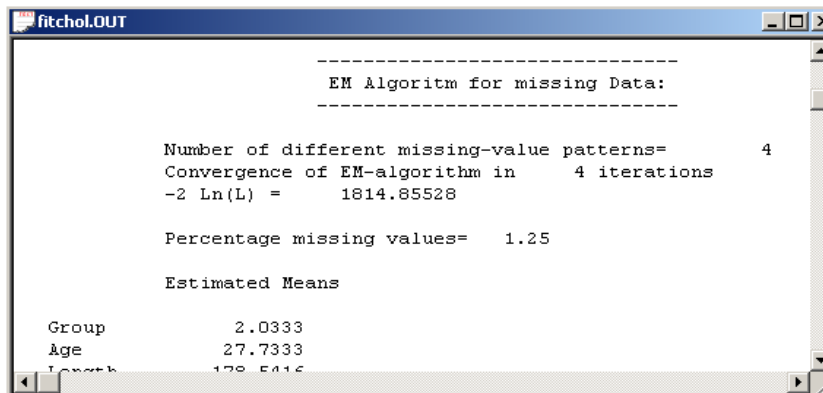
- Select all the variables from the variable list box and then click on the **Select** button to obtain the following dialog box.



- Click on the **Output Options** button to obtain the **Output** dialog box.
- Check the **Save the transformed data to file** check box in the **Data** section.
- Enter the name **fitcholim.psf** in the corresponding string field to produce the following dialog box.



- Click on the **OK** button to return to the **Multiple Imputation** dialog box.
- Click on the **Run** button on the **Multiple Imputation** dialog box to produce the text editor window for **fitchol.out**. A portion of this file follows.



After the multiple imputation, the data file **fitcholimp.psf** has 60 observations with no missing values and is saved in the same subfolder as that of **fitchol.psf**. To view the imputed data set, we proceed as follows.

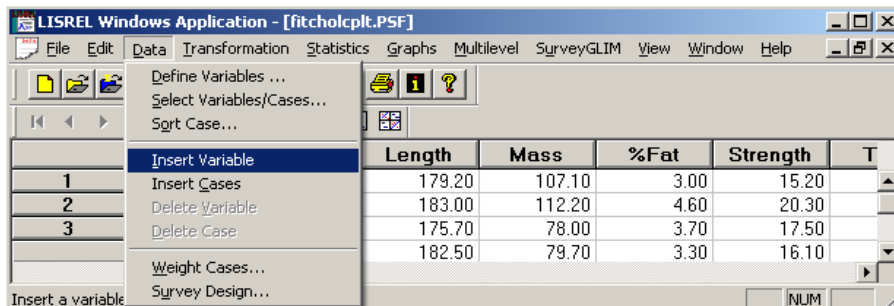
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitcholimp.psf** and select it.
- Click on the **Open** button to open the following PSF window.

	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
5	1.00	26.00	170.00	76.96	2.70	13.49	0.76	4.55
6	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57
7	1.00	26.00	169.80	78.00	1.90	10.20	0.33	3.90
8	1.00	21.00	178.60	81.10	1.50	8.70	0.48	3.91
9	1.00	23.00	179.20	83.20	1.50	8.30	1.61	4.43

Inserting a new variable

Next, we demonstrate how to insert the new variable Totchol into **fitcholcplt.psf**. This is accomplished by using the following steps.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitcholcplt.psf** and select it.
- Click on the **Open** button to open the PSF window for **fitcholcplt.psf**.
- Select the **Insert Variable** option on the **Data** menu of the PSF window as shown below



- to activate the **Insert Variables** dialog box.
- Click on the **OK** button to insert one new variable, which is named var9 by default, before Group.
- Select the **Define Variables** option on the **Data** menu to activate the **Define Variables** dialog box.
- Select the variable var9 from variable list.
- Click on the **Rename** button.
- Enter Totchol to replace var9 in the variable list.
- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to generate the following PSF window.

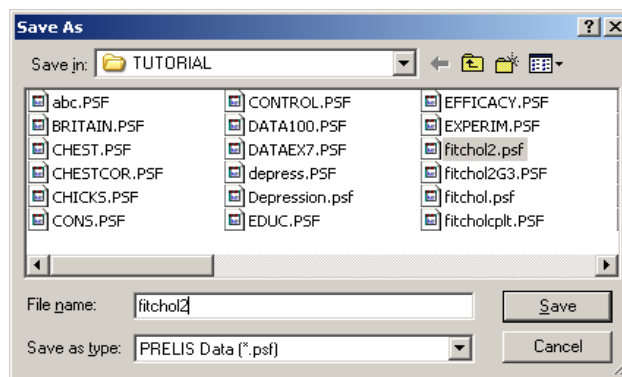
LISREL Windows Application - [fitcholplt.PSF]

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Help

	Totchol	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	0.00	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	0.00	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	0.00	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	0.00	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
	0.00	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57

Ready NUM

- Select the **Save as** option on the **File** menu to load the **Save as** dialog box.
- Enter **fitchol2** as the new file name in the **File name** string field to produce the following dialog box.



- Click on the **Save** button to open the following PSF window.

fitchol2.PSF

	Totchol	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	0.00	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	0.00	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	0.00	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	0.00	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
5	0.00	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57
6	0.00	1.00	26.00	169.80	78.00	1.90	10.20	0.33	3.90
7	0.00	1.00	21.00	178.60	81.10	1.50	8.70	0.48	3.91
8	0.00	1.00	33.00	179.20	83.20	1.50	8.30	1.61	4.43

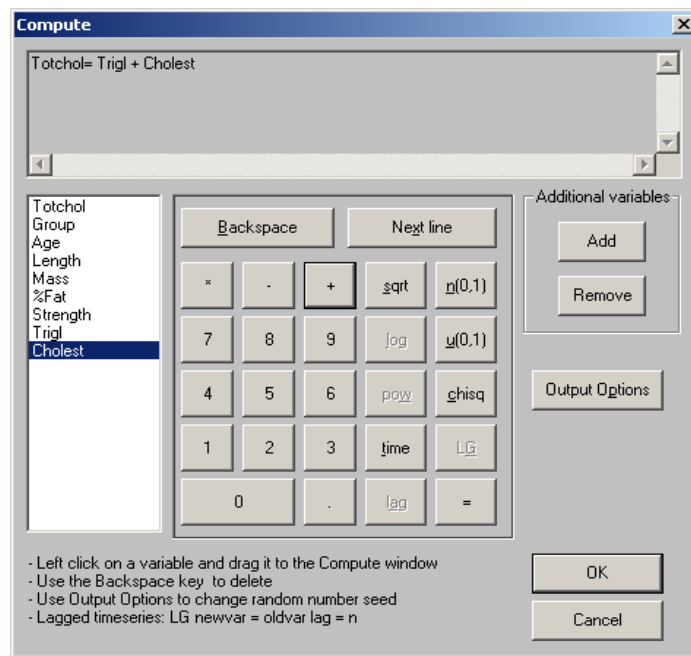
Thus, the variable Totchol has been saved to **fitchol2.psf** as a column with the value 0 assigned to each individual.

Computing values for a new variable

Next, we illustrate how to compute the values of Totchol as the sum of the Trig1 and Cholest values.

We accomplish this as follows.

- Select the **Compute** option on the **Transformation** menu to load the **Compute** dialog box.
- Select Totchol from the variable list, drag it to the top grey string field in the **Compute** dialog box, and release the mouse button.
- Click on the “=” button.
- Select and drag Trigl to the position just after the “=” sign.
- Click on the “+” button.
- Select and drag Cholest to the position just following the “+” sign to produce the following **Compute** dialog box.



- Click on the **OK** button to open the following PSF window.

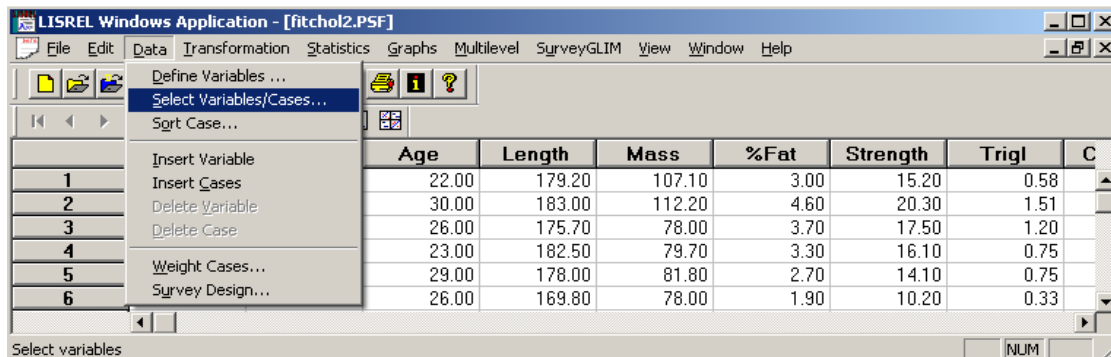
	Totchol	Group	Age	Length	Mass	%Fat	Strength	Trigl	Cholest
1	5.02	1.00	22.00	179.20	107.10	3.00	15.20	0.58	4.44
2	6.39	1.00	30.00	183.00	112.20	4.60	20.30	1.51	4.88
3	5.53	1.00	26.00	175.70	78.00	3.70	17.50	1.20	4.33
4	4.41	1.00	23.00	182.50	79.70	3.30	16.10	0.75	3.66
5	5.32	1.00	29.00	178.00	81.80	2.70	14.10	0.75	4.57
6	4.23	1.00	26.00	169.80	78.00	1.90	10.20	0.33	3.90
7	4.39	1.00	21.00	178.60	81.10	1.50	8.70	0.48	3.91
8	6.04	1.00	33.00	179.20	83.20	1.50	8.30	1.61	4.43
9	6.75	1.00	36.00	185.20	87.80	6.00	23.80	1.42	5.33
10	4.84	1.00	23.00	179.60	80.30	2.20	11.70	1.08	3.76

- Select the **Save** option on the **File** menu to save the changes to **fictchol2.psf**.

Selecting cases and creating a subset of the data

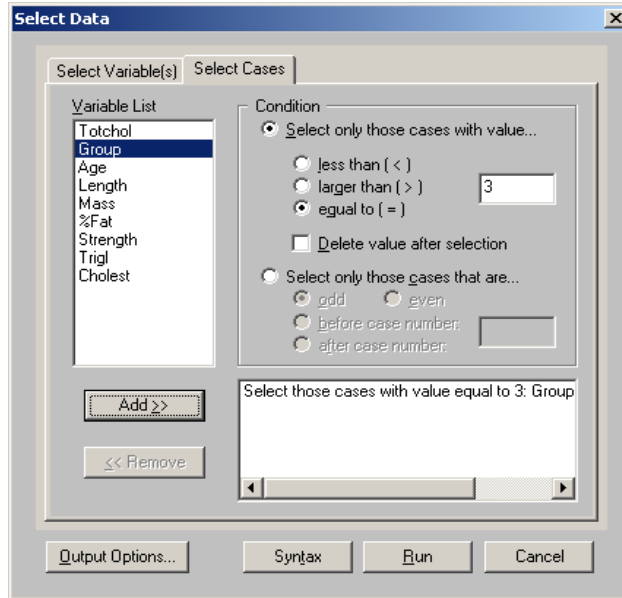
Suppose we want to focus on the subgroup of marathon athletes (Group = 3) in **fitchol2.psf**. The following steps may be used to create a data set consisting only of the marathon athletes.

- Select the **Select Variables/Cases** option on the **Data** menu of the PSF window as shown below

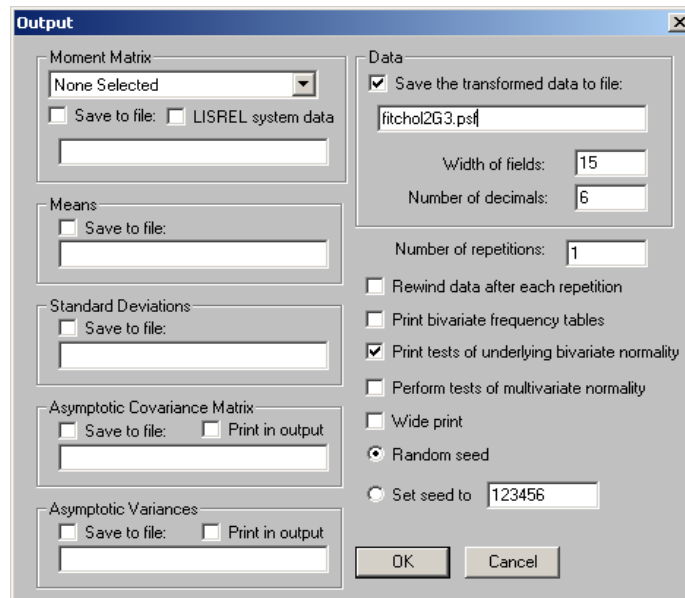


to load the **Select Data** dialog box.

- Click on the **Select Cases** tab.
- Select the variable Group from the **Variable List** list box.
- Activate the **Select only those cases with value** radio button in the **Condition** section.
- Activate the **equal to (=)** radio button.
- Enter the number 3 in the corresponding string field.
- Click on the **Add** button to produce the following dialog box.



- Click on **Output Options** button to obtain the **Output** dialog box.
- Check the **Save the transformed data to file** check box in the **Data** section
- Enter the name **fitchol2G3.psf** in the corresponding string field to generate the following **Output** dialog box.



- Click on the **OK** button to return to the **Select Data** dialog box.
- Click on the **Run** button on the **Select Data** dialog box to produce the text editor window for **fitchol2.out**. A portion of this output file is given below.

Variable	Mean	St. Dev.	T-Value	Skewness	Kurtosis	Minimum Freq.	Maximum Freq.
Totchol	5.718	1.129	22.658	0.486	0.351	3.880	8.340
Age	30.250	7.806	17.330	0.730	-0.371	18.000	45.000
Length	178.415	6.099	130.816	0.208	-0.793	169.100	190.700
Mass	71.485	7.108	44.977	0.454	1.341	58.800	89.500
%Fat	3.130	0.934	14.984	1.845	3.748	1.900	5.800
Strength	15.070	3.177	21.211	1.499	2.893	10.500	24.000
Trig1	1.127	0.951	5.298	3.795	15.802	0.380	4.970
Cholest	4.592	1.268	16.191	-0.714	2.763	1.030	7.190

The following steps may be followed to view the newly created data file **fitchol2G3.psf**.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitchol2G3.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the following PSF window.

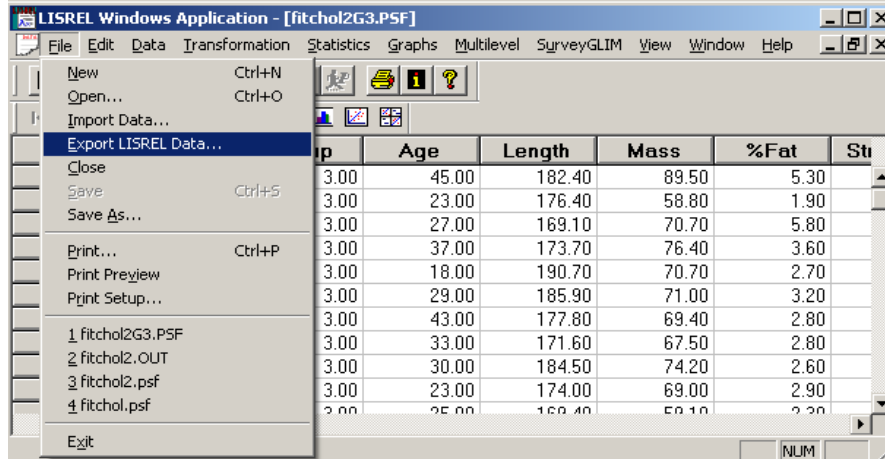
	Totchol	Group	Age	Length	Mass	%Fat
1	5.43	3.00	45.00	182.40	89.50	
2	5.50	3.00	23.00	176.40	58.80	
3	8.34	3.00	27.00	169.10	70.70	
4	6.00	3.00	37.00	173.70	76.40	
5	5.14	3.00	18.00	190.70	70.70	
6	6.32	3.00	29.00	185.90	71.00	
7	7.35	3.00	43.00	177.80	69.40	
8	7.39	3.00	33.00	171.60	67.50	

The new data set contains only the 20 cases of the marathon athletes.

Exporting the data to Excel

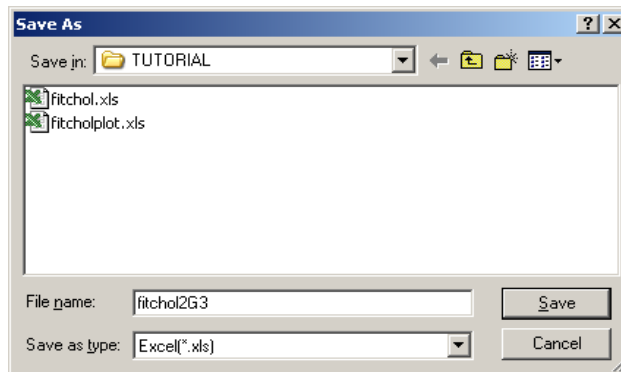
The steps below may be used to export the data in **fitchol2G3.psf** to Excel.

- Select the **Export LISREL Data** option on the **File** menu as shown below



to load the **Save As** dialog box.

- From the **Save as type** drop-down list box, select the **Comma Separated (*.csv)** option.
- Enter **fitchol2G3** in the **File name** string field to generate the following dialog box.



- Click on the **Save** button to save the new CSV data file as **fitchol2G3.csv** in the **TUTORIAL** subfolder. A portion of this file is shown in the following Excel spreadsheet window.

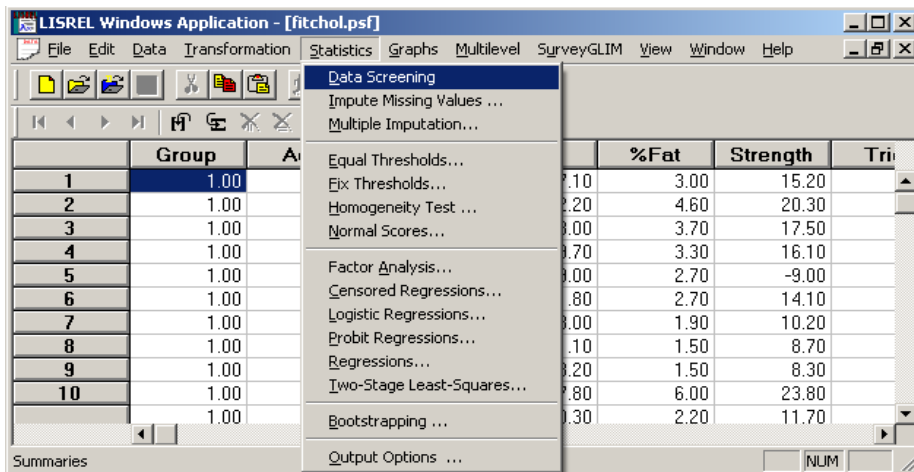
	A	B	C	D	E	F	G	H	I
1	Totchol	Group	Age	Length	Mass	%Fat	Strength	Trig	Cholest
2	5.43	3	45	182.4	89.5	5.3	21.7	0.86	4.57
3	5.5	3	23	176.4	58.8	1.9	10.5	0.84	4.66
4	8.34	3	27	169.1	70.7	5.8	24	1.15	7.19
5	6	3	37	173.7	76.4	3.6	16.6	1.12	4.88
6	5.14	3	18	190.7	70.7	2.7	13.1	0.72	4.42
7	6.32	3	29	185.9	71	3.2	15.8	1.62	4.7
8	7.35	3	43	177.8	69.4	2.8	13.9	1.23	6.12
9	7.39	3	33	171.6	67.5	2.8	14	1.22	6.17
10	6.23	3	30	184.5	74.2	2.6	13.2	1.13	5.1
11	6.08	3	23	174	69	2.9	14.6	0.88	5.1

Data screening

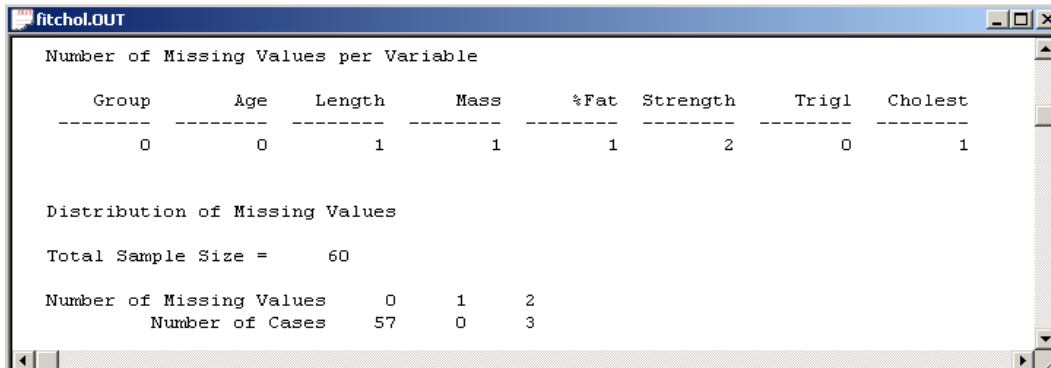
Prior to any data analysis, it may be wise to run a data screening procedure. The PRELIS data screening procedure provides information on the distribution of the missing values, univariate summary statistics and tests of the univariate normality for continuous variables. The procedure also provides information on the distribution of variables over a number of class intervals.

To screen the data in **fitchol.psf**, we proceed as follows.

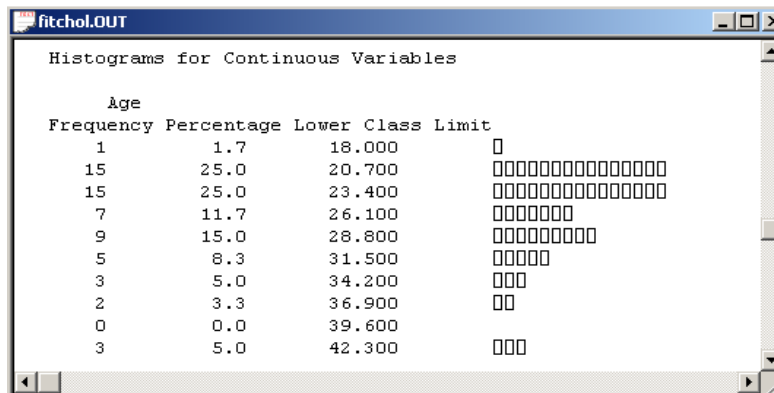
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **fitchol.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the PSF window for **fitchol.psf**.
- Select the **Data Screening** option on the **Statistics** menu of the PSF window as shown below



to produce the text editor window for **fitchol.out**. A selection of this output file is shown below.



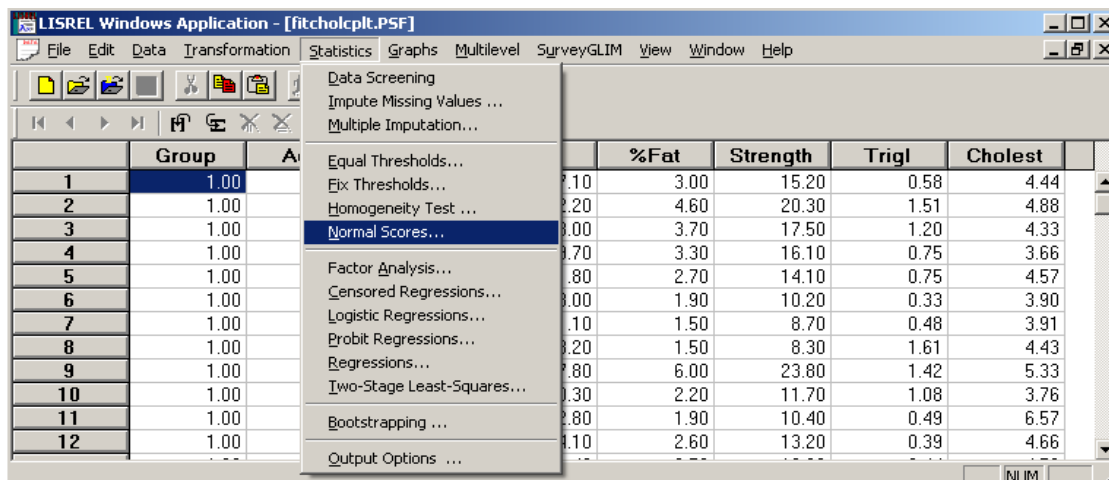
The histogram of each variable is printed in the output file. The histogram for Age is shown below.



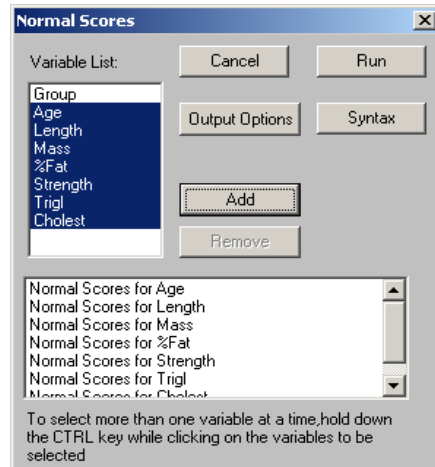
Normal scores

One possible approach to treat non-normality is to normalize the variables before the analysis. Normal scores offer an effective way of normalizing a continuous variable for which the origin and unit of measurement have no intrinsic meaning, such as test scores. Normal scores may be computed for ordinal and continuous variables. To compute the normal scores for the continuous variables in **fitcholcplt.psf**, we proceed as follows.

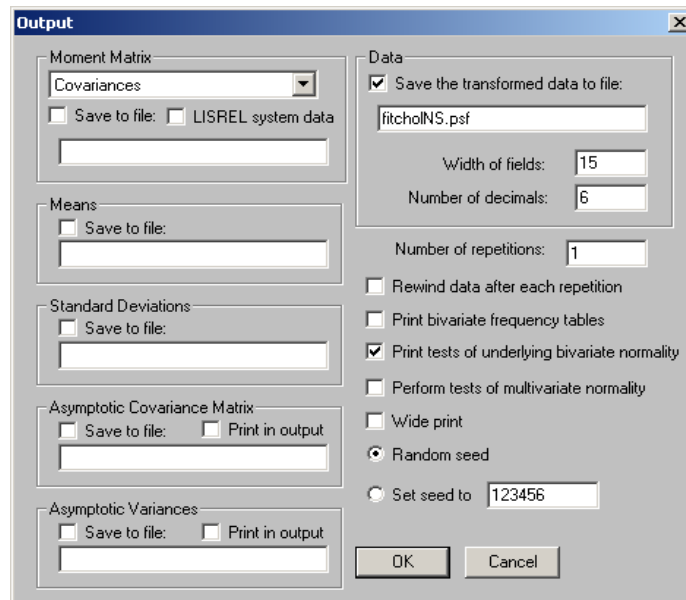
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **PRELIS Data (*.psf)** option from the **Files of type** drop-down list box.
- Browse for the file **fitcholcplt.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the PSF window for **fitcholcplt.psf**.
- Click on the **Normal Scores** option on the **Statistics** menu of the PSF window as shown below



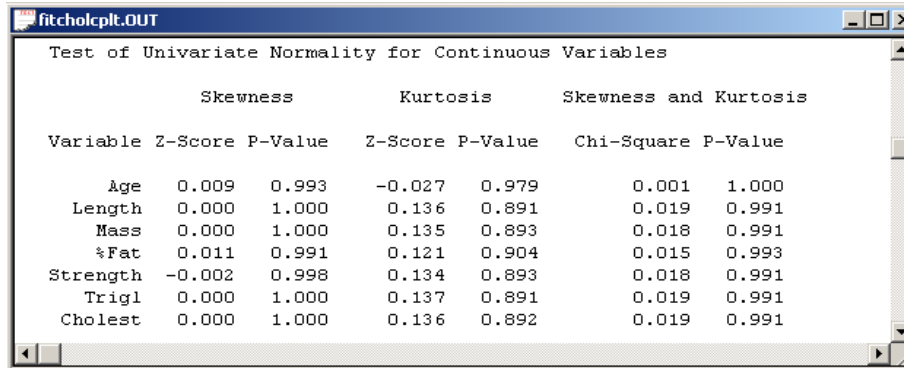
- to load the **Normal Scores** dialog box.
- Select all the continuous variables from the variable list box and the click on the **Add** button to produce the following dialog box.



- Check the **Save the transformed data to file** check box in the **Data** section.
- Enter the name **fitchoINS.psf** for the transformed data file in the corresponding string field to produce the following dialog box.



- Click on the **OK** button to return to the **Normal Scores** dialog box.
- Click on the **Run** button on the **Normal Scores** dialog box to open the text editor window for **fitchoIcplt.out**. A portion of this output file is shown below.



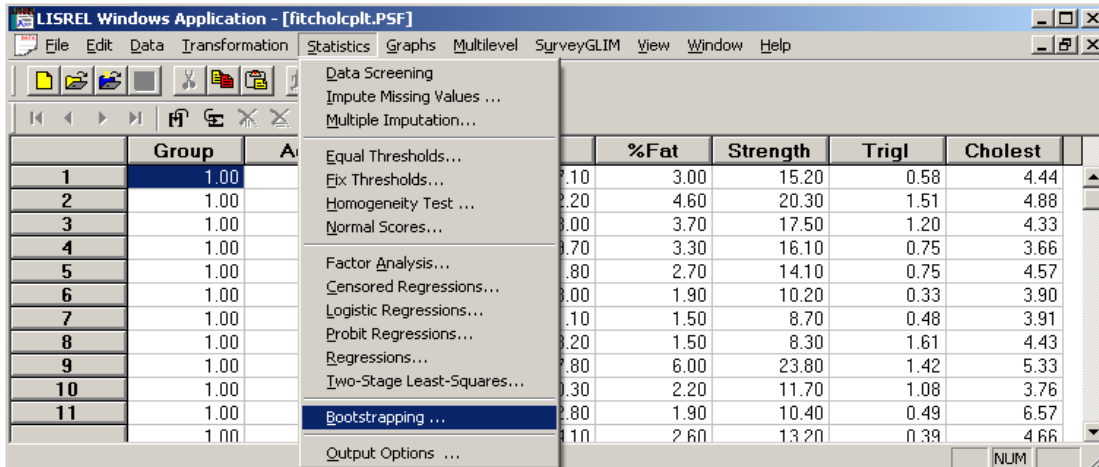
Variable	Skewness		Kurtosis		Skewness and Kurtosis	
	Z-Score	P-Value	Z-Score	P-Value	Chi-Square	P-Value
Age	0.009	0.993	-0.027	0.979	0.001	1.000
Length	0.000	1.000	0.136	0.891	0.019	0.991
Mass	0.000	1.000	0.135	0.893	0.018	0.991
%Fat	0.011	0.991	0.121	0.904	0.015	0.993
Strength	-0.002	0.998	0.134	0.893	0.018	0.991
Trig1	0.000	1.000	0.137	0.891	0.019	0.991
Cholest	0.000	1.000	0.136	0.892	0.019	0.991

Bootstrapping

Drawing random samples of data from some population is often needed for simulation studies. There are two methods for drawing the random samples, namely the bootstrap and the Monte Carlo methods. In the case of bootstrapping, the random samples are drawn from an original sample, which usually is a sample of empirical data but a set of artificial data can also be used. In Monte Carlo sampling, the samples are generated from randomly generated variables so that no real data are involved.

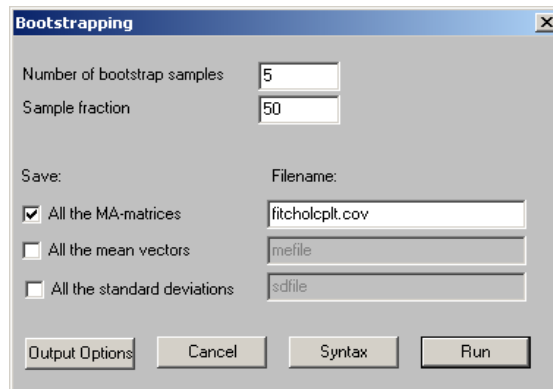
We now use **fitcholcplt.psf** to illustrate how to obtain bootstrap samples with PRELIS. Suppose we want to draw 5 samples of size 50 and study the covariance matrices of these 5 samples. This is achieved as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **fitcholcplt.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **fitcholcplt.psf** in a PSF window.
- Click the **Bootstrapping** option on the **Statistics** menu as shown below

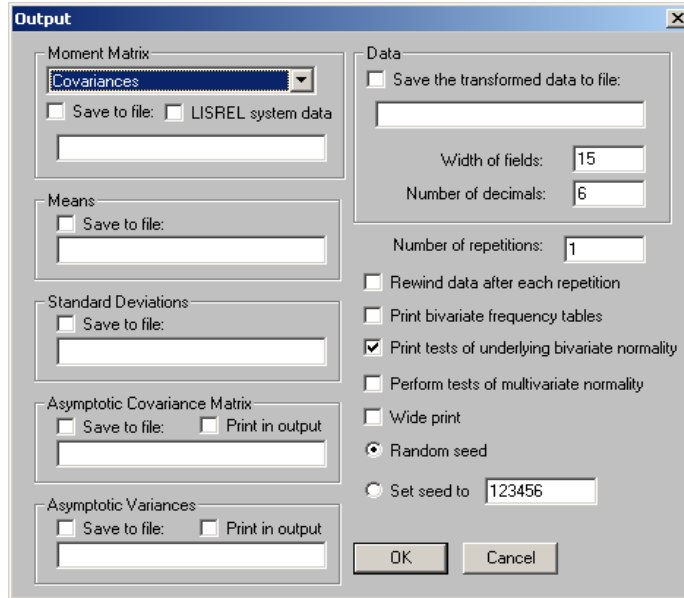


to load the **Bootstrapping** dialog box.

- Check the **All the MA-Matrices** check box.
- Enter the name **fitcholcplt.cov** in the corresponding string field to produce the following dialog box.



- Click on the **Output Options** button to load the **Output** dialog box.
- Activate the **Random seed** radio button to produce the following dialog box.



- Click on the **OK** button to return to the **Bootstrapping** dialog box.
- Click on the **Run** button to produce the text editor window for **fitcholcplt.out**. A portion of this output file is shown below.

	Group	Age	Length	Mass	%Fat	Strength
Group	1.186					
Age	0.081	28.601				
Length	0.002	4.062	27.262			
Mass	-0.748	13.062	19.012	109.232		
%Fat	-0.022	1.823	0.717	4.683	1.163	
Strength	-0.035	5.982	2.003	15.582	4.038	14.245
Trigl	0.119	0.124	-0.018	0.384	0.138	0.540
Cholest	-0.019	1.920	-0.624	0.503	0.316	1.066

The 5 covariance matrices are saved to the text file **fitcholcplt.cov**. We access this file as follows.

- Select the **Open** option on the **File** menu.
- Select the **All Files (*.*)** option on the **Files of type** drop-down list box.
- Browse the file **pasurveyBT.cov** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the text file **pasurveyBT.cov**, which includes 5 covariance matrices, each with $\frac{p \times (p+1)}{2} = 36$ elements, where $p = 8$ is the number of variables. The covariance matrices for the first two bootstrap samples are shown below.

0.36226D+00	0.89029D-01	0.24195D+02	-0.92060D-02	0.83041D+01	0.27797D+02						
-0.15394D+00	0.90389D+01	0.24928D+02	0.52209D+02	-0.12279D-01	0.12693D+01						
0.92675D+00	0.12814D+01	0.74739D+00	-0.31195D-01	0.51579D+01	0.32469D+01						
0.47501D+01	0.27794D+01	0.10461D+02	-0.70483D-02	0.39356D+00	0.85063D-01						
0.71910D+00	0.13744D+00	0.53397D+00	0.27953D+00	-0.57234D-02	0.12695D+01						
-0.49903D+00	-0.45060D+00	0.26448D+00	0.10942D+01	0.24946D+00	0.77647D+00						
0.77915D+00	-0.64819D-01	0.22667D+02	0.16070D-01	-0.71897D+01	0.30792D+02						
-0.12368D+00	0.62599D+00	0.55224D+01	0.44590D+02	0.36561D-01	-0.69163D+00						
-0.19891D+01	0.26783D+01	0.12376D+01	0.88347D-01	-0.23310D+01	-0.73332D+01						
0.96528D+01	0.43534D+01	0.15434D+02	0.20726D+00	-0.40813D+00	-0.50345D+00						
0.80721D+00	-0.31564D-02	0.56634D-01	0.12266D+01	0.26482D-01	0.22819D+01						
-0.15252D+01	-0.19041D+00	0.52279D+00	0.16532D+01	-0.84597D+00	0.18638D+01						

Computing matrices

A covariance matrix (or correlation matrix) and the corresponding estimated asymptotic covariance matrix are often needed to fit structural equation models to data. PRELIS can compute and save these matrices to files interactively. For the data in the **fitcholplt.psf**, we can compute and save the covariance matrix and the estimated asymptotic covariance matrix of the variances and covariances as follows.

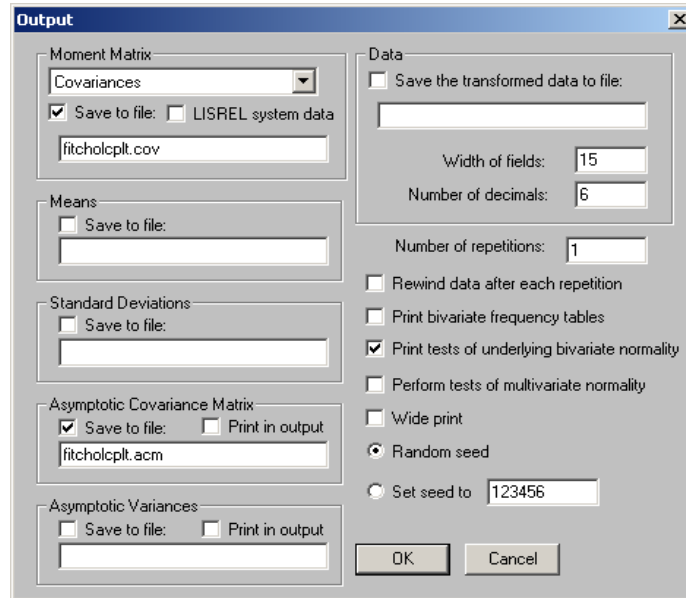
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **fitcholplt.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **fitcholplt.psf** in a PSF window.
- Select the **Output Options** option on the **Statistics** menu as shown below

	Group	A	%Fat	Strength	Trigl	Cholest	
1	1.00		2.10	3.00	15.20	0.58	4.44
2	1.00		2.20	4.60	20.30	1.51	4.88
3	1.00		3.00	3.70	17.50	1.20	4.33
4	1.00		3.70	3.30	16.10	0.75	3.66
5	1.00		3.80	2.70	14.10	0.75	4.57
6	1.00		3.00	1.90	10.20	0.33	3.90
7	1.00		3.10	1.50	8.70	0.48	3.91
8	1.00		3.20	1.50	8.30	1.61	4.43
9	1.00		3.80	6.00	23.80	1.42	5.33
10	1.00		3.30	2.20	11.70	1.08	3.76
11	1.00		2.80	1.90	10.40	0.49	6.57
12	1.00		3.10	2.60	13.20	0.39	4.66

to load the **Output** dialog box.

- Check the **Save to file** check box in the **Moment Matrix** section.
- Enter the name **fitcholplt.cov** for the computed covariance matrix of the variables in **fitcholplt.psf** in the corresponding string field.

- Check the **Save to file** check box in the **Asymptotic Covariance Matrix** section
- Enter the name **fitcholcplt.acm** in the corresponding string field to produce the following dialog box.



- Click on the **OK** button to open the text editor window for **fitcholcplt.out**. A selection of this output file is given below.

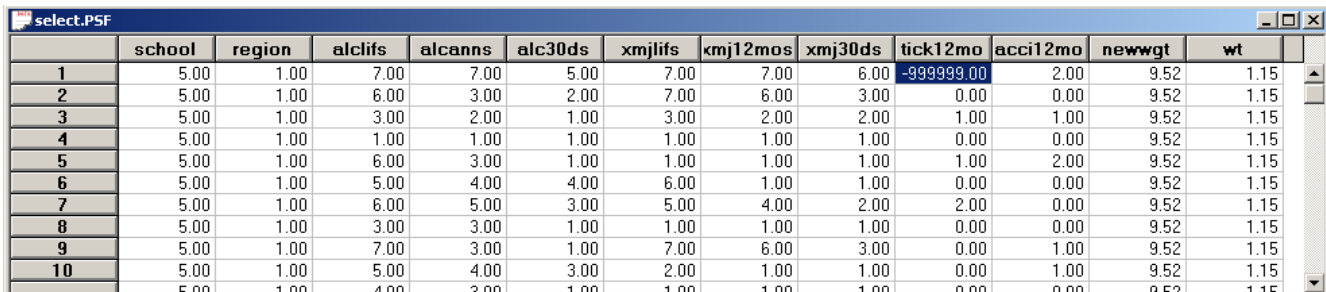
	Group	Age	Length	Mass	%Fat	Strength
Group	1.201					
Age	1.775	36.909				
Length	0.158	7.685	31.825			
Mass	-6.169	14.692	21.799	110.761		
%Fat	0.045	1.545	0.565	4.338	1.141	
Strength	0.275	4.875	1.536	14.668	3.971	14.084
Trig1	0.129	0.342	0.128	0.322	0.130	0.496
Cholest	-0.015	1.979	-1.285	-0.302	0.199	0.657

Thus, the covariance matrix is saved as a text file **fitcholcplt.cov** in the same subfolder as that of **fitcholcplt.psf**. The corresponding estimated asymptotic covariance matrix is saved as the binary file **fitcholcplt.acm** in the same subfolder.

Multiple linear regression example using drug abuse data

The data

The data consist of indicators of alcohol and marijuana usage of 1608 high school students during 2001 as recorded in the Monitoring The Future (MTF) study which is conducted by the Institute for Social Research at the University of Michigan. The raw data are listed in the file **select.psf** in the **TUTORIAL** subfolder. The first few lines of **select.psf** are shown in the following PSF window.



	school	region	alclifs	alcanns	alc30ds	xmjlifs	xmj12mos	xmj30ds	tick12mo	acci12mo	newwgt	wt
1	5.00	1.00	7.00	7.00	5.00	7.00	7.00	6.00	-999999.00	2.00	9.52	1.15
2	5.00	1.00	6.00	3.00	2.00	7.00	6.00	3.00	0.00	0.00	9.52	1.15
3	5.00	1.00	3.00	2.00	1.00	3.00	2.00	2.00	1.00	1.00	9.52	1.15
4	5.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	9.52	1.15
5	5.00	1.00	6.00	3.00	1.00	1.00	1.00	1.00	1.00	2.00	9.52	1.15
6	5.00	1.00	5.00	4.00	4.00	6.00	1.00	1.00	0.00	0.00	9.52	1.15
7	5.00	1.00	6.00	5.00	3.00	5.00	4.00	2.00	2.00	0.00	9.52	1.15
8	5.00	1.00	3.00	3.00	1.00	1.00	1.00	1.00	0.00	0.00	9.52	1.15
9	5.00	1.00	7.00	3.00	1.00	7.00	6.00	3.00	0.00	1.00	9.52	1.15
10	5.00	1.00	5.00	4.00	3.00	2.00	1.00	1.00	0.00	1.00	9.52	1.15

Note that the -999999.00 is the default missing value of this data file. The variables are:

- alclifs is the numerical response to the question “On how many occasions have you had alcoholic beverages to drink in your lifetime? ”
- alc12mos is the numerical response to the question “On how many occasions have you had alcoholic beverages to drink in the past 12 months? ”
- alc30ds is the numerical response to the question “On how many occasions have you had alcoholic beverages to drink in the past 30 days? ”
- xmjlifs is the numerical response to the question “On how many occasions have you used marijuana in your lifetime? ”
- xmj12mos is the numerical response to the question “On how many occasions have you used marijuana in the past 12 months? ”
- xmj30ds is the numerical response to the question “On how many occasions have you used marijuana in the past 30 days? ”
- tick12mo is the numerical response to the question “Within the last 12 months, how many times have you received a ticket (or been stopped and warned) for moving violations? ”
- acci12mo is the numerical response to the question “Within the last 12 months, how many times you were involved in an accident while driving? ”

More details about the data are available at

<http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/04019.xml>.

The model

The general purpose of multiple linear regression is to learn more about the relationship between several independent or predictor variables x_1, \dots, x_p and a dependent or criterion variable y . In general, then, multiple linear regression procedures will estimate a linear equation of the form:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \varepsilon$$

where β_0 is the intercept, β_1, \dots, β_p are the regression coefficients and ε is the error term. The multiple linear regression model to be fitted to the data in **select.psf** may be expressed as

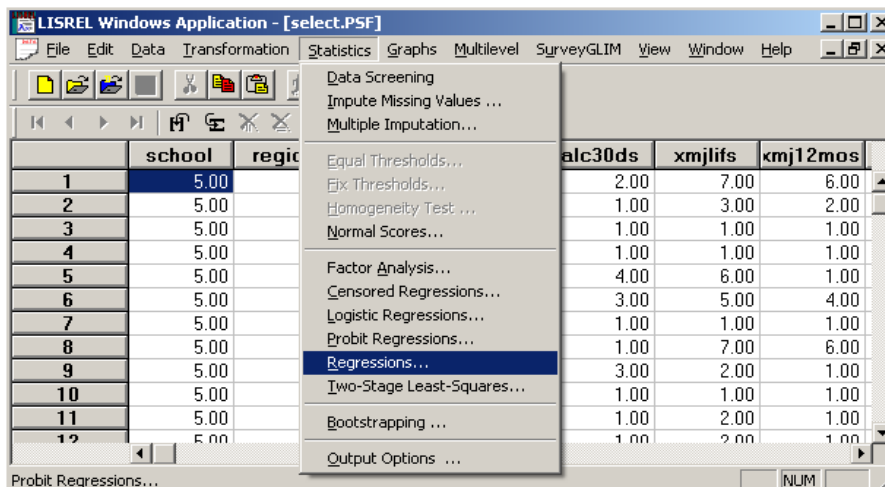
$$\text{acci12mo} = \beta_0 + \beta_1 \text{alclifs} + \beta_2 \text{alcanns} + \beta_3 \text{alc30ds} + \beta_4 \text{xmj1lifs} + \beta_5 \text{xmj12mos} + \beta_6 \text{xmj30ds} + \varepsilon$$

where β_0 denotes the mean number of accidents in the past 12 months if no alcohol or marijuana is used, β_1, \dots, β_6 are unknown regression weights and ε denotes the error term.

The analysis

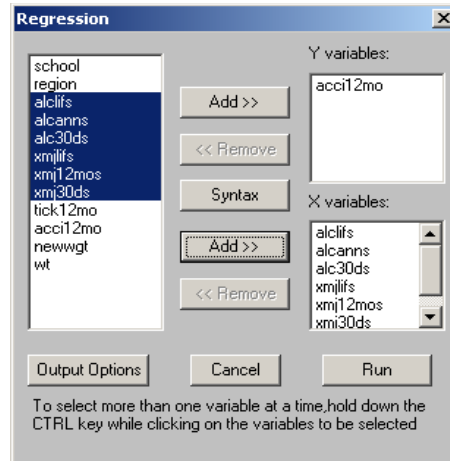
To fit the multiple linear regression model above to the data in **select.psf**, we proceed as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **select.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **select.psf** in a PSF window.
- Select the **Regressions** option on the **Statistics** menu as shown below

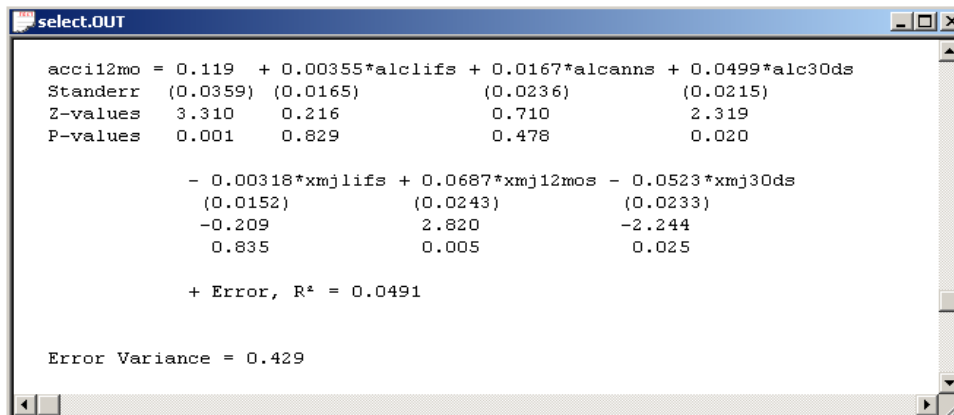


to load the **Regression** dialog box.

- Select **acci12mo** from the variable list box.
- Click on the top **Add** button to transfer the selected variable to the **Y variables** list box.
- Select the variables **alclifs**, **alcanns**, **alc30ds**, **xmj1ifs**, **xmj12mos**, and **xmj30ds** from the variable list box.
- Click on the bottom **Add** button to create the following dialog box.



- Click on the **Run** button to open the text editor window for **fitchol.out**. A selection of this output file is given below.

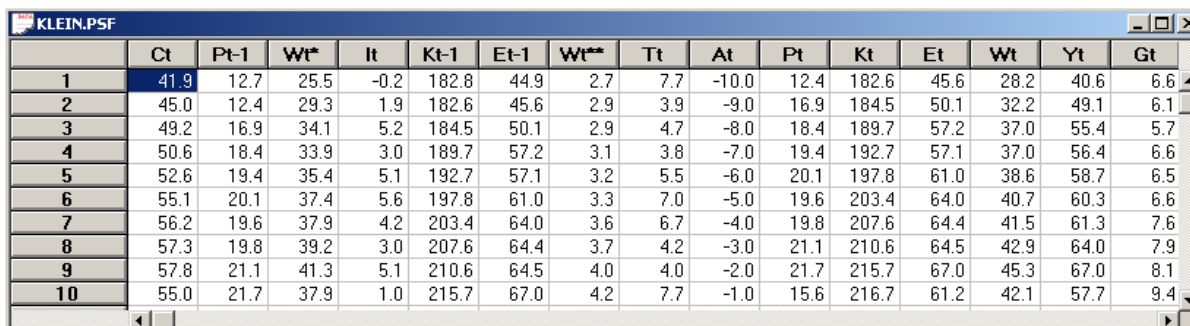


Two-stage least square example using US economy data

Two-stage Least Squares (TSLS) estimation often provides sufficient information to judge whether a structural equation model is reasonable or not. TSLS estimates and the corresponding standard error estimates may be obtained quickly, without iterations, for several typical structural equation models and are useful especially in the early stages of investigation.

The data

Klein's Model I (Klein 1950) is a classical econometric model, which has been used extensively as a benchmark problem for studying econometric methods. It is an eight-equation system based on annual data for the United States in the period between the two World Wars. It is dynamic in the sense that elements of time play important roles in the model. The data file, **klein.psf**, is located in the **TUTORIAL** subfolder of the LISREL installation. The first 10 observations of this file are shown below.



	Ct	Pt-1	Wt*	It	Kt-1	Et-1	Wt**	Tt	At	Pt	Kt	Et	Wt	Yt	Gt
1	41.9	12.7	25.5	-0.2	182.8	44.9	2.7	7.7	-10.0	12.4	182.6	45.6	28.2	40.6	6.6
2	45.0	12.4	29.3	1.9	182.6	45.6	2.9	3.9	-9.0	16.9	184.5	50.1	32.2	49.1	6.1
3	49.2	16.9	34.1	5.2	184.5	50.1	2.9	4.7	-8.0	18.4	189.7	57.2	37.0	55.4	5.7
4	50.6	18.4	33.9	3.0	189.7	57.2	3.1	3.8	-7.0	19.4	192.7	57.1	37.0	56.4	6.6
5	52.6	19.4	35.4	5.1	192.7	57.1	3.2	5.5	-6.0	20.1	197.8	61.0	38.6	58.7	6.5
6	55.1	20.1	37.4	5.6	197.8	61.0	3.3	7.0	-5.0	19.6	203.4	64.0	40.7	60.3	6.6
7	56.2	19.6	37.9	4.2	203.4	64.0	3.6	6.7	-4.0	19.8	207.6	64.4	41.5	61.3	7.6
8	57.3	19.8	39.2	3.0	207.6	64.4	3.7	4.2	-3.0	21.1	210.6	64.5	42.9	64.0	7.9
9	57.8	21.1	41.3	5.1	210.6	64.5	4.0	4.0	-2.0	21.7	215.7	67.0	45.3	67.0	8.1
10	55.0	21.7	37.9	1.0	215.7	67.0	4.2	7.7	-1.0	15.6	216.7	61.2	42.1	57.7	9.4

The 15 variables in the data set are:

- Ct is the aggregate consumption for year t.
- Pt-1 is the total profits for year t-1.
- Wt* represents the private wage bill for year t.
- It denotes the net investment for year t.
- Kt-1 is the capital stock for year t-1.
- Et-1 indicates the total production of private industry for year t-1.
- Wt** is the government wage bill for year t.
- Tt is the taxes for year t.
- At represents time in years from 1931, *i.e.* t-1931.
- Pt is the total profits for year t.
- Kt denotes the end-of-year capital stock for year t.
- Et is the total production of private industry for year t.

- W_t is the total wage bill for year t .
- Y_t is the total income for year t .
- G_t is government non-wage expenditure for year t .

All these variables except for A_t are in billions of 1934 dollars. More details about this data set are available in Klein (1950).

The model

Two-stage least squares (TSLS) is particularly useful for estimating econometric models of the form

$$\mathbf{y} = \mathbf{B}\mathbf{y} + \mathbf{\Gamma}\mathbf{x} + \mathbf{u}$$

where $\mathbf{y} = (y_1, y_2, \dots, y_p)'$ is a set of endogenous or jointly dependent variables, $\mathbf{x} = (x_1, x_2, \dots, x_q)'$ is a set of exogenous or predetermined variables uncorrelated with the error terms $\mathbf{u} = (u_1, u_2, \dots, u_p)'$, and \mathbf{B} and $\mathbf{\Gamma}$ are parameter matrices.

A typical feature of the above model is that not all y -variables and not all x -variables are included in each equation. A necessary condition for identification of each equation is that, for every y -variable on the right side of the equation, there must be at least one x -variable excluded from that equation. There is also a sufficient condition for identification, the so-called rank condition, but this is often difficult to apply in practice. For more information on the identification of interdependent systems, see, *e.g.*, Goldberger (1964, pp. 313-318).

The specific model to be fitted to the data in **klein.psf** may be expressed as

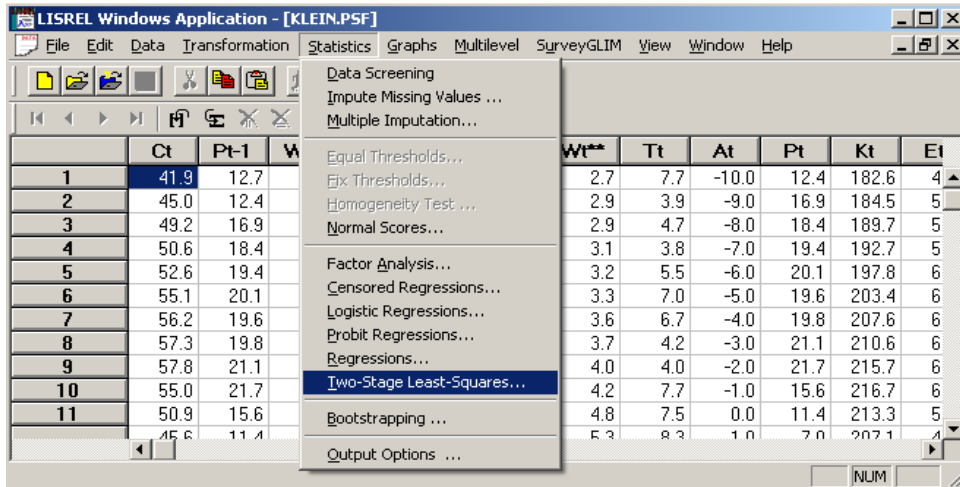
$$C_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \beta_3 W_t + \varepsilon_t$$

where β_0 denotes the expected aggregate consumption in year t , β_1 , β_2 and β_3 denote unknown regression weights and ε_t denotes an error term.

The analysis

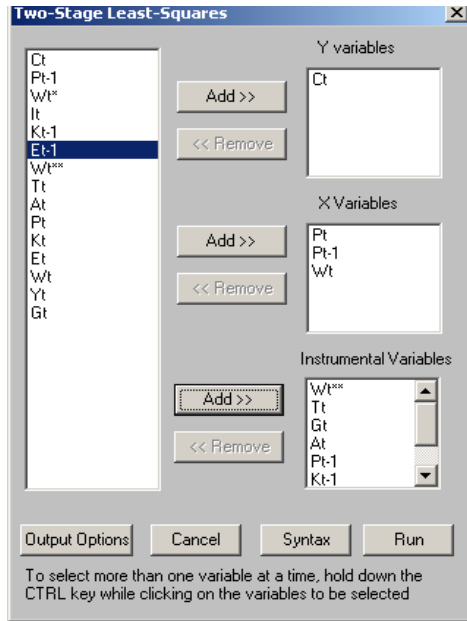
We now use the TSLS method to fit the model above to the data in **klein.psf**. We accomplish this as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **klein.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **klein.psf** in a PSF window.
- Click on the **Two-Stage Least-Squares** option on the **Statistics** menu as shown below

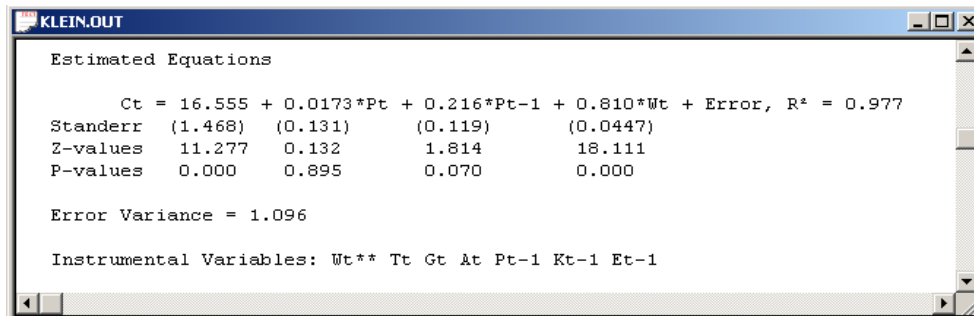


to load the **Two-Stage Least-Squares** dialog box.

- Select the label **Ct** in the variable list box.
- Click on the top **Add** button to transfer the label **Ct** to the **Y Variables** list box.
- Select the labels **Pt**, **Pt-1** and **Wt** in the variable list box.
- Click on the center **Add** button to transfer the label **Pt**, **Pt-1** and **Wt** to the **X Variables** list box.
- Select the labels **Wt****, **Tt**, **Gt**, **At**, **Pt-1**, **Kt-1** and **Et-1** in the variable list box.
- Click on the bottom **Add** button to produce the following dialog box.



- Click on the **Run** button to produce the text editor window for **klein.out**. A selection of the **klein.out** is shown in the following text editor window.

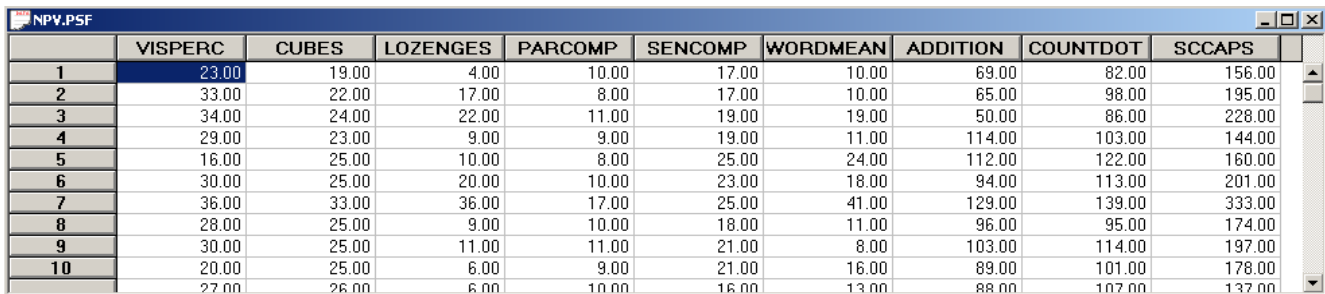


Exploratory factor analysis example using psychological data

Exploratory factor analysis (EFA) is a technique often used to detect and assess latent sources of variation and covariation in observed measurements. It is widely recognized that exploratory factor analysis can be quite useful in the early stages of experimentation or test development. In an exploratory factor analysis, one wants to explore the empirical data to discover and detect characteristic features and interesting relationships without imposing any definite model on the data. An exploratory factor analysis may be structure generating, model generating, or hypothesis generating. ML estimates of factor loadings (unrotated, VARIMAX rotated or PROMAX rotated) are provided. The number of factors may be specified manually or may be determined automatically by PRELIS. In this section, we will demonstrate how PRELIS can be used to perform an EFA.

The data

Holzinger and Swineford (1939) collected data on twenty-six psychological tests administered to 145 seventh- and eighth-grade children in the Grant-White school in Chicago. Nine of these tests are selected for this example. The data file, **npv.psf**, is located in the **TUTORIAL** subfolder of the LISREL installation. The first 10 observations of the nine selected variables are given below.



	VISPERC	CUBES	LOZENGES	PARCOMP	SENCOMP	WORDMEAN	ADDITION	COUNTDOT	SCCAPS
1	23.00	19.00	4.00	10.00	17.00	10.00	69.00	82.00	156.00
2	33.00	22.00	17.00	8.00	17.00	10.00	65.00	98.00	195.00
3	34.00	24.00	22.00	11.00	19.00	19.00	50.00	86.00	228.00
4	29.00	23.00	9.00	9.00	19.00	11.00	114.00	103.00	144.00
5	16.00	25.00	10.00	8.00	25.00	24.00	112.00	122.00	160.00
6	30.00	25.00	20.00	10.00	23.00	18.00	94.00	113.00	201.00
7	36.00	33.00	36.00	17.00	25.00	41.00	129.00	139.00	333.00
8	28.00	25.00	9.00	10.00	18.00	11.00	96.00	95.00	174.00
9	30.00	25.00	11.00	11.00	21.00	8.00	103.00	114.00	197.00
10	20.00	25.00	6.00	9.00	21.00	16.00	89.00	101.00	178.00
	27.00	26.00	6.00	10.00	16.00	13.00	88.00	107.00	137.00

- VISPERC is the visual perception test score.
- CUBES is the cubes test score.
- LOZENGES is the lozenges tests score.
- PARCOMP indicates the paragraph completion test score.
- SENCOMP represents the sentence completion test score.
- WORDMEAN is the word meaning test score.
- ADDITION is the add test score.
- COUNTDOT denotes the counting groups of dots test score.
- SCCAPS is the straight and curved capitals test score.

More information on the data in **npv.psf** may be found in Holzinger and Swineford (1939).

The mathematical model

The objective of exploratory factor analysis is to find, for a given set of response variables $\mathbf{x} = (x_1, \dots, x_q)'$, a set of underlying latent factors $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)'$, fewer in number than the observed variables. These factors are supposed to account for the intercorrelations of the response variables in the sense that when the factors are partialled out no correlations between variables should remain. The mathematical model can be expressed as

$$\mathbf{x} = \mathbf{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta}$$

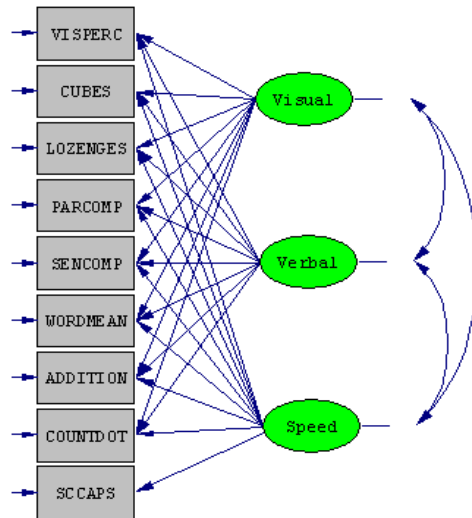
where $\mathbf{\Lambda}$ denotes an unknown parameter matrix, $\boldsymbol{\delta}$ denotes a vector of measurement errors, $E(\boldsymbol{\xi}) = \mathbf{0}$, $E(\boldsymbol{\delta}) = \mathbf{0}$ and $\boldsymbol{\delta}$ is uncorrelated with $\boldsymbol{\xi}$.

The specific model

The specific model to be fitted to the data may be expressed as

$$\begin{bmatrix} \text{VISPERC} \\ \text{CUBES} \\ \text{LOZENGES} \\ \text{PARCOMP} \\ \text{SENCOMP} \\ \text{WORDMEAN} \\ \text{ADDITION} \\ \text{COUNTDOT} \\ \text{SCCAPS} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} \\ \lambda_{51} & \lambda_{52} & \lambda_{53} \\ \lambda_{61} & \lambda_{62} & \lambda_{63} \\ \lambda_{71} & \lambda_{72} & \lambda_{73} \\ \lambda_{81} & \lambda_{82} & \lambda_{83} \\ \lambda_{91} & \lambda_{92} & \lambda_{93} \end{bmatrix} \begin{bmatrix} \text{Visual} \\ \text{Verbal} \\ \text{Speed} \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \end{bmatrix}.$$

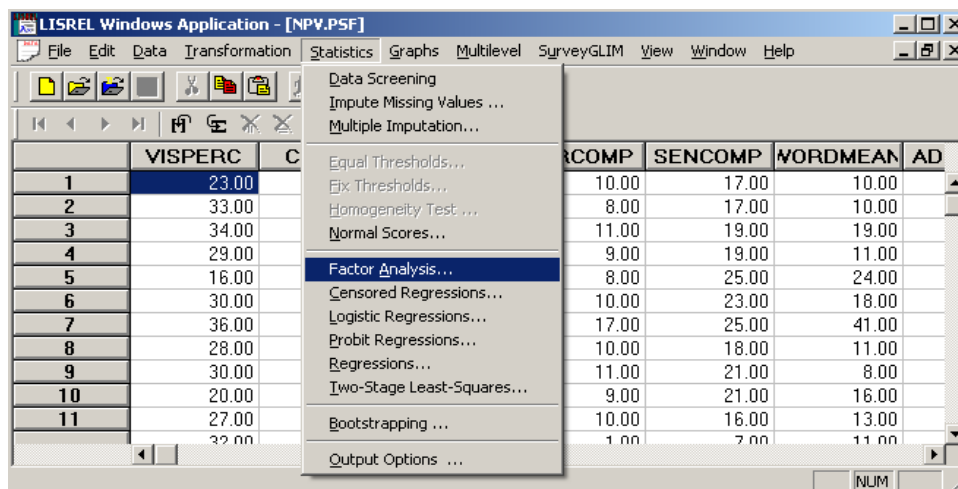
where $\lambda_{11}, \lambda_{12}, \dots, \lambda_{93}$ denote the unknown factor loadings and $\delta_1, \delta_2, \dots, \delta_9$ are the measurement errors. The specific EFA model above suggests that the psychological tests are indicators of the latent variables visual ability, verbal ability and speed. A path diagram of this EFA model follows.



The analysis

We now use the data in **npv.psf** to illustrate how to perform an exploratory factor analysis interactively.

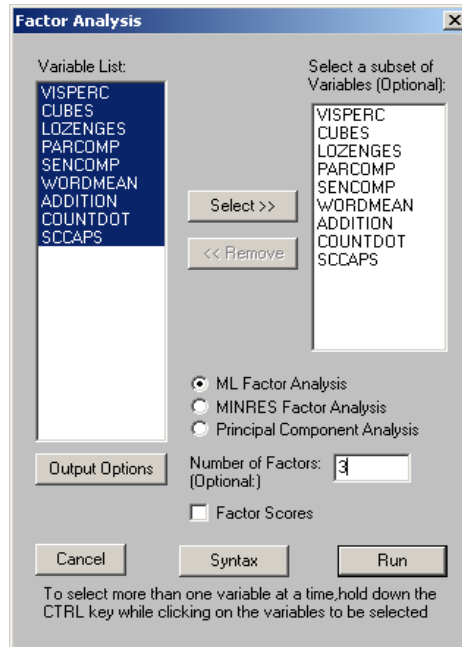
- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **npv.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **npv.psf** in a PSF window.
- Select the **Factor Analysis** option on the **Statistics** menu as shown below



to load the **Factor Analysis** dialog box.

- Select all the labels from the **Variable List** list box.

- Click on the **Select** button to transfer all the labels to **Select a subset of Variables** box.
- Activate the **ML Factor Analysis** radio button
- Enter 3 in the **Number of factors** string field to produce the following dialog box.



- Click on the **Run** button to produce the text editor window for **npv.out**. A portion of this output file is shown below.

Promax-Rotated Factor Loadings				
	Factor 1	Factor 2	Factor 3	Unique Var
VISPERC	0.678	0.039	0.031	0.499
CUBES	0.531	-0.013	-0.051	0.740
LOZENGES	0.670	0.064	-0.059	0.535
PARCOMP	0.074	0.848	-0.045	0.241
SENCOMP	0.007	0.805	0.086	0.302
WORDMEAN	0.083	0.796	-0.049	0.322
ADDITION	-0.184	0.131	0.793	0.388
COUNTDOT	0.200	-0.145	0.774	0.317
SCCAPS	0.431	0.041	0.446	0.456

Processing ordinal variables

Observations on an ordinal variable represent responses to a set of ordered categories, such as a five-category Likert scale. Ordinal variables do not have origins or units of measurements. Means, variances, and covariances of ordinal variable have no meaning (Jöreskog 2002a). In this section, we illustrate how to use PRELIS to deal with ordinal variables.

Data preparation example using political action survey data

The data

Barnes & Kaase (1979) conducted a cross-national survey to obtain information on conventional and unconventional forms of political participation in industrial societies. The survey is known as the Political Action Survey. Two surveys were conducted. The first one was conducted between 1973 and 1975; and the second was obtained during 1980-81. The SPSS data file **pasurvey.sav** in the **TUTORIAL** subfolder contains the responses to these statements for both surveys recorded for a sample of 768 United States respondents. The first portion of the data file is shown in the following **SPSS Data Editor** window.

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2	COMPLEX2	NOCARE2	TOUCH2	IN
1	2	2	1	1	1	1	9	2	2	9	2	
2	2	3	3	3	2	3	2	3	2	2	2	
3	3	2	2	3	3	3	3	2	2	2	2	
4	2	2	1	1	2	1	2	2	2	2	2	1
5	3	2	2	3	3	3	3	2	2	3	2	
6	2	2	2	2	1	2	3	2	1	3	2	
7	3	1	2	2	2	2	2	2	3	3	2	
8	2	1	2	2	1	1	3	3	1	2	9	

The operational definition of political efficacy is based on the responses to the first 12 items of the data set.

- NOSAY1 is the numerical answer to the statement “People like me have no say in what the government does” in survey 1.
- COMPLEX1 is the numerical answer to the statement “Sometimes politics and government seem so complicated that a person like me cannot really understand what is going on” in survey 1.
- NOCARE1 is the numerical answer to the statement “I don’t think that public officials care much about what people like me think” in survey 1.

- TOUCH1 is the numerical answer to the statement “Generally speaking, those we elect to Congress in Washington lose touch with people pretty quickly” in survey 1.
- INTERES1 is the numerical answer to the statement “Parties are only interested in people’s votes but not in their opinions” in survey 1.
- NOSAY2 is the numerical answer to the statement “People like me have no say in what the government does” in survey 2.
- COMPLEX2 is the numerical answer to the statement “Sometimes politics and government seem so complicated that a person like me cannot really understand what is going on” in survey 2.
- NOCARE2 is the numerical answer to the statement “I don’t think that public officials care much about what people like me think” in survey 2.
- TOUCH2 is the numerical answer to the statement “Generally speaking, those we elect to Congress in Washington lose touch with people pretty quickly” in survey 2.
- INTERES2 is the numerical answer to the statement “Parties are only interested in people’s votes but not in their opinions” in survey 2.

The respondents had to respond to the statements above with one of:

- agree strongly (recorded as 1)
- agree (recorded as 2)
- disagree (recorded as 3)
- disagree strongly (recorded as 4)
- do not know (recorded as 8)
- no answer (recorded as 9)

The 3 demographic variables in the data set are:

- AGE is the age of the respondent in 1974.
- GENDER is the gender of the respondent (1 for male and 0 for female).
- EDUCAT represents the education of the respondent (1 for compulsory level only, 2 for middle level and 3 for higher or academic level).

In this section, we use this data set to illustrate how the following tasks are accomplished interactively.

- importing the SPSS data file
- defining the variable types
- assigning labels to the categories
- defining the missing values

- performing data screening
- dealing with missing values

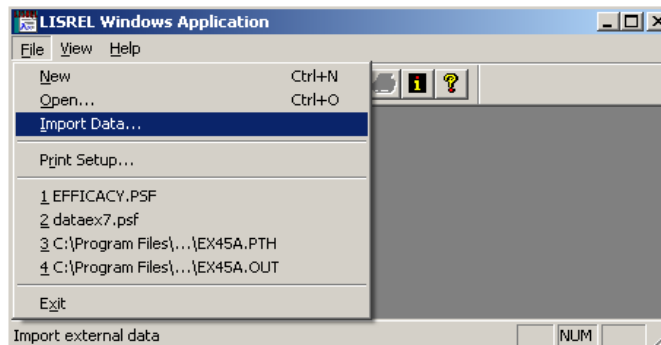
We then use the imputed data to demonstrate the following tasks.

- exporting the PSF to SPSS
- conducting an equal thresholds test and computing matrices
- conducting a fixed thresholds test and computing matrices
- generating bootstrap samples
- computing matrices

Importing the SPSS data file

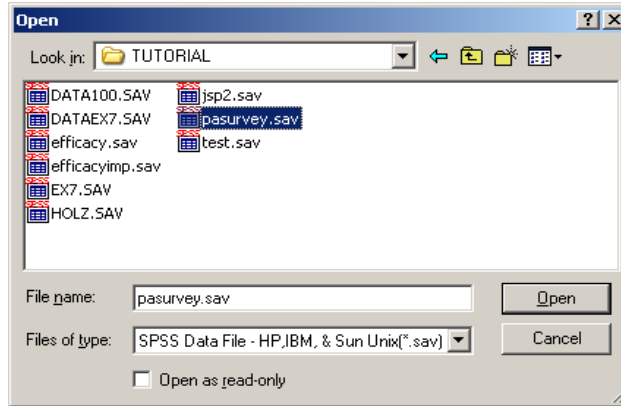
We import the SPSS data file **pasurvey.sav** in the **TUTORIAL** subfolder as follows.

- Open the main window of LISREL for Windows by using the LISREL shortcut on the LISREL program menu.
- On the **File** menu of the main window, select the **Import Data** option as shown below

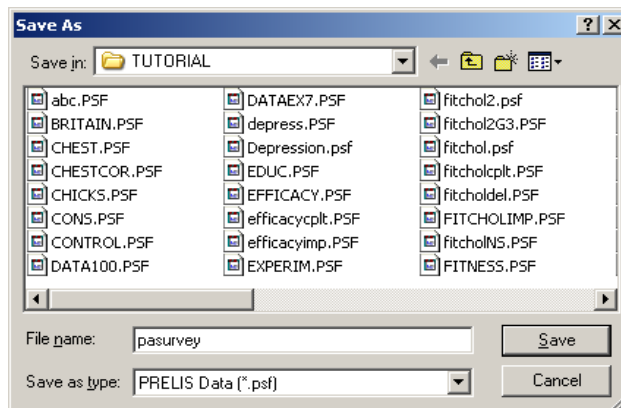


to load the **Open** dialog box.

- Select the **SPSS Data File (*.sav)** option from the **Files of type** drop-down list box.
- Browse for the file **pasurvey.sav** in the **TUTORIAL** subfolder and select it to produce the following dialog box.



- Click on the **Open** button to load the **Save As** dialog box.
- Enter the name **pasurvey** in the **File name** string field to create the following dialog box.



- Click on the **Save** button to open the following PSF window.

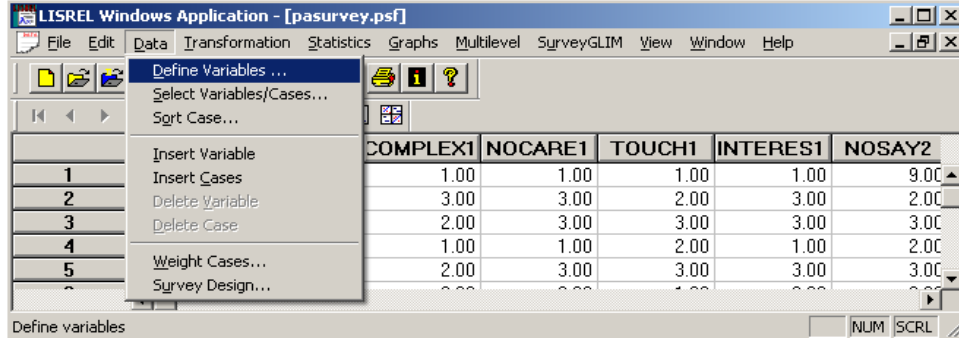
	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2	COMPLEX2	NOCARE
1	2.00	2.00	1.00	1.00	1.00	1.00	9.00	2.00	2.00	9.00
2	2.00	3.00	3.00	3.00	2.00	3.00	2.00	3.00	2.00	2.00
3	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	2.00
4	2.00	2.00	1.00	1.00	2.00	1.00	2.00	2.00	2.00	2.00
5	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	3.00
6	2.00	2.00	2.00	2.00	1.00	2.00	3.00	2.00	1.00	3.00
7	3.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00
8	2.00	1.00	2.00	2.00	1.00	1.00	3.00	3.00	1.00	2.00
9	3.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00	1.00	3.00
10	2.00	2.00	3.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00

Defining the variable types

Our next step is to specify the variable types for the variables in **pasurvey.psf**. All the variables except for AGE are ordinal. When data are imported, the default variable type is ordinal. Thus, we

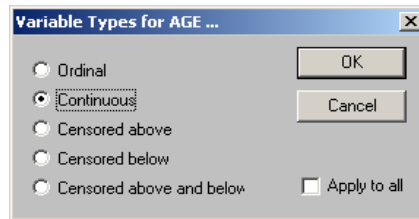
only need to define AGE as a continuous variable. This is accomplished as follows.

- Select the **Define Variables** option on the **Data** menu of the PSF window as shown below



to activate the **Define Variables** dialog box.

- Select the variable AGE from the variable list to activate all the buttons on the **Define Variables** dialog box.
- Click on the **Variable Type** button to load the **Variable Types for AGE...** dialog box.
- Activate the **Continuous** radio button to produce the following dialog box.

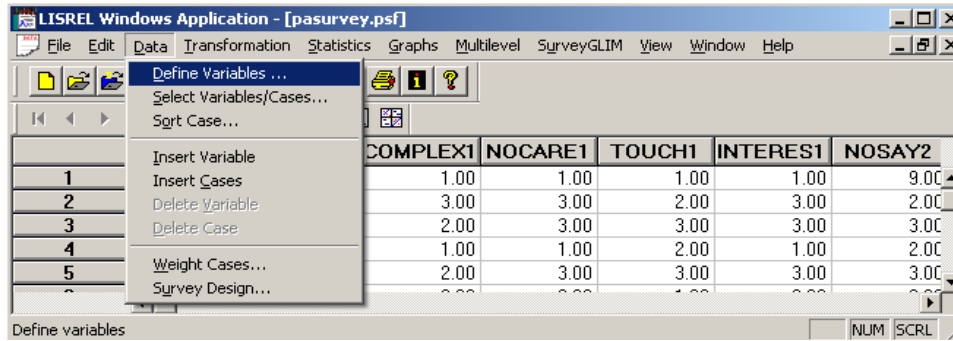


- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **pasurvey.psf**.

Assigning labels to the categories

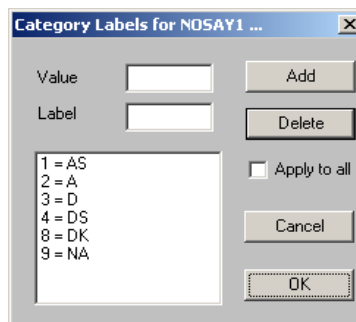
Each of the 12 political efficacy items has a maximum of six distinct numerical values. To assign category labels AS, A, D, DS, DK, and NA (AS = Agree strongly, A = Agree, D = Disagree, DS = Disagree Strongly, DK = Don't Know and NA = Not Available) to the numerical values 1, 2, 3, 4, 8 and 9 respectively, we proceed as follows.

- Select the **Define Variables** option on the **Data** menu of the PSF window as shown below



to activate the **Define Variables** dialog box.

- Select the first 12 variables from the variable list to activate all the buttons on the **Define Variables** dialog box.
- Click on the **Category Labels** button to load the **Category Labels for NOSAY1...** dialog box.
- Enter the number 1 in the string box of **Value** and enter AS in the **Label** string box.
- Click on the **Add** button to assign AS as the label for the value 1.
- Similarly assign the labels for the values 2, 3, 4, 8 and 9 respectively to obtain the following dialog box.



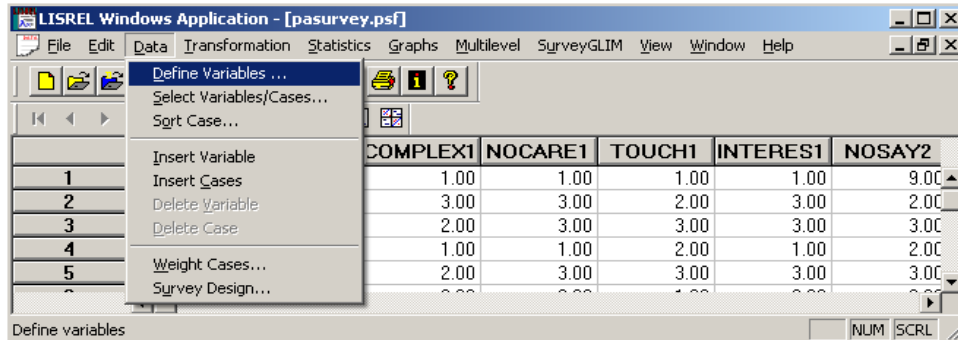
- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **pasurvey.psf**.

Dealing with the missing values

Defining the missing values

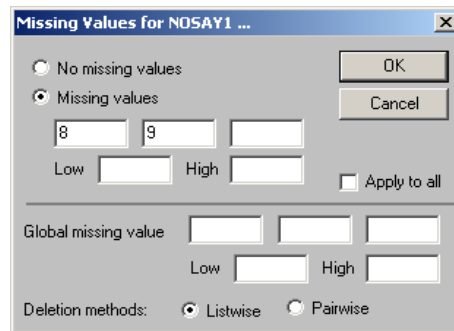
We deliberately selected the cases without any missing values in survey 1. However, the 6 political efficacy items in survey 2 contain values 8 and 9, which are actually codes to indicate missing responses. We specify this as follows.

- Select the **Define Variables** option on the **Data** menu of the PSF window as shown below



to activate the **Define Variables** dialog box.

- Select the first 12 variables and click on the **Missing Values** option to activate the **Missing Values for NOSAY1...** dialog box.
- Activate the **Missing values** radio button. Enter the values 8 and 9 in the first two string fields.
- Activate the **Listwise** radio button to produce the following dialog box.

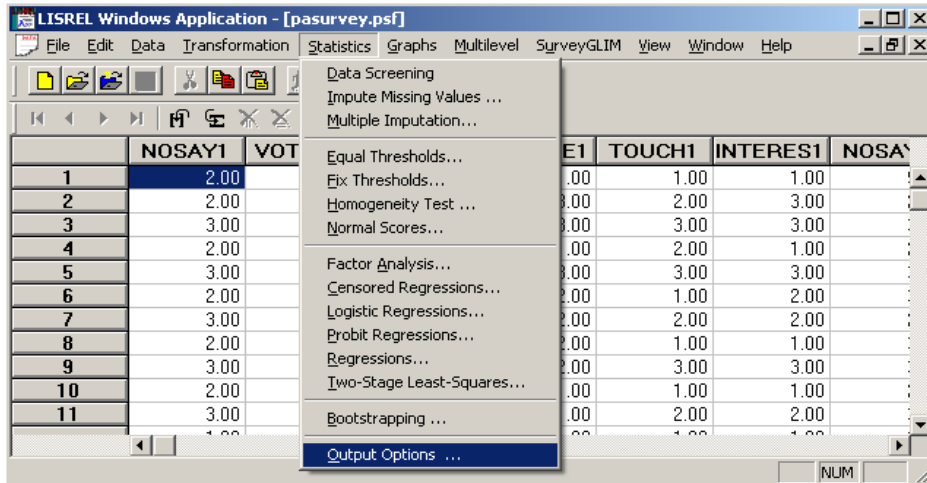


- Click on the **OK** button to return to the **Define Variables** dialog box.
- Click on the **OK** button on the **Define Variables** dialog box to return to the PSF window.
- Select the **Save** option on the **File** menu to save the changes to **pasurvey.psf**.

Listwise deletion

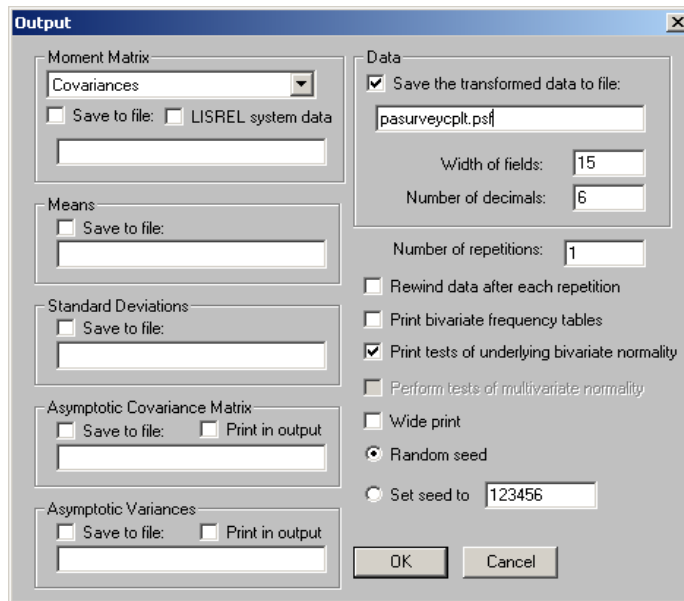
Next, we illustrate listwise deletion and the creation of a PSF which only contains the complete cases in **pasurvey.psf**. This is accomplished as follows.

- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurvey.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurvey.psf** in the following PSF window.
- Select the **Output Options** option on the **Statistics** menu as shown below



to obtain the **Output** dialog box.

- Check the **Save the transformed data to file** check box in the **Data** section.
- Enter the name **pasurveycplt.psf** in the corresponding string field to create the following dialog box.



- Click on the **OK** button to produce the text editor window for **pasurvey.out**. A selection of this output file is shown below.

```

Total Sample Size = 768

Number of Missing Values  0    1    2    3    4    5    6
      Number of Cases  718  34   12   1    0    0    3

Listwise Deletion

Total Effective Sample Size = 718

```

The file **pasurveycplt.psf**, which contains 718 observations, is saved in the same subfolder as that of **pasurvey.psf**. To view this data set, we proceed as follows.

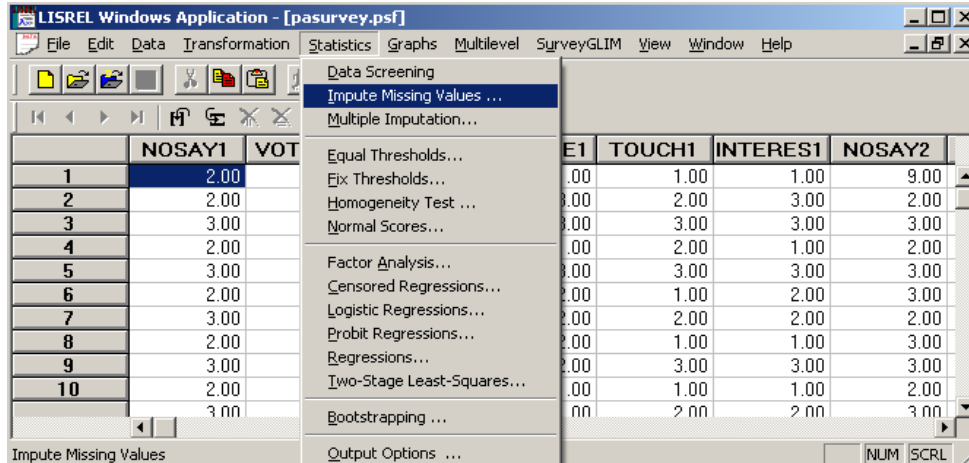
- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurveycplt.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the following PSF window.

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2	COMPLEX2	NOCARE2	TOU
1	2.00	3.00	3.00	3.00	2.00	3.00	2.00	3.00	2.00	2.00	
2	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	2.00	
3	2.00	2.00	1.00	1.00	2.00	1.00	2.00	2.00	2.00	2.00	
4	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	3.00	
5	2.00	2.00	2.00	2.00	1.00	2.00	3.00	2.00	1.00	3.00	
6	3.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	
7	3.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00	1.00	3.00	
8	2.00	2.00	3.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	
9	3.00	2.00	1.00	1.00	2.00	2.00	3.00	2.00	2.00	2.00	
10	1.00	1.00	1.00	1.00	1.00	1.00	3.00	3.00	2.00	3.00	

Imputation by matching

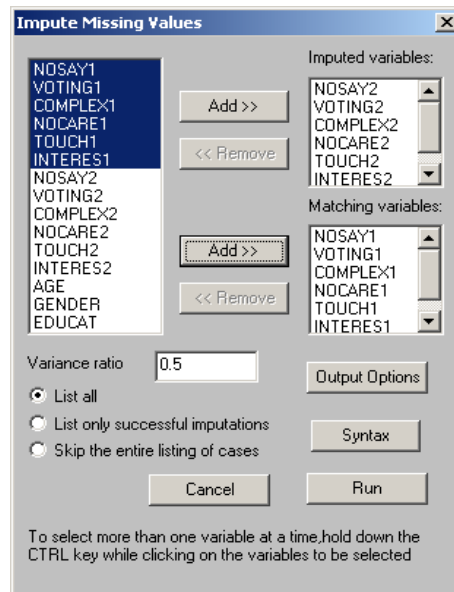
Missing values can be handled by choosing pairwise or listwise deletion. PRELIS offers another possible way to handle missing values, namely imputation by matching. The value to be substituted for the missing value for a case is obtained from another case that has a similar response pattern over a set of matching variables (see Jöreskog & Sörbom 1999, pages 153-158). We impute the missing values in the second survey by matching the patterns from the first survey in **pasurvey.psf** as follows.

- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurvey.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurvey.psf** in the following PSF window.
- Select the **Impute Missing Values** option on the **Statistics** menu of the PSF window as shown below

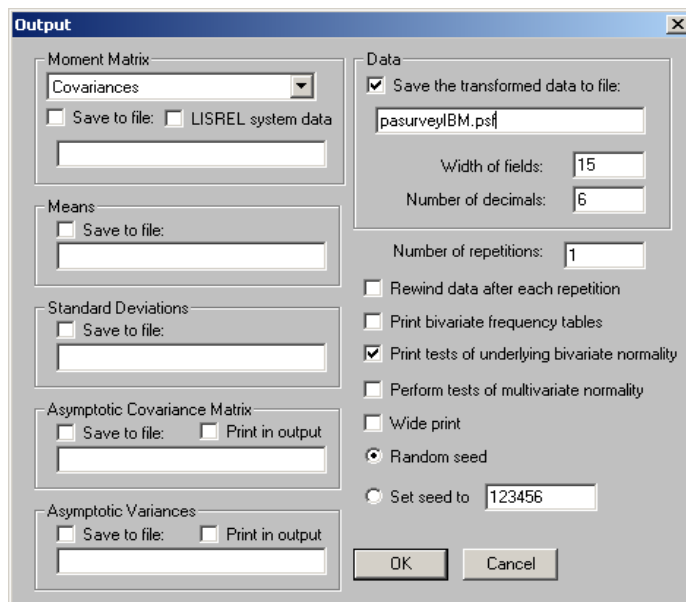


to activate the **Impute Missing Values** dialog box.

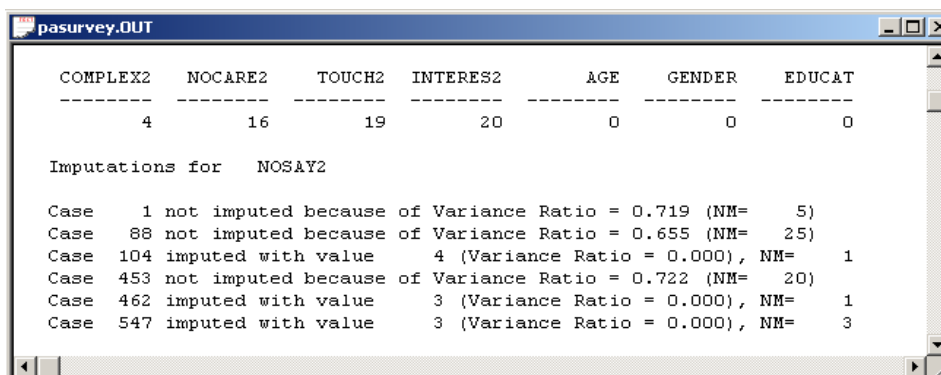
- Select the variables NOSAY2, ... , INTERES2 from the variable list box and then click on the upper **Add** button to transfer all the variables to the **Imputed variables** list box.
- Select the variables NOSAY1, ... , INTERES1 from the variable list box and then click on the lower **Add** button to obtain the following dialog box.



- Click on the **Output Options** button to obtain the **Output** dialog box.
- Check the **Save the transformed data to file** check box in the **Data** section.
- Enter the name **pasurveyIBM.psf** in the corresponding string field to produce the following dialog box.



- Click on the **OK** button to return to the **Impute Missing Values** dialog box.
- Click on the **Run** button on the **Impute Missing Values** dialog box to produce the text editor window for **pasurvey.out**. A portion of this output file is shown below.



After the imputation, the data file **pasurveyibm.psf** has 739 observations with no missing values and is saved in the same subfolder as that of **pasurvey.psf**. To view the imputed data set, we proceed as follows.

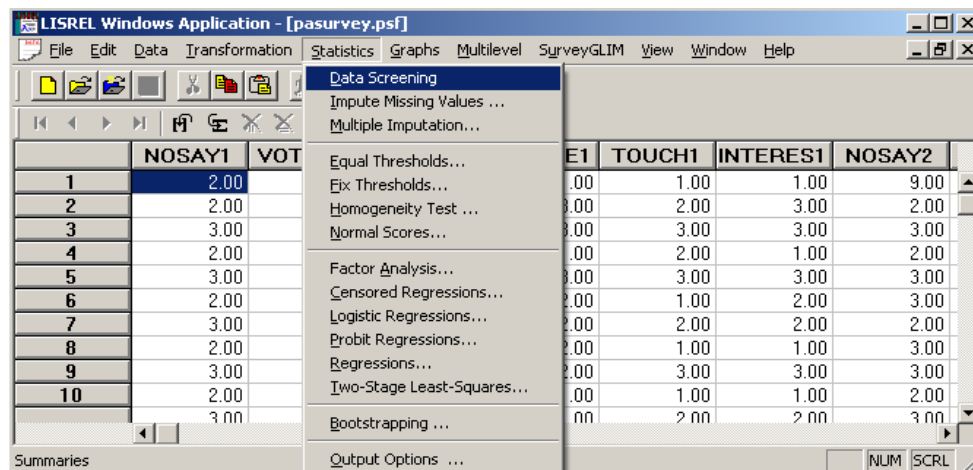
- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the following PSF window.

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2	COMPLEX2	NOCARE2
1	2.00	3.00	3.00	3.00	2.00	3.00	2.00	3.00	2.00	2.00
2	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	2.00
3	2.00	2.00	1.00	1.00	2.00	1.00	2.00	2.00	2.00	2.00
4	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00	3.00
5	2.00	2.00	2.00	2.00	1.00	2.00	3.00	2.00	1.00	3.00
6	3.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00
7	3.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00	1.00	3.00
8	2.00	2.00	3.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00
9	3.00	2.00	1.00	1.00	2.00	2.00	3.00	2.00	2.00	2.00
10	1.00	1.00	1.00	1.00	1.00	1.00	3.00	3.00	2.00	3.00

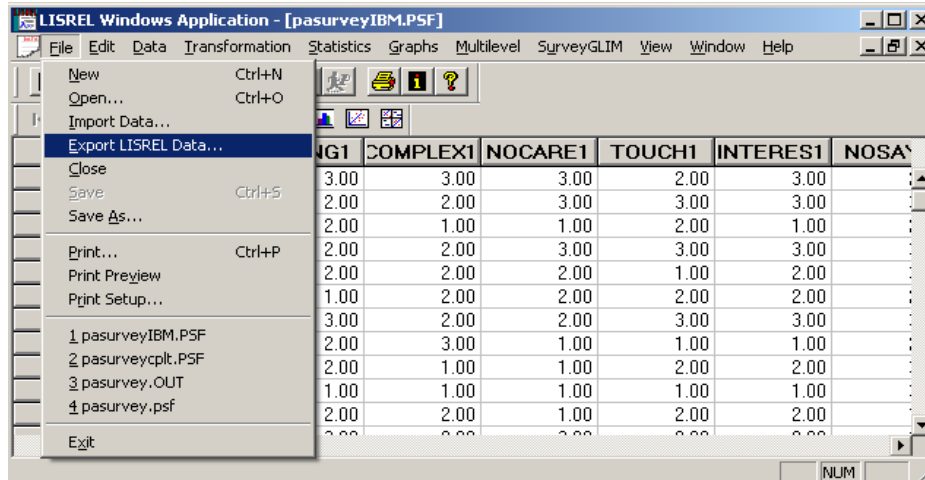
Data screening

To screen the data in **pasurvey.psf**, we proceed as follows.

- Select the **Open** option from the **File** menu.
- Select the **PRELIS Data (*.psf)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurvey.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the following PSF window.
- Click on the **Data Screening** option on the **Statistics** as shown below

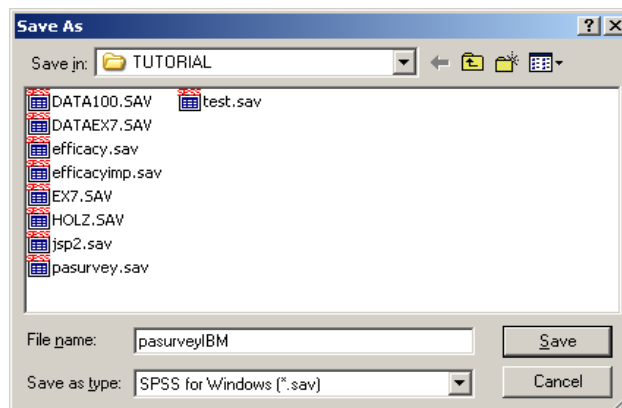


to produce the text editor window for **pasurvey.out**. A selection of this output file is shown below.



to load the **Save As** dialog box.

- From the **Save as type** drop-down list box, select the **SPSS for Windows (*.sav)** option.
- Enter **pasurveyIBM** in the **File name** string field to generate the following dialog box.



- Click on the **Save** button to save the new SPSS data file as **pasurveyibm.sav** in the **TUTORIAL** subfolder.

A portion of the SPSS data file **pasurveyibm.sav** is shown in the following **SPSS Data Editor** window.

The screenshot shows the SPSS Data Editor window for a file named 'pasurveyIBM.sav'. The window displays a data grid with 13 columns and 7 rows. The columns are labeled: NOSAY1, VOTING1, COMPLEX1, NOCARE1, TOUCH1, INTERES1, NOSAY2, VOTING2, COMPLEX2, NOCARE2, TOUCH2, and IN. The data values are as follows:

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2	COMPLEX2	NOCARE2	TOUCH2	IN
1	2	3	3	3	2	3	2	3	2	2	2	
2	3	2	2	3	3	3	3	2	2	2	2	
3	2	2	1	1	2	1	2	2	2	2	1	
4	3	2	2	3	3	3	3	2	2	3	2	
5	2	2	2	2	1	2	3	2	1	3	2	
6	3	1	2	2	2	2	2	2	3	3	2	
7	3	3	2	2	3	3	3	3	1	3	2	

Thresholds

An ordinal variable z may be regarded as a crude measurement of an underlying unobserved or unobservable continuous variable z^* . For example, a four-point ordinal scale may be conceptualized as:

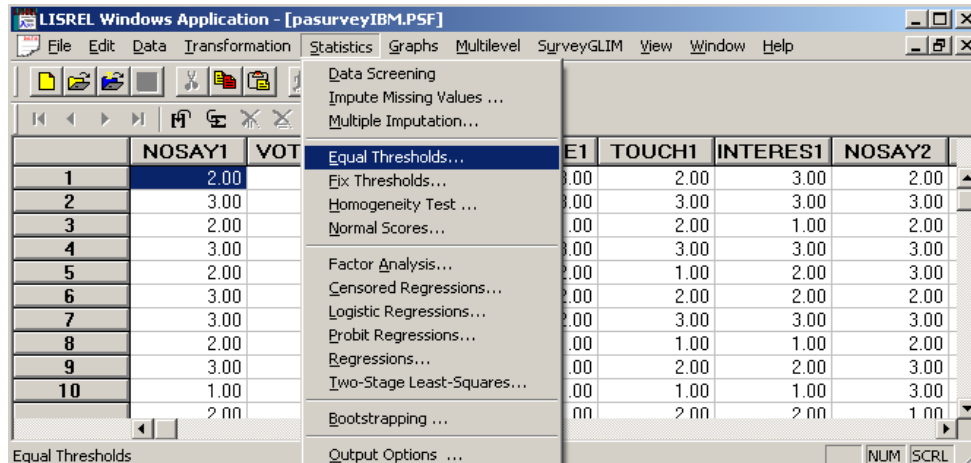
- if $z^* \leq \alpha_1$, z is scored 1,
- if $\alpha_1 < z^* \leq \alpha_2$, z is scored 2,
- if $\alpha_2 < z^* \leq \alpha_3$, z is scored 3,
- if $\alpha_3 < z^*$, z is scored 4,

where $\alpha_1 < \alpha_2 < \alpha_3$ are threshold values for z^* . It is often assumed that z^* has a standard normal distribution, in which case the thresholds can be estimated from the inverse of the normal distribution function (Jöreskog & Sörbom 1999).

Equal thresholds test and compute polychoric correlations

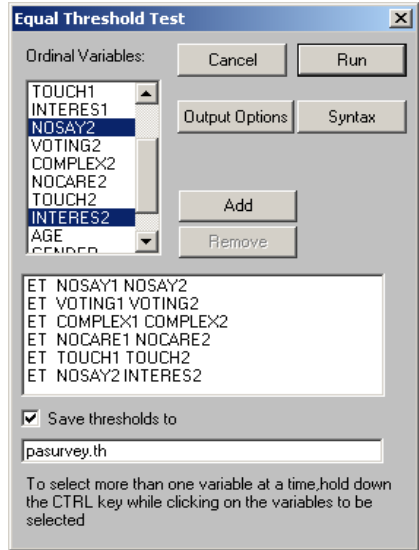
To perform an equal thresholds test for the repeated measures in **pasurveyibm.psf** and to compute the corresponding polychoric correlation matrix, we proceed as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Equal Thresholds** option on the **Statistics** menu as shown below

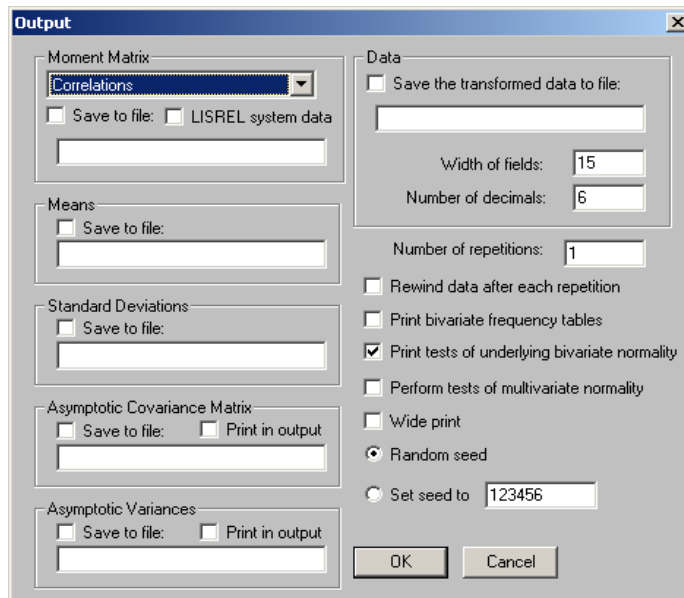


to load the **Equal Threshold Test** dialog box.

- Select the labels NOSAY1 and NOSAY2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line ET NOSAY1 NOSAY2 to the syntax box.
- Select the labels COMPLEX1 and COMPLEX2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line ET COMPLEX1 COMPLEX2 to the syntax box.
- Select the labels NOCARE1 and NOCARE2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line ET NOCARE1 NOCARE2 to the syntax box.
- Select the labels TOUCH1 and TOUCH2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line ET TOUCH1 TOUCH2 to the syntax box.
- Select the labels INTERES1 and INTERES2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line ET INTERES1 INTERES2 to the syntax box.
- Check the **Save thresholds to** check box.
- Enter the name **pasurvey.th** for the text file for the thresholds in the string field to produce the following dialog box.



- Click on the **Output Options** button to load the **Output** dialog box.
- Select the **Correlations** option from the drop-down list box in the **Moment Matrix** section to obtain the following dialog box.



- Click on the **OK** button to return to the **Equal Threshold Test** dialog box.
- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A section of this output file is shown below.

pasurveyIBM.OUT

Total Sample Size = 739

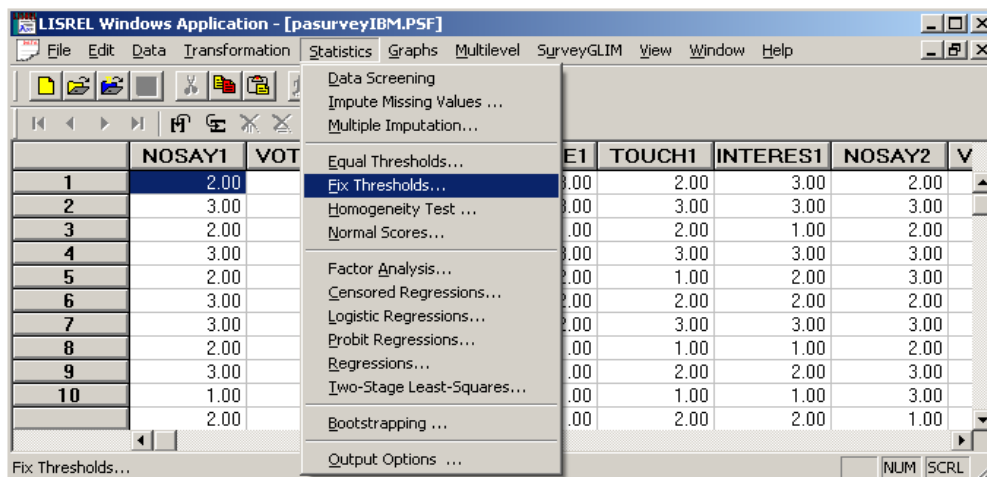
Univariate Marginal Parameters

Variable	Mean	St. Dev.	Thresholds		
NOSAY1	0.000	1.000	-1.415	-0.377	1.444
VOTING1	-0.047	1.081	-1.106	0.028	1.822
COMPLEX1	-0.032	1.050	-1.095	0.606	1.784
NOCARE1	-0.038	1.074	-1.312	0.022	1.851
TOUCH1	0.060	1.082	-1.112	0.500	2.453
INTERES1	0.000	1.000	-1.102	0.207	2.048
NOSAY2	0.336	1.059	-1.545	-0.169	1.885
VOTING2	0.047	0.912	-1.106	0.028	1.822
COMPLEX2	0.032	0.947	-1.095	0.606	1.784
NOCARE2	0.038	0.921	-1.312	0.022	1.851
TOUCH2	-0.060	0.910	-1.112	0.500	2.453
INTERES2	-0.336	0.937	-1.545	-0.169	1.885
GENDER	0.000	1.000	0.059		
EDUCAT	0.000	1.000	-0.798	0.644	

Fixed thresholds and compute polychoric correlations

To fix the thresholds of the repeated measures in **pasurveyibm.psf** and compute the corresponding polychoric correlation matrix, we proceed as follows.

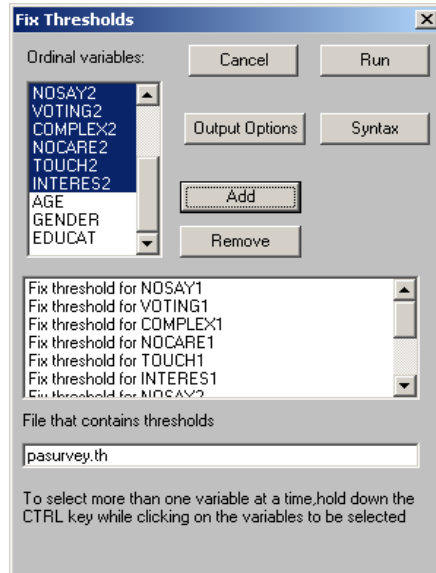
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Fix Thresholds** option on the **Statistics** menu as shown below



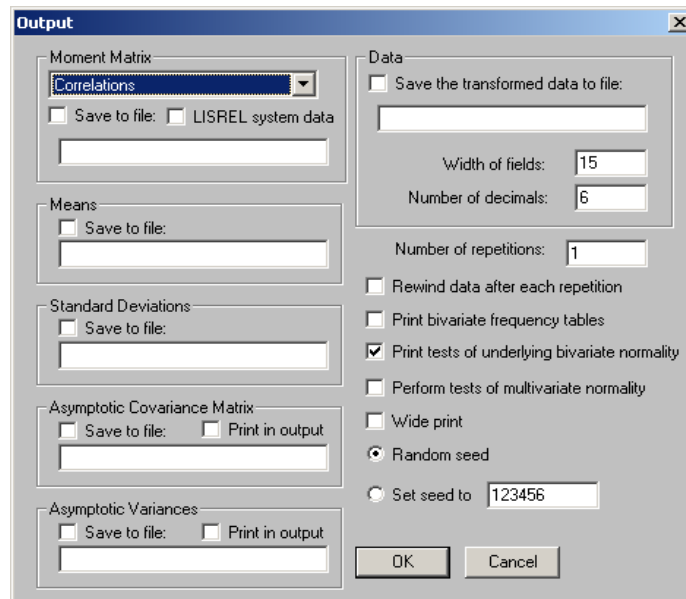
to load the **Fix Thresholds** dialog box.

- Select the first 12 variables in the **Ordinal Variables:** variable list box.

- Click on the **Add** button.
- Enter the name **pasurvey.th** for the text file that contains all the thresholds in the string field to produce the following dialog box.



- Click on the **Output Options** button to load the **Output** dialog box.
- Select the **Correlations** option from the drop-down list box in the **Moment Matrix** section to create the following dialog box.



- Click on the **OK** button to return to the **Fixed Threshold Test** dialog box.
- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A portion

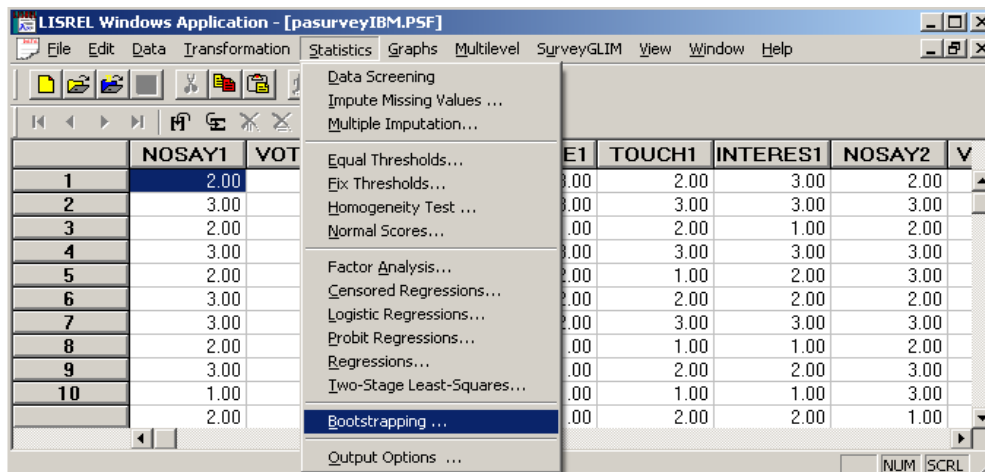
of this output file is shown below.

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1
NOSAY1	1.000					
VOTING1	0.455	1.000				
COMPLEX1	0.440	0.322	1.000			
NOCARE1	0.676	0.321	0.417	1.000		
TOUCH1	0.505	0.281	0.347	0.674	1.000	
INTERES1	0.551	0.320	0.378	0.693	0.764	1.000
NOSAY2	0.463	0.258	0.228	0.401	0.283	0.331
VOTING2	0.270	0.467	0.277	0.239	0.246	0.282
COMPLEX2	0.272	0.234	0.464	0.244	0.158	0.212
NOCARE2	0.457	0.251	0.286	0.503	0.406	0.464

Bootstrapping

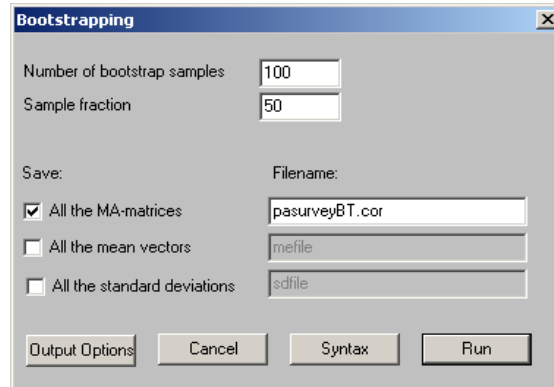
Bootstrapping is the method of drawing random samples from an original sample. We now use **pasurveyibm.psf** to illustrate how to obtain bootstrap samples with PRELIS. Suppose we want to draw 100 samples of size 50 and study the correlation matrices of these 100 samples. This is achieved as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Bootstrapping** option on the **Statistics** menu as shown below

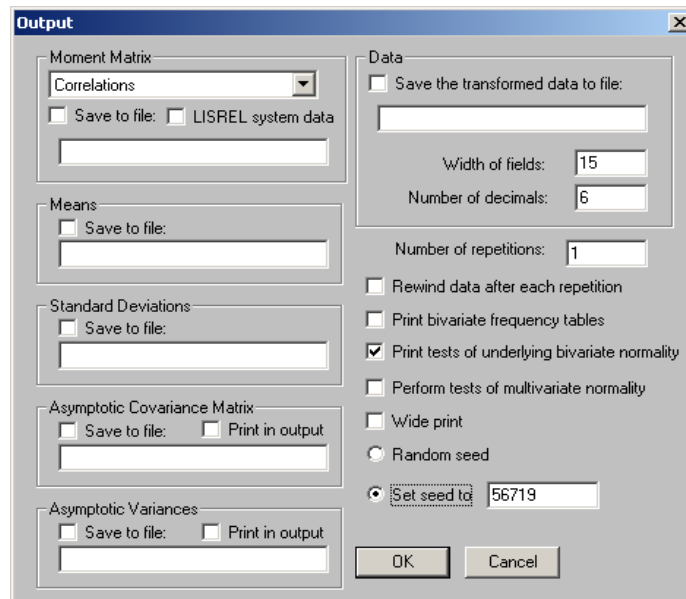


to load the **Bootstrapping** dialog box.

- Check the **All the MA-Matrices** check box.
- Enter the name **pasurveyBT.cor** in the corresponding string field to produce the following dialog box.



- Click on the **Output Options** button to load the **Output** dialog box.
- Select the **Correlations** option from the drop-down list box in the **Moment Matrix** section.
- Check the **Set seed to** radio button.
- One may set the random number generator seed, so that this run can be replicated exactly. For this example, we enter 56719 in the **Set seed to box** to produce the following dialog box.



- Click on the **OK** button to return to the **Bootstrapping** dialog box.
- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A portion of this output file is shown below.

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1
NOSAY1	1.000					
VOTING1	0.454	1.000				
COMPLEX1	0.447	0.323	1.000			
NOCARE1	0.684	0.318	0.409	1.000		
TOUCH1	0.508	0.274	0.347	0.679	1.000	
INTERES1	0.550	0.316	0.376	0.690	0.766	1.000
NOSAY2	0.473	0.265	0.238	0.402	0.294	0.338
VOTING2	0.305	0.518	0.311	0.272	0.284	0.313
COMPLEX2	0.283	0.242	0.455	0.244	0.164	0.215
NOCARE2	0.464	0.250	0.274	0.497	0.409	0.462
TOUCH2	0.305	0.225	0.161	0.295	0.381	0.418

The 100 correlation matrices are saved to the text file **pasurveybt.cor**. We access this file as follows.

- Select the **Open** option on the **File** menu.
- Select the **All Files (*.*)** option on the **Files of type** drop-down list box.
- Browse for the file **pasurveyBT.cor** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the text file **pasurveyBT.cor**, which includes 100 correlation matrices, each with $\frac{p \times (p+1)}{2} = 120$ elements, where $p = 15$ is the number of variables. The correlation matrices for the first two bootstrap samples are shown below.

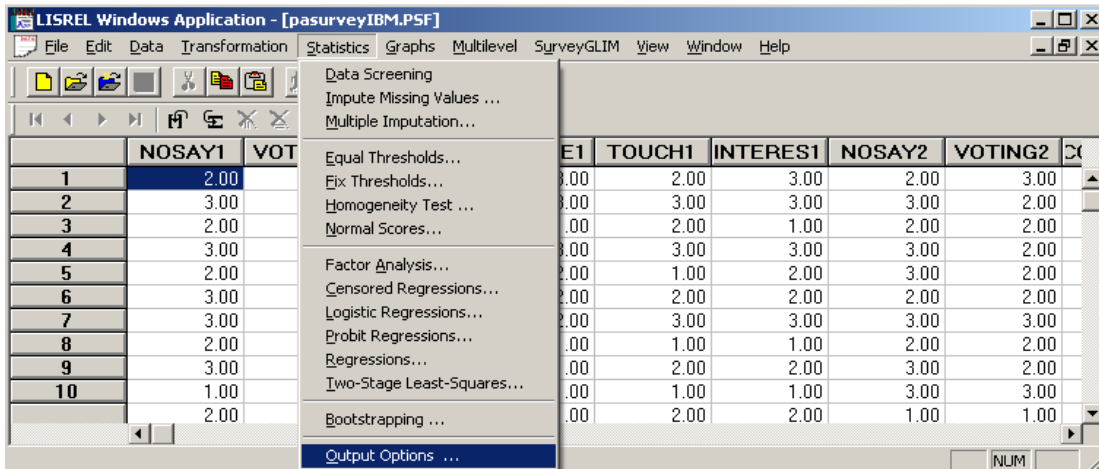
0.10000D+01	0.54648D+00	0.10000D+01	0.40371D+00	0.32589D+00	0.10000D+01
0.60325D+00	0.40848D+00	0.42232D+00	0.10000D+01	0.54275D+00	0.31578D+00
0.32806D+00	0.74785D+00	0.10000D+01	0.51217D+00	0.40059D+00	0.41302D+00
0.77023D+00	0.74810D+00	0.10000D+01	0.45305D+00	0.34825D+00	0.24834D+00
0.39803D+00	0.29455D+00	0.34400D+00	0.10000D+01	0.27844D+00	0.49726D+00
0.22495D+00	0.27679D+00	0.22638D+00	0.23351D+00	0.51564D+00	0.10000D+01
0.29169D+00	0.29335D+00	0.58041D+00	0.31884D+00	0.27845D+00	0.34568D+00
0.44971D+00	0.29300D+00	0.10000D+01	0.38043D+00	0.31060D+00	0.33012D+00
0.44481D+00	0.41270D+00	0.40884D+00	0.56075D+00	0.47618D+00	0.43027D+00
0.10000D+01	0.22471D+00	0.18436D+00	0.14473D+00	0.28211D+00	0.38506D+00
0.30428D+00	0.31088D+00	0.39635D+00	0.22178D+00	0.68533D+00	0.10000D+01
0.24081D+00	0.27990D+00	0.24275D+00	0.34763D+00	0.32829D+00	0.40379D+00

Computing matrices

The correlation matrix and the estimated asymptotic covariance matrix of the correlations can be computed and saved by PRELIS interactively. We illustrate this using the data in **pasurveyibm.psf**.

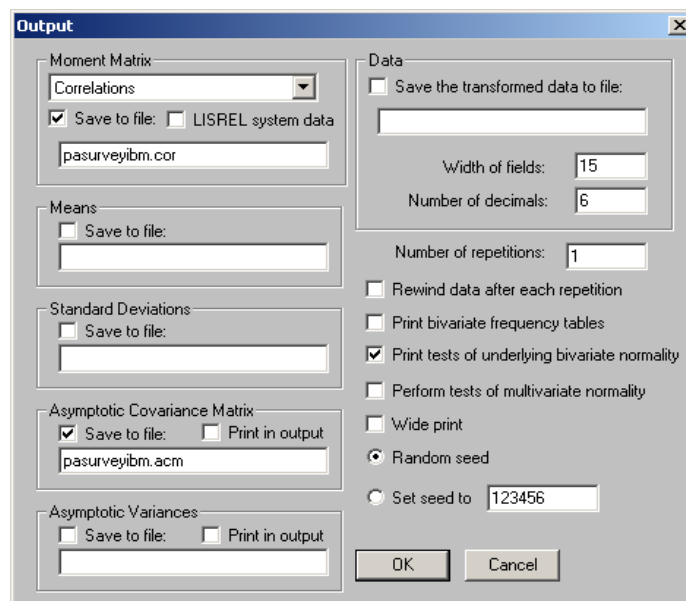
- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.

- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Output Options** option on the **Statistics** menu as shown below



to load the **Output** dialog box.

- Select the **Correlations** option from the drop-down list box in the **Moment Matrix** section.
- Check the **Save to file** check box in the **Moment Matrix** section.
- Enter the name **pasurveyibm.cor** in the corresponding string field.
- Check the **Save to file** check box in the **Asymptotic Covariance Matrix** section.
- Enter the name **pasurveyibm.acm** in the corresponding string field to produce the following dialog box.



- Click on the **OK** button to open the text editor window for **pasurveyibm.out**. A portion of

the output file is shown below.

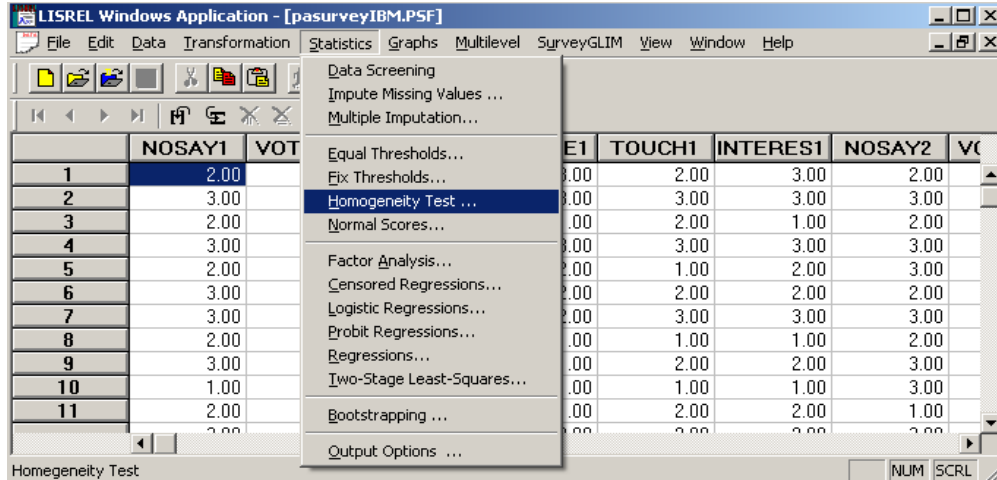
	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1			
NOSAY1	1.000								
VOTING1	0.455	1.000							
COMPLEX1	0.440	0.322	1.000						
NOCARE1	0.676	0.321	0.417	1.000					
TOUCH1	0.505	0.281	0.347	0.674	1.000				
INTERES1	0.551	0.320	0.378	0.693	0.764	1.000			
NOSAY2	0.465	0.258	0.231	0.402	0.284	0.334			
VOTING2	0.292	0.510	0.306	0.264	0.268	0.307			
COMPLEY2	0.272	0.234	0.464	0.244	0.158	0.212			

Thus, the correlation matrix is saved as the text file **pasurveyibm.cov** in the same subfolder as that of **pasurveyibm.psf**. The corresponding estimated asymptotic covariance matrix is saved as the binary file **pasurveyibm.acm** in the same subfolder.

Homogeneity test

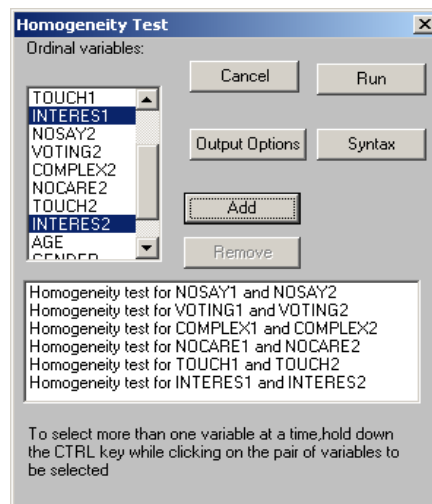
The homogeneity test (see Jöreskog & Sörbom 1999, pages 173-176) is a test of the hypothesis that the marginal distributions of two categorical variables with the same number of categories (k) are the same. The χ^2 statistic for testing this hypothesis has $k - 1$ degrees of freedom. The homogeneity test is particularly useful in longitudinal studies to test the hypothesis that the distribution of a variable has not changed from one occasion to the next. We now use the data in **pasurveyibm.psf** in the **TUTORIAL** subfolder to demonstrate the homogeneity test. This is accomplished interactively as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Homogeneity Test** option on the **Statistics** menu as shown below



to load the **Homogeneity Test** dialog box.

- Select the labels NOSAY1 and NOSAY2 in the **Ordinal Variables:** variable list.
- Click on the **Add** button to add the line Homogeneity test for NOSAY1 and NOSAY2 to the syntax box.
- Repeat this for COMPLEX1 and COMPLEX2, NOCARE1 and NOCARE2, TOUCH1 and TOUCH2 and INTERES1 and INTERES2 to produce the following dialog box.



- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A portion of this output file is shown below.

Variable vs. Variable	Chi-Squ.	D.F.	P-Value
NOSAY1 vs. NOSAY2	15.626	3	0.001
VOTING1 vs. VOTING2	26.674	3	0.000
COMPLEX1 vs. COMPLEX2	7.307	3	0.063
NOCARE1 vs. NOCARE2	16.422	3	0.001
TOUCH1 vs. TOUCH2	13.286	3	0.004
INTERES1 vs. INTERES2	13.093	3	0.004

Logistic regression example using political action survey data

In practice, it often occurs that the response variable of a study is ordinal rather than continuous. In this section, logistic regression analysis is more appropriate than multiple linear regression analysis. We now demonstrate how to use PRELIS to fit a logistic regression model to data.

The model

When the response variable follows a multinomial or Bernoulli distribution, a model that accommodates the pattern is the logistic functional form

$$y = \frac{1}{1 + e^{-(\mathbf{x}'\boldsymbol{\beta} + \varepsilon)}}.$$

The transformation to linearity is

$$\text{logit}(y) = \ln \frac{y}{1-y} = \mathbf{x}'\boldsymbol{\beta} + \varepsilon,$$

where y is an endogenous or jointly dependent variable, $\mathbf{x} = (x_1, x_2, \dots, x_q)'$ is a set of exogenous variables, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_q)$ is a set of unknown parameters and ε is the error term.

The logistic model to be fitted to the data in **pasurveyibm.psf** may be expressed as

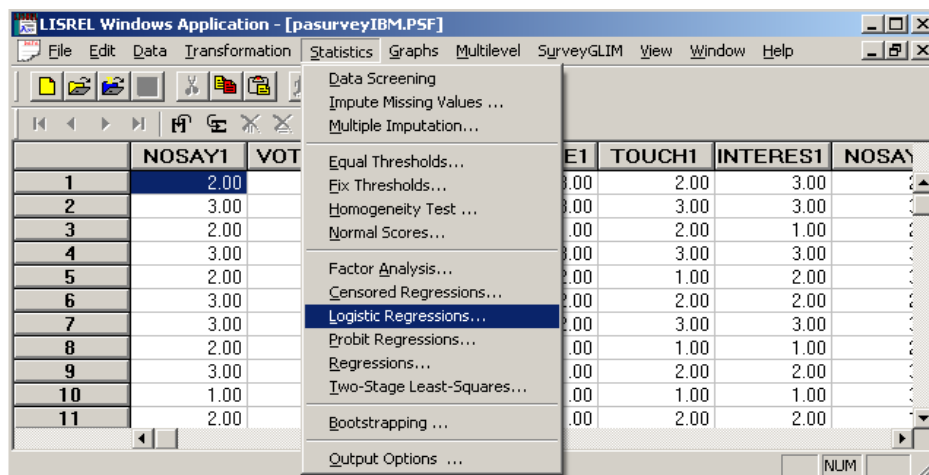
$$\text{logit}(NOSAY1) = \beta_0 + \beta_1 AGE + \beta_2 GENDER + \beta_3 EDUCAT + \varepsilon,$$

where β_0 denotes the intercept, β_1 , β_2 and β_3 denote unknown regression weights and ε denotes an error term.

The analysis

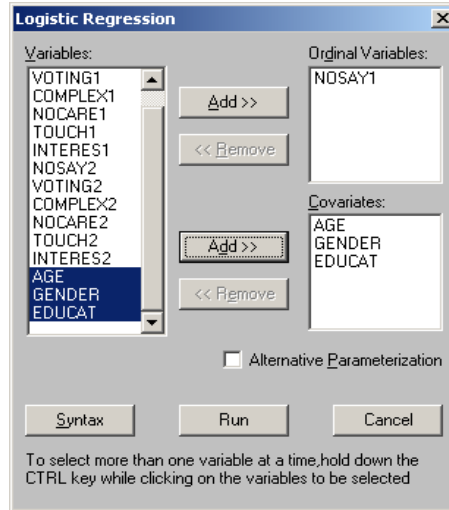
We fit the model above to the data in **pasurveyibm.psf** as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Logistic Regressions** option on the **Statistics** menu as shown below

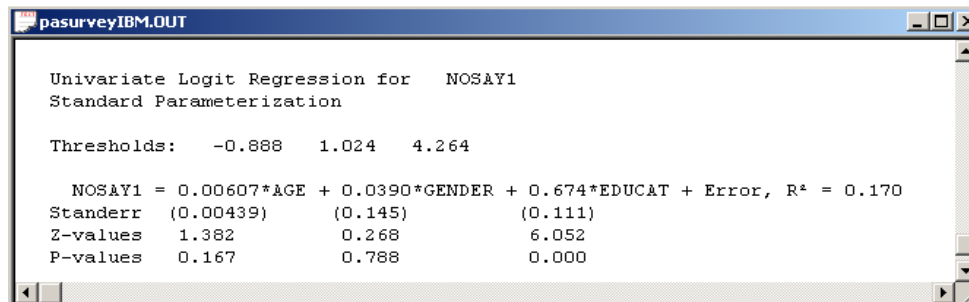


to load the **Logistic Regression** dialog box.

- Select the label **NOSAY1** in the **Variables** list box.
- Click on the top **Add** button to transfer the label **NOSAY1** to the **Ordinal Variables:** list box.
- Select the labels **AGE**, **GENDER** and **EDUCAT** in the **Variables** list box.
- Click on the bottom **Add** button to produce the following dialog box.



- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A selection of this output file is shown below.



Multivariate probit regression example using political action survey data

Probit regression analysis is very useful when the response variable is ordinal rather than continuous. In this section, we use PRELIS to fit a multivariate probit regression model to data interactively.

The model

The general probability model for the probit model is given by

$$\text{Prob}(y \leq j) = \Phi(\boldsymbol{\beta}'\mathbf{x}).$$

where y is an endogenous or jointly dependent variable, $\mathbf{x} = (x_1, x_2, \dots, x_q)'$ is a set of exogenous variables, $\boldsymbol{\beta} = (\beta_1, \beta_1, \dots, \beta_q)$ is a set of unknown parameters and Φ denotes the cumulative

distribution function of the standard Normal distribution.

The specific multivariate probit regression model to be fitted to the data in **pasurveyimp.psf** is given by

$$\text{Prob}(NOSAY1 \leq j) = \Phi(\beta_0 + \beta_1 AGE + \beta_2 GENDER + \beta_3 EDUCAT), \quad j = -0.527, 0.548, 2.425$$

and

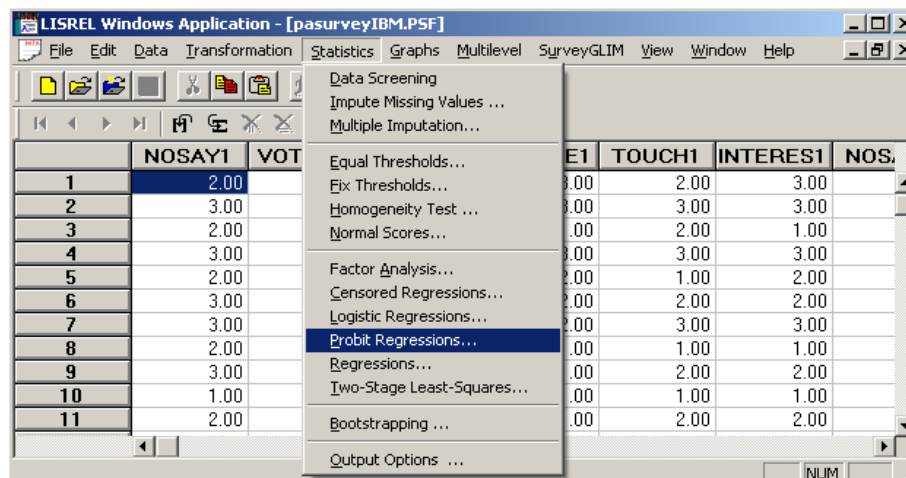
$$\text{Prob}(NOSAY2 \leq j) = \Phi(\beta_0 + \beta_1 AGE + \beta_2 GENDER + \beta_3 EDUCAT), \quad j = -0.980, 0.347, 2.324$$

where β_0 , β_1 , β_2 and β_3 denote unknown regression weights.

The analysis

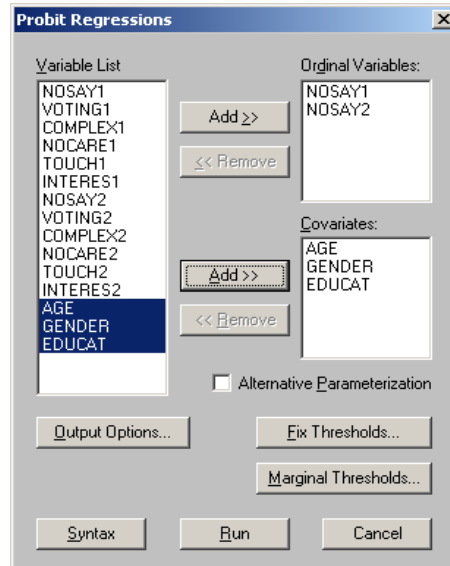
We fit the model above to the data in **pasurveyibm.psf** as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select **PRELIS data (*.psf)** from the **Files of type** drop-down list box.
- Browse for the file **pasurveyibm.psf** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open **pasurveyibm.psf** in a PSF window.
- Select the **Probit Regressions** option on the **Statistics** menu as shown below

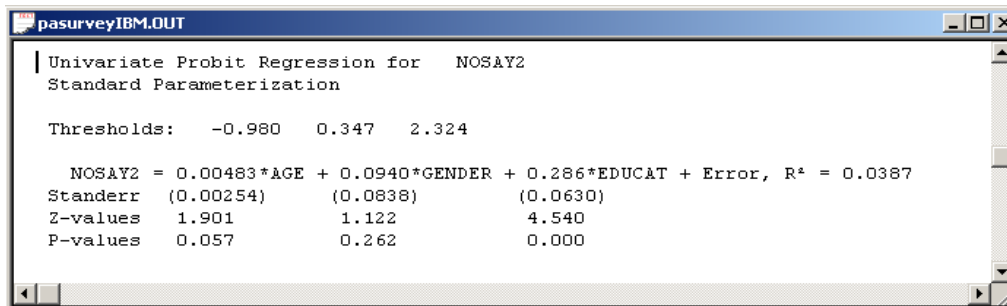


to load the **Probit Regressions** dialog box.

- Select the labels NOSAY1 and NOSAY2 in the **Variables** list box.
- Click on the top **Add** button to transfer the label NOSAY1 and NOSAY2 to the **Ordinal Variables:** list box.
- Select the labels AGE, GENDER and EDUCAT in the **Variables** list box.
- Click on the bottom **Add** button to produce following dialog box.



- Click on the **Run** button to produce the text editor window for **pasurveyibm.out**. A portion of this output file is shown in the following text editor window.



The Chi-square values are written to the text file **pasurveyibm.fit**. We access this file as follows.

- Select the **Open** option on the **File** menu to load the **Open** dialog box.
- Select the **Output (*.out; *.fit)** option from the **Files of type:** drop-down list box.
- Browse for the file **pasurveyibm.fit** in the **TUTORIAL** subfolder and select it.
- Click on the **Open** button to open the following text editor window for **pasurveyibm.fit**.

The screenshot shows a window titled "pasurveyibm.fit" with a table of fit statistics. The table has five columns: Variable, -2lnL, Chi-square, df, and Covariates. Two rows of data are shown for variables NOSAY1 and NOSAY2. The Covariates column lists AGE, GENDER, and EDUCAT for both variables.

Variable	-2lnL	Chi-square	df	Covariates
NOSAY1	1539.431	38.143	3	AGE GENDER EDUCAT
NOSAY2	1410.378	22.945	3	AGE GENDER EDUCAT

References

- Barnes, S.H. & Kaase, M. (Eds.). (1979). *Political Action: Mass Participation in Five Western Democracies*. Beverley Hills and London: Sage Publications.
- Du Toit, S.H.C., Steyn, A.G.W. & Stumpf, R.H. (1986). *Graphical Exploratory Data Analysis*. New York: Springer-Verlag.
- Du Toit, M. & Du Toit, S.H.C. (2001). *Interactive LISREL User's Guide*. Lincolnwood, IL: Scientific Software International, Inc.
- Goldberger, A.S. (1964). *Econometric theory*. Wiley: New York.
- Holzinger, K., and Swineford, F. (1939). *A study in factor analysis: The stability of a bifactor solution*. Supplementary Educational Monograph no. 48. University of Chicago Press: Chicago.
- Jöreskog, K.G. (2002a). *Structural Equation Modeling with Ordinal Variables using LISREL*. Available at <http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf>.
- Jöreskog, K.G. (2002b). *Censored variables and censored regression*. Available at <http://www.ssicentral.com/lisrel/techdocs/censor.pdf>.
- Jöreskog, K.G. (2004). *Multivariate censored regression*. Available at <http://www.ssicentral.com/lisrel/techdocs/multcens.pdf>.
- Jöreskog, K.G. & Sörbom, D. (1999). *PRELIS 2: User's Reference Guide*. Lincolnwood, IL: Scientific Software International, Inc.
- Jöreskog, K.G. & Sörbom, D. (2005). *LISREL 8.7 for Windows [Computer Software]*. Lincolnwood, IL: Scientific Software International, Inc.
- Klein, L.R. (1950). *Economic fluctuations in the United States 1921-1941*. Cowles Commission Monograph No. 11.