Introduction:

  This document ties together the four different files that make up the GRASP example for simulating missing data. In this simulation study, we made up pseudo 'missing' values by removing observed values from original datasets. The pseudo missing data were then imputed by two multiple imputation methods: sequential method vs. simultaneous method. We compared results between sequential and simultaneous methods using four criteria: (1) root mean square deviation (RMSD) = $(\sum(y-\hat{y})2/n)1/2$ where y is the true value, $\hat{y}$ is the imputed value, and n is the number of pseudo "missing" values; (2) mean absolute deviation (MAD) = $\sum|y-\hat{y}|/n$; (3) bias (BIAS) = $\sum(y-\hat{y})/n$; and (4) proportionate variance (PV) = $Var(\hat{y})/Var(y)$. The sample data used in this simulation study was extracted from the Cardiovascular Health Study (CHS), which included seven waves of observations on 5,888 adults aged 65 years or older.

Keyword Categories:

Clinical: Longitudinal study, Cardiovascular Health Study

Genetic: Not Applicable

Statistical: Missing data, multiple imputation, simulation study

Software: SAS

Related:

References:

SAS Institute. (2009). *SAS/STAT 9.2 Production: Procedures for Windows*. Cary, NC: SAS Institute Inc.

Ning Y, McAvay G, Chaudhry SI, & Allore HG. Multiple Imputation Methods for Longitudinal Data in the Cardiovascular Health Study: Sequential Approach vs. Simultaneous Approach. Unpublished manuscript.

Component Files:

a.

SAS program: missing_simulation_code.txt

b.

SAS data sample: missing_simulation_data.txt

c.

SAS output: missing_simulation_out.txt

d.

PDF file explaining entire example: missing_simulation_summary.pdf (file you are reading)

Optimal Use:

1.

Read this Summary file completely; Component d listed above.

Run the SAS program in concert with the data files; Components a & b above.

Generate output that should appear like that shown in the output.txt file; Component c listed above.

Prerequisites:

To run the example you must have a working knowledge of SAS that allows you to compile the SAS program, load the data file, and finally get results. The online SAS Help provides a number of step by step examples that can get you up to speed to run this example.

Potential Applications:

The contents here may be useful to folks performing a simulation study that compares different imputation methods. Any feedback on how to make this submission more useful is welcome. Log into the main GRASP page, i.e.

http://grasp.med.yale.edu

and choose the "forum" link in the upper menu on the right to enter feedback on any GRASP submission. Please refer to the specific file names when commenting so we can appropriately steer your suggestions.

.