Program Abstracts/Algorithms

TIMESAMPLE: A FORTRAN IV program for developing time sampling strategies

SAMUEL W. McDOWELL

University of Vermont, Burlington, Vermont 05401

TIMESAMPLE is a program to assist in developing a strategy for accumulating data based on random samples from a continuous experiment. The program is able to handle missing data and unscorable responses. It is a means of providing an estimate of the validity of sampling procedures, based on various time sampling intervals.

The program computes Pearson correlation coefficients for random time samples of subjects' behaviors, based on different categories of responses. The classification scheme may consist of two to twelve categories to describe individual responses. The basic time unit is 1/25 of the total experiment (time sampling unit), and the larger time sampling intervals are calculated by combining one to five successive units. Correlations are found using intervals of one, two, three, four, and five units. Therefore, there are 25 one-unit sampling intervals, 12 two-unit intervals, eight three-unit intervals, six four-unit intervals, and five five-unit intervals. This enables the experimenter to choose the interval that produces the highest correlation in the shortest period of time. The experimenter may choose the intervals he wishes to investigate at program execution time.

After the respective time intervals are computed, they are randomly combined. The intervals chosen are the same for each subject. For example, by using one-unit sampling intervals, there would be 25 correlations, each correlation based on successive combinations of one to 25 sampling intervals. Each correlation is based on a different random combination of sampling intervals.

The following correlations are computed for each successive combination:

- (1) correlation by subject (across categories);
- (2) correlation by category (across subjects);
- (3) correlation by total combined interval.

The means of the subject and category correlations are found after conversion to a Fisher Z, and are transformed back into Pearson correlations. There is then a single correlation for a category profile, subject profile, and group profile.

Input/output. The initial program input consists of data describing the size of the intervals to be sampled, the number of categories in the classification scheme, and information which will be used to produce headings and labels on the output. Following this information is the data which contains one subject ID, category, and interval per record.

Program output consists of a raw data matrix for each interval. The matrix is size $s \times c$, where s is the number of subjects and c is the number of categories. The matrix contains frequency counts and percentages, based on the total interval. Following this output is a listing of Pearson correlations, the Fisher Z-transformations, and the means for each size sampling interval. The Fisher correlations for the group profile are also presented graphically, providing a means of inspecting trends for varying sizes of sampling intervals and cumulative combinations of those intervals.

Computer. The program was written in FORTRAN IV, and developed on a Xerox Sigma VI computer running under UTS. Program execution time depends on the number of subjects and categories, the amount of data, and the size of the intervals to be tested.

Limitations. The program in its present form will handle two to fifteen subjects, and two to twelve categories. Program execution time will vary depending on these two variables, and the specific computer on which TIMESAMPLE is run. For three subjects and seven categories, the program used approximately 58K and one minute of CPU time.

Availability. A source listing, sample data, and sample output may be obtained from Samuel W. McDowell, Department of Psychiatry, Medical Alumni Building, University of Vermont, Burlington, Vermont 05401.