RELAN - Relation Analysis

RELAN Fact Sheet

Main Purpose

- Analyzing and decomposition of hypotheses (relations) in propositional logic notation

- Statistical testing and evaluating the effect of logical hypotheses (relations)

- Simulation of logical hypotheses and significance testing

- Chance modeling of the assumption of independence of variables

- Evaluation of causal relations between variables based on time relations

- Graph theoretical analysis of variable networks

Main Features

Relation analysis is a contingency statistically analysis by which propositional, statistical, and causal relations between a system of dichotomous variables can be demonstrated. This method allows (1) the detection of all possible relations between a set of given variables. Hypothetical relations can be (2) tested, (3) explored and (4) simulated. Relations can be evaluated by several statistical parameters, relations can be compared, simplified, and used for the purpose of prediction. Hypothetical effects of alternative causes, causal conditions, moderators, mediators and confounders can be specified, tested, and - if necessary - from the structure of the data extracted.

Statistical Calculations

- Propositional analysis of hypothesis (decomposition, tautologies, contradictions)
- Information analysis of data
- Adaptation of chance models
- Cluster Analysis (propositional basis)
- Binomial test (approximated)
- Chi-Square test
- Likelihood-ratio test (G-Test)
- Pattern analysis (Types, Antitypes)
- Precision statistic (for hypothesis)
- Normalisation of Z-Values
- Intercorrelation of (sub-)relations
- Causal effect parameters
- Simulation analysis (propositional, statistical)
- Sensitivity
- Specificity

- Positive predictive values
- Negative predictive values
- Yules Q
- Phi-Correlation
- Mutual information
- Odds ratio
- Contingency coefficient
- Cramers V
- Multiple testing correction
- Power analysis
- Graph theoretical Analysis (components, spanning trees, …)
- Elementary relation analysis (Z-values, P-values, power)
- Complex relation analysis (Z-values, P-values, power)
- Higher order relations
- Data mining

(see also RELAN Statistics)

Coding Languages

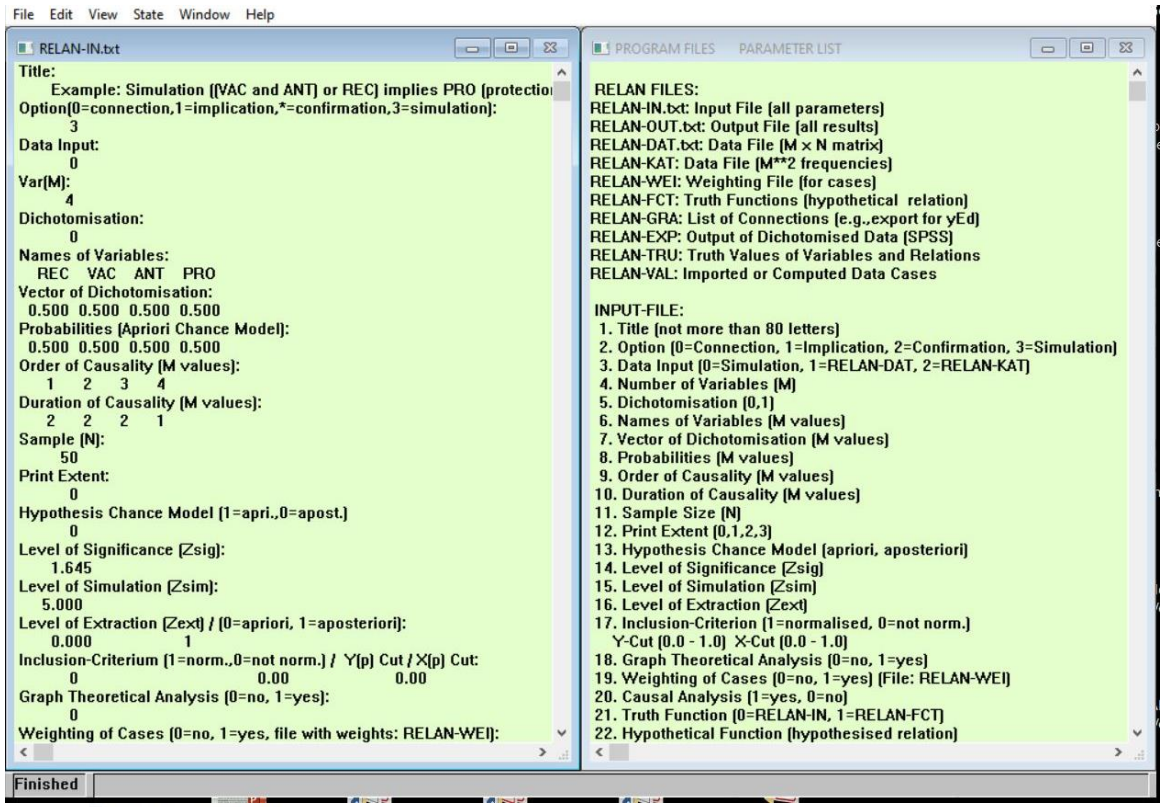
The basics of the source code was originally programmed with Fortran 77, Fortran 90 and later with DIGITAL Visual Fortran version 6.0, it has now more than 7000 statements and more than 30 subroutines.

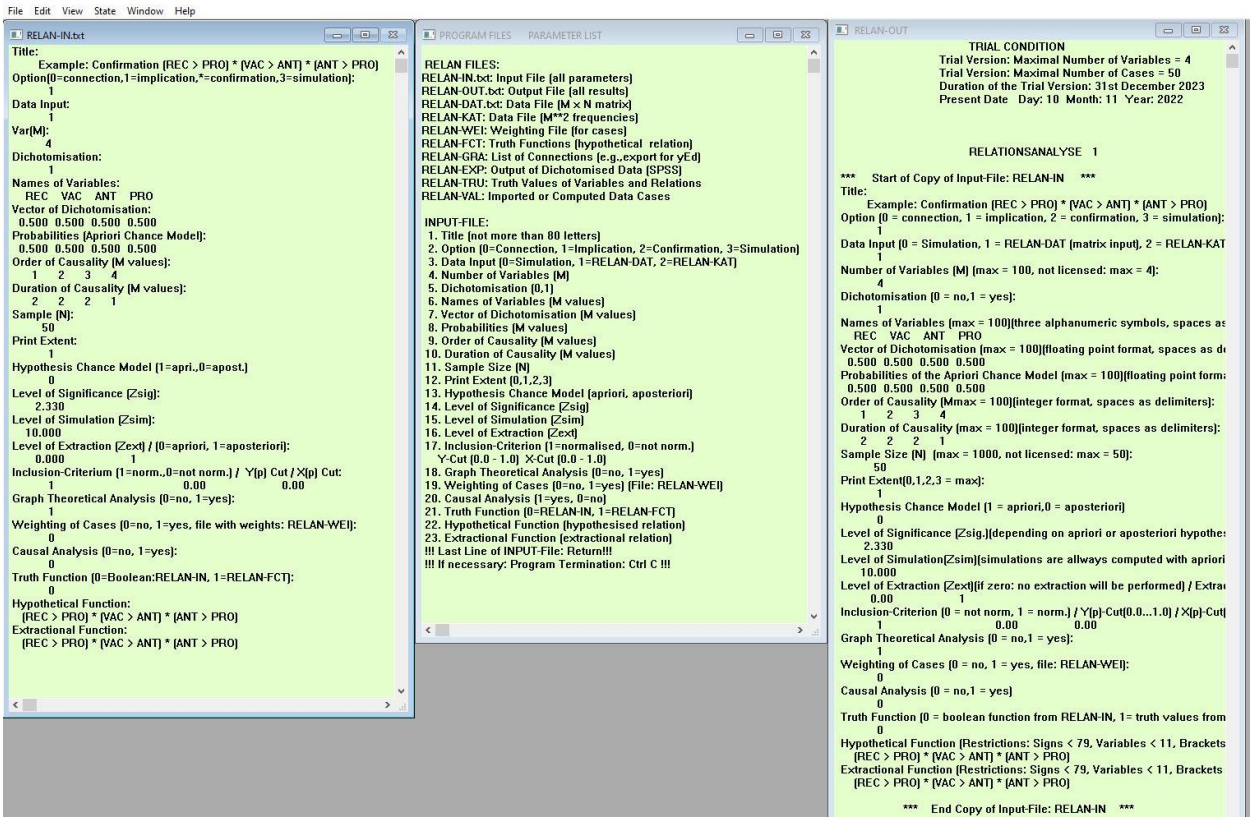
Program Flow

For the installation of the RELAN program, it is useful to create a new directory because the program creates besides the input-file RELAN-IN and the data-file RELAN-DAT another eight files: RELAN-OUT, RELAN.EXP, RELAN-FCT, RELAN-GRA, RELAN-KAT, RELAN-TRU, and RELAN-WEI. When using the Graphical User Interface (GUI), the limitations of the input parameters are shown, and only the valid parameter inputs are accepted by the program and afterwards entered into the RELAN-IN file.

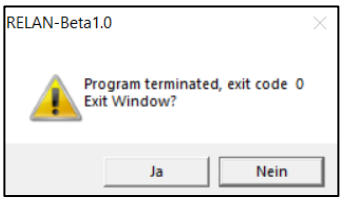
If the program is started first the license conditions are checked. Without a license only 4 variables and not more than 50 cases are processed. If the license is valid, then the computation without any restrictions begins (100 variables, 1000 cases).

After the start of the computations two light green windows are displayed. The left window shows the (running) output (RELAN-OUT), and on the top of the right window all files are listed which are available in the program and beneath of them a short description of the input parameters is offered. Moving away the upper left window then the content of the present Input file (RELAN-IN) is displaed.





These windows only should give an outline of the input, an overview of the files and the output, and they should help to detect quickly potential errors. All files have the text format (.txt) and can be modified and printed in different fonts (Recommended are non-proportional fonts: Courier New, Century Gothic, Lucida Console). If the computations are finished the following window will appear:



By pushing "Yes" all windows are closed and the program flow is terminated. If one wants to cancel the running calculations (perhaps because of an endless loop or the printout could be too extensive), then the exit can be achieved at any time by pressing the keys CTRL C together. When the program is interrupted by itself then relevant error messages are displayed in the output file (RELAN-OUT). If there happen serious calculation errors, the program flow is ended immediately.

Setting of Input Parameters

Always **all** twenty-six types of parameters in the RELAN-IN file must be given - even if they are not required in the intended evaluation run. This can be managed directly by changing the parameters already available in the provided examples of input files or indirectly by using the existing Graphical User Interface (GUI). If not all parameters are declared then the program produces serious input faults. The line before each parameter contains only descriptions and limitations of the subsequent parameter. These instructional lines can be modified as desired because they are not imported into the program. The parameters can be entered in free format with blanks as delimiters (but considering the decimal point at some parameters). In the RELAN-IN-Commentary (file) the parameters of RELAN-IN are briefly interpreted.

This input block of parameters also is printed in every output file (at the beginning) and can be copied and pasted in a new RELAN-IN. For each parameter an accompanying commentary is attached in the line before which should help to decide the correct entry.

Program Files

INPUT and DATA Files (RELAN-IN, RELAN-FCT, RELAN-WEI; RELAN-DAT, RELAN-KAT)

The parameters (RELAN-IN) contain all important control inputs for the program run, including the propositional logic functions to be checked or simulated; however, if the function is not entered as a propositional logical formula, it can be stated in the form of its logical values, but then the file RELAN-FCT (truth function file) must be used. The values have to be in the sequence of the correct order of minterms, and the number of values is 2^M (M = number of variables). In the file RELAN-WEI the weights for all cases can be entered if the parameter “Weighting of cases” is set to “yes”. For the most incorrect entries, error messages will appear in the output file which are often followed by a program stop.

The data can be entered as a matrix (RELAN-DAT), which consists of columns

(variables) and rows (cases), or as a sequence of K numbers (RELAN-KAT), which contains the frequencies of cases for each minterm of the relation ($K = 2^M$). The input of the minterm frequencies must be done in the order in which the minterms are ordered ascending as a digital number (e.g., for three variables of 0-0-0, 0-0-1, ... , 1-1-0, 1-1-1; see also RELAN-OUT). When entering data in a matrix, weights of credibility can also be assigned for the individual cases (in RELAN-WEI).

> !!! All entries in the Input files must be completed with a return (line feed) at the end, otherwise the error message "end-of-file during read" follows with a program termination!!!

OUTPUT Files (RELAN-OUT, RELAN-EXP, RELAN-GRA, RELAN-TRU, RELAN-VAL)

The output quantity in the file RELAN-OUT can be modified by the control parameter Print Extent (0 = min., 1, 2, 3 = max.). The simulated or empirically recorded data of the variables are exported (matrix) in the file RELAN-EXP (for further processing e.g., in SPSS, Statista, SAS, R) and purely in the file RELAN-VAL. In order to be able to create graphical representations of variable networks by external programs (e.g., yEd Graph Editor, Decision Explorer), all variable pairs which are significantly linked are written to the file RELAN-GRA. The file RELAN-TRU contains the truth functions of all variables, those of the subrelations as well as those of the ultimate relation to give the opportunity for further external checks.

Causal Indices

The Effect Order Index (Order of Causality) and the Effect Duration Index (Duration of Causality) describe the causal interaction assumptions between the variables. The histograms in the output file show the postulated effectiveness of all variables. By this means it is visualized which variable has effect priority and between which variables causal interactions are possible. If causal relations between variables are postulated all the positive truth values of the tested relations must be enclosed in the fields of causal interactions.

Information Analysis

In order to be able to control the variables and the cases for abnormal characteristics, information-theoretical values are calculated for the original variables and for the dichotomized variables (options: 0,1,2). If the parameter “Print Extent” is at least 1, then for each variable, the number of categories, their frequencies of belonging cases, and their informational values are printed; also, for each case, the number of different categories and the resulting informational statistics are identified. Depending on the informational analyses of the original and the dichotomized variables, it is also possible to estimate the percentage of the loss of information which is caused by the dichotomisation of the variables.

Implications and Connections

These analyses are carried out (options: 0,1) only for those pairs of variables whose logical relationship can be confirmed at the selected significance level. In the option of connection analysis, the variables are checked for seven significant relations (conjunction, adjunction, disjunction, implication, bijunction, respectively NAND, and NOR). Regarding the implication analysis, the variables are only tested for implications, which are of particular importance for asymmetric causal relationships. Concerning to the significance calculations for the various bivariate relations, however, it must be taken into account that the evaluation of significance of relations with predominantly positive truth values (e.g., adjunction, implication, nonjunction) is underestimated compared to those with predominantly non-positive truth values (e.g., conjunction, NOR). Because of this fact at the statistical evaluation the normalised z-values also should be taken into account.

REFERENCE

Maderthaner, R. (2022). RELATIONSANALYSE (RELAN) - Aussagenlogische, statistische und kausale Analyse von Daten. Springer Spektrum, Heidelberg.