

RELAN - Relation Analysis

User Manual - Deutsch

Diese Anleitung soll interessierten Laien und empirischen „Praktikern“ einen groben Überblick verschaffen, wie die verschiedenen Einsatzmöglichkeiten des Programmes aufzurufen sind, welche Eingaben für welche Problemstellungen nötig sind und welche Ausgaben bei den verschiedenen Optionen erwartet werden können. Mit den vorliegenden Probefiles (Input, Daten) und der freien Software des Programmes (RELAN Beta 1.0) können sofort Erprobungen mit RELAN durchgeführt werden. Für detaillierte theoretische Begründungen oder methodische Erklärungen muss die Referenzliteratur herangezogen werden:

Maderthaner, R. (2022): Relationsanalyse (RELAN) - Aussagenlogische, statistische und kausale Analyse von Daten.

Für einen Testlauf des Programmes ([RELAN Program](#)) sind nur vier Files unmittelbar wichtig, nämlich RELAN.exe, RELAN-IN, RELAN-OUT und eventuell RELAN-DAT, welche gemeinsam mit anderen Beispielen im Verzeichnis Testing Files vorliegen ([Testing Files](#)). Am schnellsten kann man sich mit dem Programm vertraut machen, indem man die freie Software RELAN_Beta 1.0 startet, sodass sie mit dem im Verzeichnis vorhandenen Input File (RELAN-IN) läuft und sich anhand des entstehenden Outputs (RELAN-OUT) mittels Erläuterungen (im Manual: OUTPUT) einen Eindruck verschafft. Wenn erweiterte Erfahrungen gewünscht sind, dann können für weitere Durchläufe die Input Parameter direkt im RELAN-IN File oder mit dem Graphical User Interface (GUI) oder durch Kopieren der Input-Blöcke in den Output-Beispielen ([Testing Files](#)) modifiziert werden.

Im Programm können bei den unterschiedlichen Optionen insgesamt zehn Files genutzt werden, welche im [RELAN Fact Sheet](#) genauer charakterisiert sind (RELAN-IN, RELAN-OUT, RELAN-DAT, RELAN-KAT, RELAN-WEI, RELAN-FCT, RELAN-GRA, RELAN-EXP, RELAN-TRU, RELAN-VAL). Das Programm sollte in einem eigenen Verzeichnis laufen, weil dort automatisch alle genannten Files kreiert (bzw. überschrieben) werden.

ZWECK

- Es können auf aussagenlogischer Basis **natürlich-sprachliche Hypothesen** statistisch getestet werden, indem deren sprachliche Komponenten mit logischen Bindewörtern (UND, ODER, WENN-DANN, ...) verknüpft werden, wodurch viele, bisher in der Statistik kaum berücksichtigte Beziehungen zwischen Variablen analysierbar sind. Beispielsweise kann die Hypothese, dass eine Impfung (VAC) UND Antikörperbildung (ANT) ODER eine überstandene COVID-Erkrankung (REC) einen Schutz (PRO) vor Erkrankung hervorrufen, aussagenlogisch formuliert und statistisch getestet werden: $((VAC * ANT) + REC) > PRO$ (siehe auch [RELAN Examples](#)).
- Es lassen sich in einer Hypothese verschiedene **Arten von Einflüssen** berücksichtigen, indem Ursache-, Wirkungs-, Moderator-, Mediator, sowie Bedingungsvariablen differenzierbar sind.
- Für die Auswertungen können verschiedene **Zufallsmodelle** für die Erwartungswertbestimmung der Variablenausprägungen gewählt werden, sodass Ergebnisse aus verschiedenen Untersuchungen fairer verglichen werden können.
- Es ist möglich, **Verfälschungen** (konfundierende Variablen oder Variablengruppen) aus Effekten von Hypothesen herauszurechnen.
- Es können **Kausalanalysen** durchgeführt werden, bei denen die für Kausalprozesse geforderten logischen und zeitlichen Voraussetzungen tatsächlich gegeben sind (WENN-DANN-Relationen zwischen Variablen, zeitliche Koinzidenz von prognostizierten Wirkungen).
- Überdies sind mit dem Programm umfassende explorative Analysen („data mining“) möglich, indem alle vorhandenen logischen Gesetzmäßigkeiten (Relationen) in einem gegebenen Set von (zweiwertigen) Variablen ausfindig gemacht werden können.

INPUT

Im Input File RELAN-IN können alle Eingaben direkt im Text-File vorgenommen werden oder aber - bequemer und übersichtlicher - mittels des sogenannten „Graphical User Interface“ (GUI). Um schnell einen Eindruck vom Leistungsumfang des Programmes zu bekommen, ist es auch möglich, die Input-Blöcke der verfügbaren Beispiel-Files (siehe [RELAN Examples](#), [Testing Files](#)) zu verwenden, welche auf der ersten Seite

eines jeden Output-Files vorhanden sind und in RELAN-IN hineinkopiert werden können.

Wichtig: Es müssen für jede Art von Auswertung (Options) immer alle Parameter im RELAN-IN gegeben sein, auch jene, die im gerade ausgewählten Auswertungslauf nicht benötigt werden.

Als Kurzinstruktion kann auch das kommentierte RELAN-IN-File im PDF-Format ([RELAN-IN-Commentary](#)) dienen, in dem für jeden Parameter kurze Erklärungen beigefügt sind. Jeder Parameterart (Titel, Option, Input, Number of Variables, ...) geht im Input-File jeweils eine Zeile voraus, in der die Namen und die Bedeutungen der Parameter vermerkt sind und diese Zeile kann auch für individuelle Kommentare verwendet werden kann, weil sie im Programm nicht eingelesen wird. Um die nachfolgenden Erläuterungen über die Parameter besser zuordnen zu können, sollte währenddessen auch das Probefile [RELAN-IN](#) geöffnet sein.

TITEL: Benennung der Auswertung (nicht mehr als 80 Buchstaben)

OPTION: Hier wird über die Art der im Programm gewünschten Auswertung entschieden:

SIMULATION = 3: Simulation und statistische Evaluation von aussagenlogisch formulierten Hypothesen (Relationen), welche die Namen der Variablen (z.B. REC, VAC, ANT, ...) sowie aussagenlogische Verknüpfungen enthalten, wie etwa * (UND), + (ODER), > (WENN-DANN), = Genau DANN-WENN, -* (Nicht-UND). (Siehe dazu auch [RELAN Examples](#))

CONFIRMATION = 2: Statistische Prüfung von aussagenlogischen Hypothesen mit Ausgabe von deskriptiven Statistiken, Signifikanzwerten und Effektstärken (siehe dazu auch [RELAN Statistics](#)). Die hypothetische Relation wird anhand von etwa 30 Statistiken evaluiert und kann anhand dieser Charakteristika auch mit anderen Relationen hinsichtlich Überzufälligkeit und Verallgemeinerbarkeit verglichen werden.

IMPLICATIONS = 1: Exploration von Implikationen zwischen allen Paaren von Variablen zur Prüfung, ob zwischen ihnen linksseitige oder rechtsseitige Wenn-Dann-Relationen gegeben sind. Für alle signifikanten Relationen werden mehr als 30 gebräuchliche Statistiken ausgegeben (siehe dazu auch [RELAN Statistics](#)). Zusätzlich kann eine

graphentheoretische Analyse der gefundenen Implikationen angefordert werden, wodurch die verschiedenartigen Vernetzungen zwischen den Variablen aufgezeigt werden (Ausgangs- Vermittlungs- Endknoten, Verzweigungen etc.). Eine Kausalanalyse kann Wirkungszusammenhänge innerhalb eines Variablenetzes aufklären.

CONNECTIONS = 0: Exploration von Konnektionen (sieben aussagenlogischen Verknüpfungen) zwischen den Variablen. Es werden insgesamt sieben mögliche Relationen zwischen jeweils zwei Variablen überprüft und statistisch evaluiert. (siehe dazu [RELAN Statistics](#)). Auch bei dieser Auswertungsoption ist eine graphentheoretische Analyse vorgesehen.

DATA INPUT: Die Daten können entweder in Matrix-Format (im File RELAN-DAT) oder in Vektor-Format (im File RELAN-KAT) eingelesen werden. Beim Matrix-Format sind die Daten (wie üblich) so geordnet, dass in jeder Zeile ein Fall mit seinen Variablenwerten (getrennt durch Leerzeichen) eingegeben wird (siehe etwa [RELAN-DAT](#)). Beim seltener verwendeten Vektor-Format müssen für alle (2^M) Ausprägungskombinationen (Minterme) der gegebenen (M) Variablen alle zugehörigen Fallzahlen angegeben werden. Die Reihenfolge der Eingabe orientiert sich an der Wahrheitstafel der beteiligten Variablen (z.B. 0-0-0, 0-0-1, 0-1-0 etc.).

NUMBER OF VARIABLES (M): Der einzige Unterschied zwischen der freien und der lizenzierten Version besteht in der Anzahl möglicher Variablen und im Umfang möglicher Fälle (N), die für die Analysen verwendet werden können. In der freien Version sind bis zu 4 Variablen und bis zu 50 Fälle möglich, in der lizenzierten Version 100 Variablen und 1000 Fälle.

DICHOTOMISIERUNG: Hier ist festzulegen, ob die Daten bereits dichotomisiert sind (d.h. nur Werte von 0 und 1 haben) oder ob für jede Variable beim VECTOR OF DICHOTOMISATION Trennwerte (cutting scores) genutzt werden müssen. Jeder Wert, der in den Daten gleich groß oder größer als der zugehörige Trennwert ist, wird in 1 (Eins) umgewandelt, die übrigen Werte bekommen 0 (Null) zugewiesen.

NAMES OF VARIABLES: Alle Variablen haben dreistellige Namen, welche durch ein Leerzeichen (oder mehrere) getrennt einzugeben sind.

VEKTOR OF DICHOTOMISIERUNG: Eingabe der Trennwerte für alle Variablen. Wenn der Wert einer Variablen gleich oder größer als ihr Trennwert ist, bekommt er den Wert 1 zugeordnet, ansonsten den Wert 0. Bei vorgesehener Dichotomisierung ist es günstig die Trennwerte (cutting scores) für die Variablen genau unter den Variablennamen einzugeben, damit eine verlässliche Zuordnung gesichert ist. Trennwerte können mit oder ohne Dezimalpunkt eingegeben werden.

PROBABILITIES OF THE APRIORI CHANCE MODEL: Für jede Variable sind Schätzungen von Wahrscheinlichkeiten ihrer positiven Ausprägungen einzugeben (Default-Wert: $p = 0.5$). Die Wahrscheinlichkeiten der Variablen können entweder hypothetisch festgelegt sein, aus anderen Untersuchungen stammen oder von statistischen Datenbanken übernommen werden. Diese gelten dann als APRIORI CHANCE MODEL. Wenn dieses für die Berechnungen herangezogen werden soll, dann ist beim Parameter HYPOTHESIS CHANCE MODEL der Wert „apriori“ auszuwählen. Sollen hingegen die relativen Häufigkeiten der Variablen aus der Stichprobe als Wahrscheinlichkeiten für die Berechnungen genutzt werden, dann ist beim HYPOTHESIS CHANCE MODEL „aposteriori“ auszuwählen.

ORDER OF CAUSALITY: Für jede Variable wird bei Kausalanalysen mittels einer natürlichen Zahl (z.B. 1, 2, 3, ...) symbolisiert, ab welchen Zeitpunkt die (als Ereignisse interpretierten) Ausprägungen der Variablen eine Auswirkung auf andere Variablen haben können.

DURATION OF CAUSALITY: Ebenso wird für jede Variable bei Kausalanalysen mit einer natürlichen Zahl (z.B. 2, 2, 3, ...) festgelegt, über wie viele Zeiteinheiten die mögliche Auswirkung einer jeden Variable als wirksam angesehen wird (siehe dazu auch [RELAN Examples](#)).

PRINT EXTENT: Mit diesem Parameter wird der Umfang der Ausgabe von Berechnungen im RELAN-OUT festgelegt (0 = minimal, 3 = maximal).

HYPOTHESIS CHANCE MODEL: Hier wird das gewünschte Zufallsmodell ausgewählt (Nullhypothese), nach dem die Signifikanzberechnungen durchgeführt werden. Wenn 1 (= apriori) ausgewählt wird, dann werden

die theoretisch bestimmten Wahrscheinlichkeiten herangezogen, und bei 0 (= aposteriori) werden die aus der Stichprobe errechneten relativen Häufigkeiten der Variablen (für die Berechnung der Erwartungswerte) verwendet.

LEVEL OF SIGNIFICANCE (Zsig): Festlegung der Irrtumswahrscheinlichkeit (Alpha-Fehler) für die Signifikanzprüfungen. Für einseitige Fragestellungen (z.B. ob Frequenzen größer sind als der Erwartungswert) beträgt dieser Wert 1.645 bei 5-prozentiger und 2.330 bei 1-prozentiger Irrtumswahrscheinlichkeit.

LEVEL OF SIMULATION (Zsim): Festlegung des gewünschten Signifikanzniveaus für eine Simulation einer aussagenlogischen Hypothese in einer fiktiven Stichprobe. Je größer dieser Wert gewählt wird, desto stärker werden die Daten an die postulierte aussagenlogische Hypothese angepasst. Für extrem prägnante Anpassungen der Daten an eine Hypothese können sogar Zsim-Werte bis 30 verwendet werden.

LEVEL OF EXTRACTION (Zext): Festlegung des gewünschten Signifikanzniveaus für eine (konfundierende) Relation die aus der Hypothesen-Funktion herausgerechnet werden soll. Je größer dieser Wert ist, desto mehr weicht die herauszurechnende Funktion von den Erwartungswerten ab. Wird hier 0.0 eingegeben, dann wird keine Extraktion gerechnet.

INCLUSION CRITERION: Bei Exploration von Relationen zwischen jeweils zwei Variablen (Optionen: CONNECTIONS, IMPLICATIONS) können beim INCLUSION CRITERION nicht-normierte und normierte Z-Werte gewählt werden. Da die Minima und Maxima von Z-Werten auch von der Anzahl an Wahrheitswerten einer Relation abhängen, werden im Programm neben den üblichen Z-Werten auch normierte Z-Werte ausgegeben, nämlich solche, die für Relationen mit gleicher Anzahl von Wahrheitswerten zustande kämen. Bei Eingabe von 0 (nicht normiert) werden übliche Signifikanzberechnungen für die Relationen gerechnet, bei Eingabe von 1 (normiert), werden die Z-Werte so berechnet, als hätten alle Relationen gleich viele positive und negative Wahrheitswerte.

Y(P)-CUT und X(P)-CUT: Da bei paarweisen Prüfungen von Variablenrelation sehr viele Kombinationen entstehen können, besteht im

Programm die Möglichkeit den Output zweifach zu begrenzen, indem man einen Grenzwert für die relative Häufigkeit der ersten Variable X und/oder der zweiten Variable Y angibt, sodass nur jene Resultate ausgedruckt werden, bei denen die relativen Häufigkeiten über dem jeweiligen Grenzwert liegen. Damit wird auch der praktischen Bedeutsamkeit von Variablen Rechnung getragen.

GRAPH THEORETICAL ANALYSIS: Bei den Optionen CONNECTIONS und IMPLICATIONS können für signifikante Relationen auch Analysen auf Basis der Graphentheorie durchgeführt werden. Damit werden Relationsketten identifiziert, deren Ausgangs- und Endvariablen bestimmt und zirkuläre Variablenbeziehungen aufgezeigt. Um die Ausgaben in Graphikprogrammen (z.B. yEd, Decision Explorer) weiterverarbeiten zu können, werden die signifikanten Variablenrelationen im File RELAN-GRA mit Variablenpaaren, Relationsgewichten, und Relationsarten ausgegeben.

WEIGHTING OF CASES: Wenn Zweifel an der empirischen Glaubwürdigkeit mancher Fälle bestehen (schlechte Erhebungsbedingungen, unsichere Messungen, missverständliche Anweisungen, ...), dann kann dies mit einem Glaubwürdigkeitsindex zwischen 0.0 (nicht verwertbar) und 1.0 (volle Verwertbarkeit) ausgedrückt werden. Bei einer Dateneingabe im Matrix-Format (mittels RELAN-DAT) können die einzelnen Fälle entsprechend ihrer empirischen Glaubwürdigkeit mit Index-Werten zwischen 0.0 und 1.0 gewichtet werden. Diese Gewichtungswerte müssen für alle Fälle in vorhandener Reihenfolge im RELAN-WEI File eingetragen werden (Zahlenreihe mit Leerzeichen als Zwischenraum).

CAUSAL ANALYSIS: Wenn eine Kausalanalyse durchzuführen ist (1 = yes), dann wird davon ausgegangen, dass die Ausprägungen von Variablen als „Ereignisse“ zu interpretieren sind, denen ab eines bestimmten Zeitpunktes und für eine bestimmte Dauer Wirkungen zugeschrieben werden. Für die Annahme kausaler Interaktionen zwischen den Variablenausprägungen (symbolischen Ereignissen) müssen diese in der gleichen Zeitebene realisiert sein. Das Ursache-Ereignis muss mit dem Wirkungs-Ereignis zu einem bestimmten Zeitpunkt zusammentreffen, damit eine Wirkung zustande kommen kann. Ob diese Wirkungskoinzidenz anzunehmen ist, wird im Programm anhand der definierten Wirkungsindikatoren (ORDER und DURATION OF CAUSALITY) geprüft.

Eine Relation wird dann als kausal zusammenhängend kategorisiert, wenn alle ihre logischen Komponenten in zumindest einer (durch die Indizes definierten) zeitlichen Wirkungsspanne realisiert sind.

TRUTH FUNCTION: Aussagenlogische Relationen können sowohl in Form von Funktionen (z.B. $(VAC + REC) > PRO$) als auch als in Form von Wahrheitstabellen (z.B. 1 0 0 1 1 0 etc.) charakterisiert werden. Der Parameter ist 0 (= no), wenn die Relation der HYPOTHETICAL FUNCTION als aussagenlogische Formel eingegeben wird (in RELAN-IN), und 1 (= yes), wenn über das File RELAN-FCT die Wahrheitsfunktion der hypothetischen Relation importiert wird (1 = wahr, 0 = nicht wahr). Alle in der Auswertung entwickelten Wahrheitsfunktionen werden im File RELAN-TRU ausgegeben.

HYPOTHETICAL FUNCTION: Diese Parameter-Eingabe enthält eine aussagenlogische Funktion (z.B. $VAC + REC > PRO$), die im Ausmaß des LEVEL OF SIMULATION simuliert oder im Ausmaß des LEVEL OF SIGNIFICANCE getestet wird. Wenn mehrere unverbundene Funktionen (zu aussagenlogischen Vergleichszwecken) eingegeben werden, dann wird nur die letzte Funktion statistisch evaluiert. Die Funktion darf nicht mehr als 80 Zeichen einnehmen (einschließlich Leerzeichen).

EXTRACTIIONAL FUNCTION: Diese Parameter-Eingabe enthält eine aussagenlogische Funktion, die im Ausmaß des LEVEL OF EXTRACTION simuliert und danach aus der HYPOTHETICAL FUNCTION herausgerechnet (reduziert) wird. Die Funktion darf nicht mehr als 80 Zeichen einnehmen (einschließlich Leerzeichen). Wird nur der Effekt einer einzigen Variablen „auspartialisiert“, dann muss diese als Funktion eingegeben werden (z.B. $REC * REC$).

DATEN

Am häufigsten werden in Programmen Daten in konventioneller Form, nämlich als Matrix eingegeben, wobei jeder Fall in einer Zeile mit der entsprechenden Anzahl an Variablenwerten positioniert ist. Für jeden Fall sind in der RELAN gleich viele Variablenwerte einzugeben, ein Ersetzen fehlender Werte („Imputation“) ist in der jetzigen Programmversion nicht vorgesehen. Die Variablenausprägungen sind im freien Format einzugeben, d.h. sie können Dezimalpunkte enthalten und müssen durch

ein Leerzeichen (oder mehrere) getrennt sein (siehe [RELAN-DAT](#)). Eine weitere, wahrscheinlich viel seltener gewählte Eingabeart ist jene in Form einer Zahlenreihe („Vektor“), bei der die auf jede Ausprägungskombination der Variablen entfallenden Fallzahlen einzugeben sind ([RELAN-KAT](#)). Die Reihenfolge der Eingabe orientiert sich an der Wahrheitstafel für die entsprechende Variablenmenge und ist der einschlägigen Tabelle aus jedem Output-File zu entnehmen. Da bei Vektoreingabe der Daten diese bereits aggregiert vorliegen (Summe der Fälle je Ausprägungskombination), muss hier die optionale Fallgewichtung entfallen. Im File [RELAN-EXP](#) werden die durch Simulation erhaltenen Daten oder jene, die eingelesen wurden, exportiert, sodass diese nicht nur in anderen Programmen (z.B. SPSS, SAS) verarbeitet werden, sondern (nach Löschung der Variablennamenzeile) auch für weitere Analysen in das RELAN-DAT File importiert werden können.

PROGRAM FLOW

Wenn das Programm gestartet wird, öffnen sich drei (grüne) Fenster. Das rechte Fenster zeigt die im Verzeichnis vorhandenen oder gebildeten Arten von Files an, das erste linke Fenster illustriert den ablaufenden Output (RELAN-OUT), das zweite linke Fenster (dahinter) bildet die Input-Parameter ab (RELAN-IN). Danach erscheint bei fehlerlosem Ablauf eine Box mit dem Text „Program terminated, exit code 0 – Exit Window“. Mit „yes“ schließt man die Fenster, um das File RELAN-OUT aufzurufen. Den gleichen Effekt erzielt man mit „CTRL C“, welches auch bei zu langen Wartezeiten für einen Abbruch der Berechnungen von RELAN genutzt werden kann. Sollte ein frühzeitiger Abbruch des Programmes erfolgt sein, erscheint bei den meisten Fehlern eine Begründung im Output File (RELAN-OUT). Bei inhaltlichen Problemen kann eine Mail an rrm@relan.at und bei technischen Problemen an support@relan.at gerichtet werden.

Außer dem RELAN EXE-File sind alle Files Text-Files (.txt), für die die Schriften Courier New, Century Gothic oder Lucida Console empfohlen werden. Bei Schriftgröße Punkt 8 passen die meisten Ausgaben in der Breite auf eine A4-Seite.

OUTPUT

Anschließend sollen kurz die wichtigsten Ausgaben in Output-Files erläutert werden, wofür zusätzlich das [RELAN-OUT](#) geöffnet werden sollte.

Am Beginn eines jeden Output-Files wird informiert, ob die TRIAL CONDITION (Probeversion) oder die LICENCE CONDITION läuft.

Zwischen den am Beginn des RELAN-OUT Files vorkommenden Zeilen befinden sich die Parameter des verwendeten Input-Files [RELAN-IN](#), welche - kopiert oder verändert - in ein neues RELAN-IN File eingetragen werden können:

```
*** Start of Copy of Input-File: RELAN-IN ***
```

```
--- Input Parameter ---
```

```
*** End of Copy of Input-File: RELAN-IN ***
```

CHECK OF VARIABLE NAMES

Prüfung, ob bei der Formulierung der Hypothese oder der Extraktionsrelation korrekte Variablennamen verwendet wurden.

CONVERSION OF RELATIONS

Transformation der dreistelligen Variablennamen der Hypothese sowie der Extraktionsrelation in kürzere Variablensymbole (A, B, C, ...).

CHANCE MODELS

Ausgabe, von welchen Zufallsmodellen (apriori/aposteriori) bei den statistischen Berechnungen für die Hypothese und für die Extraktion ausgegangen wurde.

ORDER AND INDICES OF VARIABLES

Ausgabe der gewählten Kausalindizes (Kausalordnung, Kausaldauer) für die Variablen. In der Darstellung CAUSAL RANGE OF VARIABLES wird der zeitliche Beginn einer möglichen Kausalwirkung einzelner Variablen sowie deren Wirkungsdauer graphisch illustriert.

DATA FILE

Name des Datenfiles, welches für die Auswertungen verwendet wurde (RELAN-DAT oder RELAN-KAT).

DATA VALUES

Mit dem höchsten Wert (= 3) des Parameters PRINT EXTENT werden alle berechneten Ergebnisse im RELAN-OUT ausgegeben.

INFORMATIONAL ANALYSIS

Informationsanalyse wird nur bei zehn (ganzzahligen) Variablen durchgeführt. Es werden folgende Kennwerte ausgegeben:
Fallbesetzungen in den Variablenkategorien (max = 30),
Informationsgehalte für Fälle und Variablen, Gesamtinformation für die Originaldaten und die dichotomisierten Daten. Das wichtigste Ergebnis ist das Verhältnis der Gesamtinformation für die Originaldaten und jene für die dichotomen Daten, welches den Informationsverlust aufzeigt, der durch die Dichotomisierung der Daten entstanden ist.

MINTERM FREQUENCIES

Die Minterme repräsentieren alle möglichen Kombinationen von Ausprägungen der beteiligten Variablen. In einer den Mintermen zugeordneten Zahlenreihe werden die Häufigkeiten an Fällen ausgegeben, welche mit den entsprechenden Ausprägungskombinationen in den Daten vorkommen.

APOSTERIORI FREQUENCIES

Dies sind die relativen Häufigkeiten, mit denen die positiven Ausprägungen der einzelnen Variablen in der Stichprobe vorkommen.

SIMULATION ANALYSIS and CONFIRMATION ANALYSIS

CHARACTERISTICS OF THE BOOLEAN FORMULA

Hier werden die in der Formel enthaltenen Variablen kontrolliert und in Kurzbezeichnung angeführt.

TRANSMISSION OF THE FORMULA

Darstellung der aufeinanderfolgenden Transformationen, in welcher Weise die Komponenten einer hypothetischen Relation zu Subrelationen verrechnet werden und schließlich zu einer Gesamtrelation führen.

TRUTH FUNCTIONS (Variables + Subrelations)

In dieser Tabelle sind die Wahrheitswerte aller Variablen, aller resultierenden Subfunktionen sowie der hypothetischen oder simulierten Relation dargestellt.

PROBABILITIES OF CHANCE MODELS

Wahrscheinlichkeiten für die Zufallsmodelle (Apriori, Aposteriori), welche durch die relativen Häufigkeiten des Auftretens der positiven Variablenausprägungen definiert sind.

ELEMENTARY RELATIONS

Für alle Minterme (Ausprägungskombinationen der Variablen) werden die statistischen Kennwerte berechnet (siehe [RELAN Statistics](#)). Die signifikanten Minterme heißen Elementarrelationen („Typen“ oder „Antitypen“ in der Konfigurationsfrequenzanalyse, KFA).

INTERCORRELATIONS

Für die Einschätzung der Ähnlichkeiten zwischen den Variablen, Subrelationen und der Hypothesen- bzw. Simulationsrelation werden über alle gegebenen Minterme Produkt-Moment-Korrelationen gerechnet, und zwar einerseits für die Wahrheitswerte der Relationen und andererseits für deren Fallbesetzungen in den Mintermen. Geringe Ähnlichkeit einer Relation mit anderen Relationen ist ein wichtiges Indiz für ihre spezifische Funktion bei der Erklärung der Resultate.

SUBRELATIONS AND HYPOTHETICAL RELATION

In dieser Tabelle sind die wichtigsten statistischen Kennwerte für alle Variablen, alle Subrelationen und für die simulierte oder hypothetische Relation zusammengefasst (siehe [RELAN Statistics](#)). Diese statistischen Kennwerte für alle aus der Zerlegung der hypothetischen Relation stammenden Subrelationen sollten unter Einbezug ihrer Interkorrelationen interpretiert werden.

HYPOTHETICAL RELATION

Diese Tabelle enthält 30 Signifikanz- und Effektstatistiken für die simulierte oder geprüfte Relation (siehe [RELAN Statistics](#)).

EXTRACTATIONAL RELATION

Wenn die hypothetische oder simulierte Relation um eine andere Relation (konfundierende Variablen bzw. Error-Relationen) reduziert werden soll, dann werden in der Folge die Charakteristika sowohl für die Extraktionsrelation als auch jene der Residualrelation (reduzierte Relation) ausgegeben.

HIGHER ORDER RELATIONS

Den Abschluss dieser beiden Auswertungsprozeduren (Simulation, Confirmation) bilden extensive explorative Analysen („data mining“), bei denen für alle Variablen getestet wird, ob auch Relationen „höherer Ordnung“ zwischen den gegebenen Variablen vorhanden sind. Darunter werden Elementarrelationen verstanden, die als Funktion einer kleineren Menge von Variablen zustande kommen, als in der getesteten oder simulierten Relation gegeben sind. Auf diese Weise können Variablenrelationen gefunden werden, denen zumindest eine prädiktive Funktionen für eine Datenstruktur zugeschrieben werden kann ([RELAN-OUT-Confirmation-1](#)). (Dieser Output wird nur bei Print Extent = 3 ausgegeben.)

SIGNIFICANT HIGHER RELATIONS

Zusammenfassende Ausgabe aller signifikanten Relationen höherer Ordnung mit zugehörigen Statistiken (siehe [RELAN-OUT](#)). (Dieser Output wird nur bei Print Extent = 2 und 3 ausgegeben.)

IMPLICATION ANALYSIS and CONNECTION ANALYSIS

Diese Analysen sind explorativ und haben den Zweck, zwischen jeweils zwei Variablen prädikatenlogische Beziehungen auf Signifikanz zu prüfen. Für diese Art von Auswertung muss der Parameter INCLUSION CRITERION entweder auf „not normalised“ oder auf „normalised“ gestellt sein. Bei der ersten Einstellung werden die Signifikanzwerte für eine Relation auf

übliche Weise berechnet, bei der zweiten werden diese so normiert, als würden sie bei gleich vielen positiven und negativen Wahrheitswerten einer Relation zustande kommen. Damit können Relationen mit unterschiedlichen Wahrheitswerten (z.B. Konjunktionen und Implikationen) fairer miteinander verglichen werden. Mittels der Zusatzparameter $Y(p)$ und/oder $X(p)$ ist es möglich, die Ausgaben so zu begrenzen, dass nur Relationen mit höheren relativen Häufigkeiten der beiden Variablen ausgegeben werden. Außerdem kann durch den Parameter CAUSAL ANALYSIS die Ausgabe auf nur jene Relationen begrenzt werden, die mit den kausalen Zeitvorgaben übereinstimmen.

IMPLICATIONS/CONNECTIONS

Ausgegeben werden nur jene Implikationen oder Konnektionen, die sich je nach Parametereinstellung als signifikant erweisen (siehe [RELAN-OUT-Implications-2](#), [RELAN-OUT-Connections-2](#)). Insgesamt werden für jede signifikante Relation vierzig statistische Kennwerte berechnet (siehe [RELAN Statistics](#)).

ANALYSIS OF GRAPHS

Wird dieser Parameter ausgewählt, dann werden alle signifikanten Relationen zwischen Paaren von Variablen einer Graphenanalyse unterzogen (siehe [RELAN-OUT-Implications-2](#)). Das heißt, es werden im Variablensystem „Wege“, „Pfade“, „Komponenten“, „Bäume“ und „spannende Bäume“ gesucht, aufgrund derer Anfangspunkte, Endpunkte, Verzweigungen, Verbindungen in einem Variablennetzwerk gefunden werden können (siehe [RELAN Statistics](#)).

REFERENCE

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