



Latent Structure Analysis

Hermann Denz

Almo Statistik-System

www.almo-statistik.de

holm@almo-statistik.de

kurt.holm@jku.at

2014

Das vorliegende Dokument ist ein Kapitel aus dem Almo-Handbuch "Sozialwissenschaftliche Skalierungsverfahren". Kurt Holm hat dieses vom sehr früh verstorbenen Hermann Denz verfasste Kapitel etwas überarbeitet und auch durch (ebenfalls leicht überarbeitete) Teile des Buchs „H. Denz: Analyse Latenter Strukturen“ ergänzt.

Siehe insbesondere auch das Almo-Dokument Nr. 7 "Ordinales Rasch-Modell.PDF" (Holm/Potuschak) und das Almo-Dokument Nr. 19 „Guttman- und Mokken-Skalierung.PDF“ (Joachim Gerich)

Weitere Almo-Dokumente

Die folgenden Dokumente können alle kostenlos von der Handbuchseite in

www.almo-statistik.de

heruntergeladen werden

0. Arbeiten_mit_Almo.PDF (1 MB)
1. Zwei- und drei-dimensionale Tabellierung.PDF (1.1 MB)
2. Beliebige-dimensionale Tabellierung.PDF (1.7 MB)
3. Nicht-parametrische Verfahren.PDF (0.9 MB)
4. Kanonische Analysen.PDF (1.8 MB)
Diskriminanzanalyse.PDF (1.8 MB)
enthält: Kanonische Korrelation, Diskriminanzanalyse, bivariate Korrespondenzanalyse, optimale Skalierung
5. Korrelation.PDF (1.4 MB)
6. Allgemeine multiple Korrespondenzanalyse.PDF (1.5 MB)
7. Allgemeines ordinales Rasch-Modell.PDF (0.6 MB)
- 7a. Wie man mit Almo ein Rasch-Modell rechnet.PDF (0.2 MB)
8. Tests auf Mittelwertsdifferenz, t-Test.PDF (1,6 MB)
9. Logitanalyse.pdf (1,2MB) enthält Logit- und Probitanalyse
10. Koeffizienten der Logitanalyse.PDF (0,06 MB)
11. Daten-Fusion.PDF (1,1 MB)
12. Daten-Imputation.PDF (1,3 MB)
13. ALM Allgemeines Lineares Modell.PDF (2.3 MB)
- 13a. ALM Allgemeines Lineares Modell II.PDF (2.7 MB)
14. Ereignisanalyse: Sterbetafel-Methode, Kaplan-Meier-Schätzer, Cox-Regression.PDF (1,5 MB)
15. Faktorenanalyse.PDF (1,6 MB)
16. Konfirmatorische Faktorenanalyse.PDF (0,3 MB)
17. Clusteranalyse.PDF (3 MB)
18. Pisa 2012 Almo-Daten und Analyse-Programme.PDF (17 KB)
19. Guttman- und Mokken-Skalierung.PFD (0.8 MB)
20. Latent Structure Analysis.PDF (1 MB)
21. Statistische Algorithmen in C (80 KB)
22. Conjoint-Analyse (PDF 0,8 MB)
23. Ausreisser entdecken (PDF 170 KB)
24. Statistische Datenanalyse Teil I, Data Mining I
25. Statistische Datenanalyse Teil II, Data Mining II
26. Statistische Datenanalyse Teil III, Arbeiten mit Almo-Datenanalyse-System
27. Mehrfachantworten, Tabellierung von Fragen mit Mehrfachantworten (0.8 MB)
28. Metrische multidimensionale Skalierung (MDS) (0,4 MB)
29. Metrisches multidimensionales Unfolding (MDU) (0,6 MB)
30. Nicht-metrische multidimensionale Skalierung (MDS) (0,4 MB)
31. Pfadanalyse als wiederholte Regressionsanalyse (0,7 MB)
32. Datei-Operationen mit Almo (1,1 MB)

Inhaltsverzeichnis

Latent Structure Analysis.....	1
P15 Skalierung nach der "latent structure analysis"	4
P15.1 Das Modell der "latent structure analysis"	4
P15.1.1 Die Aufgabe	4
P15.1.2 Dichotome Variable	4
P15.1.3 Die Modelle der LSA	4
P15.2 Die Annahmen	6
P15.2.1 Lokale stochastische Unabhängigkeit	6
P15.2.2 LSA ist probabilistisch.....	6
P15.2.3 LSA misst nicht spezifisch objektiv.....	7
P15.3 Das Ogiven-Modell.....	7
P15.3.1 Das Modell.....	8
P15.3.2 Items für das Ogiven-Modell	11
P15.3.3 Test des Ogiven-Modells	12
P15.3.4 Programm-Maske Prog15ma für das Ogiven-Modell.....	13
P15.3.5 Ergebnisse aus Prog15ma (Ogiven-Modell)	21
P15.4 Programm-Maske Prog15m1 für alle Modelle der LSA.....	26
P15.5 Ergebnisse für die Modellen der LSA aus Prog15m1.....	30
P15.5.1 Lineare Itemcharakteristik (Modell 1)	30
P15.5.2 Latent-distance Modell (Modell 3)	33
P15.5.3 Ogiven-Modell (Modell 5).....	34
P15.5.4 Latent-Class-Analyse (Modell 6 und 7).....	34
Literatur	38

P15 Skalierung nach der "latent structure analysis"

P15.1 Das Modell der "latent structure analysis"

Die Latent Structure Analysis (kurz: LSA) wurde schon sehr früh durch die Beiträge einer Vielzahl von Autoren entwickelt. Systematische Darstellungen sind z.B. enthalten in

W.S. Torgerson: Theory and Method of Scaling, Wiley, 1960, besonders Kap. 13
P.F. Lazarsfeld & N.W. Henry: Latent Structure Analysis, Boston, 1968

Eine Kleinste-Quadrate-Schätzung für das Ogiven-Modell der LSA wurde insbesondere entwickelt von

L.R. Tucker: A level of proficiency scale for a unidimensional skill,
Amer. Psychologist, 7, 408

Eine umfangreiche Literaturliste ist bei Denz (1982) zu finden.

P15.1.1 Die Aufgabe

In den Sozialwissenschaften sind viele Variable nicht direkt messbar: Autoritarismus, Intelligenz, Extraversion usw. Es kann aber angenommen werden, dass aufgrund des theoretischen Vorverständnisses eine Nominaldefinition gegeben werden kann und in einem weiteren Schritt Items formuliert werden können, welche die angezielte Messdimension treffen. Die Antworten (manifeste Variable) können als empirische Realisierungen der nicht direkt messbaren Eigenschaft (latente Variable) interpretiert werden. Dies ist im Grunde eine allgemeine Beschreibung jeder Art von Skalierung. Erst durch die Präzisierung der Annahmen, von denen die Analyse latenter Strukturen ausgeht und die anders sind als bei anderen Verfahren, gewinnt sie ihre Stellung als eigenständiges Modell.

P15.1.2 Dichotome Variable

Für den Fall, dass die manifesten Variablen quantitativ sind, gibt es eine Reihe von Skalierungsverfahren, für dichotome manifeste Variable nur wenige. Das führt dazu, dass sehr häufig schematisch z.B. die Faktorenanalyse auf dichotome Items angewendet wird, was (oft) zu falschen Ergebnissen führt (vgl. Denz 1982, S. 10ff). Die LSA wurde speziell für den Fall dichotomer Items entwickelt, obwohl sie grundsätzlich auch auf polytome erweiterbar wäre (in Almo nur dichotom).

P15.1.3 Die Modelle der LSA

Die große Verbreitung linearer Modelle (eben z.B. Faktorenanalyse) hat dazu geführt, dass die Form der Beziehung zwischen manifesten und latenten Variablen nicht mehr als theoretisch/empirisches Problem gesehen wird. Da die Linearitätsannahme nicht in jedem Fall sinnvoll ist, muss für diese Fälle auf andere Modelle (z.B. Rasch-Modell) zurückgegriffen werden. Hier erlaubt die LSA die Spezifizierung sehr unterschiedlicher auch nicht-linearer Funktionen (nicht-linearer "trace lines"). In Almo ist nur ein Teil der möglichen "trace lines" realisiert, aber diese decken einen großen Anwendungsbereich ab).

$p_{ij} = \int_{-\infty}^{\infty} f_i(x)f_j(x)g(s)dx$ folgt aus der lokalen stochastischen Unabhängigkeit (siehe P15.2.1)

$$p_{ijk\dots} = \int_{-\infty}^{\infty} f_i(x)f_j(x)f_k(x)\dots g(s)dx$$

Wird nun die latente Variable als diskret betrachtet, wird aus der Integration über den möglichen Wertebereich die Summe aller diskreten Punkte und der Dichtefunktion entsprechen die relativen Häufigkeiten der Ausprägungen:

$p_i = \sum_{l=1}^m p_{il} * n_l$ m p_{il} und n_l sind die latente Wahrscheinlichkeit für Item i in der Klasse l und die relative Häufigkeit dieser Klasse. (Summe über alle Klassen $1\dots m$)

$$p_{ij} = \sum_l p_{il} * p_{jl} * r_l \quad \sum_l r_l = 1$$

$$p_{ijk} = \sum_l p_{il} * p_{jl} * p_{kl} * r_l$$

Typ 1: Spezifizierung der Funktion (die Modelle)

a) lineare Funktion

$$f_i(x) = a_{0i} + a_{1i}x$$

b) latent-distance Modell

$$f_i(x) = a_i - b_i \text{ wenn } x < x_i$$

$$f_i(x) = a_i + b_i \text{ wenn } x \geq x_i$$

ba) Guttman-Skala

$$a_i = b_i = 0.5$$

bb) Restringiertes latent-distance-Modell

$$a_i = 0.5$$

bc) Allgemeines latent-distance-Modell

c) Ogiven-Modell

$$f_i(x) = \frac{1}{\sqrt{2\pi}} * \int_{-\infty}^{\frac{x_i - M_i}{S_i}} \exp\left(-\frac{z^2}{2}\right)$$

Typ 2: Spezifizierung der latenten Variablen Die latente Variable ist nicht mehr stetig, sondern es wird eine Konzentration der Befragten an mehreren Punkten angenommen: *latent-class-Modell*

$$p_i = p_{i1} * n_1 + p_{i2} * n_2 + \dots + p_{im} * n_m$$

P15.2 Die Annahmen

P15.2.1 Lokale stochastische Unabhängigkeit

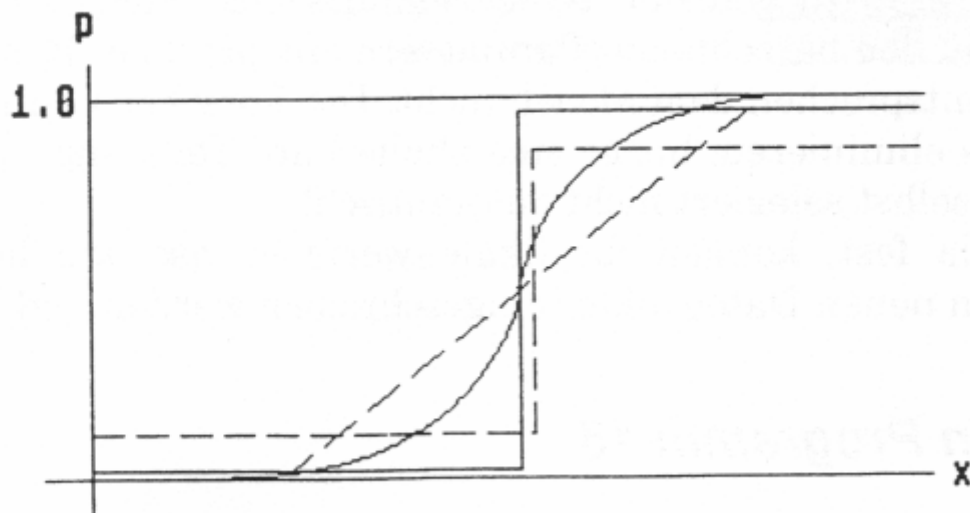
Zentrale Annahme ist das Prinzip der lokalen stochastischen Unabhängigkeit. Das bedeutet, dass die empirische Koinzidenz bestimmter Ereignisse (z.B. eine Person antwortet auf zwei verschiedene Fragen gleich) dadurch erklärt werden kann, dass die empirischen Ergebnisse Realisierung derselben latenten Variablen sind. Oder umgekehrt formuliert: die Antwort(en) einer Person können, gegeben die latente Klasse, als Produkt der latenten Wahrscheinlichkeit prognostiziert werden. Es braucht nun die latente Variable nicht notwendigerweise dichotomisch sein, sie kann einerseits beliebig viele Ausprägungen haben (jedoch immer weniger als die Anzahl der Items ist, weil sonst ja keine Reduktion der Daten vorhanden wäre), andererseits stetig sein. Die Logik der lokalen stochastischen Unabhängigkeit bleibt die gleiche.

P15.2.2 LSA ist probabilistisch

Sie ist ein probabilistisches Modell. Selbst unter der Annahme, dass mehrere Items genau in derselben Weise dasselbe messen, wird ein Befragter nicht jedesmal die identische Antwort geben (vgl. Lazarsfeld 1954, S. 354 oder als "Fehler im Befragten" interpretiert: Holm, 1976, S. 371); dazu kommt nun, dass diese Fiktion auch empirisch nicht aufrechterhalten werden kann, die verschiedenen Items werden nur annähernd dasselbe messen. Das bedeutet, dass in den Ergebnissen Zufallsfehler enthalten sein werden, für die nur die Bedingung gilt, dass sie nicht mit der latenten Variablen und untereinander korrelieren dürfen (noch anschaulicher ist die Differenzierung von Coombs, der eine Einteilung danach trifft, ob ein Modell Inkonsistenzen zulässt und dadurch die verschiedenen Formen der Analyse latenter Strukturen mit stetiger latenter Variable den Modellen mit hoher Toleranz für Inkonsistenzen zuordnet: Coombs 1967, S. 280). Die probabilistische Formulierung

in diesem Sinn lässt sich am besten an der Modifizierung der Guttman-Skala zeigen:

Graphik: Guttman Skala (durchgezogene Linie) und Analyse latenter Strukturen (strichlierte Linien)



Unter der Annahme einer Guttman Skala müssen alle Personen, deren Skalenwert kleiner a ist, auf die eingezeichnete Frage mit "nein" antworten (=Wahrscheinlichkeit einer Ja-Antwort ist Null), alle mit einem höheren Skalenwert mit ja (Wahrscheinlichkeit ist Eins). Dass dieses Modell unrealistisch ist, zeigt sich auch daran, dass die Guttman-Skala empirisch kaum verwendet wird. Dieses Modell kann nun umformuliert werden: alle Personen mit einem Skalenwert kleiner a beantworten diese Frage mit der Wahrscheinlichkeit p_1 mit "ja", alle mit einem größeren Skalenwert mit p_2 .

P15.2.3 LSA misst nicht spezifisch objektiv

An diesem Beispiel können noch zwei weitere wichtige Voraussetzungen gezeigt werden: Personen und Items können nicht unabhängig voneinander gemessen werden, die Analyse latenter Strukturen misst nicht "spezifisch objektiv" (Rasch 1966, S. 104), damit sind die Itemparameter abhängig von der jeweils gewählten Stichprobe.

P15.3 Das Ogiven-Modell

Das in der bisherigen Praxis am häufigsten angewendete Modell - das Ogivenmodell - soll zuerst und etwas ausführlicher erörtert werden.

Für die Lösung des Verfahrens müssen homogene Subgruppen gebildet werden; im Programm werden die Subgruppen nach dem Gesamtpunktwert gebildet. Im Gegensatz zum Rasch-Modell ist der Gesamtpunktwert hier keine erschöpfende Statistik des Skalenwerts der Personen, aber er sollte eine monotone Funktion dieses Skalenwertes sein. Zuerst wird im Programm überprüft, ob die Wahrscheinlichkeit von ja-Antworten (mit Code =1) auf ein Item eine monotone Funktion des Gesamtpunktwertes ist. Trifft dies nicht zu, wird eine entsprechende Warnung ausgedruckt. Dies kann also auch schon ein erstes Kriterium für die Auswahl geeigneter Items sein.

Der Begriff der „Item-Monotonie“ soll an einem Beispiel kurz erläutert werden

Relative Häufigkeit einer positiven Antwort
für jeden Gesamtpunktwert (ohne Extremgruppen 0 und 5)

Gesamt Punkte	Item 1	Item 2	Item 3	Item 4	Item 5
1	0.700	0.200	0.100	0.000	0.000
2	0.667	0.667	0.333	0.333	0.000
3	0.750	0.833	0.667	0.583	0.167
4	0.900	0.900	1.000	0.900	0.300

Almo meldet hier:

Die Wahrscheinlichkeit einer Ja-Antwort auf folgende Items ist eine nicht-monotone Funktion des Gesamtpunkts.

Item 1

Personen mit einem Gesamtpunkt 1 (aus der Beantwortung der 5 Items) haben zu 70 % Item 1 positiv beantwortet. Es darf dann nicht sein, dass Personen, die mehr Gesamtpunkte erzielt

haben, mit einer Wahrscheinlichkeit von weniger als 70 % Item 1 positiv beantworten. Das ist im Beispiel jedoch bei 2 Gesamtpunkten der Fall. Die Wahrscheinlichkeit beträgt dort 67 %

Für ein Item besteht Monotonie, wenn mit steigendem Gesamtpunkt die relative Häufigkeit einer positiven Antwort grösser wird oder mindestens gleich bleibt. Die Wahrscheinlichkeit einer positiven Antwort muss monoton steigen.

Die Monotonie des Items ist gestört, wenn die relative Häufigkeit einer positiven Antwort für den Gesamtpunkt x kleiner ist als die für den vorausgehenden Gesamtpunkt $x-1$.

Das Ogiven-Modell liefert letztendlich einen Skalenwert (einen „Fähigkeitswert“) je Gesamtpunkt. Dabei gilt die Grundannahme: Der Skalenwert ist eine monotone Funktion des Gesamtpunktwertes. Trifft das nicht zu, dann sind die Ergebnisse unbrauchbar. Es sollte dann eine neue Analyse mit anderen oder weniger Items gerechnet werden.

Ein weiteres Ergebnis sind die beiden Parameter der Ogive m_i (der Mittelwert, interpretiert als der Skalenwert des Items) und s_i (die Varianz, interpretiert als die Diskriminationsfähigkeit des Items). Ist s_i hoch (Erfahrungswert: über 1.0), bedeutet dies, dass das Item zwischen Ja- und Nein-Sagern schlecht differenziert, also dass viele Messfehler enthalten sind (vgl. Graphik 1). Es sollten deshalb Items mit großer Varianz eliminiert werden. Ein drittes Kriterium ist der Signifikanztest: es wird getestet, ob die empirischen Werte von denen signifikant abweichen, die genau den berechneten Parametern entsprechen. Items mit einem hohen Chi-Quadrat-Wert entsprechen dem Modell nicht. Der Forscher kann nun anhand dieser drei Kriterien Items eliminieren, bis er eine skalierbare Teilmenge von Items gefunden hat; das Programm selbst trifft keine Auswahl.

Steht nun die Skala fest, können die Skalenwerte ebenso wie beim Rasch-Modell berechnet und in den Datenvektor jeder Person als zusätzliche Variable eingeschrieben werden.

P15.3.1 Das Modell

Die folgenden Ausführungen sind eine von Kurt **Holm** gekürzte und leicht überarbeitete Version des Kapitels 2.4 Das Ogiven-Modell von Hermann Denz: Analyse latenter Strukturen (1982)

Die Wahrscheinlichkeit p , dass eine Person g ein Item i positiv beantwortet, ergibt sich gemäß nachfolgender Gleichung (1). Als Funktion wird dabei die Ogive der Normalverteilung (deren Integral) angenommen. Die Anteilswerte p werden als Abszissenwerte der Normalverteilungs-Ogive interpretiert. Dabei wird angenommen, dass p nur die Funktion einer einzigen latenten Variablen x ist.

$$(1) \quad p_{ig} = \frac{1}{\sqrt{2\pi}} * \int_{-\infty}^{z_{ig}} e^{(-\frac{z^2}{2})} dz$$

Dabei ist

p_{ig} = Wahrscheinlichkeit, dass die Person mit dem Gesamtpunktwert g das Item i positiv beantwortet

Der Index i kennzeichnet das Item und g den Gesamtpunkt.

Das gilt für alle folgenden Formeln

$$(2) \quad z_{ig} = \frac{x_g - M_i}{s_i}$$

wobei x =Fähigkeitswert der Person mit Gesamtpunktwert g in der latenten Variablen X

M =Schwierigkeitswert des Items i

s =Standardabweichung des Items i ;

das ist die Fähigkeit des Items i die Personen zu trennen (Diskriminierungsfähigkeit von i)

Je grösser s umso schlechter die

Diskriminierungsfähigkeit des Items - und umgekehrt

Die Parameter x und M sind inhaltlich als Personen- und Itemparameter definiert. s kann als Diskriminierungsfähigkeit des Items interpretiert werden (der Reziprokwert $a=1/s$ ist die Steigung der Ogive in ihrem Wendepunkt). Ein großer Wert in s bedeutet, dass das Item sehr wenig differenziert.

Der im folgenden dargestellte Algorithmus zur Schätzung von x , M und s (bzw. $1/s$) wurde von L.R. Tucker (1952) entwickelt. Für die Berechnung wird obige Gleichung (2) umformuliert

$$(3) \quad z_{ig} = a_i * x_g + b_i$$

wobei (3a) $a_i = 1/s_i$

$$(3b) \quad b_i = -M_i/s_i$$

Es wird unterstellt, dass der Skalenwert einer Person eine monotone Funktion des Gesamtpunktwertes ist. Wird von dieser Annahme ausgegangen, dann werden entsprechend nicht—monotone Items als nichtmodellkonform ausgewiesen

Um eine statistisch eindeutige Lösung zu finden, muss jedoch nachgewiesen werden, dass der Gesamtpunktwert tatsächlich eine erschöpfende Statistik des Skalenwertes ist. Ein allgemeiner Beweis dafür existiert nicht. Für das logistische Rasch-Modell besteht ein solcher Beweise (Fischer 1974, S.193ff, Fischer 1981).

Um diese Aufgabe lösen zu können werden Gewichte w eingeführt

$$(4) \quad w_{ig} = \frac{10 * e^{-z_{ig} * z_{ig}}}{p_{ig} * (1 - p_{ig})}$$

p_{ig} = das ist der empirische Anteil der Personen mit Gesamtpunkt g , die das Item i positiv beantworten.

z_{ig} = das ist der Abszissen-Wert (Wert auf der x-Achse) von p_{ig} aus der Standard-Normalverteilung

Die Überlegung, die hinter dieser Gewichtung steht, ist folgende: Die Zuverlässigkeit der z-Werte nimmt an den Extremen ab. Deshalb wird eine Gewichtung eingeführt, die den Fehler über alle Werte von z vereinheitlicht. Die verschiedenen vorgeschlagenen Gewichte sind immer proportional zur quadrierten Ordinate, die dem jeweiligen z-Wert entspricht

Die gesuchten Parameter a , b , x werden so geschätzt, dass der nachfolgende Ausdruck F ein Minimum wird:

$$F = \sum_i \sum_g w_{ig} (z_{ig} - (a_i x_g + b_i))^2$$

F ist also die gewichtete Quadratsumme der Differenz von empirischem und theoretischem z-Wert.

p , z und w sind aus den vorhandenen Daten bekannt, bzw. können leicht berechnet werden

Die Parameter a , b , x werden wie folgt aus den Daten geschätzt

$$(5a) \quad a_i = (A * D - B * C) / (A * E - B * B)$$

$$(5b) \quad b_i = (C - a * B) / A$$

$$(5c) \quad x_g = (F - G) / H$$

Wobei

$$A = \sum w_{ig}$$

$$B = \sum w_{ig} * x_g$$

$$C = \sum w_{ig} * z_{ig}$$

$$D = \sum w_{ig} * z_{ig} * x_g$$

$$E = \sum w_{ig} * x_g^2$$

$$F = \sum w_{ig} * z_{ig} * a_i$$

$$G = \sum w_{ig} * a_i * b_i$$

$$H = \sum w_{ig} * a_i^2$$

Dieses Gleichungssystem ist nicht direkt lösbar, sondern nur iterativ. Für den Vektor x werden Startwerte angenommen, in a und b eingesetzt und daraus neue x-Werte berechnet bis die gewünschte Stabilität erreicht ist. Dann enthält der x-Vektor die Personparameter; aus b können die Itemparameter M (Schwierigkeitsparameter) berechnet werden; a ist die Steigung der Ogive im Wendepunkt.

P15.3.2 Items für das Ogiven-Modell

Für die mit dem Ogiven-Modell zu skalierenden Variablen gelten folgende Möglichkeiten und Bedingungen:

1. In die Analyse dürfen nur dichotome Items aufgenommen werden.
2. Die Items müssen 0 - 1 kodiert sein

Ist Bedingung 1 und/oder 2 für ein Item nicht erfüllt, dann muss es umkodiert werden.

Beispiel: Ein Item V3 ist kodiert mit 3 und 4. Dann muss geschrieben werden

V3 (3=0; 4=1)

Beispiel: Ein Item V3 ist kodiert mit 3,4,5. Dann muss geschrieben werden

V3 (3,4=0; 5=1) oder V3 (3=0; 4,5=1)

3. Die Items müssen in Richtung der Messdimension kodiert sein.
Die Antwort "1" bedeutet, dass die Person höher auf der Messdimension liegt - als wenn sie mit "0" geantwortet hätte.
Ist Bedingung nicht erfüllt, dann muss das Item "umgedreht" werden
Beispiel: Item V3 ist im Vergleich zu den anderen Items invers kodiert
Dann muss geschrieben werden: V3 (0=1; 1=0)
Beachte Bedingung 5 und 6.

4. Die Items einer "Item-Batterie" dürfen nur auf einer Dimension messen und alle müssen auf derselben Dimension messen. (Bedingung der Eindimensionalität)

5. Alle Items einer Item-Batterie müssen in dieselbe Richtung kodiert sein.
Beispiel: Bei einem Wissenstest bedeutet "0" üblicherweise "Aufgabe nicht gelöst" und "1" bedeutet "Aufgabe gelöst". Dies muss für alle Items gelten.
6. Es ist jedoch zulässig, dass alle Items invers kodiert sind. Dann ermittelt das Ogiven-Modell die "Leichtigkeit" der Items und nicht deren "Schwierigkeit". In einem solchen Fall empfiehlt es sich jedoch, alle Items "umzudrehen", da im Programm-Ergebnis nur von "Schwierigkeit" die Rede ist und somit die Ergebnisse falsch interpretiert werden könnten.
BEACHTET: Bei inverser Kodierung ist die Behandlung fehlender Werte mit der Methode 1 (in Programm-Maske Prog15ma - siehe Abschnitt P15.3.4) nicht zulässig. Bei dieser Methode wird ein fehlender Wert durch 0 ersetzt. Bei inverser Kodierung würde das "Aufgabe gelöst" bedeuten.
- 7 Die Items müssen unterschiedliche "Schwierigkeitswerte" besitzen. Bei einem Wissenstest z.B. bedeutet dies, dass Wissensfragen unterschiedlich schwer zu "lösen" sind. Es ist jedoch zulässig, dass einige Items dieselbe Schwierigkeit besitzen.
8. Die Items dürfen (im Unterschied zum Rasch-Modell) ungleiche Trennschärfen (Diskriminationsfähigkeiten) besitzen. Geometrisch bedeutet dies: Die Steigungen der Ogiven können verschieden sein. Grafisch betrachtet: Die Ogiven können sich schneiden

P15.3.3 Test des Ogiven-Modells

Es wird der von Wright und Panchapakesan (1969) für das Rasch-Modell entwickelte Test verwendet wird. (vgl. auch Sixtl 1976, S.55).

$$(6) y_{ig} = (n_g * p_{ig} - n_g * r_{ig}) / \sqrt{n_g * r_{ig} * (1 - r_{ig})}$$

$$(7) Chi_i^2 = \sum y_{ig}$$

$$(8) df = g_{max} - g_0$$

n_g = Zahl der Personen mit (demselben) Gesamtpunktwert g

p_{ig} = das ist der empirische Anteil der Personen mit
(demselben)Gesamtpunkt g , die das Item i positiv
beantwortet haben

r_{ig} = theoretische Anteil (Wahrscheinlichkeit) der Personen mit
(demselben) Gesamtpunktwert g , die das Item i
positiv beantwortet haben. Das ist der Ordinaten-Wert
(y -Wert) aus der errechneten Ogive.

Der Test von Wright und Panchapakesan überprüft für den Item i , ob der empirische und der theoretische Anteil der Personen mit demselben Gesamtpunktwert, die das Item i positiv beantwortet haben, signifikant übereinstimmen. Der Test ist positiv, wenn die aus Chi-Quadrat hervorgegangene Signifikanz p größer als 0.05 bzw. wenn $(1-p)*100$ kleiner 95.0 ist.

df = Freiheitsgrade

g_{max} = maximaler Gesamtpunkt (=Zahl der Items)

g_0 = Zahl der Gesamtpunkte, die nicht besetzt sind. Da für den Gesamtpunkt 0 und den maximalen Gesamtpunkt die empirische Wahrscheinlichkeit p_{ig} notwendigerweise =0 bzw. =1 ist, muss $g_0 \geq 2$ sein

P15.3.4 Programm-Maske Prog15ma für das Ogiven-Modell

Prog15ma ist die Programm-Maske gilt für das Ogiven-Modell. Sie ist im Prinzip identisch mit der nachfolgen in Abschnitt P15.4 abgebildeten Programm-Maske Prog15m1 für alle Modelle der LSA. Zusätzlich ermöglicht sie noch das Speichern der Skalenwerte (Fähigkeitswerte) der Personen zusammen mit den schon vorhandenen Daten in eine neue Datei.

Prog15ma.Msk
Ogiven-Modell der Latent Structure Analysis
mit Speichern der Skalenwerte der Probanden in eine Datei

Beispiel:
Die Bereitschaft von Studenten zu nichtkonventionellen politischen Aktivitäten soll erfragt werden. Die Fragen lauten:

Item 1: Würden Sie sich an einer Unterschriftenaktion beteiligen
Item 2: ... einen Brief an einen Politiker schreiben
Item 3: ... an einer Kundgebung teilnehmen
Item 4: ... in einer Bürgerinitiative mitarbeiten
Item 5: ... an der Besetzung eines öffentlichen Gebäudes teilnehmen

BEACHTEN:

- Die Variablen müssen 0-1 kodiert sein. Dabei bedeutet 1 = ja und 0 = nein. Ist dies nicht der Fall, dann müssen die Items entsprechend umkodiert werden
- Die Variablen müssen in Richtung der Messdimension kodiert sein. 1 bedeutet also: "liegt höher auf der Messdimension" als 0. Ist dies nicht der Fall, dann müssen die Items in ihrer Kodierungsrichtung umgedreht werden

Die für jeden Probanden ermittelten Skalenwerte können in eine neue (Skalenwert-) Variable gegeben werden. Der um die Skalenwert-Variable verlängerte Datensatz kann dann in eine neue Datei geschrieben werden

Programm-Bedienung ---> Hilfe

Speicher fuer x Variable

Vereinbare Variable=

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

Variablennamen

Datei der Variablennamen

zeige = Namensdatei in Output zeigen
leer = nicht zeigen

Freie Namensfelder

Leere alle Eingabefelder dieser Sub-Box

<input type="button" value="↔"/>	<input type="text" value="Name9=Unterschrift"/>
<input type="button" value="↔"/>	<input type="text" value="Name10=Brief"/>
<input type="button" value="↔"/>	<input type="text" value="Name11=Kundgebung"/>
<input type="button" value="↔"/>	<input type="text" value="Name12=Buergerinitiative"/>
<input type="button" value="↔"/>	<input type="text" value="Name13=Besetzung"/>

erzeuge zusätzliche Namensfelder

Variablennamen in Datei speichern

Eingabefeld leer = nicht speichern

Datei aus der gelesen wird

bei Datei-Problemen

Format der Daten

der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_V

Wenn Dateiformat FIX oder Nicht-Standard-FREI




die zu zu skalierenden Variablen

Option: Ein- und Ausschliessen von Untersuchungseinheiten

BEACHTEN:

Für die zu skalierenden Variablen gelten folgende 2 Bedingungen:

1. Die Variablen müssen 0-1 kodiert sein
Ist dies nicht der Fall, dann müssen die Variablen in der nachfolgenden Umkodierungs-Box entsprechend umkodiert werden
2. Die Variablen müssen in Richtung der Messdimension kodiert sein. 1 bedeutet also: "liegt höher auf der Messdimension" als 0
Ist dies nicht der Fall, dann müssen die Variablen in ihrer Kodierungsrichtung umgedreht werden

	Option: Umkodierungen und Kein-Wert-Angaben
	Option: Skalenwerte ermitteln und speichern
	Grafik-Optionen
Programmende	

P15.3.4.1 Erläuterungen zu den Eingabeboxen


Fast alle nachfolgenden Eingabeboxen sind in aller Ausführlichkeit im Almo-Dokument Nr. 0 "P0 Arbeiten mit Almo" erklärt - so dass sie hier relativ kurz erläutert werden können.

Eingabebox: Speicher für x Variable

Speicher fuer x Variable	Hilfe
Vereinbare Variable= <input type="text" value="20"/>	

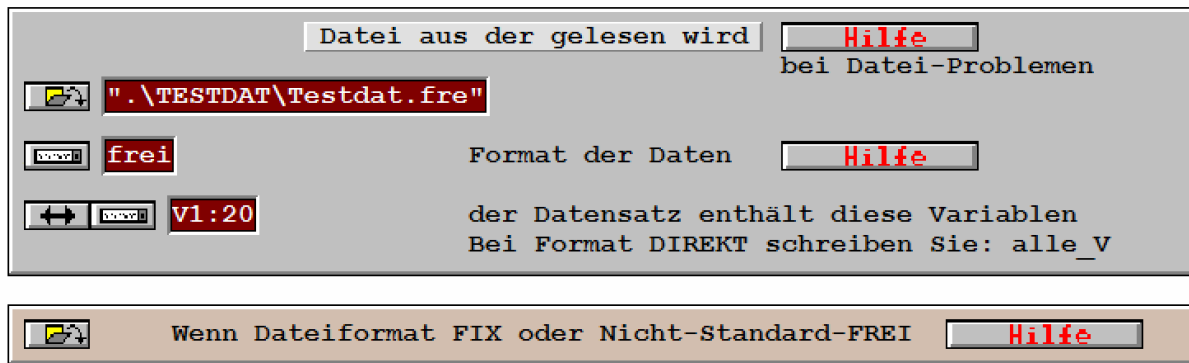
Geben Sie mindestens so viele Variable an, wie der Datensatz umfasst; eher mehr. Siehe dazu das oben erwähnte Almo-Dokument Nr. 0 "Arbeiten mit Almo", Abschnitt P0.1.

Eingabebox: Vereinbarungen

	Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert
---	--

Diese Eingabebox nur öffnen, wenn Almo nach einem 1. Programmlauf dazu auffordert. Siehe dazu das oben erwähnte Almo-Dokument Nr. 0, Abschnitt P0.2

Eingabebox: Datei aus der gelesen wird



Geben Sie hier den Namen der Datei an, in dem sich die zu analysierenden Daten befinden. Schreiben Sie den vollen Pfad- und Dateinamen. Zulässig ist auch die Windows Kurzform, bei der bis einschliesslich *Almo* der Pfad durch einen Punkt vertreten werden kann (wie oben in der Abbildung). Die häufigsten Probleme, die Benutzer mit Almo haben treten an dieser Stelle auf. Lesen Sie insbesondere die Hilfe zum "Format". Klicken Sie dazu auf den Hilfefknopf bei der Format-Anweisung.

In Prog15m1 werden als Beispiel die Daten "Testdat.fre" nach dem Ogiven-Modell ausgewertet. Die Daten wurden nicht empirisch gewonnen. Sie wurden simuliert. Es wurden 61 Datensätze erzeugt. Die ersten und die letzten Datensätze sind folgende

```

1 1 1 1 4 4 2 4 1 1 1 1 1 2 3 3 3 1 3 0
1 1 1 1 5 3 1 6 1 1 1 1 0 1 3 3 1 3 3 6
1 1 1 2 4 2 3 1 1 1 1 0 1 3 3 2 3 3 3 1
1 1 1 2 2 1 4 2 1 1 1 1 0 1 3 3 3 2 2 2
1 1 2 1 4 3 1 8 1 1 1 1 1 3 3 3 1 3 3 7
1 1 2 1 4 5 1 9 1 1 1 1 1 3 1 3 1 3 3 9
      . . . . .
      . . . . .
      . . . . .
1 2 2 2 7 1 4 5 0 0 0 0 0 3 3 3 2 1 3 5
1 2 2 2 2 5 3 6 0 0 0 0 0 1 2 3 3 1 3 6
1 2 2 2 2 6 3 4 0 0 0 0 0 3 3 1 2 1 3 3
1 2 2 2 6 3 7 2 0 0 0 0 0 1 3 2 3 2 3 2
1 2 2 2 4 8 1 6 0 1 0 0 0 3 3 3 2 3 1 6
1 2 2 2 9 7 2 5 0 0 0 0 0 3 3 3 1 1 1 5

```

Die Datei umfasst 20 Variable (Spalten) und 61 Datensätze (Zeilen). Für die Analyse werden die Variablen V9 bis V13 ausgewählt.

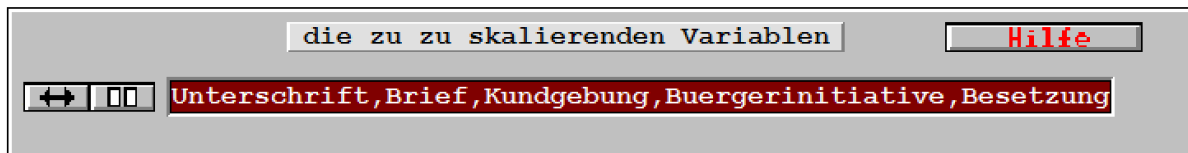
Die Daten befinden sich im Format FREI. Die Zahlenwerte sind durch ein Blank (oder Komma) getrennt. V9 bis V13 sind dichotom und mit 0,1 kodiert. Der Benutzer kann die Daten anschauen, wenn er in der Dateibox auf den Öffne-Knopf vor dem Dateinamen klickt oder in der Knopfleiste am Oberrand des Almo-Fensters den Knopf "Zeige Daten" klickt.

Empirische Daten enthalten in der Regel "fehlende Werte" (in Almo auch "*KeinWert*" genannt). Ist für einen Item "*KeinWert*" mit einem vom Benutzer festgelegten Wert kodiert (z.B. mit der Zahl -9) dann muss in der nachfolgend beschriebenen "Umkodierungsbox" eine entsprechende Deklaration erfolgen, z.B. so

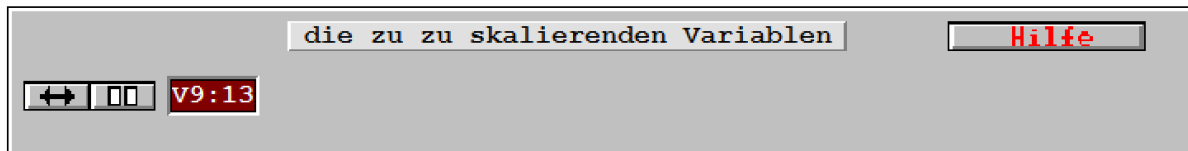
V4 (-9 = KeinWert)

Siehe in der Umkodierungsbox den Hilfefknopf "wie man die Kein-Wert-Angabe schreibt"

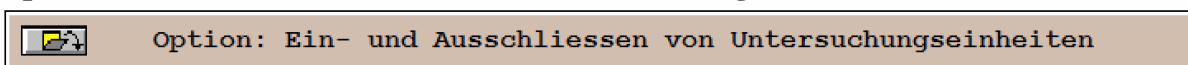
Eingabebox: Die zu skalierenden Items



In der Eingabebox "Variablennamen" wurde den zu skalierenden Variablen V9 bis V13 Namen gegeben. Hier können nun wahlweise diese Variablennamen eingesetzt werden oder auch sehr einfach geschrieben werden:



Optionsbox: Ein- und Ausschliessen von Untersuchungseinheiten.



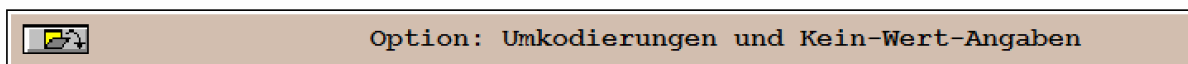
Siehe die ausführliche Darstellung dieser Optionsbox im Almo-Dokument Nr.0 „Arbeiten mit Almo“, Abschnitt P0.7.

Optionsbox: Umkodierungen und Kein-Wert-Angaben

BEACHTEN:

Für die zu skalierenden Variablen gelten folgende 2 Bedingungen:

1. Die Variablen müssen 0-1 kodiert sein
Ist dies nicht der Fall, dann müssen die Variablen in der nachfolgenden Umkodierungs-Box entsprechend umkodiert werden
2. Die Variablen müssen in Richtung der Messdimension kodiert sein. 1 bedeutet also: "liegt höher auf der Messdimension" als 0
Ist dies nicht der Fall, dann müssen die Variablen in ihrer Kodierungsrichtung umgedreht werden



Siehe dazu die ausführliche Erläuterung im Almo-Dokument Nr. 0 "Arbeiten mit Almo", Abschnitt P0.5.

Sind Items nicht 0,1 kodiert dann müssen sie hier umkodiert werden. Siehe dazu auch oben die ausführliche Beschreibung der Kodierung von Items in Abschnitt P15.3.2.

Nicht vergessen werden darf, dass empirische Daten in der Regel auch "fehlende Werte" besitzen (in Almo "KeinWert" genannt). Ist für ein Item "KeinWert" mit einem vom Benutzer festgelegten Wert kodiert (z.B. mit der Zahl -9) dann muss in der "Umkodierungsbox" eine entsprechende Deklaration erfolgen, z.B. so

V4 (-9 = KeinWert)

Wird die Optionsbox geöffnet, dann sieht man folgendes:

X Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)
 ↔ Leere alle Eingabefelder

Umkodierungen und Kein-Wert-Angaben

Bedienung der Box
wie man umkodiert
wie man Kein-Wert-Angabe schreibt

Inkludiere eine Datei mit Umkodierungen und Kein-Wert-Angaben

Eingabefelder für Umkodierungen und Kein-Wert-Angaben

Kontrollieren, ob Umkodierung so erfolgt wie gewünscht

Es werden 3 Sub-Boxen präsentiert. Nun muss noch die mittlere Sub-Box geöffnet werden. Dies ist ein Ausschnitt aus der geöffneten mittleren Sub-Box:

X Loesche wieder diese Sub-Box (Voreinstellungen wieder gueltig)

Eingabefelder für Umkodierungen und Kein-Wert-Angaben

↔	↓	V1:10 (-9=KeinWert)
↔	↓	V3 (3=0; 4=1)
↔	↓	V4 (3,4=0; 5=1)
↔	↓	V5 (0=1; 1=0)
↔	↓	V6 (1:5=0; 6:10=1)

erzeuge zusätzliche Felder für Umkodierungen / Kein_Wert-Angaben

In diesem Beispiel wird unterstellt, dass die Variablen V1 bis V10 dem Ogiven-Verfahren unterworfen werden. Fehlende Werte in diesen Variablen sind beim Schreiben der Daten mit -9 kodiert worden.

V3 ist dichotom, wurde aber mit 3 und 4 kodiert. V4 ist trichotom mit den Werten 3,4,5. V5 ist invers kodiert. V6 besitzt 10 Ausprägungen. Die Variablen V7 bis V10 sind korrekt 0,1 kodiert.

Beispielsweise könnten die Umkodierungsanweisungen so sein, wie oben eingetragen.

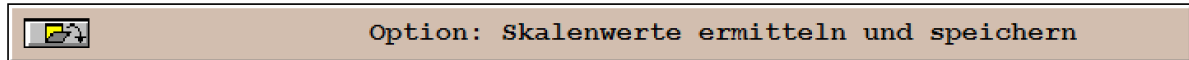
1. Eingabefeld: Zuerst wird deklariert, dass der Wert -9 für *KeinWert* steht.
2. Eingabefeld: Das Item V3 ist in den Daten mit 3 und 4 kodiert. Es wird auf 0,1 umkodiert.
3. Eingabefeld: Item V4 besitzt 3 Ausprägungen, die mit 3,4,5 kodiert sind. Durch die Umkodierungsanweisung werden 3 und 4 zu 0 sammengefasst und 5 wird auf 1 gesetzt.
4. Eingabefeld: V5 ist im Vergleich zu den anderen Items *invers* kodiert. Die Antwort "1" bedeutet: Die Person liegt niedriger auf der Messdimension. "0" bedeutet liegt höher. Dann muss 0 zu 1 und 1 zu 0 umkodiert werden
5. Eingabefeld: Ein Item V6 mit vielen Ausprägungen wird *dichotomisiert*. Hier ist Vorsicht am Platze. Das Ergebnis kann wesentlich durch die Lage des Trennpunkts beeinflusst werden.

Alternative Kein-Wert-Behandlung

Ist auch nur ein Item gleich *KeinWert*, dann wird von Almo der gesamte Datensatz ausgeschlossen. Hier bietet sich jedoch eine alternative Kein-Wert-Behandlung an. Fehlt ein Wert, dann kann der Forscher gelegentlich so interpretieren: "Die befragte Person konnte die durch das Item gestellte Aufgabe nicht lösen". Wenn diese Interpretation zulässig ist, dann kann man *KeinWert* auf 0 kodieren. Man schreibt also die obige 1. Anweisung so:

V1:10(-9=0)

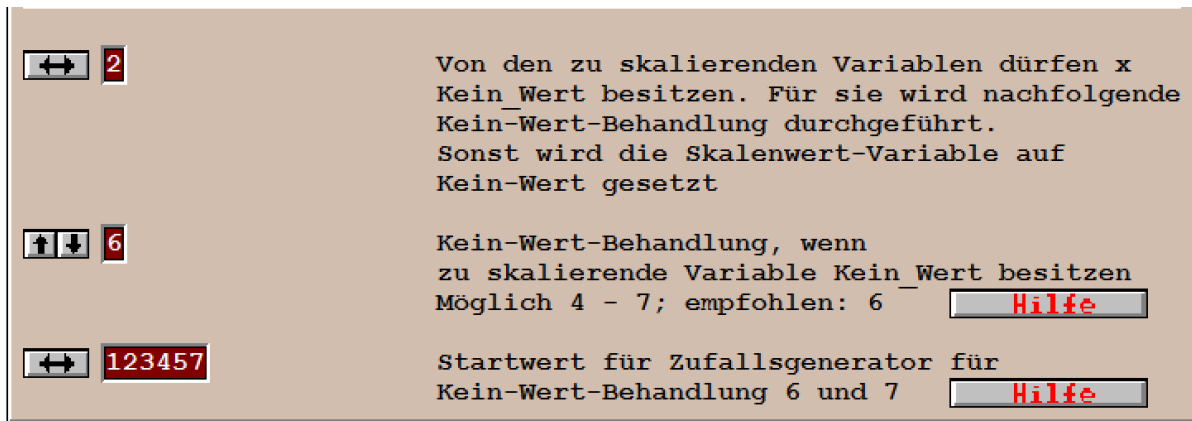
Optionsbox: Skalenwerte ermitteln und speichern



Die geöffnete Optionsbox:



The dialog box has a title bar with the text "Option: Skalenwerte ermitteln und speichern" and a "Hilfe" button on the right. Below the title bar is a close button (X) and the text "Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)". Below that is a path field containing ".\PROGS\Skalwert" with a file explorer icon to its left. The main area contains instructions: "Geben Sie einen neuen Dateinamen ohne Erweiterung an. Almo erzeugt 3 Dateien:" followed by a numbered list: 1. "eine nicht lesbare Almo-Arbeitsdatei mit der Erweiterung __.dir", 2. "eine anschaulbare Datei im freien Format mit der Erweiterung __.fre", 3. "eine Datei der Variablennamen mit der Erweiterung __.nam". Below the list, it explains the contents of the files: "In den unter 1. und 2. angegebenen neuen Dateien sind nun enthalten: - die Variablen aus der alten Datei - die Skalenwertvariable, die als letzte Variable hinter die Variablen der alten Datei gestellt wurde". "In der unter 3. angegebenen neuen Datei der Variablennamen sind nun enthalten: - die Variablennamen aus der alten Datei einschliesslich der in der Box 'Freie Namensfelder' angegebenen Namen - der Name 'Skala' für die neue angehängte Skalenwertvariable". At the bottom, there is a field with "Runde .001" and a double-headed arrow icon, with the text "die Skalenwert-Variable transformieren z.B. auf 3 Dezimalstellen runden das ist nicht obligatorisch" to its right.



Im folgenden werden die Begriffe "Skalenwert" und "Fähigkeit" gleichbedeutend verwendet.

1. Eingabefeld:

Wenn Sie einen Dateinamen in das Eingabefeld schreiben, dann

1. erzeugt Almo zwei Dateien mit diesem Namen einmal im Format FREI und einmal im Format DIREKT
2. und speichert in diese die Variablen aus der alten Datei
3. und speichert die Skalenwertvariable (bzw. Fähigkeitsvariable) als letzte Variable hinter die Variablen aus der alten Datei
4. weiterhin erzeugt Almo eine Datei der Variablennamen. Diese enthält
 - die Variablennamen aus der alten Datei einschliesslich der in der Box "Freie Namensfelder" angegebenen (oder eventuell modifizierten) Namen
 - den Name "Skala.." für die neue angehängte Skalenwertvariable, wobei Almo anstelle der 2 Punkte die Variablennummer der Skalenwertvariablen einsetzt.

Der Benutzer kann die Datei der Variablennamen in ein Fenster laden und die Variablen- und Ausprägungsnamen beliebig verändern. Danach muss wieder gespeichert werden.

Almo hat beispielsweise folgenden Namen geschrieben:

```
Name 21=Skala21;
```

Die Ziffer hinter "Skala.." (im Beispiel: 21) ist die Variablennummer der Skalenwertvariablen. Der Benutzer kann diesen Variablenname beliebig verändern, z.B. so

```
Name 21=Leistung;
```

Der Name muss eindeutig sein, d.h. er darf kein 2. Mal auftreten.

2. Eingabefeld:

Die Skalenwertvariable kann umkodiert werden. In das Eingabefeld kann jede der in Almo angebotenen Umkodierungsanweisungen eingetragen werden. Dabei darf die öffnende und schliessende runde Klammer *nicht* geschrieben werden. Es wird nur die Anweisung selbst geschrieben.

In der Regel wird man sich aber mit einer der 3 folgenden Anweisungen begnügen,

1. Die Skalenwertvariable wird gerundet (auf- bzw. abgerundet).
Bleibt das Eingabefeld leer, dann schreibt Almo einen Zahlenwert mit ca. 4 bis 6 Kommastellen. Beispiel: Es soll auf 3 Kommastellen gerundet werden. Dann schreibt man in das

Eingabefeld:

Runde .001 (eine EINS an der 3. Stelle nach dem Dezimalpunkt)

Aus dem Skalenwert z.B. 1.2348 wird dann 1.235. Zum Runden siehe die ausführliche Darstellung im Almo-Handbuch "Teil2: Almo-Programmiersprache", Abschnitt 16.4

2. Die Skalenwertvariable wird durch + - * / (plus, minus, mal, dividiert) transformiert
Man schreibt z.B.: *100
Der Skalenwert wird mit 100 multipliziert. Aus 1.2348 wird 123.48

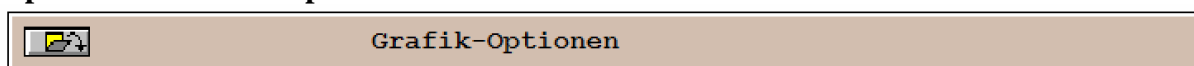
3. Die beiden Transformationen können auch kombiniert werden
Beispiel. Man schreibt: *100; Runde 1
BEACHTEN zwischen die beiden Anweisungen muss ein Semikolon geschrieben werden.
Zuerst wird mit 100 multipliziert, dann wird auf die 1. Stelle vor dem Dezimalpunkt gerundet. Aus 1.2348 wird zuerst 123.48 und schließlich 123.

Der untere Abschnitt der Optionsbox

In ihm kann der Benutzer steuern, wie Almo verfahren soll, wenn einzelne Items oder der Skalenwert selbst *KeinWert* besitzen. Hier kann jedoch ein Problem auftreten. Almo schliesst einen Datensatz aus dem Ogiven-Verfahren aus, wenn auch nur ein Item *KeinWert* ist. In einem 2. Schritt erzeugt Almo dann die neue Datei - bestehend aus den seitherigen (unveränderten) Variablen und den angehängten Skalenwerten (bzw. Fähigkeitswerten). Hier wäre es nun konsequent, einer Person als Skalenwert *KeinWert* zuzuordnen, wenn sie bei der Berechnung des Ogivenmodells gar nicht dabei war. Will der Benutzer dies, dann muss im 1. Eingabefeld des unteren Abschnitts 0 eingesetzt werden.

Es ist aber auch zulässig, den Fähigkeitswert einer Person zu *schätzen*, die nur ein oder nur wenige Items nicht beantwortet hat. Almo bietet hier an, für den fehlenden Wert einen Ersatzwert zu ermitteln, mit diesem den Gesamtpunktwert zu bilden und den Skalenwert für die fragliche Person einzusetzen, der diesem Gesamtpunkt entspricht.

Optionsbox: Grafik-Optionen



Zu dieser Optionsbox siehe Almo-Dokument Nr 0 "Arbeiten mit Almo" Abschnitt P0.7.4

P15.3.5 Ergebnisse aus Prog15ma (Ogiven-Modell)

Ausgegeben werden:

- Die Untersuchungsvariablen.
- Anzahl der analysierten Personen.
- Matrix der relativen Häufigkeiten von "ja"-Antworten für jedes Item und jeden Gesamtpunktwert (=Anzahl von "ja" je Befragten), ohne die beiden Extremgruppen, die kein oder alle Items mit "ja" beantwortet haben.
- Anzahl der Iterationen, die benötigt werden, um das Ergebnis mit der im Programm angegebenen Genauigkeit zu errechnen (im Programm wird als Maximalzahl für die Iterationen 100 angenommen, die Genauigkeit mit 0.001).
- Die Parameter der Ogive: Das Ergebnis kann in zwei Formen dargestellt werden: werden die Wahrscheinlichkeiten einer "ja"-Antwort als Funktion des Skalenwertes auf der latenten Variablen dargestellt, beschreiben die Parameter m_1 (Mittelwert) und s_1 (Standardabweichung) die Funktion; werden jedoch

bereits in z-Werte transformierte Wahrscheinlichkeiten als Ordinate gewählt, beschreibt eine einfache Geradengleichung mit den Parametern a_i (Ordinatenabstand) und b_i (Steigung) den Zusammenhang. Beide Ergebnisse werden ausgegeben.

f) Für jeden Gesamtpunktwert wird der Wert auf der latenten Variablen (Skalenwert für jede Person) berechnet. Die beiden Extremgruppen können nicht berechnet werden, ihre Skalenwerte sind per definitionem -u bzw. +u (für diese beiden Gruppen werden linear extrapolierte Werte ausgegeben).

g) Signifikanztest:

Jedes Item und das Gesamtmodell wird darauf getestet, ob die empirischen Daten durch die berechneten Parameter tatsächlich reproduziert werden können. Es wird derselbe Chi-Quadrat-Test berechnet wie im Rasch-Modell.

Fuer Analyse aus Datenvektor ausgewaehlte Variable

- 9 Unterschrift
- 10 Brief
- 11 Kundgebung
- 12 Buergerinitiative
- 13 Besetzung

Analyse latenter Strukturen

Zahl der eingelesenen Datenvektoren = 61

Zahl der verarbeiteten Datenvektoren = 61

Normal-Ogive als Itemcharakteristik

Variable	Relative Haeufigkeit von Ja-Antworten
9 Unterschrift	0.607
10 Brief	0.541
11 Kundgebung	0.475
12 Buergerinitiativ	0.426
13 Besetzung	0.213

Relative Haeufigkeit von Ja-Antworten

fuer jeden Gesamtpunktwert ohne Extremgruppen

	Untersc V9	Brief V10	Kundgeb V11	Buerger V12	Besetzu V13
1	0.700	0.200	0.100	0.000	0.000
2	0.667	0.667	0.333	0.333	0.000
3	0.750	0.833	0.667	0.583	0.167
4	0.900	0.900	1.000	0.900	0.300

Die Wahrscheinlichkeit einer Ja-Antwort auf folgende Items ist eine nicht-monotone Funktion des Gesamtpunkts.

V9 Unterschrift

Für ein Item besteht Monotonie, wenn mit steigendem Gesamtpunkt die relative Häufigkeit einer positiven Antwort grösser wird oder mindestens gleich bleibt.

Die Monotonie des Items ist gestört, wenn die relative Häufigkeit einer positiven Antwort für den Gesamtpunkt x kleiner ist als die für den vorausgehenden Gesamtpunkt $x-1$. Dieser Fall ist bei V9 gegeben. Personen mit 2 Gesamtpunkten beantworten mit $p=0.667$ V9 positiv, während Personen mit nur 1 Gesamtpunkt mit $p=0.700$ V9 positiv beantworten können.

Anzahl der Iterationen = 13

Trennschaerfe

Schwierigkeits-

Variable	(=Steilheit)	b(i)	parameter	s(i)
	a(i)		m(i)	
9 Unterschrift	0.606	0.620	-1.022	1.650
10 Brief	2.241	0.197	-0.088	0.446
11 Kundgebung	2.418	-0.383	0.158	0.414
12 Buergerinitiativ	3.134	-0.618	0.197	0.319
13 Besetzung	1.579	-1.437	0.910	0.633

a(i) = 1/s(i) s(i)= Standardabweichung des Items i
b(i) = -m(i)/s(i)

Gesamtpunktwert	Skalenwert (=Faehigkeitswert der Probanden)	Haeufigkeit
0	-0.8622 *	15
1	-0.4102	10
2	0.0419	6
3	0.2972	12
4	0.5777	10
5	0.8583 *	8

* = Skalenwert linear extrapoliert

Chi-Quadrat-Test

Variable	Chi-Quadrat	df	Signifikanz	
			p	(1-p)*100
9 Unterschrift	0.7233	3	0.623	37.66
10 Brief	0.3581	3	0.730	27.03
11 Kundgebung	2.0115	3	0.413	58.71
12 Buergerinitiativ	0.4108	3	0.711	28.87
13 Besetzung	0.7483	3	0.617	38.25
Gesamt	4.2521	12		44.86

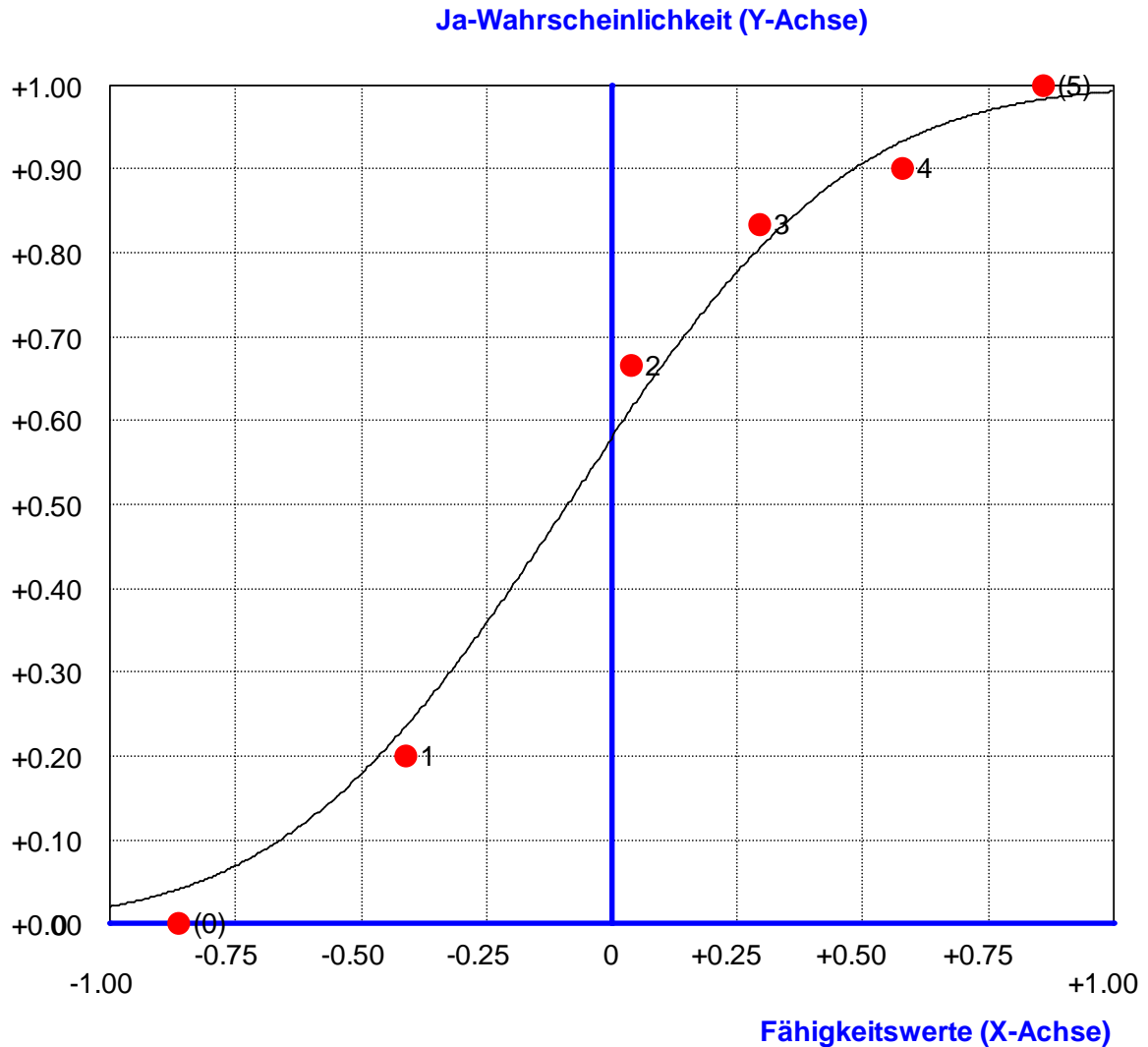
Almo gibt für jedes einzelne Item folgende Detailanalyse und folgende Grafik aus. Wir zeigen die Ausgabe für Item V10

Detailanalyse fuer Variable V10 Brief

Gesamtpunktwert	Skalenwert Faehigkeitswert	"trace line" Ja-Sage-Wahrscheinlichkeit bei jeweiligem Gesamtpunktwert	
		empirisch	theoretisch
0	-0.8622 *)	0.0000	0.0414
1	-0.4102	0.2000	0.2351
2	0.0419	0.6667	0.6144
3	0.2972	0.8333	0.8059
4	0.5777	0.9000	0.9321
5	0.8583 *)	1.0000	0.9830

*) Skalenwert linear extrapoliert

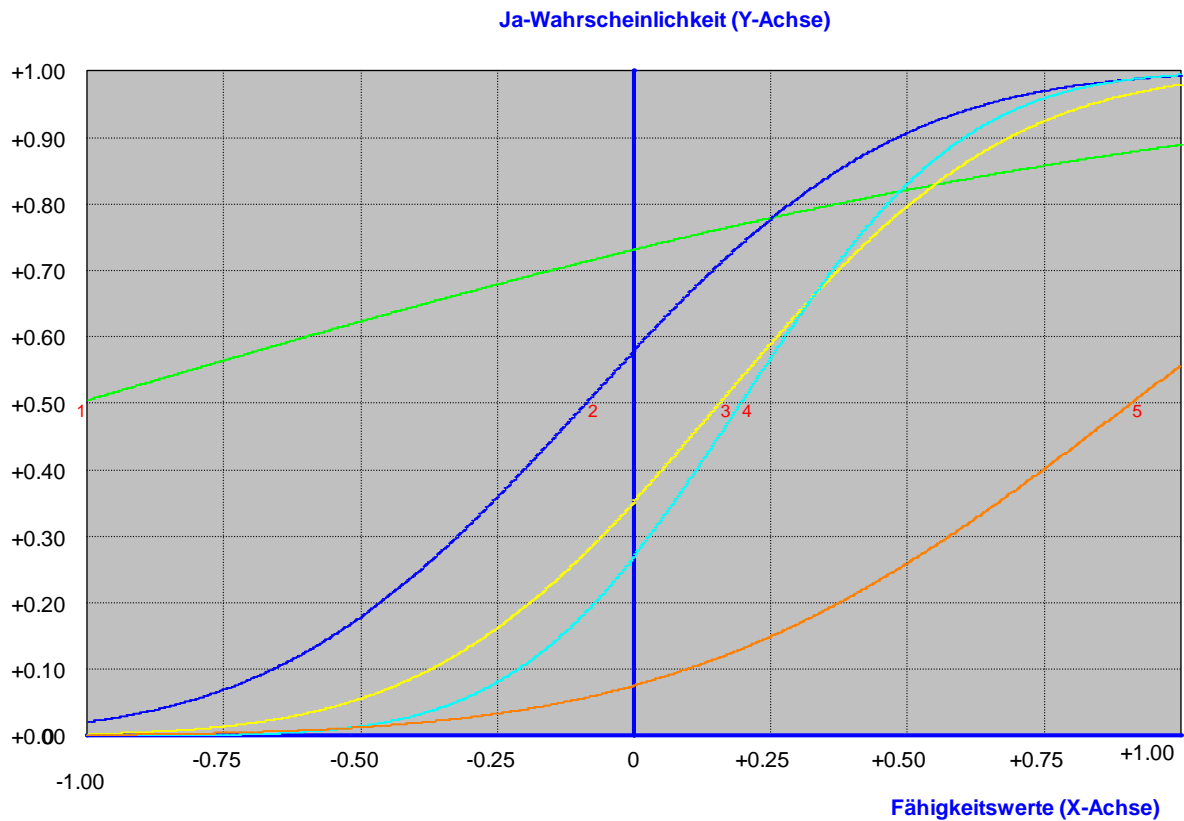
Ogive $Y = \text{Integral}(2.241/2.51 * e^{-(2.241*(X+0.088))*(2.241*(X+0.088))/2})$
 rote Punkte=empirische Ja-Wahrscheinlichkeit bei einem Gesamtpunktwert von x
 Die Ges.punktswerte stehen bei den Punkten. Ges.punktswerte in Klammern geschätzt



Der rote Gesamtpunkt 1 hat eine empirische Wahrscheinlichkeit einer positiven Antwort von genau 0.2000 und eine theoretische Wahrscheinlichkeit (das ist der darüber liegende Punkt auf der Kurve) von 0.2351.

Abschliessend werden alle Kurven in einer gemeinsamen Grafik gezeigt:

Ogiven



Formeln der Kurven:

Ogive: $Y = \text{Integral}(c/W * e^{-(c*(X-b))*(c*(X-b))/2})$
 W=Wurzel aus $2*\pi = 2.5066$
 Y=Wahrscheinlichkeit einer positiven Antwort auf Item i
 X=Faehigkeitswert der Person
 c=Steigung der Ogive von Item i (Steigung im Wendepunkt)
 b=Schwierigkeitsparameter des Item i

Reihenfolge der Kurven von links nach rechts (von kleinstem zu groesstem Schwierigkeitswert)

```

1 gruen    V9  Untersc Y=Integral( 0.606/W * e**(-(0.606*(X+1.022))*(0.606*(X+1.022))/2) )
2 blau    V10 Brief Y=Integral( 2.240/W * e**(-(2.240*(X+0.088))*(2.240*(X+0.088))/2) )
3 gelb    V11 Kundgeb Y=Integral( 2.418/W * e**(-(2.418*(X-0.158))*(2.418*(X-0.158))/2) )
4 tuerkis V12 Buerger Y=Integral( 3.134/W * e**(-(3.134*(X-0.197))*(3.134*(X-0.197))/2) )
5 orange  V13 Besetzu Y=Integral( 1.579/W * e**(-(1.579*(X-0.910))*(1.579*(X-0.910))/2) )
    
```

W ist Wurzel aus $2*\pi = 2.5066$, gekuerzt =2.51

Rangfolge der Items geordnet nach Schwierigkeit von leicht nach schwer

	Trennschaerfe (=Steilheit)	Schwierigkeits- parameter
V9 Unterschrift	0.6062	-1.0220
V10 Brief	2.2405	-0.0878
V11 Kundgebung	2.4182	+0.1584
V12 Buergeriniti	3.1340	+0.1973
V13 Besetzung	1.5790	+0.9098

Skalenwert je Person aus dem Ogiven-Modell:

In der Programm-Maske Prog15ma kann das Berechnen und das Speichern von Skalenwerten (Fähigkeitswerten) für alle Personen angefordert werden. Die Vorgehensweise ist dieselbe wie beim Rasch-Verfahren. Siehe die ausführliche Darstellung im Almo-Handbuch „Sozialwissenschaftliche Skalierungsverfahren“, Abschnitt P14.6.

P15.4 Programm-Maske Prog15m1 für alle Modelle der LSA

Prog 15m1
Latent Structure Analysis (nach Lazarsfeld)

Prog15m1 enthält alle Verfahren der Analyse latenter Strukturen nach Lazarsfeld. Diese können in zwei Typen eingeteilt werden:

"Latent-Trait"-Modelle:

- Modell 1 = lineare Itemcharakteristik
- Modell 2 = Guttman-Skala
- Modell 3 = latent distance Modell (2 Parameter)
- Modell 4 = latent distance Modell (3 Parameter)
- Modell 5 = Ogivenmodell - das "beste" Modell

"Latent-Class"-Modelle:

- Modell 6 = latent class Modell (nur 2 Klassen)
- Modell 7 = latent class Modell (allgemeine Lösung)

Der Unterschied zwischen beiden Modelltypen besteht darin, dass die "latent-trait"-Modelle eine kontinuierliche latente Dimension, z.B. den Grad der politischen Protestbereitschaft oder den Grad der sozialen Distanz, unterstellen.

Bei den "latent-class"-Modellen dagegen wird von einer nominalskalierten latenten Dimension ausgegangen. Die berechneten Klassen können als Typen (Klassen) bezeichnet werden. Zwischen ihnen muss keine Ordnungsrelation bestehen, wie dies für "latent-trait"-Modelle der Fall ist.

Beide Modelltypen setzen dichotome Items voraus. Für die "latent-trait"-Modelle müssen diese in Richtung der Messdimension kodiert sein.

Die "latent-trait"-Modelle setzen ferner u.a. Eindimensionalität der Items voraus.

Das Ogiven-Modell entspricht in etwa dem Modell der Raschskalierung (siehe Prog14ma und Prog14m1) mit dem Unterschied, dass die Items unterschiedliche Diskriminationsfähigkeit besitzen dürfen.

Was ist ein Kurzprogramm ? -->
Bedienung -->

1

Vereinbare Variable= ;

2 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

"C:\Almo7\Testdat\Varnamen.nam"

 zeige = Namensdatei in Output zeigen
leer = nicht

4

Name9=Unterschrift;
 Name10=Brief;
 Name11=Kundgebung;
 Name12=Buergerinitiative;
 Name13=Besetzung;

erzeuge zusätzliche Namensfelder

5 **Datei aus der gelesen wird** **Hilfe**
 bei Datei-Problemen
 frei **Format der Daten** **Hilfe**
 der Datensatz enthält diese Variablen
 Bei Format DIREKT schreiben Sie: alle_U

6 **Wenn Dateiformat FIX oder Nicht-Standard-FREI** **Hilfe**

7 **die zu zu skalierenden Variablen** **Hilfe**

8 **Modell-Angabe**

 1= lineare Itemcharakteristik
 2= Guttman-Skala
 3= latent distance Modell (2 Parameter)
 4= latent distance Modell (3 Parameter)
 5= Ogivenmodell - das beste Modell
 6= latent class Modell (nur 2 Klassen)
 7= latent class Modell (allgemeine Lösung)

9 **Option: Ein- und Ausschliessen von Untersuchungseinheiten**

BEACHT:

Für die zu skalierenden Variablen gelten folgende 2 Bedingungen:

1. Die Variablen müssen 0-1 kodiert sein
Ist dies nicht der Fall, dann müssen die Variablen in der nachfolgenden Umkodierungs-Box entsprechend umkodiert werden
2. Die Variablen müssen in Richtung der Messdimension kodiert sein. 1 bedeutet also: "liegt höher auf der Messdimension" als 0
Ist dies nicht der Fall, dann müssen die Variablen in ihrer Kodierungsrichtung umgedreht werden

10 **Option: Umkodierungen und Kein-Wert-Angaben**

11 **Grafik-Optionen**

12 **Programmende**

Erläuterung zu den Eingabe-Boxen

Die einzelnen Eingabe-Boxen sind im Almo-Dokument Nr. 0 "Arbeiten mit Almo.PDF" im Detail erklärt.

Die Eingabe-Box **Modell-Angabe**

Es werden folgende Modell unterschieden:

- MODELL=1 lineare Itemcharakteristik
- MODELL=2 Guttman-Skala
- MODELL=3 restringiertes latent- distance-Modell (2 Parameter)
- MODELL=4 allgemeines latent- distance-Modell (3 Parameter)
- MODELL=5 Ogivenmodell
- MODELL=6 Latent-class-Modell mit vereinfachter Rotation (nur bei 2 Klassen möglich)
- MODELL=7 Latent-class-Modell (allgemeine Lösung)

Beachte:

- 1) Es wird davon ausgegangen, dass dichotome Daten vorliegen, 1 wird als "ja" interpretiert, alle anderen Ausprägungen als "nein". Ist eine Variable anders kodiert, dann muß sie durch eine Umkodierungsanweisung in eine solche Form gebracht werden, z.B.

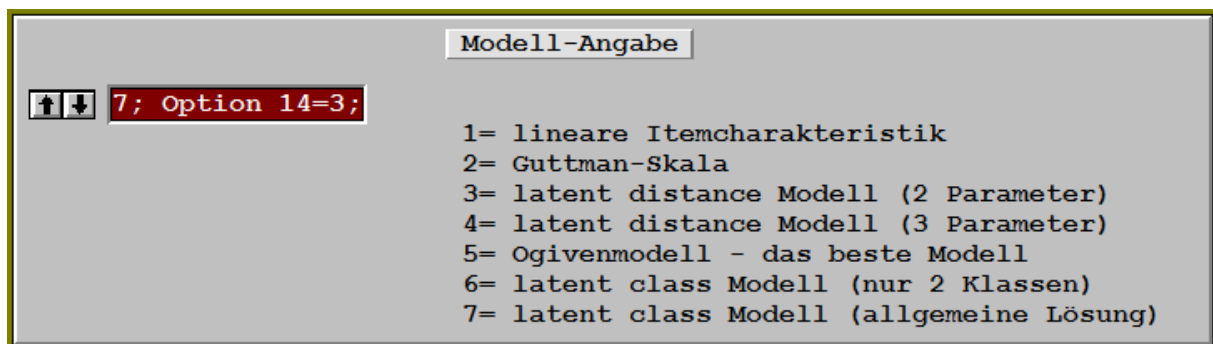
V47(1:3 = 1; 4:6 = 0)

Solchen Umkodierungen bei denen mehrere Ausprägungen zu 0 bzw. 1 zusammengefasst werden, müssen mit Vorsicht durchgeführt werden. Der Schwierigkeitswert der Variablen wird durch die Trennstelle der Dichotomisierung mitbestimmt.

- 2) Bei Modell 4: Die Items müssen fallend nach p (=Prozentsatz der Ja-Antworten) sortiert sein und in dieser Reihenfolge in die Box „die zu skalierenden Variablen“ eingeschrieben werden.

Optionen:

Bei Modell 7 kann mit **Option 14 = n** ; die Zahl der latenten Klassen auf die Zahl n festgelegt werden. In der Eingabebox „Modell-Angabe“ wird dann z.B. eingesetzt:



Nach der mit "7;" angegebenen Modell-Spezifizierung wird dann noch geschrieben
Option 14=... ;

Es können jedoch maximal nicht mehr Klassen erzeugt werden, als Eigenwerte der Matrix p der Kreuzprodukte vorhanden sind

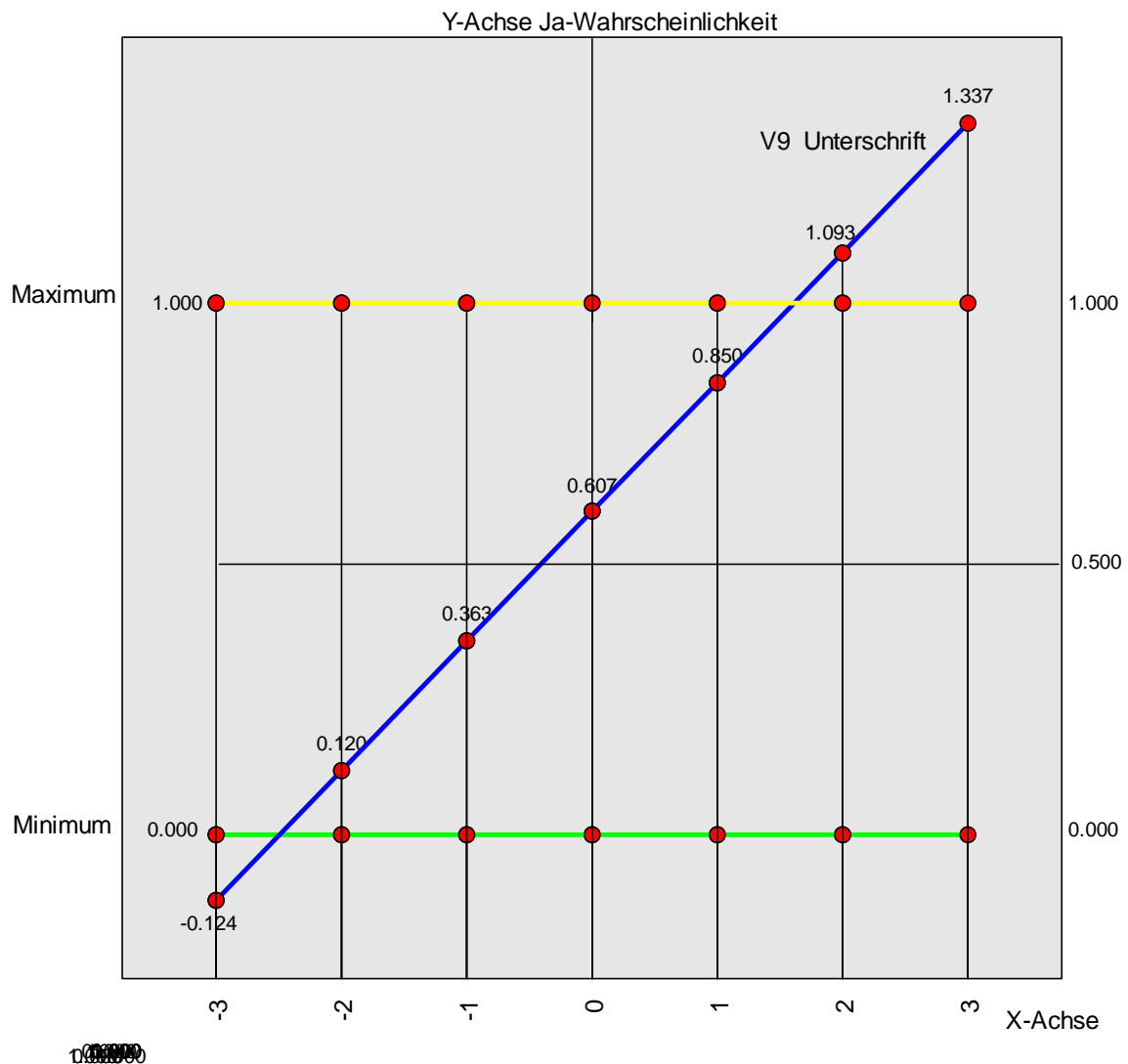
P15.5 Ergebnisse für die Modellen der LSA aus Prog15m1.

Die Ausgabe ist bei jedem Modell anders, da andere Parameter berechnet werden müssen. Vier Ausgabetypen sollen näher beschrieben werden (Modell 1, 3, 5 und 7).

P15.5.1 Lineare Itemcharakteristik (Modell 1)

Betrachten wir aus dem Beispiel die Variable V9 Unterschrift. Also gibt für diese Variable folgende trace-line aus:

theoretische Trace-Line V9



Die Gleichung der Geraden ist

$$Y = a + b \cdot x$$

a = das ist die Ordinate bei $x=0$. Bei V9 ist dies 0.607

b = das ist die Steigung der Geraden. Bei V9 ist $b=0.243$ (bei Kleinster-Quadrate-Schätzung) bzw. $b=0.251$ (bei der Mittelwert-Lösung)

Ein Schwierigkeitswert w für V9 wird von Almo nicht ausgegeben. Eine Möglichkeit bestünde darin, den X-Wert bei einer Ja-Wahrscheinlichkeit von $Y=0.5$ in der Grafik abzulesen. Er wäre ca. $x = -0.4$. Oder gemäß folgender Formel exakt auszurechnen

$$w = (0.5-a) / b \quad \text{Für V9 würde sich ergeben: } w = (0.5-0.607) / 0.243 = -0.440$$

Personenparameter (Fähigkeitswerte) der Personen werden von Almo nicht berechnet. In der Literatur gibt es einen Vorschlag zu Berechnung der Personenparameter von Lazarsfeld und von Fischer. Beide Vorschläge wurden jedoch nicht weiter verfolgt. (Siehe dazu Denz, 1982, S. 32). Üblich ist es, den Gesamtpunkt als Personenparameter zu verwenden. Damit wird jedoch das Modell der linearen Itemcharakteristik banal – es sei denn, es wird zur Item-Selektion verwendet

Almo-Ausgabe aus Modell 1

Ausgegeben werden:

- a) Die Untersuchungsvariablen.
- b) Anzahl der analysierten Personen.
- c) Kovarianzmatrix und die relativen Häufigkeiten je Variable.
- d) Kleinste-Quadrate- und Mittelwert-Lösung für die Parameter der Geradengleichungen:
 - a(i) ist der Ordinatenabstand,
 - b(i) die Steigung der Geraden.

Bei der Mittelwert-Lösung werden zusätzlich die kleinste mögliche, die größte und die Streuung aller möglichen Steigungen angegeben.

- e) Für beide Lösungen wird der Definitionsbereich der latenten Variablen x berechnet, der sich dadurch ergibt, dass keine Wahrscheinlichkeit kleiner 0 oder größer 1 sein darf.
- f) Modelltest: aus den Parametern des Modells werden erwartete Häufigkeiten berechnet und mit Hilfe eines Chi-Quadrat-Tests wird die Signifikanz der Abweichung von dem empirischen Häufigkeiten überprüft. Zum zweiten wird der relative Fehler (Determinante) berechnet und daraus eine Gesamterklärung.

Fuer Analyse aus Datenvektor ausgewaehlte Variable

- 9 Unterschrift
- 10 Brief
- 11 Kundgebung
- 12 Buergerinitiative
- 13 Besetzung

Analyse latenter Strukturen

Zahl der eingelesenen Datenvektoren = 61

Zahl der verarbeiteten Datenvektoren = 61

Modell unter der Annahme linearer Itemcharakteristiken

Kovarianzmatrix $c(i,j)$

	Untersc V9	Brief V10	Kundgeb V11	Buerger V12	Besetzu V13
9 Untersc	0.000	0.098	0.072	0.086	0.067
10 Brief	0.098	0.000	0.153	0.114	0.065
11 Kundgeb	0.072	0.153	0.000	0.158	0.079
12 Buerger	0.086	0.114	0.158	0.000	0.089
13 Besetzu	0.067	0.065	0.079	0.089	0.000

Relative Haeufigkeit je Item

0.607 0.541 0.475 0.426 0.213

Kleinste-Quadrate-Schaetzung

Variable	a(i)	b(i)
9 Unterschrift	0.607	0.243
10 Brief	0.541	0.349
11 Kundgebung	0.475	0.385
12 Buergerinitiativ	0.426	0.370
13 Besetzung	0.213	0.223

Definitionsbereich von x -0.956 bis 1.316

Chi-Quadrat = 0.5279 6 Freiheitsgrade
Signifikanz (1-p)*100 = 23.183
Gesamterklaerung = 0.9683

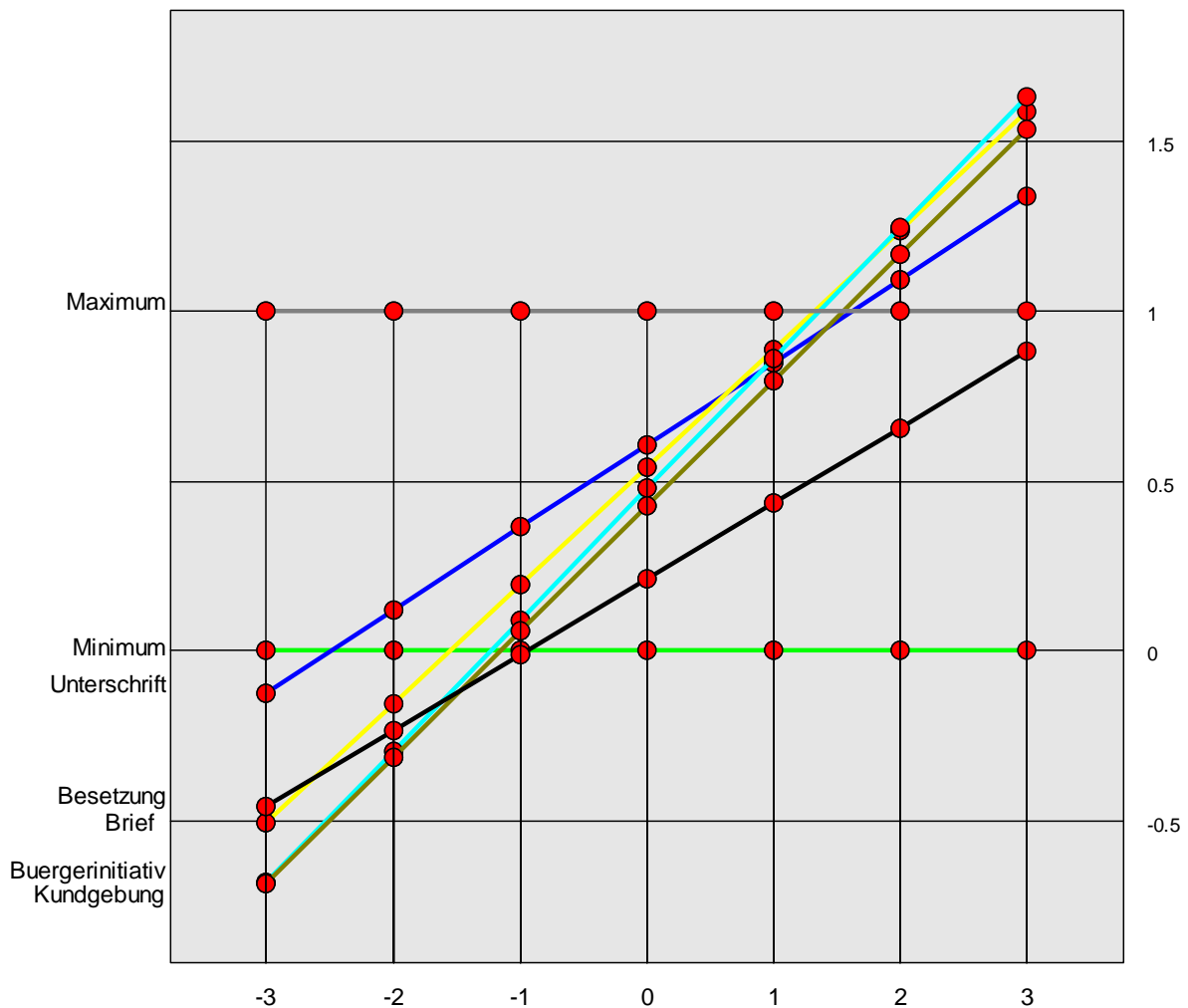
Mittelwert-Loesung

Variable	a(i)	b(i)	b(i)min	b(i)max	b(i)st.abw.
9 Unterschrift	0.607	0.251	0.198	0.319	0.0390
10 Brief	0.541	0.349	0.287	0.455	0.0536
11 Kundgebung	0.475	0.376	0.291	0.461	0.0565
12 Buergerinitiativ	0.426	0.374	0.315	0.433	0.0449
13 Besetzung	0.213	0.228	0.183	0.272	0.0311

Definitionsbereich von x -0.934 bis 1.314

Erläuterung: Das Modell kann nur Personen messen, deren Wert auf der latenten Variablen zwischen -0.934 und +1.314 liegt. Siehe nachfolgende Grafik. Wird die Variable „Besetzung“ (die schwarze Gerade in nachfolgender Grafik) herausgenommen, dann wird der untere Minimalwert geringfügig kleiner, der Definitionsbereich also minimal verbreitert.

Chi-Quadrat = 0.5821 6 Freiheitsgrade
Signifikanz (1-p)*100 = 24.334
Gesamterklaerung = 0.9692



Die Grafik der traces lines wird von Almo zunächst 3-dimensional mit umgebenden Wänden ausgegeben. Das ist nicht immer ideal. Klicken Sie auf den Grafik-Knopf. Dadurch gelangen Sie in den Grafik-Editor. Dort klicken Sie in der linken Randleiste auf „Diverse Positionen“. Es werden verschiedene alternative Darstellungsformen (derselben Grafikdaten) angeboten. Klicken Sie in der oberen Reihe auf das 3. Bild. Es entsteht dann eine 2-dimensionale Darstellung. Diese kann weiter bearbeitet werden. Z.B. sollte man die roten Punkte verkleinern.

P15.5.2 Latent-distance Modell (Modell 3)

Ausgegeben werden:

- a) Die Untersuchungsvariablen.
- b) Anzahl der analysierten Personen.
- c) Kovarianzmatrix und die relativen Häufigkeiten je Variable.
- d) p_{-kl} ist die Wahrscheinlichkeit einer "ja"-Antwort für alle Befragten, deren Skalenwert kleiner als der Skalenwert des Items ist, p_{-gr} für alle Befragten, deren Skalenwert größer ist. $x(i)$ ist der Skalenwert des Items. c_{jj} ist das geschätzte Diagonalglied der Kovarianzmatrix (Minimum, Maximum und Standardabweichung); bei der Guttman-skala entfällt dieser letzte Teil (weil nur ein Parameter berechnet werden muß), bei der allgemeinen Form muß ein

weiteres Korrekturglied geschätzt werden (u_j), wobei dann zusätzlich auch für dieses Minimum, Maximum und Standardabweichung ausgegeben werden.

- e) Für beide Lösungen wird der Definitionsbereich der latenten Variablen x berechnet, der sich dadurch ergibt, dass keine Wahrscheinlichkeit kleiner 0 oder größer 1 sein darf.
- f) Modelltest: aus den Parametern des Modells werden erwartete Häufigkeiten berechnet und mit Hilfe eines Chi-Quadrat-Tests wird die Signifikanz der Abweichung von den empirischen Häufigkeiten überprüft. Zum zweiten wird der relative Fehler (Determinante) berechnet und daraus eine Gesamterklärung.

Fuer Analyse aus Datenvektor ausgewaehlte Variable

- 9 Unterschrift
- 10 Brief
- 11 Kundgebung
- 12 Buergerinitiative
- 13 Besetzung

Analyse latenter Strukturen

Zahl der eingelesenen Datenvektoren = 61

Zahl der verarbeiteten Datenvektoren = 61

Latent-Distance-Modell

Zwei-Parameter-Version

Kovarianzmatrix $c(i,j)$

	Untersc V9	Brief V10	Kundgeb V11	Buerger V12	Besetzu V13
9 Untersc	0.000	0.098	0.072	0.086	0.067
10 Brief	0.098	0.000	0.153	0.114	0.065
11 Kundgeb	0.072	0.153	0.000	0.158	0.079
12 Buerger	0.086	0.114	0.158	0.000	0.089
13 Besetzu	0.067	0.065	0.079	0.089	0.000

Relative Haeufigkeit je Item

0.607 0.541 0.475 0.426 0.213

Variable	p-kl	x(i)	p-gr	cjj/min	cjj/max	cjj/st.abw.
10 Brief	0.118	0.446	0.882	0.095	0.207	0.0470
11 Kundgeb	0.107	0.531	0.893	0.085	0.212	0.0490
12 Buerger	0.106	0.594	0.894	0.114	0.179	0.0271
9 Untersc	0.161	0.343	0.839			
13 Besetzu	0.023	0.800	0.977			

Chi-Quadrat = 0.6054 6 Freiheitsgrade

Signifikanz $(1-p)*100$ = 24.814

Gesamterklaerung = 0.9976

P15.5.3 Ogiven-Modell (Modell 5)

Die Ein- und Ausgabe für dieses Modell wurde oben in Abschnitt Prog15.3.4 und Prog15.3.5 ausführlich dargestellt.

P15.5.4 Latent-Class-Analyse (Modell 6 und 7)

Ausgegeben werden:

- a) Die Untersuchungs-Variablen.

- b) Anzahl der analysierten Personen.
- c) Matrix der p_{ij} (Kreuzproduktematrix; es wird zusätzlich ein Nullitem eingeführt, das immer die Ausprägung 1 annimmt. Es wird mit ‚0‘ oder ‚V0‘ bezeichnet. Das bedeutet, dass in der ersten Zeile und Spalte der Matrix die p_i stehen. Die Diagonalelemente sind nicht empirisch definiert, sie werden iterativ geschätzt, und zwar im ersten Durchlauf aus den empirischen p_{ij} , in den weiteren aus der letzten Lösung.
- d) Unrotierte Eigenvektoren der in c) beschriebenen Matrix.
- e) Aus den empirischen Daten wird zusätzlich die p_{ijk} -Matrix für jedes Item und über alle Items summiert. Alle Elemente, in denen zwei oder drei Mal dasselbe Subskript auftaucht, müssen geschätzt werden - und zwar im ersten Durchgang aus den p_{ij} . In den folgenden Durchläufen werden sie immer aus der vorhergehenden Lösung berechnet.
- f) Aus der p_{ijk} -Matrix wird eine Transformationsmatrix berechnet, durch die die Eigenvektoren der p_{ij} -Matrix rotiert werden. Die Ergebnisse c bis f werden für die erste und dann nach jeweils 5 weiteren Iterationen ausgedruckt.
- g) Lösung der Latent-Class Parameter und die Anzahl der Iterationen, die dafür benötigt wurden (die Anzahl der Iterationen ist im Programm auf maximal 50 begrenzt). Für jedes Item und jede Klasse wird die Wahrscheinlichkeit einer "ja"-Antwort angegeben; die letzte Zeile gibt die relative Häufigkeit der verschiedenen Klassen an.
- h) Überprüfung des Modells: Chi-Quadrat-Test und Determinantenberechnung (in diesem Beispiel ist der Determinantenkalkül nicht berechenbar).
- i) Personparameter: Es werden alle möglichen Antwortvektoren gebildet und für jeden die Wahrscheinlichkeit berechnet, mit der er zu einer der gebildeten latenten Klassen gehört - der jeweils höchste Wert wird durch * gekennzeichnet.

Fuer Analyse aus Datenvektor ausgewaehlte Variable

- 9 Unterschrift
- 10 Brief
- 11 Kundgebung
- 12 Buergerinitiative
- 13 Besetzung

Analyse latenter Strukturen

Zahl der eingelesenen Datenvektoren = 61
 Zahl der verarbeiteten Datenvektoren = 61

Latent-Class-Analysis
 Matrix der $p(i,j)$
 Diagonalelemente geschaetzt nach 50 Iterationen

	Untersc	Brief	Kundgeb	Buerger	Besetzu	
	V0	V9	V10	V11	V12	V13
0	1.000	0.607	0.541	0.475	0.426	0.213
9 Untersc	0.607	0.433	0.426	0.361	0.344	0.197
10 Brief	0.541	0.426	0.415	0.410	0.344	0.180
11 Kundgeb	0.475	0.361	0.410	0.396	0.361	0.180
12 Buerger	0.426	0.344	0.344	0.361	0.320	0.180
13 Besetzu	0.213	0.197	0.180	0.180	0.180	0.091

Unrotierte Eigenvektoren

	1	2
0	0.931	0.361
9 Untersc	0.653	0.007
10 Brief	0.633	-0.121
11 Kundgeb	0.588	-0.218
12 Buerger	0.530	-0.194

```

13 Besetzu | 0.278 -0.117
-----*-----

```

Matrix der p(i,j,k)
Nicht definierte Zellen geschaeztzt aus den p(i,j)

	Untersc	Brief	Kundgeb	Buerger	Besetzu	
	V0	V9	V10	V11	V12	V13
0	2.262	1.761	1.776	1.707	1.549	0.829
9 Untersc	1.761	1.357	1.464	1.368	1.269	0.732
10 Brief	1.776	1.464	1.523	1.559	1.320	0.705
11 Kundgeb	1.707	1.368	1.559	1.609	1.415	0.737
12 Buerger	1.549	1.269	1.320	1.415	1.302	0.707
13 Besetzu	0.829	0.732	0.705	0.737	0.707	0.380

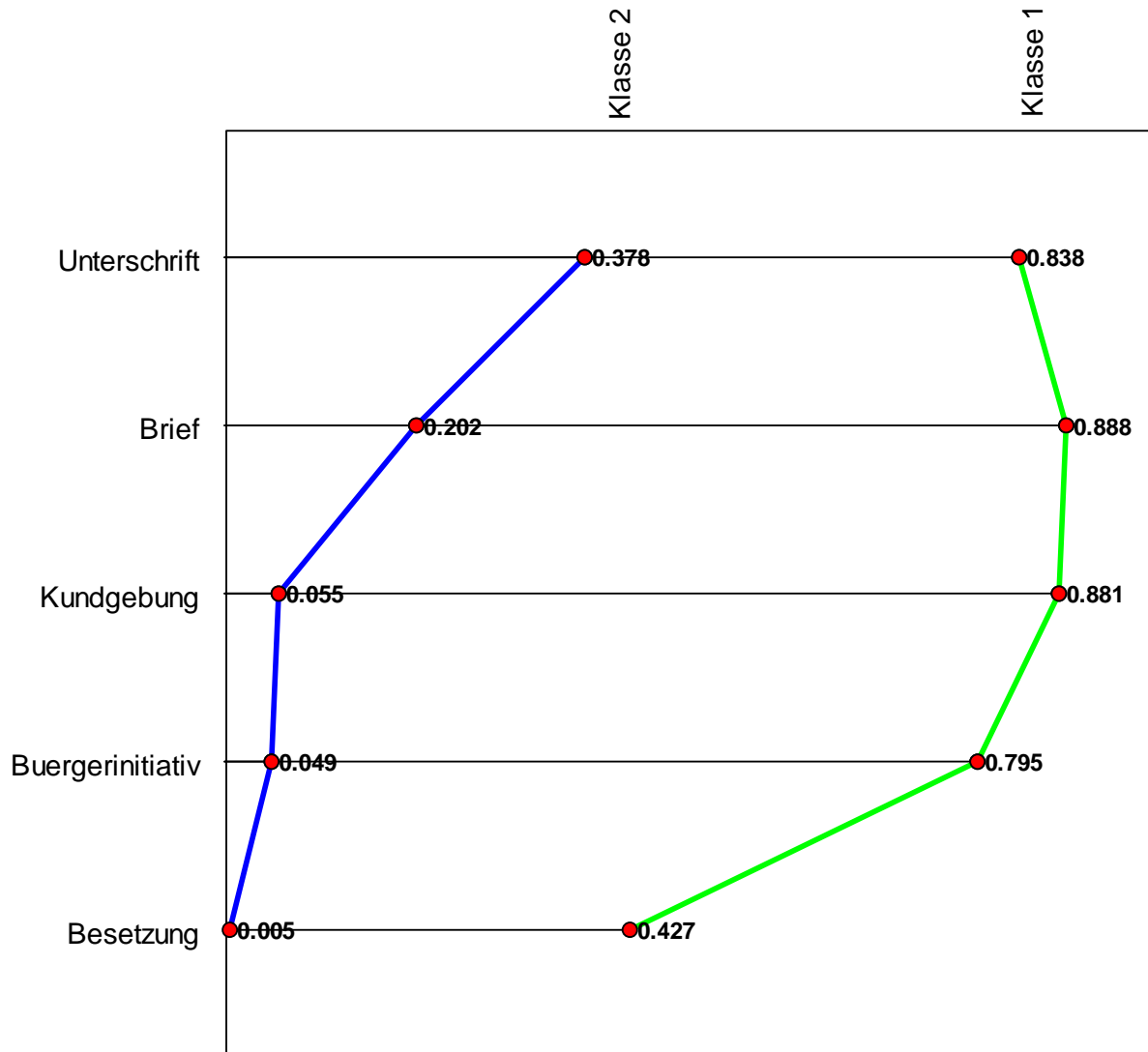
Lc-Parameter *** Loesung *** 7 Iterationen

	1	2
0	1.000	1.000
9 Untersc	0.838	0.378
10 Brief	0.888	0.202
11 Kundgeb	0.881	0.055
12 Buerger	0.795	0.049
13 Besetzu	0.427	0.005
Rel. I	0.502	0.496

Chi-Quadrat = 0.4630 10 Freiheitsgrade
Signifikanz (1-p)*100 = 16.912
Gesamterklaerung = 0.9999

Die Antwortwahrscheinlichkeiten je Variable und Klasse können grafisch dargestellt werden.

Antwortwahrscheinlichkeiten



Almo gibt dann noch die Antwort-Wahrscheinlichkeiten der einzelnen Personen aus. Hier werden nur die ersten 15 Personen gezeigt. Diese Grafik wird von Almo zunächst 3-dimensional mit umgebenden Wänden ausgegeben. Das ist wenig anschaulich. Wie oben bei der Grafik des Modells 1 sollte im Grafik-Editor eine Änderung vorgenommen werden.

Personenparameter

Variable					latente Klassen	
9	10	11	12	13	1	2
0	0	0	0	0	0.001	0.999*
0	0	0	0	1	0.079	0.921*
0	0	0	1	0	0.041	0.959*
0	0	0	1	1	0.864*	0.136
0	0	1	0	0	0.068	0.932*
0	0	1	0	1	0.915*	0.085
0	0	1	1	0	0.845*	0.155
0	0	1	1	1	0.999*	0.001
0	1	0	0	0	0.018	0.982*
0	1	0	0	1	0.728*	0.272
0	1	0	1	0	0.575*	0.425
0	1	0	1	1	0.995*	0.005
0	1	1	0	0	0.697*	0.303

0	1	1	0	1	0.997*	0.003
0	1	1	1	0	0.994*	0.006

Literatur

Coombs, C.H.: A theory of data, New York-London 1967

Denz, H.: Analyse latenter Strukturen - ein sozialwissenschaftliches Meßmodell, Francke, UTB 1198, München 1982

Lazarsfeld, P.F.: A conceptual introduction to latent structure analysis, in: Lazarsfeld, P.F.(ed.): Mathematical thinking in the social sciences, New York 1954

Rasch, G.: An individualistic approach to item analysis, in: Lazarsfeld, P.F. Henry, N.W. (eds.): Reading in mathematical social sciences, Chicago 1966

W. S. Torgerson: Theory and Method of Scaling, John Wiley, 1960, Kap. 13

Tucker, L.R.: A level of proficiency scale for a unidimensional skill, Amer. Psychologist, 7, 408