



Conjoint-Analyse

Analyse verbundener Messungen

Johann Bacher

Almo Statistik-System

www.almo-statistik.de

johann.bacher@jku.at

2013

Der vorliegende Text ist eine leicht überarbeitete und gekürzte Fassung des Kapitels P44 Conjoint-Analyse aus dem Almo-Handbuchs „Teil4: Fortgeschrittene Verfahren“

Nachfolgend wird häufig auf das Dokument **P0** Bezug genommen. Dabei handelt es sich um das Almo-Dokument "Arbeiten mit Almo.PDF". Es kann in Almo heruntergeladen werden

Weitere Almo-Dokumente

Die folgenden Dokumente können alle 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 ordinale 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)
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)
22. Conjoint-Analyse (0,8 MB)
23. Ausreisser entdecken (170 KB)

Inhaltsverzeichnis

P44 Conjoint-Analyse (Analyse verbundener Messungen)	4
<i>P44.1 Kurzbeschreibung des Verfahrens</i>	4
<i>P44.2 Beispiel für ein unvollständiges Design</i>	17
<i>P44.3 Programm-Maske Prog44m2 (unvollständiges Design)</i>	19
P44.3.1 Erläuterungen zu den Eingabe-Boxen	25
P44.3.1.1 Ideal- oder Anti-Idealpunkt.....	28
P44.3.1.2 Schreiben der Stimuluskonfigurationen bei unvollständigen Designs	32
<i>P44.4 Programm-Maske Prog44ma (vollständiges Design)</i>	34
<i>P44.5 Ergebnis aus Almo</i>	34
<i>P44.6 Exkurs: Erstellen unvollständiger Designs</i>	53
<i>Literatur</i>	56

P44 Conjoint-Analyse (Analyse verbundener Messungen)

P44.1 Kurzbeschreibung des Verfahrens

Ziel der verbundenen Messung ("conjoint analysis") ist die Ermittlung individueller Nutzenstrukturen (Nutzenfunktionen) aus Präferenzurteilen.

Dabei wird von folgenden Annahmen ausgegangen:

- Jeder Stimulus s ist durch bestimmte Merkmalsausprägungen A_i, B_j, C_k, \dots in den Merkmalen A, B, C, \dots gekennzeichnet.
- Die Merkmalsausprägungen A_i, B_j, C_k, \dots besitzen für jede Person g einen unterschiedlichen Nutzen, d.h. sie werden in einem unterschiedlichen Ausmaß gewünscht oder abgelehnt.
- Die Präferenz einer Person g für einen Stimulus s hängt von diesen individuellen Nutzenwerten ab. Bezeichnen wir die individuellen Nutzenwerte einer Person g mit $a_{ig}, b_{jg}, c_{kg}, \dots$ und die Präferenz der Person g für den Stimulus i mit y_{sg} , so wird bei der verbundenen Messung angenommen, daß sich die Präferenz als Summe der individuellen Nutzenwerte zusammensetzt. Formal ausgedrückt:

$$y_{sg} = y_{0g} + a_{ig} + b_{jg} + c_{kg} + \dots + e_{sg}$$

mit

y_{sg} = Präferenz der Person g für den Stimulus s

y_{0g} = konstanter Nutzenwert der Person g

a_{ig} = Nutzenwert der Person g für die Merkmalsausprägung i von A in s

b_{jg} = Nutzenwert der Person g für die Merkmalsausprägung j von B in s

c_{kg} = Nutzenwert der Person g für die Merkmalsausprägung k von C in s

e_{sg} = Fehlerterm der Person g für den Stimulus s

Betrachten wir zur Verdeutlichung ein Beispiel. Zur Leistungsbeurteilung in Seminaren können u. a. drei Kriterien (Merkmale) verwendet werden:

Merkmal A = fortlaufende Gruppenarbeiten mit den Ausprägungen $A_1 = \text{ja}$ und $A_2 = \text{nein}$

Merkmal B = Einzelreferat mit den Ausprägungen $B_1 = \text{ja}$ und $B_2 = \text{nein}$

Merkmal C = Klausur mit den Ausprägungen $C_1 = \text{Klausur am Ende des Seminars,}$

$C_2 = \text{zwei Teilklausuren}$

Eine Beurteilungsform (ein Stimulus) s könnte somit wie folgt aussehen:

$s = \text{Gruppenarbeiten (A}_1\text{) mit zwei Teilklausuren (C}_2\text{), aber kein Einzelreferat (B}_2\text{).}$
--

Eine Person g könnte den einzelnen Beurteilungsformen folgende Lerneffektivität (Lernerfolg) zuschreiben:

konstanter Nutzenwert	$y_{0g} = 0$	
Nutzenwerte in Bezug auf A	$a_{1g} = 1$	$a_{2g} = 0$
Nutzenwerte in Bezug auf B	$b_{1g} = 0.5$	$b_{2g} = 0$
Nutzenwerte in Bezug auf C	$c_{1g} = 0$	$c_{2g} = 0$

wobei "0" "kein Nutzen" (kein Lernerfolg) und "1" "maximaler Nutzen" (maximaler Lernerfolg) bedeuten. Die Person g schreibt also allgemein einer Leistungsbeurteilung keinen Nutzen (im Sinne eines Lernerfolges) zu. (Der konstante Nutzenwert ist gleich Null) Den größten Lernerfolg schreibt sie fortlaufenden Gruppenarbeiten zu ($a_{1g}=1.0$), einen mittleren Lernerfolg einem Einzelreferat ($b_{1g}=0.5$). Einer schriftlichen Leistungsüberprüfung in Form einer Abschlußklausur oder in Form von Teilklausuren schreibt sie keinen Nutzen zu. Unter Vernachlässigung des Fehlerterms e_{sg} ergibt sich folgende Präferenz für die Beurteilungsform s ($= A_1, B_2, C_2$).

$$\begin{aligned}
 y_{sg} &= y_{0g} + a_{1g} + b_{2g} + c_{2g} \\
 &= 0 + 1 + 0 + 0 \\
 &= 1
 \end{aligned}
 \quad (s = A_1, B_2, C_2)$$

Im Unterschied dazu, würde sich für eine Beurteilungsform, in der nur ein Einzelreferat und eine Abschlußklausur eingeht, folgende Präferenz ergeben:

$$\begin{aligned}
 y_{s'g} &= y_g + a_{2g} + b_{1g} + c_{1g} \\
 &= 0 + 0 + 0.5 + 0 \\
 &= 0.5
 \end{aligned}
 \quad (s' = A_2, B_1, C_1)$$

Für eine Beurteilungsform auf der Basis fortlaufender Gruppenarbeiten, eines Einzelreferats und von Teilklausuren würde sich folgender Nutzenwert ergeben:

$$\begin{aligned}
 y_{s''g} &= y_g + a_{1g} + b_{1g} + c_{2g} \\
 &= 0 + 1 + 0.5 + 0 \\
 &= 1.5
 \end{aligned}
 \quad (s'' = A_1, B_1, C_2)$$

Bei der verbundenen Messung werden nun aber nicht die individuellen Nutzenwerte $a_{ig}, b_{jg}, c_{kg}, ..$ einer Person g für die einzelnen Merkmalsausprägungen erhoben, sondern die Präferenzurteile y_{sg} . Aufgabe der verbundenen Messung ist es, aus den Präferenzurteilen unter der Annahme einer linearen Nutzenfunktion die individuellen Nutzenwerte zu berechnen.

In unserem Beispiel gibt es insgesamt 8 Stimuluskonfigurationen, nämlich:

	Stimulus- konfigurationen	verbale Beschreibung
1	A ₁ B ₁ C ₁	fortlaufende Gruppenarbeiten, Einzelreferat, Abschlußklausur
2	A ₁ B ₁ C ₂	fortlaufende Gruppenarbeiten, Einzelreferat, Teilklausuren
3	A ₁ B ₂ C ₁	fortlaufende Gruppenarbeiten, kein Einzelreferat, Abschlußklausur
4	A ₁ B ₂ C ₂	fortlaufende Gruppenarbeit, kein Einzelreferat, Teilklausuren
5	A ₂ B ₁ C ₁	keine fortlaufende Gruppenarbeit, Einzelreferat, Abschlußklausur
6	A ₂ B ₁ C ₂	keine fortlaufende Gruppenarbeit, Einzelreferat, Teilklausuren
7	A ₂ B ₂ C ₁	keine fortlaufende Gruppenarbeit, kein Einzelreferat, Abschlußklausur
8	A ₂ B ₂ C ₂	keine fortlaufende Gruppenarbeit, kein Einzelreferat, Teilklausuren

Den Befragten können die acht Stimuluskonfigurationen zur Beurteilung vorgelegt werden als:

- Rating- bzw. Likertskala. (Jede Stimuluskonfiguration wird unabhängig von den anderen beurteilt.)
- Rangordnung. (Die Stimuluskonfigurationen werden in eine Rangreihe gebracht.)

Nachfolgend sind beide Möglichkeiten dargestellt.

Erhebung von Präferenzurteilen durch Rating- bzw. Likertskala:

"Leistungsbeurteilungen dienen der Überprüfung des Lernerfolgs. Nachfolgend werden acht unterschiedliche Formen von Leistungsbeurteilung angeführt. Geben Sie bitte bei jeder Beurteilungsform an, ob sie aus Ihrer Sicht einen sehr großen, großen, mittleren, geringen oder sehr geringen Beitrag zum Lernerfolg leistet.

Beurteilungsform 1:	<ul style="list-style-type: none"> • fortlaufende Gruppenarbeiten • Einzelreferat • Abschlußklausur
---------------------	--

<u>leistet...</u>	sehr großen	
(zutreffendes ankreuzen)	großen	
	mittleren	
	geringen	
	sehr geringen	... <u>Beitrag zum Lernerfolg.</u>

Beurteilungsform 2:	<ul style="list-style-type: none"> • fortlaufende Gruppenarbeiten • Einzelreferat • Teilklausuren
---------------------	--

<u>leistet...</u>	sehr großen	
	großen	
	mittleren	
	geringen	
	sehr geringen	... <u>Beitrag zum Lernerfolg.</u>

Usw.

Erhebung von Präferenzurteilen durch Rangordnung:

"Leistungsbeurteilungen dienen der Überprüfung des Lernerfolgs. Nachfolgend werden acht unterschiedliche Formen von Leistungsbeurteilung angeführt. Bringen Sie diese bitte in eine Rangreihe. Geben Sie jener Leistungsbeurteilung mit dem größten Lernerfolg die Rangziffer 1, jener mit dem zweitgrößten Lernerfolg die Rangziffer 2 usw.

		Rangziffern von 1 bis 8 vergeben (1=größter Lernerfolg, 8=geringster Lernerfolg)
Beurteilungsform 1:	<ul style="list-style-type: none">• fortlaufende Gruppenarbeiten• Einzelreferat• Abschlußklausur	
Beurteilungsform 2:	<ul style="list-style-type: none">• fortlaufende Gruppenarbeiten• Einzelreferat• Teilklausuren	
usw.	<ul style="list-style-type: none">• usw.	

In Backhaus et al. (1994: 530) wird die Erfassung von Präferenzurteilen durch eine Rating- bzw. Likertskala als Präferenzwertmethode bezeichnet. Backhaus et al. führen neben der Methode des Rangordnens und der Präferenzwertmethode ein weiteres Verfahren der Datengewinnung an: die Methode der Rangverteilung. Diese unterscheidet sich von der Methode des Rangordnens dadurch, daß jeder Stimuluskombination unabhängig von den anderen Konfigurationen ein Rangwert gegeben wird. Im Unterschied zur Verwendung einer Likert- bzw. Ratingskala werden keine Antwortkategorien vorgegeben. Neben diesen drei Verfahren können noch unvollständige Rangordnungsverfahren oder Auswahlverfahren verwendet werden (Green 1982: 450:451). Bei Auswahlverfahren werden die Befragten gebeten, die Stimuluskonfigurationen in eine bestimmte Anzahl von Gruppen zu teilen, z.B. in die Gruppe der Stimuluskonfigurationen mit einer sehr starker Präferenz, mit einer starken Präferenz usw. Bei unvollständigen Rangordnungsverfahren werden die Befragten ersucht eine bestimmte Anzahl von sehr stark präferierten Stimuluskonfigurationen zu nennen und diese in eine Rangordnung zu bringen. Unvollständige Rangordnungsverfahren stellen also eine Kombination von Auswahlverfahren und Rangordnungsverfahren dar. Nachfolgende Abbildung faßt die unterschiedlichen Verfahren zusammen.

Methode	Vorgehen	Meßniveau
Rating- bzw. Likertverfahren (Präferenzwertmethode)	Jede Stimuluskonfiguration wird mit Hilfe einer vorgegebenen Antwortskala bewertet.	strenggenommen ordinal, werden i.d.R. als quantitativ betrachtet.
Methode des Rangordnens	Die Stimuluskonfigurationen werden in eine Rangreihe gebracht.	ordinal, echte Rangreihe ohne Bindung
Methode der Rangverteilung	Jeder Stimuluskonfiguration wird ein Rangwert vergeben. Im Unterschied zum Likert- bzw. Ratingverfahren sind die Rangwerte inhaltlich (semantisch) nicht definiert.	ordinal, keine echten Rangreihen, da Bindungen (gleiche Rangplätze) auftreten können, Behandlung als "quantitative" Variablen ist problematisch, da die Rangwerte inhaltlich nicht definiert sind.
Auswahlverfahren	Die Stimuluskonfigurationen werden zu Gruppen zusammengefaßt. Die Gruppen können dabei den Antwortkategorien der Likert- bzw. Ratingskala entsprechen.	Strenggenommen ordinal, werden i.d.R. als quantitativ betrachtet. Ausprägungen sind inhaltlich definiert.
unvollständige Rangordnung	Nur ein Teil der Stimuluskonfigurationen wird in eine Rangreihe gebracht.	ordinal, keine echten Rangreihen, da i.d.R. Bindungen (gleiche Rangplätze) auftreten

Allgemein empfehlen wir die Anwendung des Rating-/Likertverfahrens, da durch diese beiden Methoden "quantitative" Präferenzurteile erfaßt werden können.

Die fiktiven Antworten von zehn Befragten bei Anwendung des Rating- bzw. Likertverfahrens könnten für unser Beispiel wie folgt aussehen:

Person	Stimuluskonfiguration							
	1	2	3	4	5	6	7	8
V1	V2	V3	V4	V5	V6	V7	V8	V9
1	1	1	2	2	4	4	5	5
2	2	2	2	2	4	4	4	4
3	1	1	1	1	5	5	5	5
4	3	3	3	3	5	5	5	5
5	2	2	2	2	3	3	3	3
6	1	1	3	3	2	2	4	5
7	5	5	5	5	5	5	5	5
8	1	1	2	2	3	3	4	4
9	2	2	3	3	4	4	5	5
10	2	2	4	3	1	2	4	4

Die Kodierungen für die Stimuluskonfigurationen bedeuten: "1" = sehr großer Beitrag bis "5" = sehr geringer Beitrag.

Aufgrund der Antworten der Personen werden bei der verbundenen Messung für jede Person g die individuellen Nutzenwerte mit Hilfe der Methode der Kleinsten-Quadrate geschätzt. Die Schätzfunktionen für eine Person g für unser Beispiel lautet

$$KQ(g) = \sum_s (y_{sg} - \tilde{y}_{sg})^2 \rightarrow \text{Minimum}$$

mit

$$\tilde{y}_{sg} = y_{0g} + a_{ig} + b_{jg} + c_{kg} \quad \text{für} \quad s = A_i, B_j, C_k$$

\tilde{y}_{sg} ist der aufgrund der berechneten Nutzenfunktion der Person g prognostizierte Wert für die Stimuluskonfiguration s .

Bei der Schätzung der Nutzwerte einer Person g wird in ALMO nach der Logik des allgemeinen linearen Modells (siehe Abschnitte P20.6 und P20.7 des Handbuchs) vorgegangen. D.h.:

1. Die unabhängigen nominalen Variablen (Stimulusmerkmale) werden in Dummies aufgelöst, indem die 1/0-Dummy-Kodierung (Methode der "fitting constants") oder die -1/0/1-Dummy-Kodierung (Methode der "weighted means of squares") verwendet wird.
2. Die letzte Dummy-Variable wird gestrichen.
3. Die Dummies werden wie quantitative Variablen behandelt und ihre Regressionskoeffizienten werden mit Hilfe der Methode der Kleinsten-Quadrate berechnet.
4. Aus den Regressionskoeffizienten werden die Effekte berechnet. Diese werden in der verbundenen Messung als Teilnutzenwerte, aber auch als Effekte bezeichnet.

In unserem Beispiel ergeben sich folgende Teilnutzenwerte für die Person 1, wobei zur leichteren Interpretation die Kodierung der Beurteilungen der Stimuluskonfigurationen umgedreht wurde, so daß ein größerer Zahlenwert eine stärkere Präferenz ausdrückt.

Effekte (Teilnutzenwerte):

Variable	Effekt (Teilnutzenwert)
Konstante	3.000
V11 Gruppenarbei	
1 = ja	1.500
2 = nein	-1.500
V12 Einzelrefera	
1 = ja	0.500
2 = nein	-0.500
V13 Klausur	
1 = Abschlusskla	0.000
2 = Teilklausure	0.000

Interpretation: Im Durchschnitt schreibt die erste Person der Leistungsbeurteilung einen mittleren Lernerfolg zu (Konstante=3). Von den erfaßten Beurteilungskriterien wird der Gruppenarbeit der größte Nutzen zugeschrieben (Effekt=1.5), gefolgt von der Erstellung eines Einzelreferats (Effekt=0.5). Einer Klausur wird kein Lernbeitrag zugeschrieben. Beide Effekte sind gleich Null.

Zur Überprüfung, wie gut die für eine Person berechnete lineare Nutzenfunktion den Präferenzurteilen dieser Person angepaßt ist, werden bei der verbundenen Messung der Pearsonsche Korrelationskoeffizient und der Kendallsche Korrelationskoeffizient τ_b zwischen den empirischen und den aufgrund der berechneten Nutzenfunktion prognostizierten Werte berechnet. Bei einer perfekten Übereinstimmung zwischen berechneten (prognostizierten) Nutzenwerten und empirischen Nutzenwerte müssen beiden Koeffizienten den Wert 1.0 annehmen. Bei Kendalls τ_b ist dies nur der Fall, wenn die empirischen Daten keine Bindungen auftreten.

In unserem Beispiel ist der Pearsonsche Korrelationskoeffizient für die Person 1 gleich 1.00. Tatsächlich stimmen die prognostizierten und prognostizierten Werte auch perfekt überein, wie nachfolgende Ausgabe zeigt:

Gegenueberstellung von empirischen und prognostisierten Werten:

Stim.-Konfig. = 1	y = 5.000	y(prognostiziert) = 5.000
Stim.-Konfig. = 2	y = 5.000	y(prognostiziert) = 5.000
Stim.-Konfig. = 3	y = 4.000	y(prognostiziert) = 4.000
Stim.-Konfig. = 4	y = 4.000	y(prognostiziert) = 4.000
Stim.-Konfig. = 5	y = 2.000	y(prognostiziert) = 2.000
Stim.-Konfig. = 6	y = 2.000	y(prognostiziert) = 2.000
Stim.-Konfig. = 7	y = 1.000	y(prognostiziert) = 1.000
Stim.-Konfig. = 8	y = 1.000	y(prognostiziert) = 1.000

Berechnung des Pearsonschen Korrelationskoeffizienten R

$$R = \frac{\text{Kovarianz zwischen empirischen und prognostizierten Werten}}{\text{Standardabweichung der empirischen Werte} * \text{Standardabweichung der prognostizierten Werte}}$$

$$R = \frac{2.500}{1.581 * 1.581} = 1.000$$

Signifikanz von R

$$F = \frac{R^{*2} / (k - 1)}{(1 - R^{*2}) / (n - k)}$$

$$F = \frac{1.000 / 2.000}{(1 - 1.000) / 5.000} = KW$$

$$\text{Signifikanz } 100*(1-p) = KW$$

Berechnung von tau-b:

$$\text{tau-b} = \frac{Ns - Nd}{Sx * Sy} = \frac{Ns - Nd}{\text{Wurzel}(Ns + Nd + Tx) * \text{Wurzel}(Ns + Nd + Ty)}$$

Ns = Zahl gleichgerichteter Beziehungen in empirischen u. progn. Werten

Nd = Zahl entgegengerichteter Beziehungen in empirischen u. progn. Werten

Tx = Bindungen in empirischen Werten

Ty = Bindungen in prognostizierten Werten

$$\text{tau-b} = \frac{24.000 - 0.000}{5.292 * 5.292} = 0.857$$

$$Ns = 24.000$$

$$Nd = 0.000$$

$$Tx = 4.000$$

$$Ty = 4.000$$

Signifikanz von tau-b

$$F = \frac{(ns - nd)**2}{\text{vs (=Varianz von F)}}$$

$$F = \frac{(24.000 - 0.000)**2}{57.905} = 9.947$$

$$\text{Signifikanz } 100*(1-p) = 99.823$$

Kendalls tau_b beträgt in dem Beispiel dagegen nur 0.67, da Bindungen in den empirischen Werten vorliegen.

Allgemein empfehlen wir - wegen des Problems der Bindungen - Kendalls tau_b nur dann zu verwenden, wenn echte Rangreihen ohne Bindungen vorliegen, wenn also die Methode des Rangordnens zur Datenerhebung verwendet wird. In diesem Fall (bei echten Rangdaten) kann Kendalls tau_b der "Ordinalität" der Daten besser Rechnung tragen.

Schwellenwerte für beide Korrelationskoeffizienten, ab denen von einer guten Modellanpassung gesprochen werden kann, fehlen. Allerdings ist es möglich zu überprüfen, ob die Korrelationskoeffizienten signifikant von Null verschieden sind. Zu beachten ist allerdings, daß die Tests nur auf einer kleinen Fallzahl basieren, die gleich der Zahl der erhobenen Stimuluskonfigurationen ist.

In dem Beispiel kann für den Pearsonschen Korrelationskoeffizient keine Signifikanz berechnet werden, da er exakt 1.0 ist. Ein Wert von 1.00 bedeutet aber eine perfekte Übereinstimmung und kann immer als signifikant bezeichnet werden. Kendalls tau_b ist - wenn von einem Schwellenwert von 95 % ausgegangen wird - ebenfalls signifikant von Null verschieden.

Für alle zehn Befragten ergeben sich folgende Teilnutzenwerte (Effekte) und Korrelationskoeffizienten:

Matrix der Effekte
=====

DA	0	1	2	3	4	5	6	R	TAU
1	3.00	1.50	-1.50	0.50	-0.50	0.00	0.00	1.00	0.86
2	3.00	1.00	-1.00	0.00	0.00	0.00	0.00	1.00	0.57
3	3.00	2.00	-2.00	0.00	0.00	0.00	0.00	1.00	0.57
4	2.00	1.00	-1.00	0.00	0.00	0.00	0.00	1.00	0.57
5	3.50	0.50	-0.50	0.00	0.00	0.00	0.00	1.00	0.57
6	3.38	0.62	-0.62	1.12	-1.12	0.12	-0.12	0.98	0.94
7	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	3.50	1.00	-1.00	0.50	-0.50	0.00	0.00	1.00	0.86
9	2.50	1.00	-1.00	0.50	-0.50	0.00	0.00	1.00	0.86
10	3.25	0.00	0.00	1.00	-1.00	0.00	0.00	0.92	0.64
MW	2.81	0.86	-0.86	0.36	-0.36	0.01	-0.01	0.89	0.64
SA	0.75	0.59	0.59	0.41	0.41	0.04	0.04	0.30	0.26
n	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00

Die Spalten bedeuten:

- 0 = Konstante
- 1 = V11 (=Gruppenarbeit), Auspr. = 1 (=ja)
- 2 = V11 (=Gruppenarbeit), Auspr. = 2 (=nein)
- 3 = V12 (=Einzelreferat), Auspr. = 1 (=ja)
- 4 = V12 (=Einzelreferat), Auspr. = 2 (=nein)
- 5 = V13 (=Klausur), Auspr. = 1 (=Abschlussklausur)
- 6 = V13 (=Klausur), Auspr. = 2 (=Teilklausuren)
- R = Pearsonscher Korrelationskoeffizient
- TAU = Kendalls tau-b

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA

Interpretation: Mit Ausnahme der Person 7 und 10 schreiben alle anderen Personen einer Gruppenarbeit einen positiven Nutzen (Lernerfolg) zu. Für die Personen 7 und 10 haben Gruppenarbeiten weder einen positiven noch einen negativen Lernerfolg, da die Effekte gleich 0 sind. Einzelreferaten schreiben fünf Personen (Person 1, 6, 8, 9 und 10) einen positiven Lernerfolg zu, alle anderen keinen Lernerfolg. Klausuren schließlich schreibt nur eine Person (Person 6) einen Lernerfolg zu, wobei eine Abschlußklausur einen positiven Lernerfolg besitzt, Teilklausuren dagegen einen negativen.

Im Durchschnitt (siehe Zeile MW) haben Gruppenarbeiten den stärksten positiven Effekt (=0.86), Einzelreferate den zweit stärksten (=0.36). Die durchschnittlichen Effekte von Abschlußklausuren und von Teilklausuren liegen nahe bei Null (0.01 bzw. -0.01). Sie leisten im Durchschnitt also weder einen positiven noch einen negativen Beitrag. Umgekehrt reduziert das Fehlen von Gruppenarbeiten die Lernerfolgseinschätzung am stärksten (= -0.86), das Fehlen eines Einzelreferats am zweit stärksten (= -0.36).

Der Pearsonsche Korrelationskoeffizient ist für alle Befragten mit Ausnahme der Person 7, bei der alle Effekt gleich 0 sind, größer 0.90, also nahe dem Maximalwert von 1.00. Auf tau_b trifft dies nicht zu. Seine Verwendung ist aber bei Likert- bzw. Ratingdaten wegen des Problems der Bindungen problematisch.

Mitunter ist man insbesondere bei Rangdaten bei der konjunkten Messung aber nicht an den (absoluten) Teilnutzenwerten sondern an **relativen, normierten oder partiellen Teilnutzenwerten** interessiert. Diese geben an, welchen Nutzen eine Merkmalsausprägung im Vergleich zu allen anderen Merkmalsausprägungen besitzt. Dabei wird angenommen:

1. Die normierten Teilnutzenwerte sollen alle zwischen 0 und 1 liegen.
2. Für die Stimuluskonfiguration mit der stärksten Präferenz soll der prognostizierte Gesamtnutzen gleich 1.0 sein.
3. Für die Stimuluskonfiguration mit der geringsten Präferenz soll der prognostizierte Gesamtnutzen gleich 0.0 sein.

Während diese Normierung bei Rangdaten berechtigt und mitunter notwendig ist, führt sie bei Rating- bzw. Likertdaten zu einem Informationsverlust. In unserem Beispiel erhalten u.a. durch die Normierung die Personen 2 bis 4 dieselben relativen Nutzenwerte, obwohl sie sich in ihrem durchschnittlichen Gesamtnutzen und in ihren absoluten Teilnutzenwerten deutlich unterscheiden. Die Person 4 beispielsweise schreibt allgemein der Leistungsbeurteilung nur eine geringe Bedeutung (Konstante=2) zu, Person 4 dagegen eine mittlere bis starke Bedeutung (Konstante=3.5). Bei Person 4 führen fortlaufende Gruppenarbeiten nur zu einer geringfügigen Verbesserung der Lernerfolgseinschätzung (Effekt=0.5), bei der Person 3 ist der Effekt (=2.0) stärker ausgeprägt. Bei Rangdaten dagegen ist die Normierung berechtigt, da Rangplätze für die einzelnen Personen unterschiedliche Bedeutung haben können. Für eine Person kann beispielsweise ein Rangplatz von "1" einen sehr großen Nutzen bedeuten, für eine andere dagegen einen sehr geringen. Dieser Effekt wird durch die dargestellte Normierung beseitigt.

Matrix der normierten Teilnutzenwerte
=====

DA	1	2	3	4	5	6	SUM
1	0.75	0.00	0.25	0.00	0.00	0.00	1.00
2	1.00	0.00	0.00	0.00	0.00	0.00	1.00
3	1.00	0.00	0.00	0.00	0.00	0.00	1.00
4	1.00	0.00	0.00	0.00	0.00	0.00	1.00
5	1.00	0.00	0.00	0.00	0.00	0.00	1.00
6	0.33	0.00	0.60	0.00	0.07	0.00	1.00
7	KW	KW	KW	KW	KW	KW	KW
8	0.67	0.00	0.33	0.00	0.00	0.00	1.00
9	0.67	0.00	0.33	0.00	0.00	0.00	1.00
10	0.00	0.00	1.00	0.00	0.00	0.00	1.00
MW	0.71	0.00	0.28	0.00	0.01	0.00	1.00
SA	0.33	0.00	0.32	0.00	0.02	0.00	0.00
n	9.00	9.00	9.00	9.00	9.00	9.00	9.00

Die Spalten bedeuten:

- 1 = V11 (=Gruppenarbeit), Auspr. = 1 (=ja)
- 2 = V11 (=Gruppenarbeit), Auspr. = 2 (=nein)
- 3 = V12 (=Einzelreferat), Auspr. = 1 (=ja)
- 4 = V12 (=Einzelreferat), Auspr. = 2 (=nein)
- 5 = V13 (=Klausur), Auspr. = 1 (=Abschlussklausur)
- 6 = V13 (=Klausur), Auspr. = 2 (=Teilklausuren)

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA
- KW = kein Wert vorhanden

Interpretation: Bei Person 1 leistet eine Gruppenarbeit - im Vergleich zu den anderen Ausprägungen - den stärksten Beitrag zu einer Erhöhung des Lernerfolgs, ein Einzelreferat den zweit stärksten bei.

Bei den Personen 2 bis 5 leistet ausschließlich eine Gruppenarbeit einen Beitrag. Für Person 6 ergibt sich ein anderes Bild: Ein Einzelreferat leistet den größten Beitrag, gefolgt von Gruppenarbeiten. Die normierten Teilnutzenwerte für die Person 7 sind nicht berechenbar, da alle Effekte (mit Ausnahme der Konstanten) gleich Null sind usw.

Beachte: Nur bei dichotomen Variablen, also bei Variablen mit zwei Ausprägungen, sind die Zeilensummen gleich 1.00.

Aus den normierten Teilnutzenwerten läßt sich in einem weiteren Schritt der Analyse die relative Wichtigkeit der untersuchten Merkmale (Variablen) berechnen. Die relative Wichtigkeit eines Merkmals gibt an, welche Bedeutung diesem Merkmal im Vergleich zu den anderen Merkmalen für die Präferenzbildung einer Person zukommt, wie gut also ein Merkmal die Präferenzurteile einer Person erklärt. In unserem Beispiel ergeben sich folgende Werte:

Matrix der relativen Wichtigkeit der Variablen
 =====

DA	1	2	3	SUM
1	0.75	0.25	0.00	1.00
2	1.00	0.00	0.00	1.00
3	1.00	0.00	0.00	1.00
4	1.00	0.00	0.00	1.00
5	1.00	0.00	0.00	1.00
6	0.33	0.60	0.07	1.00
7	0.00	0.00	0.00	0.00
8	0.67	0.33	0.00	1.00
9	0.67	0.33	0.00	1.00
10	0.00	1.00	0.00	1.00
MW	0.64	0.25	0.01	0.90
SA	0.38	0.32	0.02	0.30
n	10.00	10.00	10.00	10.00

Die Spalten bedeuten:

- 1 = V11 (=Gruppenarbeit)
- 2 = V12 (=Einzelreferat)
- 3 = V13 (=Klausur)

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA

Interpretation: Mit Ausnahme der Personen 6, 7 und 10 leistet die Variable "Gruppenarbeit" mit den Ausprägungen "ja" und "nein" den größten Beitrag zur Erklärung der Lernerfolgseinschätzungen. Die relative Wichtigkeit dieser Variablen variiert dabei zwischen 0.67 (bzw. 67%) und 1.00 (bzw. 100%). Für Person 7 sind die relativen Wichtigkeiten nicht berechenbar. Sie werden in ALMO gleich Null gesetzt.

Wir können somit das Vorgehen bei der verbundenen Messung wie folgt zusammenfassen:

1. Auswahl von zu untersuchenden Stimulusmerkmalen aufgrund theoretischer, empirischer und/oder praktischer Überlegungen und Erfordernisse. Die Stimulusmerkmale können nominal oder quantitativ sein. Für quantitative Stimulusmerkmale können dabei sogenannte Ideal- bzw. Anti-Idealpunktfunktionen definiert werden.

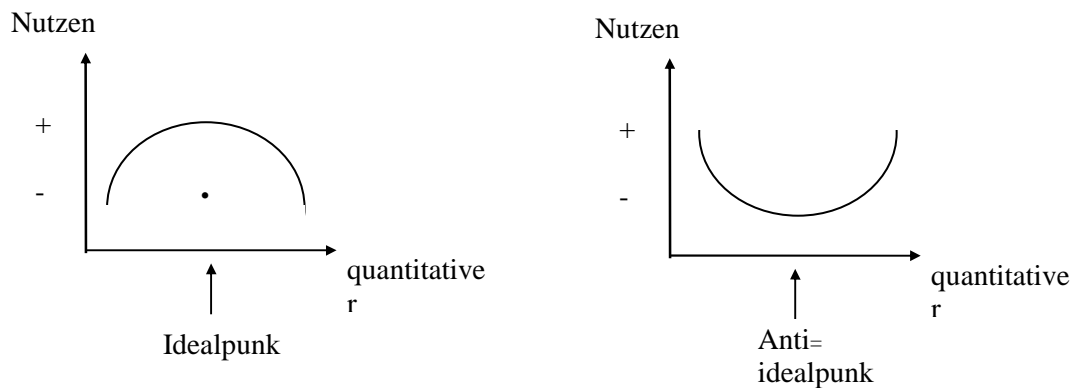
2. Bilden von Stimuluskonfigurationen aus den Stimulusmerkmalen. Dabei bestehen zwei Möglichkeiten:
 - a) Bilden eines vollständigen Designs. Alle Merkmalsausprägungen werden wie in unserem Beispiel systematisch kombiniert. Bei vielen Stimulusmerkmalen bzw. bei Stimulusmerkmalen mit einer großen Ausprägungszahl entsteht dadurch ein beträchtlicher Erhebungsaufwand. Deshalb wird häufig auf die zweite Möglichkeit zurückgegriffen.
 - b) Bilden eines unvollständigen Designs (siehe dazu Abschnitt P44.4). Es wird nur eine Auswahl aller möglichen Merkmalskombinationen zur Beantwortung vorgelegt. Die Auswahl muß dabei bestimmte Bedingungen erfüllen.
3. Erhebung der Präferenzurteile entweder in Rating- bzw. Likertform, in Rangreihenform oder nach einem anderen der dargestellten Verfahren.
4. Berechnung der (absoluten) Teilnutzenwerte für jede Person.
5. Überprüfung der Brauchbarkeit des Modells der konjunkten Messung für jede Person durch Berechnung des Pearsonschen Korrelationskoeffizienten und des Kendallschen Korrelationskoeffizienten taub. Zur Überprüfung der Brauchbarkeit können darüber hinaus noch sogenannte Holdout-Karten definiert werden. Es handelt sich dabei um empirisch erhobene Stimuluskonfigurationen, die aber nicht zur Schätzung verwendet werden.
6. Berechnung der (normierten) Teilnutzenwerte für jede Person, insbesondere bei Rangreihendaten.
7. Berechnung der relativen Wichtigkeit der untersuchten Variablen (Stimulusmerkmale) für jede Person.

ALMO berechnet automatisch Durchschnittswerte für alle Ergebnisgrößen, also z.B. durchschnittliche Effekte. Diese Durchschnittswerte sind nicht dann nicht aussagekräftig, wenn sie hohe Standardabweichungen haben. In diesem Fall kann man beispielsweise durch eine anschließende Clusteranalyse versuchen, homogene Teilgruppen mit ähnlichen Nutzenstrukturen zu bilden.

Die berechneten Effekte können aber auch dazu verwendet werden, um sogenannte Simulationskarten zu untersuchen. Simulationskarten sind Stimuluskonfigurationen, die empirisch nicht erhoben wurden. Für sie soll aufgrund der berechneten Effekte untersucht werden, wie sie im Vergleich zu den erfaßten Konfigurationen beurteilt werden, ob sie also beispielsweise gegenüber allen erhobenen Stimuluskonfigurationen präferiert werden.

Es bestehen somit über das bisherige Beispiel hinausgehend folgende **Möglichkeiten**:

1. Neben nominalen Variablen können auch quantitative unabhängige Variablen (Stimulusmerkmale) erfaßt werden.
2. Für die quantitative Variablen können Ideal- oder Antiidealfunktionen definiert werden. Bei der Idealpunktfunktion wird angenommen, daß es für einen Befragten g auf der quantitativen Variablen einen idealen Punkt mit einem maximalen Nutzen gibt. Umso größer die (quadratische) Abweichung einer untersuchten Merkmalsausprägung von diesem Idealpunkt ist, desto geringer ist der zugeschriebene Nutzen. Bei der Anti-Idealpunktfunktion wird der inverse Zusammenhang angenommen. Der Anti-Idealpunkt ist jene Merkmalsausprägung mit dem geringsten Nutzen. Nachfolgende Abbildung stellt beide Konzepte graphisch dar.



Formal lassen sich beide Funktionen als quadratische Funktion modellieren:

$$y_{sg} = y_{og} + b_{xig} (x_i - x_g)^2 + \dots$$

mit

$$x_g = \text{Ideal} - \text{bzw. Antiidealpunkt}$$

Auflösen der Funktion ergibt folgenden Ausdruck, der zur Schätzung verwendet wird.

$$\begin{aligned} y_{sg} &= y_{og} + b_{xig} (x_i - x_g)^2 + \dots \\ &= y_{og} + b_{xig} (x_i^2 - 2 \cdot x_g \cdot x_i + x_g^2) + \dots \\ &= y_{og} + b_{xig} \cdot x_i^2 - 2 \cdot b_{xig} \cdot x_g \cdot x_i + b_{xig} \cdot x_g^2 + \dots \\ &= (y_{og} + b_{xig} \cdot x_g^2) + (-2 \cdot b_{xig} \cdot x_g) \cdot x_i + b_{xig} \cdot x_i^2 + \dots \\ &= y_{og}^* + c_{xig} \cdot x_g + d_{xig} \cdot x_i^2 \end{aligned}$$

mit

$$y_{og}^* = y_{og} + b_{xig} \cdot x_g^2$$

$$c_{xig} = -2 \cdot b_{xig} \cdot x_g$$

$$d_{xig} = b_{xig}$$

D der Ideal- oder Antiidealpunkt x_g der Person g eine Konstante in allen Stimuluskonfigurationen ist. Es wird also ein Regressionskoeffizient für x_i und dem quadratischen Term x_i^2 berechnet.

Sind die Präferenzwerte in Richtung eines höheren Nutzens kodiert (ein höherer Variablenwert bedeutet eine stärkere Präferenz, einen größeren Nutzen), dann ist beim Vorliegen einer Idealpunktfunktion b_{xig} bzw. d_{xig} negativ, beim Vorliegen einer Antiidealpunktfunktion b_{xig} bzw. d_{xig} positiv.

- Es können auch nicht vollständige (unvollständige) Designs analysiert werden. Bei einem unvollständigen Design werden nicht alle möglichen Stimuluskombinationen zur Beurteilung vorgelegt, sondern nur eine Auswahl.

3. Die erfaßten Beurteilungen können Rangdaten sein.
4. Für die Validitätsprüfung können bestimmte empirisch erhobene Stimuluskonfigurationen als Holdout-Karten definiert werden. Holdout-Karten sind Stimuluskonfigurationen, die zwar empirisch erhoben, aber nicht zur Schätzung der Effekte (Nutzenwert) verwendet werden.
5. Es können die Präferenzen für Simulationskarten berechnet werden. Simulationskarten sind Stimuluskonfigurationen, die empirisch nicht erhoben wurden.
6. Die Ergebnisse können weiter analysiert werden, z.B. durch eine Clusteranalyse.

Um diese zusätzlichen Möglichkeiten zu verdeutlichen soll im folgenden ein weiteres Beispiel betrachtet werden. Wir werden dabei die Eingabe in Almo mit dem Maskenprogramm Prog44m2 vorführen.

P44.2 Beispiel für ein unvollständiges Design

Das Beispiel ist Backhaus et al. (1994: 52) entnommen: In einer Untersuchung soll die relative Wichtigkeit folgender Eigenschaften von Margarineprodukten sowie die Teilnutzenwerte der Ausprägungen durch Eigenschaften bestimmt werden.

Eigenschaft A: Geschmack (1=nach Butter; 2=pflanzlich)

Eigenschaft B: Verwendung (1=Brotaufstrich; 2=Kochen, Backen, Braten; 3 = universell)

Eigenschaft C: Kaloriengehalt (1=kalorienarm; 2=normaler Kaloriengehalt)

Eigenschaft D: Preis (1=2,50 - 3,00 DM; 2=2,00 - 2,99 DM; 3=1,50 - 1,99 DM)

Für die Datenerhebung wurde folgendes unvollständige Design verwendet.

Stimuluskonfiguration	Merkmalskombination
1	A1B3C2D1
2	A2B2C1D1
3	A2B1C2D2
4	A1B1C1D3
5	A1B1C1D1
6	A2B3C1D3
7	A1B2C1D2
8	A1B3C1D2
9	A1B2C2D3
10	A2B2C2D2
11	A1B1C2D1

Es wurden 40 Personen befragt. Diese wurden aufgefordert, die elf Konfigurationen in eine Rangreihe zu bringen. Sie gaben folgende Antworten:

Stimulskonfigurationen

DA	1	2	3	4	5	6	7	8	9	10	11
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
01	03	10	11	01	04	09	08	02	07	06	05
02	10	06	08	01	11	02	03	07	04	05	09
03	08	11	06	03	10	05	07	02	04	01	09
04	09	10	07	02	08	01	06	04	03	05	11
05	04	11	09	01	03	10	05	07	08	06	02
06	09	08	07	03	04	02	05	01	10	06	11
07	08	07	11	03	04	05	02	01	09	06	10
08	09	11	06	01	10	05	08	02	04	03	07
09	07	06	09	08	10	02	03	05	01	04	11
10	08	11	10	01	05	09	03	04	02	06	07
11	11	05	08	03	07	04	02	06	01	09	10
12	08	09	07	03	10	02	05	06	01	04	11
13	06	11	10	04	08	09	05	02	03	01	07
14	06	07	11	01	08	09	04	03	02	05	10
15	08	10	06	03	04	02	07	01	11	05	09
16	03	10	11	02	08	09	06	07	01	04	05
17	03	10	06	09	11	02	07	04	05	01	08
18	10	11	05	02	09	07	01	03	06	04	08
19	09	08	11	02	03	04	05	01	10	06	07
20	10	06	07	01	03	04	05	02	11	08	09
21	10	09	06	02	01	11	07	04	05	08	03
22	03	11	10	05	06	09	08	01	07	02	04
23	09	11	03	01	02	07	10	05	08	06	04
24	03	11	02	06	05	08	10	07	09	04	01
25	10	06	08	01	04	03	05	07	02	11	09
26	07	06	09	08	10	01	03	04	02	05	11
27	08	10	11	03	06	09	02	01	05	04	07
28	11	02	03	04	07	01	06	05	09	08	10
29	03	11	10	05	07	04	08	01	09	02	06
30	10	08	03	02	09	01	07	05	04	06	11
31	09	08	07	06	10	05	02	03	01	04	11
32	07	03	01	08	11	02	10	06	05	04	09
33	05	10	11	06	08	09	02	01	03	04	07
34	10	09	08	01	05	04	03	02	07	06	11
35	07	11	10	01	08	04	06	03	02	05	09
36	03	11	08	04	06	07	09	01	10	02	05
37	11	06	09	01	05	02	04	03	08	07	10
38	04	10	11	06	08	09	03	02	05	10	07
39	03	08	06	04	02	11	10	09	05	07	01
40	10	06	09	01	02	05	04	03	08	11	07

Eine kleinere Zahl in der Datenmatrix drückt eine stärkere Präferenz aus. Person 1 präferiert also am stärksten die Stimuluskonfiguration 4 (V5=1), am zweit stärksten die Stimuluskonfiguration 8 (V9=2) usw.

Für die Analyse sollen folgende Spezifikationen getroffen werden:

1. Die Kodierung der Präferenzurteile soll umgedreht werden, so daß ein höherer Zahlenwert eine stärkere Präferenz ausdrückt.
2. Die Variablen "Kaloriengehalt" und "Preis" sollen als quantitative Variable behandelt werden. Die Kategorien der Variablen Preis sollen dazu quantifiziert werden, die Ausprägung 1 soll also den Wert 2,75 DM erhalten, die Ausprägung 2 den Wert 2,25 DM und die Ausprägung 3 den Wert 1,75 DM.

3. Für den Preis soll eine Ideal- bzw. Antiidealfunktion berechnet werden.
4. Die drei letzten Karten sollen zur Validitätsprüfung als Holdout-Karten definiert werden.
5. Zusätzlich soll der Präferenzwert für folgende zwei Simulationskarten bestimmt werden: A2B3C2D3 und A1B2C1D1.
6. Die relativen Teilnutzenwerte sollen für eine anschließende Clusteranalyse zwischengespeichert werden.

P44.3 Eingabe mit Maskenprogramm Prog44m2 (unvollständiges Design)

Klicken Sie auf den Knopf „Verfahren“. Also präsentiert dann eine alphabetisch geordnete Liste aller vorhandenen Verfahren. Klicken Sie in dieser auf „Conjoint-Analyse“. Also präsentiert dann eine Auswahl von 2 Maskenprogrammen:

Prog44m2: Conjoint-Analyse mit unvollständigem Design
Prog44ma: Conjoint-Analyse mit vollständigem Design

Prog44ma wird kurz in Abschnitt P44.5 behandelt

Klicken Sie auf das 1. Programm. Sie sehen dann folgendes:

Prog44m2.Msk
Verbundene Messungen (Conjoint Analysis)
mit unvollständigem Design
mit verschiedenen Optionen

Ziel der verbundenen Messung ist die Ermittlung der relativen Wichtigkeit von Merkmalen und der Nutzenwerte von Merkmalsausprägungen für Präferenzurteile.

Im nachfolgenden Programm soll untersucht werden, ob die Präferenz für eine bestimmte Magarinesorte stärker vom Geschmack, von der Verwendung, vom Kaloriengehalt oder vom Preis abhängt und ob beispielsweise Magarinesorten mit Buttergeschmack gegenüber solchen mit pflanzlichem Geschmack bevorzugt werden.

In nachfolgendem Programm werden noch "holdout"-Karten und Simulationskarten miteinbezogen. Außerdem wird für eine Variable eine Ideal- oder Antiidealfunktion definiert

Beispiel und Daten entnommen aus:
 Backhaus, K., Erichson, B., Plinke, W., Weiber, R.:
 Multivariate Analysemethoden. Berlin-Heidelberg-New York,
 1994, S.498-554

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

Vereinbare Variable = ;

1

die Stimuluskonfigurationen müssen in eine Tabelle geschrieben werden. Siehe Programmende

Tabelle_A = , # Zeilen der Tabelle #
 ; # Spalten der Tabelle #

2

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

zeige = Namensdatei in Output zeigen
 leer = nicht

4

Freie Namensfelder **Hilfe**

Namen für die Merkmale, die zur Bildung der Stimulus-
konfigurationen verwendet werden (=optional)

BEACHT: Als Nummern für diese Merkmale
muss eine freie, sonst nicht verwendete
Variablennummer verwendet werden

Name21=Geschmack:Butter,pflanzlich;
 Name22=Verwendung:Brotaufstrich,Kochen/Backen/Braten,universell;
 Name23=Kaloriengehalt:kalorienarm,normal;
 Name24=Preis:2.50-3.00 DM,2.00-2.49 DM,1.50-1.99 DM;

erzeuge zusätzliche Namensfelder

5

Datei aus der gelesen wird **Hilfe**

"C:\Almo7\Testdat\Conj.fre" bei Datei-Problemen

frei Format der Daten **Hilfe**

U1:12 der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_U

6 Wenn Dateiformat FIX oder Nicht-Standard-FREI

7 U2:12
 die abhängigen Variablen
 Präferenzurteile der Stimuluskonfigurationen die aus der Datei eingelesen werden

8 Geschmack, Verwendung
 2, 3
 die unabhängigen nominalen Variablen
 Zahl ihrer Ausprägungen

9 Kaloriengehalt, Preis
 2, 3
 die unabhängigen quantitativen Variablen
 Zahl ihrer Ausprägungen

10 Option: Ein- und Ausschliessen von Untersuchungseinheiten

11 Loesche wieder diese Box
 Umkodierungen und Kein-Wert-Angaben
 Umkodierungen
 Kein_Wert-Angabe
 die Präferenzurteile werden in ihrer Kodierungsrichtung umgedreht
 U2:12 (1 bis 11 = Umdrehen)

 erzeuge zusätzliche Felder für Umkodierungen / Kein_Wert-Angaben
 Kontrollieren, ob Umkodierung so erfolgt wie gewünscht
 diese Variablen ...
 ... aus diesen Datensätzen vor und nach der Umkodierung zur Kontrolle anzeigen

12 w_squares_of_means Verfahren = w_squares_of_means oder = fitting_constants
 13 Zahl der Stimuluskonfigurationen (<=Zahl der Zeilen der Tabelle_A)
 5 Zahl der Spalten der Tabelle_A der Stimuluskonfigurationen

13

Optionen

Preis

Für eine oder mehrere der unabhängigen quantitativen Variablen kann eine Ideal- oder Antiidealfunktion definiert werden

1

1 = Individualwerte ausgeben
0 = nicht

1

1 = Darstellung des Kalküls für 1. Person
0 = nicht

14

Speichern der Nutzenwerte
(für eine anschließende Clusteranalyse)

"C:\Almo7\Progs\Nutz.fre"

Wenn Sie hier keinen Dateinamen angeben, dann wird nicht gespeichert

1

0 = absolute Teilnutzenwerte speichern
1 = relative Teilnutzenwerte speichern
2 = relative Wichtigkeiten speichern

15

Grafik-Optionen

Almo

Almo = Almo-Grafik ausgeben
0 = nicht

16

Schreiben Sie hier dahinter die Tabelle der Stimuluskonfigurationen

Die Spalten bedeuten in unserem Beispiel

U21						(=Geschmack)
U22						(=Verwendung)
U23						(=Kaloriengehalt)
U24						(=Preis)

Status der Stimuluskonfiguration

	0	= "normale" Karte
	1	= Holdout-Karte
	2	= Simulations-Karte

1	3	2	2.75	0	---	In jeder Zeile steht eine Stimulus-
2	2	1	2.75	0		konfiguration. Die Werte für die erste
2	1	2	2.25	0		Stimuluskonfiguration bedeuten:
1	1	1	1.75	0		U21 besitzt die Ausprägung 1
1	1	1	2.75	0		U22 die Ausprägung 3,
2	3	1	1.75	0		U23 die Ausprägung 2 etc.
1	2	1	2.25	0		Der Status der Stimuluskonfig. ist 0
1	3	1	2.25	0		Es handelt sich also um eine "normale"
1	2	2	1.75	1		Karte, d.h. die Karte wird zur Berechnung
2	3	2	2.25	1		der Nutzenwerte und der relativen
1	1	2	1.75	1		Wichtigkeit verwendet
2	3	2	1.75	2		
1	2	1	2.75	2		

Schalten Sie dazu die Schreibsperrung aus

Schreibsperrung <--- EIN : rot
AUS : grau

1	3	2	2.75	0
2	2	1	2.75	0
2	1	2	2.25	0
1	1	1	1.75	0
1	1	1	2.75	0
2	3	1	1.75	0
1	2	1	2.25	0
1	3	1	2.25	0
1	2	2	1.75	1
2	3	2	2.25	1
1	1	2	1.75	1
2	3	2	1.75	2
1	2	1	2.75	2

P44.3.1 Erläuterungen zu den Eingabe-Boxen

Box 1: Vereinbare

Speicher fuer x Variable Hilfe

Vereinbare Variable = 24 ;

die Stimuluskonfigurationen müssen in eine Tabelle geschrieben werden. Siehe Programmende

Tabelle_A = 13 ; # Zeilen der Tabelle #
5 ; # Spalten der Tabelle #

Eingabefeld 1: Siehe Almo-Dokument 0 "Arbeiten mit Almo", Abschnitt P0.1.

Eingabefeld 2: Die untersuchten Stimuluskonfigurationen müssen in die Almo-Tabelle TABELLE_A eingegeben werden. In der nachfolgenden Box 4 (Freie Namensfelder) erhalten die für die Stimuluskonfigurationen benötigten Variablen Namen. Die Werte für die Stimuluskonfigurationen werden am Ende des Maskenprogramms eingegeben. Hier, in der 1. Box muß die Tabelle_A im VEREINBARE-Teil definiert werden. Die Zahl der Zeilen muß gleich der Zahl der Stimuluskonfigurationen sein. Die Zahl der Spalten muß gleich der Zahl der Variablen (nominale und quantitative Variable) plus 1 sein. Die zusätzliche Spalte ist zur Definition des Status der Stimuluskonfigurationen erforderlich.

Box 2: Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert.
Siehe P0.2.

Box 3: Datei der Variablennamen
Siehe P0.3.

Box 4: Freie Namensfelder

Freie Namensfelder Hilfe

Namen für die Merkmale, die zur Bildung der Stimuluskonfigurationen verwendet werden (=optional)

BEACHTEN: Als Nummern für diese Merkmale muss eine freie, sonst nicht verwendete Variablennummer verwendet werden

↔ Name21=Geschmack:Butter,pflanzlich;
↔ Name22=Verwendung:Brotaufstrich,Kochen/Backen/Braten,universell;
↔ Name24=Preis:2.50-3.00 DM,2.00-2.49 DM,1.50-1.99 DM;

[...] erzeuge zusätzliche Namensfelder

Siehe auch P0.3.

Die zur Bildung der Stimuluskonfigurationen benötigten Variablen erhalten Namen. In unserem Beispiel sind dies die Variablen:

Geschmack
Verwendung
Kaloriengehalt
Preis

Beachte: Als Nummern für diese Variable müssen freie, sonst nicht verwendete Variablennummern verwendet werden. In unserem Beispiel werden die Variablen V1 bis V12 als Präferenzurteile aus einer Datei eingelesen. Sie sind also bereits belegt. Vereinfacht wurden (in der 1 Box) 24 Variable. Es sind also die Variablen 13 bis 24 frei. Wir haben die Variablen 21 bis 24 verwendet.

Box 5: Datei aus der gelesen wird.
Siehe P0.4.

Box 6: Wenn Dateiformat FIX oder nicht Standard-FREI erforderlich
Siehe P0.4.

Box 7: Die abhängigen Variablen (die Präferenzurteile)



Betrachten wir die Datei ("Conj.fre") aus der die Präferenzurteile eingelesen werden. In unserem Beispiel ist sie

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
--	--	--	--	--	--	--	--	--	---	---	---
01	03	10	11	01	04	09	08	02	07	06	05
02	10	06	08	01	11	02	03	07	04	05	09
03	08	11	06	03	10	05	07	02	04	01	09
04	09	10	07	02	08	01	06	04	03	05	11
05	04	11	09	01	03	10	05	07	08	06	02
.
.
.

In V1 steht eine fortlaufende Nummer, die wir eigentlich nicht brauchen. In V2 bis V12 stehen die Präferenzurteile hinsichtlich der 11 Stimuluskonfigurationen, die wir den Befragten zur Beurteilung vorgelegt hatten. Deswegen muss in die Box eingetragen werden: V2:12. Die Reihenfolge der abhängigen Variablen muß mit der Reihenfolge der Stimuluskonfigurationen übereinstimmen. Siehe dazu die Darstellung in Abschnitt P44.2.1.4.

Box 8 und 9: Die unabhängigen nominalen und quantitativen Variablen
Die "Stimulusvariablen"

die unabhängigen nominalen Variablen

↔ Geschmack, Verwendung

↔ Zahl ihrer Ausprägungen

die unabhängigen quantitativen Variablen

↔ Kaloriengehalt, Preis

↔ Zahl ihrer Ausprägungen

Geben Sie hier die Variablen an, die für die Stimuluskonfigurationen verwendet wurden. Die Reihenfolge in der die Variablen im Eingabefeld stehen, ist nicht beliebig. Ihre Reihenfolge und die Eingabe in TABELLE_A (am Programmende) müssen aufeinander abgestimmt sein. Siehe dazu die ausführliche Darstellung in Abschnitt P44.2.1.1.

Box 10: Option: Ein- und Ausschliessen von Untersuchungseinheiten
Siehe P0.7.

Box 11: Kein_Wert-Angabe und Umkodierungen

↓ Loesche wieder diese Box

Umkodierungen und Kein-Wert-Angaben

Umkodierungen
Kein_Wert-Angabe

die Präferenzurteile werden in ihrer
Kodierungsrichtung umgedreht

↔ ↕ U2:12 <1 bis 11 = Umdrehen>

erzeuge zusätzliche Felder für Umkodierungen / Kein_Wert-Angaben

Kontrollieren, ob Umkodierung so erfolgt wie gewünscht

diese Variablen ...

↔

↔

... aus diesen Datensätzen
vor und nach der Umkodierung
zur Kontrolle anzeigen

Siehe P0.5.

Im 1. Eingabefeld haben wir die Präferenzurteile V2 bis V12 "umgedreht". Der Grund dafür ist folgender:
Üblicherweise wird der höchsten Präferenz der Zahlenwert 1, der 2. höchsten der Zahlenwert 2 usw. zugeordnet. In unserem Beispiel ist dies auch so. In diesem Falle ist es für die Analyse günstig, die Kodierungsrichtung umzudrehen. Die Folgewirkung dieser Umkodierung ist, dass die Koeffizienten für die Nutzenwerte umso höher sind, je höher der Nutzen ist. Würde dieses Umdrehen der Kodierungsrichtung unterbleiben, dann würden höhere Koeffizienten einem niederen Nutzenwert entsprechen.

Box 12: Weitere Modellparameter

weitere Modellparameter	
<input type="text" value="w_squares_of_means"/>	Verfahren = w_squares_of_means oder = fitting_constants
<input type="text" value="13"/>	Zahl der Stimuluskonfigurationen (=Zahl der Zeilen der Tabelle_A)
<input type="text" value="5"/>	Zahl der Spalten der Tabelle_A der Stimuluskonfigurationen

1. *Eingabefeld*: Geben Sie hier das Schätzverfahren an, mit dem Almo rechnen soll. Möglich sind folgende:

```
w_squares_of_means (=weighted squares of means)
fitting_constants
```

Die Ergebnisse, die die beiden Verfahren liefern, sind beim unvollständigen Design weitgehend gleich, aber nicht identisch. Siehe dazu Abschnitt P44.2.1.6 und die Ausführungen im Handbuch zu P20 "Allgemeines Lineares Modell", Abschnitt P20.7.

2. *Eingabefeld*: Geben Sie hier die Zahl der Stimuluskonfigurationen an. Das ist die Zahl der Zeilen der TABELLE_A, die am Programmende eingegeben wird.

3. *Eingabefeld*: Geben Sie hier die Zahl der Spalten der TABELLE_A der Stimuluskonfigurationen an, also die Zahl der Spalten der TABELLE_A, die am Programmende eingegeben wird.

Die Zahl der Zeilen und Spalten der TABELLE_A wurde bereits in Box 1 "Vereinbare" angegeben. Sie muß hier in Box 12 aus programmtechnischen Gründen wiederholt werden.

Box 13: Optionen

Optionen	
<input type="text" value="Preis"/>	Für eine oder mehrere der unabhängigen quantitativen Variablen kann eine Ideal- oder Antiidealfunktion definiert werden
<input type="text" value="1"/>	1 = Individualwerte ausgeben 0 = nicht
<input type="text" value="1"/>	1 = Darstellung des Kalküls für 1. Person 0 = nicht

Wir erläutern zunächst die Eingabefelder 2 und 3

2. *Eingabefeld*: Individualwerte ausgeben. Wird "1" eingetragen, dann wird für jede einzelne Person deren individuelle Werte ausgegeben; bei "0" nicht.

3. *Eingabefeld*: Darstellung des Kalküls für 1. Person. Wird "1" eingetragen, dann wird der Kalkül der conjoint-analysis für die 1. Person vorgeführt.

P44.3.1.1 Ideal- oder Anti-Idealpunkt

1. *Eingabefeld*: Für eine oder mehrere der unabhängigen quantitativen Variablen

kann eine Ideal- oder Antiidealfunktion definiert werden. Unabhängige nominale Variablen dürfen nicht verwendet werden. Das Eingabefeld kann selbstverständlich auch leer gelassen werden. Siehe Abschnitt P44.2.1.3 und P44.1.

Beispiel:

Für die quantitative Variable "Preis" wird eine Ideal- bzw. Antiidealfunktion angefordert. Das bedeutet, daß für diese Variable kein einfacher linearer Zusammenhang vermutet wird

Almo gibt folgendes Ergebnis aus

Matrix der Ideal- oder Antiidealpunkte
=====

Person	1	2
1	2.28	6.00
2	2.95	6.67
3	1.96	-8.67
4	-2.50	-0.67
5	2.38	8.67
6	2.17	-8.00
7	2.14	-4.67
8	1.80	-6.67
9	1.00	-1.33
10	1.33	-2.00
usw.	.	.
MW	1.91	-2.29
SA	1.07	6.09
n	39.00	39.00

Die Spalten bedeuten:

- 1 = Ideal- oder Antiidealfunkt der Variablen "Preis"
- 2 = quadratischer Term von Preis

Der durchschnittliche ideale Preis beträgt MW = 1.91 Es handelt sich dabei um einen Idealpunkt. Preise, die ober- oder unterhalb dieses Punktes liegen, führen zu einer geringeren Präferenz. Das bedeutet, daß es keinen einfachen linearen Zusammenhang gibt.

Sind die Präferenzurteile in Richtung eines höheren Nutzen kodiert (ein höherer Variablenwert drückt eine stärkere Präferenz, d.h. einen höheren Nutzen aus), dann liegt ein Idealpunkt vor, wenn das Vorzeichen des quadratischen Terms negativ ist, andernfalls ein Anti-Idealpunkt.

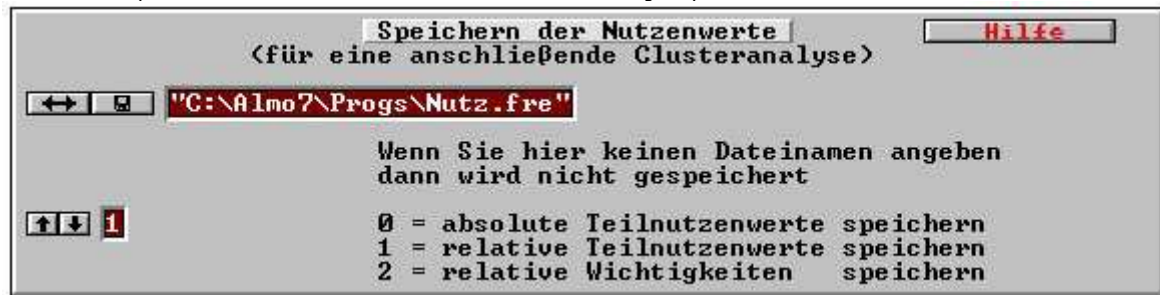
Sind die Präferenzurteile in die entgegengesetzte Richtung kodiert, ist ein Idealpunkt durch ein positives Vorzeichen im quadratischen Term gekennzeichnet und ein Antiidealfunkt durch ein negatives.

Entsprechend der angeführten Regel liegt für Person 1 und 2 ein Antiidealfunkt vor, für Person 3 und 4 ein Idealpunkt. Der Idealpunkt für Person 3 besitzt allerdings keinen realistischen Wert, da er einen Preis von "minus" 8.67 entspricht.

2. *Eingabefeld*: Individualwerte ausgeben. Wird "1" eingetragen, dann wird für jede einzelne Person deren individuelle Werte ausgegeben; bei "0" nicht.

3. *Eingabefeld*: Darstellung des Kalküls für 1. Person. Wird "1" eingetragen, dann wird der Kalkül der conjoint-analysis für die 1. Person vorgeführt.

Box 14: Speichern der Nutzenwerte
(für eine anschließende Clusteranalyse)



1. *Eingabefeld:* Wenn Sie hier einen Dateinamen angeben, dann macht Almo folgendes:

1. Es speichert die Nutzenwerte in eine Datei
2. Es schreibt ein Almo-Programm für eine hierarchische Clusteranalyse
 - a. Dieses Programm wird in der Ergebnisliste ausgegeben und zusätzlich
 - b. wird es in eine Datei geschrieben. Dabei wird der Name der Datei der Nutzenwerte mit der Endung .alm verwendet. In unserem Beispiel wird das Programm in die Datei "C:\Almo\Progs\Nutz.alm" gespeichert.

Wenn Sie das 1. Eingabefeld leer lassen, dann geschieht nichts

Im Beispiel erzeugt Almo folgendes Syntax-Programm für eine Clusteranalyse

```
#AlmPrg      #
VEREINBARE
  Variable=100;

ANFANG
Name1=flnd.Nummer;
Name2=rT_Geschmack_Butter;
Name3=rT_Geschmack_pflanzlich;
Name4=rT_Verwendung_Brotaufstrich;
Name5=rT_Verwendung_Kochen/Backen/Braten;
Name6=rT_Verwendung_universell;
Name7=rT_Kaloriengehalt_1;
Name8=rT_Kaloriengehalt_2;
Name9=rT_Preis_1.75;
Name10=rT_Preis_2.25;
Name11=rT_Preis_2.75;
ENDE

ANFANG
  Programm = 36;           # hierarchische Clusteranalyse #
  A_Quantitative_V = V2:11;
  Modell = Ward_Linkage;
  Distanzmass = quad_euklid;
  Objekte = 40;
  Min_Clusterzahl = 2;
  Max_Clusterzahl = 2;
Ende_Programmparameter

Lese V1:11 aus Datei 2 "C:\Almo15\Progs\Nutz.fre" Format frei
  leerzu Ende;
Gehe_in_Programm
Gehe_zu Lese
ENDE
```

Der Benutzer lädt dieses Syntax-Programm aus der angegebenen Datei. Durch Klick auf den Knopf "Rechne" (in der Knopfleiste unter der Menüleiste) wird es gerechnet. Das Clusteranalyse-Programm kann vom Benutzer beliebig verändert werden. Die Clusterzahl wird von Almo automatisch auf 2 gesetzt. Almo schreibt:

```
Min_Clusterzahl = 2;      (=minimale Clusterzahl)
Max_Clusterzahl = 2;      (=maximale Clusterzahl)
```

Selbstverständlich kann der Benutzer, bevor er mit diesem automatisch erzeugten Programm eine Clusteranalyse rechnet, die Zahl hinter Max_Clusterzahl = ... erhöhen.

2. Eingabefeld: Hier geben Sie an, welche Art der Nutzenwerte gespeichert werden sollen (und damit einer Clusteranalyse unterworfen werden sollen)

0 = absolute Teilnutzenwerte speichern

1 = relative Teilnutzenwerte speichern (empfohlen)

2 = relative Wichtigkeiten speichern

Zum Speichern der Nutzenwerte für eine anschließende Clusteranalyse siehe auch P44.2.1.8.

Box 15: Grafik-Optionen

Siehe P0.10.

Box 16: Schreiben der Stimuluskonfiguration

Schreiben Sie hier dahinter die Tabelle der Stimuluskonfigurationen

Die Spalten bedeuten in unserem Beispiel

U21	(=Geschmack)
: U22	(=Verwendung)
: : U23	(=Kaloriengehalt)
: : : U24	(=Preis)
: : : : :	
: : : : :	Status der Stimuluskonfiguration
: : : : :	0 = "normale" Karte
: : : : :	1 = Holdout-Karte
: : : : :	2 = Simulations-Karte
: : : : :	

1 3 2 2.75 0 ---> In jeder Zeile steht eine Stimulus-
2 2 1 2.75 0 konfiguration. Die Werte für die erste
2 1 2 2.25 0 Stimuluskonfiguration bedeuten:
1 1 1 1.75 0 U21 besitzt die Ausprägung 1
1 1 1 2.75 0 U22 die Ausprägung 3,
2 3 1 1.75 0 U23 die Ausprägung 2 etc.
1 2 1 2.25 0 Der Status der Stimuluskonfig. ist 0
1 3 1 2.25 0 Es handelt sich also um eine "normale"
1 2 2 1.75 1 Karte, d.h. die Karte wird zur Berechnung
2 3 2 2.25 1 der Nutzenwerte und der relativen
1 1 2 1.75 1 Wichtigkeit verwendet
2 3 2 1.75 2
1 2 1 2.75 2

Schalten Sie dazu die Schreibsperre aus

[Schreibsperre] <--- EIN : rot
 AUS : grau

1	3	2	2.75	0
2	2	1	2.75	0
2	1	2	2.25	0
1	1	1	1.75	0
1	1	1	2.75	0
2	3	1	1.75	0
1	2	1	2.25	0
1	3	1	2.25	0
1	2	2	1.75	1
2	3	2	2.25	1
1	1	2	1.75	1
2	3	2	1.75	2
1	2	1	2.75	2

P44.3.1.2 Schreiben der Stimuluskonfigurationen bei unvollständigen Designs

In unserem Beispiel wurde in der 1. Box, der Box "Speicher für x Variable" die **Tabelle_A** mit 13 Zeilen und 5 Spalten angelegt. Sie nimmt die vom Benutzer zu schreibenden Stimuluskonfiguration auf. In der Box "Weitere Modellparameter" wurden nochmals die Zeilen und Spalten dieser Tabelle angegeben.

Es gelten folgende Regeln:

- Die Zahl der Zeilen muß gleich der Zahl der Stimuluskonfigurationen sein. Die Zahl der Spalten muß gleich der Zahl der Variablen (nominale und quantitative Variable) *plus 1* sein. Die zusätzliche Spalte ist zur Definition des Status der Stimuluskonfigurationen erforderlich. In unserem Beispiel sieht die TABELLE_A also folgendermaßen aus.

1	3	2	2.75	0
2	2	1	2.75	0
2	1	2	2.25	0
1	1	1	1.75	0
1	1	1	2.75	0
2	3	1	1.75	0
1	2	1	2.25	0
1	3	1	2.25	0
1	2	2	1.75	1
2	3	2	2.25	1
1	1	2	2.75	1
2	3	2	1.75	2
1	2	1	2.75	2

In der 1. Spalte stehen die Ausprägungen der Variable "Geschmack", in der 2. die Ausprägungen von "Verwendung", in der 3. die von "Kaloriengehalt", in der 4. die von "Preis". In der letzten Spalte steht der Status der Stimuluskonfiguration. Siehe nachfolgend Punkt 2.

- a) Werden nominale und quantitative Stimulusvariablen untersucht, dann gilt folgende Regel: Die nominalen Variablen müssen in der TABELLE_A vor den quantitativen Variablen stehen. In unserem Beispiel müssen entsprechend dieser Regel somit die nominalen Variablen V21 (=Geschmack), V22 (=Verwendung) die ersten beiden Spalten bilden, die quantitativen Variablen V23 (=Kaloriengehalt), V24 (=Preis) die beiden folgenden Spalten.
- b) Für die Reihenfolge innerhalb einer Variablengruppe, also innerhalb der Eingabe-Box "unabhängige nominale Variable" und innerhalb der Box "unabhängige quantitative Variable" gilt folgende Regel: Die Variablen müssen in derselben Reihenfolge in der TABELLE_A stehen wie in der entsprechenden Box. Wird also in der Box der nominalen Variablen

v21,22;

geschrieben, dann muß V21 vor V22 in der TABELLE_A stehen. Würde dagegen

v22,21;

geschrieben, dann müßte V22 vor V21 in der TABELLE_A stehen.

2. Die letzte Spalte der TABELLE_A muß immer die Statusdefinitionen der untersuchten Stimuluskonfigurationen enthalten. Die Zahlenwerte bedeuten:

0 = normale Konfiguration
 1 = Holdout-Konfiguration
 2 = Simulationskonfiguration

P44.4 Programm-Maske Prog44ma (vollständiges Design)

Alle Stimulusmerkmale müssen nominalskaliert sein. Quantitative Variable wie bei Prog44m2 sind nicht erlaubt.

Die Stimuluskonfiguration für das vollständige Design wird automatisch von ALMO selbst erzeugt. Der Benutzer muss es nicht schreiben.

Es ist jedoch möglich Prog44m2 für ein vollständiges Design einzusetzen und dann auch quantitative Variable zu verwenden. Der Preis dafür ist, dass der Benutzer die vollständige Stimuluskonfiguration selbst schreiben muss.

ALMO erzeugt in Prog44ma das vollständige Design nach folgendem Prinzip: Die zuletzt in die Eingabe-Box geschriebene unabhängige nominale Variable wird als erste variiert, dann die zweitletzte usw. Wird beispielsweise in die Eingabebox "unabhängige nominale Variable" folgendes geschrieben

The screenshot shows a graphical user interface for defining stimulus configurations. At the top, a label reads "die unabhängigen nominalen Variablen". Below this, there are two rows of input fields. The first row contains a double-headed arrow icon, a small square icon, and a red box with the text "v10, 13, 15". The second row contains a double-headed arrow icon and a red box with the text "2, 2, 3". To the right of these inputs, the text "Zahl ihrer Ausprägungen" is displayed.

dann erzeugt ALMO folgende Stimuluskonfiguration

V10	V13	V15
1	1	1
1	1	2
1	1	3
1	2	1
1	2	2
1	2	3
2	1	1
2	1	2
2	1	3
2	2	1
2	2	2
2	2	3

Die Eingabe in Prog44ma ist ansonsten dieselbe wie bei Prog44m2, so dass hier auf ihre Darstellung und Erläuterung verzichtet werden kann.

P44.5 Ergebnis aus Almo

Die Ausgabe hängt von den für die Ausgabe und die Zwischenspeicherung verwendeten Optionen ab. Die in dem Beispiel gesetzten Optionen führen zu einer sehr umfassenden Ausgabe.

ALMO protokolliert zunächst die definierten Programmparameter:

```
Ergebnisse aus ALMO
```

```
-----
```

```
Fuer Analyse aus Datenvektor ausgewaehlte Variable
```

```
unabhaengige nominale Variable:
```

V21	Geschmack	UG = 1	OG = 2
V22	Verwendung	UG = 1	OG = 3

unabhaengige quantitative Variable:

V23	Kaloriengehalt	UG = 1	OG = 2	
V24	Preis	UG = 1	OG = 3	Ideal- oder Antiideal-Funktion

Zahl der Ideal- oder Antiideal-Funktionen = 1

Die Variablen "Geschmack" und "Verwendung" werden entsprechend den Programmparameterdefinitionen als nominalskaliert betrachtet, die Variablen "Kaloriengehalt" und "Preis" als quantitativ. Für die Variable Preis wird eine Ideal- bzw. Antiidealfunktion berechnet, da im Programmparameterblock für die Variable Preis das Gewicht gleich 1 (GEWICHT 24=1;) gesetzt wurde. Es wird also vermutet, daß es einen idealen Preis für Margarineprodukte mit einem maximalen Nutzen gibt und quadratische Abweichungen von diesem Preis zu einer Reduktion des Nutzen führen (Annahme einer Idealpunktfunktion).

Formal betrachtet wird folgendes Modell untersucht:

$$y_{sg} = y_{0g} + a_{ig} + b_{jg} + c_{kg} + d_{lg} + e_{sg}$$

mit

$$s = A_i B_j C_k D_l$$

- a_{ig} = Effekt der Ausprägung i in der nominalen Variablen A (=V21) bei Person g
- A_i = Ausprägung von A in der Stimuluskonfiguration s
- b_{jk} = Effekt der Ausprägung j in der nominalen Variablen B (=V22) bei Person g
- B_j = Ausprägung von B in der Stimuluskonfiguration s
- c_{kg} = Effekt der Ausprägung k in der quantitativen Variablen C (=V23) bei Person g
- C_k = Ausprägung von C in der Stimuluskonfiguration s
- d_{lg} = Effekt der Ausprägung l in der quantitativen Variablen D bei Person g
- D_e = Ausprägung von e in der Stimuluskonfiguration s
- e_{sg} = Fehlerterm

Für die erste Stimuluskonfiguration (s=1) mit den Ausprägungen

A1B3C2D1

lautet die Gleichung somit

Rangplatz (y_{1g}) der ersten Stimuluskonfiguration bei Person g

- = durchschnittlicher Gesamtnutzen von g (y_{0g})
- + Nutzen der Person (a_{1g}), wenn die Margarine nach Butter schmeckt (A_1)
- + Nutzen der Person g (b_{3g}), wenn die Margarine universell verwendbar ist (B_3)
- + Nutzen der Person g (c_{2g}), wenn die Margarine normalen Kaloriengehalt besitzt (C_2)
- + Nutzen der Person g (d_{1g}), wenn die Margarine 2,50 bis 3,00 DM kostet (D_1)
- + Fehlerterm bei Person g

abhaengige quantitative Variable:

V2 V8
V3 V9
V4 V10
V5 V11
V6 V12
V7

Zu Kontrollzwecken werden die abhängigen quantitativen Variablen (=Präferenzurteile) ausgegeben.

Verfahren = weighted squares of means

Für die Schätzung der Effekte der nominalen Variablen soll die Methode der "weighted squares of means" verwendet werden.

Design (Modell) = Stimuluskonfig. von Benutzer in Tabelle geschrieben

Zahl der Objete = 40

Die Stimuluskonfigurationen wurden über die TABELLE_A eingegeben und als Objektzahl 40 vereinbart.

Zahl eingelesener Datensätze = 40

Zahl eingelesener Objekte = 40

Es wurden 40 Datensätze eingelesen und an das Programm übergeben. Wird eine WENN-Anweisung verwendet, kann die Zahl der eingelesenen Datensätze größer als die Zahl der Objekte, die in die konjunkte Messung einbezogen werden, sein. Die Zahl der eingelesenen Objekte ist gleich der Zahl der Objekte (Personen usw.), die an das Programm zur Schätzung übergeben werden.

untersuchte Stimuluskonfigurationen:

Stimulus	Status	V21	V22	V23	V24	Praef.Urteil in V..
1	0	1	3	2	3	V2
2	0	2	2	1	3	V3
3	0	2	1	2	2	V4
4	0	1	1	1	1	V5
5	0	1	1	1	3	V6
6	0	2	3	1	1	V7
7	0	1	2	1	2	V8
8	0	1	3	1	2	V9
9	1	1	2	2	1	V10
10	1	2	3	2	2	V11
11	1	1	1	2	3	V12
12	2	2	3	2	1	-
13	2	1	2	1	3	-

Die Variablen bedeuten:
 Status = Status des Stimuluskonfiguration
 0 = Konfiguration zur Schaetzung
 1 = Holdout-Konfiguration
 2 = Simulationskonfiguration
 V21 = Geschmack
 1 = Butter
 2 = pflanzlich
 V22 = Verwendung
 1 = Brotaufstrich
 2 = Kochen/Backen/Braten
 3 = universell
 V23 = Kaloriengehalt
 1 = 1.000
 2 = 2.000
 V24 = Preis
 1 = 1.750
 2 = 2.250
 3 = 2.750

untersuchte Konfigurationen:
 Konfiguration zur Schaetzung = 8
 Holdout-Konfiguration = 3
 Simulationskonfiguration = 2

In der bisher verwendeten Schreibweise ist die erste Stimuluskonfiguration gleich A1B3C2D3, die zweite gleich A2B2C1D3 usw. Bei quantitativen Variablen berechnet ALMO die "quantitativen" Ausprägungen. Die erste Ausprägung der Variablen Preis ist gleich DM 1.75, die zweite gleich DM 2.25 und die dritte gleich DM 2.75.

Beachte: Die ausgegebenen Stimuluskonfigurationen sollten unbedingt dahingehend überprüft werden, ob den einzelnen Konfigurationen die richtige Präferenzvariable zugeordnet wurde. In dem Beispiel steht das Präferenzurteil der ersten Konfiguration in der Variablen V2, das der zweiten in der Variablen V3 usw.

Designmatrix X:

Stim-Konfig = 1		2.0	2.8	1.0	-1.0	-1.0	1.0	7.6
Stim-Konfig = 2		1.0	2.8	-1.0	0.0	1.0	1.0	7.6
Stim-Konfig = 3		2.0	2.2	-1.0	1.0	0.0	1.0	5.1
Stim-Konfig = 4		1.0	1.8	1.0	1.0	0.0	1.0	3.1
Stim-Konfig = 5		1.0	2.8	1.0	1.0	0.0	1.0	7.6
Stim-Konfig = 6		1.0	1.8	-1.0	-1.0	-1.0	1.0	3.1
Stim-Konfig = 7		1.0	2.2	1.0	0.0	1.0	1.0	5.1
Stim-Konfig = 8		1.0	2.2	1.0	-1.0	-1.0	1.0	5.1
Stim-Konfig = 9		2.0	1.8	1.0	0.0	1.0	1.0	3.1
Stim-Konfig = 10		2.0	2.2	-1.0	-1.0	-1.0	1.0	5.1
Stim-Konfig = 11		2.0	2.8	1.0	1.0	0.0	1.0	7.6
Stim-Konfig = 12		2.0	1.8	-1.0	-1.0	-1.0	1.0	3.1
Stim-Konfig = 13		1.0	2.8	1.0	0.0	1.0	1.0	7.6

Die von ALMO intern gebildete Designmatrix **X** wird nur ausgegeben, wenn die Option ZWISCHERGEB = 1; gesetzt wurde. Die Designmatrix ist allgemein wie folgt aufgebaut:

Werte in den ersten Spalten = quantitative Variablen.
 Werte in den weiteren Spalten = notwendigen Dummies der nominalen Variablen.
 Werte in der nächsten Spalte = Konstante
 Werte in den weiteren Spalten = quadratische Terme bei Ideal- und Antiidealfunktionen
 Spaltene

In unserem Beispiel sind die Spalten der Designmatrix **X** wie folgt definiert:

- 1. Spalte = Werte von V23
- 2. Spalte = Werte von V24 (gerundet), 2.8 steht für 2.75, 2.2 für 2.25, 1.8 für 1.75
- 3. Spalte = 1. Dummy von V21. Da V21 nur zwei Ausprägungen besitzt, gibt es nur eine notwendige Dummy-Variable
- 4. und 5 Spalte = notwendige Dummies von V22. Da V22 drei Ausprägungen hat, gibt es zwei notwendige Dummy-Variablen
- 6. Spalte = Konstante diese hat immer den Wert 1
- 7. Spalte = quadratischer Term von V24. In der ersten Zeile steht der gerundete Wert von $2.75^2 = 7.56$

Bei der konjunkten Messung besitzen alle Personen immer dieselbe Designmatrix. Sie unterscheiden sich nur in den Präferenzen für die einzelnen Stimuluskonfigurationen. Entsprechend dem der verbundenen Messung zugrundeliegenden Kalkül des allgemeinen linearen Modells lautet die Schätzgleichung für eine Person *g* in Matrixschreibweise:

$$\mathbf{y}_g = \mathbf{X} \cdot \boldsymbol{\beta}_g + \mathbf{e}_g$$

\mathbf{y}_g = Präferenzwerte der Person *g*

X = Designmatrix

$\boldsymbol{\beta}_g$ = Regressionskoeffizienten der Person *g*

\mathbf{e}_g = Residuen der Person *g*

In die Schätzgleichung gehen nur die "normalen" Karten ein, in unserem Beispiel also die ersten acht Stimuluskonfigurationen. Da zur Effektberechnung die Methode der "weighted squares of means" verwendet wird, wird für die nominalen Variablen die -1/0/1-Kodierung verwendet. Die notwendigen Dummies sind wie folgt definiert

	notwendige Dummy			notwendige Dummies	
Geschmack (V21)	a_i		Verwendung (V22)	b_i	b_2
A ₁	1		B ₁	1	0
A ₂	-1		B ₂	0	1
			B ₃	-1	-1

Die Schätzgleichung, mit der ALMO rechnet, lautet für die erste Stimuluskonfiguration, wenn die Variablen fortlaufend mit X_i , numeriert werden

$$y_{1g} = X_1 (= 2.0) \cdot b_1 + X_2 (= 2.75) \cdot b_2 + X_3 (= 1.0) \cdot b_3 + X_4 (= -1.0) \cdot b_4 + X_5 (= -1.0) \cdot b_5 + X_6 (= +1.0) \cdot b_6 + X_7 (= 2.75^2 = 7.5) \cdot b_7 + e_{1g}$$

mit

X_1 = V23 (Kaloriengehalt)

X_2 = V24 (Preis)

X_3 = notwendige Dummy a_1 von V21

X_4 = 1. notwendige Dummy b_1 von V22

X_5 = 2. notwendige Dummy b_1 von V22

X_6 = Konstanteneffekt

X_7 = quadratischer Term von V24

Die Schätzwerte für die Regressionskoeffizienten einer Person *g* sind entsprechend der Methode der

kleinsten Quadrate:

$$\mathbf{b}_g = (\mathbf{X}' \cdot \mathbf{X})^{-1} \cdot (\mathbf{X}' \cdot \mathbf{y}_g)$$

Bei ZWISCHERGERB = 1; wird die Berechnung der Regressionskoeffizienten für die erste Person wie folgt dokumentiert: Als erstes wird die Kreuzproduktmatrix ($\mathbf{X}' \cdot \mathbf{X}$) ausgegeben:

Quadratsummenmatrix bzw. Kreuzproduktmatrix:

```
i = 1 | 14.00
i = 2 | 23.50 44.00
i = 3 | 2.00 5.00 8.00
i = 4 | 0.00 0.00 0.00 6.00
i = 5 | -2.00 -1.75 -1.00 3.00 5.00
i = 6 | 10.00 18.50 2.00 0.00 -1.00 8.00
i = 7 | 56.62 107.28 12.62 0.00 -3.06 44.00 267.22
```

Daran anschließend wird die Inverse ($\mathbf{X}' \cdot \mathbf{X})^{-1}$ der Kreuzproduktmatrix ausgegeben:

Inverse der Quadratsummenmatrix bzw. Kreuzproduktmatrix:

```
i = 1 | 1.17
i = 2 | -6.67 229.33
i = 3 | 0.17 -1.67 0.17
i = 4 | -0.22 2.22 -0.06 0.30
i = 5 | 0.44 -4.44 0.11 -0.26 0.52
i = 6 | 6.64 -250.81 1.78 -2.44 4.87 276.20
i = 7 | 1.33 -49.33 0.33 -0.44 0.89 53.78 10.67
```

Daran anschließend werden für den ersten Datensatz die Regressionskoeffizienten ($=\mathbf{b}_{g=1}$) ausgegeben

Regressionskoeffizienten:

Variable	Regressions- koeffizient
Konstante	36.458
V21 Geschmack	
1 Butter	3.000
V22 Verwendung	
1 Brotaufs	0.667
2 Kochen/B	-2.000
V23 Kaloriengeha	-1.000
V24 Preis	-27.333
quadratischer Term =	6.000
Idealpunkt =	2.278

Der Ideal- bzw. Antiidealpunkt berechnet sich allgemein wie folgt:

$$\text{Idealpunkt} = - \text{Regressionskoeffizient} / 2 * \text{Regressionskoeffizient des quadratischen Term}$$

In unserem Beispiel ergibt sich ein Wert von $- (-27.333) / 2 * 6.000 = 2.278$. Da das Vorzeichen des quadratischen Terms positiv ist und die Präferenzurteile in Richtung eines höheren Nutzens kodiert sind, liegt ein Antiidealpunkt vor (siehe Abschnitt P44.1). Die erste Person bevorzugt also Produkte, deren Preis entweder nach oben oder nach unten von DM 2.28 abweicht.

Aus den Regressionskoeffizienten werden die Effekte nach folgendem Prinzip berechnet:

Effekt der Ausprägung einer quantitativen Variablen, für die eine lineare Funktion (also keine Ideal- oder Antiidealfunktion) definiert ist:

$$\text{Effekt} = \text{Regressionskoeffizient} * \text{Variablenwert}$$

Effekt der Ausprägung einer quantitativen Variablen, für die eine Ideal- oder Antiidealpunktfunktion definiert ist:

$$\text{Effekt} = \text{Regressionskoeffizient} * \text{Variablenwert} + \text{Regressionskoeffizient für den quadrat. Variablenwert} * \text{quadratierter Variablenwert}$$

Effekt der Ausprägungen einer nominalen Variablen:

bei der Methode der "weighted squares of sums":

$$\begin{aligned} \text{Effekt der notwendigen Ausprägungen (alle Ausprägungen ohne der letzten)} &= \text{Regressionskoeffizient der entsprechenden notwendigen Dummy} \\ \text{Effekt der letzten Ausprägungen} &= \text{Summe der Effekte der notwendigen Ausprägungen mal "-1"}. \end{aligned}$$

bei der Methode der "fitting constants" (siehe Holm, K., 1979: Die Befragung 6. München, S. 75-78).

Für die Person 1 ergeben sich aus den berechneten Regressionskoeffizienten folgende Effekte (in Klammern Darstellung der Berechnung):

Effekte (Teilnutzenwerte):

Variable		Effekt (Teilnutzenwert)	
Konstante		36.458	(= 36.458)
V21	Geschmack		
	1 = Butter	3.000	(= 3.000)
	2 = pflanzlich	-3.000	(= -(3.000))
V22	Verwendung		
	1 = Brotaufstrich	0.667	(= 0.667)
	2 = Kochen/Backe	-2.000	(= -2.000)
	3 = universell	1.333	(= -(0.667 + (-2.000)))
V23	Kaloriengeha		
	1 = 1.000	-1.000	(= -1.000 · 1.000)
	2 = 2.000	-2.000	(= -1.000 · 2.000)
V24	Preis		
	1 = 1.750	-29.458	(= -27.333 · 1.75 + 6.00 · 1.75 ²)
	2 = 2.250	-31.125	(= -27.333 · 2.25 + 6.00 · 2.25 ²)
	3 = 2.750	-29.792	(= -27.333 · 2.75 + 6.00 · 2.75 ²)

Die Interpretation der Effekte ist abhängig von der Kodierrichtung der abhängigen Variablen.

In unserem Beispiel wurden diese so umkodiert, daß ein größerer Variablenwert eine stärkere Präferenz ausdrückt. Ein positives Vorzeichen bedeutet eine Erhöhung des Nutzen, ein negatives Vorzeichen eine Reduktion. Bei Person 1 erhöht der Buttergeschmack den Nutzen. Dies trifft auch auf die Verwendung als Brotaufstrich sowie auf eine universelle Verwendung zu. Die Effekte der quantitativen Variablen sind alle negativ. Sie reduzieren den Gesamtnutzen.

Nach der Darstellung der Berechnung der Effekte werden bei ZWISCHERGE = 1; die aufgrund der Regressionsgleichung berechneten Prognosewerte ausgegeben. Diese werden zur Berechnung des Pearsonschen Korrelationskoeffizienten R und von Kendalls tau_b zur Modellprüfung benötigt. Formal sind die Prognosewerte wie folgt definiert:

$$\tilde{y}_g = \mathbf{X} \cdot \mathbf{b}_g$$

Für den ersten Befragten ergeben sich folgende Prognosewert:

Stim.-Konfig. = 1		y = 9.000	y(prognostiziert) = 9.000
Stim.-Konfig. = 2		y = 2.000	y(prognostiziert) = 0.667
Stim.-Konfig. = 3		y = 1.000	y(prognostiziert) = 1.000
Stim.-Konfig. = 4		y = 11.000	y(prognostiziert) = 9.667
Stim.-Konfig. = 5		y = 8.000	y(prognostiziert) = 9.333
Stim.-Konfig. = 6		y = 3.000	y(prognostiziert) = 4.333
Stim.-Konfig. = 7		y = 4.000	y(prognostiziert) = 5.333
Stim.-Konfig. = 8		y = 10.000	y(prognostiziert) = 8.667

Berechnung des Pearsonschen Korrelationskoeffizienten R
(zur genaueren Information siehe Abschnitt P44.1)

$$R = \frac{12.167}{3.488 * 3.674} = 0.949$$

Es besteht kein perfekter Zusammenhang zwischen empirischen und prognostisierten Werten.

$$\text{Signifikanz } 100*(1-p) = 77.240$$

Der Pearsonsche Korrelationskoeffizient weicht nicht signifikant von Null (kein Zusammenhang) ab, wenn ein Niveau von 90%, 95% oder 99% angenommen wird. Das Modell der konjunkten Messung ist für den ersten Datensatz nicht angemessen, wenn zur Modellbeurteilung der Pearsonsche Korrelationskoeffizient verwendet wird.

Berechnung von Tau-b (zur genaueren Information siehe Abschnitt P44.1):

$$\text{tau-b} = \frac{24.000 - 4.000}{5.292 * 5.292} = 0.714$$

$$\begin{aligned} N_s &= 24.000 \\ N_d &= 4.000 \\ T_x &= 0.000 \\ T_y &= 0.000 \end{aligned}$$

Es treten insgesamt 24 Paare mit gleichgerichteten Beziehungen auf und 4 mit entgegengerichteten Beziehungen. Bindungen liegen nicht vor.

Signifikanz von tau-b

$$\text{Signifikanz } 100*(1-p) = 98.669$$

Im Unterschied zum Pearsonschen Korrelationskoeffizienten ist Kendalls tau_b signifikant von Null verschieden. Beide Koeffizienten führen also zu unterschiedlichen Ergebnissen. Wie sich bei der Analyse der Holdout-Karten zeigen wird, sprechen mehr Befunde dafür, daß für den ersten Befragten das untersuchte Modell nicht angemessen ist.

Nach der Darstellung des Kalküls werden die Matrix der Effekte, die Signifikanzen und die Idealpunkte ausgegeben:

Matrix der Effekte

=====

DA	0	1	2	3	4	5	6	7	8	9
1	36.46	3.00	-3.00	0.67	-2.00	1.33	-1.00	-2.00	-29.46	-31.13
2	60.19	0.33	-0.33	-1.78	3.22	-1.44	0.33	0.67	-48.42	-54.75
3	-25.40	0.67	-0.67	0.11	-1.56	1.44	-0.33	-0.67	32.96	32.63
4	19.60	0.17	-0.17	0.11	-1.22	1.11	-1.83	-3.67	-7.88	-10.88
5	50.57	3.33	-3.33	1.22	0.22	-1.44	1.83	3.67	-45.79	-49.13
6	-22.50	0.00	0.00	1.00	-2.67	1.67	-5.50	-11.00	36.17	37.50
7	-7.49	1.67	-1.67	-0.56	-0.22	0.78	-5.33	-10.67	20.71	21.37
8	-12.69	0.67	-0.67	0.78	-1.89	1.11	-0.83	-1.67	21.58	20.25
9	8.03	-0.17	0.17	-3.11	1.89	1.22	-0.67	-1.33	0.58	-0.75
10	5.46	3.00	-3.00	0.67	0.33	-1.00	-1.50	-3.00	3.21	1.87

usw.

MW	2.87	0.97	-0.97	0.11	-0.65	0.55	-1.50	-3.01	6.02	5.23
SA	29.43	1.45	1.45	1.70	1.78	1.32	2.68	5.37	29.64	31.43
n	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00

DA	10	R	TAU
1	-29.79	0.95	0.71
2	-57.75	1.00	1.00
3	27.96	0.96	0.93
4	-14.21	1.00	1.00
5	-48.13	1.00	1.00
6	34.83	1.00	1.00
7	19.71	1.00	1.00
8	15.58	0.93	0.86
9	-2.75	0.98	1.00
10	-0.46	1.00	1.00

usw.

MW	3.32	0.99	0.97
SA	30.32	0.02	0.07
n	40.00	40.00	40.00

Die Spalten bedeuten:

- 0 = Konstante
 - 1 = V21 (=Geschmack), Auspr. = 1 (=Butter)
 - 2 = V21 (=Geschmack), Auspr. = 2 (=pflanzlich)
 - 3 = V22 (=Verwendung), Auspr. = 1 (=Brotaufstrich)
 - 4 = V22 (=Verwendung), Auspr. = 2 (=Kochen/Backen/Braten)
 - 5 = V22 (=Verwendung), Auspr. = 3 (=universell)
 - 6 = V23 (=Kaloriengehalt), Auspr. 1 = 1.000
 - 7 = V23 (=Kaloriengehalt), Auspr. 2 = 2.000
 - 8 = V24 (=Preis), Auspr. 1 = 1.000
 - 9 = V24 (=Preis), Auspr. 2 = 2.000
 - 10 = V24 (=Preis), Auspr. 3 = 1.750
- R = Pearsonscher Korrelationskoeffizient
TAU = Kendalls tau-b

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA

Die Interpretation soll am Beispiel der durchschnittlichen (mittleren Effekte) dargestellt werden. Die

Effekte der einzelnen Personen lassen sich analog interpretieren, mit dem Unterschied, daß auf die Höhe der Standardabweichung nicht geachtet werden muß. Die über alle Personen gemittelten Effekte (durchschnittliche Effekte) stehen in der Zeile MW, die Standardabweichungen in der Spalte SA. Die gemittelten Effekte weisen relativ hohe Streuungen (Zeile SA) auf. Der durchschnittliche Effekt der Ausprägung "Brotaufstrich" beträgt beispielsweise 0.11, die Standardabweichung ist mit einem Wert von 1.70 mehr als zehnmal so groß. Es liegt somit mit einer hohen Wahrscheinlichkeit keine homogene Population vor. Die berechneten Durchschnittswerte sind somit wenig aussagekräftig. Sie sollen dennoch hier zu Illustrationszwecken interpretiert werden:

Durchschnittliche Effekte der nominalen Variablen V21 ("Geschmack") und V22 ("Verwendung"): Ihre Interpretation wurde bereits in Abschnitt P44.1 behandelt. In dem Beispiel erhöht sich die Präferenz für eine Magarine allgemein, wenn sie nach Butter schmeckt (durchschnittlicher Effekt = 0.97) und universell verwendbar (durchschnittlicher Effekt = 0.54). Die Präferenz erhöht sich ferner schwach, wenn sie als Brotaufstrich verwendet werden kann (durchschnittlicher Effekt = 0.106).

Durchschnittlicher Effekt der quantitativen Variablen V23 ("Kaloriengehalt"): Die Präferenz nimmt im Durchschnitt mit dem Kaloriengehalt ab, da der durchschnittliche Effekt der Ausprägung "2" (=normaler Kaloriengehalt) stärker negativ ist als jener für die Ausprägung "1" (=kalorienreduziert).

Durchschnittlicher Effekt der quantitativen Variablen V24 ("Preis"): Bei der Interpretation der Effekt der Variablen "Preis" muß berücksichtigt werden, daß für diese Variable eine Ideal- bzw. Antiidealpunktfunktion definiert wurde. Das. bedeutet, daß es keinen einfachen linearen Zusammenhang gibt. Der durchschnittliche ideale Preis beträgt DM 1.91 (siehe Ausgabe weiter unten). Es handelt sich dabei um einen Idealpunkt. Preise, die ober- oder unterhalb dieses Punktes liegen, führen zu einer geringeren Präferenz.

Signifikanz von R und TAU:

DA	R	Sign.	TAU	Sign.
1	0.949	77.240	0.714	98.669
2	0.997	98.405	1.000	99.938
3	0.956	79.998	0.929	99.855
4	0.999	99.515	1.000	99.938
5	0.999	99.593	1.000	99.938
6	0.999	99.381	1.000	99.938
7	1.000	KW	1.000	99.938
8	0.925	68.150	0.857	99.690
9	0.976	88.344	1.000	99.938
10	0.999	99.603	1.000	99.938

usw.

Sign. = Signifikanzniveau $100 \cdot (1-p)$

p = Fehlerwahrscheinlichkeit

Der Pearsosche Korrelationskoeffizient R ist für die erste Person nicht signifikant von Null verschieden. Es liegt also keine gute Modellanpassung vor, τ_b ist dagegen signifikant von Null verschieden. (Da echte Rangreihen vorliegen, tritt das Problem der Bindungen nicht auf. Die Verwendung von τ_b ist zulässig. Sie hat gegenüber der Verwendung des Pearsonschen Korrelationskoeffizienten R den Vorteil, daß der Ordinalität der empirischen Rangdaten "besser" Rechnung getragen wird, siehe dazu auch Abschnitt P44.1). Für die zweite Person sind beide Koeffizienten signifikant von Null verschieden.

Matrix der Ideal- oder Antiidealpunkte

=====

DA	1	2
1	2.28	6.00
2	2.95	6.67
3	1.96	-8.67
4	-2.50	-0.67
5	2.38	8.67
6	2.17	-8.00
7	2.14	-4.67
8	1.80	-6.67
9	1.00	-1.33
10	1.33	-2.00
usw.		
MW	1.91	-2.29
SA	1.07	6.09
n	39.00	39.00

Die Spalten bedeuten:

- 1 = Ideal- oder Antiidealpunkt von V24 (=Preis)
- 2 = quadratischer Term von V24 (=Preis)

Beachte: Sind die Praeferenzurteile in Richtung eines hoeheren Nutzen kodiert (ein hoeherer Variablenwert drueckt eine staerkere Praeferenz, einen hoeheren Nutzen aus), dann liegt ein Idealpunkt vor, wenn das Vorzeichen des quadratischen Terms negativ ist, andernfalls ein Anit-idealpunkt. Sind die Praeferenzurteile in die entgegengesetzte Richtung kodiert, ist ein Idealpunkt durch ein positives Vorzeichen im quadratischen Term gekennzeichnet und ein Antiidealpunkt durch ein negatives.

Entsprechend der angeführten Regel liegt für Person 1 und 2 ein Antiidealpunkt vor, für Person 3 und 4 ein Idealpunkt vor. Der Idealpunkt für Person 3 besitzt allerdings keinen realistischen Wert, da er einen Preis von "minus" DM 8.67 entspricht.

Nach der Ausgabe der absoluten Effekte, der Signifikanzen und gegebenenfalls der Idealpunkte werden die Holdout-Karten untersucht. Holdout-Karten dienen der Validitätsprüfung. Für den ersten Befragten ergibt sich folgendes Bild:

```
Prognosewert fuer Stimuluskonfiguration 9 (=erste Holdout-Karte)
= 36.458
+ -1.000 * 2.000
+ -27.333 * 1.750
+ 3.000 * 1.000
+ 0.667 * 0.000
+ -2.000 * 1.000
+ 6.000 * 3.062
= 6.000
```

```
Prognosewert fuer Stimuluskonfiguration 10
= 36.458
+ -1.000 * 2.000
+ -27.333 * 2.250
+ 3.000 * -1.000
+ 0.667 * -1.000
+ -2.000 * -1.000
+ 6.000 * 5.062
= 1.667
```

```
Prognosewert fuer Stimuluskonfiguration 11
= 36.458
+ -1.000 * 2.000
+ -27.333 * 2.750
+ 3.000 * 1.000
```

```

+   0.667 *   1.000
+  -2.000 *   0.000
+   6.000 *   7.562
=   8.333

```

Gegenueberstellung von empirischen und prognostizierten Werten:

```

Stim.-Konfig. = 9 | y = 5.000 y(prognostiziert) = 6.000
Stim.-Konfig. = 10 | y = 6.000 y(prognostiziert) = 1.667
Stim.-Konfig. = 11 | y = 7.000 y(prognostiziert) = 8.333

```

```

          0.778
R = ----- = 0.345
      2.762 * 0.816

```

Signifikanz $100*(1-p) = 23.269$

```

      2.000 - 1.000
tau-b = ----- = 0.333
      1.732 * 1.732

```

Signifikanz $100*(1-p) = 39.861$

Die prognostizierten und empirischen Werte für die Holdout-Karten stimmen für den ersten Befragten nur in einem geringen Ausmaß überein. Sowohl der Pearsonsche Korrelationskoeffizient als auch Kendalls τ_b sind nicht signifikant von 0 verschieden. Das Modell für den ersten Befragten besitzt nur eine geringe Validität, wenn zur Validitätsprüfung die Holdout-Karten verwendet werden.

Matrix der Residuenwerte

=====

DA	1	2	3	R	TAU
1	-1.00	4.33	-1.33	0.34	0.33
2	-8.00	2.67	1.33	0.78	1.00
3	2.00	3.67	0.33	1.00	1.00
4	2.00	1.00	-1.00	1.00	1.00
5	-8.00	5.67	-0.67	0.09	-0.33
6	2.00	0.33	-1.33	0.82	0.33
7	-1.00	3.67	-0.67	-0.45	-0.33
8	2.00	2.67	2.33	0.99	1.00
9	2.00	0.67	0.33	0.99	1.00
10	1.00	5.67	-0.67	0.66	0.33
usw.					

MW	0.38	2.06	0.02	0.73	0.63
SA	3.07	2.78	1.26	0.35	0.45
n	40.00	40.00	40.00	40.00	40.00

Die Spalten bedeuten:

- 1 = Residuenwert fuer Holdout-Karte 1 (Stim.-Konfig=9)
- 2 = Residuenwert fuer Holdout-Karte 2 (Stim.-Konfig=10)
- 3 = Residuenwert fuer Holdout-Karte 3 (Stim.-Konfig=11)
- R = Pearsonscher Korrelationskoeffizient
- TAU = Kendalls tau-b

Im Unterschied zum ersten Befragten werden für die Befragten 2 bis 4 bei Verwendung von τ_b perfekte Validitätswerte erzielt. Wegen der geringen Fallzahl (diese entspricht den drei Karten) ist τ_b allerdings nicht signifikant von Null verschieden.

Signifikanz von R und TAU:

```
-----
```

DA	R	Sign.	TAU	Sign.
1	0.345	23.269	0.333	39.861
2	0.777	56.437	1.000	88.292
3	0.995	93.678	1.000	88.292
4	0.999	96.755	1.000	88.292
5	0.085	5.386	-0.333	39.861
6	0.818	60.719	0.333	39.861
7	-0.454	30.607	-0.333	39.861
8	0.987	89.394	1.000	88.292
9	0.994	93.172	1.000	88.292
10	0.663	46.253	0.333	39.861

usw.

Nach der Analyse der Holdout-Karten werden die Simulationskarten untersucht. Für den ersten Befragten ergeben sich folgende Prognosewerte für die beiden Simulationskarten:

```
Prognosewert fuer Stimuluskonfiguration 12 (=erste Simulationskarte)
= 36.458
+ -1.000 * 2.000
+ -27.333 * 1.750
+ 3.000 * -1.000
+ 0.667 * -1.000
+ -2.000 * -1.000
+ 6.000 * 3.062
= 3.333
```

```
Prognosewert fuer Stimuluskonfiguration 13
= 36.458
+ -1.000 * 1.000
+ -27.333 * 2.750
+ 3.000 * 1.000
+ 0.667 * 0.000
+ -2.000 * 1.000
+ 6.000 * 7.562
= 6.667
```

Der prognostizierte Rangplatz für die Konfiguration 12 (=erste Simulationskarte) ist gleich 3.333, jener der zweiten gleich 6.667. Würde eine der beiden Simulationskarten gegenüber allen empirisch erhobenen bevorzugt werden, müßte sie einen Prognosewert größer 11 haben. Bei Person 1 ist dies nicht der Fall. Insgesamt ergeben sich folgende Prognosewerte:

Matrix der empirischen Werte bzw. der Prognosewerte
=====

DA	11	12	13
1	7.00	3.33	6.67
2	3.00	10.67	6.33
3	3.00	7.67	1.33
4	1.00	9.00	2.50
usw.			
17	4.00	12.33	2.17
usw.			
MW	4.28	5.46	5.00
SA	2.85	3.37	2.23
n	40.00	40.00	40.00

Die Spalten bedeuten:

- 11 = empirischer Wert fuer Stimuluskarte 11
- 12 = Prognosewert fuer Simulationskarte 12
- 13 = Prognosewert fuer Simulationskarte 13

Beim 17. Befragten tritt für die erste Simulationskarte ein Prognoswert größer 11 auf. Sie erhält daher auch den höchsten Rangplatz, also die stärkste Präferenz.

Matrix der Rangplaetze
=====

DA	1	2
1	4.00	8.00
2	12.00	7.00
3	9.00	2.00
4	10.00	3.00
usw.		
17	13.00	3.00
usw.		
MW	6.47	6.03
SA	3.43	2.69
n	40.00	40.00

Die Rangplaetze bedeuten: 1 = kleinster Rang, 13 = hoechster Rang

Die Spalten bedeuten:

1 = Rangplatz fuer Simulationskarte 1 (Stim.-Konf. =12)

2 = Rangplatz fuer Simulationskarte 2 (Stim.-Konf. =13)

Beim Befragten 17 erhält die Simulationskarte 1 den höchsten Rangplatz. Sie wird gegenüber allen anderen Stimuluskonfigurationen bevorzugt. Ihre Bevorzugswahrscheinlichkeit ist daher gleich 1.

Matrix der Bevorzugswahrscheinlichkeiten
=====

DA	1	2
1	0.00	0.00
2	0.00	0.00
3	0.00	0.00
4	0.00	0.00
5	0.00	0.00
usw.		
17	1.00	0.00
usw.		
MW	0.05	0.00
SA	0.22	0.00
n	40.00	40.00

Die Spalten bedeuten:

1 = Bevorzugswahrsch.keit fuer Simulationskarte 1 (Stim.-Konf. = 12)

2 = Bevorzugswahrsch.keit fuer Simulationskarte 2 (Stim.-Konf. = 13)

Interpretation: Insgesamt (im Durchschnitt) wird die Simulationskarte 1 nur von 5 Prozent der Befragten gegenüber den erhobenen Konfigurationen bevorzugt, die Simulationskarte 2 von 0 Prozent.

Nach der Analyse der Simulationskarten werden aus den absoluten Teilnutzenwerten die relativen Teilnutzenwerte berechnet. Bei Verwendung von der Option ZWISCHERGEB = 1; stellt ALMO exemplarisch die Berechnung für den ersten Datensatz dar.

Die Teilnutzenwerte jeder Ausprägung werden zunächst justiert mit

$$\begin{aligned} \text{justierter Teilnutzenwert} & \quad \text{Teilnutzenwert} & \quad \text{minimaler Teilnutzenwert} \\ \text{der Ausprägung } k \text{ der} & = \text{ der Ausprägung } k - \text{ der Variablen } i \\ \text{der Variablen } i & \quad \text{der Variablen } i \end{aligned}$$

$$\text{minimaler Teilnutzenwert von V21 Geschmack} = -3.0000 \text{ (Ausprägung 2)}$$

Die minimalen Teilnutzenwerte der anderen Variablen sind:

$$\begin{aligned} V22 &= -2.000 \\ V23 &= -2.000 \\ V24 &= -31.125 \end{aligned}$$

Die justierten Teilnutzenwerte von V21 sind:

$$\begin{aligned} 3.0000 - (-3.0000) &= 6.0000 \\ -3.0000 - (-3.0000) &= 0.0000 \end{aligned}$$

justierte Teilnutzenwerte (in Klammern Darstellung der Berechnung):

Variable		justierter Teilnutzen- wert	
V21	Geschmack		
	1 = Butter	6.000	(Effekt-minimaler Teilnutzenwert = $-3.000 - (-3.000) = 6.000$)
	2 = pflanzlich	0.000	(Effekt-minimaler Teilnutzenwert = $-3.000 - (-3.000) = 0.000$)
V22	Verwendung		
	1 = Brotaufstrich	2.667	($0.667 - (-2.000) = 2.667$)
	2 = Kochen/Backe	0.000	($2.000 - (-2.000) = 0$)
	3 = universell	3.333	($1.333 - (-2.000) = 3.333$)
V23	Kaloriengeha		
	1 = 1.000	1.000	($-1.000 - (-2.000) = 1.000$)
	2 = 2.000	0.000	($-1.000 - (-2.000) = 0.000$)
V24	Preis		
	1 = 1.750	1.667	($-29.458 - (-31.125) = 1.667$)
	2 = 2.250	0.000	($-31.125 - (-31.125) = 0.000$)
	3 = 2.750	1.333	($-29.792 - (-31.125) = 1.333$)

Aus den justierten Teilnutzenwerten werden die normierten Teilnutzenwerte berechnet mit

$$\begin{aligned} \text{norm. Teilnutzen} & \quad \text{justierter Teilnutzen der Ausprägung } k \text{ der Var. } i \\ \text{der Ausprägung } k & = \frac{\text{-----}}{\text{Summe der maximalen justierten Teilnutzen aller Var.}} \\ \text{der Variablen } i & \end{aligned}$$

maximale justierte Teilnutzenwerte der Variablen =

$$\begin{aligned} & 6.0000 \text{ (von V21)} \\ + & 3.3333 \text{ (von V22)} \\ + & 1.0000 \text{ (von V23)} \\ + & 1.6667 \text{ (von V24)} \\ = & 12.0000 \end{aligned}$$

normierte Teilnutzenwerte von V21

$$\begin{aligned} 6.0000 / 12.0000 &= 0.5000 \\ 0.0000 / 12.0000 &= 0.0000 \end{aligned}$$

normierte Teilnutzenwerte:

Variable		normierter Teilnutzen- wert	
V21	Geschmack		
	1 = Butter	0.500	($6.000/12=0.500$)
	2 = pflanzlich	0.000	($0.000/12=0.000$)
V22	Verwendung		

	1 = Brotaufstrich	0.222	(2.667/12=0.222)
	2 = Kochen/Backen	0.000	(0.000/12=0.000)
	3 = universell	0.278	(3.333/12=0.278)
V23	Kaloriengehalt		
	1 = 1.000	0.083	(1.000/12=0.083)
	2 = 2.000	0.000	(0.000/12=0.000)
V24	Preis		
	1 = 1.750	0.139	(1.667/12=0.139)
	2 = 2.250	0.000	(0.000/12=0.000)
	3 = 2.750	0.111	(1.333/12=0.111)

Befragter 1 schreibt dem butterähnlichen Geschmack den relativ größte Teilnutzen zu.

Matrix der normierten Teilnutzenwerte

```
=====
```

DA	1	2	3	4	5	6	7	8	9	10
1	0.50	0.00	0.22	0.00	0.28	0.08	0.00	0.14	0.00	0.11
2	0.04	0.00	0.00	0.33	0.02	0.00	0.02	0.61	0.20	0.00
3	0.14	0.00	0.17	0.00	0.31	0.03	0.00	0.52	0.48	0.00
4	0.03	0.00	0.12	0.00	0.22	0.17	0.00	0.58	0.31	0.00
5	0.46	0.00	0.18	0.11	0.00	0.00	0.13	0.23	0.00	0.07
6	0.00	0.00	0.29	0.00	0.35	0.44	0.00	0.11	0.21	0.00
7	0.29	0.00	0.00	0.03	0.11	0.46	0.00	0.09	0.14	0.00
8	0.12	0.00	0.24	0.00	0.27	0.07	0.00	0.54	0.42	0.00
9	0.00	0.04	0.00	0.54	0.46	0.07	0.00	0.36	0.21	0.00
10	0.47	0.00	0.13	0.10	0.00	0.12	0.00	0.29	0.18	0.00
usw.										
MW	0.19	0.03	0.15	0.10	0.20	0.16	0.03	0.25	0.19	0.02
SA	0.18	0.11	0.16	0.14	0.17	0.18	0.08	0.19	0.14	0.06
n	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00

```
-----
```

Die Spalten bedeuten:

- 1 = V21 (=Geschmack), Auspr. = 1 (=Butter)
- 2 = V21 (=Geschmack), Auspr. = 2 (=pflanzlich)
- 3 = V22 (=Verwendung), Auspr. = 1 (=Brotaufstrich)
- 4 = V22 (=Verwendung), Auspr. = 2 (=Kochen/Backen/Braten)
- 5 = V22 (=Verwendung), Auspr. = 3 (=universell)
- 6 = V23 (=Kaloriengehalt), Auspr. 1 = 1.000
- 7 = V23 (=Kaloriengehalt), Auspr. 2 = 2.000
- 8 = V24 (=Preis), Auspr. 1 = 1.000
- 9 = V24 (=Preis), Auspr. 2 = 2.000
- 10 = V24 (=Preis), Auspr. 3 = 1.750

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA

Die normierten Teilnutzenwerte weisen im Durchschnitt (Zeile MW) - wie die absoluten Teilnutzenwerte (=Effekte) - eine hohe Standardabweichung (Zeile SA) auf. Bei ihrer Interpretation ist daher Vorsicht angebracht.

Aus den normierten Teilnutzenwerten wird im letzten Schritt der Analyse die relative Wichtigkeit der Variablen berechnet. Es ergeben sich folgende Werte:

Matrix der relativen Wichtigkeit der Variablen

=====

DA	1	2	3	4	SUM
1	0.50	0.28	0.08	0.14	1.00
2	0.04	0.33	0.02	0.61	1.00
3	0.14	0.31	0.03	0.52	1.00
4	0.03	0.22	0.17	0.58	1.00
5	0.46	0.18	0.13	0.23	1.00
6	0.00	0.35	0.44	0.21	1.00
7	0.29	0.11	0.46	0.14	1.00
8	0.12	0.27	0.07	0.54	1.00
9	0.04	0.54	0.07	0.36	1.00

usw.

MW	0.22	0.29	0.20	0.29	1.00
SA	0.18	0.15	0.16	0.17	0.00
n	40.00	40.00	40.00	40.00	40.00

Die Spalten bedeuten:

- 1 = V21 (=Geschmack)
- 2 = V22 (=Verwendung)
- 3 = V23 (=Kaloriengehalt)
- 4 = V24 (=Preis)

Die Zeilen bedeuten:

- DA = Datensatznummer
- MW = Durchschnitt (Mittelwert)
- SA = Standardabweichung
- n = Fallzahl zur Berechnung von MW und SA

Zur Erklärung der Präferenzbildung der Person 1 kommt dem Geschmack der Margarine die größte Bedeutung zu (relative Wichtigkeit von V21 = 0.50 bzw. 50%), dem Kaloriengehalt die geringste Bedeutung (relative Wichtigkeit von V23 = 0.08 bzw. 8%). Die Präferenzbildung der Person 2 wird vor allem durch den Preis (relative Wichtigkeit von V24 = 0.61 bzw. 61%) und die Verwendung (relative Wichtigkeit von V22 = 0.33 bzw. 33%) erklärt. Im Durchschnitt leisten alle Variablen in etwa denselben Beitrag zur Erklärung der Präferenzbildung der befragten Personen. Von etwas größerer Bedeutung sind die Verwendung und der Preis. Aber auch die durchschnittlichen relativen Wichtigkeiten weisen hohe Streuungen auf.

Zusammenfassend wird man die untersuchten Befragtengruppe als inhomogen bezeichnen. Es wurden Personen mit unterschiedlichen individuellen Nutzenfunktionen befragt. In einem nächsten Analyseschritt wird man daher versuchen, homogene Teilgruppen von Befragten zu bilden. Dazu erzeugt ALMO automatisch bei Verwendung von OPTION 8 ein ALMO-Programm für eine anschließende hierarchische Clusteranalyse (PROGRAMM P36). Die Daten werden in der nach dem Ist-Gleich-Zeichen von OPTION 8 spezifizierten Datei gespeichert. Welche Daten gespeichert werden, kann durch OPTION 18 gesteuert werden. Die maximale Clusterzahl für das anschließende Clusteranalyseprogramm kann durch OPTION 17 = ...; spezifiziert werden. In dem Beispiel wurde OPTION 18 auf 1 gesetzt. Dies führt dazu, daß die nominierten Teilnutzenwerte gespeichert werden. Da OPTION 17 = 6; gewählt wurde, wird im Clusteranalyseprogramm die maximale Clusterzahl gleich 6 gesetzt.

Reihenfolge, in der die Variablen in die Datei geschrieben werden:

```
V1 = fortlaufende Nummer
V2 = relativer Teilnutzen von V21 (=Geschmack), Auspr. = 1 (=Butter)
V3 = relativer Teilnutzen von V21 (=Geschmack), Auspr. = 2 (=pflanzlich)
V4 = relativer Teilnutzen von V22 (=Verwendung), Auspr. = 1 (=Brotaufstrich)
V5 = relativer Teilnutzen von V22 (=Verwendung), Auspr. = 2 (=Kochen/Backen
    /Braten)
V6 = relativer Teilnutzen von V22 (=Verwendung), Auspr. = 3 (=universell)
V7 = relativer Teilnutzen von V23 (=Kaloriengehalt), Auspr. 1 = 1.000
V8 = relativer Teilnutzen von V23 (=Kaloriengehalt), Auspr. 2 = 2.000
V9 = relativer Teilnutzen von V24 (=Preis), Auspr. 1 = 1.750
V10 = relativer Teilnutzen von V24 (=Preis), Auspr. 2 = 2.250
V11 = relativer Teilnutzen von V24 (=Preis), Auspr. 3 = 2.750
```

Beispiel fuer das Einlesen:

```
ANFANG
Name1=flnd.Nummer;
Name2=rT_Geschmack_Butter;
Name3=rT_Geschmack_pflanzlich;
Name4=rT_Verwendung_Brotaufstrich;
Name5=rT_Verwendung_Kochen/Backen/Braten;
Name6=rT_Verwendung_universell;
Name7=rT_Kaloriengehalt_1;
Name8=rT_Kaloriengehalt_2;
Name9=rT_Preis_1.75;
Name10=rT_Preis_2.25;
Name11=rT_Preis_2.75;
ENDE
```

```
ANFANG
    Programm = 36;
    A_Quantitative_V = V2:11;
    Modell = Ward_Linkage;
    Distanzmass = quad_euklid;
    Objekte = 40;
    Min_Clusterzahl = 2;
    Max_Clusterzahl = 6;
Ende_Programmparameter
```

```
Lese V1:11 aus Datei 2 'c:\almo6\progs\nutz.fre' Format frei
    leerzu Ende;
Gehe_in_Programm
Gehe_zu Lese
ENDE
```

```
***** MITTEILUNG
    Das ALMO-Programm wurde unter dem Namen 'c:\almo6\progs\nutz.alm'
    auf der Platte abgespeichert
```

Das ALMO-Programm zur Clusteranalyse wird in dem Beispiel unter dem Namen "C:\ALMO\PROGS\NUTZ.ALM" abgespeichert. Der Dateiname wird von ALMO automatisch aus der definierten Datendatei ("C:\ALMO\PROGS\NUTZ.FRE") erzeugt, indem die Namensweiterung durch "ALM" ersetzt wird.

Durchführen der Clusteranalyse erbringt in unserem Beispiel folgende Ergebnisse (Zur Interpretation der Ergebnisse der Clusteranalyse siehe Abschnitt P36.7): Das Mojena-I-Kriterium weist eine 10-Clusterlösung als signifikant aus, das Mojena-II-Kriterium eine 7-Clusterlösung. Die 7-Clusterlösung, die erst nach einem erneuten Programmdurchlauf berechnet wird, indem MIN_CLUSTERZAHL=7; und MAX_CLUSTERZAHL=7; gesetzt wird, sieht folgendermaßen aus:

Clusterzugehoerigkeit der Elemente bei 7 Clustern

```
-----
Cluster 1 (n= 6)    1
                   22
                   24
                   29
                   36
                   39
Cluster 2 (n= 8)    2
                   3
                   4
                   8
                   12
```

usw.

Masszahlen fuer Klassifikationsvariablen im Clustern 1:

Variable	n=	Min.	Max.	MA	SA	z-Wert
2 rTGeschm	6	0.07	0.50	0.29	0.14	1.69
3 rTGeschm	6	0.00	0.00	0.00	0.00	-99.99
4 rTVerwen	6	0.14	0.49	0.29	0.12	2.58
5 rTVerwen	6	0.00	0.00	0.00	0.00	-99.99
6 rTVerwen	6	0.03	0.58	0.36	0.19	1.96
7 rTKalori	6	0.00	0.09	0.04	0.04	-7.55
8 rTKalori	6	0.00	0.38	0.09	0.14	0.91
9 rTPreis1	6	0.00	0.14	0.06	0.06	-7.17
10 rTPreis2	6	0.00	0.11	0.05	0.04	-7.69
11 rTPreis2	6	0.00	0.31	0.10	0.10	1.60

Cluster 1 ist dadurch charakterisiert, daß der Verwendung als Brotaufstrich (=V4) ein überdurchschnittlich hoher Nutzen zugeschrieben wird (z-Wert von V4 ist größer 2!!).

Masszahlen fuer Klassifikationsvariablen im Clustern 2:

Variable	n=	Min.	Max.	MA	SA	z-Wert
2 rTGeschm	8	0.00	0.26	0.09	0.08	-3.21
3 rTGeschm	8	0.00	0.21	0.03	0.07	-0.31
4 rTVerwen	8	0.00	0.24	0.10	0.09	-1.58
5 rTVerwen	8	0.00	0.33	0.12	0.13	0.57
6 rTVerwen	8	0.00	0.31	0.16	0.10	-1.13
7 rTKalori	8	0.00	0.17	0.05	0.06	-5.09
8 rTKalori	8	0.00	0.08	0.01	0.03	-2.16
9 rTPreis1	8	0.40	0.68	0.55	0.08	10.43
10 rTPreis2	8	0.20	0.55	0.39	0.11	4.57
11 rTPreis2	8	0.00	0.00	0.00	0.00	-99.99

Im Unterschied zum Cluster C1 kommt im Cluster C2 dem Preis eine besondere Bedeutung zu. Preiswertere Margarineprodukte mit Preisen bis zu DM 2.25 (V9 und V10) werden bevorzugt. Die weiteren Cluster können analog interpretiert werden.

P44.6 Exkurs: Erstellen unvollständiger Designs

In der Forschungspraxis ist oft die Verwendung eines vollständigen Designs nicht möglich, da dadurch ein zu großer Erhebungsaufwand verbunden mit einer möglichen Überforderung des Befragten entstehen würde. Sollen z.B. vier Stimulusmerkmale mit jeweils vier Ausprägungen untersucht werden, würde ein vollständiges Design 256 Konfigurationen umfassen, bei fünf Stimulusmerkmalen mit ebenfalls jeweils vier Ausprägungen wären dies bereits 1024 Konfigurationen. In der Praxis ist man daher häufig auf die Anwendung nicht vollständiger, unvollständiges oder reduzierter Designs (alle drei Bezeichnungen werden synonym verwendet) angewiesen, sofern eine Dichotomisierung der Merkmale nicht möglich ist.

Bei der Konstruktion unvollständiger Designs ist die formale Grundregel zu beachten, daß in dem unvollständigen Design keine linearen Abhängigkeiten auftreten, daß also die Effekte aller untersuchten Ausprägungen geschätzt werden können. Treten lineare Abhängigkeiten auf, gibt ALMO eine Fehlermeldung ("Matrix nicht invertierbar") aus und bricht die Berechnung ab.

Eine Methode zur Konstruktion unvollständiger Designs besteht in der Verwendung von sogenannten Lateinischen Quadraten bzw. allgemein von Quadraten mit mehreren Alphabeten (Backhaus et al. 1994: 509, Cox 1958: 205-218). Diese sind allgemein dadurch gekennzeichnet, daß die untersuchten Merkmale unabhängig, also unkorreliert sind. Interaktionseffekte können allerdings nicht geschätzt werden. Voraussetzung für die Konstruktion Lateinischer Quadrate ist, daß alle Merkmale dieselbe Anzahl von Ausprägungen besitzen, also z.B. alle Merkmale 2, 3, 4 oder mehr Ausprägungen. Dies wird als symmetrischer Fall bezeichnet. Besitzen dagegen die Merkmale eine unterschiedliche Anzahl von Ausprägungen, z.B. Merkmal 1 zwei Ausprägungen, Merkmal 2 drei Ausprägungen usw., liegt ein asymmetrischer Fall vor. Unvollständige Designs für den asymmetrischen lassen sich allerdings durch Transformationen aus Lateinischen Quadraten bzw. allgemein aus Quadraten mit mehreren Alphabeten erzeugen. Die Orthogonalitätseigenschaft kann dabei aber mitunter verloren gehen. Nachfolgend soll die Konstruktion von Lateinischen Quadraten (Quadrate mit einem Alphabet), von Quadraten mit mehreren Alphabeten und das Erzeugen von unvollständigen Designs für den asymmetrischen Fall aus Lateinischen Quadraten bzw. allgemein aus Quadraten mit mehreren Alphabeten dargestellt werden.

Lateinische Quadrate (Quadrate mit einem Alphabet)

Lateinische Quadrate werden wie folgt gebildet (Cox 1958: 205-207):

1. Es wird eine Tabelle mit den Merkmalen A und B gebildet.
2. In die Zellen der Tabelle wird das Merkmal C eingetragen. Die Ausprägungen des Merkmals C werden dabei in jeder Zeile systematisch variiert: In der ersten Zeile wird mit C1 begonnen, in der zweiten mit C2, in der dritten mit C3 usw. (Auch andere systematische Variationen sind möglich!)

Nachfolgende Beispiele verdeutlichen die Konstruktion:

3 x 3 Lateinisches Quadrat

	B1	B2	B3
A1	C1	C2	C3
A2	C2	C3	C1
A3	C3	C1	C2

Stimuluskonfig.		
A1	B1	C1
A1	B2	C2
A1	B3	C3
A2	B1	C2
A2	B2	C3
A2	B3	C1
A3	B1	C3
A3	B2	C1
A3	B3	C2

Im Unterschied zu einem vollständigen Design müssen anstelle von 27 Stimuluskonfigurationen nur 9 erfragt werden.

4 x 4 Lateinisches Quadrat

	B1	B2	B3	B4	Stimuluskonfig.			
A1	C1	C2	C3	C4	A1	B1	C1	Im Unterschied zu einem vollständigen Design müssen anstelle von 64 Stimuluskonfigurationen nur 16 erfragt werden.
A2	C2	C3	C4	C1	A1	B2	C2	
A3	C3	C4	C1	C2	A1	B3	C3	
A4	C4	C1	C2	C3	A1	B4	C4	
					A2	B1	C2	
					A2	B2	C3	
					usw.			
					A4	B4	C3	

6 x 6 Lateinisches Quadrat

	B1	B2	B3	B4	B5	B6	Stimuluskonfig.			
A1	C1	C2	C3	C4	C5	C6	A1	B1	C1	Im Unterschied zu einem vollständigen Design müssen anstelle von 216 Stimuluskonfigurationen nur 36 erfragt werden.
A2	C2	C3	C4	C5	C6	C1	A1	B2	C2	
A3	C3	C4	C5	C6	C1	C2	A1	B3	C3	
A4	C4	C5	C6	C1	C2	C3	A1	B4	C4	
A5	C5	C6	C1	C2	C3	C4	A1	B5	C5	
A6	C6	C1	C2	C3	C4	C5	A1	B6	C6	
							usw.			
							A6	B6	C5	

Griechisch-Lateinische Quadrate (Quadrate mit zwei Alphabeten)

Liegen an Stelle von drei Merkmalen vier vor, lässt sich die Logik der Lateinischen Quadrate erweitern. Die dabei entstehenden Versuchspläne werden als Griechisch-Lateinische Quadrate bzw. von Quadraten mit zwei Alphabeten bezeichnet. Das Konstruktionsprinzip besteht aus folgenden Regeln:

1. Es wird eine Tabelle mit den Merkmalen A und B gebildet.
2. In die Zellen der Tabelle wird das Merkmal C eingetragen. Die Ausprägungen des Merkmals C werden in jeder Zeile systematisch variiert: In der ersten Zeile wird mit C1 begonnen, in der zweiten mit C2, in der dritten mit C3 usw. (Auch andere systematische Variationen sind möglich!)
3. In die Zellen der Tabelle wird das Merkmal D eingetragen. Die Ausprägungen des Merkmals D werden in jeder Zeile systematisch variiert, allerdings nach einem anderen Prinzip als das Merkmal C. In der ersten Zeile wird mit D1 begonnen, in der zweiten Zeile allerdings nicht mit D2, sondern mit D3 oder D4.

Beispiele sollen wiederum die Konstruktion verdeutlichen:

3 x 3 Griechisch-Lateinisches Quadrat

	B1	B2	B3	Stimuluskonfiguratur.				
A1	C1D1	C2D2	C3D3	A1	B1	C1	D1	Im Unterschied zu einem vollständigen Design müssen anstelle von 27 Stimuluskonfigurationen nur 9 erfragt werden.
A2	C2D3	C3D1	C1D2	A1	B2	C2	D2	
A3	C3D2	C1D3	C2D1	A1	B3	C3	D3	
				A2	B1	C2	D3	
				A2	B2	C3	D1	
				A2	B3	C1	D2	
				A3	B1	C3	D2	
				A3	B2	C1	D3	
				A3	B3	C2	D1	

4 x 4 Griechisch-Lateinisches Quadrat

	B1	B2	B3	B4	Stimuluskonfig.				
A1	C1D1	C2D2	C3D3	C4D4	A1	B1	C1	D1	Im Unterschied zu einem vollständigen Design müssen anstelle von 256 Stimuluskonfigurationen nur 16 erfragt werden.
A2	C2D4	C3D1	C4D2	C1D3	A1	B2	C2	D2	
A3	C3D3	C4D4	C1D1	C2D2	A1	B3	C3	D3	
A4	C4D2	C1D3	C2D4	C3D1	A1	B4	C4	D4	
					A2	B1	C2	D4	
					A2	B2	C3	D1	
					usw.				
					A4	B4	C3	D1	

Quadrate mit mehr als zwei Alphabeten

Das Prinzip kann auch auf mehr als vier Variablen verallgemeinert werden. Allgemein ist folgende Regel zu beachten (Cox 1958: 212-213): Besitzen die Variablen m Ausprägungen, so können maximal aus $m+1$ Variablen nach dem Prinzip der Lateinischen Quadrate Designs gebildet werden. Bei drei Ausprägungen können also aus maximal 4 Variablen (mit jeweils drei Ausprägungen) "Lateinische Quadrate" (=Quadrate mit zwei Alphabeten) gebildet werden, bei vier Ausprägungen aus maximal 5 Variablen (=Quadrate mit drei Alphabeten) usw.

Erstellen von unvollständigen Designs für den asymmetrischen Fall aus Lateinischen Quadraten bzw. allgemein aus Quadraten mit mehreren Alphabeten.

Das Vorgehen (Backhaus et al. 1994: 524-526) besteht aus folgenden Schritten:

1. Bilde ein Lateinisches Quadrat bzw. ein Quadrat mit mehreren Alphabeten mit der höchsten Ausprägungszahl
2. Definiere Zusammenfassungenregeln, z.B. "D1 und D2 sollen zu D1 werden"
3. Führe die Zusammenfassungen durch.

Das Vorgehen soll anhand eines Beispiels verdeutlicht werden: Für die Variablen A (3 Ausprägungen), B (3 Ausprägungen), C (2 Ausprägungen) und D (4 Ausprägungen) soll ein unvollständiges Design nach der Logik der Lateinischen Quadrate gebildet werden. Die höchste Ausprägungszahl ist vier. Wir bilden daher ein 4 x 4 Griechisch-Lateinisches Quadrat (Quadrat mit zwei Alphabeten):

	B1	B2	B3	B4	Stimuluskonfig.			
A1	C1D1	C2D2	C3D3	C4D4	A1	B1	C1	D1
A2	C2D4	C3D1	C4D2	C1D3	A1	B2	C2	D2
A3	C3D3	C4D4	C1D1	C2D2	A1	B3	C1	D3
A4	C4D2	C1D3	C2D4	C3D1	A1	B1	C2	D4
					A2	B1	C2	D4
					A2	B2	C1	D1
					A2	B3	C2	D2
					A2	B1	C1	D3
					A3	B1	C1	D3
					A3	B2	C2	D4
					A3	B3	C1	D1
					A3	B1	C2	D2
					A1	B1	C2	D2
					A1	B2	C1	D3
					A1	B3	C2	D4
					A1	B1	C1	D1

Folgende Umkodierregeln sollen angewendet werden:
 A4 soll A1 werden.
 B4 soll B1 werden
 C3 soll C1 werden
 C4 soll C2 werden

Daraus ergeben sich nebenstehende Stimuluskonfigurationen:

Literatur

Backhaus, K., Erichson, B., Plinke, W., Weiber, R., 1994: Multivariate Analyseverfahren. Eine anwendungsorientierte Einführung. 7. Auflage. Berlin - Heidelberg - New York.

Cox, D.R., 1958: Planing of Experiments. New York - London - Sydney.

Green, P.E., Tull, D.S., 1982: Methoden und Techniken der Marktforschung. 4. Auflage. Stuttgart

SPSS Inc., 1990: SPSS Categories. Chicago.