



Beliebig-dimensionale Tabellierung

P11

Kurt Holm

Almo Statistik-System
www.almo-statistik.de
holm@almo-statistik.de
kurt.holm@jku.at

Autor: em. Prof. Dr. Kurt Holm, Universität Linz, Österreich

Siehe auch die beiden Almo-Dokumente Nr. 27 "Eindimensionale Häufigkeitsverteilung" und Nr. 1 "Zwei- und drei-dimensionale Tabellierung".

Im Text wird häufig auf das Dokument **P0** Bezug genommen. Dabei handelt es sich um das Almo-Dokument "Arbeiten mit Almo.PDF" (Dokument 0).

Weitere Almo-Dokumente

Die folgenden Dokumente können alle kostenlos von der Handbuchseite in www.almo-statistik.de heruntergeladen werden. Alle Dokumente befinden sich im PDF-Format.

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

INHALTSVERZEICHNIS

P11 PROGRAMM 11: Beliebige-dimensionale Tabellierung	4
P11.1 Eingabe in Programm-Maske Prog11m	6
P11.2 Programm-Maske mit Optionen Prog11m1	19
P11.4 Ausgabe.....	25
P11.4.1 Ausgabe für 2-dimensionale Tabelle	25
P11.4.2 Ausgabe für Mehrfach-Tabelle	28
P11.4.3 Ausgabe für Partial- bzw. Interaktions-Tabelle	31
P11.4.4 Ausgabe für multidimensionale Kontingenztabelle.....	34
P11.6 Eingabe einer fertigen "multidimensionalen Kontingenztabelle" mit Prog11m4....	38
P11.6.1 Erläuterungen zu den Boxen	42
P11.7 Der "OLAP-Würfel"	44
P11.7.1 Programm-Maske zum Erstellen eines "OLAP-Würfels"	46
P11.7.2 Die Auswertung des "OLAP-Würfels" mit Prog11m3	49
P11.7.3 Ergebnisse aus Prog11m3	54
Literatur	54

Das vorliegende Dokument besteht aus überarbeiteten Auszügen aus dem Almo-Handbuch „Teil 3b: Grundlegende Verfahren“. Querverweise auf den Abschnitte P10 beziehen sich das Almo-Dokument Nr. 1 "Zwei- und drei-dimensionale Tabellierung". Querverweise auf den Abschnitt P0 beziehen sich auf das Almo-Dokument Nr. 0 "Arbeiten mit Almo“

P11 PROGRAMM 11: Beliebige-dimensionale Tabellierung

Programm 11 wurde von Kurt Holm geschrieben. Es verwendet zwei Unterprogramme (Einlesen der Daten und Berechnung des Chi-Quadrat-Wertes für 2-dimensionale Tabellen), die von Hermann Denz programmiert wurden.

Programm 11 erstellt:

1. Zweidimensionale Tabellen (wie in Prog 10).

Beispiel:

	Schulbildung	
	niedrig	hoch
Geschl. männlich	16	34
weiblich	8	11

Die Zahl der Ausprägungen der beiden Variablen ist beliebig

2. Zweidimensionale Mehrfachtabellen.

Beispiel:

	Schulbildung	
	niedrig	hoch
Geschl. männlich	16	34
weiblich	8	11

Wohnort	Stadt	Land
	12	15
	12	30

Beruf	Arbeiter	Angestel
	12	12
	12	33

Diese entstehen dadurch, dass verschiedene unabhängige Variable jeweils gegen dieselbe abhängige Variable tabelliert werden. Berechnet wird der Chi-Quadrat-Test, der Kontingenzkoeffizient C (cor) und Cramers V.
Die Zahl der Ausprägungen der Variablen ist beliebig

3. Interaktionstabellen bzw. Partialtabellen .

Beispiel:

Geschl	Ort	Beruf	Schulbildung		Summe
			niedrig	hoch	
männl	Stadt	Arbeit	4	4	8
		Angest	4	4	8
	Land	Arbeit	4	4	8
		Angest	4	22	26
weibl	Stadt	Arbeit	2	2	4
		Angest	2	5	7
	Land	Arbeit	2	2	4
		Angest	2	2	4
Summe			24	45	69

Die Interaktion mehrerer Variablen wird als unabhängige Variable betrachtet, die gegen eine abhängige Variable tabelliert wird. Gerechnet wird der Chi-Quadrat-Test, C (cor) und Cramers V.

Die Zahl der Ausprägungen der Variablen ist beliebig

4. Beliebige-dimensionale Kontingenztabelle.

Beispiel:

Geschl	Ort	Beruf	Fälle
männl	Stadt	Arbeit	8
		Angest	8
	Land	Arbeit	8
		Angest	26
weibl	Stadt	Arbeit	4
		Angest	7
	Land	Arbeit	4
		Angest	4
Summe			69

Beliebig viele Variable werden kombiniert. Für ihre Kombination wird die Zahl der Fälle ausgezählt. Die Zahl der Ausprägungen der Variablen ist beliebig
Gerechnet werden

- Chi-Quadrat-Test auf allseitige Abhängigkeit der an der Kontingenztabelle beteiligten Variablen.
- 2I-Test auf allseitige Abhängigkeit
- Multidimensionale Konfigurationsfrequenzanalyse (nach Lienert) mit exaktem

Binomialtest

Weitere Tabellen-Typen werden in Almo-Dokument Nr.1 "Zwei- und drei-dimensionale Tabellierung" dargestellt.

P11.1 Eingabe in Programm-Maske Prog11m

**Prog11m.Msk
Tabellierung**

1. 2-dimensionale Tabelle
2. 2-dimensionale Mehrfach-Tabelle
3. Partial- bzw. Interaktions-Tabelle
4. multidimensionale Kontingenztabelle

Berechnet werden folgende Koeffizienten:
Chi-Quadrat, Chi-Quadrat-Test und 2I-Test auf allseitige
Abhängigkeit, Kontingenzkoeffizient C(cor)

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

1
Vereinbare Variable= ;

2 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3
 "C:\Almo7\Testdat\Uarnamen.nam"
 zeige = Namensdatei in Output zeigen
leer = nicht

4

 erzeuge zusätzliche Namensfelder

5 bei Datei-Problemen
 "C:\Almo7\TESTDAT\TESTDAT.FRE"
 Format der Daten
 der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_U

6 Wenn Dateiformat FIX oder Nicht-Standard-FREI

7
Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11

erzeuge zusätzliche Felder für Tabellen-Angaben

8

2-dimensionale "Mehrfach-Tabelle" **Hilfe**

BEACHTIE: Nur 1 Spaltenvariable erlaubt

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11
Wohnort	Stadt	12	15
	Land	12	30
Beruf	Arbeiter	12	12
	Angestel	12	33

Geschlecht + Wohnort + Beruf mit Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

9

Partielltabelle
Interaktionstabelle **Hilfe**

BEACHTIE: Nur 1 Spaltenvariable erlaubt

Beispiel:

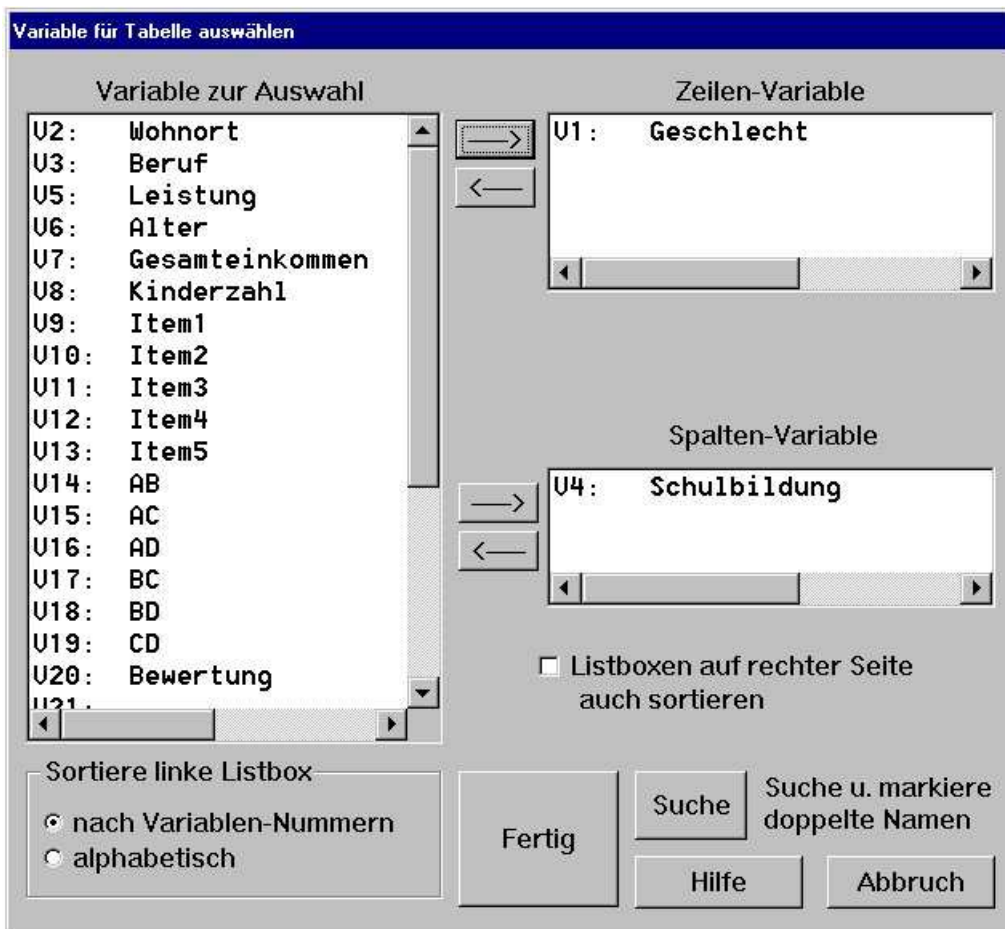
Geschl	Ort	Beruf	Schulbildung		Summe
			niedrig	hoch	
männl	Stadt	Arbeit	4	4	8
		Angest	4	4	8
	Land	Arbeit	4	4	8
		Angest	4	22	26
weibl	Stadt	Arbeit	2	2	4
		Angest	2	5	7
	Land	Arbeit	2	2	4
		Angest	2	2	4
Summe			24	45	69

Geschlecht * Wohnort * Beruf mit Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

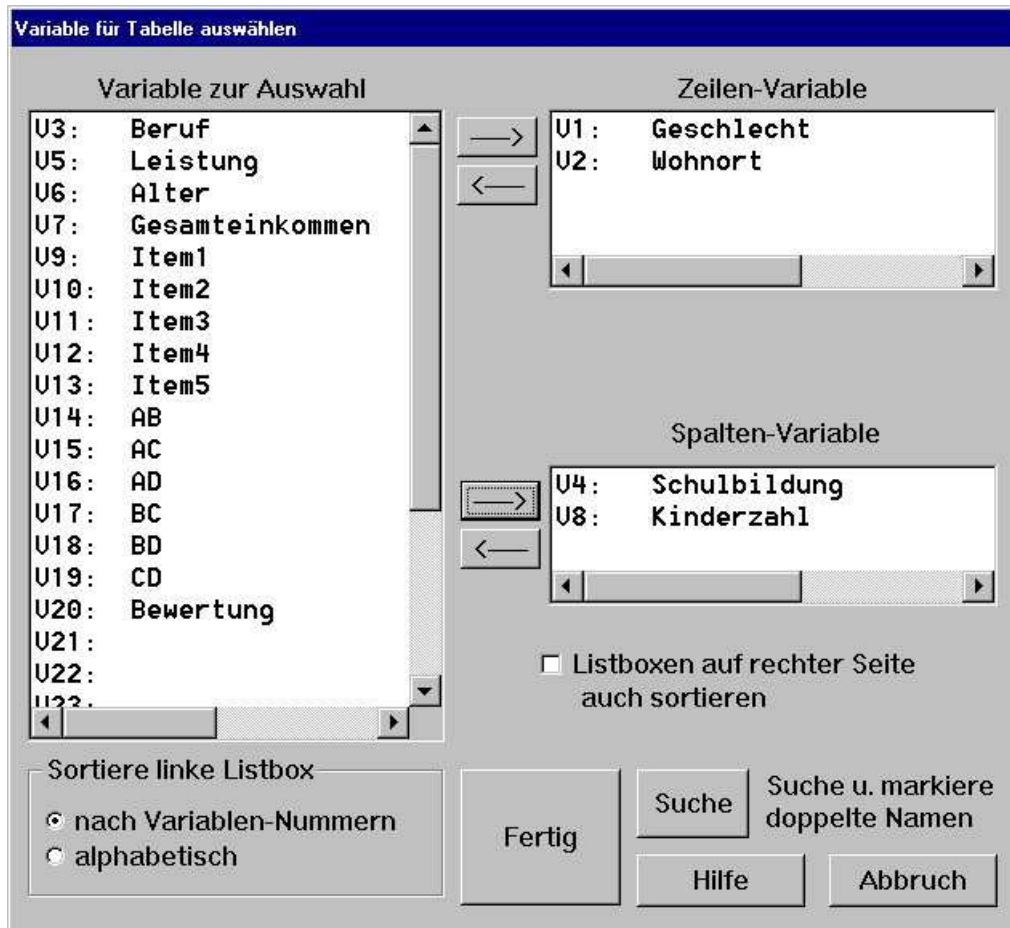


Wenn Sie auf den Knopf mit den 2 Fenstersymbolen klicken, dann wird die "Variablen-Auswahl-Box" geöffnet. In Ihr geben Sie an, welche Variable als Zeilenvariable und welche als Spaltenvariable die Tabelle bilden sollen.



Almo bildet die Tabelle "Geschlecht mit Schulbildung"

Es können auch mehrere 2-dimensionale Tabellen gleichzeitig angegeben werden.
Beispiel:



Almo bildet die Tabellen

	Geschlecht mit Schulbildung
	Geschlecht mit Kinderzahl
	Wohnort mit Schulbildung
	Wohnort mit Kinderzahl

Aus den beiden Beispielen wird ersichtlich:

Alle Variable in der Box "Zeilenvariable" werden mit allen Variablen aus der Box "Spaltenvariable" gepaart. Für jedes Paar wird eine 2-dimensionale Tabelle gebildet.

Sie können auch direkt in das Eingabefeld hineinschreiben.

Wenn Sie z.B. schreiben:



dann ist V1 die Zeilen- und V2 die Spalten-Variable. V1 steht vorne in der Tabelle und V2 oben rüber in der Tabelle.

Sie müssen auch nicht unbedingt Variablennamen verwenden. Sie können auch, wie in obigem Beispiel auch nur Variablennummern (mit einem V davor) schreiben.

BEACHTTE:

Sie müssen dann in Box 11 auf den Knopf "Variable aus Tabellenangaben" klicken. Almo registriert dann die Variablen, die zur Tabellenbildung verwendet werden

Sie können in das Eingabefeld auch 2 oder mehrere Tabellenangaben schreiben, z.B. so



Zwischen die Tabellenangaben muss also ein Schrägstrich geschrieben werden.

Vor und hinter "mit" können mehrere (durch Beistrich getrennte) Variable stehen. Folgende Angaben für 2-dimensionale Tabellen sind beispielsweise möglich:

V1	mit	V3	tabelliert wird:	V1 mit V3
V1,2	mit	V3	tabelliert wird:	V1 mit V3, V2 mit V3
V1	mit	V3,4	tabelliert wird:	V1 mit V3, V1 mit V4
V1,2	mit	V3,4	tabelliert wird:	V1 mit V3, V1 mit V4 V2 mit V3, V2 mit V4

P11.1.3 Box 8: 2-dimensionale Mehrfachtabellen



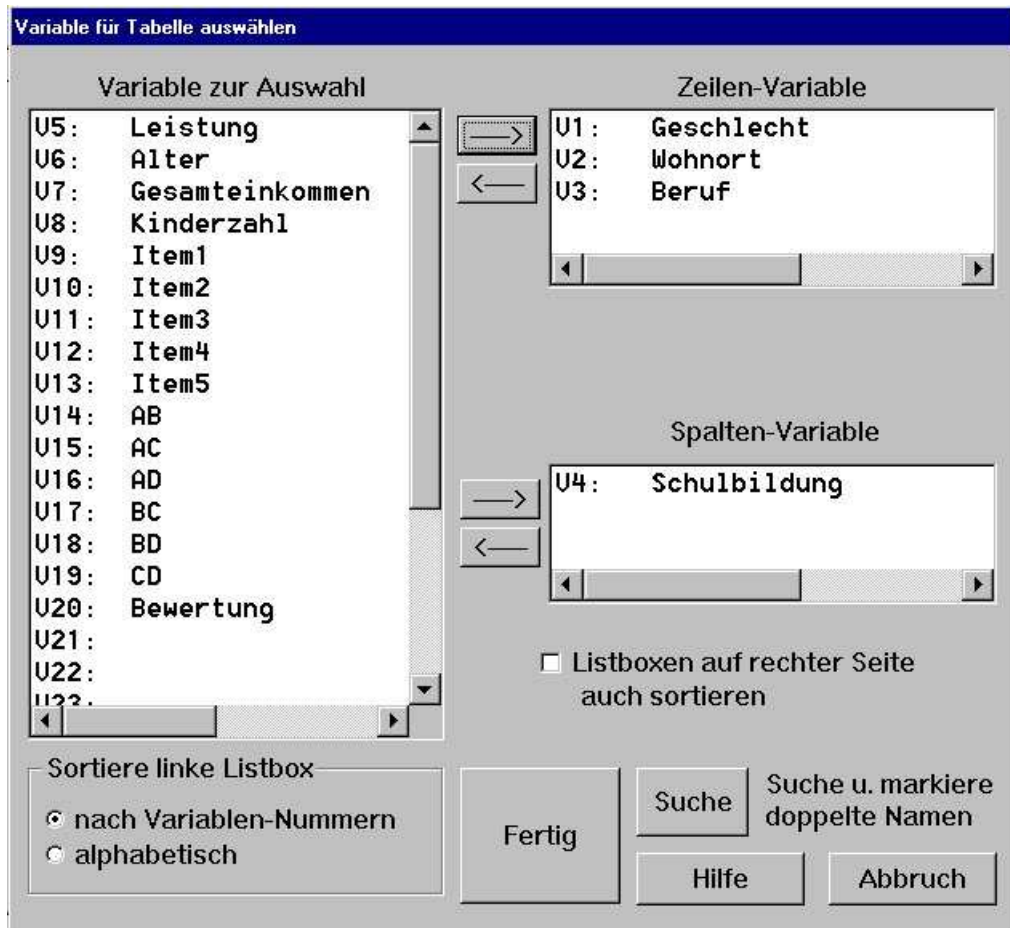
Eine Mehrfachtafel besteht aus beliebig vielen 2-dimensionalen Tabellen - mit der Besonderheit, dass die Spaltenvariable immer dieselbe ist und die Zeilenvariablen jeweils andere sind.

Es ist also nur 1 Spaltenvariable, aber beliebig viele Zeilenvariablen erlaubt.

Die Spaltenvariable kann im Prinzip beliebig viele Ausprägungen besitzen. Die Übersichtlichkeit geht jedoch verloren, wenn so viele Ausprägungen vorhanden sind, dass Almo die Tabelle umbrechen muss. Bei den Optionen in Box 13, die das "Aussehen" der auszugebenden Tabelle steuern, besteht die Möglichkeit, die Breite der Ausgabe einzustellen. Wenn Sie hier beispielsweise 120 eingeben, dann wird die Tabelle mit einer Breite von 120 Zeichen dargestellt. Ist sie breiter, dann muss Almo umbrechen.

Wenn Sie auf den Knopf mit den 2 Fenstersymbolen klicken, dann wird die "Variablen-Auswahl-Box" geöffnet. In Ihr geben Sie an, welche Variable als Zeilenvariable und welche als Spaltenvariable die Tabelle bilden sollen.

Beispiel:



Almo bildet die Mehrfachtablelle

V1 + V2 + V3 MIT V8
 Geschlecht + Wohnort + Beruf MIT Kinderzahl

Die Zahl der Zeilenvariablen muss mindestens zwei sein. Die Spaltenvariable darf nur eine sein.

Sie können auch direkt in das Eingabefeld hineinschreiben.
 Wenn Sie z.B. schreiben:



dann sind V1,V2,V3 die Zeilen- und V8 die Spalten-Variable.

Sie müssen auch nicht unbedingt Variablennamen verwenden. Sie können auch, wie in obigem Beispiel auch nur Variablennummern (mit einem V davor) schreiben.

BEACHTEN:

Sie müssen dann in Box 11 auf den Knopf "Variable aus Tabellenangaben" klicken. Almo registriert dann die Variablen, die zur Tabellenbildung verwendet werden.

Sie können in das Eingabefeld auch 2 oder mehrere Tabellenangaben schreiben, z.B. so



Zwischen die Tabellenangaben muss also ein Schrägstrich geschrieben werden.

P11.1.4 Box 9: Partialtabellen bzw. Interaktionstabellen

Partialtabellen Hilfe

Interaktionstabelle

BEACHTEN: Nur 1 Spaltenvariable erlaubt

Beispiel:

			Schulbildung		Summe
			niedrig	hoch	
Geschl	Ort	Arbeit	4	4	8
		Angest	4	4	8
männl	Stadt	4	4	8	
	Land	4	22	26	
weibl	Stadt	2	2	4	
	Land	2	2	4	
Summe			24	45	69

Geschlecht * Wohnort * Beruf mit Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

Betrachten wir folgendes Beispiel:

Rückzahlungsbereitschaft im Versandhandel
Die Tabelle ist zeilenweise prozentuiert

Geschlecht Produkt		Rückzahlung		Summe	
		nein	ja		
männlich	Kleidung	55	45	100	Partialtabelle für die Männer
	Möbel	29	71	100	
	Technik	23	77	100	
weiblich	Kleidung	35	65	100	Partialtabelle für die Frauen
	Möbel	32	68	100	
	Technik	20	80	100	
Summe		35	65	100	

In dieser Tabelle wird untersucht, wie das "Produkt", das auf Kredit im Versandgeschäft gekauft wurde, die Bereitschaft zur "Rückzahlung" (des Kredits) bestimmt - dabei betrachten wir Männer und Frauen getrennt.

Die Tabelle besteht also aus 2 "Partialtabellen", aus einer Partialtabelle für die Männer und einer für die Frauen.

Die Variable des Geschlechts, für die die Partialtabellen gebildet werden, wird häufig als "Kontrollvariable" bezeichnet.

Wir bezeichnen die Variablen, die die Zeilen der Tabelle bilden auch als "interagierende" Zeilenvariable.

Es können beliebig viele "interagierende" Zeilenvariable eingesetzt werden. D.h. vor

dem Almo-Schlüsselwort "mit" können beliebig viele Variable angegeben werden. Hinter "mit" darf nur eine stehen.

Wir erkennen, dass die Rückzahlungsbereitschaft bei Kleidung am schlechtesten und bei technischen Produkten am besten ist. Möbel liegen in der Mitte. Dabei ist bei weiblichen Kunden die Rückzahlungsbereitschaft insgesamt besser als bei Männern.

Auffallend ist, dass die Rückzahlungsbereitschaft bei Männern besonders schlecht ist, wenn das (auf Kredit) gekaufte Produkt Kleidung ist.

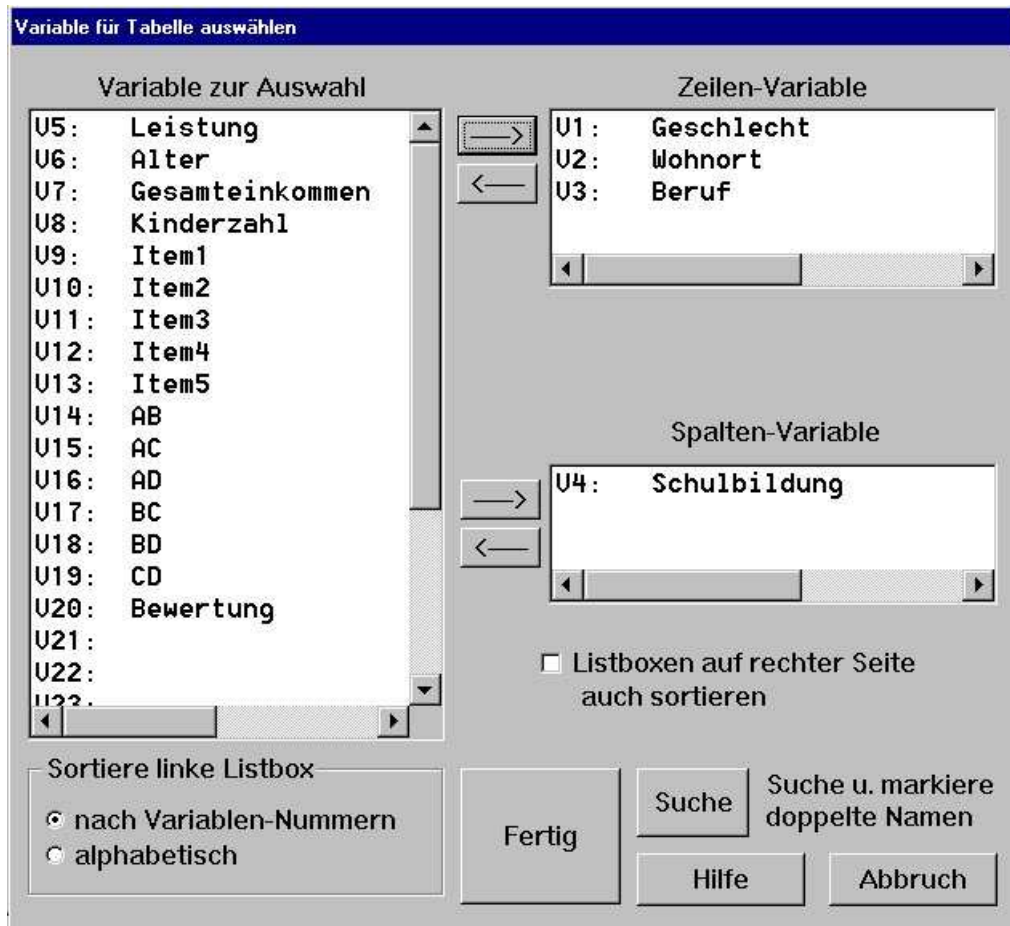
Wir können auch sagen, dass wir untersuchen, wie die "Interaktion" von Geschlecht und Produkt die Rückzahlungsbereitschaft bestimmt.

Deswegen bezeichnen wir diese Tabelle auch als Interaktionstabelle.

Die Spaltenvariable in der Partial- bzw. Interaktionstabelle kann im Prinzip beliebig viele Ausprägungen besitzen. Die Übersichtlichkeit geht jedoch verloren, wenn so viele Ausprägungen vorhanden sind, dass Almo die Tabelle umbrechen muss. Bei den Optionen in Box 13, die das "Aussehen" der auszugebenden Tabelle steuern (siehe P0.9) besteht die Möglichkeit, die Breite der Ausgabe einzustellen. Wenn Sie hier beispielsweise 120 eingeben, dann wird die Tabelle mit einer Breite von 120 Zeichen dargestellt. Ist sie breiter, dann muss Almo umbrechen.

Wenn Sie auf den Knopf mit den 2 Fenstersymbolen klicken, dann wird die "Variablen-Auswahl-Box" geöffnet. In Ihr geben Sie an, welche Variable als (interagierende) Zeilenvariable und welche als Spaltenvariable die Tabelle bilden sollen.

Beispiel:



Almo bildet die Partial- bzw. Interaktionstabelle

V1 * V2 * V3 MIT V8
 Geschlecht * Wohnort * Beruf MIT Kinderzahl

Die Zahl der Zeilenvariablen muss mindestens zwei sein. Die Spaltenvariable darf nur eine sein.

Sie können auch direkt in das Eingabefeld hineinschreiben.
 Wenn Sie z.B. schreiben:



dann sind V1,V2,V3 die interagierenden Zeilenvariablen, die durch ein Multiplikations-zeichen zu verbinden sind und V8 die Spaltenvariable. Sie müssen auch nicht unbedingt Variablennamen verwenden. Sie können auch, wie in obigem Beispiel nur Variablennummern (mit einem V davor) schreiben.

BEACHTTE:

Sie müssen dann in Box 11 auf den Knopf "Variable aus Tabellenangaben" klicken. Almo registriert dann die Variablen, die zur Tabellenbildung verwendet werden.

Sie können in das Eingabefeld auch 2 oder mehrere Tabellenangaben schreiben, z.B. so



Zwischen die Tabellenangaben muss also ein Schrägstrich geschrieben werden.

P11.1.5 Box 10: Multidimensionale Kontingenztabelle

Multidimensionale Kontingenztabelle Hilfe

Beispiel:

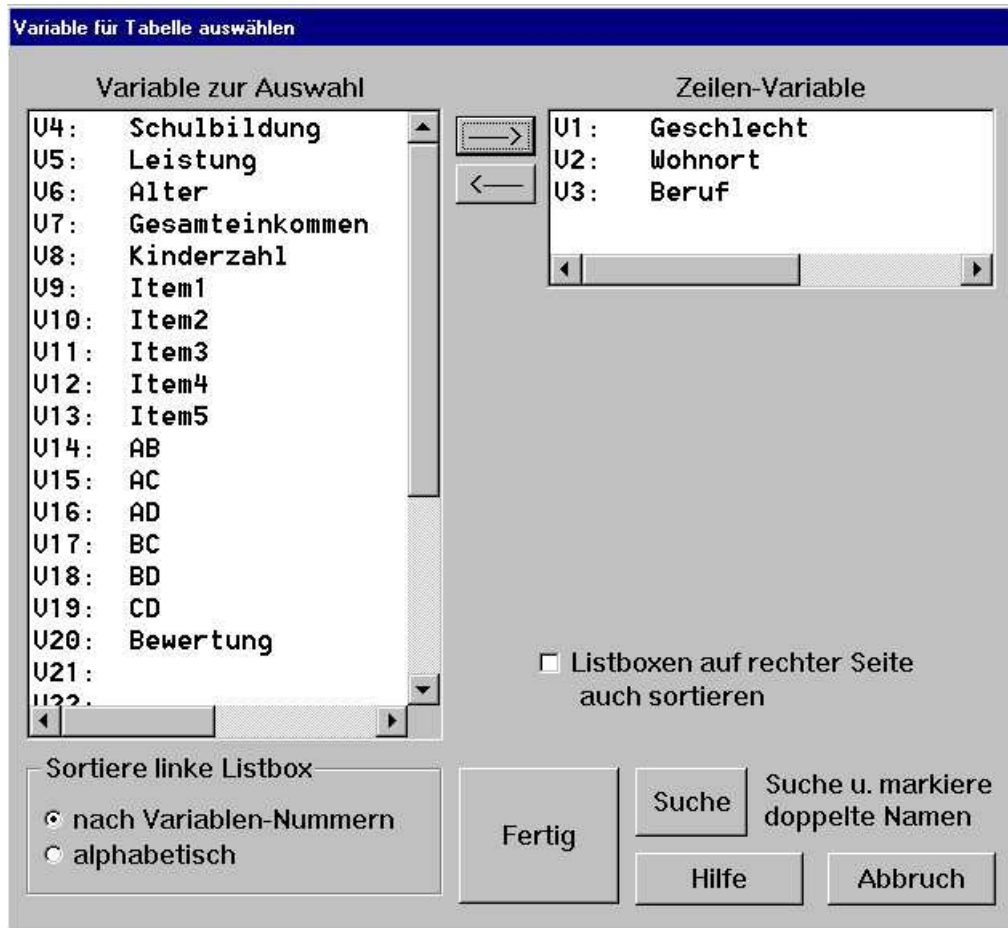
Geschl	Ort	Beruf	Fälle
männl	Stadt	Arbeit	8
		Angest	8
	Land	Arbeit	8
		Angest	26
weibl	Stadt	Arbeit	4
		Angest	7
	Land	Arbeit	4
		Angest	4
Summe			69

BEACHTTE: Kein "mit" vorhanden

An der Beispieltabelle, die in der Box abgebildet wird, ist ersichtlich, was wir unter einer "multidimensionale Kontingenztabelle" verstehen: Zwei oder mehrere Variable werden kombiniert. Dann zählt Almo aus, wieviel Fälle es in den Daten für die jeweilige Kombination gefunden hat.

Wenn Sie auf den Knopf mit den 2 Fenstersymbolen klicken, dann wird die "Variablen-Auswahl-Box" geöffnet. In Ihr geben Sie an, welche Variable kombiniert werden sollen.

Beispiel:



Almo bildet die multidimensionale Kontingenztabelle

V1 * V2 * V3
 Geschlecht * Wohnort * Beruf

Die Zahl der Zeilenvariablen muss mindestens zwei sein. Spaltenvariable gibt es keine

Sie können auch direkt in das Eingabefeld hineinschreiben.
 Wenn Sie z.B. schreiben:



dann sind V1,V2,V3 die zu kombinierenden Variablen.

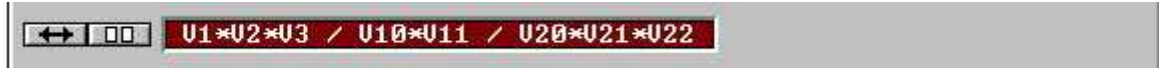
Sie müssen auch nicht unbedingt Variablennamen verwenden. Sie können auch, wie in obigem Beispiel auch nur Variablennummern (mit einem V davor) schreiben.

BEACHTTE: Es gibt kein "mit".

BEACHTTE:

Sie müssen dann in Box 11 auf den Knopf "Variable aus Tabellenangaben" klicken. Almo registriert dann die Variablen, die zur Tabellenbildung verwendet werden.

Sie können in das Eingabefeld auch 2 oder mehrere Tabellenangaben schreiben, z.B. so



Zwischen die Tabellenangaben muss also ein Schrägstrich geschrieben werden.

Box 11: Variable für Tabellen

Siehe Erläuterungen zu Prog10m1.Msk, P10.0.4.

Box 12: Kein_Wert-Angabe und Umkodierung, siehe P0.5 und P10.0.5.

Box 13: Optionen, die das „Aussehen“ der Tabellen steuern

Sie können das Erscheinungsbild der Tabellen in der Almo-Ausgabe beeinflussen. Siehe dazu P0.9.

Box 14: Grafik-Optionen

Siehe dazu P0.10.

P11.2 Programm-Maske mit Optionen Prog11m1

Prog1m1.Msk
Tabellierung
mit Optionen

1. 2-dimensionale Tabelle
2. 2-dimensionale Mehrfach-Tabelle
3. Partial- bzw. Interaktions-Tabelle
4. multidimensionale Kontingenztabelle
(optional mit Konfigurationsfrequenzanalyse)

Berechnet werden folgende Koeffizienten:
Chi-Quadrat, Chi-Quadrat-Test und 2I-Test auf allseitige Abhängigkeit
Kontingenzkoeffizient C(cor)

Zu 4. kann die Tabelle der Erwartungswerte zum Chi-Quadrat und
die Tabelle der Chi-Quadrat-Beiträge ermittelt werden.

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

1

Vereinbare Variable= ;

2 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

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

 zeige = Namensdatei in Output zeigen
leer = nicht

4

5

bei Datei-Problemen

"C:\Almo7\TESTDAT\TESTDAT.FRE"

 Format der Daten

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

6 Wenn Dateiformat FIX oder Nicht-Standard-FREI

7

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11

2-dimensionale "Mehrfach-Tabelle"

Hilfe

BEACHTEN: Nur 1 Spaltenvariable erlaubt

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11
Wohnort	Stadt	12	15
	Land	12	30
Beruf	Arbeiter	12	12
	Angestel	12	33

↔ **Geschlecht + Wohnort + Beruf mit Schulbildung**

erzeuge zusätzliche Felder für Tabellen-Angaben

8

Partialtabellen / Interaktionstabelle

Hilfe

BEACHTEN: Nur 1 Spaltenvariable erlaubt

Beispiel:

			Schulbildung		Summe
			niedrig	hoch	
Geschl	männl	Stadt Arbeit	4	4	8
		Stadt Angest	4	4	8
	Land	Arbeit	4	4	8
		Angest	4	22	26
weibl	Stadt	Arbeit	2	2	4
		Stadt Angest	2	5	7
	Land	Arbeit	2	2	4
		Land Angest	2	2	4
Summe			24	45	69

↔ **Geschlecht * Wohnort * Beruf mit Schulbildung**

erzeuge zusätzliche Felder für Tabellen-Angaben

9

Multidimensionale Kontingenztabelle **Hilfe**

Beispiel:

			Fälle
Geschl	Ort	Beruf	
männl	Stadt	Arbeit	8
		Angest	8
Land	Land	Arbeit	8
		Angest	26
weibl	Stadt	Arbeit	4
		Angest	7
Land	Land	Arbeit	4
		Angest	4
Summe			69

BEACHTE: Kein "mit" vorhanden

Geschlecht * Wohnort * Beruf * Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

10

Variable aus Tabellenangaben

<----- keine Benutzereingabe
 Also ermittelt die Variablen, die
 in obige Tabellen eingegangen sind
 selbst. Sie können aber auch auf
 diesen Knopf klicken

11

Option: Nur für multidimensionale Kontingenztabelle

12

Option: Ein- und Ausschliessen von Untersuchungseinheiten

13

Option: Umkodierungen und Kein-Wert-Angaben

14

Option: Untersuchungseinheiten gewichten

15

Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

16

Grafik-Optionen

17

Programmende

18

Erläuterungen zu den Optionen:

Gegenüber dem vorausgehend dargestellten Programm enthält Progl1m1 noch zusätzlich folgende drei Boxen:

Box 13: Ein- und Ausschließen von Untersuchungseinheiten, siehe dazu Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.7.

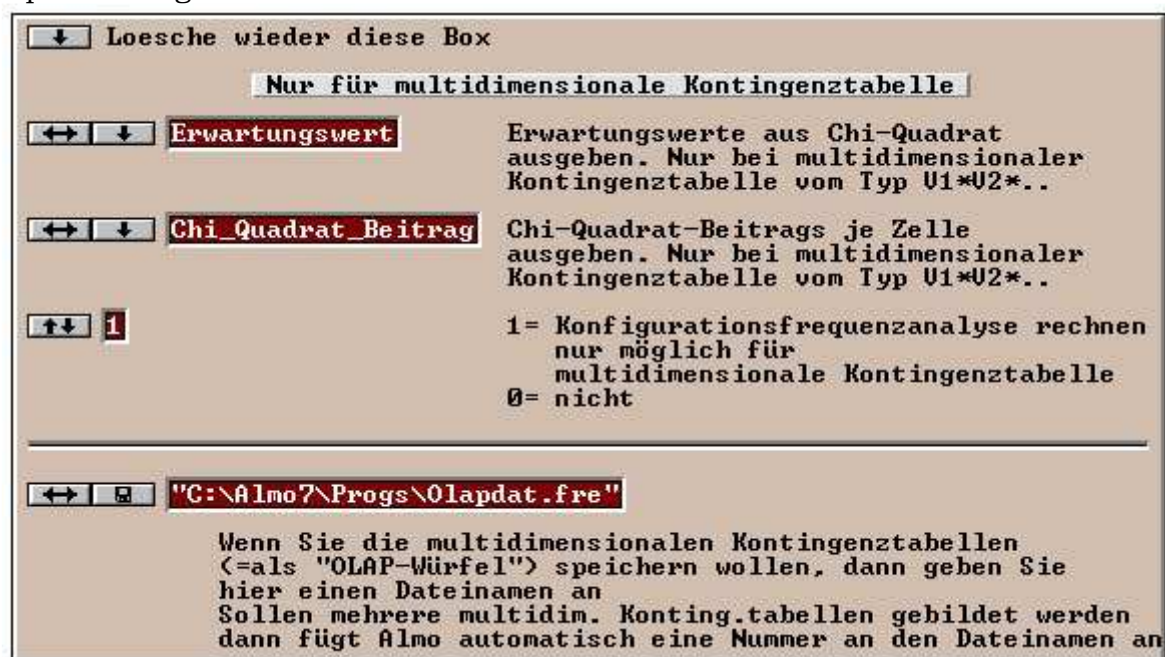
Box 15: Gewichtung

Siehe dazu Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.8.

Box 12: Optionen: Nur für multidimensionaler Kontingenztabelle



Optionsbox geöffnet:



In dieser Box werden Ihnen weitere Tests und Koeffizienten angeboten. Die Berechnung der Erwartungswerte, der Chi-Quadrat-Beiträge und die Konfigurationsfrequenzanalyse haben wir in P10.4.1 und P10.7.1 dargestellt.

Im unteren Teil dieser Optionsbox wird dem Benutzer angeboten, die multidimensionale Kontingenztabelle als „Olap-Würfel“ zu speichern. Wir werden in Abschnitt P11.6 diese Möglichkeit erklären.

Wird im Eingabefeld ein Dateiname angegeben, dann erzeugt Almo 2 Dateien:

1. eine nicht lesbare Datei mit der Erweiterung `__.dir` in unserem Beispiel: "C:\Almo\Progs\Olapdat.dir"
2. eine Datei der Variablen- und Ausprägungsnamen der Variablen des "OLAP-Würfels" mit der Erweiterung `__.nam` in unserem Beispiel: "C:\Almo\Progs\Olapdat.nam"

Beachte: Damit Also eine Datei der Variablen- und Ausprägungsnamen erstellen und speichern kann, sollten Sie zumindest den Variablen, die für die multidimensionalen Kontingenztabellen verwendet werden, Variablen- und möglichst auch Ausprägungsnamen geben.

Werden zu einem späteren Zeitpunkt mit Prog11m3 Sub-Tabellen gebildet, so müssen diese beiden Dateien in Prog11m3 in der Box

Datei aus der der "OLAP-Würfel" eingelesen wird

und in der Box

Datei der Variablennamen

eingesetzt werden.

P11.4 Ausgabe

Aus dem Maskenprogramm Prog11m.Msk erhalten wir folgende Ausgabe, die wir hier gekürzt darstellen. (Die Gewichtung und das Ein- bzw. Ausschließen von Datensätzen im Maskenprogramm ist herausgenommen).

P11.4.1 Ausgabe für 2-dimensionale Tabelle

Tabelle 1: V1 Geschlecht
mit
V4 Schulbildung

Bezeichnung der Variablen auf Stirnseite der Tabelle (Zeilenvariable)

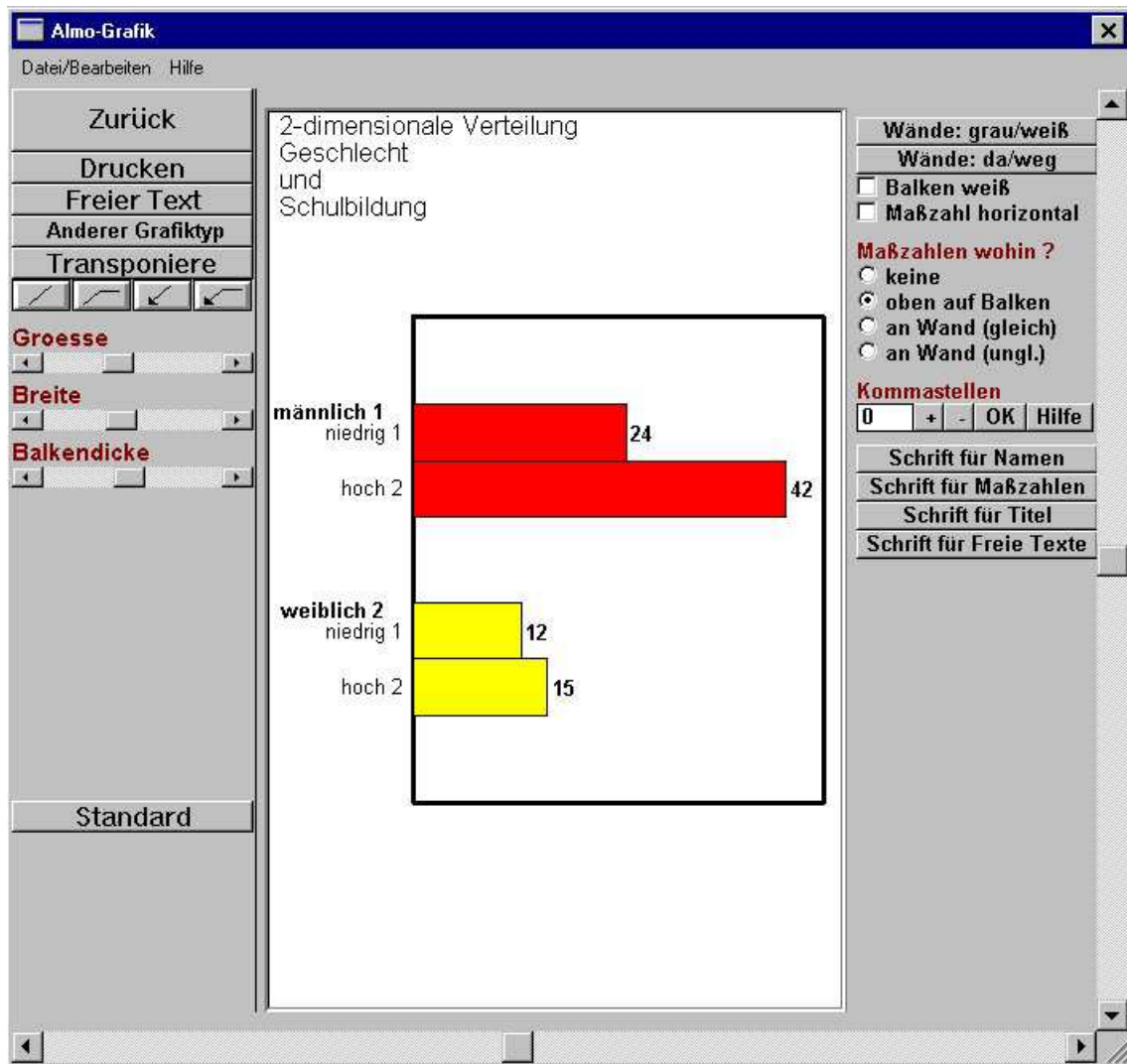
V1 Geschlecht
1 männlich
2 weiblich

Bezeichnung der Variablen im Kopf der Tabelle (Spaltenvariable)

V4 Schulbildung
1 niedrig
2 hoch

	Schulbildung		Summe
	niedrig 1	hoch 2	
Geschlecht männlich	12	22	34
weiblich	12	15	27
Summe	24	37	61

Almo erzeugt für diese Tabelle folgendes Balkendiagramm

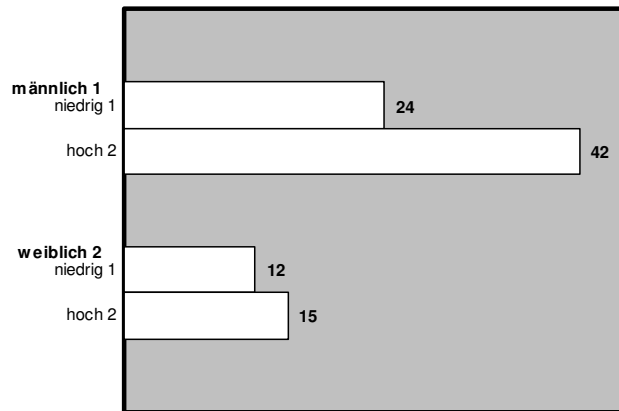


Auf der linken Seite des Grafikeditors befindet sich der Knopf

Transponiere

Nach Klick auf diesen Knopf entsteht folgende Grafik:

2-dimensionale Verteilung
Geschlecht
und
Schulbildung

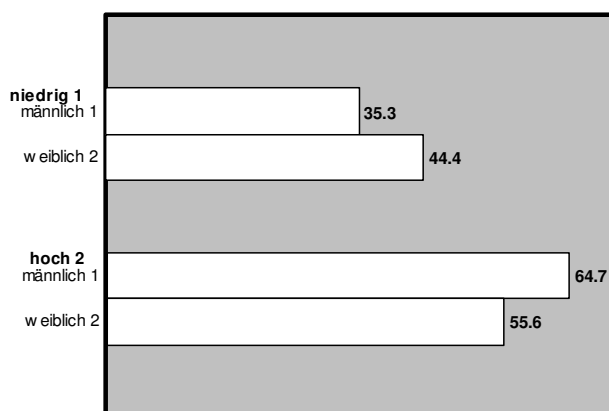


zeilenweise prozentuiert

		Schulbildung		Summe
		niedrig 1	hoch 2	
Geschlecht	männlich	35.29	64.71	100.00
	weiblich	44.44	55.56	100.00
Summe		39.34	60.66	100.00

Almo erzeugt für diese Tabelle folgendes Balkendiagramm

2-dimensionale Verteilung
Geschlecht
und
Schulbildung



Das Balkendiagramm zeigt, dass in die Kategorie "niedrige Schulbildung"

von den 100 % Männern nur 35 % fallen
von den 100 % Frauen jedoch 44 % fallen

Ein Klick auf den Transponiere-Knopf ist hier nicht sinnvoll, da in der Tabelle zeilenweise auf 100 % prozentuiert wurde.

Koeffizienten
Chi-Quadrat = 0.5280 df = 1 Signifikanz (1-p)*100 = 53.240

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Kontingenzkoeffizient C(cor) = 0.131
Cramers V = 0.093

******* Erläuterung:**

Es besteht kein signifikanter Zusammenhang zwischen Geschlecht und Schulbildung. Wären die Tabelle signifikant, dann würde der Korrelationskoeffizient die Stärke des Zusammenhangs ausdrücken. Der "beste" Korrelationskoeffizient ist Cramers V. Auch wenn er in aller Regel niedriger ist als der Kontingenzkoeffizient sollte man ihn im Forschungsbericht verwenden. Siehe dazu im Handbuch zu P20, Abschnitt P20.9.5.

P11.4.2 Ausgabe für Mehrfach-Tabelle

Tabelle 2: V1 Geschlecht + V2 Wohnort + V3 Beruf
mit
V4 Schulbildung

Bezeichnung der Variablen auf Stirnseite der Tabelle (Zeilenvariable)

V1 Geschlecht
 1 männlich
 2 weiblich
V2 Wohnort
 1 Stadt
 2 Land
V3 Beruf
 1 Arbeiter
 2 Angestellter
 3 Selbständiger

Bezeichnung der Variablen im Kopf der Tabelle (Spaltenvariable)

V4 Schulbildung
 1 niedrig
 2 hoch

		Schulbildung		Summe
		niedrig 1	hoch 2	
Geschlecht	männlich	12	22	34
	weiblich	12	15	27
Summe		24	37	61

Wohnort	Stadt	12	15	27
	Land	12	22	34
Summe		24	37	61

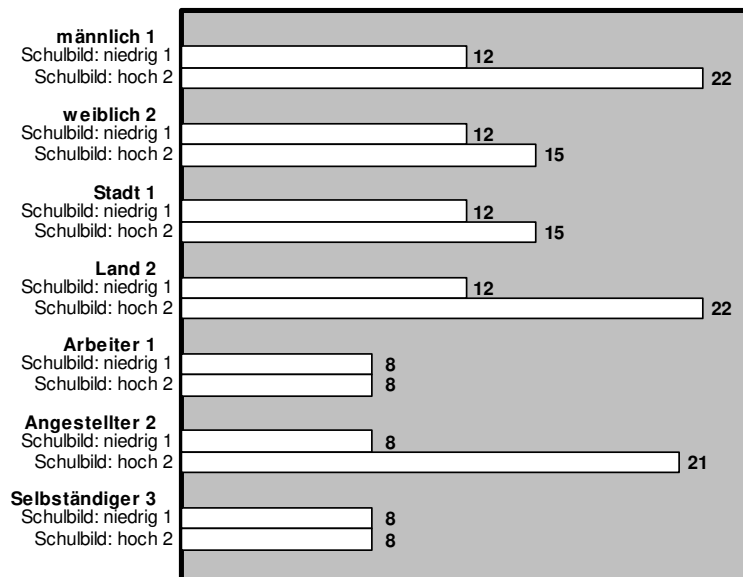
Beruf	Arbeiter	8	8	16
	Angestellter	8	21	29
	Selbständige	8	8	16
Summe		24	37	61

***** **Erläuterung:**

Es entstehen 3 Teiltabellen, die zu einer einzigen Tabelle zusammengefasst sind. Dadurch entsteht eine sehr kompakte und übersichtliche Ausgabe.

Almo erzeugt für diese Tabelle folgendes Balkendiagramm:

2-dimensionale Verteilung



zeilenweise prozentuiert

		Schulbildung		Summe
		niedrig 1	hoch 2	
Geschlecht	männlich	35.29	64.71	100.00
	weiblich	44.44	55.56	100.00
Summe		39.34	60.66	100.00

Wohnort		Schulbildung		Summe
		niedrig 1	hoch 2	
Stadt		44.44	55.56	100.00
Land		35.29	64.71	100.00
Summe		39.34	60.66	100.00

Beruf	Schulbildung		Summe
	niedrig 1	hoch 2	
Arbeiter	50.00	50.00	100.00
Angestellter	27.59	72.41	100.00
Selbständige	50.00	50.00	100.00
Summe	39.34	60.66	100.00

Auch für diese Tabelle erzeugt Almo ein Balkendiagramm, das wir hier nicht abbilden.

```

Koeffizienten fuer Teiltabelle 1          V1 Geschlecht
                                           mit
                                           V4 Schulbildung

Chi-Quadrat = 0.5280          df = 1          Signifikanz (1-p)*100 = 53.240

                                           Ein Wert ueber ca. 95 bedeutet:
                                           Zwischen den Variablen besteht
                                           ein signifikanter Zusammenhang

Kontingenzkoeffizient C(cor) = 0.131
Cramers V = 0.093
  
```

```

=====

Koeffizienten fuer Teiltabelle 2          V2 Wohnort
                                           mit
                                           V4 Schulbildung

Chi-Quadrat = 0.5280          df = 1          Signifikanz (1-p)*100 = 53.240

                                           Ein Wert ueber ca. 95 bedeutet:
                                           Zwischen den Variablen besteht
                                           ein signifikanter Zusammenhang

Kontingenzkoeffizient C(cor) = 0.131
Cramers V = 0.093
  
```

```

=====

Koeffizienten fuer Teiltabelle 3          V3 Beruf
                                           mit
                                           V4 Schulbildung
  
```

Chi-Quadrat = 3.2025 df = 2 Signifikanz (1-p)*100 = 80.025

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Kontingenzkoeffizient C(cor) = 0.316
Cramers V = 0.229

******* Erläuterung:**

Alle 3 Teiltabellen sind nicht signifikant. Wären die Tabelle signifikant, dann würde der Korrelationskoeffizient die Stärke des Zusammenhangs ausdrücken. Der "beste" Korrelationskoeffizient ist Cramers V. Auch wenn er in aller Regel niedriger ist als der Kontingenzkoeffizient sollte man ihn im Forschungsbericht verwenden. Siehe dazu im Handbuch zu P20, Abschnitt P20.9.5.

P11.4.3 Ausgabe für Partial- bzw. Interaktions-Tabelle

Tabelle 3: V1 Geschlecht * V2 Wohnort * V3 Beruf
mit
V4 Schulbildung

Bezeichnung der Variablen auf Stirnseite der Tabelle (Zeilenvariable)

- V1 Geschlecht
 - 1 männlich
 - 2 weiblich
- V2 Wohnort
 - 1 Stadt
 - 2 Land
- V3 Beruf
 - 1 Arbeiter
 - 2 Angestellter
 - 3 Selbständiger

Bezeichnung der Variablen im Kopf der Tabelle (Spaltenvariable)

- V4 Schulbildung
 - 1 niedrig
 - 2 hoch

			Schulbildung		Summe	
			niedrig 1	hoch 2		
Geschlecht	Wohnort	Beruf				
männlich	Stadt	Arbeiter	2	2	4	Partialtabelle für Männer in der Stadt
		Angestellt	2	2	4	
		Selbständi	2	2	4	
	Land	Arbeiter	2	2	4	Partialtabelle für Männer auf dem Land
		Angestellt	2	12	14	
		Selbständi	2	2	4	
weiblich	Stadt	Arbeiter	2	2	4	Partialtabelle für Frauen in der Stadt
		Angestellt	2	5	7	
		Selbständi	2	2	4	
	Land	Arbeiter	2	2	4	Partialtabelle für Frauen auf dem Land
		Angestellt	2	2	4	
		Selbständi	2	2	4	
Summe			24	37	61	

zeilenweise prozentuiert

Geschlecht	Wohnort	Beruf	Schulbildung		Summe
			niedrig 1	hoch 2	
männlich	Stadt	Arbeiter	50.00	50.00	100.00
		Angestellt	50.00	50.00	100.00
		Selbständi	50.00	50.00	100.00
	Land	Arbeiter	50.00	50.00	100.00
		Angestellt	14.29	85.71	100.00
		Selbständi	50.00	50.00	100.00
weiblich	Stadt	Arbeiter	50.00	50.00	100.00
		Angestellt	28.57	71.43	100.00
		Selbständi	50.00	50.00	100.00
	Land	Arbeiter	50.00	50.00	100.00
		Angestellt	50.00	50.00	100.00
		Selbständi	50.00	50.00	100.00
Summe			39.34	60.66	100.00

Koeffizienten fuer Partialtabelle 1 V3 Beruf
mit
V4 Schulbildung
fuer
V1 Geschlecht: 1 männlich
V2 Wohnort: 1 Stadt

Chi-Quadrat = 0.0000 df = 2 Signifikanz (1-p)*100 = 0.000

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
kleiner 5 =100% der Zellen

Kontingenzkoeffizient C(cor) = 0.000
Cramers V = 0.000

******* Erläuterung:**

In dieser Partialtabelle sind die 3 Berufsgruppen (der in der Stadt wohnenden Männer) mit gleicher Häufigkeit auf die niedrige und hohe Schulbildung verteilt. Demzufolge ist der Chi-Quadrat-Wert und die Signifikanz =0.

Auch für diese Tabelle erzeugt Almo ein Balkendiagramm, das wir hier nicht abbilden.

=====
Koeffizienten fuer Partialtabelle 2 V3 Beruf
mit
V4 Schulbildung
fuer
V1 Geschlecht: 1 männlich
V2 Wohnort: 2 Land

Chi-Quadrat = 3.2738 df = 2 Signifikanz (1-p)*100 = 80.737

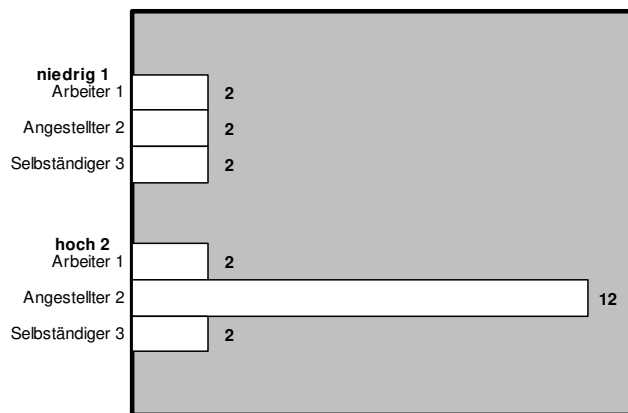
Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
kleiner 5 = 83% der Zellen

Kontingenzkoeffizient C(cor) = 0.509
Cramers V = 0.386

Almo erzeugt für diese Tabelle folgendes Balkendiagramm

2-dimensionale Verteilung
Beruf
mit
Schulbildung
für
....Geschlecht : männlich
....Wohnort : Land



=====
Koeffizienten fuer Partialtabelle 3 V3 Beruf
mit
V4 Schulbildung
fuer
 V1 Geschlecht: 2 weiblich
 V2 Wohnort: 1 Stadt

Chi-Quadrat = 0.7143 df = 2 Signifikanz (1-p)*100 = 29.499

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
kleiner 5 =100% der Zellen

Kontingenzkoeffizient C(cor) = 0.302
Cramers V = 0.218

Auch für diese Tabelle erzeugt Almo ein Balkendiagramm, das wir hier nicht abbilden.

=====
Koeffizienten fuer Partialtabelle 4 V3 Beruf
mit
V4 Schulbildung
fuer

V1 Geschlecht: 2 weiblich
V2 Wohnort: 2 Land

Chi-Quadrat = 0.0000 df = 2 Signifikanz (1-p)*100 = 0.000

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
 kleiner 5 =100% der Zellen

Kontingenzkoeffizient C(cor) = 0.000
Cramers V = 0.000

Auch für diese Tabelle erzeugt Almo ein Balkendiagramm, das wir
hier nicht abbilden.

=====

Koeffizienten fuer gesamte Interaktionstabelle

Beachte: Die Merkmalskombinationen der unabhängigen Variablen Geschlecht,
Wohnort, Beruf werden wie die Ausprägungen *einer* unabhängigen Variablen
betrachtet.

Chi-Quadrat = 5.9273 df =11 Signifikanz (1-p)*100 = 12.161

Ein Wert ueber ca. 95 bedeutet:
Zwischen den Variablen besteht
ein signifikanter Zusammenhang

Erwartungswerte kleiner 1 = 0%
 kleiner 5 = 92% der Zellen

Kontingenzkoeffizient C(cor) = 0.421
Cramers V = 0.312

P11.4.4 Ausgabe für multidimensionale Kontingenztabelle

Tabelle 4: V1 Geschlecht * V2 Wohnort * V3 Beruf

Bezeichnung der Variablen auf Stirnseite der Tabelle (Zeilenvariable)

V1 Geschlecht
 1 männlich
 2 weiblich
V2 Wohnort
 1 Stadt
 2 Land
V3 Beruf
 1 Arbeiter
 2 Angestellter
 3 Selbständiger

			Fälle
Geschlecht	Wohnort	Beruf	
männlich	Stadt	Arbeiter	4
		Angestellt	4
		Selbständi	4
	Land	Arbeiter	4
		Angestellt	14
		Selbständi	4
weiblich	Stadt	Arbeiter	4
		Angestellt	7
		Selbständi	4
	Land	Arbeiter	4
		Angestellt	4
		Selbständi	4
Summe			61

Matrix der Erwartungswerte

			Werte
Geschlecht	Wohnort	Beruf	
männlich	Stadt	Arbeiter	3.9473
		Angestellt	7.1545
		Selbständi	3.9473
	Land	Arbeiter	4.9707
		Angestellt	9.0094
		Selbständi	4.9707
weiblich	Stadt	Arbeiter	3.1346
		Angestellt	5.6815
		Selbständi	3.1346
	Land	Arbeiter	3.9473
		Angestellt	7.1545
		Selbständi	3.9473
Summe			61.0000

***** **Erläuterung:**

Zum Begriff des "Erwartungswertes" siehe P10.4.1.

Matrix der Chi-Quadrat-Beitraege

Geschlecht Wohnort Beruf			Werte
männlich	Stadt	Arbeiter	0.0007
		Angestellt	1.3909
		Selbständi	0.0007
	Land	Arbeiter	0.1896
		Angestellt	2.7644
		Selbständi	0.1896
weiblich	Stadt	Arbeiter	0.2389
		Angestellt	0.3060
		Selbständi	0.2389
	Land	Arbeiter	0.0007
		Angestellt	1.3909
		Selbständi	0.0007
Summe			6.7119

******* Erläuterung:**

Zum Begriff des "Chi-Quadrat-Beitrages" siehe P10.4.1.

Chi-Quadrat-Test auf allseitige Abhaengigkeit

 Chi-Quadrat = 6.7119 df = 7 Signifikanz (1-p)*100 = 53.932

2I-Test auf allseitige Abhaengigkeit

 2I-Wert = 6.8089 df = 7 Signifikanz (1-p)*100 = 54.982

Der Chi-Quadrat- und der 2I-Test sind pauschale Tests auf allseitige Un- bzw. Abhängigkeit. Ein signifikanter Wert bedeutet also, dass irgendwelche Variable - nicht notwendigerweise alle - zusammenhängen.

******* Erläuterung:**

Mit der Konfigurationsfrequenzanalyse wird untersucht, ob einzelne Chi-Quadrat-Beiträge signifikant sind. Siehe hierzu auch P10.7.1.

Konfigurationsfrequenzanalyse (KFA)

Schranke fuer	bei Signifikanz
Chi-Quadrat-Beitrag	(1-p)*100
5.027256	85.0
5.738647	90.0
<u>6.956879</u>	<u>95.0</u>
8.238113	97.5
10.089935	99.0
13.923828	99.9

Bei der KFA werden zuerst die Chi-Quadrat-Schranken angegeben. So muß etwa ein einzelner Chi-Quadrat-Beitrag größer sein als 6.956 um mit 95 % Sicherheit auf einen Typus hinzuweisen. In unserem Beispiel ist das bei keinem einzigen Wert der Fall.

Matrix der p-Werte aus KFA-Binomialtest

Geschlecht Wohnort Beruf			Werte
männlich	Stadt	Arbeiter	0.562460
		Angestellt	0.142930
		Selbständi	0.562460
	Land	Arbeiter	0.438336
		Angestellt	0.058517
		Selbständi	0.438336
weiblich	Stadt	Arbeiter	0.383658
		Angestellt	0.340018
		Selbständi	0.383658
	Land	Arbeiter	0.562460
		Angestellt	0.142930
		Selbständi	0.562460
Summe			-

Schranke fuer p-WERT in Tabelle	bei Signifikanz (1-p)*100
0.012500	85.0
0.008333	90.0
0.004167	95.0
0.002083	97.5
0.000833	99.0
0.000083	99.9

Dann werden die p-Werte aus dem exakten Binomialtest angegeben und die für sie gültigen Schranken. So muß ein p-Wert niedriger sein als 0.004167 um mit 95 % Sicherheit auf einen Typus zu verweisen. Auch das ist in unserem Beispiel bei keinem Tabellenwert der Fall. Der exakte Binomialtest ist zur Interpretation zu verwenden, wenn Erwartungswerte kleiner 5 auftreten. Zur multidimensionalen KFA siehe Krauth/Lienert(1973) und Lienert (1978, S.790-795).

Hierarchische KFA

Für die hierarchische KFA, wie sie Lienert (1978, S.795ff) oder Lohse u.a. (1982, S.405) beschreiben, ist es notwendig mehrere KFAs zu rechnen und zu vergleichen. Bei Lienert und Lohse u.a. wird dieser Vergleich mittels der Fisherschen z-Approximation durchgeführt. Anstelle der z-Approximation kann in ALMO genauer über die (1-p)*100-Signifikanzwerte verglichen werden, die für den allseitigen Chi-Quadrat- oder 2I- Test ermittelt wurden.

P11.6 Eingabe einer fertigen "multidimensionalen Kontingenztabelle" mit Prog11m4

Die fertige Tabelle muss von Hand eingegeben werden. Das Programm entspricht Prog11m3: OLAP II - mit dem Unterschied, dass dort die Tabelle aus einer Datei eingelesen wird.

Aus der fertigen Tabelle können "ausgelöst" werden

1. 2-dimensionale Tabellen
2. 2-dimensionale Mehrfach-Tabellen
3. Partial- bzw. Interaktions-Tabellen
4. multidimensionale Kontingenztabelle

Berechnet werden standarmäßig folgende Koeffizienten:

Chi-Quadrat, Chi-Quadrat-Test und 2I-Test auf allseitige Abhängigkeit
Kontingenzkoeffizient C(cor), Cramers V

Zu 4. kann (optional) die Tabelle der Erwartungswerte zum Chi-Quadrat, die Tabelle der Chi-Quadrat-Beiträge sowie eine Konfigurationsfrequenzanalyse gerechnet werden

Prog11m4.Msk

Eingabe einer fertigen "multidimensionalen Kontingenztabelle"

Das Programm entspricht OLAP II - mit dem Unterschied, dass die Tabelle nicht aus einer Datei eingelesen wird, sondern vom Benutzer von Hand einzugeben ist

Aus der fertigen Tabelle können "ausgelöst" werden

1. 2-dimensionale Tabellen
2. 2-dimensionale Mehrfach-Tabellen
3. Partial- bzw. Interaktions-Tabellen
4. multidimensionale Kontingenztabellen

Berechnet werden standarmäßig folgende Koeffizienten:
 Chi-Quadrat, Chi-Quadrat-Test und 2I-Test auf allseitige Abhängigkeit
 Kontingenzkoeffizient C(cor), Cramers U

Zu 4. kann (optional) die Tabelle der Erwartungswerte zum Chi-Quadrat, die Tabelle der Chi-Quadrat-Beiträge sowie eine Konfigurationsfrequenzanalyse gerechnet werden

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

1

Vereinbare Variable= ;

2 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

Name1=Geschlecht:männlich,weiblich;
 Name2=Wohnort:Stadt,Land;
 Name3=Beruf:Arbeiter,Angestellter,Selbständiger;
 Name4=Schulbildung:niedrig,hoch;
 erzeuge zusätzliche Namensfelder

4

in der Reihenfolge, in der sie in der Tabelle aufeinander folgen

Geschlecht, Wohnort, Beruf, Schulbildung

Im 1. Eingabefeld die Variablennamen bzw. -nummern eingeben
 im 2. die Zahl der Ausprägungen der Variablen

5

aus eingegebener Tabelle auslösen

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11

Geschlecht mit Schulbildung
 erzeuge zusätzliche Felder für Tabellen-Angaben

2-dimensionale "Mehrfach-Tabelle"
aus eingegebener Tabelle auslösen

Hilfe

BEACHTEN: Nur 1 Spaltenvariable erlaubt

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11
Wohnort	Stadt	12	15
	Land	12	30
Beruf	Arbeiter	12	12
	Angestel	12	33

↔ □□ **Geschlecht + Wohnort + Beruf mit Schulbildung**

[...] erzeuge zusätzliche Felder für Tabellen-Angaben

6

Partialtabellen
Interaktionstabelle
aus eingegebener Tabelle auslösen

Hilfe

BEACHTEN: Nur 1 Spaltenvariable erlaubt

Beispiel:

Geschl	Ort	Beruf	Schulbildung		Summe
			niedrig	hoch	
männl	Stadt	Arbeit	4	4	8
		Angest	4	4	8
	Land	Arbeit	4	4	8
		Angest	4	22	26
weibl	Stadt	Arbeit	2	2	4
		Angest	2	5	7
	Land	Arbeit	2	2	4
		Angest	2	2	4
Summe			24	45	69

↔ □□ **Geschlecht * Wohnort * Beruf mit Schulbildung**

[...] erzeuge zusätzliche Felder für Tabellen-Angaben

7

Multidimensionale Kontingenztabelle
aus eingegebener Tabelle auslösen **Hilfe**

Beispiel:

			Fälle
Geschl	Ort	Beruf	
männl	Stadt	Arbeit	8
		Angest	8
		Land	8
		Angest	26
weibl	Stadt	Arbeit	4
		Angest	7
		Land	4
		Angest	4
Summe			69

BEACHTTE: Kein "mit" vorhanden

Geschlecht * Wohnort * Beruf

erzeuge zusätzliche Felder für Tabellen-Ängaben

8

Variable aus Tabellenangaben

<----- keine Benutzereingabe
Almo ermittelt die Variablen, die
in obige Tabellen eingegangen sind
selbst. Sie können aber auch auf
diesen Knopf klicken

9

Option: Nur für multidimensionale Kontingenztabelle

10

Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

11

Grafik-Optionen

12

Tabelle schreiben

Schreiben Sie hier dahinter die Tabelle

Die fertige Tabelle ist in
folgender Weise einzugeben ---> **Hilfe**

Schalten Sie dazu die Schreibsperre aus

Schreibsperre <--- EIN : rot
AUS : grau

13

```

1 1 1 1 52
1 1 1 2 112
1 1 2 1 22
1 1 2 2 42
1 1 3 1 22
1 1 3 2 32
1 2 1 1 31
1 2 1 2 27
1 2 2 1 8
1 2 2 2 12
1 2 3 1 82
1 2 3 2 33
2 1 1 1 41
2 1 1 2 17
2 1 2 1 35
2 1 2 2 51
2 1 3 1 31
2 1 3 2 18
2 2 1 1 19
2 2 1 2 44
2 2 2 1 82
2 2 2 2 35
2 2 3 1 13
2 2 3 2 22

```

P11.6.1 Erläuterungen zu den Boxen

Wir beginnen mit der letzten Box "Tabelle schreiben"

Tabelle schreiben

Schreiben Sie hier dahinter die Tabelle

Die fertige Tabelle ist in folgender Weise einzugeben ---->

Schalten Sie dazu die Schreibsperre aus

Schreibsperre <--- EIN : rot
AUS : grau

Betrachten wir ein Beispiel einer 4-dimensionalen Kontingenztabelle:

diese Tabelle soll eingegeben werden

in dieser Weise ist
sie am Ende des
Maskenprogramms
zu schreiben

Geschlec	Wohnort	Beruf	Schulbil	Fälle	
1 männli	1 Stadt	1 Arbeit	1 niedri	52	1 1 1 1 52
			2 hoch	112	1 1 1 2 112
		2 Angest	1 niedri	22	1 1 2 1 22
			2 hoch	42	1 1 2 2 42
	3 Selbst	1 niedri	2 hoch	22	1 1 3 1 22
			2 hoch	32	1 1 3 2 32
	2 Land	1 Arbeit	1 niedri	31	1 2 1 1 31
			2 hoch	27	1 2 1 2 27
		2 Angest	1 niedri	8	1 2 2 1 8
2 hoch			12	1 2 2 2 12	
3 Selbst		1 niedri	2 hoch	82	1 2 3 1 82
			2 hoch	33	1 2 3 2 33
2 weibli	1 Stadt	1 Arbeit	1 niedri	41	2 1 1 1 41
			2 hoch	17	2 1 1 2 17
		2 Angest	1 niedri	35	2 1 2 1 35
		2 hoch	51	2 1 2 2 51	
	3 Selbst	1 niedri	2 hoch	31	2 1 3 1 31
			2 hoch	18	2 1 3 2 18
	2 Land	1 Arbeit	1 niedri	19	2 2 1 1 19
			2 hoch	44	2 2 1 2 44
		2 Angest	1 niedri	82	2 2 2 1 82
2 hoch			35	2 2 2 2 35	
3 Selbst		1 niedri	2 hoch	13	2 2 3 1 13
			2 hoch	22	2 2 3 2 22

Box 1 und 2: Siehe Almo-Dokument Nr. 0 "Arbeiten mit Almo", PO.1

Box 3: Namen der Variable, die die fertige Tabelle bilden

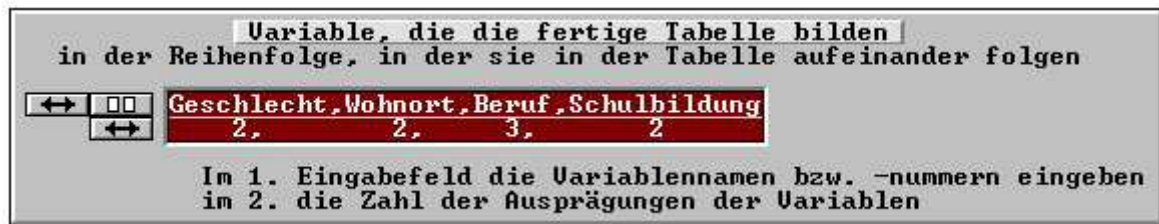


In unserem Beispiel sind es 4 Variable, die die Tabelle bilden und von Almo als V1 bis V4 begriffen werden. Sie erhalten die Namen:

Geschlecht, Wohnort, Beruf, Schulbildung

Box 4: Variable, die die fertige Tabelle bilden in der Reihenfolge, in der sie in der

Tabelle aufeinander folgen



Im 1. Eingabefeld sind die Variablennamen bzw. -nummern einzugeben, im 2. die Zahl der Ausprägungen der Variablen.

Box 5: 2-dimensionale Tabelle aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.2

Box 6: 2-dimensionale "Mehrfach-Tabelle" aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.3

Box 7: Partialtabellen (Interaktionstabelle) aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.4

Box 8: Multidimensionale Kontingenztabelle aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.5

Box 9: Option: Nur für multidimensionale Kontingenztabelle
Siehe Abschnitt P11.2 Erläuterung zu Box 12.

Hier haben Sie die Möglichkeit, die von Ihnen geschriebene Tabelle in eine Datei zu speichern. Geben Sie in der Box 8 alle in der Box 4 enthaltenen Variablen als multidimensionale Kontingenztabelle an. In der Optionsbox 9 teilen Sie einen Dateinamen mit.

Box 10: Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix
Siehe Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.9

Box 11: Grafik-Optionen
Siehe P0.10

P11.7 Der "OLAP-Würfel"

Wir übernehmen den Begriff des "OLAP-Würfels" aus der Wirtschaftsinformatik. Das Wort "OLAP" ist aus den Anfangsbuchstaben von "online analytic processing" gebildet. In Almo ist ein Olap-Würfel nichts anderes als eine multidimensionale Kontingenztabelle. Betrachten wir ein Beispiel. Wir bilden eine multidimensionale Kontingenztabelle aus folgenden Variablen

V1 * V2 * V3 * V4
Geschlecht * Wohnort * Beruf * Schulbildung

Die 4 Variablen besitzen folgende Ausprägungen:

- V1 Geschlecht
 - 1 männlich
 - 2 weiblich
- V2 Wohnort
 - 1 Stadt
 - 2 Land
- V3 Beruf
 - 1 Arbeiter
 - 2 Angestellter
 - 3 Selbständiger
- V4 Schulbildung
 - 1 niedrig
 - 2 hoch

Mit Progl1m oder Progl1m1 ermitteln wir folgende multidimensionale Kontingenztabelle

Die Tabelle kann auch in folgender Form geschrieben werden, wobei die Ausprägungen der Variablen durch ihre Code-Ziffern angegeben werden

				Fälle
Geschl	Wohnor	Beruf	Schulb	
männl.	Stadt	Arbeit	niedri	52
			hoch	112
		Angest	niedri	22
			hoch	42
		Selbst	niedri	22
			hoch	32
	Land	Arbeit	niedri	31
			hoch	27
		Angest	niedri	8
			hoch	12
		Selbst	niedri	82
			hoch	33
weibl.	Stadt	Arbeit	niedri	41
			hoch	17
		Angest	niedri	35
			hoch	51
		Selbst	niedri	31
			hoch	18
	Land	Arbeit	niedri	19
			hoch	44
		Angest	niedri	82
			hoch	35
		Selbst	niedri	13
			hoch	22
Summe				883

Geschl	Wohnor	Beruf	Schulb	Fälle
1	1	1	1	52
1	1	1	2	112
1	1	2	1	22
1	1	2	2	42
1	1	3	1	22
1	1	3	2	32
1	2	1	1	31
1	2	1	2	27
1	2	2	1	8
1	2	2	2	12
1	2	3	1	82
1	2	3	2	33
2	1	1	1	41
2	1	1	2	17
2	1	2	1	35
2	1	2	2	51
2	1	3	1	31
2	1	3	2	18
2	2	1	1	19
2	2	1	2	44
2	2	2	1	82
2	2	2	2	35
2	2	3	1	13
2	2	3	2	22

Den auf der rechten Seite stehenden Zahlenkörper dieser Tabelle bezeichnen wir als "OLAP-Würfel".

Ein OLAP_Würfel kann mit Prog11m1 und Prog11m4 über die Option "Nur für multidimensionale Kontingenztabelle" und mit dem nachfolgend dargestellten Programm Prog11m2 hergestellt und in eine Datei gespeichert werden. Danach können mit Prog11m3 aus dem Olap-Würfel Sub-Tabellen jeglicher Art "herausgelöst" werden, also

1. 2-dimensionale Tabellen
2. 2-dimensionale Mehrfach-Tabellen
3. Partial- bzw. Interaktions-Tabellen
4. multidimensionale Kontingenztabelle

Für diese Sub-Tabellen können dann verschiedene Koeffizienten errechnet werden. Dies ist der Sinn des Olap-Würfels. Selbstverständlich könnten die Sub-Tabellen, die der Benutzer sehen möchte, auch unmittelbar mit Prog11m oder Prog11m1 aus der Datei der Untersuchungseinheiten gebildet werden. Ob die Rechenzeit insgesamt kürzer wird, wenn ein Olap-Würfel gebildet wird und aus diesem Sub-Tabellen herausgelöst werden, hängt von der Größe der Daten-Datei und der Größe des Olap-Würfels ab. Besteht dieser aus vielen Variablen mit vielen Ausprägungen, so kann die Datei, die der Olap-Würfel einnimmt, sehr groß werden. Dann ist es ökonomischer die gewünschten Tabellen mit Prog11m oder Prog11m1 zu bilden.

P11.7.1 Programm-Maske zum Erstellen eines "OLAP-Würfels"

Prog11m2Msk
OLAP I
 OLAP-Würfel aus Daten bilden
 (<=multidimensionale Kontingenztabelle bilden)
 OLAP-Würfel können auch mit Prog11m1 über die Option
 "Nur für multidimensionale Kontingenztabelle" gebildet werden
 Oder doppelklicken Sie gleich auf: [".\Almo_Msk\Prog11m1.Msk"]

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

1

Vereinbare Variable= ;

2 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

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

zeige zeige = Namensdatei in Output zeigen
 leer = nicht

4

erzeuge zusätzliche Namensfelder

5

bei Datei-Problemen

"C:\Almo7\TESTDAT\TESTDAT.FRE"

frei Format der Daten

U1:20 der Datensatz enthält diese Variablen
 Bei Format DIREKT schreiben Sie: alle_U

6 Wenn Dateiformat FIX oder Nicht-Standard-FREI

7

(<=multidimensionale Kontingenztabelle)
aus nachfolgenden Variablen bilden

Geschlecht * Wohnort * Beruf * Schulbildung

 <--- keine Benutzereingabe

8

9

10

11

12

13

Erläuterungen zu den Boxen:

Box 1 bis 6: Siehe Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.1 bis P0.4

Box 7: OLAP-Würfel (=multidimensionale Kontingenztabelle) aus nachfolgenden Variablen bilden

Die Eingabe ist dieselbe wie bei Prog11m in Box 10. Siehe Abschnitt P11.1.5

Box 8: Datei, in die OLAP-Würfel gespeichert werden soll

Geben Sie einen Namen für die Datei an, in die der OLAP-Würfel gespeichert werden soll.

Wird im Eingabefeld ein Dateiname angegeben, dann erzeugt Almo 2 Dateien:

1. eine nicht lesbare Datei

mit der Erweiterung `__.dir`
in unserem Beispiel: "C:\Almo\Progs\Olapdat.dir"

2. eine Datei der Variablen- und Ausprägungsnamen der Variablen des "OLAP-Würfels" mit der Erweiterung `__.nam`
in unserem Beispiel: "C:\Almo\Progs\Olapdat.nam"

Beachte: Damit Almo eine Datei der Variablen- und Ausprägungsnamen erstellen und speichern kann, sollten Sie zumindest den Variablen, die für die multidimensionalen Kontingenztabelle verwendet werden, Variablen- und möglichst auch Ausprägungsnamen geben.

Werden zu einem späteren Zeitpunkt mit Prog11m3 Sub-Tabellen gebildet, so müssen diese beiden Dateien in Prog11m3 in der Box

 Datei aus der der "OLAP-Würfel" eingelesen wird

und in der Box

 Datei der Variablennamen

eingesetzt werden.

Box 9: Option: Ein- und Ausschließen von Untersuchungseinheiten
siehe Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.7

Box 10: Option: Umkodierungen und Kein-Wert-Angaben
siehe P0.5

Box 11: Option: Untersuchungseinheiten gewichten
siehe P0.8

Box 12: Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix
siehe P0.9

P11.7.1.1 Ausgabe aus Prog11m2

Almo liefert die multidimensionale Kontingenztabelle - so wie diese oben abgebildet ist.

P11.7.2 Die Auswertung des "OLAP-Würfels" mit Prog11m3

Ein OLAP-Würfel wird dadurch ausgewertet, dass Sub-Tabellen aus ihm herausgelöst werden und für diese verschiedene Koeffizienten errechnet werden. Almo enthält zwei Programme die dies leisten, die sich lediglich dadurch unterscheiden, dass bei dem einen, der Olap-Würfel "von Hand" eingegeben werden muß (Prog11m4) und beim anderen der OLAP-Würfel aus einer Datei eingelesen wird (Prog11m3). Wir werden im folgenden Prog11m3 vorstellen.

Prog11m3.Msk
OLAP II

Aus dem OLAP-Würfel können folgende Sub-Tabellen "herausgelöst" werden

1. 2-dimensionale Tabellen
2. 2-dimensionale Mehrfach-Tabellen
3. Partial- bzw. Interaktions-Tabellen
4. multidimensionale Kontingenztabellen

Der OLAP-Würfel muss ursprünglich mit Prog11m1 oder Prog11m2 oder Prog11m4 gebildet worden sein

Berechnet werden standardmäßig folgende Koeffizienten:
Chi-Quadrat, Chi-Quadrat-Test und 2I-Test auf allseitige Abhängigkeit
Kontingenzkoeffizient C(cor), Cramers U

Zu 4. kann (optional) die Tabelle der Erwartungswerte zum Chi-Quadrat, die Tabelle der Chi-Quadrat-Beiträge sowie eine Konfigurationsfrequenzanalyse gerechnet werden

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

1

Vereinbare Variable= ;

2

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

3

"C:\Almo7\TESTDAT\Olapdat.nam"

Die Namensdatei muss zuvor mit Prog11m1, Prog11m2, Prog11m4 gebildet worden sein

zeige = Namensdatei in Output zeigen
leer = nicht

4

erzeuge zusätzliche Namensfelder

5

"C:\Almo7\TESTDAT\Olapdat.dir"

Die Datei muss zuvor mit Prog11m1, Prog11m2, Prog11m4 gebildet worden sein

6

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11

erzeuge zusätzliche Felder für Tabellen-Angaben

7

2-dimensionale "Mehrfach-Tabelle" **Hilfe**
 aus OLAP-Würfel auslösen

BEACHTTE: Nur 1 Spaltenvariable erlaubt

Beispiel:

		Schulbildung	
		niedrig	hoch
Geschl.	männlich	16	34
	weiblich	8	11
Wohnort	Stadt	12	15
	Land	12	30
Beruf	Arbeiter	12	12
	Angestel	12	33

Beruf + Geschlecht + Wohnort mit Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

8

Partielltabelle **Hilfe**
 Interaktionstabelle
 aus OLAP-Würfel auslösen

BEACHTTE: Nur 1 Spaltenvariable erlaubt

Beispiel:

Geschl	Ort	Beruf	Schulbildung		Summe
			niedrig	hoch	
männl	Stadt	Arbeit	4	4	8
		Angest	4	4	8
	Land	Arbeit	4	4	8
		Angest	4	22	26
weibl	Stadt	Arbeit	2	2	4
		Angest	2	5	7
	Land	Arbeit	2	2	4
		Angest	2	2	4
Summe			24	45	69

Beruf * Geschlecht * Wohnort mit Schulbildung

erzeuge zusätzliche Felder für Tabellen-Angaben

Erläuterungen zu den Boxen:

Box 1 und 2: Siehe Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.1

Box 3: Datei der Variablennamen



Siehe dazu auch P0.3

Geben Sie hier den Namen der Datei an, den Sie in Prog11m1 bzw. Prog11m2 bzw. Prog11m4 für das Speichern der multiplen Kontingenztabelle in der Box "Option: Nur für multiple Kontingenztabellen" eingegeben haben. Dabei muß die Endung ---.nam hinzugefügt werden. In Prog11m1 haben Sie beispielsweise als Dateiname eingegeben

"C:\Almo\Progs\Olapdat"

Dann müssen Sie hier als Dateiname für die Variablennamen schreiben

"C:\Almo\Progs\Olapdat.nam"

Box 4: Freie Namensfelder

Siehe dazu auch Almo-Dokument Nr. 0 "Arbeiten mit Almo", P0.3

Hier können Sie die in der Datei "C:\Almo\Progs\Olapdat.nam" (in unserem Beispiel) enthaltenen Variablen- und Ausprägungsnamen verändern. Zusätzliche Namen für Variable, die nicht im OLAP-Würfel enthalten sind, sollten nicht eingesetzt werden. Sie machen keinen Sinn.

Box 5: Datei aus der der "OLAP-Würfel" eingelesen wird



Siehe dazu auch P0.4

Hier gilt entsprechendes wie oben zu Box 3. Geben Sie hier den Namen der Datei an, den Sie in Prog11m1 bzw. Prog11m2 bzw. Prog11m4 für das Speichern der multiplen Kontingenztabelle in der Box "Option: Nur für multiple Kontingenztabellen" eingegeben haben. Dabei muß die Endung ---.dir hinzugefügt werden. In Prog11m1 haben Sie beispielsweise als Dateiname eingegeben

"C:\Almo\Progs\Olapdat"

Dann müssen Sie hier als Dateiname für den "OLAP-Würfel" schreiben

"C:\Almo\Progs\Olapdat.dir"

Box 6: 2-dimensionale Tabelle aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.2

Box 7: 2-dimensionale "Mehrfach-Tabelle" aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.3

Box 8: Partialtabellen (Interaktionstabelle) aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.4

Box 9: Variable aus Tabellenangabe
Hier ist keine Eingabe durch den Benutzer erforderlich. Also setzt die Variablennummern in dem Augenblick, in dem der Benutzer auf den Knopf „Rechne“ klickt, selbst ein.

Box 10: Multidimensionale Kontingenztabelle aus eingegebener Tabelle auslösen
Siehe Abschnitt P11.1.5

Box 11: Option: Nur für multidimensionale Kontingenztabelle
Siehe Abschnitt P11.2 Erläuterung zu Box 12

Box 12: Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix
Siehe P0.9

Box 13: Grafik-Optionen
Siehe P0.10

P11.7.3 Ergebnisse aus Prog11m3

Siehe hierzu die Abschnitte P11.4.1 bis P11.4.4.

Literatur

Bortz J., Lienert G. A, Boehnke K.: Verteilungsfreie Methoden in der Biostatistik;
Springer Verlag: Berlin-Heidelberg 1990

G.A. Lienert: Verteilungsfreie Methoden in der Biostatistik, Meisenheim am Glan,
1978

Krauth, J./ Lienert, G.A.: KFA - Die Konfigurationsfrequenzanalyse, Freiburg,
München, 1973

Krauth, J.: Einführung in die Konfigurationsfrequenzanalyse, Beltz Verlag,
Weinheim, 1993

Lohse, Ludwig,Röhr: Statistische Verfahren, Berlin (Ost), 1986

M.J.A. Berry, G.Linoff: Data Mining Techniques, Wiley, New York, 1997, Chapter
16: Where does OLAP fit in ?