CENTRE FOR INTERNATIONAL STUDENT ASSESSMENT

1 Technische Universität München

2 Universität der Bundeswehr München

Parametrische und non-parametrische Klassifikation

Testung der Modellgültigkeit von Latent-Class-Modellen mit der Konfigurations-Frequenz-Anlalyse

Joerg-Henrik Heine [1] & Christian Tarnai [2]

Einleitung

Mit der Anwendung der Latent-Class-Analyse (LCA - z.B. Forman, 1984) und der Konfigurations-Frequenz-Anlayse (KFA - z.B. Lienert, 1971; Stemmler, 2014), werden bei der Daten Auswertung ähnliche Ziele verfolgt. Bei beiden Verfahren geht es darum innerhalb einer Stichprobe, charakteristische Subgruppen (bei der LCA) bzw. Typen (bei der KFA) anhand beobachteter Merkmale zu klassifizieren. Die Vorteile von LCA und KFA, gegenüber anderen Verfahren, liegen dabei in deren Formulierung als statistisch testbare Modelle nach dem Kriterium der lokalen stochastischen Unabhängigkeit (Forman, 1984, S. 205; Stemmler, 2014). Trotz ihrer Ähnlichkeit hinsichtlich der verfolgten Analyseziele, werden beide Verfahren (LCA und KFA) jedoch in der praktischen Anwendung selten komplementär zur Datenanalyse eingesetzt (vgl. Lautsch & Plichta, 2005).
In der vorliegenden Arbeit soll, über einen solchen komplementären Einsatz der beiden Verfahren hinausgehend, die Anwendbarkeit der KFA als eine alternative Methode zur Modelltestung beziehungweise Validierung von LCA-Modellen untersucht werden.

Forschungsfragen

1. Methodenvergleich (KFA - LCA) zur Identifikation von Personenheterogenität bzw. Identifikation bedeutsamer Subgruppen.
2. Lassen sich LCA Modell anhand des Kriteriums der lokalen stochastischen Unabhängigkeit mit der KFA validieren?

Methode

Stichprobe

- Studenten der Universität der Bundeswehr (UniBW - Jahrgänge 2007 - 2011); Sozialwissenschaftliche Fächer; (n = 1343).
- 77 % männlich; Alter mit 8 Kategorien erfasst *Mo* = 23 Jahre.

Instrument

- Big Five Inventory (BFI): Persönlichkeitsinventar (UniBW Version - Schmolck, 2006) vergleichbar mit BFI-K (Rammstedt & John, 2005).

Analysen

Forschungsfrage 1 I) Berechnung von LCA-Modellen für die einzelnen Skalen mit dem R-Pakte ***poLCA*** (Linzer & Lewis, 2011). II) Berechnung der Haupteffekt-KFA für die einzelnen Skalen mit dem R-Pakte ***confreq*** (Heine, 2015). III) Kreuztabulierung der Klassifikationsergebnisse.

Forschungsfrage 2 Überprüfen der lokalen stochastischen Unabhängigkeit innerhalb der latenten Klassen durch Haupteffekt und funktionale KFA-Modelle (Victor, 1989) mit dem R-Paket ***confreq*** (Heine, 2015).

Deskriptive Ergebnisse – Pattern-Häufigkeiten

Die Betrachtung der absoluten Häufigkeiten der Antwort Pattern für jede der 5 Dimensionen zeigt, dass von den jeweils 256 möglichen, einige Pattern in der Stichprobe nicht beobachtbar sind.

Table 1: Anzahl unterschiedlicher Pattern aus 256 möglichen für 5 Dimensionen der Persönlichkeit des BFI, n = 1340.

	Verträglichkeit (A)	Gewissenhaftigkeit (C)	Extraversion (E)	Neurotizismus (N)	Offenheit (O)
Nicht beobachtet	70	120	85	51	71
Beobachtet	186	136	171	205	185
Mögliche	256	256	256	256	256

Ergebnisse

Forschungsfrage 1

Der relative Modellvergleich bei der LCA (1-9 Klassen) resultiert für die 5 BFI-Skalen nach dem Kriterium BIC in 2 (Verträglichkeit) bis 4 (Extraversion) latenten Klassen. Die mittlere modale Zuordnungswahrscheinlichkeit des Klassifikationsergebnisses reicht von $r_{modr\,class}$ = .67 bis $r_{modr\,class}$ = .97; vgl. Tabelle 2.

Table 2: Relativer Modellvergleich: LCA-Modelle; n = 1340.

	BIC Anzahl Klassen	Klassenlösung nicht konvergiert	Klasse 1 $r_{modr\,class}$	Klasse 1 n	Klasse 2 $r_{modr\,class}$	Klasse 2 n	Klasse 3 $r_{modr\,class}$	Klasse 3 n	Klasse 4 $r_{modr\,class}$	Klasse 4 n
A	2	7; 9	.86	584	.83	756	-	-	-	-
C	3	6	.87	280	.86	675	.90	385	-	-
E	4	-	.83	493	.87	287	.89	229	.67	331
N	3	8	.87	207	.87	463	.79	670	-	-
O	3	9	.97	463	.81	316	.85	561	-	-

Anmerkungen: Berechnung der 1 bis 9 Klassenlösung für jede BFI-Dimension.

Bei den Berechnungen der Haupteffekt KFA (jeweils für die 5 BFI-Skalen) resultieren überwiegend Typen (8 Verträglichkeit bis 21 Extraversion) und keine oder eine geringe Anzahl von Antitypen (0 Verträglichkeit bis 4 Gewissenhaftigkeit). Tabelle 3 gibt die Ergebnisse der Kreuztabulierung der gefundenen Typen und Antitypen aus der KFA (Spalten) mit dem Klassifikationsergebnis der LCA (Zeilen) wieder. Die jeweiligen Latenten Klassen lassen sich mit den über die KFA gefundenen Typen anhand der typischerweise gewählten Antwortkategorien charakterisieren (vgl. Tab. 4 für die Dimension Verträglichkeit, A).

Tabelle 3: Anzahl unterschiedlicher Pattern (KFA) nach latenten Klassen (LCA).

	A		C		E		N		O	
	T	A	T	A	T	A	T	A	T	A
Klasse 1	5	0	7	0	3	0	3	0	7	0
Klasse 2	3	0	3	3	7	0	6	0	6	0
Klasse 3	-	-	5	1	7	0	2	1	3	1
Klasse 4	-	-	-	-	4	0	-	-	-	-

Anmerkungen: T = Typ; A = Antityp; ns. = nicht signifikant

Tabelle 4: Anzahl unterschiedlicher Pattern (KFA) nach latenten Klassen (LCA) für die BFI-Dimension *Verträglichkeit* (A).

Pattern	1 2 1 1 1	2 1 1 1 1	2 2 1 1 1	2 2 2 1 1	3 1 2 1 1	3 3 3 3 3	4 3 3 4	4 4 4 3	x x x x x
Klasse 1	7	9	12	22	16	0	0	0	518
Klasse 2	0	0	0	0	0	41	36	14	665
Mean Prob.	1.00	1.00	0.99	0.99	0.99	0.95	0.98	0.99	0.83

Forschungsfrage 2

Zur Überprüfung der lokalen stochastischen Unabhängigkeit innerhalb der latenten Klassen werden für jede der nach modaler Klassenzuordnungswahrscheinlichkeit definierten Personenklassen Haupteffekt-KFA (Tab. 5) und funktionale KFA-Modelle (Victor, 1989) (Tab. 6) für die BFI-Dimension *Verträglichkeit* (A) mit dem R-Paket ***confreq*** (Heine, 2015) gerechnet.
Bei der funktionalen KFA werden die durch die Anwendung der Haupteffekt-KFA auf die Gesamtstichprobe gefundenen, die latenten Klassen charakterisierenden, signifikanten Typen (vgl. Forschungsfrage 1, und Tab 4) bei der Modelltestung herausgenommen (Gleichsetzen von erwarteten und beobachteten Häufigkeiten).

Tabelle 5: Haupteffekt KFA für BFI-Dimension *Verträglichkeit*.

	pat.	obs.	exp.	Type	Chi	df	pChi	ex.bin.test	z.Chi	z.pChi	z.Bin	z.pBin
Klasse 1	2 3 4 1	10	1,393832	$+^2$	53,138514	1	0,000000	0,000002	7,289617	0,000000	7,298331	0,000000
	3 2 2 3	15	3,2345	$+^2$	42,797024	1	0,000000	0,000001	6,541943	0,000000	6,560135	0,000000
	3 2 3 2	27	6,980601	$+^2$	57,412867	1	0,000000	0,000000	7,577128	0,000000	7,622823	0,000000
Klasse 2	1 1 1 1	0	0,000127	$-^1$	0,000127	1	0,991003	0,000127	-0,011276	0,495502	-0,011276	0,495502
	2 3 3 2	19	5,392828	$+^2$	34,333592	1	0,000000	0,000004	5,859487	0,000000	5,880499	0,000000
	3 3 1 3	11	2,393407	$+^2$	30,948955	1	0,000000	0,000040	5,563179	0,000000	5,572006	0,000000
	3 3 3 1	32	15,815736	$+^2$	16,561378	1	0,000047	0,000190	4,069567	0,000024	4,112815	0,000020

Anmerkungen: Typen / Antitypen nach exaktem binomial Test; Bonferroni adj. Alpha: α_{adj} = 0.0001953125;
1) Antityp; 2) Typ.

Tabelle 6: Funktionale KFA für BFI-Dimension *Verträglichkeit* - Typen aus KFA für Gesamtstichprobe gleichgesetzt (erwartet = beobachtet).

	pat.	obs.	exp.	Typ	Chi	df	pChi	ex.bin.test	z.Chi	z.pChi	z.Bin	z.pBin
Klasse 1	2 3 4 1	10	1,418129	$+^2$	51,933596	1	0,000000	0,000002	7,206497	0,000000	7,215263	0,000000
	3 2 2 3	15	3,318808	$+^2$	41,114242	1	0,000000	0,000002	6,412039	0,000000	6,430336	0,000000
	3 2 3 2	27	7,114782	$+^2$	55,577516	1	0,000000	0,000000	7,455033	0,000000	7,500864	0,000000
	1 2 1 1	7	7	b^3	-	-	-	-	-	-	-	-
	2 1 1 1	9	9	b^3	-	-	-	-	-	-	-	-
	2 2 1 1	12	12	b^3	-	-	-	-	-	-	-	-
	2 2 2 1	22	22	b^3	-	-	-	-	-	-	-	-
	3 1 2 1	16	16	b^3	-	-	-	-	-	-	-	-
Klasse 2	1 1 1 1	0	0,000138	$-^1$	0,000138	1	0,990635	0,000138	-0,011738	0,495317	-0,011738	0,495317
	2 3 3 2	19	5,483045	$+^2$	33,322377	1	0,000000	0,000005	5,772554	0,000000	5,793602	0,000000
	3 3 1 3	11	2,292027	$+^2$	33,083726	1	0,000000	0,000027	5,751845	0,000000	5,760585	0,000000
	3 3 3 1	32	15,621934	$+^2$	17,170796	1	0,000034	0,000154	4,143766	0,000017	4,187254	0,000014
	3 3 3 3	41	41	b^3	-	-	-	-	-	-	-	-
	4 3 3 4	36	36	b^3	-	-	-	-	-	-	-	-
	4 4 4 3	14	14	b^3	-	-	-	-	-	-	-	-

Anmerkungen: Typen / Antitypen nach exaktem binomial Test; Bonferroni adj. Alpha: α_{adj} = 0.0001953125;
1) Antityp; 2) Typ; b) nicht getestet (erwartet = beobachtet).

Diskussion

- Forschungsfrage 1:
  - Die mit der KFA gefundenen signifikanten, *primären* Typen lassen sich eindeutig jeweils einer der mit der LCA gefundene Klassen zuordnen und charakterisieren diese durch die jeweils typischerweise gewählten Antwortkategorien.
- Forschungsfrage 2:
  - Durch Anwendung der Haupteffekt- und funktionalen KFA lassen sich „innerhalb“ der latenten Klassen nach dem exakten Binomial-Test andere signifikante, zusätzliche *sekundäre* Typen nachweisen.
  - Die zuvor für die gesamte Stichprobe gefundenen primären Typen (vgl. Forschungsfrage 1), werden bei der Haupteffekt KFA „innerhalb“ der latenten Klassen nicht mehr signifikant.
  - Der Befund zusätzlicher *sekundäre* Typen bestätigt die auch von anderen Autoren (z.B. Stemmler, 2014, S. 68) gefundene „erhöhte Sensitivität“ der KFA.

Literatur

Heine, J.-H. (2015). *confreq: Configural Frequencies Analysis Using Log-linear Modeling (V 1.3-1)*. Abgerufen von http://cran.r-project.org/web/packages/confreq/index.html

Krauth, J., & Lienert, G. A. (1973). *Die Konfigurationsfrequenzanalyse (KFA) und ihre Anwendung in Psychologie und Medizin: ein multivariates nichtparametrisches Verfahren zur Aufdeckung von Typen und Syndromen; mit 70 Tabellen*. Freiburg; München: Alber Karl.

Lienert, G. A. (1971). Die Konfigurationsfrequenzanalyse: I. Ein neuer Weg zu Typen und Syndromen. *Zeitschrift für Klinische Psychologie und Psychotherapie, 19*(2), 99–115.

Linacre, J. M. (1998). Detecting multidimensionality: which residual data-type works best? *Journal of outcome measurement, 2*, 266–283.

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software, 42*(10), 1–29.

Stemmler, M. (2014). *Person-Centered Methods*. Cham: Springer International Publishing.

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): *Diagnostica, 51*(4), 195–206. http://doi.org/10.1026/0012-1924.51.4.195

Victor, N. (1989). An Alternativ Approach to Configural Frequency Analysis. *Methodika, 3*, 61–73.

Schmolck, P. (2006). *ESF–Projekt WT06 Individuelle Testergebnisse* (unveröffentlichter Projektbericht No. 3). Neubiberg: Universität der Bundeswehr München.

Kontakt: joerg.heine@tum.de