

Diskussion

Zur Anwendungsindikation von nonparametrischen Prädiktionsverfahren

Diskussionsbeitrag zu SCHULZE (1978): «Ein Verfahren zur multivariaten Analyse der Bedingungen von Rangvariablen: Hierarchische Rangvarianzanalyse»

HANS-PETER KRÜGER

Fachbereich Erziehungs- und Kulturwissenschaften der Universität Erlangen-Nürnberg, Lehrstuhl Psychologie

SCHULZE schlug 1978 an dieser Stelle ein Verfahren zur hierarchischen Prädiktion eines ordinalen Prädikanden vor. Seine Methode wird an einem Beispiel in Abhebung von zwei anderen Verfahren mit nur nominalen Voraussetzungen diskutiert. Es zeigt sich, daß die Anwendung des Verfahrens sensu SCHULZE mit einigen Restriktionen verbunden ist. Vor allem ist ein Satz von mehrdimensional orthogonalen Prädiktoren Voraussetzung für eine interpretierbare Auswertung. Da dies in praxi in aller Regel nicht gegeben ist, wird zur Bestimmung von optimalen Prädiktorensätzen die Assoziationsstrukturanalyse nach KRAUTH & LIENERT (1973) empfohlen.

SCHULZE (1978) suggested a new method to predict an ordinal variable out of a set of predictors in hierarchical manner. His method is contrasted with two other methods demanding only nominal scaled variables. It is shown that the method sensu SCHULZE is restricted to predictors having no two or higher-dimensional interactions. Such data hardly could be found in empirical work. The association structure analysis (KRAUTH & LIENERT, 1973) is recommended as more appropriate.

1. Fragestellung

SCHULZE (1978) hat jüngst an dieser Stelle ein Verfahren vorgeschlagen, mit dem aus einem Satz von Prädiktoren ein ordinaler Prädikand bestimmt werden kann (Beispiel: Voraussage des Schulerfolgs aus den Prädiktoren Intelligenz, Geschlecht und Sozialschicht). SCHULZE siedelt sein Verfahren zwischen den parametrischen Verfahren mit intervallskaliertem Prädikanden und den Verfahren an, die nur nominales Skalenniveau des Prädikanden verlangen. Prädiktionsverfahren dieser Art sind von großem forschungspraktischen Interesse, da sie neben der Möglichkeit der Hypothesenprüfung auch die Chance einer gerichteten Heuristik eröffnen. Eine Inspektion des Verfahrens von SCHULZE erbringt einige Restriktionen in der Anwendung. Diese Restriktionen sollen im Vergleich zu zwei anderen Verfahren diskutiert werden, die nur nominale Voraussetzungen auf der Seite des Prädikanden benötigen:

- a) der Prädiktions-Konfigurationsfrequenzanalyse (PKFA) von LIENERT & KRAUTH (1973) und
- b) der hierarchischen Konfigurationsfrequenzanalyse bzw. der Assoziationsstrukturanalyse (ASA) von KRAUTH & LIENERT (1973a).

Zugrundegelegt wird ein Datensatz, wie ihn SCHULZE selbst hypothetisch vorschlägt (a. a. O., p. 137). TENT et al. (1974) haben die Abhängigkeit der Schulnoten von Schülermerkmalen an Volksschülern untersucht. Im folgenden werden aus diesem Material (in Anlehnung an das Beispiel bei LIENERT & KRAUTH, 1973) ausgewählt:

- als Prädikand: die Note im Fach Deutsch mit 3 Stufen (1-2 → N+, 3-4 → N0, 5-6 → N-)
- als Prädiktoren: das Geschlecht (M = männlich, W = weiblich) der Sozialstatus des Vaters (S+ → hoch, S0 → mittel und S- → nieder)
- die Leistung im Rechtschreibtest AST-4 mit R+ → hoch und R- → gering

In die Erhebung gehen $N = 1839$ Volksschüler ein. Die Ausgangslage zeigt Tabelle 1a, wobei hier nur die Häufigkeiten in Fettschrift (linker Teil der

Tabelle) interessieren. Die Tabellen 1b-1h zeigen die Untertafeln, die entstehen, wenn einzelne Prädiktoren weggelassen werden (Hierarchie).

Tab. 1: Prädiktion der Deutschnote N aus den Prädiktoren Geschlecht G, Sozialschicht S und Rechtschreibleistung R.

Tab. 1a: Prädiktion von N aus G, S und R.

Konfigu	N+ beobachtete Werte	N0	N- Su	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2			
M S+ R-	11	90	36	137	40.0 21.0	74.8 3.1	22.2 8.6	42.7 23.5	79.8 1.2	23.7 6.3	e X^2
M S+ R+	89	78	2	169	49.3 2.2	92.3 23.5	27.4 55.8	43.0 0.0	80.4 20.0	23.8 47.2	e X^2
M S0 R-	12	162	93	267	78.0 1.8	145.8 57.2	43.3 15.8	69.2 8.2	129.4 77.5	38.4 1.8	e X^2
M S0 R+	81	92	6	179	52.3 15.8	97.7 0.3	29.0 18.2	69.7 1.8	130.3 11.3	38.7 27.6	e X^2
M S- R-	2	65	52	119	34.7 30.9	65.0 0.0	19.3 55.5	24.9 21.0	46.6 7.2	13.8 105.2	e X^2
M S- R+	27	41	2	70	20.4 2.1	38.2 0.2	11.3 7.7	25.1 0.1	46.9 0.7	13.9 10.2	e X^2
W S+ R-	14	66	10	90	26.3 5.7	49.1 5.8	14.6 1.4	40.7 17.5	76.2 1.3	22.6 7.0	e X^2
W S+ R+	116	62	0	178	52.0 78.9	97.2 12.7	28.8 28.8	41.0 136.6	76.8 2.8	22.7 22.7	e X^2
W S0 R-	17	142	56	215	62.8 32.4	117.4 5.2	34.8 12.9	66.0 36.4	123.4 2.7	36.6 10.2	e X^2
W S0 R+	138	125	6	269	78.5 45.0	146.9 3.3	43.6 32.4	66.5 76.6	124.4 0.0	36.9 25.9	e X^2
W S- R-	8	46	34	88	25.7 12.2	48.0 0.1	14.3 27.3	23.7 10.4	44.4 0.0	13.2 32.7	e X^2
W S- R+	22	35	1	58	16.9 1.5	31.7 0.4	9.4 7.5	23.9 0.1	44.8 2.1	13.3 11.3	e X^2

Tab. 1b: Prädiktion von N aus G und S.

Konfigu	N+ beobachtete Werte	N0	N- Su	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2			
M S+ .	100	168	38	306	89.4 1.3	167.1 0.0	49.6 2.7	85.7 2.3	160.3 0.3	47.5 1.9	e X^2
M S0 .	93	254	99	446	130.2 10.6	243.5 0.5	72.3 9.9	138.9 15.1	259.8 0.1	77.1 6.2	e X^2
M S- .	29	106	54	189	55.2 12.4	103.2 0.1	30.6 17.8	50.0 8.8	93.5 1.6	27.7 24.7	e X^2
W S+ .	130	128	10	268	78.3 34.2	146.3 2.3	43.4 25.7	81.8 28.3	153.0 4.0	45.4 27.6	e X^2
W S0 .	155	267	62	484	141.3 1.3	264.2 0.0	78.4 3.4	132.6 3.7	247.9 1.4	73.5 1.8	e X^2
W S- .	30	81	35	146	42.6 3.7	79.7 0.0	23.7 5.4	47.7 6.6	89.3 0.7	26.5 2.7	e X^2

Tab. 1c: Prädiktion von N aus G und R.

Konfigu	N+ beobachtete Werte	N0	N-	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2			
M . R-	25	317	181	523	152.7 106.8	285.5 3.5	84.7 109.3	136.8 91.4	255.8 14.5	75.9 145.2	e X^2
M . R+	197	211	10	418	122.1 46.0	228.2 1.3	67.7 49.2	137.9 25.3	257.8 8.5	76.5 57.8	e X^2
W . R-	39	254	100	393	114.8 50.0	214.6 7.3	63.7 20.7	130.6 64.2	244.1 0.3	72.4 10.4	e X^2
W . R+	276	222	7	505	147.5 112.0	275.7 10.5	81.8 68.4	131.6 158.4	246.0 2.3	73.0 59.7	e X^2

Tab. 1d: Prädiktion von N aus S und R.

Konfigu	N+ beobachtete Werte	N0	N-	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2			
. S+ R-	25	156	46	227	66.3 25.7	123.9 8.3	36.8 2.3	83.4 40.9	156.0 0.0	46.3 0.0	e X^2
. S+ R+	205	140	2	347	101.3 106.1	189.4 12.9	56.2 52.3	84.1 173.6	157.2 1.8	46.6 42.7	e X^2
. S0 R-	29	304	149	482	140.7 88.7	263.1 6.3	78.1 64.3	135.2 83.4	252.8 10.3	75.0 72.8	e X^2
. S0 R+	219	217	12	448	130.8 59.4	244.6 3.1	72.6 50.6	136.2 50.1	254.8 5.6	75.6 53.5	e X^2
. S- R-	10	111	86	207	60.4 42.1	113.0 0.0	33.5 82.6	48.7 30.7	91.0 4.3	27.0 128.5	e X^2
. S- R+	49	76	3	128	37.4 3.6	69.9 0.5	20.7 15.2	49.0 0.0	91.7 2.7	27.2 21.5	e X^2

Tab. 1e: Prädiktion von N aus G.

Konfigu	N+ beobachtete Werte	N0	N-	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2				
M . .	222	528	191	941	274.7 10.1	513.7 0.3	152.4 9.7	wie bei der Prädiktions-KFA				e X^2
W . .	315	476	107	898	262.2 10.6	490.2 0.4	145.5 10.1	wie bei der Prädiktions-KFA				e X^2

Tab. 1f: Prädiktion von N aus S.

Konfigu	N+ beobachtete Werte	N0	N-	Su	Prädiktions-KFA Erwartung und X^2			Hierarchische KFA Erwartung und X^2				
. S+ .	230	296	48	574	167.6 23.2	313.3 0.9	93.0 21.7	wie bei der Prädiktions-KFA				e X^2
. S0 .	248	521	161	930	271.5 2.0	507.7 0.3	150.7 0.7	wie bei der Prädiktions-KFA				e X^2
. S- .	59	187	89	335	97.8 15.4	182.8 0.0	54.2 22.2	wie bei der Prädiktions-KFA				e X^2

Tab. 1g: Prädiktion von N aus R.

Konfigu	N+ beobachtete Werte	N0	N-	Su	Prädiktions-KFA Erwartung und X ²			Hierarchische KFA Erwartung und X ²	e X ²
R-	64	571	281	916	267.4 <i>154.7</i>	500.0 <i>10.0</i>	184.4 <i>118.3</i>	wie bei der Prädiktions-KFA	e X ²
R+	473	433	17	923	269.5 <i>153.6</i>	503.9 <i>9.9</i>	149.5 <i>117.4</i>	wie bei der Prädiktions-KFA	e X ²

Tab. 1h: Häufigkeiten von G, S, R und N.

	Stufen				
	1	2	3	Σ	
G	941	898		1839	Geschlecht
S	574	930	335	1839	Sozialschicht des Vaters
R	916	923		1839	Rechtschreiben im AST-4
N	537	1004	298	1839	Deutschnote

(«Konfigu» = Prädiktor-Konfiguration. Die beobachteten Werte sind in fetter Schrift, die Erwartungswerte in Normalschrift, die X²-Werte in Kursivschrift in die Tabellen eingetragen.)

2. Die Auswertung nach den drei Verfahren

2.1 Auswertung nach der Hierarchischen Rangvarianzanalyse (HRVA)

Der Grundgedanke der HRVA ist einfach: die

Tab. 2: Rangvarianzanalytische Zerlegung der Gruppe der schlechten Rechtschreiber (R-) nach dem Prädiktor Sozialschicht (S+, S0 und S-).

Konfigu	N+	N0	N-	Σ
R- S+	25	156	46	227
R- S0	29	304	149	482
R- S-	10	111	86	207
Σ	64	571	281	916
mittlerer Rang	32.5	350	776	

$$V_{R-S} = \sum_j n_{.j} R_j^2 - \frac{\sum (n_{.j} R_j)^2}{n_{..}} = (64 \cdot 32.5^2 + 571 \cdot 350^2 + 281 \cdot 776^2) - [(64 \cdot 32.5)^2 + (571 \cdot 350)^2 + (281 \cdot 776)^2] / 916 = 46\ 662\ 975$$

$$V_{R-S_i} = \sum_j n_{ij} R_j^2 - \frac{\sum (n_{ij} R_j)^2}{n_{.i}} \text{ (für alle } i)$$

für S+: $V_{R-S+} = (25 \cdot 32.5^2 + 156 \cdot 350^2 + 46 \cdot 776^2) - [(25 \cdot 32.5)^2 + (156 \cdot 350)^2 + (46 \cdot 776)^2] / 227 = 10\ 269\ 283.0$

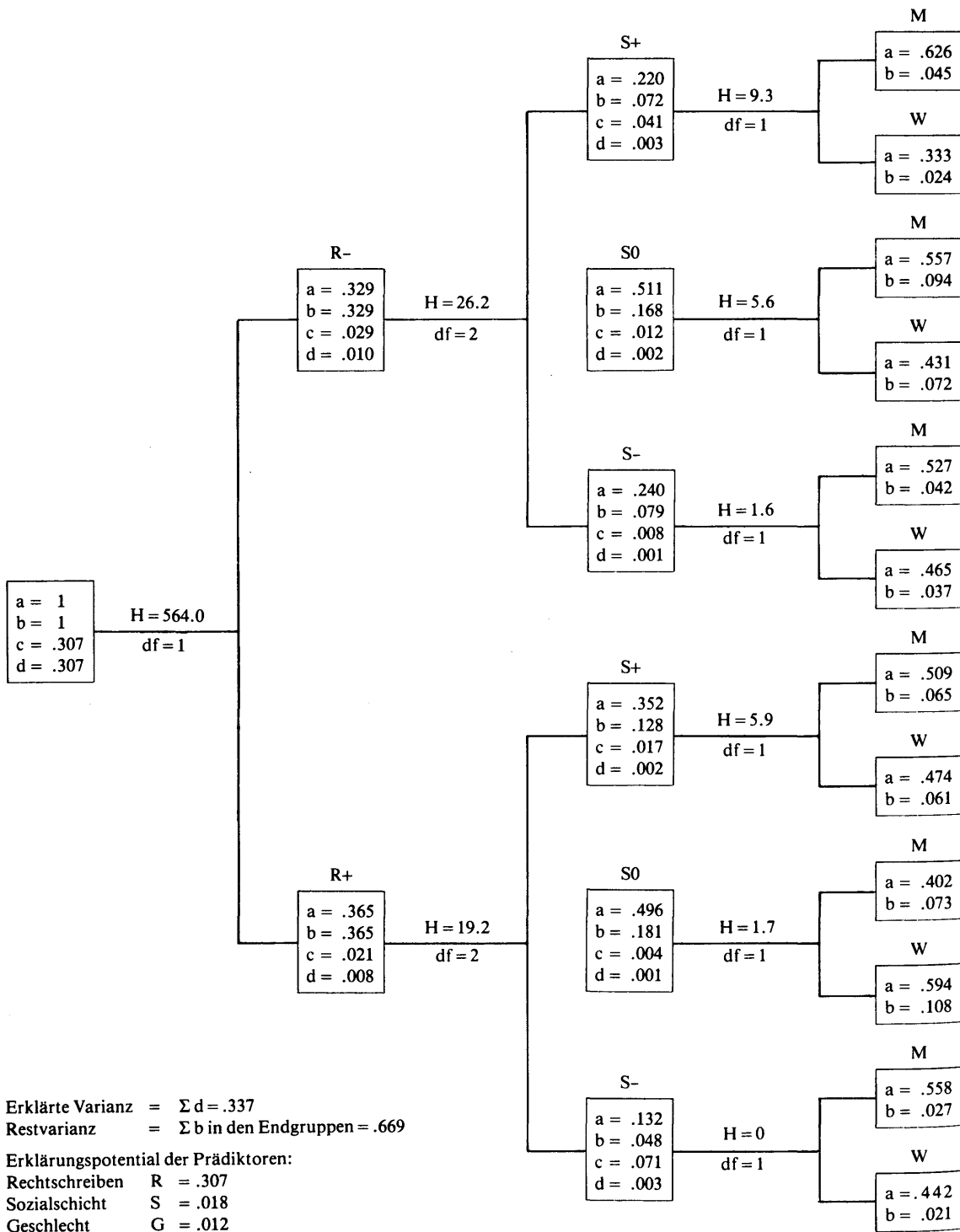
Zusammenstellung der Varianzen:

Total V _{R-S}	46 662 975.0
für S+ V _{R-S+}	10 269 283.0
für S0 V _{R-S0}	23 853 651.7
für S- V _{R-S-}	11 206 116.8
innerhalb	45 329 051.4
zwischen = Total - Σ V _{R-S_i}	1 333 923.6
erklärte Varianz = zwischen/Total	0.028586
H = erklärte Varianz x (n _. - 1)	26.156 df = 2

Stichprobe wird nach einer Sequenz von Prädiktoren zerlegt. Über jede dieser Zerlegungen wird ein H-Test nach KRUSKAL & WALLIS (1952) gerechnet. Ein Beispiel zeigt Tabelle 2. Aus Tabelle 1d wird die Gruppe der schlechten Rechtschreiber (R-) nach der Sozialschicht aufgeteilt. Für diese Untertafel werden zuerst die mittleren Ränge berechnet. Bezeichnet man die Spaltensummen mit n_{.j}, errechnet sich der mittlere Rang R_j für Spalte i nach der Formel

$$R_i = \sum_j^{i-1} n_j + \frac{n_i + 1}{2}$$

Tab. 3: Tafel der hierarchischen Rangvarianzanalyse – Prädiktorenfolge Rechtschreiben (R), Sozialschicht (S) und Geschlecht (G).



Beispiel: für N_0 ergibt sich in Tabelle 2 der mittlere Rang zu $R_{N_0} = 64 + (571 + 1)/2 = 350$

Nach den in Tabelle 2 angegebenen Formeln werden mit diesen mittleren Rängen die Varianzen errechnet. Es bedeuten

V_{R-S} das Gesamt der Varianz bei den schlechten Rechtschreibern

V_{R-S_i} der Anteil der Varianz, der durch die Aufteilung dieser Gruppe nach der Sozialschicht S_i erklärt wird.

Hinter dem Verfahren verbirgt sich so nichts anderes als eine ANOVA aus mittleren Rängen. Bekanntlich ist bei Rangwerten die Prüfgröße

$$[SS_{\text{treatment}}/SS_{\text{total}}] \cdot df_{\text{total}}$$

der H-Wert von KRUSKAL & WALLIS, der asymptotisch χ^2 -verteilt ist mit $k-1$ Freiheitsgraden (siehe etwa WINER, 1971, p. 849).

Für jede Hierarchiestufe (für jede Unterteilung einer Gruppe nach einem neuen Prädiktor) wird eine neue einfaktorielle ANOVA berechnet. Das bedeutet, daß die Rangmittel für die verschiedenen Untergruppen verschieden sind. *Damit aber entfällt per definitionem eine Prüfung der Wechselwirkungen.*

In Tabelle 3 steht das Ergebnis der HRVA für die Prädiktorenfolge Rechtschreiben - Sozialschicht - Geschlecht. Das Berechnungsbeispiel der Tabelle 2 erbrachte für die Gruppe R-S ein $H = 26.156$, das als $H = 26.2$ in Tabelle 3 zu finden ist. Die einzelnen Abkürzungen in der Tabelle 3 bedeuten:

$a = g_{U(O)} = V_U/V_O =$ Anteil der Varianz einer Untergruppe an der Varianz der zugehörigen Obergruppe

$$\text{Beispiel: für } V_{R-S^+} \text{ ist } g_{U(O)} = \frac{10\ 269\ 283}{46\ 662\ 975} = 0.220$$

$b = g_{X(G)} = \text{prod } a_k =$ Anteil der Varianz der Teilgruppe X an der Varianz der Gesamtgruppe. Dabei gibt a_k alle a 's an, die vor der Aufteilung nach dem neuen Kriterium liegen (anschaulich: alle a 's, denen man auf dem Weg zurück in der Hierarchie begegnet).

$$\text{Beispiel: für } V_{R-S^+} \text{ ist } b = 0.220 \cdot 0.329 \cdot 1 = 0.072$$

$c =$ erklärte $V_{X(X)} = V_{\text{zwischen}}/V_{\text{Gesamt}} =$ durch die Aufteilung von Gruppe X erklärter Anteil der Rangvarianz, bezogen auf die Rangvarianz der Obergruppe X

Beispiel: die Aufteilung der schlechten Rechtschreiber nach Sozialschicht in Tabelle 2 erklärt 0.028586 der Varianz der Obergruppe R-. Folglich ist dieser Wert in Tabelle 3 bei R- unter c zu finden (diese Gruppe wurde ja aufgeteilt).

$d =$ erklärte $V_{X(G)} =$ erklärte $V_{X(X)} \cdot g_{X(G)} = c \cdot b =$ durch die Aufteilung der Gruppe X erklärte Varianz, bezogen auf die Varianz der Gesamtgruppe.

Das Gesamt der durch die Aufteilung erklärten Varianz ergibt sich aus der Σd , die sich mit der Σb in den Endgruppen zur Gesamtvarianz von 1 aufsummieren muß. Die einzelnen Rechenschritte und Formeln sind bei SCHULZE zu finden¹. In Tabelle 3 werden 33,7% der Gesamtvarianz der Deutschnote durch die Prädiktoren erklärt, wovon 30,7% auf Kosten des Rechtschreibens, 1,8% auf Kosten der Sozialschicht und 1,2% auf Kosten des Geschlechts gehen. Die einzelnen H-Werte in der Tabelle 3 zeigen:

- massiven präzisierenden Wert hat das Rechtschreiben ($H = 564.0$), weniger starken Sozialschicht ($H = 26.2$) und Geschlecht ($H = 19.2$).
- die weitere Aufteilung der bereits nach Rechtschreiben und Sozialschicht getrennten Gesamtgruppe bringt nicht grundsätzlich weitere Klärung. Nicht signifikant sind bei der weiteren Aufteilung nach Geschlecht die Gruppen
 - b_1 : R-S- ($H = 1.6$)
 - b_2 : R + S0 ($H = 1.7$)
 - b_3 : R + S- ($H = 0$)

Das heißt, daß das Geschlecht nur dann noch prädiktive Wirkung hat, wenn bei guten Recht-

¹ Im Berechnungsbeispiel von SCHULZE (a. a. O., p. 138f.) sind einige Fehler unterlaufen.

- Auf p. 139 links oben fehlt bei der Berechnung von V_{AA} auf dem Bruchstrich das Quadrat der Klammer.
- Darunter muß $V_{AD} = 1596685$ (statt 1600149) sein. Dadurch verändern sich in der Folge V_{Ain} zu 8361617 (statt 8365082) und V_{Azw} zu 2018677 (statt 2015212).
- Auf p. 139 rechts oben wird durch die Berichtigungen $V_{A(A)} = .194472$ (statt .194138). Darunter ist zwar der H-Wert mit 99.56 richtig angegeben, doch entsteht er nicht aus dem Produkt $0.194138 \cdot 512$, sondern aus $0.194472 \cdot 512$.

schreibern hohe Sozialschicht, bei schlechten Rechtschreibern hohe oder mittlere Sozialschicht vorliegt. Nach den Entscheidungsregeln von SCHULZE deuten verschiedene Erklärungsanteile des gleichen Prädiktors auf eine Wechselwirkung hin. Eine Prüfung ist wie erwähnt wegen der Technik der Rangvergabe nicht möglich.

2.2 Auswertung nach der Prädiktions-KFA (PKFA)

Die PKFA beruht auf lediglich nominalen Voraussetzungen bei der Skalierung des Prädikanden. Aus der Kontingenztafel der Tabelle 1a lassen sich für die einzelnen Zellen unter der Hypothese der Unabhängigkeit zwischen Prädiktorenkonfiguration und Prädikanden die Erwartungswerte

$$e = \text{Zeilensumme} \cdot \text{Spaltensumme} / N$$

berechnen. Das ist für alle Untertafeln im Mittelteil der Tabellen 1a-1g geschehen.

Beispiel: Für die Konfiguration MS+R- und die Prädikandenstufe N+ berechnet sich der Erwartungswert aus der Zeilen-

Tab. 4: Ergebnisse aus der hierarchischen Prädiktions-KFA und aus der hierarchischen KFA.

Prädiktorkonfiguration	Prädiktions-KFA Hierarchische KFA		
	X ²	X ²	X ² _{corr}
G S . .	-	8.59 ₂	8.59 ₂
G . R .	-	25.66 ₁	25.66 ₁
G . . N	41.49 ₂	41.49 ₂	41.49 ₂
. S R .	-	44.93 ₂	44.93 ₂
. S . N	86.76 ₄	86.76 ₄	86.76 ₄
. . R N	564.34 ₂	564.34 ₂	564.34 ₂
G S R .	-	81.84 ₇	2.67 ₂
G S . N	131.54 ₁₀	138.68 ₁₂	1.84 ₄
G . R N	585.02 ₆	638.55 ₇	7.06 ₂
. S R N	623.64 ₁₀	723.28 ₁₂	27.25 ₄
G S R N	650.28 ₂₂	816.27 ₂₉	5.69 ₄
		$\Sigma X^2_{\text{corr}} =$	816.27 ₂₉

G = Geschlecht
S = Sozialschicht
R = Rechtschreiben
N = Deutschnote

Die angehängten Zahlen bedeuten die Zahl der Freiheitsgrade.

summe 137 und der Spaltensumme 537 (Zahl der N+) zu $e = 137 \cdot 537 / 1839 = 40.0$.

Als Prüfgröße wird das übliche X² mit (b-e)²/e eingeführt, wobei b die Zahl der tatsächlich beobachteten Fälle in dieser Zelle bedeutet. Das X² für das Beispiel beträgt $X^2 = (11-40.0)/40.0 = 21.0$ und ist kursiv im Mittelteil der Tabelle 1a zu finden. Werden die X² über alle Zellen summiert, ergibt sich ein Gesamt-X² mit (k-1)(l-1) Freiheitsgraden, wobei k die Zahl der Prädiktorenkonfigurationen, l die Zahl der Prädikandenstufen angibt.

Dieses Verfahren kann hierarchisch für alle möglichen Prädiktorenkonfigurationen angewendet werden. Das Ergebnis zeigt Tabelle 4, mittlerer Teil. Für das Geschlecht ergibt sich ein X² = 41.49 mit 2 Freiheitsgraden usw. Alle X²-Werte sind hochsignifikant, so daß von einer hohen Aussagekraft der Prädiktoren für den Prädikanden ausgegangen werden kann. Den höchsten Wert als Einzelprädiktor erhält das Rechtschreiben mit X² = 564.34, was mit den Ergebnissen aus der HRVA korrespondiert.

2.3 Die Auswertung nach der Assoziationsstrukturanalyse (ASA)

Zur Untersuchung höherwertiger Zusammenhänge zwischen Merkmalen haben KRAUTH & LIERNERT (1973a) die ASA entwickelt, die es erlaubt, für jede Interaktion auf jeder Stufe der Hierarchie eine X²-Komponente angeben zu können. Im Unterschied zur PKFA werden bei der ASA die Erwartungswerte unter der *Annahme der totalen Unabhängigkeit* auch zwischen den Prädiktoren berechnet (nicht wie bei der PKFA nur unter der Annahme der Unabhängigkeit zwischen Prädiktorenkonfiguration und Prädikand). Der Erwartungswert ergibt sich dann aus dem Produkt der Auftretenswahrscheinlichkeiten der beteiligten Variablenstufen und wird zusammen mit der beobachteten Häufigkeit b zu einer X²-Komponente verrechnet.

Beispiel: für die Konfiguration M S+ R- N+ in Tabelle 1a errechnet sich der Erwartungswert zu

$$e = \frac{h_M}{N} \cdot \frac{h_{S+}}{N} \cdot \frac{h_{R-}}{N} \cdot \frac{h_{N+}}{N} \cdot N =$$

$$= (941/1839 \cdot 574/1839 \cdot 916/1839 \cdot 537/1839) \cdot 1839 = 42.7.$$

(Die Häufigkeiten sind Tabelle 1h zu entnehmen). X^2 wird damit

$$X^2 = (11 - 42.7)^2 / 42.7 = 23.5$$

Die Erwartungs- und X^2 -Werte sind für alle Konfigurationen der Hierarchie in den Tabellen 1a-1h im rechten Teil zu finden. Die X^2 -Komponenten werden zu einem Gesamt- X^2 aufsummiert. Die Zahl der Freiheitsgrade ergibt sich für k Merkmale mit je k_i Ausprägungen zu

$$df = \text{prd } k_i - \text{sum } k_i + (k - 1)$$

Im Beispiel der Tabelle 1a sind $k=4$ Merkmale mit $k_1=2$, $k_2=3$, $k_3=2$ und $k_4=3$ Ausprägungen, so daß

$$df = 2 \cdot 3 \cdot 2 \cdot 3 - (2 + 3 + 2 + 3) + (4 - 1) = 29$$

In Tabelle 4, rechte Seite, sind die Gesamt- X^2 aus den Tabellen 1a-1g zu finden. In der ASA kann nun nach dem Verfahren von SUTCLIFFE (1957) das Gesamt- X^2 einer höheren Interaktion so korrigiert werden, daß die X^2 aller niederen Interaktionen abgezogen werden. Der verbleibende Rest stellt eine Abschätzung des Erklärungspotentials dar, das durch die Interaktion der Merkmale erbracht wird. Im Beispiel: das Gesamt- X^2 der Merkmale S R N (Tabelle 1d) beträgt 723.28 mit 12 Freiheitsgraden. Darin sind enthalten die Einzelzusammenhänge zwischen S und R, S und N und R und N. Als tatsächlicher Betrag der Interaktion bleibt so die Differenz

$$\begin{aligned} X^2_{SRN} - X^2_{SR} - X^2_{SN} - X^2_{RN} &= \\ = 723.28 - 44.93 - 86.76 - 564.34 &= \\ = 27.25 \end{aligned}$$

Die Zahl der Freiheitsgrade verhält sich ebenfalls additiv, so daß diese Restkomponente mit $df = 12 - 2 - 4 - 2 = 4$ df zu interpretieren ist. Die Tabelle 4 enthält in der letzten Spalte die so korrigierten X^2 -Werte, deren Summe sich zu dem Gesamt- X^2 der Tafel G S R N (Tabelle 1a) aufsummieren muß.

Die ASA zeigt auf, daß

- die Prädiktoren untereinander hoch korreliert sind. Mädchen sind bessere Rechtschreiber (GR), höhere Sozialschicht impliziert ebenfalls höhere Rechtschreibleistung (SR). Bedenklich in bezug auf die Stichprobe ist der Zusammenhang zwischen Geschlecht und Sozialschicht (GS): in den höheren Sozialschichten finden sich mehr Mädchen.
- neben dem von den vorher erwähnten Verfah-

ren bereits bekannten prädiktiven Potential der Einzelprädiktoren auch eine signifikante höherdimensionale Interaktion vorliegt. Die Prädiktoren Sozialschicht und Rechtschreiben erklären *zusammen* einen Teil der Notenzvarianz, *der nicht bereits durch diese beiden Prädiktoren einzeln erklärt wurde* (SRN in Tabelle 4).

Zur Demonstration dieser höheren Interaktion werden die Notenhäufigkeiten der Tabelle 1d in Mittelwerte umgewandelt, wobei N+ als +1, N0 als 0 und N- als -1 gesetzt wird. Die Gruppen erhalten dann folgende Mittelwerte:

S+R-:	(25 (+1) + 156 (0) + 46 (-1))/227	= -.093
S+R+:		= .585
S0R-:		= -.248
S0R+:		= -.462
S-R-:		= -.367
S-R+:		= .359

Es zeigt sich, daß die Notendifferenz zwischen guten und schlechten Rechtschreibern mit sinkender Sozialschicht zunimmt. Sie beträgt für S+ = (-.093 - .585) = -.678, für S0 = (-.248 - .462) = -.710 und für S- = (-.367 - .359) = -.726. Mangelnde Rechtschreibleistung schlägt sich also umso stärker in der Deutschnote wieder, je geringer die Sozialschicht der Kinder ist.

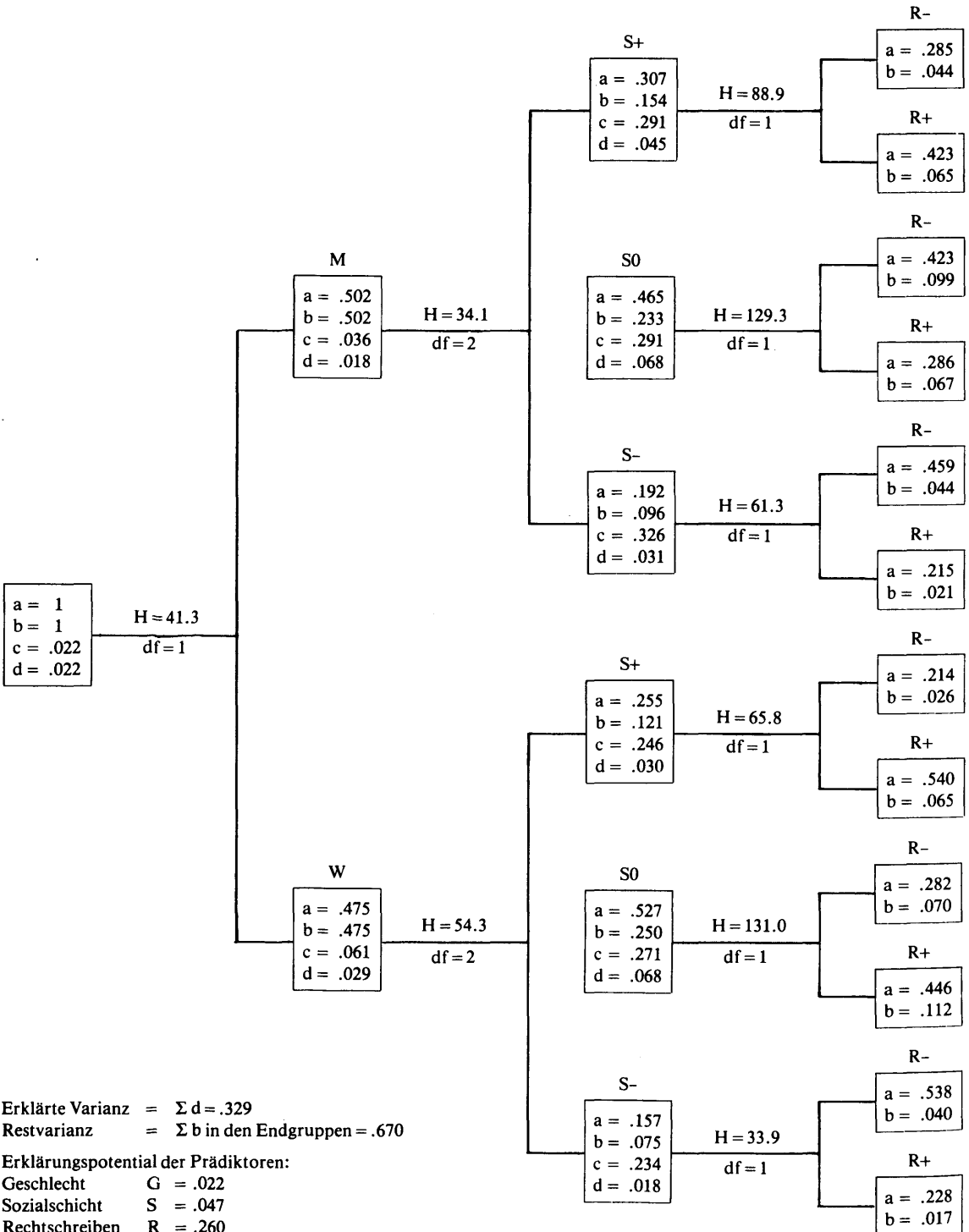
3. Diskussion

HRVA und PKFA berücksichtigen beide nicht die Interaktion zwischen den Prädiktoren. Damit sind die H- bzw. X^2 -Werte in diesen Analysen auch nicht additiv. Das bedeutet auch, daß die PKFA keine Abschätzung der durch die Hinzunahme eines neuen Prädiktors gewonnenen Erklärungspotentials erlaubt. Es ist nicht zu klären, inwieweit eine Erhöhung des Gesamt- X^2 bei der Hinzunahme eines neuen Prädiktors durch die Interaktion des neuen mit den bereits vorhandenen Prädiktoren (anstatt mit dem Prädikanden) bedingt ist. Das resultiert daraus, daß die Erwartungswerte aus den Randsummen der Kontingenztafel berechnet werden. Diese Randsummenverteilung spiegelt aber die Interaktion zwischen den Prädiktoren wieder. Zum ändern ist ebenfalls nicht zu klären, ob eine eventuelle Erhöhung des X^2 -Wertes darauf zurückzuführen ist, daß

- der neue Prädiktor *allein* die Prädiktionsgüte erhöht, oder daß
- die *Interaktion* des neuen Prädiktors mit den bereits vorhandenen Prädiktoren den eigentlichen Erklärungswert hat.

Folgerichtig verzichten LIENERT & KRAUTH

Tab. 5: Tafel der hierarchischen Rangvarianzanalyse – Prädiktorenfolge Geschlecht (G), Sozialschicht (S) und Rechtschreiben (R).



auch auf eine Prädiktorenhierarchie. Sieschlagen vor, innerhalb der Prädiktions-Kontingenztafel einzelne Prädiktionstypen sachlogisch zu Agglutinationstypen zusammenzufassen. Da sich beobachtete und erwartete Werte additiv verhalten, muß für diese Agglutinationstypen lediglich das X^2 neu berechnet werden.

Die gleichen Restriktionen gelten auch für die HRVA. Bei ihr kann zwar über den H-Wert bestimmt werden, ob die Aufteilung einer Gruppe nach einem neuen Prädiktor neue Erklärungsvarianz bringt. Insoweit scheint die HRVA gegenüber der PKFA im Vorteil zu sein (abgesehen von der schärferen Prüfung auf Ordinalbasis). Doch erweist sich die Algorithmik als gefährlich. Da bei der Aufteilung einer Gruppe in Untergruppen diese Untergruppen neue Rangwerte erhalten, werden mögliche Interaktionen *per definitionem gleich Null gesetzt, aber nicht eliminiert*. Das hat bei hoch interagierenden Prädiktorenssets gravierende Folgen.

In Tabelle 5 wurde der gleiche Datensatz im gleichen Verfahren der HRVA noch einmal durchgerechnet. Lediglich die Folge der Prädiktoren wurde vertauscht. Sie war in Tabelle 3 R-S-G und ist jetzt in Tabelle 5 G-S-R. In Tabelle 5 sind mit der neuen Prädiktorenfolge *alle H-Werte signifikant*. Erklärbar ist das daraus, daß z. B. Recht schreiben als erster Prädiktor verwendet aufgrund seiner Abhängigkeit zur Sozialschicht bereits Varianz abschöpft, die eigentlich der Sozialschicht zuzuordnen ist. Je nach Abfolge sind dann auch die Summen der erklärten Varianz wie das Erklärungspotential der einzelnen Prädiktoren unterschiedlich (vergleiche die Aufstellungen am Fuß der Tabellen 3 und 5). Das heißt aber, daß lediglich bei mehrdimensional orthogonalen Prädiktoren die Abfolge der Prädiktoren am Ergebnis der HRVA nichts ändert. Eine solche mehrdimensionale Orthogonalität liegt jedoch bei empirischen Daten in aller Regel nicht vor.

SCHULZE sieht dieses Problem wohl, wenn er zur Aufdeckung von Interaktionseffekten empfiehlt, zuerst diejenige Variable, «von der man vermutet, daß von ihrer Ausprägung die Wirksamkeit der anderen Variablen abhängt» (a. a. O., p. 134), in die Analyse einzuführen. Diese tentative Einführung von Prädiktoren ist aber bei inferentieller Prüfung nicht zulässig. Mit ihr werden implizit neue Hypothesen getestet. Damit wird das α -Risiko in undurchschaubarer Weise verändert. Wie bei der PKFA und der ASA muß aber

auch in der HRVA das α -Risiko bei inferentieller Prüfung auf die Zahl der durchgeführten k Tests zu $\alpha^\circ = \alpha/k$ adjustiert werden (siehe dazu KRAUTH & LIENERT, 1973b). Bei der Festlegung der Zahl durchzuführenden Einzeltests stehen nur zwei Wege offen:

- a) entweder wird *theoriegeleitet vor der Datenerhebung* festgelegt, welche Prädiktionskonfigurationen geprüft werden sollen (sachlogische ex ante - Definition);
- b) oder bei der Prüfung ex post wird k als die Zahl aller möglichen Tests definiert (hypothesenfreie ex post-Definition).

Es ist unmittelbar evident, daß die zweite Vorgehensweise zu extrem konservativen Entscheidungen führt, da k mit der Zahl der Prädiktoren exponentiell anwächst. Jedes andere Vorgehen (vor allem, wenn überhaupt nicht adjustiert wird), führt dazu, daß lediglich heuristische «Signifikanzen» zu gewinnen sind, die an einem *neuen* Datensatz inferentiell überprüft werden müssen.

SCHULZE schlägt zur Aufdeckung von Interaktionseffekten weiter vor, die Varianzanteile $V_{X(X)} = c$ jeweils auf der gleichen Hierarchiestufe zu vergleichen. Sind diese unterschiedlich, sei eine Interaktion wahrscheinlich (a. a. O., p. 137). Dieser Vergleich ist bei korrelierten Prädiktoren nicht möglich. Die Verletzung der Unabhängigkeit macht diese Werte inkompatibel. Die Technik der neuen Rangvergabe bei der Aufteilung in Untergruppen hat als weitere Konsequenz, daß die Produktbildung $V_{X(X)} \cdot g_{X(G)} = c \cdot b = d$ bei korrelierten Prädiktoren auf falschen Voraussetzungen beruht (Multiplikationssatz der Wahrscheinlichkeit). Damit sind aber die Erklärungsvarianzen wie deren Summe nicht zu interpretieren. Das erweist sich daran, daß in Tabelle 3 und 5

- a) sowohl die erklärte Varianz verschieden ist (in Tabelle 3 sind es 33,7%, in Tabelle 5 32,9%),
- b) wie auch die $g_{X(G)} = b$ der Endgruppen in beiden Tabellen unterschiedlich sind,

obwohl genau die gleichen Prädiktoren, lediglich in anderer Reihung verwendet wurden.

Gegenüber beiden Verfahren erlaubt die ASA eine Abschätzung der Interaktionen. Die Prädiktion aus Einzelprädiktoren (G, S und R) erbringt dieselben Werte wie die PKFA. Zusätzlich läßt sich aber auch das Erklärungspotential von Prädiktoreninteraktionen abschätzen (im Beispiel

SRN). Damit ist die ASA gegenüber den anderen Verfahren immer dann im Vorteil, wenn es um die optimale Auswahl der Prädiktoren geht.

Der Vorteil der erhöhten Effizienz der ordinalen Prüfung sensu SCHULZE wird eingeschränkt durch die methodisch nicht aufgefangene Prüfung von Interaktionen höherer Ordnung. Wie die PKFA ist die HRVA nur dort zu empfehlen, wo

- a) die Interaktionen zwischen den Prädiktoren gleich Null oder nahe Null sind, und
- b) es um die Globalbeurteilung eines Prädiktorensatzes geht. Dieser sollte theoretisch begründbar sein, da eine empirische Prüfung des Prädiktionsbeitrags im Höherdimensionalen nicht möglich ist.

Ist dagegen eine Interaktion zwischen den Prädiktoren anzunehmen, sollte die ASA eingesetzt werden. Der Verlust an Schärfe wegen der nur nominalen Voraussetzungen wird in aller Regel aufgefangen durch die klare Interpretationsmöglichkeit. Vor allem bei ungenügender theoretischer Grundlegung kann der optimale Satz von Prädiktoren empirisch bestimmt werden. Wird gleichzeitig das α -Risiko adjustiert, muß auch bei einer solchen Prüfung nicht auf inferenzstatistische Aussagen verzichtet werden.

Literatur

- KRAUTH, J. & LIENERT, G. A. 1973a. Die Konfigurationsfrequenzanalyse und ihre Anwendung in Psychologie und Medizin. Freiburg/München: Alber.
- KRAUTH, J. & LIENERT, G. A. 1973b. Nichtparametrischer Nachweis von Syndromen durch simultane Binomialtests. *Biometrische Zeitschrift* 15, 13–20.
- KRUSKAL, W. H. & WALLIS, W. A. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47, 583–621.
- LIENERT, G. A. & KRAUTH, J. 1974. Die Konfigurationsfrequenzanalyse als Prädiktionsmodell in der angewandten Psychologie. In: Eckensberger, L. H., Eckensberger, U. S., (Eds.): Bericht über den 28. Kongreß der DGfP. Band 2. Göttingen: Hogrefe, 219–229.
- SCHULZE, G. 1978. Ein Verfahren zur multivariaten Analyse der Bedingungen von Rangvariablen: Hierarchische Rangvarianzanalyse. *Zeitschrift für Sozialpsychologie* 9, 129–141.
- SUTCLIFFE, J. P. 1957. A general method of analysis of frequency data for multiple classification designs. *Psychological Bulletin* 54, 134–137.
- TENT, L., FINGERHUT, W. & LANGFELDT, H.-P. 1976. Quellen des Lehrerurteils. Weinheim: Beltz.
- WINER, B. J. 1971². *Statistical principles in experimental design*. New York: McGraw-Hill.

