

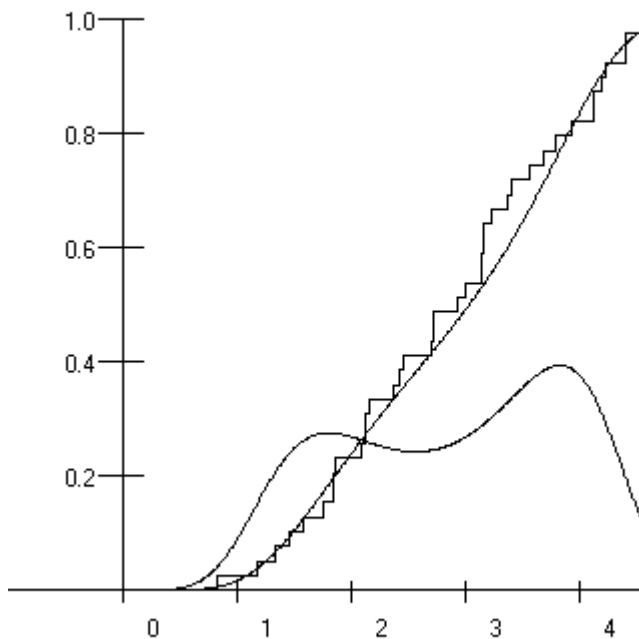
Statistik für Informationsmanager

Klausur vom 21. Juli 2003

Musterlösung

Für die Teilaufgaben hat es folgende Punktzahlen gegeben:

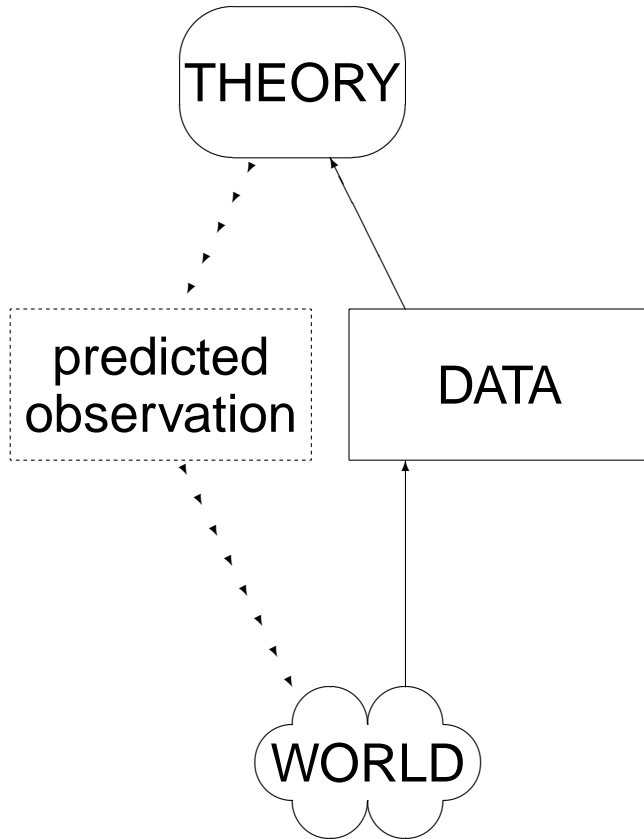
Aufgabe	Maximal erreichbar	Im Mittel wurden erreicht	Standardabweichung
1	$2*(1 + 2)$	3.74	1.98
2	$8*(3/8 + 3/8)$	4.70	0.98
3	$3*(1 + 1)$	5.33	0.87
4.1	$5*1$	2.98	1.17
4.2	$3*1 + 3$	3.13	1.71
4.3	$5*1$	3.16	1.68
5	$4*2 + 1$	4.99	3.00
6	$5*1 + 2*2$	7.97	1.49
Summe	52	36.01	8.73
Note	1.0	2.96	1.22



Die Verteilung der Noten (die hier verwendeten Noten sind nicht die endgültigen, sondern eine Lineartransformation der Gesamtpunktzahl!) ist deutlich zweigipflig: Ein Modus liegt bei etwa 1.7, der andere bei etwa 3.7. Dazwischen liegt ein Minimum bei etwa 2.7 — es spricht vieles dafür, dass diejenigen 16, die besser als 2.7 abgeschnitten haben, an den Übungen regelmäßig teilgenommen haben, während diejenigen 23, die darauf schon mal häufiger verzichtet haben, schlechter als 2.7 abgeschnitten haben: Die eine Gruppe hätte dann mit guten Noten bestanden, die andere hat das Nichtbestehen riskiert (und zu einem knappen Drittel denn auch realisiert).

1. Grundlagen

Erinnern Sie sich an die mehrfach gezeigte Abbildung;



Zeichnen Sie ein, wohin

- a. Amtliche Statistik (1P) zwischen Welt und Daten
 - b. Zeitreihenanalyse (1P) an allen anderen drei Stellen
- hingehören! Geben Sie jeweils eine kurze Begründung!

a. Amtliche Statistik gehört dorthin, weil ... (2P) ihre wesentliche Aufgaben darin besteht, Daten zu erheben und der Forschung (aber auch der Politik, der Verwaltung und der Wirtschaft) in Form von Statistischen Berichten, Statistischen Jahrbüchern oder auf Datenträgern zugänglich zu machen.

b. Zeitreihenanalyse gehört dorthin, weil ... (2P) sie b1. Daten modelliert (Daten → Theorie), b2. Vorhersagen erlaubt (Theorie → Vorhersagen) und b3. Vorhersagen mit neuen Beobachtungen vergleicht (Vorhersagen → Welt).

2. Skalenniveaus

Geben Sie jeweils für die folgenden Variablen an, ob sie nominal, ordinal, intervall- oder ratioskaliert sind (jeweils 3/8 P) und geben Sie jeweils eine kurze Begründung (jeweils weitere 3/8 P). (Eine gute Begründung für eine unserer Ansicht nach falsche Einordnung gibt eventuell ebenfalls Punkte, keine Begründung für eine unserer Ansicht nach richtige Einordnung gibt einen Punktabzug!)

	Nominal	Ordinal	Intervall	Ratio	Begründung
ECTS-Punkte für ein Modul eines Studienganges				X	Misst den Arbeitsaufwand, der zur Erbringung der Leistung erforderlich ist: doppelt so viel Aufwand, doppelt so viel Punkte
Note in einem Leistungsnachweis		X	(X)		Nur fiktiv ist der Unterschied zwischen 1.0 und 1.3 dem zwischen 3.7 und 4.0 gleich — und der Unterschied zwischen 1.0 und 2.0 ist bekanntlich viel kleiner als der zwischen 4.0 und 5.0! Trotzdem berechnet man häufig Mittelwerte.
EAN (Europäische Artikelnummer)	X				Nur ein Label, Name, Etikett ...
Wechselkurs zwischen zwei Währungen				X	Die Null kommt zwar nicht vor, aber der Dollar kann im Vergleich zum Euro um p% fallen oder steigen → Ratioskala!
Längengrad oder Breitengrad (gilt für beide dasselbe??)			L:X	B:X	Auf dem gleichen Breitenkreis ist die Entfernung zwischen 10°O und 20°O genau doppelt so groß wie die zwischen 6°O und 11°O; bei den Breitenkreisen ist der Äquator eine natürliche Null; aber Vorsicht, man muss die Längengrade auf [-180, +180] und die Breitengrade auf [-90, +90] beschränken!
Listenplatz auf einer Kandidatenliste		X			Die Plätze sind aufsteigend nummeriert, aber nicht linear zu den Stimmzahlen (oder der Qualität der Kandidaten ...).
Marktanteil eines Produkts				X	... ein Anteil, also eine Ratio ...
Nummer eines Fachbereichs	X				Ein Name in Gestalt einer Nummer!

3. Univariate Statistik

Berechnen Sie (jeweils falls sinnvoll!) den Modus, den Median und den Mittelwert für folgende Variablen (richtige Berechnung[en] jeweils 1P, richtige Begründung jeweils noch 1P, wenn bei einer Teilaufgabe richtigerweise alle drei Werte berechnet worden sind, bedarf es keiner Begründung, und es gibt trotzdem 2P):

1. Gesamtpunktzahlen von 20 Studierenden bei einer Klausur:

118.2	80.4	133.8	92.8	102.0	90.8	68.4	58.4	55.1	81.9
120.6	41.2	118.6	83.8	105.2	131.5	99.3	83.3	139.0	145.2
Mittelwert:			97.475						
Median			Alles zwischen 92.8 und 99.3, vor allem 96.05						
Modus			Der Modus dürfte genau genommen dem Mittelwert gleich sein, wenn die Werte normalverteilt sind (was sie zu sein scheinen!), sonst: irgendwo in der dichtesten Gruppe ...						
gegebenenfalls Begründung für nicht berechnete Parameter:			Beim Modus: nicht trivial aus den Daten zu ermitteln!						

2. Gesamtnoten der gleichen Studierenden in der gleichen Klausur:

3.7	5.0	3.0	5.0	4.0	5.0	5.0	5.0	5.0	5.0
3.3	5.0	3.7	5.0	4.0	3.0	5.0	5.0	2.7	2.7
Mittelwert:			4.255 dürfte man ausrechnen, wenn man oben Noten für intervallskaliert gehalten hat, sonst (eigentlich) nicht!						
Median			5.0						
Modus			5.0						
gegebenenfalls Begründung für nicht berechnete Parameter:			Siehe Mittelwert						

3. Geschlecht von zwanzig Studierenden:

W	M	M	M	W	W	M	M	W	W
M	M	W	W	M	M	M	W	M	W
Mittelwert:			Gibt es nicht.						
Median			Gibt es auch nicht.						
Modus			M						
gegebenenfalls Begründung für nicht berechnete Parameter:			Dies ist sicher eine kategoriale Variable, bei der nur der Modus berechnet werden kann.						

4. Bivariate Statistik

4.1 Nominale Variable

Bei der Bundestagswahl 2002 war in 125 Wahlkreisen die CDU oder die CSU nach Erststimmen stärker als die SPD, in den übrigen 174 Wahlkreisen war die SPD stärker als die CDU bzw. als die CSU. 1998 hatte die SPD nach diesem Kriterium noch besser abgeschnitten. Untersucht man, welche Partei vor der anderen lag, getrennt nach größeren Regionen in Deutschland, so ergeben sich die nachstehenden beiden Kreuztabellen:

		NSO Nord/Süd/Ost			Gesamt
		1.00 Nord	2.00 Süd (BY, BW)	3.00 Ost	
CDU/CSU > SPD	Anzahl	36	73	16	125
	% von NSO Nord/Süd/Ost	23.7 %	90.1 %	24.2 %	41.8 %
SPD > CDU/CSU	Anzahl	116	8	50	174
	% von NSO Nord/Süd/Ost	76.3 %	9.9 %	75.8 %	58.2 %
Gesamt	Anzahl	152	81	66	299
	% von NSO Nord/Süd/Ost	100.0%	100.0%	100.0%	100.0%

		NSO Nord/Süd/Ost			Gesamt
		1.00 Nord	2.00 Süd (BY, BW)	3.00 Ost	
CDU/CSU > SPD	Anzahl	31	64	14	109
	% von NSO Nord/Süd/Ost	20.4 %	79.0 %	21.2 %	36.5 %
SPD > CDU/CSU	Anzahl	121	17	52	190
	% von NSO Nord/Süd/Ost	79.6 %	21.0 %	78.8 %	63.5 %
Gesamt	Anzahl	152	81	66	299
	% von NSO Nord/Süd/Ost	100.0%	100.0%	100.0%	100.0%

Alle Zusammenhangsmaße sind signifikant von Null verschieden. Sie lauten im einzelnen:

		Wert 2002	Wert 1998
Lambda	Symmetrisch	0.375	0.313
	Siegerpartei abhängig	0.520	0.431
	Nord/Süd/Ost abhängig	0.252	0.224
Goodman-und-Kruskal-Tau	Siegerpartei abhängig	0.357	0.291
	Nord/Süd/Ost abhängig	0.180	0.146
Unsicherheitskoeffizient	Symmetrisch	0.224	0.173
	Siegerpartei abhängig	0.282	0.222
	Nord/Süd/Ost abhängig	0.186	0.141

Beurteilen Sie an Hand der Kreuztabelle die Verteilung der Parteichancen in den verschiedenen Teilen Deutschlands (1P). Hier genügte es herauszufinden, dass im Süden die CDU/CSU in den meisten Wahlkreise stärker war als die SPD und dass es im Norden und Osten umgekehrt war, dies bei beiden Wahlen, nur dass die SPD 2002 etwas schlechter abgeschnitten hat, und zwar in allen drei Teilregionen.

Welche Angaben aus der Kreuztabelle — oder welche Koeffizienten — haben Sie dazu am besten benutzen können? (1P). Die Koeffizienten nützen da nicht besonders viel, die abgedruckten Prozentzahlen, vor allem die Spaltenmaxima waren hilfreicher.

Bitte bilden Sie einen Satz, mit dem Sie die Bedeutung der Zahl 0.520 (Lambda mit Siegerpartei 2002 als abhängiger Variable) erläutern (1P). Bei Kenntnis der Region lässt sich der Name der stärkeren Partei um 52 % leichter raten als ohne Kenntnis der Region; die erwartete Fehlerhäufigkeit beim Raten der stärkeren Partei reduziert sich bei Kenntnis der Region des zu ratenden Wahlkreises um 0.520 ...

Wie beurteilen Sie, dass die Koeffizienten für 2002 durchgängig größer sind als die für 1998? (1P) Die regionalen Unterschiede haben sich verstärkt, was daran liegen kann, dass 2002 die beiden Kanzlerkandidaten aus besonders weit auseinander liegenden Regionen (Hannover / Wolftratshausen) stammten.

Was können Sie daraus schließen, dass die Koeffizienten für "Siegerpartei abhängig" durchgängig größer sind als die für "Nord/Süd/Ost abhängig"? (1P) ..., dass die Parteichancen von der Region abhängen und nicht umgekehrt.

4.2 Ordinale Variable

Aus der Konsumentenbefragung 1995 kann man eine Antwort auf die Frage ableiten, ob Haushalte um so preisbewusster sind, je mehr Kinder sie haben. Aus der Kreuztabelle der Variablen Preisbewusstsein (mit den Ausprägungen 1=nicht preisbewusst, 2=wenig preisbewusst, 3=überwiegend preisbewusst und 4=sehr preisbewusst) und der Variablen Anzahl der Kinder (0 bis 9) haben sich die folgenden Zusammenhangsmaße ergeben:

		Alte Länder	Neue Länder
Somers-d	Symmetrisch	.143	.126
	V19 Anzahl der Kinder unter 18 Jahren abhängig	.120	.106
	V20 Preisbewusstsein abhängig	.178	.154
Kendall-Tau-b		.146	.128
Kendall-Tau-c		.109	.103
Gamma		.268	.214
Anzahl der gültigen Fälle		3711	927

Fragen:

Warum ist Gamma größer als die anderen Koeffizienten? (1P) Alle Koeffizienten sind Brüche, ihre Zähler sind gleich (C-D), aber Gamma hat den kleinsten Nenner (C+D), bei den anderen Koeffizienten stehen im Nenner noch weitere Summanden (die Anzahlen der verbundenen Paare in verschiedenen Kombinationen).

Lässt sich aus dem Unterschied zwischen den beiden "asymmetrischen" Somers' D eine Schlussfolgerung ziehen? (0.5P) Wenn Ja, welche, wenn nein, warum nicht? (gute Begründungen für Antworten, die wir für falsch halten, geben weitere 0.5 Punkte, fehlende Begründungen auch für richtige Antworten geben zwar keinen Punktabzug, aber auch keinen Zuschlag!) Ja, im Zweifel ist eher die Kinderzahl die Ursache und das Preisbewusstsein die Wirkung, und nicht umgekehrt.

Wäre die Verwendung eines nominalen Maßes (Phi, Lambda, ...) hier auch angemessen? (0.5P) Wenn Ja, warum, wenn nein, warum nicht? (wie oben) Die nominalen Maße würden zwar auch etwas über die Stärke des Zusammenhangs, aber nichts über die Richtung des Zusammenhangs sagen, und das letztere interessiert bei ordinalen Variablen.

Beurteilen Sie die Unterschiede in den Koeffizienten zwischen den neuen und den alten Bundesländern (3 P). In den alten Bundesländern sind alle Koeffizienten höher als ihre Entsprechungen aus den neuen Bundesländern. Daraus folgt nicht, dass das Preisbewusstsein oder die Kinderzahl in den neuen Bundesländern niedriger wäre als in den alten, nur die Abhängigkeit ist größer.

4.3 Metrische Variable

In den letzten beiden Klausuren im Vordiplom Wirtschaftsinformatik, die aus je vier Teilklausuren bestanden haben, habe ich folgende Korrelationen (Produkt-Moment-Korrelationskoeffizienten) zwischen den Punktzahlen für die Teilklausuren, der Gesamtpunktzahl und der Note gefunden:

.	Emp. Grdl.	O+M	BWL	KLR	Punkte ges.	Note
Punktzahl in der Teilklausur Empirische Grundlagen	1.000					
Punktzahl in der Teilklausur Organisation und Management	0.318	1.000				
Punktzahl in der Teilklausur Einführung in die Betriebswirtschaftslehre	0.304	0.438	1.000			
Punktzahl in der Teilklausur Kosten- und Leistungsrechnung	0.437	0.315	0.588	1.000		
Punktzahl insgesamt	0.741	0.652	0.768	0.796	1.000	
Note für die gesamte Klausur	-0.735	-0.574	-0.706	-0.677	-0.915	1.000

Fragen:

Ist die Verwendung des Produkt-Moment-Korrelationskoeffizienten hier überhaupt gerechtfertigt? (0.5P) Wenn Ja, warum, wenn nein, warum nicht? (gute Begründungen für Antworten, die wir für falsch halten, geben weitere 0.5 Punkte, fehlende Begründungen auch für richtige Antworten geben zwar keinen Punktabzug, aber auch keinen Zuschlag!) Ja, alle tun es, und bei den Punktzahlen kann man wohl von Intervallskaliertheit (wenn nicht gar Ratioskaliertheit!) ausgehen, jedenfalls gibt man sich im allgemeinen alle Mühe, die Punktzahlen nach der Schwierigkeit der Aufgaben festzusetzen. Aber selbst wenn man die Punktzahlen und die Noten eigentlich für ordinal hält, liegt man mit dem PMK als Maßzahl nicht gar so falsch.

Beurteilen Sie an Hand der Korrelationskoeffizienten für die vier Teilklausuren die "Verwandtschaftsverhältnisse" zwischen den vier Teilgebieten und prüfen Sie, ob diese Beurteilung mit Ihrem Vorverständnis übereinstimmt! (1P) Die engste "Verwandtschaft" haben wir zwischen KLR und BWL (beides sehr BWL-spezifisch); für das Paar O+M / BWL gilt Ähnliches, EmpGrdl und KLR scheinen über die Notwendigkeit, zu rechnen, verwandt zu sein.

Wie erklären Sie sich, dass die Korrelationen der Gesamtpunktzahl mit den Einzelpunktzahl am höchsten sind? (1P) Schließlich ist die Gesamtpunktzahl einfach die Summe der vier Teilpunktzahlen.

Wie erklären Sie sich, dass die Korrelationen der Note mit allen anderen Variablen negativ sind? (1P) Je mehr Punkte, desto besser (und in Deutschland: kleiner) die Note!

Warum konnte ich die Zahlen in der oberen Dreiecksmatrix der Tabelle oben einfach weglassen? (1P) Weil die Matrix symmetrisch ist (und auch der PMK!).

5. Multivariate Statistik / Multiple Regression

Bei der Vordiplomsklausur Wirtschaftsinformatik vor einigen Wochen wurden die vier Teilklausuren getrennt korrigiert; für jede Teilklausur gab es maximal 50 Punkte — das ergab die vier in die nachstehende Tabelle als unabhängig aufgenommenen Variablen „Organisation und Management“, „Einführung in die BWL“, „Empirische Grundlagen“ und „Kosten- und Leistungsrechnung“. Abhängige Variable ist die Gesamtnote.

Koeffizienten					
	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
	b	Standardfehler	β		
(Konstante)	6.839	.232		29.460	.000
Punktzahl in der Teilklausur Empirische Grundlagen (x_1)	-.037	.005	-.477	-7.290	.000
Punktzahl in der Teilklausur Organisation und Management (x_2)	-.024	.008	-.205	-3.122	.003
Punktzahl in der Teilklausur Einführung in die Betriebswirtschaftslehre (x_3)	-.035	.007	-.358	-4.751	.000
Punktzahl in der Teilklausur Kosten- und Leistungsrechnung (x_4)	-.018	.007	-.193	-2.560	.014
Abhängige Variable: NOTE (y)					

Wie lässt sich nach dem Ergebnis der Regression die Gesamtnote aus den einzelnen Punktzahlen der Teilklausuren berechnen? Schreiben Sie dazu — unter Verwendung der Symbole y , x_1 , x_2 , x_3 und x_4 sowie der Zahlenwerte von b_0 , b_1 , b_2 , b_3 und b_4 — die Regressionsgleichung auf (2P). Wer den Zirkumflex über dem y weggelassen hat, hat einen halben Punkt eingebüßt, wer die Zahlenwerte nicht hingeschrieben hat, hat einen Punkt abgezogen bekommen (es sei denn, die nächste Frage war richtig beantwortet).

$$\hat{y}_i = 6.839 + (-0.037) x_{i1} + (-0.024) x_{i2} + (-0.035) x_{i3} + (-0.018) x_{i4}$$

Berechnen Sie mit Hilfe der soeben aufgeschriebenen Regressionsgleichung die geschätzte Gesamtnote von jemand, der in allen vier Teilklausuren 25 Punkte erreicht hat (2P). 3.989 war die Antwort — es musste schon irgendetwas in der Nähe von 4 herausgekommen!

Welche Teilklausur(en) hat/haben sich am stärksten auf die Gesamtnote ausgewirkt? (2P) Die mit den höchsten Regressionskoeffizienten (genau genommen kam es nur auf die standardisierten an, aber hier machte das keinen Unterschied).

Kann man bei der kleinen Stichprobe (bei lediglich 50 Teilnehmern) davon ausgehen, dass die Ergebnisse über den Zusammenhang zwischen den Punktzahlen der Teilklausuren und der Gesamtnote auf die Grundgesamtheit übertragbar ist? (2P) Hier kann man einerseits sagen: Jeder Jahrgang ist wieder anders, hat bei anderen Lehrenden Vorlesungen gehört — dann musste man sagen: Dies ist gar keine Stichprobe, sondern eine Aussage über die Grundgesamtheit von 50 Studierenden des Jahrgangs 2001, und dann gibt es keine Signifikanzüberlegungen. Oder man schlug sich auf die Seite derer, die annehmen, dass jede Klausur und ihre Ergebnisse für alle Studierenden der Wirtschaftsinformatik am Anfang des 21. Jahrhunderts sind — und dann musste man feststellen, dass alle Koeffizienten sich signifikant von Null unterscheiden (und dass selbst bei der KLR-Teilklausur die Wahrscheinlichkeit einen größeren Koeffizienten in einer gleichgroßen Stichprobe vorzufinden, obwohl in der Grundgesamtheit keine Korrelation herrscht, bloß 1.4 % ist (0.014 entspricht 1.4 %, nicht 14 %!)).

Beschreiben Sie die Grundgesamtheit, aus der diese Stichprobe entnommen ist (1P). Siehe oben!

6. Multivariate Statistik / Faktorenanalyse

Ein Teil der Fragen aus dem Verbraucherpanel bezog dich auf Einkaufs- und Ernährungsgewohnheiten. Aus diesen 41 Fragen ließ sich die Faktorladungsmatrix auf der nächsten Seite gewinnen. Koeffizienten kleiner als 0.2 sind zur Erleichterung des Auffindens einer Einfachstruktur weggelassen. Interpretieren Sie die vier Faktoren, indem Sie ihnen möglichst prägnante Namen geben. Begründen Sie Ihre Namensgebung.

	Faktoren				
	1	2	3	4	5
V78 Achten auf fettarme Ernährung	0.717				
V79 Achten auf schonende reizarme Kost	0.713				
V80 schlankheitsbewusste Ernährung	0.659				
V92 Rücksichtnahme auf Gesundheit bei Ernährung	0.658				
V66 Achten auf Figur	0.655				
V96 Achten auf Kalorien	0.642				
V102 Vermeidung von gesundheitsschädlicher Ernährung	0.611		0.249		
V81 Kauf von zusatzstofffreien Lebensmitteln	0.557		0.382		-0.215
V72 vollwertige Ernährung	0.541				
V93 Information über umweltbelastete Lebensmittel- kein Kauf mehr	0.495		0.312		
V99 häufige Verwendung von Getreidekörnern	0.419			-0.373	
V65 Ablehnung von Konservierungsstoffen in Produkten	0.410		0.374		
V90 regelmäßige Verwendung von Vitamin- und Mineralstoffpräparaten	0.359	0.249			0.333
V70 Vorzug von edlen Speisen und Getränken		0.703			
V73 Delikatessen		0.687			
V98 sehr anspruchsvoll beim Essen und Trinken		0.663	0.215		
V86 Verwöhnen mit gutem Essen		0.651			
V82 Vorliebe für ausgefallene Speisen und Gerichte		0.646	-0.201		-0.263
V94 Spaß am Ausprobieren fremdländischer Spezialitäten		0.581	-0.414	-0.252	
V67 Begeisterung für ausländische Spezialitäten		0.511	-0.378	-0.234	
V103 bei Wahl Kauf von deutschen Lebensmitteln	0.273		0.642		
V83 beste Qualität der deutschen Nahrungsmittel	0.235		0.631		
V100 Lebensmittel bekannter Marken besser als die unbekannter Marken		0.253	0.615		
V68 Gleichgültigkeit gegenüber Herkunftsland der Lebensmittel		0.228	-0.596		
V89 kein Vertrauen zu Nahrungsmitteln ohne Markenbezeichnung			0.596		
V75 Preis, nicht Marke, ist für Kauf entscheidend		-0.215	-0.426	0.214	
V71 Vorzug von Hausmannskost		-0.239	0.208	0.619	
V91 zuviel Wirbel um Ernährung				0.589	
V87 Vorliebe für herzhaftere, deftige Mahlzeiten				0.549	
V77 Überschätzung des Einflusses der Ernährung auf die Gesundheit				0.536	
V105 Kein Essen und Trinken mehr bei Annahme aller Ernährungstips				0.531	
V97 Kochen von altbewährten Gerichten		-0.264	0.351	0.506	
V85 vegetarische Ernährung	0.327			-0.407	
V76 Multivitaminsäfte als wichtige Ergänzung	0.270			0.280	0.259
V95 Vorzug von Kochen einfacher Gerichte		-0.231			0.706
V84 Vorliebe für schnell kochbare Gerichte		-0.222			0.701
V74 kein Kochen ohne Fertigprodukte					0.593
V69 viel Zeit für Kochen	0.217	0.336		0.353	-0.553
V101 Konserven genauso gut wie Frisches				0.349	0.497
V88 Kauf von ausschließlich frischen Lebensmittel	0.336		0.355		-0.401
V104 bei normaler Kost keine Ergänzungspräparate notwendig				0.231	-0.293

Faktor 1 beschreibt die latente Eigenschaft ... (1 P), sich gesundheitsbewusst ernähren zu wollen.

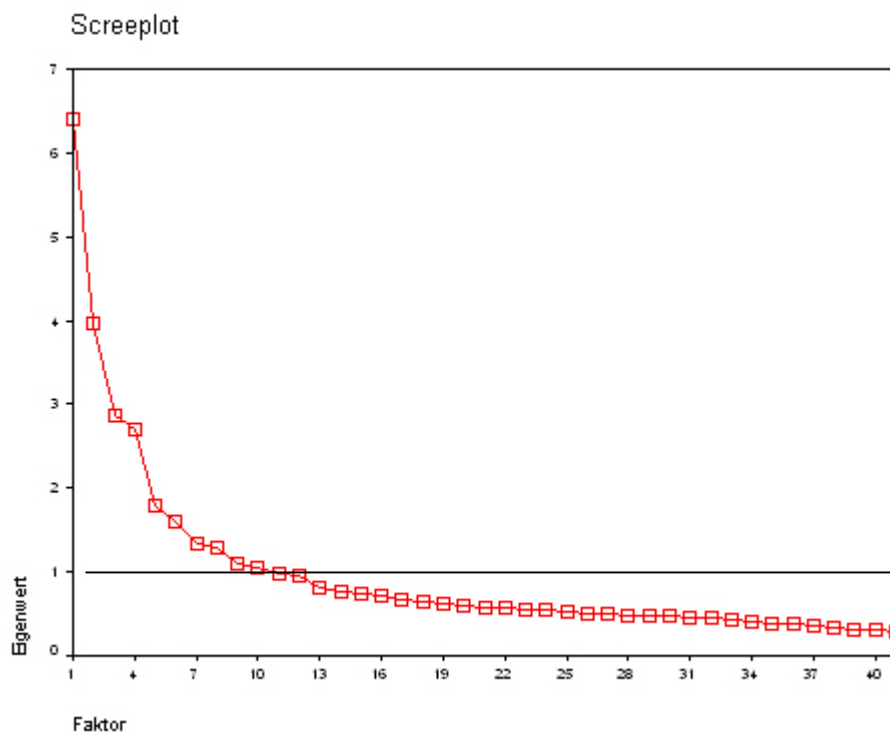
Faktor 2 ... (1 P), sich beim Essen Luxus gönnen zu wollen.

Faktor 3 ... (1 P), äußere Merkmale wie Marke oder Herkunftsland zur Entscheidungsgrundlage zu machen.

Faktor 4 ... (1 P), ohne Rücksicht auf Mode und Gesundheit, beim Essen zu bevorzugen, was man immer schon gegessen hat.

Faktor 5 ... (1 P), ohne Rücksicht auf Geschmack, Mode und Gesundheit zu essen, was gerade zur Hand ist und ohne Aufwand zubereitet werden kann.

Hier finden Sie den Scree-Plot zu der vorstehenden Auswertung:



Halten Sie danach die Extraktion von fünf Faktoren noch für gerechtfertigt? (Ja / Nein / Kommt darauf an)
Wenn ja: warum? (2 P) Sowohl der Screeplot als auch die gelungene Einfachstruktur der Ladungsmatrix im Zusammenspiel mit den mühelos interpretierbaren Faktoren legen nahe, dass fünf Faktoren eine gute Wahl war.

Wenn nein: warum nicht? Wie viele hätte man statt dessen extrahieren sollen? (2 P) Man kann aber im Screeplot auch den großen Knick nach Faktor drei oder nach Faktor 8 sehen oder nach dem Kaiser-Kriterium 11 Faktoren extrahieren.

(Antworten auf beide Unterfragen werden gewertet! Wenn Sie dezidiert der Auffassung sind, dass nicht drei, sondern zwei oder vier oder fünf oder sechs oder sieben oder acht ... Faktoren hätten extrahiert werden sollen, müssen Sie auf „wenn nein: warum nicht?“ besonders ausführlich antworten; ebenso müssen Sie, wenn Sie dezidiert der Auffassung sind, dass drei Faktoren die richtige Wahl waren, auf „wenn ja, warum?“ besonders ausführlich antworten.)