

Empirische Forschungsmethoden

Klausur vom 3. August 2006

Name:

Matrikelnummer:

Erlaubte Hilfsmittel: Außer Schreibwerkzeug diesmal auch Taschenrechner (wird aber nicht nötig sein); Antworten bitte im Anschluss an die Fragen, Rückseiten können für Antworten und Entwürfe genutzt werden; Entwürfe bitte anschließend durchstreichen!

Diese Klausur soll — nicht — als Freiversuch gewertet werden.

.....
Unterschrift

Die Klausur zur Veranstaltung “Empirische Forschungsmethoden” ist korrigiert. Entsprechend den Ankündigungen ist die Bewertung so vorgenommen worden, dass sich die Gesamtnote aus den beiden besten abgegebenen Übungsblättern und der Klausur errechnet. In den beiden besten Übungsblättern konnten bis zu 40 Punkte erreicht werden, in der Klausur bis zu 80. Die Gesamtpunktzahl errechnete sich aus der Summe der Punkte der beiden besten Übungsblätter und der halbierten Summe der in der Klausur erreichten Punkte. Die schlechteste Punktzahl, die erreicht wurde, war die 41, die beste war 71.75. Ab 68.75 Punkten gab es eine 1.7, 66.5 bis 68.5 Punkte ergaben eine 2.0, 60 bis 64 Punkte ergaben eine 2.3, 55.5 bis 59.5 Punkte ergaben eine 2.7, 52 bis 54.5 Punkte ergaben eine 3.0, 47 bis 51 Punkte ergaben eine 3.3, und schließlich gab es für weniger als 42 Punkte eine 4.0.

Aufgabe	Maximal erreichbar		Im Mittel wurden erreicht	
1.1	1	12	0.93	7.49
1.2	2		1.11	
1.3	2		1.21	
1.4	4		1.82	
1.5	1		0.50	
1.6	2		1.39	
2	7	7	5.25	5.25
3.1	5	14	3.38	9.16
3.2	5		3.86	
3.3	4		1.92	
4.1	8	16	6.34	10.09
4.2	4		2.11	
4.3	4		1.63	
5.1	5	19	1.37	7.32
5.2	4		2.26	
5.3	4		1.58	
5.4	6		2.10	
6.1	4	12	2.70	7.84
6.2	2		1.17	
6.3	4		2.66	
6.4	2		1.32	
Summe	80	80		47.14

1. Grundlagen

1.1 Fußballfreunde wissen: Heimspiele werden häufiger gewonnen als Auswärtsspiele. Wie können Sie diese Hypothese überprüfen?

Das scheint denn doch eine schwierige Aufgabe gewesen zu sein. Kaum jemand hat herausgefunden, dass man nicht sagen kann, ob ein Spiel ein Heimspiel ist oder ein Auswärtsspiel. Manche haben Spieltage als Fälle genommen, andere Mannschaften, andere Spiele; vielfach wurde bei 1.2 bis 1.4 von Spielen, bei 1.5 und 1.6 aber von Mannschaften ausgegangen oder umgekehrt (in solchen Fällen gab es pauschal sechs Punkte ohne weitere Differenzierung, das erschien gerechtfertigt, weil in diesen Fällen ansonsten manches Richtige geschrieben wurde.

Alle Daten findet man übrigens unter <http://www.bundesliga.de/de/statistik/vergleich/index.php>

1.2 Welches sind die „Fälle“ („Objekte“), für die Sie (eine oder mehrere?) Variablen („Attribute“) erheben müssen?

Am zweckmäßigsten sind die Fälle die einzelnen Spiele, dann braucht man zunächst nur eine einzige Variable (außer der Bezeichnung des Spiels, z.B. „Bayern München gegen München 1860 am 24.04.2004“ — das übrigens 1:0 ausgegangen ist und als Heimsieg gilt — nur den Spielausgang — „Heimsieg“, „unentschieden“ oder „Auswärtssieg“ oder stattdessen die Tordifferenz zugunsten des Gastgebers, positive Zahlen sind dann Heimsiege, negative Auswärtssiege und die 0 steht für unentschieden; bei dieser Variante kann man auch noch ausrechnen, um wie viel günstiger die Tordifferenz für den Gastgeber ist). Natürlich kann man auch folgende Variablen wählen: Datum, Heimmannschaft, Gastmannschaft, Tore des Gastgebers, Tore des Gastes (woraus man dann die brauchbaren Variablen ableiten kann).

1.3 Welches Skalenniveau und welche Wertebereiche haben Ihre Variablen?

Die Anzahlen der Tore sind natürlich ratioskaliert, die Tordifferenz auch. Die oben definierte Variable „Spielausgang“ ist ordinal (Heimsieg > unentschieden > Auswärtssieg oder umgekehrt) oder dichotom (Heimsieg oder nicht).

1.4 Füllen Sie in dem folgenden Datenblatt die Überschrift und die Vorspalte und einige Zellen, die letzteren natürlich fiktiv, aus:

	Spiel	Tordifferenz	Entfernung (km)	Variable D	...
Fall 1	Bayern München gegen München 1860 am 24.04.2004	+1	0		
Fall 2	Bayern München gegen München 1860 am 10.09.2002	+2	0		
Fall 3	Stuttgart Kickers gegen FC St. Pauli am 09.02.2001	-2	653		
Fall 4					
...					

1.5 Welches ist die Grundgesamtheit, wie sähe eine Stichprobe aus?

Eigentlich ist die Grundgesamtheit die Menge aller Fußballspiele, die jemals gespielt wurden; falls man das beschränken möchte, kann man auch alle jemals ausgetragenen Bundesligaspiele nehmen. Eine Stichprobe bräuchte man im zweiten Fall eigentlich nur für Aufgabe 1.6, da für die Ausgangsfrage keine weiteren Variablen erforderlich sind, aber wenn man alle weltweiten Spiele erfassen will oder wenn man für alle Bundesliga-Paarungen die Entfernung (oder sonstige Merkmale) erheben will, dann muss man wohl eine Stichprobe ziehen, z.B. alle Spiele der nächsten Spielzeit oder jedes dritte Spiel der nächsten drei Jahre oder eine zufällige Auswahl von 3 % aller dokumentierten Spiele einer bestimmten Liga.

1.6 Zur Erklärung des Heimvorteils gibt es eine Reihe von Hypothesen. Eine lautet: Die Mannschaften, die auswärts spielen müssen, sind von der langen Anreise angestrengt oder übermüdet und spielen deshalb weniger konzentriert. Welche weitere Variable könnten Sie für die in Aufgabe 1.2 beschriebenen Fälle erheben, um die Hypothese über den Zusammenhang zwischen Heimvorteil und Anreise zu überprüfen?

Mindestens braucht man die Entfernung in Kilometern oder die tatsächlich benötigte Fahrzeit und das verwendete Verkehrsmittel.

Alternative Hypothesen: Fahrtzeit der Fans, Eigenschaften des Spielfelds etc. waren hier nicht zu untersuchen. Nur alle Spiele einer Mannschaft zu nehmen, wie verschiedentlich vorgeschlagen, wäre eine wenig repräsentative Stichprobe.

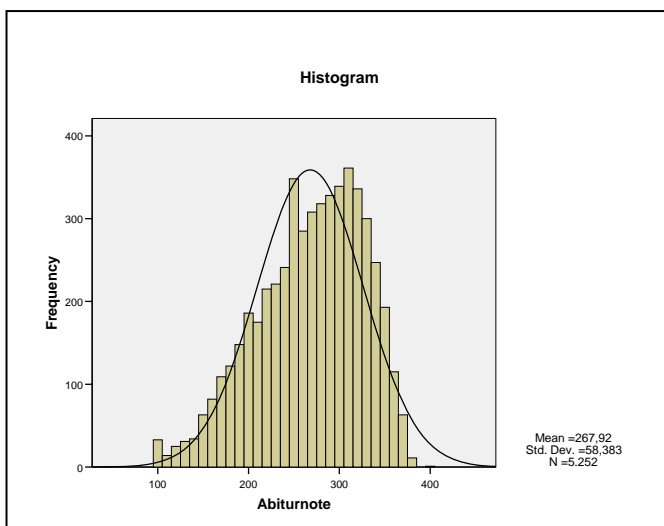
2. Skalenniveaus

In den Aufgaben dieser Klausur werden verschiedene Variable benutzt. Bitte geben Sie hier schon zusammenfassend an, welches Skalenniveau nach Ihrer Einschätzung die einzelnen Variablen haben und begründen Sie Ihre Einschätzung. Sie können auch einer Variablen verschiedene Skalenniveaus zuweisen, müssen dann aber auch das begründen.

	Dichotom	Nominal	Ordinal	Intervall	Ratio	Begründung
Ausgang eines Fußballspiels	X		X		X	Je nachdem, ob das in Tordifferenzen (ratio) oder nur kategorial definiert wird, im letzteren Fall entweder dichotom (Heimsieg oder nicht) oder ordinal (Heimsieg, unentschieden, Auswärtssieg)
Studiengang		X				klarer Fall
Geschlecht	X					nominal wurde auch als richtig gewertet; falls man als drittes Geschlecht (wie geschehen) die Transsexuellen einführt, wird die Variable allerdings ordinal: weiblich > transsexuell > männlich!
Abiturnote			X	X		Ratio ist problematisch: dann wäre der Unterschied zwischen einer 1 und 2 der gleiche wie der zwischen einer 2 und einer 4!
Fragen zum Einwanderungsproblem			X			Hier und bei den beiden nächsten Variablen musste man sich die späteren Aufgaben anschauen, dann war die Antwort eigentlich klar!
Die Faktoren aus Aufgabe 5				X		Ratio ist hier auch problematisch: Ist jemand mit dem Faktorwert +1.6 doppelt so optimistisch, wie jemand mit -0.8 pessimistisch ist?
Die Einschätzungen der Auswirkungen von Technologien aus Aufgabe 6			X			siehe oben

3. Univariate Statistik und Mittelwertvergleiche

Die Abiturnoten der Studierenden am Campus Landau (Sommersemester 2006) folgen der im Histogramm beschriebenen Verteilung. Parameter der Verteilung finden Sie unten ebenfalls.

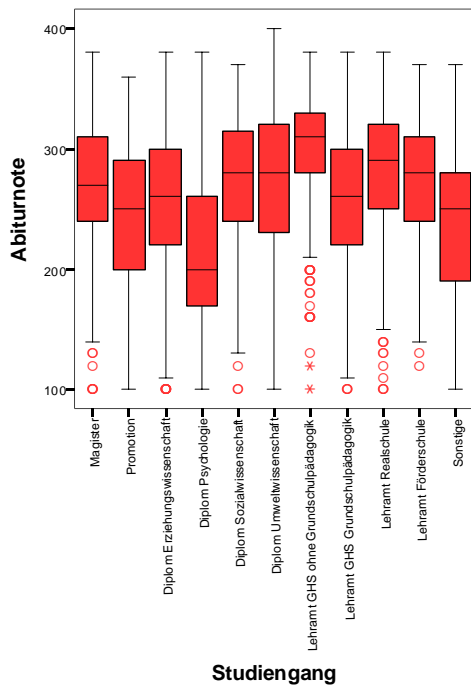


Abiturnote		
Mean		267.92
Median		270.00
Mode		310
Std. Deviation		58.383
Variance		3408.610
Skewness		-.485
Std. Error of Skewness		.034
Kurtosis		-.342
Std. Error of Kurtosis		.068
Range		300
Minimum		100
Maximum		400
Percentiles	25	230.00
	50	270.00
	75	310.00

3.1 Erläutern Sie, was Sie am Histogramm und an den Parametern erkennen können (beachten Sie, dass die Abiturnoten mit 100 multipliziert sind!).

Der Modus, der Median und der Mittelwert mussten miteinander verglichen werden, die Schiefe und (weniger wichtig) die Wölbung (Kurtosis) sollten beurteilt werden; ein Vergleich mit der Normalverteilung war gefragt, außerdem die Beurteilung der Streuung (Vergleich von Standardabweichung, Quartilsabstand und Wertebereich).

3.2 Die folgende Abbildung zeigt vereinfacht die Verteilung der Abiturnoten für Studierende verschiedener Studiengänge. Was entnehmen Sie dieser graphischen Darstellung? Zur Erinnerung: die „Boxes“ umfassen die mittleren beiden Quartile, die „Whiskers“ sind höchstens anderthalb mal so lang wie die „Boxes“, die Kreise und Sternchen sind „Ausreißer“.



Hier war der große Unterschied zwischen den Studierenden des Diplomstudiengangs Psychologie und derer für Lehramt GHS ohne GP herauszuarbeiten: bester gegen schlechtesten Median, längster gegen kürzesten Quartilsabstand. Als Begründung konnte gut auf den NC verwiesen werden (auch im Blick auf GHS mit/ohne GP). Dafür gab es vier Punkte, wer außerdem noch zu anderen Studiengängen etwas zu sagen wusste, bekam noch einen weiteren Punkt.

3.3 Eine Stichprobe von 500 Landauer Studierenden ergibt folgende Mittelwerte für die Abiturnoten der männlichen und der weiblichen Studierenden:

Group Statistics	geschl	N	Mean	Std. Deviation	Std. Error Mean
Abiturnote	M	137	281.97	53.808	4.597
	W	317	268.14	57.534	3.231

t-test for Equality of Means, Equal variances assumed							
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Abiturnote	2.397	452	.017	13.832	5.771	2.492	25.172

Kann man aus dieser Stichprobe schließen, dass die Landauer Studentinnen bessere Abiturnoten haben als ihre männlichen Kommilitonen? (Beachten Sie wieder, dass die Abiturnoten mit 100 multipliziert sind!) Um wie viel im Mittel? Was können Sie mit den Zahlen in den letzten beiden Zellen der unteren Tabelle anfangen? Was sagt Ihnen die Zahl 0.017?

Ja, definitiv kann man das schließen! In der Stichprobe sind die Frauen um 0.13832 Noten besser als die Männer, und das ist auf 1.7 %-Niveau signifikant, d.h. wenn man eine weitere gleich große Stichprobe aus den Landauer Studierenden zieht, dann ist die Wahrscheinlichkeit ganze 1.7 %, dass man einen noch größeren Unterschied zwischen Männern und Frauen herausbekommt, wenn zwischen den beiden in der Grundgesamtheit kein Unterschied besteht. In der Grundgesamtheit sind die Frauen übrigens sogar um 0.20027 besser. Dass rund 2.5 mal so viele Frauen wie Männer in der Stichprobe waren, spielt überhaupt keine Rolle — wer so argumentiert hat, konnte keinesfalls die volle Punktzahl bekommen (das Zahlenverhältnis in der Grundgesamtheit ist übrigens 1513 zu 3739). Wenn man übrigens eine

Stichprobe mit 381 Männern und 352 Frauen zieht (ebenfalls zufällig), dann ist die Notendifferenz auch 0.12534 zugunsten der Frauen.

Mit 0.02492 und 0.25172 konnte kaum jemand etwas anfangen: Dies sind die Grenzen des Intervalls, in die bei 95 % aller denkbaren gleichgroßen Stichproben aus der gleichen Grundgesamtheit die Notendifferenzen zugunsten der Frauen fallen werden, wenn man unsere ursprüngliche Stichprobe zu Grunde legt.

Richtig ist natürlich die Beobachtung, dass eine so geringe Notendifferenz inhaltlich kaum eine Bedeutung hat, erstens sind beide Mittelwerte gleichermaßen „3+“, und außerdem sind Abiturnoten auch keine lupenrein intervallskalierten Variablen (trotz Punktesystem in der Oberstufe).

4. Bivariate Statistik

Im Eurobarometer Oktober-Dezember 2004 wurden den Befragten zwei Fragen zum Einwanderungsproblem vorgelegt:

QB9 For each of the following propositions, tell me if you...?						
(SHOW CARD)						
	(READ OUT)	TOTALLY AGREE	TEND TO AGREE	TEND TO DISAGREE	TOTALLY DISAGREE	DK
(242)	1 We need immigrants to work in certain sectors of our economy	1	2	3	4	5
(243)	2 The arrival of immigrants in Europe can efficiently solve the problem of Europe's aging population	1	2	3	4	5

Count		IMMIGRATION - NEEDED FOR ECONOMY				Total
		totally agree	tend to agree	tend to disagree	totally disagree	
IMMIGRATION - SOLVES AGING PROBLEM	totally agree	1477	501	113	163	2254
	tend to agree	1075	3414	976	617	6082
	tend to disagree	469	2167	3561	1172	7369
	totally disagree	253	791	1184	5030	7258
Total		3274	6873	5834	6982	22963

4.1 Formulieren Sie an Hand dieser Kreuztabelle einige Sätze, die den Zusammenhang zwischen den beiden Variablen beschreiben! Welche Hypothese hätten Sie schon vor Betrachten dieser Tabelle gehabt bzw. welche Hypothese können Sie aus ihr ableiten?

Man könnte an Hand der Frageformulierungen erwarten, dass beide Fragen ähnlich beantwortet werden würden. Das lässt sich aus der Kreuztabelle auch recht gut ablesen (siehe die fett gedruckten Zahlen in der Diagonalen, die sowohl Zeilen- als auch Spaltenmaxima sind). Außerdem kann man an den Randverteilungen sehen, dass die Mehrheiten beiden Aussagen eher oder völlig widersprechen.

4.2 Für diesen Zusammenhang ist Phi 0.811 und Cramers' V 0.468. Erklären Sie den Unterschied dieser beiden eng verwandten Maße. Sind diese beiden Maßzahlen hier hilfreich?

Phi ist die Wurzel aus Φ^2 , Cramer's V^2 ist Φ^2 dividiert durch (hier) 3. Φ^2 ist also $0.811^2 = 0.657721$, dies durch 3 dividiert ergibt 0.21924033, und die Wurzel daraus ist 0.46823107. Aber so genau musste es nicht sein. Die beiden Maßzahlen sind nur begrenzt hilfreich, denn sie werden eher für nominal skalierte Variable genommen und sagen also nichts über die Richtung, sondern nur über die Stärke des Zusammenhangs aus. Beide sind übrigens keine PRE-Maße, sondern hängen beide an χ^2 .

4.3 Für diesen Zusammenhang ist Gamma 0.704 und Kendalls tau-b 0.553. Erklären Sie den Unterschied auch dieser beiden eng verwandten Maße (es genügt zu sagen, warum Gamma größer ist). Sind diese beiden Maßzahlen hier hilfreich?

Gamma hat im Nenner nur die Summe der Zahlen der konkordanten und der diskordanten Paare, tau-b auch noch die der verbundenen Paare, darum muss es praktisch immer kleiner sein. Beide sind hier hilfreich, denn beide sind für ordinal skalierte Variablen entwickelt worden.

5. Multivariate Statistik / Multiple Regression

Im Eurobarometer vom September und Oktober 2002 gibt es zwei Fragebatterien, die sich mit dem Grad der Informiertheit der Befragten im Bereich Biologie, Biotechnologie und Umwelt befassen, und eine weitere, bei der gefragt wird, wie viel Sorgen sich die Befragten um verschiedene Aspekte der Umwelt und des Umweltschutzes machen. Aus diesen drei Fragebatterien ließen sich mit Hilfe einer Faktorenanalyse (die hier nicht dokumentiert ist) insgesamt vier Faktoren extrahieren, die die Befragten einigermaßen genau quantitativ beschreiben. Diese vier Faktoren haben folgende Parameter:

		Wissen über Fragen der Biotechnologie	Wissen über allgemeine Fragen der Biologie und der Genetik	Eigene Einschätzung der Informiertheit in Umweltfragen	Grad der Besorgtheit über den Zustand der Umwelt
N	Valid	16067	16067	13473	12794
	Missing	0	0	2594	3273
Mean		.0000000	.0000000	.0000000	.0000000
Median		-.0243092	.1511354	.0082692	-.0183264
Std. Deviation		1.0000000	1.0000000	1.0000000	1.0000000
Skewness		.187	-.710	.130	.649
Kurtosis		-.898	.075	.108	.585

Wir benutzen nun die drei ersten Faktoren, um herauszufinden, wie stark der Grad der Besorgtheit über den Zustand der Umwelt durch den Grad an Informiertheit und Wissen über Umwelt, Biologie, Genetik und Biotechnologie determiniert wird. Die Regression ergibt folgende Koeffizienten:

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-.001	.009		-.113	.910
Eigene Einschätzung der Informiertheit in Umweltfragen	.202	.009	.202	22.094	.000
Wissen über Fragen der Biotechnologie	-.065	.009	-.066	-7.227	.000
Wissen über allgemeine Fragen der Biologie und der Genetik	.076	.009	.074	8.103	.000

Dependent Variable: Grad der Besorgtheit über den Zustand der Umwelt

5.1 Wie kommt es, dass bei dieser Analyse die standardisierten und die unstandardisierten Koeffizienten (nahezu) identisch sind?

Da wir hier Faktoren als unabhängige und abhängige Variablen genommen haben, sind sie alle standardisiert.

5.2 Haben Sie — nach Ihrer Antwort auf 5.1 — eine Erklärung dafür, warum die Konstante (b_0) nicht exakt gleich 0.0 ist (Hinweis: Beachten Sie die kursiv gedruckte Zeile „Missing“ in der oberen Tabelle)?

Der Mittelwert der beiden ersten Variablen (Faktoren) für all die Fälle, die bei den beiden anderen Faktoren nicht missing sind, ist offenbar nicht genau 0, so dass für die Konstante nun doch ein von 0 leicht verschiedener Wert herauskommt.

5.3 Wie können Sie sich die unterschiedlichen Vorzeichen der beiden Koeffizienten der beiden „Wissen über“-Faktoren erklären?

Schwierige Frage: Je mehr man über Biotechnologie weiß, desto weniger folgt man den Argumenten der Gegner gentechnischer Experimente. Allgemeines Wissen über Biologie und Genetik reicht zur Immunität gegen diese Argumente offenbar nicht aus. Das mag daran liegen, dass die möglicherweise wenigen Spezialisten der Biotechnologie nur die positiven Argumente für Gentechnologie verinnerlicht haben.

5.4 Was bedeutet die Zahl 0.910 am Ende der ersten Zeile der Tabelle?

Kurz gefasst: -0.001 unterscheidet sich nicht signifikant von 0. Genauer: in 91 % aller gleich großen Stichproben aus der gleichen Grundgesamtheit, für die $\beta_0=0$ angenommen werden kann, wird man ein b_0 finden, dass nach seinem Absolutbetrag größer ist als 0.001.

6. Multivariate Statistik / Faktorenanalyse

Im Eurobarometer vom September und Oktober 2002 gibt es eine Fragebatterie, in der gefragt wird, ob bestimmte moderne Technologien das Leben der Befragten in den nächsten 20 Jahren verbessern werden (mit den Antwortmöglichkeiten „wird sich verbessern“, „kein Effekt“, „macht alles nur schlimmer“). Eine der Fragen (kursiv gedruckt) ist in zwei Varianten gestellt worden, wie sich aus den Tabellen unten ergibt (d.h. ein Teil der Befragten hat die linke, der Rest die rechte Variante vorgelegt bekommen). Koeffizienten <0.25 sind weggelassen, die Matrizen sind nach der jeweils höchsten Ladung sortiert, weswegen die Reihenfolgen in den beiden Matrizen nicht identisch sind.

Rotated Component Matrix(a)

	Component	
	1	2
Q9 TECHNOL CHANGE: COMPUTERS/INFO TECH	.765	
Q9 TECHNOL CHANGE: TELECOMMUNICATIONS	.715	.288
Q9 TECHNOL CHANGE: INTERNET	.700	.330
Q9 TECHNOL CHANGE: SOLAR ENERGY	.643	
Q9 TECHNOL CHANGE: NUCLEAR ENERGY		.806
Q9 TECHNOL CHANGE: NANOTECHNOLOGY	.295	.734
Q9 TECHNOL CHANGE: BIOTECHNOLOGY (A)		.689
Q9 TECHNOL CHANGE: SPACE EXPLORATION	.410	.586
Q9 TECHNOL CHANGE: MOBILE PHONES	.491	.518

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 3 iterations.

Rotated Component Matrix(a)

	Component	
	1	2
Q9 TECHNOL CHANGE: COMPUTERS/INFO TECH	.752	
Q9 TECHNOL CHANGE: TELECOMMUNICATIONS	.745	
Q9 TECHNOL CHANGE: INTERNET	.698	.299
Q9 TECHNOL CHANGE: SOLAR ENERGY	.657	
Q9 TECHNOL CHANGE: MOBILE PHONES	.574	.435
Q9 TECHNOL CHANGE: NUCLEAR ENERGY		.802
Q9 TECHNOL CHANGE: GENETIC ENGINEERING (B)		.749
Q9 TECHNOL CHANGE: NANOTECHNOLOGY	.298	.695
Q9 TECHNOL CHANGE: SPACE EXPLORATION	.411	.613

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 3 iterations.

6.1 Beurteilen Sie die Unterschiede zwischen den beiden Ladungsmatrizen. Hat die Ersetzung des Begriffs „Biotechnology“ durch „Genetic engineering“ einen Einfluss auf das Ergebnis gehabt?

Offensichtlich ja, aber nicht nur hinsichtlich der Ladung der beiden Alternativformulierungen, auch einige der anderen Technologien werden offenbar in einem anderen Licht gesehen, wenn „Biotechnologie“ durch „Gentechnik“ ersetzt wird.

6.2 Geben Sie den beiden Faktoren Namen:

Hier haben viele gesagt, dieser Faktor beschreibe die Einstellung zu den guten Technologien, der andere die zu den bösen. Das ist natürlich Ansichtssache. Andere haben gemeint, dieser Faktor beschreibe die positive Einstellung zu modernen Technologien, der andere die negative — diese Benennung geht in die Irre, denn in dieser Variante müssten die beiden Faktoren höchst negativ korreliert sein; sie sind aber unkorreliert. Besser hat mir gefallen, wenn der erste

Faktor als die Einstellung zu den allgemein bekannten und von vielen persönlich genutzten Technologien, der andere als die Einstellung zu den noch in der Entwicklung befindlichen und kaum individuell genutzten Technologien.

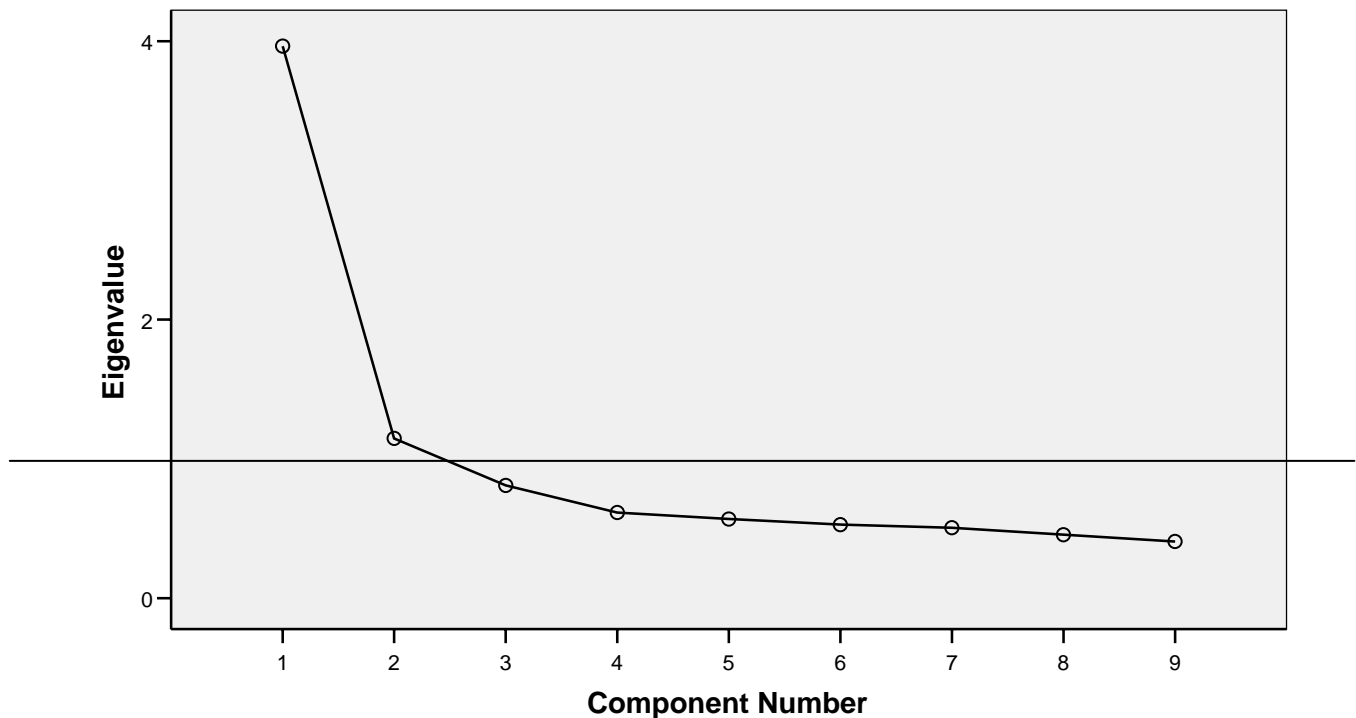
6.3 Wie beurteilen Sie die Einfachstruktur? Sind die Faktoren sauber genug getrennt?

Mobiltelefone und Raumfahrt laden in beiden Faktorenanalysen auf beide Faktoren, im übrigen ist die Trennung einigermaßen sauber. Aber eine bessere Einfachstruktur ist wohl nicht möglich (da gibt es nichts „auszuwählen“, wie einige gemeint haben, sondern nur zu optimieren, und wenn das optimierte Ergebnis noch unbefriedigend ist, dann liegt das in der Natur der Dinge und nicht an der Methode.

Die Ladung von Raumfahrt auf beiden Faktoren kann man sich dann so erklären, dass viele Leute Satelliten nutzen (zum Fernsehen, auch zum Telefonieren und zum Navigieren, viele andere Leute aber an Astronauten in Raumstationen und Space Shuttles gedacht haben. Bei den Mobiltelefonen könnte man argumentieren, dass hier viele Leute diese Technologie benutzen, zugleich aber auch sehen, welcher immenser Fortschritt von Jahr zu Jahr festzustellen ist.

6.4 Hier finden Sie den Scree-Plot zu einer der beiden vorstehenden Auswertungen:

Scree Plot



Analysis weighted by WEIGHT EURO 15

Halten Sie danach die Extraktion von zwei Faktoren noch für gerechtfertigt? (Ja / Nein / Kommt darauf an)

Wenn ja: warum? (2 P)

Hier konnte man das Kaiser-Kriterium anführen, aber auch den deutlichen Knick bei der 2.

Wenn nein: warum nicht? Wie viele hätte man statt dessen extrahieren sollen? (2 P)

Dass der Knick bei der 2 der deutlichste ist, konnte man aber auch zur Argumentation verwenden, dass ein Faktor ausgereicht hätte (der hätte dann eine allgemeine Einstellung zur Technik und zum technischen Fortschritt dargestellt). Auch mit drei Faktoren konnte man argumentieren: die Punkte 4 bis 9 liegen nun wirklich fast exakt auf einer Geraden.

(Antworten auf beide Unterfragen werden gewertet! Wenn Sie dezidiert der Auffassung sind, dass nicht zwei, sondern einer oder drei 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.)