

Statistik für Informationsmanager

Nachklausur vom 2. Dezember 2003

Name:

Matrikelnummer:

Erlaubte Hilfsmittel: Außer Schreibwerkzeug diesmal auch Taschenrechner; Antworten bitte im Anschluss an die Fragen, Rückseiten können für Antworten und Entwürfe genutzt werden; Entwürfe bitte anschließend durchstreichen!

Mit der elektronischen Veröffentlichung meiner Klausurnote zusammen mit meiner Matrikelnummer bin ich einverstanden:

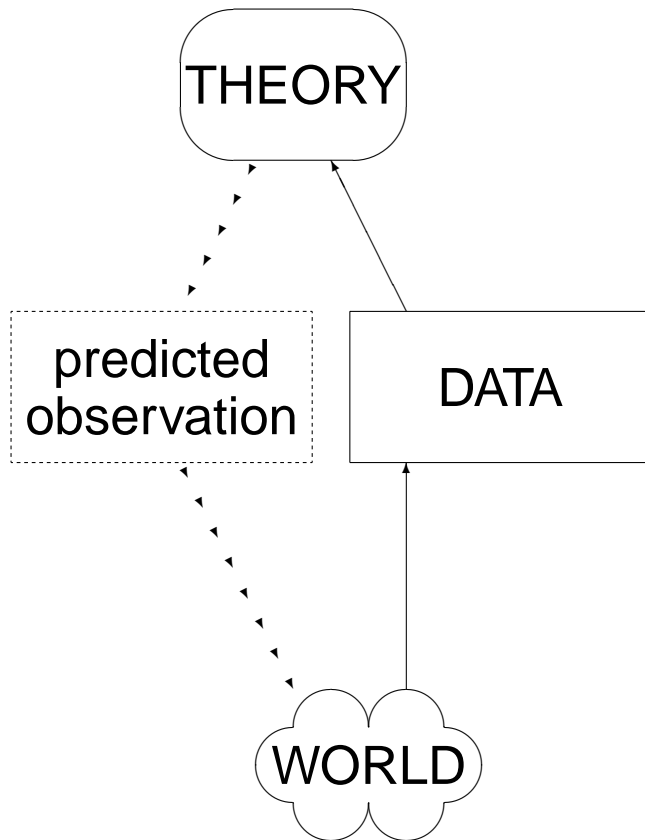
Ja Unterschrift
 Nein (keine Unterschrift heißt nein)

Für die Teilaufgaben wird es folgende Punktzahlen geben:

Aufgabe	Maximal erreichbar	Sie haben erreicht
1	$2 \cdot (1 + 2) = 6$	
2	$8 \cdot (3/8 + 3/8) = 6$	
3	$3 \cdot (1 + 1) = 6$	
4.1	$3 \cdot 2 = 6$	
4.2	$3 \cdot 2 = 6$	
4.3	3	
4.4	3	
5	$4 \cdot 2 + 1 = 9$	
6	$2 \cdot 1 + 3 + 2 \cdot 2 = 9$	
Summe	54	

1. Grundlagen

Erinnern Sie sich an die mehrfach gezeigte Abbildung;



Zeichnen Sie ein, wohin

a. Feldexperimente (1P)

b. Teilnehmende Beobachtung (1P)

hingehören! Geben Sie jeweils eine kurze Begründung!

a. Feldexperimente gehören dorthin, weil ... (2P)

b. Teilnehmende Beobachtung gehört dorthin, weil ... (2P)

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
Arbeitsaufwand eines/r Studierenden für ein Modul eines Studienganges in Std.					
Gesamtnote auf einem Abschluss- oder Zwischenzeugnis im IM-Studiengang					
Bestellnummer					
Aktienkurs					
Uhrzeit					
Rangplatz in einer Bundesliga-Tabelle					
Dauer einer Aktivität					
Haarfarbe					

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. Wartezeiten von 20 am Nachmittag zufällig befragten Studierenden an der Essensausgabe der Mensa (jeweils vom Zeitpunkt der Ankunft am Ende der Warteschlange bis zum Finden eines Sitzplatzes (in Minuten):

2	2	2	2.5	2.5	2.5	2.5	3	3	3
4	5	6	6	7	7	7	8	8	9
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

2. Zufriedenheit der gleichen Studierenden mit dem Mensaessen dieses Tages (auf einer Skala von 1=sehr unzufrieden bis 5=sehr zufrieden):

5	4	4	4	3	3	4	5	2	1
3	4	5	3	4	3	4	2	1	1
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

3. Studiengang der gleichen Studierenden:

Lehramt	Magister	Lehramt	IM BSc	Pädagogik	Magister	Lehramt	Informatik	CV	Magister
Informatik	IM BSc	CV	CV	Pädagogik	Magister	Informatik	CV	IM BSc	Informatik
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

4. Bivariate Statistik

Betrachten Sie bitte noch einmal die (natürlich fiktiven!) Daten aus Aufgabe 3. Dort sind 20 Studierende mit den drei Variablen „Wartezeit in der Mensa“, „Zufriedenheit mit dem Mensaessen“ und „Studiengang“ beschrieben (in allen drei Listen sind die 20 fiktiven Studierenden natürlich in der gleichen Reihenfolge aufgeführt). Der Übersichtlichkeit ist hier noch einmal eine zusammenfassende Tabelle.

Stud. Nr.	Wartezeit	Zufriedenheit	Studiengang
1	2	5	Lehramt
2	2	4	Magister
3	2	4	Lehramt
4	2.5	4	IM BSc
5	2.5	3	Pädagogik
6	2.5	3	Magister
7	2.5	4	Lehramt
8	3	5	Informatik
9	3	2	CV
10	3	1	Magister
11	4	3	Informatik
12	5	4	IM BSc
13	6	5	CV
14	6	3	CV
15	7	4	Pädagogik
16	7	3	Magister
17	7	4	Informatik
18	8	2	CV
19	8	1	IM BSc
20	9	1	Informatik

Fragen:

4.1 Können Sie zwischen den drei Variablen Zusammenhänge entdecken? (Z.B.: „Je länger jemand wartet, desto (un)zufriedener ist er/sie.“ oder „Studierende des Studiengangs X warten länger als die anderen / sind leichter zufrieden zu stellen als die anderen.“) Schreiben Sie alle Zusammenhänge auf, die Ihnen aufgefallen sind. (Für jeden entdeckten Zusammenhang 2 P, maximal 6 P — denn zwischen drei Variablen gibt es höchstens drei Zusammenhänge!)

4.2 Beschreiben Sie kurz für jeden Zusammenhang, den Sie entdeckt haben, wie Sie ihn entdeckt haben. (Je 2 P, maximal 6 P)

4.3 Erscheinen Ihnen die statistischen Zusammenhänge, die Sie entdeckt haben, plausibel? Begründen Sie Ihre Einschätzung. (3 P)

4.4 Welche Maßzahlen würden Sie berechnen, um die Stärke der jeweils entdeckten Zusammenhänge zahlenmäßig ausdrücken zu können? (Achtung! Beachten Sie, dass die drei Variablen möglicherweise unterschiedliches Skalenniveau haben und Sie Zusammenhangsmaße nur für zwei Variable gleichen Skalenniveaus kennen gelernt haben! Beschreiben Sie also gegebenenfalls auch, wie Sie die Variablen transformieren müssen, bevor sie Koeffizienten berechnen können!) (bis zu 3 P)

5. Multivariate Statistik / Multiple Regression

In der ALLBUS-Studie 1998 wurde u.a. das Nettoeinkommen erfragt. Außerdem sollten sich alle Befragten zu ihrem Freizeitverhalten äußern (siehe nächste Aufgabe! — Beachten Sie, welche Bedeutung die Ziffern 1 bis 5 zur Beschreibung der Freizeitaktivitäten haben!). Die folgende Tabelle listet die Koeffizienten einer multiplen Regression auf 12 der 24 Freizeitvariablen auf (die Hinzunahme der anderen 12 hätte das Regressionsergebnis — $R^2 = 0.129$, $R = 0.359$ — nicht signifikant verbessert).

Koeffizienten

	Nicht standardisierte Koeffizienten B	Standardisierte Koeffizienten Beta	Signifikanz
(Konstante)	4032.21 DEM		8.3349E-39
V27 Freizeit: Essen oder trinken gehen	-331.17 DEM	-0.211	7.6703E-15
V40 Freizeit: Klassische Konzerte, Theater etc.	-282.48 DEM	-0.143	2.5032E-07
V38 Freizeit: Besuch von Sportveranstaltungen	-228.87 DEM	-0.148	4.9511E-09
V39 Freizeit: Kino, Pop- und Jazzkonzerte, Tanzen	151.63 DEM	0.091	0.00277469
V24 Freizeit: Einfach nichts tun, faulenzen	102.55 DEM	0.086	0.00029288
V28 Freizeit: Besuch Nachbarn, Freunde, Bek.	146.69 DEM	0.085	0.000603
V18 Freizeit: Zeitschriften lesen	113.69 DEM	0.091	0.00029187
V17 Freizeit: Bücher lesen	-113.56 DEM	-0.098	0.00016892
V26 Freizeit: Yoga, Meditation, autogenes Training	137.18 DEM	0.093	0.00023274
V21 Freizeit: Mit dem Computer beschäftigen	-88.10 DEM	-0.088	0.00070377
V19 Freizeit: LP, CD, Kassetten hören	85.16 DEM	0.078	0.00457858
V34 Freizeit: Kirchgang	-85.06 DEM	-0.060	0.01271998

Abhängige Variable: V318 Befr.: Nettoeinkommen, Offene Abfrage

Was bedeuten die beiden Zahlen **4032.21 DEM** und **-331.17 DEM** in den ersten beiden Zahlen der Tabelle? (2P).

Berechnen Sie das zu erwartende Nettoeinkommen eines/r Befragten, der/die alle 12 Freizeitaktivitäten „täglich“ betreibt (d.h. jeweils „1“ angegeben hat) im Vergleich zu jemand, der/die alle 12 Freizeitaktivitäten „nie“ betreibt (d.h. jeweils „5“ angegeben hat). (2 P).

Nennen Sie die drei Freizeitaktivitäten, die sich nach diesem Analyseergebnis am stärksten auf das Nettoeinkommen auswirken, in der richtigen Reihenfolge. Beachten Sie dabei den Unterschied zwischen standardisierten und unstandardisierten Koeffizienten! (2 P)

Drücken Sie in einem ganzen Satz aus, was die Zahl **0.000603** (in der Mitte der letzten Spalte) bzw. ihr Kehrwert 1658 bedeutet! (3 P)

6. Multivariate Statistik / Faktorenanalyse

In der ALLBUS-Studie von 1998 wurden u.a. Fragen zum Freizeitverhalten gestellt. Die Fragen lauteten:

Nun einige Fragen zu Ihrer Freizeit.

Geben Sie bitte zu jeder der Tätigkeiten auf dieser Liste an, wie oft Sie das in Ihrer Freizeit machen: täglich (1), mindestens einmal jede Woche (2), mindestens einmal jeden Monat (3), seltener (4) oder nie (5).

Aus diesen 24 Fragen ließ sich die nachstehende Faktorladungsmatrix gewinnen. Koeffizienten kleiner als 0.26 sind zur Erleichterung des Auffindens einer Einfachstruktur weggelassen, außerdem sind die Variablen nach ihren höchsten Faktorladungen sortiert. Interpretieren Sie die beiden Faktoren, indem Sie ihnen möglichst prägnante Namen geben. Begründen Sie Ihre Namensgebung.

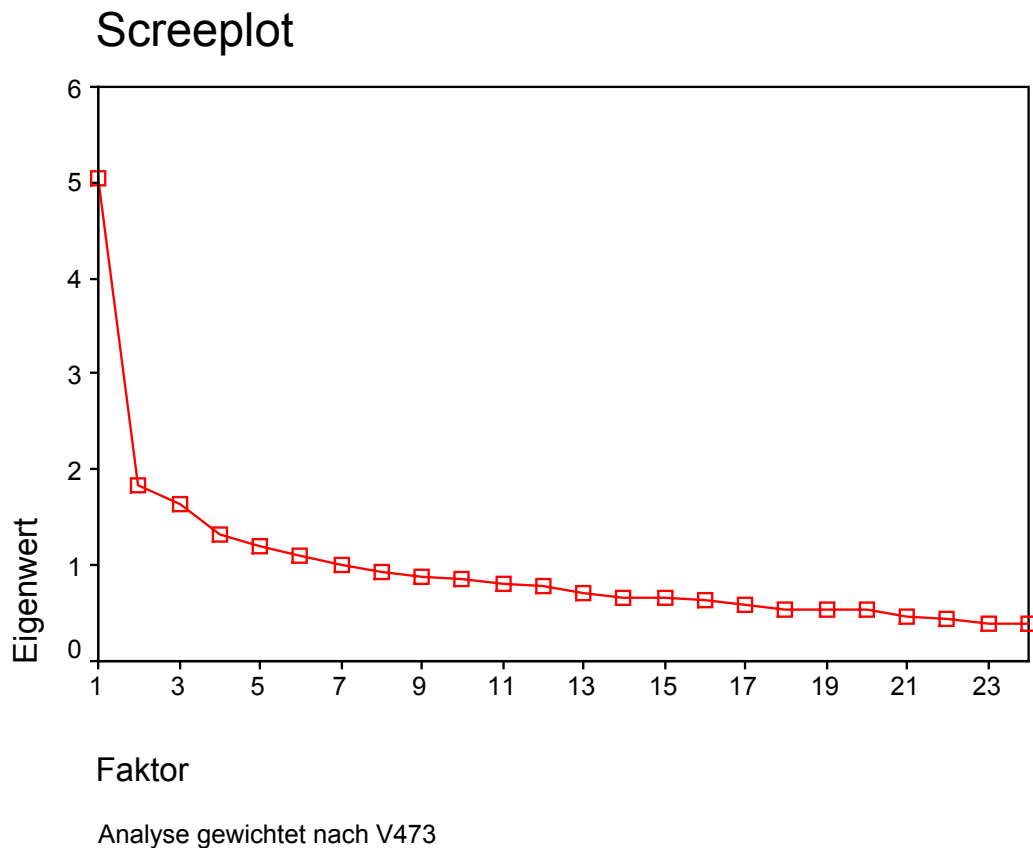
	Komponente	
	1	2
V39 Freizeit: Kino, Pop- und Jazzkonzerte, Tanzen	.729	
V21 Freizeit: Mit dem Computer beschäftigen	.658	
V19 Freizeit: LP, CD, Kassetten hören	.634	
V20 Freizeit: Videokassetten anschauen	.615	
V23 Freizeit: Sich privat weiterbilden	.605	
V37 Freizeit: Aktive sportliche Betätigung	.602	.316
V27 Freizeit: Essen oder trinken gehen	.554	
V22 Freizeit: Internet, Online-Dienste nutzen	.532	
V38 Freizeit: Besuch on Sportveranstaltungen	.467	
V24 Freizeit: Einfach nichts tun, faulenzeln		
V35 Freizeit: Kunst, musische Tätigkeiten	.323	.514
V33 Freizeit: Ehrenamtliche Tätigkeiten		.497
V40 Freizeit: Klassische Konzerte, Theater etc.	.425	.478
V17 Freizeit: Bücher lesen		.473
V34 Freizeit: Kirchgang	-.322	.449
V31 Freizeit: Ausflüge oder kurze Reisen	.325	.431
V26 Freizeit: Yoga, Meditation, Autogenes Training	.331	.422
V29 Freizeit: Besuch Familie, Verwandtschaft		.422
V30 Freizeit: Gesellschaftsspiele in Familie		.421
V25 Freizeit: Spazieren gehen, wandern		.408
V32 Freizeit: Beteiligung in Politik		.408
V36 Freizeit: Basteln, Reparaturen		.322
V28 Freizeit: Besuch Nachbarn, Freunde, Bekannte		.311
V18 Freizeit: Zeitschriften lesen		.257

Faktor 1 beschreibt die latente Eigenschaft ... (1 P)

Faktor 2 ... (1 P)

Wie beurteilen Sie die Einfachstruktur? Widersprechen die mehrfachen hohen Faktorladungen bei V37, V35, V40, V17, V34, V31 und V26 Ihrer Interpretation oder stützen sie sie? Warum? (3 P)

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)

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

(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.)