

Statistik für Informationsmanager

Nachklausur vom 18. Dezember 2002

Name:

Matrikelnummer:

Erlaubte Hilfsmittel: Außer Schreibwerkzeug keine; 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:

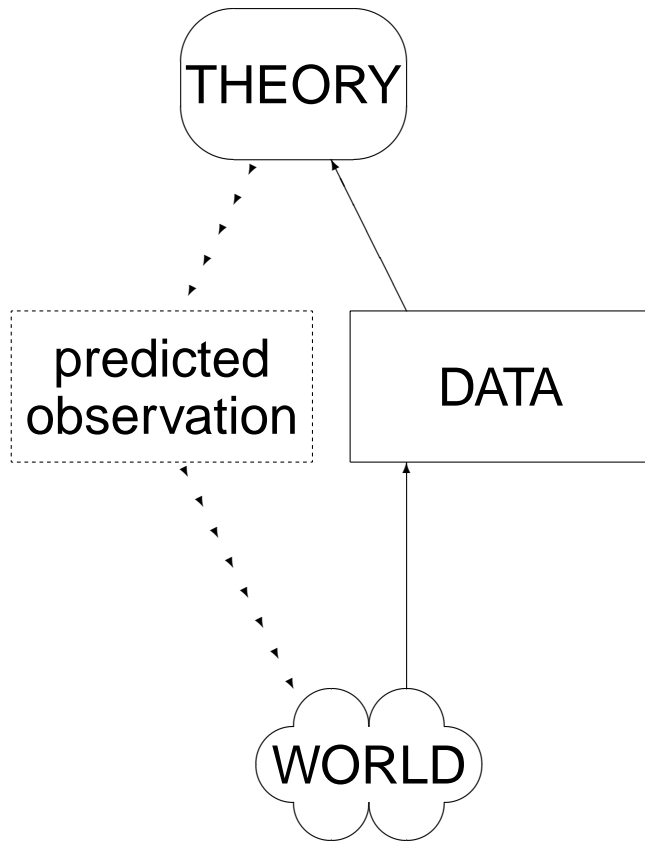
Ja Unterschrift
Nein (keine Unterschrift heißt nein)

Für die Teilaufgaben wird es folgende Punktzahlen geben:

Aufgabe	Maximal erreichbar	Sie haben erreicht
1	$2*(1 + 2)$	
2	$8*(3/8 + 3/8)$	
3	$3*(1 + 1)$	
4.1	$4*1$	
4.2	$2*1 + 3$	
4.3	$5*1$	
5	$4*2 + 1$	
6	$3*1.5 + 2 + 2*1.75$	
Summe	51	

1. Grundlagen

Erinnern Sie sich an die mehrfach gezeigte Abbildung;



Zeichnen Sie ein, wohin

- a. Fragebogenerhebung (1P)
 - b. Bivariate Statistik mit PRE-Maßen (1P)
- hingehören! Geben Sie jeweils eine kurze Begründung!
- a. Fragebogenerhebung gehört dorthin, weil ... (2P)

- b. Bivariate Statistik mit PRE-Maßen 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
Stimmenanteil der CDU bei der Bundestagswahl 2002					
Lebensalter					
Klausurnote					
Fachsemesterzahl					
Bestellnummer					
Einschätzung des Risikos, abends überfallen zu werden, auf einer Skala von 1= „ganz unwahrscheinlich“ bis 5= „sehr wahrscheinlich“					
Gewählte(r) Kandidat(in) auf einem Stimmzettel					
Fruchtbarkeitsrate (Anzahl der Lebendgeburten auf 10.000 Frauen im Alter von 15 bis 45 Jahren)					

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. Preis von 25 CD-RWs in zwanzig verschiedenen Geschäften:

17.90	16.90	20.00	14.99	17.95	16.80	17.50	22.22	15.80	13.99
12.50	14.49	19.99	18.80	17.50	17.50	17.50	16.80	15.55	18.88
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

2. Fachsemesterzahl von 20 Informatik-Absolventen:

27	25	14	13	21	13	37	17	11	25
25	31	13	25	11	10	29	21	13	25
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

3. Wochentage, an denen diese 20 Informatik-Absolventen geboren wurden:

Mittwoch	Mittwoch	Samstag	Freitag	Dienstag	Freitag	Montag	Dienstag	Donnerstag	Freitag
Sonntag	Donnerstag	Donnerstag	Sonntag	Dienstag	Samstag	Donnerstag	Samstag	Montag	Dienstag
Mittelwert:									
Median									
Modus									
gegebenenfalls Begründung für nicht berechnete Parameter:									

4. Bivariate Statistik

4.1 Nominale Variable

In der Politbarometerbefragung vom September 1998 ergab sich für die Befragten nach Bundesland und Konfession folgende Kreuztabelle („Nord“ fasst die Länder der „alten“ Bundesrepublik ohne Bayern und Baden-Württemberg, aber mit Berlin-West, zusammen, „Ost“ fasst die alte DDR zusammen):

		NSO Nord/Süd/Ost			Gesamt
		1.00 Nord	2.00 Süd (BY, BW)	3.00 Ost	
1.00 katholisch	Anzahl	239	194	18	451
	% von NSO Nord/Süd/Ost	34.9%	57.7%	7.1%	35.5%
2.00 evangelisch	Anzahl	318	107	72	497
	% von NSO Nord/Süd/Ost	46.5%	31.8%	28.6%	39.1%
3.00 sonstige oder konfessionslos	Anzahl	127	35	162	324
	% von NSO Nord/Süd/Ost	18.6%	10.4%	64.3%	25.5%
	Anzahl	684	336	252	1272
	% von NSO Nord/Süd/Ost	100.0%	100.0%	100.0%	100.0% <input type="checkbox"/>

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

		Wert
Lambda	Symmetrisch	.156
	KONF Konfession abhängig	.228
	NSO Nord/Süd/Ost abhängig	.060
Goodman-und-Kruskal-Tau	KONF Konfession abhängig	.113
	NSO Nord/Süd/Ost abhängig	.099
Unsicherheitskoeffizient	Symmetrisch	.114
	KONF Konfession abhängig	.110
	NSO Nord/Süd/Ost abhängig	.119

Beurteilen Sie an Hand der Kreuztabelle die Konfessionsverteilung in den verschiedenen Teilen Deutschlands (1P).

Welche Angaben aus der Kreuztabelle — oder welche Koeffizienten — haben Sie dazu am besten benutzen können? (1P)

Bitte bilden Sie einen Satz, mit dem Sie die Bedeutung der Zahl 0.228 (Lambda mit Konfession als abhängiger Variable) erläutern (1P).

Warum sind die beiden asymmetrischen Lambda-Koeffizienten so verschieden (0.228 gegenüber 0.060), obwohl die beiden anderen Paare asymmetrischer Koeffizienten (0.113 / 0.099 bzw. 0.110 / 0.119) keine solchen Unterschiede aufweisen? (1P)

4.2 Ordinale Variable

Eine weitere Auswertung desselben Datensatzes fragt nach dem Zusammenhang zwischen der Stärke des Interesses an Politik bei den Befragten und der Größe der Gemeinde, in der sie leben. Die Stärke des Politikinteresses wurde auf einer fünfstufigen Skala mit den Punkten 1=sehr stark, 2=stark, 3=etwas, 4=kaum, 5=gar nicht erhoben, die Größe der Gemeinden war gruppiert: 1: < 2000, 2: 2000 bis 5000, ..., 8: > 500.000 Einwohner.

Dabei haben sich die folgenden Zusammenhangsmaße ergeben:

		Wert
Cramer's V		0.127
Lambda	Symmetrisch	0.011
Uncertainty coefficient	Symmetrisch	0.018
Somers-d	Symmetrisch	-0.131
	V5 EINWOHNER DER GEMEINDE abhängig	-0.147
	V149 STÄRKE POLITIKINTERESSE abhängig	-0.118
Kendall-Tau-b		-0.132
Kendall-Tau-c		-0.128
Gamma		-0.169
Anzahl der gültigen Fälle		1254

Fragen:

Ist die Verwendung der Maße Gamma und Somers' D für ordinale Variable hier angemessen? (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!)

Ist die Verwendung der nominalen Maße hier überhaupt angemessen? (0.5P) Wenn Ja, warum, wenn nein, warum nicht? (wie oben)

Formulieren Sie den Zusammenhang in einem Satz: „**Je größer die Gemeinde, desto das Politikinteresse.**“ Achten Sie auf die Bedeutung der oben erläuterten Zahlencodes (3P).

4.3 Metrische Variable

Für den Lehrbericht 1999/2000 habe ich auf der Basis derjenigen Studierenden, für die wir die Abiturnote und die Vordiplomnote kennen, folgende Korrelationen (Produkt-Moment-Korrelationskoeffizienten) zwischen der Abiturnote und den Vordiplom-Gesamt- und Einzelnoten gefunden:

.	Abi	VD ges.	Math	Anw	TI	PrInf	ThInf
Abiturnote	1.000						
Vordiplomnote gesamt	0.483	1.000					
Mathematik	0.437	0.804	1.000				
Anwendungsfach	0.265	0.615	0.384	1.000			
Technische Informatik	0.362	0.677	0.431	0.194	1.000		
Praktische Informatik	0.409	0.718	0.472	0.287	0.378	1.000	
Theoretische Informatik	0.330	0.773	0.588	0.355	0.367	0.441	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!)

Wie beurteilen Sie, dass die Anwendungsfachnote mit den meisten anderen Variablen relativ niedrig korreliert? (1P)

Wie erklären Sie sich, dass die Abiturnote mit der Mathematiknote im Vordiplom am höchsten korreliert? (1P)

Wie erklären Sie sich, dass die höchsten Korrelationen überhaupt diejenigen zwischen der Gesamtnote im Vordiplom und den Einzelnoten im Vordiplom sind? (1P)

Warum konnte ich die Zahlen in der oberen Dreiecksmatrix der Tabelle oben einfach weglassen? (1P)

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)	7.236	.349		20.705	.000
Organisation und Management (x_1)	-.032	.011	-.270	-2.830	.010
Einführung in die BWL (x_2)	-.048	.011	-.450	-4.427	.000
Empirische Grundlagen (x_3)	-.039	.008	-.447	-4.697	.000
Kosten- und Leistungsrechnung (x_4)	-.015	.010	-.157	-1.497	.149
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).

Berechnen Sie die geschätzte Gesamtnote von jemand, der in allen vier Teilklausuren 25 Punkte erreicht hat (2P).

Welche Teilklausur(en) hat/haben sich am stärksten auf die Gesamtnote ausgewirkt? (2P)

Kann man bei der kleinen Stichprobe (bei lediglich 26 Teilnehmern) davon ausgehen, dass die Ergebnisse über den Zusammenhang zwischen den Punktzahlen der Teilklausuren und der Gesamtnote auf die Grundgesamtheit übertragbar ist? (2P)

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

6. Multivariate Statistik / Faktorenanalyse

Die Befragung von ca. 750 Einwohnern der Verbandsgemeinde Asbach im Westerwald bezog sich u.a. auf die Beschreibung von Eigenschaften des Ortsteils, in dem die Befragten wohnen. Es wurden drei Faktoren extrahiert und rotiert, es ergab sich folgende Ladungsmatrix:

	Component		
	1	2	3
F0801 abwechslungsreich	.469	.500	
F0802 kinderfreundlich	.342	.346	.271
F0803 provinziell	-.216		.434
F0804 gastlich	.579	.275	.206
F0805 gemütlich	.641		.351
F0806 geschäftig, arbeitsam		.626	
F0807 modern, fortschrittlich	.322	.708	
F0808 farblos, trist	-.647		
F0809 langweilig	-.720	-.225	
F0810 lebendig, aktiv	.513	.535	
F0811 traditionsverbunden		.319	.531
F0812 überfüllt, hektisch	-.375	.425	-.435
F0813 überschaubar			.573
F0814 unpersönlich	-.684		
F0815 weltoffen, tolerant, aufgeschlossen	.460	.453	
F0816 schmutzig	-.331		-.290
F0817 erwachsenenbetont	-.409		
F0818 sympathisch	.608	.353	.336
F0819 umweltbewusst		.471	.283
F0820 behindertenfreundlich		.702	
F0821 ruhig			.670
F0822 sicher			.583
F0823 interessant	.523	.574	
F0824 fremdenfreundlich	.441	.239	
F0825 hohe Wohnqualität	.450	.230	.474

Dies war der Text der Frage:

Welche der folgenden Eigenschaften würden Sie Ihrer Stadt/Ortsgemeinde zuordnen ?

Bitte bei den folgenden Eigenschaften die **Skalen** von 1 – 4 **ankreuzen**:

- 1** = „trifft völlig zu“;
- 2** = „trifft teilweise zu“;
- 3** = „trifft kaum zu“;
- 4** = „trifft nicht zu“

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.

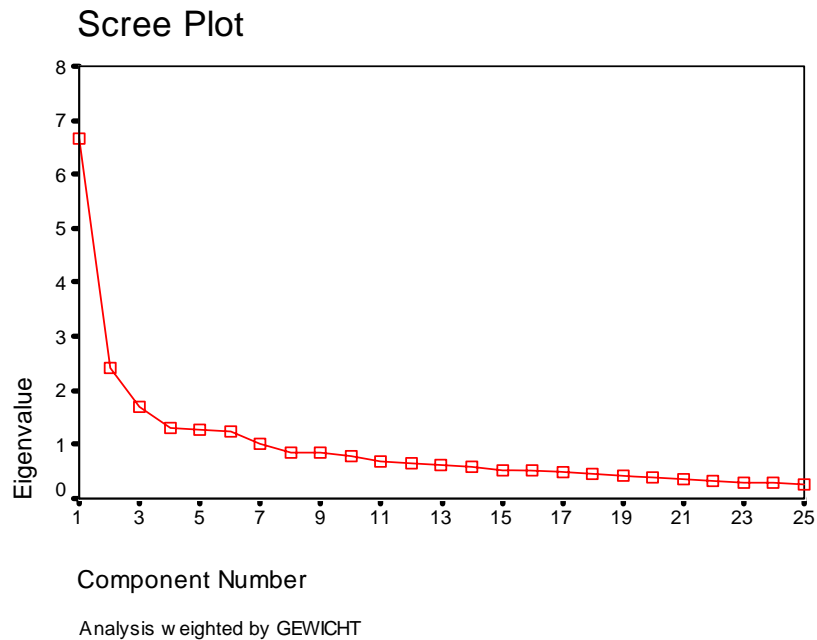
Faktor 1 beschreibt den Ortsteil unter dem Gesichtspunkt ... (1.5P)

Faktor 2 ... (1.5P)

Faktor 3 ... (1.5P)

Die Mittelwerte der drei Faktorwerte, berechnet aus den Faktorwerten der Bewohner des Ortsteils Krautscheid, ergaben sich zu 0.523, -0.580 und 0.316. Wie sehen die Krautscheider ihren Ortsteil? (2P)

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



Halten Sie danach die Extraktion von drei Faktoren noch für gerechtfertigt? (Ja / Nein / Kommt darauf an)
Wenn ja: warum? (1.75P)

Wenn nein: warum nicht? Wie viele hätte man statt dessen extrahieren sollen? (1.75P)

(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.)