



**KLAUSUR: MULTIVARIATE ANALYSIS**  
**PRÜFER: PROF. DR. B. ERICHSON**

WS 2007/08

PRÜFUNGSNUMMER: 2624

Als Hilfsmittel sind zugelassen: Taschenrechner lt. Aushang des Prüfungsamtes.

Die folgenden Aufgaben sind alle zu bearbeiten. (Die Sollbearbeitungszeit ist für jede Aufgabe angegeben. Wobei ein Punkt einer Minute entspricht.) Der Klausuraufgabenbogen umfasst drei Seiten. Fehlende Werte bitte gleich in den Aufgabenboden eintragen und diesen mit abgeben!

**Aufgabe 1: Grundwissen (9 Punkte)**

- Warum sind multivariate Analysemethoden für das Marketing von besonderer Wichtigkeit?
- Welche zwei Gruppen von multivariaten Analysemethoden lassen sich unterscheiden? Nennen Sie beispielhaft jeweils zwei Methoden für jede Gruppe!
- Bezüglich der Anwendbarkeit von Statistiken und Analysemethoden ist das Skalenniveau der Daten zu beachten. Welche Skalenniveaus werden unterschieden?
- Die fundamentale Formel der Streuungszerlegung läßt sich wie folgt schreiben:

$$\sum_{k=1}^K (y_k - \bar{y})^2 = \sum_{k=1}^K (\hat{y}_k - \bar{y})^2 + \sum_{k=1}^K (y_k - \hat{y}_k)^2$$

Interpretieren Sie diese Gleichung!

**Aufgabe 2: Faktorenanalyse (12 Punkte)**

- Erläutern Sie kurz die Aufgabenstellung der Faktorenanalyse! Worin lag ihre erste Anwendung?
- Die Grundgleichung der Faktorenanalyse lautet:  $Z = A \cdot P$ . Beschreiben Sie die Elemente dieser Gleichung kurz!
- Was sind Kommunalitäten? Berechnen Sie diese für die erste 1 Variable nach der Extraktion von zwei Faktoren!
- Berechnen Sie die fehlenden Eigenwerte der Faktoren und geben Sie deren Varianzerklärungsanteil an.
- Zur Bestimmung der Faktorenanzahl lässt sich das Kaiserkriterium nutzen. Was besagt dieses?

Faktor	Eigenwert	% der Varianz	Faktorladungen		
			Variablen	Faktor 1	Faktor 2
1	2,5	50	Fettgehalt	0,92	-0,02
2			Kaloriengehalt	0,89	0,06
3	0,7	14	Süße	0,28	-0,7
4			Knackigkeit	0,03	0,8
5	0,14	2,8	Preis	0,89	0,2

Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax

**Aufgabe 3: Regressionsanalyse (22 Punkte)**

Herr Brand, Produktmanager bei Proper&Clean, möchte untersuchen, wie er den Absatz des von ihm betreuten Produktes "Softy" beeinflussen kann. Zu diesem Zweck hat er Daten über 30 Beobachtungsperioden zusammengestellt und eine Regressionsanalyse mit SPSS durchgeführt, die nachfolgend wiedergegeben ist:

Leider sind Herrn Brand infolge eines Defekts an seiner Festplatte alle Daten sowie auch einige Zahlen im SPSS-Output verloren gegangen. Können Sie ihm helfen?

*Bitte Seite wenden!*



**Deskriptive Statistik**

	N	Mittelwert	Standardabweichung
Absatz	42	87,4786	15,6
Preis	42	24,9536	1,9
Distribution	42	62,2595	11

**Modellzusammenfassung**

Modell	R	R-Quadrat	Standardfehler des Schätzers
1	,706(a)		11,36682

a Einflußvariablen : (Konstante), Distribution, Preis

**ANOVA(b)**

	Quadratsumme	df
Regression	5014,5	2
Residuen	5038,9	39
Gesamt	10053,4	41

a Einflußvariablen : (Konstante), Distribution, Preis

b Abhängige Variable: Absatz

**Koeffizienten(a)**

	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	
	B	Standardfehler	Beta	T
(Konstante)	137,4	36,3		3,8
Preis	-3,3	1,1	-,4	
Distribution	,53	,2		

a Abhängige Variable: Absatz

- a) Wie lautet das Modell von Herrn Brand und wie die von ihm geschätzte Regressionsfunktion?
- b) Beurteilen Sie die Güte des Modells! Berechnen Sie dazu das Bestimmtheitsmaß zur Sicherheit auf zwei verschiedene Arten! Interpretieren Sie das Ergebnis kurz!
- c) Zur Prüfung der Güte empfiehlt sich auch die Berechnung der F-Statistik. Warum?
- d) Führen Sie den F-Test durch! (Irrtumswahrscheinlichkeit  $\alpha = 5\%$ , F-Tabelle in Aufgabe 4)
- e) Untersuchen Sie mittels der Beta-Werte, welche der beiden Marketing-Variablen stärker wirkt?
- f) Prüfen Sie die Signifikanz der Marketing-Variablen (Irrtumswahrscheinlichkeit  $\alpha = 5\%$  )!
- g) Einen primären Anwendungsbereich der Regressionsanalyse bildet die Untersuchung von Kausalbeziehungen. Bildet die Signifikanz einer "unabhängigen Variablen" einen Beweis dafür, daß diese einen kausalen Einfluß auf die abhängige Variable ausübt? Begründen Sie Ihre Antwort!

**Aufgabe 4: Varianzanalyse (8 Punkte)**

Im Rahmen der Mediaanalyse der AG.MA werden jährlich von mehreren Instituten etwa 30.000 Personen zur Ermittlung der Reichweiten von Zeitschriften befragt.

Die nachfolgende Tabelle zeigt in Form von Indexzahlen die durchschnittlichen Reichweiten für die einzelnen Institute über die Zeit.

Institute ↓	Jahr →	1994	1993	1992	1991	1990	1989	1988	Mittelwerte
Infratest		-	97	100	102	106	100	110	102.50
MMA		101	103	89	98	110	109	107	102.43
GFM Getas		99	95	106	93	101	107	104	100.71
Ifak		100	-	107	91	90	85	81	92.33
Sample		113	104	-	-	-	-	-	108.50
Marplan		95	106	97	104	95	96	94	98.14
Basisresearch		92	95	101	106	96	102	105	99.57
Mittelwerte		100.00	100.00	100.00	99.00	99.67	99.83	100.17	99.81

Es soll geklärt werden, ob die Unterschiede bezüglich der Reichweiten, die die Institute ermitteln, signifikant sind. Diese Frage ist für die Medien von hoher Brisanz, da von den Reichweiten die Preise für die Schaltung von Werbeanzeigen abhängen.

- a) Welches Analyseverfahren würden Sie zur Untersuchung der vorliegenden Fragestellung heranziehen. Begründen Sie Ihre Wahl!
- b) Formulieren Sie das Prüfkriterium des hier anzuwendenden Analyseverfahrens!

**Seite 3 beachten!**



- c) Berechnen Sie das Prüfkriterium unter Verwendung der unten gegebenen SPSS-Auswertung! Interpretieren Sie das Ergebnis! Nutzen Sie zur Bestimmung des kritischen Wertes die F-Tabelle!
- d) Wie groß ist das Bestimmtheitsmaß? Berechnen Sie!

**ONEWAY ANOVA**  
Reichweite

	Quadratsumme	df
Zwischen den Gruppen	603,4	6
Innerhalb der Gruppen	1421	35

**F-Tabelle (95 % Vertrauenswahrscheinlichkeit)**

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10

$v_1$  = Zahl erklärender Variablen im Zähler (J),  $v_2$  = df im Nenner (K-J-1)

**Aufgabe 5: Clusteranalyse (9 Punkte)**

Ein Kommilitone, der zur Zeit ein Praktikum in einer Zoohandlung absolviert, soll dort eine Segmentierung des reichhaltigen Katzensortiments durchführen. Folgende Daten stehen zur Verfügung:

Katze	Pflege	Futter	Kosten
MainCoon	5	2	3
Perser	7	4	3
Tiger	2	7	6
Panther	1	6	7

$$d_{k,l} = \left[ \sum_{j=1}^J |x_{kj} - x_{lj}|^r \right]^{\frac{1}{r}}$$

Da er nicht weiterkommt, bittet er Sie um Hilfe.

- a) Vervollständigen Sie die nachfolgende Distanzmatrix mittels obiger Metrik, wobei  $r=2$  ! Wie heißt das verwendete Distanzmaß? **Tragen Sie Ihre Ergebnisse in die untenstehende Tabelle ein!**

Distanzen	Main Coon	Perser	Tiger	Panther
Main Coon	0	2,83	6,56	6,93
Perser	2,83	0	6,56	
Tiger	6,56	6,56	0	
Panther	6,93			0

- b) Führen Sie nach dem Single-Linkage-Verfahren einen ersten Fusionierungsschritt durch und berechnen Sie die neuen Distanzen! **Tragen Sie Ihre Ergebnisse in die untenstehende Tabelle ein!**

Distanzen			
	0		
		0	
			0

- c) Zur Prüfung von Cluster-Lösungen wird u.a. die folgende Formel verwendet. Wozu dient sie? Interpretieren Sie die Formel!  $t = \frac{\bar{X}(J,G) - \bar{X}(J)}{S(J)}$

**Viel Erfolg!**