

Diplomprüfung Evaluation und Forschungsmethodik SS 2008

*Alle Fragen müssen beantwortet werden. Eine Möglichkeit zur Auswahl besteht nicht.
ACHTUNG: Alle Berechnungsschritte müssen deutlich werden!*

1 FAKTORENANALYSE

- (1) Vor Durchführung einer Faktorenanalyse muss die Güte der Korrelationsmatrix geprüft werden. Sie erhalten ein Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin von 0.264. Der Bartlett-Test auf Sphärizität ergibt: $\chi^2(df=28)=32.62$, $p=.248$. Erläutern Sie, ob Ihre Daten für die Auswertung mit Hilfe der Faktorenanalyse geeignet sind. (2P)
- (2) Bei einer Faktorenanalyse wurden nach dem Kaiser-Kriterium 3 Faktoren extrahiert. Dabei ergab sich die folgende Tabelle (vor der Rotation).

Faktor	Eigenwert	% Varianz	kumulierte % Varianz
1	4.681	33.438	33.438
2	3.344	19.671	53.109
3	2.538	18.129	71.238

Warum ist die Ausgabe fehlerhaft? (2P)

- (3) Sie lassen sich Faktorwerte ausgeben. Wie sind diese Faktorwerte grundsätzlich zu interpretieren? (1P)

2 CONJOINT ANALYSE

- (1) In der vorliegenden Tabelle sind die Teilnutzenwerte als Ergebnis einer Conjoint-Analyse eingetragen. Vervollständigen Sie die Tabelle. (Nicht auf dem Aufgabenblatt!) (2P)

Faktor	Teilnutzenwert	Transformierter Teilnutzen
A1	-2.00	
A2	-1.50	
A3	0.00	
A4	3.50	
B1	-0.25	
B2	0.50	
B3	-0.25	
C1	-6.00	
C2	-0.50	
C3	3.00	
C4	3.50	

- (2) Welche 2 Aussagen über mögliche Präferenzveränderungen und Anteile am Gesamtnutzen kann man der Ergebnistabelle entnehmen? (2P)

3 QUANTITATIVES VORGEHEN II - GRAPHISCHE ANALYSEN

- (1) Die Varianzanalyse verlangt verschiedene Voraussetzungen. Geben Sie für jede der folgenden Voraussetzungen an, wie Sie diese grafisch anhand von Datenanalysegraphen kontrollieren können. Geben Sie jeweils an, was auf x- bzw. y-Achse aufgetragen wird.
- (a) Normalverteilung
 - (b) Varianzhomogenität (2P)

4 QUANTITATIVES VORGEHEN III - INFERENZSTATISTISCHE ANALYSEN

- (1) Was ist in der Inferenzstatistik der Unterschied zwischen der „Teststärke“ und der „Effektstärke“? Definieren Sie diese Begriffe kurz. (2P)
- (2) Nennen Sie 3 Faktoren, die die „Teststärke“ beeinflussen und skizzieren Sie kurz, wie dieser Einfluss aussieht. (3P)

5 VARIANZANALYSE & ANCOVA/MANOVA

Sie interessieren sich für die Veränderung der Konzentrationsleistung beim Abbau von Blutalkohol. Hierzu bilden Sie per Randomisierung 5 unabhängige Stichproben freiwilliger Teilnehmer. Die ersten 4 Gruppen müssen so viel Alkohol konsumieren, dass – je nach Stichprobe – 0.2, 0.4, 0.8, oder 1.6 Promille Blutalkoholkonzentration (BAK) erreicht werden; die fünfte Gruppe erhält nur alkoholfreie Getränke. Alle Teilnehmer müssen zu 6 Messzeitpunkten (vor dem Trinken, 1, 2, 3, 4, 5 Stunden nach Einnahme) einen Standard-Aufmerksamkeitstest (d2) bearbeiten. Sie wissen, dass die Leistung bei hohen BAKs schlecht wird und von einer Abbaurate von 0.1 Promille pro Stunde auszugehen ist. Ihre Forschungsfrage ist, ob sich die Leistung bei hohen BAKs langsamer erholt als bei niedrigen BAKs. Die ANOVA weist beide Haupteffekte und die Interaktion als signifikant aus.

- (1) Um welches Versuchsdesign handelt es sich (uni/multivariat, ein/zwei/dreifaktoriell, mit/ohne Messwiederholung)? (1P)
- (2) Skizzieren Sie ein (ideales) Interaktionsdiagramm für diesen Versuch. (3P)
- (3) Ein Reviewer wirft Ihnen vor, der signifikante Interaktionseffekt belege Ihre Forschungsfrage nicht, da er aufgrund des Versuchsdesigns aus 2 Gründen trivial sei – und er hat recht. Welche 2 Punkte müssen Sie am Versuchsdesign ändern, damit ein signifikanter Interaktionseffekt nur wegen unterschiedlicher Erholungsgeschwindigkeiten signifikant wird? (2P)

6 LOGISTISCHE REGRESSION

- (1) Nach der Durchführung einer Logistischen Regression sollte vom Untersucher die Güte der Modellschätzung beurteilt werden. Erläutern Sie die Güte der Modellschätzung anhand des folgenden SPSS-Outputs und gehen Sie dabei auf folgende Kriterien und deren Bedeutung ein:
 - (a) Omnibus-Tests der Modellkoeffizienten
 - (b) Hosmer-Lemeshow-Tests (4P)

Omnibus-Tests der Modellkoeffizienten

		Chi-Quadrat	df	Sig.
Schritt 1	Schritt	28.104	3	.000
	Block	28.104	3	.000
	Modell	28.104	3	.000

Hosmer-Lemeshow-Test

Schritt	Chi-Quadrat	df	Sig.
1	8.864	8	.354

7 EPIDEMIOLOGIE

- (1)
 - (a) Was versteht man unter der „Chance“ einer Krankheit und was unter dem „Risiko“ einer Krankheit? Erläutern Sie beide Konzepte kurz in eigenen Worten und geben Sie die zugehörige Formel an.
 - (b) Was versteht man unter dem „Odds Ratio“ und was unter dem „Relativen Risiko“? Erläutern Sie beide Konzepte kurz in eigenen Worten und geben Sie die zugehörige Formel an. (2P)
- (2) Erläutern Sie anhand einer Formel, wann das Odds Ratio ähnliche Werte annimmt wie das Relative Risiko. (1P)
- (3) Erläutern Sie den Zusammenhang zwischen der Strenge eines diagnostischen Kriteriums sowie der Sensitivität und der Spezifität eines diagnostischen Tests. (1P)

8 DISKRIMINANZANALYSE

- (1) Welche Daten brauchen Sie für eine Diskriminanzanalyse? (1P)
- (2) Beschreiben Sie 3 Gemeinsamkeiten/Ähnlichkeiten von Diskriminanzanalyse und Faktorenanalyse. (3P)

9 KONFIGURATIONSFREQUENZANALYSE

In einer Studie zur Wirkung von Public Health Maßnahmen wurde an einer Zufallsstichprobe von 300 Personen die Intelligenz ($IQ > 100 = "+"$, $IQ < 100 = "-"$), das Alter ($> 40 = "+"$, $< 40 = "-"$) und der Cholesterinwert ($> 200 \text{ mg/dl} = "+"$; $< 200 \text{ mg/dl} = "-"$) erhoben. In der Tabelle finden Sie die Verteilung der 300 Personen auf die Symptomkombinationen.

Intelligenz	Alter	Cholesterin	f	e	χ^2
"+"	"+"	"+"	58	53.36	0.40
"+"	"+"	"-	51	48.60	0.12
"+"	"-	"+"	31	41.36	2.60
"+"	"-	"-	41	37.67	0.29
"-	"+"	"+"	29	35.08	1.05
"-	"+"	"-	31	31.95	0.03
"-	"-	"+"	39	27.19	5.13
"-	"-	"-	20	24.77	0.92

- (1) Gibt es irgendeinen Zusammenhang der Variablen?
 - (a) Berechnen Sie die entsprechende Statistik.
 - (b) Mit welchem kritischen Wert müssten Sie diese Statistik vergleichen? Geben Sie die Bezeichnung der Tabelle, die Freiheitsgrade und den Wahrscheinlichkeitswert an, unter denen Sie nachsehen würden.
 - (c) An der entsprechenden Stelle im Statistikbuch finden Sie den Wert 9.49. Wie lautet Ihre Folgerung?
 - (d) Würde die klassische KFA ein Syndrom identifizieren und wenn ja, welches? (4P)
- (2) In einer klassischen KFA wird zunächst ein Test berechnet, ob überhaupt ein Zusammenhang zwischen den Variablen besteht und anschließend für jede Symptomkombination betrachtet, ob ein Syndrom auffällig von der Erwartung abweicht. Welches Problem ergibt sich bei diesem Vorgehen und wie kann man ihm begegnen? (2P)

10 MULTIPLE REGRESSIONSANALYSE

Sie möchten gerne prognostizieren, wie lange ein Patient in einer Klinik zur Behandlung verweilen wird. Dazu stehen Ihnen für eine große Stichprobe von Patienten die Verweildauer und verschiedene Daten aus der Eingangsuntersuchung zur Verfügung: der Schweregrad der Erkrankung bei Einlieferung, das Alter, ein einfacher Bluttest und 3 weitere Laborwerte. Die letzten 3 Laborwerte kosten extra und sollen gegebenenfalls wegfallen.

- (1) Wie prüfen Sie statistisch korrekt, ob die 3 zusätzlichen Laborwerte die Prognose verbessern? (1P)
- (2) Welche Kenngrößen aus der Auswertung benutzen Sie, um zu bewerten, welches Symptom für die Vorhersage am wichtigsten ist? (1P)
- (3) Welche Kenngrößen benutzen Sie, um für einen neuen Patienten zu prognostizieren, wie lange er/sie wahrscheinlich in der Klinik bleiben wird? (1P)

11 LINEARE STRUKTURANALYSE

In einer Untersuchung soll unter Verwendung der Methode der Linearen Strukturanalyse die „Belastung der Schüler“ als ein hochkomplexes Konstrukt untersucht werden, das von unterschiedlichen Einflussgrößen beeinflusst wird. Für die Modellerstellung werden folgende 19 Variablen definiert:

Latente endogene Variable: Belastung der Schüler (BELAST)

Latente exogene Variablen: Schule als Institution (SCHULE), unterrichtender Lehrer (LEHRER), Mutter und/oder Vater (ELTERN)

Indikatorvariablen von BELAST: physische Beeinträchtigungen (PB), reaktive Depressivität (RD), Angst (AN), Stress (ST)

Indikatorvariablen von SCHULE: schlechte lernorganisatorische Bedingungen (LO), schlechtes AG- u. Freizeitangebot (AG), negative Sicht auf Schule (NS)

Indikatorvariablen von LEHRER: leistungsfordernder Lehrer (LF), nicht lernunterstützender Lehrer (LU), nicht schülerzentrierter Lehrer (SZ), kränkender Lehrer (KR)

Indikatorvariablen von ELTERN: schlechtes Verhältnis zu Vater (VV), schlechtes Verhältnis zu Mutter (VM), mangelndes Schulinteresse der Eltern (SI), Streit wegen Schule (SS)

- (1) Erstellen Sie für die skizzierte Fragestellung das entsprechende Pfaddiagramm. (4P)

12 ALLGEMEIN

Im Rahmen eines Praktikums geraten Sie in die "data Mining"-Abteilung eines großen Konzerns, der Daten über das Kaufverhalten seiner Kunden per Kundenkarte sammelt. Sie sollen herausfinden, ob es Zusammenhänge zwischen Geschlecht der Käufer, geographischer Region, Jahreszeit, Erwerb eines Produkts A und Erwerb eines Produkts B gibt, da man die beiden Produkte evtl. bündeln will.

- (1) Welche 2 Auswertungsverfahren würden Sie vorschlagen? Begründen Sie Ihre Wahl. (2P)
 (2) Welches Verfahren würden Sie wählen, wenn Ihnen in diesem Konzern keines der üblichen Statistikpakete zur Verfügung steht? (1P)

ANHANG

Tabelle der Standardnormalverteilung

p	z	y
0.01	-2.33	0.03
0.02	-2.05	0.05
0.03	-1.88	0.07
0.04	-1.75	0.09
0.05	-1.64	0.10
0.06	-1.55	0.12
0.07	-1.48	0.13
0.08	-1.41	0.15
0.09	-1.34	0.16
0.10	-1.28	0.18
0.11	-1.23	0.19
0.12	-1.17	0.20
0.13	-1.13	0.21
0.14	-1.08	0.22
0.15	-1.04	0.23
0.16	-0.99	0.24
0.17	-0.95	0.25
0.18	-0.92	0.26
0.19	-0.88	0.27
0.20	-0.84	0.28
0.21	-0.81	0.29
0.22	-0.77	0.30
0.23	-0.74	0.30
0.24	-0.71	0.31
0.25	-0.67	0.32

p	z	y
0.26	-0.64	0.32
0.27	-0.61	0.33
0.28	-0.58	0.34
0.29	-0.55	0.34
0.30	-0.52	0.35
0.31	-0.50	0.35
0.32	-0.47	0.36
0.33	-0.44	0.36
0.34	-0.41	0.37
0.35	-0.39	0.37
0.36	-0.36	0.37
0.37	-0.33	0.38
0.38	-0.31	0.38
0.39	-0.28	0.38
0.40	-0.25	0.39
0.41	-0.23	0.39
0.42	-0.20	0.39
0.43	-0.18	0.39
0.44	-0.15	0.39
0.45	-0.13	0.40
0.46	-0.10	0.40
0.47	-0.08	0.40
0.48	-0.05	0.40
0.49	-0.03	0.40
0.50	0.00	0.40

p	z	y
0.51	0.03	0.40
0.52	0.05	0.40
0.53	0.08	0.40
0.54	0.10	0.40
0.55	0.13	0.40
0.56	0.15	0.39
0.57	0.18	0.39
0.58	0.20	0.39
0.59	0.23	0.39
0.60	0.25	0.39
0.61	0.28	0.38
0.62	0.31	0.38
0.63	0.33	0.38
0.64	0.36	0.37
0.65	0.39	0.37
0.66	0.41	0.37
0.67	0.44	0.36
0.68	0.47	0.36
0.69	0.50	0.35
0.70	0.52	0.35
0.71	0.55	0.34
0.72	0.58	0.34
0.73	0.61	0.33
0.74	0.64	0.32
0.75	0.67	0.32

p	z	y
0.76	0.71	0.31
0.77	0.74	0.30
0.78	0.77	0.30
0.79	0.81	0.29
0.80	0.84	0.28
0.81	0.88	0.27
0.82	0.92	0.26
0.83	0.95	0.25
0.84	0.99	0.24
0.85	1.04	0.23
0.86	1.08	0.22
0.87	1.13	0.21
0.88	1.17	0.20
0.89	1.23	0.19
0.90	1.28	0.18
0.91	1.34	0.16
0.92	1.41	0.15
0.93	1.48	0.13
0.94	1.55	0.12
0.95	1.64	0.10
0.96	1.75	0.09
0.97	1.88	0.07
0.98	2.05	0.05
0.99	2.33	0.03