

Diplomprüfung Evaluation und Forschungsmethodik WS 2008/09

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

1 FAKTORENANALYSE

- (1)
 - (a) Welches Problem der Faktorenanalyse wird mit der Wahl der Extraktionsmethode gelöst?
 - (b) Nennen Sie 2 unterschiedliche Extraktionsmethoden und beschreiben Sie deren methodischen Hauptunterschied.
 - (c) Welche Folgen hat die Wahl der Extraktionsmethode für die Interpretation der Faktorlösung? (4P)
- (2) Erklären Sie kurz die Begriffe „Eigenwert“ und „Kommunalität“. (2P)

2 PROBABILISTISCHE SKALIERUNG

In einem Seminar zur Methodenlehre bittet der Dozent die Studenten einzuschätzen, inwieweit sich 3 Politiker für das Amt des Bundeskanzlers eignen. Die Beurteilung wurde in einem Dominanzpaarvergleich durchgeführt. Dabei ergab sich folgende Matrix der Präferenzwahrscheinlichkeiten:

	Politiker A	Politiker B	Politiker C
Politiker A	50%	16%	30%
Politiker B	84%	50%	68%
Politiker C	70%	32%	50%

- (1) Berechnen Sie die Skalenniveaus für die Eignung der 3 Politiker nach dem „Law of Comparative Judgement“. Stellen Sie die einzelnen Zwischenschritte Ihres Rechenvorgangs dar. (3P)
- (2) Interpretieren Sie das Ergebnis Ihrer Berechnung. Machen Sie dabei eine Aussage auf dem höchstmöglichen Skalenniveau. (2P)
- (3)
 - (a) Was versteht man unter zirkulären Triaden?
 - (b) Nennen Sie ein Beispiel, wie es zu einem gehäuften Auftreten von zirkulären Triaden kommen kann. (2P)

3 LOGISTISCHE REGRESSION

- (1) Erläutern Sie kurz 2 Voraussetzungen, die erfüllt sein müssen, damit sich Daten für eine Auswertung mit Hilfe der logistischen Regression eignen. (2P)

In einer Untersuchung wurde die Schwierigkeit einer Skipiste [Intervallskala von 1 bis 10] und die Jahreszeit [1=Herbst, 2=Winter, 3=Frühjahr; Referenzkategorie: Herbst] als Risikofaktoren für einen Sturz [0=kein Sturz, 1= Sturz] während der Fahrt untersucht.

Die folgende Tabelle zeigt das Ergebnis der Auswertung mit Hilfe der Logistischen Regression:

	Regressions- koeffizient B	Standard- fehler	p- Wert	Differenz für OR	OR
Konstante	-0.849	2.179	0.442		0.428
Schwierigkeit	0.226	0.139	0.042	1 Stufe	1.248
Jahreszeit	-1.346	0.204	0.036	Winter vs. Herbst	0.260
	-0.928	0.234	0.039	Frühjahr vs. Herbst	0.396

- (2) Interpretieren Sie das Ergebnis. (3P)

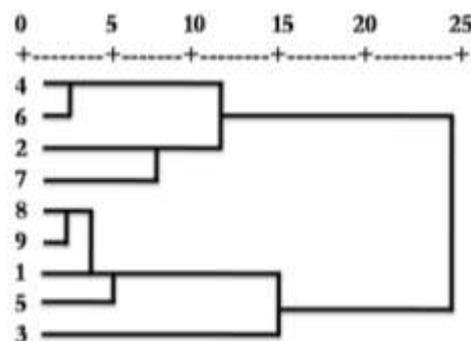


4 CLUSTERANALYSE

- (1) (a) Welche Anforderung stellt die hierarchische Clusteranalyse an das Datenniveau?
(b) Welche Möglichkeiten bestehen, einen gegebenen Datensatz an diese Anforderungen anzupassen? (3P)
- (2) Nehmen sie auf Grundlage der folgenden Zuordnungsübersicht sowie des Dendrogramms Stellung zu einer möglichen Clusterlösung (3P)

Zuordnungsübersicht

Schritt	Zusammengeführte Cluster		Koeffizienten	Erstes Vorkommen des Clusters		Nächster Schritt
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	4	6	41.932	0	0	6
2	8	9	57.304	0	0	3
3	1	8	69.175	0	2	4
4	1	5	76.844	3	0	7
5	2	7	104.990	0	0	6
6	2	4	162.896	5	1	8
7	1	3	184.810	4	0	8
8	1	2	294.468	7	6	0



5 VARIANZANALYSE

Sie untersuchen die Wirkung eines neuen Konzentrationstrainings auf die Leistungsfähigkeit von Schülern, indem Sie per Zufall je 30 Schüler einer von 3 Gruppen zuweisen: Eine Gruppe erhält das neue Konzentrationstraining, die zweite Gruppe ein klassisches Entspannungstraining, die dritte Gruppe keinerlei Training.

Sie messen die Wirkung der Manipulation durch den standardisierten IBF-Test („Intelligenz-Basis-Faktoren“), der laut Testmanual aus den faktorenanalytisch ermittelten und voneinander unabhängigen Dimensionen „sprachgebundene kognitive Funktionen“, „zahlgebundene kognitive Funktionen“, „räumliches Vorstellungsvermögen“ und „Merkfähigkeit“ besteht.

- (1) Um welches Versuchsdesign handelt es sich hier? (1P)
- (2) Der statistische Test wird nicht signifikant. Ein Kommilitone schlägt Ihnen vor, lediglich die „sprachgebundenen kognitiven Funktionen“ als abhängige Variable in eine Analyse einzubeziehen und, um „die Signifikanz zu erhöhen“, die „zahlgebundenen kognitiven Funktionen“ als Kovariate zu verwenden. Warum ist der Vorschlag keine gute Idee? (2P)
- (3) Sie erheben nochmals je 30 Schüler; außerdem verwenden Sie für die Auswertung tatsächlich nur die Skalenwerte der „sprachgebundenen kognitiven Funktionen“. Durch die Vergrößerung der Stichprobe wird der statistische Test nun signifikant. Als Sie Ihrem Betreuer stolz die Ergebnisse präsentieren, verlangt er (1) dass Sie ein nichtparametrisches Verfahren verwenden sollen und (2) dass Sie die Konzentrationsgruppe mit der Entspannungsgruppe vergleichen sollen.
 - (a) Welche(s) Verfahren verwenden Sie klassisch für (1), welches für (2)?
 - (b) Gibt es Alternativen? (3P)

6 KONFIGURATIONSFREQUENZANALYSE

Sie untersuchen den Zusammenhang dreier dichotomer Variablen mit einer KFA.

A	B	C	f	e
+	+	+	10	13.48
+	+	-	50	47.19
+	-	+	20	9.85
+	-	-	25	34.48
-	+	+	10	15.41
-	+	-	60	53.93
-	-	+	10	11.26
-	-	-	40	39.41

Der χ^2 -Wert für die gesamte Tabelle ist bereits berechnet beträgt 16.86. Die χ^2 -Werte für die Vierfeldertafeln A x B und A x C betragen 0.03 und 4.59.

- (1) Prüfen Sie, ob der trivariate Zusammenhang zwischen A, B und C signifikant ist. Dazu werden Sie etwas rechnen müssen (2 Nachkommastellen). (4P)

7 EPIDEMIOLOGIE

- (1) Nennen Sie 3 Ziele der epidemiologischen Forschung (3P)
- (2) Herr K. wird mit einem neuen Test positiv auf Diabetes getestet. Der Arzt sagt, der neue Test könne sehr gut Gesunde von Kranken unterscheiden. Herr K. möchte nun wissen, wie wahrscheinlich es ist, dass er tatsächlich Diabetes hat.
- (a) Welchen Kennwert würden Sie berechnen?
- (b) Skizzieren Sie die zugehörige Tafel mit Beschriftung und geben Sie unter Berücksichtigung Ihrer Beschriftung die Formel des von Ihnen gewählten Kennwertes an. (3P)

8 METAANALYSE

- (1) Nennen Sie 3 Kriterien, die einen Einfluss auf die Güte einer Metaanalyse haben. Beschreiben Sie jeweils kurz, wie dieser Einfluss aussieht („je ... desto...“). (3P)
- (2) Cook und Campbell (1979) führen an, dass bei einer Bewertung von Primärstudien die Validität des statistischen Schlusses und die Konstruktvalidität zu beachten sind. Beschreiben Sie kurz, was unter diesen Validitätsbereichen zu verstehen ist und geben Sie jeweils 2 Beispiele an, die bei den jeweiligen Validitätsarten als einschlägig gelten. (3P)

9 DISKRIMINANZANALYSE

- (1) Bitte notieren Sie die korrekte Version der folgenden Beschreibung der Diskriminanzanalyse auf Ihrem Antwortbogen (nicht hier auf der Angabe!):
- (a) „In der Diskriminanzanalyse betrachtet man metrische/ordinale/kategoriale unabhängige Variablen, die Merkmale/Prädiktoren/Kriterien genannt werden und eine/mehrere metrische/ordinale/kategoriale abhängige Variable/n.“
- (b) Zweck der Diskriminanzanalyse ist die Vorhersage der Gruppenzugehörigkeit einer Person. Dazu benutzt man Diskriminanzfunktionen/Klassifikationsfunktionen.
- (c) Werden in der Eichstichprobe 4 Gruppen untersucht, kann es maximal Diskriminanzfunktionen und ... Klassifikationsfunktionen geben.“ (3P)
- (2) Welches alternative Verfahren wäre ebenfalls zur Vorhersage der Gruppenzugehörigkeit einer Person geeignet und auf welcher Art von Daten operiert dieses Verfahren üblicherweise? (2P)



10 LOGLINEARE MODELLE

Aus internationalen Patientenakten von Krebspatienten extrahieren Sie die Art der Erkrankung (Brust-, Lungen-, Haut-, Prostata-, Darmkrebs), das Geschlecht (m/w), die Art der Berufstätigkeit (schwere körperliche Arbeit/leichte körperliche Arbeit/Bürotätigkeit) sowie das Herkunftsland (Schweden/Spainien/Indien/Japan). Das Studienprotokoll schreibt vor, dass Sie aus jedem Herkunftsland N=500 Personenakten ziehen.

- (1) Wieviele Werte des Effektparameters $\mu_{\text{Erkrankung}}$ können frei variieren? (1P)
- (2) Im Modelltest wird der Effekt $\mu_{\text{Krebs} \times \text{Geschlecht}}$ zwangsläufig signifikant. Warum? (2P)
- (3) Bei exploratorischen Untersuchungen ohne a priori Hypothesen verwendet man häufig ein automatisches Verfahren zur Auswahl des am besten passenden hierarchischen loglinearen Modells. Beschreiben Sie dieses Verfahren kurz. (2P)

11 LINEARE STRUKTURANALYSE

- (1) Erläutern Sie, inwieweit die Lineare Strukturanalyse eine Erweiterung der Pfadanalyse darstellt? (1P)

In einer Untersuchung soll unter Verwendung der Methode der Linearen Strukturanalyse die „Lesefähigkeit der Schüler“ als ein Konstrukt untersucht werden, das von unterschiedlichen Einflussgrößen beeinflusst wird. Für die Modellerstellung werden folgende 9 Variablen definiert:

Latente endogene Variable: Lesefähigkeit der Schüler (LESE)

Latente exogene Variablen: Einstellung zum Fernsehen (EIN_TV), Einstellung zu Büchern (EIN_BUCH)

Indikatorvariablen von EIN_TV: tägliche Zeit vor dem Fernseher (TVZEIT), Anzahl der Fernsehgeräte im Haushalt (TVZAHL)

Indikatorvariablen von EIN_BUCH: tägliche Lesezeit (BZEIT), Anzahl eigener Bücher (BZAHL)

Indikatorvariablen von LESE: Deutschnote (NOTE), Lesetest (TEST)

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

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