

**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 0.854. Der Bartlett-Test auf Sphärizität ergibt: $\chi^2(df=28)=554.63$, $p<.001$. Erläutern Sie, ob Ihre Daten für die Auswertung mit Hilfe der Faktorenanalyse geeignet sind. (2P)
- (2) An unterschiedlichen Stellen während einer Faktorenanalyse beeinflussen Entscheidungen der auswertenden Person das Ergebnis der Analyse. Erläutern Sie drei Entscheidungen, die der Anwender während der Auswertung treffen muss und stellen Sie jeweils zwei mögliche Entscheidungsalternativen dar. (3P)
- (3) Bei einer Faktorenanalyse wurden nach dem Kaiser-Kriterium drei Faktoren extrahiert. Dabei ergibt sich nach erfolgter Rotation die folgende Tabelle. Ist die Ausgabe zwangsläufig fehlerhaft? Begründen Sie Ihre Antwort. (2P)

Faktor	Eigenwert nach Rotation	erklärte Varianz [%]	kumulierte erklärte Varianz [%]
1	4.681	33.438	33.438
2	3.344	23.884	57.321
3	0.923	6.593	63.915

2 SIGNALENTDECKUNGSTHEORIE

- (1) Welche Parameter liefert die SDT? Wie werden diese Parameter berechnet? Wie sind die Parameter zu interpretieren? (5P)
- (2) In einem Signalentdeckungsexperiment werden 100mal SN (Signal) und 200mal N (Noise) dargeboten. Die Versuchsperson erzielt 53 Treffer und gibt 64mal Falschen Alarm. Normalverteilung und gleiche Varianzen vorausgesetzt: Wie groß sind d' und β ? (2P)

3 CONJOINT-ANALYSE

- (1) Die Regressionsanalyse wird als kompositionelles Verfahren bezeichnet. Grenzen Sie das Vorgehen der Conjoint-Analyse demgegenüber ab. (2P)
- (2) Welche Voraussetzungen machen Regressionsanalyse und die traditionelle Conjoint-Analyse bezüglich des Datenniveaus der erhobenen Urteile? (2P)

4 VARIANZANALYSE UND VERLAUFSKURVEN

- (1) In einer Untersuchung wurde der Einfluss von Tageszeit (Morgen vs. Abend) und Alkoholkonsum (0.0 vs. 1.0 Promille) auf die Fahrleistung gemessen. Jede Versuchsperson fährt unter genau einer Bedingung. Die Fahrleistung einer VP setzt sich dabei aus zwei Komponenten zusammen: der Qualität der Längsregelung (Maß: minimaler Abstand zum Vorderfahrzeug) und der Querregelung (Maß: Anzahl Spurüberschreitungen). Der Versuch wird varianzanalytisch ausgewertet.
Handelt es sich um ein
a) ein- oder mehrfaktorielles,
b) uni- oder multivariates Design
c) mit unabhängiger oder abhängiger Messung?
d) In welche Bestandteile wird die Varianz der Messwerte zerlegt? (4P)
- (2) Welche Voraussetzungen macht die parametrische Analyse von Split-Plot-Designs, wenn jeder Faktor nur zwei Stufen hat? Welche Voraussetzung kommt hinzu, wenn jeder Faktor drei (oder mehr) Stufen hat? Wie kann man die ANOVA retten, wenn diese zusätzliche Voraussetzung verletzt ist? (3P)

5 PROBABILISTISCHE SKALIERUNG

- (3) Erläutern Sie die Ausgangslage und das schrittweise Vorgehen beim BTL-Modell in eigenen Worten. (4P)
- (4) N=50 Studenten verglichen fünf Duftstoffe paarweise auf der Dimension „Schärfe“. Der Duft „Tanne“ erhielt nach Anwendung des BTL einen Skalenwert von 1.2, der Duftstoff „Essig“ einen Skalenwert von 4. Wie viele Studenten haben bei Gültigkeit des BTL „Essig“ für schärfer als „Tanne“ befunden? (1P)

6 LOGISTISCHE REGRESSION

Folgende Tabelle zeigt das Ergebnis einer Logistischen Regression. In der zugrunde liegenden Untersuchung wurden das Alter [in Jahren], das Geschlecht [m vs. w] sowie schlafbezogene Atmungsstörungen (sog. Schlafapnoe) [ja vs. nein] als Risikofaktoren für einen Müdigkeitsunfall untersucht.

	Regressions- koeffizient B	Standard- fehler	p-Wert	Differenz für OR	OR
Alter	0.196	0.052	0.000	1 Jahr	1.217
Schlafapnoe	1.923	0.842	0.022	ja vs. nein	6.844
Geschlecht	1.496	0.719	0.029	männl. vs. weibl.	5.976
Konstante	-4.447	1.806	0.014		0.012

- (1) Interpretieren Sie das Ergebnis. (3P)
- (2) Nach der Durchführung einer Logistischen Regression sollte vom Untersucher die Güte der Modellschätzung beurteilt werden. Erläutern Sie die grundsätzliche Interpretation des
(a) Omnibus-Tests der Modellkoeffizienten,
(b) eines Pseudo-R²-Maßes und
(c) des Hosmer-Lemeshow-Tests. (3P)

7 METAANALYSE

- (1) Welche vier Hauptproblembereiche für die Bewertung von Metaanalysen werden häufig angeführt? (2P)
- (2) Im Zusammenhang mit dem sog. Publication Bias werden Funnel-Plots und das sog. Fail-Safe-N diskutiert. Beschreiben und erläutern Sie, was unter den Begriffen
 - (a) „Publication Bias“,
 - (b) „Funnel-Plot“ und
 - (c) „Fail-Safe-N“ verstanden wird? (3P)

8 LOGLINEARE MODELLE

Im Rahmen von "best practice"-Untersuchungen zur Behandlung von Depressionen stellt sich die Frage, wie der Therapieerfolg (nach 3 Monaten geheilt/nicht geheilt) mit der Art der Therapie (A/B/C), dem Geschlecht (m/w) und dem Bildungsgrad (ohne/Qualifizierender Hauptschulabschluss/Mittlere Reife/Abitur/Studium) zusammenhängt. Dazu analysieren Sie Datenbestände mehrerer Krankenkassen.

Sie vermuten (auf Basis der Literatur) folgende Zusammenhänge:

1. Patienten mit Studium wählen eher Therapie A
2. Weibliche Patienten sprechen besonders gut auf Therapie C an, haben also unter Therapie C höhere Heilungschancen

- (1) Notieren Sie zunächst das saturierte loglineare Modell für diese Daten. Bitte verwenden Sie dabei die Anfangsbuchstaben der Faktoren als Indizes. (1.5P)
- (2) Wie würde ein (hierarchisches) loglineares Modell aussehen, wenn nur die vermuteten Zusammenhänge und keine anderen gelten? Streichen Sie einfach die entsprechenden Terme aus dem saturierten Modell. (1.5P)

9 MULTIPLE REGRESSIONSANALYSE

Ziel der multiplen Regressionsanalyse ist die Identifikation eines Modells, das mit möglichst wenig Variablen die Daten möglichst gut beschreibt und inhaltlich interpretierbar ist.

Bei der Suche nach diesem Modell gibt es drei verschiedene Ansätze: (1) Das sog. Einschlussverfahren, bei dem alle Prädiktoren gleichzeitig ins Modell kommen, (2) die sog. sequentielle oder hierarchische Regression, bei der der Forscher festlegt, in welcher Reihenfolge die Prädiktoren eingebaut werden und (3) die statistische oder schrittweise Regression, bei der die Prädiktoren nach statistischen Kriterien aufgenommen und ggf. wieder ausgeschlossen werden.

- (1) Sie erhalten bei einem Datensatz mit dem Einschlussverfahren vs. mit dem sequentiellen Verfahren im Allgemeinen unterschiedliche Modellgleichungen. Erklären Sie, warum das so ist, indem Sie das Prinzip des sequentiellen Verfahrens mit dem Einschlussverfahren vergleichen (evtl. mit Venn-Diagrammen). (3P)
- (2) Was muss für die Prädiktoren gelten, damit Sie bei beiden Herangehensweisen dieselben Gleichungen erhalten? (1P)
- (3) Nennen Sie einen forschungspraktischen Grund, warum man eine sequentielle Regression durchführt. (1P)

10 QUALITATIVE ANALYSEN

- (1) Wie bewerten Sie die häufig geführte Diskussion zur Abgrenzung qualitativer und quantitativer Forschung? Begründen Sie Ihre Antwort kurz. (1P)
- (2) Was versteht man unter dem Begriff der "Methoden-Triangulation" im Zusammenhang mit qualitativer Forschung? (2P)

11 ALLGEMEIN

Im Zuge der Entwicklung eines Auswahlverfahrens für Studienanfänger schlägt jemand folgendes zweistufige Vorgehen vor:

1. Schritt: Wir nehmen die Abiturzeugnisse, fassen die Noten in den Sprachen (Deutsch, alle Fremdsprachen) per Mittelung zu einer Sprachnote zusammen und die naturwissenschaftlichen Fächer zu einer Naturwissenschaftsnote; das mathematisch/abstrakte Denken erfassen wir über die Mathematiknote; aus dem Rest machen wir eine Restnote.

2. Schritt: Dann zählen wir Sprachnote, Naturwissenschaftsnote und Mathematiknote zusammen und nehmen die Leute mit den höchsten Punktwerten; die Restnote wird ignoriert.

Wenn man diesen Vorschlag mit statistischen Standardverfahren vergleicht:

- (1) a) Welches statistische Verfahren entspricht Schritt 1?
b) Wie würde die Sprachnote eines einzelnen Bewerbers dann üblicherweise bezeichnet? (2P)
- (2) Welches statistische Verfahren würden Sie bei Schritt 2 wählen, wenn Ihnen ein quantitatives Bewährungskriterium (Punkte in Zwischenprüfung) zur Verfügung steht? Welche Komponenten dieses Verfahrens legt der Vorschlag in Schritt 2 implizit fest? (2P)
- (3) Welchen Vorteil hätte die Verwendung des statistischen Verfahrens für Schritt 1 für die weitere Auswertung mit dem statistischen Verfahren in Schritt 2? (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

ENDE