

FAKTORENANALYSE ALS DIMENSIONSANALYSE

Zielsetzung:

Unter dem Ökonomieprinzip betrachtet geht es um eine Dimensionsanalyse
Verschiedene manifest (beobachtbar) gewordene Variablen werden auf weniger (nicht beobachtbare) latente Variablen zurückgeführt.

Personen mit möglichst wenigen Charakteristika möglichst genau beschreiben

Alternative zur FA ist nicht wirklich gegeben, eventuell das Rasch-Modell
(Annahme: 1 Test misst nur bestimmte Dimensionen und das wird geprüft, also mit Versuchs- und Irrtumsverfahren mit dem Rasch-Modell vorgegeben)

FA führt auf wenig entscheidende Faktoren zurück (FA ist ein Modell, und zwar ein formales)

Modell = Annahme, die richtig oder falsch sein kann, also Gedankenspiel
allgemeine Variable = Funktion von lg, abhängig von verschiedenen Faktoren = latente Variablen (Dimensionen)

Modell $X_i = lg (F_1, F_2, F_3, \dots F_j, \dots F_n)$

Dieses Modell ist tautologisch.

Keine näheren Aussagen über $F_1, \dots, F_m$, etc.

unprüfbare Spezifikationsgleichung -> von lg (= der Funktion)
(ist nicht gleich dem Rasch-Modell); jede Variable x_i wird erklärt als gewichtete Summe voneinander unabhängigen Faktoren

$$x_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{ij}F_j \dots + a_{im}F_m + a_{is_j} [\text{spezifischer Anteil}] + e_i$$

$$a_{is_j} + e_i = \text{uniqueness}$$

FA = unüberprüfbares Modell, das behauptet, dass jede beobachtbare Variable als Summe latenter Variablen erklärbar ist. Modell besteht in Addition und Gewichtung.

spezifischer Faktor + Fehler = uniqueness (= das für eine bestimmte Variable Typische)

wesentlich: **Jede manifeste Variable ist erklärbar durch eine Reihe latenter unabhängiger Variablen.**

Behauptung: **Faktoren wirken additiv und voneinander unabhängig (d.h. es kommen keine Wechselwirkungen vor). Diese Behauptung ist unüberprüfbar**; Aufschaukelungsprozess wird geleugnet. Daher keine Eigenschaftsableitung, die nachher zu prüfen wäre.

FA = praktikables Modell der Psychologie, Medizin, Soziologie, aber unüberprüfbar!
Wenn es bezweifelt wird, braucht man die Ergebnisse ja nicht zu glauben... (ist NICHT so beim Rasch-Modell!)

Idee der Diagnostik:

$$x_e = a_{e1}F_1 + a_{e2}F_2 + \dots$$

$$x_j = \dots$$

⇒ man könnte jetzt Korrelationen rechnen und INDIREKT auf a_i bzw. m (= Zahl der Faktoren) schließen

vgl. T-Shirts.... Es gibt weniger Variationen als Kombinationen Länge + Breite. Bei starker Korrelation liegt meist 1 Länge und Breite zugrunde. Statt vieler Beschreibungen genügen wenige elementare.

Rechnung = aufwändig; Idee = simpel x_i bzw. x_j = Kriterium, das aus bestimmten Größen wahres Ergebnis (x_i) vorhersagen kann. Wenn man die entsprechenden Größen kennt $|(\hat{x}_i - x_i)| \rightarrow$ Differenz soll ein Minimum werden... Ist im Prinzip eine Extremwertaufgabe, d.h. 1. Ableitung Null setzen...

FA geht so vor, dass sie die beobachtbaren Korrelationen minus die geschätzten zu einem Minimum macht (also über die Abweichungen der Korrelationen)

→ **Eigenwertproblem** = Technik der Mathematik, um zu einer Lösung zu kommen = ein Minimum zu finden

Eigenwertproblem nur lösbar, wenn m (= Zahl der Faktoren) bekannt ist

→ Detailprobleme der Diagnostik

Für Praxis bedeutet das, durch Heuristiken, Probieren, etc. das nun bekannt zu machen (z.B. 4 oder 5 Faktoren)

Wie kommt man auf die Zahl der Faktoren?

→ subjektiv festgelegt; nach eigenen Richtlinien; Eigenwertdiagramm wesentlich:

als Voraussetzung:

Variablen und Faktoren sind standardisiert $\bar{x} = 0, s = 1$

$$\sigma^2(x_i) = 1$$

2 Faktoren, uniqueness = $0 = u_i$

Varianz einer Summe = $u * u$, wenn sie unabhängig sind

wesentlich bei Varianz
 a_i = Konstante
Variable = $F_1, F_2, \dots$

$$\sigma^2(x_i) = 1 = \sigma^2(F_1) [= 1] + \sigma^2(F_2) [= 1] \quad \sigma^2(F_2) = \text{multiplikative Konstante}$$

additive wirkt zum Quadrat
 $a_{i1}^2 + a_{i2}^2$

$\sigma^2(x_e) = 1 = a_{e1}^2 + \dots + a_{ek}^2$ = die Summe der quadrierten Ladungen (Gewichte), der Eigenwert, d.h. die Summe der quadrierten Ladungen wird Eigenwert genannt

a_i^2 und $a_{e1}^2 F_1$ ist ein Teil der Varianz, und zwar des Faktors F_1 an der Variablen i

λ = die insgesamt Varianz, also des Faktors F an allen Variablen (Varianz, die erklärt wird), d.h. die der eine Faktor an allen Variablen beträgt.

λ = **Eigenwert**

sollte groß sein, maximale Anzahl der Variable (k); hier 2 Variablen; sollte maximal k betragen (hier daher $2k$)

Eigenwert sollte mindestens so viel wie eine Variable erklären.

Von Anzahl der Faktoren m wählt man Zahl der Eigenwerte, die größer als 1 sind, d.h. also so viele Faktoren extrahieren, wie Eigenwerte größer als 1 sind.
Varianz = Variabilität (= Unterschiedlichkeit) kann erklärt werden.

Eigenwertdiagramm -> Scree-Test

Idee: man muss wissen, welches m man wählt und bestimmt die Eigenwerte. Es gibt so viele Eigenwerte wie m , sie werden der Größe nach geordnet vom p -ten Eigenwert zum $p-1$ ten Eigenwert

p kleiner als $p-1$

wenn Unterschied klein ist, etwa gleich viel erklärt; $EW > 1$

Faustregel: Eigenwert EW größer als 1; die Zahl dieser suchen

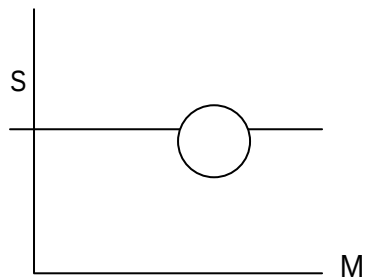
Scree-Test sucht deutlichen Knick (beinhaltet Willkür in der Beurteilung)
d.h. **Scree-Test ist manchmal deutlich bei der Festlegung der Eigenwerte**

Bei 16 Dimensionen Abbruch (nicht nachvollziehbar), d.h. wir bekommen es mit Manualen und Verfahren zu tun, von denen alles Mögliche behauptet wird (zum Vergleich 16PF-R Scree-Plot -> er hat nur Korrelationsmatrix)

FA = Modell, nicht prüfbar -> demnach ist **16PF ein A-priori-Fragebogen**
16 unabhängig gewonnene Eigenwerte, werden dann noch revidiert

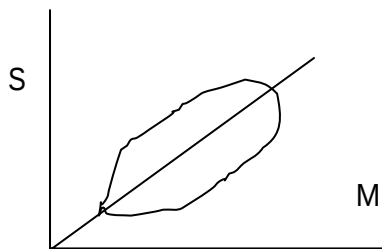
Stichprobenabhängigkeit der FA (ist nicht so beim Rasch-Modell)

motorische Fähigkeiten und Sprachkompetenz bei 3-Jährigen = entwicklungspsychologisch unabhängige Faktoren



weil kein Zusammenhang zwischen S und M
-> Punktschwarm; Korrelation = 0, weil Regressions-
gerade parallel zur x-Achse

Bei 3-Jährigen und 4-Jährigen

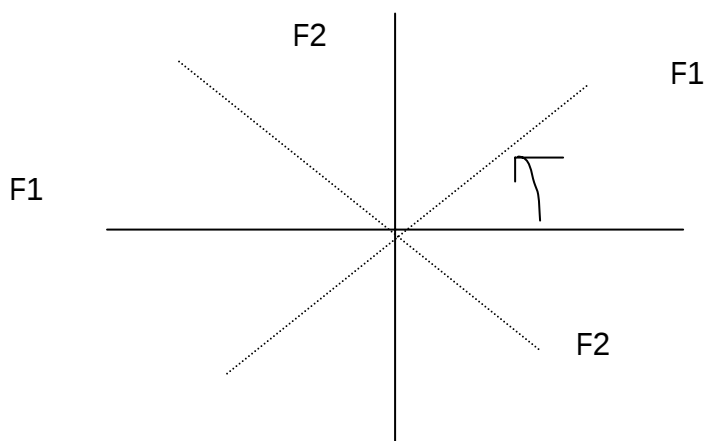


hier gibt es eine Korrelation -> unabhängige
Faktoren

Unabhängigkeit der Faktoren:

- 1) Stichprobenabhängigkeit
- 2) Rotationsproblem

1. EW = größer; 2. EW = kleiner (2-Faktorenlösung);
- 2 Faktoren korrelieren nicht = unabhängig, d.h. im Plus auftragbar



F1 erklärt viel; Ladungen des F2 sind relativ belanglos -> Rotation -> nachher kann
man F2 als gleichwertig interpretieren

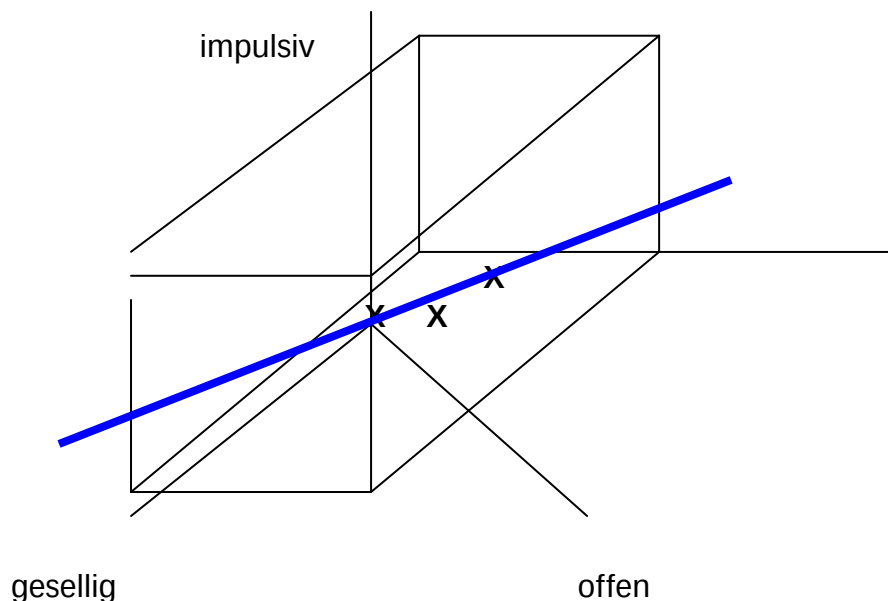
ideal = simple structure:

Einfachstruktur wird angestrebt (in jedem Faktor soll möglichst nur 1 hohe und sonst niedrige Ladungen auftreten) -> 1 Marker-Variable ist typisch für 1 Faktor

gerechnet mit Kaisers Varimax-Methode -> VARIMAX (Ladungen pro Faktor sollen Maximum sein)

gelingt die simple-structure-Lösung -> nächstes Problem

3. „Faktortaufen“ impulsiv, offen, gesellig



3-dimensionale Diagonale ergibt sich -> die 3 verschiedenen Punkte / Personen liegen auf dieser Geraden / Kontinuum.

Die 3 Eigenwerte sind abhängig und durch eine Dimension erklärbar, z.B. alle anderen taufen -> extravertiert. Dieses „Taufen“ ist die reinste Willkür (dort wo transparent ist in Wissenschaft manchmal nötig)

Textautor meint mit extravertiert bestimmte Begrifflichkeit, nämlich „offen“

Vorher orthogonale (rechtwinkelige) Rotation durchgeführt, um simple-structure-Prinzip zu realisieren; wenn **schiefwinkelig rotiert -> führt zu besserer Lösung**. Frage: ist das noch systemimmanent? die beiden Faktoren korrelieren ja (ist über den Winkel der Regressionsgeraden zu errechnen) -> damit Verletzung der Voraussetzung, passt nicht zu Ausgangsüberlegung!

Aufgabe:

2 Faktoren der Teilleistungsstörung

erster Faktor misst akustische Speicherschwäche

zweiter Faktor sprachliche Artikulation

Kind ist akustisch $T = 25$ und gleichzeitig $T = 3$

Beide Fähigkeiten korrelieren mit 0,6 (= korrelierende Faktoren -> sind nicht zu interpretieren)

Nicht Förderung in beiden Faktoren, sondern eine Maßnahme genügt, weil Faktoren abhängig sind!

Schiefwinkelige haben Risiko, dass sie eine Mehrfachindikation andeuten, die nicht gegeben ist und Maßnahmen nahe legen, die so nicht nötig sind!

Besser nochmals F^* und Faktoren 2. Ordnung finden (aus den Faktoren erster Ordnung). Scree-Test hätte auch genügt (erste Faktoren); mangelnde Methodenkenntnis; Testautoren geben für alle Faktoren Interpretationsmöglichkeit

Fazit: Das Ideal besser realisieren oder man lernt bei Fallanwendung die schiefwinkelige Rotation zu interpretieren

auf schiefwinkelige 16 Faktoren wurde nochmals FA angewendet -> Ergebnis = 5 Faktoren

Problem: wenn wir Fragebogen haben, der die Items identifiziert, zu welchen Faktoren gehören sie?

z.B. Items aufsummieren; nicht jede Person lädt in jedem Faktor gleich hoch
Einzelne Items laden ja nicht gleich in jedem Faktor; es schaut nur so aus, als würden sie gleich hoch laden. Es müsste über Faktorscore für jede Person errechnet werden.

Nicht jede Variable lädt in jedem Faktor gleich hoch. Variablen werden Faktoren zugeordnet, sind aber nicht ident. **Antworten entsprechen nicht dem Testergebnis, sind nicht fair.**

Verrechnungsschema wird unterlegt, das nicht zu methodischem Ergebnis passt. [bezieht sich alles auf 16-PF]

Problem: Wir erfassen 2 **Items mit dichotomer Antwortmöglichkeit** (FA nur für intervallskalierte Daten gedacht)

Produkt Hauptdiagonale minus Produkt Nebendiagonale

Beispiel: 2 Items, die Extraversion messen: $r = 0,5$; idealer Zusammenhang, trotzdem ist $r = 0,5$

		1		
		+	-	
2	+			50
	-			50
		20	80	

Vierfelderkorrelation ist Korrelation von Randsummen.

Korrelationen zwischen Items hängen wesentlich von der Schwierigkeit (Attraktivität, ja / nein zu sagen) ab.

Man bekommt nur hohe Korrelationen, wenn Items gleiche Schwierigkeit haben (gleiche Ranghäufigkeit), unabhängig vom Inhalt.

FA auf dichotome Items angewendet führt regelmäßig zu Schwierigkeitsfaktoren und als solchen interpretierbaren Ergebnissen. FA auf dichotome Items angewendet ergibt artifizielle Faktoren!