

Seminar: Pfadanalyse mit latenten und manifesten Variablen
 Dozent: Tilo Görl
 Referent: Thomas Breitner
 Datum: 08/06/01

Meßmodelle – Konfirmatorische Faktorenanalyse

Allgemeines zur Faktorenanalyse:

Ziel der Faktorenanalyse ist es, Variablen (Merkmale) gemäß ihrer korrelativen Beziehungen in voneinander unabhängige Gruppen zu klassifizieren.

Folgende Steckbrief-Eigenschaften kann man der Faktorenanalyse zuschreiben

- Die Faktorenanalyse ist ein „datenreduzierendes“ Verfahren, da eine Vielzahl von Variablen auf eine geringere Zahl von Faktoren abgebildet wird.
- Die Faktorenanalyse ist ein heuristisches, hypothesengenerierendes Verfahren, da sowohl die Faktorenanzahl (z.T.) und deren Interpretation in der Hand des Untersuchers liegt.
 - vgl. die konfirmatorische Faktorenanalyse, die ein hypothesentestendes Verfahren darstellt.

Randbedingungen der PCA

- PCA-Faktoren sind unkorreliert.
- PCA-Faktoren klären sukzessiv maximale Varianz in den Variablen $X_1 \dots X_m$ auf.

Durchführung einer Faktorenanalyse:

- (1) Variablenauswahl und Errechnung der Korrelationsmatrix
- (2) Extraktion der Faktoren
- (3) Bestimmung der Kommunalitäten
- (4) Zahl der Faktoren
- (5) Faktorinterpretation
- (6) Bestimmung der Faktorwerte

zu (1):

Variablenauswahl und Berechnung der Korrelationsmatrix

- um zu verlässlichen Aussagen zu kommen, ist eine homogene Stichprobe wichtig
- systematische Fehler bei der Erhebung, z.B. durch Ja-Sager-Mentalität oder vermeintlich sozial erwünschter Antworten, sind zu vermeiden

Korrelation der Indikatoren:

$$r_{x_1, x_2} = \frac{\sum_{k=1}^K (x_{k1} - \bar{x}_1) \cdot (x_{k2} - \bar{x}_2)}{\sqrt{\sum_{k=1}^K (x_{k1} - \bar{x}_1)^2 \cdot \sum_{k=1}^K (x_{k2} - \bar{x}_2)^2}}$$

Berechnung der Korrelationskoeffizienten über alle Zeilen/Spalten ergibt die Korrelationsmatrix.
eventl. Standardisierung der Indikator-Variablen :

$$z_{kj} = \frac{x_{kj} - \bar{x}_j}{s_j}$$

Vorteile: Vergleichbarkeit der Indikatoren auch über verschiedene Maßeinheiten hinweg

Somit vereinfacht sich auch die Berechnung der Korrelationen, da die Kovarianz zwischen zwei z-standardisierten Variablen gleich der Produkt-Moment-Korrelation der unstandardisierten Variablen ist (die Standardabweichungen im Nenner sind jeweils 1):

$$r_{x_1, x_2} = \frac{s_{x_1, x_2}}{s_{x_1} s_{x_2}} \quad \text{mit: } s_{x_1, x_2} = \frac{1}{K-1} \sum_k (x_{k1} - \bar{x}_1)(x_{k2} - \bar{x}_2)$$

Prüfung der Korrelationsmatrix auf Verwendbarkeit für die Faktorenanalyse

Hier geht es um die Frage, ob die empirische Verteilung der Stichprobe auch in gewissen Grenzen mit der angenommenen Verteilung in der Grundgesamtheit übereinstimmt. Ob sich also Erkenntnisse aus der Stichprobe auf die Grundgesamtheit generalisieren lassen können. Dazu existieren eine Reihe von Tests:

- Signifikanzniveau der Korrelationen

Jede Korrelation wird darauf überprüft, mit welcher Wahrscheinlichkeit sie auch zufällig zustande hätte kommen können. Der entsprechende Test ist für die Produkt-Moment-Korrelation wie für die Punktbiseriale Korrelation (ein dichotomes, ein intervallskaliertes Merkmal) der t-Test. Es wird ein empirischer t-Wert für jede Korrelation berechnet und mit einem kritischen t-Wert verglichen. Unter den Bedingungen der Nullhypothese ist die Korrelation zwischen zwei

Variablen x und y in der Grundgesamtheit Null. Ist nun der ermittelte t -Wert größer als der kritische t -Wert, so kann die Nullhypothese mit der entsprechend abgetesteten Wahrscheinlichkeit (meist 1% oder 5%) verworfen werden und die H_1 angenommen werden. Für die Freiheitsgrade gilt: $df = n - 2$. Auf diese Weise ist zu überprüfen, ob es zwischen den Indikatoren überhaupt Unterschiede gibt.

- weitere Tests:
 - Inverse der Korrelationsmatrix
 - Bartlett-Test (test of sphericity)
 - Anti-Image-Kovarianz-Matrix
 - Kaiser-Meyer-Olkin-Kriterium

zu (2):

Extraktion der Faktoren

Jeder Indikator als eine Linearkombination von Faktoren beschreiben:

a = Gewichte = Faktorladungen

p = Faktoren

$$z_{kj} = a_{j1} \cdot p_{k1} + a_{j2} \cdot p_{k2} + \dots + a_{jq} \cdot p_{kq} = \sum_{q=1}^Q a_{jq} \cdot p_{kq}$$

Für diese Gleichung lassen sich im Grunde unendlich viele Lösungen finden. Eine Lösung führt zu den Faktoren der PCA; mit folgenden Eigenschaften:

- sie sind voneinander unabhängig
- sie erklären sukzessiv („nach und nach“) maximale Varianz
je mehr Faktoren, je höher die erklärte Varianz im Modell

Graphische Interpretation von Faktoren:

Aus der Korrelationsmatrix läßt sich ein Vektordiagramm erstellen. Und zwar entspricht der inverse Cosinus des Korrelationskoeffizienten einem Winkel: $r_{xy} = 0,89 \Rightarrow 27,1^\circ$

Die Zahl der benötigten Achsen entspricht der Zahl der Faktoren.

Unkorrelierte Faktoren stehen rechtwinklig zueinander ($\text{inv cos } 90^\circ \Rightarrow r = 0$)

Wie findet man nun die Lage der Faktoren zu den Variablen?

Angenommen man habe nur zwei Variablen, so ließen sich diese annäherungsweise auch durch einen dritten, „winkelhalbierenden“ Vektor zwischen Variable 1 und Variable 2 beschreiben. Dies entspräche dann einem Faktor. \Rightarrow Winkel wieder als Korrelation zu interpretieren

Für unser menschliches Vorstellungsvermögen gibt's hier noch das Problem der Multidimensionalität (mehr als drei Achsen), das entsteht, wenn mehrere Indikator-Vektoren im korrekten Winkelverhältnis zueinander gebracht werden müssen.

zu (3):

Bestimmung der Kommunalitäten

Hier geht man davon aus, daß nicht die gesamte Varianz der Indikatoren durch die Faktoren aufgeklärt werden kann. Die Gesamtvarianz einer Variablen setzt sich also aus der durch den ersten Faktor erklärten Varianz, der durch den zweiten Faktor erklärten Varianz usw. und einer Restvarianz, die durch die Faktoren nicht erklärt wird, zusammen.

Problem: Schätzung der Kommunalitäten \Rightarrow Erfordert Sachkenntnis des Forschers

Je höher die Anzahl der erhobenen Variablen, desto unwichtiger werden die geschätzten Kommunalitäten.

Kommunalität $h_j^2 = 0,8$ [$a_{j1}^2 = 0,7 + a_{j2}^2 = 0,1$] $s_{j\text{Rest}}^2 = 0,2$ (a = Faktorladungen)

zu (4):

Zahl der zu extrahierenden Faktoren

Die Anzahl der Faktoren hängt vom subjektiven Empfinden des Untersuchers ab. Allerdings gibt es eine Reihe von statistischen Kriterien:

- Kaiser-Kriterium:

„Die Zahl der zu extrahierenden Faktoren ist gleich der Zahl der Faktoren mit Eigenwerten größer eins.“

Die Eigenwerte sind die Summe der quadrierten Faktorladungen eines Faktors über alle Variablen. Die Eigenwerte sind ein Indikator dafür, wie viel Varianz in den Ausgangswerten durch einen Faktor erklärt werden kann. Wenn ein Eigenwert eines Faktors kleiner als 1 ist, so erklärt er weniger Varianz als eine einzelne Variable, die, da sie standardisiert ist, gerade schon eine Varianz von 1 hat.

- Scree-Test

Rangordnung der Eigenwerte der Faktoren und Suche nach dem „Knick“ im Graphen

zu (5):

Faktorinterpretation

Zur Faktorinterpretation behilft man sich mit den Faktorladungen. Die Interpretation soll anhand

eines Beispiels aus dem Backhaus verdeutlicht werden:

Es wurden fünf Marken von Aufstrichfetten untersucht. Die Vpn sollten jeweils Eischätzungen bezüglich

1. Anteil ungesättigter Fettsäuren
2. Kaloriengehalt
3. Vitamingehalt
4. Haltbarkeit
5. Preis

machen. Anhand der Faktorladungsmatrix lassen sich Bündel von Variablen bestimmen, die durch den jeweiligen, hoch ladenden Faktor beschrieben werden.

So ergaben sich z.B. die Faktoren „Gesundheit“ für die ersten drei Eigenschaften und „Wirtschaftlichkeit“ für die letzten zwei Eigenschaften.

Hier entsteht auch die Frage, ab welcher Ladung man eine Variable überhaupt einem Faktor zuordnet. Eine Konvention hilft: ab 0,5 geht man von hohen Ladungen aus. Lädt eine Variable auf verschiedenen Faktoren hoch, dann muß sie auch bei jeder Interpretation der betroffenen Faktoren berücksichtigt werden.

Eine Rotation der Achsen dient der besseren Interpretierbarkeit. Die Aussagekraft wird dabei nicht verändert.

zu (6):

Bestimmung der Faktorwerte

Die Faktorenwerte geben an, wie stark die in einem Faktor zusammengefaßten Merkmale in einer Variable ausgeprägt sind. Zur Berechnung siehe Bortz oder Backhaus.

- negativer Faktorwert = ein Objekt ist in Bezug auf diesen Faktor im Vgl. zu allen anderen Objekten unterdurchschnittlich ausgeprägt.
- Faktorwert = 0: ein Objekt ist in Bezug auf diesen Faktor durchschnittlich ausgeprägt
- positiver Faktorwert: ein Produkt ist in Bezug auf diesen Faktor im Vgl. zu allen anderen Objekten überdurchschnittlich ausgeprägt

Zusammenfassung der Kennziffern der Faktorenanalyse:

- Faktorladung
Korrelation zwischen Faktor und Indikator
- Faktorwert
s.o.
- Kommunalität

Wieviel Varianz eines Indikators wird durch die Faktoren erklärt?

- Eigenwert (eines Faktors)

Wieviel von der Gesamtvarianz aller Variablen können durch diesen Faktor erklärt?

Konfirmatorische Faktorenanalyse

Zusammenfassung:

Wir nehmen an, die Indikatoren seien in der GG normalverteilt. Unbekannt sind uns aber die Mittelwerte, Varianzen und Kovarianzen der Indikatoren. Im Maximum-Likelihood-Ansatz der FA werden in der Population gültige Varianzparameter der Indikatoren gesucht, die die Wahrscheinlichkeit des Zustandekommens der empirisch gefundenen Korrelationsmatrix maximieren. Mit der konf. FA können Hypothesen über die Faktorenstruktur eines Datensatzes getestet werden. Mit Anpassungstests wird dann überprüft, ob die Abweichungen der empirischen von der unterstellten Korrelationsmatrix zufällig oder statistisch bedeutsam ist.

Die konfirmatorische Faktorenanalyse geht hypothesengeleitet vor.

Nach Kline stehen wir vor vier Aufgaben:

1. unser theoretisches Konstrukt muß operationalisiert werden
2. Indikatoren müssen gefunden werden für dieses Konstrukt (diese Faktoren)
3. die Indikatoren, die für ein bestimmtes Konstrukt dienen sollen, sollen auch so ungefähr das selbe messen – sie sollten miteinander korrelieren
4. wenn die Daten nicht mit den aus ihnen gewonnen Vorhersagen übereinstimmen, ist das Modell in Frage zu stellen / zu verwerfen.

Principles of Measurement

zur Meßtheorie: die Indikatoren sollen frei von Meßfehlern sein: reliable and valid:

- Die Reliabilität betrifft die Genauigkeit oder Präzision einer Messung.
- Die Validität einer Messung hängt davon ab, ob wirklich das gemessen wird, was eigentlich inhaltlich gemessen werden soll.
- Objektivität ist das dritte, von Kline nicht so ausgetretene Kriterium, das angibt, in welchem Ausmaß eine Messung unabhängig vom Untersucher ist.

Für Kline gibt es vier grundlegende Arten von Reliabilität:

1. Test-retest reliability
Wiederholbarkeit / Reproduzierbarkeit der Ergebnisse (meist zeitlich gesehen)

Soll ein Test immer das selbe messen, so müssen ihre Ergebnisse miteinander korrelieren. Der Intervall ist vom Untersuchungsgegenstand abhängig.

4. Alternate forms reliability

Reproduktion der Ergebnisse über verschiedenen Methoden

6. Interrater reliability

Bewertung der Ergebnisse durch unabhängige Experteninstanzen. Vor allem sinnvoll bei sehr subjektiven Erhebungen. Die Werte der Vpn werden verschiedenen Experten vorgelegt, die diese auf ihre Reliabilität bewerten. Weichen die Bewertungen sehr voneinander ab, sollte das Meßinstrument geändert werden.

Kline führt einen Reliabilitätskoeffizienten (r_{xx}) ein, der von 0 (überhaupt nicht verlässlich) bis 1 (total verlässlich) reicht.

Kline unterscheidet ebenfalls verschiedene Arten von Validität:

Messinstrumente können sehr exakt immer das Falsche messen; dann sind sie zwar reliabel, aber nicht valide.

1. content validity

Die Inhaltsvalidität bedeutet, daß die Gültigkeit der Messung mehr oder weniger für jedermann einsichtig aus den einzelnen Teilen des Meßinstruments hervorgeht.

Letztlich beruht sie auf der Kenntnis von 'Experten' über den betreffenden Gegenstand (wobei u.U. sehr viele oder alle Leute Experten sein können). So wird z.B. bei jeder Prüfung I. unterstellt: Eine Mathematikprobe sagt etwas über die 'Mathematik-Fähigkeit', weil ein Mathematiklehrer oder -professor in der Lage sein sollte zu beurteilen, was gute und was schlechte Mathematik-Fähigkeiten sind. Die Behauptung, ein Meßinstrument habe I., bedeutet in der Forschungspraxis aber oft nichts anderes, als daß der Entwickler des Instruments selbst glaubt, das Instrument sei valide. Ehrliche Personen gebrauchen hier den Begriff der 'face validity', d.h. der 'augenscheinlichen Validität'.

Dennoch ist die Idee der I. sehr wichtig: Es geht letztlich eben darum, daß eine Messung das relevante Phänomen möglichst in allen Aspekten erfaßt, und dies kann nur durch Forschen, Nachdenken und Kommunikation zwischen Wissenschaftlern herausgefunden werden und nicht durch bestimmte, immer funktionierende 'Techniken'.

5. criterion-related validity und concurrent validity

Bei der Kriteriumsvalidität geht es um die Übereinstimmung eines Meßinstruments mit anderen relevanten Merkmalen (sog. Außenkriterien). Genauer unterscheidet man hier zwischen der Übereinstimmungsvalidität (engl: concurrent validity) (das Außenkriterium wird gleichzeitig erhoben) und der Vorhersagevalidität (engl.: predictive validity), bei der das Außenkriterium

erst später gemessen wird.

Übereinstimmungsvalidität wird z.B. erhoben, wenn die Messung mit dem gleichen Merkmal in Beziehung gesetzt wird, wie es durch eine andere Messung ermittelt wurde (Beispiel: Ein kürzerer Intelligenztest wird mit einem längeren verglichen)

8. construct validity

Konstruktvalidität ist ein komplexes Vorgehen, bei dem man eine Reihe von plausiblen oder sogar bestätigten Hypothesen prüft, die sich u.a. auf das Konstrukt beziehen, dessen Validität geprüft werden soll.

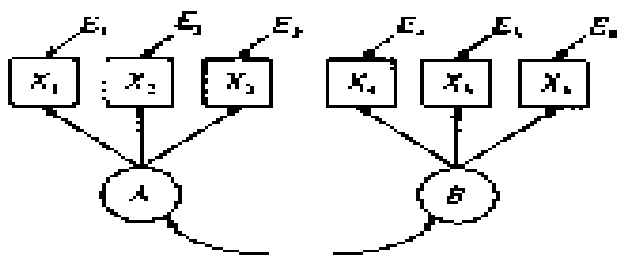
Wenn sich diese Hypothesen auch jetzt bestätigen, so ist anzunehmen, daß das fragliche Meßinstrument auch gültig ist. Eine Nicht-Bestätigung der Hypothesen kann allerdings auch bedeuten, daß die angeblich plausiblen oder bestätigten Hypothesen eben doch falsch waren, oder daß die anderen Variablen mit nicht validen Instrumenten gemessen wurden.

Kline führt einen Validitätskoeffizienten (r_{xy}) ein.

Zusammenfassend läßt sich wiederholen, daß es wichtig ist, verschiedene Indikatoren für ein Konstrukt zu finden, um z.B. folgendes Szenario zu vermeiden: Ein Forscher will etwas über das Konstrukt „Intelligenz“ erheben. Dazu benutzt er drei verschiedene Arten von Fragebögen. Die Auswertung ergibt, daß die gefundenen Werte hoch miteinander korrelieren. Dies kann allerdings auch an der Gleichartigkeit des Erhebungsinstrumentes und der somit herausgeforderten Gleichartigkeit der Antworten liegen.

CFA Models and Their Specification

Hier mal ein Standard CFA Modell; wobei X die Indikatoren darstellen, E die Störterme (der Varianzanteil der Indikatoren, der nicht durch die Konstrukte erklärt wird) und A bzw. B die latenten Variablen.



Solche Standard-CFA-Modelle können durch folgende drei Aussagen beschrieben werden:

1. Jeder Indikator hat zwei Ursachen: einen Faktor und einen Meßfehlerterm, in dem alle anderen varianzgenerierenden Quellen vereinigt sind
2. Die Fehlerterme sind voneinander und von den Faktoren unabhängig
Würden z.B. E1 und E2 korrelieren, so bedeutet dies, daß sie etwas gemeinsames messen, was im Modell nicht vorhanden ist.
4. Die Verbindungen zwischen den Faktoren bleiben unanalysiert

Die Pfeile, die von den Faktoren auf die Indikatoren zeigen repräsentieren den vermuteten direkten Effekt der latenten Variablen auf die beobachteten Werte. Die statistisch geschätzten Werte dieser Effekte heißen Faktorladungen.

Das Zustandekommen des Meßfehlerterms kann auf zwei Ursachen zurückzuführen sein:

1. zufällige Fehler
2. systematische Fehler (siehe letzte Woche)

Die Annahmen 2 und 3 des genannten Standardmodells werden auch als *unidimensionale Modelle* beschrieben. Sollten die Fehlerterme nicht voneinander unabhängig sein oder wenn ein Indikator auf mehr als einem Faktor lädt, spricht man von *multidimensionalen Modellen*. Angenommen, die Fehlerterme zweier Indikatoren korrelieren, dann heißt das, daß sie etwas messen, was im Modell nicht repräsentiert ist.

Neben uni- und multidimensionalen Modellen gibt es auch multihierarchische Modelle. Hier wären z.B. für ein Intelligenzmodell auf unterster Ebene die Indikatoren, dann einige Faktoren von intelligenten Fähigkeiten wie Sprache, räumliches Denken etc, und auf oberster Ebene könnte man einen allgemeinen Intelligenzfaktor ansiedeln.