EC330
Tratamento da Informação

Visão além do alcance: uma introdução à análise fatorial

Prof. Dr. Maurício U. Kleinke

PECIM – PPG em Ensino de Ciências e Matemática

Grupo de Ensino e Avaliação

DFA/IFGW/Unicamp

Visão além do alcance: uma introdução à análise fatorial

Dalson Brito Figueiredo Filho

José Alexandre da Silva Júnior

IFGW
Instituto de Física Gleb Wataghin

Start

- *"What is important to study cannot be measured and what can be measured is not important to study" (CONVERSE, 1964)*

- *"Concepts without percepts are empty (...) percepts without concepts are blind" (KANT)*

- “ainda que o desenvolvimento da teoria seja importante em si mesmo, eu acredito que os mais sérios e importantes problemas que requerem nossa imediata e forte atenção são aqueles de conceitualização e mensuração” (BLALOCK, 1979: 882).

Proxi

- Uma das possibilidades para tentar capturar esse conceito é eleger um *proxy*. Por exemplo, alguns estudos escolhem uma linha de pobreza específica como indicador de vulnerabilidade.

Proxi

- Operacionalmente, o pesquisador pode escolher o *proxy* como representativo do seu conceito/variável de interesse, analisar como ele se relaciona com outras variáveis e utilizá-lo como variável dependente ou independente em estágios mais avançados de seu desenho de pesquisa.

Variáveis subjacentes

- Um processo alternativo de mensuração é identificar variáveis que “caminham juntas”, ou seja, variáveis que apresentam a mesma estrutura subjacente (TABACHINICK e FIDELL, 2007).
- Tecnicamente, isso pode ser implementado através da análise fatorial. A principal função das diferentes técnicas de análise fatorial é reduzir uma grande quantidade de variáveis observadas a um número reduzido de fatores.

Dimensões latentes e construtos

- Os fatores representam as dimensões latentes (construtos) que resumem ou explicam o conjunto de variáveis observadas (HAIR et al, 2005).
- Ao resumir dados, a análise fatorial obtém dimensões latentes que descrevem os dados em um número menor de conceitos do que as variáveis individuais originais (HAIR et al, 2006, p. 91).

Variáveis não observáveis

Duas formas de mensurar variáveis não observáveis:

- escolher um *proxy;*
- reduzir um conjunto de variáveis a poucas dimensões através de análise fatorial.

Por que os pesquisadores de CH não utilizam análise fatorial?

(1) A resistência dos cientistas sociais brasileiros aos métodos quantitativos;

(2) o grau de complexidade matemática envolvido na operacionalização das diferentes técnicas de análise fatorial;

- Os benefícios associados à sua utilização (AF) ainda não são evidentes do ponto de vista conceitual.

Breve revisão da literatura

- “A análise fatorial é uma técnica amplamente estatística, mas seus fundamentos teóricos são um tanto obscuros e sujeitos a disputa” (BARTHOLOMEW, 1984, p. 221).
- “No modelo da análise fatorial, há muitas variáveis observadas cujo objetivo é gerar fatores subjacentes não observados” (KING, 2001, p. 682).

Breve revisão da literatura

- “a análise fatorial supõe que as variáveis observadas (medidas) são combinações lineares de algumas variáveis-fonte subjacentes (fatores). Isto é, ela supõe a existência de um sistema de fatores subjacentes e um sistema de variáveis observadas”
- (KIM e MUELLER, 1978, p. 7-8)3.

Análise fatorial: exploratória e confirmatória

- Análise Fatorial Exploratória:
  - Busca de padrões em um sistema, não tem um construto a ser analisado/avaliado.

- Análise Fatorial Confirmatória:
  - Existe um construto a ser conferido, a ser confirmada ou negada a presença no grupo que está sendo analisado.

A análise fatorial exploratória (AFE)

- A análise fatorial exploratória (AFE) geralmente é utilizada nos estágios mais embrionários da pesquisa, no sentido de literalmente explorar os dados.

- Além disso, a AFE pode ser utilizada para criar variáveis independentes ou dependentes que podem ser utilizadas posteriormente em modelos de regressão.

Análise fatorial confirmatória

- A análise fatorial confirmatória (AFC) é utilizada para testar hipóteses.

- Nesse caso, o pesquisador guiado por alguma teoria testa em que medida determinadas variáveis são representativas de um conceito/dimensão.

Planejamento da análise fatorial em três estágios

Procedimento	O que deve ser observado
Verificar a adequabilidade da base de dados	Nível de mensuração das variáveis, tamanho da amostra, razão entre o número de casos e a quantidade de variáveis e o padrão de correlação entre as variáveis.
Determinar a técnica de extração e o número de fatores a serem extraídos	O tipo de extração (principal components, principal factors, image factoring; maximum likelihood factoring; alpha factoring; unweighted least squares; generalized least squares).
Decidir o tipo de rotação dos fatores	Se for ortogonal (Varimax, Quartimax, Equamax), se for oblíqua (direct oblimin, Promax).

Tamanho da amostra

- Em relação ao número de casos, quanto maior, melhor.
- Hair et al (2006) sugerem que a amostra deve ser superior a 50 observações, sendo aconselhável no mínimo 100 casos para assegurar resultados mais robustos (HAIR et al, 2006).
- A razão entre o número de casos e a quantidade de variáveis deve exceder cinco para um ou mais (HAIR et al, 2006).

Síntese do primeiro estágio do planejamento da AF

Nível de mensuração	Variáveis contínuas e discretas
Amostra	Amostras mínimas entre 50 e 100 casos; razão entre o número de observações e a quantidade de variáveis igual ou superior a cinco.
Correlação	Maior parte dos coeficientes de correlação deve apresentar valores acima de 0,30.
KMO	Quanto maior melhor, tendo 0,50 como o patamar mínimo de adequabilidade (Hair et al 2006)
BTS	$p<0,05$

Análise de componentes principais (ACP) *versus* análise fatorial (AF)

"Se você estiver interessado numa solução teórica não contaminada por variabilidade de erro, a análise fatorial deve ser sua escolha.

Se você quiser simplesmente um resumo empírico do conjunto de dados, a análise de componentes principais é uma escolha melhor"

- (TABACHINICK e FIDELL, 2007, p. 608).

Análise de componentes principais (ACP) *versus* análise fatorial (AF)

"a análise de componentes principais é em geral preferida para fins de redução de dados (traduzindo o espaço das variáveis num espaço ótimo de fatores), enquanto a análise fatorial é em geral preferida quando o objetivo da pesquisa é detectar a estrutura dos dados ou a modelagem causal".

Síntese do segundo estágio do planejamento da AF

Tipo de extração	Determinar a técnica de extração dos fatores (principal components, principal factors, image factoring; maximum likelihood factoring; alpha factoring; unweight least square; generalized least square)
Regra de Kaiser	Devem ser extraídos apenas os fatores com valor do eigenvalue acima de 1
Scree test	Analisar graficamente a dispersão do número de fatores até que curva da variância individual de cada fator se tornar horizontal ou sofrer uma queda abrupta
Variância acumulada	> 60%
Razão teórica	O pesquisador deve justificar teoricamente como as variáveis se relacionam com os fatores extraídos.

Conclusões

- Esse artigo apresentou a lógica da análise fatorial a partir de uma abordagem intuitiva. Isso porque grande parte da literatura apresenta um grau proibitivo de complexidade, o que dificulta a compreensão. Em termos substantivos, acreditamos que problemas de mensuração constituem um dos principais obstáculos ao desenvolvimento do conhecimento.

Conclusões

- De forma direta, acreditamos que variáveis mal medidas comprometem a validade das inferências. Quanto mais erros em nossas medidas, tanto piores serão nossas conclusões a respeito dos fenômenos de interesse.