A very short introduction to Mokken Scaling

Edps/Soc 584 & Psych 594

Carolyn J. Anderson

Department of Educational Psychology

ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Spring 2017

Overview

- First proposed by Mokken (1971)
- Non-parametric scaling procedure for dichotomous & polytomous items.
- Many parametric IRT models are special cases of Mokken.
- Mainly used for scaling test and questionnaire data.
- For our purposes, we'll use it for item analysis.

References:

- Sijtsma, K., & Molenaar, W.I. (1987). *Introduction to Nonparametric Item Response Theory*. Sage: Thousand Oaks, CA.
- van der Ark, L.A. (2010). Getting started with Mokken scale analysis in R. *Journal of Statistical Software, 46*. www.jstatsoft.org., CRAN.R-project.org/package=mokken.
- van der Ark, L.A. (2012). New developments in Mokken scale analysis in R. *Journal of Statistical Software, 48*. www.jstatsoft.org.

Components

- An automated item selection procedure (AISP)
  - Partitions of a set of ordinal items into scales or "Mokken Scale".
  - Some items might not be selected. These are "unscalable".

- Methods to check goodness-of-fit of non-parametric IRT model for each Mokken Scale or the strength of scales.

Notation

Notation:

- θ_a is a value of the latent variable.
- X_j is item j and x is the response to variable j.
- $P(X_j = x_j|\theta)$ equals to probability that response x_j is made to i given the value on the latent variable.

Assumptions

The most general non-parametric model is the "Monotone Homogeneity Model":

- Unidimensionality: There is only one latent being measured (i.e., one underlying variable explaining association between responses to items).
- Local independence.

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_J = x_J|\theta) = \prod_{j=1}^{J} P(X_j = x_j|\theta)$$

- Monotonicity:

$P(X_1 \geq x|\theta)$ is a nondcreasing function of θ

Assumptions (continued)

The addition of this fourth assumption yields the "double monotonicity model".

Nonintersection: For fixed value of θ_o, If

$$P(X_i \geq x|\theta_o) \geq P(X_j \geq x|\theta_o)$$

Then

$$P(X_i \geq x|\theta) \geq P(X_j \geq x|\theta)$$

for all θ.

The more general model has a special cases parametric IRT models: Rasch, 2 and 3 parameter logistic, graded response.

Scalability Coefficients

These are computed for pairs of items (i.e, H_{ij}), each item (i.e.,. H_i) and the total (i.e., H)

The item pair scalability coefficients are

- Based on the frequencies of responses in a cross-classification of responses to two items.
- Coefficients are weighted by responses inconsistent with Guttman's model; that is, "Guttman errors".
- Are defined as

$$H_{ij} = \frac{\text{cov}(X_i, X_j)}{\text{cov}(X_i, X_j)^{\text{max}}}$$

- $-\infty \geq H_{ij} \geq 1$
- If there are no Guttman errors, then $H_{ij} = 1$.
- The Monotone model implies that $0 \geq H_{ij} \geq 1$.
- If $H_{ij} < 0$, then an item does not fit the model.

Scalability Coefficient for Item and Scale

- The item scalability coefficient H_i equals

$$H_i = \frac{\text{cov}(X_i, R_{(i)})}{\text{cov}(X_i, R_{(i)})^{\text{max}}},$$

where $R_{(i)}$ is a rest-score.

- The test-scalability coefficient is H

$$H = \frac{\sum_{i=1}^{I} \text{cov}(X_i, R_{(i)})}{\sum_{i=1}^{I} \text{cov}(X_i, R_{(i)})^{\text{max}}}$$

Rules of Thumb

If $H = 1$, the you have a perfect Guttman scale.

Rules of thumb:

- If $H < .3$ the items are unscalable.
- If $.3 \geq H \geq .4$, a scale is weakly scalable.
- If $.4 \geq H \geq .5$, a scale is moderately scalable.
- If $.5 \geq H$, a scale is strongly scaleable.

R for Victimization

R code:

1. Install `mokken` package from your favorite site
2. library(mokken)
3. Read in data.
4. X = cbind(v1, v2, v3, v4)
5. coefh(X) (this yields the H coefficients)

Output for Victimization

Hij

	v1	se	v2	se	v3	se	v4	
v1			0.891	(0.026)	0.771	(0.040)	0.541	(0.
v2	0.891	(0.026)			0.843	(0.030)	0.530	(0.
v3	0.771	(0.040)	0.843	(0.030)			0.507	(0.
v4	0.541	(0.065)	0.530	(0.070)	0.507	(0.067)		

Hi			H	
Item	H	se	H	(se)
v1	0.759	(0.030)	0.712	(0.033)
v2	0.784	(0.027)		
v3	0.732	(0.033)		
v4	0.526	(0.061)		

Partitioning into Scales

Algorithm for the automated Item Selection Procedure or AISP.

There are 2 that have been implemented in `mokken`:

- Hierarchical clustering algorithm:
  - Starts by takeing 2 items having largest value of H_{ij} (significantly different from 0).
  - Items adding that meet criteria until no more can be added.
  - Takes unselected items and starts another Mokken scale
  - Continues until no more scales can be created.
- Genetic algorithm

R for Partitioning into Scales

Using victimization and fight items:

1. Take a look at these: coefH(vfX)
2. Using the default (HCA): hca $\longleftarrow$ aisp(vfX)
3. Using the genetic algorithm: ga $\longleftarrow$ asip(vfX, search="ga")
4. Combine to can compare: cbind(hc,ga)

Demonstrate in R. . . . Do we have 3 scales?