MAXENT-Modeling: A framework for IRT-Modeling?

Dipl.-Psych. Georg Hosoya

FU-Berlin
Cluster Languages of Emotion

10. Februar 2012

Overview

Discrete MAXENT-Models and RMs

- Conceptual foundation and relationships of the MAXENT-framework
- The canonical MAXENT-distribution is structurally similar to Rasch's modeling framework (Rasch, 1961)
- Example: PCM
- Model generation: Is it possible to apply the MAXENT-modeling framework to solve practical problems with IRT?

Overview

A RM for the assessment of intraindividual variability

- Are interindividual differences in intraindividual variability measurable?
- Definition of a model for multivariate, discrete time series (ambulatory assessment)
- Model properties, transition matrix, CRFs, difference of logits of adjacent categories for two persons

Application to ambulatory assessment data

- Estimation of model parameters via MCMC
- Heuristic assessment of model fit via standardized residuals
- Andrich-Reliability

MAXENT-Models and RMs

Principle of Maximum-Entropy

Reasoning under uncertainty: given a set of constraints, choose the one distribution (model) that a.) is in congruence with the available information (constraints) and b.) has maximum information entropy. (see e.g. Jaynes, 1957a, 1957b and 2003, chap. 11).

MAXENT-Models and RMs

Solution for the discrete case (details see Jaynes, 2003, chap. 11)

$$p(X = x) = \frac{1}{Z(\lambda_1, \ldots, \lambda_m)} \cdot \exp\left\{-\sum_{j=1}^{m} \lambda_j f_j(x)\right\} \tag{1}$$

Partition function:

$$Z(\lambda_1, \ldots, \lambda_m) = \sum_{i=1}^{n} \exp\left\{-\sum_{j=1}^{m} \lambda_j f_j(x)\right\} \tag{2}$$

MAXENT-Models and RMs

Expectation of $f_k(x)$ under the model

$$\langle f_k \rangle = -\frac{\partial \log Z}{\partial \lambda_k} \tag{3}$$

(Exponential family → sufficient statistics [Expectations should fit the acutual data F_k])

Variance of $f_k(x)$

$$\langle f_k^2 \rangle - \langle f_k \rangle^2 = -\frac{\partial^2 \log Z}{\partial \lambda_k^2}. \tag{4}$$

Useful for assessing test-information and model-fit (standardized residuals)

MAXENT-Models and RMs

Compatible modeling frameworks:

- Undirected graphical models (fusion of probability theory and graph theory)
- Markov random fields
- Boltzmann distribution / Gibbs distribution
- Logistic Regression / Perzeptron
- Multinomial Logit Regression
- Bayes Networks
- HMMs
- ...

MAXENT-Models and RMs

Rasch's (1961) and Jayne's framework (1957)

$$P(X_{vi} = x) = \frac{\exp[\phi_x \theta_v + \psi_x \alpha_i + \chi_x \theta_v \alpha_i + \omega_x]}{\sum_{l=0}^{m} \exp[\phi_l \theta_v + \psi_l \alpha_i + \chi_l \theta_v \alpha_i + \omega_l]} \tag{5}$$

$$p(X = x) = \frac{\exp\left\{-\sum_{j=1}^{m} \lambda_j f_j(x)\right\}}{\sum_{i=1}^{n} \exp\left\{-\sum_{j=1}^{m} \lambda_j f_j(x)\right\}} \tag{6}$$

Structurally equivalent, only the notation differs.

Example PCM (Masters, 1982)

$$p(X_{vi} = x_{vi}) = \frac{\exp\{\theta_v x_{vi} + \beta_{ix}\}}{\sum_{l=1}^{m} \exp\{\theta_v l + \beta_{il}\}}, \tag{7}$$

$$Z = \sum_{l=1}^{m} \exp\{\theta_v l + \beta_{il}\}. \tag{8}$$

Expectation of x_{vi} under the model

$$\frac{\partial logZ}{\partial \theta_v} = \frac{\sum_{l=1}^{m} l \cdot \exp\{\theta_v \cdot l + \beta_{il}\}}{\sum_{l=1}^{m} \exp\{\theta_v \cdot l + \beta_{il}\}} \tag{9}$$

$$= \langle x_{vi} \rangle \tag{10}$$

Variance of x_{vi} under the model

$$\frac{\partial^2 logZ}{\partial \theta_v^2} = \frac{\sum_{l=1}^{m} l^2 \cdot \exp\{\theta_v l + \beta_{il}\}}{\sum_{l=1}^{m} \exp\{\theta_v l - \beta_{il}\}} - \left[\frac{\sum_{l=1}^{m} l \cdot \exp\{\theta_v l + \beta_{il}\}}{\sum_{l=1}^{m} \exp\{\theta_v l + \beta_{il}\}}\right]^2 \tag{11}$$

$$= \langle x_{vi}^2 \rangle - \langle x_{vi} \rangle^2. \tag{12}$$

Intermediate conlcusions

- RMs and MAXENT-Models are structurally very similar, if not equivalent
- MAXENT-Models fit into the broader framework of undirected graphical models
- The MAXENT-framework gives a reason rooted in information theory for the form of a certain model
- Probabilistic psychometrics is not an ivory tower, very broad (mostly unexplored) connections to other disciplines via the modeling technique

Next step: Exploration of the framework for the formulation of new probabilistic IRT-models for research questions in psychometrics

A RM for the assessment of intraindividual variability

Model based on absolute sucessive difference

$$p(X_{vi[t]} = x_{vit} | x_{vi[t-1]}) = \frac{\exp\{|x_{vi[t]} - x_{vi[t-1]}| \cdot \eta_v - \beta_{ix}\}}{\sum_{l=1}^{m} \exp\{|l - x_{vi[t-1]}| \cdot \eta_v - \beta_{il}\}}$$

- x_{vit}: response of a person v on item i at timepoint t
- $x_{vi[t-1]}$: response of a person v on item i at timepoint $t-1$
- η_v: variability parameter of Person v
- β_{ix}: easiness of category l of item i
- sum zero norming over item-specific category parameters

Parameters (one item, 4 response categories)

$\eta_v = 0,\ \beta = [-0.5, 0.5, 0.5, -0.5]$

Transition matrix

	$x_{vi[t]}$=1	$x_{vi[t]}$=2	$x_{vi[t]}$=3	$x_{vi[t]}$=4
$x_{vi[t-1]}$=1	.13	.37	.37	.13
$x_{vi[t-1]}$=2	.13	.37	.37	.13
$x_{vi[t-1]}$=3	.13	.37	.37	.13
$x_{vi[t-1]}$=4	.13	.37	.37	.13

Independence of x_{vit} from $x_{vi[t-1]}$.

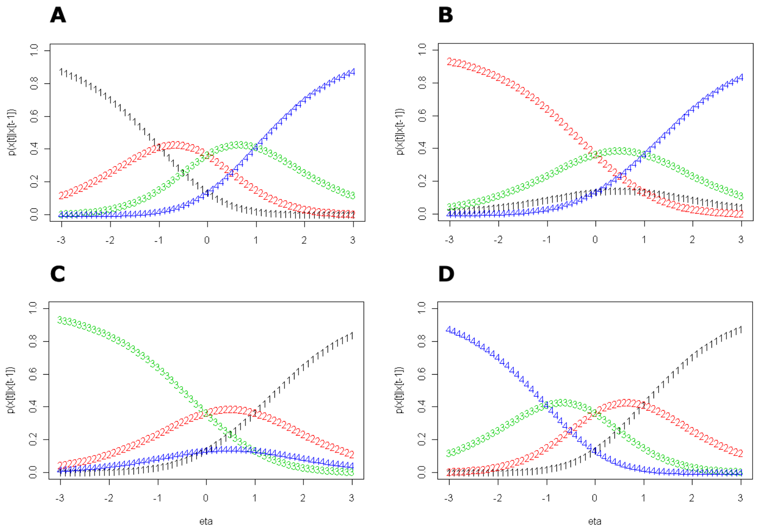
Parameters (one item, 4 response categories)

$\eta_v = -1.5,\ \beta = [0, 0, 0, 0]$

Relatively stable transition matrix

	$x_{vi[t]}$=1	$x_{vi[t]}$=2	$x_{vi[t]}$=3	$x_{vi[t]}$=4
$x_{vi[t-1]}$=1	.57	.35	.08	.00
$x_{vi[t-1]}$=2	.06	.76	.17	.01
$x_{vi[t-1]}$=3	.01	.17	.76	.06
$x_{vi[t-1]}$=4	.00	.08	.35	.57

4 conditional CRFs per Item



Difference of logits for adjacent category probabilities for two persons

$$\frac{\log\left(\frac{p(x_{vit}+1|x_{vi[t-1]})}{p(x_{vit}|x_{vi[t-1]})}\right) - \log\left(\frac{p(x_{vit}+1|x_{vi[t-1]})}{p(x_{vit}|x_{vi[t-1]})}\right)'}{[f(x_{vit}+1) - f(x_{vit})]} = \eta_v - \eta_v' \quad (13)$$

$$f(x_{vit}) = |x_{vit} - x_{vi[t-1]}| \quad (14)$$

$$f(x_{vit}+1) = |(x_{vit}+1) - x_{vi[t-1]}| \quad (15)$$

Application to ambulatory assessment data

Dataset: Crayen et al. 2012

- 165 students at Freie Universität Berlin
- study of mood regulation
- 2 weeks of ambulatory assessment with handheld device (7-8 signals/day)
- relevant here: MDBF short-form (Steyer, Schwenkmezger, Nostitz & Eid, 1997)
- n=64760

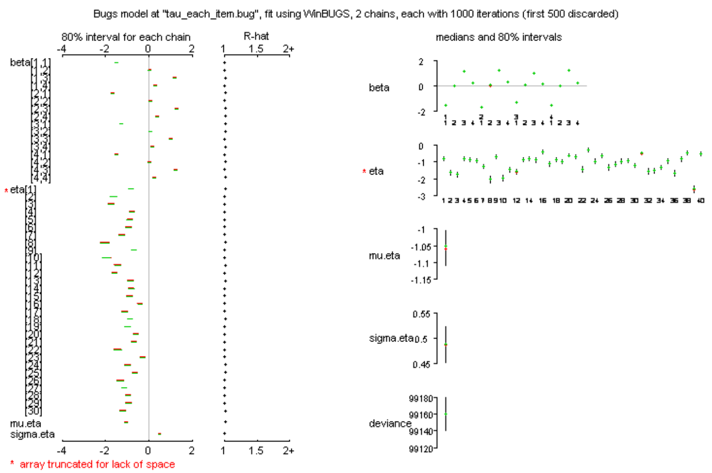
Application to ambulatory assessment data

Subscale: pleasant-unpleasent mood

- unwell-well
- bad-good
- unhappy-happy
- discontent-content
- 4-point rating scale

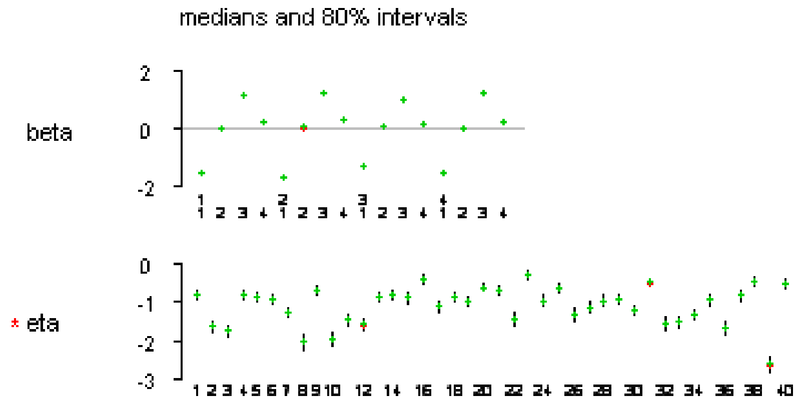
Application to ambulatory assessment data

MCMC-results

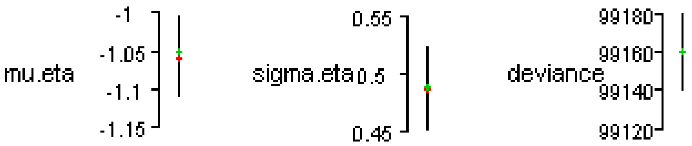


Application to ambulatory assessment data

Posterior distributions of category- and person parameters



Paremeters of latent score distribution



Reliability

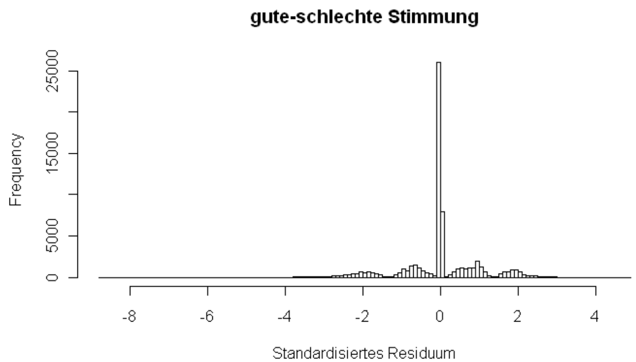
Andrich:

$$Rel = \frac{var(\hat{\eta}) - var(e)}{var(\hat{\eta})} = \frac{0.225 - 0.011}{0.225} = 0.95 \qquad (16)$$

Based on latent score distribution:

$$Rel = \frac{var(\hat{\eta})}{var(\eta)} = \frac{0.225}{0.238} = 0.95 \qquad (17)$$

Standardized residuals



Item Outfit

Item	Outfit	χ^2	df	p
well	1.06	17217.63	16190	0.00
good	1.00	16239.95	16190	0.39
contempt	0.99	16068.96	16190	0.75
happy	0.88	14232.33	16190	1.00

Conclusion

- MAXENT-modeling seems to be a framework that is applicable to research-questions in psychometrics
- Very well worked out formalism in for example artificial intelligence and machine learning (probabilistic graphical models)
- Complicated matter, further exploaration and cross-disciplinary dialogue necessary.

Contact: georg.hosoya@fu-berlin.de

Entry points into the literature

IRT

Fischer, G.H. & Molenaar, I.W. (Eds.) (1995). *Rasch models - recent developments and applications.* New York: Springer.

MAXENT

Jaynes, E.T. (2003). *Probability theory - the logic of science.* Cambridge: Cambridge University Press. (especially chapter 11)
Jaynes, E.T. (1957a). *Information theory and statistical mechanics.* Physical Review 106(4), 620-630.
Jaynes, E.T. (1957b). *Information theory and statistical mechanics II.* Physical Review, 108(2), 171-190.

Entry points into the literature

Probabilistic graphical models

Koller, D. & Friedman, N. (2009). *Probabilistic graphical models - principles and techniques.* Cambridge, MA: MIT Press.

GLAMM

Skondral, A. & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling - multilevel, longitudinal and structural equation models.* Boca Raton: Chapman & Hall.

Intraindividual variability

Ram, N. & Gerstorf, D. (2009). *Time-structured and net intraindividual variability: tools for examining the development of characteristics and processes.* Psychology and Aging, 24(4), 856-862.