Das Mixed Rasch-Modell

Kathrin Jahreis

Vortrag im Rahmen des Masterseminars
Psychometrische Modelle: Theorie und Anwendungen
am Institut für Statistik (LMU)

06./07.06.2014

Das Mixed Rasch-Modell (MRM)

- Flexible Mischverteilungs-Modellklasse für kategoriale Daten mit großer heuristischer Kraft.
- Das MRM (Rost, 1990) resultiert aus der Überwindung der Limitationen des RM und der Latent-Class-Analysis (LCA).
- Es ermöglicht die Modellierung von Differential Item Functioning (DIF). Es können z.B. kognitive Strukturen oder Lösungsstrategien aus den Daten erkannt werden.
- Das MRM kann als Modellgeltungstest für das RM fungieren.

Limitationen des RM

- Annahme **paralleler** ICCs setzt voraus, dass Itemparameter (IP) σ_i für alle Personen mit Fähigkeit θ_v konstant sind.
- In der Testpraxis ist diese Annahme selten gerechtfertigt:
  - Schlecht konstruierte Items (ROST, 1990)
  - Items mit für unterschiedliche Subgruppen variierender Schwierigkeit zu diagnostischen Zwecken
- Da bei DIF das RM die Komplexität der Testsituation nicht erfasst, wird seine Passung von Modellgeltungstest kategorisch verworfen.

Die klassische LCA

- LCA (Lazarsfeld, 1950; Lazarsfeld & Henry, 1968) verwendet kategoriales θ_g statt stetiges θ_v.
- Qualitative Zuordnung der getesteten Individuen in G diskrete Klassen (disjunkt und exhaustiv).
- Spezifische Antwortmuster je Klasse, welche jedoch zwischen Klassen divergieren.
- Konstante, klassenspezifische LWS π_{ig} ($g = 1$, 2,...,G), wobei $\pi_{ig} \in [0, 1]$ (Wskparameter).

Die klassische LCA

- Annahme, dass Klassen bzgl. ihres Antwortverhaltens unabhängig voneinander sind. In Klassen gilt LSU der Item-Responses.
- Klassengrößen (KG) π_g sind unbekannt (Wskparameter). Es gilt Normierung $\sum_{g=1}^{G} \pi_g = 1$.
- Unbedingte PWS ist **diskrete Mischung** der bedingten PWS:

$$p(\underline{x}) = \sum_{g=1}^{G} \pi_g p(\underline{x}|g) \tag{1}$$

- Anzahl latenter Klassen G ist **kein**, auf g bedingte PWS sind Modellparameter (MP).
- Anzahl gelöster Aufgaben **nicht** suffizient für θ_v.

Limitationen der LCA

- LCA ist voraussetzungsarme, flexible Modellklasse (Annahme über G als einzige Annahme bzgl. Beschaffenheit der latenten Klassen).
- LCA bzgl. Annahme von Personenhomogenität weniger restriktiv als RM (Klassen können sich durch unterschiedliche IP auszeichnen).
- Jedoch fordert LCA konstante LWS aller Klassenmitglieder für ein Item. Es ist keine Variabilität von θ in Klassen möglich.

Lokalisierte Klassen

- Lokalisierte Klassen (Lazarsfeld & Henry, 1968) setzen Existenz von stetigem θ voraus, dessen konkrete Ausprägungen sich jedoch an Verdichtungspunkten häufen.
- Eigentlich stetiges θ reduziert sich auf begrenzte Anzahl von disjunkten und exhaustiven, latenten Klassen.
- Da Klassen dennoch auf Kontinuum lokalisiert sind, können klassenspezifische LWS durch eine auf das gesamte Kontinuum bezogene ICC definiert werden.

Die Linear-logistische LCA

- Linear-logistische LCA nach Formann (1985) steht in direkten Zusammenhang mit lokalisierten Klassen.
- Reparametrisierung der MP der LCA (π_g, π_{ig}) durch logistische Trafo:

$$\alpha_{ig} = \log\left(\frac{\pi_{ig}}{1-\pi_{ig}}\right) \Rightarrow \pi_{ig} = \frac{\exp(\alpha_{ig})}{1+\exp(\alpha_{ig})} \qquad (2)$$

- Parameter α_{ig}, α_g können Werte zwischen $-\infty$ und $+\infty$ annehmen.
- Nun Präzisierung der LCA durch Restriktionen (additiv-lineare Zerlegungen von α_{ig}, α_{ig}) möglich.

Die Linear-logistische LCA

- Bei Zerlegung von α_{ig} in Klassenparameter θ_g und IP σ_i resultiert **Modell mit lokalisierten Klassen**, dessen **Itemfunktionen diejenigen des RM** sind:

$$p(x_{vi}) = \sum_{g=1}^{G} \pi_g \frac{\exp(x_{vi}(\theta_g - \sigma_i))}{1 + \exp(\theta_g - \sigma_i)} \tag{3}$$

- Im Vgl. zu RM ist θ nicht kontinuierlich, sondern diskret verteilt (Anzahl Fähigkeitsausprägungen entspricht G).
- Approximation des RM durch Modell lokalisierter Klassen erfordert nach Formann (1995):

$$G = \begin{cases} (k+1)/2, & \text{wenn } k \text{ ungerade,} \\ k/2+1, & \text{wenn } k \text{ gerade.} \end{cases} \tag{4}$$

Die informelle Formulierung des dichotomen MRM

- **Grundidee des MRM** nach Rost (1990):
  Teile getestete Personenpopulation derart in G exhaustive und disjunkte, latente Klassen mit maximal usl. IP auf, dass klassenintern jeweils das RM gilt.
- **Fazit auf Basis der bisherigen Kenntnisse:**
  Variierende Itemschwierigkeit für verschiedene Klassen UND Zulässigkeit von Fähigkeitsunterschieden innerhalb der Rasch-skalierbaren Klassen entsprechen Koexistenz der beiden Grundideen von RM und LCA.

Das formale Modell des dichotomen MRM

- **Notation:**
  - θ_{vg}: Fähigkeit von Person v aus Klasse g.
  - σ_{ig}: Leichtigkeit von Item i für Individuen aus Klasse g.
  - p_{vig}: Wsk, dass Person v mit Fähigkeit θ_{vg} Item i löst.
- **Annahmen:**
  - $\sum_{g=1}^{G} \pi_g = 1$: KG $\pi_g \in [0, 1]$ ergeben in Summe 1.
  - $\sum_{i=1}^{k} \sigma_{ig} = 0 \; \forall g$: IP einer Klasse sind summennormiert.

Das formale Modell des dichotomen MRM

- Die **klassenspezifische LWS** p_{vig} ergibt sich wie folgt:

$$p_{vig} = p(X_{vi} = 1|g) = \frac{\exp(\theta_{vg} + \sigma_{ig})}{1 + exp(\theta_{vg} + \sigma_{ig})} \quad (5)$$

- Wegen Annahme exhaustiver und disjunkter Klassen ergibt sich **unbedingte LWS** p_{vi} als:

$$p_{vi} = \sum_{g=1}^{G} \pi_g p_{vig} = \sum_{g=1}^{G} \pi_g \frac{exp(\theta_{vg} + \sigma_{ig})}{1 + \exp(\theta_{vg} + \sigma_{ig})} \quad (6)$$

- MRM ist Mischverteilung mit klassenspezifischen LWS p_{vig} als Rasch-skalierte **Verteilungen** der latenten PV und den KG π_g als **Mischungskomponenten**.

Das MRM als Metamodell von RM und LCA

MRM bildet gemeinsames **Metamodell** von RM und LCA:

- **Szenario I:**
  - Annahme, dass $G = 1$.
  - Wegen Normierung $\sum_{g=1}^{G} \pi_g = 1$ folgt, dass $\pi_g = \pi_1 = 1$.
  - Es resultiert Modellgleichung des RM.
- **Szenario II:**
  - Annahme, dass es in Klassen keine Varianz in θ_{vg} gibt.
  - $\theta_{vg}{=}\theta_g$ (LWS für Item in g für alle Personen konstant).
  - Gleichung (6) kann mit (7) reparametrisiert werden:

$$p_{vig} = \frac{\exp(\theta_g + \sigma_{ig})}{1 + exp(\theta_g + \sigma_{ig})} = \pi_{ig} \tag{7}$$

  - Es resultiert Modell der LCA für LWS für Item i.

Die bedingte Likelihood des MRM

- Auf (6) basierende **joint Likelihood** des MRM enthält π_g, σ_{ig} und θ_{vg} als MP.
- Reparametrisierung der Likelihood (L) durch Verwendung der **bedingten** statt der **unbedingten PWS** bietet sich an.
- Inzidentelle θ_{vg} sind in **bedingter L** nicht mehr enthalten.
- Eigenschaft stellt Vorteil für Parameterschätzung dar.

Herleitung der bedingten Likelihood des MRM

- **Klassenunspezifische PWS** von $\underline{x} = (x_1, x_2, ..., x_k)$:

$$p(\underline{x}) = \sum_{g=1}^{G} \pi_g p(\underline{x}|g) = \sum_{g=1}^{G} \pi_g p(\underline{x}|r, g) p(r|g) \tag{8}$$

- Unter Verwendung des Satzes der totalen WSK kann $p(\underline{x}|g)$ zusätzlich auf Summenwert r von $\underline{x}$ bedingt werden.
- **Klassenspezifische SWS** $p(r|g)$ werden mit π_{rg} reparametrisiert und stellen die neuen MP dar.
- Für sie gilt $\sum_{r=1}^{k-1} \pi_{rg} = 1 \; \forall \; g$.

Die bedingte Likelihood des MRM

- **Bedingte PWS** $p(\underline{x}|r,g)$ sind der Anteil, den die WSK für bestimmtes Pattern mit Score r an der GesamtWSK aller Pattern mit Score r hat:

$$p(\underline{x}|r,g) = \frac{p(\underline{x}|g)}{\sum\limits_{\underline{x}|r,g} p(\underline{x}|g)} \tag{9}$$

- **Herleitung von** $p(\underline{x}|g)$ ergibt ausgehend von (5):

$$p(\underline{x}|g) = \frac{\exp(r\ \theta_{vg})\exp(\sum\limits_{i=1}^{k} x_i\ \sigma_{ig})}{d_{vg}} \tag{10}$$

Die bedingte Likelihood des MRM

- Die Summe der WSK aller Pattern $\underline{x}$ mit Summenscore r ergibt sich unter Verwendung von (10) als:

$$\sum_{\underline{x}|r,g} p(\underline{x}|g) = \sum_{\underline{x}|r,g} \frac{\exp(r\ \theta_{vg}) \exp(\sum_{i=1}^{k} x_i\ \sigma_{ig})}{d_{vg}} \tag{11}$$

- Einsetzen von (10) und (11) in (9) ergibt:

$$p(\underline{x}|r,g) = \frac{\exp(\sum_{i=1}^{k} x_i\ \sigma_{ig})}{\sum_{\underline{x}|r,g} \exp(\sum_{i=1}^{k} x_i\ \sigma_{ig})} \tag{12}$$

Die bedingte Likelihood des MRM

- (12) ist Funktion ohne θ_{vg}.
- Nenner von (12) kann umgeschrieben werden zu:

$$\sum_{\underline{x}|r,g} \exp(\sum_{i=1}^{k} x_i \ \sigma_{ig}) = \sum_{\underline{x}|r,g} \prod_{i=1}^{k} (\exp x_i \ \sigma_{ig}) = \gamma_r(\exp(\sigma)) \tag{13}$$

- Aus (12) und (13) folgt die **bedingte PWS** als:

$$p(\underline{x}|r,g) = \frac{exp(\sum_{i=1}^{k} x_i \ \sigma_{ig})}{\gamma_r(\exp(\sigma))} \tag{14}$$

Die bedingte Likelihood des MRM

- **Bedingte PWS** können als schlichte Funktion der IP und ihrer symmetrischen Grundfunktionen r-ter Ordnung γ_r dargestellt werden.
- γ_r ist Funktion von delogarithmierten IP und Score r. Es umfasst alle möglichen Antwortpattern, welche zu einem Score r führen und ist unabhängig von konkreten Pattern $\underline{x}_v$ einer Person.
- Setzt man (14) in (8) ein, ergibt sich **Modellgleichung auf Ebene der PWS**:

$$p(\underline{x}) = \sum_{g=1}^{G} \pi_g \pi_{rg} \frac{exp(\sum_{i=1}^{k} x_i \ \sigma_{ig})}{\gamma_r(\exp(\sigma))} \tag{15}$$

Die bedingte Likelihood des MRM

- Wegen Unabhängigkeit des Antwortverhaltens von Klassenmitgliedern und Klassen untereinander, ergibt sich **bedingte L der gesamten Testdatenmatrix** zu:

$$L = \prod_{\underline{x}} \left(\sum_{g=1}^{G} \pi_g \pi_{rg} \frac{exp(\sum_{i=1}^{k} x_i \ \sigma_{ig})}{\gamma_r(\exp(\sigma))} \right)^{n(\underline{x})} \tag{16}$$

- IP können geschätzt werden, ohne dass PP bekannt sind oder Annahmen bzgl. ihrer Verteilung in den Klassen getroffen werden müssen.
- Da L keine inzidentellen Parameter mehr enthält, resultieren konsistente, asymptotisch normalverteilte und asymptotisch erwartungstreue $\hat{\sigma}_{ig}$.

Parameterschätzung im MRM

- Schätzung erfolgt mit dem **EM-Algorithmus** (Dempster, Laird & Rubin, 1977).
- Flexibler, iterativer Algorithmus zur Berechnung von ML-Schätzungen bei unvollständigen Daten.
- **Grundidee:** Schätzproblem auf vollständige Likelihood und somit auf Max vollständiger Daten zurückführen.
- Eine Iteration besteht aus **Expectation-Schritt** und **Maximization-Schritt**.
- Iteration bis zur Konvergenz des unbekannten Parameters.

Parameterschätzung im MRM

Bedingte klassenspezifischen Log-L des MRM:

$$L_g = \prod_{\underline{x}} \left(\pi_{rg} \frac{exp(\sum_{i=1}^{k} x_i \ \sigma_{ig})}{\gamma_r(\exp(\sigma))} \right)^{n(\underline{x})} \tag{17}$$

- G ist kein Standard-MP (wird nicht geschätzt).
- π_g, σ_{ig} und π_{rg} sind unabhängige MP.
- Scores $r = 0$ und $r = k$ sind explizit von Schätzung ausgeschlossen. Wegen bedingter ML-Schätzung beeinflusst ihr Ausschluss Schätzungen der IP nicht.

Parameterschätzung im MRM

E-Schritt: Berechne **erwartete Patternhäufigkeiten** (PH) $\hat{n}(\underline{x}, g)$ für jede Klasse:

$$\hat{n}(\underline{x}, g) = n(\underline{x}) \frac{\pi_g p(\underline{x}|g)}{\sum\limits_{g=1}^{G} \pi_g p(\underline{x}|g)} \tag{18}$$

- $n(\underline{x})$ sind global beobachtete Häufigkeiten von Pattern $\underline{x}$.
- $\hat{n}(\underline{x}, g)$ hängen nur von MP π_g, σ_{ig} und π_{rg} ab. Verwende vorläufige Schätzungen oder Startwerte.
- $\hat{n}(\underline{x}, g)$ bilden Basis, um im **M-Schritt** bessere Schätzungen für π_g, π_{rg} und σ_{ig} zu berechnen.

Schätzung der IP im MRM

π_g, π_{rg} und σ_{ig} können durch Max. der **klassenspezifischen Log-**L_g für jede Klasse g separat geschätzt werden:

$$\log L_g = \sum_{\underline{x}} \hat{n}(\underline{x}, g) \left(log(\pi_{rg}) + \sum_{i=1}^{k} x_i \sigma_{ig} - log(\gamma_r(\exp(\sigma))) \right) \tag{19}$$

Erste partielle Abl. von (19) nach σ_{ig} gleich Null gesetzt. Es resultieren Schätzgleichungen:

$$\hat{\sigma}_{ig} = \log \frac{n_{ig}}{\sum_{r=0}^{k} (n_{rg} \gamma_{r-1,i} / \gamma_r)} \tag{20}$$

$\gamma_{r-1,i}$ (sym. GF (r-1)-ter Ordnung aller IP mit Ausnahme von Item i) ist erste partielle Abl. von γ_r.

Parameterschätzung im MRM

- n_{ig} sind die vorläufigen Schätzungen für die Anzahl an Personen in Klasse g, die Item i gelöst haben.
- n_{rg} sind Schätzer für die Anzahl von Personen mit Score r in Klasse g.
- n_{ig} und n_{rg} lassen sich mittels einfacher Summen von $\hat{n}(\underline{x}, g)$ aus E-Schritt berechnen. Erwartete Counts müssen nicht ganzzahlig sein.
- Die symmetrischen GF auf rechter Seite von (20) werden mit vorläufigen IP-Schätzungen berechnet.

Parameterschätzung im MRM

Schätzer für die SWS und die KG:

$$\hat{\pi}_{rg} = \frac{n_{rg}}{n_g} \tag{21}$$

$$\hat{\pi}_g = \frac{\sum_{\underline{x}} \hat{n}(\underline{x}, g)}{N} \tag{22}$$

- $\sum_{\underline{x}} \hat{n}(\underline{x}, g)$ sind **Klassenanteile** der beob. klassenspezifischen PH.

Schätzung der PP im MRM

θ_{vg} wurden durch Bedingen auf r aus L entfernt, können aber unter Verwendung von $\hat{\sigma}_{ig}$ und Schätzungen der Scoreparameter durch Max von L_g bzgl. θ_{vg} (hängt nur von Score r_v des Individuum S_v ab) geschätzt werden.

Intra-class L:

$$L_g = \prod_v p(\underline{x}_v|g) = \prod_v \frac{\exp(\sum_{i=1}^{k} x_{vi}(\theta_{vg} + \sigma_{ig}))}{\prod_{i=1}^{k} [1 + exp(\theta_{vg} + \sigma_{ig})]} \tag{23}$$

$$\log L_g = \sum_v r_v \, \theta_{vg} + \sum_i n_{ig} \, \sigma_{ig} - \sum_v \sum_i \log(1 + \exp(\theta_{vg} + \sigma_{ig})) \tag{24}$$

Schätzung der PP im MRM

- n_{ig} ist Anzahl von Personen, die Item i in Klasse g löst (muss nicht bekannt sein, da n_{ig} durch Ableiten von (24) nach θ_{vg} herausfällt).
- Schätzgleichungen für θ_{vg} ergeben sich durch Ableiten von (24) nach θ_{vg} und auflösen der Ableitung nach r_v:

$$r_v = \sum_{i=1}^{k} \frac{\exp(\theta_{vg} + \sigma_{ig})}{1 + \exp(\theta_{vg} + \sigma_{ig})} \tag{25}$$

- Wie in RM erhalten Personen mit demselben Summenscore auch denselben PP. Hier PP aber klassenspezifisch, d.h. jede Person erhält G θ_{vg}.
- Schätzungen können als bedingt angesehen werden (Individuum S_v hat Parameter θ_{vg} sofern es zu g gehört).

Parameterschätzung im MRM

- Varianz von θ_{vg} zw. Klassen hängt nur von $\hat{\sigma}_{ig}$ ab und nicht davon, welche Items von Individuum gelöst wurden.
- Umgekehrt hängt wahrscheinlichste Klassenzugehörigkeit von Individuum wenig von r_v, sondern von $\underline{x}$ (Welche Items wurden gelöst?) ab.
- **FAZIT:** MRM erlaubt es, quantitativen und qualitativen Aspekt eines Response-Patterns voneinander zu trennen. Stetige Mischvariable hängt davon ab, wie viele Items gelöst wurden (RM) und die diskrete Mischvariable davon, welche Items gelöst wurden (LCA).

Anwendung des MRM

- Vielzahl von Einsatzgebieten (Psychologie, Sprachwissenschaft, Genetik...).
- Zwar muss Klassenanzahl a priori vorgegeben werden, über Beschaffenheit ist aber kein Vorwissen notwendig.
- A posteriori folgt ggf. substanzwissenschaftliche Untersuchung der vom MRM identifizierten Klassen.
- **Beispiele aus Literatur:**
  - Praktische versus theoretische Kompetenzen von Schülern in Physik (Rost, Häußler & Hoffmann,1989)
  - Identifizierung von Lösungsstrategien (Mislevy & Verhelst, 1990)

Simulationsdaten von Rost (1990) in R

- Simulationsdaten von Rost (1990) sind in R-Paket **psychomix** als Funktion *simRaschmix()* implementiert:
- Betrachte drei DGPs mit je N=1800 ($r = 0$ und $r = k$ ausgeschlossen) und 10 Items.
  - **DGP1**: $G = 1$ (Rasch-skalierbar) mit IP-Vektor $\underline{\beta}_1$
  - **DGP2**: $G = 2$, $\underline{\beta}_1 = -\underline{\beta}_2$ (umgekehrte RF der IP)
  - **DGP3**: $G = 3$, $\underline{\beta}_3$ (fast keine Varianz der IP)
- Schätzung von MRM erfolgt mit Funktion *raschmix()* aus **psychomix**.

Simulationsdaten von Rost (1990) in R

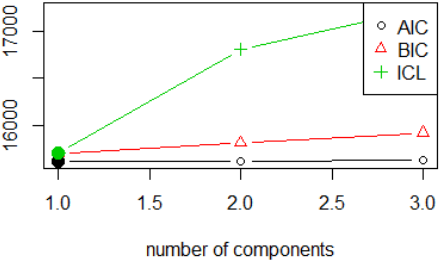

Figure: Modellwahl bei DGP1 mittels Informationskriterien

Simulationsdaten von Rost (1990) in R

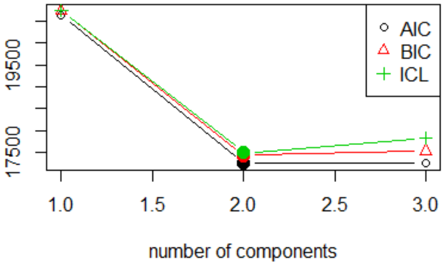

Figure: Modellwahl bei DGP2 mittels Informationskriterien

Simulationsdaten von Rost (1990) in R

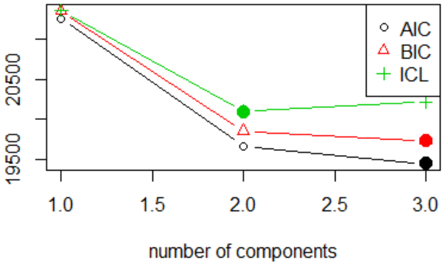

Figure: Modellwahl bei DGP3 mittels Informationskriterien

Simulationsdaten von Rost (1990) in R

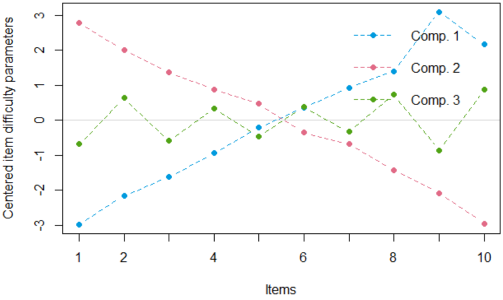

Figure: Schätzungen der IP bei DGP3

Fazit

- Bei MRM ist Annahme von Eindimensionalität in Gesamtpopulation der getesteten Personen nicht notwendig. Es ist somit ein zuverlässiges Verfahren zur Modellierung bzw. Identifizierung von DIF.
- MRM setzt a priori weder Annahmen über Ordnung der Items in Klassen noch bzgl. Klassengröße voraus.
- Anzahl latenter Klassen G ist (wie bei LCA) kein Modellparameter und muss a priori festgesetzt werden.
- Wie im RM gibt es Möglichkeit, IP verteilungsfrei zu schätzen.
- Es gibt auch eine Erweiterung des MRM auf mehrstufige ordinale Daten (Rost, 1991).

Literatur

- Baghaei, P. & Carstensen, C.H. (2013). Fitting the Mixed Rasch Model to a Reading Comprehension Test: Identifying Reader Types.*Practical Assesment, Research & Evaluation*, 18, 5, 1-8.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM-algorithm.*Journal of the Royal Statistical Society*, 39, 1, 1-38.
- Fischer, G.H. (1974). Einführung in die Theorie psychologischer Tests. Bern: Huber.
- Formann, A.K. (1985). Constrained Latent Class Models: Theory and Applications.*British Journal of Mathematical and Statistical Psychology*, 38, 87-111.
- Formann, A.K. (1995). Linear Logistic Latent Class Analysis and the Rasch Model. In G.H. Fischer & I.W. Molenaar (Hrsg.),*Rasch models: Foundations, recent developments, and applications* (S.239-255). Berlin: Springer.
- Frick, H., Strobl, C., Leisch, F., Zeileis, A. (2011). Flexible Rasch Mixture Models with Package psychomix.*Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics*. Universität Innsbruck.
- Lazarsfeld, P.F. (1950). The Logical and Mathematical Foundation of Latent Structure Analysis, in: S. Stouffer (Hrsg.), Measurement and Prediction (S. 362-412). NJ: Princeton University Press.
- Lazarsfeld, P. F., Henry, N. W. (1968). Latent Structure Analysis. Boston: Houghton Mifflin.
- Mislevy, R.J., Verhelst, N.D. (1990). Modeling Item Responses when different subjects employ different solution strategies. Psychometrika Vol. 55, Nr. 2, S. 195-215.
- Pawitan, Y. (2001). In All Likelihood. Statistical Modelling and Inference Using Likelihood. Oxford: Clarendon Press.

Literatur

- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis, Applied Psychological Measurement 14(3), S. 271.
- Rost, J. (1991). A logistic mixture distribution model for polychotomous item responses, British Journal of Mathematical and Statistical Psychology 44(1), S.75-92.
- Rost, J. (1996). Lehrbuch Testtheorie, Testkonstruktion. Bern: Huber.
- Rost, J., Langeheine, R. (1997). A guide through latent structure models for categorical data. In J. Rost & Langeheine, R. (Hrsg.), *Applications of Latent Trait and Latent Class Models in the Social Sciences* (S.13 -37). Münster: Waxmann.
- Rost, J. (2006). Item-Response-Theorie, in: Petermann, F., Eid, M. (Hrsg.). Handbuch der Psychologischen Diagnostik (Band 4), S. 261 274. Göttingen: Hogrefe.
- Rost, J. (2006). Latent-Class-Analyse, in: Petermann F., Eid, M. (Hrsg.). Handbuch der Psychologischen Diagnostik (Band 4), S. 275-287. Göttingen: Hogrefe.
- Rost, J., Von Davier, M. (1995). Mixture Distribution Rasch Models. In G.H. Fischer & I.W. Molenaar (Hrsg.),*Rasch models: Foundations, recent developments, and applications* (S.257-268). Berlin: Springer.
- Rost, J., Eid, M. (2009): Mischverteilungsmodelle. In: H. Holling (Hrsg.). Grundlagen und statistische Methoden der Evaluationsforschung (Enzyklopädie der Psychologie, Themenbereich B, Serie IV, Band 1),14. Kapitel, S.483-524. Göttingen: Hogrefe.
- Schmid, H. (1992). Psychologische Tests: Theorie und Konstruktion. Freiburg (Schweiz): Universitätsverlag.
- Rost, J., von Davier, M. (1995). Mixture distribution Rasch Models, in: Fischer, G., Molenaar, I. (Hrsg.). Rasch Models: Foundations, recent developments, and applications, Springer, S.257-268.
- Smit, A., Kelderman, H., van der Flier, H. (1999). Collateral information and Mixed Rasch models. Methods of Psychological Research Online, Vol. 4, No. 3.