

Übung zur Vorlesung - Theorien Psychometrischer Tests II

N. Rose

6. Übung (11.12.2008)



- **Agenda**
 - graphische Modellgeltungskontrolle
 - Likelihood Ratio Test
 - Bootstrap-Inferenz



Testbarkeit im Rasch-Modell

- Was bedeutet **Testbarkeit**?

... Frage der Testbarkeit, beschäftigt sich mit Konsequenzen, die aus den Definitionen eines Modells abgeleitet werden können und als Grundlage der Modellgeltungskontrolle dienen können!

Testbarkeit im Rasch-Modell

- **Testbare Konsequenzen im Rasch-Modell:**

1. Gleichheit der Rangreihe der unbedingten Lösungswahrscheinlichkeiten über die Subpopulationen (s):

$$P^{(s)}(Y_i = 1) \leq P^{(s)}(Y_j = 1) \quad \forall s \subset \Omega_U$$

2. Gleichheit der Itemparameter über Subpopulationen (s und t):

$$\beta_i^{(s)} = \beta_i^{(t)}$$

Testbarkeit im Rasch-Modell

- Testbare Konsequenzen im Rasch-Modell:

3. Gleichheit bestimmter Wahrscheinlichkeitsverhältnisse:

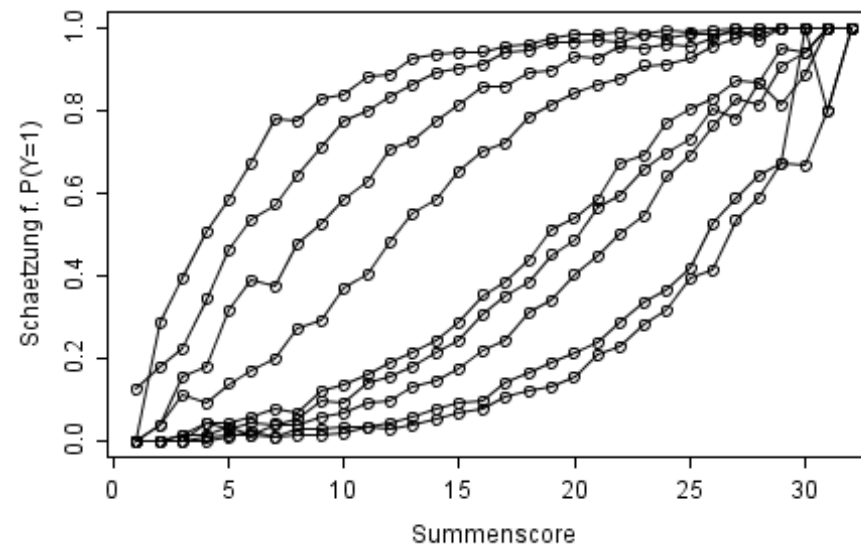
$$\frac{P^{(s)}(Y_i = 1 \cap Y_j = 0)}{P^{(s)}(Y_i = 0 \cap Y_j = 1)} = \frac{P^{(t)}(Y_i = 1 \cap Y_j = 0)}{P^{(t)}(Y_i = 0 \cap Y_j = 1)}$$

graphische Modellgeltungskontrolle

- **Nennen Sie Möglichkeiten der graphische Modellgeltungskontrolle im Rasch-Modell.**

1. überschneidungsfreie nicht-parametrische itemcharakteristische Kurven (*ICC*) für alle Items eines Tests

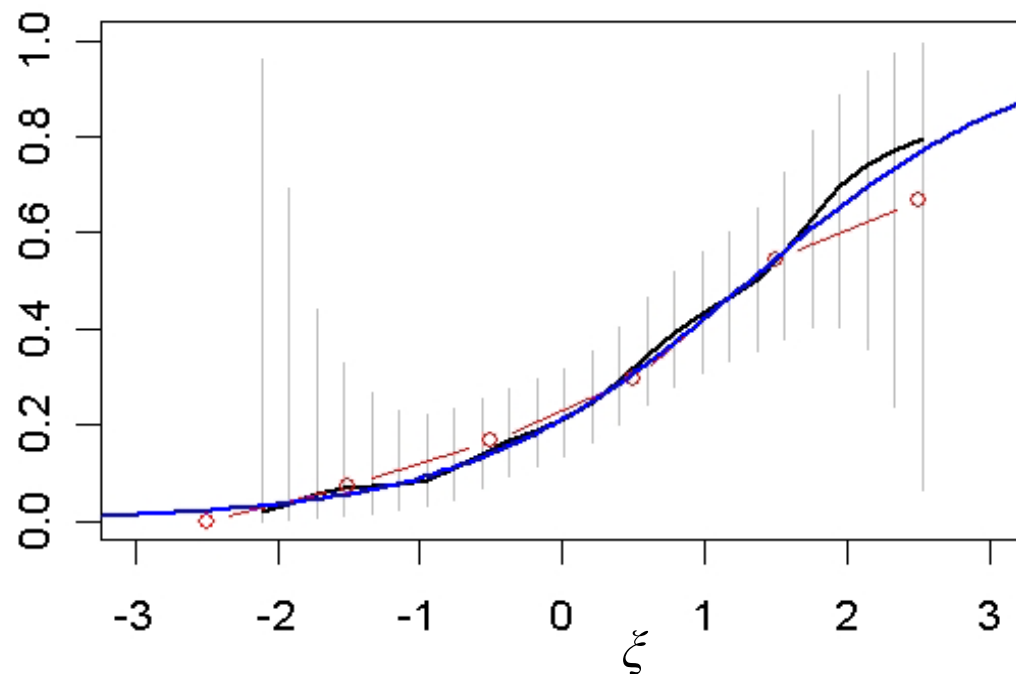
(Liniendiagramm: Summenscore vs. relative Lösungshäufigkeiten gegeben dem Summenscore)



graphische Modellgeltungskontrolle

2. graphischer Vergleich nicht-parametrische *ICC* und modellimplizierte *ICC* unter Annahme der Gültigkeit des Rasch-Modells

$$P(Y_i = 1 | \xi)$$



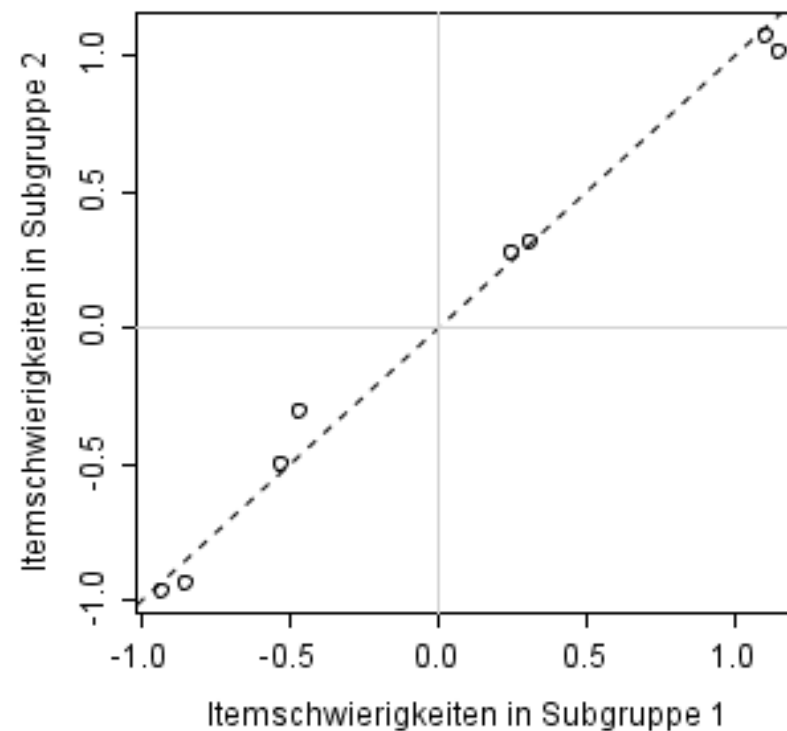
ICC unter Annahme
des Rasch-Modells

nicht-parametrische
ICC (Ramsey-Curve)

nicht-parametrische
ICC (Mittelwerte –
kategorisierte lat.
Variable)

graphische Modellgeltungskontrolle

3. graphischer Vergleich der geschätzten Itemschwierigkeiten aus zwei verschiedenen Subpopulationen
→ graphisch müssen die Schwierigkeitsschätzungen auf der winkelhalbierenden Geraden liegen!



Testbarkeit im Rasch-Modell

- **Was ist das Prinzip des Likelihood Ratio Tests (*LR*-Test)?**

Inferenzstatistischer Vergleich eines restriktiven Modells M_0 (Modell der Nullhypothese) und einem weniger restriktiven Modells M_1 (Modell der Alternativhypothese), anhand der Likelihoods L_0 und L_1 der beiden Modelle.

→ Allgemeiner *LR*-Test auf Modellgeltung im Rasch-Modell:

$L_{sat}(\mathbf{Y})$ = Likelihood ohne Annahme eines Modells

$L_{RA}(\mathbf{Y})$ = Likelihood unter Annahme des Rasch-Modells

Testbarkeit im Rasch-Modell

- **Wie kann die Daten-Likelihood $L_{sat}(Y)$ ohne Annahme eines Modells geschätzt werden?**

$L_{sat}(Y)$ kann durch das Produkt der relativen Antwortmusterhäufigkeiten (exponenziert mit ihren absoluten Auftretenshäufigkeiten) geschätzt werden.

- **Wie wird die Daten Likelihood $L_{RA}(Y)$ unter Gültigkeit des Rasch-Modells geschätzt.**

$L_{RA}(Y)$ als Produkt aller Antwortmuster unter Verwendung der Rasch-Modellgleichung.

Testbarkeit im Rasch-Modell

- Berechnung von $L_{sat}(\mathbf{Y})$

- Berechnung von $L_{RA}(\mathbf{Y})$

Allgemeine Daten-Likelihood: $L(\mathbf{Y}) = \prod_{u=1}^N P(\underline{\mathbf{Y}}(u))$

$$L_{sat}(\mathbf{Y}) = \prod_{q=1}^Q \hat{P}(\underline{\mathbf{Y}}_q)^{n(\underline{\mathbf{Y}}_q)}$$

$$L_{RA}(\mathbf{Y}) = \prod_{u=1}^N P(\underline{\mathbf{Y}} | \xi = \xi(u), \boldsymbol{\beta})$$

- mit: $\hat{P}(\underline{\mathbf{Y}}_q) = \left(\frac{n(\underline{\mathbf{Y}}_q)}{N} \right)$

- bei k Items:

$$P(\underline{\mathbf{Y}} | \xi = \xi(u), \boldsymbol{\beta}) = \prod_{i=1}^k \left(\frac{\exp(\xi(u) - \beta_i)}{1 + \exp(\xi(u) - \beta_i)} \right)^{y_{iu}} \cdot \left(\frac{1}{1 + \exp(\xi(u) - \beta_i)} \right)^{1 - y_{iu}}$$

Q = Zahl der Antwortmuster

$\underline{\mathbf{Y}}_q$ = bestimmtes Antwortmuster

$\hat{P}(\underline{\mathbf{Y}}_q)$ = geschätzte Wahrscheinlichkeit
des Antwortmusters

Testbarkeit im Rasch-Modell

- **Was ist der Likelihood Ratio?**

Ein Quotient zweier Likelihoods: $LR = \frac{L_0(Y)}{L_1(Y)}$

- **Inwiefern kann der Likelihood Ratio zur inferenzstatistischen Prüfung eines Messmodells verwendet werden?**

Die Verteilung des logarithmierten und mit -2 multiplizierten LR ist bekannt!

χ^2 -Verteilung!

Testbarkeit im Rasch-Modell

- **Welche allgemeine Nullhypothese wird bei der Prüfung von IRT-Messmodellen (z. B. Rasch-Modell) mittels dem *LR*-Test geprüft:**

H_0 : Die wahren und die modellimplizierten Antwortmusterwahrscheinlichkeiten sind gleich!

Bootstrap-basierte Modellgeltungskontrolle

- **Wann und warum sollte der bootstrap-basierte Modellgeltungstest angewendet werden?**

Wenn die Zahl der möglichen Antwortmuster, die Zahl der beobachteten Antwortmuster in einer Stichprobe übersteigt!

Wenn die Zahl der möglichen und beobachteten Antwortmuster differiert, ist die Verteilung der Prüfgröße unbekannt.

→ Berechnung des p -Wertes *nicht* möglich!

- **Was ist das Grundprinzip der Bootstrap-Inferenz?**

Die Verteilung der Prüfgröße wird empirisch ermittelt!

Vergleich der Prüfgröße aus der Originalstichprobe mit der empirisch ermittelten Verteilung der Prüfgröße aus den Bootstrap-samples.