

# Übung zur Vorlesung - Theorien Psychometrischer Tests II

N. Rose

8. Übung (08.01.2008)



- **Agenda**
  - Vergleich Rasch-Modell vs. 2-parametrisches logistisches Modell nach Birnbaum
  - 2PL-Modelle in Mplus

# Vergleich 1PL- vs. 2PL-Modell

- Welche Modelle der klassischen Testtheorie entsprechen dem 1PL- bzw. 2PL-Modell auf Logitebene?

**A:** Auf Logitebene kann das 1PL-Modell als Modell essentiell  $\tau$ -äquivalenter Variablen, das 2PL-Modell als Modell  $\tau$ -kongenerischer Variablen betrachtet werden.

- Welche Bedeutung hat der Summenscore im 1PL- bzw. 2PL-Modell?

**A:** Im Raschmodell ist der Summenscore eine suffiziente Statistik bzgl. der latenten Personenvariable. Dies gilt nicht im 2PL-Modell!

**Folge:** Im 2PL-Modell ist nicht nur wichtig wie viele Items, sondern auch welche Items gelöst werden (Antwortmuster wichtig!).

# Vergleich 1PL- vs. 2PL-Modell

- Wie drückt sich die **Insuffizienz** des Summenscores bzgl. der zu schätzenden latenten Personenvariable, im 2PL-Modell in der Antwortmuster-Likelihood aus?

**A:** Die Antwortmuster-Likelihoods verschiedener Antwortmuster bei gleichem Summenscore, haben ihr Maxima bei verschiedenen Werten von  $\xi$ !

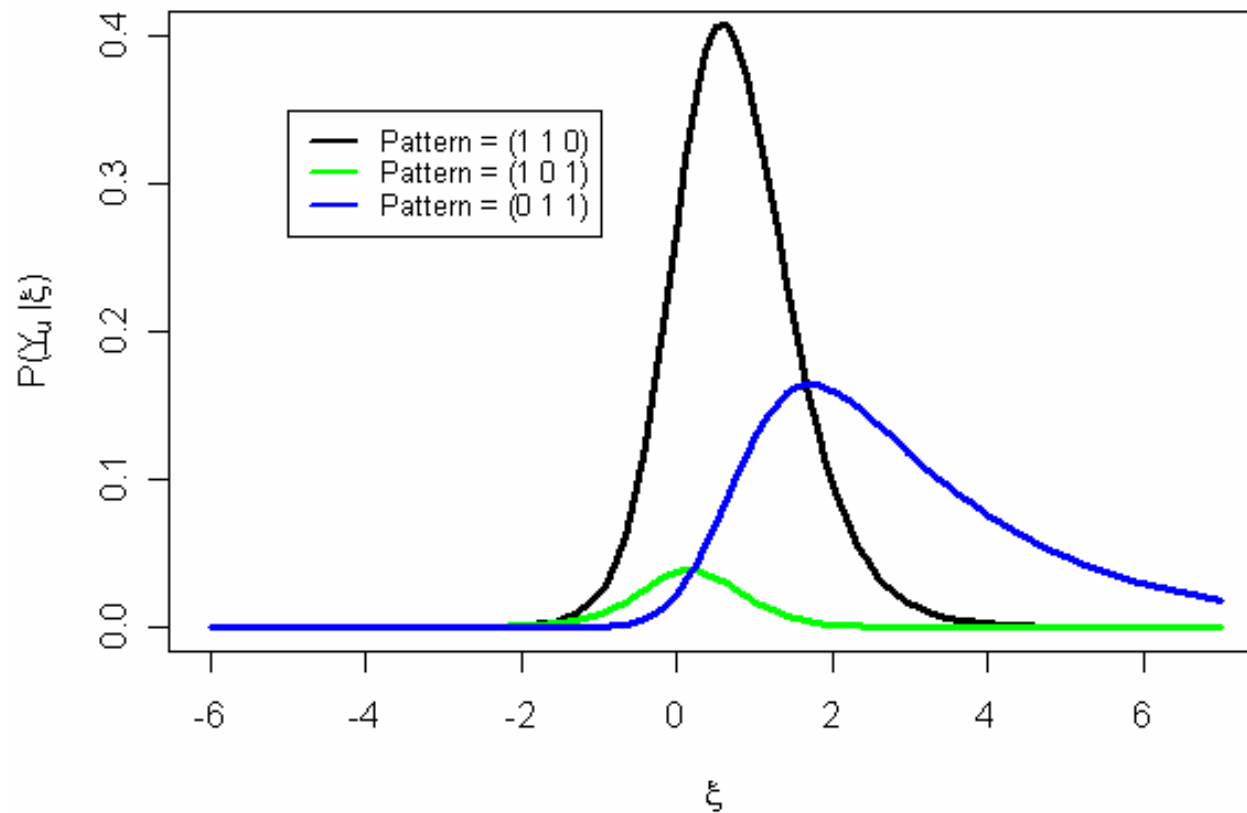
**Folge:** Unterschiedliche Parameterschätzungen für  $\xi$  bei gleichem Summenscore aber unterschiedlichen Antwortmuster.

**Tip 1:** Lernen mit Rabix 0.99!

**Tip 2:** Graphiken sprachlich formal möglichst exakt beschreiben!

# Vergleich 1PL- vs. 2PL-Modell

**Antwortmusterwahrscheinlichkeiten -2PL-Modell**  
**3 Items & Summenscore = 2**



3 Items:

$$\alpha_i = 0.5, \beta_1 = -1$$

$$\alpha_2 = 3, \beta_2 = 0$$

$$\alpha_3 = 2, \beta_3 = 1$$

# Vergleich 1PL- vs. 2PL-Modell

- Worin unterscheiden sich die Iteminformationsfunktionen des 1PL- bzw. 2PL-Modells?

**A:** Im 1PL-Modell ist die Iteminformationsfunktion  $I(Y_i | \xi)$  gleich der bedingten Varianzfunktion  $Var(Y_i | \xi)$ ! Dies gilt nicht im 2PL-Modell.

**Erinnerung:** Iteminformationsfunktion im 2PL-Modell

$$\begin{aligned} I(Y_i | \xi) &= \alpha_i^2 \cdot Var(Y_i | \xi) \\ &= \alpha_i^2 \cdot P(Y_i = 1 | \xi) \cdot P(Y_i = 0 | \xi) \end{aligned}$$

**A:** Im 1PL-Modell ist die Iteminformationsfunktion  $I(Y_i | \xi)$  und folglich die Standardfehler nur von der Itemschwierigkeit abhängig, im 2PL-Modell zusätzlich von der Itemdiskrimination  $\alpha_i$ !

# Vergleich 1PL- vs. 2PL-Modell

- Was ist der Vorteil, und was der Nachteil bei Items mit hohen Itemdiskriminationen  $\alpha_i$  im 2PL-Modells?

**Vorteil:** Hohe Werte von  $\alpha_i$  bedeuten hohe Werte bzgl. der Iteminformationsfunktion sofern  $\xi$  nahe  $\beta_i$  liegt. Dies bedeutet niedrige Standardfehler.

**Nachteil:** Der oben genannte Vorteil ist mit größer werdenden Werten von  $\alpha_i$  auf einen zunehmend kleineren Bereich von  $\xi$  beschränkt!

**Nachteil:** hohe Messgenauigkeit in einem kleinen Bereich von  $\xi$ !

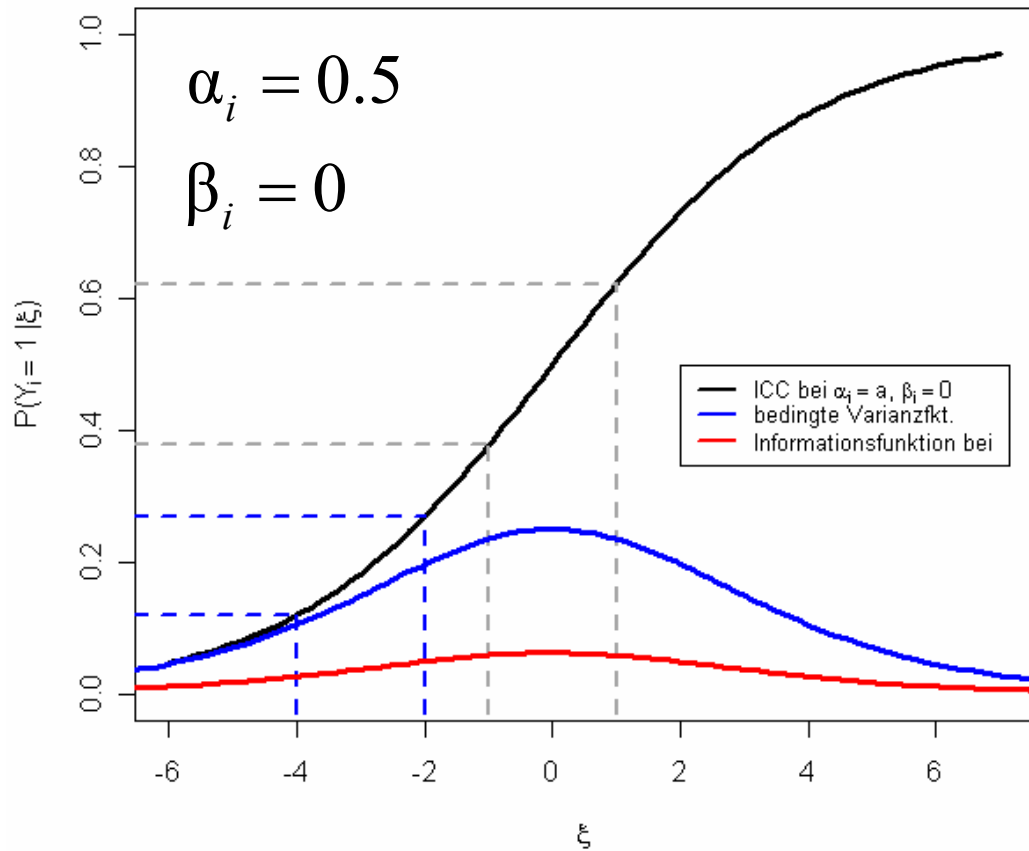
# Vergleich 1PL- vs. 2PL-Modell

ICC

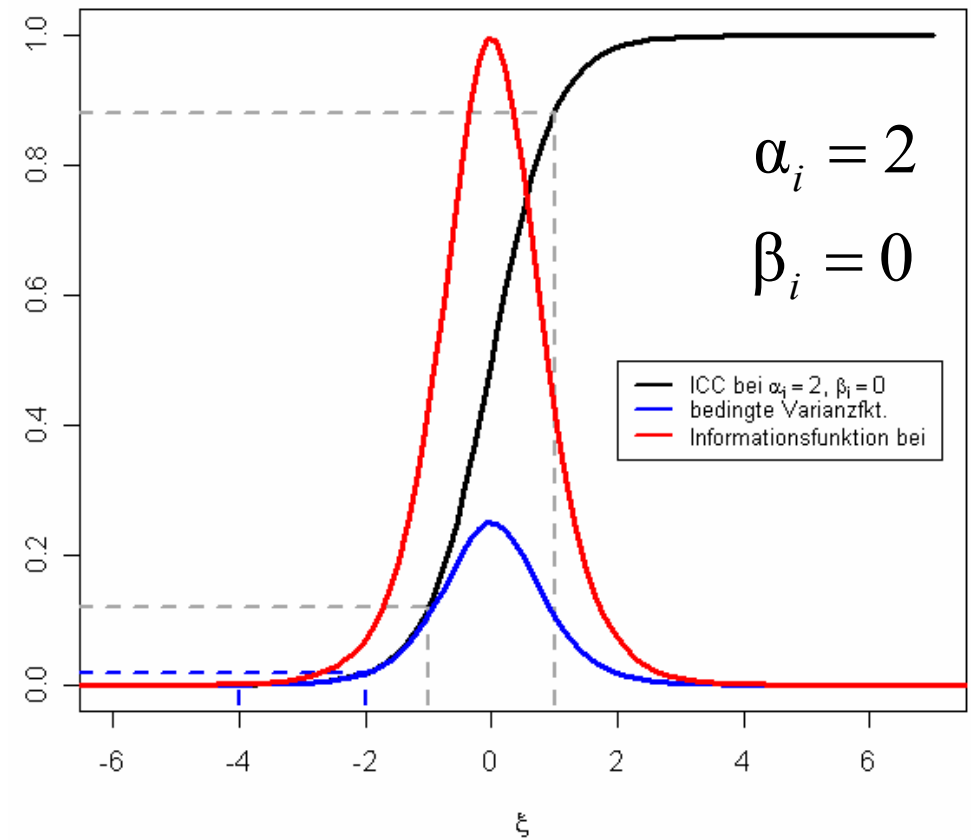
$Var(Y_i | \xi)$

$I(Y_i | \xi)$

ICC - 2PL-Modell n. Birnbaum



ICC - 2PL-Modell n. Birnbaum





# Vergleich 1PL- vs. 2PL-Modell



# 2PL-Modell nach Birnbaum - Mplus

- Modellspezifikation in Mplus:

```
Mplus_16items_2p.inp
TITLE:      16 Items 1 latente Dimension 2PL;
DATA:      FILE IS 16items_2pl_1latent.dat;
           TYPE IS INDIVIDUAL;
VARIABLE:  NAMES ARE i1-i16;
           USEVARIABLES ARE i1-i16;
           CATEGORICAL ARE i1-i16;
ANALYSIS:  Estimator=MLR;
MODEL:     XI BY i1-i16*;
           XI@1;
           [XI@0];
OUTPUT:    TECH1;
SAVE:     FILE IS person_est.dat;
           SAVE=FSCORES;
```

Logistisches IRT-Modell

BY – Statement für  
Itemdiskriminationen

Speichern der Personenparameter  
(EAP)

# 2-Parametrische Modelle in Mplus

- Output

```
MODEL RESULTS
```

|            |        | Estimates | S.E.  | Est./S.E. |
|------------|--------|-----------|-------|-----------|
| XI         | BY     |           |       |           |
|            | I1     | 2.194     | 0.284 | 7.728     |
|            | I2     | 2.008     | 0.238 | 8.442     |
|            | I3     | 0.520     | 0.100 | 5.176     |
|            | ...    | ...       | ...   | ...       |
|            | I16    | 1.143     | 0.159 | 7.184     |
| Means      |        |           |       |           |
|            | XI     | 0.000     | 0.000 | 0.000     |
| Thresholds |        |           |       |           |
|            | I1\$1  | -2.489    | 0.246 | -10.117   |
|            | I2\$1  | -2.021    | 0.193 | -10.448   |
|            | I3\$1  | -0.533    | 0.083 | -6.388    |
|            | ...    | ...       | ...   | ...       |
|            | I16\$1 | 1.595     | 0.132 | 12.095    |
| Variances  |        |           |       |           |
|            | XI     | 1.000     | 0.000 | 0.000     |

Itemdiskriminationsparameter  
in der logistischen Metrik

Transformierte Itemschwierig-  
keitsparameter

# 2-Parametrische Modelle in Mplus

- Output (Fortsetzung):

```
IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC
WHERE THE LOGIT IS 1.7*DISCRIMINATION*(THETA - DIFFICULTY)
```

| Item Discriminations |    |       |       |       |
|----------------------|----|-------|-------|-------|
| XI                   | BY |       |       |       |
| I1                   |    | 1.291 | 0.167 | 7.728 |
| I2                   |    | 1.181 | 0.140 | 8.442 |
| ...                  |    | ...   | ...   | ...   |
| I16                  |    | 0.672 | 0.094 | 7.184 |

| Means |  |       |       |       |
|-------|--|-------|-------|-------|
| XI    |  |       |       |       |
| XI    |  | 0.000 | 0.000 | 0.000 |

| Item Difficulties |  |        |       |         |
|-------------------|--|--------|-------|---------|
|                   |  |        |       |         |
| I1\$1             |  | -1.134 | 0.092 | -12.279 |
| I2\$1             |  | -1.007 | 0.088 | -11.426 |
| ...               |  | ...    | ...   | ...     |
| I16\$1            |  | 1.395  | 0.163 | 8.558   |

Itemdiskriminationsparameter in der Metrik des Normal-Ogiven-Modells (sog. Probit-Modelle)

Itemschwierigkeitsparameter in der Metrik der latenten Variable (Logit UND Probit)

# 2-Parametrische Modelle in Mplus

- **Was ist Was im Mplus-Output?**

→ ausgehend von der Grundgleichung:

$$P(Y_i = 1 | \xi) = \frac{\exp[\alpha_i \cdot (\xi - \beta_i)]}{1 + \exp[\alpha_i \cdot (\xi - \beta_i)]}$$

$\beta_i$  = **Item Difficulties in Mplus!**

$\alpha_i$  = **Estimates in Mplus!**

→ kann transformiert werden:

```
IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC  
WHERE THE LOGIT IS 1.7*DISCRIMINATION*(THETA - DIFFICULTY)
```

$$P(Y_i = 1 | \xi) = \frac{\exp[1.7\alpha_i \cdot (\xi - \beta_i)]}{1 + \exp[1.7\alpha_i \cdot (\xi - \beta_i)]}$$

## 2-Parametrische Modelle in Mplus

- es folgt: 
$$P(Y_i = 1 | \xi) = \frac{\exp[1.7\alpha_i \cdot (\xi - \beta_i)]}{1 + \exp[1.7\alpha_i \cdot (\xi - \beta_i)]}$$
$$= \frac{\exp[1.7\alpha_i\xi - 1.7\alpha_i\beta_i]}{1 + \exp[1.7\alpha_i\xi - 1.7\alpha_i\beta_i]}$$

$$\alpha_i^* = 1.7\alpha_i = \text{Item Discrimination in Mplus!}$$

$$\beta_i^* = 1.7\alpha_i\beta_i = \text{Threshold in Mplus!}$$

**Ausblick:** Der transformierte Itemdiskriminationsparameter  $\alpha_i^* = 1.7\alpha_i$  entspricht approximativ dem Itemdiskriminationsparameter unter Verwendung der kumulierten Standardnormalverteilung als Linkfunktion!

## 2-Parametrische Modelle in Mplus

- In Mplus wird die so genannte Marginal Maximum Likelihood (*MML*) Schätzung verwendet. Welche Annahme wird in Mplus bzgl. der latenten Variable  $\xi$  gemacht?

**A:** Bei der *MML*-Schätzung werden nicht die Personenparameter für jede Beobachtung berechnet, sondern nur die Parameter welche die Verteilung von  $\xi$  beschreiben.

**Nachteil:** Es muss eine Verteilungsannahme gemacht werden!

**A:** Im Mplus wird angenommen, dass  $\xi$  normalverteilt ist.

## 2-Parametrische Modelle in Mplus

- Welche Personenparameterschätzungen können in *Mplus* berechnet werden, welche Besonderheit haben diese.

```
OUTPUT:      TECH1;  
SAVE:        FILE IS person_est.dat;  
             SAVE=FSCORES;
```

**A:** Es sind so genannte Expected A Posteriori (*EAP*) Schätzer.

**A:** Obwohl es sich um fehlerbehaftete Parameterschätzer handelt ist ihre Varianz geringer als die Varianz der wahren Personenvariable!