Lecture
„Item Response Theory“

3rd Lecture

6. November 2016

Dichotomous manifest variables (items)

- Binary or dichotomous items are often used, in particular in the measurement of abilities
(e.g. PISA)
- the manifest random variables Y_i, $i = 1, \dots, m$, (in this context, also called *items*) are defined as follows

$$Y_i(\omega) = \begin{cases} 0, & \text{if item } i \text{ is not solved} \\ 1, & \text{if item } i \text{ is solved} \end{cases}$$

Data matrix $\mathbf{Y}$ for dichotomous variables

Persons	y_1	y_2	y_3	...	y_m
u_1	0	1	0	...	0
u_2	1	1	1	...	0
u_3	0	1	0	...	0
u_4	1	1	1	...	1
...	...	...	...	...	...
...	...	...	...	...	...
u_N	0	0	1	...	1

Rows of the data matrix are also called *answer patterns* or *response patterns*

Row sums are the sum scores/test scores of the persons

Column sums are the absolute frequencies of answers in category 1 for the items $Y_1 \rightarrow Y_m$

Questions

Are correlations appropriate for quantifying the strength of the dependencies between ...

... dichotomous variables?

... metric and dichotomous variables?

Correlation and probabilities to solve an item

		Y_2		
		1	0	
Y_1	1	.90	0	.90
	0	0	.10	.10
	gesamt	.90	.10	1.00

Pearson correlation $\rho = 1$

		Y_2		
		1	0	
Y_1	1	.50	0	.50
	0	.40	.10	.50
	total	.90	.10	1.00

Pearson correlation $\rho = 0.\overline{3}$

		Y_2		
		1	0	
Y_1	1	.50	0	.50
	0	.45	.5	.50
	total	.95	.5	1.00

Pearson correlation $\rho = 0.229$

- True score variable of a dichotomous item Y_i:

$$E(Y_i \mid U) = P(Y_i = 1 \mid U). \tag{1}$$

- Hence, for a dichotomous item Y_i, the conditional expectation $E(Y_i \mid U)$ (i.e., the true score variable) is identical to the U-conditional probability $P(Y_i = 1 \mid U)$ of the event $\{Y_i = 1\}$ of solving the item.

If $Y_i : \Omega \to \mathbb{R}$ is dichotomous with values 0 and 1, then the co-domain of the conditional expectation $E(Y_i \mid U)$ is the interval $[0, 1]$

Rules for logarithm and exponential function

1. Definitions: $x = log_b a \Leftrightarrow b^x = a, \quad x = ln\, a = log_e a \Leftrightarrow e^x = a$, where $e \approx 2.718281\ldots$
2. $\log_b(x_1 \cdot x_2) = \log_b(x_1) + \log_b(x_2), \quad$ where $b, x_1, x_2 > 0$ and $b \neq 1$
3. $\log_b(1) = 0$
4. $\log_b(b) = 1$
5. $\log_b\left(\frac{x_1}{x_2}\right) = \log_b(x_1) - \log_b(x_2)$
6. $\log_b(x^a) = a \cdot \log_b(x)$

1. Definition: $exp(x) = e^x$
2. $\exp(\ln x) = x, \quad$ where $x > 0$
3. $\exp(0) = 1$
4. $\exp(x_1 + x_2) = \exp(x_1) \cdot \exp(x_2)$
5. $\exp(x_1 - x_2) = \frac{\exp(x_1)}{\exp(x_2)}$
6. $\exp(a \cdot x) = \exp(x)^a$

Relationship between logit and probability

We presume $0 < P(Y_i = 1 \mid U) < 1$ and define: $logit_i := \ln\left(\frac{P(Y_i = 1 \mid U)}{1 - P(Y_i = 1 \mid U)}\right)$ (2)

The following equations are equivalent to this definition

$$\exp(logit_i) = \frac{P(Y_i = 1 \mid U)}{1 - P(Y_i = 1 \mid U)} \tag{3}$$

$$\exp(logit_i) \cdot (1 - P(Y_i = 1 \mid U)) = P(Y_i = 1 \mid U) \tag{4}$$

$$\exp(logit_i) - \exp(logit_i) \cdot P(Y_i = 1 \mid U) = P(Y_i = 1 \mid U) \tag{5}$$

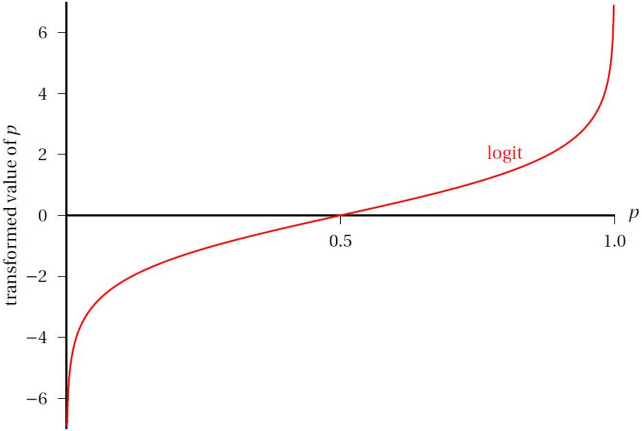
$$\frac{\exp(logit_i)}{P(Y_i = 1 \mid U)} - \exp(logit_i) = 1 \tag{6}$$

$$\frac{\exp(logit_i)}{P(Y_i = 1 \mid U)} = 1 + \exp(logit_i) \tag{7}$$

$$\frac{P(Y_i = 1 \mid U)}{\exp(logit_i)} = \frac{1}{1 + \exp(logit_i)} \tag{8}$$

$$P(Y_i = 1 \mid U) = \frac{\exp(logit_i)}{1 + \exp(logit_i)} \tag{9}$$

Relationship between logit and probability



Assumptions defining the des Rasch model

1 „Rasch homogeneity“

$$logit_i = logit_j - \beta_{ij}, \quad \forall i, j = 1, \ldots, m, \quad \beta_{ij} \in \mathbb{R}. \tag{10}$$

The logits are parallel to each other, that is, each logit is a translation of each other logit. In other words, two logits differ only by an additive constant, say β_{ij}.

2 Local (i.e., U-conditional) stochastic independence

$$\forall\, i = 1, \ldots, m: \; P(Y_i = 1 \mid U, Y_1, Y_2, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) = P(Y_i = 1 \mid U) \tag{11}$$

The probabilities to solve an item only depends on the person variable U, not on the answers to the other items.

The following proposition is equivalent to (11):

$$P(Y_{i_1} = 1, \ldots, Y_{i_k} = 1 \mid U) = P(Y_{i_1} = 1 \mid U) \cdot \ldots \cdot P(Y_{i_k} = 1 \mid U), \tag{12}$$

$$\forall \{i_1, \ldots, i_k\} \subset I = \{1, \ldots, m\}. \tag{13}$$

A proposition that is equivalent to Rasch homogeneity

Rasch homogeneit:

$$logit_i = logit_j - \beta_{ij}, \quad \forall\, i, j = 1, ..., m. \tag{14}$$

Definition of the latent variable ξ:

$$\xi := logit_1 = \ln \frac{P(Y_1 = 1 \mid U)}{1 - P(Y_1 = 1 \mid U)}. \tag{15}$$

Definition of the difficulties β_i: For $j = 1$, Equations (14) and (15) imply:

$$logit_i = \xi - \beta_i, \quad \text{where } \beta_i := \beta_{i1}, \quad \forall\, i = 1, \ldots, m. \tag{16}$$

Equations (9), (14), (15) and (16) imply the model equation:

$$P(Y_i = 1 \mid U) = \frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)}, \quad \forall\, i = 1, \ldots, m. \tag{17}$$

This equation is equivalent to (14). Hence, the probability to solve item Y_i given the person variable U only depends on the difference between the ability variable ξ and the item difficulty β_i.

Proof of $P(Y_i = 1 \mid \xi) \; = \; P(Y_i = 1 \mid U)$:

$$\begin{aligned}
P(Y_i = 1 \mid \xi) &= E(Y_i \mid \xi) \\
&= E\big(E(Y_i \mid U) \,\big|\, \xi\big) \qquad \left[\xi = \ln \frac{P(Y_1 = 1 \mid U)}{1 - P(Y_1 = 1 \mid U)}\right] \\
&= E\big(P(Y_i = 1 \mid U) \,\big|\, \xi\big) \\
&= E\left(\frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)} \,\middle|\, \xi\right) \\
&= \frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)}.
\end{aligned} \tag{18}$$

Rasch model

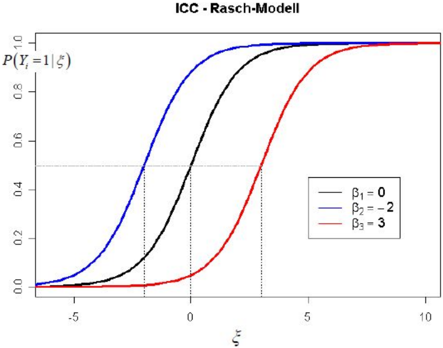
- **Uniqueness**
  The item parameter β_i and the latent variable ξ are not uniquely defined by the definition of the Rasch model, because:

$$\xi - \beta_i = (\xi + \gamma) - (\beta_i + \gamma) = \xi^* - \beta_i^*, \quad \gamma \in \mathbb{R}.$$

- **admissible transformations** are translations
  The item parameter β_i and the latent variable ξ are **difference scaled.**
- **meaningful propositions**:
  - about differences between values of the latent variable ξ
  - about differences between item parameters β_i
  - about the variance $Var(\xi)$.
- **Fixing the scale of ξ and the difficulties β_i:**
  - Fixing an item difficulty, for example, $\beta_1 = 0$ (this is done in Eq. (16))
  - Norming the sum of the item parameters, for example, $\sum_{i=1}^{m} \beta_i = 0$
  - Fixing the expectation of the latent variable, for example, $E(\xi) = 0$.

Rasch-Modell

- graphical presentation of the item characteristics

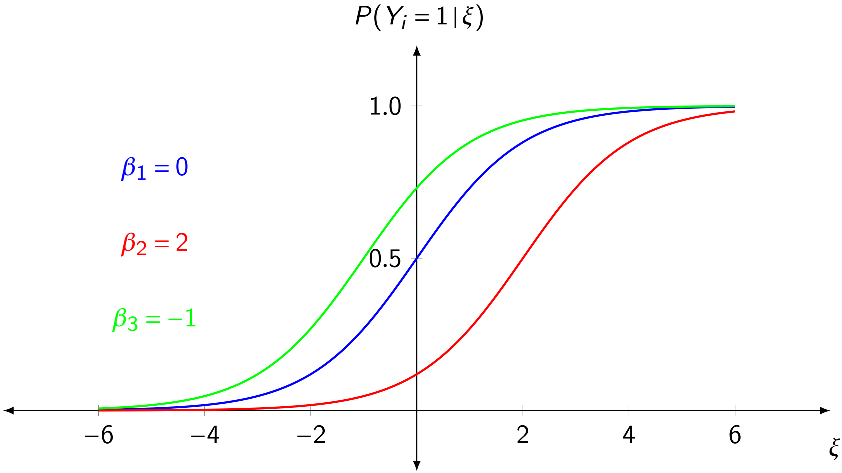


$$P(Y_1 = 1 \mid \xi) = \frac{\exp(\xi - 0)}{1 + \exp(\xi - 0)}$$

$$P(Y_2 = 1 \mid \xi) = \frac{\exp(\xi - (-2))}{1 + \exp(\xi - (-2))}$$

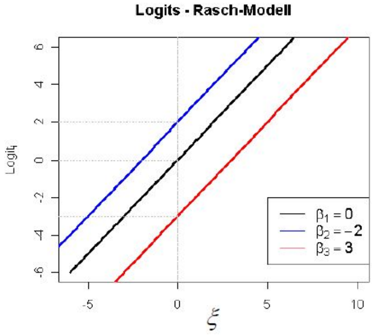
$$P(Y_3 = 1 \mid \xi) = \frac{\exp(\xi - 3)}{1 + \exp(\xi - 3)}$$

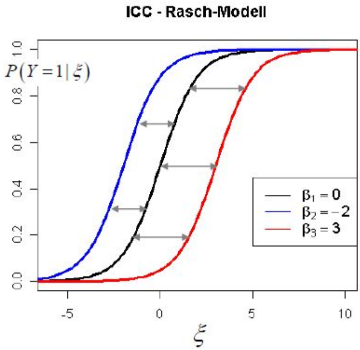
Logistic functions



Rasch model

- Logits **and** item characteristics are parallel to each other!





The Rasch model

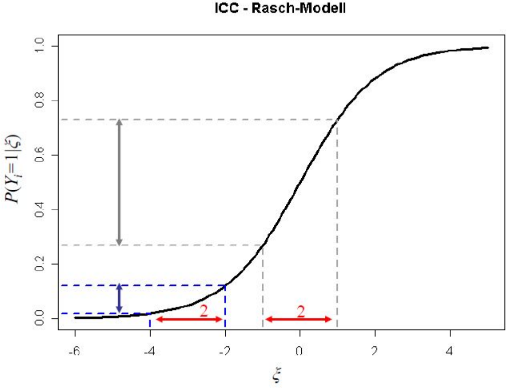
1. Personen parameters
   - are the values of he latent variable ξ

$$\xi = \ln\left(\frac{P(Y_1 = 1 \mid U)}{1 - P(Y_1 = 1 \mid U)}\right) \quad \text{i.e. } \xi = f(U)$$

2. Item difficulty β_i
   - is the value of ξ for which the probability to solve item Y_i is .5, that is, $P(Y_i = 1 \mid \xi = \beta_i) = .5$
   - Hence, the item difficulty β_i describes the location of the item Y_i on the scale of ξ

The Rasch model

- **Item difficulty** β_i **and** ***ICC***



The **discrimination** of an item has its maximum at $\xi = \beta_i$

The Rasch model

- Variance of the item Y_i:

$$Var(Y_i) = P(Y_i = 1) \cdot [1 - P(Y_i = 1)]$$
$$= P(Y_i = 1) \cdot P(Y_i = 0)$$

- In general:

$$Var(Y_i \mid \xi) = E\left([Y_i - E(Y_i \mid \xi)]^2 \mid \xi\right) = E(\epsilon_i^2 \mid \xi)$$

- *xi*-conditional variance of Y_i:

$$Var(Y_i \mid \xi) = P(Y_i = 1 \mid \xi) \cdot [1 - P(Y_i = 1 \mid \xi)]$$
$$= P(Y_i = 1 \mid \xi) \cdot P(Y_i = 0 \mid \xi)$$

- in the Rasch model:

$$Var(Y_i \mid \xi) = I_{Y_i}(\xi) \equiv \frac{[P_i'(\xi)]^2}{P_i(\xi) \cdot [1 - P_i(\xi)]},$$

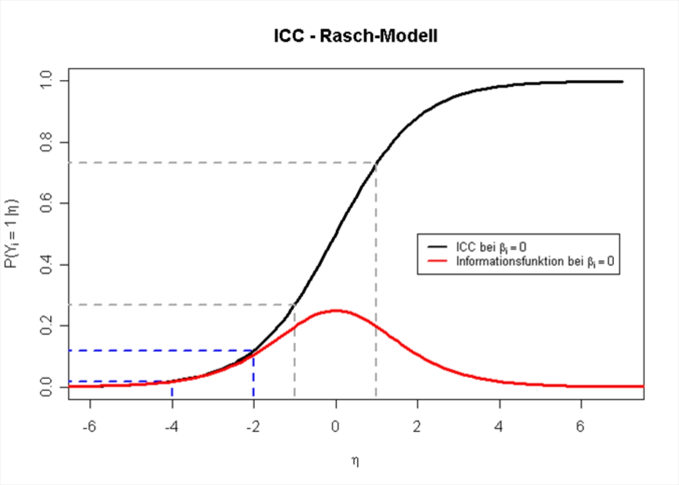
  where $I_{Y_i}(\xi)$ is the item information function, and $P_i'(\xi)$ the first derivative of $P_i(\xi) \equiv \frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)}$ to ξ.

The Rasch model

- The ξ-conditional variance $Var(Y_i \mid \xi)$, that is, the item information function $I_{Y_i}(\xi)$, quantify the degree of insecurity with respect to the value of (answer to) the item.

- The item information function $I_{Y_i}(\xi)$ quantifies the amount of information about the value of ξ that is provided by a value of the item Y_i.

The Rasch model



Identification of the item difficulties

The two assumptions of the Rasch model imply:

$$\beta_i - \beta_j = \ln\left(\frac{P(Y_i = 0, Y_j = 1)}{P(Y_i = 1, Y_j = 0)}\right). \tag{19}$$

Fixing the scale of ξ by $\beta_1 = 0$, the equation implies

$$\beta_i = \ln\left(\frac{P(Y_i = 0, Y_1 = 1)}{P(Y_i = 1, Y_1 = 0)}\right). \tag{20}$$

Precision of the ability estimation

Test information function

$$I_{Y_1,\ldots,Y_m}(\xi) = \sum_{i=1}^{m} I_{Y_i}(\xi) \tag{21}$$

Estimated standard error of the Person parameter (for a large number m of items)

$$SE_{\widehat{\xi}}(\xi) \approx \sqrt{\frac{1}{I_{Y_1,\ldots,Y_m}(\xi)}} \tag{22}$$

Confidence interval of the estimator of the person parameter

$$\widehat{\xi} \pm z_{\alpha/2} \cdot SE_{\widehat{\xi}}(\xi) \tag{23}$$

where $\widehat{\xi}$ is an estimator of ξ.

Reliability of the person parameter estimator

$$Rel(\widehat{\xi}) = \frac{Var[E(\widehat{\xi} \mid U)]}{Var(\widehat{\xi})} = 1 - \frac{Var(\varepsilon)}{Var(\widehat{\xi})}, \tag{24}$$

where

$$\varepsilon := \widehat{\xi} - E(\widehat{\xi} \mid U). \tag{25}$$

Remember:

$$\begin{aligned} Var(\varepsilon) = E(\varepsilon^2) = E\big(E(\varepsilon^2 \mid \xi)\big) &= E\big(Var(\widehat{\xi} \mid \xi)\big) \\ &= E\big(Var(\varepsilon \mid \xi)\big) = E\big(SD(\varepsilon \mid \xi)^2\big). \end{aligned} \tag{26}$$

$SD(\varepsilon \mid \xi)$ is estimated by $SE_{\widehat{\xi}}(\xi)$. Hence,

$$\widehat{Var}(\varepsilon) = \widehat{E}[SE_{\widehat{\xi}}(\xi)^2]. \tag{27}$$