

2 Item-Response-Modellierung

Grundlage für die der Arbeit vorliegenden Modelle sind die Konzepte von Georg Rasch zur Auswertung von Intelligenztests [Rasch, 1960]. Seine initiale Idee ist, dass das Ergebnis eines Intelligenztests einer Person nur von zwei Komponenten abhängt, nämlich einem Faktor für die Fähigkeit der Person und einem Faktor für die Schwierigkeit des Tests. In dieser Arbeit wird jeweils die Schwierigkeit einzelner Testitems betrachtet.

Beide von Rasch identifizierten Komponenten werden durch latente Parameter ausgedrückt, die nur relativ zu einem festgelegten Referenzwert interpretiert werden können. Im Kapitel „A structural model for items of a test“ stellt Rasch [Rasch, 1960] Überlegungen an, wie man die Messungen für die Personen-Fähigkeit und die Item-Schwierigkeit auf einer Verhältnisskala ausdrücken kann.

2.1 Das klassische Rasch-Modell

Sei ξ der Parameter für die Fähigkeit der Person und λ der Parameter für die Schwierigkeit des Items eines Tests, so gelte für zwei Personen und zwei Items folgende Relation:

$$\left. \begin{array}{l} \xi_1 = \nu \xi_2 \\ \lambda_1 = \nu \lambda_2 \end{array} \right\} \Rightarrow \frac{\xi_1}{\lambda_1} = \frac{\xi_2}{\lambda_2} \quad (2.1)$$

Inhaltlich bedeutet Gleichung (2.1), dass sich sowohl die Fähigkeit von Person 1 und Person 2 als auch die Schwierigkeit von Item 1 und Item 2 um den Faktor ν unterscheidet. Die Wahrscheinlichkeit, dass Person 1 Item 1 löst, ist also genauso groß wie die Wahrscheinlichkeit, dass Person 2 Item 2 löst [Rasch, 1960].

Damit obige Aussage als gültig angesehen werden kann, sollten die Relationen (2.1) auch auf alle weiteren Items und alle anderen Personen übertrag-

bar sein. Sei $\nu > 1$, dann sollte Person 1 bei allen Items um den Faktor ν besser abschneiden als Person 2, und Item 1 sollte für alle Personen um den Faktor ν schwieriger sein als Item 2. Es ist somit sinnvoll, die Wahrscheinlichkeit, dass eine Person ein Item korrekt löst, als Funktion des Verhältnisses $\zeta = \frac{\xi}{\lambda}$ zu modellieren [Rasch, 1960]. Um Personen jeder Fähigkeit und Items jeder Schwierigkeit zu berücksichtigen, sollte der Wertebereich von ζ zwischen 0 und $+\infty$ liegen. Eine Transformation von ζ in den Wertebereich zwischen 0 und 1 ist die naheliegende Transformation [Rasch, 1960]:

$$\frac{\zeta}{1 + \zeta} = \frac{\xi}{\xi + \lambda} \in (0, 1), \text{ falls } \zeta \in (0, +\infty) \quad (2.2)$$

Für die Wahrscheinlichkeit π_{pi} , dass Person p Item i korrekt löst, gilt dann:

$$\pi_{pi} = \frac{\xi_p}{\xi_p + \lambda_i} \Leftrightarrow \log\left(\frac{\pi_{pi}}{1 - \pi_{pi}}\right) = \log(\xi_p) - \log(\lambda_i) \quad (2.3)$$

Die rechte Seite von Gleichung (2.3) entspricht dem bekannten binären Rasch-Modell. Sei $y_{pi} \in \{0, 1\}$ der Indikator, ob Person p Item i korrekt löst, so gilt für dessen Wahrscheinlichkeit [Strobl, 2010]:

$$\pi_{pi} = P(y_{pi} = 1 | \theta_p, \beta_i) = \frac{\exp(\theta_p - \beta_i)}{1 + \exp(\theta_p - \beta_i)} \text{ mit } p = 1, \dots, P, i = 1, \dots, I, \quad (2.4)$$

wobei θ_p für den Personen-Parameter und β_i für den Item-Parameter steht. Eine alternative Formulierung des Modells lautet:

$$\text{logit}(P(y_{pi} = 1 | \theta_p, \beta_i)) = \log\left(\frac{P(y_{pi} = 1 | \theta_p, \beta_i)}{1 - P(y_{pi} = 1 | \theta_p, \beta_i)}\right) = \theta_p - \beta_i \quad (2.5)$$

Mit $\theta_p = \log(\xi_p)$ und $\beta_i = \log(\lambda_i)$ entsprechen sich die Gleichungen (2.3) und (2.5).

Wie einleitend angedeutet wurde, sind die zu schätzenden Personen- und Itemparameter nur relativ interpretierbar. Modell (2.5) ist in dieser Form nicht eindeutig lösbar. Es ist notwendig, vor der Schätzung einen Parameter festzusetzen, der als Referenzwert dient. Gewählt wird der Personen-Parameter $\theta_p = 0$. Dies macht eine einfache Darstellung des Modells in Abschnitt 2.4 möglich [Tutz und Schauburger, 2013].

2.2 Rasch-Modell mit itemmodifizierenden Effekten

Im Rasch-Modell (2.5) wird die Schwierigkeit von Item i allein durch den Parameter β_i modelliert. Im Folgenden wird zusätzlich der Effekt des Differential Item Functioning berücksichtigt. Um zuzulassen, dass die Schwierigkeit bestimmter Items von Kovariablen abhängen kann, wird Modell (2.5) um itemmodifizierende Effekte γ_i erweitert.

Sei \mathbf{x}_p der Vektor an Kovariablen von Person p , so wird der Item-Parameter β_i um den linearen Prädiktor $\mathbf{x}_p^T \gamma_i$ ergänzt. Es gilt zu beachten, dass \mathbf{x}_p einen personenspezifischen und γ_i einen itemspezifischen Parameter darstellt [Tutz und Schauburger, 2013].

Das vollständige Modell in äquivalenter Form zu (2.5) lautet:

$$\begin{aligned} \text{logit}(P(y_{pi} = 1 | \theta_p, \beta_i, \mathbf{x}_p)) &= \log \left(\frac{P(y_{pi} = 1 | \theta_p, \beta_i, \mathbf{x}_p)}{1 - P(y_{pi} = 1 | \theta_p, \beta_i, \mathbf{x}_p)} \right) = \\ &= \theta_p - (\beta_i + \mathbf{x}_p^T \gamma_i) \end{aligned} \tag{2.6}$$

Mit Modell (2.6) ist es möglich, Unterschiede in der Beantwortung einzelner Items eines Tests zwischen Subgruppen zu modellieren, die durch die Kovariablen \mathbf{x} gebildet werden. Im einfachsten Fall ist x_p die Realisierung einer binären Variable, z. B. Geschlecht. Sei $x_p = 1$ für eine männliche Person und $x_p = 0$ für eine weibliche Person. Für den Fall, dass ein Unterschied

zwischen Männern und Frauen besteht, erhält man als Item-Parameter

$$\begin{aligned} \beta_i + \gamma_i & \text{ für Männer und} \\ \beta_i & \text{ für Frauen.} \end{aligned} \tag{2.7}$$

Der Parameter γ_i entspricht in diesem Beispiel dem Unterschied der Schwierigkeit von Item i zwischen Männern und Frauen [Tutz und Schauberger, 2013].

Die Stärke von Modell (2.6) ist, dass im Allgemeinen auch metrische oder mehrkategoriale Kovariablen x_p ins Modell aufgenommen werden können und es weiterhin seine Gültigkeit behält. Des Weiteren ist die Anzahl an Kovariablen des Modells beliebig wählbar, ohne, dass Modell (2.6) an Gültigkeit verliert. Nimmt man Linearität in den Logits an, so lautet der Item-Parameter für die stetige Kovariable Alter:

$$\beta_i + \text{Alter} \cdot \gamma_i \tag{2.8}$$

Falls γ_i gleich Null ist, so ist die Schwierigkeit von Item i in jedem Alter gleich [Tutz und Schauberger, 2013].

Im Weiteren bezeichne Q die Anzahl an Kovariablen im Modell, mit $q = 1, \dots, Q$. Ist ein Parameter γ_{iq} ungleich Null, bedeutet es, dass das Item von den Personen der Gruppen, die durch die Kovariable q gebildet werden, unterschiedlich beantwortet wird. Modell (2.6) gibt also nicht nur an, welche Items dem klassischen Rasch-Modell (2.5) nicht genügen, sondern auch explizit, durch welche Kovariablen die Item-Parameter β_i modifiziert werden [Tutz und Schauberger, 2013].

2.3 Das logistische Regressionsmodell

Ein gängiges, statistisches Modell für die Modellierung einer binären Zufallsvariable in Abhängigkeit anderer Einflussgrößen ist das logistische Regressionsmodell. Eine ausführliche Darstellung der Theorie zu generalisierten Regressionsmodellen findet man in [Fahrmeir et al., 2009]. Gegeben seien die Daten (y_i, \mathbf{x}_i) , $i = 1, \dots, n$, wobei $y_i \in \{0, 1\}$ und $E(y_i | \mathbf{x}_i) = P(y_i = 1 | \mathbf{x}_i) = \pi_i$.

Damit lautet das vollständige Modell:

1. Zufallskomponente

$$y_i | \pi_i \sim B(\pi_i)$$

2. Linearer Prädiktor

$$\eta_i = \mathbf{x}_i^T \boldsymbol{\delta}$$

3. Link-Funktion

$$\pi_i = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\delta})}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\delta})} \Rightarrow g(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \mathbf{x}_i^T \boldsymbol{\delta}$$

Im Folgenden werden die Modelle (2.5) und (2.6) in das vorgestellte Framework des logistischen Regressionsmodells eingebettet.

2.4 Rasch-Modell als logistisches Regressionsmodell

Zur Schätzung der Rasch-Modelle aus Abschnitt 2.1 und 2.2 sollen Algorithmen verwendet werden, die auf Maximum-Likelihood-Schätzungen basieren. Dazu ist es hilfreich, die Modelle (2.5) und (2.6) in die bekannte Form eines logistischen Regressionsmodells, welches im vorherigen Abschnitt 2.3 vorgestellt wurde, zu bringen. Die Darstellung der Modelle ist entnommen aus [Tutz und Schaubberger, 2013].

Gegeben seien die Daten (y_{pi}, \mathbf{x}_p) , $p = 1, \dots, P$, $i = 1, \dots, I$. Mit Wahrscheinlichkeit $\pi_{pi} = P(y_{pi} = 1 | \mathbf{z}_{pi})$ gilt für die Linkfunktion des logistischen Regressionsmodells:

$$g(\pi_{pi}) = \mathbf{z}_{pi}^T \boldsymbol{\delta}, \quad (2.9)$$

wobei \mathbf{z}_{pi} den Designvektor für Person p bzgl. Item i darstellt. Diesen gilt es klar vom Kovariablen-Vektor \mathbf{x}_p für Person p zu unterscheiden.

Mit Vektor $\boldsymbol{\delta}^T = (\boldsymbol{\theta}^T, \boldsymbol{\beta}^T)$ lässt sich Modell (2.5) folgendermaßen schreiben:

$$\begin{aligned} \log \left(\frac{P(y_{pi} = 1 | \mathbf{z}_{pi})}{1 - P(y_{pi} = 1 | \mathbf{z}_{pi})} \right) &= \theta_p - \beta_i = \mathbf{1}_{P(p)}^T \boldsymbol{\theta} - \mathbf{1}_{I(i)}^T \boldsymbol{\beta} = \\ &= \left(\mathbf{1}_{P(p)}^T, -\mathbf{1}_{I(i)}^T \right) \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\beta} \end{pmatrix} = \mathbf{z}_{pi}^T \boldsymbol{\delta} \end{aligned} \quad (2.10)$$

In Modellgleichung (2.10) gilt:

$$\begin{aligned} \mathbf{1}_{P(p)}^T &= (0, \dots, 0, 1, 0, \dots, 0), \text{ mit Länge } P-1 \text{ und } 1 \text{ an Position } p \\ \boldsymbol{\theta} &= (\theta_1, \dots, \theta_{P-1}) \\ \mathbf{1}_{I(i)}^T &= (0, \dots, 0, 1, 0, \dots, 0), \text{ mit Länge } I \text{ und } 1 \text{ an Position } i \\ \boldsymbol{\beta} &= (\beta_1, \dots, \beta_I) \end{aligned}$$

Für das Modell gilt die Restriktion $\theta_P = 0$. Nur durch Festsetzen eines Parameters ist das Modell eindeutig lösbar.

Mit Vektor $\boldsymbol{\delta}^T = (\boldsymbol{\theta}^T, \boldsymbol{\beta}^T, \boldsymbol{\gamma}_1^T, \dots, \boldsymbol{\gamma}_I^T)$ lässt sich Modell (2.6) folgendermaßen schreiben:

$$\begin{aligned} \log \left(\frac{P(y_{pi} = 1 | \mathbf{z}_{pi})}{1 - P(y_{pi} = 1 | \mathbf{z}_{pi})} \right) &= \theta_p - \beta_i - \mathbf{x}_p^T \boldsymbol{\gamma}_i = \\ &= \mathbf{1}_{P(p)}^T \boldsymbol{\theta} - \mathbf{1}_{I(i)}^T \boldsymbol{\beta} - \mathbf{x}_p^T \boldsymbol{\gamma}_i = \mathbf{z}_{pi}^T \boldsymbol{\delta} \end{aligned} \quad (2.11)$$

Der gesamte Designvektor in (2.11) mit Komponente $-\mathbf{x}_p^T$ bezüglich Parameter γ_i lautet $\mathbf{z}_{pi}^T = (\mathbf{1}_{P(p)}^T, -\mathbf{1}_{I(i)}^T, 0, \dots, -\mathbf{x}_p^T, \dots, 0)$.

Für einen kleinen Datensatz mit zwei Personen und zwei Items sehen die vollständigen Komponenten von Modell (2.11) wie folgt aus:

$$\mathbf{y} = \begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & -1 & 0 & -\mathbf{x}_1^T & 0 \\ 1 & 0 & -1 & 0 & -\mathbf{x}_1^T \\ 0 & -1 & 0 & -\mathbf{x}_2^T & 0 \\ 0 & 0 & -1 & 0 & -\mathbf{x}_2^T \end{pmatrix} \quad \text{und} \quad \boldsymbol{\delta} = \begin{pmatrix} \theta_1 \\ \beta_1 \\ \beta_2 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} \quad (2.12)$$

In dieser Darstellung hat die Matrix \mathbf{Z} genau $P \cdot I$ Zeilen, was der Anzahl an Beobachtungen und $P + 2 \cdot I - 1$ Spalten, was der Anzahl zu schätzender Parameter in Modell (2.11) entspricht. Die Parametervektoren der itemmodifizierenden Effekte $\gamma_1, \dots, \gamma_I$ stellen in dieser Form des Modells jeweils einen Parameter dar. Für die weitere Notation gilt:

- $N = P \cdot I \hat{=}$ Anzahl an Beobachtungen
- $J = P + 2 \cdot I - 1 \hat{=}$ Anzahl an Parametern in Modell (2.11)

Es ist zu beachten, dass die Anzahl zu schätzender Parameter in Modell (2.11) im Allgemeinen nicht größer ist als die Anzahl an Beobachtungen. Sie ist jedoch so groß, dass eine regularisierte Schätzung des Modells notwendig ist (siehe Kapitel 3).

2.5 Modell mit zusätzlichem Populationseffekt

In Modell (2.6) wird implizit angenommen, dass Unterschiede zwischen den Gruppen, die durch die Kovariablen \mathbf{x} gebildet werden, nur bezüglich bestimmter Items eines Test bestehen. Tatsächlich kann es jedoch sein, dass ein grundsätzlicher Fähigkeitsunterschied zwischen den Gruppen besteht.

Dies führe dazu, dass das Ergebnis der betrachteten Gruppen bezüglich des gesamten Tests unterschiedlich gut ausfällt. Dieser Effekt ist klar vom Effekt des Differential Item Functioning zu unterscheiden, welcher durch die itemmodifizierenden Effekte $\gamma_1, \dots, \gamma_I$ modelliert wird.

Um grundsätzliche Fähigkeitsunterschiede zu berücksichtigen, kann Modellgleichung (2.11) um eine weitere Kovariable γ erweitert werden [Tutz und Schaubерger, 2013].

Der lineare Prädiktor des logistischen Regressionsmodells lässt sich dann folgendermaßen schreiben:

$$\eta_{pi} = \mathbf{z}_{pi}^T \boldsymbol{\delta} = \mathbf{1}_{P(p)}^T \boldsymbol{\theta} - \mathbf{1}_{I(i)}^T \boldsymbol{\beta} - \mathbf{x}_p^T \boldsymbol{\gamma} - \mathbf{x}_p^T \boldsymbol{\gamma}_i, \quad (2.13)$$

$$\begin{aligned} \text{mit } \mathbf{z}_{pi}^T &= (\mathbf{1}_{P(p)}^T, -\mathbf{1}_{I(i)}^T, -\mathbf{x}_p^T, 0, \dots, -\mathbf{x}_p^T, \dots, 0) \\ \text{und } \boldsymbol{\delta}^T &= (\boldsymbol{\theta}^T, \boldsymbol{\beta}^T, \boldsymbol{\gamma}^T, \boldsymbol{\gamma}_1^T, \dots, \boldsymbol{\gamma}_I^T). \end{aligned}$$

Der Parameter γ entspricht dem Effekt der Kovariablen \mathbf{x}_p bezüglich des Ergebnisses im gesamten Test. Im Fall einer binären Kovariable wird durch den Parameter γ modelliert, ob eine Gruppe den Test besser absolviert als die andere Gruppe [Tutz und Schauberg, 2013].

2.6 Identifizierbarkeit der Modelle

Der lineare Prädiktor des Rasch-Modells mit itemmodifizierenden Effekten (2.11) lautet $\eta_{pi} = \theta_p - \beta_i - \mathbf{x}_p^T \boldsymbol{\gamma}_i$ für Person p und Item i . Bei genauerer Betrachtung sieht man, dass der lineare Prädiktor η_{pi} mit einer Konstante \mathbf{c} auf geschickte Art umparametrisiert werden kann und Modell (2.11) in der bisherigen Form nicht eindeutig lösbar ist [Tutz und Schauberg, 2013].

Genauer gilt:

$$\begin{aligned}
 \eta_{pi} &= \theta_p - \beta_i - \mathbf{x}_p^T \boldsymbol{\gamma}_i \\
 &= \theta_p - \beta_i - \mathbf{x}_p^T (\boldsymbol{\gamma}_i - \mathbf{c}) - \mathbf{x}_p^T \mathbf{c} \\
 &= \tilde{\theta}_p - \beta_i - \mathbf{x}_p^T \tilde{\boldsymbol{\gamma}}_i,
 \end{aligned} \tag{2.14}$$

mit $\tilde{\theta}_p = \theta_p - \mathbf{x}_p^T \mathbf{c}$ und $\tilde{\boldsymbol{\gamma}}_i = \boldsymbol{\gamma}_i - \mathbf{c}$.

Die Parameter $\boldsymbol{\delta}^T = (\boldsymbol{\theta}^T, \boldsymbol{\beta}^T, \boldsymbol{\gamma}_1^T, \dots, \boldsymbol{\gamma}_I^T)$ und $\tilde{\boldsymbol{\delta}}^T = (\tilde{\boldsymbol{\theta}}^T, \boldsymbol{\beta}^T, \tilde{\boldsymbol{\gamma}}_1^T, \dots, \tilde{\boldsymbol{\gamma}}_I^T)$ ergeben exakt dasselbe Modell. Die Parameter θ_p sind genau um den Wert $\mathbf{x}_p^T \mathbf{c}$ und die Parameter $\boldsymbol{\gamma}_i$ um den Wert \mathbf{c} verschoben [Tutz und Schauberger, 2013].

Unter folgenden Restriktionen ist Modell (2.11) eindeutig identifizierbar:

1. $\beta_I = 0, \boldsymbol{\gamma}_I^T = (0, \dots, 0)$ (oder für ein beliebiges, anderes Item).
2. Die gewöhnliche Designmatrix mit Zeilen $(1, \mathbf{x}_1^T), \dots, (1, \mathbf{x}_P^T)$ hat vollen Rang.

Um die Identifizierbarkeit von Modell (2.11) zu garantieren, müssen lt. Bedingung 1 der Koeffizient β_i und die Koeffizienten γ_{iq} eines Items festgesetzt werden. Im Allgemeinen kann hierfür jedes beliebige Item gewählt werden. Bedingung 2 ist eine allgemeine Bedingung, wie sie in ähnlicher Form auch in üblichen Regressionsmodellen benötigt wird. Den Beweis und weitere Details zu den Bedingungen findet sich in [Tutz und Schauberger, 2013]. Die eingeführten Bedingungen sind unabhängig von der ursprünglichen Restriktion des Modells $\theta_P = 0$.

Modelliert man Daten durch ein Rasch-Modell mit itemmodifizierenden Effekten (2.6), so geht man grundsätzlich davon aus, dass für die meisten Items das einfache Rasch-Modell (2.5) gültig ist und nur für wenige Items die Koeffizienten $\boldsymbol{\gamma}_i$ ungleich Null sind. Es ist wünschenswert, dass

bei der Schätzung die maximale Anzahl an Items bestimmt wird, für die das Rasch-Modell Gültigkeit besitzt. Welches Item in Bedingung 1 gewählt wird, hängt genau von dieser Zielsetzung ab. Wie die Restriktionen bei der Schätzung explizit umgesetzt werden, um am Ende eine eindeutig identifizierbare Lösung vorliegen zu haben, wird in Abschnitt 3.5 erläutert.

Betrachtet man das erweiterte Modell mit zusätzlichem Populationseffekt (2.13), so stößt man auf ein weiteres Identifikationsproblem. Trotz der Festsetzung von $\beta_i = 0$ und $\gamma_i^T = (0, \dots, 0)$ für ein Item i ist es nicht möglich, die Parameter γ und θ_p ohne weitere Restriktion klar voneinander zu unterscheiden. Der lineare Prädiktor η_{pi} kann folgendermaßen umparametrisiert werden [Tutz und Schauburger, 2013]:

$$\begin{aligned}\eta_{pi} &= \theta_p - \beta_i - \mathbf{x}_p^T \gamma - \mathbf{x}_p^T \gamma_i \\ &= \theta_p + \mathbf{x}_p^T \mathbf{c} - \beta_i - \mathbf{x}_p^T (\gamma + \mathbf{c}) - \mathbf{x}_p^T \gamma_i \\ &= \tilde{\theta}_p - \beta_i - \mathbf{x}_p^T \tilde{\gamma} - \mathbf{x}_p^T \gamma_i,\end{aligned}\tag{2.15}$$

mit $\tilde{\theta}_p = \theta_p + \mathbf{x}_p^T \mathbf{c}$ und $\tilde{\gamma} = \gamma + \mathbf{c}$.

Nachdem Vektor \mathbf{c} beliebig wählbar ist, kann immer $\gamma = (0, \dots, 0)$ festgelegt werden, und man erhält wieder die ursprüngliche Form des linearen Prädiktors (2.11). In dieser Form ist also nicht eindeutig, welcher Teil der Fähigkeit der Person durch die Zugehörigkeit zur jeweiligen Gruppe erklärt werden kann [Tutz und Schauburger, 2013].

Eine Möglichkeit, dies zu identifizieren, ist ein zweistufiges Schätzverfahren:

1. Schätze das Modell (2.11) ohne zusätzlichen Populationseffekt.
2. Berechne eine Regression der geschätzten Parameter $\hat{\theta}_p$ auf den Kovariablenvektor \mathbf{x}_p .

Das Regressionsmodell in Schritt 2 gibt schließlich an, welcher Teil der Va-

Boosting-Techniken zur Modellierung
itemmodifizierender Effekte

Eine Erweiterung klassischer Item-Response-Modelle

Berger, M.

2015, IX, 125 S. 45 Abb., Softcover

ISBN: 978-3-658-08704-3