

Methode voor schatting van itemmoeilijkheden

Item response theory

Om de moeilijkheid van de opgaven precies te schatten, wordt gebruik gemaakt van *item response theory* (IRT). Opgaven worden binnen deze theorie items genoemd. De antwoorden van leerlingen worden responsen genoemd. De kans dat een bepaalde kandidaat een bepaald item goed beantwoordt, is afhankelijk van de vaardigheid van de persoon en van kenmerken van het item, zoals de moeilijkheid ervan.

One parameter logistic model

De vaardigheid van een kandidaat wordt met de Griekse letter θ aangeduid. De kans dat een kandidaat met een vaardigheid θ een item goed maakt, wordt omschreven met een wiskundige formule of functie. Van de familie van modellen die binnen de IRT vallen, wordt in dit geval het *one parameter logistic model* (OPLM) gebruikt (Glas & Verhelst, 1989; Verhelst & Glas, 1993; Verhelst, Glas & Verstralen, 1993; Eggen & Verhelst, 2011).

De itemresponsfunctie van het OPLM is gegeven door:

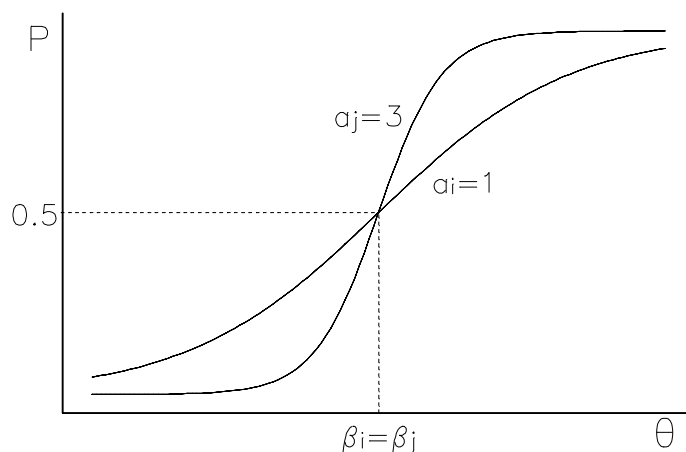
$$f_i(\theta) = \frac{\exp [a_i(\theta - \beta_i)]}{1 + \exp [a_i(\theta - \beta_i)]}$$

waarin:

a_i = de zogenaamde discriminatie-index van het item

β_i = de moeilijkheidsparameter van item i

In onderstaand figuur zijn de itemresponscurven weergegeven van twee items i en j , die even moeilijk zijn maar verschillend discrimineren. Als de vaardigheid van de kandidaat gelijk is aan de moeilijkheid van een opgave, dan heeft hij/zij een kans van 50% om de opgave goed te maken. Opgaven met een hoge discriminatie-index onderscheiden beter tussen hoog- en laag-vaardige kandidaten, ofwel de kans om een opgave goed te maken neemt hierbij snel toe met θ . Merk op dat discriminatie hier dus niets met geslacht of afkomst te maken heeft. De index a_i wordt ook wel de hellingsparameter genoemd.



Twee itemresponscurven in het OPLM: zelfde moeilijkheid, verschillende discriminatie-index

Conditional maximum likelihood

Een vaak toegepaste schattingsmethode voor de moeilijkheidsparameters β_i is de 'conditionele grootste aannemelijkheidsmethode', in het Engels: *conditional maximum likelihood* (CML). Deze methode maakt gebruik van het feit dat in het model een afdoende steekproefgrootte (*sufficient statistic*) bestaat voor de latente variabele θ , namelijk de ruwe score of het aantal correct beantwoorde items. Dat betekent grofweg dat, indien de itemparameters bekend zijn, alle informatie die het antwoordpatroon over de vaardigheid bevat, kan worden samengevat in de ruwe score; het doet er dan verder niet meer toe welke opgaven goed en welke fout zijn gemaakt. Hieruit vloeit voort dat de conditionele kans op een juist antwoord op item i , gegeven de ruwe score, een functie is die alleen afhankelijk is van de itemparameters en onafhankelijk van de waarde van θ (Verhelst, 1992). De CML-schattingsmethode maakt van deze functie gebruik. Deze methode maakt geen enkele vooronderstelling over de verdeling van de vaardigheid in de populatie, en is ook onafhankelijk van de wijze waarop de steekproef is getrokken. Door de indices a_i te beperken tot (positieve) gehele getallen, en door ze a priori als constanten in te voeren, is het mogelijk CML-schattingen van de itemparameters β_i te maken.

Literatuur

De literatuurreferenties in dit document zijn terug te vinden in het hoofdstuk [Literatuur](#) in de toetsspecial [Normering met een vaardigheidsschaal](#).