

Practicum 9

Het Rasch-model via PCA

9.1 Rasch-model

Zie Meerling II pagina 160-166

Werkboek M&T 2B: ND Verhelst (1999) *Wat is moderne testtheorie?*

9.2 Opdrachten.

9.2.1 Kopieer via de internet-snelkoppeling File Copy het bestand REKENEN2.SAV.

Het bestand bevat dezelfde 259 leerlingen als rekenen.sav. De antwoorden op 10 rekensommen (z.o.z.) zijn reeds gehercodeerd (1 = fout, 2 is goed). De variabele GOED bevat het aantal goed gemaakte sommen per leerling. De PCAscores van de 2-dimensionale PCAanalyse op de 10 sommen zijn toegevoegd.

We kunnen de centrale vraag van de vorige week, nmlk:

vormen de sommen een 1-dimensionale schaal?

ook via het Rasch-model benaderen. Helaas zijn er (nog) geen simpele Rasch-programa's. Er kan echter bewezen worden dat je het Rasch-model kunt benaderen met PCA.

Eerst het Rasch-model:

- het is een 1-dimensionaal model voor de beantwoording van een aantal items.
- het is een probabilistisch model, dat de kans geeft dat een item-g goed wordt beantwoord: P_g (goed) als functie van twee andere dingen (parameters):
 - 1^e: de schaalwaarde θ_i (theta) van leerling-i op het 1-dimensionale model; deze schaalwaarde moet geschat worden.
 - 2^e: de moeilijkheidsgraad van het betreffende item: b_g (de b van item g, waarbij g vandaag loopt van 1 naar 10).
- De functie levert P-waarden op die liggen tussen 0 en 1, en P stijgt zachtjes en gestaag (monotoon) als θ toeneemt van laag naar hoog.

Hier is de functie:

$$P_{gi}(\text{goed} \mid \theta_i, b_g) = \text{EXP}(\theta_i - b_g) / (1 + \text{EXP}(\theta_i - b_g))$$

hiermee bereken je de kans dat item g door persoon i goed wordt beantwoord. Als $\theta_i = b_g$ dan geldt dat $P = 1 / (1 + 1) = 0.50$ (omdat dan $\theta_i - b_g = 0$ en omdat $\text{EXP}(0) = 1$)

Met PCA kunnen we zowel θ als b_g benaderen:

- de PCAscores op de 1^e dimensie zijn een goede schatting van de θ -waarden voor de leerlingen
- de b_g van een item kan worden geschat uit de categoriekwantificaties (= gemiddelde score) op de eerste PCA as van de goed- en foutcategorien van het item:
$$b_g = (CQ_{\text{goed}} + CQ_{\text{fout}}) / 2$$
(zelf uit te rekenen voor het betreffende item). Deze schatting van b_g kun je beschouwen als een soort grenswaarde tussen goede en foute antwoorden op een item. Op die grens is de kans op een goed antwoord 0.50 en dat klopt met de definitie van b_g .

Nu berekenen we voor som 6 de theoretische waarden P, op basis van bovenstaande formule voor het Rasch-model. Hiervoor moeten we eerst de categoriekwantificaties op de eerste dimensie voor 'goed' en 'fout' weten.

9.2.2 Bepaal deze categoriekwantificaties via Means. Gebruik vervolgens in de Compute opdracht voor het bepalen van P de SPSS functie EXP() om een e-macht te berekenen. Noem de resultaatvariabele Prash_6.

De volgende stap is om deze modelfunctie laten tekenen. Gewenst resultaat: een item-karakteristieke curve, zie Meerling pag. 162), met op de assen:

horizontaal: θ = objectscores van de leerlingen

vertikaal: P = kansen van de leerlingen om som 9 goed te beantwoorden.

- 9.2.3 Maak nu een spreidingsdiagram (scatter) van PCAScores op de x-as en Prsch_6 op de y-as. Maak de range van de y-as 0-1. Controleer of de vorm van de curve globaal overeenkomt met die in Figuur 6 (Meerling, p.162). Druk de grafiek van de Rasch-curve landscape en paginagroot af.

Klopt deze Rasch-model-curve voor som 6 met de data? Hieronder volgt één van de mogelijke aanpakken om dit te onderzoeken.

Rasch geeft kansen op een goed antwoord op som 6 voor iedere waarde van θ , volgens het model tenminste. De vraag kan gesteld worden hoeveel leerlingen met een bepaalde θ dit item daadwerkelijk goed hebben, en hoeveel hebben het fout? Klopt de proportie goede antwoorden met de modelkansen? Dit is de idee.

De praktische uitwerking gaat als volgt. Neem geen afzonderlijke θ -waarden, maar deel de objectscores van de 1^e PCAdimensie in ongeveer gelijke stukjes. Dit kan worden gedaan met de variabele GOED.

- 9.2.4
- Maak een kruistabel van GOED (rij) en som 6 (kolom) en vraag om rij-percentages: deze kruistabel geeft het aantal en het % leerlingen met een goedscore in een categorie van GOED. Zet deze percentages om in kansen (8% → kans van .08).
 - Bereken de gemiddelde modelkansen (Prsch_6) voor de leerlingen in de verschillende categorieën van 'goed' (via Means).
 - Bereken de gemiddelde θ (obsco1_1) voor de leerlingen in de verschillende categorieën van 'goed' (via Means).

Vergelijk de kruistabel-kansen met de gemiddelde modelkansen en de gemiddelde θ 's. Print de kruistabel en gemiddeldentabellen.

- 9.2.5 Bereken analoog aan opgave 9.2.2 voor item 3 de variabele Prsch_3. Maak analoog aan opgave 9.2.3 een spreidingsdiagram, maar nu via overlay, van de objectscores op de x-as met Prsch_3 en Prsch_6 op de y-as (die weer van 0 - 1 loopt).

Werkgroeprepresentatie

NB Neem de geprinte SPSS uitvoer mee naar de werkgroep.

De **centrale vraag** van dit practicum was: vormen de 10 items een 1-dimensionale schaal voor deelvaardigheid?

Geef hierop in de presentaties antwoord op basis van A) de Rasch-benadering B) een uitbreiding op het Rasch model.

Vorbereiding voor beide presentaties

Zet in het spreidingsdiagram van opgave 9.2.3 de kruistabel-kansen en de gemiddelde modelkansen op de Y-as uit tegen de gemiddelde θ per categorie van de variabele GOED op de X-as. Merk op dat de modelkanspunten steeds midden in een puntengroepje terechtkomen, dat hoort bij een bepaalde waarde van GOED.

- A) Beschrijf op basis van a) de kruistabel en de gemiddeldentabel en b) het nieuwe spreidingsdiagram het verschil tussen de modelkansen en de empirische kansen. Doe dit in termen van moeilijkheid en discriminatie van dit item.
- B) Bekijk op basis van het nieuwe spreidingsdiagram met welke parameter het Rasch-model uitgebreid moet worden om deze verschillen (deels) verhelpen?
De helling van de curve op het punt $\theta = 0$ is in het Rasch-model altijd gelijk aan 1 en deze helling op dat punt is gelijk aan 0.25 maal de waarde van de nieuwe parameter. Maak een schatting van deze nieuwe parameter voor item 6.