**Item Response Models**

*P* items are developed to measure a single ability or trait, and these items are administered to *J* test takers. In this model, the items are scored as correct (1) or incorrect (0). The Rasch model, a one-parameter item response model, is specified to estimate the log-odds of a correct response, as a function of person *j* ability and item *p* difficulty.

The typical formulation of the Rasch (1 parameter) model is a function of person *j* ability θ and item *i* difficulty *b*.

$$P\left(Y\_{ij}=1\right)=\frac{e^{θ\_{j}-b\_{j}}}{1+e^{θ\_{j}-b\_{j}}}=\frac{1}{1+e^{-(θ\_{j}-b\_{j})}}$$

Working with logits (log-odds) through the HGLM framework, we reorganize the model

$$log\left[P\left(Y\_{ij}=1\right)\right]=η\_{ij}=θ\_{j}-b\_{i}$$

Notice that β*i* is now –*bj*. In the logit link function notation of HLM, we have the log-odds link as a function of the difference between person ability and item difficulty

η*jp* = π0*j* – π*qj*

To define the scale in logits, we set persons or items to have a mean of 0. Then, the probability that examinee *j* correctly answers item *p* is

$$Prob\left(Y\_{ij}=1\right)=φ\_{ij}=\frac{1}{1+exp⁡\{-η\_{ij}\}}$$

This is the same formula we saw earlier in the case of regressing binomial outcomes, where we employ the logit link function (the log-odds of c success), η*ij* , and want to turn that log-odds into a probability.

In a standard Rasch IRT model, person abilities and item difficulties are fixed. In the HLM formulation of the Rasch model, person abilities vary randomly over a population of test takers.

*Level-1 Model*

The sampling model is Bernoulli (a binomial sampling model where there is *m* = 1 trial). *Yij* = 1 with probability φ*ij*. We use the logit link function

$$η\_{ij}=log\left(\frac{φ\_{ij}}{1-φ\_{ij}}\right)$$

The level-1 model is then

$$η\_{ij}=π\_{0j}+\sum\_{p=1}^{P-1}π\_{pj}X\_{pij}$$

π0*j* is person *j* ability

*Xpij* is a dummy indicator variable that = 1 if response *i* for person *j* is to item *p*, 0 otherwise

π*pj* is the difference in log-odds of a correct response between item *p* and a reference item (omitted due to redundancy) for person *j*

This level-1 equation can be reduced to η*jp* = π0*j* – π*qj* for item *i* associated with the *q*th dummy variable. In this way, the probability that person *j* correctly answers item *i* is

$$Prob\left(Y\_{ij}=1\right)=φ\_{ij}=\frac{1}{1+exp⁡\{-η\_{ij}\}}$$

The difficulty of the reference item is arbitrarily set to 0.

The only reason why there are two subscripts is to account for the instance that person *j* might not have a respond *i* for every item *p*.

*Level-2 Model*

We have two sets of parameters estimated at level-1 including person abilities (π0*j*) and item difficulties or item effects (π*pj* ). To emulate the Rasch model, we fix all of the item effects

π0*j* = γ0*j* + *u*0*j* where *u*0*j* ~ N(0, τ00)

π1*j* = γ10

π2*j* = γ20

…

π(*p*-1)*j* = γ(*p*-1)0

τ00 is the variance of abilities in the population.

Item effects are fixed (invariant) across examinees.

The log-odds that person *j* will respond correctly to item *p* is β00 + *u*0*j* for the reference item and β00 + β*p*0 + *u*0*j* for any other item *p*.

Returning to the Rasch model η*jp* = θ*j* – δ*p*

Person *j* ability is θ*j* = π0*j* = β00 + *u*0*j*

 Item difficulty of the reference item is δ*p* = 0

 and for all other *p* items, δ*p* = – β0*p*

* Differences in item difficulties are given by differences between regression coefficients
* Differences in person abilities are given by differences in random intercepts
* Item difficulties and person abilities are on an interval scale defined in logit metric.

To run these models in HLM, create the MDM as a HLM2 or HLM3, depending on number of levels. In the MDM, level-1 includes the item responses (a single variable in long form) and the item indicators; level-2 and others include relevant variables for those levels. When designing the model, identify all item indicators except one (reference item) and all item effects must be fixed at level-2; the intercept is random (the *u*0*j* is the person ability relative to the mean). Select the Basic Model Specifications for the distribution of outcome variables, Bernoulli for dichotomous outcomes, Ordinal for ordinal outcomes, etc.

**Minnesota Student Survey**

Example of Teacher School Support 6 items, item responses dichotomized (0=1, 2; 1=3, 4, 5)

|  |  |  |  |
| --- | --- | --- | --- |
| Y21d | Overall, adults at my school treat students fairly. | {1, Strongly disagree; to 4, Strongly agree} | TSS |
| Y21e | Adults at my school listen to the students. | {1, Strongly disagree; to 4, Strongly agree} | TSS |
| Y21f | The school rules are fair. | {1, Strongly disagree; to 4, Strongly agree} | TSS |
| Y21g | At my school, teachers care about students. | {1, Strongly disagree; to 4, Strongly agree} | TSS |
| Y21h | Most teachers at my school are interested in me as a person. | {1, Strongly disagree; to 4, Strongly agree} | TSS |
| Y59d | How much do you feel teachers/other adults at school care about you? | {1, Not at all; 2, A little; 3, Some; 4, Quite a bit; 5, Very much} | TSS |

**Level-1 Model**

    Prob(*RESP345mj*=1|ψ*j*) = ϕ*mj*
    log[ϕ*mj*/(1 - ϕ*mj*)] = η*mj*
    η*mj* = ψ0*j* + ψ1*j*\*(*IND1mj*) + ψ2*j*\*(*IND2mj*) + ψ3*j*\*(*IND3mj*) + ψ4*j*\*(*IND4mj*) + ψ5*j*\*(*IND5mj*)

**Level-2 Model**

    ψ0*j* = γ00 + *u*0*j*
    ψ1*j* = γ10
    ψ2*j* = γ20
    ψ3*j* = γ30
    ψ4*j* = γ40
    ψ5*j* = γ50

Level-1 variance = 1/[ϕ*mj*(1-ϕ*mj*)]

*u*0*j* is the person levels of TSS

γ00 is the item location (in logits) for the reference item (item 6 in our example)

The item locations for the remaining items are obtained by subtracting each item’s effect from the reference item effect (–γ*i*0 – γ00)

*Note*: In a three-level model, the trait level of persons are the aggregated level-2 and level-3 residuals, the person residual (trait level) and the residual for their group (level-3 residual).

**Final estimation of fixed effects: (Unit-specific model)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fixed Effect |  Coefficient |  Standarderror |  *t*-ratio |  Approx.*d.f.* |  *p*-value |
| For INTRCPT1, *ψ0*  |
|     INTRCPT2, *γ00*  | 1.551164 | 0.036655 | 42.318 | 7639 | <0.001 |
| For IND1 slope, *ψ1*  |
|     INTRCPT2, *γ10*  | 0.530704 | 0.047733 | 11.118 | 38002 | <0.001 |
| For IND2 slope, *ψ2*  |
|     INTRCPT2, *γ20*  | 0.281744 | 0.046264 | 6.090 | 38002 | <0.001 |
| For IND3 slope, *ψ3*  |
|     INTRCPT2, *γ30*  | -0.082213 | 0.044577 | -1.844 | 38002 | 0.065 |
| For IND4 slope, *ψ4*  |
|     INTRCPT2, *γ40*  | 1.055976 | 0.051815 | 20.380 | 38002 | <0.001 |
| For IND5 slope, *ψ5*  |
|     INTRCPT2, *γ50*  | -0.476906 | 0.043173 | -11.046 | 38002 | <0.001 |

A comparison to Rasch estimates of item location using a full ordinal partial-credit model (Winsteps):

**A Model for Impact (e.g., gender difference in trait level)**

    ψ0*j* = γ00 + γ01(sex)*j* +*u*0*j*
    ψ1*j* = γ10
    ψ2*j* = γ20
    ψ3*j* = γ30
    ψ4*j* = γ40
    ψ5*j* = γ50

In this model, ψ0*j* is common to all items in level-1, as it is the location of the reference item to identify the scale; thus it is the overall difference between males and females on all items – the difference in ability between male and female respondents.

**Final estimation of fixed effects: (Unit-specific model)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fixed Effect |  Coefficient |  Standarderror |  *t*-ratio |  Approx.*d.f.* |  *p*-value |
| For INTRCPT1, *ψ0*  |
|     INTRCPT2, *γ00*  | 1.644895 | 0.043713 | 37.629 | 7638 | <0.001 |
|      FEMALE, *γ01*  | -0.181614 | 0.045773 | -3.968 | 7638 | <0.001 |
| For IND1 slope, *ψ1*  |
|     INTRCPT2, *γ10*  | 0.530804 | 0.047740 | 11.119 | 38002 | <0.001 |
| For IND2 slope, *ψ2*  |
|     INTRCPT2, *γ20*  | 0.281827 | 0.046271 | 6.091 | 38002 | <0.001 |
| For IND3 slope, *ψ3*  |
|     INTRCPT2, *γ30*  | -0.082237 | 0.044585 | -1.844 | 38002 | 0.065 |
| For IND4 slope, *ψ4*  |
|     INTRCPT2, *γ40*  | 1.056250 | 0.051822 | 20.382 | 38002 | <0.001 |
| For IND5 slope, *ψ5*  |
|     INTRCPT2, *γ50*  | -0.477160 | 0.043183 | -11.050 | 38002 | <0.001 |

**A Model for DIF**

**Final estimation of fixed effects: (Unit-specific model)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fixed Effect |  Coefficient |  Standarderror |  *t*-ratio |  Approx.*d.f.* |  *p*-value |
| For INTRCPT1, *ψ0*  |
|     INTRCPT2, *γ00*  | 1.652380 | 0.053370 | 30.961 | 7638 | <0.001 |
|      FEMALE, *γ01*  | -0.195052 | 0.073457 | -2.655 | 7638 | 0.008 |
| For IND1 slope, *ψ1*  |
|     INTRCPT2, *γ10*  | 0.535299 | 0.069964 | 7.651 | 37997 | <0.001 |
|      FEMALE, *γ11*  | -0.008013 | 0.095727 | -0.084 | 37997 | 0.933 |
| For IND2 slope, *ψ2*  |
|     INTRCPT2, *γ20*  | 0.350346 | 0.068304 | 5.129 | 37997 | <0.001 |
|      FEMALE, *γ21*  | -0.126609 | 0.092908 | -1.363 | 37997 | 0.173 |
| For IND3 slope, *ψ3*  |
|     INTRCPT2, *γ30*  | -0.172108 | 0.064635 | -2.663 | 37997 | 0.008 |
|      FEMALE, *γ31*  | 0.171673 | 0.089290 | 1.923 | 37997 | 0.055 |
| For IND4 slope, *ψ4*  |
|     INTRCPT2, *γ40*  | 1.006111 | 0.075437 | 13.337 | 37997 | <0.001 |
|      FEMALE, *γ41*  | 0.095013 | 0.103819 | 0.915 | 37997 | 0.360 |
| For IND5 slope, *ψ5*  |
|     INTRCPT2, *γ50*  | -0.459143 | 0.063078 | -7.279 | 37997 | <0.001 |
|      FEMALE, *γ51*  | -0.034561 | 0.086560 | -0.399 | 37997 | 0.690 |