Using Rasch Measures For Rasch Model Fit Analysis

George Karabatsos, Ph.D.

In Rasch fit analysis, Z_{ni} is used to measure the fit of a single person-item response, while mean-square (MS) statistics analyze the fit of response sets, and ZSTD tests the significance of a particular MS value.

Most analysts find the Rasch model person measures and item calibrations easier to understand and communicate than the Z_{ni}, MS, and ZSTD statistics. For instance, only through the necessary calculations do we know how much logit-misfit is involved for a given Z_{ni} or MS value. Furthermore, Z_{ni}, MS, and ZSTD are nonlinear functions of Rasch model values (e.g., B_n–D_i).

This paper introduces a Rasch model fit statistic that enables the analyst to interpret fit of a response on the same scale as person measures and item calibrations. Essentially, this is accomplished by explicitly incorporating the logistic Rasch model in the fit statistics.

RESPONSE-FIT INDEX FOR DICHOTOMOUS CHOICES

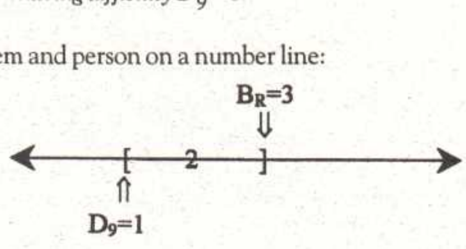
Let K_{ni} denote the logit-fit of person n's response to item i, calculated by: $K_{ni} = f_{ni}(B_n - D_i)$ [1]

where f_{ni} classifies the model-fit of a person-item response

$f_{ni} = 0$ for a response that fits the model
($X_{ni}=1$ when $B_n \geq D_i$, or $X_{ni}=0$ when $B_n \leq D_i$)
$f_{ni} = -1$ for a response that misfits the model
($X_{ni}=1$ when $B_n < D_i$, or $X_{ni}=0$ when $B_n > D_i$).

Example 1. *Richard with ability $B_R=3$ encounters "item 9" having difficulty $D_9=1$.*

Map item and person on a number line:

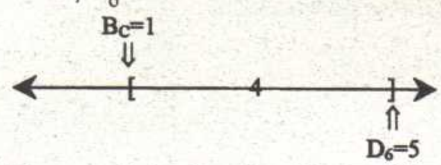


Expected Response Rule: Since $B_n > D_i$, then $\{X_{ni}=1\}$ is the expected response.

Two Possible Scenarios:

Response	Fit result	Interpretation
$\{X_{ni}=1\}$	$K_n = 0(3-1) = 0$	Response fits measurement model.
$\{X_{ni}=0\}$	$K_n = -1(3-1) = -2$	Richard responded 2 logits *below* expectation.

Example 2. Cindy with ability $B_C=1$ encounters "item 6" having difficulty $D_6=5$.

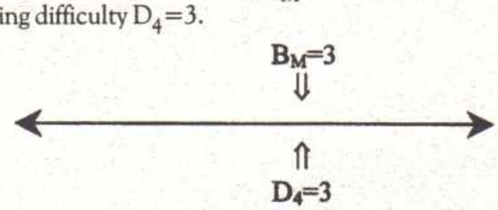


Expected Response Rule: Since $B_n < D_i$, then $\{X_{ni}=0\}$ is the expected response.

Two Possible Scenarios:

Response	Fit result	Interpretation
$\{X_{ni}=1\}$	$K_{ni} = -1(1-5) = 4$	Cindy responded 4 logits above expectation.
$\{X_{ni}=0\}$	$K_{ni} = 0(1-5) = 0$	Response fits measurement model.

Example 3. Mary with ability $B_M=3$ encounters "item 4" having difficulty $D_4=3$.



Expected Response Rule: Since $B_n = D_i$, then $\{X_{ni}=0\}$ and $\{X_{ni}=1\}$ have equal probability ($P_{ni1}=.50$, therefore $P_{ni0}=.50$). So by definition, *neither response misfits the model.*

Two Possible Scenarios:

Response	Fit result	Interpretation
$\{X_{ni}=1\}$	$K_{ni} = 0(3-3) = 0$	Response fits measurement model.
$\{X_{ni}=0\}$	$K_{ni} = 0(3-3) = 0$	Response fits measurement model.

RESPONSE-FIT INDEX FOR POLYTOMOUS CHOICES

Since all Rasch models reduce to the dichotomous-response model, Equation 1 can be extended to analyze the fit of a rating-scale response. For an item with m response categories, there are m–1 adjacent-category steps, where each step j is denoted by the parameter F_j. A person's rating scale response to that item indicates a certain number of "advanced" steps, and a certain number of "unadvanced" steps. Each "advanced" versus "unadvanced" step response is a dichotomy, and therefore, there are j dichotomous responses within a single rating scale response.

The fit calculation of a single rating scale response involves calculating $f_{ni}(B_n-D_i-F_j)$ for each of the steps, and letting K_{ni} equal the calculation that differs the most from zero. The K_{ni} for a single rating scale response is therefore calculated by:

$$K_{ni} = |\max| \ [f_{nij} \ (B_n-D_i-F_j)] \qquad [2]$$

where,

$|\max|$ maximum in absolute value

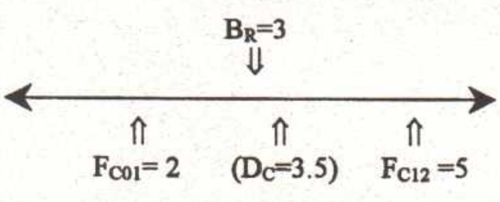
$f_{nij} = 0$ for a step-response that fits the model

$f_{nij} = -1$ for a step-response that misfits the model

In the case of dichotomous response choices, there is only one threshold j, in which case equation [2] reduces to equation [1].

Here is an example of an item with a three category (m=3) rating scale, where $X_{ni}=\{0,1,2\}$, rendering m–1=2 steps. Let F_{01} denote the parameter for the step to category 1 from 0, and F_{12} for the step to category 2 from 1.

Example 4. Bob with ability $B_B=3$ encounters "item C" having difficulty $D_C=3.5$, where $F_{01}=?1.5$ and $F_{12} = +1.5$ relative to D_C.



Expected Response Rule: Since $B_n>F_{i01}$ and $B_n<F_{i12}$, , $\{X_{ni}=1\}$ is the expected response.

Three Possible Scenarios:

Response	Fit result	Interpretation
$\{X_{ni}=2\}$	$K_{ni}= \lvert\max\rvert \ [0(3-2),-1(3-5)] = 2$	Bob responded 2 logits above expectation.
$\{X_{ni}=1\}$	$K_{ni}= \lvert\max\rvert \ [0(3-2),0(3-5)] = 0$	Response fits measurement model.
$\{X_{ni}=0\}$	$K_{ni}= \lvert\max\rvert \ [-1(3-2),0(3-5)] = -1$	Bob responded 1 logit below expectation.

FIT ANALYSIS OF RESPONSE SETS

Analyzing response sets is straightforward. The average of the absolute value of $|K_{ni}|$ values can be taken across all responses of interest:

$$|\overline{\mathbf{K}}_{\mathbf{ni}}| = \frac{\sum |\mathbf{K}_{\mathbf{ni}}|}{\mathbf{N}_{\{\mathbf{X_{ni}=x}\}}} \qquad [3]$$

to obtain the "average logit noise," where $N_{\{Xni = x\}}$ denotes the total number of responses. Person $|\overline{K}_{ni}|$ is obtained by applying Equation 3 for all person responses; item $|\overline{K}_{ni}|$ is calculated for all item responses.

It is also informative to take the average of certain response subsets. Examples include (1) the subset of "negative" K_{ni} values, and (2) the subset of "positive" K_{ni} values. Subset (1) indicates the magnitude of surprising "low" responses (e.g., occurring from sleeping, carelessness, etc.), and subset (2) indicates the magnitude of surprising "high" responses (e.g., lucky-guessing).

The accuracy of K_{ni} depends on parameter values estimated from the data, *but we know we estimate parameters from noisy data in the first place* (Z_{ni}, *MS*, *ZSTD*, and all parameter-dependent fit methods suffer this uncertainty). When data noise is high, we cannot trust the accuracy of parameter estimates, and therefore can no longer trust the accuracy of K_{ni} and other parameter-dependent fit statistics. In cases where data is too noisy for the parameter-dependent fit statistics to be useful, an alternative is a an estimate of Guttman fit:

$$G = \frac{N_{|K_{ni}|>0}}{N_{\{X_{ni}=x\}}} \qquad [4]$$

which is the proportion of unexpected responses across the relevant response set. G is linearized by the transformation log(G/(1–G).

It is also informative to change the numerator of Equation [4] to calculate the proportion of surprising "low" responses ($N_{K<0}$) and "high" responses ($N_{K>0}$).

G interprets Kni values as ordinal (possible values: either $K_{ni}=0$ or $|K_{ni}|>0$), which renders it more robust than $|\overline{K}_{ni}|$ (and Z_{ni}, MS, ZSTD) to inaccurate parameter estimations. Hence, G can be considered a parameter-free fit statistic.