SAS® GLOBAL FORUM 2018

USERS PROGRAM

%Q-Index: A SAS® Macro for a Conditional Item-Fit Index for the Rasch Model

Samantha Estrada M.S., Sergio Nava-Muñoz M. S., Rafael Perez Abreu, PhD [2]

April 8 – 11 | Denver, CO

#SASGF

%Q-Index: A SAS® Macro for a Conditional Item-Fit Index for the Rasch Model

Samantha Estrada M.S. [1], Sergio Nava-Muñoz M. S.[2], Rafael Perez Abreu, PhD [2]

[1] University of Northern Colorado; [2] Centro de Investigación en Matemáticas, A.C. Unidad Aguascalientes

ABSTRACT

- Rasch analysis is commonly used in educational and psychological testing; the method is also popular in the measurement of health status and evaluation outcomes (Christensen, 2013).
- The purpose of fit statistics is to screen misfitting items, which is an important issue in Rasch analysis in order to evaluate the items consistently through indicators.
- If fit statistics are incorrect, a misfitting item might not be located correctly, or a good item might be identified incorrectly as a misfitting item.
- The Q-Index has desirable characteristics, which could provide a solution to applied researchers concerned with the limitations of current fit indices.
- In SAS, the Rasch model can be estimated utilizing PROC LOGISTIC or a variety of macros such as the GLIMMIX_Rasch (Li, Chen & Kromrey, 2013).
- In this poster, the researchers use SAS to estimate the Rasch model and compute the Q-Index.

WHY IS THIS MACRO USEFUL?

- In PROC LOGISTIC or %GLIMMIX_Rasch the item fit statistics are not available which can be calculated utilizing %Q-Index
- In a situation where Rasch specialized software is unavailable.
- Rasch analysis is a popular psychometric procedure and it is used in a variety of fields such as psychology, education and health sciences

METHODS

The Rasch model can be estimated by

$$P(X_{vi} = 1 \mid \Theta_v = \theta_v) = \frac{e^{(\theta_v - \beta_i)}}{1+e^{(\theta_v - \beta_i)}} \qquad (1)$$

Equation 1 is the original formulation of the model according to Rasch (1980). Where X_{vi} is a random variable indicating success or failure. $X = 1$ indicates success, for example a correct response, while $X = 0$ indicates failure or an incorrect response on the item. The subscript v represents the person while the subscript i represents the item. The probability of a correct response increases as the ability parameter increase toward infinity.

METHODS cont'd

The rating scale model (RSM) is a type of polytomous Rasch model (Bond & Fox, 2015). The RSM is defined as follows:

$$P(X_{vi} = x \mid \Theta = \theta_v) = \frac{e^{(\theta_v x + \psi_{ix})}}{\sum_{h=0}^{m_i} e^{(\theta_v h + \psi_{ih})}} \qquad (2)$$

Where $\theta_v x$ is the person parameter and ψ_{ix} is the ith threshold location parameter of item x.

If the responses by examinees are denoted as X_{vi} the possible responses are coded as $X_{vi} = 0, 1, 2, \dots, m_i$ where the number of response categories for any given item i is $m_i + 1$. Higher ratings should indicate higher levels on the latent trait of interest (Engelhard, 2013).

The equation for the Q-Index index is as follows:

$$Q_i = \frac{\sum_v (x_{vi} - x_{v.G})\theta_v}{\sum_v (x_{v.A} - x_{v.G})\theta_v} \qquad (3)$$

Where the θ parameters can be estimated three different ways: (a) estimated by using all items, (b) estimated by using all items except i, or (c) using other tests which measure the same trait. The Guttman and anti-Guttman pattern response for each examinee, v, conditioned on the given item score distribution, is obtained by ordering examinees according to their ability level, θ as well as assigning the $n_o, n_1, \dots n_m$ response categories $0, 1. \dots, m$ to the examinees in either ascending or descending order (Rost & von Davier, 1994).

MACRO OPTIONS

Options	Descriptions
%path	Specify the location of the data file
%datafile	Specify the name of the data file of interest
%nitms=	Specify the number of items in the survey or inventory
%nperson	Specify the number of persons in the sample
%itmstart	Specify what row the items begin
%idlen	Specify the length of the ID for the persons in the sample

%Q-Index: A SAS® Macro for a Conditional Item-Fit Index for the Rasch Model

Samantha Estrada M.S. [1] Sergio Nava-Muñoz M. S.[2], Rafael Perez Abreu, PhD [2]

[1] University of Northern Colorado; [2] Centro de Investigación en Matemáticas, A.C. Unidad Aguascalientes

EXCERPT CODE

```
PROC IML;
USE X;
  READ ALL VAR _ALL_ INTO X;
CLOSE X;
USE G;
  READ ALL VAR _ALL_ INTO G;
CLOSE G;
USE A;
  READ ALL VAR _ALL_ INTO A;
CLOSE A;
USE B;
  READ ALL VAR {ABILITY} INTO B;
CLOSE PERSON_PAR;

NUM = (X-G)#B;
DEN = (A-G)#B;
Q = NUM[+,]/DEN[+,];
*PRINT X, G, A, B, NUM, DEN, Q;
QT = T(Q);

TITLE "Q-INDEX";
PRINT QT [LABEL="Q-INDEX"];

CREATE QINDEX VAR {QT};
/** CREATE DATA SET **/
APPEND;
/** WRITE DATA IN VECTORS **/
CLOSE QINDEX;
/** CLOSE THE DATA SET **/
QUIT;
PROC PRINT DATA=QINDEX;
RUN;
```

RESULTS

%Qindex	Winmira
0.1776	0.1847
0.2214	0.2299
0.1407	0.1462
0.2146	0.2217
0.3671	0.3773
0.1519	0.1563
0.1323	0.1357
0.1817	0.1850
0.2195	0.2227
0.2014	0.2020

SELECTED REFERENCES

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences.* 711 Third Avenue New York, NY 10017: Routledge.

Li, I., Chen, Y., & Kromrey, J. (2013). (2013). Evaluating the performance of the SAS® GLIMMIX procedure for the dichotomous rasch model: A simulation study . Paper presented at the *South East SAS User Group,* Myrtle Beach, SC. Retrieved from http://analytics.ncsu.edu/sesug/2013/SD-10.pdf

Olsbjerg, M., & Christensen, K. B. (2015). % lrasch_mml: A SAS macro for marginal maximum likelihood estimation in longitudinal polytomous rasch models. *Journal of Statistical Software, 67*(c02)

Pan, T., & Chen, Y. (2011). Using proc logistic to estimate the Rasch model. *Proceedings of the SAS Global Forum 2011 Conference, April 4-7*

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.

Rost, J., & von Davier, M. (1994). A conditional item-fit index for rasch models. *Applied Psychological Measurement, 18*(2), 171-182. doi:https://doi.org/10.1177/014662169401800206

SAS® GLOBAL FORUM 2018

April 8 – 11 | Denver, CO
Colorado Convention Center

#SASGF

Paper 2395-2018

Q-Index: A SAS® Code for a Conditional Item-Fit Index for the Rasch Model

Samantha Estrada, University of Northern Colorado
Sergio Nava-Muños, Centro de Investigación en Matemáticas, A.C. Unidad Aguascalientes
Rafael Perez Abreu, Centro de Investigación en Matemáticas, A.C. Unidad Aguascalientes

ABSTRACT

The use of Rasch analysis has increased in the past decades. Rasch analysis is commonly used in educational and psychological testing; the method is also popular in the measurement of health status and evaluation outcomes (Christensen, 2013). Moreover, the purpose of fit statistics is to screen misfitting items, which is an important issue in Rasch analysis in order to evaluate the items consistently through indicators. If fit statistics are incorrect, a misfitting item may not be located correctly, or a good item may be identified incorrectly as a misfitting item. More importantly, the properties and benefits of using a certain model, in this case the Rasch model, will only hold if the data fit the model. The Q-Index has desirable characteristics, which could provide a solution to applied researchers concerned with the limitations of current fit indices. However, little research has been performed regarding the robustness of Q-Index (Ostini & Nering, 2006). This may be due to the lack of availability of the Q-Index in popular Rasch software such as WINSTEPS. In fact, the Q-Index is only available in the Rasch software WINMIRA that has not been updated since 2001 (von Davier, 2001). In this paper, the researchers utilize SAS to estimate the Rasch model using PROC LOGISTIC (Pan & Chen, 2011) and compute the Q-Index. Finally, this paper describes a SAS macro that fits the Rasch model for dichotomous and polytomous data.

INTRODUCTION

The field of psychometrics is vast, and interdisciplinary. Scholars and researchers in a wide variety of fields such as psychology, education, business and health sciences make use of surveys, tests, and inventories to collect data and conduct their research. At its core, good research design means that the measures utilized in the research should have good psychometric properties; in other words, the measures should reflect the construct or trait of scientific interest and have minimal measurement error. A method to evaluate the psychometric properties of a survey is Rasch analysis.

RASCH ANALYSIS

Rasch (1980) initially developed the model for dichotomous data, since the model has been adapted for different types of measurement such as rating scale data. In the Rasch model, two conditions are part of the model: (a) the trait possessed by the person and (b) the difficulty necessary to provide a certain level of response. These are often refered to as "person measures" and "item difficulties" (Boone, Staver, & Yale, 2014).The following function represents the probability of success for an examinee's response on a dichotomous item:

$$P(X_{vi} = 1 \mid \Theta_v = \theta_v) = \frac{e^{(\theta_v - \beta_i)}}{1+e^{(\theta_v - \beta_i)}} \qquad (1)$$

Equation 1 is the original formulation of the model according to Rasch (1980). Where X_{vi} is a random variable indicating success or failure. $X = 1$ indicates success, for example a correct response, while $X = 0$ indicates failure or an incorrect response on the item. The subscript v represents the person while the subscript i represents the item. The probability of a correct response increases as the ability parameter increase toward infinity. For example, in an educational testing setting the higher the ability of the student and the easier the item, the probability of a correct response is larger. Likewise, in a health science example, the person parameter could represent the level of depression, or pain, while the item parameters would represent the risk of experiencing certain symptoms related to the trait. Consider a dichotomous item constructed to measure depression: Do you have difficulty sleeping in the last two weeks? Or, did your appetite decreased in the last two weeks? According to the Rasch model, the level of depression measured by these items is measured by the person parameter. The Θ_v represents an unobservable, or latent trait and θ is the person parameter which denotes the examinee's location on the latent trait scale. The β represents the item difficulty or item location parameter on the same latent trait scale, and is an item parameter. Both θ and β are on a logit scale (Christensen et al., 2013). Equation 1 is a function of the difference between the examinee's ability and the item difficulty (Wu & Adams, 2013).

$$P(X_{vi} = 0 \mid \Theta_v = \theta_v) = 1 - P(X = 1) = \frac{1}{1+e^{(\theta_v - \beta_i)}} \qquad (2)$$

In Equation 2 responses are coded as 1 for a correct response and 0 for an incorrect response. The logit function of the probability of a positive response is:

$$logit(P(X_{vi} = 1 \mid \Theta_v = \theta_v)) = \theta_v - \beta_i \qquad (3)$$

For this reason, both θ_v and β_i are said to be measured on a logit scale (Christensen et al., 2013). Logit is also known as log-odds. Linacre and Wright (1989) define logit as "the distance along the line of the variable that increases the odds of observing the event specified in the measurement model by a factor of 2.718.., the value of 'e'" (para. 7).

RATING SCALE MODEL

The rating scale model (RSM) is a type of polytomous Rasch model (Bond & Fox, 2015). The assumptions of the polytomous Rasch model are: (a) the latent trait θ is a scalar; thus, the latent trait is unidimensional, (b) the examinees are independent, and (c) the items are locally independent. In other words, the items are conditionally independent given the latent trait.

The RSM is defined as follows:

$$P(X_{vi} = x \mid \Theta = \theta_v) = \frac{e^{(\theta_v x + \psi_{ix})}}{\sum_{h=0}^{m_i} e^{(\theta_v h + \psi_{ih})}} \qquad (4)$$

Where $\theta_v x$ is the person parameter and ψ_{ix} is the ith threshold location parameter of item x.

If the responses by examinees are denoted as X_{vi} the possible responses are coded as $X_{vi} = 0, 1, 2, \dots, m_i$ where the number of response categories for any given item i is $m_i + 1$. Higher ratings should indicate higher levels on the latent trait of interest (Engelhard, 2013). The scoring of ordered categories, with ordered integers such that 0,1…m, implies that the distance between these categories is in equal intervals. For example, the distance between 1 and 2 is the same distance as between 2 and 3 (Engelhard, 2013).

In SAS, Rasch analysis can be done via the PROC LOGISTIC procedure. In fact, a SUGI paper exists on the topic (Pan & Chen, 2011). Additionally, there exist macros to estimate the Rasch model (Chen, Li, & Kromrey, 2013; Olsbjerg & Christensen, 2015). In the Rasch model, as in other psychometric models, it is important to assess the fit of the model. If the data do not fit the model then the benefits and properties only hold if the data fit the model.

ITEM FIT IN THE RASCH MODEL

From the introduction of Rasch analysis item fit was studied by researchers (Gustafsson,1980; Rasch, 1980). In Rasch analysis, the purpose of item fit statistics is to screen misfitting items. If fit statistics are incorrect, a misfitting item may not be located correctly, or a good item may be identified incorrectly as a misfitting item. Imagine a test in an Introduction to Statistics course, a misfitting item would be an item that is answered correctly by low performing students. Likewise, a misfitting item could be an item that is too easy but that students answered incorrectly though they performed well on the test.

In statistics, Chi-square statistics are typically utilized to determine the association between two groups, variables or criteria. For example, in Rasch analysis, the criteria would be if the association of the data and the Rasch model; specifically, how the data fits the Rasch model (Boone, Staver & Yale, 2014). In Rasch analysis is possible to assess both person fit and item fit in addition to model fit. While following the procedures of Pan and Chen's (2011) paper it is possible to assess the model fit utilizing a Chi-square statistic currently there is no procedure to estimate item fit within the SAS software. However, utilizing the %GLIMMIX_Rasch macro by Chen, Li, and Kromrey (2013) one can obtain item fit statistics though research suggest these popular fit statistics known as Infit and Outfit may be flawed.

Popular Rasch software such as Winsteps and RUMM provide item fit statistics based on residuals named Infit, Outfit and standardized Infit and Outfit. However, problems with item fit statistics based on residuals are varied. First, Rost and von Davier (1994) state that the power of fit tests depends on the variance of the fit statistic. In other words, when the item trait locations are close to the examinee's ability level it is difficult to know if there is lack of fit (Ostini & Nering, 2006). Ostini and Nering (2006) state that this problem is common knowledge in the Rasch and item response theory (IRT) fields, but there is no clear solution. Additionally, statistical inference tests are sensitive to sample sizes (Ostini & Nering, 2006). For this reason, researchers suggest the cutoff values for Infit and Outfit should be reevaluated according to sample size (Wang & Chen, 2005; Wu & Adams, 2013). However, these proposed cutoff values have yet to be established which leads to confusion among researchers in their use of residual item fit statistics (Wu & Adams, 2013). The Q-Index, developed by Rost and von Davier provides an alternative to residual item fit statistics.

Q-INDEX

The Q-Index has desirable characteristics which could provide a solution to applied researchers concerned with the limitations of current fit indices. Rost and von Davier (1994) designed the Q-Index which is available in the Rasch software Winmira (von Davier, 2001). Though the software has not been updated since 2001, researchers have utilized the Q-Index in a variety of topics, including using the Q-Index in addition to Infit and Outfit in their studies regarding superitems (items where participants must fill in the blanks in a text) (Eckes, 2011); as a standalone fit statistic for studying motor competence in early childhoold (Utesch et al. 2016); fitting the mixed Rasch model to a reading comprehension test in order to identify types of readers (Baghaei & Carstensen, 2013); assessing the psychometric properties of a sleeping deprivation measure (Janssen, Phillipson, O'Connor & Johns, 2017). The Q-Index has desirable characteristics, which could provide a solution to applied researchers concerned with the limitations of current fit indices.

The Q-Index makes use of the statistical properties of the Rasch model, namely, parameter separability and conditional inference. Parameter separability refers to the form in which the parameters in the Rasch model occur, see equation 3 ($\theta_v - \beta_i$), linear and without interactions, which allows likelihood equations in which the relation between the person ability and data are contained in an equation which is separate from an equation which contains the data and item difficulty parameters. This occurs due to the algebraic separation of parameters specified within the Rasch model, this in turn, allows "derivation of conditional estimation equations" for either item difficulty or person ability (p. 27). In other words, the equations used to estimate item difficulties do not involve the person abilities parameters and vice versa (Wright & Stone, 1999).

Q-Index does not require estimation of the item parameters for any given item but it is conditioned on the score distribution of said item (Rost & von Davier, 1994). In other words, the fit of an item i is evaluated with regard to the conditional probability of its observed response vector. Rost and von Davier's Q-Index is currently available in the Rasch software Winmira (von Davier, 2001). The Q-Index can be utilized with any unidimensional Rasch model, for example, the Rasch dichotomous model, the rating scale model (Wright & Masters, 1982), the equidistance model (Andrich, 1982), the partial credit model (Masters, 1982), continuous rating scale model (Müller, 1987), or the dispersion model (Rost, 1988).

When testing the significance of the fit of an item, the item parameters are estimated first and then utilized to derive the sampling distribution for the item parameter (Rost & von Davier, 1994). Unlike the chi-square fit statistics, the Q-Index is not based on the differences between observed and expected response scores. For this reason, the Q-Index does not suffer from problems caused by the discrete nature of the response scores (Rost & von Davier, 1994). Furthermore, the Q-Index is based on the likelihood of observed response patterns and it utilizes the likelihood of an item pattern conditioning on the score of the item. This results in an item fit index that is essentially free of the item parameter.

Additionally, the Q-Index utilizes the concept behind the Guttman pattern. The Guttman pattern was named after sociologist Lois Guttman and it is sometimes called the dominance model (Van Schuur, 2011). The dominance model is also known as cumulative scale analysis, implicational scale analysis, and Guttman scaling. Guttman scaling is a type of unidimensional measurement. Lois Guttman purpose for this type of scale was to assess “attitudes” more specifically assess the morale of American soldiers in World War II. Currently, Guttman scale is still used for attitude scales. The idea behind Guttman scaling is to have a scale with dichotomous answers Yes/No to answer a set of questions which increase in specificity. The person answering the questions would advance a certain question and then stop when he or she no longer agrees (or disagrees) with the topic. For example, in a five item questionnaire regarding attitudes towards statistics, if a person gets to question three and then stops then the implication is that the person does not agree with questions four and five. Thus, the Guttman pattern this answer would create would look as follows: 11100. In a sample, people will choose different stopping points in the survey, which allows the ranking of their attitudes toward statistics.

Finally, the equation for the Q-Index index is as follows:

$$Q_i = \frac{\sum_v (x_{vi} - x_{v.G})\theta_v}{\sum_v (x_{v.A} - x_{v.G})\theta_v} \qquad (5)$$

Where the θ parameters can be estimated three different ways: (a) estimated by using all items, (b) estimated by using all items except i, or (c) using other tests which measure the same trait. The Guttman and anti-Guttman pattern response for each examinee, v, conditioned on the given item score distribution, is obtained by ordering examinees according to their ability level, θ as well as assigning the $n_o, n_1, \ldots n_m$ response categories $0, 1. \ldots, m$ to the examinees in either ascending or descending order (Rost & von Davier, 1994).

The Q-Index is available in the Rasch software Winmira and the Continuous Rating Scale Model program (CRSM) (von Davier, 2001; version 1.3; Müller, 1999). However, the Winmira software has not been updated since 2001, and the CRSM is only available upon request. The Q-Index is standardized, and ranges from 0 to 1 with a midpoint of .5. A value of 0 indicates perfect fit while a value of 1 indicates the item is misfitting (Rost & von Davier, 1994). The midpoint of .5 indicate the independence of the item and the latent trait. Rost and von Davier (1994) state that Q-Index is “derived for the ordinal Rasch model” unlike the majority of the current fit statistics which were developed for the dichotomous Rasch model (p. 174).

SAS CODE FOR Q-INDEX

First a simulated data set with 10 items and 100 persons was generated and entered into SAS using the following code (only the first five respondents are shown):

```
DATA sampledata;

   INPUT ID i1 i2 i3 i4 i5 i6 i7 i8 i9 i10;

CARDS;

p1 1101111110
p2 1110111010
p3 1111111010
p4 0111010101
p5 1010001010


RUN;
```

To estimate the Rasch model, the authors recommend the use of the %GLIMMIX_Rasch macro by Chen, Li, Kromrey (2013). Additionally, the SUGI paper by Pan and Chen (2011) describe how to obtain the Rasch model estimates using PROC LOGISTIC. The following code can be used to estimate the Q-Index by using the Chen, Li, Kromrey macro by including it within the SAS macro.

```
/* SAS  2: Q-Index */

%LET path=/folders/myshortcuts/Myfolders/QIndex;
LIBNAME orion "&path";
DM 'output; clear; log; clear;';
OPTIONS NODATE NONUMBER;
TITLE;

* Function to create Guttman and Anti-Guttman arrays;

%MACRO Create_G_and_AG (nitms= , nperson =);

* Calculate Guttman;
DATA Guttman;
SET IMPORT;
DROP ability;
run;

%DO j=1 %TO &nitms;
%LET v=i&j;
PROC SORT DATA=Gutman OUT=temp (KEEP=&v);
    BY &v;
RUN;

DATA Guttman;
    SET Guttman;
    SET temp;
RUN;
%END;

* Calculate Anti-Gutman;

DATA AGuttman;
SET IMPORT;
DROP ability;
RUN;

%DO j=1 %TO &nitms;
%LET v=i&j;
PROC SORT DATA=AGuttman OUT=temp (KEEP=&v);
    BY descending &v;
RUN;

DATA AGuttman;
    SET AGuttman;
    SET temp;
RUN;
%END;
%MEND;
```

```
/************************************
  Read a Excel Files into Item data
  The Example2Original must have
  the X items design matrix and the ability
  of persons
************************************/;


PROC IMPORT OUT = import
DATAFILE= "&path/Example2Original.xls"
DBMS=xls REPLACE;
RUN;

DATA one;
SET IMPORT;
person = RIGHT(person);
DROP ability;
RUN;

%Create_G_and_AG (nitms= 10, nperson =100);

* Extract ability from file;
DATA b;
SET IMPORT;
KEEP ability;
RUN;

* Sort X by ability;
DATA X;
SET IMPORT;
SET b;
RUN;

PROC SORT DATA=X;
BY ability;
RUN;

DATA X;
SET X;
DROP person ability;
RUN;

DATA G;
SET Gutman;
DROP person;
RUN;

DATA A;
SET AGutman;
DROP person;
RUN;

PROC SORT DATA=B;
```

```
by ability;
RUN;

PROC IML;
use X;
  read all VAR _ALL_ INTO X;
close X;
use G;
  read all VAR _ALL_ INTO G;
close G;
use A;
  read all VAR _ALL_ INTO A;
close A;
use B;
  read all VAR {ability} INTO B;
close Person_par;

Num = (X-G)#B;
Den = (A-G)#B;
Q = Num[+,]/Den[+,];
*print X, G, A, B, Num, Den, Q;
QT = T(Q);

TITLE "Q-index";
PRINT QT [LABEL="Q-Index"];

CREATE QIndex VAR {QT};    /** create data set **/
APPEND;                    /** write data in vectors **/
close QIndex;              /** close the data set **/
QUIT;

PROC PRINT DATA=QIndex;
RUN;
```

Table 1 represents the output given by SAS when calculating the Q-Index using the %GLIMMIX_Rasch macro and in the second column is the output of the Rasch specialized software Winmira (von Davier, 2001).

	SAS Q-Index	WINMIRA
Item 1	0.1776	0.1847
Item 2	0.2214	0.2299
Item 3	0.1407	0.1462
Item 4	0.2146	0.2217
Item 5	0.3671	0.3773
Item 6	0.1519	0.1563
Item 7	0.1323	0.1357
Item 8	0.1817	0.1850

Item 9	0.2195	0.2227
Item 10	0.2014	0.2020

Table 2 shows the side by side comparison of the estimation of the Q-Index via SAS and Winmira. However, this example data set was rating scale data with responses ranging from 1 to 5.

	SAS Q-Index	WINMIRA
Item 1	0.1751	0.1751
Item 2	0.1138	0.1138
Item 3	0.1257	0.1257
Item 4	0.1236	0.1236
Item 5	0.0848	0.0848
Item 6	0.1034	0.1034
Item 7	0.1205	0.1205
Item 8	0.1216	0.1216
Item 9	0.1146	0.1146
Item 10	0.0985	0.0985

CONCLUSION

There exist tools available to estimate the Rasch model in SAS such as Pan and Chen's (2011) PROC LOGISTIC tutorial and as we used in this paper the %GLIMMIX_Rasch macro. The latter provides the user with the calculation of popular fit statistics such as Infit, Outfit and their standardized form. The code provided in this paper utilizes the %GLIMMIX_Rasch macro in calculating yet another item fit statistic known as the Q-Index. As described in the literature review, the Q-Index has desirable characteristics in contrast with item fit statistics such as Infit and Outfit (or can be used in combination with these item fit statistics in order to make decisions during the item calibration stage). Furthermore, the Q-Index was specifically designed for ordinal data which has become more popular in recent years. The code proposed in this paper can be used for both dichotomous and rating scale data.

REFERENCES

References

Andersen, E. B. (1997). The rating scale model. In van der Linden, W J, & R. K. Hambleton (Eds.), (pp. 67-84). New York: Springer.

Andrich, D. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika, 47*(1), 105-113.

Baghaei, P., & Carstensen, C. H. (2013). Fitting the mixed Rasch model to a reading comprehension test: Identifying reader types. *Practical Assessment, Research & Evaluation, 5*(18) Retrieved from http://pareonline.net/getvn.asp?v=18&n=5

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences.* 711 Third Avenue New York, NY 10017: Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences* Springer.

Chen, Y-H, Li, I., Kromrey, J.D. (2013). (2013). GLIMMIX_RASCH: A SAS® macro for fitting the dichotomous Rasch model. Paper presented at the *Proceedings of the Annual South East SAS Users Group Conference,* Cary, NC: SAS Institute Inc. Retrieved from http://analytics.ncsu.edu/sesug/2013/SD-10.pdf

Christensen, K. B. (2013). Conditional maximum likelihood estimation in polytomous Rasch models using SAS. *ISRN Computational Mathematics, 2013*

Christensen, K. B. (2013). Conditional maximum likelihood estimation in polytomous Rasch models using sas. *ISRN Computational Mathematics, 2013*

Christensen, K. B., Kreiner, S., & Mesbah, M. (2013). *Rasch models in health.* Wiley Online Library.

Davier, M. V. (2001). WINMIRA: A program system for analyses with the Rasch model, with the latent class analysis and with the mixed Rasch model. *Kiel: IPN,*

Davier, M. v. (2001). Winmira 2001. *Kiel: Institute for Science Education,*

Eckes, T. (2011). Item banking for C-tests: A polytomous Rasch modeling approach. *Psychological Test and Assessment Modeling, 53*(4), 414-439.

Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences* Routledge.

Gustafsson, J. (1980). Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology, 33*(2), 205-233.

Janssen, K. C., Phillipson, S., O'Connor, J., & Johns, M. W. (2017). Validation of the Epworth sleepiness scale for children and adolescents using Rasch analysis. *Sleep Medicine, 33*, 30-35.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.

Müller, H. (1987). A Rasch model for continuous ratings. *Psychometrika, 52*(2), 165-181.

Müller, H. (1999). *Probabilistische testmodelle für diskrete und kontinuierliche ratingskalen: Einführung in die item-response-theorie für abgestufte und kontinuierliche items [probabilistic test models for* *discrete and continuous rating scales: An introduction to item response theory for graded and continuous items].* Bern: Huber:

Olsbjerg, M., & Christensen, K. B. (2015). % lRasch_mml: A SAS macro for marginal maximum likelihood estimation in longitudinal polytomous Rasch models. *Journal of Statistical Software, 67*(c02)

Ostini, R., & Nering, M. L. (2006). *Polytomous item response theory models* Sage.

Pan, T., & Chen, Y. (April, 2011). Using PROC LOGISTIC to estimate the Rasch model. Paper presented at the *Proceedings of the SAS Global Forum 2011 Conference.*

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.

Rost, J. (1988). Measuring attitudes with a threshold model drawing on a traditional scaling concept. *Applied Psychological Measurement, 12*(4), 397-409.

Van Schuur, W. H. (2011). *Ordinal item response theory: Mokken scale analysis.* Sage.

Wang, W., & Chen, C. (2005). Item parameter recovery, standard error estimates, and fit statistics of the WINSTEPS program for the family of Rasch models. *Educational and Psychological Measurement, 65*(3), 376-404.

Wright, B. D., & Masters,G.N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

Wright, B. D., & Stone, M. H. (1999). *Measurement essentials* (2nd ed.). Wilmington, DE: Wide Range.

Wu, M., & Adams, R. J. (2013). Properties of Rasch residual fit statistics. *Journal of Applied Measurement, 14*(4), 339-355.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Samantha Estrada
University of Northern Colorado
501 20th St, Greeley, CO 80639
970-351-2807
samantha.estrada@unco.edu

Rafael Perez Abreu Ph.D
Centro de Investigación en Matemáticas A.C.
Unidad Aguascalientes, Mexico, 20256
(449) 9185061
rabreu@cimat.mx
http://www.cimat.mx/es/Aguascalientes