Paper PO-10

GLIMMIX_Rasch: A SAS® Macro for Fitting the Dichotomous Rasch Model

Yi-Hsin Chen, University of South Florida, Tampa, FL
Isaac Li, University of South Florida, Tampa, FL
Jeffrey D. Kromrey, University of South Florida, Tampa, FL

ABSTRACT

For research areas in education, psychology, public health, sociology, etc., the Rasch measurement models provide a framework of design and analysis tools that differ from classical true score theory (CTST). General statistical software packages (e.g., SPSS or STATA) have been considered either not suitable or difficult to program for the purpose of implementing this complex approach. To obtain more accurate, valid, and reliable measurement, researchers often resort to specialized commercial Rasch programs at additional cost and have to spend time and efforts learning how to operate them. The latest version of SAS (SAS 9.3) procedure GLIMMIX can be employed to perform the Rasch measurement models easily. This article documents a SAS macro that fits the dichotomous Rasch model, estimates item and person parameters, and calculates unstandardized and standardized fit (INFIT and OUTFIT) indices associated with item difficulty and person ability. All parameter estimates and values for fit indices from the GLIMMIX_Rasch macro are almost identical to those from the WINSTEP software, a popular computer program for the Rasch models.

Keywords: RASCH, GLIMMIX, BASE SAS, SAS/STAT

INTRODUCTION

Since a measurement model proposed by Georg Rasch in 1960, the Rasch model, named after Georg Rasch, has been extensively used in various fields (e.g., education, psychology, public health, sociology) and expanded into a family of the Rasch models for diverse research purposes. For instance, the linear logistic test model (LLTM; Fischer, 1973) incorporates item stimulus features into consideration to predict item difficulties. The Rasch models can deal with dichotomous and polytomous data. The current paper is focused on the Rasch model for dichotomous data, called the dichtomous Rasch model. Mathematically speaking, the dichotomous Rasch model is a special case of non-linear mixed-effect models. The dichotomous Rasch model is a mixed-effect model because the item parameter (i.e., difficulty) is fixed but the person parameter is treated as random. Furthermore, this model is non-linear because the link function between the response probability and the parameters is non-linear (Sheu, Chen, Su, & Wang, 2005). For more details of the connection between common item response theory models and non-linear mixed-effect model, the reader is recommended to refer to Rijmen, Tuerlinckx, De Boeck, and Kuppens' (2003) article.

Several packages of commercial or free-of-charge computer software (e.g., BIGSTEPS, WINSTEPS, ConQuest, RASCAL, and RUMM2020) have been designed to conduct the dichotomous Rasch model. These diverse software packages have their own interfaces, use different estimation approaches, and create different statistics for examination of model-data fit. Some of them are recognized to be particularly hard to learn and use (Hays, Morales, & Reise, 2000). In contrast, SAS provides a popular and familiar platform to be used for estimating parameters and statistics of the dichotomous Rasch model (Sheu, Chen, Su, & Wang, 2005). The PROC NLMIXED procedure in SAS is a commonly used approach for many non-linear mixed-effect models, such as the dichotomous Rasch model as one of them. Its generality and flexibility make PROC NLMIXED more attractive to the large body of researchers and practitioners. Despite these advantages, the nature of PROC NLMIXED presents a shortcoming, which is the often time-consuming estimation algorithms. It is very common to take hours to obtain parameter estimates. To overcome this problem, the procedure of PROC GLIMMIX is an alternative.

The PROC GLIMMIX procedure is a new approach in SAS/STAT in the latest SAS 9.3 version. Originally, this function was provided through a macro in SAS. PROC GLIMMIX fits generalized linear mixed models (GLMMs) based on linearizations. The default estimation method in PROC GLIMMIX for models containing random effects is a technique known as restricted pseudo-likelihood (RPL) estimation. Both marginal quasi-likelihood (MQL) and penalized quasi-likelihood (PQL) can be implemented for GLMMs as well. PROC GLIMMIX extends the SAS mixed model tools in a number of ways including fitting models to multivariate data in which observations do not all have the same distribution or link. One of advantages of PROC GLIMMIX over PROC NLMIXED is the time-consuming issue. PROC GLIMMIX can save SAS running time 10 times over PROCNLMIXED. Thus, the focus of this study is on developing a SAS macro to fit a dichotomous Rasch model with PROC GLIMMIX. A comparison of parameter estimates from this SAS macro and from the commercial software called WINSTEP is made.

MATHEMATICAL EQUATIONS OF THE DICHOTOMOUS RASCH MODEL

The logistic function is applied for the dichotomous Rasch model and the equation is presented as follows (Embretson & Reise, 2000):

(1) $P_{ij} = P(X_{ij} = 1 \mid \theta_j, b_i) = \dfrac{\exp(\theta_j - b_i)}{1 + \exp(\theta_j - b_i)}$

(2) $Q_{ij} = 1 - P_{ij}$

where P_{ij} represents the probability of the correct response and Q_{ij} is the probability of the incorrect response, θ_j is the ability level of person j, and b_i is the item difficulty of item i.

The more general equation to represent the probability of the response (either correct or incorrect) for an item is:

(3) $P(U_{ij} = X \mid \theta_j, b_i) = P_{ij}^{X} Q_{ij}^{1-X}$

where X equals 1, the correct response, or 0, the incorrect response.

Thus, the probability of a response pattern for person j based on equations (1) and (3), conditional on person j ability level and the vector of item parameters is the product of the response probability of each item (i.e., local independence assumption), as follows:

(4) $P(U_{1j}, \ldots, U_{Ij} \mid \theta_j, \underline{b}) = \prod_{i=1}^{I} P_{ij}^{X} Q_{ij}^{1-X}$

For simplicity,

(5) $P(U_{1j}, \ldots, U_{nj} \mid \theta_j, \underline{b}) = P(\underline{U}_j \mid \theta_j, \underline{b}) = \prod_{i=1}^{I} P_{ij}^{X} Q_{ij}^{1-X}$

Mathematically, the probability of a particular response pattern for a random sample from the population (i.e., a normally distributed sample) can be derived by integrating Equation (5), as follows:

(6) $P(\underline{U}_j \mid \theta_j, \underline{b}) = \int_{\theta} \prod_{i=1}^{I} P_{ij}^{X} Q_{ij}^{1-X} g(\theta) \partial\theta$

Finally, the likelihood of the entire data can be obtained by multiplying the likelihood of a response pattern across persons, as follows:

(7) $L(U) = \prod_{j=1}^{J} P(\underline{U}_j \mid \theta_j, \underline{b}) = \prod_{j=1}^{J} \prod_{i=1}^{I} P_{ij}^{X} Q_{ij}^{1-X}$

UNSTANDARDIZED AND STANDARDIZED OUTFIT AND INFIT STATISTICS

In addition to estimates of item difficulty and person parameters, fit statistics, including unstandardized and standardized INFIT and OUTFIT statistics, for item difficulties and person ability estimates are used to screen misfit items and persons for dichotomous Rasch model analyses. The following are the brief descriptions of fit statistics and the necessary equations for this macro to compute these statistics (also refer to Wang & Chen, 2005).

The expected score of person j on item i for the dichotomous Rasch model can be represented as follows:

(8) $E_{ij} = \sum_{k=0}^{K-1} k \times P_{ijk}$ or $E_{ij} = P(X_{ij} = 1)$ for the dichotomous model

where k is the number of response categories and P_{ijk} is the probability of person j who scores k on item i. Let us use X_{ij} to represent the observed score of person j on item i. The variance of the observed score X_{ij} is

(9) $V_{ij} = \sum_{k=0}^{K-1} (k - E_{ij})^2 \times P_{ijk}$ or $V_{ij} = P(X_{ij} = 1) \times Q(X_{ij} = 0)$ for the dichotomous model

The score residual (Y_{ij}) is

(10) $Y_{ij} = X_{ij} - E_{ij}$

The standardized residual of an observed score X_{ij} is then

(11) $Z_{ij} = \frac{Y_{ij}}{\sqrt{V_{ij}}} = \frac{(X_{ij} - E_{ij})}{\sqrt{V_{ij}}}$

OUTFIT stands for the outlier-sensitive fit statistic and is defined as the average of the squared standardized residuals over the number of persons or items. Thus, OUTFIT is also called the unweighted mean squared error (MNSQ). ITEM OUTFIT can be obtained by averaging the squared standardized residuals over J persons and PERSON OUTFIT by averaging the squared standardized residuals over I items. The equations are as follows:

(12) $OUTFIT_{item\ i} = \frac{\sum_{j=1}^{J} Z_{ij}^2}{J}$ and $OUTFIT_{person\ j} = \frac{\sum_{i=1}^{I} Z_{ij}^2}{I}$, respectively.

Item and person OUTFIT statistics have an expected value of unity (1) and a variance of

(13) $s_{out\ i}^2 = \sum_{j=1}^{J} \frac{(C_{ij} / V_{ij}^2)}{J^2} - \frac{1}{J}$ and $s_{out\ j}^2 = \sum_{i=1}^{I} \frac{(C_{ij} / V_{ij}^2)}{I^2} - \frac{1}{I}$, respectively, where C_{ij} is the kurtosis of the observed score X_{ij} and C_{ij} is

(14) $C_{ij} = \sum_{k=0}^{K-1} (k - E_{ij})^4 \times P_{ijk}$

The unweighted mean squared error ($OUTFIT_{MNSQ}$) can be converted into the normally-distributed standardized *t* statistic by the Wilson-Hilferty cube root transformation.

(15) $t_{outfit\_item\ i} = (OUTFIT_{item\ i}^2 - 1)(3 / s_{out\ i}) + (s_{out\ i} / 3)$ and $t_{outfit\_person\ j} = (OUTFIT_{person\ j}^2 - 1)(3 / s_{out\ j}) + (s_{out\ j} / 3)$

INFIT stands for the inlier pattern sensitive fit statistic and is defined as the weighted mean squared error The equations of the INFIT statistics for item i and person j are:

(16) $INFIT_{item\ i} = \frac{\sum_{j=1}^{J} V_{ij} Z_{ij}^2}{\sum_{j=1}^{J} V_{ij}}$ and $INFIT_{person\ j} = \frac{\sum_{i=1}^{I} V_{ij} Z_{ij}^2}{\sum_{i=1}^{I} V_{ij}}$, respectively, which has an expected value of 0 and a variance of

(17) $s_{in\ i}^2 = \frac{\sum_{j=1}^{J} (C_{ij} - V_{ij}^2)}{(\sum_{j=1}^{J} V_{ij})^2}$ and $s_{in\ j}^2 = \frac{\sum_{i=1}^{I} (C_{ij} - V_{ij}^2)}{(\sum_{i=1}^{I} V_{ij})^2}$

The weighted mean squared error ($INFIT_{MNSQ}$) can be also converted into the normally-distributed standardized *t* statistic by the Wilson-Hilferty cube root transformation.

(18) $t_{infit\_item\,i} = (InFIT^{2}_{item\,i} - 1)(3/s_{in\,i}) + (s_{in\,i}/3)$ and $t_{infit\_person\,j} = (OUTFIT^{2}_{person\,j} - 1)(3/s_{in\,j}) + (s_{in\,j}/3)$

MACRO GLIMMIX_RASCH DETAILS

The GLIMMIX_Rasch SAS macro provides the estimates of item difficulty and person ability based on the dichotomous Rasch model. In addition, the unstandardized and standardized fit indices (INFIT and OUTFIT) mentioned above are also computed in this macro using PROC SQL. The GLIMMIX_Rasch macro is written in BASE SAS. For running the macro, an external dataset needs to be called in. The arguments to the macro involve the path of the external dataset (path=), the name of the dataset (datafile=), the length of ID in the dataset (idlen=), the starting column of the first item in the dataset (itmstart=), and the number of items (nitms=). The output tables that combine item difficulty or person ability estimates with their corresponding fit indices are provided from PROC PRINT. Using PROC MEANS, the means, standard deviations, minimum values, and maximum values of estimated parameters and fit indices are also presented in other tables through this macro.

```
/*generic session options;*/
options mprint;

%macro glimmix_rasch (path= , datafile= , idlen=, itmstart=, nitms= );

 * +----------------------------------------------------------------------------+
    call in the external dataset
 * +----------------------------------------------------------------------------+;
data one;
  infile "&path\&datafile" end = last;
  input person $ 1-&idlen @&itmstart (i01 - i&nitms) (1.);
  person = RIGHT(person);
  if last then call symputx('nperson',_n_);
run;
%put &nperson;

* +----------------------------------------------------------------------------+
   transpose the original data into the format for multilevel analysis
* +----------------------------------------------------------------------------+;
proc sort; by person;
proc transpose data=one out=longform name=item prefix=response;
  by person;
run;

* +----------------------------------------------------------------------------+
   create the design matrix (or indictors) for the dichotomous rasch model
* +----------------------------------------------------------------------------+;
data two;
   set longform;
    array items{&nitms} item1-item&nitms;
     do j=1 to &nitms;
      items{j}=0;
      items{compress(item,'i')}=1;
     end;
    rename response1 = response;
   drop j ;
run;

 * +----------------------------------------------------------------------------+
    use sas glimmix to run the dichotomous rasch model
    use output and ods to obtain the statistics and parameters we want
 * +----------------------------------------------------------------------------+;
proc glimmix data = two noclprint noitprint noprofile;
  class item person;
  model response(descending) = item person/ s noint link=logit dist=binary;
  lsmeans item / cl ilink;
  lsmeans person / cl ilink;
  output out=fit_statistics pred(ilink)=prob resid(ilink)=residual variance(ilink)=variance;
  ods output lsmeans=lsmeans_item_effects;
```

```
run;

* +--------------------------------------------------------------------------+
   extract and create the output for item difficulty
* +--------------------------------------------------------------------------+;

* obtain the mean of item difficulty;
proc means noprint data=lsmeans_item_effects;
  var estimate;
  where effect = 'item';
  output out=a mean=mn_estimate;
run;

* obtain item parameter estimates;
data item_par;
  if _n_ = 1 then set a;
  retain mn_estimate;
  length item $ 4;
  set lsmeans_item_effects;
   if effect = 'item';
   difficulty= (estimate-mn_estimate)*-1; *make the mean of item difficulty 0;
   se_d=stderr;
   ci_lower=lower;
   ci_upper=upper;
   prop_correct=mu;
   se_p=stderrmu;
   keep item difficulty se_d prop_correct se_p ci_lower ci_upper;
run;

* +--------------------------------------------------------------------------+
   extract and create the output for person parameters
* +--------------------------------------------------------------------------+;
data person_par;
  set lsmeans_item_effects;
  if effect='person';
  ability=estimate*1;
  se_a=stderr*1;
  prop_correct=mu;
  se_p=stderrmu;
  ci_lower=lower;
  ci_upper=upper;
  keep person ability se_a prop_correct se_p ci_lower ci_upper;

* +--------------------------------------------------------------------------+
   create and compute item and person infit and outfit statistics
* +--------------------------------------------------------------------------+;
data fit_statistics;
  set fit_statistics;
  p=prob;
  q=1-p;/*probability of failure*/
  var=variance; /*equals pxq only for binary models*/
  std_resid=residual/sqrt(variance);
  kur=p**4*q+q**4*p;/*kurtosis*/
run;

 * compute item fit statistics;
proc sql;
  create table item_fit_stats as
  select person, item format=$4. length = 4, response, residual, variance, var, std_resid, kur,
         (sum(kur-var**2))/(sum(var))**2 as var_infit, /*varaince infit mean squares*/
         sum(var*std_resid**2)/sum(var) as ms_infit, /*mean square infit*/
         (calculated ms_infit**(1/3)-1)*(3/sqrt(calculated var_infit))+(sqrt(calculated
var_infit)/3) as zstd_infit,/*infit standardized z*/
         sum(kur/var**2)/(&nperson**2)-(1/&nperson) as var_outfit, /*variance of item outfit mean
squares*/
         sum(std_resid**2)/&nperson as ms_outfit, /*mean square outfit*/
         (calculated ms_outfit**(1/3)-1)*(3/calculated var_outfit**(1/2))+(calculated
var_outfit**(1/2)/3) as zstd_outfit /*item outfit standardized z (zstd)*/
  from fit_statistics
  group by item;
quit;
```

```
* compute person fit statistics;
proc sql;
  create table person_fit_stats as
  select person, item format=$4. length = 4, response, residual, variance, var, std_resid, kur,
         sum(kur-var**2)/sum(var)**2 as var_infit_person, /*varaince infit mean squares for
person;*/
         sum(var*std_resid**2)/sum(var) as ms_infit_person, /*estimated person infit mean squares
for person*/
         (calculated ms_infit_person**(1/3)-1)*(3/calculated var_infit_person**(1/2))+(calculated
var_infit_person**(1/2)/3) as zstd_infit_person, /*person infit standardized z (zstd)*/
         sum(kur/(var)**2)/(&nitms**2)-(1/&nitms) as var_outfit_person, /*variance outfit mean
squares for person*/
         sum(std_resid**2)/&nitms as ms_outfit_person, /*estimated person outfit mean squares*/
         (calculated ms_outfit_person**(1/3)-1)*(3/calculated
var_outfit_person**(1/2))+(calculated var_outfit_person**(1/2)/3) as zstd_outfit_person
/*estimated person outfit standardized z (zstd)*/
  from fit_statistics
  group by person;
quit;

* +-----------------------------------------------------------------------------+
   create a table of item related parameters and statistics
* +-----------------------------------------------------------------------------+;

*merge two item-related statistics;
proc sort data=item_fit_stats nodupkey;  by item;
data item_statistics;
  merge item_par item_fit_stats;
  by item;
keep item difficulty se_d ms_infit zstd_infit ms_outfit zstd_outfit;
*generate report table;
proc print data=item_statistics;
       format difficulty se_d ms_infit zstd_infit ms_outfit zstd_outfit 5.2;
       title 'This is a table for item related parmaters and statistics';
proc means data=item_statistics maxdec=2;
       title 'The means of item related parameters and Statistics';
run;

* +-----------------------------------------------------------------------------+
   create a table of person related parameters and statistics
* +-----------------------------------------------------------------------------+;

*merge two person-related statistics;
proc sort data=person_fit_stats nodupkey; by person;
data person_statistics;
  merge person_par person_fit_stats;
  by person;
keep person ability se_a ms_infit_person zstd_infit_person ms_outfit_person zstd_outfit_person;
*generate report tabls;
proc print data=person_statistics;
  format ability se_a ms_infit_person zstd_infit_person ms_outfit_person zstd_outfit_person 5.2;
  title 'This is a table for person related parameters and Statistics';
run;
proc means data=person_statistics maxdec = 2;
  title 'The means of person related parameters and Statistics';
run;
%mend;
```

INVOKING THE MACRO

To demonstrate how to use this macro, a widely used dataset from Kikumi Tatsuoka's (1990) fraction subtraction is used. The fraction subtraction dataset consists of 20 items with 536 examinees. The data are available at www.blackwellpublishing.com/rss/Volumes/Cv51p3.htm from the *Journal of the Royal Statistical Society*. It should be noted that based on the Rasch model the examinees with 0% and 100% of the items correct and the items with 0% and 100% of the examinees correct will not provide any contribution to parameter estimation. Thus, these examinees and items need to be removed from the dataset. The researcher should screen the data before using this macro. In our case, there are 43 examinees and no items removed from the dataset. A total of 493 examinees remain in the data.

The data format in the macro is set up as follows. There is no space between items.

```
input person $ 1-&idlen @&itmstart (i01 - i&nitms) (1.);
```

```
10211101111111111111111
10300001100100001010000
10401111111111111111111
```

If the researcher has a dataset that contains a space between items, what he/she needs to modify in the macro is to change the item format (1.) located at the end of the input command to (2.).

```
input person $ 1-&idlen @&itmstart (i01 - i&nitms) (2.);
```

```
102 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
103 0 0 0 0 1 1 0 0 1 0 0 0 0 1 0 1 0 0 0 0
104 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

The arguments used for the fraction subtraction data are shown below. The arguments indicate that the data are located in the folder called “GLIMMIX” in the C drive (path=c:\GLIMMIX) and the file name is called “fractiondata” (datafile=fractiondata). The first three columns contain examinees’ ID number (idlen=3). The first item starts at the fourth column in the dataset (itmstart=4) and the total number of items is 20 (nitms=20).

```
%glimmix_rasch (path=C:\GLIMMIX, datafile=fractiondata, idlen=3, itmstart=4,
nitms=20);
```

OUTPUT FROM THE MACRO

There are four tables created from the macro: two for item estimates and two for person estimates. The content of Output 1 includes item difficulty (difficulty), standard error of estimates for item difficulty (se_d), unstandardized INFIT (ms_infit), standardized INFIT (zstd_infit), unstandardized OUTFIT (ms_outfit), and standardized OUTFIT (zstd_outfit). Output 3 presents the same content regarding persons and there are only 20 examinees shown in the table. Outputs 2 and 4 present the average values of item and person parameters and their corresponding fit indices, respectively.

Obs	item	difficulty	se_d	ms_infit	zstd_ infit	ms_outfit	zstd_ outfit
1	i01	0.03	0.13	0.95	-0.71	0.99	-0.02
2	i02	-0.35	0.13	0.73	-4.42	0.67	-2.57
3	i03	0.16	0.13	0.91	-1.37	0.84	-1.16
4	i04	0.06	0.13	1.26	3.54	1.93	5.43
5	i05	-0.41	0.13	1.51	6.73	1.81	4.53
6	i06	-2.32	0.15	0.75	-3.63	1.44	1.23
7	i07	1.14	0.13	0.94	-0.81	0.69	-1.87
8	i08	-1.85	0.14	1.36	4.70	1.82	2.49
9	i09	-0.94	0.13	1.78	9.90	2.59	6.33
10	i10	1.28	0.13	0.84	-2.41	0.71	-1.67
11	i11	0.51	0.13	0.80	-3.11	0.87	-0.88
12	i12	-1.67	0.13	0.95	-0.74	1.37	1.40
13	i13	2.07	0.14	0.92	-1.26	0.55	-1.85
14	i14	-1.60	0.13	0.81	-3.09	1.24	1.01
15	i15	0.70	0.13	0.85	-2.18	0.83	-1.16
16	i16	-1.36	0.13	0.93	-1.03	1.17	0.83
17	i17	0.85	0.13	0.77	-3.50	0.60	-2.96
18	i18	0.55	0.13	0.98	-0.34	0.94	-0.37
19	i19	1.93	0.13	0.79	-3.39	0.47	-2.48
20	i20	1.20	0.13	0.78	-3.38	0.63	-2.26

Output 1. Example Output of Item Parameters and Fit Statistics from the Rasch_GLIMMIX Macro

Variable	N	Mean	Std Dev	Minimum	Maximum
difficulty	20	0.00	1.29	-2.32	2.07
se_d	20	0.13	0.01	0.13	0.15
ms_infit	20	0.98	0.28	0.73	1.78
zstd_infit	20	-0.52	3.81	-4.42	9.90
ms_outfit	20	1.11	0.56	0.47	2.59
zstd_outfit	20	0.20	2.72	-2.96	6.33

Output 2. Example Output of the Means of Item Parameters and Fit Statistics from the Rasch_GLIMMIX Macro

Obs	person	ability	se_a	ms_infit_ person	zstd_ infit_ person	ms_outfit_ person	zstd_ outfit_ person
1	1	0.57	0.52	1.76	2.94	2.28	2.65
2	2	2.73	0.79	1.34	0.75	2.03	1.06
3	3	-0.24	0.52	1.13	0.62	1.06	0.30
4	4	1.14	0.55	0.90	-0.35	0.89	-0.05
5	5	-1.81	0.63	0.79	-0.58	0.62	-0.38
6	6	-1.81	0.63	0.67	-1.00	0.41	-0.85
7	7	-2.79	0.80	1.01	0.20	0.55	-0.05
8	8	-2.79	0.80	1.11	0.39	1.09	0.48
9	9	-2.24	0.69	0.97	0.05	0.52	-0.34
10	10	-1.81	0.63	1.00	0.10	0.65	-0.33
11	11	0.85	0.53	0.73	-1.25	0.58	-1.00
12	12	1.45	0.57	1.05	0.27	1.58	1.01
13	13	-1.11	0.56	1.07	0.32	1.12	0.41
14	14	-2.24	0.69	1.03	0.21	0.82	0.09
15	15	-2.24	0.69	0.82	-0.36	0.46	-0.46
16	16	-1.81	0.63	0.90	-0.19	1.35	0.68
17	17	-1.81	0.63	1.00	0.12	0.81	-0.05
18	18	-2.24	0.69	1.09	0.35	0.69	-0.08
19	19	-1.81	0.63	0.71	-0.87	0.43	-0.80
20	20	-3.61	1.06	0.93	0.20	0.34	0.14
....							

Output 3. Example Output of Person Parameters and Fit Statistics from the Rasch_GLIMMIX Macro

Variable	N	Mean	Std Dev	Minimum	Maximum
ability	493	0.15	2.12	-3.61	3.54
se_a	493	0.69	0.18	0.52	1.06
ms_infit_person	493	1.00	0.26	0.45	2.13
zstd_infit_person	493	0.02	0.96	-2.41	3.91
ms_outfit_person	493	1.11	1.11	0.23	12.79
zstd_outfit_person	493	0.12	0.96	-2.07	3.68

Output 4. Example Output of the Means of Person Parameters and Fit Statistics from the Rasch_GLIMMIX Macro

A COMPARISON OF GLIMMIX ESTIMATION WITH THE WINSTEP SOFTWARE

The software called WINSTEP is the most popular program for the Rasch models. The following two tables show the outputs for item- and person-related parameter estimates and fit indices from the WINSTEP software. “MEASURE” is item difficulty or person ability. “MODEL s.e.” represents the standard error of estimates for item difficulty or person ability. “MNSQ” and “ZSTD” represent unstandardized and standardized indices, respectively. At the bottom of the tables are the average values of these parameters and fit statistics. As you can see, the values from GLIMMIX_Rasch macro are almost identical to those from the WINSTEP software.

ENTRY NUMBER	RAW SCORE	COUNT	**MEASURE**	**MODEL S.E.**	**INFIT MNSQ**	**INFIT ZSTD**	**OUTFIT MNSQ**	**OUTFIT ZSTD**	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	fraction
1	255	493	.03	.12	.95	-.7	.99	.0	.69	85.0	81.5	I0001
2	279	493	-.34	.12	.73	-4.4	.67	-2.6	.75	88.2	81.4	I0002
3	246	493	.16	.12	.91	-1.4	.84	-1.2	.71	84.4	81.7	I0003
4	253	493	.06	.12	1.26	3.5	1.92	5.4	.57	77.3	81.6	I0004
5	283	493	-.41	.12	1.51	6.7	1.80	4.5	.50	70.8	81.4	I0005
6	396	493	-2.31	.14	.75	-3.7	1.43	1.3	.59	91.1	85.5	I0006
7	184	493	1.14	.13	.94	-.8	.69	-1.9	.69	80.1	82.2	I0007
8	371	493	-1.84	.13	1.35	4.7	1.81	2.5	.46	75.9	83.7	I0008
9	317	493	-.94	.13	1.78	9.9	2.57	6.3	.37	65.1	81.4	I0009
10	175	493	1.28	.13	.84	-2.4	.70	-1.7	.70	85.2	82.4	I0010
11	224	493	.51	.12	.80	-3.1	.87	-.9	.73	87.0	81.8	I0011
12	361	493	-1.66	.13	.95	-.8	1.36	1.4	.60	85.6	83.1	I0012
13	128	493	2.07	.13	.91	-1.3	.55	-1.9	.63	84.0	83.2	I0013
14	357	493	-1.59	.13	.81	-3.1	1.24	1.0	.64	88.8	82.9	I0014
15	212	493	.69	.13	.85	-2.2	.83	-1.2	.72	84.0	82.0	I0015
16	343	493	-1.36	.13	.93	-1.0	1.17	.8	.63	84.2	82.1	I0016
17	202	493	.85	.13	.77	-3.5	.59	-3.0	.74	87.8	82.0	I0017
18	221	493	.55	.12	.97	-.4	.94	-.4	.68	84.0	81.8	I0018
19	136	493	1.93	.13	.79	-3.4	.47	-2.5	.68	86.4	83.1	I0019
20	180	493	1.20	.13	.78	-3.4	.63	-2.3	.72	86.2	82.3	I0020
MEAN	256.2	493.0	.00	.13	.98	-.5	1.10	.2		83.1	82.4	
S.D.	78.1	.0	1.25	.00	.27	3.7	.55	2.7		6.2	1.0	

Output 5. Example Output of Item Parameters and Fit Statistics from the WINSTEP

ENTRY NUMBER	RAW SCORE	COUNT	**MEASURE**	**MODEL S.E.**	**INFIT MNSQ**	**INFIT ZSTD**	**OUTFIT MNSQ**	**OUTFIT ZSTD**	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	middle
1	12	20	.57	.52	1.76	2.9	2.27	2.6	-.07	45.0	71.7	1
2	18	20	2.73	.78	1.33	.7	2.02	1.1	-.05	90.0	90.0	2
3	9	20	-.24	.52	1.13	.6	1.06	.3	.43	65.0	73.1	3
4	14	20	1.14	.55	.90	-.3	.89	-.1	.50	70.0	74.5	4
5	4	20	-1.81	.63	.79	-.6	.62	-.4	.58	90.0	81.3	5
6	4	20	-1.81	.63	.67	-1.0	.41	-.8	.67	90.0	81.3	6
7	2	20	-2.78	.80	1.01	.2	.55	.0	.37	90.0	90.0	7
8	2	20	-2.78	.80	1.11	.4	1.09	.5	.24	90.0	90.0	8
9	3	20	-2.24	.69	.97	.0	.52	-.3	.47	80.0	85.2	9
10	4	20	-1.81	.63	.99	.1	.64	-.3	.49	80.0	81.3	10
11	13	20	.85	.53	.73	-1.2	.58	-1.0	.66	85.0	72.8	11
12	15	20	1.45	.57	1.05	.3	1.57	1.0	.33	75.0	77.6	12
13	6	20	-1.11	.56	1.06	.3	1.12	.4	.44	75.0	77.5	13
14	3	20	-2.24	.69	1.03	.2	.82	.1	.39	80.0	85.2	14
15	3	20	-2.24	.69	.82	-.4	.46	-.5	.54	90.0	85.2	15
16	4	20	-1.81	.63	.90	-.2	1.34	.7	.44	80.0	81.3	16
17	4	20	-1.81	.63	1.00	.1	.81	.0	.44	90.0	81.3	17
18	3	20	-2.24	.69	1.09	.3	.69	-.1	.38	80.0	85.2	18
19	4	20	-1.81	.63	.71	-.9	.43	-.8	.65	80.0	81.3	19
20	1	20	-3.60	1.06	.93	.2	.34	.1	.34	95.0	95.0	20
....					...		...		...	...		...
MEAN	10.4	20.0	.15	.69	1.00	.0	1.10	.1		83.1	82.4	
S.D.	6.1	.0	2.12	.18	.26	1.0	1.04	1.0		11.1	7.9	

Output 6. Example Output of Person Parameters and Fit Statistics from the WINSTEP

CONCLUSION

Since the family of the Rasch models has been widely used in various research fields and the commercial software for these models sometimes is not available or is difficult to use, this macro provides an easy access to conducting the dichotomous Rasch model analysis for researchers and practitioners using the popular statistical software, SAS. The GLIMMIX_Rasch macro yields item and person parameter estimates. It also computes unstandardized and standardized INFIT and OUTFIT indices for evaluating item and person fit. All parameter estimates and values for fit indices from the GLIMMIX_Rasch macro are almost identical to those from the WINSTEP software, a popular computer program for the Rasch models. Furthermore, in SAS both GLIMMIX and NLMIXED procedures can be used for the analyses of the Rasch models. However, one of the GLIMMIX's advantages over the NLMIXED is time-saving. For instance, the time used in this macro for the data with 20 items and around 500 examinees was just two seconds using a computer with CPU of 3.40 GHz and memory of 8 GB.

REFERENCES

Fischer, G. H. (1973). Linear logistic test model as an instrument in educational research. *Acta Psychologica, 37*, 359-374.

Sheu, C.-F., Chen, C.-T., Su, Y.-H., & Wang, W.-C. (2005). Using SAS PROC NLMIXED to fit item response theory models. *Behavior Research Methods, 37*, 202-218.

Rijmen, F., Tuerlinckx, F., De Boeck, P. & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods, 8,* 185-205.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care, 38(suppl II)*, II-28–II-42.

Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 453-488). Hillsdale, NJ: Erlbaum.

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*, 376-404.

CONTACT INFORMATION

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

Yi-Hsin Chen
University of South Florida
4202 East Fowler Ave., EDU 105
Tampa, FL 33620
Fax: (813) 974-4495
Email: ychen5@usf.edu