

# Package ‘iarm’

October 13, 2022

**Type** Package

**Title** Item Analysis in Rasch Models

**Version** 0.4.3

**Description** Tools to assess model fit and identify misfitting items for Rasch models (RM) and partial credit models (PCM). Included are item fit statistics, item characteristic curves, item-restscore association, conditional likelihood ratio tests, assessment of measurement error, estimates of the reliability and test targeting as described in Christensen et al. (Eds.) (2013, ISBN:978-1-84821-222-0).

**License** GPL-2

**Depends** R (>= 3.5.0), eRm, ggplot2

**Imports** gridExtra, Hmisc, psychotools, stats, utils, vcdExtra,

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.2.1

**NeedsCompilation** no

**Author** Marianne Mueller [aut, cre],  
Pedro Henrique Ribeiro Santiago [ctb]

**Maintainer** Marianne Mueller <marianne.mueller@math.ethz.ch>

**Repository** CRAN

**Date/Publication** 2022-08-27 13:00:02 UTC

## R topics documented:

iarm-package . . . . .	2
amts . . . . .	3
boot_fit . . . . .	4
clr_tests . . . . .	4
desc2 . . . . .	5
ICCplot . . . . .	6
item_obsexp . . . . .	8
item_restscore . . . . .	9

item_target . . . . .	10
out_infit . . . . .	11
partgam . . . . .	12
partgam_DIF . . . . .	13
partgam_LD . . . . .	14
person_estimates . . . . .	15
print.bootfit . . . . .	16
print.outfit . . . . .	16
score_groups . . . . .	17
test_prop . . . . .	17

<b>Index</b>	<b>19</b>
--------------	-----------

---

iarm-package	<i>iarm: A package for item analysis in Rasch models</i>
--------------	--

---

## Description

Tools to assess model fit and identify misfitting items for Rasch models (RM) and partial credit models (PCM). Included are item fit statistics, item-restscore association, conditional likelihood ratio tests, assessment of measurement error, estimates of the reliability and test targeting.

## Item Fit statistics

Item fit statistics are used to assess whether individual items fit the Rasch model. Outfit and infit mean squares are well-known and much used statistics. They summarize standardized response residuals comparing observed responses to items to the expected responses. To avoid bias expected responses are calculated under the conditional distribution of responses given the total score. Parametric bootstrapping is used to assess the significance of misfitting items. The item restscore gamma coefficient is used to assess differential item discrimination.

## Conditional likelihood ratio tests (CLR)

The conditional likelihood ratio test of Andersen is an overall test of fit of data to the model. The test compares conditional maximum likelihood estimates of item parameters in different subgroups to the estimates for the complete sample of persons. Subgroups are defined by outcomes of the total score (test of homogeneity) or by outcomes of an exogenous variable (test of no differential item functioning, DIF).

## References

- Andersen, E. B. (1973) A goodness of fit test for the Rasch model. *Psychometrika*, 38, 123-140.
- Kreiner, S. & Christensen, K. B. (2011) Exact evaluation of Bias in Rasch model residuals. *Advances in Mathematics Research*, 12, 19-40.
- Mueller, M. & Kreiner, S. (2015) Item Fit Statistics in Common Software for Rasch Analysis. Research Report 15-06, Department of Biostatistics, University of Copenhagen.

---

amts

*Abbreviated Mental Test Score (AMTS)*

---

### Description

A dataset containing the responses of 197 persons to the ten questions of the Abbreviated Mental Test Score (AMTS). The AMTS is used to identify patients with dementia. One point is given for each correct answer, a score of 6 or less suggests that the patient has some mental impairment.

### Format

A data frame with 197 rows and 13 variables.

**id** id number of the patient.

**agegrp** a factor with levels 16-65, 66-75, 76-85, 86+ for the age of the patient.

**sex** a factor with levels male, female of the patient.

**age** age of patient, with 1 if the respondent knows his/her own age and 0 otherwise.

**time** time (nearest hour), with 1 if correct and 0 otherwise.

**address** address, with 1 if correct and 0 otherwise.

**name** name of hospital (or area of town if at home) , with 1 if correct and 0 otherwise.

**year** current year, with 1 if correct and 0 otherwise.

**dob** date of birth of patient, with 1 if correct and 0 otherwise.

**month** month, with 1 if correct and 0 otherwise.

**firstww** date of first world war, with 1 if correct and 0 otherwise.

**monarch** name of monarch, with 1 if correct and 0 otherwise.

**countbac** count backwards 20-1, with 1 if correct and 0 otherwise.

### References

Slade, A., Fear, J. & Tennant, A. (2006) Identifying patients at risk of nursing home admission: The Leeds Elderly Assessment Dependency Screening tool (LEADS). *BMC Health Services Research*, 6:31.

### Examples

```
data(amts)
str(amts)
```

---

<code>boot_fit</code>	<i>Computes Bootstrapping P Values for Outfit and Infit Statistics</i>
-----------------------	--

---

**Description**

Computes Bootstrapping P Values for Outfit and Infit Statistics

**Usage**

```
boot_fit(
  object,
  B,
  p.adj = c("BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none")
)
```

**Arguments**

<code>object</code>	an object of class "Rm" (output of RM or PCM) or class "pcmodel"
<code>B</code>	Number of replications.
<code>p.adj</code>	Correction method for multiple testing. The methods are "BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none". See <a href="#">p.adjust</a> .

**Value**

object of class bootfit with outfit and infit statistics and corresponding p values.

---

<code>clr_tests</code>	<i>Conditional Likelihood Ratio Tests (CLR)</i>
------------------------	---

---

**Description**

The conditional likelihood ratio tests compare item parameters in low and high score groups for an overall test of homogeneity, and in groups defined by the levels of exogenous factors for tests of no differential item functioning (DIF).

**Usage**

```
clr_tests(dat.items, dat.exo = NULL, model = c("RM", "PCM"))
```

**Arguments**

<code>dat.items</code>	A data frame with the responses to the items.
<code>dat.exo</code>	A single factor or a data frame consisting of one or more exogenous factor variables.
<code>model</code>	If <code>model="RM"</code> a Rasch model will be fitted, if <code>model="PCM"</code> a partial credit model for polytomous items is used.

**Value**

matrix with test statistics, df and p values.

**Author(s)**

Marianne Mueller

**References**

Andersen, E.B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, 38, 123-140.

**Examples**

```
#CLR overall test and test of no DIF for agegrp and sex
clr_tests(ams[,4:13],ams[,2:3])
```

---

desc2

*Depression Screening DESC-II*

---

**Description**

A dataset containing the responses of 799 patients (indication group psychiatry, otolaryngology, cardiology, neurology) to the short form DESC-II with 10 items. There are 5 response categories from 0 = never to 4 = always. A higher score is supposed to mean a higher depression.

**Format**

A data frame with 799 rows and 14 variables.

**code** id number of the patient

**group** a factor with levels psychiatry, otolaryngology, cardiology, neurology for the indication group of the patient.

**gender** a factor with levels female, male of the patient.

**agegroup** a factor with levels 18-34, 35-49, 50-59, 60-87 for the age of the patient.

**DESC\_2\_1** feeling not to be needed

**DESC\_2\_2** loss of interest in other people

**DESC\_2\_3** disheartened

**DESC\_2\_4** no pleasure doing things

**DESC\_2\_5** feeling to be no good

**DESC\_2\_6** uninspired

**DESC\_2\_7** pessimistic

**DESC\_2\_8** discouraged

**DESC\_2\_9** withdrawal

**DESC\_2\_10** thinking of taking one's life

## References

Forkmann et al. (2009) Development and validation of the Rasch-based Depression Screening (DESC) using Rasch analysis and structural equation modelling. *J Behav Ther Exp Psychiatry*, 40(3): 468-78.

## Examples

```
data(desc2)
str(desc2)
```

---

 ICCplot

---

*Item Characteristic Curves*


---

## Description

Plots Item Characteristic Curves for dichotomous and polytomous items. The plot can display observed scores as total scores (method="score") or as average scores within adjacent class intervals (method="cut"). Class intervals can be useful when the sample size is not large enough to contain an adequate number of respondents with the same total score for each possible total score. The function includes the option to plot observed scores according to values of an exogenous variable to evaluate differential item functioning (dif="yes").

## Usage

```
ICCplot(
  data,
  itemnumber,
  pallete = "Paired",
  xticks = 1,
  yticks = 0.5,
  thetain = -6,
  thetaend = 6,
  method = "score",
  grid = "yes",
  cinumber = 6,
  itemdescrip = "",
  axis.rumm = "yes",
  dif = "no",
  difvar = NA,
  diflabels = c("Group1", "Group 2", "Group 3", "Group 4", "Group5"),
  difstats = "yes",
  title = "Item Characteristic Curve",
  icclabel = "yes",
  xaxistitle = "Theta",
  yaxistitle = "Item Score"
)
```

**Arguments**

<code>data</code>	An object of class "data.frame" containing the items (include all items present in the model). The variables need to be numeric.
<code>itemnumber</code>	A numeric vector indicating the columns of the data (the items) which ICCs are going to be plotted. Maximum of four items per plot.
<code>pallette</code>	An object of class "character". Choose a pre-made color pallette from package RColorBrewer. Only available for <code>dif="no"</code> .
<code>xticks</code>	A numeric scalar. Specify x-axis tick values.
<code>yticks</code>	A numeric scalar. Specify y-axis tick values.
<code>thetain</code>	A numeric scalar. Specify minimum theta values for person parameters.
<code>thetaend</code>	A numeric scalar. Specify maximum theta values for person parameters.
<code>method</code>	The method for displaying observed scores. Choose "score" to plot total scores. Choose "cut" to plot class intervals.
<code>grid</code>	Chooses whether the background grid should be displayed. Options are "yes" or "no".
<code>cinumber</code>	A numeric scalar. The number of adjacent class intervals in which participants will be divided. Notice that the number of class intervals cannot be higher than the number of total scores.
<code>itemdescrip</code>	A character vector indicating the description of the plotted items. Maximum of four descriptions (one description per item plotted).
<code>axis.rumm</code>	Configures whether the plot should display the entire trait range or solely the trait range close to the observed scores (similar to private software RUMM2030). Options are "yes" or "no".
<code>dif</code>	Configures whether the observed scores will be plotted according to values of an exogenous variable to evaluate differential item function. Options are "yes" or "no".
<code>difvar</code>	Chooses the variable which will be used to evaluate differential item functioning. Only necessary when <code>dif="yes"</code> .
<code>diflabels</code>	A character vector indicating the labels to values of the variable chosen to evaluate differential item functioning. Only necessary when <code>dif="yes"</code> .
<code>difstats</code>	Displays the partial gamma coefficient to indicate the magnitude of differential item functioning. Options are "yes" or "no". Only necessary when <code>dif="yes"</code> .
<code>title</code>	A character vector. The title of the plot.
<code>icclabel</code>	Displays the labels of Expected Item Score and Observed Item Score. Options are "yes" or "no".
<code>xaxistitle</code>	A character vector. The x-axis title.
<code>yaxistitle</code>	A character vector. The y-axis title.

**Author(s)**

Pedro Henrique Ribeiro Santiago <pedro.ribeirosantiago@adelaide.edu.au>, Marianne Mueller

## Examples

```
## Not run: # Creates a plot for Item 1 using total scores
ICCplot(desc2[,5:13], itemnumber=1, method="score", itemdescrip="Item 1")

# Creates a plot for Item 1 using 8 class intervals
ICCplot(desc2[,5:13], itemnumber=1, method="cut", cinumber=8, itemdescrip="Item 1")

# Creates a plot for Item 1 using 8 class intervals without RUMM style axis
ICCplot(desc2[,5:13], itemnumber=1, method="cut", cinumber=8, itemdescrip="Item 1", axis.rumm="no")

# Creates a plot for Item 3 using 8 class intervals and evaluating DIF according to gender
ICCplot(desc2[,5:13], itemnumber=3, method="cut", cinumber=8, itemdescrip="Item 3",
dif="yes", difvar=desc2$gender, diflabels=c("Men", "Women"))

# Creates a plot with three items using 5 class intervals and evaluating DIF according to gender
ICCplot(desc2[,5:13], itemnumber=1:3, method="cut", cinumber=5,
itemdescrip=c("Item 1","Item 2","Item 3"), dif="yes"
difvar=desc2$gender, diflabels=c("Men", "Women"))

## End(Not run)
```

---

item\_obsexp

*Observed and Expected Item Mean Scores*

---

## Description

Homogeneity of item responses in the low and high score groups is analyzed by looking at observed and expected item mean scores together with standardized residuals. If the Andersen's CLR test has shown some evidence against homogeneity, this comparison can indicate which items might be responsible.

## Usage

```
item_obsexp(object)
```

## Arguments

**object** An object of class "Rm", a fitted Rasch model or partial credit model using the functions RM or PCM in package eRm, or an object of class "pcmodel", a fitted partial credit model using the function pcmodel in package psychotools.

## Value

list with observed and expected mean scores together with standardized residuals for the two score groups.

## Author(s)

Marianne Mueller



**Examples**

```

rm.mod <- RM(amts[,4:13])
item_obsexp(rm.mod)
## Not run:
pc.mod <- PCM(desc2[,5:14])
item_obsexp(pc.mod)

## End(Not run)

```

---

item_restscore	<i>Item Restscore Association</i>
----------------	-----------------------------------

---

**Description**

The observed Gamma coefficient between the score of a single item and the total score of the remaining items is compared with the corresponding expected Gamma coefficient under the Rasch model.

**Usage**

```

item_restscore(
  object,
  p.adj = c("BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none")
)

```

**Arguments**

object	An object of class "Rm", a fitted Rasch model or partial credit model using the functions RM or PCM in package eRm, or an object of class "pcmodel", a fitted partial credit model using the function pcmodel in package psychotools.
p.adj	Correction method for multiple testing. The methods are "BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none". See <a href="#">p.adjust</a> .

**Value**

a matrix containing:

observed	observed gamma coefficients
expected	expected gamma coefficients
se	standard errors
pvalue	p values (under normal distribution assumption)
padj	adjusted p values if selected
sig	significance stars: 0 " *** " 0.001 " ** " 0.01 " * " 0.05 " . " 0.1 " " 1

**Author(s)**

Marianne Mueller

## References

Kreiner, S. (2011). A note on item-restscore association in Rasch models. *Applied Psychological Measurement*, 35, 557-561.

## Examples

```
rm.mod <- RM(amts[,4:13])
item_restscore(rm.mod)
```

---

item\_target

*Computation of Item Targets for Polytomous Models*

---

## Description

The item target is the value of the person parameter where item information is maximized.

## Usage

```
item_target(obj)
```

## Arguments

obj                    An object of class "eRm" (but not "dRm"), a fitted partial credit model using the function PCM in package eRm or of class "pcmodel" (from package psychotools).

## Value

vector with item targets.

## Author(s)

Marianne Mueller

## Examples

```
## Not run:
pc.mod <- PCM(desc2[, 5:14])
item_target(pc.mod)

## End(Not run)
```

---

out\_infit

*Item Outfit and Infit Statistics*


---

### Description

To avoid bias observed item responses are compared to expected responses under the conditional distribution of responses given the total score. This leads to standardized residuals which can be summarized to outfit and infit statistics in the usual way.

### Usage

```
out_infit(
  object,
  se = TRUE,
  p.adj = c("BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none")
)
```

### Arguments

object	An object of class "Rm", a fitted Rasch model or partial credit model using the functions RM or PCM in package eRm, or an object of class "pcmodel", a fitted partial credit model using the function pcmodel in package psychotools.
se	If TRUE the standard errors will be included.
p.adj	Correction method for multiple testing. The methods are "BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none". See <a href="#">p.adjust</a> .

### Details

The fit statistics and their standard errors are calculated as described in Christensen et al. P values are based on the normal distribution of the standardized fit statistics.

### Value

an object of class outfit containing:

outfit	outfit statistics
outfit.se	standard errors of outfit statistics
out.pvalue	p values of outfit statistics
out.pvalue.adj	adjusted p values of outfit statistics if selected
infit	infit statistics
infit.se	standard errors of infit statistics
in.pvalue	p values of infit statistics
in.pvalue.adj	adjusted p values of infit statistics if selected
p.adj	adjustment method

**Author(s)**

Marianne Mueller

**References**

Christensen, K. B. , Kreiner, S. & Mesbah, M. (Eds.) *Rasch Models in Health*. Iste and Wiley (2013), pp. 86 - 90.

Kreiner, S. & Christensen, K. B. (2011) Exact evaluation of Bias in Rasch model residuals. *Advances in Mathematics Research*, 12, 19-40.

**Examples**

```
rm.mod <- RM(amts[,4:13])
out_infit(rm.mod)
```

---

partgam

*Conditional and Partial Gamma Coefficients*

---

**Description**

Calculates conditional and partial Gamma coefficients for x and y given z with confidence intervals.

**Usage**

```
partgam(x, y, z, conf.level = 0.95)
```

**Arguments**

x, y, z            Three numeric vectors or factors.  
conf.level        Confidence level for the returned confidence interval.

**Value**

data frame with estimates, standard errors and confidence interval limits.

**Author(s)**

Marianne Mueller

**References**

Davis, J. A. A Partial coefficient for Goodman and Kruskal's Gamma. *Journal of the American Statistical Association*, 62 (317), 1967, pp. 189-193.

**See Also**

[partgam\\_DIF](#), [partgam\\_LD](#)

partgam\_DIF

*Partial Gamma to detect Differential Item Functioning (DIF)***Description**

Items should function in the same way for all subgroups of persons. An item shows differential item functioning (DIF) if there is a significant association between the item score and an exogenous variable, controlling for the scale score. Partial Gamma coefficients are used as test statistics.

**Usage**

```
partgam_DIF(
  dat.items,
  dat.exo,
  p.adj = c("BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none")
)
```

**Arguments**

<code>dat.items</code>	A data frame with the responses to the items.
<code>dat.exo</code>	A single grouping factor or a data frame consisting of several exogenous factor variables.
<code>p.adj</code>	Correction method for multiple testing. The methods are "BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none". See <a href="#">p.adjust</a> .

**Value**

data frame with Gamma coefficients, standard errors, p values, adjusted p values if an adjustment method has been chosen, and confidence limits for every pair of an item and an exogenous variable.

**Author(s)**

Marianne Mueller

**References**

Bjorner, J., Kreiner, S., Ware, J., Damsgaard, M. and Bech, P. Differential item functioning in the Danish translation of the SF-36. *Journal of Clinical Epidemiology*, 51 (11), 1998, pp. 1189-1202.

**See Also**

[partgam\\_LD](#)

**Examples**

```
partgam_DIF(ams[, 4:13], ams[, 2:3])
```

partgam\_LD

*Partial Gamma to detect Local Dependence (LD)***Description**

Rasch models assume locally independent items. There should be no substantial correlation left between two items once the underlying factor has been taken into account. Partial Gamma coefficients between pairs of items controlled for the rest score can be used to assess this requirement. The rest score is calculated as the score without the second item.

**Usage**

```
partgam_LD(
  dat.items,
  p.adj = c("BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none")
)
```

**Arguments**

`dat.items`      A data frame with the responses to the items.

`p.adj`            Correction method for multiple testing. The methods are "BH", "holm", "hochberg", "hommel", "bonferroni", "BY", "none". See [p.adjust](#).

**Details**

Because it matters which of the two items of a pair is subtracted from the total score to give the rest score, calculations are done for each pair in both ways. Results are stored in two different data frames.

**Value**

list of two data frames with Gamma coefficients, standard errors, p values, adjusted p values if an adjustment method has been chosen, and confidence limits for every pair of items.

**Author(s)**

Marianne Mueller

**References**

Christensen, K. B. , Kreiner, S. & Mesbah, M. (Eds.) *Rasch Models in Health*. Iste and Wiley (2013), pp. 133 - 135.

**See Also**

[partgam\\_DIF](#)

**Examples**

```
partgam_LD(amts[,4:13])
```

---

person_estimates	<i>Person Estimates with MLE and WLE</i>
------------------	--

---

**Description**

Computes Person estimates with maximum likelihood estimation (MLE) and weighted likelihood estimation (WLE) for raw scores 0 to m.

**Usage**

```
person_estimates(object, properties = F, allperson = F)
```

**Arguments**

object	An object of class "Rm", a fitted Rasch model or partial credit model using the functions RM or PCM in package eRm, or an object of class "raschmodel" or "pcmodel", a fitted Rasch model or partial credit model using the functions raschmodel or pcmodel in package psychotools.
properties	If TRUE additional properties of the estimates are given (see below).
allperson	If TRUE person estimates (MLE and WLE) for all persons in the data set are delivered.

**Value**

If properties = False a matrix containing:

Raw score	raw score
MLE	MLE of person parameters
WLE	WLE of person parameters

If properties = TRUE a list with two components, one for MLE and the other for WLE. Each component contains:

Raw score	raw score
MLE or WLE	person estimates
SEM	standard error of measurement
Bias	bias
RMSE	root mean square error
Score.SEM	score sem

**Author(s)**

Marianne Mueller





---

score_groups	<i>Generate two Score Groups</i>
--------------	----------------------------------

---

**Description**

Creates a grouping variable which divides the sample in two groups (high and low scorers) of roughly equal size, without taking into account persons with extreme scores.

**Usage**

```
score_groups(dat.items, label = FALSE)
```

**Arguments**

dat.items	A data frame with the responses to the items.
label	If TRUE the levels of the group factor are named according to the split used, if FALSE (default) the group factor has levels 1 and 2.

**Details**

The score groups are used for tests of item homogeneity.

**Value**

Score group variable, a factor with two levels.

---

test_prop	<i>Properties of the Test</i>
-----------	-------------------------------

---

**Description**

Information summarizing measurement quality of the test and test targeting.

**Usage**

```
test_prop(object)
```

**Arguments**

object	An object of class "Rm", a fitted Rasch model or partial credit model using the functions RM or PCM in package eRm, or an object of class "pcmodel", a fitted partial credit model using the function pcmodel in package psychotools.
--------	---

**Value**

a list containing:

Separation reliability

the person separation reliability as calculated in package eRm for objects of class "Rm".

Test difficulty

person value with an expected score equal to half of the maximum score.

Test target

person value where test information is maximized.

Test information

maximal value of the test information

**Author(s)**

Marianne Mueller

**References**

Christensen, K. B. , Kreiner, S. & Mesbah, M. (Eds.) *Rasch Models in Health*. Iste and Wiley (2013), pp. 63 - 70.

**Examples**

```
rm.mod <- RM(amtS[,4:13])
test_prop(rm.mod)
```

# Index

## \* datasets

amts, 3  
desc2, 5

amts, 3

boot\_fit, 4

clr\_tests, 4

desc2, 5

iarm (iarm-package), 2

iarm-package, 2

ICCplot, 6

item\_obsexp, 8

item\_restscore, 9

item\_target, 10

out\_infit, 11

p.adjust, 4, 9, 11, 13, 14

partgam, 12

partgam\_DIF, 12, 13, 14

partgam\_LD, 12, 13, 14

person\_estimates, 15

print.bootfit, 16

print.outfit, 16

score\_groups, 17

test\_prop, 17