

# Package ‘PerFit’

June 1, 2018

**Type** Package

**Title** Person Fit

**Version** 1.4.2

**Date** 2018-05-31

**Author** Jorge N. Tendeiro

**Maintainer** Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**Description** Several person-fit statistics (PFSs) are offered. These statistics allow assessing whether individual response patterns to tests or questionnaires are (im)plausible given the other respondents in the sample or given a specified item response theory model. Some PFSs apply to dichotomous data, such as the likelihood-based PFSs ( $Iz$ ,  $Iz^*$ ) and the group-based PFSs (personal biserial correlation, caution index, (normed) number of Guttman errors, agreement/disagreement/dependability statistics,  $U3$ ,  $ZU3$ ,  $NCI$ ,  $Ht$ ). PFSs suitable to polytomous data include extensions of  $Iz$ ,  $U3$ , and (normed) number of Guttman errors.

**Imports** stats, graphics, fda, Hmisc, irtoys, MASS, Matrix

**Depends** ltm, mirt

**License** GPL (>= 2)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2018-06-01 14:05:10 UTC

## R topics documented:

PerFit-package	2
A, D, E statistics	6
cutoff	9
flagged.resp	11
G, Gnormed	13
Gpoly, Gnormed.poly	16
Ht statistic	19
Inadequacy scale data	22

Intelligence data . . . . .	23
lz, lzstar . . . . .	23
lzpoly . . . . .	27
NCI statistic . . . . .	29
PerFit.PFS . . . . .	32
PerFit.SE . . . . .	33
Personal biserial statistic . . . . .	34
Physical Functioning . . . . .	37
plot PerFit . . . . .	37
PRFplot . . . . .	40
print PerFit . . . . .	43
summary PerFit . . . . .	44
The caution statistic . . . . .	45
U3, ZU3 . . . . .	49
U3poly . . . . .	52

<b>Index</b>	<b>55</b>
--------------	-----------

---

PerFit-package	<i>Person Fit</i>
----------------	-------------------

---

## Description

Person fit consists of a set of techniques aimed at detecting unusual responses to tests or questionnaires. There are several person-fit statistics available in the literature, see Karabatsos (2003) and Meijer and Sijtsma (2001) for comprehensive reviews. Both dichotomous and polytomous types of items are considered. This R-package outputs the values of the chosen person-fit statistic, the IDs of the respondents that were flagged, and plots the sample distribution of the scores of the person-fit statistic. Nonparametric person response functions (Sijtsma and Meijer, 2001) may also be requested in order to help interpreting individual answering behaviors (dichotomous data only).

## Details

Package:	PerFit
Type:	Package
Title:	Person Fit
Version:	1.4.2
Date:	2018-05-31
Author:	Jorge N. Tendeiro
Maintainer:	Jorge N. Tendeiro <j.n.tendeiro@rug.nl>
Description:	Several person-fit statistics (PFSs) are offered. These statistics allow assessing whether individual response pat
Imports:	stats, graphics, fda, Hmisc, irtoys, MASS, Matrix
Depends:	ltm, mirt
License:	GPL (>=2)

Index of help topics:

A.KB	Agreement, disagreement, and dependability statistics
Cstar	C.Sato, Cstar person-fit statistics
G	Number of Guttman errors
Gpoly	Number of Guttman errors for polytomous items
Ht	Ht person-fit statistic
InadequacyData	The NPV-J inadequacy scale data
IntelligenceData	Intelligence data (number completion)
NCI	NCI person-fit statistic
PRFplot	Person response function (PRF)
PerFit-package	Person Fit
PerFit.PFS	Compute several person-fit statistics
PerFit.SE	Compute standard errors for person fit statistics
PhysFuncData	The SF-36 physical functioning data
U3	U3, ZU3 person-fit statistics
U3poly	U3poly person-fit statistic
cutoff	Compute a cutoff value given the scores of a person-fit statistic
flagged.resp	Find (potentially) aberrant response patterns
lz	lz and lzstar person-fit statistics
lzpoly	lzpoly person-fit statistic
plot.PerFit	Plot method for objects of class "PerFit"
print.PerFit	Print method for objects of class "PerFit"
r.pbis	Personal biserial statistic
summary.PerFit	Summary method for objects of class "PerFit"

The **PerFit** package contains several person-fit functions. The goal is to detect response vectors that seem to be strange in terms of the sample of respondents or in terms to an item response theory (IRT) model.

There are many person-fit statistics available in the literature. Statistics are typically categorized according to the type of items (Dicho = dichotomous, Poly = polytomous) and the type of IRT model (NParam=nonparametric, Param=parametric) that they apply to. The current version of **PerFit** includes the following statistics:

Person-fit statistic (R function)	Reference	Type item	Type IRT model
<a href="#">r.pbis</a>	Donlon and Fisher (1968)	Dicho	NParam
<a href="#">C.Sato</a>	Sato (1975)	Dicho	NParam
<a href="#">G, Gnormed</a>	van der Flier (1977), Meijer (1994)	Dicho	NParam
<a href="#">A.KB, D.KB, E.KB</a>	Kane and Brennan (1980)	Dicho	NParam
<a href="#">U3, ZU3</a>	van der Flier (1980, 1982)	Dicho	NParam
<a href="#">Cstar</a>	Harnisch and Linn (1981)	Dicho	NParam
<a href="#">NCI</a>	Tatsuoka and Tatsuoka (1982, 1983)	Dicho	NParam
<a href="#">lz</a>	Drasgow, Levine, and Williams (1985)	Dicho	Param
<a href="#">lzpoly</a>	Drasgow, Levine, and Williams (1985)	Poly	Param
<a href="#">Ht</a>	Sijtsma (1986)	Dicho	NParam
<a href="#">Gpoly</a>	Molenaar (1991)	Poly	NParam
<a href="#">Gnormed.poly</a>	Molenaar (1991), Emons (2008)	Poly	NParam
<a href="#">lzstar</a>	Snijders (2001)	Dicho	Param

U3poly Emons (2008) Poly NParam

All functions above have an output of class `PerFit`.

The package provides other functions that help analyzing the data when conducting person-fit analyses:

Function	Description
<code>cutoff</code>	Estimate cutoff values for the person-fit statistics, to be used as decision rules.
<code>flagged.resp</code>	Identify which respondents were flagged according to the chosen cutoff.
<code>plot</code> (class <code>PerFit</code> )	Plot the distribution of person-fit scores with the cutoff superimposed.
<code>PRFplot</code>	Plot the nonparametric person response function (Sijtsma and Meijer, 2001).

More person-fit statistics will be added to the package in future updates.

### Versions

- Version 1.0 (April 2014)
- Version 1.1 (May 2014)
  - Functions `plot.PerFit` and `PRFplot` now allow the user to edit the axes labels and the titles.
- Version 1.2 (August 2014)
  - Some output values of some functions were renamed for the sake of consistency. The package documentation was adapted accordingly.
- Version 1.3 (March 2015)
  - The package underwent a major revision:
    - Class `PerFit` now consists of a list with 12 objects.
    - New methods for objects of class `PerFit` were added (`summary`, `print`).
    - Routines accomodating for missing values were added.
    - Function `cutoff` was updated. Now, model-fitting item response patterns are generated in order to find the cutoff value.
    - Function `plot.PerFit` now allows displaying a bootstrap percentile confidence interval for the cutoff statistic, as well as ticks marking the flagged respondents.
    - Person response functions are now approximated by functional data objects computed by means of the `fda` package. The functional data objects are returned to the user.
    - Standard errors for the person-fit statistics are now available (see function `PerFit.SE`).
    - Many control checks were added throughout the entire code.
- Version 1.4 (July 2015)
  - The default missing values approach is now pairwise elimination. The imputation methods introduced with version 1.3.1 are still available.
  - A bug was removed from function `lzpoly()` (many thanks to Marco Bressan for spotting it!).
  - The `PerFit-package.Rd` file was updated (version 1.4.2).

**Author(s)**

Jorge N. Tendeiro

Maintainer: Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**References**

- Donlon, T. F., and Fischer, F. E. (1968) An index of an individual's agreement with group-defined item difficulties. *Educational and Psychological Measurement*, **28(1)**, 105–113.
- Drasgow, F., Levine, M. V., and Williams, E. A. (1985) Appropriateness measurement with poly-chotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, **38(1)**, 67–86.
- Emons, W. M. (2008) Nonparametric person-fit analysis of polytomous item scores. *Applied Psychological Measurement*, **32(3)**, 224–247.
- Harnisch, D. L., and Linn, R. L. (1981) Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. *Journal of Educational Measurement*, **18(3)**, 133–146.
- Kane, M. T., and Brennan, R. L. (1980) Agreement coefficients as indices of dependability for domain-referenced tests. *Applied Psychological Measurement*, **4(1)**, 105–126.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Meijer, R. R. (1994) The number of Guttman errors as a simple and powerful person-fit statistic. *Applied Psychological Measurement*, **18(4)**, 311–314.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Molenaar, I. W. (1991) A weighted Loewinger H-coefficient extending Mokken scaling to multicategory items. *Kwantitatieve Methoden*, **12(37)**, 97–117.
- Sato, T. (1975) *The construction and interpretation of S-P tables*. Tokyo: Meiji Tosho.
- Sijtsma, K. (1986) A coefficient of deviance of response patterns. *Kwantitatieve Methoden*, **7**, 131–145.
- Sijtsma, K., and Meijer, R. R. (2001) The person response function as a tool in person-fit research. *Psychometrika*, **66(2)**, 191–207.
- Snijders, T. B. (2001) Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika*, **66(3)**, 331–342.
- Tatsuoka, K. K., and Tatsuoka, M. M. (1982) Detection of aberrant response patterns and their effect on dimensionality. *Journal of Educational Statistics*, **7(3)**, 215–231.
- Tatsuoka, K. K., and Tatsuoka, M. M. (1983) Spotting erroneous rules of operation by the individual consistency index. *Journal of Educational Measurement*, **20(3)**, 221–230.
- Tendeiro, J. N., Meijer, R. R., and Niessen, A. S. M. (2016). PerFit: An R Package for Person-Fit Analysis in IRT. *Journal of Statistical Software*, **74(5)**, 1–27.
- van der Flier, H. (1977) Environmental factors and deviant response patterns. In Y. H. Poortinga (Ed.), *Basic problems in cross-cultural psychology*. Amsterdam: The Netherlands.
- van der Flier, H. (1980) *Vergelijkbaarheid van individuele testprestaties* [Comparability of individual test performance]. Lisse: The Netherlands.

van der Flier, H. (1982) Deviant response patterns and comparability of test scores. *Journal of Cross-Cultural Psychology*, **13**(3), 267–298.

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# As an example, compute the Ht person-fit scores:
Ht.out <- Ht(InadequacyData)
# Ht.out$PFscores

# Compute the cutoff value at 1% level:
set.seed(124) # To fix the random seed generator.
Ht.cut <- cutoff(Ht.out, Blvl=.01)

# Plot the sample distribution of the Ht scores with the above cutoff superimposed:
plot(Ht.out, cutoff.obj=Ht.cut)

# Determine which respondents were flagged by Ht at 1% level:
flagged.resp(Ht.out, cutoff.obj=Ht.cut, scores=FALSE)
# Flagged respondents: 30, 37, 46, 49,...

# Plot the person response function of respondent 30 (flagged as aberrant):
Resp30 <- PRFplot(InadequacyData, respID=30)
# Plot the person response function of respondent 35 (not flagged as aberrant):
Resp35 <- PRFplot(InadequacyData, respID=35)
```

---

A, D, E statistics      *Agreement, disagreement, and dependability statistics*

---

### Description

Kane and Brennan's person-fit statistics.

### Usage

```
A.KB(matrix,
      NA.method = "Pairwise", Save.MatImp = FALSE,
      IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
      mu = 0, sigma = 1)

D.KB(matrix,
      NA.method = "Pairwise", Save.MatImp = FALSE,
      IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
      mu = 0, sigma = 1)

E.KB(matrix,
```

```

NA.method = "Pairwise", Save.MatImp = FALSE,
IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
mu = 0, sigma = 1)

```

### Arguments

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
<code>IRT.PModel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PModel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when method="BM". Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

### Details

Kane and Brennan (1980) discussed the agreement, disagreement, and dependability statistics. Assume that the items are ordered in increasing difficulty order (i.e., according to decreasing proportion-correct score). The agreement statistic for respondent  $n$  is

$$A.KB = \sum_i X_i p_i,$$

where  $X_i$  is the 0-1 score of respondent  $n$  on item  $i$  and  $p_i$  is the proportion-correct score of item  $i$ .

The disagreement statistic is

$$D.KB = A.KB(max) - A.KB,$$

where  $A.KB(max)$  is the maximum value of A.KB given respondent  $n$ 's total score.

The dependability statistic is

$$E.KB = A.KB/A.KB(max).$$

Small values of A.Kb and E.KB (i.e., in the left tail of the sampling distribution) are (potentially) indicative of aberrant response behavior. Large values of D.Kb (i.e., in the right tail of the sampling distribution) are (potentially) indicative of aberrant response behavior. These statistics are not computed for rows of `matrix` that consist of only 0s or only 1s (NA values are returned instead).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL", "2PL", or "3PL"`). Item parameters (IP) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

## Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
<code>ID.all0s</code>	Row indices of all-0s response vectors removed from the analysis (if applicable).
<code>ID.all1s</code>	Row indices of all-1s response vectors removed from the analysis (if applicable).
<code>matrix</code>	The data matrix after imputation of missing values was performed (if applicable).
<code>Ncat</code>	The number of response categories (2 in this case).
<code>IRT.PModel</code>	The parametric IRT model used in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>IP</code>	The $I \times 3$ matrix of estimated item parameters in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>Ability.PModel</code>	The method used to estimate abilities in case <code>NA.method="PModel"</code> , otherwise NULL.



Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Kane, M. T., and Brennan, R. L. (1980) Agreement coefficients as indices of dependability for domain-referenced tests. *Applied Psychological Measurement*, **4**(1), 105–126.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16**(4), 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25**(2), 107–135.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the A.KB, D.KB, and E.KB scores:
A.out <- A.KB(InadequacyData); A.out
D.out <- D.KB(InadequacyData); D.out
E.out <- E.KB(InadequacyData); E.out
```

---

cutoff *Compute a cutoff value given the scores of a person-fit statistic*

---

### Description

Compute a cutoff value given the scores of a person-fit statistic.

### Usage

```
cutoff(x,
       ModelFit = "NonParametric", Nreps=1000,
       IP=x$IP, IRT.PModel=x$IRT.PModel, Ability=x$Ability,
       Ability.PModel=x$Ability.PModel, mu=0, sigma=1,
       Blvl = 0.05, Breps = 1000, CIlvl = 0.95,
       UDlvl = NA)
```

**Arguments**

x	Object of class "PerFit".
ModelFit	Method required to compute model-fitting item score patterns. The options available are "NonParametric" (default) and "Parametric".
Nreps	Number of model-fitting item score patterns generated. Default is 1000.
IP	Matrix with previously estimated item parameters. Default is x\$IP.
IRT.PModel	Parametric IRT model (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is x\$IRT.PModel.
Ability	Matrix with previously estimated item parameters. Default is x\$Ability.
Ability.PModel	Method to use in order to estimate the latent ability parameters (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is x\$Ability.PModel.
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.
Blvl	Significance level for bootstrap distribution (value between 0 and 1). Default is 0.05.
Breps	Number of bootstrap resamples. Default is 1000.
CI1vl	Level of bootstrap percentile confidence interval for the cutoff statistic.
UD1vl	User-defined cutoff level.

**Details**

This function computes a reference value (referred to as a 'cutoff') associated to the values of a person-fit statistic computed from a sample. The idea is to create a decision rule: Individual person-fit values at or more extreme than the cutoff result in flagging the corresponding respondents as (potentially) displaying aberrant response behavior.

Depending on the person-fit statistic, an "extreme" score might be a very small (e.g., for Ht) or a very large (e.g., for G) value. The cutoff function routinely reports of which type the person-fit statistic being used is (Tail="lower" or Tail="upper", respectively).

The procedure consists of generating Nreps model-fitting item response vectors based on parametric model parameters (when ModelFit="Parametric") or on proportion of respondents per answer category (when ModelFit="NonParametric"). This allows computing a sample of Nreps values of the person fit statistic corresponding to model-fitting item response patterns. A bootstrap procedure is then used to approximate the sampling distribution of the quantile of level Blvl (resp., 1-Blvl) for "lower" (resp. "upper") types of person fit statistics, based on Breps resamples. The cutoff (and its standard error) is given by the median (standard deviation) of this bootstrap distribution. Alternatively, the cutoff can be manually entered by the user (e.g., when it is available from prior data calibration) by means of UD1vl.

**Value**

An object of class "PerFit.cutoff", which is a list with 5 elements:

Cutoff	Numeric. Value of the computed cutoff.
--------	--

Cutoff.SE	Numeric. Bootstrap estimated standard error of the cutoff value.
Prop.flagged	Numeric. Proportion of respondents flagged (that is, with person-fit scores at or more extreme than the cutoff).
Tail	String with values "lower" or "upper". It indicates the type of person-fit statistic.
Cutoff.CI	Numeric. Percentile bootstrap (CI <sub>lv</sub> )% confidence interval for the cutoff value.

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**See Also**

[flagged.resp](#), [plot.PerFit](#), [PRFplot](#)

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# As an example, compute the Ht person-fit scores:
Ht.out <- Ht(InadequacyData)

# Compute the cutoff value at 1% level:
cutoff(Ht.out, Blvl=.01)
```

---

flagged.resp

*Find (potentially) aberrant response patterns*

---

**Description**

Find which respondents in the sample were flagged by the specified person-fit statistic.

**Usage**

```
flagged.resp(x, cutoff.obj=NULL,
             scores=TRUE, ord=TRUE,
             ModelFit="NonParametric", Nreps=1000,
             IP=x$IP, IRT.PModel=x$IRT.PModel, Ability=x$Ability,
             Ability.PModel=x$Ability.PModel, mu=0, sigma=1,
             Blvl = 0.05, Breps = 1000, CIlv = 0.95,
             UDlv=NA)
```

**Arguments**

<code>x</code>	Object of class "PerFit".
<code>cutoff.obj</code>	Object of class "PerFit.cutoff".
<code>scores</code>	Logical: Should item scores of flagged respondents be shown in the output? Default is TRUE.
<code>ord</code>	Logical: Should items be ordered in increasing order of difficulty (i.e., in decreasing proportion-correct order)? Default is TRUE. Only used if <code>scores=TRUE</code> .
<code>ModelFit</code>	Method required to compute model-fitting item score patterns. The options available are "NonParametric" (default) and "Parametric".
<code>Nreps</code>	Number of model-fitting item score patterns generated. Default is 1000.
<code>IP</code>	Matrix with previously estimated item parameters. Default is <code>x\$IP</code> .
<code>IRT.PModel</code>	Parametric IRT model (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is <code>x\$IRT.PModel</code> .
<code>Ability</code>	Matrix with previously estimated item parameters. Default is <code>x\$Ability</code> .
<code>Ability.PModel</code>	Method to use in order to estimate the latent ability parameters (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is <code>x\$Ability.PModel</code> .
<code>mu</code>	Mean of the apriori distribution. Only used when <code>method="BM"</code> . Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when <code>method="BM"</code> . Default is 1.
<code>Blvl</code>	Significance level for bootstrap distribution (value between 0 and 1). Default is 0.05.
<code>Breps</code>	Number of bootstrap resamples. Default is 1000.
<code>CIlvl</code>	Level of bootstrap percentile confidence interval for the cutoff statistic.
<code>UDlvl</code>	User-defined cutoff level.

**Details**

This function finds the respondents in the dataset that were flagged by the person-fit statistic. This statistic is specified by means of the "PerFit" class object `x` (`x$PFStatistic`).

The cutoff score may be provided by means of the `cutoff.obj` object, otherwise it is internally computed (for which the function parameters `ModelFit` through `CIlvl` are required; see [cutoff](#) for more details).

If `scores=TRUE` then the respondents' item scores will be shown in the output, either in the original item order (`ord=FALSE`) or in increasing difficulty order (`ord=TRUE`).

**Value**

If `scores=FALSE` the output is a list with 3 elements:

<code>PFscores</code>	A two-column matrix with the row index and the value of the person-fit statistic for the flagged respondents.
<code>Cutoff.lst</code>	The corresponding <code>PerFit.cutoff</code> object.

PFS                    The person-fit statistic.

If scores=TRUE the output is a list with four elements:

Scores                Matrix with columns: FlaggedID, item scores (It\*\*), and PFscores.

MeanItemValue      The items mean value (which is nothing more than the proportion-correct for dichotomous items).

Cutoff.lst          The corresponding PerFit.cutoff object.

PFS                    The person-fit statistic.

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### See Also

[cutoff](#), [plot.PerFit](#), [PRFplot](#)

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# As an example, compute the Ht person-fit scores:
Ht.out <- Ht(InadequacyData)
Ht.out$PFscores

# Estimate the cutoff value at 1% level:
Ht.cut <- cutoff(Ht.out, Blvl=.01)

# Determine which respondents were flagged by Ht at 1% level:
flagged.resp(Ht.out, Ht.cut, scores=FALSE)$PFscores
```

---

G, Gnormed

*Number of Guttman errors*

---

### Description

van der Flier's statistics based on the number of Guttman errors.

### Usage

```
G(matrix,
  NA.method = "Pairwise", Save.MatImp = FALSE,
  IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
  mu = 0, sigma = 1)
```

```
Gnormed(matrix,
```

```
NA.method = "Pairwise", Save.MatImp = FALSE,
IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
mu = 0, sigma = 1)
```

### Arguments

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
<code>IRT.PModel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PModel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when method="BM". Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

### Details

Consider the items' proportion-correct scores,  $p_1, p_2, \dots, p_I$  ( $I$  = number of items). A Guttman error consists of an item score pair  $(X_i, X_j) = (0, 1)$  with  $p_i > p_j$ . Hence, there is a Guttman error when an easier item is answered incorrectly and a more difficult item is answered correctly.

G counts the number of (0,1) pairs given that the items are ordered in decreasing proportion-correct scores order. However, G depends on the total number of items for a given number-correct score. In particular, for a number-correct  $S$ , G has maximum equal to  $S(I - S)$ . Gnormed was created to bound G between 0 and 1 by dividing it by its maximum (conditional on the number-correct score). Hence, (potentially) aberrant response behavior is indicated by large values of  $G/Gnormed$  (i.e., in the right tail of the sampling distribution).

Gnormed is perfectly linearly related to Tatsuoka and Tatsuoka's (1982, 1983) NCI statistic ( $NCI = 1 - 2Gnormed$ ).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (IRT.PModel = "1PL", "2PL", or "3PL"). Item parameters (IP) and ability parameters (Ability) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

PFscores	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
PFstatistic	The person-fit statistic used.
PerfVects	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
ID.all0s	Row indices of all-0s response vectors removed from the analysis (if applicable).
ID.all1s	Row indices of all-1s response vectors removed from the analysis (if applicable).
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories (2 in this case).
IRT.PModel	The parametric IRT model used in case NA.method="PModel", otherwise NULL.
IP	The $I \times 3$ matrix of estimated item parameters in case NA.method="PModel", otherwise NULL.
Ability.PModel	The method used to estimate abilities in case NA.method="PModel", otherwise NULL.
Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

## References

- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Meijer, R. R. (1994) The number of Guttman errors as a simple and powerful person-fit statistic. *Applied Psychological Measurement*, **18(4)**, 311–314.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Tatsuoka, K. K., and Tatsuoka, M. M. (1982) Detection of aberrant response patterns and their effect on dimensionality. *Journal of Educational Statistics*, **7(3)**, 215–231.
- Tatsuoka, K. K., and Tatsuoka, M. M. (1983) Spotting erroneous rules of operation by the individual consistency index. *Journal of Educational Measurement*, **20(3)**, 221–230.
- van der Flier, H. (1977) Environmental factors and deviant response patterns. In Y. H. Poortinga (Ed.), *Basic problems in cross-cultural psychology*. Amsterdam: The Netherlands.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32(6)**, 466–479.

## See Also

[NCI](#)

## Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the G scores:
G.out <- G(InadequacyData)

# Compute the Gnormed scores:
Gnormed.out <- Gnormed(InadequacyData)
```

---

Gpoly, Gnormed.poly    *Number of Guttman errors for polytomous items*

---

## Description

Molenaar (1991) and Emons (2008) statistics, based on the number of Guttman errors, for polytomous items.

## Usage

```
Gpoly(matrix, Ncat,
       NA.method = "Pairwise", Save.MatImp = FALSE,
       IP = NULL, IRT.PModel = "GRM", Ability = NULL, Ability.PModel = "EAP")

Gnormed.poly(matrix, Ncat,
```



```
NA.method = "Pairwise", Save.MatImp = FALSE,
IP = NULL, IRT.PModel = "GRM", Ability = NULL, Ability.PModel = "EAP")
```

### Arguments

matrix	A data matrix of polytomous item scores: Persons as rows, items as columns, item scores are integers between 0 and (Ncat-1), missing values allowed.
Ncat	Number of answer options for each item.
NA.method	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
Save.MatImp	Logical. Save (imputed) data matrix to file? Default is FALSE.
IP	Matrix with previously estimated item parameters: One row per item. The first (Ncat-1) columns contain the between-categories threshold parameters (for the GRM) or the item step difficulties (for the PCM and the GPCM). The last, Ncat-th, column has the slopes. In case no item parameters are available then IP=NULL.
IRT.PModel	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "PCM", "GPCM", and "GRM" (default).
Ability	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of matrix. In case no ability parameters are available then Ability=NULL.
Ability.PModel	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "EB", "EAP" (default), and "MI".

### Details

Molenaar (1991) generalized the G person-fit statistic to polytomous items, Gpoly. The idea is based on the so-called item-step difficulty, which is the probability of moving from answer category ( $c$ ) to answer category ( $c + 1$ ) ( $c = 0, \dots, Ncat - 2$ ).

Just like G, Gpoly depends on the test length. Emons (2008) developed Gnormed.poly, which is a normalized version of Gpoly.

Aberrant response behavior is (potentially) indicated by large values of Gpoly/Gnormed.poly (i.e., in the right tail of the sampling distribution).

The number of answer options, Ncat, is the same for all items.

Gpoly reduces to G, and Gnormed.poly reduces to Gnormed, when Ncat=2.

Missing values in matrix are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (NA.method = "Hotdeck"), nonparametric model imputation (NA.method = "NPMModel"), and parametric model imputation (NA.method = "PModel"); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The

donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.

- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities estimated by means of parametric IRT models (IRT.PModel = "PCM", "GPCM", or "GRM"). Item parameters (IP) and ability parameters (Ability) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

PFscores	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
PFstatistic	The person-fit statistic used.
PerfVects	Not applicable.
ID.all0s	Not applicable.
ID.all1s	Not applicable.
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories.
IRT.PModel	The parametric IRT model used.
IP	The $I \times N_{cat}$ matrix of estimated item parameters.
Ability.PModel	The method used to estimate abilities used.
Ability	The vector of $N$ estimated ability parameters.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Emons, W. M. (2008) Nonparametric person-fit analysis of polytomous item scores. *Applied Psychological Measurement*, **32**(3), 224–247.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16**(4), 277–298.
- Meijer, R. R. (1994) The number of Guttman errors as a simple and powerful person-fit statistic. *Applied Psychological Measurement*, **18**(4), 311–314.

Molenaar, I. W. (1991) A weighted Loevinger H-coefficient extending Mokken scaling to multcategory items. *Kwantitatieve Methoden*, **12**(37), 97–117.

Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

### See Also

[G](#), [Gnormed](#)

### Examples

```
# Load the physical functioning data (polytomous item scores):
data(PhysFuncData)

# Compute the Gpoly scores:
Gpoly.out <- Gpoly(PhysFuncData, Ncat=3)

# Compute the Gnormed.poly scores:
Gnormedpoly.out <- Gnormed.poly(PhysFuncData, Ncat=3)
```

---

Ht statistic

*Ht person-fit statistic*

---

### Description

Sijtsma's Ht person-fit statistic.

### Usage

```
Ht(matrix,
    NA.method = "Pairwise", Save.MatImp = FALSE,
    IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
    mu = 0, sigma = 1)
```

### Arguments

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.

IRT.PModel	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
Ability	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of matrix. In case no ability parameters are available then Ability=NULL.
Ability.PModel	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

### Details

Sijtsma (1986) adapted a statistic introduced by Mokken (1971) that originally allowed assessing the scalability of an item to the Guttman (1944, 1950) model. The same statistic was applied by Sijtsma to the *transposed* data in order to detect respondents that would not comply with the Guttman model. Assume, without loss of generality, that the rows of the data matrix are ordered by increasing order of total score  $S_n$  ( $n = 1, \dots, N$ ). The statistic formula is given by the ratio

$$Ht = \frac{cov(x_n, r_{(n)})}{cov_{max}(x_n, r_{(n)})},$$

where  $r_{(n)}$  is the vector of total item scores computed *excluding* respondent  $n$  and the denominator is the maximum covariance given the marginal. Hence, Ht is actually similar to Sato's [C.Sato](#).

Ht is maximum 1 for respondent  $n$  when no respondent with a total score smaller/larger than  $S_n$  can answer an item correctly/incorrectly that respondent  $n$  has answered incorrectly/correctly, respectively. Ht equals zero when the average covariance of the response pattern of respondent  $n$  with the other response patterns equals zero. Hence, (potentially) aberrant response behavior is indicated by small values of Ht (i.e., in the left tail of the sampling distribution). The Ht statistic was shown to perform relatively well in several simulation studies (Karabatsos, 2003; Sijtsma, 1986; Sijtsma and Meijer, 1992, Tendeiro and Meijer, 2014).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).

- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (IRT.PModel = "1PL", "2PL", or "3PL"). Item parameters (IP) and ability parameters (Ability) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

PFscores	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
PFstatistic	The person-fit statistic used.
PerfVects	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
ID.all0s	Row indices of all-0s response vectors removed from the analysis (if applicable).
ID.all1s	Row indices of all-1s response vectors removed from the analysis (if applicable).
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories (2 in this case).
IRT.PModel	The parametric IRT model used in case NA.method="PModel", otherwise NULL.
IP	The $I \times 3$ matrix of estimated item parameters in case NA.method="PModel", otherwise NULL.
Ability.PModel	The method used to estimate abilities in case NA.method="PModel", otherwise NULL.
Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Guttman, L. (1944) A basis for scaling qualitative data. *American Sociological Review*, **9**, 139-150.
- Guttman, L. (1950) The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star & J. A. Claussen (Eds.), *Measurement and precision* (pp. 60-90). Princeton NJ: Princeton University Press.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277-298.
- Mokken, R. J. (1971) *A theory and procedure of scale analysis*. Berlin, Germany: De Gruyter.
- Sijtsma, K. (1986) A coefficient of deviance of response patterns. *Kwantitatieve Methoden*, **7**, 131-145.

Sijtsma, K., and Meijer, R. R. (1992) A method for investigating the intersection of item response functions in Mokken's nonparametric IRT model. *Applied Psychological Measurement*, **16**(2), 149-157.

Tendeiro, J. N., and Meijer, R. R. (2014) Detection of Invalid Test Scores: The Usefulness of Simple Nonparametric Statistics. *Journal of Educational Measurement*, **51**(3), 239-259.

Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466-479.

### See Also

[C. Sato](#)

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the Ht scores:
Ht.out <- Ht(InadequacyData)
```

---

Inadequacy scale data *The NPV-J inadequacy scale data*

---

### Description

The NPV-J (Dutch: Junior Nederlandse Persoonlijkheidsvragenlijst, NPV-J; Luteijn, van Dijk, and Barelds, 2005) is a large Dutch personality inventory. The NPV-J consists of 105 mostly positively formulated items and is intended to determine how adolescents between 9 and 15 years of age judge their own behavior. The NPV-J has five subscales; the `InadequacyScale` data concern scores of 806 adolescents on 28 items measuring inadequacy (one of the subscales). The data are dichotomous (0 = Disagree, 1 = Agree). The original sample consisted of 866 respondents, however 60 all-0s or all-1s response vectors were removed from the data.

### Usage

```
data(InadequacyData)
```

### Format

A 806x28 matrix of dichotomous item scores.

### References

Luteijn, F., van Dijk, H., and Barelds, D. P. H. (2005) *NPV-J: Junior Nederlandse Persoonlijkheidsvragenlijst*. Herziene handleiding 2005. Amsterdam: Harcourt Assessments B.V..

Meijer, R. R., and Tendeiro, J. N. (2012) The use of the  $I_z$  and  $I_z^*$  person-fit statistics and problems derived from model misspecification. *Journal of Educational and Behavioral Statistics*, **37**(6), 758-766.

**Examples**

```
data(InadequacyData)
```

---

Intelligence data	<i>Intelligence data (number completion)</i>
-------------------	--

---

**Description**

The data are dichotomous scores of a Dutch intelligence test on number completion (Dutch: "Cijferreeksen", Drenth and Hoolwerf, 1970). The file consists of archival data that were collected in a high-stakes personnel selection context around 1990.

**Usage**

```
data(IntelligenceData)
```

**Format**

A 1000x26 matrix of dichotomous item scores.

**References**

Drenth, P. J. D., and Hoolwerf, G. (1970) *Numerieke aanleg test - Cijferreeksen (NAT-Cijferreeksen)* [Numerical ability test]. Amsterdam: The Netherlands.

**Examples**

```
data(IntelligenceData)
```

---

lz, lzstar	<i>lz and lzstar person-fit statistics</i>
------------	--

---

**Description**

Compute the lz (Drasgow, Levine, and Williams, 1985) and the lzstar (Snijders, 2001) person-fit statistics.

**Usage**

```
lz(matrix,
  NA.method = "Pairwise", Save.MatImp = FALSE,
  IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
  mu = 0, sigma = 1)
```

```
lzstar(matrix,
  NA.method = "Pairwise", Save.MatImp = FALSE,
  IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
  mu = 0, sigma = 1)
```

## Arguments

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMoDel" (default), and "PMoDel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then <code>IP=NULL</code> .
<code>IRT.PMoDel</code>	Specify the IRT model to use in order to estimate the item parameters (only if <code>IP=NULL</code> ). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then <code>Ability=NULL</code> .
<code>Ability.PMoDel</code>	Specify the method to use in order to estimate the latent ability parameters (only if <code>Ability=NULL</code> ). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when <code>method="BM"</code> . Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when <code>method="BM"</code> . Default is 1.

## Details

Drasgow et al. (1985) introduced one of the most used person-fit statistics, `Iz`. This statistic is the standardized log-likelihood of the respondent's response vector. `Iz` is (supposed to be) asymptotically standard normally distributed.

The computation of `Iz` requires that both item and ability parameters are available. Function `Iz` allows to user to enter his/her own item and ability parameter estimates (variables `IP` and `Ability`, respectively). Alternatively, `Iz` relies on functions available through the `irtoys` package for estimating the parameters. Specifically, the user can choose one from three possible IRT models to fit the data: `IRT.PMoDel="1PL"`, `IRT.PMoDel="2PL"`, or `IRT.PMoDel="3PL"`. As for estimating the ability parameters there are three possible methods: `Ability.PMoDel="ML"` (maximum likelihood), `Ability.PMoDel="BM"` (Bayes modal), or `Ability.PMoDel="WL"` (weighted likelihood).

It was later observed by several researchers (e.g., Molenaar and Hoijtink, 1990) that the asymptotic approximation only holds when true ability values are used. This limitation was overcome by Snijders (2001), who further developed `Iz` into the `Izstar` statistic. An accessible paper that thoroughly explains the basic principles behind `Izstar` is Magis, Raiche, and Beland (2012). It is important to realize that not all item and/or ability estimation procedures can be used when computing `Izstar`. In particular, the estimation of the ability parameters is constrained (see Snijders, 2001, Equation 5). The `Izstar` algorithm internally estimates the ability parameters accordingly for one of three possible methods: `Ability.PMoDel="ML"` (maximum likelihood), `Ability.PMoDel="BM"` (Bayes modal), or `Ability.PMoDel="WL"` (weighted likelihood), see Magis et al. (2012). The user may provide his or her own ability estimates in case they are available by means of other software.



In this case it is necessary to specify the method that was used for the estimation (ML, BM, or WL) using the argument `Ability.PModel`.

Aberrant response behavior is (potentially) indicated by small values of `Iz/Izstar` (i.e., in the left tail of the sampling distribution).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL", "2PL", or "3PL"`). Item parameters (IP) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

## Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
<code>ID.all0s</code>	Row indices of all-0s response vectors removed from the analysis (if applicable).
<code>ID.all1s</code>	Row indices of all-1s response vectors removed from the analysis (if applicable).
<code>matrix</code>	The data matrix after imputation of missing values was performed (if applicable).
<code>Ncat</code>	The number of response categories (2 in this case).
<code>IRT.PModel</code>	The parametric IRT model used.
<code>IP</code>	The $I \times 3$ matrix of estimated item parameters.
<code>Ability.PModel</code>	The method used to estimate abilities used.
<code>Ability</code>	The vector of $N$ estimated ability parameters.
<code>NAs.method</code>	The imputation method used (if applicable).

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**References**

- Drasgow, F., Levine, M. V., and Williams, E. A. (1985) Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, **38**(1), 67–86.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16**(4), 277–298.
- Magis, D., Raiche, G., and Beland, S. (2012) A didactic presentation of Snijders’s  $I_z$  index of person fit with emphasis on response model selection and ability estimation. *Journal of Educational and Behavioral Statistics*, **37**(1), 57–81.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25**(2), 107–135.
- Molenaar, I. W., and Hoijtink, H. (1990) The many null distributions of person fit indices. *Psychometrika*, **55**(1), 75–106.
- Snijders, T. B. (2001) Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika*, **66**(3), 331–342.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

**See Also**

[lzpoly](#)

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the lz scores using a subsample of the first 200 response vectors:
lz.out <- lz(InadequacyData[1:200,])
# Use parameters estimated externally (in this case item parameters estimated by mirt):
mod <- mirt(InadequacyData[1:200,], 1)
ip.mirt <- coef(mod, IRTpars = TRUE, simplify = TRUE, digits = Inf)$items[,c('a', 'b', 'g')]
lz.out2 <- lz(InadequacyData[1:200,], IP = ip.mirt)

# Compute the lzstar scores using a subsample of the first 200 response vectors:
lzstar.out <- lzstar(InadequacyData[1:200,])
```

---

Izpoly	<i>Izpoly person-fit statistic</i>
--------	------------------------------------

---

### Description

Compute the Izpoly (Drasgow, Levine, and Williams, 1985) person-fit statistic.

### Usage

```
Izpoly(matrix, Ncat,
       NA.method = "Pairwise", Save.MatImp = FALSE,
       IP = NULL, IRT.PModel = "GRM", Ability = NULL, Ability.PModel = "EAP")
```

### Arguments

<code>matrix</code>	A data matrix of polytomous item scores: Persons as rows, items as columns, item scores are integers between 0 and (Ncat-1), missing values allowed.
<code>Ncat</code>	Number of answer options for each item.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item. The first (Ncat-1) columns contain the between-categories threshold parameters (for the GRM) or the item step difficulties (for the PCM and the GPCM). The last, Ncat-th, column has the slopes. In case no item parameters are available then IP=NULL.
<code>IRT.PModel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "PCM", "GPCM", and "GRM" (default).
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PModel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "EB", "EAP" (default), and "MI".

### Details

Statistic Izpoly is the natural extension of `lz` to polytomously scores items. In this case the user can choose one from three possible IRT models to fit the data: The partial credit model (IRT.PModel="PCM"), the generalized partial credit model (IRT.PModel="GPCM"), or the graded response model (IRT.PModel="GRM"). Ability parameters can be estimated by means of one of three methods: Empirical Bayes (Ability.PModel="EB"), expected a posteriori (Ability.PModel="EAP"), or multiple imputation (Ability.PModel="MI").

The estimation of the model parameters is based on the `ltm` package. This function will estimate the item and ability parameters when both sets of parameters are missing. It will also estimate

one set of parameters in case only the other set is provided. It is possible that some estimation convergence problems occur that may break the function. In this case it is advisable to estimate the model parameters externally and then to run this function with those estimates provided via the commands `IP` and `Ability`.

Aberrant response behavior is (potentially) indicated by small values of `Izpoly` (i.e., in the left tail of the sampling distribution).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "PCM", "GPCM", or "GRM"`). Item parameters (`IP`) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

## Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	Not applicable.
<code>ID.all0s</code>	Not applicable.
<code>ID.all1s</code>	Not applicable.
<code>matrix</code>	The data matrix after imputation of missing values was performed (if applicable).
<code>Ncat</code>	The number of response categories.
<code>IRT.PModel</code>	The parametric IRT model used.
<code>IP</code>	The $I \times N_{cat}$ matrix of estimated item parameters.
<code>Ability.PModel</code>	The method used to estimate abilities used.
<code>Ability</code>	The vector of $N$ estimated ability parameters.
<code>NAs.method</code>	The imputation method used (if applicable).

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**References**

- Drasgow, F., Levine, M. V., and Williams, E. A. (1985) Appropriateness measurement with polytomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, **38(1)**, 67–86.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Magis, D., Raiche, G., and Beland, S. (2012) A didactic presentation of Snijders's  $I_z$  index of person fit with emphasis on response model selection and ability estimation. *Journal of Educational and Behavioral Statistics*, **37(1)**, 57–81.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Molenaar, I. W., and Hoijtink, H. (1990) The many null distributions of person fit indices. *Psychometrika*, **55(1)**, 75–106.
- Snijders, T. B. (2001) Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika*, **66(3)**, 331–342.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32(6)**, 466–479.

**See Also**

[lz, lzstar](#)

**Examples**

```
# Load the physical functioning data (polytomous item scores):
data(PhysFuncData)

# Compute the lzpoly scores:
lzpoly.out <- lzpoly(PhysFuncData, Ncat=3)
```

---

NCI statistic

*NCI person-fit statistic*

---

**Description**

Tatsuoka and Tatsuoka's NCI statistic.

**Usage**

```
NCI(matrix,
      NA.method = "Pairwise", Save.MatImp = FALSE,
      IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
      mu = 0, sigma = 1)
```

**Arguments**

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMoDel" (default), and "PMoDel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then <code>IP=NULL</code> .
<code>IRT.PMoDel</code>	Specify the IRT model to use in order to estimate the item parameters (only if <code>IP=NULL</code> ). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then <code>Ability=NULL</code> .
<code>Ability.PMoDel</code>	Specify the method to use in order to estimate the latent ability parameters (only if <code>Ability=NULL</code> ). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when <code>method="BM"</code> . Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when <code>method="BM"</code> . Default is 1.

**Details**

The NCI person-fit statistic was introduced by Tatsuoka and Tatsuoka (1982, 1983). It is perfectly linearly related to van der Flier's (1977) Gnormed statistic ( $NCI = 1 - 2G_{normed}$ ), see [G](#) for mathematical details.

NCI equals 1 for perfect Guttman vectors (i.e., when only the easiest items are answered correctly, given the total score) and equals -1 for reversed Guttman vectors (i.e., when only the hardest items are answered correctly, given the total score). Hence, (potentially) aberrant response behavior is indicated by small values of NCI (i.e., in the left tail of the sampling distribution).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMoDel"`), and parametric model imputation (`NA.method = "PMoDel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.

- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (IRT.PModel = "1PL", "2PL", or "3PL"). Item parameters (IP) and ability parameters (Ability) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

PFscores	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
PFstatistic	The person-fit statistic used.
PerfVects	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
ID.all0s	Row indices of all-0s response vectors removed from the analysis (if applicable).
ID.all1s	Row indices of all-1s response vectors removed from the analysis (if applicable).
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories (2 in this case).
IRT.PModel	The parametric IRT model used in case NA.method="PModel", otherwise NULL.
IP	The $I \times 3$ matrix of estimated item parameters in case NA.method="PModel", otherwise NULL.
Ability.PModel	The method used to estimate abilities in case NA.method="PModel", otherwise NULL.
Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Tatsuoka, K. K., and Tatsuoka, M. M. (1982) Detection of aberrant response patterns and their effect on dimensionality. *Journal of Educational Statistics*, **7(3)**, 215–231.

Tatsuoka, K. K., and Tatsuoka, M. M. (1983) Spotting erroneous rules of operation by the individual consistency index. *Journal of Educational Measurement*, **20**(3), 221–230.

van der Flier, H. (1977) Environmental factors and deviant response patterns. In Y. H. Poortinga (Ed.), *Basic problems in cross-cultural psychology*. Amsterdam: The Netherlands.

Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

## See Also

[G](#)

## Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the NCI scores:
NCI.out <- NCI(InadequacyData)
```

---

PerFit.PFS

*Compute several person-fit statistics*

---

## Description

Compute several person-fit statistics.

## Usage

```
PerFit.PFS(matrix, method=NULL, simplified=TRUE,
           NA.method = "Pairwise", Save.MatImp = FALSE,
           IP = NULL, IRT.PModel = NULL, Ability = NULL, Ability.PModel = NULL,
           mu = 0, sigma = 1)
```

## Arguments

matrix	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
method	Vector of person-fit statistics to be computed.
simplified	Logical. If FALSE, a list of PerFit objects is returned. If TRUE (default), a matrix with person-fit scores per respondent (one column per statistic) is reported.
NA.method	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
Save.MatImp	Logical. Save (imputed) data matrix to file? Default is FALSE.



IP	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
IRT.PModel	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL).
Ability	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of matrix. In case no ability parameters are available then Ability=NULL.
Ability.PModel	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL).
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

### Details

Function `PerFit.PFS` is a wrapper allowing to compute more than one person-fit statistic simultaneously.

### Value

If `simplified=TRUE`, a N-by-m data frame is returned, where N is the number of respondents and m is the number of methods.

If `simplified=FALSE` a list of m `PerFit` objects is returned.

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the lzstar, U3, and Ht scores:
PerFit.PFS(InadequacyData, method=c("lzstar", "U3", "Ht"))
```

---

PerFit.SE

*Compute standard errors for person fit statistics*

---

### Description

Compute standard errors for person fit statistics.

**Usage**

```
PerFit.SE(x)
```

**Arguments**

x                    Object of class "PerFit".

**Details**

Function `PerFit.SE` computes jackknife standard errors for the scores of the person fit statistic in object `x`.

**Value**

A matrix with two columns: `PFscores` shows the values of the person-fit statistic and `PFscores.SE` shows the estimated standard errors.

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the Ht scores:
Ht.out <- Ht(InadequacyData)

# Compute the SEs:
Ht.SE <- PerFit.SE(Ht.out)
Ht.SE
```

---

Personal biserial statistic

*Personal biserial statistic*

---

**Description**

Donlon and Fischer's personal biserial statistic.

**Usage**

```
r.pbis(matrix,
        NA.method = "Pairwise", Save.MatImp = FALSE,
        IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
        mu = 0, sigma = 1)
```

**Arguments**

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMoel" (default), and "PMoel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
<code>IRT.PMoel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PMoel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when <code>method="BM"</code> . Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when <code>method="BM"</code> . Default is 1.

**Details**

Donlon and Fischer (1968) suggested to use the correlation between a respondent's score vector and the item proportion-correct scores in the sample as a measure of person fit. Low values should be indicative of misfit of the response vector with respect to the group of respondents.

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMoel"`), and parametric model imputation (`NA.method = "PMoel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).

- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL", "2PL", or "3PL"`). Item parameters (IP) and ability parameters (Ability) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
<code>ID.all0s</code>	Row indices of all-0s response vectors removed from the analysis (if applicable).
<code>ID.all1s</code>	Row indices of all-1s response vectors removed from the analysis (if applicable).
<code>matrix</code>	The data matrix after imputation of missing values was performed (if applicable).
<code>Ncat</code>	The number of response categories (2 in this case).
<code>IRT.PModel</code>	The parametric IRT model used in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>IP</code>	The $I \times 3$ matrix of estimated item parameters in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>Ability.PModel</code>	The method used to estimate abilities in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>Ability</code>	The vector of $N$ estimated ability parameters in case <code>NA.method="PModel"</code> , otherwise NULL.
<code>NAs.method</code>	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Donlon, T. F., and Fischer, F. E. (1968) An index of an individual's agreement with group-defined item difficulties. *Educational and Psychological Measurement*, **28(1)**, 105–113.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32(6)**, 466–479.

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the r.pbis scores:
rpbis.out <- r.pbis(InadequacyData)
```

---

Physical Functioning    *The SF-36 physical functioning data*

---

**Description**

These data are from the Physical Functioning scale of the SF-36 (Ware and Sherbourne, 1992). Data consist of scores of 714 respondents on 10 polytomously scored items (0 = no, not limited at all; 1 = limited a little; 2 = limited a lot).

**Usage**

```
data(PhysFuncData)
```

**Format**

A 714x10 matrix of polytomous item scores (scores 0, 1, and 2).

**References**

Ware, J. E., Jr., and Sherbourne, C. D. (1992) The MOS 36-item short-form health survey (SF-36): Conceptual framework and item selection. *Medical Care*, **30**, 473–483.

**Examples**

```
data(PhysFuncData)
```

---

plot PerFit                      *Plot method for objects of class "PerFit"*

---

**Description**

Plot method for objects of class "PerFit".

**Usage**

```
## S3 method for class 'PerFit'
plot(x, cutoff.obj=NULL,
      ModelFit="NonParametric", Nreps=1000,
      IP=x$IP, IRT.PModel=x$IRT.PModel, Ability=x$Ability,
      Ability.PModel=x$Ability.PModel, mu=0, sigma=1,
      Blvl = 0.05, Breps = 1000, CIlvl = 0.95,
      UDlvl = NA,
      Type="Density", Both.scale=TRUE, Cutoff=TRUE, Cutoff.int=TRUE,
      Flagged.ticks = TRUE,
      Xlabel=NA, Xcex=1.5, title=NA, Tcex=1.5,
      col.area="lightpink", col.hist="lightblue", col.int="darkgreen",
      col.ticks="red", ...)
```

**Arguments**

x	Object of class "PerFit".
cutoff.obj	Object of class "PerFit.cutoff".
ModelFit	Method required to compute model-fitting item score patterns. The options available are "NonParametric" (default) and "Parametric".
Nreps	Number of model-fitting item score patterns generated. Default is 1000.
IP	Matrix with previously estimated item parameters. Default is x\$IP.
IRT.PModel	Parametric IRT model (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is x\$IRT.PModel.
Ability	Matrix with previously estimated item parameters. Default is x\$Ability.
Ability.PModel	Method to use in order to estimate the latent ability parameters (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is x\$Ability.PModel.
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.
Blvl	Significance level for bootstrap distribution (value between 0 and 1). Default is 0.05.
Breps	Number of bootstrap resamples. Default is 1000.
CIlvl	Level of bootstrap percentile confidence interval for the cutoff statistic.
UDlvl	User-defined cutoff level.
Type	Type of plot: "Histogram", "Density" (default), "Both".
Both.scale	Logical: Should the y-axis be adjusted so that both the histogram and the density graphics are completely visible? Default is TRUE.
Cutoff	Logical: Should the estimated cutoff be added to the plot? Default is TRUE.
Cutoff.int	Logical: Should an approximated (1-Blvl)% bootstrap confidence interval be added to the plot? Default is TRUE.

Flagged.ticks	Logical: Should ticks representing the flagged respondents be added to the plot? Default is TRUE.
Xlabel	Label of x-axis, otherwise a default label is shown.
Xcex	Font size of the label of x-axis.
title	Title of the plot, otherwise a default title is shown.
Tcex	Font size of the title of the plot.
col.area	Color of "flagging region".
col.hist	Color of histogram.
col.int	Color of bootstrap confidence interval.
col.ticks	Color of the ticks marking the flagged respondents.
...	Extra graphical parameters to be passed to plot.

### Details

This function plots the empirical distribution of the scores of the person-fit statistic specified by the "PerFit" class object `x`. A histogram, density, or a combination of both displays is possible.

The cutoff score may be provided by means of the `cutoff.obj` object, otherwise it is internally computed (for which the function parameters `ModelFit` through `CILvl` are required; see [cutoff](#) for more details). The value of the cutoff is superimposed to the plot when `Cutoff=TRUE`. In this case, the adequate "flagging region" is colored, thus indicating the range of values for which the person-fit statistic flags respondents as potentially displaying aberrant behavior. The option `Both.scale` was introduced to help to better tune the scale of the y-axis. Furthermore, the percentile confidence interval for the cutoff value (with confidence level defined by the `cutoff.obj`) is displayed in the x-axis, and ticks marking the flagged respondents are display on the top of the plot.

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### See Also

[cutoff](#), [flagged.resp](#), [PRFplot](#)

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the ZU3 scores:
ZU3.out <- ZU3(InadequacyData)

# Plot the sampling distribution of the ZU3 scores, with cutoff value based on a nominal 5% level,
# and 90% confidence interval:
plot(ZU3.out, Type="Both", Blvl=.05, CILvl = 0.90)
```

PRFplot

*Person response function (PRF)***Description**

Plot the nonparametric person response function with variability bands.

**Usage**

```
PRFplot(matrix, respID, h=.09, N.FPts=101,
        VarBands=FALSE, VarBands.area=FALSE, alpha=.05,
        Xlabel=NA, Xcex=1.5, Ylabel=NA, Ycex=1.5, title=NA, Tcex=1.5,
        NA.method="Pairwise", Save.MatImp=FALSE,
        IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
        mu = 0, sigma = 1,
        message = TRUE)
```

**Arguments**

matrix	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
respID	Vector specifying the respondents for whom PRFs are to be computed.
h	Bandwidth value. Default is 0.09.
N.FPts	Number of (equidistant) focal points in the [0,1] interval. Default is 101.
VarBands	Logical: Draw the (1-alpha) variability bands? Default is FALSE.
VarBands.area	Logical: Draw the area between the (1-alpha) variability bands? Default is FALSE.
alpha	Significance level for the variability bands. Default is 0.05.
Xlabel	Define label of x-axis, otherwise a default label is shown.
Xcex	Font size of the label of x-axis.
Ylabel	Define label of y-axis, otherwise a default label is shown.
Ycex	Font size of the label of y-axis.
title	Define the title of the plot, otherwise a default title is shown.
Tcex	Font size of the title of the plot.
NA.method	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
Save.MatImp	Logical. Save (imputed) data matrix to file? Default is FALSE.
IP	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.



IRT.PModel	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
Ability	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of matrix. In case no ability parameters are available then Ability=NULL.
Ability.PModel	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.
message	Display prompt message (one per person)? Default is TRUE.

### Details

Function PRFplot displays the so-called nonparametric person response functions (PRFs; Emons, Sijtsma, and Meijer, 2004; Sijtsma and Meijer, 2001), for the respondents identified in respID. The PRF relates item difficulty (0-1 range on the x-axis) with the associated probability of correct response (on the y-axis). The PRF is typically nonincreasing, implying that the probability of answering increasingly difficult items should (at least) not increase. The code is based on nonparametric kernel smoothing (Emons et al., 2004). The value of the PRF at each focal point (representing a difficulty parameter between 0 and 1) is estimated as a weighted sum score, where scores pertaining to items with difficulty close to the focal point are given the largest weights. The weights are functions of the Gaussian kernel function. It is necessary to specify a bandwidth value ( $h$ ) in order to compute the weights. The  $h$  value controls the trade-off between bias and sampling variation (Emons et al., 2004). Small  $h$  values reduce bias but increase variance, leading to PRFs that capture too much measurement error. Large  $h$  values, on the other hand, increase bias which renders PRFs that are often too flat, thus missing potentially relevant misfitting response behavior. Therefore, it is important to carefully specify the value  $h$ . Emons et al. (2004, pp. 10-13), after a simulation study, advised that " $h$  values between 0.07 and 0.11 are reasonable choices".

Moreover, variability bands of level  $\alpha$  (0.05 by default) can also be added to the plot. These bands are computed following the jackknife procedure explained in Emons et al. (2004).

The PRFs and variability bands for each respondent are approximated by means of functional data objects (e.g., Ramsay, Hooker, and Graves, 2009), with the help of the fda package. This procedure follows two steps:

- Compute a B-splines basis system. This basis consists of a set of (thirteen) piecewise polynomials, all of degree three/order four (i.e., cubic polynomial segments), with one knot per break point. This allows any two consecutive splines,  $sp1$  and  $sp2$ , with common break point BP, verifying  $sp1(BP) = sp2(BP)$ ,  $sp1'(BP) = sp2'(BP)$ , and  $sp1''(BP) = sp2''(BP)$ . At 0 and 1 (extremes of the x-range), four (= order) knots are used.
- Specify coefficients  $c$  for the B-splines basis system computed above and then create functional data objects. Based on smoothing using regression analysis (Ramsay et al., 2009, section 4.3).

Missing values in matrix are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (NA.method = "Hotdeck"), nonparametric model imputation (NA.method = "NPMModel"), and parametric model imputation (NA.method = "PModel"); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL", "2PL", or "3PL"`). Item parameters (`IP`) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

PRFplot returns three functional data objects (for the PRFs, lower-bound of the variability bands, and upper-band of the variability bands) for all respondents in the sample.

### Value

The output is a list with three functional data objects of class `fd` (see package `fda`):

<code>PRF.FDO</code>	Functional data object of the PRFs for the entire sample.
<code>VarBandsLow.FDO</code>	Functional data object of the lower-bound of the variability bands for the entire sample.
<code>VarBandsHigh.FDO</code>	Functional data object of the upper-bound of the variability bands for the entire sample.

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Emons, W. M., Sijtsma, K., and Meijer, R. R. (2004) Testing hypotheses about the person-response function in person-fit analysis. *Multivariate Behavioral Research*, **39**(1), 1–35.
- Ramsay, J. O., Hooker, G., and Graves, S. (2009) *Functional data analysis with R and MATLAB*. New York: US.
- Sijtsma, K., and Meijer, R. R. (2001) The person response function as a tool in person-fit research. *Psychometrika*, **66**(2), 191–207.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

**See Also**

[cutoff](#), [plot.PerFit](#), [flagged.resp](#)

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# As an example, compute the Ht person-fit scores:
Ht.out <- Ht(InadequacyData)
Ht.out$PFscores

# Determine which respondents were flagged by Ht at 1% level:
set.seed(124) # To fix the random seed generator.
Ht.flagged <- flagged.resp(Ht.out, Blvl=.01, scores=FALSE)
Ht.flagged <- Ht.flagged$PFscores[,1]
# Flagged respondents: 30 37 46 49 137 216 531.

# Plot the PRFs of the first three flagged respondents:
Flagged <- PRFplot(InadequacyData, respID=Ht.flagged[1:3])
# Plot the person response function of respondent 35 (not flagged as aberrant):
PRFplot(InadequacyData, respID=35)
# Plot the PRFs of all respondents:
plot(Flagged$PRF.FDO)
```

---

print PerFit	<i>Print method for objects of class "PerFit"</i>
--------------	---

---

**Description**

Print method for objects of class "PerFit".

**Usage**

```
## S3 method for class 'PerFit'
print(x, ...)
```

**Arguments**

x	Object of class "PerFit".
...	Additional arguments to be passed to print.

**Details**

For a given object of class `PerFit`, this function displays the scores of the person-fit statistic.

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**See Also**

[cutoff](#), [flagged.resp](#), [plot.PerFit](#), [summary.PerFit](#)

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the ZU3 scores:
ZU3.out <- ZU3(InadequacyData)

print(ZU3.out)
```

---

summary PerFit

*Summary method for objects of class "PerFit"*


---

**Description**

Summary method for objects of class "PerFit".

**Usage**

```
## S3 method for class 'PerFit'
summary(object, cutoff.obj=NULL,
        ModelFit="NonParametric", Nreps=1000,
        IP=object$IP, IRT.PModel=object$IRT.PModel, Ability=object$Ability,
        Ability.PModel=object$Ability.PModel, mu=0, sigma=1,
        Blvl = 0.05, Breps = 1000, Cilvl = 0.95,
        Ud1vl = NA, ...)
```

**Arguments**

object	Object of class "PerFit".
cutoff.obj	Object of class "PerFit.cutoff".
ModelFit	Method required to compute model-fitting item score patterns. The options available are "NonParametric" (default) and "Parametric".
Nreps	Number of model-fitting item score patterns generated. Default is 1000.
IP	Matrix with previously estimated item parameters. Default is object\$IP.
IRT.PModel	Parametric IRT model (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is object\$IRT.PModel.
Ability	Matrix with previously estimated item parameters. Default is object\$Ability.

Ability.PModel	Method to use in order to estimate the latent ability parameters (required if "ModelFit=Parametric" or if the person fit statistic is parametric). Default is object\$Ability.PModel.
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.
Blvl	Significance level for bootstrap distribution (value between 0 and 1). Default is 0.05.
Breps	Number of bootstrap resamples. Default is 1000.
UDlvl	User-defined cutoff level.
CIlvl	Level of bootstrap percentile confidence interval for the cutoff statistic.
...	Additional arguments to be passed to summary.

### Details

For a given object of class `PerFit`, this function prints: The PFS used, the cutoff value, the tail of the distribution of the person-fit statistic associated to misfit, the proportion of flagged respondents in the sample, and their row indices.

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### See Also

[cutoff](#), [flagged.resp](#), [plot.PerFit](#), [summary.PerFit](#)

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the ZU3 scores:
ZU3.out <- ZU3(InadequacyData)

summary(ZU3.out)
```

---

The caution statistic *C.Sato, Cstar person-fit statistics*

---

### Description

Computes the caution statistic *C.Sato* and the modified caution statistic *Cstar*.

**Usage**

```

C.Sato(matrix,
        NA.method = "Pairwise", Save.MatImp = FALSE,
        IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
        mu = 0, sigma = 1)

Cstar(matrix,
        NA.method = "Pairwise", Save.MatImp = FALSE,
        IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
        mu = 0, sigma = 1)

```

**Arguments**

matrix	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
NA.method	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
Save.MatImp	Logical. Save (imputed) data matrix to file? Default is FALSE.
IP	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
IRT.PModel	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
Ability	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of matrix. In case no ability parameters are available then Ability=NULL.
Ability.PModel	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
mu	Mean of the apriori distribution. Only used when method="BM". Default is 0.
sigma	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

**Details**

The C.Sato statistic (also referred to as C in the literature) was proposed by Sato (1975):

$$C.Sato = 1 - \frac{cov(x_n, p)}{cov(x_n^*, p)},$$

where  $x_n$  is the 0-1 response vector of respondent  $n$ ,  $p$  is the vector of item proportions-correct, and  $x_n^*$  is the so-called Guttman vector containing correct answers for the easiest items (i.e., with the largest proportion-correct values) only. C.Sato is zero for Guttman vectors and its value tends to increase for response vectors that depart from the group's answering pattern, hence warning

the researcher to be *cautious* about interpreting such item scores. Therefore, (potentially) aberrant response behavior is indicated by large values of C.Sato (i.e., in the right tail of the sampling distribution).

Harnisch and Linn (1981) proposed a modified version of the caution statistic which bounds the caution statistic between 0 and 1 (also referred to as C\* or MCI in the literature):

$$Cstar = \frac{cov(x_n^*, p) - cov(x_n, p)}{cov(x_n^*, p) - cov(x'_n, p)},$$

where  $x'_n$  is the reversed Guttman vector containing correct answers for the hardest items (i.e., with the smallest proportion-correct values) only. Cstar is sensitive to the so-called Guttman errors. A Guttman error is a pair of scores (0,1), where the 0-score pertains to the easiest item and the 1-score pertains to the hardest item. Cstar ranges between 0 (perfect Guttman vector) and 1 (reversed Guttman error), thus larger values indicate potential aberrant response behavior.

These statistics are not computed for rows of `matrix` that consist of only 0s or only 1s (NA values are returned instead).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL"`, `"2PL"`, or `"3PL"`). Item parameters (`IP`) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

## Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
<code>ID.all0s</code>	Row indices of all-0s response vectors removed from the analysis (if applicable).

ID.all1s	Row indices of all-1s response vectors removed from the analysis (if applicable).
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories (2 in this case).
IRT.PModel	The parametric IRT model used in case NA.method="PModel", otherwise NULL.
IP	The $I \times 3$ matrix of estimated item parameters in case NA.method="PModel", otherwise NULL.
Ability.PModel	The method used to estimate abilities in case NA.method="PModel", otherwise NULL.
Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

**Author(s)**

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

**References**

- Harnisch, D. L., and Linn, R. L. (1981) Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. *Journal of Educational Measurement*, **18(3)**, 133–146.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16(4)**, 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25(2)**, 107–135.
- Sato, T. (1975) *The construction and interpretation of S-P tables*. Tokyo: Meiji Tosho.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32(6)**, 466–479.

**See Also**

[Ht](#)

**Examples**

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the C.Sato scores:
C.out <- C.Sato(InadequacyData)

# Compute the Cstar scores:
Cstar.out <- Cstar(InadequacyData)
```



U3, ZU3

*U3, ZU3 person-fit statistics***Description**

van der Flier's U3 and ZU3 person-fit statistics.

**Usage**

```
U3(matrix,
    NA.method = "Pairwise", Save.MatImp = FALSE,
    IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
    mu = 0, sigma = 1)

ZU3(matrix,
    NA.method = "Pairwise", Save.MatImp = FALSE,
    IP = NULL, IRT.PModel = "2PL", Ability = NULL, Ability.PModel = "ML",
    mu = 0, sigma = 1)
```

**Arguments**

<code>matrix</code>	Data matrix of dichotomous item scores: Persons as rows, items as columns, item scores are either 0 or 1, missing values allowed.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item, and three columns ([,1] item discrimination; [,2] item difficulty; [,3] lower-asymptote, also referred to as pseudo-guessing parameter). In case no item parameters are available then IP=NULL.
<code>IRT.PModel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "1PL", "2PL" (default), and "3PL".
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PModel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "ML" (default), "BM", and "WL".
<code>mu</code>	Mean of the apriori distribution. Only used when method="BM". Default is 0.
<code>sigma</code>	Standard deviation of the apriori distribution. Only used when method="BM". Default is 1.

### Details

Suppose the items are ordered in decreasing proportion-correct score,  $p_1 > p_2 > \dots > p_I$  ( $I$  = number of items). Given response vector  $(X_1, X_2, \dots, X_I)$  with total score  $S = \sum_i X_i$ , van der Flier (1980, 1982) defined the U3 statistic as

$$\frac{\sum_{i=1}^S p_i - \sum_{i=1}^I p_i}{\sum_{i=1}^S p_i - \sum_{i=I-S+1}^I p_i}.$$

U3 varies from 0 for perfect Guttman response vectors (i.e., with only the easiest items correct) through 1 for reversed Guttman response vectors (i.e., with only the hardest items correct). Hence, increasingly large U3 values provide stronger indications of answering misfit.

U3 scores are dependent on the number-correct score, hence van der Flier proposed ZU3 as a standardization (formulas to compute  $E(U3)$  and  $Var(U3)$  can be found in van der Flier, 1982). ZU3 is supposed to be asymptotically approximated by the standard normal distribution, but this approximation is not without problems (see Emons, Meijer, and Sijtsma, 2002).

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMModel"`), and parametric model imputation (`NA.method = "PModel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from Bernoulli distributions with probabilities estimated by means of parametric IRT models (`IRT.PModel = "1PL"`, `"2PL"`, or `"3PL"`). Item parameters (`IP`) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

### Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	A message indicating whether perfect response vectors (all-0s or all-1s) were removed from the analysis.
<code>ID.all0s</code>	Row indices of all-0s response vectors removed from the analysis (if applicable).

ID.all1s	Row indices of all-1s response vectors removed from the analysis (if applicable).
matrix	The data matrix after imputation of missing values was performed (if applicable).
Ncat	The number of response categories (2 in this case).
IRT.PModel	The parametric IRT model used in case NA.method="PModel", otherwise NULL.
IP	The $I \times 3$ matrix of estimated item parameters in case NA.method="PModel", otherwise NULL.
Ability.PModel	The method used to estimate abilities in case NA.method="PModel", otherwise NULL.
Ability	The vector of $N$ estimated ability parameters in case NA.method="PModel", otherwise NULL.
NAs.method	The imputation method used (if applicable).

### Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

### References

- Emons, W. M., Meijer, R. R., and Sijtsma, K. (2002). Comparing simulated and theoretical sampling distributions of the U3 person-fit statistic. *Applied Psychological Measurement*, **26**(1), 88–108.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16**(4), 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25**(2), 107–135.
- van der Flier, H. (1980) *Vergelijkbaarheid van individuele testprestaties* [Comparability of individual test performance]. Lisse: The Netherlands.
- van der Flier, H. (1982) Deviant response patterns and comparability of test scores. *Journal of Cross-Cultural Psychology*, **13**(3), 267–298.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

### See Also

[U3poly](#)

### Examples

```
# Load the inadequacy scale data (dichotomous item scores):
data(InadequacyData)

# Compute the U3 scores:
U3.out <- U3(InadequacyData)

# Compute the ZU3 scores:
ZU3.out <- ZU3(InadequacyData)
```

U3poly

*U3poly person-fit statistic***Description**

Generalization of van der Flier's U3 person-fit statistic to polytomously scored items.

**Usage**

```
U3poly(matrix, Ncat,
        NA.method = "Pairwise", Save.MatImp = FALSE,
        IP = NULL, IRT.PModel = "GRM", Ability = NULL, Ability.PModel = "EAP")
```

**Arguments**

<code>matrix</code>	A data matrix of polytomous item scores: Persons as rows, items as columns, item scores are integers between 0 and (Ncat-1), missing values allowed.
<code>Ncat</code>	Number of answer options for each item.
<code>NA.method</code>	Method to deal with missing values. The default is pairwise elimination ("Pairwise"). Alternatively, simple imputation methods are also available. The options available are "Hotdeck", "NPMModel" (default), and "PModel".
<code>Save.MatImp</code>	Logical. Save (imputed) data matrix to file? Default is FALSE.
<code>IP</code>	Matrix with previously estimated item parameters: One row per item. The first (Ncat-1) columns contain the between-categories threshold parameters (for the GRM) or the item step difficulties (for the PCM and the GPCM). The last, Ncat-th, column has the slopes. In case no item parameters are available then IP=NULL.
<code>IRT.PModel</code>	Specify the IRT model to use in order to estimate the item parameters (only if IP=NULL). The options available are "PCM", "GPCM", and "GRM" (default).
<code>Ability</code>	Vector with previously estimated latent ability parameters, one per respondent, following the order of the row index of <code>matrix</code> . In case no ability parameters are available then Ability=NULL.
<code>Ability.PModel</code>	Specify the method to use in order to estimate the latent ability parameters (only if Ability=NULL). The options available are "EB", "EAP" (default), and "MI".

**Details**

Emons (2008) generalized the U3 statistic (van der Flier, 1980, 1982) to polytomous items. The idea is based on the so-called item-step difficulty, which is the probability of moving from answer category ( $c$ ) to answer category ( $c + 1$ ) ( $c = 0, \dots, Ncat - 2$ ).

U3poly varies from 0 (no misfit) through 1 (extreme misfit). Hence, increasingly large U3poly values provide stronger indications of answering misfit.

The number of answer options, Ncat, is the same for all items.

U3poly reduces to U3 when Ncat=2.

Missing values in `matrix` are dealt with by means of pairwise elimination by default. Alternatively, single imputation is also available. Three single imputation methods exist: Hotdeck imputation (`NA.method = "Hotdeck"`), nonparametric model imputation (`NA.method = "NPMoDel"`), and parametric model imputation (`NA.method = "PMoDel"`); see Zhang and Walker (2008).

- Hotdeck imputation replaces missing responses of an examinee ('recipient') by item scores from the examinee which is closest to the recipient ('donor'), based on the recipient's non-missing item scores. The similarity between nonmissing item scores of recipients and donors is based on the sum of absolute differences between the corresponding item scores. The donor's response pattern is deemed to be the most similar to the recipient's response pattern in the group, so item scores of the former are used to replace the corresponding missing values of the latter. When multiple donors are equidistant to a recipient, one donor is randomly drawn from the set of all donors.
- The nonparametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities defined by donors with similar total score than the recipient (based on all items except the NAs).
- The parametric model imputation method is similar to the hotdeck imputation, but item scores are generated from multinomial distributions with probabilities estimated by means of parametric IRT models (`IRT.PMoDel = "PCM", "GPCM", or "GRM"`). Item parameters (`IP`) and ability parameters (`Ability`) may be provided for this purpose (otherwise the algorithm finds estimates for these parameters).

## Value

An object of class "PerFit", which is a list with 12 elements:

<code>PFscores</code>	A list of length $N$ (number of respondents) with the values of the person-fit statistic.
<code>PFstatistic</code>	The person-fit statistic used.
<code>PerfVects</code>	Not applicable.
<code>ID.all0s</code>	Not applicable.
<code>ID.all1s</code>	Not applicable.
<code>matrix</code>	The data matrix after imputation of missing values was performed (if applicable).
<code>Ncat</code>	The number of response categories.
<code>IRT.PMoDel</code>	The parametric IRT model used.
<code>IP</code>	The $I \times N_{cat}$ matrix of estimated item parameters.
<code>Ability.PMoDel</code>	The method used to estimate abilities used.
<code>Ability</code>	The vector of $N$ estimated ability parameters.
<code>NAs.method</code>	The imputation method used (if applicable).

## Author(s)

Jorge N. Tendeiro <j.n.tendeiro@rug.nl>

## References

- Emons, W. M. (2008) Nonparametric person-fit analysis of polytomous item scores. *Applied Psychological Measurement*, **32**(3), 224–247.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement In Education*, **16**(4), 277–298.
- Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, **25**(2), 107–135.
- van der Flier, H. (1980) *Vergelijkbaarheid van individuele testprestaties* [Comparability of individual test performance]. Lisse: The Netherlands.
- van der Flier, H. (1982) Deviant response patterns and comparability of test scores. *Journal of Cross-Cultural Psychology*, **13**(3), 267–298.
- Zhang, B., and Walker, C. M. (2008) Impact of missing data on person-model fit and person trait estimation. *Applied Psychological Measurement*, **32**(6), 466–479.

## See Also

[U3](#), [ZU3](#)

## Examples

```
# Load the physical functioning data (polytomous item scores):  
data(PhysFuncData)  
  
# Compute the U3poly scores:  
U3poly.out <- U3poly(PhysFuncData, Ncat=3)
```

# Index

## \*Topic **datasets**

Inadequacy scale data, [22](#)  
Intelligence data, [23](#)  
Physical Functioning, [37](#)

## \*Topic **univar**

A, D, E statistics, [6](#)  
cutoff, [9](#)  
flagged.resp, [11](#)  
G, Gnormed, [13](#)  
Gpoly, Gnormed.poly, [16](#)  
Ht statistic, [19](#)  
lz, lzstar, [23](#)  
lzpoly, [27](#)  
NCI statistic, [29](#)  
PerFit.PFS, [32](#)  
PerFit.SE, [33](#)  
Personal biserial statistic, [34](#)  
plot PerFit, [37](#)  
PRFplot, [40](#)  
print PerFit, [43](#)  
summary PerFit, [44](#)  
The caution statistic, [45](#)  
U3, ZU3, [49](#)  
U3poly, [52](#)

A, D, E statistics, [6](#)

A.KB, [3](#)

A.KB (A, D, E statistics), [6](#)

C.Sato, [3](#), [20](#), [22](#)

C.Sato (The caution statistic), [45](#)

Cstar, [3](#)

Cstar (The caution statistic), [45](#)

cutoff, [4](#), [9](#), [12](#), [13](#), [39](#), [43–45](#)

D.KB, [3](#)

D.KB (A, D, E statistics), [6](#)

E.KB, [3](#)

E.KB (A, D, E statistics), [6](#)

flagged.resp, [4](#), [11](#), [11](#), [39](#), [43–45](#)

G, [3](#), [19](#), [30](#), [32](#)

G (G, Gnormed), [13](#)

G, Gnormed, [13](#)

Gnormed, [3](#), [19](#)

Gnormed (G, Gnormed), [13](#)

Gnormed.poly, [3](#)

Gnormed.poly (Gpoly, Gnormed.poly), [16](#)

Gpoly, [3](#)

Gpoly (Gpoly, Gnormed.poly), [16](#)

Gpoly, Gnormed.poly, [16](#)

Ht, [3](#), [48](#)

Ht (Ht statistic), [19](#)

Ht statistic, [19](#)

Inadequacy scale data, [22](#)

InadequacyData (Inadequacy scale data),  
[22](#)

Intelligence data, [23](#)

IntelligenceData (Intelligence data), [23](#)

irtoys, [24](#)

lz, [3](#), [27](#), [29](#)

lz (lz, lzstar), [23](#)

lz, lzstar, [23](#)

lzpoly, [3](#), [26](#), [27](#)

lzstar, [3](#), [29](#)

lzstar (lz, lzstar), [23](#)

NCI, [3](#), [16](#)

NCI (NCI statistic), [29](#)

NCI statistic, [29](#)

PerFit (PerFit-package), [2](#)

PerFit-package, [2](#)

PerFit.PFS, [32](#)

PerFit.SE, [33](#)

Personal biserial statistic, [34](#)

PhysFuncData (Physical Functioning), [37](#)

Physical Functioning, [37](#)  
plot, [4](#)  
plot PerFit, [37](#)  
plot.PerFit, [11](#), [13](#), [43–45](#)  
plot.PerFit(plot PerFit), [37](#)  
PRFplot, [4](#), [11](#), [13](#), [39](#), [40](#)  
print PerFit, [43](#)  
print.PerFit(print PerFit), [43](#)

r.pbis, [3](#)  
r.pbis(Personal biserial statistic), [34](#)

summary PerFit, [44](#)  
summary.PerFit, [44](#), [45](#)  
summary.PerFit(summary PerFit), [44](#)

The caution statistic, [45](#)

U3, [3](#), [54](#)  
U3 (U3, ZU3), [49](#)  
U3, ZU3, [49](#)  
U3poly, [4](#), [51](#), [52](#)

ZU3, [3](#), [54](#)  
ZU3 (U3, ZU3), [49](#)