Package 'difR'

June 13, 2012

Type Package

Title Collection of methods to detect dichotomous differential item
functioning (DIF) in psychometrics

Version 4.3

Date 2012-06-13

Author David Magis (U Liege), Sebastien Beland (UQAM) and Gilles Raiche (UQAM)

Maintainer David Magis <david.magis@ulg.ac.be>

Depends R (>= 2.14.0), lme4, ltm

Description The difR package contains several traditional methods to detect DIF in dichoto-
mously scored items. Both uniform and non-uniform DIF effects can be detected, with meth-
ods relying upon item response models or not. Some methods deal with more than one focal group.

License file LICENSE

LazyLoad yes

R topics documented:

---

difR-package   *Collection of methods to detect dichotomous differential item function-*
                      ing (DIF) in psychometrics

Description

The difR package contains several traditional methods to detect DIF in dichotomously scored items. Both uniform and non-uniform DIF effects can be detected, with methods relying upon item response models or not. Some methods deal with more than one focal group.

Methods currently available are:

1. Transformed Item Difficulties (TID) method (Angoff and Ford, 1973)
2. Mantel-Haenszel (Holland and Thayer, 1988)
3. Standardization (Dorans and Kullick, 1986)
4. Breslow-Day (Aguerri et al., 2009; Penfield, 2003)
5. Logistic regression (Swaminathan and Rogers, 1990)
6. Lord's chi-square test (Lord, 1980)
7. Raju's area (Raju, 1990)
8. Likelihood-ratio test (Thissen, Steinberg and Wainer, 1988)
9. Generalized Mantel-Haenszel (Penfield, 2001)
10. Generalized logistic regression (Magis, Raiche, Beland and Gerard, 2011)
11. Generalized Lord's chi-square test (Kim, Cohen and Park, 1995).

The difR package is further described in Magis, Beland, Tuerlinckx and De Boeck (2010).

Details

Package:	difR
Type:	Package
Version:	4.2
Date:	2012-03-29
Depends:	R (>= 2.8.0), lme4, ltm
License:	see file LICENSE

Author(s)

Sebastien Beland

Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)

Universite du Quebec a Montreal

<sebastien.beland.1@hotmail.com>

David Magis

Research Group of Quantitative Psychology and Individual Differences

Katholieke Universiteit Leuven

<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/

Gilles Raiche

Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)

Universite du Quebec a Montreal

<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Aguerri, M.E., Galibert, M.S., Attorresi, H.F. and Maranon, P.P. (2009). Erroneous detection of nonuniform DIF using the Breslow-Day test in a short test. *Quality and Quantity, 43*, 35-44.

Angoff, W. H., and Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. Journal of Educational Measurement, 95-106.

Dorans, N. J. and Kullick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*, 355-368.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Dirs.), *Test validity*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2011). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

Penfield, R. D. (2001). Assessing differential item functioning among multiple groups: a comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.

Penfield, R.D. (2003). Application of the Breslow-Day test of trend in odds ratio heterogeneity to the detection of nonuniform DIF. *Alberta Journal of Educational Research, 49*, 231-243.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*, 197-207.

Swaminathan, H. and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.

Thissen, D., Steinberg, L. and Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer and H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

See Also

Other useful packages can be found in the *R Psychometric* task view. See http://cran.stat.sfu.ca/web/views/Psychometrics.html for further details.

---

breslowDay   *Breslow-Day DIF statistic*

Description

Computes Breslow-Day statistics for DIF detection.

Usage

```
breslowDay(data, member, anchor=1:ncol(data), BDstat="BD")
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and one entries only. See **Details**.
anchor	a vector of integer values specifying which items (all by default) are currently considered as anchor (DIF free) items. See **Details**.
BDstat	character specifying the DIF statistic to be used. Possible values are "BD" (default) and "trend". See **Details**.

Details

breslowDay computes one of the Breslow-Day statistics (1980) in the specific framework of differential item functioning. It forms the basic command of [difBD](#) and is specifically designed for this call.

The data are supplied by the data argument, with one row per subject and one column per item. Missing values are allowed but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership, specified by the member argument, must hold only zeros and ones, a value of zero corresponding to the reference group and a value of one to the focal group.

Option anchor sets the items which are considered as anchor items for computing Breslow-Day DIF statistics. Items other than the anchor items and the tested item are discarded. anchor must hold

integer values specifying the column numbers of the corresponding anchor items. It is primarily designed to perform item purification.

Two test statistics are available: the usual Breslow-Day statistic for testing homogeneous association (Aguerri, Galibert, Attorresi and Maranon, 2009) and the trend test statistic for assessing some monotonic trend in the odss ratios (Penfield, 2003). The DIF statistic is supplied by the BDstat argument, with values `"BD"` (default) for the usual statistic and `"trend"` for the trend test statistic.

Value

A list with two arguments:

res   A matrix with one row per item and three columns: the first one contains the Breslow-Day statistic values, the second column indicates the degrees of freedom, and the last column displays the asymptotic *p*-values.

BDstat   the value of the BDstat argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Aguerri, M.E., Galibert, M.S., Attorresi, H.F. and Maranon, P.P. (2009). Erroneous detection of nonuniform DIF using the Breslow-Day test in a short test. *Quality and Quantity, 43*, 35-44.

Breslow, N.E. and Day, N.E. (1980). *Statistical methods in cancer research, vol. I: The analysis of case-control studies*. Scientific Publication No 32. International Agency for Research on Cancer, Lyon, France.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R.D. (2003). Application of the Breslow-Day test of trend in odds ratio heterogeneity to the detection of nonuniform DIF. *Alberta Journal of Educational Research, 49*, 231-243.

See Also

difBD, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # With all items as anchor items
 breslowDay(verbal[,1:24], verbal[,26])

 # With all items as anchor items and trend
 # test statistic
 breslowDay(verbal[,1:24], verbal[,26], BDstat = "trend")

 # Removing item 3 from the set of anchor items
 breslowDay(verbal[,1:24], verbal[,26], anchor=c(1:5,7:24))

## End(Not run)
```

---

contrastMatrix	*Contrast matrix for computing generalized Lord's chi-squared DIF statistic*

Description

This command sets the appropriate contrast matrix C for computing the generalized Lord's chi-squared statistics in the framework of DIF detection among multiple groups.

Usage

```
contrastMatrix(nrFocal, model)
```

Arguments

nrFocal   numeric: the number of focal groups.

model   character: the logistic model to be fitted (either ″1PL″, ″2PL″, ″3PL″ or ″3PLc″).
               See **Details**.

Details

The contrast matrix C is necessary to calculate the generalized Lord's chi-squared statistic. It is designed to perform accurate tests of equality of item parameters accross the groups of examinees (see Kim, Cohen and Park, 1995). This is a subroutine for the command genLordChi2 which returns the DIF statistics.

The number of focal groups has to be specified by the argument nrFocal. Moreover, four logistic IRT models can be considered: the 1PL, 2PL and 3PL models can be set by using their acronyms (e.g. ″1PL″ for 1PL model, and so on). It is also possible to consider the constrained 3PL model, where all pseudo-guessing values are equal across the groups of examinees and take some predefined values which do not need to be supplied here. This model is specified by the value ″3PLc″ for argument model.

Value

A contrast matrix designed to test equality of item parameter estimates from the specified `model` and with `nrFocal` focal groups. The output matrix has a number of rows equal to `nrFocal` times the number of tested parameters (one for 1PL model, two for 2PL and constrained 3PL models, three for 3PL model). The number of columns is equal to (`nrFocal`+1) times the number of tested parameters. See Kim, Cohen and Park (1995) for further details.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

genLordChi2, difGenLord

Examples

```
## Not run:

 # Contrast matrices with 1PL model and several focal groups
 contrastMatrix(2, "1PL")
 contrastMatrix(3, "1PL")
 contrastMatrix(4, "1PL")

 # Contrast matrices with 2PL, constrained and unconstrained 3PL models,
 # and three focal groups
 contrastMatrix(3, "2PL")
 contrastMatrix(3, "3PLc")
 contrastMatrix(3, "3PL")

## End(Not run)
```

---

dichoDif   *Comparison of DIF detection methods*

Description

This function compares the specified DIF detection methods with respect to the detected items.

Usage

```
dichoDif(Data, group, focal.name, method, props=NULL, thrTID=1.5,
   alpha=0.05, MHstat="MHChisq", correct=TRUE, exact=FALSE,
   stdWeight="focal", thrSTD=0.1, BDstat="BD", type="both",
   criterion="LRT", model="2PL", c=NULL, engine="ltm", discr=1,
   irtParam=NULL, same.scale=TRUE, signed=FALSE, purify=FALSE,
   nrIter=10, save.output=FALSE, output=c("out","default"))
## S3 method for class 'dichoDif'
print(x, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within Data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
method	character vector specifying the different methods to be compared. See **Details**.
props	either NULL (default) or a two-column matrix with proportions of success in the reference group and the focal group. See **Details**.
thrTID	numeric: the threshold for detecting DIF items with TID method (default is 1.5).
alpha	numeric: significance level (default is 0.05).
MHstat	character: specifies the DIF statistic to be used for DIF identification. Possible values are "MHChisq" (default) and "logOR". See **Details**.
correct	logical: should the Mantel-Haenszel continuity correction be used? (default is TRUE).
exact	logical: should an exact test be computed? (default is FALSE).
stdWeight	character: the type of weights used for the standardized P-DIF statistic. Possible values are "focal" (default), "reference" and "total". See **Details**.
thrSTD	numeric: the threshold (cut-score) for standardized P-DIF statistic (default is 0.10).
BDstat	character specifying the DIF statistic to be used. Possible values are "BD" (default) and "trend". See **Details**.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	a character string specifying which DIF statistic is computed. Possible values are "LRT" (default) or "Wald". See **Details**.

model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL"). Default is "2PL".
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.
signed	logical: should the Raju's statistics be computed using the signed (TRUE) or unsigned (FALSE, default) area? See **Details**.
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	result from a dichoDif class object.
...	other generic parameters for the print function.

Details

dichoDif is a generic function which calls one or several DIF detection methods and summarize their output. The possible methods are: "TID" for Transformed Item Difficulties (TID) method (Angoff and Ford, 1973), "MH" for mantel-Haenszel (Holland and Thayer, 1988), "Std" for standardization (Dorans and Kulick, 1986), "Logistic" for logistic regression (Swaminathan and Rogers, 1990), "BD" for Breslow-Day method (Penfield, 2003), "Lord" for Lord's chi-square test (Lord, 1980), "Raju" for Raju's area method (Raju, 1990), and "LRT" for likelihood-ratio test method (Thissen, Steinberg and Wainer, 1988).

If method has a single component, the output of dichoDif is exactly the one provided by the method itself. Otherwise, the main output is a matrix with one row per item and one column per method. For each specified method and related arguments, items detected as DIF and non-DIF are respectively encoded as "DIF" and "NoDIF". When printing the output an additional column is added, counting the number of times each item was detected as functioning differently (Note: this is just an informative summary, since the methods are obviously not independent for the detection of DIF items).

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from either the computation of the sum-scores, the fitting of the logistic models or the IRT models (according to the method).

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

With the TID method, one can alternatively provide the matrix of proportions of success in for each item in each group. This matrix must have the same format as that provided to the trItemDiff function; see the corresponding help file for further details.

For Lord and Raju methods, one can specify either the IRT model to be fitted (by means of model, c, engine and discr arguments), or the item parameter estimates with arguments irtParam and same.scale. See difLord and difRaju for further details.

The threshold for detecting DIF items depends on the method. For standardization it has to be fully specified (with the thr argument), as well as for the TID method (through the thrTID argument). For the other methods it is depending on the significance level set by alpha.

For Mantel-Haenszel method, the DIF statistic can be either the Mantel-Haenszel chi-square statistic or the log odds-ratio statistic. The method is specified by the argument MHstat, and the default value is "MHChisq" for the chi-square statistic. Moreover, the option correct specifies whether the continuity correction has to be applied to Mantel-Haenszel statistic. See difMH for further details.

By default, the asymptotic Mantel-Haenszel statistic is computed. However, the exact statistics and related P-values can be obtained by specifying the logical argument exact to TRUE. See Agresti (1990, 1992) for further details about exact inference.

The weights for computing the standardized P-DIF statistics are defined through the argument stdWeight, with possible values "focal" (default value), "reference" and "total". See stdPDIF for further details.

For Breslow-Day method, two test statistics are available: the usual Breslow-Day statistic for testing homogeneous association (Aguerri, Galibert, Attorresi and Maranon, 2009) and the trend test statistic for assessing some monotonic trend in the odss ratios (Penfield, 2003). The DIF statistic is supplied by the BDstat argument, with values "BD" (default) for the usual statistic and "trend" for the trend test statistic.

For logistic regression, the argument type permits to test either both uniform and nonuniform effects simultaneously (type="both"), only uniform DIF effect (type="udif") or only nonuniform DIF effect (type="nudif"). The criterion argument specifies the DIF statistic to be computed, either the likelihood ratio test statistic (by setting criterion="LRT") or the Wald test (by setting criterion="Wald"). See Logistik for further details.

For Raju's method, the type of area (signed or unsigned) is fixed by the logical signed argument, with default value FALSE (i.e. unsigned areas). See RajuZ for further details.

Item purification can be requested by specifying purify option to TRUE. Recall that item purification process is slightly different for IRT and for non-IRT based methods. See the corresponding methods for further information.

The output of the dichoDif function can be stored in a text file by fixing save.output and output appropriately. See the help file of selectDif function (or any other DIF method) for further information.

Value

Either the output of one of the DIF detection methods, or a list of class "dichoDif" with the following arguments:

DIF	a character matrix with one row per item and whose columns refer to the different specified detection methods. See **Details**.
props	the value of the props argument.
thrTID	the value of the thrTID argument.

correct	the value of `correct` argument.
exact	the value of `exact` argument.
alpha	the significance level `alpha`.
MHstat	the value of the `MHstat` argument.
stdWeight	the value of the `stdWeight` argument.
thrSTD	the value of `thrSTD` argument.
BDstat	the value of the `BDstat` argument.
type	the value of the `type` argument.
criterion	the value of the `criterion` argument.
model	the value of `model` argument.
c	the value of `c` argument.
engine	The value of the `engine` argument.
discr	the value of the `discr` argument.
irtParam	the value of `irtParam` argument.
same.scale	the value of `same.scale` argument.
purification	the value of `purify` argument.
nrPur	an integer vector (of length equal to the number of methods) with the number of iterations in the purification process. Returned only if `purify` is TRUE.
convergence	a logical vector (of length equal to the number of methods) indicating whether the iterative purification process converged. Returned only if `purify` is TRUE.
save.output	the value of the `save.output` argument.
output	the value of the `output` argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7*, 131-177.

Aguerri, M.E., Galibert, M.S., Attorresi, H.F. and Maranon, P.P. (2009). Erroneous detection of nonuniform DIF using the Breslow-Day test in a short test. *Quality and Quantity, 43*, 35-44.

Angoff, W. H., and Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 2*, 95-106.

Dorans, N. J. and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*, 355-368.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Dirs.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R.D. (2003). Application of the Breslow-Day test of trend in odds ratio heterogeneity to the detection of nonuniform DIF. *Alberta Journal of Educational Research, 49*, 231-243.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*, 197-207.

Swaminathan, H. and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.

Thissen, D., Steinberg, L. and Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer and H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

See Also

difTID, difMH, difStd, difBD, difLogistic, difLord, difRaju, difLRT

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Comparing TID, Mantel-Haenszel, standardization and logistic regression
 # TID threshold 1.0
 # Standardization threshold 0.08
 # no continuity correction,
 # with item purification
 dichoDif(verbal, group=25, focal.name=1, method=c("TID","MH","Std",
   "Logistic"), correct=FALSE, thrSTD=0.08, thrTID=1, purify=TRUE)

 # Same analysis, but saving the output into the 'dicho' file
 dichoDif(verbal, group=25, focal.name=1, method=c("TID","MH","Std",
   "Logistic"), correct=FALSE, thrSTD=0.08, thrTID=1, purify=TRUE,
   save.output = TRUE, output = c("dicho","default"))

 # Comparing Lord and Raju results with 2PL model and
```

```
  # with item purification
 dichoDif(verbal, group=25, focal.name=1, method=c("Lord","Raju"),
   model="2PL", purify=TRUE)

 ## End(Not run)
```

---

difBD   *Breslow-Day DIF method*

Description

Performs DIF detection using Breslow-Day method.

Usage

```
difBD(Data, group, focal.name, BDstat="BD", alpha=0.05,
   purify=FALSE, nrIter=10, save.output=FALSE,
   output=c("out","default"))
## S3 method for class 'BD'
print(x, ...)
## S3 method for class 'BD'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
   save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within Data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
BDstat	character specifying the DIF statistic to be used. Possible values are "BD" (default) and "trend". See **Details**.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a BD class object.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).

save.plot   logical: should the plot be saved into a separate file? (default is FALSE).

save.options   character: a vector of three components. The first component is the name of the
                 output file, the second component is either the file path or "default" (default
                 value), and the third component is the file extension, either "pdf" (default) or
                 "jpeg". See **Details**.

...   other generic parameters for the plot or the print functions.

Details

The method of Breslow-Day (1980) allows for detecting non-uniform differential item functioning
without requiring an item response model approach.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addi-
tion, Data can hold the vector of group membership. If so, group indicates the column of Data
which corresponds to the group membership, either by specifying its name or by giving the column
number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA
values. They are discarded from sum-score computation.

The vector of group membership must hold only two different values, either as numeric or character.
The focal group is defined by the value of the argument focal.name.

Two test statistics are available: the usual Breslow-Day statistic for testing homogeneous associa-
tion (Aguerri, Galibert, Attorresi and Maranon, 2009) and the trend test statistic for assessing some
monotonic trend in the odds ratios (Penfield, 2003). The DIF statistic is supplied by the BDstat
argument, with values "BD" (default) for the usual statistic and "trend" for the trend test statistic.

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-
squared distribution with lower-tail probability of one minus alpha, and the degrees of freedom
depend on the DIF statistic. With the usual Breslow-Day statistic (BDstat=="BD"), it is the number
of partial tables taken into account (Aguerri *et al.*, 2009). With the trend test statistic, the degrees
of freedom are always equal to one (Penfield, 2003).

Item purification can be performed by setting purify to TRUE. Purification works as follows: if at
least one item was detected as functioning differently at the first step of the process, then the data
set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested
item (if necessary). The process stops when either two successive applications of the method yield
the same classifications of the items (Clauser and Mazor, 1998), or when nrIter iterations are run
without obtaining two successive identical classifications. In the latter case a warning message is
printed.

The output of the difBD, as displayed by the print.BD function, can be stored in a text file provided
that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case,
the name of the text file must be given as a character string into the first component of the output
argument (default name is "out"), and the path for saving the text file can be given through the
second component of output. The default value is "default", meaning that the file will be saved
in the current working directory. Any other path can be specified as a character string: see the
Examples section for an illustration.

The plot.BD function displays the DIF statistics in a plot, with each item on the X axis. The type
of point and the colour are fixed by the usual pch and col arguments. Option number permits to
display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or
JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the
components of save.options. The first two components perform similarly as those of the output
argument. The third component is the figure format, with allowed values "pdf" (default) for PDF
file and "jpeg" for JPEG file.

Value

A list of class "BD" with the following arguments:

BD	a matrix with one row per item and three columns: the first one contains the Breslow-Day statistic value, the second column indicates the degrees of freedom, and the last column displays the asymptotic *p*-values.
alpha	the significance level for DIF detection.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
BDstat	the value of the BDstat argument.
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIter of allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Aguerri, M.E., Galibert, M.S., Attorresi, H.F. and Maranon, P.P. (2009). Erroneous detection of nonuniform DIF using the Breslow-Day test in a short test. *Quality and Quantity, 43*, 35-44.

Breslow, N.E. and Day, N.E. (1980). *Statistical methods in cancer research, vol. I: The analysis of case-control studies*. Scientific Publication No 32. International Agency for Research on Cancer, Lyon.

Clauser, B.E. and Mazor, K.M. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice, 17*, 31-44.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R.D. (2003). Application of the Breslow-Day test of trend in odds ratio heterogeneity to the detection of nonuniform DIF. *Alberta Journal of Educational Research, 49*, 231-243.

See Also

breslowDay, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Three equivalent settings of the data matrix and the group membership
 difBD(verbal, group=25, focal.name=1)
 difBD(verbal, group="Gender", focal.name=1)
 difBD(verbal[,1:24], group=verbal[,25], focal.name=1)

 # With the BD trend test statistic
 difBD(verbal, group=25, focal.name=1, BDstat="trend")

 # With item purification
 difBD(verbal, group="Gender", focal.name=1, purify=TRUE)
 difBD(verbal, group="Gender", focal.name=1, purify=TRUE, nrIter=5)

 # Saving the output into the "BDresults.txt" file (and default path)
 r <- difBD(verbal, group=25, focal.name=1, save.output = TRUE,
            output = c("BDresults","default"))

 # Graphical devices
 plot(r)

 # Plotting results and saving it in a PDF figure
 plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

 # Changing the path, JPEG figure
 path <- "c:/Program Files/"
 plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difGenLogistic   *Generalized logistic regression DIF method*

Description

Performs DIF detection among multiple groups using generalized logistic regression method.

Usage

```
difGenLogistic(Data, group, focal.names, type="both",
   criterion="LRT", alpha=0.05, purify=FALSE, nrIter=10,
   save.output=FALSE, output=c("out","default"))
## S3 method for class 'genLogistic'
print(x, ...)
## S3 method for class 'genLogistic'
plot(x, plot="lrStat", item=1, itemFit="best",pch=8, number=TRUE,
   col="red", colIC=rep("black",length(x$focal.names)+1),
   ltyIC=1:(length(x$focal.names)+1), title=NULL, save.plot=FALSE,
   save.options=c("plot","default","pdf"), ref.name=NULL, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.names	numeric or character vector indicating the levels of group which correspond to the focal groups.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	character: the type of test statistic used to detect DIF items. Possible values are "LRT" (default) and "Wald". See **Details**.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a Logistik class object.
plot	character: the type of plot, either "lrStat" or "itemCurve". See **Details**.
item	numeric or character: either the number or the name of the item for which logistic curves are plotted. Use only when plot="itemCurve".

`itemFit`	character: the model to be selected for drawing the item curves. Possible values are `"best"` (default) for drawing from the best of the two models, and `"null"` for using fitted parameters of the null model M_0. Not used if `"plot"` is `"lrStat"`. See **Details**.
`pch, col`	type of usual `pch` and `col` graphical options.
`number`	logical: should the item number identification be printed (default is `TRUE`).
`colIC, ltyIC`	vectors of elements of the usual `col` and `lty` arguments for logistic curves. Used only when `plot="itemCurve"`.
`title`	either a character string with the title of the plot, or `NULL` (default), for which a specific title is automatically displayed.
`save.plot`	logical: should the plot be saved into a separate file? (default is `FALSE`).
`save.options`	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or `"default"` (default value), and the third component is the file extension, either `"pdf"` (default) or `"jpeg"`. See **Details**.
`ref.name`	either `NULL`(default) or a character string for the name of the reference group (to be used instead of "Reference" in the legend). Ignored if `plot` is `"lrStat"`.
`...`	other generic parameters for the `plot` or the `print` functions.

Details

The generalized logistic regression method (Magis, Raiche, Beland and Gerard, 2010) allows for detecting both uniform and non-uniform differential item functioning among multiple groups without requiring an item response model approach. It consists in fitting a logistic model with the test score, the group membership and an interaction between both as covariates. The statistical significance of the parameters related to group membership and the group-score interaction is then evaluated by means of the usual likelihood-ratio test. The argument `type` permits to test either both uniform and nonuniform effects simultaneously (`type="both"`), only uniform DIF effect (`type="udif"`) or only nonuniform DIF effect (`type="nudif"`). The identification of DIF items can be performed with either the Wald test or the likelihood ratio test, by setting the `criterion` argument to `"Wald"` or `"LRT"` respectively. See [genLogistik](#) for further details.

The `Data` is a matrix whose rows correspond to the subjects and columns to the items. In addition, `Data` can hold the vector of group membership. If so, `group` indicates the column of `Data` which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, `group` must be a vector of same length as `nrow(Data)`.

Missing values are allowed for item responses (not for group membership) but must be coded as `NA` values. They are discarded from the fitting of the logistic models (see [glm](#) for further details).

The vector of group membership must hold at least three values, either as numeric or character. The focal groups are defined by the values of the argument `focal.names`. If there is a unique focal group, then `difGenLogistic` returns the output of [difLogistic](#).

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus `alpha` and with *J* (if `type="udif"` or `type="nudif"`) or *2J* (if `type="both"`) degrees of freedom (*J* is the number of focal groups).

Item purification can be performed by setting `purify` to `TRUE`. Purification works as follows: if at least one item is detected as functioning differently at the first step of the process, then the data set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested item (if necessary). The process stops when either two successive applications of the method yield the same classifications of the items (Clauser and Mazor, 1998), or when `nrIter` iterations are run

without obtaining two successive identical classifications. In the latter case a warning message is printed.

The measures of effect size are provided by the difference ΔR^2 between the R^2 coefficients of the two nested models (Nagelkerke, 1991; Gomez-Benito, Dolores Hidalgo and Padilla, 2009). The effect sizes are classified as "negligible", "moderate" or "large". Two scales are available, one from Zumbo and Thomas (1997) and one from Jodoin and Gierl (2001). The output displays the ΔR^2 measures, together with the two classifications.

The output of the difGenLogistic, as displayed by the print.genLogistic function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

Two types of plots are available. The first one is obtained by setting plot="lrStat" and it is the default option. The likelihood ratio statistics are displayed on the Y axis, for each item. The detection threshold is displayed by a horizontal line, and items flagged as DIF are printed with the color defined by argument col. By default, items are spotted with their number identification (number=TRUE); otherwise they are simply drawn as dots whose form is given by the option pch.

The other type of plot is obtained by setting plot="itemCurve". In this case, the fitted logistic curves are displayed for one specific item set by the argument item. The latter argument can hold either the name of the item or its number identification. If the argument itemFit takes the value "best", the curves are drawn according to the output of the best model among M_0 and M_1. That is, two curves are drawn if the item is flagged as DIF, and only one if the item is flagged as non-DIF. If itemFit takes the value "null", then the two curves are drawn from the fitted parameters of the null model M_0. See genLogistik for further details on the models. The colors and types of traits for these curves are defined by means of the arguments colIC and ltyIC respectively. These are set as vectors of length $J + 1$, the first element for the reference group and the others for the focal groups. Finally, the ref.name argument permits to display the name if the reference group (instead of "Reference") in the legend.

Both types of plots can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "genLogistic" with the following arguments:

genLogistik	the values of the generalized logistic regression statistics.
logitPar	a matrix with one row per item and $2 + J * 2$ columns, holding the fitted parameters of the best model (among the two tested models) for each item.
parM0	the matrix of fitted parameters of the null model M_0, as returned by the Logistik command.
covMat	a 3-dimensional matrix of size p x p x K, where p is the number of estimated parameters and K is the number of items, holding the p x p covariance matrices of the estimated parameters (one matrix for each tested item).
deltaR2	the differences in Nagelkerke's R^2 coefficients. See **Details**.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.

DIFitems	either the column indicators for the items which were detected as DIF items, or "No DIF item detected".
type	the value of type argument.
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number of nrItem allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
focal.names	the value of focal.names argument.
criterion	the value of the criterion argument.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Gomez-Benito, J., Dolores Hidalgo, M. and Padilla, J.-L. (2009). Efficacy of effect size measures in logistic regression: an application for detecting DIF. *Methodology, 5*, 18-25.

Hidalgo, M. D. and Lopez-Pina, J.A. (2004). Differential item functioning detection and effect size: a comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement, 64*, 903-915.

Jodoin, M. G. and Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2010). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78,* 691-692.

Zumbo, B. D. and Thomas, D. R. (1997). A measure of effect size for a model-based approach for studying DIF. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science.

See Also

genLogistik, genDichoDif, subtestLogistic

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender ("Man" or "Woman") and
 # trait anger score ("Low" or "High")
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
 group[Anger<=20 & Gender==1]<-"ManLow"
 group[Anger>20 & Gender==1]<-"ManHigh"

 # New data set
 Verbal<-cbind(verbal[,1:24],group)

 # Reference group: "WomanLow"
 names<-c("WomanHigh","ManLow","ManHigh")

 # Testing both types of DIF effects
 # Three equivalent settings of the data matrix and the group membership
 r<-difGenLogistic(Verbal, group=25, focal.names=names)
 difGenLogistic(Verbal, group="group", focal.name=names)
 difGenLogistic(Verbal[,1:24], group=Verbal[,25], focal.names=names)

 # Using the Wald test
 difGenLogistic(Verbal, group=25, focal.names=names, criterion="Wald")

 # With item purification
 difGenLogistic(Verbal, group=25, focal.names=names, purify=TRUE)
 difGenLogistic(Verbal, group=25, focal.names=names, purify=TRUE,
   nrIter=5)

 # Testing for nonuniform DIF effect
 difGenLogistic(Verbal, group=25, focal.names=names, type="nudif")

 # Testing for uniform DIF effect
 difGenLogistic(Verbal, group=25, focal.names=names, type="udif")

 # Saving the output into the "GLresults.txt" file (and default path)
 r <- difGenLogistic(Verbal, group=25, focal.name=names,
                save.output = TRUE, output = c("GLresults","default"))

 # Graphical devices
```

```
plot(r)
plot(r, plot="itemCurve", item=1)
plot(r, plot="itemCurve", item=1, itemFit="best")
plot(r, plot="itemCurve", item=6)
plot(r, plot="itemCurve", item=6, itemFit="best")

# Plotting results and saving it in a PDF figure
plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

# Changing the path, JPEG figure
path <- "c:/Program Files/"
plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difGenLord   *Generalized Lord's chi-squared DIF method*

Description

Performs DIF detection among multiple groups using generalized Lord's chi-squared method.

Usage

```
difGenLord(Data, group, focal.names, model, c=NULL, engine="ltm",
    discr=1, irtParam=NULL, nrFocal=2, same.scale=TRUE,
    alpha=0.05, purify=FALSE, nrIter=10, save.output=FALSE,
    output=c("out","default"))
## S3 method for class 'GenLord'
print(x, ...)
## S3 method for class 'GenLord'
plot(x, plot = "lordStat", item = 1, pch = 8,
    number = TRUE, col = "red", colIC = rep("black",
    length(x$focal.names)+1), ltyIC = 1:(length(x$focal.names)
    + 1), save.plot=FALSE, save.options=c("plot","default","pdf"),
        ref.name=NULL, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within Data) of group membership. See **Details**.
focal.names	numeric or character vector indicating the levels of group which correspond to the focal groups.
model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL").
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.

engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
nrFocal	numeric: the number of focal groups (default is 2).
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is TRUE). See **Details**.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a GenLord class object.
plot	character: the type of plot, either "lordStat" or "itemCurve". See **Details**.
item	numeric or character: either the number or the name of the item for which ICC curves are plotted. Used only when plot="itemCurve".
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
colIC, ltyIC	vectors of elements of the usual col and lty arguments for ICC curves. Used only when plot="itemCurve".
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
ref.name	either NULL(default) or a character string for the name of the reference group (to be used instead of "Reference" in the legend). Ignored if plot is "lordStat".
...	other generic parameters for the plot or the print functions.

Details

The generalized Lord's chi-squared method (Kim, Cohen and Park, 1995), also referred to as *Qj* statistic, allows for detecting uniform or non-uniform differential item functioning among multiple groups by setting an appropriate item response model. The input can be of two kinds: either by displaying the full data, the group membership, the focal groups and the model, or by giving the item parameter estimates (with the option irtParam). Both can be supplied, but in this case only the parameters in irtParam are used for computing generalized Lord's chi-squared statistic.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded for item parameter estimation.

The vector of group membership must hold at least three different values, either as numeric or character. The focal groups are defined by the values of the argument focal.names.

If the model is not the 1PL model, or if engine is equal to "ltm", the selected IRT model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006). Otherwise, the 1PL model is fitted as a generalized linear mixed model, by means of the glmer function of the lme4 package (Bates and Maechler, 2009).

With the "1PL" model and the "ltm" engine, the common discrimination parameter is set equal to 1 by default. It is possible to fix another value through the argumentdiscr. Alternatively, this common discrimination parameter can be estimated (though not returned) by fixing discr to NULL.

The 3PL model can be fitted either unconstrained (by setting c to NULL) or by fixing the pseudo-guessing values. In the latter case, the argument c is either a numeric vector of same length of the number of items, with one value per item pseudo-guessing parameter, or a single value which is duplicated for all the items. If c is different from NULL then the 3PL model is always fitted (whatever the value of model).

The irtParam matrix has a number of rows equal to the number of groups (reference and focal ones) times the number of items J. The first J rows refer to the item parameter estimates in the reference group, while the next sets of J rows correspond to the same items in each of the focal groups. The number of columns depends on the selected IRT model: 2 for the 1PL model, 5 for the 2PL model, 6 for the constrained 3PL model and 9 for the unconstrained 3PL model. The columns of irtParam have to follow the same structure as the output of itemParEst command (the latter can actually be used to create the irtParam matrix). The number of focal groups has to be specified with argument nrFocal (default value is 2).

In addition to the matrix of parameter estimates, one has to specify whether items in the focal groups were rescaled to those of the reference group. If not, rescaling is performed by equal means anchoring (Cook and Eignor, 1991). Argument same.scale is used for this choice (default option is TRUE and assumes therefore that the parameters are already placed on a same scale).

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus alpha and p degrees of freedom. The value of p is the product of the number of focal groups by the number of item parameters to be tested (1 for the 1PL model, 2 for the 2PL model or the constrained 3PL model, and 3 for the unconstrained 3PL model).

Item purification can be performed by setting purify to TRUE. In this case, the purification occurs in the equal means anchoring process: items detected as DIF are iteratively removed from the set of items used for equal means anchoring, and the procedure is repeated until either the same items are identified twice as functioning differently, or when nrIter iterations have been performed. In the latter case a warning message is printed. See Candell and Drasgow (1988) for further details.

The output of the difGenLord, as displayed by the print.GenLord function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

Two types of plots are available. The first one is obtained by setting plot="lordStat" and it is the default option. The chi-squared statistics are displayed on the Y axis, for each item. The detection threshold is displayed by a horizontal line, and items flagged as DIF are printed with the color defined by argument col. By default, items are spotted with their number identification (number=TRUE); otherwise they are simply drawn as dots whose form is given by the option pch.

The other type of plot is obtained by setting plot="itemCurve". In this case, the fitted ICC curves are displayed for one specific item set by the argument item. The latter argument can hold either the name of the item or its number identification. The item parameters are extracted from the itemParFinal matrix if the output argument purification is TRUE, otherwise from the itemParInit matrix and after a rescaling of the item parameters using the [itemRescale](#) command. A legend is displayed in the upper left corner of the plot. The colors and types of traits for these curves are defined by means of the arguments colIC and ltyIC respectively. These are set as vectors of length 2, the first element for the reference group and the second for the focal group. Finally, the ref.name argument permits to display the name if the reference group (instead of "Reference") in the legend.

Both types of plots can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "GenLord" with the following arguments:

genLordChi	the values of the generalized Lord's chi-squared statistics.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
df	the degrees of freedom of the asymptotic null distribution of the statistics.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIterof allowed iterations. Returned only if purify is TRUE.
model	the value of model argument.
c	The value of the c argument.
engine	The value of the engine argument.
discr	the value of the discr argument.
itemParInit	the matrix of initial parameter estimates, with the same format as irtParam either provided by the user (through irtParam) or estimated from the data (and displayed after rescaling).
itemParFinal	the matrix of final parameter estimates, with the same format as irtParam, obtained after item purification. Returned only if purify is TRUE.
estPar	a logical value indicating whether the item parameters were estimated (TRUE) or provided by the user (FALSE).
names	the names of the items.

focal.names	the value of the focal.names argument.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Candell, G.L. and Drasgow, F. (1988). An iterative procedure for linking metrics and assessing item bias in item response theory. *Applied Psychological Measurement, 12*, 253-260.

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemParEst

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender ("Man" or "Woman") and trait
 # anger score ("Low" or "High")
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
```

```
group[Anger<=20 & Gender==1]<-"ManLow"
group[Anger>20 & Gender==1]<-"ManHigh"

# New data set
Verbal<-cbind(verbal[,1:24],group)

# Reference group: "WomanLow"
names<-c("WomanHigh","ManLow","ManHigh")

# Three equivalent settings of the data matrix and the group membership
# 1PL model, "ltm" engine
r <- difGenLord(Verbal, group=25, focal.names=names, model="1PL")
difGenLord(Verbal, group="group", focal.name=names, model="1PL")
difGenLord(Verbal[,1:24], group=Verbal[,25], focal.names=names,
model="1PL")

# 1PL model, "ltm" engine, estimated common discrimination
r <- difGenLord(Verbal, group=25, focal.names=names, discr=NULL)

# 1PL model, "lme4" engine
difGenLord(Verbal, group="group", focal.name=names, model="1PL",
engine="lme4")

# With item purification
difGenLord(Verbal, group=25, focal.names=names, model="1PL",
   purify=TRUE)

# Saving the output into the "GLresults.txt" file (and default path)
r <- difGenLord(Verbal, group=25, focal.names=names, model="1PL",
        save.output = TRUE, output = c("GLresults","default"))

# Splitting the data into the four subsets according to "group"
data0<-data1<-data2<-data3<-NULL
for (i in 1:nrow(verbal)){
 if (group[i]=="WomanLow") data0<-rbind(data0,as.numeric(verbal[i,1:24]))
 if (group[i]=="WomanHigh") data1<-rbind(data1,as.numeric(verbal[i,1:24]))
 if (group[i]=="ManLow") data2<-rbind(data2,as.numeric(verbal[i,1:24]))
 if (group[i]=="ManHigh") data3<-rbind(data3,as.numeric(verbal[i,1:24]))
 }

# Estimation of the item parameters (1PL model)
m0.1PL<-itemParEst(data0, model="1PL")
m1.1PL<-itemParEst(data1, model="1PL")
m2.1PL<-itemParEst(data2, model="1PL")
m3.1PL<-itemParEst(data3, model="1PL")

# Merging the item parameters WITHOUT rescaling
irt.noscale<-rbind(m0.1PL,m1.1PL,m2.1PL,m3.1PL)
rownames(irt.noscale)<-rep(colnames(verbal[,1:24]),4)

# Merging the item parameters WITH rescaling
irt.scale<-rbind(m0.1PL, itemRescale(m0.1PL,m1.1PL),
itemRescale(m0.1PL,m2.1PL) ,itemRescale(m0.1PL,m3.1PL))
rownames(irt.scale)<-rep(colnames(verbal[,1:24]),4)

# Equivalent calculations
difGenLord(irtParam=irt.noscale, nrFocal=3, same.scale=FALSE)
```

```
difGenLord(irtParam=irt.scale, nrFocal=3, same.scale=TRUE)

# With item purification
difGenLord(irtParam=irt.noscale, nrFocal=3, same.scale=FALSE,
  purify=TRUE)

# Graphical devices
plot(r)
plot(r, plot="itemCurve", item=1)
plot(r, plot="itemCurve", item=6)
plot(r, plot="itemCurve", item=6, ref.name="WomanHigh")

# Plotting results and saving it in a PDF figure
plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

# Changing the path, JPEG figure
path <- "c:/Program Files/"
plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difGMH   *Generalized Mantel-Haenszel DIF method*

Description

Performs DIF detection among multiple groups using the generalized Mantel-Haenszel method.

Usage

```
difGMH(Data, group, focal.names, alpha=0.05, purify=FALSE,
    nrIter=10, save.output=FALSE, output=c("out","default"))
## S3 method for class 'GMH'
print(x, ...)
## S3 method for class 'GMH'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
    save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within Data) of group membership. See **Details**.
focal.names	numeric or character vector indicating the levels of group which correspond to the focal groups.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).

save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a GMH class object.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
...	other generic parameters for the plot or the print functions.

Details

The generalized Mantel-Haenszel statistic (Somes, 1986) can be used to detect uniform differential item functioning among multiple groups, without requiring an item response model approach (Penfield, 2001).

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership must hold at least three value, either as numeric or character. The focal groups are defined by the values of the argument focal.names. If there is a unique focal group, then difGMH returns the output of [difMH](**without** continuity correction).

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus alpha and with as many degrees of freedom as the number of focal groups.

Item purification can be performed by setting purify to TRUE. Purification works as follows: if at least one item detected as functioning differently at the first step of the process, then the data set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested item (if necessary). The process stops when either two successive applications of the method yield the same classifications of the items (Clauser and Mazor, 1998), or when nrIter iterations are run without obtaining two successive identical classifications. In the latter case a warning message is printed.

The output of the difGMH, as displayed by the print.GMH function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

The plot.GMH function displays the DIF statistics in a plot, with each item on the X axis. The type of point and the colour are fixed by the usual pch and col arguments. Option number permits to display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the

components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "GMH" with the following arguments:

GMH	the values of the generalized Mantel-Haenszel statistics.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIter of allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
focal.names	the value of focal.names argument.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Clauser, B. E. and Mazor, K. M. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice, 17*, 31-44.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R. D. (2001). Assessing differential item functioning among multiple groups: a comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.

Somes, G. W. (1986). The generalized Mantel-Haenszel statistic. *The American Statistician, 40*, 106-108.

See Also

difGMH, difMH

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender ("Man" or "Woman") and
 # trait anger score ("Low" or "High")
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
 group[Anger<=20 & Gender==1]<-"ManLow"
 group[Anger>20 & Gender==1]<-"ManHigh"

 # New data set
 Verbal<-cbind(verbal[,1:24],group)

 # Reference group: "WomanLow"
 names<-c("WomanHigh","ManLow","ManHigh")

 # Three equivalent settings of the data matrix and the group membership
 difGMH(Verbal, group=25, focal.names=names)
 difGMH(Verbal, group="group", focal.name=names)
 difGMH(Verbal[,1:24], group=Verbal[,25], focal.names=names)

 # With item purification
 difGMH(Verbal, group=25, focal.names=names, purify=TRUE)
 difGMH(Verbal, group=25, focal.names=names, purify=TRUE, nrIter=5)

 # Saving the output into the "GMHresults.txt" file (and default path)
 r <- difGMH(Verbal, group=25, focal.name=names, save.output = TRUE,
            output = c("GMHresults","default"))

 # Graphical devices
 plot(r)

 # Plotting results and saving it in a PDF figure
 plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

 # Changing the path, JPEG figure
 path <- "c:/Program Files/"
 plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difLogistic   *Logistic regression DIF method*

Description

Performs DIF detection using logistic regression method.

Usage

```
difLogistic(Data, group, focal.name, type="both", criterion="LRT",
    alpha=0.05, purify=FALSE, nrIter=10,save.output=FALSE,
    output=c("out","default"))
## S3 method for class 'Logistic'
print(x, ...)
## S3 method for class 'Logistic'
plot(x, plot="lrStat", item=1, itemFit="best",
    pch=8, number=TRUE, col="red", colIC=rep("black",2),
    ltyIC=c(1,2),  save.plot=FALSE,
    save.options=c("plot","default","pdf"),
    group.names=NULL, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	a character string specifying which DIF statistic is computed. Possible values are "LRT" (default) or "Wald". See **Details**.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process. (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a Logistik class object.
plot	character: the type of plot, either "lrStat" (default) or "itemCurve". See **Details**.
item	numeric or character: either the number or the name of the item for which logistic curves are plotted. Used only when plot="itemCurve".

itemFit	character: the model to be selected for drawing the item curves. Possible values are "best" (default) for drawing from the best of the two models, and "null" for using fitted parameters of the null model M_0. Not used if "plot" is "lrStat". See **Details**.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
colIC, ltyIC	vectors of two elements of the usual col and lty arguments for logistic curves. Used only when plot="itemCurve".
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
group.names	either NULL (default) or a vector of two character strings giving the names of the reference group and the focal group (in this order) for display in the legend. Ignored if plot is "lrStat".
...	other generic parameters for the plot or the print functions.

Details

The logistic regression method (Swaminathan and Rogers, 1990) allows for detecting both uniform and non-uniform differential item functioning without requiring an item response model approach. It consists in fitting a logistic model with the test score, the group membership and an interaction between both as covariates. The statistical significance of the parameters related to group membership and the group-score interaction is then evaluated by means of either the likelihood-ratio test or the Wald test. The argument type permits to test either both uniform and nonuniform effects simultaneously (type="both"), only uniform DIF effect (type="udif") or only nonuniform DIF effect (type="nudif"). The argument criterion permits to select either the likelihood ratio test (criterion=="LRT") or the Wald test (criterion=="Wald"). See [Logistik](#) for further details.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from the fitting of the logistic models (see [glm](#) for further details).

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus alpha and with one (if type="udif" or type="nudif") or two (if type="both") degrees of freedom.

Item purification can be performed by setting purify to TRUE. Purification works as follows: if at least one item is detected as functioning differently at the first step of the process, then the data set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested item (if necessary). The process stops when either two successive applications of the method yield the same classifications of the items (Clauser and Mazor, 1998), or when nrIter iterations are run without obtaining two successive identical classifications. In the latter case a warning message is printed.

The measures of effect size are provided by the difference ΔR^2 between the R^2 coefficients of the two nested models (Nagelkerke, 1991; Gomez-Benito, Dolores Hidalgo and Padilla, 2009). The

effect sizes are classified as "negligible", "moderate" or "large". Two scales are available, one from Zumbo and Thomas (1997) and one from Jodoin and Gierl (2001). The output displays the ΔR^2 measures, together with the two classifications.

The output of the difLogistic, as displayed by the print.Logistic function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

Two types of plots are available. The first one is obtained by setting plot="lrStat" and it is the default option. The likelihood ratio statistics are displayed on the Y axis, for each item. The detection threshold is displayed by a horizontal line, and items flagged as DIF are printed with the color defined by argument col. By default, items are spotted with their number identification (number=TRUE); otherwise they are simply drawn as dots whose form is given by the option pch.

The other type of plot is obtained by setting plot="itemCurve". In this case, the fitted logistic curves are displayed for one specific item set by the argument item. The latter argument can hold either the name of the item or its number identification. If the argument itemFit takes the value "best", the curves are drawn according to the output of the best model among M_0 and M_1. That is, two curves are drawn if the item is flagged as DIF, and only one if the item is flagged as non-DIF. If itemFit takes the value "null", then the two curves are drawn from the fitted parameters of the null model M_0. See [Logistik](#) for further details on the models. The colors and types of traits for these curves are defined by means of the arguments colIC and ltyIC respectively. These are set as vectors of length 2, the first element for the reference group and the second for the focal group. Finally, the argument group.names permits to display the names of the reference and focal groups (instead of "Reference" and "Focal") in the legend.

Both types of plots can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "Logistic" with the following arguments:

Logistik	the values of the logistic regression statistics.
logitPar	a matrix with one row per item and four columns, holding the fitted parameters of the best model (among the two tested models) for each item.
parM0	the matrix of fitted parameters of the null model M_0, as returned by the [Logistik](#) command.
deltaR2	the differences in Nagelkerke's R^2 coefficients. See **Details**.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the column indicators for the items which were detected as DIF items, or "No DIF item detected".
type	the value of type argument.
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.

difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if `purify` is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number of `nrItem` allowed iterations. Returned only if `purify` is TRUE.
names	the names of the items.
criterion	the value of the `criterion` argument.
save.output	the value of the `save.output` argument.
output	the value of the `output` argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Clauser, B.E. and Mazor, K.M. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice, 17*, 31-44.

Finch, W.H. and French, B. (2007). Detection of crossing differential item functioning: a comparison of four methods. *Educational and Psychological Measurement, 67*, 565-582.

Gomez-Benito, J., Dolores Hidalgo, M. and Padilla, J.-L. (2009). Efficacy of effect size measures in logistic regression: an application for detecting DIF. *Methodology, 5*, 18-25.

Hidalgo, M. D. and Lopez-Pina, J.A. (2004). Differential item functioning detection and effect size: a comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement, 64*, 903-915.

Jodoin, M. G. and Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691-692.

Swaminathan, H. and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.

Zumbo, B.D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): logistic regression modelling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Zumbo, B. D. and Thomas, D. R. (1997). A measure of effect size for a model-based approach for studying DIF. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science.

See Also

Logistik, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Testing both DIF effects simultaneously
 # Three equivalent settings of the data matrix and the group membership
 r <- difLogistic(verbal, group=25, focal.name=1)
 difLogistic(verbal, group="Gender", focal.name=1)
 difLogistic(verbal[,1:24], group=verbal[,25], focal.name=1)

 # Testing both DIF effects with the Wald test
 r2 <- difLogistic(verbal, group=25, focal.name=1,
            criterion="Wald")

 # Testing nonuniform DIF effect
 difLogistic(verbal, group=25, focal.name=1, type="nudif")

 # Testing uniform DIF effect
 difLogistic(verbal, group=25, focal.name=1, type="udif")

 # With item purification
 difLogistic(verbal, group="Gender", focal.name=1, purify=TRUE)
 difLogistic(verbal, group="Gender", focal.name=1, purify=TRUE,
            nrIter=5)

 # Saving the output into the "Lresults.txt" file (and default path)
 r <- difLogistic(verbal, group=25, focal.name=1,
          save.output = TRUE, output = c("Lresults","default"))

 # Graphical devices
 plot(r)
 plot(r2)
 plot(r, plot="itemCurve", item=1)
 plot(r, plot="itemCurve", item=1, itemFit="null")
 plot(r, plot="itemCurve", item=6)
 plot(r, plot="itemCurve", item=6, itemFit="null")

 # Plotting results and saving it in a PDF figure
```

```
    plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

    # Changing the path, JPEG figure
    path <- "c:/Program Files/"
    plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difLord   *Lord's chi-squared DIF method*

Description

Performs DIF detection using Lord's chi-squared method.

Usage

```
difLord(Data, group, focal.name, model, c=NULL, engine="ltm",
    discr=1, irtParam=NULL, same.scale=TRUE, alpha=0.05,
    purify=FALSE, nrIter=10, save.output=FALSE,
    output=c("out","default"))
## S3 method for class 'Lord'
print(x, ...)
## S3 method for class 'Lord'
plot(x, plot = "lordStat", item = 1, pch = 8, number = TRUE,
    col = "red", colIC = rep("black", 2), ltyIC = c(1, 2),
    save.plot=FALSE, save.options=c("plot","default","pdf"),
    group.names=NULL, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL").
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.

alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a Lord class object.
plot	character: the type of plot, either "lordStat" or "itemCurve". See **Details**.
item	numeric or character: either the number or the name of the item for which ICC curves are plotted. Used only when plot="itemCurve".
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
colIC, ltyIC	vectors of two elements of the usual col and lty arguments for ICC curves. Used only when plot="itemCurve".
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
group.names	either NULL (default) or a vector of two character strings giving the names of the reference group and the focal group (in this order) for display in the legend. Ignored if plot is "lordStat".
...	other generic parameters for the plot or the print functions.

Details

Lord's chi-squared method (Lord, 1980) allows for detecting uniform or non-uniform differential item functioning by setting an appropriate item response model. The input can be of two kinds: either by displaying the full data, the group membership and the model, or by giving the item parameter estimates (through the option irtParam). Both can be supplied, but in this case only the parameters in irtParam are used for computing Lord's chi-squared statistic.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded for item parameter estimation.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

If the model is not the 1PL model, or if engine is equal to "ltm", the selected IRT model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006). Otherwise, the 1PL model is fitted as a generalized linear mixed model, by means of the glmer function of the lme4 package (Bates and Maechler, 2009).

With the "1PL" model and the "ltm" engine, the common discrimination parameter is set equal to 1 by default. It is possible to fix another value through the argument discr. Alternatively, this common discrimination parameter can be estimated (though not returned) by fixing discr to NULL.

The 3PL model can be fitted either unconstrained (by setting c to NULL) or by fixing the pseudo-guessing values. In the latter case, the argument c holds either a numeric vector of same length of the number of items, with one value per item pseudo-guessing parameter, or a single value which is duplicated for all the items. If c is different from NULL then the 3PL model is always fitted (whatever the value of model).

The irtParam matrix has a number of rows equal to twice the number of items in the data set. The first J rows refer to the item parameter estimates in the reference group, while the last J ones correspond to the same items in the focal group. The number of columns depends on the selected IRT model: 2 for the 1PL model, 5 for the 2PL model, 6 for the constrained 3PL model and 9 for the unconstrained 3PL model. The columns of irtParam have to follow the same structure as the output of itemParEst command (the latter can actually be used to create the irtParam matrix).

In addition to the matrix of parameter estimates, one has to specify whether items in the focal group were rescaled to those of the reference group. If not, rescaling is performed by equal means anchoring (Cook and Eignor, 1991). Argument same.scale is used for this choice (default option is TRUE and assumes therefore that the parameters are already placed on the same scale).

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus alpha and p degrees of freedom ($p=1$ for the 1PL model, $p=2$ for the 2PL model or the 3PL model with constrained pseudo-guessing parameters, and $p=3$ for the unconstrained 3PL model).

Item purification can be performed by setting purify to TRUE. In this case, the purification occurs in the equal means anchoring process. Items detected as DIF are iteratively removed from the set of items used for equal means anchoring, and the procedure is repeated until either the same items are identified twice as functioning differently, or when nrIter iterations have been performed. In the latter case a warning message is printed. See Candell and Drasgow (1988) for further details.

Under the 1PL model, the displayed output also proposes an effect size measure, which is -2.35 times the difference between item difficulties of the reference group and the focal group (Penfield and Camilli, 2007, p. 138). This effect size is similar Mantel-Haenszel's Δ_{MH} effect size, and the ETS delta scale is used to classify the effect sizes (Holland and Thayer, 1985).

The output of the difLord, as displayed by the print.Lord function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

Two types of plots are available. The first one is obtained by setting plot="lordStat" and it is the default option. The chi-squared statistics are displayed on the Y axis, for each item. The detection threshold is displayed by a horizontal line, and items flagged as DIF are printed with the color defined by argument col. By default, items are spotted with their number identification (number=TRUE); otherwise they are simply drawn as dots whose form is given by the option pch.

The other type of plot is obtained by setting plot="itemCurve". In this case, the fitted ICC curves are displayed for one specific item set by the argument item. The latter argument can hold either the name of the item or its number identification. The item parameters are extracted from the itemParFinal matrix if the output argument purification is TRUE, otherwise from the itemParInit matrix and after a rescaling of the item parameters using the itemRescale command. A legend is displayed in the upper left corner of the plot. The colors and types of traits for these

curves are defined by means of the arguments colIC and ltyIC respectively. These are set as vectors of length 2, the first element for the reference group and the second for the focal group. Finally, the argument group.names permits to display the names of the reference and focal groups (instead of "Reference" and "Focal") in the legend.

Both types of plots can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "Lord" with the following arguments:

LordChi	the values of the Lord's chi-squared statistics.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the i-th row refer to items which were classified respectively as non-DIF and DIF items at the (i-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIterof allowed iterations. Returned only if purify is TRUE.
model	the value of model argument.
c	The value of the c argument.
engine	The value of the engine argument.
discr	the value of the discr argument.
itemParInit	the matrix of initial parameter estimates,with the same format as irtParam either provided by the user (through irtParam) or estimated from the data (and displayed without rescaling).
itemParFinal	the matrix of final parameter estimates, with the same format as irtParam, obtained after item purification. Returned only if purify is TRUE.
estPar	a logical value indicating whether the item parameters were estimated (TRUE) or provided by the user (FALSE).
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Candell, G.L. and Drasgow, F. (1988). An iterative procedure for linking metrics and assessing item bias in item response theory. *Applied Psychological Measurement, 12*, 253-260.

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Holland, P. W. and Thayer, D. T. (1985). An alternative definition of the ETS delta scale of item difficulty. *Research Report RR-85-43*. Princeton, New-Jersey: Educational Testing Service.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R. D., and Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao and S. Sinharray (Eds.), *Handbook of Statistics 26: Psychometrics* (pp. 125-167). Amsterdam, The Netherlands: Elsevier.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemParEst, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]
```

```
# Three equivalent settings of the data matrix and the group membership
# (1PL model, "ltm" engine)
r <- difLord(verbal, group=25, focal.name=1, model="1PL")
difLord(verbal, group="Gender", focal.name=1, model="1PL")
difLord(verbal[,1:24], group=verbal[,25], focal.name=1, model="1PL")

# 1PL model, "lme4" engine
difLord(verbal, group=25, focal.name=1, model="1PL", engine="lme4")

# 2PL model
difLord(verbal, group="Gender", focal.name=1, model="2PL")

# 3PL model with all pseudo-guessing parameters constrained to 0.05
difLord(verbal, group="Gender", focal.name=1, model="3PL", c=0.05)

# Same models, with item purification
difLord(verbal, group=25, focal.name=1, model="1PL", purify=TRUE)
difLord(verbal, group="Gender", focal.name=1, model="2PL", purify=TRUE)
difLord(verbal, group="Gender", focal.name=1, model="3PL", c=0.05,
purify=TRUE)

# Saving the output into the "LordResults.txt" file (and default path)
r <- difLord(verbal, group=25, focal.name=1, model="1PL",
     save.output = TRUE, output = c("LordResults","default"))

# Splitting the data into reference and focal groups
nF<-sum(Gender)
nR<-nrow(verbal)-nF
data.ref<-verbal[,1:24][order(Gender),][1:nR,]
data.focal<-verbal[,1:24][order(Gender),][(nR+1):(nR+nF),]

## Pre-estimation of the item parameters (1PL model, "ltm" engine)
item.1PL<-rbind(itemParEst(data.ref, model="1PL"),
itemParEst(data.focal, model="1PL"))
difLord(irtParam=item.1PL, same.scale=FALSE)

## Pre-estimation of the item parameters (1PL model, "lme4" engine)
item.1PL<-rbind(itemParEst(data.ref, model="1PL", engine="lme4"),
itemParEst(data.focal, model="1PL", engine="lme4"))
difLord(irtParam=item.1PL, same.scale=FALSE)

## Pre-estimation of the item parameters (2PL model)
item.2PL<-rbind(itemParEst(data.ref, model="2PL"),
itemParEst(data.focal, model="2PL"))
difLord(irtParam=item.2PL, same.scale=FALSE)

## Pre-estimation of the item parameters (constrained 3PL model)
item.3PL<-rbind(itemParEst(data.ref, model="3PL", c=0.05),
itemParEst(data.focal, model="3PL", c=0.05))
difLord(irtParam=item.3PL, same.scale=FALSE)

# Graphical devices
plot(r)
plot(r, plot="itemCurve", item=1)
plot(r, plot="itemCurve", item=6)

# Plotting results and saving it in a PDF figure
```

```
plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

# Changing the path, JPEG figure
path <- "c:/Program Files/"
plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difLRT   *Likelihood-Ratio Test DIF method*

Description

Performs DIF detection using Likelihood Ratio Test (LRT) method.

Usage

```
difLRT(Data, group, focal.name, alpha=0.05, purify=FALSE,
    nrIter=10, save.output=FALSE, output=c("out","default"))
## S3 method for class 'LRT'
print(x, ...)
## S3 method for class 'LRT'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
    save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a LRT class object.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).

save.options   character: a vector of three components. The first component is the name of the
                 output file, the second component is either the file path or `"default"` (default
                 value), and the third component is the file extension, either `"pdf"` (default) or
                 `"jpeg"`. See **Details**.

...   other generic parameters for the `plot` or the `print` functions.

Details

The likelihood-ratio test method (Thissen, Steinberg and Wainer, 1988) allows for detecting uniform differential item functioning by fitting a closed-form Rasch model and by testing for extra interactions between group membership and item response. Currently only the Rasch model can be used, so only uniform DIF can be detected. Moreover, items are tested one by one and the other items act as anchor items.

The `Data` is a matrix whose rows correspond to the subjects and columns to the items. Missing values are allowed but must be coded as `NA` values. In addition, `Data` can hold the vector of group membership. If so, `group` indicates the column of `Data` which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, `group` must be a vector of same length as `nrow(Data)`.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument `focal.name`.

The function `glmer` from package `lme4` (Bates and Maechler, 2009) is used to fit the closed-form Rasch model. More precisely, the probability that response Y_{ijg} of subject i from group g (focal or reference) to item j is modeled as

$$logit(Pr(Y_{ijg} = 1) = \theta_{ig} + \gamma_g - \beta_j$$

where θ_i is subject's ability, β_j is the item difficulty and γ_g is the difference mean ability level between the focal and the reference groups. Subject abilities are treated as random effects, while item difficulties and γ_g are treated as fixed effects. Each item is tested by incorporating an interaction term, δ_{gj}, and by testing its statistical significance using the traditional likelihood-ratio test.

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the chi-squared distribution with lower-tail probability of one minus `alpha` and one degree of freedom.

Item purification can be performed by setting `purify` to `TRUE`. In this case, items detected as DIF are iteratively removed from the set of tested items, and the procedure is repeated (using the remaining items) until no additional item is identified as functioning differently. The process stops when either there is no new item detected as DIF, or when `nrIter` iterations are run and new DIF items are nevertheless detected. In the latter case a warning message is printed.

The output of the `difLRT`, as displayed by the `print.LRT` function, can be stored in a text file provided that `save.output` is set to `TRUE` (the default value `FALSE` does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the `output` argument (default name is `"out"`), and the path for saving the text file can be given through the second component of `output`. The default value is `"default"`, meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

The `plot.LRT` function displays the DIF statistics in a plot, with each item on the X axis. The type of point and the color are fixed by the usual `pch` and `col` arguments. Option `number` permits to display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing `save.plot` to `TRUE` allows this process. The figure is defined through the components of `save.options`. The first two components perform similarly as those of the `output` argument. The third component is the figure format, with allowed values `"pdf"` (default) for PDF file and `"jpeg"` for JPEG file.

Value

A list of class "LRT" with the following arguments:

LRT	the values of the likelihood-ratio statistics.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number of allowed iterations (10 by default). Returned only if purify is TRUE.
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Note

Because of the fitting of the modified Rasch model with glmer, the process can be very time consuming.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Thissen, D., Steinberg, L. and Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer and H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

See Also

 LRT, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Keeping the first 5 items and the first 50 subjects
 # (this is an artificial simplification to reduce the computational time)
 verbal<-verbal[1:50,c(1:5,25)]

 # Three equivalent settings of the data matrix and the group membership
 r <- difLRT(verbal, group=6, focal.name=1)
 difLRT(verbal, group="Gender", focal.name=1)
 difLRT(verbal[,1:5], group=verbal[,6], focal.name=1)

 # With item purification
 difLRT(verbal, group=6, focal.name=1, purify=TRUE)

 # Saving the output into the "LRTresults.txt" file (and default path)
 r <- difLRT(verbal, group=6, focal.name=1, save.output = TRUE,
             output = c("LRTresults","default"))

 # Graphical devices
 plot(r)

 # Plotting results and saving it in a PDF figure
 plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

 # Changing the path, JPEG figure
 path <- "c:/Program Files/"
 plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

 # WARNING: do not trust the results above since they are based on
 # a selected subset of the verbal data set!

## End(Not run)
```

---

difMH   *Mantel-Haenszel DIF method*

Description

Performs DIF detection using Mantel-Haenszel method.

Usage

```
difMH(Data, group, focal.name , MHstat="MHChisq", correct=TRUE, exact=FALSE,
    alpha=0.05, purify=FALSE, nrIter=10, save.output=FALSE,
    output=c("out","default"))
## S3 method for class 'MH'
print(x, ...)
## S3 method for class 'MH'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
    save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
MHstat	character: specifies the DIF statistic to be used for DIF identification. Possible values are "MHChisq" (default) and "logOR". See **Details** .
correct	logical: should the continuity correction be used? (default is TRUE)
exact	logical: should an exact test be computed? (default is FALSE).
alpha	numeric: significance level (default is 0.05).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a MH class object.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
...	other generic parameters for the plot or the print functions.

Details

The method of Mantel-Haenszel (1959) allows for detecting uniform differential item functioning without requiring an item response model approach.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data

which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

The DIF statistic is specified by the MHstat argument. By default, MHstat takes the value "MHChisq" and the Mantel-Haenszel chi-square statistic is used. The other optional value is "logOR", and the log odds-ratio statistic (that is, the log of alphaMH divided by the square root of varLambda) is used. See Penfield and Camilli (2007), Philips and Holland (1987) and mantelHaenszel help file.

By default, the asymptotic Mantel-Haenszel statistic is computed. However, the exact statistics and related P-values can be obtained by specifying the logical argument exact to TRUE. See Agresti (1990, 1992) for further details about exact inference.

The threshold (or cut-score) for classifying items as DIF depends on the DIF statistic. With the Mantel-Haenszel chi-squared statistic (MHstat=="MHChisq"), it is computed as the quantile of the chi-square distribution with lower-tail probability of one minus alpha and with one degree of freedom. With the log odds-ratio statistic (MHstat=="logOR"), it is computed as the quantile of the standard normal distribution with lower-tail probability of 1-alpha/2. WIth exact inference, it is simply the alpha level since exact P-values are returned.

By default, the continuity correction factor -0.5 is used (Holland and Thayer, 1988). One can nevertheless remove it by specifying correct=FALSE.

In addition, the Mantel-Haenszel estimates of the common odds ratios α_{MH} are used to measure the effect sizes of the items. These are obtained by $\Delta_{MH} = -2.35 \log \alpha_{MH}$ (Holland and Thayer, 1985). According to the ETS delta scale, the effect size of an item is classified as negligible if $|\Delta_{MH}| \leq 1$, moderate if $1 \leq |\Delta_{MH}| \leq 1.5$, and large if $|\Delta_{MH}| \geq 1.5$. The values of the effect sizes, together with the ETS classification, are printed with the output. Note that this is returned only for asymptotic tests, i.e. when exact is FALSE.

Item purification can be performed by setting purify to TRUE. Purification works as follows: if at least one item was detected as functioning differently at some step of the process, then the data set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested item (if necessary). The process stops when either two successive applications of the method yield the same classifications of the items (Clauser and Mazor, 1998), or when nrIter iterations are run without obtaining two successive identical classifications. In the latter case a warning message is printed.

The output of the difMH, as displayed by the print.MH function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

The plot.MH function displays the DIF statistics in a plot, with each item on the X axis. The type of point and the color are fixed by the usual pch and col arguments. Option number permits to display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file. Note that no plot is returned for exact inference.

Value

A list of class "MH" with the following arguments:

MH	the values of the Mantel-Haenszel DIF statistics (either exact or asymptotic).
alphaMH	the values of the mantel-Haenszel estimates of common odds ratios. Returned only if exact is FALSE.
varLambda	the values of the variances of the log odds-ratio statistics. Returned only if exact is FALSE.
Pval	the exact P-values of the exact MH test. Returned only if exact is TRUE.
MHstat	the value of the MHstat argument. Returned only if exact is FALSE.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection. Returned only if exact is FALSE.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
correct	the value of correct option.
exact	the value of exact option.
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIter of allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7*, 131-177.

Holland, P. W. and Thayer, D. T. (1985). An alternative definition of the ETS delta scale of item difficulty. *Research Report RR-85-43*. Princeton, NJ: Educational Testing Service.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Ed.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719-748.

Penfield, R. D., and Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao and S. Sinharray (Eds.), *Handbook of Statistics 26: Psychometrics* (pp. 125-167). Amsterdam, The Netherlands: Elsevier.

Philips, A., and Holland, P. W. (1987). Estimators of the Mantel-Haenszel log odds-ratio estimate. *Biometrics, 43*, 425-431.

Raju, N. S., Bode, R. K. and Larsen, V. S. (1989). An empirical assessment of the Mantel-Haenszel statistic to detect differential item functioning. *Applied Measurement in Education, 2*, 1-13.

Uttaro, T. and Millsap, R. E. (1994). Factors influencing the Mantel-Haenszel procedure in the detection of differential item functioning. *Applied Psychological Measurement, 18*, 15-25.

See Also

mantelHaenszel, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Excluding the "Anger" variable
 verbal <- verbal[colnames(verbal)!="Anger"]

 # Three equivalent settings of the data matrix and the group membership
 r <- difMH(verbal, group=25, focal.name=1)
 difMH(verbal, group="Gender", focal.name=1)
 difMH(verbal[,1:24], group=verbal[,25], focal.name=1)

 # With log odds-ratio statistic
 r2 <- difMH(verbal, group=25, focal.name=1, MHstat="logOR")

 # With exact inference
 difMH(verbal, group=25, focal.name=1, exact=TRUE)

 # With item purification
 difMH(verbal, group="Gender", focal.name=1, purify=TRUE)
```

```
    difMH(verbal, group="Gender", focal.name=1, purify=TRUE, nrIter=5)

    # Without continuity correction and with 0.01 significance level
    difMH(verbal, group="Gender", focal.name=1, alpha=0.01, correct=FALSE)

    # Saving the output into the "MHresults.txt" file (and default path)
    r <- difMH(verbal, group=25, focal.name=1, save.output = TRUE,
               output = c("MHresults","default"))

    # Graphical devices
    plot(r)
    plot(r2)

    # Plotting results and saving it in a PDF figure
    plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

    # Changing the path, JPEG figure
    path <- "c:/Program Files/"
    plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difRaju   *Raju's area DIF method*

Description

Performs DIF detection using Raju's area method.

Usage

```
difRaju(Data, group, focal.name, model, c=NULL, engine="ltm", discr=1,
    irtParam=NULL,  same.scale=TRUE, alpha=0.05, signed=FALSE, purify=FALSE,
    nrIter=10, save.output=FALSE, output=c("out","default"))
## S3 method for class 'Raj'
print(x, ...)
## S3 method for class 'Raj'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
    save.options=c("plot","default","pdf"), ...)
```

Arguments

Data
: numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.

group
: numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.

focal.name
: numeric or character indicating the level of group which corresponds to the focal group.

model
: character: the IRT model to be fitted (either "1PL", "2PL" or "3PL").

c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used only if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.
alpha	numeric: significance level (default is 0.05).
signed	logical: should the Raju's statistics be computed using the signed (TRUE) or unsigned (FALSE, default) area? See **Details**.
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a Raj class object.
pch, col	type of usual pch and col graphical options.
number	logical: should the item number identification be printed (default is TRUE).
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
...	other generic parameters for the plot or the print functions.

Details

Raju's area method (Raju, 1988, 1990) allows for detecting uniform or non-uniform differential item functioning by setting an appropriate item response model. The input can be of two kinds: either by displaying the full data, the group membership and the model, or by giving the item parameter estimates (with the option irtParam). Both can be supplied, but in this case only the parameters in irtParam are used for computing Raju's statistic.

By default, the Raju's *Z* statistics are obtained by using the *unsigned* areas between the ICCs. However, these statistics can also be computed using the *signed* areas, by setting the argument signed to TRUE (default value is FALSE). See RajuZ for further details.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded for item parameter estimation.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

If the model is not the 1PL model, or if engine is equal to "ltm", the selected IRT model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006). Otherwise, the 1PL model is fitted as a generalized linear mixed model, by means of the glmer function of the lme4 package (Bates and Maechler, 2009).

With the "1PL" model and the "ltm" engine, the common discrimination parameter is set equal to 1 by default. It is possible to fix another value through the argumentdiscr. Alternatively, this common discrimination parameter can be estimated (though not returned) by fixing discr to NULL.

The 3PL model can be fitted either unconstrained (by setting c to NULL) or by fixing the pseudo-guessing values. In the latter case, the argument c holds either a numeric vector of same length of the number of items, with one value per item pseudo-guessing parameter, or a single value which is duplicated for all the items. If c is different from NULL then the 3PL model is always fitted (whatever the value of model).

The irtParam matrix has a number of rows equal to twice the number of items in the data set. The first J rows refer to the item parameter estimates in the reference group, while the last J ones correspond to the same items in the focal group. The number of columns depends on the selected IRT model: 2 for the 1PL model, 5 for the 2PL model, 6 for the constrained 3PL model and 9 for the unconstrained 3PL model. The columns of irtParam have to follow the same structure as the output of itemParEst command (the latter can actually be used to create the irtParam matrix).

In addition to the matrix of parameter estimates, one has to specify whether items in the focal group were rescaled to those of the reference group. If not, rescaling is performed by equal means anchoring (Cook and Eignor, 1991). Argument same.scale is used for this choice (default option is TRUE and assumes therefore that the parameters are already placed on the same scale).

The threshold (or cut-score) for classifying items as DIF is computed as the quantile of the stanard normal distribution with lower-tail probability of 1-alpha/2.

Item purification can be performed by setting purify to TRUE. In this case, the purification occurs in the equal means anchoring process. Items detected as DIF are iteratively removed from the set of items used for equal means anchoring, and the procedure is repeated until either the same items are identified twice as functioning differently, or when nrIter iterations have been performed. In the latter case a warning message is printed. See Candell and Drasgow (1988) for further details.

Under the 1PL model, the displayed output also proposes an effect size measure, which is -2.35 times the difference between item difficulties of the reference group and the focal group (Penfield and Camilli, 2007, p. 138). This effect size is similar Mantel-Haenszel's Δ_{MH} effect size, and the ETS delta scale is used to classify the effect sizes (Holland and Thayer, 1985).

The output of the difRaju, as displayed by the print.Raj function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

The plot.Raj function displays the DIF statistics in a plot, with each item on the X axis. The type of point and the color are fixed by the usual pch and col arguments. Option number permits to display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "Raj" with the following arguments:

RajuZ	the values of the Raju's statistics.
alpha	the value of alpha argument.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
signed	the value of the signed argument.
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIterof allowed iterations. Returned only if purify is TRUE.
model	the value of model argument.
c	The value of the c argument.
engine	The value of the engine argument.
discr	the value of the discr argument.
itemParInit	the matrix of initial parameter estimates,with the same format as irtParam either provided by the user (through irtParam) or estimated from the data (and displayed without rescaling).
itemParFinal	the matrix of final parameter estimates, with the same format as irtParam, obtained after item purification. Returned only if purify is TRUE.
estPar	a logical value indicating whether the item parameters were estimated (TRUE) or provided by the user (FALSE).
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Candell, G.L. and Drasgow, F. (1988). An iterative procedure for linking metrics and assessing item bias in item response theory. *Applied Psychological Measurement, 12*, 253-260.

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Holland, P. W. and Thayer, D. T. (1985). An alternative definition of the ETS delta scale of item difficulty. *Research Report RR-85-43*. Princeton, NJ: Educational Testing Service.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R. D., and Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao and S. Sinharray (Eds.), *Handbook of Statistics 26: Psychometrics* (pp. 125-167). Amsterdam, The Netherlands: Elsevier.

Raju, N.S. (1988). The area between two item characteristic curves. *Psychometrika, 53*, 495-502.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*, 197-207.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

RajuZ, itemParEst, dichoDif

Examples

```
## Not run:

# Loading of the verbal data
data(verbal)
attach(verbal)

# Excluding the "Anger" variable
verbal<-verbal[colnames(verbal)!="Anger"]

# Three equivalent settings of the data matrix and the group membership
# (1PL model, "ltm" engine)
difRaju(verbal, group=25, focal.name=1, model="1PL")
difRaju(verbal, group="Gender", focal.name=1, model="1PL")
difRaju(verbal[,1:24], group=verbal[,25], focal.name=1, model="1PL")

# With signed areas
difRaju(verbal, group=25, focal.name=1, model="1PL", signed=TRUE)

# (1PL model, "lme4" engine)
difRaju(verbal, group="Gender", focal.name=1, model="1PL",
engine="lme4")

# 2PL model, signed and unsigned areas
difRaju(verbal, group="Gender", focal.name=1, model="2PL")
difRaju(verbal, group="Gender", focal.name=1, model="2PL", signed=TRUE)
```

```
# 3PL model with all pseudo-guessing parameters constrained to 0.05
# Signed and unsigned areas
difRaju(verbal, group="Gender", focal.name=1, model="3PL", c=0.05)
difRaju(verbal, group="Gender", focal.name=1, model="3PL", c=0.05,
  signed=TRUE)

# Same models, with item purification
difRaju(verbal, group="Gender", focal.name=1, model="1PL", purify=TRUE)
difRaju(verbal, group="Gender", focal.name=1, model="2PL", purify=TRUE)
difRaju(verbal, group="Gender", focal.name=1, model="3PL", c=0.05,
purify=TRUE)

# With signed areas
difRaju(verbal, group="Gender", focal.name=1, model="1PL", purify=TRUE,
  signed=TRUE)
difRaju(verbal, group="Gender", focal.name=1, model="2PL", purify=TRUE,
  signed=TRUE)
difRaju(verbal, group="Gender", focal.name=1, model="3PL", c=0.05,
purify=TRUE, signed=TRUE)

## Splitting the data into reference and focal groups
nF<-sum(Gender)
nR<-nrow(verbal)-nF
data.ref<-verbal[,1:24][order(Gender),][1:nR,]
data.focal<-verbal[,1:24][order(Gender),][(nR+1):(nR+nF),]

## Pre-estimation of the item parameters (1PL model, "ltm" engine)
item.1PL<-rbind(itemParEst(data.ref,model="1PL"),
itemParEst(data.focal,model="1PL"))
difRaju(irtParam=item.1PL,same.scale=FALSE)

## Pre-estimation of the item parameters (1PL model, "lme4" engine)
item.1PL<-rbind(itemParEst(data.ref, model="1PL", engine="lme4"),
itemParEst(data.focal, model="1PL", engine="lme4"))
difRaju(irtParam=item.1PL, same.scale=FALSE)

## Pre-estimation of the item parameters (2PL model)
item.2PL<-rbind(itemParEst(data.ref, model="2PL"),
itemParEst(data.focal, model="2PL"))
difRaju(irtParam=item.2PL, same.scale=FALSE)

## Pre-estimation of the item parameters (constrained 3PL model)
item.3PL<-rbind(itemParEst(data.ref, model="3PL", c=0.05),
itemParEst(data.focal, model="3PL", c=0.05))
difRaju(irtParam=item.3PL, same.scale=FALSE)

# Saving the output into the "RAJUresults.txt" file (and default path)
r <- difRaju(verbal, group=25, focal.name=1, model="1PL",
        save.output = TRUE, output = c("RAJUresults","default"))

# Graphical devices
plot(r)

# Plotting results and saving it in a PDF figure
plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))
```

```
 # Changing the path, JPEG figure
 path <- "c:/Program Files/"
 plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

 ## End(Not run)
```

---

difStd	*Standardization DIF method*

Description

Performs DIF detection using standardization method.

Usage

```
difStd(Data, group, focal.name, stdWeight="focal", thrSTD=0.1,
    purify=FALSE, nrIter=10, save.output=FALSE,
    output=c("out","default"))
## S3 method for class 'PDIF'
print(x, ...)
## S3 method for class 'PDIF'
plot(x, pch=8, number=TRUE, col="red", save.plot=FALSE,
    save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
stdWeight	character: the type of weights used for the standardized P-DIF statistic. Possible values are "focal" (default), "reference" and "total". See **Details**.
thrSTD	numeric: the threshold (cut-score) for standardized P-DIF statistic (default is 0.10).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	the result from a PDIF class object.
pch, col	type of usual pch and col graphical options.

number	logical: should the item number identification be printed (default is TRUE).
save.plot	logical: should the plot be saved into a separate file? (default is FALSE).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or "default" (default value), and the third component is the file extension, either "pdf" (default) or "jpeg". See **Details**.
...	other generic parameters for the plot or the print functions.

Details

The method of standardization (Dorans and Kulick, 1986) allows for detecting uniform differential item functioning without requiring an item response model approach.

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument focal.name.

The threshold (or cut-score) for classifying items as DIF has to be set by the user by the argument thrSTD. Default value is 0.10 but Dorans (1989) also recommends value 0.05. For this reason it is not possible to provide asymptotic *p*-values.

The weights for computing the standardized P-DIF statistics are defined through the argument stdWeight, with possible values "focal" (default value), "reference" and "total". See [stdPDIF](#) for further details.

In addition, two types of effect sizes are displayed. The first one is obtained from the standardized P-DIF statistic itself. According to Dorans, Schmitt and Bleistein (1992), the effect size of an item is classified as negligible if $|St - P - DIF| \leq 0.05$, moderate if $0.05 \leq |St - P - DIF| \leq 0.10$, and large if if $|St - P - DIF| \geq 0.10$. The second one is based on the transformation to the ETS Delta Scale (Holland and Thayer, 1985) of the standardized 'alpha' values (Dorans, 1989; Holland, 1985). The values of the effect sizes, together with the Dorans, Schmitt and Bleistein (DSB) and the ETS Delta scale (ETS) classification, are printed with the output.

Item purification can be performed by setting purify to TRUE. Purification works as follows: if at least one item was detected as functioning differently at some step of the process, then the data set of the next step consists in all items that are currently anchor (DIF free) items, plus the tested item (if necessary). The process stops when either two successive applications of the method yield the same classifications of the items (Clauser and Mazor, 1998), or when nrIter iterations are run without obtaining two successive identical classifications. In the latter case a warning message is printed.

The output of the difStd, as displayed by the print.PDIF function, can be stored in a text file provided that save.output is set to TRUE (the default value FALSE does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the output argument (default name is "out"), and the path for saving the text file can be given through the second component of output. The default value is "default", meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

The plot.PDIF function displays the DIF statistics in a plot, with each item on the X axis. The type of point and the color are fixed by the usual pch and col arguments. Option number permits

to display the item numbers instead. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "PDIF" with the following arguments:

PDIF	the values of the standardized P-DIF statistics.
stdAlpha	the values of the standardized alpha values (for effect sizes computation).
alpha	the value of alpha argument.
thr	the value of the thrSTD argument.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIter of allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
stdWeight	the value of the stdWeight argument.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Clauser, B.E. and Mazor, K.M. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice, 17*, 31-44.

Dorans, N. J. (1989). Two new approaches to assessing differential item functioning. Standardization and the Mantel-Haenszel method. *Applied Measurement in Education, 2*, 217-233.

Dorans, N. J. and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*, 355-368.

Dorans, N. J., Schmitt, A. P. and Bleistein, C. A. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement, 29*, 309-319.

Holland, P. W. (1985, October). *On the study of differential item performance without IRT*. Paper presented at the meeting of Military Testing Association, San Diego (CA).

Holland, P. W. and Thayer, D. T. (1985). An alternative definition of the ETS delta scale of item difficulty. *Research Report RR-85-43*. Princeton, NJ: Educational Testing Service.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

stdPDIF, dichoDif

Examples

```
## Not run:

# Loading of the verbal data
data(verbal)

# Excluding the "Anger" variable
verbal<-verbal[colnames(verbal)!="Anger"]

# Three equivalent settings of the data matrix and the group membership
difStd(verbal, group=25, focal.name=1)
difStd(verbal, group="Gender", focal.name=1)
difStd(verbal[,1:24], group=verbal[,25], focal.name=1)

# With other weights
difStd(verbal, group="Gender", focal.name=1, stdWeight="reference")
difStd(verbal, group="Gender", focal.name=1, stdWeight="total")

# With item purification
difStd(verbal, group="Gender", focal.name=1, purify=TRUE)
difStd(verbal, group="Gender", focal.name=1, purify=TRUE, nrIter=5)

# With detection threshold of 0.05
difStd(verbal, group="Gender", focal.name=1, thrSTD=0.05)

# Saving the output into the "STDresults.txt" file (and default path)
r <- difStd(verbal, group=25, focal.name=1, save.output = TRUE,
            output = c("STDresults","default"))

# Graphical devices
```

```
plot(r)

# Plotting results and saving it in a PDF figure
plot(r, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

# Changing the path, JPEG figure
path <- "c:/Program Files/"
plot(r, save.plot = TRUE, save.options = c("plot", path, "jpeg"))

## End(Not run)
```

---

difTID   *Transformed Item Difficulties (TID) DIF method*

Description

Performs DIF detection using Transformed Item Difficulties (TID) method.

Usage

```
difTID(Data, group, focal.name, props = NULL,
    thrTID = 1.5, purify = FALSE, nrIter = 10,
    save.output=FALSE, output=c("out","default"))
## S3 method for class 'TID'
print(x, ...)
## S3 method for class 'TID'
plot(x, plot="dist", pch=8, number=TRUE, col="red",
    save.plot=FALSE, save.options=c("plot","default","pdf"), ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
props	either NULL (default) or a two-column matrix with proportions of success in the reference group and the focal group. See **Details** .
thrTID	numeric: the threshold for detecting DIF items (default is 1.5).
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.

x	the result from a `TID` class object.
plot	character: either `"dist"` (default) to display the perpendicular distances, or `"delta"` for the Delta plot. See **Details**.
pch, col	type of usual `pch` and `col` graphical options.
number	logical: should the item number identification be printed (default is `TRUE`).
save.plot	logical: should the plot be saved into a separate file? (default is `FALSE`).
save.options	character: a vector of three components. The first component is the name of the output file, the second component is either the file path or `"default"` (default value), and the third component is the file extension, either `"pdf"` (default) or `"jpeg"`. See **Details**.
...	other generic parameters for the `plot` or the `print` functions.

Details

The Transformed Item Difficulties (TID) method, also known as Angoff's Delta method (Angoff, 1982; Angoff and Ford, 1973) allows for detecting uniform differential item functioning without requiring an item response model approach.

The `Data` is a matrix whose rows correspond to the subjects and columns to the items. In addition, `Data` can hold the vector of group membership. If so, `group` indicates the column of `Data` which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, `group` must be a vector of same length as `nrow(Data)`.

Missing values are allowed for item responses (not for group membership) but must be coded as `NA` values. They are discarded from the computation of proportions of success.

The vector of group membership must hold only two different values, either as numeric or character. The focal group is defined by the value of the argument `focal.name`.

Alternatively, one can provide the matrix of proportions of success in for each item in each group. This matrix must have the same format as that provided to the [trItemDiff](#) function; see the corresponding help file for further details.

The threshold (or cut-score) for classifying items as DIF must be supplied through the `thrTID` argument. The default value is 1.5, as being one of the most commonly used values (e.g. Facon and Nuchadee, 2010; Muniz, Hambleton, and Xing, 2001; Robin, Sirecci, and Hambleton, 2003). Other values can be specified instead.

Item purification can be performed by setting `purify` to `TRUE`. Purification works as follows: if at least one item was detected as functioning differently at some step of the process, then the intercept and slope parameters of the major axis are re-computed by discarding all items previously flagged as DIF. All perpendicular distances are then re-computed for all items. The process stops when either two successive applications of the method yield the same classifications of the items, or when `nrIter` iterations are run without obtaining two successive identical classifications. In the latter case a warning message is printed.

The output of the `difTID`, as displayed by the `print.TID` function, can be stored in a text file provided that `save.output` is set to `TRUE` (the default value `FALSE` does not execute the storage). In this case, the name of the text file must be given as a character string into the first component of the `output` argument (default name is `"out"`), and the path for saving the text file can be given through the second component of `output`. The default value is `"default"`, meaning that the file will be saved in the current working directory. Any other path can be specified as a character string: see the **Examples** section for an illustration.

Two types of plots are available through the `plot.TID` function. If the argument `plot` is set to `"dist"` (the default value), then the perpendicular distances are represented on the Y axis of a

scatter plot, with each item on the X axis. If plot is set to "delta", the Delta plot is returned, that is, the scatter plot of pairs of Delta scores for each item, with the reference group on the X axis and the focal group on the Y axis. The type of point and the color are fixed by the usual pch and col arguments. Option number permits to display the item numbers instead. Detection thresholds are also printed. Also, the plot can be stored in a figure file, either in PDF or JPEG format. Fixing save.plot to TRUE allows this process. The figure is defined through the components of save.options. The first two components perform similarly as those of the output argument. The third component is the figure format, with allowed values "pdf" (default) for PDF file and "jpeg" for JPEG file.

Value

A list of class "TID" with the following arguments:

Dj	the values of the perpendicular distances.
prop	the matrix of proportions of success.
delta	the matrix of Delta scores, in the same format as the prop matrix.
axisPar	a vector of length two with the intercept and slope parameters of the major axis of Delta points.
thr	the threshold (cut-score) for DIF detection.
DIFitems	either the column indicators of the items which were detected as DIF items, or "No DIF item detected".
purification	the value of purify option.
nrPur	the number of iterations in the item purification process. Returned only if purify is TRUE.
difPur	a binary matrix with one row per iteration in the item purification process and one column per item. Zeros and ones in the *i*-th row refer to items which were classified respectively as non-DIF and DIF items at the (*i*-1)-th step. The first row corresponds to the initial classification of the items. Returned only if purify is TRUE.
convergence	logical indicating whether the iterative item purification process stopped before the maximal number nrIter of allowed iterations. Returned only if purify is TRUE.
names	the names of the items.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Department of Mathematics
University of Liege, Belgium
<david.magis@ulg.ac.be>

References

Angoff, W. H. (1982). Use of difficulty and discrimination indices for detecting item bias. In R. A. Berck (Ed.), *Handbook of methods for detecting item bias* (pp. 96-116). Baltimore, MD: Johns Hopkins University Press.

Angoff, W. H., and Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 2*, 95-106.

Facon, B., and Nuchadee, M.-L. (2010). An item analysis of Raven's Colored Progressive Matrices among participants with Down syndrome. *Research in Deevelopmental Disabilities, 31*, 243-249.

Muniz, J., Hambleton, R. K., and Xing, D. (2001). Small sample studies to detect flaws in item translations. *International Journal of Testing, 1*, 115-135.

Robin, F., Sirecci, S. G., and Hambleton, R. K. (2003). Evaluating the equivalence of different language versions of a credentialing exam. *International Journal of Testing, 3*, 1-20.

See Also

trItemDiff, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Excluding the "Anger" variable
 verbal <- verbal[colnames(verbal)!="Anger"]

 # Three equivalent settings of the data matrix and the group membership
 r <- difTID(verbal, group=25, focal.name=1)
 difTID(verbal, group="Gender", focal.name=1)
 difTID(verbal[,1:24], group=verbal[,25], focal.name=1)

 # With item purification and threshold 1
 r2 <- difTID(verbal, group="Gender", focal.name=1, purify=TRUE, thrTID=1)

 # Saving the output into the "TIDresults.txt" file (and default path)
 difTID(verbal, group=25, focal.name=1, save.output = TRUE,
   output = c("TIDresults","default"))

 # Providing the proportions of success only
 props <- cbind(colMeans(verbal[verbal[,25]==0,1:24]),
   colMeans(verbal[verbal[,25]==1,1:24]) )
 difTID(prop=props)

 # Graphical devices
 plot(r2)
 plot(r2, plot="delta")

 # Plotting results and saving it in a PDF figure
 plot(r2, save.plot = TRUE, save.options = c("plot", "default", "pdf"))

 # Changing the path, JPEG figure
 path <- "c:/Program Files/"
 plot(r2, save.plot = TRUE, save.options = c("plot", path, "jpeg"))
```

```
## End(Not run)
```

---

genDichoDif   *Comparison of DIF detection methods among multiple groups*

Description

This function compares the specified DIF detection methods among multiple groups, with respect to the detected items.

Usage

```
genDichoDif(Data, group, focal.names, method, type="both",
    criterion="LRT", alpha=0.05, model="2PL", c=NULL,
    engine = "ltm", discr=1, irtParam=NULL, nrFocal=2,
    same.scale=TRUE, purify=FALSE, nrIter=10,
    save.output=FALSE, output=c("out","default"))
## S3 method for class 'genDichoDif'
print(x, ...)
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.names	numeric or character vector indicating the levels of group which correspond to the focal groups.
method	character: the name of the selected methods. See **Details**.
type	a character string specifying which DIF effects must be tested (default is "both"). See **Details**.
criterion	character: the type of test statistic used to detect DIF items with generalized logistic regression. Possible values are "LRT" (default) and "Wald". See **Details**.
alpha	numeric: significance level (default is 0.05).
model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL"). Default is "2PL".
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
nrFocal	numeric: the number of focal groups (default is 2).

same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.
x	result from a genDichoDif class object.
...	other generic parameters for the print function.

Details

genDichoDif is a generic function which calls one or several DIF detection methods among multiple groups, and summarize their output. The possible methods are: "GMH" for Generalized Mantel-Haenszel (Penfield, 2001), "genLogistic" for generalized logistic regression (Magis, Raiche Beland and Gerard, 2010) and "genLord" for generalized Lord's chi-square test (Kim, Cohen and Park, 1995).

If method has a single component, the output of genDichoDif is exactly the one provided by the method itself. Otherwise, the main output is a matrix with one row per item and one column per method. For each specified method and related arguments, items detected as DIF and non-DIF are respectively encoded as "DIF" and "NoDIF". When printing the output an additional column is added, counting the number of times each item was detected as functioning differently (Note: this is just an informative summary, since the methods are obviously not independent for the detection of DIF items).

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from either the computation of the sum-scores, the fitting of the logistic models or the IRT models (according to the method).

The vector of group membership must hold at least three different values, either as numeric or character. The focal groups are defined by the values of the argument focal.names.

For the generalized logistic regression method, the argument type permits to test either both uniform and nonuniform effects simultaneously (with type="both"), only uniform DIF effect (with type="udif") or only nonuniform DIF effect (with type="nudif"). Furthermore, the argument criterion defines which test must be used, either the Wald test ("Wald") or the likelihood ratio test ("LRT").

For generalized Lord method, one can specify either the IRT model to be fitted (by means of model, c, engine and discr arguments), or the item parameter estimates with arguments irtParam and same.scale. See [difGenLord](#) for further details.

The threshold for detecting DIF items depends on the method and is depending on the significance level set by alpha.

Item purification can be requested by specifying purify option to TRUE. Recall that item purification process is slightly different for IRT and for non-IRT based methods. See the corresponding methods for further information.

The output of the genDichoDif function can be stored in a text file by fixing save.output and output appropriately. See the help file of selectGenDif function (or any other DIF method) for further information.

Value

Either the output of one of the DIF detection methods, or a list of class "genDichoDif" with the following arguments:

DIF	a character matrix with one row per item and whose columns refer to the different specified detection methods. See **Details**.
alpha	the significance level alpha.
method	the value of methodargument.
type	the value of type argument.
criterion	the value of the criterion argument.
model	the value of model option.
c	the value of c option.
engine	The value of the engine argument.
discr	the value of the discr argument.
irtParam	the value of irtParam option.
same.scale	the value of same.scale option.
purification	the value of purify option.
nrPur	an integer vector (of length equal to the number of methods) with the number of iterations in the purification process. Returned only if purify is TRUE.
convergence	a logical vector (of length equal to the number of methods) indicating whether the iterative purification process converged. Returned only if purify is TRUE.
save.output	the value of the save.output argument.
output	the value of the output argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2010). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

Penfield, R. D. (2001). Assessing differential item functioning among multiple groups: a comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.

See Also

difGMH, difGenLogistic, difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender ("Man" or "Woman") and trait
 # anger score ("Low" or "High")
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
 group[Anger<=20 & Gender==1]<-"ManLow"
 group[Anger>20 & Gender==1]<-"ManHigh"

 # New data set
 Verbal<-cbind(verbal[,1:24],group)

 # Reference group: "WomanLow"
 names<-c("WomanHigh","ManLow","ManHigh")

 # Comparing the three available methods
 # with item purification
 genDichoDif(Verbal, group=25, focal.names=names, method=c("GMH",
  "genLogistic","genLord"), purify=TRUE, save.output = TRUE,
   output = c("dicho","default"))

 # Same analysis, but saving the output into the 'genDicho' file
 genDichoDif(Verbal, group=25, focal.names=names, method=c("GMH",
  "genLogistic","genLord"), purify=TRUE)

## End(Not run)
```

---

genLogistik   *Generalized logistic regression DIF statistic*

Description

Calculates the "generalized logistic regression" likelihood-ratio or Wald statistics for DIF detection among multiple groups.

Usage

```
genLogistik(data, member, anchor=1:ncol(data), type="both",
   criterion="LRT")
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and positive integer entries only. See **Details**.
anchor	a vector of integer values specifying which items (all by default) are currently considered as anchor (DIF free) items. See **Details**.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	character: the type of test statistic used to detect DIF items. Possible values are "LRT" (default) and "Wald". See **Details**.

Details

This command computes the generalized logistic regression statistic (Magis, Raiche, Beland and Gerard, 2010) in the specific framework of differential item functioning among $(J + 1)$ groups and J is the number of focal groups. It forms the basic command of difGenLogistic and is specifically designed for this call.

The three possible models to be fitted are:

$$M_0 : logit(\pi_i) = \alpha + \beta X + \gamma_i + \delta_i X$$

$$M_1 : logit(\pi_i) = \alpha + \beta X + \gamma_i$$

$$M_2 : logit(\pi_i) = \alpha + \beta X$$

where π_i is the probability of answering correctly the item in group i ($i = 0, ..., J$) and X is the sum score. Parameters α and β are the common intercept and the slope of the logistic curves, while γ_i and δ_i are group-specific parameters. For identification reasons the parameters γ_0 and δ_0 of the reference group are set to zero. The set of parameters $\{\gamma_i : i = 1, ..., J\}$ of the focal groups ($g = i$) represents the uniform DIF effect across all groups, and the set of parameters $\{\delta_i : i = 1, ..., n\}$ is used to model nonuniform DIF effect across all groups. The models are fitted with the [glm](glm) function.

Two tests are available: the Wald test and the likelihood ratio test. With the likelihood ratio test, two nested models are fitted and compared by means of Wilks' Lambda (or likelihood ratio) statistic (Wilks, 1938). With the Wald test, the model parameters are statistically tested using an appropriate contrast matrix. Each test is set with the criterion argument, with the values "LRT" and "Wald" respectively.

The argument type determines the type of DIF effect to be tested. The three possible values of type are: type="both" which tests the hypothesis $H_0 : \gamma_i = \delta_i = 0$ for all i; type="nudif" which tests the hypothesis $H_0 : \delta_i = 0$ for all i; and type="udif" which tests the hypothesis $H_0 : \gamma_i = 0 | \delta_i = 0$ for all i. In other words, type="both" tests for DIF (without distinction

between uniform and nonuniform effects), while type="udif" and type="nudif" test for uniform and nonuniform DIF, respectively. Whatever the tested DIF effects, this is a simultaneous test of the equality of focal group parameters to zero.

The data are passed through the data argument, with one row per subject and one column per item. Missing values are allowed but must be coded as NA values. They are discarded from the fitting of the logistic models (see glm for further details).

The vector of group membership, specified with member argument, must hold only zeros and positive integers. The value zero corresponds to the reference group, and each positive integer value corresponds to one focal group. At least two different positive integers must be supplied.

Option anchor sets the items which are considered as anchor items for computing the logistic regression DIF statistics. Items other than the anchor items and the tested item are discarded. anchor must hold integer values specifying the column numbers of the corresponding anchor items. It is mainly designed to perform item purification.

In addition to the results of the fitted models (model parameters, covariance matrices, test statistics), Nagelkerke's R^2 coefficients (Nagelkerke, 1991) are computed for each model and the output returns the differences in these coefficients. Such differences are used as measures of effect size by the difGenLogistic command; see Gomez-Benito, Dolores Hidalgo and Padilla (2009), Jodoin and Gierl (2001) and Zumbo and Thomas (1997).

Value

A list with four components:

lrt	the values of the generalized logistic regression DIF statistics (that is, the likelihood ratio test statistics).
deltaR2	the differences between Nagelkerke's R^2 coefficients of the tested models. See **Details**.
parM0	a matrix with one row per item and $2 + J * 2$ columns (where J is the number of focal groups), holding successively the fitted parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}_i$ and $\hat{\delta}_i$ ($i = 1, ..., J$) of the "full" model (M_0 if type="both" or type="nudif", M_1 if type="udif").
parM1	the same matrix as parM0 but with fitted parameters for the "simpler" model (M_1 if type="nudif", M_2 if type="both" or type="udif").
covMat	a 3-dimensional matrix of size p x p x K, where p is the number of estimated parameters and K is the number of items, holding the p x p covariance matrices of the estimated parameters (one matrix for each tested item).
criterion	the value of the criterion argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)

Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Gomez-Benito, J., Dolores Hidalgo, M. and Padilla, J.-L. (2009). Efficacy of effect size measures in logistic regression: an application for detecting DIF. *Methodology, 5*, 18-25.

Jodoin, M. G. and Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2010). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691-692.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics, 9*, 60-62.

Zumbo, B. D. and Thomas, D. R. (1997). A measure of effect size for a model-based approach for studying DIF. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science.

See Also

difGenLogistic, genDichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender (0 or 1) and trait anger score
 # ("Low" or "High")
 # Reference group: women with low trait anger score (<=20)
 group<-rep(0,nrow(verbal))
 group[Anger>20 & Gender==0]<-1
 group[Anger<=20 & Gender==1]<-2
 group[Anger>20 & Gender==1]<-3

 # Testing both types of DIF simultaneously
 # With all items
 genLogistik(verbal[,1:24], group)
 genLogistik(verbal[,1:24], group, criterion="Wald")

 # Removing item 6 from the set of anchor items
 genLogistik(verbal[,1:24], group, anchor=c(1:5,7:24))
 genLogistik(verbal[,1:24], group, anchor=c(1:5,7:24), criterion="Wald")

 # Testing nonuniform DIF effect
 genLogistik(verbal[,1:24], group, type="nudif")
 genLogistik(verbal[,1:24], group, type="nudif", criterion="Wald")
```

```
 # Testing uniform DIF effect
 genLogistik(verbal[,1:24], group, type="udif")
 genLogistik(verbal[,1:24], group, type="udif", criterion="Wald")

## End(Not run)
```

---

genLordChi2 *Generalized Lord's chi-squared DIF statistic*

Description

Calculates the generalized Lord's chi-squared statistics for DIF detection among multiple groups.

Usage

```
genLordChi2(irtParam, nrFocal)
```

Arguments

irtParam numeric: the matrix of item parameter estimates. See **Details**.

nrFocal numeric: the number of focal groups.

Details

This command computes the generalized Lord's chi-squared statistic (Kim, Cohen and Park, 1995), also called the *Qj* statistics, in the specific framework of differential item functioning with multiple groups. It forms the basic command of [difGenLord](#) and is specifically designed for this call.

The irtParam matrix has a number of rows equal to the number of groups (reference and focal ones) times the number of items *J*. The first *J* rows refer to the item parameter estimates in the reference group, while the next sets of *J* rows correspond to the same items in each of the focal groups. The number of columns depends on the selected IRT model: 2 for the 1PL model, 5 for the 2PL model, 6 for the constrained 3PL model and 9 for the unconstrained 3PL model. The columns of irtParam have to follow the same structure as the output of itemParEst command (the latter can actually be used to create the irtParam matrix).

In addition, the item parameters of the reference group and the focal groups must be placed on the same scale. This can be done by using [itemRescale](#) command, which performs equal means anchoring between two groups of item estimates (Cook and Eignor, 1991).

The number of focal groups has to be specified with argument nrFocal.

Value

A vector with the values of the generalized Lord's chi-squared DIF statistics.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

itemParEst, itemRescale, difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender ("Man" or "Woman") and
 # trait anger score ("Low" or "High")
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
 group[Anger<=20 & Gender==1]<-"ManLow"
 group[Anger>20 & Gender==1]<-"ManHigh"

 # Splitting the data into the four subsets according to "group"
 data0<-data1<-data2<-data3<-NULL
 for (i in 1:nrow(verbal)){
 if (group[i]=="WomanLow") data0<-rbind(data0,as.numeric(verbal[i,1:24]))
 if (group[i]=="WomanHigh") data1<-rbind(data1,as.numeric(verbal[i,1:24]))
 if (group[i]=="ManLow") data2<-rbind(data2,as.numeric(verbal[i,1:24]))
 if (group[i]=="ManHigh") data3<-rbind(data3,as.numeric(verbal[i,1:24]))
 }

 # Estimation of the item parameters (1PL model)
```

```
m0.1PL<-itemParEst(data0, model="1PL")
m1.1PL<-itemParEst(data1, model="1PL")
m2.1PL<-itemParEst(data2, model="1PL")
m3.1PL<-itemParEst(data3, model="1PL")

# merging the item parameters with rescaling
irt.scale<-rbind(m0.1PL,itemRescale(m0.1PL,m1.1PL),
itemRescale(m0.1PL,m2.1PL),itemRescale(m0.1PL,m3.1PL))

# Generalized Lord's chi-squared statistics
genLordChi2(irt.scale, nrFocal=3)

## End(Not run)
```

---

genMantelHaenszel   *Generalized Mantel-Haenszel DIF statistic*

Description

Calculates the generalized Mantel-Haenszel statistics for DIF detection among multiple groups.

Usage

```
genMantelHaenszel(data, member, anchor=1:ncol(data))
```

Arguments

data   numeric: the data matrix (one row per subject, one column per item).

member   numeric: the vector of group membership with zero and positive integer entries
                only. See **Details**.

anchor   a vector of integer values specifying which items (all by default) are currently
                considered as anchor (DIF free) items. See **Details**.

Details

This command computes the generalized Mantel-Haenszel statistic (Somes, 1986) in the specific
framework of differential item functioning. It forms the basic command of difGMH and is specifi-
cally designed for this call.

The data are passed through the data argument, with one row per subject and one column per item.
Missing values are allowed but must be coded as NA values. They are discarded from sum-score
computation.

The vector of group membership, specified with member argument, must hold only zeros and pos-
itive integers. The value zero corresponds to the reference group, and each positive integer value
corresponds to one focal group. At least two different positive integers must be supplied.

Option anchor sets the items which are considered as anchor items for computing generalized
Mantel-Haenszel statistics. Items other than the anchor items and the tested item are discarded.
anchor must hold integer values specifying the column numbers of the corresponding anchor items.
It is primarily designed to perform item purification.

Value

A vector with the values of the generalized Mantel-Haenszel DIF statistics.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R. D. (2001). Assessing differential item functioning among multiple groups: a comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.

Somes, G. W. (1986). The generalized Mantel-Haenszel statistic. *The American Statistician, 40*, 106-108.

See Also

difGMH

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender (0 or 1) and trait anger
 # score ("Low" or "High")
 # Reference group: women with low trait anger score (<=20)
 group<-rep(0,nrow(verbal))
 group[Anger>20 & Gender==0]<-1
 group[Anger<=20 & Gender==1]<-2
 group[Anger>20 & Gender==1]<-3

 # Without continuity correction
 genMantelHaenszel(verbal[,1:24], group)

 # Removing item 6 from the set of anchor items
 genMantelHaenszel(verbal[,1:24], group, anchor=c(1:5,7:24))
```

```
## End(Not run)
```

---

itemPar1PL   *Item parameter estimation for DIF detection using Rasch (1PL) model*

Description

Fits the Rasch (1PL) model and returns related item parameter estimates.

Usage

```
itemPar1PL(data, engine="ltm", discr=1)
```

Arguments

data   numeric: the data matrix.

engine   character: the engine for estimating the 1PL model, either "ltm" (default) or
                    "lme4".

discr   either NULL or a real positive value for the common discrimination parameter
                    (default is 1). Not used if engine is "lme4". See **Details**.

Details

itemPar1PL permits to get item parameter estimates from the Rasch or 1PL model. The output
is ordered such that it can be directly used with the general [itemParEst](#) command, as well as the
methods of Lord ([difLord](#)) and Raju ([difRaju](#)) and Generalized Lord's ([difGenLord](#)) to detect
differential item functioning.

The data is a matrix whose rows correspond to the subjects and columns to the items.

Missing values are allowed but must be coded as NA values. They are discarded for item parameter
estimation.

The estimation engine is set by the engine argument. By default (engine="ltm"), the Rasch model
is fitted using marginal maximum likelihood, by means of the function rasch from the ltm package
(Rizopoulos, 2006). The other option, engine="lme4", permits to fit the Rasch model as a general-
ized linear mixed model, by means of the glmer function of the lme4 package (Bates and Maechler,
2009).

With the "ltm" engine, the common discrimination parameter is set equal to 1 by default. It is
possible to fix another value through the argumentdiscr. Alternatively, this common discrimination
parameter can be estimated (though not returned) by fixing discr to NULL. See the functionalities
of [rasch](#) command for further details.

Value

A matrix with one row per item and two columns, the first one with item parameter estimates and
the second one with the related standard errors.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemPar2PL, itemPar3PL, itemPar3PLconst, itemParEst, difLord, difRaju,

difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Getting item parameter estimates ('ltm' engine)
 itemPar1PL(verbal[,1:24])

 # Estimating the common discrimination parameter instead
 itemPar1PL(verbal[,1:24], discr=NULL)

 # Getting item parameter estimates ('lme4' engine)
 itemPar1PL(verbal[,1:24], engine="lme4")

## End(Not run)
```

itemPar2PL | *Item parameter estimation for DIF detection using 2PL model*

Description

Fits the 2PL model and returns related item parameter estimates.

Usage

```
itemPar2PL(data)
```

Arguments

data | numeric: the data matrix.

Details

itemPar2PL permits to get item parameter estimates from the 2PL model. The output is ordered such that it can be directly used with the general itemParEst command, as well as the methods of Lord (difLord) and Raju (difRaju) and Generalized Lord's (difGenLord) to detect differential item functioning.

The data is a matrix whose rows correspond to the subjects and columns to the items.

Missing values are allowed but must be coded as NA values. They are discarded for item parameter estimation.

The 2PL model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006).

Value

A matrix with one row per item and five columns: the estimates of item discrimination *a* and difficulty *b* parameters, the related standard errors *se(a)* and *se(b)*, and the covariances *cov(a,b)*, in this order.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

`itemPar1PL`, `itemPar3PL`, `itemPar3PLconst`, `itemParEst`, `difLord`, `difRaju`, `difGenLord`

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Getting item parameter estimates
 itemPar2PL(verbal[,1:24])

## End(Not run)
```

---

itemPar3PL   *Item parameter estimation for DIF detection using 3PL model*

Description

Fits the 3PL model and returns related item parameter estimates.

Usage

```
itemPar3PL(data)
```

Arguments

data   numeric: the data matrix.

Details

`itemPar3PL` permits to get item parameter estimates from the 3PL model. The output is ordered such that it can be directly used with the general `itemParEst` command, as well as the methods of Lord (`difLord`) and Raju (`difRaju`) and Generalized Lord's (`difGenLord`) to detect differential item functioning.

The output consists of nine columns which are displayed in the following order. The first three columns hold the estimates of item discrimination a, difficulty b and pseudo-guessing c parameters. In the next three columns one can find the related standard errors $se(a)$, $se(b)$ and $se(c)$. Eventually, the last three columns contain the covariances between item parameters, respectively $cov(a,b)$, $cov(a,c)$ and $cov(b,c)$.

The `data` is a matrix whose rows correspond to the subjects and columns to the items.

Missing values are allowed but must be coded as `NA` values. They are discarded for item parameter estimation.

The 3PL model is fitted using marginal maximum likelihood by means of the functions from the `ltm` package (Rizopoulos, 2006).

Value

A matrix with one row per item and nine columns. See **Details**.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemPar1PL, itemPar2PL, itemPar3PLconst, itemParEst, difLord, difRaju,

difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Getting item parameter estimates
 # itemPar3PL(verbal[,1:24])

## End(Not run)
```

itemPar3PLconst	*Item parameter estimation for DIF detection using constrained 3PL model*

Description

Fits the 3PL model with constrained pseudo-guessing values and returns related item parameter estimates.

Usage

```
itemPar3PLconst(data, c=rep(0,ncol(data)))
```

Arguments

data	numeric: the data matrix.
c	numeric value or vector of constrained pseudo-guessing parameters. See **Details**.

Details

itemPar3PLconst permits to get item parameter estimates from the 3PL model for which the pseudo-guessing parameters are constrained to some fixed values. The output is ordered such that it can be directly used with the general itemParEst command, as well as the methods of Lord (difLord) and Raju (difRaju) and Generalized Lord's (difGenLord) to detect differential item functioning.

The output is similar to that of itemPar2PL method to fit the 2PL model; an additional column is included and holds the fixed pseudo-guessing parameter values.

The data is a matrix whose rows correspond to the subjects and columns to the items.

Missing values are allowed but must be coded as NA values. They are discarded for item parameter estimation.

The argument c can be either a single numeric value or a numeric vector of the same length of the number of items. In the former case, the pseudo-guessing parameters are considered to be all identical to the given c value; otherwise c is directly used to constraint these parameters.

The constrained 3PL model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006).

Value

A matrix with one row per item and six columns: the item discrimination *a* and difficulty estimates *b*, the corresponding standard errors *se(a)* and *se(b)*, the covariances *cov(a,b)* and the constrained pseudo-guessing values *c*.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis

Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemPar1PL, itemPar2PL, itemPar3PL, itemParEst, difLord, difRaju,

difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Constraining all pseudo-guessing parameters to be equal to 0.05
 itemPar3PLconst(verbal[,1:24], c=0.05)

 # Constraining pseudo-guessing values to  0.1 for the first 10 items,
 # and to 0.05 for the remaining ones
 itemPar3PLconst(verbal[,1:24], c=c(rep(0.1,10),rep(0.05,14)))

## End(Not run)
```

---

itemParEst   *Item parameter estimation for DIF detection*

Description

Fits a specified logistic IRT model and returns related item parameter estimates.

Usage

```
itemParEst(data, model, c=NULL, engine="ltm", discr=1)
```

Arguments

`data`	numeric: the data matrix.
`model`	character: the IRT model to be fitted (either ″1PL″, ″2PL″ or ″3PL″).
`c`	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
`engine`	character: the engine for estimating the 1PL model, either ″ltm″ (default) or ″lme4″.
`discr`	either NULL or a real positive value for the common discrimination parameter (default is 1). Used only if model is ″1PL″ and engine is ″ltm″. See **Details**.

Details

`itemParEst` permits to get item parameter estimates of some pre-specified logistic IRT model, together with estimates of the standard errors and the covariances between item parameters, if any. The output is ordered such that it can be directly used with the methods of Lord ([difLord](#)) and Raju ([difRaju](#)) and Generalized Lord's ([difGenLord](#)) to detect differential item functioning.

The data is a matrix whose rows correspond to the subjects and columns to the items.

Missing values are allowed but must be coded as NA values. They are discarded for item parameter estimation.

If the model is not the 1PL model, or if engine is equal to ″ltm″, the selected IRT model is fitted using marginal maximum likelihood by means of the functions from the ltm package (Rizopoulos, 2006). Otherwise, the 1PL model is fitted as a generalized linear mixed model, by means of the glmer function of the lme4 package (Bates and Maechler, 2009). With the ″ltm″ engine, the common discrimination parameter can be either fixed to a constant value using the discr argument, or it can be estimated (though not returned) by specifying discr to NULL. The default value of the common discrimination is 1.

The 3PL model can be fitted either unconstrained or by fixing the pseudo-guessing values. In the latter case the argument c holds either a numeric vector of same length of the number of items, with one value per item pseudo-guessing parameter, or a single value which is duplicated for all the items. If c is different from NULL then the 3PL model is always fitted (whatever the value of model).

Each row of the output matrix corresponds to one item of the data set; the number of columns depends on the fitted model. At most, nine columns are produced, with the unconstrained 3PL model. The order of the columns is the following: first, the estimates of item discrimination *a*, difficulty *b* and pseudo-guessing *c*; second, the corresponding standard errors *se(a)*, *se(b)* and *se(c)*; finally, the covariances between the item parameters, *cov(a,b)*, *cov(a,c)* and *cov(b,c)*.

If the 2PL model is fitted, only five columns are displayed: *a*, *b*, *se(a)*, *se(b)* and *cov(a,b)*. In case of the 1PL model, only *b* and *se(b)* are returned. If the constrained 3PL is considered, the output matrix holds six columns, the first five being identical to those from the 2PL model, and the last one holds the fixed pseudo-guessing parameters.

Value

A matrix with one row per item and at most nine columns, with item parameter estimates, standard errors and covariances, if any. See **Details**.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)

Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software, 17*, 1-25. URL: http://www.jstatsoft.org/v17/i05/

See Also

itemPar1PL, itemPar2PL, itemPar3PL, itemPar3PLconst, difLord, difRaju, difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Estimation of the item parameters (1PL model, "ltm" engine)
 items.1PL<-itemParEst(verbal[,1:24],model="1PL")

 # Estimation of the item parameters (1PL model, "ltm" engine,
 # estimated common discrimination parameter)
 items.1PL<-itemParEst(verbal[,1:24],model="1PL", discr=NULL)

 # Estimation of the item parameters (1PL model, "lme4" engine)
 items.1PL<-itemParEst(verbal[,1:24],model="1PL", engine="lme4")

 # Estimation of the item parameters (2PL model)
 items.2PL<-itemParEst(verbal[,1:24],model="2PL")

 # Estimation of the item parameters (3PL model)
 # items.3PL<-itemParEst(verbal[,1:24],model="3PL")

 # Constraining all pseudo-guessing values to be equal to 0.05
 items.3PLc<-itemParEst(verbal[,1:24],model="3PL",c=0.05)

## End(Not run)
```

itemRescale	*Rescaling item parameters by equal means anchoring*

Description

Rescale the item parameters from one data set to the scale of the parameters from another data set, using equal means anchoring.

Usage

```
itemRescale(mR, mF, items=1:nrow(mR))
```

Arguments

mR numeric: a matrix of item parameter estimates (one row per item) which constitutes the reference scale. See **Details**.

mF numeric: a matrix of item parameter estimates (one row per item) which have to be rescaled. See **Details**.

items a numeric vector of integer values specifying which items are used for equal means anchoring. See **Details**.

Details

The matrices mR and mF must have the same format as the output of the command `itemParEst` and one the possible models (1PL, 2PL, 3PL or constrained 3PL). The number of columns therefore equals two, five, nine or six, respectively.

Rescaling is performed by equal means anchoring (Cook and Eignor, 1991). The items involved in the anchoring process are specified by means of their row number in either mR or mF, and are passed through the items argument.

itemRescale primarily serves as a routine for item purification in Lord (`difLord`) and Raju (`difRaju`) Generalized Lord's (`difGenLord`) methods of DIF identification (Candell and Drasgow, 1988).

Value

A matrix of the same format as mF with the rescaled item parameters.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Candell, G.L. and Drasgow, F. (1988). An iterative procedure for linking metrics and assessing item bias in item response theory. *Applied Psychological Measurement, 12*, 253-260.

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

itemPar1PL, itemPar2PL, itemPar3PL, itemPar3PLconst, difLord, difRaju,

difGenLord

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Splitting the data set into reference and focal groups
 nF<-sum(Gender)
 nR<-nrow(verbal)-nF
 data.ref<-verbal[,1:24][order(Gender),][1:nR,]
 data.focal<-verbal[,1:24][order(Gender),][(nR+1):(nR+nF),]

 # Estimating item parameters in each data set with 1PL model
 mR<-itemPar1PL(data.ref)
 mF<-itemPar1PL(data.focal)

 # Rescaling focal group item parameters, using all items for anchoring
 itemRescale(mR, mF)

 # Rescaling focal group item parameters, using the first 10 items for
 # anchoring
 itemRescale(mR, mF, items=1:10)

 # Estimating item parameters in each data set with 2PL model
 mR<-itemPar2PL(data.ref)
 mF<-itemPar2PL(data.focal)

 # Rescaling focal group item parameters, using all items for anchoring
 itemRescale(mR, mF)

## End(Not run)
```

Logistik	*Logistic regression DIF statistic*

Description

Calculates the "logistic regression" likelihood-ratio statistics and effect sizes for DIF detection.

Usage

```
Logistik(data, member, anchor=1:ncol(data), type="both",
    criterion="LRT")
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and one entries only. See **Details**.
anchor	a vector of integer values specifying which items (all by default) are currently considered as anchor (DIF free) items. See **Details**.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	a character string specifying which DIF statistic is computed. Possible values are "LRT" (default) or "Wald". See **Details**.

Details

This command computes the logistic regression statistic (Swaminathan and Rogers, 1990) in the specific framework of differential item functioning. It forms the basic command of difLogistic and is specifically designed for this call.

The three possible models to be fitted are:

$$M_0 : logit(\pi_g) = \alpha + \beta X + \gamma_g + \delta_g X$$

$$M_1 : logit(\pi_g) = \alpha + \beta X + \gamma_g$$

$$M_2 : logit(\pi_g) = \alpha + \beta X$$

where π_g is the probability of answering correctly the item in group g and X is the sum score. Parameters α and β are the intercept and the slope of the logistic curves (common to all groups), while γ_g and δ_g are group-specific parameters. For identification reasons the parameters γ_0 and δ_0 for reference group ($g = 0$) are set to zero. The parameter γ_1 of the focal group ($g = 1$) represents the uniform DIF effect, and the parameter δ_1 is used to model nonuniform DIF effect. The models are fitted with the glm function.

Two types of DIF statistics can be computed: the likelihood ratio test statistics, obtained by comparing the fit of two nested models, and the Wald statistics, obtained with an appropriate contrast matrix for testing the model parameters (Johnson and Wichern, 1998). These are specified by the argument criterion, with respective values "LRT" and "Wald". By default, the LRT statistics are computed.

If `criterion` is `"LRT"`, the argument `type` determines the models to be compared by means of the LRT statistics. The three possible values of `type` are: `type="both"` (default) which tests the hypothesis $H_0 : \gamma_1 = \delta_1 = 0$ by comparing models M_0 and M_2; `type="nudif"` which tests the hypothesis $H_0 : \delta_1 = 0$ by comparing models M_0 and M_1; and `type="udif"` which tests the hypothesis $H_0 : \gamma_1 = 0$ by comparing models M_1 and M_2 (assuming that $\delta_1 = 0$). In other words, `type="both"` tests for DIF (without distinction between uniform and nonuniform effects), while `type="udif"` and `type="nudif"` test for uniform and nonuniform DIF, respectively.

If `criterion` is `"Wald"`, the argument `type` determines the logistic model to be considered and the appropriate contrast matrix. If `type=="both"`, the considered model is model M_0 and the contrast matrix has two rows, (0,0,1,0) and (0,0,0,1). If `type=="nudif"`, the considered model is also model M_0 but the contrast matrix has only one row, (0,0,0,1). Eventually, if `type=="udif"`, the considered model is model M_1 and the contrast matrix has one row, (0,0,1).

The data are passed through the `data` argument, with one row per subject and one column per item. Missing values are allowed but must be coded as `NA` values. They are discarded from the fitting of the logistic models (see `glm` for further details).

The vector of group membership, specified with `member` argument, must hold only zeros and ones, a value of zero corresponding to the reference group and a value of one to the focal group.

Option `anchor` sets the items which are considered as anchor items for computing logistic regression DIF statistics. Items other than the anchor items and the tested item are discarded. `anchor` must hold integer values specifying the column numbers of the corresponding anchor items. It is mainly designed to perform item purification.

The output contains: the selected DIF statistics (either the LRT or the Wald statistic) computed for each item, and two matrices with the parameter estimates of both models, for each item. In addition, Nagelkerke's R^2 coefficients (Nagelkerke, 1991) are computed for each model and the output returns the differences in these coefficients. Such differences are used as measures of effect size by the `difLogistic` command; see Gomez-Benito, Dolores Hidalgo and Padilla (2009), Jodoin and Gierl (2001) and Zumbo and Thomas (1997). The `criterion` argument is also returned.

Value

A list with four components:

`stat`	the values of the logistic regression DIF statistics.
`deltaR2`	the differences between Nagelkerke's R^2 coefficients of the tested models. See **Details**.
`parM0`	a matrix with one row per item and four columns, holding successively the fitted parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}_1$ and $\hat{\delta}_1$ of the "full" model (M_0 if `type="both"` or `type="nudif"`, M_1 if `type="udif"`).
`parM1`	the same matrix as `parM0` but with fitted parameters for the "simpler" model (M_1 if `type="nudif"`, M_2 if `type="both"` or `type="udif"`).
`criterion`	the value of the `criterion` argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences

Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Gomez-Benito, J., Dolores Hidalgo, M. and Padilla, J.-L. (2009). Efficacy of effect size measures in logistic regression: an application for detecting DIF. *Methodology, 5*, 18-25.

Jodoin, M. G. and Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.

Johnson, R. A. and Wichern, D. W. (1998). *Applied multivariate statistical analysis (fourth edition)*. Upper Saddle River, NJ: Prentice-Hall.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691-692.

Swaminathan, H. and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.

Zumbo, B. D. and Thomas, D. R. (1997). A measure of effect size for a model-based approach for studying DIF. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science.

See Also

difLogistic, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Testing both types of DIF simultaneously
 # With all items
 Logistik(verbal[,1:24], verbal[,26])

 # Testing both types of DIF simultaneously
 # With all items and Wald test
 Logistik(verbal[,1:24], verbal[,26], criterion="Wald")

 # Removing item 6 from the set of anchor items
 Logistik(verbal[,1:24], verbal[,26], anchor=c(1:5,7:24))

 # Testing for nonuniform DIF
 Logistik(verbal[,1:24], verbal[,26], type="nudif")
```

```
 # Testing for uniform DIF
 Logistik(verbal[,1:24], verbal[,26], type="udif")

## End(Not run)
```

LordChi2   *Lord's chi-square DIF statistic*

Description

Calculates the Lord's chi-square statistics for DIF detection.

Usage

```
LordChi2(mR, mF)
```

Arguments

mR	numeric: the matrix of item parameter estimates (one row per item) for the reference group. See **Details**.
mF	numeric: the matrix of item parameter estimates (one row per item) for the focal group. See **Details**.

Details

This command computes the Lord's chi-square statistic (Lord, 1980) in the specific framework of differential item functioning. It forms the basic command of difLord and is specifically designed for this call.

The matrices mR and mF must have the same format as the output of the command itemParEst with one the possible models (1PL, 2PL, 3PL or constrained 3PL). The number of columns therefore equals two, five, nine or six, respectively. Moreover, item parameters of the focal must be on the same scale of that of the reference group. If not, make use of e.g. equal means anchoring (Cook and Eignor, 1991) and itemRescale to transform them adequately.

Value

A vector with the values of the Lord's chi-square DIF statistics.

Note

WARNING: the previous versions of LordChi2 were holding an error: under the 3PL model, the covariance matrices Sig_1 and Sig_2 were wrongly computed as the variance of the pseudo-guessing parameters were replaced by the parameter estimates. This has been fixed from since version 4.0 of difR. Many thanks to J. Patrick Meyer (Curry School of Education, University of Virginia) for having discovered this mistake.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

itemParEst, itemRescale, difLord, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Splitting the data into reference and focal groups
 nF<-sum(Gender)
 nR<-nrow(verbal)-nF
 data.ref<-verbal[,1:24][order(Gender),][1:nR,]
 data.focal<-verbal[,1:24][order(Gender),][(nR+1):(nR+nF),]

 # Pre-estimation of the item parameters (1PL model)
 mR<-itemParEst(data.ref,model="1PL")
 mF<-itemParEst(data.focal,model="1PL")
 mF<-itemRescale(mR, mF)
 LordChi2(mR, mF)

 # Pre-estimation of the item parameters (2PL model)
 mR<-itemParEst(data.ref,model="2PL")
 mF<-itemParEst(data.focal,model="2PL")
 mF<-itemRescale(mR, mF)
 LordChi2(mR, mF)
```

```
# Pre-estimation of the item parameters (constrained 3PL model)
mR<-itemParEst(data.ref,model="3PL",c=0.05)
mF<-itemParEst(data.focal,model="3PL",c=0.05)
mF<-itemRescale(mR, mF)
LordChi2(mR, mF)

## End(Not run)
```

LRT   *Likelihood-Ratio Test DIF statistic*

Description

Calulates Likelihoo-Ratio Test (LRT) statistics for DIF detection.

Usage

```
LRT(data, member)
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and one entries only. See **Details**.

Details

This command computes the likelihood-ratio test statistic (Thissen, Steinberg and Wainer, 1988) in the specific framework of differential item functioning. It forms the basic command of difLRT and is specifically designed for this call.

The data are passed through the data argument, with one row per subject and one column per item. Missing values are allowed but must be coded as NA values.

The vector of group membership, specified with member argument, must hold only zeros and ones, a value of zero corresponding to the reference group and a value of one to the focal group.

The LRT DIF statistic is computed for each item separately, using all other items as anchor items.

Value

A vector with the values of the LRT DIF statistics.

Note

Because of the fitting of the modified Rasch model with glmer the process can be very time consuming (see the **Details** section of difLRT).

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-31. http://CRAN.R-project.org/package=lme4

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Thissen, D., Steinberg, L. and Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer and H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

See Also

difLRT, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Keeping the first 5 items and the first 50 subjects
 # (this is an artificial simplification to reduce the computational time)
 # Sixth column holds the group membership
 verbal<-verbal[1:50,c(1:5,25)]

 # Likelihood-ratio statistics
 LRT(verbal[,1:5],verbal[,6])

## End(Not run)
```

---

mantelHaenszel   *Mantel-Haenszel DIF statistic*

Description

Calculates Mantel-Haenszel statistics for DIF detection.

Usage

```
mantelHaenszel(data, member, correct=TRUE, exact=FALSE,
  anchor=1:ncol(data))
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and one entries only. See **Details**.
correct	logical: should the continuity correction be used? (default is TRUE).
exact	logical: should an exact test be computed? (default is FALSE).
anchor	a vector of integer values specifying which items (all by default) are currently considered as anchor (DIF free) items. See **Details**.

Details

This command basically computes the Mantel-Haenszel (1959) statistic in the specific framework of differential item functioning. It forms the basic command of `difMH` and is specifically designed for this call.

The data are passed through the data argument, with one row per subject and one column per item.

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership, specified with member argument, must hold only zeros and ones, a value of zero corresponding to the reference group and a value of one to the focal group.

By default, the continuity correction factor -0.5 is used (Holland and Thayer, 1988). One can nevertheless remove it by specifying correct=FALSE.

By default, the asymptotic Mantel-Haenszel statistic is computed. However, the exact statistics and related P-values can be obtained by specifying the logical argument exact to TRUE. See Agresti (1990, 1992) for further details about exact inference.

Option anchor sets the items which are considered as anchor items for computing Mantel-Haenszel statistics. Items other than the anchor items and the tested item are discarded. anchor must hold integer values specifying the column numbers of the corresponding anchor items. It is primarily designed to perform item purification.

In addition to the Mantel-Haenszel statistics to identify DIF items, mantelHaenszel computes the estimates of the common odds ratio α_{MH} which are used for measuring the effect size of the items (Holland and Thayer, 1985, 1988). They are returned in the resAlpha argument of the output list. Moreover, the logarithm of α_{MH}, say λ_{MH}, is asymptotically distributed and its variance is computed and returned into the varLambda argument. Note that this variance is the one proposed by Philips and Holland (1987), since it seems the most accurate expression for the variance of λ_{MH} (Penfield and Camilli, 2007).

Value

A list with several arguments:

resMH	the vector of the Mantel-Haenszel DIF statistics (either asymptotic or exact).
resAlpha	the vector of the (asymptotic) Mantel-Haenszel estimates of the common odds ratios. Returned only if exact is FALSE.
varLambda	the (asymptotic) variance of the λ_{MH} statistic. Returned only if exact is FALSE. See **Details**.
Pval	the exact P-values of the MH test. Returned only if exact is TRUE.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7*, 131-177.

Holland, P. W. and Thayer, D. T. (1985). An alternative definition of the ETS delta scale of item difficulty. *Research Report RR-85-43*. Princeton, NJ: Educational Testing Service.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Ed.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719-748.

Penfield, R. D., and Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao and S. Sinharray (Eds.), *Handbook of Statistics 26: Psychometrics* (pp. 125-167). Amsterdam, The Netherlands: Elsevier.

Philips, A., and Holland, P. W. (1987). Estimators of the Mantel-Haenszel log odds-ratio estimate. *Biometrics, 43*, 425-431.

See Also

difMH, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # With and without continuity correction
 mantelHaenszel(verbal[,1:24], verbal[,26])
 mantelHaenszel(verbal[,1:24], verbal[,26], correct=FALSE)

 # Exact test
 mantelHaenszel(verbal[,1:24], verbal[,26], exact=TRUE)

 # Removing item 6 from the set of anchor items
 mantelHaenszel(verbal[,1:24], verbal[,26], anchor=c(1:5,7:24))

## End(Not run)
```

RajuZ   *Raju's area DIF statistic*

Description

Calculates the Raju's statistics for DIF detection.

Usage

```
RajuZ(mR, mF, signed=FALSE)
```

Arguments

mR	numeric: the matrix of item parameter estimates (one row per item) for the reference group. See **Details**.
mF	numeric: the matrix of item parameter estimates (one row per item) for the focal group. See **Details**.
signed	logical: should the *signed* area be computed, or the *unsigned* (i.e. in absolute value) ara? Default is FALSE, i.e. the unsigned area. See **Details**.

Details

This command computes the Raju's area statistic (Raju, 1988, 1990) in the specific framework of differential item functioning. It forms the basic command of `difRaju` and is specifically designed for this call.

The matrices mR and mF must have the same format as the output of the command `itemParEst` and one the possible models (1PL, 2PL or constrained 3PL). The number of columns therefore equals two, five or six, respectively. Note that the unconstrained 3PL model cannot be used in this method: all pseudo-guessing parameters must be equal in both groups of subjects. Moreover, item parameters of the focal must be on the same scale of that of the reference group. If not, make use of e.g. equal means anchoring (Cook and Eignor, 1991) and `itemRescale` to transform them adequately.

By default, the *unsigned* area, given by Equation (57) in Raju (1990), is computed. It makes use of Equations (14), (15), (23) and (46) for the numerator, and Equations (17), (33) to (39), and (52) for the denominator of the *Z* statistic. However, the *signed* area, given by Equation (56) in Raju (1990), can be used instead. In this case, Equations (14), (21) and (44) are used for the numerator, and Equations (17), (25) and (48) for the denominator. The choice of the type of area is fixed by the logical *signed* argument, with default value `FALSE`.

Value

A list with two components:

res	a matrix with one row per item and three columns, holding respectively Raju's area between the two item characteristic curves, its standard error and the Raju DIF statistic (the latter being the ratio of the first two columns).
signed	the value of the `signed` argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Cook, L. L. and Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational Measurement: Issues and Practice, 10*, 37-45.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Raju, N.S. (1988). The area between two item characteristic curves. *Psychometrika, 53*, 495-502.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*, 197-207.

See Also

itemParEst, itemRescale, difRaju, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
```

```
attach(verbal)

# Splitting the data into reference and focal groups
nF<-sum(Gender)
nR<-nrow(verbal)-nF
data.ref<-verbal[,1:24][order(Gender),][1:nR,]
data.focal<-verbal[,1:24][order(Gender),][(nR+1):(nR+nF),]

# Pre-estimation of the item parameters (1PL model)
mR<-itemParEst(data.ref,model="1PL")
mF<-itemParEst(data.focal,model="1PL")
mF<-itemRescale(mR, mF)

# Signed and unsigned Raju statistics
RajuZ(mR, mF)
RajuZ(mR, mF, signed=TRUE)

# Pre-estimation of the item parameters (2PL model)
mR<-itemParEst(data.ref,model="2PL")
mF<-itemParEst(data.focal,model="2PL")
mF<-itemRescale(mR, mF)

# Signed and unsigned Raju statistics
RajuZ(mR, mF)
RajuZ(mR, mF, signed=TRUE)

# Pre-estimation of the item parameters (constrained 3PL model)
mR<-itemParEst(data.ref,model="3PL",c=0.05)
mF<-itemParEst(data.focal,model="3PL",c=0.05)
mF<-itemRescale(mR, mF)

# Signed and unsigned Raju statistics
RajuZ(mR, mF)
RajuZ(mR, mF, signed=TRUE)

## End(Not run)
```

---

selectDif   *Selection of one of the DIF detection methods*

Description

This function performs DIF detection for one pre-specified method.

Usage

```
selectDif(Data, group, focal.name, method, props=NULL, thrTID=1.5,
    alpha=0.05, MHstat="MHChisq", correct=TRUE, exact=FALSE,
    stdWeight="focal", thrSTD=0.1, BDstat="BD",type="both",
    criterion="LRT", model="2PL", c=NULL, engine="ltm", discr=1,
    irtParam=NULL, same.scale=TRUE, signed=FALSE, purify=FALSE,
    nrIter=10, save.output=FALSE, output=c("out","default"))
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.name	numeric or character indicating the level of group which corresponds to the focal group.
method	character: the name of the selected method. See **Details**.
props	either NULL (default) or a two-column matrix with proportions of success in the reference group and the focal group. See **Details** .
thrTID	numeric: the threshold for detecting DIF items with TID method (default is 1.5).
alpha	numeric: significance level (default is 0.05).
MHstat	character: specifies the DIF statistic to be used for DIF identification. Possible values are "MHChisq" (default) and "logOR". See **Details** .
correct	logical: should the continuity correction be used? (default is TRUE).
exact	logical: should an exact test be computed? (default is FALSE).
stdWeight	character: the type of weights used for the standardized P-DIF statistic. Possible values are "focal" (default), "reference" and "total". See **Details**.
thrSTD	numeric: the threshold (cut-score) for standardized P-DIF statistic (default is 0.10).
BDstat	character specifying the DIF statistic to be used. Possible values are "BD" (default) and "trend". See **Details**.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	a character string specifying which DIF statistic is computed. Possible values are "LRT" (default) or "Wald". See **Details**.
model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL"). Default is "2PL".
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.
signed	logical: should the Raju's statistics be computed using the signed (TRUE) or unsigned (FALSE, default) area? See **Details**.
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).

save.output   logical: should the output be saved into a text file? (Default is FALSE).

output   character: a vector of two components. The first component is the name of the
               output file, the second component is either the file path or "default" (default
               value). See **Details**.

Details

This is a generic function which calls one of the DIF detection methods and displays its output. It
is mainly used as a routine for dichoDif command.

The possible methods are: "TID" for Transformed Item Difficulties (TID) method (Angoff and Ford,
1973), "MH" for mantel-Haenszel (Holland and Thayer, 1988), "Std" for standardization (Dorans
and Kulick, 1986), "Logistic" for logistic regression (Swaminathan and Rogers, 1990), "BD" for
Breslow-Day method (Penfield, 2003), "Lord" for Lord's chi-square test (Lord, 1980), "Raju" for
Raju's area method (Raju, 1990), and "LRT" for likelihood-ratio test method (Thissen, Steinberg
and Wainer, 1988).

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addi-
tion, Data can hold the vector of group membership. If so, group indicates the column of Data
which corresponds to the group membership, either by specifying its name or by giving the column
number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA
values. They are discarded from either the computation of the sum-scores, the fitting of the logistic
models or the IRT models (according to the method).

The vector of group membership must hold only two different values, either as numeric or character.
The focal group is defined by of the argument focal.name.

With the TID method, one can alternatively provide the matrix of proportions of success in for each
item in each group. This matrix must have the same format as that provided to the trItemDiff
function; see the corresponding help file for further details.

For Lord and Raju methods, one can specify either the IRT model to be fitted (by means of model,
c, engine and discr arguments), or the item parameter estimates with arguments irtParam and
same.scale. See difLord and difRaju for further details.

The threshold for detecting DIF items depends on the method. For standardization it has to be fully
specified (with the thr argument), as well as for the TID method (through the thrTID argument).
For the other methods it is depending on the significance level set by alpha.

For Mantel-Haenszel method, the DIF statistic can be either the Mantel-Haenszel chi-square statistic
or the log odds-ratio statistic. The method is specified by the argument MHstat, and the default
value is "MHChisq" for the chi-square statistic. Moreover, the option correct specifies whether the
continuity correction has to be applied to Mantel-Haenszel statistic. See difMH for further details.

By default, the asymptotic Mantel-Haenszel statistic is computed. However, the exact statistics and
related P-values can be obtained by specifying the logical argument exact to TRUE. See Agresti
(1990, 1992) for further details about exact inference.

The weights for computing the standardized P-DIF statistics are defined through the argument
stdWeight, with possible values "focal" (default value), "reference" and "total". See stdPDIF
for further details.

For Breslow-Day method, two test statistics are available: the usual Breslow-Day statistic for test-
ing homogeneous association (Aguerri, Galibert, Attorresi and Maranon, 2009) and the trend test
statistic for assessing some monotonic trend in the odds ratios (Penfield, 2003). The DIF statistic is
supplied by the BDstat argument, with values "BD" (default) for the usual statistic and "trend" for
the trend test statistic.

For logistic regression, the argument `type` permits to test either both uniform and nonuniform effects simultaneously (type="both"), only uniform DIF effect (type="udif") or only nonuniform DIF effect (type="nudif"). The `criterion` argument specifies the DIF statistic to be computed, either the likelihood ratio test statistic (with criterion="LRT") or the Wald test (with criterion="Wald"). See `Logistik` for further details.

For Raju's method, the type of area (signed or unsigned) is fixed by the logical `signed` argument, with default value FALSE (i.e. unsigned areas). See `RajuZ` for further details.

Item purification can be requested by specifying `purify` option to TRUE. Recall that item purification is slightly different for IRT and for non-IRT based methods. See the corresponding methods for further information.

The output of the selected method can be stored in a text file by fixing `save.output` and `output` appropriately. See the help file of the corresponding method for further information.

Value

The output of the selected DIF detection method.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7*, 131-177.

Aguerri, M.E., Galibert, M.S., Attorresi, H.F. and Maranon, P.P. (2009). Erroneous detection of nonuniform DIF using the Breslow-Day test in a short test. *Quality and Quantity, 43*, 35-44.

Angoff, W. H., and Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 2*, 95-106.

Dorans, N. J. and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*, 355-368.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Dirs.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Penfield, R.D. (2003). Application of the Breslow-Day test of trend in odds ratio heterogeneity to the detection of nonuniform DIF. *Alberta Journal of Educational Research, 49*, 231-243.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*, 197-207.

Swaminathan, H. and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.

Thissen, D., Steinberg, L. and Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer and H. Braun (Eds.), *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

See Also

[difTID](), [difMH](), [difStd](), [difBD](), [difLogistic](), [difLord](), [difRaju](), [difLRT](), [dichoDif]()

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Excluding the "Anger" variable
 verbal<-verbal[colnames(verbal)!="Anger"]

 # Calling Mantel-Haenszel
 selectDif(verbal, group=25, focal.name=1, method="MH")

 # Calling Mantel-Haenszel and saving output in 'MH.txt' file
 selectDif(verbal, group=25, focal.name=1, method="MH",
    save.output=TRUE,output=c("MH","default"))

 # Calling Lord method
 # 2PL model, with item purification
 selectDif(verbal, group=25, focal.name=1, method="Lord",model="2PL",
 purify=TRUE)

## End(Not run)
```

---

selectGenDif   *Selection of one of the DIF detection methods among multiple groups*

Description

This function performs DIF detection among multiple groups for one pre-specified method.

Usage

```
selectGenDif(Data, group, focal.names, method, type="both",
   criterion="LRT", alpha=0.05, model="2PL", c=NULL,
   engine = "ltm", discr=1, irtParam=NULL, nrFocal=2,
   same.scale=TRUE, purify=FALSE, nrIter=10,
   save.output=FALSE, output=c("out","default"))
```

Arguments

Data	numeric: either the data matrix only, or the data matrix plus the vector of group membership. See **Details**.
group	numeric or character: either the vector of group membership or the column indicator (within data) of group membership. See **Details**.
focal.names	numeric or character vector indicating the levels of group which correspond to the focal groups.
method	character: the name of the selected method. See **Details**.
type	a character string specifying which DIF effects must be tested. Possible values are "both" (default), "udif" and "nudif". See **Details**.
criterion	character: the type of test statistic used to detect DIF items with generalized logistic regression. Possible values are "LRT" (default) and "Wald". See **Details**.
alpha	numeric: significance level (default is 0.05).
model	character: the IRT model to be fitted (either "1PL", "2PL" or "3PL"). Default is "2PL".
c	optional numeric value or vector giving the values of the constrained pseudo-guessing parameters. See **Details**.
engine	character: the engine for estimating the 1PL model, either "ltm" (default) or "lme4".
discr	either NULL or a real positive value for the common discrimination parameter (default is 1). Used onlky if model is "1PL" and engine is "ltm". See **Details**.
irtParam	matrix with *2J* rows (where *J* is the number of items) and at most 9 columns containing item parameters estimates. See **Details**.
nrFocal	numeric: the number of focal groups (default is 2).
same.scale	logical: are the item parameters of the irtParam matrix on the same scale? (default is "TRUE"). See **Details**.
purify	logical: should the method be used iteratively to purify the set of anchor items? (default is FALSE).
nrIter	numeric: the maximal number of iterations in the item purification process (default is 10).
save.output	logical: should the output be saved into a text file? (Default is FALSE).
output	character: a vector of two components. The first component is the name of the output file, the second component is either the file path or "default" (default value). See **Details**.

Details

This is a generic function which calls one of the DIF detection methods for multiple groups, and displays its output. It is mainly used as a routine for genDichoDif command.

There are three possible methods currently implemented: ″GMH″ for Generalized Mantel-Haenszel (Penfield, 2001), ″genLogistic″ for generalized logistic regression (Magis, Raiche, Beland and Gerard, 2010) and ″genLord″ for generalized Lord's chi-square test (Kim, Cohen and Park, 1995).

The Data is a matrix whose rows correspond to the subjects and columns to the items. In addition, Data can hold the vector of group membership. If so, group indicates the column of Data which corresponds to the group membership, either by specifying its name or by giving the column number. Otherwise, group must be a vector of same length as nrow(Data).

Missing values are allowed for item responses (not for group membership) but must be coded as NA values. They are discarded from either the computation of the sum-scores, the fitting of the logistic models or the IRT models (according to the method).

The vector of group membership must hold at least three different values, either as numeric or character. The focal groups are defined by the values of the argument focal.names.

For the generalized logistic regression method, the argument type permits to test either both uniform and nonuniform effects simultaneously (with type="both"), only uniform DIF effect (with type="udif") or only nonuniform DIF effect (with type="nudif"). Furthermore, the argument criterion defines which test must be used, either the Wald test (″Wald″) or the likelihood ratio test (″LRT″).

For generalized Lord method, one can specify either the IRT model to be fitted (by means of model, c, engine and discr arguments), or the item parameter estimates with arguments irtParam, nrFocal and same.scale. See difGenLord for further details.

The threshold for detecting DIF items depends on the method and is depending on the significance level set by alpha.

Item purification can be requested by specifying purify option to TRUE. Recall that item purification is slightly different for IRT and for non-IRT based methods. See the corresponding methods for further information.

The output of the selected method can be stored in a text file by fixing save.output and output appropriately. See the help file of the corresponding method for further information.

Value

The output of the selected DIF detection method.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Kim, S.-H., Cohen, A.S. and Park, T.-H. (1995). Detection of differential item functioning in multiple groups. *Journal of Educational Measurement, 32*, 261-276.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2010). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

Penfield, R. D. (2001). Assessing differential item functioning among multiple groups: a comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.

See Also

difGMH, difGenLogistic, difGenLord

Examples

```
## Not run:

# Loading of the verbal data
data(verbal)
attach(verbal)

# Creating four groups according to gender ("Man" or "Woman") and trait
# anger score ("Low" or "High")
group<-rep("WomanLow",nrow(verbal))
group[Anger>20 & Gender==0]<-"WomanHigh"
group[Anger<=20 & Gender==1]<-"ManLow"
group[Anger>20 & Gender==1]<-"ManHigh"

# New data set
Verbal<-cbind(verbal[,1:24],group)

# Reference group: "WomanLow"
names<-c("WomanHigh","ManLow","ManHigh")

# Calling generalized Mantel-Haenszel
selectGenDif(Verbal, group=25, focal.names=names, method="GMH")

# Calling generalized Mantel-Haenszel and saving output in 'GMH.txt'
# file
selectGenDif(Verbal, group=25, focal.name=names, method="GMH",
   save.output=TRUE,output=c("GMH","default"))

# Calling generalized logistic regression
selectGenDif(Verbal, group=25, focal.names=names, method="genLogistic")

# Calling generalized Lord method (2PL model)
selectGenDif(Verbal, group=25, focal.names=names, method="genLord",
model="2PL")

## End(Not run)
```

---

stdPDIF   *Standardization DIF statistic*

Description

Calculates standardized P-difference statistics for DIF detection.

Usage

```
stdPDIF(data, member, anchor=1:ncol(data), stdWeight="focal")
```

Arguments

data	numeric: the data matrix (one row per subject, one column per item).
member	numeric: the vector of group membership with zero and one entries only. See **Details**.
anchor	a vector of integer values specifying which items (all by default) are currently considered as anchor (DIF free) items. See **Details**.
stdWeight	character: the type of weights used for the standardized P-DIF statistic. Possible values are "focal" (default), "reference" and "total". See **Details**.

Details

This command computes the standardized P-DIF statistic in the specific framework of differential item functioning (Dorans and Kulick, 1986). It forms the basic command of [difStd](#) and is specifically designed for this call. In addition, the standardized alpha values (Dorans, 1989) are also computed as a basis for effect size calculation.

The standardized P-DIF statistic is a weighted average of the difference in proportions of successes in the reference group and in the focal group. The average is computed across the test score strata. The weights can be of three kinds (Dorans, 1989; Dorans and Kulick, 1986) and are specified through the stdWeight argument: the proportion of focal groups examinees within each stratum (stdWeight="focal"), the proportion of reference group examinees within each stratum (stdWeight="reference"), and the proportion of examinees (from both groups) within each stratum (stdWeight="total"). By default, the weights are built from the focal group.

Similarly to the 'alpha' estimates of the common odds ratio for the Mantel-Haenszel method (see [mantelHaenszel](#)), the *standardized alpha values* can be computed as rough measures of effect sizes, after a transformation to the Delta Scale (Holland, 1985). See Dorans (1989, p.228, Eqn.15) for further details.

The data are passed through the data argument, with one row per subject and one column per item. Missing values are allowed but must be coded as NA values. They are discarded from sum-score computation.

The vector of group membership, specified with member argument, must hold only zeros and ones, a value of zero corresponding to the reference group and a value of one to the focal group.

Option anchor sets the items which are considered as anchor items for computing standardized P-DIF statistics. Items other than the anchor items and the tested item are discarded. anchor must hold integer values specifying the column numbers of the corresponding anchor items. It is mainly designed to perform item purification.

Value

A list with two arguments:

resStd   the vector of the standardized P-DIF statistics.

resAlpha   the vector of standardized alpha values.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, <http://ppw.kuleuven.be/okp/home/>
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, <http://www.er.uqam.ca/nobel/r17165/>

References

Dorans, N. J. (1989). Two new approaches to assessing differential item functioning: standardization and the Mantel-Haenszel method. *Applied Measurement in Education, 2*, 217-233.

Dorans, N. J. and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*, 355-368.

Holland, P. W. (1985, October). *On the study of differential item performance without IRT*. Paper presented at the meeting of Military Testing Association, San Diego (CA).

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

See Also

difStd, dichoDif, mantelHaenszel

Examples

```
## Not run:
# Loading of the verbal data
data(verbal)

# All items as anchor items
stdPDIF(verbal[,1:24], verbal[,26])

# All items as anchor items, reference group weights
stdPDIF(verbal[,1:24], verbal[,26], stdWeight="reference")

# All items as anchor items, both groups' weights
stdPDIF(verbal[,1:24], verbal[,26], stdWeight="total")
```

```
  # Removing item 6 from the set of anchor items
  stdPDIF(verbal[,1:24], verbal[,26], anchor=c(1:5,7:24))

## End(Not run)
```

---

subtestLogistic   *Testing for DIF among subgroups with generalized logistic regression*

Description

Performs the Wald test to identify DIF items among a subset of groups of examinees, using the results of generalized logistic regression for all groups.

Usage

```
subtestLogistic(x, items, groups, alpha=0.05)
## S3 method for class 'subLogistic'
print(x, ...)
```

Arguments

x	an object of class "genLogistic", typically the output of the difGenLogistic command.
items	numeric or character: a vector of items to be tested. See **Details**.
groups	numeric or character: a vector of groups of examinees to be compared. See **Details**.
alpha	numeric: the significance level (default is 0.05).
...	other generic parameters for the print function.

Details

This command makes use of the results from the generalized logistic regression to perform subtests between two or more groups of examinees (Magis, Raiche, Beland and Gerard, 2010). The Wald test is used with an appropriate contrast matrix.

The subtestLogistic command requires a preliminary output of the generalized logistic regression with all groups of examinees, preferable with the [difGenLogistic](#) command. The object x is an object of class "genLogistic" from which subtests can be performed. The same DIF effect (either uniform, nonuniform, or both types) is tested among the subset of groups of examinees as the one tested with all groups. It is provided b y the argument type argument of x.

The argument items is a vector of the names of the items to be tested, or their number in the data set. A single item can be specified.

The argument groups specifies which groups of examinees are considered in this subtest routine. It is a vector of either group names or integer values. In the latter case, the reference group is specified with the 0 (zero) value, while the focal groups are set up by their rank in the x$focal.names argument. At least two groups must be specified, and all groups can be included (which leads back to the generalized logistic regression with the Wald test).

The output provides, among others, the Wald statistics, the degrees of freedom and related asymptotic *p*-values for each tested item, as well as the contrast matrix.

Value

A list of class "subLogistic" with the following components:

stats	a table with as many rows as tested items, and four columns: the item number, the Wald statistic, the degrees of freedom and the asymptotic *p*-value.
contrastMatrix	the contrast matrix used for testing DIF among the groups set up by groups.
items	the value of the items argument.
groups	the value of the groups argument.
type	the value of the x$type argument.
purification	the value of the x$purification argument.
alpha	the value of the alpha argument.

Author(s)

Sebastien Beland
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<sebastien.beland.1@hotmail.com>
David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Gilles Raiche
Centre sur les Applications des Modeles de Reponses aux Items (CAMRI)
Universite du Quebec a Montreal
<raiche.gilles@uqam.ca>, http://www.er.uqam.ca/nobel/r17165/

References

Magis, D., Raiche, G., Beland, S. and Gerard, P. (2010). A logistic regression procedure to detect differential item functioning among multiple groups. Unpublished manuscript.

See Also

difGenLogistic, genDichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)
 attach(verbal)

 # Creating four groups according to gender (0 or 1) and trait anger score
 # ("Low" or "High")
 # Reference group: women with low trait anger score (<=20)
 group<-rep("WomanLow",nrow(verbal))
 group[Anger>20 & Gender==0]<-"WomanHigh"
 group[Anger<=20 & Gender==1]<-"ManLow"
 group[Anger>20 & Gender==1]<-"ManHigh"
```

```
# New data set
Verbal<-cbind(verbal[,1:24],group)

# Reference group: "WomanLow"
names<-c("WomanHigh","ManLow","ManHigh")

# Testing all types of DIF with all items
rDIF<-difGenLogistic(Verbal, group=25, focal.names=names)
rUDIF<-difGenLogistic(Verbal, group=25, focal.names=names, type="udif")
rNUDIF<-difGenLogistic(Verbal, group=25, focal.names=names, type="nudif")

# Subtests between the reference group and the first two focal groups
# for item "S2WantShout" (item 6) and the three types of DIF
subGroups <- c("WomanLow","WomanHigh","ManLow")
subtestLogistic(rDIF, items=6, groups=subGroups)
subtestLogistic(rUDIF, items=6, groups=subGroups)
subtestLogistic(rNUDIF, items=6, groups=subGroups)

# Subtests between the reference group and the first focal group
# for items "S2WantShout" (item 6) and "S3WantCurse" (item 7)
# (only both DIF effects)
subGroups <- c("WomanLow","WomanHigh")
items1 <- c("S2WantShout","S3WantCurse")
items2 <- 6:7
subtestLogistic(rDIF, items=items1, groups=subGroups)
subtestLogistic(rDIF, items=items2, groups=subGroups)

## End(Not run)
```

---

trItemDiff   *Transformed Item Difficulties Perpendicular Distances*

Description

Computes the perpendicular distances for DIF detection with Transformed Item Difficulties (TID) approach.

Usage

```
trItemDiff(prop, anchor=1:nrow(prop))
```

Arguments

prop   numeric: a matrix with one row per subject and two columns: the first column
            with proportions of success in the reference group, the second column with pro-
            portions of success in the focal group.

anchor   a vector of integer values specifying which items (all by default) are currently
            considered as anchor (DIF free) items. See **Details**.

Details

This command basically computes the perpendicular distances from each Delta point to the major axis (Angoff, 1982; Angoff and Ford, 1973). It forms the basic command of `difTID` and is specifically designed for this call.

The data are passed through the `prop` argument, a matrix with one row per item and two columns. The first column holds the proportions of success (correct responses) for each item in the reference group, and the second column provides the same information but for the focal group. Missing values are not allowed. Moreover, these proportions are internally constrained into the interval $[0.001; 0.999]$ to ensure valid computation of Delta scores.

The computation is made in three steps: (a) the proportions of success are transformed into Delta scores; (b) the major axis of the ellipse made by the Delta points is determined and its intercept and slope parameters are obtained; (c) the perpendicular distances (between each Delta point and the major axis) are computed. See Angoff (1982) and Angoff and Ford (1973) for further details.

Option `anchor` sets the items which are considered as anchor items for computing the perpndicular distances. Only anchor items are used to compute the intercept and slope parameters of the major axis. anchor must hold integer values specifying the column numbers of the corresponding anchor items. It is primarily designed to perform item purification.

Value

A list with four arguments:

prop   the value of the `prop` argument.

delta   the matrix of Delta scores, in the same format as `prop`.

pars   a vector of length two with the intercept and slope parameters of the major axis of Delta points.

dist   a numeric vector of perpendicular distances.

Author(s)

David Magis
Research Group of Quantitative Psychology and Individual Differences
Katholieke Universiteit Leuven
<David.Magis@psy.kuleuven.be>, http://ppw.kuleuven.be/okp/home/
Department of Mathematics
University of Liege, Belgium
<david.magis@ulg.ac.be>

References

Angoff, W. H. (1982). Use of difficulty and discrimination indices for detecting item bias. In R. A. Berck (Ed.), *Handbook of methods for detecting item bias* (pp. 96-116). Baltimore, MD: Johns Hopkins University Press.

Angoff, W. H., and Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 2*, 95-106.

See Also

difTID, dichoDif

Examples

```
## Not run:

 # Loading of the verbal data
 data(verbal)

 # Computing proportions of success
 props <- cbind(colMeans(verbal[verbal[,26]==0,1:24]),
   colMeans(verbal[verbal[,26]==1,1:24]) )

 # Perpendicular distances
 trItemDiff(props)

 # Removing item 6 from the set of anchor items
 trItemDiff(props, anchor=c(1:5,7:24))

## End(Not run)
```

---

verbal	*Verbal Aggression Data Set*

Description

The Verbal Aggression data set comes from Vansteelandt (2000) and is made of the responses of 316 subjects (243 women and 73 men) to a questionnaire of 24 items, about verbal aggression. All items describe a frustrating situation together with a verbal aggression response. A correct answer responses is coded as 0 and 1, a value of one meaning that the subject would (want to) respond to the frustrating situation in an aggressive way. In addition, the *Trait Anger* score (Spielberger, 1988) was computed for each subject.

Format

The verbal matrix consists of 316 rows (one per subject) and 26 columns.

The first 24 columns hold the responses to the dichotomously scored items. The 25th column holds the trait anger score for each subject. The 26th column is vector of the group membership; values 0 and 1 refer to women and men, respectively.

Each item name starts with S followed by a value between 1 and 4, referring to one of the situations below:

S1: A bus fails to stop for me.

S2: I miss a train because a clerk gave me faulty information.

S3: The grocery store closes just as I am about to enter.

S4: The operator disconnects me when I had used up my last 10 cents for a call.

The second part of the name is either *Want* or *Do*, and indicates whether the subject wanted to respond to the situation or actually did respond.

The third part of the name is one of the possible aggressive responses, either *Curse*, *Scold* or *Shout*.

For example, item S1WantShout refers to the sentence: "a bus fails to stop for me. I want to shout". The corresponding item response is 1 if the subject agrees with that sentence, and 0 if not.

Source

The Verbal aggression data set is taken originally from Vansteelandt (2000) and has been used as an illustrative example in De Boeck (2008), De Boeck and Wilson (2004) and Smits, De Boeck and Vansteelandt (2004), among others. The following URL *http://bear.soe.berkely.edu/EIRM/* permits to get access to the full data set.

References

De Boeck, P. (2008). Random item IRT models. *Psychometrika, 73*, 533-559.

De Boeck, P. and Wilson, M. (2004). *Explanatory item response models: a generalized linear and nonlinear approach*. New York: Springer.

Magis, D., Beland, S., Tuerlinckx, F. and De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.

Smits, D., De Boeck, P. and Vansteelandt, K. (2004). The inhibition of verbal aggressive behavior. *European Journal of Personality, 18*, 537-555.

Spielberger, C.D. (1988). *State-trait anger expression inventory research edition. Professional manual*. Odessa, FL: Psychological Assessment Resources.

Vansteelandt, K. (2000). *Formal models for contextualized personality psychology*. Unpublished doctoral dissertation, K.U. Leuven, Belgium.

Index