Package ‘tcl’

October 6, 2025

Type Package

Title Testing in Conditional Likelihood Context

Version 1.0.1

Date 2025-10-05

Author Clemens Draxler [aut, cre],
Andreas Kurz [aut]

Maintainer Clemens Draxler <clemens.draxler@umit-tirol.at>

Description An implementation of hypothesis testing in an extended Rasch modeling framework, including sample size planning procedures and power computations. Provides 4 statistical tests, i.e., gradient test (GR), likelihood ratio test (LR), Rao score or Lagrange multiplier test (RS), and Wald test, for testing a number of hypotheses referring to the Rasch model (RM), linear logistic test model (LLTM), rating scale model (RSM), and partial credit model (PCM). Three types of functions for power and sample size computations are provided. Firstly, functions to compute the sample size given a user-specified (predetermined) deviation from the hypothesis to be tested, the level alpha, and the power of the test. Secondly, functions to evaluate the power of the tests given a user-specified (predetermined) deviation from the hypothesis to be tested, the level alpha of the test, and the sample size. Thirdly, functions to evaluate the so-called post hoc power of the tests. This is the power of the tests given the observed deviation of the data from the hypothesis to be tested and a user-specified level alpha of the test. Power and sample size computations are based on a Monte Carlo simulation approach. It is computationally very efficient. The variance of the random error in computing power and sample size arising from the simulation approach is analytically derived by using the delta method. Additionally, functions to compute the power of the tests as a function of an effect measure interpreted as explained variance are provided.
Draxler, C., & Alexandrowicz, R. W. (2015), <doi:10.1007/s11336-015-9472-y>.

License GPL-2

Depends R (>= 3.5.0)

Imports eRm, psychotools, ltm, numDeriv, graphics, grDevices, stats,
methods, MASS, splines, Matrix, lattice, rlang

Suggests knitr, rmarkdown

Encoding UTF-8

LazyLoad true

NeedsCompilation no

RoxygenNote 7.3.3

VignetteBuilder knitr

Repository CRAN

Date/Publication 2025-10-06 08:20:02 UTC

Contents

`change_test` *Tests in context of measurement of change using LLTM.*

Description

Computes gradient (GR), likelihood ratio (LR), Rao score (RS) and Wald (W) test statistics for hypotheses on parameters expressing change between two time points.

Usage

```
change_test(data)
```

Arguments

`data`	Data matrix containing the responses of n persons to 2k binary items. Columns 1 to k contain the responses to k items at time point 1, and columns (k+1) to 2k the responses to the same k items at time point 2.

Details

Assume all items be presented twice (2 time points) to the same persons. The data matrix X has n rows (number of persons) and 2k columns considered as virtual items. Assume a constant shift of item difficulties of each item between the 2 time points represented by one parameter. The shift parameter is the only parameter of interest. Of interest is the test of the hypothesis that the shift parameter equals 0 against the two-sided alternative that it is not equal to zero.

Value

A list of class `tcl` of test statistics, degrees of freedom, and p-values.

`test`	A numeric vector of gradient (GR), likelihood ratio (LR), Rao score (RS), and Wald test statistics.
`df`	Degrees of freedom.
`pvalue`	A vector of corresponding p-values.
`data`	Data matrix.
`call`	The matched call.

References

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch models: Foundations, Recent Developments, and Applications (pp. 131-155). New York: Springer.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. Psychometrika, 48(1), 3-26.

See Also

`invar_test`, and `LLTM_test`.

Examples

```
## Not run:
# Numerical example with 400 persons and 4 items
# presented twice, thus 8 virtual items

# Data y generated under the assumption that shift parameter equals 0
# (no change from time point 1 to 2)

# design matrix W used only for example data generation
#     (not used for estimating in change_test function)
W <- rbind(c(1,0,0,0,0),
  c(0,1,0,0,0),
  c(0,0,1,0,0),
  c(0,0,0,1,0),
  c(1,0,0,0,1),
  c(0,1,0,0,1),
  c(0,0,1,0,1),
  c(0,0,0,1,1))
```

```
# eta Parameter, first 4 are nuisance, i.e. , easiness parameters of the 4 items
# at time point 1, last one is the shift parameter.
eta <- c(-2,-1,1,2,0)

y <- eRm::sim.rasch(persons = rnorm(400), items = colSums(eta * t(W)))

res <- change_test(data = y)

res$test # test statistics
res$df # degrees of freedoms
res$pvalue # p-values


## End(Not run)
```

---

cml_power *Power and Power Curve Functions*

Description

Functions to compute the power of χ^2 tests, i.e., Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test, and to plot power curves as functions of effect size and sample size.

`cml_power()` computes the power of the tests given a specified effect size, type I error prob. alpha, informative sample size, and degrees of freedom.

`p_curve()` generates a power curve as a function of effect size.

`p_ncurve()` generates a power curve as a function of sample size.

Usage

```
cml_power(obj, effect = 0.03, alpha = 0.05, n = "auto", df = "auto")

p_curve(obj, alpha = 0.05, n = 300, df = "auto", from = 0, to = 0.2, ...)

p_ncurve(
  obj,
  effect = 0.03,
  alpha = 0.05,
  df = "auto",
  from = 0,
  to = 600,
  ...
)
```

Arguments

`obj`	An object of class 'tcl_sa_size', typically containing information such as degrees of freedom (`df`) and informative sample size (`n`). If missing, values for `df` and `n` need to be set manually.
`effect`	Numeric value representing the effect size. A real number between 0 and 1, interpreted as a proportion of pseudo-variance between persons with different covariate values (but the same person parameter). Default is 0.03.
`alpha`	Type I error probability. Default is 0.05.
`n`	Informative sample size (excluding persons with a score of 0 or highest possible score). Default is `"auto"`, in which case the value is extracted from `obj`.
`df`	Degrees of freedom. Default is `"auto"`, in which case the value is extracted from `obj`.
`from`	Lower bound of the effect or sample size range (default is 0).
`to`	Upper bound of the effect or sample size range (default is 0.2 for effect size, and 600 for sample site).
`...`	Additional graphical arguments passed to `plot` (e.g., `col`, `lwd`, `ylim`) via `p_curve` and `p_ncurve`.

Details

The effect is interpreted as a pseudo R^2-like measure of explained variance (as in linear models). It is 0 when persons with the same person parameters yield the same response probabilities. If two persons with the same person parameter but different covariate values yield different response probabilities, an additional variance component is introduced and the effect is greater than 0

The power of the tests is computed from the cumulative distribution function of the non-central χ^2 distribution, where the respective non-centrality parameter is obtained by multiplying the effect with the informative sample size. This is only an approximation based on results of asymptotic theory. The approximation may be poor when the informative sample size is small and/or the effect is large.

Value

- `cml_power()`: Numeric vector of power values.
- `p_curve()`, `p_ncurve()`: A power curve plotted to the active graphics device.

References

Draxler, C., & Kurz, A. (2025). Testing measurement invariance in a conditional likelihood framework by considering multiple covariates simultaneously. Behavior Research Methods, 57(1), 50.

See Also

`sa_sizeRM`, `sa_sizePCM`, and `sa_sizeChange`

Examples

```
## Not run:
##### Sample size of Rasch Model #####

res <-  sa_sizeRM(local_dev = list( c(0, -0.5, 0, 0.5, 1) , c(0, 0.5, 0, -0.5, 1)))

cml_power(obj = res)

p_curve(obj = res)
p_curve(obj = res, col = "red", lwd = 2, ylim = c(0, 1))

p_ncurve(obj = res)
p_ncurve(obj = res, col = "red", lwd = 2, ylim = c(0, 1))


## End(Not run)
```

---

discr_test   *Testing item discriminations*

Description

Computes Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test of hypothesis of equal item discriminations against the alternative that at least one item discriminates differently (only for binary data).

Usage

```
discr_test(data)
```

Arguments

`data`	Data matrix.

Details

The tests are based on the following model suggested in Draxler, Kurz, Gürer, and Nolte (2024)

$$\text{logit}\big(E(Y)\big) = \tau + \alpha + \delta(r-1),$$

where $E(Y)$ ist the expected value of a binary response (of a person to an item), $r = 1, \ldots, k-1$ is the person score, i.e., number of correct responses of that person when responding to k items, τ is the respective person parameter and α and δ are two parameters referring to the respective item. The parameter α represents a baseline, i.e., the easiness or attractiveness of the respective item in person score group $r = 1$. The parameter δ denotes the constant change of the attractiveness of that item between successive person score groups. Thus, the model assumes a linear effect of the person score r on the logit of the probability of a correct response.

The four test statistics are derived from a conditional likelihood function in which the τ parameters are eliminated by conditioning on the observed person scores. The hypothesis to be tested is formally given by setting all δ parameters equal to 0. The alternative assumes that at least one δ parameter is not equal to 0.

Value

A list of class `tcl` of test statistics, degrees of freedom, and p-values.

`test`	A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR)statistics.
`df`	A numeric vector of corresponding degrees of freedom.
`pvalue`	A vector of corresponding p-values.
`data`	Data matrix.
`call`	The matched call.

References

Draxler, C., Kurz, A., Guerer, C., & Nolte, J. P. (2024). An Improved Inferential Procedure to Evaluate Item Discriminations in a Conditional Maximum Likelihood Framework. *Journal of Educational and Behavioral Statistics*, 49(3), 403-430.

See Also

`invar_test`, `change_test`, and `LLTM_test`.

Examples

```
## Not run:
##### Dataset PISA Mathematics data.pisaMath {sirt} #####

library(sirt)
data(data.pisaMath)
y <- data.pisaMath$data[, grep(names(data.pisaMath$data), pattern = "M" )]

res <- discr_test(data = y)
# $test
#      W     LR     RS     GR
# 72.470 73.032 76.725 73.430
#
# $df
# W LR RS  GR
# 10 10 10 10
#
# $pvalue
#       W         LR         RS          GR
# "< 0.001" "< 0.001" "< 0.001" "< 0.001"
#
# $call
# discr_test(X = y)
```

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

`get_eRm_arg` *Extract Arguments from an eRm Object*

Description

This function extracts specific arguments from an object of class ''LR'' from the 'eRm' package. Depending on the selected argument, it retrieves degrees of freedom ('df'), local deviations ('local_dev'), or informative sample size ('n_info').

Usage

```
get_eRm_arg(obj, arg = c("df", "local_dev, n_info"))
```

Arguments

`obj`	An object of class ''LR'', typically created using functions from the 'eRm' package.
`arg`	A character string specifying the argument to extract. Options are: • '"df"' (default): Extracts the degrees of freedom. • '"local_dev"': Extracts item parameters for the two person groups from the model. If more than two split groups are available, only the first two are selected. • '"n_info"': Computes and returns the informative sample size using 'n_info()'.

Details

If multiple argument values are provided, '"df"' is selected by default. If an invalid 'arg' is provided, the function throws an error.

Value

The extracted argument value:

- A numeric value if 'arg = "df"'.
- A list containing local deviation parameters if 'arg = "local_dev"'.
- A computed sample size if 'arg = "n_info"'.

Note

If 'obj' contains more than two split groups, only the first two will be selected for '"local_dev"', with a message notifying the user.

Examples

```
## Not run:
  # Example usage with an LR object
  dat = eRm::sim.rasch(1000,10)
  mod = eRm::RM(dat)

  obj <- eRm::LRtest(mod) # Create an LR object
  get_eRm_arg(obj, "df")      # Extract degrees of freedom
  get_eRm_arg(obj, "local_dev")  # Extract local deviations
  get_eRm_arg(obj, "n_info")  # Extract informative sample size

## End(Not run)
```

`invar_test` *Test of invariance of item parameters between two groups.*

Description

Computes Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test statistics for hypothesis of equality of item parameters between two groups of persons against a two-sided alternative that at least one item parameter differs between the two groups.

Usage

```
invar_test(data, splitcr = "median", model = "RM")
```

Arguments

`data` Data matrix.

`splitcr` Split criterion which is either "mean", "median" or a numeric vector x.

- **"mean"** Corresponds to division of the sample according to the mean of the person score.
- **"median"** Corresponds to division of the sample according to the median of the person score.
- **x** Has length equal to number of persons and contains zeros and ones. It indicates group membership for every person.

`model` RM, PCM, RSM

Details

Note that items are excluded for the computation of GR, LR, and W due to inappropriate response patterns within subgroups and for computation of RS due to inappropriate response patterns in the total data. If the model is identified from the total data but not from one or both subgroups only RS will be computed. If the model is not identified from the total data, no test statistic is computable.

Value

A list of class `tcl` of test statistics, degrees of freedom, and p-values.

`test`	A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test statistics.
`df`	A numeric vector of corresponding degrees of freedom.
`pvalue`	A vector of corresponding p-values.
`deleted_items`	A list with numeric vectors of item numbers that were excluded before computing corresponding test statistics.
`sample_size_informative`	Informative sample size of data omitting persons with min. and max score.
`effect`	Numeric value for each test representing the effect size. A real number between 0 and 1, interpreted as a proportion of pseudo -variance between the two groups of persons considered.
`data`	Data matrix.
`call`	The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

`change_test`, and `LLTM_test`.

Examples

```
## Not run:
##### Rasch Model #####
y <- eRm::sim.rasch(persons = rnorm(400), c(0,-3,-2,-1,0,1,2,3))
x <- c(rep(1,200),rep(0,200))

res <- invar_test(data = y, splitcr = x, model = "RM")
```

```
res$test # test statistics
res$df # degrees of freedoms
res$pvalue # p-values
res$deleted_items # excluded items

$test
     W     LR     RS     GR
14.972 14.083 13.678 12.492

$df
W LR RS GR
7  7  7  7

$pvalue
    W       LR       RS      GR
"0.073" "0.050" "0.057" "0.043"

$deleted_items
 $deleted_items$GR
 [1] "none"

 $deleted_items$LR
 [1] "none"

 $deleted_items$RS
 [1] "none"

 $deleted_items$W
 [1] "none"

$sample_size_informative
[1] 395

$effect
    W    LR    RS    GR
0.014 0.014 0.014 0.014

$call
invar_test(X = y, splitcr = x, model = "RM")


## End(Not run)
```

`LLTM_test` *Testing linear restrictions on parameter space of item parameters of RM.*

Description

Computes Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test statistics for hypotheses defined by linear restrictions on parameter space of the item parameters of RM.

Usage

```
LLTM_test(data, W)
```

Arguments

data	Data matrix.
W	Design matrix of LLTM.

Details

The RM item parameters are assumed to be linear in the LLTM parameters. The coefficients of the linear functions are specified by a design matrix W. In this context, the LLTM is considered as a more parsimonious model than the RM. The LLTM parameters can be interpreted as the difficulties of certain cognitive operations needed to respond correctly to psychological test items. The item parameters of the RM are assumed to be linear combinations of these cognitive operations. These linear combinations are defined in the design matrix W.

Value

A list of class `tcl` of test statistics, degrees of freedom, and p-values.

test	A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test statistics.
df	Degrees of freedom.
pvalue	A vector of corresponding p-values.
data	Data matrix.
call	The matched call.

References

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch models: Foundations, Recent Developments, and Applications (pp. 131-155). New York: Springer.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. Psychometrika, 48(1), 3-26.

See Also

`change_test`, and `invar_test`.

Examples

```
## Not run:
# Numerical example assuming no deviation from linear restriction

# design matrix W defining linear restriction
W <- rbind(c(1,0), c(0,1), c(1,1), c(2,1))
```

```
# assumed eta parameters of LLTM for data generation
eta <- c(-0.5, 1)

# assumed vector of item parameters of RM
b <- colSums(eta * t(W))

y <- eRm::sim.rasch(persons = rnorm(400), items = b - b[1])  # sum0 = FALSE

res <- LLTM_test(data = y, W = W )

res$test # test statistics
res$df # degrees of freedoms
res$pvalue # p-values


## End(Not run)
```

---

mix_mod *Mixed model considering the effects of multiple covariates .*

Description

Estimates and tests linear effects of multiple covariates on item parameters of the Rasch model (RM) simultaneously.

Usage

```
mix_mod(data, Xcov)
```

Arguments

data	Data matrix consisting of binary responses, i.e., 0s and 1s. Missing responses are NAs.
Xcov	Covariate matrix. Persons in rows and covariates in columns, e.g., age, gender, drug dosage, etc. In case of one covariate Xcov must be a one-column matrix.

Details

The underlying model is a mixed-effects logit model with random person effects and fixed item and covariates effects, i.e.,

$$\log \frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)} = \tau_i + \alpha_j + \sum_{p=1}^{q} x_{ip}\delta_{jp}, \quad i=1,\dots,n,\ j=1,\dots,k,\ p=1,\dots,q,$$

where $Y_{ij} \in \{0,1\}$, τ_i is a person parameter, α_j is a baseline effect of item j, δ_{jp} is an effect of covariate p on item j, and x_{ip} is a covariate value observed for person i and covariate p. For identifiability, $\alpha_1 = 0$, $\delta_{1p} = 0 \ \forall p$. Setting all δ parameters $(\forall j, p)$ to 0 yields the RM as a special case (with the αs as the item parameters of the RM).

The α and δ parameters are estimated using a conditional maximum likelihood (CML) approach and four different tests based on the conditional likelihood and derived from asymptotic theory are provided, i.e., likelihood ratio (LR), Rao score (RS), Wald (W), and gradient (GR) test. The hypothesis of interest is $\delta_{jp} = 0 \ \forall j, p$ against the alternative that at least one δ parameter is not equal to 0. Furthermore, Z test statistics (i.e., standard normal distribution when the true effect of a covariate on an item is 0) are computed for each item and covariate separately.

Value

A list of class `tcl` with the following components:

`CML_estimates`	Conditional maximum likelihood (CML) estimates of item (α and δ) parameters (easiness, attractiveness). The effects of the first item are set to 0 for identifiability.
`SE`	Standard errors of CML estimates.
`Z_statistics`	Z test statistics for each single parameter (α and δ), i.e., testing the hypothesis that the true value of the respective parameter is 0 against the alternative of $\neq 0$.
`pvalue`	A matrix of two-sided p-values for the Z tests.
`loglik`	Conditional log-likelihood.
`tests`	A table summarizing the results of four tests (W, LR, RS, GR) of the hypothesis that the effects of all covariates on all items are all 0. The table contains the test statistic (`stat`), degrees of freedom (`df`), and two-sided p-value (`pvalue`) for each test.
`information_criteria`	AIC, BIC
`call`	The matched call.

References

Draxler, C., & Kurz, A. (2025). Testing measurement invariance in a conditional likelihood framework by considering multiple covariates simultaneously. Behavior Research Methods, 57(1), 50.

See Also

`invar_test`, `change_test`, and `LLTM_test`.

Examples

```
## Not run:
##### Rasch Model #####
dat <- eRm::raschdat3
x1 <- c(rep(0,250), rep(1,250))
x2 <- runif(500,min = 0, max = 1)
X <- cbind(x1,x2)

res <- mix_mod(data = dat, Xcov = X)
```

```
# $CML_estimates
#      1      2      3      4      5      6
# base 0 -0.596 -1.152 -1.804 -1.846 -2.353
# 1    0 -0.380 -0.403  0.072 -0.121 -0.452
# 2    0  0.814  0.780  0.612 -0.277  0.069
#
# $SE
#       1     2     3     4     5     6
# base NA 0.356 0.347 0.349 0.353 0.369
# 1    NA 0.314 0.303 0.301 0.307 0.320
# 2    NA 0.564 0.545 0.541 0.548 0.572
#
# $Z_statistics
#       1      2      3      4      5      6
# base NA -1.675 -3.320 -5.175 -5.223 -6.377
# 1    NA -1.210 -1.330  0.240 -0.396 -1.413
# 2    NA  1.443  1.431  1.132 -0.505  0.121
#
# $pvalue
#       1     2     3     4     5     6
# base NA 0.094 0.001 0.000 0.000 0.000
# 1    NA 0.226 0.183 0.810 0.692 0.158
# 2    NA 0.149 0.153 0.258 0.613 0.904
#
# $loglik
# [1] -993.8575
#
# $tests
#      stat df pvalue
# W  14.339 10  0.158
# LR 14.462 10  0.153
# RS 14.507 10  0.151
# GR 14.499 10  0.151
#
# $information_criteria
#           AIC      BIC
# [1,] 2007.715 2049.432
#
# $call
# mix_mod(data = dat, Xcov = X)
#
# attr(,"class")
# [1] "tcl"

## End(Not run)
```

---

opt_n *Computes the optimal sample size for item parameter invariance tests.*

Description

Computes the informative sample size given an effect of interest and type I and II error probabilities (alpha and beta) for Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test. The routine supports two modes: Either provide the return object of a previous call to `invar_test()` or provide the effect size of interest along with the degrees of freedom.

Usage

```
opt_n(
  invar_obj = NULL,
  effect = NULL,
  df = NULL,
  alpha = 0.05,
  beta = 0.05,
  n_range = 10:10000
)
```

Arguments

`invar_obj`	Return object of a previous call to `invar_test()`. Default is `NULL`. If missing, values for `effect` and `df` need to be set manually.
`effect`	Numeric value representing the effect size. A real number between 0 and 1, interpreted as a proportion of pseudo-variance between persons with different covariate values (but the same person parameter). Default is `NULL`.
`df`	Degrees of freedom of the test. Default is `NULL`.
`alpha`	Type I error probability. Default is 0.05.
`beta`	Type II error probability. Default is 0.05.
`n_range`	A numeric vector specifying the sample sizes to be evaluated. Default is `10:10000`).

Details

The informative sample size is the number of observations realizing a score greater than zero and less than the maximum possible score, as these two values are not informative for the tests.

Providing the return object of a previous call to `invar_test()` allows using the results of a pilot study to obtain an empirical estimate of parameter differences between the groups.

The default search range of `10:10000` should suffice for most applications. However, if the maximum is reached, a warning is given.

If `effect` and `df` are provided, the sample sizes of all four tests will be equal due to their asymptotic equivalence. If an `invar_obj` is provided, the sample sizes will usually differ slightly.

Note: The `invar_test()` function currently only supports a two-group split.

Value

A list of two elements:

`opt_n`	The required sample sizes for the four tests.

real_pow	The realized power, as the sample sizes are rounded to the next integer.
call	The matched call.

References

Draxler, C., & Kurz, A. (2025). Testing measurement invariance in a conditional likelihood framework by considering multiple covariates simultaneously. *Behavior Research Methods*, 57(1), 50.

See Also

invar_test, p_curve, p_ncurve

Examples

```
## Not run:
# --- a priori mode:

  opt_n(effect=0.3,df=20)       # n=102

  opt_n(effect=0.001,df=300)    # Warning!
  opt_n(effect=0.001,df=300,n_range=1000:100000) # Warning disappears, n=91087

# --- pilot sample mode:

  library(eRm)
  opt_n(invar_test(raschdat1))

# --- typical problem: items eliminated

  ex2 = invar_test(pcmdat,model="PCM")

# The following items were excluded for the computation of GR,LR, and W
# due to inappropriate response patterns within subgroups:
I2 I4 I1 I5

> opt_n(ex2)

# Parameters:
# alpha = 0.05
# beta = 0.05
# power = 0.95
# df = 7 7 19 7                 # note the different df!

# Observed effects
#     GR     LR     RS      W
# 0.1295 0.1284 0.8462 0.1226   # note the effect differences!

# Optimal Sample Size
#  GR  LR  RS   W
# 169 170  36 178               # note the different sample sizes!
#
# Realized Power
```

```
#    GR    LR    RS     W
# 0.950 0.950 0.952 0.950

## End(Not run)
```

---

post_hocChange   *Power analysis of tests in context of measurement of change using LLTM*

Description

Returns post hoc power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given data and probability of error of first kind α. The hypothesis to be tested states that the shift parameter quantifying the constant change for all items between time points 1 and 2 equals 0. The alternative states that the shift parameter is not equal to 0. It is assumed that the same items are presented at both time points. See function `change_test`.

Usage

```
post_hocChange(data, alpha = 0.05)
```

Arguments

`data`	Data matrix as required for function `change_test`.
`alpha`	Probability of error of first kind.

Details

The power of the tests (Wald, LR, score, and gradient) is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with $df = 1$ and noncentrality parameter λ. In case of evaluating the post hoc power, λ is assumed to be given by the observed value of the test statistic. Given the probability of the error of the first kind α the post hoc power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, let $q_{1-\alpha}$ be the $1-\alpha$ quantile of the central χ^2 distribution with df = 1. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with $df = 1$ and λ equal to the observed value of the test statistic.

Value

A list of results of class `tcl_post_hoc`.

- `test` A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test statistics.
- `power` Posthoc power value for each test.
- `dev_obs` CML estimate of shift parameter expressing observed deviation from the hypothesis to be tested.
- `score_dist` Relative frequencies of person scores. Uninformative scores, i.e., minimum and maximum scores, are omitted. Note that the person score distribution also influences the power of the tests.
- `df` Degrees of freedom df.
- `ncp` Noncentrality parameter λ of χ^2 distribution from which power is determined. It equals the observed value of the test statistic.
- `call` The matched call.

References

Draxler, C., & Alexandrowicz, R. W. (2015). Sample size determination within the scope of conditional maximum likelihood estimation with special focus on testing the Rasch model. Psychometrika, 80(4), 897-919.

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch models: Foundations, Recent Developments, and Applications (pp. 131-155). New York: Springer.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. Psychometrika, 48(1), 3-26.

See Also

`sa_sizeChange`, and `powerChange`.

Examples

```
## Not run:
# Numerical example with 200 persons and 4 items
# presented twice, thus 8 virtual items

# Data y generated under the assumption that shift parameter equals 0.5
# (change from time point 1 to 2)

# design matrix W used only for exmaple data generation
#     (not used for estimating in change_test function)
W <- rbind(c(1,0,0,0,0), c(0,1,0,0,0), c(0,0,1,0,0), c(0,0,0,1,0),
           c(1,0,0,0,1), c(0,1,0,0,1), c(0,0,1,0,1), c(0,0,0,1,1))

# eta parameter vector, first 4 are nuisance, i.e., item parameters at time point 1.
# (easiness parameters of the 4 items at time point 1),
# last one is the shift parameter
```

```
eta <- c(-2,-1,1,2,0.5)

y <- eRm::sim.rasch(persons=rnorm(150), items=colSums(-eta*t(W)))

res <- post_hocChange(data = y, alpha = 0.05)

# > res
# $test
#     W     LR     RS     GR
# 9.822 10.021  9.955 10.088
#
# $power
#     W    LR    RS    GR
# 0.880 0.886 0.884 0.888
#
# $dev_obs #`observed deviation (estimate of shift parameter)`
# [1] 0.504
#
# $score_dist #`person score distribution`
#
#     1     2     3     4     5     6     7
# 0.047 0.047 0.236 0.277 0.236 0.108 0.047
#
# $df #`degrees of freedom`
# [1] 1
#
# $ncp # `noncentrality parameter`
#     W     LR     RS     GR
# 9.822 10.021  9.955 10.088
#
# $call
# post_hocChange(alpha = 0.05, data = y)

## End(Not run)
```

`post_hocPCM` *Power analysis of tests of invariance of item parameters between two groups of persons in partial credit model*

Description

Returns post hoc power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given data and probability of error of first kind α. The hypothesis to be tested assumes equal item-category parameters of the partial credit model between two predetermined groups of persons. The alternative states that at least one of the parameters differs between the two groups.

Usage

```
post_hocPCM(data, splitcr, alpha = 0.05)
```

Arguments

`data`	Data matrix with item responses (in ordered categories starting from 0).
`splitcr`	A numeric vector of length equal to number of persons that contains zeros and ones indicating group membership of the persons.
`alpha`	Probability of error of first kind.

Details

The power of the tests (Wald, LR, score, and gradient) is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with df equal to the number of free item-category parameters in the partial credit model and noncentrality parameter λ. In case of evaluating the post hoc power, λ is assumed to be given by the observed value of the test statistic. Given the probability of the error of the first kind α the post hoc power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, let $q_{1-\alpha}$ be the $1-\alpha$ quantile of the central χ^2 distribution with df equal to the number of free item-category parameters. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with df equal to the number of free item-category parameters and λ equal to the observed value of the test statistic.

Value

A list of results of class `tcl_post_hoc`.

`test`	A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test statistics.
`power`	Post hoc power value for each test.
`dev_global`	Observed global deviation from the hypothesis to be tested, represented by a single number. It is obtained by dividing the test statistic by the informative sample size, which excludes persons with minimum or maximum person scores.
`dev_local`	CML estimates of free item-category parameters in both groups of persons, representing observed deviation from the hypothesis to be tested locally per item and response category.
`score_dist_group1`	Relative frequencies of person scores in group 1. Uninformative scores, i.e., minimum and maximum scores, are omitted. Note that the person score distribution also influences the power of the tests.
`score_dist_group2`	Relative frequencies of person scores in group 2. Uninformative scores, i.e., minimum and maximum scores, are omitted. Note that the person score distribution also influences the power of the tests.
`df`	Degrees of freedom df.

`ncp` Noncentrality parameter λ of the χ^2 distribution from which power is determined. It equals the observed value of the test statistic.

`call` The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

`sa_sizePCM`, and `powerPCM`.

Examples

```
## Not run:
# Numerical example for post hoc power analysis for PCM

y <- eRm::pcmdat2
n <- nrow(y) # sample size
x <- c( rep(0,n/2), rep(1,n/2) ) # binary covariate

res <- post_hocPCM(data = y, splitcr = x, alpha = 0.05)

# > res
# $test
#      W     LR     RS     GR
# 11.395 11.818 11.628 11.978
#
# $power
#     W    LR    RS    GR
# 0.683 0.702 0.694 0.709
#
# $dev_global #`observed global deviation`
#     W    LR    RS    GR
# 0.045 0.046 0.045 0.047
#
```

```
# $ dev_local #`observed local deviation`
#        I1-C2 I2-C1 I2-C2  I3-C1  I3-C2  I4-C1  I4-C2
# group1 2.556 0.503 2.573 -2.573 -2.160 -1.272 -0.683
# group2 2.246 0.878 3.135 -1.852 -0.824 -0.494  0.941
#
# $score_dist_group1 #`person score distribution in group 1`
#
#     1     2     3     4     5     6     7
# 0.016 0.097 0.137 0.347 0.121 0.169 0.113
#
# $score_dist_group2 #`person score distribution in group 2`
#
#     1     2     3     4     5     6     7
# 0.015 0.083 0.136 0.280 0.152 0.227 0.106
#
# $df #`degrees of freedom`
# [1] 7
#
# $ncp #`noncentrality parameter`
#      W     LR     RS     GR
# 11.395 11.818 11.628 11.978
#
# $call
# post_hocPCM(alpha = 0.05, data = y, x = x)

## End(Not run)
```

---

post_hocRM — *Power analysis of tests of invariance of item parameters between two groups of persons in binary Rasch model*

Description

Returns post hoc power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given data and probability of error of first kind α. The hypothesis to be tested assumes equal item parameters between two predetermined groups of persons. The alternative states that at least one of the parameters differs between the two groups.

Usage

```
post_hocRM(data, splitcr, alpha = 0.05)
```

Arguments

data	Binary data matrix.
splitcr	A numeric vector of length equal to number of persons containing zeros and ones indicating group membership of the persons.
alpha	Probability of error of first kind.

Details

The power of the tests (Wald, LR, score, and gradient) is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with df equal to the number of items minus 1 and noncentrality parameter λ. In case of evaluating the post hoc power, λ is assumed to be given by the observed value of the test statistic. Given the probability of the error of the first kind α the post hoc power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, let $q_{1-\alpha}$ be the $1-\alpha$ quantile of the central χ^2 distribution with df equal to the number of items minus 1. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with df equal to the number of items reduced by 1 and λ equal to the observed value of the test statistic.

Value

A list of results of class `tcl_post_hoc`.

`test` A numeric vector of Wald (W), likelihood ratio (LR), Rao score (RS), and gradient (GR) test statistics.

`power` Post hoc power value for each test.

`dev_global` Observed global deviation from hypothesis to be tested represented by a single number. It is obtained by dividing the test statistic by the informative sample size. The latter does not include persons with minimum or maximum person score.

`dev_local` CML estimates of free item parameters in both groups of persons (first item parameter set to 0 in both groups) representing observed deviation from hypothesis to be tested locally per item.

`score_dist_group1` Relative frequencies of person scores in group 1. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`score_dist_group2` Relative frequencies of person scores in group 2. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`df` Degrees of freedom df.

`ncp` Noncentrality parameter λ of χ^2 distribution from which power is determined. It equals observed value of test statistic.

`call` The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2020). The Gradient Test and its Finite Sample Size Properties in a Conditional Maximum Likelihood and Psychometric Modeling Context. Communications in Statistics-Simulation and Computation, 1-19.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

sa_sizeRM, and powerRM.

Examples

```
## Not run:
# Numerical example for post hoc power analysis for Rasch Model

y <- eRm::raschdat1
n <- nrow(y) # sample size
x <- c( rep(0,n/2), rep(1,n/2) ) # binary covariate

res <-  post_hocRM(data = y, splitcr = x, alpha = 0.05)

# > res
# $test
#      W     LR     RS     GR
# 29.241 29.981 29.937 30.238
#
# $power
#     W    LR    RS    GR
# 0.890 0.900 0.899 0.903
#
# $dev_global #`observed global deviation`
#     W    LR    RS    GR
# 0.292 0.300 0.299 0.302
#
# $dev_local #`observed local deviation`
#           I2    I3    I4    I5    I6    I7    I8    I9   I10   I11
# group1 1.039 0.693 2.790 2.404 1.129 1.039 0.864 1.039 2.790 2.244
# group2 2.006 0.945 2.006 3.157 1.834 0.690 0.822 1.061 2.689 2.260
#          I12   I13   I14   I15   I16   I17   I18   I19   I20   I21
```

```
# group1 1.412 3.777 3.038 1.315 2.244 1.039 1.221 2.404 0.608 0.608
# group2 0.945 2.962 4.009 1.171 2.175 1.472 2.091 2.344 1.275 0.690
#          I22   I23   I24   I25   I26   I27   I28   I29   I30
# group1 0.438 0.608 1.617 3.038 0.438 1.617 2.100 2.583 0.864
# group2 0.822 1.275 1.565 2.175 0.207 1.746 1.746 2.260 0.822
#
# $score_dist_group1 #`person score distribution in group 1`
#
#    1    2    3    4    5    6    7    8    9   10   11   12   13
# 0.02 0.02 0.02 0.06 0.02 0.10 0.10 0.06 0.10 0.12 0.08 0.12 0.12
#   14   15   16   17   18   19   20   21   22   23   24   25   26
# 0.06 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
#   27   28   29
# 0.00 0.00 0.00
#
# $score_dist_group2 #`person score distribution in group 2`
#
#    1    2    3    4    5    6    7    8    9   10   11   12   13
# 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
#   14   15   16   17   18   19   20   21   22   23   24   25   26
# 0.08 0.12 0.10 0.16 0.06 0.04 0.10 0.12 0.08 0.02 0.02 0.02 0.08
#   27   28   29
# 0.00 0.00 0.00
#
# $df #`degrees of freedom`
# [1] 29
#
# $ncp #`noncentrality parameter`
#      W     LR     RS     GR
# 29.241 29.981 29.937 30.238
#
# $call
# post_hocRM(alpha = 0.05, data = y, x = x)


## End(Not run)
```

---

powerChange — *Power analysis of tests in context of measurement of change using LLTM*

Description

Returns power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probability of error of first kind α, sample size, and a deviation from the hypothesis to be tested. The latter states that the shift parameter quantifying the constant change for all items between time points 1 and 2 equals 0. The alternative states that the shift parameter is not equal to 0. It is assumed that the same items are presented at both time points. See function `change_test`.

Usage

```
powerChange(n_total, eta, alpha = 0.05, persons = rnorm(10^6))
```

Arguments

`n_total`	Total sample size for which power shall be determined.
`eta`	A vector of eta parameters of the LLTM. The last element represents the constant change or shift for all items between time points 1 and 2. The other elements of the vector are the item parameters at time point 1. A choice of the eta parameters constitutes a scenario of deviation from the hypothesis of no change.
`alpha`	Probability of the error of first kind.
`persons`	A vector of person parameters (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the power of the tests is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with $df = 1$ and noncentrality parameter λ. The latter depends on a scenario of deviation from the hypothesis to be tested and a specified sample size. Given the probability of the error of the first kind α the power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

As regards the concept of sample size a distinction between informative and total sample size has to be made since the power of the tests depends only on the informative sample size. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value of the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons.

In particular, the determination of λ and the power of the tests, respectively, is based on a simple Monte Carlo approach. Data (responses of a large number of persons to a number of items presented at two time points) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. The hypothesis to be tested assumes no change between time points 1 and 2. A scenario of a deviation is given by a choice of the item parameters at time point 1 and the shift parameter, i.e., the LLTM eta parameters, as well as the person parameters (to be drawn randomly from a specified distribution). The shift parameter represents a constant change of all item parameters from time point 1 to time point 2. A test statistic T (Wald, LR, score, or gradient) is computed from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. The power of the tests can be determined given a user-specified total sample size denoted by n_{total}. The noncentrality parameter λ can then be expressed by $\lambda = n_{total}*(n_{infsim}/n_{totalsim})*e$, where $n_{totalsim}$ denotes the total number of persons in the simulated data and $n_{infsim}/n_{totalsim}$ is the proportion of informative persons in the sim. data. Let $q_{1-\alpha}$ be the $1-\alpha$ quantile of the central χ^2 distribution with $df = 1$. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with $df = 1$ and $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$. Thereby, it is assumed that n_{total} is composed of a frequency distribution of person scores that is proportional to the observed distribution of person scores in the simulated data.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realized value of a random variable E. The same holds true for λ as well as the power of the tests. Thus, the power is a realized value of a random variable that shall be denoted by P. Consequently, the (realized) value of the power of the tests need not be equal to the exact power that follows from the user-specified n_{total}, α, and the chosen item parameters and shift parameter used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters the power of the tests will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing the power of the tests can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution. Thus, $Var(T) = 2(df + 2\lambda)$, with $df = 1$ and $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. The power of the tests is a function of e which is given by $F_{df,\lambda}(q_\alpha)$, where $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$ and $df = 1$. Then, by the delta method one obtains (for the variance of P)

$$Var(P) = Var(E) * (F'_{df,\lambda}(q_\alpha))^2,$$

where $F'_{df,\lambda}$ is the derivative of $F_{df,\lambda}$ with respect to e. This derivative is determined numerically and evaluated at e using the package numDeriv. The square root of $Var(P)$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of power.

Value

A list of results of class `tcl_power`.

`power`	Power value for each test.
`mc_err_power`	Monte Carlo error of power computation for each test.
`dev_est_shift`	Shift parameter estimated from the simulated data, representing the constant shift of item parameters between time points 1 and 2.
`score_dist`	Relative frequencies of person scores observed in simulated data. Uninformative scores, i.e., minimum and maximum scores, are omitted. Note that the person score distribution also has an influence on the power of the tests.

df	Degrees of freedom (df).
ncp	Noncentrality parameter (λ) of the χ^2 distribution from which power is determined.
call	The matched call.

References

Draxler, C., & Alexandrowicz, R. W. (2015). Sample size determination within the scope of conditional maximum likelihood estimation with special focus on testing the Rasch model. Psychometrika, 80(4), 897-919.

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch models: Foundations, Recent Developments, and Applications (pp. 131-155). New York: Springer.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. Psychometrika, 48(1), 3-26.

See Also

sa_sizeChange, and post_hocChange.

Examples

```
## Not run:

# Numerical example: 4 items presented twice, thus 8 virtual items

# eta Parameter, first 4 are nuisance
# (easiness parameters of the 4 items at time point 1),
# last one is the shift parameter
eta <- c(-2,-1,1,2,0.5)
res <- powerChange(n_total = 150, eta = eta, persons=rnorm(10^6))

# > res
# $power
#     W    LR    RS    GR
# 0.905 0.910 0.908 0.911
#
# $mc_err_power #`MC error of power`
#     W    LR    RS    GR
# 0.002 0.002 0.002 0.002
#
# $dev_est_shift #`deviation (estimate of shift parameter)`
# [1] 0.499
#
# $score_dist #`person score distribution`
#
#     1     2     3     4     5     6     7
# 0.034 0.093 0.181 0.249 0.228 0.147 0.068
#
# $df #`degrees of freedom`
```

```
# [1] 1
#
# $ncp #`noncentrality parameter`
#      W     LR     RS     GR
# 10.692 10.877 10.815 10.939
#
# $call
# powerChange(alpha = 0.05, n_total = 150, eta = eta, persons = rnorm(10^6))
#

## End(Not run)
```

---

powerPCM *Power analysis of tests of invariance of item parameters between two groups of persons in partial credit model*

Description

Returns power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probability of error of first kind α, sample size, and a deviation from the hypothesis to be tested. The hypothesis to be tested assumes equal item-category parameters of the partial credit model between two predetermined groups of persons. The alternative states that at least one of the parameters differs between the two groups.

Usage

```
powerPCM(
  n_total,
  obj = NULL,
  local_dev = NULL,
  alpha = 0.05,
  persons1 = rnorm(10^6),
  persons2 = rnorm(10^6)
)
```

Arguments

`n_total` Total sample size for which power shall be determined.

`obj` An object of class ‘LR‘ from the ‘eRm‘ package. If provided, ‘local_dev‘ is extracted automatically. If missing, ‘local_dev‘ must be set manually.

`local_dev` A list consisting of two lists. One list refers to group 1, the other to group 2. Each of the two lists contains a numeric vector per item, i.e., each list contains as many vectors as items. Each vector contains the free item-cat. parameters of the respective item. The number of free item-cat. parameters per item equals the number of categories of the item minus 1.

`alpha` Probability of error of first kind.

`persons1` A vector of person parameters in group 1 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

`persons2` A vector of person parameters in group 2 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the power of the tests is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with df equal to the number of free item-category parameters and noncentrality parameter λ. The latter depends on a scenario of deviation from the hypothesis to be tested and a specified sample size. Given the probability of the error of the first kind α the power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

As regards the concept of sample size a distinction between informative and total sample size has to be made since the power of the tests depends only on the informative sample size. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value of the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons.

In particular, the determination of λ and the power of the tests, respectively, is based on a simple Monte Carlo approach. Data (responses of a large number of persons to a number of items) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. A scenario of a deviation is given by a choice of the item-cat. parameters and the person parameters (to be drawn randomly from a specified distribution) for each of the two groups. Such a scenario may be called local deviation since deviations can be specified locally for each item-category. The relative group sizes are determined by the choice of the number of person parameters for each of the two groups. For instance, by default 10^6 person parameters are selected randomly for each group. In this case, it is implicitly assumed that the two groups of persons are of equal size. The user can specify the relative group sizes by choosing the length of the arguments persons1 and persons2 appropriately. Note that the relative group sizes do have an impact on power and sample size of the tests. The next step is to compute a test statistic T (Wald, LR, score, or gradient) from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. The power of the tests can be determined given a user-specified total sample size denoted by `n_total`. The noncentrality parameter λ can then be expressed by $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$, where $n_{totalsim}$ denotes the total number of persons in the simulated data and $n_{infsim}/n_{totalsim}$ is the proportion of informative persons in the sim. data. Let $q_{1-\alpha}$ be the $1 - \alpha$ quantile of the central χ^2 distribution with df equal to the number of free item-category parameters. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with df equal to the number of free item-category parameters and $\lambda = n_{total}(n_{infsim}/n_{totalsim}) * e$. Thereby, it is assumed that n_{total} is composed of a frequency distribution of person scores that is proportional to the observed distribution of person scores in the simulated data. The same holds true in respect of the relative group sizes, i.e., the relative frequencies of the two person groups in a sample of size n_{total} are assumed to be equal to the relative frequencies of the two groups in the simulated data.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realized value of a random variable E. The same holds true for λ as well as the power of the tests. Thus, the power is a realized value of a random variable that shall be denoted by P. Consequently, the (realized) value of the power of the tests need not be equal to the exact power that follows from the user-specified n_{total}, α, and the chosen item-category parameters used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters the power of the tests will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing the power of the tests can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution with df equal to the number of free item-category parameters and noncentrality parameter λ. Thus, $Var(T) = 2(df + 2\lambda)$, with $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. The power of the tests is a function of e which is given by $F_{df,\lambda}(q_\alpha)$, where $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$ and df equal to the number of free item-category parameters. Then, by the delta method one obtains (for the variance of P).

$$Var(P) = Var(E) * (F'_{df,\lambda}(q_\alpha))^2,$$

where $F'_{df,\lambda}$ is the derivative of $F_{df,\lambda}$ with respect to e. This derivative is determined numerically and evaluated at e using the package numDeriv. The square root of $Var(P)$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of power.

Value

A list of results of class `tcl_power`.

`power`	Power value for each test.
`mc_error_power`	Monte Carlo error of power computation for each test.
`dev_global`	Global deviation computed from simulated data for each test. See Details.

`dev_local` CML estimates of free item-category parameters in both groups of persons obtained from the simulated data expressing a deviation from the hypothesis to be tested locally per item and response category.

`score_dist_group1`
Relative frequencies of person scores in group 1 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`score_dist_group2`
Relative frequencies of person scores in group 2 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`df` Degrees of freedom df.

`ncp` Noncentrality parameter λ of χ^2 distribution from which power is determined.

`call` The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

`sa_sizePCM`, and `post_hocPCM`.

Examples

```
## Not run:
#Numerical example of power analysis for the PCM model

# free item-category parameters for group 1 and 2  with 5 items, with 3 categories each
local_dev <-  list (  list(c( 0, 0), c( -1, 0), c( 0, 0),  c( 1, 0), c( 1, 0.5)) ,
                      list(c( 0, 0), c( -1, 0), c( 0, 0),  c( 1, 0), c( 0, -0.5))  )
```

```
res <-  powerPCM(n_total = 200, local_dev = local_dev)

# > res
# $power
#     W    LR    RS    GR
# 0.863 0.885 0.876 0.892
#
# $mc_error_power #`MC error of power`
#     W    LR    RS    GR
# 0.002 0.002 0.002 0.002
#
# $dev_global #`global deviation`
#     W    LR    RS    GR
# 0.102 0.107 0.105 0.109
#
# $dev_local #`local deviation`
#         I1-C2  I2-C1  I2-C2  I3-C1  I3-C2 I4-C1 I4-C2  I5-C1  I5-C2
# group1  0.002 -0.997 -0.993  0.006  0.012 1.002 1.007  1.006  1.508
# group2 -0.007 -1.005 -1.007 -0.006 -0.009 0.993 0.984 -0.006 -0.510
#
# $score_dist_group1 # `person score distribution in group 1`
#
#     1     2     3     4     5     6     7     8     9
# 0.112 0.130 0.131 0.129 0.122 0.114 0.101 0.091 0.070
#
# $score_dist_group2 #`person score distribution in group 2`
#
#     1     2     3     4     5     6     7     8     9
# 0.091 0.108 0.117 0.122 0.122 0.121 0.115 0.110 0.093
#
# $df #`degrees of freedom`
# [1] 9
#
# $ncp #`noncentrality parameter`
#      W     LR     RS     GR
# 18.003 19.024 18.596 19.403
#
# $call
# powerPCM(alpha = 0.05, n_total = 200, persons1 = rnorm(10^6),
#          persons2 = rnorm(10^6), local_dev = local_dev)


# Numerical example of power analysis for the PCM model
# extracting local_dev from an eRm object
ppar = rnorm(10000)
ipar = list(-2:2,-2:2+0.2,-2:2-0.3,-2:2+0.1,-2:2-0.8)
dat2 = psychotools::rpcm(theta = ppar, delta = ipar)

mod2 = PCM(dat2$data)
obj2 = eRm::LRtest(mod2)

res <- powerPCM(n_total = 200, obj = obj2)
```

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

---

powerRM — *Power analysis of tests of invariance of item parameters between two groups of persons in binary Rasch model*

Description

Returns power of Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probability of error of first kind α, sample size, and a deviation from the hypothesis to be tested. The latter assumes equality of the item parameters in the Rasch model between two predetermined groups of persons. The alternative states that at least one of the parameters differs between the two groups.

Usage

```
powerRM(
  n_total,
  obj = NULL,
  local_dev = NULL,
  alpha = 0.05,
  persons1 = rnorm(10^6),
  persons2 = rnorm(10^6)
)
```

Arguments

`n_total`	Total sample size for which power shall be determined.
`obj`	An object of class 'LR' from the 'eRm' package. If provided, 'local_dev' is extracted automatically. If missing, 'local_dev' must be set manually.
`local_dev`	A list of two vectors containing item parameters for the two person groups representing a deviation from the hypothesis to be tested locally per item. Note that the 'reference category', i.e. the first item parameter, also needs to be listed and set to zero.
`alpha`	Probability of error of first kind.
`persons1`	A vector of person parameters in group 1 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.
`persons2`	A vector of person parameters in group 2 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the power of the tests is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with df equal to the number of items minus 1 and noncentrality parameter λ. The latter depends on a scenario of deviation from the hypothesis to be tested and a specified sample size. Given the probability of the error of the first kind α the power of the tests can be determined from λ. More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

As regards the concept of sample size a distinction between informative and total sample size has to be made since the power of the tests depends only on the informative sample size. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value of the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons.

In particular, the determination of λ and the power of the tests, respectively, is based on a simple Monte Carlo approach. Data (responses of a large number of persons to a number of items) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. A scenario of a deviation is given by a choice of the item parameters and the person parameters (to be drawn randomly from a specified distribution) for each of the two groups. Such a scenario may be called local deviation since deviations can be specified locally for each item. The relative group sizes are determined by the choice of the number of person parameters for each of the two groups. For instance, by default 10^6 person parameters are selected randomly for each group. In this case, it is implicitly assumed that the two groups of persons are of equal size. The user can specify the relative group sizes by choosing the length of the arguments persons1 and persons2 appropriately. Note that the relative group sizes do have an impact on power and sample size of the tests. The next step is to compute a test statistic T (Wald, LR, score, or gradient) from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. The power of the tests can be determined given a user-specified total sample size denoted by `n_total`. The noncentrality parameter λ can then be expressed by $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$, where $n_{totalsim}$ denotes the total number of persons in the simulated data and $n_{infsim}/n_{totalsim}$ is the proportion of informative persons in the sim. data. Let $q_{1-\alpha}$ be the $1-\alpha$ quantile of the central χ^2 distribution with df equal to the number items minus 1. Then,

$$power = 1 - F_{df,\lambda}(q_{1-\alpha}),$$

where $F_{df,\lambda}$ is the cumulative distribution function of the noncentral χ^2 distribution with df equal to the number of items minus 1 and $\lambda = n_{total}(n_{infsim}/n_{totalsim}) * e$. Thereby, it is assumed that n_{total} is composed of a frequency distribution of person scores that is proportional to the observed distribution of person scores in the simulated data. The same holds true in respect of the relative group sizes, i.e., the relative frequencies of the two person groups in a sample of size n_{total} are assumed to be equal to the relative frequencies of the two groups in the simulated data.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realized value of a random variable E. The same holds true for λ as well as the power of the tests. Thus, the power is a realized value of a random variable that shall be denoted by P. Consequently, the (realized) value of the power of the tests need not be equal to the exact power that follows from the user-specified n_{total}, α, and the chosen item parameters used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters the power of the tests will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing the power of the tests can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution with df equal to the number of free item parameters and noncentrality parameter λ. Thus, $Var(T) = 2(df + 2\lambda)$, with $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. The power of the tests is a function of e which is given by $F_{df,\lambda}(q_\alpha)$, where $\lambda = n_{total} * (n_{infsim}/n_{totalsim}) * e$ and df equal to the number of free item parameters. Then, by the delta method one obtains (for the variance of P).

$$Var(P) = Var(E) * (F'_{df,\lambda}(q_\alpha))^2,$$

where $F'_{df,\lambda}$ is the derivative of $F_{df,\lambda}$ with respect to e. This derivative is determined numerically and evaluated at e using the package numDeriv. The square root of $Var(P)$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of power.

Value

A list of results of class `tcl_power`.

`power` Power value for each test.

`mc_error_power` Monte Carlo error of power computation for each test.

`dev_global` Global deviation computed from simulated data for each test. See Details.

`dev_local` CML estimates of item parameters in both groups of persons obtained from the simulated data expressing a deviation from the hypothesis to be tested locally per item.

`score_dist_group1` Relative frequencies of person scores in group 1 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`score_dist_group2` Relative frequencies of person scores in group 2 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the power of the tests.

`df` Degrees of freedom df.

`ncp` Noncentrality parameter λ of χ^2 distribution from which power is determined.

`call` The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

sa_sizeRM, and post_hocRM.

Examples

```
## Not run:
# Numerical example of power analysis
# for the Rasch model with beta_1 restricted to 0

res <- powerRM(n_total = 130, local_dev = list( c(0, -0.5, 0, 0.5, 1) , c(0, 0.5, 0, -0.5, 1)))

# > res
# $power
#     W    LR    RS    GR
# 0.824 0.840 0.835 0.845
#
# $mc_error_power #`MC error of power`
#     W    LR    RS    GR
# 0.002 0.002 0.002 0.002
#
# $dev_global #`global deviation`
#     W    LR    RS    GR
# 0.118 0.122 0.121 0.124
#
# $dev_local #`local deviation`
```

```
#         Item2 Item3  Item4 Item5
# group1 -0.499 0.005  0.500 1.001
# group2  0.501 0.003 -0.499 1.003
#
# $score_dist_group1 #`person score distribution in group 1`
#
#     1     2     3     4
# 0.249 0.295 0.269 0.187
#
# $score_dist_group2 #`person score distribution in group 2`
#
#     1     2     3     4
# 0.249 0.295 0.270 0.186
#
# $df #`degrees of freedom`
# [1] 4
#
# $ncp #`noncentrality parameter`
#      W     LR     RS     GR
# 12.619 13.098 12.937 13.264
#
# $call
# powerRM(n_total = 130, local_dev = list(c(0, -0.5, 0, 0.5, 1),
#                                         c(0, 0.5, 0, -0.5, 1)))

# Numerical example of power analysis for the Rasch model
# extracting local_dev from an eRm object
dat = eRm::sim.rasch(1000,10)
mod = eRm::RM(dat)

obj <- eRm::LRtest(mod)
res <- powerRM(n_total = 130, obj = obj)

## End(Not run)
```

`sa_sizeChange` *Sample size planning for tests in context of measurement of change using LLTM*

Description

Returns sample size for Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probabilities of errors of first and second kinds α and β as well as a deviation from the hypothesis to be tested. The hypothesis to be tested states that the shift parameter quantifying the constant change for all items between time points 1 and 2 equals 0. The alternative states that the shift parameter is not equal to 0. It is assumed that the same items are presented at both time points. See function `change_test`.

Usage

```
sa_sizeChange(eta, alpha = 0.05, beta = 0.05, persons = rnorm(10^6))
```

Arguments

`eta`	A vector of eta parameters of the LLTM. The last element represents the constant change or shift for all items between time points 1 and 2. The other elements of the vector are the item parameters at time point 1. A choice of the eta parameters constitutes a scenario of deviation from the hypothesis of no change.
`alpha`	Probability of error of first kind.
`beta`	Probability of error of second kind.
`persons`	A vector of person parameters (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the sample size is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with $df = 1$ and noncentrality parameter λ. The latter is, inter alia, a function of the sample size. Hence, the sample size can be determined from the condition $\lambda = \lambda_0$, where λ_0 is a predetermined constant which depends on the probabilities of the errors of the first and second kinds α and β (or power). More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, the determination of λ and the sample size, respectively, is based on a simple Monte Carlo approach. As regards the concept of sample size a distinction between informative and total sample size has to be made. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value of the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons. The Monte Carlo approach used in the present problem to determine λ and informative (and total) sample size can briefly be described as follows. Data (responses of a large number of persons to a number of items presented at two time points) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. The hypothesis to be tested assumes no change between time points 1 and 2. A scenario of a deviation is given by a choice of the item parameters at time point 1 and the shift parameter, i.e., the LLTM eta parameters, as well as the person parameters (to be drawn randomly from a specified distribution). The shift parameter represents a constant change of all item parameters from time point 1 to time point 2. A test statistic T (Wald, LR, score, or gradient) is computed from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. Let the informative sample size sought be denoted by n_{inf} (thus, this is not the informative sample size observed in the sim. data). The noncentrality parameter λ can be expressed by the product $n_{inf} * e$. Then, it follows from the condition $\lambda = \lambda_0$ that

$$n_{inf} * e = \lambda_0$$

and

$$n_{inf} = \lambda_0/e.$$

Note that the sample of size n_{inf} is assumed to be composed only of persons with informative person scores, where the relative frequency distribution of these informative scores is considered to be equal to the observed relative frequency distribution of the informative scores in the simulated data. The total sample size n_{total} is then obtained from the relation $n_{inf} = n_{total} * pr$, where pr is the proportion or relative frequency of persons observed in the simulated data with a minimum or maximum score. Basing the tests given a level α on an informative sample of size n_{inf} the probability of rejecting the hypothesis to be tested will be at least $1 - \beta$ if the true global deviation $\geq e$.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realized value of a random variable E. Consequently, n_{inf} is also a realization of a random variable N_{inf}. Thus, the (realized) value n_{inf} need not be equal to the exact value of the informative sample size that follows from the user-specified (predetermined) α, β, and scenario of a deviation from the hypothesis to be tested, i.e., the selected item parameters and shift parameter used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters n_{inf} will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure of n_{inf} based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing n_{inf} can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution. Thus, $Var(T) = 2(df + 2\lambda)$, with $df = 1$ and $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. Since $n_{inf} = f(e) = \lambda_0/e$ one obtains by the delta method (for the variance of the corresponding random variable N_{inf})

$$Var(N_{inf}) = Var(E) * (f'(e))^2,$$

where $f'(e) = -\lambda_0/e^2$ is the derivative of $f(e)$. The square root of $Var(N_{inf})$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of informative sample size.

Value

A list of results of class `tcl_sa_size`.

`sample_size_informative`
: Informative sample size for each test, omitting persons with min. and max score.

`mc_error_sample_size`
: Monte Carlo error of sample size computation for each test.

`dev`
: Shift parameter estimated from the simulated data representing the constant shift of item parameters between time points 1 and 2.

`score_dist`	Relative frequencies of person scores observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the sample size.
`df`	Degrees of freedom df.
`ncp`	Noncentrality parameter λ of χ^2 distribution from which sample size is determined.
`sample_size_total`	Total sample size for each test. See Details.
`call`	The matched call.

References

Draxler, C., & Alexandrowicz, R. W. (2015). Sample size determination within the scope of conditional maximum likelihood estimation with special focus on testing the Rasch model. Psychometrika, 80(4), 897-919.

Fischer, G. H. (1995). The Linear Logistic Test Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch models: Foundations, Recent Developments, and Applications (pp. 131-155). New York: Springer.

Fischer, G. H. (1983). Logistic Latent Trait Models with Linear Constraints. Psychometrika, 48(1), 3-26.

See Also

powerChange, and post_hocChange.

Examples

```
## Not run:
# Numerical example 4 items presented twice, thus 8 virtual items

# eta Parameter, first 4 are nuisance
# (easiness parameters of the 4 items at time point 1),
# last one is the shift parameter
eta <- c(-2,-1,1,2,0.5)

res <- sa_sizeChange(eta = eta)

# > res
# $sample_size_informative #`informative sample size`
#   W  LR  RS  GR
# 177 174 175 173
#
# $mc_error_sample_size #`MC error of sample size`
#     W    LR    RS    GR
# 1.321 1.287 1.299 1.276
#
# $dev #`deviation (estimate of shift parameter)`
# [1] 0.501
#
```

```
# $score_dist #`person score distribution`
#
#     1     2     3     4     5     6     7
# 0.034 0.094 0.181 0.249 0.227 0.147 0.068
#
# $df #`degrees of freedom`
# [1] 1
#
# $ncp #`noncentrality parameter`
# [1] 12.995
#
# $sample_size_total #`total sample size`
#   W  LR  RS  GR
# 182 179 180 178
#
# $call
# sa_sizeChange(eta = eta)

## End(Not run)
```

`sa_sizePCM` *Sample size planning for tests of invariance of item-category parameters between two groups of persons in partial credit model*

Description

Returns sample size for Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probabilities of errors of first and second kinds α and β as well as a deviation from the hypothesis to be tested. The hypothesis to be tested assumes equal item-category parameters in the partial credit model between two predetermined groups of persons. The alternative assumes that at least one parameter differs between the two groups.

Usage

```
sa_sizePCM(
  obj = NULL,
  local_dev = NULL,
  alpha = 0.05,
  beta = 0.05,
  persons1 = rnorm(10^6),
  persons2 = rnorm(10^6)
)
```

Arguments

`obj`	An object of class ‘LR‘ from the ‘eRm‘ package. If provided, ‘local_dev‘ is extracted automatically. If missing, ‘local_dev‘ must be set manually.

`local_dev`	A list consisting of two lists. One list refers to group 1, the other to group 2. Each of the two lists contains a numerical vector per item, i.e., each list contains as many vectors as items. Each vector contains the free item-cat. parameters of the respective item. The number of free item-cat. parameters per item equals the number of categories of the item minus 1.
`alpha`	Probability of the error of first kind.
`beta`	Probability of the error of second kind.
`persons1`	A vector of person parameters for group 1 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details. .
`persons2`	A vector of person parameters for group 2 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the sample size is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with $df = l$, where l is the number of free item-category parameters in the partial credit model, and noncentrality parameter λ. The latter is, inter alia, a function of the sample size. Hence, the sample size can be determined from the condition $\lambda = \lambda_0$, where λ_0 is a predetermined constant which depends on the probabilities of the errors of the first and second kinds α and β (or power). More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, the determination of λ and the sample size, respectively, is based on a simple Monte Carlo approach. As regards the concept of sample size a distinction between informative and total sample size has to be made. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value of the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons. The Monte Carlo approach used in the present problem to determine λ and informative (and total) sample size can briefly be described as follows. Data (responses of a large number of persons to a number of items) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. The hypothesis to be tested assumes equal item-category parameters between the two groups of persons. A scenario of a deviation is given by a choice of the item-cat. parameters and the person parameters (to be drawn randomly from a specified distribution) for each of the two groups. Such a scenario may be called local deviation since deviations can be specified locally for each item-category. The relative group sizes are determined by the choice of the number of person parameters for each of the two groups. For instance, by default 10^6 person parameters are selected randomly for each group. In this case, it is implicitly assumed that the two groups of persons are of equal size. The user can specify the relative groups sizes by choosing the length of the arguments `persons1` and `persons2` appropriately. Note that the relative group sizes do have an impact on power and sample size of the tests. The next step is to compute a test statistic T (Wald, LR, score, or gradient) from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. Let the informative sample size sought be denoted by n_{inf} (thus, this is not the informative sample size observed in the sim. data). The noncentrality parameter λ can be expressed by the product $n_{inf} * e$. Then, it follows from the condition $\lambda = \lambda_0$ that

$$n_{inf} * e = \lambda_0$$

and

$$n_{inf} = \lambda_0 / e.$$

Note that the sample of size n_{inf} is assumed to be composed only of persons with informative person scores in both groups, where the relative frequency distribution of these informative scores in each of both groups is considered to be equal to the observed relative frequency distribution of informative scores in each of both groups in the simulated data. Note also that the relative sizes of the two person groups are assumed to be equal to the relative sizes of the two groups in the simulated data. By default, the two groups are equal-sized in the simulated data, i.e., one yields $n_{inf}/2$ persons (with informative scores) in each of the two groups. The total sample size n_{total} is obtained from the relation $n_{inf} = n_{total} * pr$, where pr is the proportion or relative frequency of persons observed in the simulated data with a minimum or maximum score. Basing the tests given a level α on an informative sample of size n_{inf} the probability of rejecting the hypothesis to be tested will be at least $1 - \beta$ if the true global deviation $\geq e$.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realization of a random variable E. Consequently, n_{inf} is also a realization of a random variable N_{inf}. Thus, the (realized) value n_{inf} need not be equal to the exact value of the informative sample size that follows from the user-specified (predetermined) α,β, and scenario of a deviation from the hypothesis to be tested, i.e., the selected item-category parameters used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters n_{inf} will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data, i.e., the lengths of the vectors persons1 and persons2, is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure of n_{inf} based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing n_inf can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution. Thus, $Var(T) = 2(df + 2\lambda)$, with $df = l$ and $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. Since $n_{inf} = f(e) = \lambda_0/e$ one obtains by the delta method (for the variance of the corresponding random variable N_{inf})

$$Var(N_{inf}) = Var(E) * (f'(e))^2,$$

where $f'(e) = -\lambda_0/e^2$ is the derivative of $f(e)$. The square root of $Var(N_{inf})$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of informative sample size.

Value

A list of results of class `tcl_sa_size`.

`sample_size_informative`
: Informative sample size for each test, omitting persons with min. and max score.

`mc_error_sample_size`
: Monte Carlo error of informative sample size for each test.

`dev_global`
: Global deviation computed from simulated data. See Details.

`dev_local`
: CML estimates of free item-category parameters in both group of persons obtained from the simulated data expressing a deviation from the hypothesis to be tested locally per item and response category.

`score_dist_group1`
: Relative frequencies of person scores in group 1 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the sample size.

`score_dist_group2`
: Relative frequencies of person scores in group 2 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the sample size.

`df`
: Degrees of freedom df.

`ncp`
: Noncentrality parameter λ of χ^2 distribution from which sample size is determined.

`sample_size_total_group1`
: Total sample size in group 1 for each test. See Details.

`sample_size_total_group2`
: Total sample size in group 2 for each test. See Details.

`call`
: The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

powerPCM, and post_hocPCM.

Examples

```
## Not run:
##### Sample size of PCM Model #####

# free item-category parameters for group 1 and 2  with 5 items, with 3 categories each
local_dev <-  list (  list(c( 0, 0), c( -1, 0), c( 0, 0),  c( 1, 0), c( 1, 0.5)) ,
                      list(c( 0, 0), c( -1, 0), c( 0, 0),  c( 1, 0), c( 0, -0.5))  )

res <- sa_sizePCM(local_dev = local_dev, alpha = 0.05, beta = 0.05, persons1 = rnorm(10^6),
                  persons2 = rnorm(10^6))

# > res
# $sample_size_informative #`informative sample size`
#   W  LR  RS  GR
# 234 222 227 217
#
# $mc_error_sample_size #`MC error of sample size`
#     W    LR    RS    GR
# 1.105 1.018 1.053 0.988
#
# $dev_global  #`global deviation`
#     W    LR    RS    GR
# 0.101 0.107 0.104 0.109
#
# $dev_local #`local deviation`
#          I1-C2  I2-C1  I2-C2  I3-C1  I3-C2 I4-C1 I4-C2 I5-C1  I5-C2
# group1 -0.001 -1.000 -1.001 -0.003 -0.011 0.997 0.998 0.996  1.492
# group2  0.001 -0.998 -0.996 -0.007 -0.007 0.991 1.001 0.004 -0.499
#
# $score_dist_group1 #`person score distribution in group 1`
#
#     1     2     3     4     5     6     7     8     9
# 0.111 0.130 0.133 0.129 0.122 0.114 0.101 0.091 0.070
#
# $score_dist_group2 #`person score distribution in group 2`
#
#     1     2     3     4     5     6     7     8     9
# 0.090 0.109 0.117 0.121 0.121 0.121 0.116 0.111 0.093
#
# $df #`degrees of freedom`
```

```
# [1] 9
#
# $ncp #`noncentrality parameter`
# [1] 23.589
#
# $sample_size_total_group1 #`total sample size in group 1`
#   W  LR  RS  GR
# 132 125 128 123
#
# $sample_size_total_group2 #`total sample size in group 2`
#   W  LR  RS  GR
# 133 126 129 123
#
# $call
# sa_sizePCM(alpha = 0.05, beta = 0.05, persons1 = rnorm(10^6),
#            persons2 = rnorm(10^6), local_dev = local_dev)


# Sample size of of PCM
# extracting local_dev from an eRm object
ppar = rnorm(10000)
ipar = list(-2:2,-2:2+0.2,-2:2-0.3,-2:2+0.1,-2:2-0.8)
dat2 = psychotools::rpcm(theta = ppar, delta = ipar)

mod2 = PCM(dat2$data)
obj2 = eRm::LRtest(mod2)

res <- sa_sizePCM(obj = obj2)

## End(Not run)
```

`sa_sizeRM` *Sample size planning for tests of invariance of item parameters between two groups of persons in binary Rasch model*

Description

Returns sample size for Wald (W), likelihood ratio (LR), Rao score (RS) and gradient (GR) test given probabilities of errors of first and second kinds α and β as well as a deviation from the hypothesis to be tested. The hypothesis to be tested assumes equal item parameters between two predetermined groups of persons. The alternative assumes that at least one parameter differs between the two groups.

Usage

```
sa_sizeRM(
  obj = NULL,
  local_dev = NULL,
  alpha = 0.05,
```

```
  beta = 0.05,
  persons1 = rnorm(10^6),
  persons2 = rnorm(10^6)
)
```

Arguments

`obj` An object of class 'LR' from the 'eRm' package. If provided, 'local_dev' is extracted automatically. If missing, 'local_dev' must be set manually.

`local_dev` A list consisting of two vectors containing item parameters for the two person groups representing a deviation from the hypothesis to be tested locally per item. Note that the 'reference category', i.e. the first item parameter, also needs to be listed and set to zero.

`alpha` Probability of the error of first kind.

`beta` Probability of the error of second kind.

`persons1` A vector of person parameters for group 1 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

`persons2` A vector of person parameters for group 2 (drawn from a specified distribution). By default 10^6 parameters are drawn at random from the standard normal distribution. The larger this number the more accurate are the computations. See Details.

Details

In general, the sample size is determined from the assumption that the approximate distributions of the four test statistics are from the family of noncentral χ^2 distributions with df equal to the number of items minus 1, and noncentrality parameter λ. The latter is, inter alia, a function of the sample size. Hence, the sample size can be determined from the condition $\lambda = \lambda_0$, where λ_0 is a predetermined constant which depends on the probabilities of the errors of the first and second kinds α and β (or power). More details about the distributions of the test statistics and the relationship between λ, power, and sample size can be found in Draxler and Alexandrowicz (2015).

In particular, the determination of λ and the sample size, respectively, is based on a simple Monte Carlo approach. As regards the concept of sample size a distinction between informative and total sample size has to be made. In the conditional maximum likelihood context, the responses of persons with minimum or maximum person score are completely uninformative. They do not contribute to the value o f the test statistic. Thus, the informative sample size does not include these persons. The total sample size is composed of all persons. The Monte Carlo approach used in the present problem to determine λ and informative (and total) sample size can briefly be described as follows. Data (responses of a large number of persons to a number of items) are generated given a user-specified scenario of a deviation from the hypothesis to be tested. The hypothesis to be tested assumes equal item parameters between the two groups of persons. A scenario of a deviation is given by a choice of the item parameters and the person parameters (to be drawn randomly from a specified distribution) for each of the two groups. Such a scenario may be called local deviation since deviations can be specified locally for each item. The relative group sizes are determined by the choice of the number of person parameters for each of the two groups. For instance, by default

10^6 person parameters are selected randomly for each group. In this case, it is implicitly assumed that the two groups of persons are of equal size. The user can specify the relative groups sizes by choosing the lengths of the arguments `persons1` and `persons2` appropriately. Note that the relative group sizes do have an impact on power and sample size of the tests. The next step is to compute a test statistic T (Wald, LR, score, or gradient) from the simulated data. The observed value t of the test statistic is then divided by the informative sample size n_{infsim} observed in the simulated data. This yields the so-called global deviation $e = t/n_{infsim}$, i.e., the chosen scenario of a deviation from the hypothesis to be tested being represented by a single number. Let the informative sample size sought be denoted by n_{inf} (thus, this is not the informative sample size observed in the sim. data). The noncentrality parameter λ can be expressed by the product $n_{inf} * e$. Then, it follows from the condition $\lambda = \lambda_0$ that

$$n_{inf} * e = \lambda_0$$

and

$$n_{inf} = \lambda_0/e.$$

Note that the sample of size n_{inf} is assumed to be composed only of persons with informative person scores in both groups, where the relative frequency distribution of these informative scores in each of both groups is considered to be equal to the observed relative frequency distribution of informative scores in each of both groups in the simulated data. Note also that the relative sizes of the two person groups are assumed to be equal to the relative sizes of the two groups in the simulated data. By default, the two groups are equal-sized in the simulated data, i.e., one yields $n_{inf}/2$ persons (with informative scores) in each of the two groups. The total sample size n_{total} is obtained from the relation $n_{inf} = n_{total} * pr$, where pr is the proportion or relative frequency of persons observed in the simulated data with a minimum or maximum score. Basing the tests given a level α on an informative sample of size n_{inf} the probability of rejecting the hypothesis to be tested will be at least $1 - \beta$ if the true global deviation $\geq e$.

Note that in this approach the data have to be generated only once. There are no replications needed. Thus, the procedure is computationally not very time-consuming.

Since e is determined from the value of the test statistic observed in the simulated data it has to be treated as a realization of a random variable E. Consequently, n_{inf} is also a realization of a random variable N_{inf}. Thus, the (realized) value n_{inf} need not be equal to the exact value of the informative sample size that follows from the user-specified (predetermined) α, β, and scenario of a deviation from the hypothesis to be tested, i.e., the selected item parameters used for the simulation of the data. If the CML estimates of these parameters computed from the simulated data are close to the predetermined parameters n_{inf} will be close to the exact value. This will generally be the case if the number of person parameters used for simulating the data, i.e., the lengths of the vectors `persons1` and `persons2`, is large, e.g., 10^5 or even 10^6 persons. In such cases, the possible random error of the computation procedure of n_{inf} based on the sim. data may not be of practical relevance any more. That is why a large number (of persons for the simulation process) is generally recommended.

For theoretical reasons, the random error involved in computing n_{inf} can be pretty well approximated. A suitable approach is the well-known delta method. Basically, it is a Taylor polynomial of first order, i.e., a linear approximation of a function. According to it the variance of a function of a random variable can be linearly approximated by multiplying the variance of this random variable with the square of the first derivative of the respective function. In the present problem, the variance of the test statistic T is (approximately) given by the variance of a noncentral χ^2 distribution. Thus, $Var(T) = 2(df + 2\lambda)$, with df equal to the number of items minus 1 and $\lambda = t$. Since the global deviation $e = (1/n_{infsim}) * t$ it follows for the variance of the corresponding random variable E that $Var(E) = (1/n_{infsim})^2 * Var(T)$. Since $n_{inf} = f(e) = \lambda_0/e$ one obtains by the delta method (for the variance of the corresponding random variable N_{inf})

$$Var(N_{inf}) = Var(E) * (f'(e))^2,$$

where $f'(e) = -\lambda_0/e^2$ is the derivative of $f(e)$. The square root of $Var(N_{inf})$ is then used to quantify the random error of the suggested Monte Carlo computation procedure. It is called Monte Carlo error of informative sample size.

Value

A list of results of class `tcl_sa_size`.

`sample_size_informative`
: Informative sample size for each test omitting persons with min. and max score.

`mc_error_sample_size`
: Monte Carlo error of informative sample size for each test.

`dev_global`
: Global deviation computed from simulated data. See Details.

`dev_local`
: CML estimates of free item parameters in both groups obtained from the simulated data. First item parameter set 0 in both groups.

`score_dist_group1`
: Relative frequencies of person scores in group 1 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the sample size.

`score_dist_group2`
: Relative frequencies of person scores in group 2 observed in simulated data. Uninformative scores, i.e., minimum and maximum score, are omitted. Note that the person score distribution does also have an influence on the sample size.

`df`
: Degrees of freedom df.

`ncp`
: Noncentrality parameter λ of χ^2 distribution from which sample size is determined.

`sample_size_total_group1`
: Total sample size in group 1 for each test. See Details.

`sample_size_total_group2`
: Total sample size in group 2 for each test. See Details.

`call`
: The matched call.

References

Draxler, C. (2010). Sample Size Determination for Rasch Model Tests. Psychometrika, 75(4), 708–724.

Draxler, C., & Alexandrowicz, R. W. (2015). Sample Size Determination Within the Scope of Conditional Maximum Likelihood Estimation with Special Focus on Testing the Rasch Model. Psychometrika, 80(4), 897–919.

Draxler, C., Kurz, A., & Lemonte, A. J. (2022). The gradient test and its finite sample size properties in a conditional maximum likelihood and psychometric modeling context. Communications in Statistics-Simulation and Computation, 51(6), 3185-3203.

Glas, C. A. W., & Verhelst, N. D. (1995a). Testing the Rasch Model. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 69–95). New York: Springer.

Glas, C. A. W., & Verhelst, N. D. (1995b). Tests of Fit for Polytomous Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), Rasch Models: Foundations, Recent Developments, and Applications (pp. 325-352). New York: Springer.

See Also

powerRM, and post_hocRM.

Examples

```
## Not run:
##### Sample size of Rasch Model #####

res <-  sa_sizeRM(local_dev = list( c(0, -0.5, 0, 0.5, 1) , c(0, 0.5, 0, -0.5, 1)))

# > res
# $sample_size_informative #`informative sample size`
#   W  LR  RS  GR
# 159 153 155 151
#
# $mc_error_sample_siz #`MC error of sample size`
#     W    LR    RS    GR
# 0.721 0.682 0.695 0.670
#
# $dev_global #`global deviation`
#     W    LR    RS    GR
# 0.117 0.122 0.120 0.123
#
# $dev_local #`local deviation`
#         Item2  Item3  Item4 Item5
# group1 -0.502 -0.005  0.497 1.001
# group2  0.495 -0.006 -0.501 0.994
#
# $score_dist_group1 #`person score distribution in group 1`
#
#     1     2     3     4
# 0.249 0.295 0.268 0.188
#
# $score_dist_group2 #`person score distribution in group 2`
#
#     1     2     3     4
# 0.249 0.295 0.270 0.187
#
# $df #`degrees of freedom`
# [1] 4
```

```
#
# $ncp #`noncentrality parameter`
# [1] 18.572
#
# $sample_size_total_group1 #`total sample size in group 1`
#  W LR RS GR
# 97 93 94 92
#
# $sample_size_total_group2 #`total sample size in group 2`
#  W LR RS GR
# 97 93 94 92
#
# $call
# sa_sizeRM(local_dev = list(c(0, -0.5, 0, 0.5, 1),
#                            c(0, 0.5, 0, -0.5, 1)))
#
##### Sample size of Rasch Model #####
# extracting local_dev from an eRm object
dat = eRm::sim.rasch(1000,10)
mod = eRm::RM(dat)

obj <- eRm::LRtest(mod)
res <- sa_sizeRM(obj = obj)

## End(Not run)
```

---

tcl_hessian *Computation of Hessian matrix.*

Description

Uses function `hessian()` from numDeriv package to compute (approximate numerically) Hessian matrix evaluated at arbitrary values of item easiness parameters.

Usage

```
tcl_hessian(data, eta, W, model = "RM")
```

Arguments

`data`	data matrix.
`eta`	numeric vector of item easiness parameters.
`W`	design matrix.
`model`	RM, PCM, RSM, LLTM. Default is set to "RM".

Value

Hessian matrix evaluated at eta.

References

Gilbert, P., Gilbert, M. P., & Varadhan, R. (2016). numDeriv: Accurate Numerical Derivatives. R package version 2016.8-1.1. url: https://CRAN.R-project.org/package=numDeriv

Examples

```
## Not run:
# Rasch model with beta_1 restricted to 0
y <- eRm::raschdat1
res <- eRm::RM(X = y, sum0 = FALSE)
mat <- tcl_hessian(data = y, eta = res$etapar, model = "RM")


## End(Not run)
```

---

tcl_scorefun   *Computation of score function.*

Description

Uses function jacobian() from numDeriv package to compute (approximate numerically) score function (first order partial derivatives of conditional log likelihood function) evaluated at arbitrary values of item easiness parameters.

Usage

```
tcl_scorefun(data, eta, W, model = "RM")
```

Arguments

data	data matrix.
eta	numeric vector of item easiness parameters.
W	design matrix.
model	RM, PCM, RSM, LLTM. Default is set to "RM".

Value

Score function evaluated at eta.

References

Gilbert, P., Gilbert, M. P., & Varadhan, R. (2016). numDeriv: Accurate Numerical Derivatives. R package version 2016.8-1.1. url: https://CRAN.R-project.org/package=numDeriv

Examples

```
## Not run:
# Rasch model with beta_1 restricted to 0
y <- eRm::raschdat1
res <- eRm::RM(X = y, sum0 = FALSE)
scorefun <- tcl_scorefun(data = y, eta = res$etapar, model = "RM")

## End(Not run)
```

Index