Package 'pln'

October 16, 2024

Type Package

Title Polytomous Logit-Normit (Graded Logistic) Model Estimation

Version 0.2-3

Date 2024-10-16

Maintainer Carl F. Falk <cffalk@gmail.com>

Description Performs bivariate composite likelihood and full information maximum likelihood estimation for polytomous logit-normit (graded logistic) item response theory (IRT) models.

Depends R (>= 2.9.0)

License GPL-3

LazyLoad yes

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation yes

Author Carl F. Falk [aut, cre], Harry Joe [aut]

Repository CRAN

Date/Publication 2024-10-16 21:30:02 UTC

Contents

pln-package *Polytomous Logit-Normit (Graded Logistic) Model Estimation*

Description

Performs bivariate composite likelihood and full information maximum likelihood estimation for polytomous logit-normit (graded logistic) item response theory (IRT) models.

Details

Package:	pln
Type:	Package
Version:	0.2-2
Date:	2020-07-28
License:	GPL-3
LazyLoad:	yes

This package currently contains several functions performing estimation of unidimensional (single latent trait) polytomous logit-normit models (also known graded logistic) using bivariate composite likelihood and full information maximum likelihood estimation.

Acknowledgment

Some code from the ltm package (version 0.9-7) was modified for counting the frequency of response patterns.

Author(s)

Carl Falk <cffalk@gmail.com>, Harry Joe

Maintainer: Carl F. Falk <cffalk@gmail.com>

References

Bartholomew, D., Knott, M., and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*, 3rd Edition. Wiley.

Maydeu-Olivares, A., and Joe, H. (2005). Limited and full information estimation and goodness-of-fit testing in 2^n contingency tables: A unified framework. *Journal of the American Statistical Association, 100*, 1009-1020.

Maydeu-Olivares, A., and Joe, H. (2006). Limited information and goodness-of-fit testing in multidimensional contingency tables. *Psychometrika, 71*, 713-732.

Varin, C., Reid, N. and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica, 21*, 5-42.

item5fr *5-item Test Data Set*

Description

A simulated data set with a 5 items, each with 3 ordinal categories labeled as 0,1,2. Each row corresponds to a response pattern with the sixth column containing the frequency of each response pattern.

Usage

```
data(item5fr)
```

item9cat5 *9 Item Test Data Set*

Description

A simulated data set with a 9 items, each with 5 ordinal categories labeled as 0,1,2,3,4. Each row corresponds to a single individual's response on the 9 items.

Usage

```
data(item9cat5)
```

nrmlepln *Full information maximum likelihood and bivariate composite likelihood estimation for polytomous logit-normit (graded logistic) model*

Description

Full information maximum likelihood and bivariate composite likelihood estimation for polytomous logit-normit and Rasch models, via Newton Raphson iterations.

Usage

```
nrmlepln(x, ncat, nitem=NULL, alphas=NULL, betas=NULL, abound=c(-10,10),
         bbound=c(-1,10), nq=48, mxiter=200, m2=TRUE, iprint=FALSE)
nrmlerasch(x, ncat, nitem=NULL, alphas=NULL, abound=c(-10,10),
         bbound=c(-1,10), nq=48, mxiter=200, m2=TRUE, iprint=FALSE)
nrbcpln(x, ncat, nitem=NULL, alphas=NULL, betas=NULL, abound=c(-10,10),
         bbound=c(-1,10), nq=48, mxiter=200, se=TRUE, iprint=FALSE)
```

Arguments

x	A data matrix. Data can be in one of two formats: 1) raw data where the number of rows corresponds to an individual's response and each column represents an item, and 2) a matrix of dimensions nrec×(nitem+1) where each row corresponds to a response pattern and the last column is the frequency of that response pattern. A data matrix of the second type requires input for nitem and nrec.
ncat	Number of ordinal categories for each item, coded as 0,...,(ncat-1). Currently supported are items that have the same number of categories.
nitem	Number of items. If omitted, it is assumed that x contains a data matrix of the first type (raw data) and the number of columns in x will be selected as the number of items.
alphas	A vector of length nitem×(ncat-1) corresponding to starting values for the (decreasing) cutpoints for the items. If omitted, these will be computed from the function startalphas.
betas	A vector of length nitem corresponding to starting values for the beta vectors of slopes. If omitted, these will be computed from the function startbetas. For the polytomous logit-normit, there is one slope for each item; for the Rasch model, there is a common slope beta for all of the items.
abound	Vector of length 2 that sets upper and lower bounds on parameter estimation for alphas. Currently experimental; changing defaults it not recommended. Estimation problems are more likely solved by changing starting values.
bbound	Vector of length 2 that sets upper and lower bounds on parameter estimation for betas. Currently experimental; changing defaults it not recommended. Estimation problems are more likely solved by changing starting values.
nq	Number of quadrature points to use during estimation. This argument is currently experimental. It is recommended to use the default of 48.
mxiter	Maximum number of iterations for estimation.
m2	Logical. If TRUE, computes goodness-of-fit statistics from Maydeu-Olivares and Joe (2005, 2006; i.e., M_2).
iprint	Logical. Enables debugging / diagnostic information from C code that conducts estimation.
se	Logical. If TRUE, calculates standard errors for the bivariate composite likelihood method.

Details

Estimation of graded logistic models is performed under the following parameterization:

$$Pr(y_i = k_i|\eta) = \begin{cases} 1 - \Psi(\alpha_{i,k} + \beta_i\eta) & \text{if } k_i = 0 \\ \Psi(\alpha_{i,k} + \beta_i\eta) - \Psi(\alpha_{i,k+1} + \beta_i\eta) & \text{if } 0 < k_i < m-1 \\ \Psi(\alpha_{i,k+1} + \beta_i\eta) & \text{if } k_i = m-1 \end{cases}$$

Where the items are $y_i, i = 1, \ldots, n$, and response categories are $k = 0, \ldots, m-1$. η is the latent trait, Ψ is the logistic distribution function, α is an intercept (cutpoint) parameter, and β

is a slope parameter. When the number of categories for the items is 2, this reduces to the 2PL parameterization:

$$Pr(y_i = 1|\eta) = \Psi(\alpha_1 + \beta_i\eta)$$

Both `nrmlepln` and `nrbcpln` perform estimation under these parameterizations, via Newton Raphson iterations, using full information maximum likelihood (`nrmlepln`) and bivariate composite likelihood (`nrbcpln`). See Maydeu-Olivares and Joe (2005, 2006) for more information on bivariate composite likelihood estimation (see also Varin, Reid, and Firth, 2011). Under `nrmlerasch` a common β parameter is estimated for all items.

Value

A list containing the following slots.

Slots

`alphas` A vector of parameter estimates for alphas. Length is `nitem`×(`ncat`-1). Estimates are in order by item, e.g., all alphas for item 1, followed by all alphas for item 2, and so on.

`betas` A vector of parameter estimates for betas. Length is `nitem`.

`nllk` Negative (composite) log-likelihood for polytomous logit-normit (or Rasch) model.

`conv` Integer indicating whether estimation converged. Currently only returned for composite likelihood estimation.

`sealphas` A vector of standard errors for the alpha estimates.

`sebetas` A vector of standard errors for the beta estimates.

`invhes` Inverse Hessian matrix for the MLE estimates.

`vcov` Asymptotic covariance matrix for the composite likelihood estimates.

`teststat` Value of M_2.

`df` Degrees of freedom for M_2.

`pval` P-value for M_2.

Author(s)

Carl F. Falk <cffalk@gmail.com>, Harry Joe

References

Bartholomew, D., Knott, M., and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*, 3rd Edition. Wiley.

Maydeu-Olivares, A., and Joe, H. (2005). Limited and full information estimation and goodness-of-fit testing in 2^n contingency tables: A unified framework. *Journal of the American Statistical Association, 100*, 1009-1020.

Maydeu-Olivares, A., and Joe, H. (2006). Limited information and goodness-of-fit testing in multidimensional contingency tables. *Psychometrika, 71*, 713-732.

Varin, C., Reid, N. and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica, 21*, 5-42.

See Also

startalphas startbetas

Examples

```
### Matrix of response patterns and frequencies
data(item5fr)

## ML estimation
nrmleplnout<-nrmlepln(item5fr, ncat=3, nitem=5)
print(nrmleplnout)

## BCL estimation
nrbcplnout<-nrbcpln(item5fr, ncat=3, nitem=5)
print(nrbcplnout)

## ML Rasch estimation
nrmleraschout<-nrmlerasch(item5fr, ncat=3, nitem=5)
print(nrmleraschout)


### Raw data
data(item9cat5)

## ML estimation
nrmleplnout<-nrmlepln(item9cat5, ncat=5)
print(nrmleplnout)

## BCL estimation
nrbcplnout<-nrbcpln(item9cat5, ncat=5, se=FALSE)
print(nrbcplnout)

## ML Rasch estimation
nrmleraschout<-nrmlerasch(item9cat5, ncat=5)
print(nrmleraschout)
```

---

simulpln *Simulate data from polytomous logit-normit (graded logistic) model*

Description

Simulate data from polytomous logit-normit (graded logistic) model

Usage

```
simulpln(n, nitem, ncat, alphas, betas)
```

Arguments

n	Number of responses to generate.
nitem	Number of items.
ncat	Number of categories for the items.
alphas	A vector of length nitem×(ncat-1) corresponding to true values for the (decreasing) cutpoints for the items.
betas	A vector of length nitem corresponding to values for the beta vectors of slopes.

Details

Data from graded logistic models is generated under the following parameterization:

$$Pr(y_i = k_i|\eta) = \begin{cases} 1 - \Psi(\alpha_{i,k} + \beta_i\eta) & \text{if } k_i = 0 \\ \Psi(\alpha_{i,k} + \beta_i\eta) - \Psi(\alpha_{i,k+1} + \beta_i\eta) & \text{if } 0 < k_i < m-1 \\ \Psi(\alpha_{i,k+1} + \beta_i\eta) & \text{if } k_i = m-1 \end{cases}$$

Where the items are $y_i, i = 1, \ldots, n$, and response categories are $k = 0, \ldots, m-1$. η is the latent trait, Ψ is the logistic distribution function, α is an intercept (cutpoint) parameter, and β is a slope parameter. When the number of categories for the items is 2, this reduces to the 2PL parameterization:

$$Pr(y_i = 1|\eta) = \Psi(\alpha_1 + \beta_i\eta)$$

Value

A data matrix in which each row represents a response pattern and the final column represents the frequency of each response pattern.

Author(s)

Carl F. Falk <cffalk@gmail.com>, Harry Joe

See Also

nrmlepln nrmlerasch nrbcpln

Examples

```
n<-500;
ncat<-3;
nitem<-5
alphas=c(0,-.5,  .2,-1,  .4,-.6,  .3,-.2,  .5,-.5)
betas=c(1,1,1,.5,.5)

set.seed(1234567)
datfr<-simulpln(n,nitem,ncat,alphas,betas)
nrmleplnout<-nrmlepln(datfr, ncat=ncat, nitem=nitem)
nrmleplnout
```

`startalphas` *Starting values for polytomous logit-normit model*

Description

Computes starting values for estimation of polytomous logit-normit model.

Usage

```
startalphas(x, ncat, nitem = NULL)
startbetas(x, ncat, nitem = NULL)
```

Arguments

`x`	A data matrix. Data can be in one of two formats: 1) raw data where the number of rows corresponds to the number of raw cases and each column represents an item, and 2) a matrix of dimensions `nrec`$\times$(`nitem`+1) where each row corresponds to a response pattern and the last column is the frequency of that response pattern. A data matrix of the second type requires input for `nitem` and `nrec`.
`ncat`	Number of ordinal categories for each item, coded as 0,...,(`ncat`-1). Currently supported are items that have the same number of categories.
`nitem`	Number of items. If omitted, it is assumed that `x` contains a data matrix of the first type (raw data) and the number of columns in `x` will be selected as the number of items.

Details

`startalphas` computes starting values for the (decreasing) cutpoints for the items based on logit transformed probabilities, assuming independent items.

`startbetas` computes starting values for slopes under the polytomous logit-normit model, using a method based on values that are proportional to the average correlations of each item with all other items. Starting values are currently bounded between -.2 and 1.

Value

A vector of starting values, depending on which function was called.

Author(s)

Carl F. Falk <cffalk@gmail.com>, Harry Joe

See Also

`nrmlepln` `nrmlerasch` `nrbcpln`

Examples

```
### Raw data
data(item9cat5)

myAlphas<-startalphas(item9cat5, ncat=5)
print(myAlphas)

myBetas<-startbetas(item9cat5, ncat=5)
print(myBetas)

nrbcplnout<-nrbcpln(item9cat5, ncat=5, alphas=myAlphas, betas=myBetas, se=FALSE)
print(nrbcplnout)

## Matrix of response patterns and frequencies
data(item5fr)

myAlphas<-startalphas(item5fr, ncat=3, nitem=5)
print(myAlphas)

myBetas<-startbetas(item5fr, ncat=3, nitem=5)
print(myBetas)

nrbcplnout<-nrbcpln(item5fr, ncat=3, nitem=5, alphas=myAlphas, betas=myBetas, se=FALSE)
print(nrbcplnout)
```

Index