Package ‘extRC’

July 25, 2020

Type Package

Title Extended RC Models for Contingency Tables

Version 1.1

Date 2020-07-17

Author Francesco Bartolucci, Antonio Forcina

Maintainer Francesco Bartolucci <francesco.bartolucci@unipg.it>

Description Maximum likelihood estimation of an extended class of row-column (RC) association models for two-dimensional contingency tables, which are formulated by a condition of reduced rank on a matrix of extended association parameters; see Forcina (2019) <arXiv:1910.13848>. These parameters are defined by choosing the logit type for the row and column variables among four different options and a transformation derived from suitable divergence measures.

License GPL (>= 2)

Imports MASS

NeedsCompilation no

Repository CRAN

Date/Publication 2020-07-25 04:10:06 UTC

R topics documented:

`extRC-package` *Overview of the Package extRC*

Description

Estimation of extended RC models, which are formulate by constraining different types of association parameters to have a reduced rank.

Details

The package contains functions for maximum likelihood (ML) estimation of an extended class of row-column (RC) association models for two-dimensional contingency tables, as described in Forcina (2019). These models are formulated by a condition of reduced rank on a matrix of extended association parameters, which are defined by choosing the logit type for the row and column variables among four different options and a transformation derived from Cressie and Read (1984). Among the available alternatives, it is possible to use log-odds ratio based on different types of aggregation of the joint probabilities. The class of models generalizes that proposed in Kateri and Papaioannou (1994), Bartolucci and Forcina (2002), and Espendiller (2017), and includes the original RC association models of Goodman (1979) and the correspondence analysis model, as formulated in Goodman (1981) and Gilula et al. (1988). Maximum likelihood estimation is based on an algorithm that is an adaptation of the Aitchison and Silvey (1958) algorithm for constrained ML estimation and is related to the algorithm described in Evans and Forcina (2013) for fitting constrained marginal models.

The main function in the package is `extRC` that provides an output that may be shown by usual R commands `print`, `summary`, and `plot`.

Author(s)

Francesco Bartolucci, Antonio Forcina

Maintainer: Francesco Bartolucci <francesco.bartolucci@unipg.it>

References

Aitchison, J. and Silvey (1958). Maximum-likelihood estimation of parameters subject to restraints. *The Annals of Mathematical Statistics*, **29**, 813-828.

Bartolucci, F. and Forcina, A. (2002). Extended RC association models allowing for order restrictions and marginal modeling. *Journal of the American Statistical Association*, **97**, 1192-1199.

Cressie, N. and Read, T.R., 1984. Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society: Series B*, **46**, 440-464.

Espendiller, M., 2017. *Association in contingency tables*. Ph.D. thesis.

Evans, R.J. and Forcina, A. (2013). Two algorithms for fitting constrained marginal models. *Computational statistics & Data analysis*, **66**, 1-7.

Forcina (2019), An extended class of RC association models: definition and estimation, arXiv:1910.13848.

Gilula, Z., Krieger, A.M., and Ritov, Y., 1988. Ordinal association in contingency tables: Some interpretive aspects. *Journal of the American Statistical Association*, **83**, 540-545.

Goodman, L.A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, **74**, 537-552.

Goodman, L.A., 1981. Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, **76**, 320-334.

Kateri, M. and Papaioannou, T. (1994). *f-divergence Association Models*. University of Ioannina.

Examples

```
# load data
data(mobility)

# fit model for a single la
out = extRC(mobility,mod=c("l","l"),k=1,la=0.6)
summary(out)
```

`cuby` *Step length*

Description

Internal function that computes step length of the estimation algorithm in `extRC` by fitting a cubic polynomial.

Usage

```
cuby(g)
```

Arguments

`g`	vector of likelihood values at different step lengths

Value

`comp1`	optimal length

Author(s)

Francesco Bartolucci, Antonio Forcina

`Deta` *Computation of marginal parameters*

Description

Given a vector of canonical parameters coding distribution for an IxJ contingency table and the RC model specification in list `Model`, it computes vector of marginal and joint parameters and matrix of its derivatives with respect the canonical parameters.

Usage

```
Deta(the, Model, der = FALSE)
```

Arguments

`the`	vector of canonical parameters
`Model`	list specifying all model components
`der`	to require derivative computation (optional)

Value

`eta`	vector of marginal parameters
`Der`	derivative matrix with respect to canonical parameters

Author(s)

Francesco Bartolucci, Antonio Forcina

`dfm` *First difference matrix*

Description

It creates a matrix of first differences of order k.

Usage

```
dfm(k)
```

Arguments

`k`	size of the matrix

Value

`D`	first difference matrix

Author(s)

Francesco Bartolucci, Antonio Forcina

Examples

```
D = dfm(5)
x = runif(5)
(D%*%x)
```

Drank *Check matrix rank*

Description

Given the row vectorized matrix, it computes the vector of discrepancies with respect to a certain rank and its derivative.

Usage

```
Drank(ga, lev, k, der = FALSE)
```

Arguments

`ga`	row vectorized matrix of interaction
`lev`	vector of the number of row and column categories in the original table (the numbers of rows and columns of the input matrix must be increased by 1)
`k`	matrix rank
`der`	to require derivative

Value

`fr`	vector of discrepancies with respect to the rank
`Dfr`	derivative of fr

References

Bartolucci, F. and Forcina, A. (2002). Extended RC association models allowing for order restrictions and marginal modeling. *Journal of the American Statistical Association*, **97**, 1192-1199.

Examples

```
A = matrix(rnorm(12),4) # matrix the rank of which must be checked
a = as.vector(t(A))
out = Drank(a,c(5,4),1,der=TRUE)
(out$fr)
(out$Dfr)
```

extRC *Extended RC model*

Description

Main function that fits extended RC models based on different types of aggregation (continuation, local, global) and different divergence functions defined by a suitable value of lambda.

Usage

```
extRC(N, mod, k, la, marg.cons = c("free","equal","shift"))
```

Arguments

N	observed contingency table
mod	vector indicating the types of aggregation for row and column variables ("c" for continuation, "l" for local, "g" for global)
k	rank required for the matrix of interaction parameters
la	value of lambda parameter
marg.cons	type of constraint on the marginal distributions

Value

la	vector of lambda values (when a vector is in input)
dev	deviance of the fitted model (when only one lambda value is in input) or vector of deviances (when a vector of lambda values is in input)
df	degrees of freedom (when only one lambda value is in input)
it	number of iterations (when only one lambda value is in input)
dis	final discrepancy (when only one lambda value is in input)
pj	vector of joint probabilities under the fitted model (when only one lambda value is in input)
eta	full vector of marginal parameters (when only one lambda value is in input)
etaX	vector of row marginal parameters (when only one lambda value is in input)
etaY	vector of column marginal parameters (whenonly one lambda value is in input)
Eta	matrix of association parameters (when only one lambda value is in input)
la	vector of lambda values (when more lambda values are in input)
dev	vector of deviance values (when more lambda values are in input)

Author(s)

Francesco Bartolucci, Antonio Forcina

Examples

```
# load data
data(mobility)

# for a single value of lambda, fit model with constraints of rank 1 on
# local-local logits and without constraints on the marginal distributions
out = extRC(mobility,mod=c("l","l"),k=1,la=0.6)
summary(out)

# for a single value of lambda, fit model with constraints of rank 1 on
# local-local logits and under constrain of equal marginal distributions
out = extRC(mobility,mod=c("l","l"),k=1,la=0.6,marg.cons="equal")
summary(out)

# for a single value of lambda, fit model with constraints of rank 2 on
# global-global logits and under constraint that marginal distributions
# are equal up to a constant shift
out = extRC(mobility,mod=c("g","g"),k=2,la=0.6,marg.cons="shift")
summary(out)

# fit model for a vector of lambdas
la = seq(-1.8,0.6,length.out=10)
out1 = extRC(mobility,mod=c("l","l"),k=1,la=la)
plot(out1)
```

Hmat *Matrix algebra transformation*

Description

Internal function that performs a matrix algebra transformation that is used for estimation in `extRC`.

Usage

```
Hmat(G)
```

Arguments

`G`	input matrix

Value

`H`	transformed matrix

Author(s)

Francesco Bartolucci, Antonio Forcina

MainRC *Estimation of extended RC models*

Description

Internal function that implements the Aitchinson-Silvey algorithm to estimate extended RC models.

Usage

```
MainRC(y, Model, the0 = NULL, output = FALSE)
```

Arguments

y	row vectorized vector of frequencies of the contingency table
Model	list of model components
the0	initial vector of canonical parameters (optional)
output	to require full output (optional)

Value

dev	final deviance
df	degrees of freedom
pj	vector of joint probabilities
it	number of iterations
dis	final discrepancy

Author(s)

Francesco Bartolucci, Antonio Forcina

MatIn *Aggregation matrices*

Description

Computation of aggregation matrices for generalized interactions that are used in codeextRC to estimate extended RC models.

Usage

```
MatIn(lev, mod)
```

Arguments

`lev`	vector number of rows and columns
`mod`	type of logit for each dimension

Value

`R0`	aggregation matrix for the row margin upper level
`R1`	aggregation matrix for the row margin lower level
`C0`	aggregation matrix for the column margin upper level
`C1`	aggregation matrix for the column margin lower level
`J00`	aggregation matrix for the left upper quadrant
`J01`	aggregation matrix for the rigth upper quadrant
`J10`	aggregation matrix for the left lower quadrant
`J11`	aggregation matrix for the right lower quadrant

Author(s)

Francesco Bartolucci, Antonio Forcina

`mobility` *Social mobility data*

Description

Social mobility table of 3,500 British individuals, who are cross-classified according to their occupational status and the occupation status of their fathers.

Usage

```
data("mobility")
```

Format

The format is: num [1:5, 1:5] 50 28 11 14 3 45 174 78 150 42 ... - attr(*, "dimnames")=List of 2 ..$: chr [1:5] "F1" "F2" "F3" "F4" ... ..$: chr [1:5] "S1" "S2" "S3" "S4" ...

References

Mosteller, F. (1968). Association and estimation in contingency tables. *Journal of the American Statistical Association*, **63**, 1-28.

Bartolucci, F. and Forcina, A. (2002). Extended RC association models allowing for order restrictions and marginal modeling. *Journal of the American Statistical Association*, **97**, 1192-1199.

`plot` *Plot the output*

Description

It plots the output of codeextRC function for a vector of lambda values.

Usage

```
## S3 method for class 'extRC'
plot(x, ...)
```

Arguments

`x`	output from `extRC`
`...`	further arguments passed to or from other methods

Value

None

Author(s)

Francesco Bartolucci, Antonio Forcina

`PraD` *Discrepancy with respect to inequality constraints*

Description

Internal function that, given a vector of joint probabilities from an IxJ table (vectorized by row) and the RC model specification in list Model, computes vector of discrepancies and matrix of its derivatives with respect to the canonical parameters.

Usage

```
PraD(the, Model, der = FALSE)
```

Arguments

`the`	vector of canonical parameters
`Model`	list of model components
`der`	to require the derivative (optional)

Value

`hdis`	vector of discrepancies
`Hdis`	matrix of derivatives of discrepancies with respect to the canonical parameter (optional)

Author(s)

Francesco Bartolucci, Antonio Forcina

`print` *Print the output.*

Description

Given the output of codeextRC function, it is written in a readable form.

Usage

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

Arguments

`x`	output from `extRC`
`...`	further arguments passed to or from other methods

Value

`None`

Author(s)

Francesco Bartolucci, Antonio Forcina

summary — *Summary of extRC fits*

Description

Summary method for the output of codeextRC function.

Usage

```
## S3 method for class 'extRC'
summary(object, ...)
```

Arguments

`object`	output from `extRC`
`...`	further arguments passed to or from other methods

Value

None

Author(s)

Francesco Bartolucci, Antonio Forcina

tril — *Lower triangular matrix*

Description

Given a square matrix, it provides the lower triangular part, including the main diagonal.

Usage

```
tril(M)
```

Arguments

`M`	square matrix

Value

`N`	transformed matrix

Author(s)

Francesco Bartolucci, Antonio Forcina

Examples

```
M = matrix(1:9,3)
N = tril(M)
```

Index