Package ‘FMsmsnReg’

March 30, 2016

Type Package

Title Regression Models with Finite Mixtures of Skew Heavy-Tailed Errors

Version 1.0

Date 2016-03-29

Author Luis Benites Sanchez and Rocío Paola Maehara and Victor Hugo Lachos

Maintainer Luis Benites Sanchez <lbenitesanchez@gmail.com>

Imports mvtnorm

Description Fit linear regression models where the random errors follow a finite mixture of Skew Heavy-Tailed Errors.

License GPL (>= 2)

Repository CRAN

NeedsCompilation no

Date/Publication 2016-03-30 08:14:20

R topics documented:

FMsmsnReg-package ......................................................... 2
ais ................................................................. 3
FMsmsnReg ......................................................... 4
horses ............................................................. 6

Index 8
Description

This package contains a principal function that performs to estimate the parameters of a regression model considering an error that follows a finite mixture of Skew Heavy-Tailed Errors, using an analytically simple and efficient EM-type algorithm for iteratively computing maximum likelihood estimates of the parameters.

Details

Package: FMmsnReg
Type: Package
Version: 1.0
Date: 2016-03-30
License: GPL (>=2)

Author(s)

Luis Benites Sanchez <lbenitesanchez@gmail.com>, Rocio Paola Maehara <rmaeharaa@gmail.com> and Victor Hugo Lachos <hlachos@ime.unicamp.br>

References


See Also

FMmsnReg

Examples

#See examples for the FMmsnReg function linked above.
Description

Data on 102 male and 100 female athletes collected at the Australian Institute of Sport.

Format

This data frame contains the following columns:

- **Sex** (0 = male or 1 = female)
- **Ht** height (cm)
- **Wt** weight (kg)
- **LBM** lean body mass
- **RCC** red cell count
- **WCC** white cell count
- **Hc** Hematocrit
- **Hg** Hemoglobin
- **Ferr** plasma ferritin concentration
- **BMI** body mass index, weight/height^2
- **SSF** sum of skin folds
- **Bfat** Percent body fat
- **Label** Case Labels
- **Sport** Sport

References


Examples

```r
#Load the data
library(FMsmsnReg)
data(ais)
attach(ais)

#Set the response y and covariate x
x1 <- cbind(SSF,Ht)
y <- Bfat

#Fits a linear Regression Model with Finite Mixtures of Skew t
parST <- FMmsnReg(y, x1, g=2, get.init = TRUE, criteria = TRUE,
group = FALSE,family = "Skew.t", error = 10^-4,
```
**Description**

Performs a Finite Mixture of Scale Mixture Skew Normal Regression Model using EM-type algorithm (ECME) for iteratively computing maximum likelihood estimates of the parameters.

**Usage**

```r
fmsmsnReg(y, x1, abetas = NULL, medj= NULL, sigmaR = NULL, shape = NULL, 
pii = NULL, g = NULL, get.init = TRUE, criteria = TRUE, group = FALSE, 
family = "Skew.normal", error = 10^-4, iter.max = 5000, obs.prob= FALSE, kmeans.param = NULL, 
show.converge=FALSE, cp=0.5)
```

**Arguments**

- **y**
  - the response matrix (dimension nx1)
- **x1**
  - Matrix or vector of covariates.
- **Abetas**
  - Parameters of vector regression dimension \((p + 1)\) include intercept
- **medj**
  - a list of \(g\) arguments of vectors of initial values (dimension \(p\)) for the location parameters
- **sigma2**
  - a list of \(g\) arguments of matrices of initial values (dimension \(p \times p\)) for the scale parameters
- **shape**
  - a list of \(g\) arguments of vectors of initial values (dimension \(p\)) for the skewness parameters
- **pii**
  - Initial value for the EM algorithm. Each of them must be a vector of length \(g\). (the algorithm considers the number of components to be adjusted based on the size of these vectors)
- **g**
  - the number of cluster to be considered in fitting
- **get.init**
  - if TRUE, the initial values are generated via k-means
- **criteria**
  - It indicates if are calculated the criterion selection methods (AIC, BIC, EDC and ICL)
group if TRUE, the vector with the classification of the response is returned
family distribution family to be used in fitting ("Skew.t", "Skew.cn", "Skew.slash", "Skew.normal")
error define the stopping criterion of the algorithm
iter.max the maximum number of iterations of the EM algorithm
obs.prob if TRUE, the posterior probability of each observation belonging to one of the g groups is reported
kmeans.param a list with alternative parameters for the kmeans function when generating initial values, list(iter.max = 10, n.start = 1, algorithm = "Hartigan-Wong")
show.convergence graphics of convergence for the parameters
cp Cut Point

Value
The function returns a list with 16 elements detailed as
iter Number of iterations.
criteria Attained criteria value.
convergence Convergence reached or not.
mu Location parameter estimate.
sigma2 Scale parameter estimate.
lambda Shape parameter estimate.
pii Weight parameter estimate.
nu Estimated degrees of freedom parameter.
SE Standard Error estimates, if the output shows NA the function does not provide the standard error for this parameter.
table Table containing the inference for the estimated parameters.
loglik Log-likelihood value.
AIC Akaike information criterion.
BIC Bayesian information criterion.
EDC Efficient Determination Criterion.
ICL Information Completed Likelihood.
time Processing time.

Author(s)
Luis Benites Sanchez <lbenitesanchez@gmail.com> and Rocio Paola Maehara <rmaehara@gmail.com> and Victor Hugo Lachos <hlachos@ime.unicamp.br>
References


See Also

fmsmsnReg, ais, horses

Examples

#See examples for the fmsmsnReg function linked above.

| horses | Horse Racing at Eagle Farm data |

Description

Results of horse races at Eagle Farm, Brisbane, on 31 August 1998. The data, collected by Donald Forbes for his MS305 Data Analysis Project, give results for each horse in a sequence of 8 races.

Format

This data frame contains the following columns:

- **Position**: (Finishing position)
- **Starters**: Number of horses in race
- **Last**: Finishing position in last race
- **Since**: Days since last race
- **Number**: Identifying number of horse in race
- **Carried**: Weight carried
- **Weight**: Handicap weight
- **Barrier**: Barrier position at start of race
- **Distance**: Length of race
- **Lengths**: Number of lengths that horse finished from winner
- **Odds**: Starting odds
- **Starts**: Number of races previously started in
- **Age**: Age of horse in years
- **Ratio**: Proportion of wins in previous starts
horses

References


Examples

```r
# Load the data
library(FMsmsnReg)
data(horses)
attach(horses)

# Set the response y and covariate x
x1 <- cbind(1, Last, Carried)
y <- Position

# Fits a linear Regression Model with Finite Mixtures of Skew Contaminated Normal
parCN <- FMmsnReg(y, x1, g=2, get.init = TRUE, criteria = TRUE, group = FALSE,
                  family = "Skew.cn", error = 10^-4, iter.max = 5000, obs.prob = FALSE,
                  kmeans.param = NULL, show.converge = FALSE, cp=0.5)

# Fits a linear Regression Model with Finite Mixtures of Skew normal
parSN <- FMmsnReg(y, x1, g=2, get.init = TRUE, criteria = TRUE,
                  group = FALSE, family = "Skew.normal", error = 10^-4,
                  iter.max = 5000, obs.prob = FALSE, kmeans.param = NULL, show.converge = FALSE, cp=0.5)
```
Index

*Topic EM
  FMsmnReg, 4
*Topic Finite Mixture
  FMsmnReg, 4
*Topic datasets
  ais, 3
  horses, 6
*Topic package
  FMsmnReg, 4
  FMsmnReg-package, 2
*Topic smsn
  FMsmnReg, 4
  ais, 3, 6
  FMsmnReg, 2, 4, 6
  FMsmnReg-package, 2
  horses, 6, 6