Package ‘LMest’
February 15, 2013

Type Package

Title Fit Latent Markov models in basic versions

Version 1.0

Date 2012-11-27

Author Francesco Bartolucci, University of Perugia (IT), http://www.stat.unipg.it/bartolucci

Maintainer Francesco Bartolucci <bart@stat.unipg.it>

Description Fit certain versions of the Latent Markov model for longitudinal categorical data.

License GPL (>= 2)

Depends MASS, Matrix

Repository CRAN

Date/Publication 2012-11-28 07:17:00

NeedsCompilation no

R topics documented:

LMest-package .......................................................... 2
aggr_data ............................................................... 3
bootstrap_lm_basic ..................................................... 3
complk ................................................................. 5
data_criminal .......................................................... 6
data_drug ............................................................... 7
draw_lm_basic ........................................................... 8
est_lm_basic ............................................................ 9
invglob ................................................................. 10
lk_obs ................................................................. 11
long2wide ............................................................. 12
recursions ............................................................. 13

Index 14
LMest-package  

Fit latent Markov models

Description

Set of functions to fit latent Markov models in the basic version.

Details

Package:  LMest  
Type:  Package  
Version:  1.0  
Date:  2012-11-27  
License:  Unlimited

The most important functions is est_lm_basic which estimates the basic LM model

Author(s)

Francesco Bartolucci, University of Perugia (IT), http://www.stat.unipg.it/bartolucci  
Maintainer: Francesco Bartolucci <bart@stat.unipg.it>

References


Examples

# Example of drug consumption data  
# load data  
data(data_drug)  
data_drug = as.matrix(data_drug)  
S = data_drug[,1:5]-1  
yv = data_drug[,6]  
n = sum(yv)  
# fit of the Basic LC model  
k = 3  
out1 = est_lm_basic(S,yv,k,mod=1)  
## Not run:  
out2 = bootstrap_lm_basic(out1$piv,out1$Pi,out1$Psi,n,mod=1,B=1000)  
## End(Not run)
aggr_data

Function that aggregates data

Description
Given a matrix of configurations (covariates and responses) unit by unit, this function finds the corresponding matrix of distinct configurations and the corresponding vector of frequencies.

Usage
aggr_data(data)

Arguments
data matrix of covariate and response configurations unit by unit

Value
data_dis matrix of distinct configurations
data
freq vector of corresponding frequencies
label the index of each provided response among the distinct ones

Author(s)
Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

Examples
# draw a matrix of random responses and find distinct responses
X = matrix(sample(5,100,replace=TRUE),50,2)
out = aggr_data(X)

bootstrap_lm_basic

Parametric bootstrap for the basic LM model

Description
Function that performs bootstrap parametric resampling to compute standard errors for the parameter estimates.

Usage
bootstrap_lm_basic(piv, Pi, Psi, n, B = 100, start = 0, mod = 0, tol = 10^-6)
Arguments

- **piv**: initial probability vector
- **Pi**: probability transition matrices
- **Psi**: matrix of conditional response probabilities
- **n**: sample size
- **B**: number of bootstrap samples
- **start**: type of starting values (0 = deterministic, 1 = random)
- **mod**: model on the transition probabilities (0 for time-heter., 1 for time-homog., 2-(T-1) partial homog. of that order)
- **tol**: tolerance level for convergence

Value

- **mPsi**: average of bootstrap estimates of the conditional response matrix
- **mpiv**: average of bootstrap estimates of the initial probability vector
- **mPi**: average of bootstrap estimates of the transition probability matrices
- **sePsi**: standard errors for the conditional response matrix
- **sepiv**: standard errors for the initial probability vector
- **sePi**: standard errors for the transition probability matrices

Author(s)

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

Examples

```r
## Not run:
# Example of drug consumption data
# load data
data(data_drug)
data_drug = as.matrix(data_drug)
S = data_drug[,1:5]-1
yv = data_drug[,6]
n = sum(yv)
# fit of the Basic LC model
k = 3
out1 = est_lm_basic(S,yv,k,mod=1)
out2 = bootstrap_lm_basic(out1$piv,out1$Pi,out1$Psi,n,mod=1,B=1000)

## End(Not run)
```
complk

Complete log-likelihood of the basic latent Markov model

Description

Function that computes complete log-likelihood of the latent Markov model (internal use).

Usage

complk(S, yv, piv, Pi, Psi, k)

Arguments

S    matrix of distinct response configurations
yv   corresponding vector of frequencies
piv  vector of initial probabilities
Pi   transition probability matrix
Psi  conditional response probabilities
k    number of latent classes

Value

lk    log-likelihood
Phi   matrix of the conditional probabilities of the observed response configurations
L     matrix of the forward probabilities
pv    vector of marginal probabilities

Author(s)

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

References

Description

Dataset about crimes committed by a cohort of subjects in Galles and Wales.

Usage

data(data_criminal)

Format

A data frame with 17947 observations on the following 13 variables.

<table>
<thead>
<tr>
<th>id</th>
<th>subject id</th>
</tr>
</thead>
<tbody>
<tr>
<td>sex</td>
<td>gender of the subject</td>
</tr>
<tr>
<td>time</td>
<td>occasion of observation</td>
</tr>
<tr>
<td>y1</td>
<td>crime of type 1</td>
</tr>
<tr>
<td>y2</td>
<td>crime of type 2</td>
</tr>
<tr>
<td>y3</td>
<td>crime of type 3</td>
</tr>
<tr>
<td>y4</td>
<td>crime of type 4</td>
</tr>
<tr>
<td>y5</td>
<td>crime of type 5</td>
</tr>
<tr>
<td>y6</td>
<td>crime of type 6</td>
</tr>
<tr>
<td>y7</td>
<td>crime of type 7</td>
</tr>
<tr>
<td>y8</td>
<td>crime of type 8</td>
</tr>
<tr>
<td>y9</td>
<td>crime of type 9</td>
</tr>
<tr>
<td>y10</td>
<td>crime of type 10</td>
</tr>
</tbody>
</table>

References


Examples

data(data_criminal)
**data_drug**

---

**Dataset about marijuana consumption**

---

**Description**

Longitudinal dataset about marijuana consumption measured by ordinal variables with 3 categories.

**Usage**

```r
data(data_drug)
```

**Format**

A data frame with 51 observations on the following 6 variables.

- V1 reported drug use at the 1st occasion
- V2 reported drug use at the 2nd occasion
- V3 reported drug use at the 3rd occasion
- V4 reported drug use at the 4th occasion
- V5 reported drug use at the 5th occasion
- V6 frequency of the response configuration

**Source**


**References**


**Examples**

```r
data(data_drug)
```
**draw_lm_basic**  
*Draws samples from LM model*

**Description**

Function that draws samples from the LM model with specific parameters.

**Usage**

```r
draw_lm_basic(piv, Pi, Psi, n)
```

**Arguments**

- `piv` vector of initial probabilities of the latent Markov chain
- `Pi` set of transition matrices
- `Psi` matrix of conditional response probabilities
- `n` sample size

**Value**

- `Y` matrix of response configurations unit by unit
- `S` matrix of distinct response configurations
- `yv` corresponding vector of frequencies

**Author(s)**

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

**Examples**

```r
# draw a sample for 1000 units and only one response variable
piv = c(0.7,0.3)
Pi = matrix(c(0.9,0.1,0.1,0.9),2,2)
Pi = array(Pi,c(2,2,6))
Pi[,1] = 0
Psi = matrix(c(0.7,0.2,0.1,0.5,0.4,0.1),3,2)
Psi = array(Psi,c(3,2,1))
out = draw_lm_basic(piv, Pi, Psi, n=1000)
```
**est_lm_basic**

*Estimate basic LM model*

---

**Description**

Main function for estimating the basic LM model.

**Usage**

```r
est_lm_basic(S, yv, k, start = 0, mod = 0, tol = 10^-6)
```

**Arguments**

- **S**: matrix of available configurations (n by T by r)
- **yv**: vector of frequencies of the available configurations
- **k**: number of latent states
- **start**: type of starting values (0 = deterministic, 1 = random)
- **mod**: model on the transition probabilities (0 for time-heter., 1 for time-homog., 2-(T-1) partial homog. of that order)
- **tol**: tolerance level for convergence

**Value**

- **lk**: maximum log-likelihood
- **piv**: estimate of initial probability vector
- **Pi**: estimate of transition probability matrices
- **Psi**: estimate of conditional response probabilities
- **np**: number of free parameters
- **aic**: value of AIC for model selection
- **bic**: value of BIC for model selection
- **lkv**: log-likelihood trace at every step
- **V**: array containing the posterior distribution of the latent states for each response configuration and time occasion

**Author(s)**

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

**References**

Examples

# Example of drug consumption data
# load data
data(data_drug)
data_drug = as.matrix(data_drug)
S = data_drug[,1:5]-1
yv = data_drug[,6]
n = sum(yv)
# fit of the Basic LC model
k = 3
out = est_lm_basic(S,yv,k,mod=1)

## Not run:
# Example based on criminal data
# load criminal data
data = data(data_criminal)
out = long2wide(data_criminal,"id","time","sex",c("y1","y2","y3","y4","y5","y6","y7","y8","y9","y10"),aggr=T,full=/zero.noslash)
XX = out$XX; for(j in 1:dim(XX)[1]) XX[,j] = max(XX[,j])
YY = out$YY
freq = out$freq
n1 = sum(freq[XX[,1]==1])
n2 = sum(freq[XX[,1]==2])
nc1 = floor(n1/0.34); nc2 = floor(n2/0.085);
dd = dim(YY)
YY1 = YY; YY = array(/zero.noslash,c(dd[1]+2,dd[2],dd[3]))
YY[3:(2+dd[1]),,] = YY1
XX = rbind(rep(1,dd[2]),rep(2,dd[2]),XX)
freq = c(nc1-n1,nc2-n2,freq)
n = sum(freq)
# fit basic LM model with increasing number of states to select the most suitable
Res/zero.noslash = vector("list",7)
for(k in 1:7){
    Res/zero.noslash[[k]] = est_lm_basic(YY,freq,k,mod=1,tol=10^-4)
    save(list = ls(),file="example_criminal_temp.RData")
}
out1 = Res/zero.noslash[[6]]

## End(Not run)

invglob

Invert vector of global logits.

Description

Function to invert the marginal parametrization based on global logits (internal function).

Usage

invglob(eta)
lk_obs

Arguments
eta  vector of global logits

Value
p  vector of joint probability

Author(s)
Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

lk_obs  Compute the observable log-likelihood of the LM model

Description
Function that computes the observable log-likelihood of the LM model (not used in this version).

Usage
lk_obs(th, Am, Bm, Cm, b, k, S, yv, T, r, mod)

Arguments
th  vector of parameters
Am  design matrix for the logits
Bm  design matrix for the logits
Cm  design matrix for the logits
b  number of response categories
k  number of states
S  matrix of distinct response configurations
yv  corresponding vector of frequencies
T  number of times occasions
r  number of response variables
mod  type of model

Value
lk  log-likelihood
sc  score vector

Author(s)
Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci
long2wide

From data in the long format to data in the wide format

Description

Function that transforms data in the long format to data in the wide format.

Usage

long2wide(data, nameid, namet, colx, coly, aggr = T, full = 999)

Arguments

data matrix of data
nameid name of the id column
namet name of the t column
colx vector of the names of the columns of the covariates
coly vector of the names of the columns of the responses
aggr if wide aggregated format is required
full number to use for missing data

Value

listid list of id for every unit
listt list of the time occasions
data_wide data in wide format
XX array of the covariates
YY array of the responses
freq vector of the corresponding frequencies

Author(s)

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

Examples

# Example based on criminal data
# load criminal data
data = data(data_criminal)
# consider only the first 1000 records to shorten time
out = long2wide(data_criminal[1:1000,],"id","time","sex",c("y1","y2","y3","y4","y5","y6","y7","y8","y9","y10")
<table>
<thead>
<tr>
<th>recursions</th>
<th>Recursions used by est_lm_basic</th>
</tr>
</thead>
</table>

**Description**

Implementation of the recursions to compute manifest probability of the responses and posterior distribution of the latent states (internal function).

**Usage**

```
recursions(S, yv, Psi, piv, Pi, k, lth, Am, Bm, Cm, b)
```

**Arguments**

- **S**: matrix of distinct response configurations
- **yv**: corresponding vector of frequencies
- **Psi**: matrix of conditional response probabilities
- **piv**: initial probability vector
- **Pi**: transition probability matrices
- **k**: number of latent states
- **lth**: internal argument
- **Am**: internal argument
- **Bm**: internal argument
- **Cm**: internal argument
- **b**: internal argument

**Value**

- **lk**: log-likelihood
- **sc**: score vector
- **F1**: internal argument
- **F2**: internal argument
- **F1d**: internal argument
- **F2d**: internal argument

**Author(s)**

Francesco Bartolucci, University of Perugia, http://www.stat.unipg.it/bartolucci

**References**

Index

*Topic **datasets**
  data_criminal, 6
  data_drug, 7

*Topic **package**
  LMest-package, 2

aggr_data, 3
bootstrap_lm_basic, 3
complk, 5
data_criminal, 6
data_drug, 7
draw_lm_basic, 8
est_lm_basic, 9
invglob, 10
lk_obs, 11
LMest (LMest-package), 2
LMest-package, 2
long2wide, 12
recursions, 13