The mmlcr Package

April 19, 2006

Version 1.3.5

Date 2006-04-10

Title Mixed-Mode Latent Class Regression

Author Steve Buyske <buyske@stat.rutgers.edu>

Maintainer Steve Buyske <buyske@stat.rutgers.edu>

Depends MASS, nnet, nlme, survival

Description Mixed-mode latent class regression (also known as mixed-mode mixture model regression or mixed-mode mixture regression models) which

License GPL Version 2 or later.

URL http://www.stat.rutgers.edu/~buyske/software.html

R topics documented:

BIC.mmlcr ................................................................. 2
Sredirect ................................................................. 3
dcnorm ................................................................. 4
dnb2 ................................................................. 5
fitted.mmlcr ......................................................... 5
mmlcr ................................................................. 6
plot.mmlcronce ..................................................... 8
mmlcrObject ......................................................... 9
mmlcrclassify ......................................................... 10
mmlcrcomponentinit ................................................ 10
mmlcrdf ............................................................... 11
mmlcrfit ............................................................... 12
mmlcrlike ............................................................. 13
plot.mmlcr ............................................................. 14
plot.mmlcrlong ...................................................... 14
postprob .............................................................. 15
theta.mmmmod ....................................................... 16

Index 17
BIC.mmlcr

Bayesian Information Criterion

Description

This generic function calculates the Bayesian information criterion, also known as Schwarz’s Bayesian
criterion (SBC), for an mmlcr object for which a log-likelihood value can be obtained, according
to the formula $-2\log\text{likelihood} + n_{\text{par}} \log(n_{\text{obs}})$, where $n_{\text{par}}$ represents the number of parameters
and $n_{\text{obs}}$ the number of observations in the fitted model.

Usage

```r
## S3 method for class 'mmlcr':
BIC(object, ...)
```

Arguments

- `object` a fitted mmlcr object.
- `...` optional fitted model objects.

Value

if just one object is provided, returns a numeric value with the corresponding BIC; if more than one
object are provided, returns a `data.frame` with rows corresponding to the objects and columns
representing the number of parameters in the model (`df`) and the BIC.

References


See Also

AIC, mmlcrObject

Examples

```r
## Not run: data(mmlcrdf)
## Not run:
mmlcrdf.mmlcr2 <- mmlcr(outer = ~ sex + cov1 | id,
components = list(
  list(formula = resp1 ~ 1, class = "cnormonce", min = 0, max = 50),
  list(formula = resp2 ~ poly(age, 2) + tcov1, class = "poislong"),
  list(formula = resp3 ~ poly(age, 2), class = "multinomlong")
), data = mmlcrdf, n.groups = 2)
## End(Not run)

## Not run: BIC(mmlcrdf.mmlcr2)
```
Longitudinal Dataset of Aggression and Reading Ability in Children.

Description

The `Srcdfull` data frame has 1620 rows and 9 columns. It is a longitudinal dataset from 405 children with scores for aggression and reading ability.

Usage

data(Srcdfull)

Format

This data frame contains the following columns:

- **id**: Subject id.
- **sex**: A factor variable with the sex of each subject.
- **age**: Age of subject at time of interview.
- **anti**: Antisocial score, an integer from 0 to 12.
- **read**: Reading ability score, ranging from 0 to 8.4.
- **momage**: Mother’s age at first interview.
- **kidage**: Subject’s age at first interview.
- **homecog**: An integer score, ranging from 0 to 14, at the first interview for cognitive stimulation at home.
- **homeemo**: An integer score, ranging from 0 to 14, at the first interview for emotional support at home.

Details

The sample consists of N=405 children drawn from the Children of the National Longitudinal Survey of Youth, about half of whom are missing one or more of the repeated measures on aggression or reading ability. The measures of interest include four repeated measures of aggressive behavior taken at two-year intervals, four repeat measures of reading achievement taken at two-year intervals, initial interview measures of cognitive stimulation and social support of the child in the home, and background measures of child gender, child age, and mother age.

Source

http://www.unc.edu/~curran/srcd-docs/srcddata.txt

References

http://www.unc.edu/~curran/srcd-docs/srcdmeth.pdf
Examples

data(Srcdfull)
Srcdfull.mmlcr4 <- mmlcr(outer = ~1 | id, components = list(
  list(formula = anti ~ poly(age, 2), min = 0, max = 12,
    class = "cnormlong"),
  list(formula = read ~ poly(age, 2), class = "normlong"),
  data = Srcdfull, n.groups = 4))

dcnorm

Density function for Censored Normal Distribution

Description

Density function for the censored (not truncated!) normal distribution in a form suitable for mmlcrlike

Usage

dcnorm(x, mean = rep(0, length(x)), sigma = rep(1,
  length(x)), min = rep(-(.Machine$double.xmax),
  length(x)), max = rep(.Machine$double.xmax,
  length(x)))

Arguments

  x        vector of quantiles
  mean     vector of means
  sigma    vector of standard deviations
  min      vector of left censoring cutoffs
  max      vector of right censoring cutoffs

Value

A vector of densities.

See Also

mmlcr, mmlcrlike

Examples

dcnorm(0, mean = 0, sigma = 1)
dcnorm(0, mean = 0, sigma = 1, min = 0)
Description

The density for the negative binomial function, in a form suitable for the mmlcrlike function.

Usage

dnb2(y, mu, th)

Arguments

y Observations
mu The mean of the distribution
th The parameter $\theta$.

Details

The variance function of the distribution is $\mu + \mu^2 / \theta$.
The density is

$$f(y) = \frac{\Gamma(\theta + y)\mu^y\theta^\theta}{\Gamma(\theta)y!(\mu + \theta)!\mu + \theta)}.$$  

Value

The density.

References


See Also

mmlcr, mmlcrlike

---

fitted.mmlcr

Extract mmlcr Model Component Fitted or Residual Values

Description

Extract fitted or residual values from components of a mmlcr object.

Usage

fitted.mmlcr(object, which = 1:length(object$components), ...)
residuals.mmlcr(object, which = 1:length(object$components), ...)
Arguments

- **object**: Fitted `mmlcr` object.
- **which**: Which components to extract fit or residuals from
- **...**: Included for compatibility with generic functions.

Details

For each observation in a component, the fit is the weighted mean of the fits from the individual class fits. The weights are the posterior class membership probabilities.

Value

A list, with a vector of fits or residuals for each component in which.

**WARNING**

Although the function works on multinom components, the results are not meaningful.

### mmlcr

**Mixed Mode Latent Class Regression**

Description

This function fits a latent class regression model, where the responses may be of different types, as well as longitudinal or time-invariant.

Usage

```r
mmlcr(object = NULL, ...)  
## Default S3 method:  
mmlcr(object = NULL, outer, components,  
      data = error("data must be given (as a data.frame)"),  
      subset, n.groups = 2, prior.prob = NULL,  
      post.prob = NULL, no.fit = FALSE, max.iter = 50,  
      trace = TRUE, tol = 0.005, ...)  
## S3 method for class 'mmlcr':  
mmlcr(object, max.iter = 50, trace = TRUE, tol = 0.005, ...)
```

Arguments

- **outer**: a formula of the form `~ predictors | id` where the model statement for predictors is as for regression models and `id` is the subject factor. The expression `1` can be used for the predictors when there is no other suitable candidate.
- **components**: a list of the response components. Each element of the list is itself a list. At a minimum, that list consists of a formula of the form `resp ~ covars`, where `resp` is the response variable for that component and `covars` are any additional covariates (`resp ~ 1` is commonly used, or something like `resp ~ age` for longitudinal classes), and a class. Implemented classes include `cnormlong`, `cnormonce`, `multinomlong`, `multinomonce`, `nb1long`, `nblong`, `nbonce`, `normlong`, `normonce`, `poislong`, and `poissonc`. Some classes require additional list elements. Details can be found with the `mmlcrcomponentinit` help files.
### mmlcr

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>data</code></td>
<td>a data frame in which to interpret the variables occurring in outer and in</td>
</tr>
<tr>
<td></td>
<td>the formulas for the individual components.</td>
</tr>
<tr>
<td><code>subset</code></td>
<td>expression giving which subset of the rows of the data should be used in</td>
</tr>
<tr>
<td></td>
<td>the fit. All observations are included by default.</td>
</tr>
<tr>
<td><code>n.groups</code></td>
<td>the number of latent classes.</td>
</tr>
<tr>
<td><code>prior.prob</code></td>
<td>a data.frame giving the prior probabilities of class membership according</td>
</tr>
<tr>
<td></td>
<td>to covariate information, not “prior knowledge.” The row.names should match</td>
</tr>
<tr>
<td></td>
<td>the subject id given in outer. If no value is given, which is recommended,</td>
</tr>
<tr>
<td></td>
<td>the default is equal probabilities.</td>
</tr>
<tr>
<td><code>post.prob</code></td>
<td>a data.frame giving the posterior probabilities of class membership. The</td>
</tr>
<tr>
<td></td>
<td>row.names should match the subject id given in outer. If <code>post.prob</code> is</td>
</tr>
<tr>
<td></td>
<td>not given, starting values are assigned randomly. If one is fitting a</td>
</tr>
<tr>
<td></td>
<td>modification of a prior fit called <code>fit1</code>, perhaps by changing the</td>
</tr>
<tr>
<td></td>
<td>covariates, the recommendation is to include the term <code>post.prob = fit1$post.prob</code>.</td>
</tr>
<tr>
<td><code>no.fit</code></td>
<td>if TRUE, returns a <code>mmlcrObject</code> without fitting. The <code>mmlcrObject</code> could</td>
</tr>
<tr>
<td></td>
<td>be fit later by a call to <code>mmlcr(mmlcrObject)</code>.</td>
</tr>
<tr>
<td><code>object</code></td>
<td>an <code>mmlcr</code> object. See the <code>mmlcrObject</code> help file for details.</td>
</tr>
<tr>
<td><code>max.iter</code></td>
<td>the maximum number of iterations</td>
</tr>
<tr>
<td><code>trace</code></td>
<td>if TRUE, traces through the iterations, giving the loglikelihood, a</td>
</tr>
<tr>
<td></td>
<td>convergence speed index, the loglikelihood goal, and the current class</td>
</tr>
<tr>
<td></td>
<td>percentages.</td>
</tr>
<tr>
<td><code>tol</code></td>
<td>the tolerance between the loglikelihood and the loglikelihood goal, used as</td>
</tr>
<tr>
<td></td>
<td>a stopping criterion.</td>
</tr>
<tr>
<td>...</td>
<td>possible additional arguments.</td>
</tr>
</tbody>
</table>

### Details

The components portion of the call refers to components of the response, not components in the sense, used in much of the mixtures literature, that is here called latent classes or groups.

It is not yet possible to specify a prior, or even starting values, for the parameters of the individual latent classes. Instead, one can assign starting values to the posterior probabilities of each individual via the `post.prob` part of the function call.

In a typical use of this package, one might want to model, say, alcohol use. The idea is that there may be several model trajectories over the years, such as little or no use, heavy use, adolescent-limited use, and so on. For each class, we would like to model a different longitudinal regression (i.e., the same form for the regression, but different coefficients for each latent class). Furthermore, we would like to include covariates for class membership, so that the model looks like

\[
\text{Covariates} \implies \text{Latent Class} \implies \text{Longitudinal Trajectory},
\]

with the regression coefficients on the right independent of the covariates on the left conditional on class membership. One could potentially have a number of arrows leading off of the latent class, so that one could simultaneously model cigarette and alcohol use, for example. The first arrow is modeled with the outer formula, which the second arrow is modeled with the formula(s) in the component term.

### Value

A fitted `mmlcrObject`

### See Also

`mmlcrObject, mmlcrcomponentinit, summary.mmlcr, plot.mmlcr`
Examples

data(mmlcrdf)

mmlcrdf.mmlcr2 <- mmlcr(outer = ~ sex + cov1 | id,
components = list(
  list(formula = resp1 ~ 1, class = "cnormonce", min = 0, max = 50),
  list(formula = resp2 ~ poly(age, 2) + tcov1, class = "poislong"),
  list(formula = resp3 ~ poly(age, 2), class = "multinomlong")
), data = mmlcrdf, n.groups = 2)

mmlcrdf.mmlcr2.inter <- mmlcr(outer = ~ sex * cov1 | id,
components = list(
  list(formula = resp1 ~ 1, class = "cnormonce", min = 0, max = 50),
  list(formula = resp2 ~ poly(age, 2) + tcov1, class = "poislong"),
  list(formula = resp3 ~ poly(age, 2), class = "multinomlong")
), data = mmlcrdf, n.groups = 2,
post.prob = mmlcrdf.mmlcr2$post.prob, no.fit = TRUE)

mmlcrdf.mmlcr2.inter <- mmlcr(m.mmlcrdf.mmlcr2.inter)

---

**plot.mmlcronce**

*Plot a mmlcronce Component*

Description

Generates a plot for mmlcronce component of a mmlcr object, showing weighted means plus and minus one standard deviation for each class. Generally called by `plot.mmlcr`.

Usage

```r
plot.mmlcronce(x, post.prob, ...)
```

Arguments

- `x` A mmlcr object.
- `post.prob` Data.frame of class membership posterior probabilities.
- `...` Other arguments to be passed to `plot`.

Note

There is no component plot for the multinomonce components.

See Also

`plot.mmlcr`, `mmlcr`, `mmlcrObject`
**Description**

Classes of objects returned by fitting mixed mode latent class regression objects. This class of objects is returned by the `mmlcr` function to represent a fitted mixed mode latent class regression model. Objects of this class have methods for the functions `print`, `plot`, `summary`, `anova`, `vcov`, `fitted`, and `residuals`, among others.

**Details**

The following components must be included in a legitimate `mmlcr` object.

- **components**: a list of the response components. Each element of the list is itself a list, which includes the data for that response, the fit coefficients, the fitted response and the response residuals, among others.

- **outer**: a formula for latent class membership.

- **outer.df**: a data.frame with the variables used in outer.

- **gamma.matrix**: the coefficients from the multinomial fit for latent classes. The values are log(odds ratios) compared to the first class.

- **prior.prob**: a data.frame giving the prior probabilities (that is, based on the fit from outer) of class membership for each subject.

- **post.prob**: a data.frame giving the posterior probabilities (that is, based on the prior probabilities as well as the data from the components) of class membership for each subject.

- **loglikelihood**: the loglikelihood of the final fit.

- **BIC**: the BIC of the final fit.

- **df**: the degrees of freedom of the model.

- **resid.df**: the residual degrees of freedom of the model.

- **call**: the call to the `mmlcr` function.

**See Also**

- `mmlcr`
mmlcrclassify

**Classification of Individuals by Modal Posterior Probabilities**

**Description**

This function gives a latent class classification for each individual by giving the class to which the individual has the highest posterior probabilities.

**Usage**

```r
mmlcrclassify(object)
```

**Arguments**

- `object` an `mmlcrObject`.

**Value**

a data.frame with row.names equal to the ids and with a single column giving the classification

**See Also**

`mmlcrObject`, `mmlcr`

**Examples**

```r
## Not run:
data(mmlcrdf)

mmlcrdf.mmlcr2 <- mmlcr(outer = ~ sex + cov1 | id, 
components = list(
    list(formula = resp1 ~ 1, class = "cnormonce", min = 0, max = 50),
    list(formula = resp2 ~ poly(age, 2) + tcov1, class = "poislong"),
    list(formula = resp3 ~ poly(age, 2), class = "multinomlong")
), data = mmlcrdf, n.groups = 2)

mmlcrclassify(mmlcrdf.mmlcr2)## End(Not run)
```

mmlcrcomponentinit

**Initialize an 'mmlcr' component**

**Description**

Used internally by `mmlcr` to initialize an `mmlcr` component. The function is generic.

**Usage**

```r
mmlcrcomponentinit(object, n.groups, prob, data, grouping)
```
Arguments

object see details below
n.groups the number of classes. Provided by mmlcr.
prob a data.frame of (initial) posterior probabilities of class membership. Provided by mmlcr.
data a data.frame. Provided by mmlcr.
grouping the variable of data indicating subject id. Provided by mmlcr.

Details

The object argument is a list, originally provided to the mmlcr function as a component of a list in the components argument. At a minimum, that list consists of a formula of the form resp ~ covars, where resp is the response variable for that component and covars are any additional covariates (resp ~ 1 is commonly used), and a class. For longitudinal responses, the covariates should include the chronological variable, e.g., resp ~ poly(age, 2). Implemented classes include cnormlong, cnormonce, multinomlong, multinomonce, nblong, nbonce, normlong, normonce, poislong, and poisonce. The *long classes are for longitudinal responses, while the *once classes are for one-time responses. So that a rectangular data structure can be used for the data argument in mmlcr, the *once classes use just the first occurrence of a response for each subject.

The cnorm* classes are for censored normal distributions. They take a min and max argument for the minimum and maximum possible observed value. The multinom* classes are for multinomial, or categorical responses. The response variable should be a factor. The nb* classes are for negative binomial models (the usual overdispersed Poisson model). The norm* classes are for normal models, and the pois* classes are for Poisson models. For the cnorm* and norm* classes the minimum and maximum variances are constrained to be within a factor of 100 of each other. For the pois* classes an overdispersion estimate is calculated but not used.

Value

an mmlcr component object.

See Also

mmlcr, mmlcrObject

mmlcrdf

An artificial data frame for 'mmlcr' examples.

Description

The mmlcrdf data frame has 750 rows and 8 columns. It is a simulated data frame used for the mmlcr examples.

Usage

data(mmlcrdf)
mmlcrfit

Description

fits an mmlcr component object. Used internally by the function mmlcr

Usage

mmlcrfit(object, weights, classnumber)

Arguments

object an mmlcr component
weights a data.frame of weights, with row.names indicating subject ids.
classnumber a number indicating the latent class to use, i.e., which column of weights to use.
Value

the same \texttt{mmlcr} component object with an updated fit.

See Also

\texttt{mmlcr}, \texttt{mmlcrObject}, \texttt{mmlcrlike}

---

\texttt{mmlcrlike} \hspace{1cm} \textit{Used internally by \texttt{mmlcr} to calculate component likelihoods.}

Description

Find the likelihoods for each subject in an \texttt{mmlcr} component object. Meant to be internally by the function \texttt{mmlcr}

Usage

\texttt{mmlcrlike(object, weights, classnumber, final = FALSE)}

Arguments

\begin{itemize}
  \item \texttt{object} \hspace{1cm} a component of an \texttt{mmlcr} object
  \item \texttt{weights} \hspace{1cm} weights for the observations in object
  \item \texttt{classnumber} \hspace{1cm} the class number for the likelihoods
  \item \texttt{final} \hspace{1cm} see below
\end{itemize}

Value

if final = FALSE, a vector of likelihoods corresponding to the weights. The weights are not used except for their row.names.

if final = TRUE, the component is returned, with fitted and residual components added. The fitted and residual components will only be correct if \texttt{mmlcrlike} is called sequentially over the full set of classnumbers, as is the case when \texttt{mmlcrlike} is called by \texttt{mmlcr}.

Warning

the \texttt{mmlcrlike.multinomlong} method returns fitted and residual components that have no real meaning. It is included only for completeness.

See Also

\texttt{mmlcr}, \texttt{mmlcrObject}
### plot.mmlcr

**Plot Components of mmlcr Object**

#### Description
Generates a plot for each component of a mmlcr object. Longitudinal components are plotted against time for each class, while “once” components show means and standard deviations for each class.

#### Usage
```
plot.mmlcr(x, which = 1:length(x$components), ...)  
```

#### Arguments

- **x**: A mmlcr object.
- **which**: Which components to plot (by number).
- **...**: Arguments to be passed to component plots. In particular, a “smooth” parameter, equal to `supsmu`'s `bass` can be used in longitudinal components.

#### Details
Longitudinal components show plots of means for each class over the first covariate. Non-longitudinal components show the weighted mean plus and minus one weighted standard deviation for each class. There is no component plot for the multinom components.

#### See Also
- `mmlcr`, `plot.mmlcrlong`, `plot.mmlcronce`, `mmlcrObject`

### plot.mmlcrlong

**Plot a mmlcrlong Component**

#### Description
Generates a plot for mmlcrlong component of a mmlcr object. Class means are plotted against “time.” Generally called by `plot.mmlcr`.

#### Usage
```
plot.mmlcrlong(x, post.prob, smooth = 0, xlab = names(data)[2], ylab = names(data)[1], ylim = c(min(yy, na.rm = TRUE), max(yy, na.rm = TRUE)), cols = rep(1, dim(post.prob)[2]), pch = as.character(1:(dim(post.prob)[2])), ...)  
```


postprob

Arguments

- x: A mmlcr object.
- post.prob: Data.frame of class membership posterior probabilities.
- smooth: A smoothing parameter, equal to supsmu's bass.
- xlab: Label for the x-axis.
- ylab: Label for the y-axis.
- ylim: Limits for the y-axis.
- cols: The cols argument for par.
- pch: The pch argument for par.
- ...: Other arguments to be passed to plot.

Note

There is no component plot for the multinomlong components.

See Also

- plot.mmlcr, mmlcr, mmlcrObject

---

postprob  

Extract Posterior Probabilities from mmlcr Object

Description

Extract posterior class probabilities from an mmlcr object.

Usage

postprob(object)

Arguments

- object: A mmlcr object.

Value

A data.frame of posterior probabilities of class membership

See Also

- mmlcr, mmlcrObject
Examples

```r
## Not run:
data(mmlcrdf)

mmlcrdf.mmlcr2 <- mmlcr(outer = ~ sex + cov1 | id,
    components = list(
        list(formula = resp1 ~ 1, class = "cnormonce", min = 0, max = 50),
        list(formula = resp2 ~ poly(age, 2) + tcov1, class = "poislong"),
        list(formula = resp3 ~ poly(age, 2), class = "multinomlong"),
    ), data = mmlcrdf, n.groups = 2)

postprob(mmlcrdf.mmlcr2)
## End(Not run)
```

---

**theta.mmmod**

*Estimate theta of the Negative Binomial by Moments*

**Description**

Estimate theta of the negative binomial by moments, adapted to mmlcrfit from theta.mm of the MASS library

**Usage**

```r
theta.mmmod(y, u, dfr, wts, limit = 10, eps = .Machine$double.eps^0.25)
```

**Arguments**

- `y` Vector of observed values.
- `u` Estimated mean vector.
- `dfr` Residual degrees of freedom.
- `wts` Weights for the observations.
- `limit` Limit on iterations.
- `eps` Convergence tolerance.

**Value**

Estimate of theta.

**See Also**

theta.mm, mmlcrfit, theta.mm from the MASS library
Index

*Topic datasets
  mmlcrdf, 10
  Srcdfull, 2
*Topic distribution
  dcnorm, 3
  dnb2, 4
  theta.mmmod, 15
*Topic hplot
  plot.mmlcr, 13
  plot.mmlcrlong, 13
  plot.mmlcronce, 7
*Topic models
  BIC.mmlcr, 1
  fitted.mmlcr, 5
  mmlcr, 5
  mmlcrclassify, 9
  mmlcrcomponentinit, 9
  mmlcrfit, 11
  mmlcrlike, 12
  mmlcrObject, 8
  plot.mmlcr, 13
  plot.mmlcrlong, 13
  plot.mmlcronce, 7
  postprob, 14
  theta.mmmod, 15

AIC, 2
anova.mmlcr (mmlcr), 5
BIC.mmlcr, 1
dcnorm, 3
dnb2, 4

fitted.mmlcr, 5
formula.mmlcr (mmlcr), 5

logLik.mmlcr (mmlcr), 5

mmlcr, 4, 5, 8–14
mmlcrclassify, 9
mmlcrcomponentinit, 7, 9
mmlcrdf, 10
mmlcrfit, 11, 15
mmlcrlike, 4, 12, 12
mmlcrObject, 2, 7, 8, 8–10, 12–14
plot.mmlcr, 7, 8, 13, 14
plot.mmlcrlong, 13, 13
plot.mmlcronce, 7, 13
plot.multinomlong (plot.mmlcr), 13
plot.multinomonce (plot.mmlcr), 13
postprob, 14
print.mmlcr (mmlcr), 5
print.summary.mmlcr (mmlcr), 5
residuals.mmlcr (fitted.mmlcr), 5
Srcdfull, 2
summary.cnormlong (mmlcrcomponentinit), 9
summary.mmlcr, 7
summary.mmlcr (mmlcr), 5
summary.multinomlong (mmlcrcomponentinit), 9
summary.nblong (mmlcrcomponentinit), 9
summary.normlong (mmlcrcomponentinit), 9
summary.poislong (mmlcrcomponentinit), 9

theta.mm, 15
theta.mmmod, 15

vcov.mmlcr (mmlcr), 5

17