

# Package: libdr (via r-universe)

January 20, 2025

**Type** Package

**Title** Support Functions used by the LIBDR LCMM Analysis

**Version** 1.0.1

**Description** This package provides support functions frequently used by Constantine et al. for their paper, "Large-scale clustering of longitudinal faecal calprotectin and C-reactive protein profiles in inflammatory bowel disease". These functions are primarily used to manipulate data and visualise findings.

**License** GPL (>= 3)

**Depends** R (>= 4.1.0), splines, patchwork, magrittr, dplyr

**Imports** ComplexHeatmap, doParallel, forcats, foreach, ggplot2, knitr, lcmm, nnet, plyr, pracma, scales, tibble, tidyr, viridis

**Suggests** magick

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.3.2

**Config/pak/sysreqs** libicu-dev libpng-dev perl

**Repository** <https://nathansam.r-universe.dev>

**RemoteUrl** <https://github.com/VallejosGroup/Lothian-IBDR>

**RemoteRef** HEAD

**RemoteSha** 2cae64a89c4eac6217937ceb2951e2637c4fca96

**RemoteSubdir** libdr

## Contents

buildDT . . . . .	2
compareClustering . . . . .	2
describe_cat . . . . .	3
lcmmHeatmap . . . . .	3
makeMetrics . . . . .	4

mlrPlot . . . . .	4
plotCat . . . . .	5
rankCumulative . . . . .	5
reLabel . . . . .	6
spaghettiPlot . . . . .	6

<b>Index</b>	<b>9</b>
--------------	----------

---

buildDT	<i>Build DT::datatable objects from matrix of fit statistics</i>
---------	--

---

### Description

Build DT::datatable objects from matrix of fit statistics

### Usage

```
buildDT(hlme.metrics)
```

### Arguments

hlme.metrics	A data frame or matrix of model fit metrics for different number of assumed clusters (G)
--------------	--

---

compareClustering	<i>Create co-clustering matrix for FCAL and CRP models</i>
-------------------	--

---

### Description

Create co-clustering matrix for FCAL and CRP models

### Usage

```
compareClustering(G, models.fcal, models.crp, cutoff = FALSE, threshold = 0.8)
```

### Arguments

G	Number of clusters assumed by both models
models.fcal	List holding all hlme objects created from FCAL data
models.crp	List holding all hlme objects created from CRP data
cutoff	Logical, should only subjects assigned posterior cluster membership probabilities above a given threshold for both models be included?
threshold	Numeric. The posterior probability of cluster membership cutoff required for a subject to be included in the plot. Only used if cutoff = TRUE

---

describe_cat	<i>Produce descriptive analysis of categorical variables</i>
--------------	--

---

**Description**

Produce descriptive analysis of categorical variables

**Usage**

```
describe_cat(vars, data)
```

**Arguments**

vars	Character vector giving the names of categorical variables
data	Data frame (or tibble) holding the data of interest

---

lcmmHeatmap	<i>Produce a heatmap of LCMM cluster assignments across multiple lcmm::hlme fit objects</i>
-------------	---

---

**Description**

Produce a heatmap of LCMM cluster assignments across multiple lcmm::hlme fit objects

**Usage**

```
lcmmHeatmap(
  models,
  cores = NULL,
  png = TRUE,
  pdf = FALSE,
  file.name = "heatmap",
  return.data = FALSE
)
```

**Arguments**

models	list containing objects generated by <code>lcmm::hlme</code> .
cores	Numeric. Number of CPU cores to use. Defaults to maximum available.
png	Logical. If TRUE, the heatmap will be saved to file as a PNG.
pdf	Logical. If TRUE, the heatmap will be saved to file as a PDF.
file.name	Character. File name (without extension) for saved heatmap files. Defaults to "heatmap". Not used if png and pdf are both FALSE.
return.data	Logical. If TRUE, the averaged co-cluster matrix will be returned.

---

makeMetrics	<i>Produce table of model fit metrics (AIC, BIC, Log-likelihood) for LCMMs</i>
-------------	--

---

**Description**

Produce table of model fit metrics (AIC, BIC, Log-likelihood) for LCMMs

**Usage**

```
makeMetrics(Gs, models)
```

**Arguments**

Gs	Vector indicating which values of G should metrics be calculated for.
models	List containing LCMM model fits

---

mlrPlot	<i>Forest plot of multinomial logistic regression model</i>
---------	---

---

**Description**

Forest plot of multinomial logistic regression model

**Usage**

```
mlrPlot(
  dat,
  var,
  class = "class_combined",
  prob = "probmax",
  minprob = 0.5,
  extern = NULL,
  uni = FALSE
)
```

**Arguments**

dat	Data frame holding covariate and cluster assignment data.
var	Character. The name of the covariate(s) of interest. If given as a character vector then a multivariate model is fitted using the specified covariates.
class	Character. The name of the class assignment variable. Assumed to be "class_combined" if not manually specified.
prob	Character. The name of the variable which gives posterior probabilities for cluster membership. Assumed to be "probmax" if not manually specified.

minprob	Numeric. The minimum posterior probability for cluster membership required for a subject to be included in the analysis. Assumed to be 0.5 if not manually specified
extern	Optional data frame containing multinomial logistic regression results obtained from an external dataset. If not given then it is assumed that no external data is required.
uni	Logical. Should only a univariate model be considered? Defaults to FALSE (multivariate modelling).

---

plotCat *Percentage bar plots for categorical data*

---

### Description

Percentage bar plots for categorical data

### Usage

```
plotCat(dat, var, class = "class_combined")
```

### Arguments

dat	Data frame holding categorical and cluster assignment data.
var	Character. The name of the categorical variable of interest.
class	Character. The name of the class assignment variable. Assumed to be "class_combined" if not manually specified

### Value

A [patchwork](#) object

---

rankCumulative *Rank clusters by the area under their mean trajectories*

---

### Description

Rank clusters by the area under their mean trajectories

### Usage

```
rankCumulative(model, tmax = 7, var.time = "calpro_time")
```

**Arguments**

model	An <code>hlme</code> object
tmax	Numeric. Maximum observation period. Defaults to 10
var.time	Character. Name of variable used for time. Either "calpro_time" or "crp_time" Also used to determine if the LCMM has been fitted for FCAL or CRP.

**Value**

A data frame with columns for the original cluster number (Original), the new rank (New), and the area calculated via the trapezoidal rule (Area).

---

reLabel	<i>relabel alluvial class labels ensuring consistency.</i>
---------	--

---

**Description**

Relabel class labels in an alluvial data frame to ensure class labels are consistent with the class labels used models assuming fewer classes

**Usage**

```
reLabel(new.G, alluvial.df)
```

**Arguments**

new.G	The number of classes assumed by the LCMM which is now subject to class relabelling.
alluvial.df	Data frame containing class assignments for subjects for differing G

---

spaghettiPlot	<i>Spaghetti plots of each class</i>
---------------	--------------------------------------

---

**Description**

Spaghetti plots of each class

**Usage**

```
spaghettiPlot(
  data,
  models,
  G,
  log = TRUE,
  indi = FALSE,
  multi = TRUE,
  tmax = 10,
  ylim = NULL,
  column = FALSE,
  pprob.cutoff = NA,
  sizes = FALSE,
  mapping = NULL,
  save = FALSE,
  knots = FALSE,
  knot.type = "quantiles",
  n.knots = 3,
  grbf = FALSE,
  l = 1,
  color = NULL,
  clusters = FALSE,
  external = NULL,
  var.time = "calpro_time"
)
```

**Arguments**

<code>data</code>	Data frame used to fit the LCMM(s).
<code>models</code>	A list containing <code>hlme{lcmm}</code> S3 objects
<code>G</code>	Numeric. How many classes does the model assume?
<code>log</code>	Logical. Should plots be on log scale? Defaults to TRUE.
<code>indi</code>	Logical. Should separate plots for each class be generated? Defaults to FALSE.
<code>multi</code>	Logical. Should sub-plots for each class be plotted alongside each other? Defaults to TRUE.
<code>tmax</code>	Numeric. Maximum observation period. Defaults to 10
<code>ylim</code>	Character. What should the y-axis of the plots be scaled to fit? Either NULL (default), "data", "pred", or "conf". The default uses pre-determined values for FCAL and CRP
<code>column</code>	Logical. Should all sub-plots be in a single column? Defaults to FALSE (two columns)
<code>pprob.cutoff</code>	Posterior probability cut-off for subjects to be included as trajectories
<code>sizes</code>	Logical. Should class sizes be given as plot titles?
<code>mapping</code>	Numeric vector Giving the new ordering of sub-plots. One must take into account plots are generated by column - not row. If not specified then the default class numbers will be used.

save	Logical. Should sub-figure labels be generated?
knots	Logical. If TRUE, knot locations will be plotted as vertical lines. Defaults to FALSE
knot.type	Character. How are knots placed? Either "quantiles" (default) or "equal" (equally spaced). Not used if knots = FALSE.
n.knots	Numeric integer giving the number of knots used. Defaults to 3. Not used if knots = FALSE.
grbf	Logical. If TRUE, then the model is assumed to have been fitted using two Gaussian Radial basis functions. Defaults to FALSE (using natural cubic splines instead).
l	Numeric. Length scale constant for the Gaussian radial basis functions (not used if GRBF = FALSE)
color	Character. Name of variable in data to use to colour lines and points corresponding to subjects. If NULL then lines and points will be grey.
clusters	Logical. Are clusters given via a "class" column in data? Otherwise clusters are found from the hlme object in models[[G]]. Defaults to FALSE.
external	Optional data frame providing mean cluster estimates from <code>predictY{lcmm}</code> . Intended to be used with outputs from models applied to an external cohort. Columns should be "time", "Ypred", "CI1", and "CI2".
var.time	Character. Name of variable used for time. Either "calpro_time" or "crp_time" Also used to determine if the LCMM has been fitted for FCAL or CRP.

# Index

buildDT, 2

compareClustering, 2

describe\_cat, 3

hlme, 6, 7

lcmm: :hlme, 3

lcmmHeatmap, 3

makeMetrics, 4

mlrPlot, 4

patchwork, 5

plotCat, 5

predictY, 8

rankCumulative, 5

reLabel, 6

spaghettiPlot, 6