# Package ‘LCA’

October 12, 2022

**Version** 0.1.1

**Date** 2013-09-30

**Title** Localised Co-Dependency Analysis

**Maintainer** Ed Curry <e.curry@imperial.ac.uk>

**Depends** R (>= 2.15.0)

**Description** Performs model fitting and significance estimation for Localised Co-Dependency between pairs of features of a numeric dataset.

**License** GPL (>= 2)

**URL** https://www.r-project.org, https://www.imperial.ac.uk/people/e.curry

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2019-04-22 06:54:13 UTC

**Author** Ed Curry [aut, cre]

## R topics documented:

estimateB *ML Estimation of Laplace Beta*

## Description

Estimates initial value of parameter Beta from the PTL distribution used in LCA analysis.

## Usage

```
estimateB(x)
```

## Arguments

| | |
|---|---|
| `x` | Numeric vector of differences between the values of each feature, for a pair of objects in the dataset. |

## Details

Calculates maximum-likelihood estimate for Beta in the Laplace distribution fit to distribution of `x`.

## Value

Numeric value for initial estimate of PTL distribution parameter `Beta`

## Author(s)

Ed Curry <e.curry@imperial.ac.uk>

evaluateDiffSignificance *Evaluate Statistical Significance of an Observed Difference Between Two Objects*

## Description

Use PTL model to estimate the significance of a difference between the values of some feature of interest in two selected objects from a dataset.

## Usage

```
evaluateDiffSignificance(d,diff,PTLmodel)
```

### Arguments

| | |
|---|---|
| `d` | Numeric value specifying global dissimilarity between the selected objects |
| `diff` | Numeric value specifying magnitude of difference between the values of a selected feature of interest in the selected objects |
| `PTLmodel` | List, as returned by the function `fitPTLmodel`, with named elements `alpha`, `beta` and `gamma` specifying linear models for PTL parameter prediction. |

### Details

Evaluates statistical significance of observing as great a difference as that observed between the values of a selected feature of interest in the selected objects, given the global dissimilarity between those objects and the PTL models fitted to characterise these distributions across the whole dataset.

### Value

Numeric value giving p-value representing significance estimate of the observed difference, given the fitted models.

### Author(s)

Ed Curry `<e.curry@imperial.ac.uk>`

---

## `fitPTLmodel` *Calibrate Polynomial-Tail Laplace (PTL) model prdictions for LCA analysis*

---

### Description

Fits PTL models to randomly sampled pairs of the dataset, to enable prediction of PTL model parameter values based on hyperparameter `d`.

### Usage

```
fitPTLmodel(x,nPairs=10000)
```

### Arguments

| | |
|---|---|
| `x` | Numeric data input array, standardised to range (0,1) |
| `nPairs` | Numeric value specifying the number of samplings of pairs of objects to use to obtain hyperparameter fits |

### Details

Evaluates parameters for PTL model fits to the distributions of feature-wise differences between each of a specified (large) number of pairs of objects represented in dataset `x`. Obtains subsequent model fits explaining the individual PTL parameters `alpha`,`beta`,`gamma` in terms of the global (Euclidean) distances between the corresponding pairs of objects.

### Value

List with the following components:

- `alpha` — Object of class `lm`, which can be used to predict an appropriate value of `alpha` in the PTL distribution corresponding to a pair of objects in the dataset with a specified global dissimilarity
- `beta` — Object of class `lm`, which can be used to predict an appropriate value of `alpha` in the PTL distribution corresponding to a pair of objects in the dataset with a specified global dissimilarity
- `gamma` — Object of class `lm`, which can be used to predict an appropriate value of `alpha` in the PTL distribution corresponding to a pair of objects in the dataset with a specified global dissimilarity

### Author(s)

Ed Curry <e.curry@imperial.ac.uk>

---

## `getPTLExpectedCounts` *Predict Distribution of Feature-Wise Differences*

---

### Description

Predicts the expected number of features with a difference between two objects of a given global dissimilarity lying within a set of specified ranges.

### Usage

```
getPTLExpectedCounts(alpha,beta,gamma,bin_limits,ntrials)
```

### Arguments

- `alpha` — Numeric value specifying the parameter `alpha` in the PTL model used to estimate distribution of differences between the given objects
- `beta` — Numeric value specifying the parameter `beta` in the PTL model used to estimate distribution of differences between the given objects
- `gamma` — Numeric value specifying the parameter `gamma` in the PTL model used to estimate distribution of differences between the given objects
- `bin_limits` — Numeric vector specifying the limits of each range to be evaluated. Effectively, this gives the breakpoints between cells of the predicted histogram.
- `ntrials` — Numeric value specifying the number of features being evaluated in the dataset

### Details

Uses a PTL model with the specified parameters to estimate the expected number of features with differences between specified ranges. Used in calibration of PTL model parameter prediction to the dataset.

### Value

Numeric vector giving expected counts for numbers of features with a difference lying within the given set of specified ranges.

### Author(s)

Ed Curry <e.curry@imperial.ac.uk>

---

`getPTLparams` *Find best values of PTL parameters*

---

### Description

Finds parameters `alpha`, `beta` and `gamma` in PTL model to fit an observed distribution of differences in each feature's values between two given objects from a dataset.

### Usage

```
getPTLparams(x1,x2)
```

### Arguments

| | |
|---|---|
| `x1` | Numeric data input vector, standardised to range (0,1) |
| `x2` | Numeric data input vector, standardised to range (0,1) |

### Details

Uses iterative NLS fitting to determine parameters of PTL model to represent the distribution of the differences observed between two objects selected from the dataset being analysed with LCA.

### Value

List with the following elements:

| | |
|---|---|
| `d` | Numeric value specifying pair-wise global distance between objects `x1` and `x2` |
| `beta` | Numeric value specifying value of parameter `beta` in best PTL fit |
| `alpha` | Numeric value specifying value of parameter `alpha` in best PTL fit |
| `gamma` | Numeric value specifying value of parameter `gamma` in best PTL fit |

### Author(s)

Ed Curry <e.curry@imperial.ac.uk>

LCA *Localised Co-dependency Analysis*

## Description

Performs Localised Co-dependency Analysis

## Usage

```
LCA(x,PTLmodel,clique,seed.row,combine.method="Fisher",
adjust.method="BH",comparison.alpha=0.05)
```

## Arguments

| | |
|---|---|
| `x` | Numeric data input array, standardised to range (0,1) |
| `PTLmodel` | List with named elements `alpha`, `beta` and `gamma` specifying PTL parameters |
| `clique` | Numeric vector specifying which columns of data table represent entities defining the clique across which to evaluate co-dependency |
| `seed.row` | Numeric value specifying which row of data table to use as 'seed' feature with which to evaluate co-dependency |
| `combine.method` | Character specifying which method to use for combining individual LCD estimates. One of "Fisher" or "Inverse Product". |
| `adjust.method` | Character specifying which method to use for multiple testing adjustment of significance estimates. See `p.adjust` for further details. |
| `comparison.alpha` | Significance level threshold for including objects in the set to be used for evaluating LCD significance estimates for a given pair of features in a given clique. |

## Details

Function to evaluate LCD, within the members of `clique`, for all features in a dataset against the feature represented by `seed.row`.

## Value

List with elements:

| | |
|---|---|
| `LCD` | Data frame giving across-clique LCD significance estimates for each feature in the dataset, as both unadjusted p-value and adjusted for multiple testing. |
| `combinations` | An array detailing the individual pair-wise LCD tests performed amongst members of the clique, which were combined to give the overall significance estimates |

## Author(s)

Ed Curry <e.curry@imperial.ac.uk>

## Examples

```
suppressWarnings(RNGversion("3.5.0"))
## create a data matrix

x <- array(runif(1200),dim=c(40,12))

## implant similarity across a 'clique'

clique.cols <- sample(ncol(x),4)
x[,clique.cols] <- x[,clique.cols] + rnorm(nrow(x))

## scale x to (0,1)

x[x<0] <- 0
x[x>1] <- 1

## choose a 'seed' feature and some partner

seed.row <- sample(nrow(x),1)
partner.row <- sample(setdiff(c(1:nrow(x)),seed.row),1)

x[c(seed.row,partner.row),clique.cols] <- x[c(seed.row,partner.row),clique.cols] +
rep(rnorm(length(clique.cols)),each=2)

## calibrate PTL models to dataset

PTL.fit <- fitPTLmodel(x,nPairs=15)

## evaluate LCD between 'seed' feature and all other features

LCA.result <- LCA(x,PTLmodel=PTL.fit,clique=clique.cols,seed.row=seed.row)
## Not run: head(LCA.result$LCD)
```

---

`LCD` *Localised Co-Dependency Estimates*

---

## Description

Evaluates Statistical Significance of Localised Co-Dependency (LCD)

## Usage

```
LCD(x1,x2,seed.row,PTLmodel)
```

## Arguments

| | |
|---|---|
| `x1` | Numeric data vector giving values of all features for one selected object |
| `x2` | Numeric data vector giving values of all features for another selected object |

| | |
|---|---|
| `seed.row` | Numeric value specifying which row of data table to use as 'seed' feature with which to evaluate co-dependency |
| `PTLmodel` | List with named elements `alpha`, `beta` and `gamma` specifying PTL parameters |

## Details

Function to evaluate LCD, between two selected objects, for all features in a dataset against the feature represented by `seed.row`.

## Value

Numeric vector giving p-values for significance estimates of localised co-dependency, with the feature specified by `seed.row`, of all features in the dataset being analysed.

## Author(s)

Ed Curry <`e.curry@imperial.ac.uk`>

---

# `predictPTLparams` *Predict Values for PTL Model*

---

## Description

Generates parameter values for PTL model, based on a specified value of the hyperparameter

## Usage

```
predictPTLparams(d,PTLmodel)
```

## Arguments

| | |
|---|---|
| `d` | Numeric value specifying global dissimilarity between the selected objects |
| `PTLmodel` | List, as returned by the function `fitPTLmodel`, with named elements `alpha`, `beta` and `gamma` specifying linear models for PTL parameter prediction. |

## Details

Uses hyperparameter-based prediction linear models, calibrated to the dataset being analysed using the `fitPTLmodel` function, to estimate PTL model parameters for a pair of objects in the dataset with a global dissimilarity `d`.

## Value

List with elements:

- `alpha` Numeric value specifying the parameter `alpha` in the PTL model used to estimate distribution of differences between the given objects
- `beta` Numeric value specifying the parameter `beta` in the PTL model used to estimate distribution of differences between the given objects
- `gamma` Numeric value specifying the parameter `gamma` in the PTL model used to estimate distribution of differences between the given objects

## Author(s)

Ed Curry `<e.curry@imperial.ac.uk>`

---

`PTL` *Polynomial-Tail Laplace*

---

## Description

Probability density and distribution functions for Polynomial-Tail Laplace distribution

## Usage

```
dPTL(x,alpha,beta,gamma)
pPTL(q,alpha,beta,gamma)
```

## Arguments

- `x,q` Numeric vector of quantiles
- `alpha` Linear tail adjustment coefficient for PTL distribution
- `beta` Exponential decay term for PTL distribution, similar to `beta` parameter in Laplace distribution
- `gamma` Polynomial tail adjustment coefficient for PTL distribution

## Details

The PTL distribution has density

$$f(x) = \begin{cases} 0 & x < -2 \\ \dfrac{\alpha(\frac{x^2}{2} + 2x + 2) + \beta(e^{\frac{x}{\beta}} - e^{\frac{-2}{\beta}}) + \gamma(\frac{x^3}{3} + 4x + \frac{16}{3})}{4\alpha + 2\beta(1 - e^{\frac{-2}{\beta}}) + \frac{32\gamma}{3}} & -2 \leq x \leq 0 \\ \dfrac{\alpha(2x - \frac{x^2}{2} - 2) + \beta(e^{\frac{-2}{\beta}} - e^{\frac{x}{\beta}}) + \gamma(4x - \frac{x^3}{3} - \frac{16}{3})}{4\alpha + 2\beta(1 - e^{\frac{-2}{\beta}}) + \frac{32\gamma}{3}} & 0 < x \leq 2 \\ 1 & x > 2 \end{cases}$$

## Value

`dnorm` gives the density, `pnorm` gives the distribution function.

The length of the result is the maximum of the lengths of the numerical parameters for the other functions. The numerical parameters are recycled to the length of the result.

## Author(s)

Ed Curry <e.curry@imperial.ac.uk>

# Index