Package 'fdaPOIFD'

January 13, 2021

Type Package

Title Partially Observed Integrated Functional Depth

Version 1.0.0

Maintainer Antonio Elías <antonioefz91@gmail.com>

Description Integrated Depths for Partially Observed Functional Data (PoFD). Applications to visualization, outlier detection and classification. Software companion for Elías, Antonio, Jiménez, Raúl, Paganoni, Anna M. and Sangalli, Laura M., (2020), ``Integrated Depth for Partially Observed Functional Data".

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Depends R (>= 3.5.0)

Imports ggplot2, tibble, magrittr, reshape2, patchwork, MASS, fdapace, FastGP, stats

URL https://github.com/aefdz/fdaPOIFD

BugReports https://github.com/aefdz/fdaPOIFD

NeedsCompilation no

Author Antonio Elías [aut, cre],
Raúl Jiménez [ctb],
Anna M. Paganoni [ctb],
Laura M. Sangalli [ctb]

Repository CRAN

Date/Publication 2021-01-13 12:00:02 UTC

R topics documented:

`boxplot_PoFD` *Functional Boxplot for Partially Observed Functional Data*

Description

Plots the Functional Boxplot for PoFD and returns the magnitude and domain outliers. Magnitude outliers in blue, a dotted red indicates that the outlier situation occurs in a region with less than `fdom` proportion of the central region.

Usage

```
boxplot_PoFD(data, centralRegion = 0.5, fmag = 1.5, fdom = 0)
```

Arguments

`data`	matrix p by n, being n the number of functions and p the number of grid points.
`centralRegion`	number between 0 and 1 determining the proportion of the deepest functions that builds the central region.
`fmag`	factor to enhance the functional central region and determine the functional whiskers. Default is equal to 1.5. The whiskers provide the rule to unmask magnitude outliers.
`fdom`	factor that provides the maximum proportion of observed functions in the central region to consider a magnitude outlier as a domain outlier also. A value equals to 0 means that domain outliers are those functions that are observed on the domain where any of the functions building the central region are observed. A value equals to 1 determine as domain outlier any magnitude outlier out of the region where the central region is completely observed.

Value

a list with the functional boxplot for PoDF the magnitude outliers and the domain outliers.

References

Sun, Y. and Genton, M. G. (2011). Functional boxplots. Journal of Computational & Graphical Statistics, 20(2):316–334.

Examples

```
data(exampleData)
boxplot_PoFD(exampleData$PoFDextremes_outliers, centralRegion = 0.5, fmag = 1.5, fdom = 0)
```

exampleData *exampleData*

Description

An illustrative Functional Gaussian processes with different partially observed patterns with outliers and without outliers.

Usage

```
exampleData
```

Format

A list with three data sets (functions by columns):

PoFDintervals Partially observed functional data in intervals

PoFDextremes Partially Observed functional data with missing intervals at the extremes

PoFDextremes_outliers Same as above but including two magnitude and shape outliers

References

Elías, Antonio, Jiménez, Raúl, Paganoni, Anna M. and Sangalli, Laura M. (2020). Integrated Depths for Partially Observed Functional Data.

Examples

```
data(exampleData)
plot_PoFD(exampleData$PoFDintervals)
```

gaussian_PoFD *Gaussian Partially Observed Functional Data*

Description

Generates samples of partially observed gaussian functions following different censoring regimes.

Usage

```
gaussian_PoFD(n, p, type, observability, ninterval)
```

Arguments

n	total number of functional observations
p	total number of points observed for each function
type	type of partially observed data. Options are "sparse", "interval" and "common". See Elías et al (2020).
observability	mean observed proportion of the domain where each function is observed.
ninterval	if type = "interval", n_interval is an integer with the number of observed intervals 1, 2, 3... Large values of this parameter requires a large parameter p to guarantee the observability level.

Value

a list containing two elements 1) a functional sample and 2) the same sample of functions but partially observed following one of the schemes described in the argument type.

References

Elías, Antonio, Jiménez, Raúl, Paganoni, Anna M. and Sangalli, Laura M. (2020). Integrated Depths for Partially Observed Functional Data.

Examples

```
gaussian_pofd <- gaussian_PoFD(n=100, p=200, type="sparse", observability=0.5)
```

---

outliergram_PoFD *Outliergram for Partially Observed Functional Data*

Description

Plots the Outliergram for PoFD and returns the shape outliers.

Usage

```
outliergram_PoFD(data, fshape = 1.5, p1 = 1, p2 = 0)
```

Arguments

data	matrix p by n, being n the number of functions and p the number of grid points.
fshape	inflation of the outliergram that determine the shape outlier rule.
p1	parameter of the outliergram for resampling method. Default = 1.
p2	parameter of the outliergram for resampling method. Default = 0.

Value

a list with the functional outliergram for PoDF and the shape outliers.

References

Arribas-Gil, A. and Romo, J. (2014). Shape outlier detection and visualization for functional data: the outliergram.Biostatistics, 15(4):603–619.

Examples

```
data(exampleData)
outliergram_PoFD(exampleData$PoFDextremes_outliers, fshape = 1.5, p1 = 1, p2 = 0)
```

`plot_PoFD` *Plot Partially Observed Functional Data*

Description

Plot the sample of partially observed curves and the proportion of observed functions.

Usage

```
plot_PoFD(data)
```

Arguments

`data`	matrix p by n, being n the number of functions and p the number of grid points.

Value

Plot of the partially observed functional data and the proportion of observed functions at each time point.

Examples

```
data(exampleData)
plot_PoFD(exampleData$PoFDextremes)
```

POIFD *Integrated Depth for Partially Observed Functional Data*

Description

Compute the depth measures of a partially observed functional data set evaluated in a common grid.

Usage

```
POIFD(data, type = c("MBD", "FMD", "MHRD"), phi)
```

Arguments

data	matrix p by n, being n the number of functions and p the number of grid points. Rownames are the dense grid x and colnames the identifier of each functional data.
type	chosen depth measure. Fraiman and Muniz depth ("FMD"), Modified band depth ("MBD") or Modified Half Region Depth and Modified Epigraph/Hipograph Index "MHRD")
phi	phi function of weights for the POIFD. The default value is as in the paper, i.e. the proportion of observed functions at each time point.

Value

Ordered vector of depths from the deepest to outward. The names are the functions names (if provided) or the column position.

Examples

```
data <- gaussian_PoFD(n=100, p=200, type="sparse", observability=0.5)$pofd
poifd <- POIFD(data,  type = c("MBD"))
```

Index