Jackstrap

March 27, 2020

Type Package

Title Jackstrap Sousa & Stosic (2005) (Outliers in DEA)

Version 0.1.0

Author Kleber Morais de Sousa, Maria da Conceição Sampaio de Sousa, Paulo Aguiar do Monte

Maintainer Kleber Morais de Sousa <kleberfinancas@gmail.com>

Description This package applies the method developed by Sousa and Stosic (2005) Technical Efficiency of the Brazilian Municipalities: Correcting Nonparametric Frontier Mesurements For Outliers. Journal of Productivity Analysis, 24, 147-181.

Depends R (>= 2.15.1), fBasics, Benchmarking, dplyr, ggplot2, foreach, doParallel, doFuture, reshape, tidyr

License UFRPE, UFPB

Encoding UTF-8

LazyData TRUE

RoxygenNote 6.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

`hist_jack_ks` *Histogram with Jackstrap Efficiency Indicators: This function builds graphics with distributions of efficiency indicators without outliers and complete sample. The outliers are defined by K-S Test.*

Description

Histogram with Jackstrap Efficiency Indicators: This function builds graphics with distributions of efficiency indicators without outliers and complete sample. The outliers are defined by K-S Test.

Usage

```
hist_jack_ks(efficiency, model_hist_ks)
```

Arguments

`efficiency`	is the jackstrap object created by jackstrap function.
`model_hist_ks`	is the desired graphic model. There are four kinds: 1- Density Histogram of efficiency indicator with complete sample and without outliers by K-S test; 2 - Histogram of efficiency with complete sample and without outliers by K-S test; 3 - Histogram of efficiency without ouliers by K-S test; 4 - Histogram of efficiency with complete sample.

Value

Return the plot with efficiency indicators with complete sample and/or without outliers by combination leverage level and K-S test;

Examples

```
hist_jack_ks(efficiency_ks, 1)
hist_jack_ks(efficiency_ks, 2)
hist_jack_ks(efficiency_ks, 3)
hist_jack_ks(efficiency_ks, 4)
```

`hist_jack_step` *Histogram with Jackstrap Efficiency Indicators: This function builds a graphic with indicator distributions without outliers and complete sample. The outliers are defined by heaviside step function.*

Description

Histogram with Jackstrap Efficiency Indicators: This function builds a graphic with indicator distributions without outliers and complete sample. The outliers are defined by heaviside step function method.

Usage

```
hist_jack_step(efficiency, model_hist_step)
```

Arguments

efficiency	is the jackstrap object created by jackstrap function.
model_hist_step	is the desired graphic model. There are four kinds: 1- Density Histogram of efficiency indicators with complete sample and without outliers by heaviside step function; 2 - Histogram of efficiency with complete sample and without outliers by heaviside step function; 3 - Histogram of efficiency without ouliers by heaviside step function; 4 - Histogram of efficiency with complete sample.

Value

Return the plot with efficiency indicators with complete sample and/or without outliers by heaviside step function;

Examples

```
hist_jack_step(efficiency, 1)
hist_jack_step(efficiency, 2)
hist_jack_step(efficiency, 3)
hist_jack_step(efficiency, 4)
```

---

jackstrap *Jackstrap Method: The tool identifies outliers in Nonparametric Frontier. This function applies the developed tecnique by Sousa and Stosic (2005) Technical Efficiency of the Brazilian Municipalites: Correcting Nonparametric Frontier Meansurements for Outliers*

Description

Jackstrap Method: Tool identifies outliers in Nonparametric Frontier. This function applies the developed tecnique by Sousa and Stosic (2005) Technical Efficiency of the Brazilian Municipalites: Correcting Nonparametric Frontier Meansurements for Outliers

Usage

```
jackstrap(data, ycolumn, xcolumn, bootstrap = 1000,
  perc_sample_buble = 0.2, dea_method = "vrs",
  orientation_dea = "in", n_seed = NULL, repos = FALSE,
  num_cores = 1)
```

Arguments

data	is the dataset with input and output used to mensure efficiency; Dataset needs to have this form: 1th column: name of DMU (string); 2th column: code of DMU (integer); n columns of output variables; n columns of input variables.
ycolumn	is the quantity of y columns of dataset.
xcolumn	is the quantity of x columns of dataset.
bootstrap	is the quantity of applied resampling.
perc_sample_buble	is the percentage of sample in each bubble.

dea_method	is the DEA method: "crs" is DEA with constant returns to scale (CCR); "vrs" is DEA with varible returns to scale; and "fdh" is Free Disposal Hull (FDH) with variable returns to scale.
orientation_dea	is the direction of the DEA: "in" for focus on inputs; and "out" for focus on outputs.
n_seed	is the code as seed used to get new random samples.
repos	identify if the resampling method is with replacement TRUE or not FALSE.
num_cores	is the number of cores available to process.

Value

Return the jackstrap object with follow informations: "parameters" contain the parameters used on function; "bootstrap" is quantity of applied resampling; "mean_leverage" is leverage average for each DMU; "mean_geral_leverage" is general average of leverage and step function threshold; "sum_leverage" is accrued leverage on all resampling for each DMU; "count_dmu" is amount of each DMU selected by bootstrap. "efficiency_step_func" are efficiency indicators obtained by heaviside step function criteria; "result_kstest_method" are p-values of K-S test obtained by removing sequencially one by one the high leverage DMU; "efficiency_ks_method" are efficiency indicators obtained by K-S test criteria.

Examples

```
municipalities <- as.data.frame(jackstrap::municipalities)

efficiency <- jackstrap (data=municipalities, ycolumn=2, xcolumn=1, bootstrap=500,
                         perc_sample_bubble=0.10, dea_method="vrs",
                         orientation_dea="in", n_seed = 2000, repos=FALSE, num_cores=4)
```

---

jackstrap_ks

Jackstrap KS Method: The tool identifies outliers in Nonparametric Frontier. This function applies the developed tecnique by Sousa and Stosic (2005) Technical Efficiency of the Brazilian Municipalites: Correcting Nonparametric Frontier Meansurements for Outliers and to use the K-S test with criteria to define outliers

---

Description

Jackstrap KS Method: Tool identifies outliers in Nonparametric Frontier. This function applies the developed tecnique by Sousa and Stosic (2005) Technical Efficiency of the Brazilian Municipalites: Correcting Nonparametric Frontier Meansurements for Outliers and to use the K-S test with criteria to define outliers

Usage

```
jackstrap_ks(data, jackstrap_obj, num_cores, perc)
```

Arguments

data	is the dataset with input and output used to mensure efficiency; Dataset need to have this form: 1th column: name of DMU (string); 2th column: code of DMU (integer); n columns of output variables; n columns of input variables.
jackstrap_obj	is the object created by the function jackstrap.
num_cores	is number of cores used to process.
perc	is the percentage of sample used to K-S test.

Value

Return the jackstrap object increased with follow informations: "result_kstest_method" are p-values of K-S test obtained by removing sequencially one by one the high leverage DMU; "efficiency_ks_method" are efficiency indicators obtained by K-S test criteria.

Examples

```
efficiency_ks <- jackstrap_ks (data=municipalities, jackstrap_obj=efficiency, num_cores = 4, perc=0.80)
```

---

municipalities *Dataset of Brazilian Municipalities*

Description

Dataset of Brazilian Municipalities

Usage

```
municipalities
```

Format

A data frame with 489 rows (DMUs) and 3 variables (2 outputs and 1 inputs):

`Description of DMUs` string variable with descriptions of the each DMUs

`Code` integer variable identifies each DMU for integer code

`y1` float variable with quantity of first produced output with resources

`y2` float variable with quantity of second produced output with resources

`x1` float variable with first input (resource) used to produce outpus (y)

`plot_jackstrap_ks` *Plot Jackstrap KS: This function plots p-value of Kolmogorov-Smirnov Test in order decresing of leverage.*

Description

Plot Jackstrap KS: This function plots p-value of Kolmogorov-Smirnov Test in order decresing of leverage.

Usage

```
plot_jackstrap_ks(data_plot, model_plot)
```

Arguments

`data_plot`	is the jackstrap object created by jackstrap function.
`model_plot`	is the desired model. There are two models: 1 - The graphic shows the amount of removed DMU on x axis and p-value of K-S test on y axis; 2 - The graphic shows DMU code on x axis and p-value of K-S test on y axis.

Value

Return the plot with p-value of K-S test and removed DMU or DMU code.

Examples

```
plot_jackstrap_ks(efficiency_ks, 1)
```

`summary_jackstrap` *Summary Jackstrap: This function shows the main outcomes with outlier technique developed by Sousa and Stosic(2005).*

Description

Summary Jackstrap: This function shows the main outcomes with outlier technique developed by Sousa and Stosic(2005).

Usage

```
summary_jackstrap(object_jackstrap)
```

Arguments

`object_jackstrap`	is the jackstrap object created by jackstrap function.
`data`	is the dataset of research.

Value

Return the data frame with follow informations: "parameters" contain the parameters used on function; "outliers_by_step_func" are the outliers by heaviside step function criteria; "outliers_by_ks" are the outliers by K-S test; "dmu_efficiency_by_step_func" are DMUs evaluated as efficients by heaviside step function criteria; "dmu_inefficiency_by_step_func" are the DMUs evaluated as maximum inefficient by heaviside step function criteria; "dmu_efficiency_ks" are DMUs evaluated as efficients by K-S test criteria; "dmu_inefficiency_by_ks" are the DMUs evaluated as maximum inefficient by K-S test criteria.

Examples

```
summary_efic <- summary_jackstrap(efficiency_ks)
```

Index