Package ‘OVL.CI’

November 13, 2023

Type Package

Title Inference on the Overlap Coefficient: The Binormal Approach and Alternatives

Version 0.1.0

Maintainer Alba M. Franco-Pereira <albfranc@ucm.es>

Description Provides functions to construct confidence intervals for the Overlap Coefficient (OVL). OVL measures the similarity between two distributions through the overlapping area of their distribution functions. Given its intuitive description and ease of visual representation by the straightforward depiction of the amount of overlap between the two corresponding histograms based on samples of measurements from each one of the two distributions, the development of accurate methods for confidence interval construction can be useful for applied researchers. Implements methods based on the work of Franco-Pereira, A.M., Nakas, C.T., Reiser, B., and Pardo, M.C. (2021) <doi:10.1177/09622802211046386>.

License GPL-2

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 7.2.3

Imports ks

Depends R (>= 2.10)

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Alba M. Franco-Pereira [aut, cre, cph],
Christos T. Nakas [aut],
Benjamin Reiser [aut],
M.Carmen Pardo [aut]

Repository CRAN

Date/Publication 2023-11-13 17:43:18 UTC

R topics documented:

`kernel.e` *Epanechnikov kernel*

Description

Evaluates the Epanechnikov kernel

Usage

```
kernel.e(u)
```

Arguments

u	vector of observations

Value

evaluation of the Epanechnikov kernel

Examples

```
x = rnorm(100,1,2)
kernel.e(x)
```

kernel.e.density *Epanechnikov kernel density estimation*

Description

Estimates the density function using the Epanechnikov kernel

Usage

```
kernel.e.density(data, points, h)
```

Arguments

data	vector of observations
points	in which the function is evaluated
h	bandwidth

Value

density estimation

Examples

```
x = rnorm(100,1,2)
gridd = seq(-5,5,length.out=1000)
h = (4/3)^(1/5)*sd(x)*length(x)^(-1/5)
kernel.e.density (x,gridd,h)
```

kernel.g *Gaussian kernel*

Description

Evaluates the Gaussian kernel

Usage

```
kernel.g(u)
```

Arguments

u	vector of observations

Value

evaluation of the Gaussian kernel

Examples

```
x = rnorm(100,1,2)
kernel.g(x)
```

`kernel.g.density` *Gaussian kernel density estimation*

Description

Estimates the density function using the Gaussian kernel

Usage

```
kernel.g.density(data, points, h)
```

Arguments

`data`	vector of observations
`points`	in which the function is evaluated
`h`	bandwidth

Value

density estimation

Examples

```
x = rnorm(100,1,2)
gridd = seq(-5,5,length.out=1000)
h = (4/3)^(1/5)*sd(x)*length(x)^(-1/5)
kernel.g.density (x,gridd,h)
```

`likbox` *Likelihood function of the BoxCox transformation*

Description

Computation of the likelihood function of the BoxCox transformation

Usage

```
likbox(h, data, n)
```

Arguments

h	parameter of the Box-Cox transformation
data	joint vector of controls (first) and cases
n	length of the vector of controls

Value

the likelihood function of the BoxCox transformation

Examples

```
h=-1.6
controls=rnorm(50,6,1)
cases=rnorm(100,6.5,0.5)
likbox(h,c(controls,cases),n=length(controls))
```

OVL.BCAN *OVL.BCAN*

Description

Parametric approach using a bootstrap-based approach to estimate the variance

Usage

```
OVL.BCAN(x, y, alpha = 0.05, B = 100, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
B	bootstrap size
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.BCAN (controls,cases)
```

OVL.BCbias *OVL.BCbias*

Description

Parametric approach using a bootstrap bias-corrected approach

Usage

```
OVL.BCbias(x, y, alpha = 0.05, B = 100, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
B	bootstrap size
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.BCAN (controls,cases)
```

OVL.BCPB *OVL.BCPB*

Description

Parametric approach using a bootstrap percentil approach to estimate the variance

Usage

```
OVL.BCPB(x, y, alpha = 0.05, B = 100, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
B	bootstrap size
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.BCPB (controls,cases)
```

OVL.D *OVL.D*

Description

Parametric approach using the delta method

Usage

```
OVL.D(x, y, alpha = 0.05)
```

Arguments

x	controls
y	cases
alpha	confidence level

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.D (controls,cases)
```

OVL.DBC *OVL.DBC*

Description

Parametric approach using the delta method after the Box-Cox transformation

Usage

```
OVL.DBC(x, y, alpha = 0.05, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.DBC (controls,cases)
```

OVL.DBCL *OVL.DBCL*

Description

Parametric approach using the delta method after the Box-Cox transformation taking into account the variability of the estimated transformation parameter

Usage

```
OVL.DBCL(x, y, alpha = 0.05, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.DBCL (controls,cases)
```

OVL.K *OVL.K*

Description

Kernel approach estimating the variance via bootstrap

Usage

```
OVL.K(x, y, alpha = 0.05, B = 100, k = 1, h = 1)
```

Arguments

x	controls
y	cases
alpha	confidence level
B	bootstrap size
k	kernel. When k=1 (default value) the kernel used in the estimation is the Gaussian kernel. Otherwise, the Epanechnikov kernel is used instead.
h	bandwidth. When h=1 (default value) the cross-validation bandwidth is chosen. Otherwise, the bandwidth considered by Schmid and Schmidt (2006) is used instead.

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.K (controls,cases)
```

OVL.KPB *OVL.KPB*

Description

Kernel approach using a bootstrap percentile approach

Usage

```
OVL.KPB(x, y, alpha = 0.05, B = 100, k = 1, h = 1)
```

Arguments

x	controls
y	cases
alpha	confidence level
B	bootstrap size
k	kernel. When k=1 (default value) the kernel used in the estimation is the Gaussian kernel. Otherwise, the Epanechnikov kernel is used instead.
h	bandwidth. When h=1 (default value) the cross-validation bandwidth is chosen. Otherwise, the bandwidth considered by Schmid and Schmidt (2006) is used instead.

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.KPB (controls,cases)
```

OVL.LogitBCAN *OVL.LogitBCAN*

Description

BCAN procedure carried out in the logit scale and back-transformed

Usage

```
OVL.LogitBCAN(x, y, alpha = 0.05, B = 100, h_ini = -0.6)
```

Arguments

`x`	controls
`y`	cases
`alpha`	confidence level
`B`	bootstrap size
`h_ini`	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.LogitBCAN (controls,cases)
```

`OVL.LogitD` *OVL.LogitD*

Description

Parametric approach using the delta method after switching to a logit scale and then transforming back

Usage

```
OVL.LogitD(x, y, alpha = 0.05)
```

Arguments

`x`	controls
`y`	cases
`alpha`	confidence level

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.LogitD (controls,cases)
```

OVL.LogitDBC *OVL.LogitDBC*

Description

Parametric approach using the delta method after the Box-Cox transformation after switching to a logit scale and then transforming back

Usage

```
OVL.LogitDBC(x, y, alpha = 0.05, h_ini = -0.6)
```

Arguments

x	controls
y	cases
alpha	confidence level
h_ini	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.LogitDBC (controls,cases)
```

OVL.LogitDBCL *OVL.LogitDBCL*

Description

Parametric approach using the delta method after the Box-Cox transformation in the logit scale and back-transformed considering the variability of the estimated transformation parameter

Usage

```
OVL.LogitDBCL(x, y, alpha = 0.05, h_ini = -0.6)
```

Arguments

`x`	controls
`y`	cases
`alpha`	confidence level
`h_ini`	initial value in the optimization problem

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.LogitDBCL (controls,cases)
```

`OVL.LogitK` *OVL.LogitK*

Description

Kernel approach estimating the variance via bootstrap in the logit scale and back-transformed

Usage

```
OVL.LogitK(x, y, alpha = 0.05, B = 100, k = 1, h = 1)
```

Arguments

`x`	controls
`y`	cases
`alpha`	confidence level
`B`	bootstrap size
`k`	kernel. When k=1 (default value) the kernel used in the estimation is the Gaussian kernel. Otherwise, the Epanechnikov kernel is used instead.
`h`	bandwidth. When h=1 (default value) the cross-validation bandwidth is chosen. Otherwise, the bandwidth considered by Schmid and Schmidt (2006) is used instead.

Value

confidence interval

Examples

```
controls = rnorm(50,6,1)
cases = rnorm(100,6.5,0.5)
OVL.LogitK (controls,cases)
```

ssdd *Sample variance computation*

Description

Computes the sample variance of a vector of observations

Usage

```
ssdd(x)
```

Arguments

x	vector of observations

Value

the sample variance

Examples

```
x = rnorm(100,1,2)
ssdd(x)
```

test_data *Simulated data with normal distributions to showcase the CI'S Overlap Coefficient (OVL) calculation*

Description

Contains controls and cases data from normal distributions

Usage

```
data(test_data)
```

Format

A data frame with 100 rows and 2 variables:

controls Simulated data from a N(10,1)distribution for the control group

cases Simulated data from a N(10.5,0.5)distribution for the case group

References

This data set was artificially created for the OVL.CI package.

Examples

```
data(test_data)
```

U *Auxiliary function*

Description

Evaluates an auxiliary function

Usage

```
U(mu1, mu2, sigma1, sigma2)
```

Arguments

mu1	sample mean of a vector x
mu2	sample mean of a vector y
sigma1	sample standard deviation of a vector x
sigma2	sample standard deviation of a vector y

Value

evaluation of an auxiliary function

Examples

```
U(1,2,1,1)
```

Index