Package 'sac'

October 14, 2022

Version 1.0.2

Date 2021-01-26

Title Semiparametric Analysis of Change-Point

Author Zhong Guan <zguan@iusb.edu>

Maintainer Zhong Guan <zguan@iusb.edu>

Depends R (>= 1.4)

Description Semiparametric empirical likelihood ratio
based tests of change-point with one-change or epidemic alternatives
with data-based model diagnostic are contained.

License GPL (>= 2)

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2021-01-27 15:00:02 UTC

R topics documented:

BootsChapt 2
BootsModelTest 4
CriticalValues 5
cumsum.test 6
plots 7
Pvalues 9
schapt 11
SemiparChangePoint 13

Index **15**

BootsChapt *Bootstrap (Permutation) Test of Change-Point(s) with One-Change or Epidemic Alternative*

Description

By resampling with(out) replacement from the original sample data, we can obtain bootstrap(permutation) versions of the test statistics. The p-values of the test(s) from the original data are approximated by the p-values of the bootstrap(permutation) version statistics.

Usage

```
BootsChapt(x, stat1, stat2 = NULL, B, replace = FALSE,
    alternative = c("one.change", "epidemic"), adj.Wn = FALSE,
    tol = 1.0e-7, maxit = 50,trace = FALSE,... )
```

Arguments

x	a numeric vector or matrix containing the data, one row per observation;
stat1	test statistic Sn for "one-change" alternative or Vn for "epidemic" alternative, output of SemiparChangePoint.
stat2	test statistic Wn for "epidemic" alternative, output of SemiparChangePoint.
B	number of resamples
replace	a logical indicating whether bootstrap samples for bootstrap test of the change-point are selected with or without replacement, if replace= FALSE (default), corresponds to permutation test, otherwise, bootstrap test;
alternative	a character string specifying the alternative hypothesis, must be one of "one-change" (default) or "epidemic". You can specify just the initial letter.
adj.Wn	logical indicating if Wn should be adjusted or not for "epidemic" alternative.
tol	the desired accuracy (convergence tolerance), an argument of glm.control.
maxit	the maximum number of iterations, an argument of glm.control.
trace	logical indicating if output should be produced for each iteration, an argument of glm.control.
...	other arguments

Details

The procedure will fail when there is separation in the data in the sense of Albert \& Anderson(1984, *Biometrika*) and Santner \& Duffy (1986, *Biometrika*). In this case, the change-point(s) may be detected easily using nonparametric method based on cumsum. Now, this program does not check whether the data is separated.

Value

p.boots	bootstrap p-value of Sn for "one-change" alternative
p.boots.Vn	bootstrap p-value of Vn for "epidemic" alternative
p.boots.Wn	bootstrap p-value of Wn for "epidemic" alternative

Note

Default alternative is "one-change", even when stat2 is not NULL. If alternative = "epidemic", both stat1 and stat2 should be provided. Statistic Wn need be adjusted only for one dimensional observations and if no bootstrap test is conducted. However, if Wn is already adjusted, you have to asign adj.Wn = TRUE to calculate the p-value of Wn.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z.(2001) Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*.

Guan, Z.(2004) A semiparametric changepoint model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also

SemiparChangePoint, schapt, p.OneChange, p.Epidemic.Vn, p.Epidemic.Wn

Examples

```
require(sac) #load the package

# one-change alternative
k<-10
n<-20
x<-rnorm(n,0,1)
x[(k+1):n]<-x[(k+1):n]+1.5
T<-SemiparChangePoint(x, alternative = "one.change")$Sn
BootsChapt(x, T, B = 5)
    #Choose larger B to get better approximate p-value.
```

BootsModelTest *Bootstrap Test of the Validity of the Semiparametric Change-Point Model*

Description

Using bootstrap method to approximate the p-value of test of the model validity. Bootstrap samples are drawn from the semiparametrica empirical distribution which are estimates of the underlying population distributions.

Usage

```
BootsModelTest(x, k, m, B, Alpha, Beta, tol = 1.0e-7, maxit=50, trace=FALSE)
```

Arguments

x	a numeric vector or matrix containing the data, one row per observation;
k	the estimated change-point, output of SemiparChangePoint
m	= n the sample size for "one-change" alternative, or the estimated second change-point for "epidemic" alternative, an output of SemiparChangePoint
B	number of resamples
Alpha	estimated parameter α, output of SemiparChangePoint
Beta	estimated parameter β, output of SemiparChangePoint
tol	the desired accuracy (convergence tolerance), an argument of glm.control.
maxit	the maximum number of iterations, an argument of glm.control.
trace	logical indicating if output should be produced for each iteration, an argument of glm.control.

Value

Delta	The test statistic of the model validity
Pvalue	The bootstrapped p-value

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z.(2001) Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*.

Guan, Z.(2004) A semiparametric changepoint model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also

SemiparChangePoint, schapt

Examples

```
## Nile data with one change-point: the annual flows drop in 1898.
## It is believed to be caused by the building of the first Aswan dam.
if(! "package:stats" %in% search()) library(stats)
data(Nile)
require(sac) #load the package
Nile.res<-SemiparChangePoint(Nile, alternative = "one.change")
BootsModelTest(Nile, Nile.res$k.hat, length(Nile), B=5, Nile.res$alpha.hat,
    Nile.res$beta.hat)
    # Choose larger B to get better approximate p-value.
    # It takes longer to do bootstrap model test for large B.
```

CriticalValues *Critical Values of Tests of Change-Point(s) with One-Change or Epidemic Alternative*

Description

Return the approximate critical values of the test statistics given level alfa

Usage

```
Sn.alfa(alfa,n,d,model=c("parametric","semiparametric"),
        tol = .Machine$double.eps^0.25, maxiter = 1000)
CV.Epidemic.Vn(alfa, d, tol = 1e-10)
CV.Epidemic.Wn(alfa, tol = 1e-07)
```

Arguments

alfa	significance level
n	sample size
model	a character string specifying the model, must be one of "parametric" or "semiparametric" (default). You can specify just the initial letter
d	dimension of the data value
tol	the desired accuracy (convergence).
maxiter	the maximum number of iterations for uniroot.

Details

Function Sn.alfa returns the critical value of Sn for one-change alternative. The functions CV.Epidemic.Vn and CV.Epidemic.Wn calculate critical values for Vn and Wn.

Value

Critical values

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Csorgo, M. and Horvath, L. (1997), *Limit Theorems in Change-Point Analysis*, New York: John Wiley

See Also

schapt

Examples

```
require(sac) #load the package
alpha<-0.05
n<-20
d<-1
Sn.alfa(alpha, n, d, model="semiparametric")
CV.Epidemic.Vn(alpha, d)
CV.Epidemic.Wn(alpha)
```

`cumsum.test` *Nonparametric Test for Change-Point with One-change or Epidemic Alternative*

Description

Compute test statistic based on CUMSUM and change-point estimate

Usage

```
cumsum.test(x, alternative = c("one.change", "epidemic"))
```

Arguments

`x`	a numeric vector or matrix containing the data, one row per observation;
`alternative`	a character string specifying the alternative hypothesis, must be one of `"one-change"` (default) or `"epidemic"`. You can specify just the initial letter.

Value

`Sn`	test statistic
`k.hat`	estimated change-point
`m.hat`	the second estimated change-point for epidemic alternative

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Csorgo, M. and Horvath, L. (1997), *Limit Theorems in Change-Point Analysis*, New York: John Wiley

See Also

cumsum

Examples

```
require(sac) #load the package
# one-change alternative
k<-10
n<-30
x<-rnorm(n,0,1)
x[(k+1):n]<-x[(k+1):n]+1.5
cumsum.test(x, alternative = "one.change")
# epidemic alternative
k<-10
m<-20
n<-30
x<-rnorm(n,0,1)
x[(k+1):m]<-x[(k+1):m]+1.5
cumsum.test(x,  alternative = "epidemic")
```

---

plots *Visualized Model Diagnostic and Loglikelihood Plot*

Description

Plot and compare the empirical likelihood and semiparametric empirical likelihood distribution functions, plot loglikelihood function.

Usage

```
Graf.Diagnostic(x, k, m, Alpha, Beta, Color, LTY, xlab = "x",
    ylab = "Estimated DF's", main = "Model Diagnostic",
    OneLegend = TRUE, lgnd1, lgnd2, arw1, arw2, ...)
Plot.ll(x, ll, col, xaxis.lab = NULL, xlab = "k", ylab = "Loglikelihood",
    main = "Plot of Loglikelihood",...)
```

Arguments

x	a numeric vector or matrix containing the data, one row per observation;
ll	loglikelihood function, output of SemiparChangePoint
col	color code or character string for the loglikelihood curve
xaxis.lab	a vector of character strings or numeric values to be placed at the tickpoints as labels of axis
k	the estimated change-point, output of SemiparChangePoint
m	= n, the sample size, for "one-change" alternative, or the estimated second change-point for "epidemic" alternative, an output of SemiparChangePoint
Alpha	estimated parameter α, output of SemiparChangePoint
Beta	estimated parameter β, output of SemiparChangePoint
Color	a vector of character strings or color codes for curves of estimated distribution functions $\hat{F}$, $\tilde{F}$, $\hat{G}$ and $\tilde{G}$
LTY	vector of lty's, LTY=c(lty1, lty2, lty3, lty4), corresponds to the above color codes
xlab	character string for x-axis lable
ylab	character string for y-axis lable
main	character string for main title
OneLegend	a logical indicating whether plot one or two legend.
lgnd1	a numeric vector of two specify the position of the first legend box
lgnd2	a numeric vector of two specify the position of the second legend box, if OneLegend = FALSE
arw1	a numeric vector of four numbers indicating start and end positions of the first arrows point to curves
arw2	a numeric vector of four numbers indicating start and end positions of the second arrows point to curves
...	other arguments of function plot

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z.(2001) Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*;

Guan, Z.(2004) A semiparametric change-point model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Changepoints with Epidemic Alternatives.

See Also

schapt

Examples

```
require(sac) #load the package
k<-30
n<-80
x<-rnorm(n,0,1)
x[(k+1):n]<-x[(k+1):n]+1.5
res<-SemiparChangePoint(x, alternative = "one.change")
Plot.ll(x, res$ll, col="blue")

## Nile data with one change-point: the annual flows drop in 1898 which corresponds
## to k=28. It is believed to be caused by the building of the first Aswan dam.
if(! "package:sac" %in% search()) library(sac)
    #if package sac has not been loaded, load it.
if(! "package:stats" %in% search()) library(stats)
data(Nile)
plot(Nile, type="p")
Nile.res<-SemiparChangePoint(Nile, alternative = "one.change")
Color<-c(1,2,3,4); LTY<-c(1,2,3,4)

## Plots of estimated distribution functions
Graf.Diagnostic(Nile, Nile.res$k.hat, length(Nile), Nile.res$alpha.hat,
    Nile.res$beta.hat, Color, LTY, xlab = "x", ylab = "Estimated DF's",
    main="Model Diagnostic for Nile Data", OneLegend = FALSE, lgnd1 =
    c(1100, 0.15), lgnd2 = c(600, .99), arw1=c(780, .93, 1010, .9),
    arw2 = c(1165, .15, 1015, .24))

## Plot of loglikelihood function
Plot.ll(Nile, Nile.res$ll, col = "blue")
Plot.ll(Nile, Nile.res$ll, col = "blue", xaxis.lab = seq(1871,1970, length = 100),
    xlab = "Year")
```

`Pvalues` *The p-values of Test Statistics Based on Asymptotic Distribution*

Description

Calculate the approximate p-values of the test statistics `Tn`, `Vn` and `Wn` using limit null distributions.

Usage

```
p.OneChange(n, d, Sn)
p.Epidemic.Vn(Vn, d, tol = 1e-10)
p.Epidemic.Wn(Wn, tol = 1e-07)
```

Arguments

`Sn`	test statistic `Sn` of the one-change alternative
`Vn`	test statistic `Vn` of the epidemic alternative

Wn	test statistic Wn of the epidemic alternative
n	sample size
d	dimension of the data value
tol	the desired accuracy.

Value

p.value	p-value

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z.(2001) Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*.

Guan, Z.(2004) A semiparametric changepoint model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also

schapt, BootsChapt

Examples

```
require(sac) #load the package
# one-change alternative
k<-10
n<-30
x<-rnorm(n,0,1)
x[(k+1):n]<-x[(k+1):n]+1.5
T<-SemiparChangePoint(x, alternative = "one.change")$Sn
p.OneChange(n, d=1, T)

# epidemic alternative
k<-5
m<-10
n<-20
x<-rnorm(n,0,1)
x[(k+1):m]<-x[(k+1):m]+1.5
res<-SemiparChangePoint(x, alternative = "e")
V<-res$Vn; W<-res$Wn
p.Epidemic.Vn(V, d=1)
p.Epidemic.Wn(W)
```

schapt *Semiparametric Analysis of Changepoint*

Description

Semiparametric empirical likelihood ratio based test of changepoint with one-change or epidemic alternatives with data-based model diagnostic

Usage

```
schapt(x, n.boots = 0, replace = FALSE, alternative = c("one.change",
    "epidemic"), conf.level = 0.95, adj.Wn = FALSE, model.test = FALSE,
    n.model.boots = 0, tol=1.0e-7, maxit=50,trace=FALSE,... )
```

Arguments

x	a numeric vector or matrix containing the data, one row per observation;
n.boots	number of bootstrap samples for bootstrap test of the change-point, if n.boots =0 , do not perform bootstrap test;
replace	a logical indicating whether bootstrap samples for bootstrap test of the change-point are selected with or without replacement, if replace= FALSE (default), corresponds to permutation test, otherwise, bootstrap test;
alternative	a character string specifying the alternative hypothesis, must be one of "one-change" (default) or "epidemic". You can specify just the initial letter. Epidemic alternative is also called square wave alternative in the literature.
conf.level	confidence level.
adj.Wn	logical indicating if Wn should be adjusted or not for "epidemic" alternative.
model.test	a logical indicating whether the test of model validity is performed.
n.model.boots	number of bootstrap samples for model test, if either n.model.boots = 0 or model.test=FALSE, then model test will not be performed.
tol	the desired accuracy (convergence tolerance), an argument of glm.control.
maxit	the maximum number of iterations, an argument of glm.control.
trace	logical indicating if output should be produced for each iteration, an argument of glm.control.
...	other future arguments

Details

Model: $\log\{g(x)/f(x)\} = \exp\{\alpha + \beta' T(x)\}$, where $f(x)$ and $g(x)$ are the density (frequency) functions of the two hypothesized populations, and $T(x)$ can be chosen as $T(x) = x$ or $T(x) = (x, x^2)$. The procedure will fail when there is separation in the data in the sense of Albert \& Anderson(1984, *Biometrika*) and Santner \& Duffy (1986, *Biometrika*). In this case, the changepoint(s) may be detected easily using nonparametric method based on cumsum. Currently, this function does not check whether the data are separated.

Value

data.name	dataset name
parameter	sample size n and degree(s) of freedom of the df of Sn for "one-change" alternative
alternative	the alternative hypothesis
statistic	a list contains Sn for "one-change" alternative, Sn, Vn and Wn for "epidemic" alternative; also contains Delta if model test is performed
estimate	a list contains change-point(s) and alpha and beta
p.value	a list contains p-value(s), p(Sn), of Sn for "one-change" alternative, p(Vn) and p(Wn), of Vn and Wn, repectively, for "epidemic" alternative; also p.boots(model) of Delta if model test is performed, if bootstrap test(s) of the change-point(s) are performed, the it also containts the corresponding p-values, p.boots(Sn), p.boots(Vn) and p.boots(Wn) accordingly.

Note

Statistic Wn need be adjusted only for one dimensional observations and if no bootstrap test is conducted. If returned p-value is 0, this means that the p-value is less than 1.0e-7.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z. (2001). Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*.

Guan, Z.(2004) A semiparametric change-point model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also

Graf.Diagnostic, Plot.ll

Examples

```
require(sac) #load the package
# one-change alternative
## Nile data with one change-point: the annual flows drop in 1898.
## It is believed to be caused by the building of the first Aswan dam.
if(! "package:sac" %in% search()) library(sac)
    #if package sac has not been loaded, load it.
if(! "package:stats" %in% search()) library(stats)
data(Nile)
plot(Nile, type="p")
schapt(Nile, alternative = "one.change")
```

SemiparChangePoint *Semiparametric Test of Change-point(s) with One-change or Epidemic Alternative*

Description

Calculate test statistics, loglikelihood function and estimate unknown parameters in the semiparametric model.

Usage

```
SemiparChangePoint(x, alternative = c("one.change", "epidemic"),
    adj.Wn = FALSE, tol = 1e-07, maxit = 50, trace = FALSE, ...)
```

Arguments

x	a numeric vector or matrix containing the data, one row per observation;
alternative	a character string specifying the alternative hypothesis, must be one of "one-change" (default) or "epidemic". You can specify just the initial letter.
tol	the desired accuracy (convergence tolerance), an argument of glm.control.
adj.Wn	logical indicating if Wn should be adjusted or not for "epidemic" alternative.
maxit	the maximum number of iterations, an argument of glm.control.
trace	logical indicating if output should be produced for each iteration, an argument of glm.control.
...	other future arguments

Details

Model: $\log\{g(x)/f(x)\} = \exp\{\alpha + \beta' T(x)\}$, where $f(x)$ and $g(x)$ are the density (frequency) functions of the two hypothesized populations, and $T(x)$ can be chosen as $T(x) = x$ or $T(x) = (x, x^2)$. The procedure will fail when there is separation in the data in the sense of Albert \& Anderson(1984, *Biometrika*) and Santner \& Duffy (1986, *Biometrika*). In this case, the change-point(s) may be detected easily using nonparametric method based on cumsum. Currently, this function does not check whether the data are separated.

Value

k.hat	change-point estimate
m.hat	second change-point estimate for "epidemic" alternative
ll	loglikelihood function
Sn	likelihood ratio test statistic for "one-change" alternative
Vn	test statistic based integal of weighted likelihood ratio for "epidemic" alternative

Wn	test statistic based supremum of weighted likelihood ratio for "epidemic" alternative
alpha.hat	estimate of α
beta.hat	estimate of β

Note

Statistic Wn need be adjusted only for one dimensional observations and if no bootstrap test is conducted.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Guan, Z.(2001) Some Results About Empirical Likelihood Method, *Ph.D. Thesis, The University of Toledo*.

Guan, Z.(2004) A semiparametric change-point model, *Biometrika*, 91, 4, 849–862.

Guan, Z. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also

schapt, p.OneChange, p.Epidemic.Vn, p.Epidemic.Wn

Examples

```
require(sac) #load the package
# one-change alternative
k<-10
n<-30
x<-rnorm(n,0,1)
x[(k+1):n]<-x[(k+1):n]+1.5
SemiparChangePoint(x, alternative = "one.change")

# epidemic alternative
k<-5
m<-10
n<-20
x<-rnorm(n,0,1)
x[(k+1):m]<-x[(k+1):m]+1.5
SemiparChangePoint(x,  alternative = "epidemic")
```

Index

∗ **device**
  plots, 7
∗ **htest**
  BootsChapt, 2
  BootsModelTest, 4
  CriticalValues, 5
  Pvalues, 9
  schapt, 11
  SemiparChangePoint, 13
∗ **models**
  SemiparChangePoint, 13
∗ **nonparametric**
  cumsum.test, 6
∗ **robust**
  schapt, 11

axis, *8*

BootsChapt, 2, *10*
BootsModelTest, 4

CriticalValues, 5
cumsum, *7*
cumsum.test, 6
CV.Epidemic.Vn *(*CriticalValues*)*, 5
CV.Epidemic.Wn *(*CriticalValues*)*, 5

glm.control, *2*, *4*, *11*, *13*
Graf.Diagnostic, *12*
Graf.Diagnostic *(*plots*)*, 7

p.Epidemic.Vn, *3*, *14*
p.Epidemic.Vn *(*Pvalues*)*, 9
p.Epidemic.Wn, *3*, *14*
p.Epidemic.Wn *(*Pvalues*)*, 9
p.OneChange, *3*, *14*
p.OneChange *(*Pvalues*)*, 9
Plot.ll, *12*
Plot.ll *(*plots*)*, 7
plots, 7
Pvalues, 9

schapt, *3*, *5*, *6*, *8*, *10*, 11, *14*
SemiparChangePoint, *2–5*, *8*, 13
Sn.alfa *(*CriticalValues*)*, 5