# The sac Package

May 6, 2005

**Version** 1.0  
**Date** 2005-04-04  
**Title** Semiparametric Analysis of Changepoint  
**Author** Zhong Guan <zguan@iusb.edu>  
**Maintainer** Zhong Guan <zguan@iusb.edu>  
**Depends** R (>= 1.4)  
**Description** Semiparametric empirical likelihood ratio based test of changepoint with one-change or epidemic alternatives with data-based model diagnostic  
**License** GPL 2.0 or newer

---

## R topics documented:

- **BootsChapt** .......................................................... 1  
- **BootsModelTest** .................................................. 3  
- **CriticalValues** .................................................... 4  
- **Pvalues** ............................................................. 6  
- **SemiparChangePoint** .......................................... 7  
- **cumsum.test** ....................................................... 9  
- **plots** .............................................................. 10  
- **sac-internal** ...................................................... 12  
- **schart** ............................................................. 12

## Index

- **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 assign adj.Wn = TRUE to calculate the p-value of Wn.

Author(s)

Zhong Guan (zguan@iusb.edu)
References


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

See Also


Examples

```r
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 semiparametric empirical distribution which are estimates of the underlying population distributions.

Usage

```r
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**: 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 $\alpha$, output of `SemiparChangePoint`
- **Beta**: estimated parameter $\beta$, 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`
Critical Values

Details

Value

<table>
<thead>
<tr>
<th>Delta</th>
<th>The test statistic of the model validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pvalue</td>
<td>The bootstrapped p-value</td>
</tr>
</tbody>
</table>

Note

Author(s)

Zhong Guan (zguan@iusb.edu)

References


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

See Also

SemiparChangePoint, schapt

Examples

```r
## 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.
```

---

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
Critical Values

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

Note

Author(s)

Zhong Guan (zguan@iusb.edu)

References


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)
The \( p \)-values of Test Statistics Based on Asymptotic Distribution

**Description**

Calculate the approximate \( p \)-values of the test statistics \( T_n \), \( V_n \) and \( W_n \) using limit null distributions.

**Usage**

\[
\begin{align*}
p.\text{OneChange}(n, d, S_n) \\
p.\text{Epidemic.Vn}(V_n, d, \text{tol} = 1e-10) \\
p.\text{Epidemic.Wn}(W_n, \text{tol} = 1e-07)
\end{align*}
\]

**Arguments**

- \( S_n \) test statistic \( S_n \) of the one-change alternative
- \( V_n \) test statistic \( V_n \) of the epidemic alternative
- \( W_n \) test statistic \( W_n \) of the epidemic alternative
- \( n \) sample size
- \( d \) dimension of the data value
- \( \text{tol} \) the desired accuracy.

**Details**

**Value**

\[
\begin{align*}
p.\text{value} & \quad \text{\( p \)-value}
\end{align*}
\]

**Note**

**Author(s)**

Zhong Guan (zguan@iusb.edu)

**References**


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

**See Also**

schapt, BootsChapt
Examples

```r
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)
```

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

```r
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 $W_n$ 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 is 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 integral of weighted likelihood ratio for "epidemic" alternative
- \( Wn \): test statistic based supremum of weighted likelihood ratio for "epidemic" alternative
- \( alpha.hat \): estimate of \( \alpha \)
- \( beta.hat \): estimate of \( \beta \)

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. Semiparametric Tests for Change-points with Epidemic Alternatives.

See Also


Examples

```r
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")

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.

Details

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

Note

Author(s)
Zhong Guan ⟨zguan@iusb.edu⟩

References

See Also
cumsum
Examples

```r
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

```r
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 $\alpha$, output of `SemiparChangePoint`
- `Beta`: estimated parameter $\beta$, output of `SemiparChangePoint`
- `Color`: a vector of character strings or color codes for curves of estimated distribution functions $\hat{F}$, $\hat{G}$ and $\tilde{G}$
- `LTY`: vector of lty’s, LTY=c(lty1, lty2, lty3, lty4), corresponds to the above color codes
Author(s)

Zhong Guan (zguan@iusb.edu)

References


See Also

schapt

Examples

```r
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")
```

```r
## 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: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)
```

```r
## #Plots of estimated distribution functions
```
```r
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")
```

---

### sac-internal  
Internal sac Functions

**Description**

Internal speltest functions.

**Usage**

```r
dvalue.epidemic(x, tol = 1e-07)
Deltan.ModelTest(x, k, m, Alpha, Beta)
Z.alfa(alfa, n, d)
```

**Details**

These are not to be called by the user.

---

### 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**

```r
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 $W_n$ 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 change-
point(s) may be detected easily using nonparametric method based on cumsum. Currently, this
function does not check whether the data is separated.

Value

data.name    dataset name
parameter    sample size $n$ and degree(s) of freedom of the df of $S_n$ for "one-change"
alternative
alternative  the alternative hypothesis
statistic    a list contains $S_n$ for "one-change" alternative, $V_n$ and $W_n$ 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), $p(S_n)$, of $S_n$ for "one-change" alternative, $p(V_n)$
and $p(W_n)$, of $V_n$ and $W_n$, respectively, 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 contains the corresponding
p-values, p.boots($S_n$), p.boots($V_n$) and p.boots($W_n$) accordingly.

Note

Statistic $W_n$ 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. There is an R
package, called "strucchange", for testing structural change in linear regression models (see
sctest).

Author(s)

Zhong Guan (zguan@iusb.edu)
References


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

See Also

`Graf.Diagnostic`, `Plot.ll`

Examples

```r
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)
pert(Nile, type="p")
shapat(Nile, alternative = "one.change")
```
Index

*Topic device
  plots, 10
*Topic htest
  BootsChapt, 1
  BootsModelTest, 3
  CriticalValues, 4
  Pvalues, 6
  schapt, 12
  SemiparChangePoint, 7
*Topic internal
  sac-internal, 12
*Topic models
  SemiparChangePoint, 7
*Topic nonparametric
  cumsum.test, 9
*Topic robust
  schapt, 12

axis, 10

BootsChapt, 1, 6
BootsModelTest, 3

CriticalValues, 4
cumsum, 9
cumsum.test, 9
CV.Epidemic.Vn(CriticalValues), 4
CV.Epidemic.Wn(CriticalValues), 4

Deltan.ModelTest (sac-internal), 12
dvalue.epidemic (sac-internal), 12

G.alfa (sac-internal), 12
glm.control, 2, 3, 7, 13
Graf.Diagnostic, 14
Graf.Diagnostic (plots), 10

p.Epidemic.Vn, 3, 8
p.Epidemic.Vn (Pvalues), 6
p.Epidemic.Wn, 3, 8
p.Epidemic.Wn (Pvalues), 6
p.OneChange, 3, 8
p.OneChange (Pvalues), 6
Plot.l1 (plots), 10
plots, 10
Pvalues, 6
sac-internal, 12
schapt, 3–6, 8, 11, 12
sctest, 13
SemiparChangePoint, 2–4, 7, 10
Sn.alfa (CriticalValues), 4
Z.alfa (sac-internal), 12
Zn.alfa (sac-internal), 12