Package 'BreakPoints'

October 12, 2022

Type Package

Title Identify Breakpoints in Series of Data

Version 1.2

Author Santiago I. Hurtado [cre, aut] <santih@carina.fcaglp.unlp.edu.ar>

Maintainer Santiago I. Hurtado <santih@carina.fcaglp.unlp.edu.ar>

Description Compute Buishand Range Test, Pettit Test, SNHT, Student t-test, and Mann-Whitney Rank Test, to identify breakpoints in series. For all functions NA is allowed. Since all of the mention methods identify only one breakpoint in a series, a general function to look for N breakpoint is given. Also, the Yamamoto test for climate jump is available. Alexandersson, H. (1986) <doi:10.1002/joc.3370060607>, Buishand, T. (1982) <doi:10.1016/0022-1694(82)90066-X>, Hurtado, S. I., Zaninelli, P. G., & Agosta, E. A. (2020) <doi:10.1016/j.atmosres.2020.104955>, Mann, H. B., Whitney, D. R. (1947) <doi:10.1214/aoms/1177730491>, Pettitt, A. N. (1979) <doi:10.2307/2346729>, Ruxton, G. D., jul (2006) <doi:10.1093/beheco/ark016>, Yamamoto, R., Iwashima, T., Kazadi, S. N., & Hoshiai, M. (1985) <doi:10.2151/jmsj1965.63.6_1157>.

License GPL-3

Depends R (>= 3.1.0)

Imports MASS, zoo

ByteCompile yes

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-12 10:20:08 UTC

R topics documented:

`N_break_point` *N_break_point*

Description

Look for several breakpoints

Usage

```
N_break_point(serie, n_max = 1, n_period=10,
                        seed=FALSE, auto_select = FALSE,
                        alpha = NULL,method='SNHT',dstr='norm',
                        seed_set = 9658, change_random_seed = TRUE,
                        seed_method = 6842)
```

Arguments

`serie`	numeric vector where the breakpoint is looked for
`n_max`	integer up to how many breakpoint should be test
`n_period`	an integer specifying the minimal length of a complete period to consider
`seed`	if FALSE (default) the first breakpoints to inicialize the iteration are calculated by splitting the serie in equal separated parts. If seed is given must be a list() of length n_max where it is specified the first breakpoints to take for each iteration.
`method`	which method should be used for breakpoint detection, supported: 'student', 'mann-whitney','SNHT' (default),'buishand' and 'pettit'
`dstr`	character specifying which distribution should be used for test simulations, only used if method is SNHT or buishand. Posible distributions 'norm' (default, normal dist), 'gamma', and 'self' (bootstrap)
`auto_select`	logical, should an automatic selection of how many breakpoints are be made, default FALSE
`alpha`	numeric, critical p value to use for auto_select
`seed_set`	Either a number to used to set a seed or NULL to set no seed inside the function
`change_random_seed`	Logical, can the .Random.seed change inside the function, or must remain the same after applying the function
`seed_method`	Either a number to used to set seed inside SNHT or buishand methods or NULL to set no seed

Details

Compute homogeneity test for all possible breaks in the serie considering several breakpoints. NA values are allow. In order to guarantee same result for the same input seed_set and seed_method (if method in SNHT or buishand) must be given.

Value

N_break_point returns a list with the breakpoints index, it's p value and how many breakpoints are. If auto_select = F, a list with one list as specify for each n breakpoint tried

breaks index where the breakpoints are found

p.value p value of the test

n how many breakpoints are looked for

References

Hurtado, S. I., Zaninelli, P. G., & Agosta, E. A. (2020). A multi-breakpoint methodology to detect changes in climatic time series. An application to wet season precipitation in subtropical Argentina. Atmospheric Research, 104955.

Examples

```
# Make a serie with three jumps, same as yamamoto example
set.seed(524)
x <- c(rnorm(30,1,1),rnorm(30,2,1),rnorm(30,1,1),rnorm(20,2,1))

# Look up to 5 breaks using pettit
break_prosition <- N_break_point(serie=x, n_max = 5, method='pettit',
                                 auto_select=TRUE,alpha=0.1)

plot(x)
abline(v = break_prosition$breaks, col='red')
```

SNHT *Buishand Range Test and Standard Normal Homogeneity Test*

Description

Compute Buishand Range Test or Standard Normal Homogeneity Test for a serie, NAs allow in both Test

Usage

```
SNHT(serie,n_period=10,dstr='norm',simulations = 1000,
                seed_set = 9658, change_random_seed = TRUE)

Buishand_R(serie,n_period=10,dstr='norm',simulations = 1000,
                seed_set = 9658, change_random_seed = TRUE)
```

Arguments

serie	numeric vector where the breakpoint is looked for
n_period	an integer specifying the minimal length of a complete period to consider
dstr	character specifying which distribution should be used for test simulations, 'norm' (default; normal distribution), 'gamma', and 'self' (will compute bootstrap)
simulations	an integer specifying how many Monte Carlo simulations to perform, default is 1000.
seed_set	Either a number to used to set a seed or NULL to set no seed inside the function
change_random_seed	Logical, can the .Random.seed change inside the function, or must remain the same after applying the function

Details

SNHT compute Standard Normal Homogeneity Test where NA values are allow. In order to guarantee same result for the same input seed_set must be given.

Buishand_R Compute Buishand Range Test for Homogeneity where NA values are allow. In order to guarantee same result for the same input seed_set must be given.

Value

SNHT and Buishand_R returns a list with the breakpoint index and it's p value

breaks index where the breakpoint is found

p.value p value of the test

References

- Alexandersson, H., jan 1986. A homogeneity test applied to precipitation data. Journal of Climatology 6 (6), 661–675. URL http://doi.wiley.com/10.1002/joc.3370060607

- Buishand, T., aug 1982. Some methods for testing the homogeneity of rainfall records. Journal of Hydrology 58 (1-2), 11–27. URL https://doi.org/10.1016/0022-1694(82)90066-X

Examples

```
# Make a serie with one breakpoint
x <- c(rnorm(60,1,1),rnorm(40,2,1))

# Look for break using SNHT, Buishand_R can be used in exactly the same way
break_prosition <- SNHT(serie = x)

plot(x)
abline(v = break_prosition$breaks)
```

Student *Mann-Whitney-Wilcoxon Test, Student t-test and Pettit Test*

Description

Compute Rolling Mann-Whitney-Wilcoxon Test, Rolling Student t-test and Pettit test for homogeneity, NAs allow.

Usage

```
stu(serie,n_period=10)

man.whi(serie,n_period=10)

pettit(serie,n_period=10)
```

Arguments

serie	numeric vector where the breakpoint is looked for
n_period	an integer specifying the minimal length of a complete period to consider

Details

`man.whi` compute Mann-Whitney-Wilcoxon Test and `stu` the Student t-test for all possible breaks in the serie and get the most significant break.In both test NA values are allow.

`pettit` Compute the Pettit Test for Homogeneity. NA values are allow.

Value

`pettit`, `man.whi` and `stu` returns a list with the breakpoint index and its p value

breaks index where the breakpoint is found

p.value p value of the test

References

- Ruxton, G. D., jul 2006. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology 17 (4), 688–690. URL: http://academic.oup.com/beheco/article/17/4/688/unequal-variance-ttest-is-an-underused

- Mann, H. B., Whitney, D. R., mar 1947. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics 18 (1), 50–60. URL http://projecteuclid.org/euclid.aoms/1177730491

- Pettitt, A. N., 1979. A Non-Parametric Approach to the Change-Point Problem. Applied Statistics 28 (2), 126. URL https://www.jstor.org/stable/10.2307/2346729?origin=crossref

Examples

```
# Make a serie with one breakpoint
x <- c(rnorm(60,1,1),rnorm(40,2,1))

# Look for break using pettit(), man.whi() and stu()
break_prosition_pettit <- pettit(serie = x)
break_prosition_man.whi <- man.whi(serie = x)
break_prosition_stu <- stu(serie = x)

plot(x)
abline(v = break_prosition_pettit$breaks,col='red')
abline(v = break_prosition_man.whi$breaks,col='blue')
abline(v = break_prosition_stu$breaks,col= 'green')
```

`yamamoto` *Yamamoto*

Description

Compute Yamamoto Test for climate jumps in a serie.

Usage

```
yamamoto(serie, alpha = 0.1, n_period = 10)
```

Arguments

`serie`	numeric vector where the breakpoint is looked for
`n_period`	an integer specifying the length of the window to use, can not bet odd
`alpha`	numeric, p value to use

Details

`yamamoto` compute the Yamamoto Test.

Value

`yamamoto` returns a list with the breakpoints indexes and the amount

breaks vector of indexes where the breakpoint is found

n Amount of breakpoints

References

Yamamoto, R., Iwashima, T., Kazadi, S. N., & Hoshiai, M. (1985). Climatic jump: a hypothesis in climate diagnosis. Journal of the Meteorological Society of Japan. Ser. II, 63(6), 1157-1160.

Examples

```
# Make a serie with three jumps, same as N_break_point example
set.seed(524)
x <- c(rnorm(30,1,1),rnorm(30,2,1),rnorm(30,1,1),rnorm(20,2,1))

# Look for break using yamamoto()
break_prosition <- yamamoto(serie = x)

plot(x)
abline(v = break_prosition$breaks, col='red')
```

Index