Package 'hdthreshold'

July 22, 2025

Title Inference on Many Jumps in Nonparametric Panel Regression Models

Version 1.0.0

Description Provides uniform testing procedures for existence and heterogeneity of threshold effects in high-dimensional nonparametric panel regression models. The package accompanies the paper Chen, Keilbar, Su and Wang (2023) ``Inference on many jumps in nonparametric panel regression models". arXiv preprint <doi:10.48550/arXiv.2312.01162>.

Imports fdrtool, KernSmooth, rdrobust, stats

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.1

NeedsCompilation no

Author Georg Keilbar [aut, cre, cph],
Likai Chen [ctb],
Liangjun Su [ctb],
Weining Wang [ctb]

Maintainer Georg Keilbar <georg.keilbar@hu-berlin.de>

Repository CRAN

Date/Publication 2024-12-23 10:50:06 UTC

Contents

K *Uniform kernel function*

Description

Uniform kernel function

Usage

```
K(x)
```

Arguments

x	a vector

Value

a vector of values

Examples

```
K(0)
```

MAinf_normal *Simulate an MA infinity process with algrebraic decay*

Description

Simulate an MA infinity process with algrebraic decay

Usage

```
MAinf_normal(N, beta)
```

Arguments

N	sample size
beta	algebraic decay parameter

Value

simulated MA infinity process

Examples

```
x = MAinf_normal(100, 1.5)
```

`simulation.derivative` *Monte Carlo simulation for existence of derivative threshold effects under known threshold location*

Description

Monte Carlo simulation to study the size and power properties of the uniform test for existence of threshold effects in the first derivative under known threshold locations. Provides the Monte Carlo distribution of the test statistic and empirical rejection probabilities at 10%, 5% and 1% level.

Usage

```
simulation.derivative(
  N,
  TL,
  p,
  M,
  epsilon = c("iid", "factor"),
  running = c("iid", "factor"),
  hetero = c(0, 1)
)
```

Arguments

`N`	cross-sectional dimension
`TL`	time series length
`p`	fraction of non-zero coefficients
`M`	number of Monte Carlo runs
`epsilon`	specification of error term. If `"iid"` is selected the error term is iid standard normal. If `"factor"` is selected, the error term follows a factor model with strong cross-sectional and weak temporal dependence.
`running`	specification of running variable. If `"iid"` is selected the running variable is iid uniformly distributed. If `"factor"` is selected, the running variable follows a factor model with strong cross-sectional and weak temporal dependence.
`hetero`	if `hetero=1` the error term is heteroskedastic, if `hetero=0` the error term is homoskedastic.

Value

A list containing the value of the test statistic for each Monte Carlo run and the empirical rejection rate for a 10%, 5% and 1% confidence level.

Examples

```
result_derivative = simulation.derivative(10, 200, 0, 10, epsilon = "iid",
                    running = "iid", hetero = 0)
```

`simulation.hetero` *Monte Carlo simulation for heterogeneity of threshold effects under known threshold location*

Description

Monte Carlo simulation to study the size and power properties of the uniform test for heterogeneity of threshold effects under known threshold locations. Provides the Monte Carlo distribution of the test statistic and empirical rejection probabilities at 10%, 5% and 1% level.

Usage

```
simulation.hetero(
  N,
  TL,
  p,
  M,
  epsilon = c("iid", "factor"),
  running = c("iid", "factor"),
  hetero = c(0, 1)
)
```

Arguments

`N`	cross-sectional dimension
`TL`	time series length
`p`	fraction of non-zero coefficients
`M`	number of Monte Carlo runs
`epsilon`	specification of error term. If `"iid"` is selected the error term is iid standard normal. If `"factor"` is selected, the error term follows a factor model with strong cross-sectional and weak temporal dependence.
`running`	specification of running variable. If `"iid"` is selected the running variable is iid uniformly distributed. If `"factor"` is selected, the running variable follows a factor model with strong cross-sectional and weak temporal dependence.
`hetero`	if `hetero=1` the error term is heteroskedastic, if `hetero=0` the error term is homoskedastic.

Value

A list containing the value of the test statistic for each Monte Carlo run and the empirical rejection rate for a 10%, 5% and 1% confidence level.

Examples

```
result_hetero = simulation.hetero(10, 200, 0, 10, epsilon = "iid",
                running = "iid", hetero = 0)
```

`simulation.pooled` *Monte Carlo simulation for pooled test of existence of threshold effects*

Description

Monte Carlo simulation to study the size and power properties of the pooled test for existence of threshold effects under unknown threshold locations. The pooled test can be based on a nonparametric regression model or a linear panel threshold regression model. Provides the Monte Carlo distribution of the test statistic and empirical rejection probabilities at 10%, 5% and 1% level.

Usage

```
simulation.pooled(
  N,
  TL,
  p,
  M,
  epsilon = c("iid", "factor"),
  running = c("iid", "factor"),
  hetero = c(0, 1),
  threshold = c("uniform", "exponential", "gaussian"),
  method = c("parametric", "nonparametric")
)
```

Arguments

`N`	cross-sectional dimension
`TL`	time series length
`p`	fraction of non-zero coefficients
`M`	number of Monte Carlo runs
`epsilon`	specification of error term. If `"iid"` is selected the error term is iid standard normal. If `"factor"` is selected, the error term follows a factor model with strong cross-sectional and weak temporal dependence.
`running`	specification of running variable. If `"iid"` is selected the running variable is iid uniformly distributed. If `"factor"` is selected, the running variable follows a factor model with strong cross-sectional and weak temporal dependence.
`hetero`	if `hetero=1` the error term is heteroskedastic, if `hetero=0` the error term is homoskedastic.
`threshold`	specifies the distribution for the non-zero threshold coefficients, possible values are `"normal"` for the standard normel, `"exponential"` for an exponential distribution with parameter 1, or `"uniform"` for a uniform distribution.
`method`	method of estimation (`"nonparametric"` vs. `"parametric"`)

Value

A list containing the value of the test statistic for each Monte Carlo run and the empirical rejection rate for a 10%, 5% and 1% confidence level.

Examples

```
result_pooled = simulation.pooled(5, 400, 0, 10, epsilon = "iid", running = "iid",
                hetero = 0, threshold = "gaussian", method = "nonparametric")
```

---

simulation.threshold *Monte Carlo simulation for existence of threshold effects under known threshold location*

Description

Monte Carlo simulation to study the size and power properties of the uniform test for existence of threshold effects under known threshold locations. Provides the Monte Carlo distribution of the test statistic and empirical rejection probabilities at 10%, 5% and 1% level.

Usage

```
simulation.threshold(
  N,
  TL,
  p,
  M,
  epsilon = c("iid", "factor"),
  running = c("iid", "factor"),
  hetero = c(0, 1)
)
```

Arguments

N	cross-sectional dimension
TL	time series length
p	fraction of non-zero coefficients
M	number of Monte Carlo runs
epsilon	specification of error term. If "iid" is selected the error term is iid standard normal. If "factor" is selected, the error term follows a factor model with strong cross-sectional and weak temporal dependence.
running	specification of running variable. If "iid" is selected the running variable is iid uniformly distributed. If "factor" is selected, the running variable follows a factor model with strong cross-sectional and weak temporal dependence.
hetero	if hetero=1 the error term is heteroskedastic, if hetero=0 the error term is homoskedastic.

Value

A list containing the value of the test statistic for each Monte Carlo run and the empirical rejection rate for a 10%, 5% and 1% confidence level.

Examples

```
result_threshold = simulation.threshold(10, 200, 0, 10, epsilon = "iid",
                   running = "iid", hetero = 0)
```

`simulation.unknown` *Monte Carlo simulation for uniform test of existence of threshold effects under unknown threshold location*

Description

Monte Carlo simulation to study the size and power properties of the uniform test for existence of threshold effects under unknown threshold locations. Provides the Monte Carlo distribution of the test statistic and empirical rejection probabilities at 10%, 5% and 1% level.

Usage

```
simulation.unknown(
  N,
  TL,
  p,
  M,
  epsilon = c("iid", "factor"),
  running = c("iid", "factor"),
  hetero = c(0, 1),
  threshold = c("uniform", "exponential", "gaussian")
)
```

Arguments

`N`	cross-sectional dimension
`TL`	time series length
`p`	fraction of non-zero coefficients
`M`	number of Monte Carlo runs
`epsilon`	specification of error term. If `"iid"` is selected the error term is iid standard normal. If `"factor"` is selected, the error term follows a factor model with strong cross-sectional and weak temporal dependence.
`running`	specification of running variable. If `"iid"` is selected the running variable is iid uniformly distributed. If `"factor"` is selected, the running variable follows a factor model with strong cross-sectional and weak temporal dependence.

hetero	if hetero=1 the error term is heteroskedastic, if hetero=0 the error term is homoskedastic.
threshold	specifies the distribution for the non-zero threshold coefficients, possible values are "normal" for the standard normel, "exponential" for an exponential distribution with parameter 1, or "uniform" for a uniform distribution.

Value

A list containing the value of the test statistic for each Monte Carlo run and the empirical rejection rate for a 10%, 5% and 1% confidence level.

Examples

```
result_unknown = simulation.unknown(2, 800, 0, 10, epsilon = "iid", running = "iid",
                 hetero = 0, threshold = "gaussian")
```

---

threshold.derivative.test

Uniform test for existence of derivative threshold effects

Description

Uniform test for existence of threshold effects in the first derivative for nonparametric panel regressions. Both the known and unknown threshold location case are covered. Apart from the uniform test statistic and the corresponding p-value, a table for the results of the individual threshold estimates and test statistics is provided.

Usage

```
threshold.derivative.test(
  data,
  response,
  running,
  id,
  bw = NULL,
  C = 0,
  alpha = NULL,
  alternative = "two"
)
```

Arguments

data	a data frame containing the response, running and id variables
response	name of the dependent variable (aka response variable)
running	name of the running variable (aka forcing variable)
id	name of the id variable

bw	an optional scalar bandwidth parameter for the local linear estimation. If not specified, the bandwidth is selected by the command rdrobust::rdbwselect().
C	a scalar value for the true threshold location (for the known case) or a grid of candidate threshold locations (for the unknown case)
alpha	specifies a threshold to determine which and how many individual-specific threshold effects and test statistics are displayed in the output table. Only individuals which are significant at the alpha confidence level are selected.
alternative	specifies whether we consider a two-sided alternative (default) or left-/right-sided alternative.

Value

A list containing:

I_hat	the value of the uniform test statistic
p_value	the corresponding p-value
N	the cross-sectional dimension
Critical_values	critical values at 10%, 5%, 1%, and 0.1% confidence level
Table	a table displaying the estimation result for a selection of individuals, including the id variable, the thresh

See Also

threshold.example(), rdrobust::rdbwselect().

Examples

```
d = threshold.example(10, 200, 0.1, 2)
threshold.derivative.test(data = d, response = "y", running = "x", id = "id", C = 0)
```

---

threshold.example *Simulate an example data frame*

Description

Simulate an example data frame

Usage

```
threshold.example(N, TL, p, gamma)
```

Arguments

N	cross-sectional dimension
TL	time series length
p	fraction of non-zero coefficients
gamma	value of non-zero coefficients

Value

simulated data frame

Examples

```
d = threshold.example(10, 200, 0.1, 2)
```

---

threshold.heterogeneity.test

Uniform test for heterogeneity of threshold effects

Description

Uniform test for heterogeneity of threshold effects in a nonparametric panel regression under known threshold locations. Apart from the uniform test statistic and the corresponding p-value, a table for the results of the individual threshold estimates and test statistics is provided.

Usage

```
threshold.heterogeneity.test(
  data,
  response,
  running,
  id,
  bw = NULL,
  c = 0,
  alpha = NULL,
  alternative = "two",
  use.median = FALSE
)
```

Arguments

data	a data frame containing the response, running and id variables
response	name of the dependent variable (aka response variable)
running	name of the running variable (aka forcing variable)
id	name of the id variable

bw	an optional scalar bandwidth parameter for the local linear estimation. If not specified, the bandwidth is selected by the command rdrobust::rdbwselect().
c	a scalar value for the true threshold location
alpha	specifies a threshold to determine which and how many individual-specific threshold effects and test statistics are displayed in the output table. Only individuals which are significant at the alpha confidence level are selected.
alternative	specifies whether we consider a two-sided alternative (default) or left-/right-sided alternative.
use.median	if TRUE, the median replaces the mean as a robust alternative in the test for heterogeneity

Value

A list containing:

Q_hat	the value of the uniform test statistic
p_value	the corresponding p-value
N	the cross-sectional dimension
Critical_values	critical values at 10%, 5%, 1%, and 0.1% confidence level
Table	a table displaying the estimation result for a selection of individuals, including the id variable, the thresh

See Also

threshold.example(), rdrobust::rdbwselect().

Examples

```
d = threshold.example(10, 200, 0.1, 2)
threshold.heterogeneity.test(data = d, response = "y", running = "x", id = "id", c = 0)
```

---

threshold.test *Uniform test for existence of threshold effects*

Description

Uniform test for existence of threshold effects in a nonparametric panel regression. Both the known and unknown threshold location case are covered. Apart from the uniform test statistic and the corresponding p-value, a table for the results of the individual threshold estimates and test statistics is provided.

Usage

```
threshold.test(
  data,
  response,
  running,
  id,
  bw = NULL,
  C = 0,
  alpha = NULL,
  alternative = "two"
)
```

Arguments

data	a data frame containing the response, running and id variables
response	name of the dependent variable (aka response variable)
running	name of the running variable (aka forcing variable)
id	name of the id variable
bw	an optional scalar bandwidth parameter for the local linear estimation. If not specified, the bandwidth is selected by the command rdrobust::rdbwselect().
C	a scalar value for the true threshold location (for the known case) or a grid of candidate threshold locations (for the unknown case)
alpha	specifies a threshold to determine which and how many individual-specific threshold effects and test statistics are displayed in the output table. Only individuals which are significant at the alpha confidence level are selected.
alternative	specifies whether we consider a two-sided alternative (default) or left-/right-sided alternative.

Value

A list containing:

I_hat	the value of the uniform test statistic
p_value	the corresponding p-value
N	the cross-sectional dimension
Critical_values	critical values at 10%, 5%, 1%, and 0.1% confidence level
Table	a table displaying the estimation result for a selection of individuals, including the id variable, the thresh

See Also

threshold.example(), rdrobust::rdbwselect().

Examples

```
d = threshold.example(10, 200, 0.1, 2)
threshold.test(data = d, response = "y", running = "x", id = "id", C = 0)
```

Index