

# Package ‘md’

July 22, 2025

**Type** Package

**Title** Selecting Bandwidth for Kernel Density Estimator with Minimum Distance Method

**Version** 1.0.4

**Date** 2016-02-21

**Author** Genzo Kaga

**Maintainer** Genzo Kaga <genzo.k-1321@ruri.waseda.jp>

**Description** Selects bandwidth for the kernel density estimator with minimum distance method as proposed by Devroye and Lugosi (1996). The minimum distance method directly selects the optimal kernel density estimator from countably infinite kernel density estimators and indirectly selects the optimal bandwidth. This package selects the optimal bandwidth from finite kernel density estimators.

**License** GPL-2 | GPL-3

**LazyData** TRUE

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2016-02-21 17:02:45

## Contents

|                      |   |
|----------------------|---|
| counter . . . . .    | 2 |
| deltaboxm . . . . .  | 2 |
| dmm . . . . .        | 3 |
| fhat . . . . .       | 3 |
| fhatboxm . . . . .   | 4 |
| ker . . . . .        | 4 |
| md . . . . .         | 5 |
| nom . . . . .        | 6 |
| scheffe . . . . .    | 6 |
| subcounter . . . . . | 7 |
| xmm . . . . .        | 7 |
| xym . . . . .        | 8 |

|         |                                                                       |
|---------|-----------------------------------------------------------------------|
| counter | <i>Get 0-1 vector which is used for calculating empirical measure</i> |
|---------|-----------------------------------------------------------------------|

**Description**

Get 0-1 vector which is used for calculating empirical measure

**Usage**

counter(e2, box2)

**Arguments**

- e2

integer
- box2

matrix which has 0-1 elements related to Scheffe set

**Value**

0-1 vector

|           |                          |
|-----------|--------------------------|
| deltaboxm | <i>Calculating delta</i> |
|-----------|--------------------------|

**Description**

Calculating delta

**Usage**

deltaboxm(ij, box, box2, mu\_box2, grid)

**Arguments**

- ij

number order vector
- box

matrix which has estimated values of all kernel density estimators
- box2

matrix which has 0-1 elements related to Scheffe set
- mu\_box2

matrix which has values of all empirical measures
- grid

length of grid in domain

**Value**

delta value

---

|     |                                                    |
|-----|----------------------------------------------------|
| dmm | <i>Get number order matrix which is used in md</i> |
|-----|----------------------------------------------------|

---

**Description**

Get number order matrix which is used in md

**Usage**

```
dmm(i, prod2)
```

**Arguments**

|       |                                |
|-------|--------------------------------|
| i     | integer                        |
| prod2 | integer which is defined in md |

**Value**

matrix related to number order

---

|      |                                                                       |
|------|-----------------------------------------------------------------------|
| fhat | <i>Calculating estimated density value on some x with bandwidth h</i> |
|------|-----------------------------------------------------------------------|

---

**Description**

Calculating estimated density value on some x with bandwidth h

**Usage**

```
fhat(x, h, data)
```

**Arguments**

|      |             |
|------|-------------|
| x    | scalar      |
| h    | bandwidth   |
| data | data sample |

**Value**

estimated density value

**Examples**

```
fhat(0,0.2,rnorm(100))
```

---

|          |                                                                   |
|----------|-------------------------------------------------------------------|
| fhatboxm | <i>Get estimated values of kernel density estimator on domain</i> |
|----------|-------------------------------------------------------------------|

---

**Description**

Get estimated values of kernel density estimator on domain

**Usage**

fhatboxm(ij, data\_for\_d, h, x)

**Arguments**

- ij                    number order vector
- data\_for\_d        data sample which is split to be used for kernel density estimator
- h                    bandwidth
- x                    scalar

**Value**

estimated values of kernel density estimator on domain

---

|     |                        |
|-----|------------------------|
| ker | <i>Kernel function</i> |
|-----|------------------------|

---

**Description**

Kernel function

**Usage**

ker(x)

**Arguments**

- x                    scalar

**Value**

density

---

|    |                                                                                                                                                                                                                                                                                                                                        |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| md | <i>md selects bandwidth for kernel density estimator with minimum distance method. Minimum distance method directly selects optimal kernel density estimator in countably infinite kernel density estimators and indirectly selects optimal bandwidth. md selects optimal bandwidth in countably finite kernel density estimators.</i> |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

---

## Description

md selects bandwidth for kernel density estimator with minimum distance method. Minimum distance method directly selects optimal kernel density estimator in countably infinite kernel density estimators and indirectly selects optimal bandwidth. md selects optimal bandwidth in countably finite kernel density estimators.

## Usage

```
md(data, hnumber, ds)
```

## Arguments

|         |                                                                                                                                  |
|---------|----------------------------------------------------------------------------------------------------------------------------------|
| data    | data sample                                                                                                                      |
| hnumber | the number of bandwidth which md can select. 60 is enough. Of course, you can take it more.                                      |
| ds      | rate of data split. Minimum distance method has to split data for constructing kernel density estimators and empirical measures. |

## Value

bandwidth

## Examples

```
# select bandwidth
md(runif(100), 20, 0.6)

# select bandwidth and plot
data <- rnorm(100)
bandwidth <- md(data, 20, 0.6)
x <- seq(min(data), max(data), length=100)
plot(x, sapply(x, fhat, bandwidth, data), type="l", ylab="density")
```

---

|     |                                                    |
|-----|----------------------------------------------------|
| nom | <i>Get number order matrix which is used in md</i> |
|-----|----------------------------------------------------|

---

**Description**

Get number order matrix which is used in md

**Usage**

nom(i, length)

**Arguments**

|        |                               |
|--------|-------------------------------|
| i      | integer                       |
| length | the number of grids in domain |

**Value**

matrix related to number order

---

|         |                                 |
|---------|---------------------------------|
| scheffe | <i>Calculating Scheffe sets</i> |
|---------|---------------------------------|

---

**Description**

Calculating Scheffe sets

**Usage**

scheffe(ij, box)

**Arguments**

|     |                                                   |
|-----|---------------------------------------------------|
| ij  | number order vector                               |
| box | estimated values of all kernel density estimators |

**Value**

0-1 vector

---

|            |                                               |
|------------|-----------------------------------------------|
| subcounter | <i>Auxiliary function which is used in md</i> |
|------------|-----------------------------------------------|

---

**Description**

Auxiliary function which is used in md

**Usage**

subcounter(1, data, x)

**Arguments**

|      |             |
|------|-------------|
| 1    | integer     |
| data | data sample |
| x    | scalar      |

**Value**

integer

---

|     |                                                    |
|-----|----------------------------------------------------|
| xmm | <i>Get number order vector which is used in md</i> |
|-----|----------------------------------------------------|

---

**Description**

Get number order vector which is used in md

**Usage**

xmm(j)

**Arguments**

|   |         |
|---|---------|
| j | integer |
|---|---------|

**Value**

vector related to number order

---

|     |                                                    |
|-----|----------------------------------------------------|
| xym | <i>Get number order matrix which is used in md</i> |
|-----|----------------------------------------------------|

---

**Description**

Get number order matrix which is used in md

**Usage**

xym(i, length)

**Arguments**

- i integer
- length the number of grids in domain

**Value**

matrix related to number order

# Index

counter, [2](#)

deltaboxm, [2](#)  
dmm, [3](#)

fhat, [3](#)  
fhatboxm, [4](#)

ker, [4](#)

md, [5](#)

nom, [6](#)

scheffe, [6](#)  
subcounter, [7](#)

xmm, [7](#)  
xym, [8](#)