Package ‘ICE’

February 19, 2015

Title Iterated Conditional Expectation

Version 0.69

Author W. John Braun

Description Kernel Estimators for Interval-Censored Data

LazyLoad true

LazyData true

Depends R (>= 2.0.1), KernSmooth, stats

ZipData no

Maintainer W. John Braun <braun@stats.uwo.ca>

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2013-03-28 19:53:05

R topics documented:

`bickde` *Bandwidth choice for Interval-Censored Kernel Density Estimation*

Description

Likelihood Cross-Validation bandwidth choice for interval-censored kernel density estimates. Also computed is the direct-plug-in estimate (using the KernSmooth function dpik based on the interval midpoints.

Usage

```
bickde(data, factor)
```

Arguments

`data`	A matrix with two columns, consisting of left and right endpoints of the interval data
`factor`	A scalar factor which gives upper and lower bounds for the initial interval for the golden section search relative to a preliminary value computed by dpik. Default value is 10 which specifies an initial interval of (dpik/10, dpik*10).

Details

Maximization of the likelihood is accomplished by golden section search using the optimize() function. This can be very SLOW.

Value

A scalar value.

Author(s)

W.J. Braun

References

Braun, J., Duchesne, T. and Stafford, J.E. (2005) Local likelihood density estimation for interval censored data. Canadian Journal of Statistics 33: 39-60.

Examples

```
# Not run
# bickde(ICHemophiliac)
```

ICGG *Gentleman and Geyer's Data*

Description

Gentleman and Geyer's Data.

Usage

```
data(ICGG)
```

Format

A data frame with 6 observations on the following 2 variables.

left the left interval endpoint

right the right interval endpoint

Source

Gentleman, R. and Geyer, C. (1994) Maximum Likelihood for Interval Censored Data: Consistency and Computation. Biometrika.

Examples

```
library(KernSmooth)
tmp <- apply(ICGG, 1, mean)  # tmp now contains the interval midpoints
h <- dpik(tmp)               # direct-plug-in bandwidth selected for
                             # interval midpoints
par(mfrow=c(1,2))
estimate <- ickde(ICGG, h=h, m=200)
plot(estimate, type="l", main="One fixed point")
estimate <- ickde(ICGG, f=c(rep(1,60),rep(0,90),rep(1,60)), h=.1)
plot(estimate, type="l", main="Another fixed point")
```

ICHemophiliac *Hemophiliac Data Set*

Description

Time (interval-censored) to onset of HIV infection.

Usage

```
data(ICHemophiliac)
```

Format

a matrix with 2 columns. The first column contains the left endpoints of the intervals, and the second column contains the right endpoints.

---

ickde *Interval-Censored Kernel Density Estimation*

Description

Iterated conditional expectation kernel density estimation using a local constant. The bandwidth is assumed fixed. (See the example for a way to get a quick ballpark estimate of the bandwidth.) The gaussian, epanechnikov and biweight kernels can be used. Note that the bandwidth estimate would have to be adjusted before using with epanechnikov or biweight.

Usage

```
ickde(I, h, f, m, n.iterations = 10, x1, xm, right.limit = 10000,kernel="gaussian", old=TRUE)
```

Arguments

`I`	A matrix with two columns, consisting of left and right endpoints of the interval data
`h`	A scalar bandwidth
`f`	An initial estimate of the density at a sequence of grid points (optional; if this is used, do not specify m)
`m`	The number of (equally-spaced) grid points at which the density is to be estimated
`n.iterations`	The maximum number of iterations allowed
`x1`	The left-most grid point (optional)
`xm`	The right-most grid point (optional)
`right.limit`	For right-censored data, the value to be used as an artificial right endpoint for the intervals
`kernel`	character argument indicated choice of kernel; current choices are "gaussian", "epanechnikov", "biweight"
`old`	logical value, indicating whether denominators in conditional expectation calculation should use the previous value of the density estimate.

Value

An object of class `IC`

Author(s)

W.J. Braun

References

Braun, J., Duchesne, T. and Stafford, J.E. (2005) Local likelihood density estimation for interval censored data. Canadian Journal of Statistics 33: 39-60.

Examples

```
tmp <- apply(ICHemophiliac, 1, mean)
h <- try(dpik(tmp), silent=T) # dpik() will work if KernSmooth is loaded
if (class(h) !="numeric" ) h <- .9  # this makes the example work
                      # if KernSmooth is not loaded
estimate <- ickde(ICHemophiliac, m=200, h=h)
plot(estimate, type="l")
```

`iclllde` *Interval-Censored Local Linear Density Estimation*

Description

This is the local linear version of `ickde`.

Usage

```
iclllde(I, h, f, m, n.iterations = 10, x1, xm, right.limit = 10000, kernel="gaussian")
```

Arguments

`I`	A matrix with two columns, consisting of left and right endpoints of the interval data
`h`	A scalar bandwidth
`f`	An initial estimate of the density at a sequence of grid points (optional)
`m`	The number of (equally-spaced) grid points at which the density is to be estimated
`n.iterations`	The maximum number of iterations allowed
`x1`	The minimum grid point (optional)
`xm`	The maximum grid point (optional)
`right.limit`	For right-censored data, the value to be used as an artificial right endpoint for the intervals
`kernel`	character argument indicated choice of kernel; current choices are "gaussian", "epanechnikov", "biweight"

Value

An object of class `IC`

Examples

```
data(ICHemophiliac)
estimate <- iclde(ICHemophiliac, m=200, h=.9)
plot(estimate, type="l")
```

---

iclocpoly — *Interval-Censored Local Polynomial Regression Estimation*

Description

Local polynomial regression estimation for interval-censored data.

Usage

```
iclocpoly(x, y=NULL, y.IC, degree=0, h, niter=10, kernel="normal", gridsize=401)
```

Arguments

x	uncensored explanatory variable vector
y	uncensored portion of response vector (optional)
y.IC	two-column matrix of left and right interval endpoints for censored responses
degree	degree of local polynomial
h	bandwidth
niter	number of iterations
kernel	smoothing kernel to be used; default is "normal"; other choices as in the function locpoly
gridsize	number of gridpoints; again as in locpoly

Value

A list consisting of the explanatory variable x and the imputed responses y as well as the estimate of sigma. Usually, one would apply locpoly with a similar bandwidth to obtain the final fitted model.

Examples

```
library(KernSmooth)
data(motor.IC)
estimate <- iclocpoly(x=motor.IC$V1, y.IC = cbind(motor.IC$y.L,
motor.IC$y.R), degree=0, h=1)
plot(motor.IC$V1, motor.IC$V2)
lines(locpoly(estimate$x, estimate$y, bandwidth=1, degree=0),col=4)
```

`inmost` *Innermost Intervals for Interval-Censored Data*

Description

This function calculates the innermost intervals (Turnbull's algorithm) for interval-censored data. Right-censored data is not allowed at this point.

Usage

```
inmost(data, eps)
```

Arguments

`data`	A matrix with two columns, consisting of left and right endpoints of the interval data
`eps`	A scalar additive value which jitters the data slightly to break ties.

Value

A 2 column matrix of the left and right endpoints of the innermost intervals.

Author(s)

W.J. Braun

References

Turnbull (1976) The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society Series B, 38: 290-295.

`likelihoodcv` *Log Cross-Validated Likelihood*

Description

Log Cross-Validated Likelihood for interval-censored data. The likelihood is the product of integrals over the innermost intervals. Leave-one-out cross-validation here is accomplished by leaving out each innermost interval and re-computing the integral using the remaining data.

Usage

```
likelihoodcv(p, data, m=101)
```

Arguments

`p`	A scalar specifying the bandwidth used in the kernel density estimate.
`data`	A matrix with two columns, consisting of left and right endpoints of the interval data
`m`	numeric quantity which controls the number of grid points used in the density estimate

Value

A scalar value.

Author(s)

W.J. Braun

References

Braun, J., Duchesne, T. and Stafford, J.E. (2005) Local likelihood density estimation for interval censored data. Canadian Journal of Statistics 33: 39-60.

`motor.IC` *Motorcycle Data with artificial interval-censored responses*

Description

The original motorcycle data with exponential random variables added and subtracted to the responses to create interval-censored responses.

Usage

```
data(motor.IC)
```

Format

A data frame with 4 variables

V1 time, the explanatory vector

V2 acceleration, the response

y.L y - an exponential random variable

y.R y + an exponential random variable

Index