Package 'cenROC'

March 17, 2020

Type Package

Title Estimation of Time-Dependent ROC Curve and AUC for Censored Survival Data

Version 1.0.0

Author Kassu Mehari Beyene and Anouar El Ghouch

Maintainer Kassu Mehari Beyene <kasu.beyene@uclouvain.be>

Description Contains functions to estimate smoothed and non-smoothed (empirical) time-dependent ROC curve and the corresponding area under the ROC curve (AUC) for the right censored survival data.

Depends R(>= 3.6.0), condSURV

License GPL(>=2)

Encoding UTF-8

LazyData true

RoxygenNote 7.0.2

R topics documented:

cenROC 1
CV 3
mayo 4
NR 5
PI 6

Index **7**

`cenROC` *Estimation of time-dependent ROC curve for right censored survival data*

Description

This function computes the time-dependent ROC curve for a right censored survival data using the cumulative sensitivity and dynamic specificity definitions. The ROC curves can be either empirical (non-smoothed) or smoothed with/wtihout boundary correction. It also calculates the time-dependent area under the ROC curve (AUC).

Usage

```
cenROC(Y, M, censor, t, U = NULL, h = NULL, bw = "NR", method = "tra",
    ktype = "normal", ktype1 = "normal", plot = TRUE)
```

Arguments

Y	The numeric vector of event-times or observed times.
M	The numeric vector of marker values for which the time-dependent ROC curves is computed.
censor	The censoring indicator, 1 if event, 0 otherwise.
t	A scaler time point at which we want to compute the time-dependent ROC curve.
U	The vector of grid points where the ROC curve is estimated. The default is a sequence of 151 numbers between 0 and 1.
h	A scaler for the bandwidth of Beran's weight calculaions. The defualt is the value obtained by using the method of Sheather and Jones (1991).
bw	A character string specifying the bandwidth estimation method for the ROC itself. The possible options are "NR" for the normal reference, the plug-in "PI" and the cross-validation "CV". The default is the "NR" normal reference method. It can also be user specified numerical value.
method	The method of ROC curve estimation. The possible options are "emp" emperical metod; "untra" smooth without boundary correction and "tra" is smooth ROC curve estimation with boundary correction. The default is the "tra" smooth ROC curve estimate with boundary correction.
ktype	A character string giving the type kernel distribution to be used for smoothing the ROC curve: "normal", "epanechnikov", "biweight", or "triweight". By default, the "normal" kernel is used.
ktype1	A character string specifying the desired kernel needed for Beran weight calculation. The possible options are "normal", "epanechnikov", "tricube", "boxcar", "triangular", or "quartic". The defaults is "normal" kernel density.
plot	The logical parameter to see the ROC curve plot. The default is TRUE.

Details

The empirical (non-smoothed) ROC estimate and the smoothed ROC estimate with/without boundary correction can be obtained using this function. The smoothed ROC curve estimators require selecting two bandwidth parametrs: one for Beran's weight calculation and one for smoothing the ROC curve. For the latter, three data-driven methods: the normal reference "NR", the plug-in "PI" and the cross-validation "CV" were implemented. To select the bandwidth parameter needed for Beran's weight calculation, by default, the plug-in method of Sheather and Jones (1991) is used but it is also possible introduce a numeric value. See Beyene and El Ghouch (2020) for details.

Value

Returns the following items:

ROC The vector of the estimated ROC at U for the given t. These will be numeric numbers between 0 and 1.

U The vector of grid points used.

AUC The estimated area under the ROC curve at a given t.

bw The computed value of bandwidth. For the empirical method this is always 1.

Author(s)

Kassu Mehari Beyene, Catholic University of Louvain. <kasu.beyene@uclouvain.be>

Anouar El Ghouch, Catholic University of Louvain. <anouar.elghouch@uclouvain.be>

References

Beyene, K. M. and El Ghouch A. (2020). Smoothed time-dependent ROC curves for right-censored survival data. *submitted*.

Sheather, S. J. and Jones, M. C. (1991). A Reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*. Series B (Methodological) 53(3): 683–690.

Examples

```
library(cenROC)

data(mayo)
cenROC(Y=mayo$time, M=mayo$mayoscore5, censor=mayo$censor, t=365*6)$AUC
```

CV *The cross-validation bandwidth selection for weighted data*

Description

This function computes the data-driven bandwidth for smoothing the ROC (or distribution) function using the CV method of Beyene and El Ghouch (2020). This is an extension of the classical (unweighted) cross-validation bandwith selection method to the case of weighted data.

Usage

```
CV(X, wt, ktype = "normal")
```

Arguments

X	The numeric data vector.
wt	The non-negative weight vector.
ktype	A character string giving the type kernel to be used: "normal", "epanechnikov", "biweight", or "triweight". By default, the "normal" kernel is used.

Details

Bowman et al (1998) proposed the cross-validation bandwidth selection method for unweighted kernal smoothed distribution function. This method is implemented in the R package kerdiest. We adapted this for the case of weighted data by incorporating the weight variable into the cross-validation function of Bowman's method. See Beyene and El Ghouch (2020) for details.

Value

Returns the computed value for the bandwith parameter.

Author(s)

Kassu Mehari Beyene, Catholic University of Louvain. <kasu.beyene@uclouvain.be>

Anouar El Ghouch, Catholic University of Louvain. <anouar.elghouch@uclouvain.be>

References

Beyene, K. M. and El Ghouch A. (2020). Smoothed time-dependent ROC curves for right-censored survival data. *submitted*.

Bowman A., Hall P. and Trvan T.(1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika* 85:799-808.

Quintela-del-Rio, A. and Estevez-Perez, G. (2015). `kerdiest`: Nonparametric kernel estimation of the distribution function, bandwidth selection and estimation of related functions. `R` package version 1.2.

Examples

```
## Not run: library(cenROC)

X <- rnorm(100) # random data vector
wt <- runif(100) # weight vector

## Cross-validation bandwidth selection
CV(X = X, wt = wt)$bw


## End(Not run)
```

`mayo` *Mayo Marker Data*

Description

Two marker values with event time and censoring status for the subjects in Mayo PBC data.

Usage

```
data(mayo)
```

Format

A data frame with 312 observations and 4 variables: time (event time/censoring time), censor (censoring indicator), mayoscore4, mayoscore5. The two scores are derived from 4 and 5 covariates respectively.

References

Heagerty, P. J., and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics*, 61(1), 92-105.

NR *The normal reference bandwidth selection for weighted data*

Description

This function computes the data-driven bandwidth for smoothing the ROC (or distribution) function using the NR method of Beyene and El Ghouch (2020). This is an extension of the classical (unweighted) normal reference bandwith selection method to the case of weighted data.

Usage

```
NR(X, wt, ktype = "normal")
```

Arguments

X	The numeric data vector.
wt	The non-negative weight vector.
ktype	A character string giving the type kernel to be used: "normal", "epanechnikov", "biweight", or "triweight". By default, the "normal" kernel is used.

Details

See Beyene and El Ghouch (2020) for details.

Value

Returns the computed value for the bandwith parameter.

Author(s)

Kassu Mehari Beyene, Catholic University of Louvain. <kasu.beyene@uclouvain.be>

Anouar El Ghouch, Catholic University of Louvain. <anouar.elghouch@uclouvain.be>

References

Beyene, K. M. and El Ghouch A. (2020). Smoothed time-dependent ROC curves for right-censored survival data. *submitted*.

Examples

```
library(cenROC)

X <- rnorm(100) # random data vector
wt <- runif(100) # weight vector

## Normal reference bandwidth selection
NR(X = X, wt = wt)$bw
```

PI *The plug-in bandwidth selection for weighted data*

Description

This function computes the data-driven bandwidth for smoothing the ROC (or distribution) function using the PI method of Beyene and El Ghouch (2020). This is an extension of the classical (unweighted) direct plug-in bandwith selection method to the case of weighted data.

Usage

```
PI(X, wt, ktype = "normal")
```

Arguments

X	The numeric vector of random variable.
wt	The non-negative weight vector.
ktype	A character string giving the type kernel to be used: "normal", "epanechnikov", "biweight", or "triweight". By default, the "normal" kernel is used.

Details

See Beyene and El Ghouch (2020) for details.

Value

Returns the computed value for the bandwith parameter.

Author(s)

Kassu Mehari Beyene, Catholic University of Louvain. <kasu.beyene@uclouvain.be>

Anouar El Ghouch, Catholic University of Louvain. <anouar.elghouch@uclouvain.be>

References

Beyene, K. M. and El Ghouch A. (2020). Smoothed time-dependent ROC curves for right-censored survival data. *submitted*.

Examples

```
library(cenROC)

X <- rnorm(100) # random data vector
wt <- runif(100) # weight vector

## Plug-in bandwidth selection
PI(X = X, wt = wt)$bw
```

Index

∗Topic **datasets**
mayo, 4

cenROC, 1
CV, 3

mayo, 4

NR, 5

PI, 6