Package 'csabounds'

September 6, 2017

Title Bounds on Distributional Treatment Effect Parameters

Version 1.0.0

Description The joint distribution of potential outcomes is not typically identified under standard identifying assumptions such as selection on observables or even when individuals are randomly assigned to being treated. This package contains methods for obtaining tight bounds on distributional treatment effect parameters when panel data is available and under a Copula Stability Assumption as in Callaway (2017) <https://ssrn.com/abstract=3028251>.

Imports stats, ggplot2, BMisc, pbapply, progress, qte

Depends R (>= 3.0)

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Author Brantly Callaway [aut, cre]

Maintainer Brantly Callaway <brantly.callaway@temple.edu>

Repository CRAN

Date/Publication 2017-09-06 18:07:28 UTC

R topics documented:

attcpo *attcpo*

Description

compute the Average Treatment Effect on the Treated Conditional on the previous outcome (ATT-CPO)

Usage

```
attcpo(formla, t, tmin1, tmin2, tname, data, idname, Y0tqteobj, h = NULL,
  yseq = NULL, yseqlen = 100, se = TRUE, iters = 100,
  method = "level")
```

Arguments

formla	e.g. y ~ treat
t	the last time period
tmin1	the middle time period
tmin2	the first time period
tname	the name of the column containing time periods in the data
data	a data.frame
idname	the name of the column containing an individual identifier over time
Y0tqteobj	a qte object (from the qte package) containing the the counterfactual distribution of untreated potential outcomes for the treated group
h	optional bandwidth
yseq	optional sequence of y values, default is to use all unique yvalues in the data, though this can increase computation time
yseqlen	optional length of y values to use, aids in automatically generating yseq if desired
se	whether or not to compute standard errors
iters	how many bootstrap iterations to use if computing standard errors; default is 100.
method	should be either "levels" or "rank"; whether to compute the ATT-CPO using based on the levels of Y0tmin1 or the ranks of Y0tmin1; "levels" is the default.

Value

att-cpo

Examples

```
data(displacements)
cc <- qte::CiC(learn ~ treat,
               t=2011, tmin1=2007, tname="year",
               idname="id", panel=TRUE, data=displacements,
               probs=seq(.05,.95,.01),se=FALSE)
cc$F.treated.tmin1 <- ecdf(subset(displacements, year==2007 & treat==1)$learn)
cc$F.treated.tmin2 <- ecdf(subset(displacements, year==2003 & treat==1)$learn)
ac <- attcpo(learn ~ treat, 2011, 2007, 2003, "year", displacements,
        "id", cc, method="rank", yseqlen=10)
ac
ggattcpo(ac)
```

`csa.bounds` *csa.bounds*

Description

Compute bounds on the distribution and quantile of the treatment effect as given in Callaway (2017) under the copula stability assumption and when a first step estimator of the counterfactual distribution of untreated potential outcomes for the treated group is available.

Usage

```
csa.bounds(formla, t, tmin1, tmin2, tname, idname, data, delt.seq, y.seq,
  Y0tqteobj, F.y0 = NULL, F.y1 = NULL, h = NULL, method = c("level",
  "rank"), cl = 1)
```

Arguments

`formla`	outcomevar ~ treatmentvar
`t`	the 3rd period
`tmin1`	the 2nd period
`tmin2`	the 1st period
`tname`	the name of the column containing periods
`idname`	the name of the column containing ids
`data`	a panel data frame
`delt.seq`	the possible values to compute bounds on the distribution of the treatment effect for
`y.seq`	the possible values for y to take
`Y0tqteobj`	a qte object for obtaining the counterfactual distribution of untreated potential outcomes for the treated group in period t
`F.y0`	(optional) pre-computed distribution of counterfactual untreated outcomes for the treated group

F.y1	(optional) pre-computed distribution of treated outcomes for the treated group
h	optional bandwidth
method	"level" or "rank" determining whether method should be used conditional on ytmin1 or the rank of ytmin1
cl	(optional) number of multi-cores to use

Value

csaboundsobj

Examples

```
## Not run:
data(displacements)
delt.seq <- seq(-4,4,length.out=50)
y.seq <- seq(6.5,13,length.out=50)
cc <- qte::CiC(learn ~ treat,
               t=2011, tmin1=2007, tname="year",
               idname="id", panel=TRUE, data=displacements,
               probs=seq(.05,.95,.01),se=FALSE)
cc$F.treated.tmin2 <- ecdf(subset(displacements, year==2003 & treat==1)$learn)
cc$F.treated.tmin1 <- ecdf(subset(displacements, year==2007 & treat==1)$learn)
cb <- csa.bounds(learn ~ treat, 2011, 2007, 2003, "year", "id",
        displacements, delt.seq, y.seq, cc,
        method="level", cl=1)
cb
ggCSABounds(cb)

## End(Not run)
```

---

displacements *Job displacement data from the NLSY*

Description

A dataset with 266 observations of displaced and non-displaced workers during the great recession and their earnings in 2003, 2007, and 2011. All displaced workers in the sample are displaced in either 2008 or 2009 so 2003 and 2007 are pre-displacement periods.

Usage

```
displacements
```

Format

A data frame with 798 rows 3 columns:

id individual identifier

treat 1 for displaced individuals, 0 otherwise

year the year of earnings

learn the log of earnings in year t for individual i

Source

subset of NLSY 1979 data used in Callaway (2017)

`E.Y0` *E.Y0*

Description

a function for computing the conditional expectation of Y_0t given particular value of y_tmin1 under the Copula Stability Assumption

Usage

```
E.Y0(ytmin1val, Y0tmin1, Y0tmin2, Y0tqteobj, h = NULL, method = "level")
```

Arguments

`ytmin1val`	the value to compute the conditional expectation for
`Y0tmin1`	a vector of untreated potential outcomes for the treated group in period t-1
`Y0tmin2`	a vector of untreated potential outcomes for the treated group in period t-2
`Y0tqteobj`	a qte object which should have set F.treated.t.cf which is the counterfactual distribution of untreated potential outcomes for the treated group in period t
`h`	optional bandwidth paramater
`method`	can be "level" or "rank", whether the conditional expectation is based on the level of Y0tmin1 or its rank

Examples

```
data(displacements)
ytmin1 <- 10
Y0tmin1 <- subset(displacements, year==2007 & treat==1)$learn
Y0tmin2 <- subset(displacements, year==2003 & treat==1)$learn
cc <- qte::CiC(learn ~ treat,
               t=2011, tmin1=2007, tname="year",
               idname="id", panel=TRUE, data=displacements,
               probs=seq(.05,.95,.01),se=FALSE)
cc$F.treated.tmin2 <- ecdf(subset(displacements, year==2003 & treat==1)$learn)
```

```
cc$F.treated.tmin1 <- ecdf(subset(displacements, year==2007 & treat==1)$learn)
E.Y0(ytmin1, Y0tmin1, Y0tmin2, cc)
```

E.Y1 *E.Y1*

Description

a function for computing the conditional expectation of Y_1t given a value for Y_0tmin1

Usage

```
E.Y1(ytmin1val, Y1t, Y0tmin1, h = NULL, method = "level")
```

Arguments

ytmin1val	scalar value to compute conditional expectation for
Y1t	vector of treated potential outcomes for the treated group in period t
Y0tmin1	vector of untreated potential outcomes for the treated group in period t-1
h	optional bandwidth paramater
method	can be "level" or "rank", whether the conditional expectation is based on the level of Y0tmin1 or its rank

Value

the conditional expectation of y1 conditional on y0tmin1

Examples

```
data(displacements)
ytmin1 <- 10
Y1t <- subset(displacements, year==2011 & treat==1)$learn
Y0tmin1 <- subset(displacements, year==2007 & treat==1)$learn
E.Y1(ytmin1, Y1t, Y0tmin1)
```

F.Y0 *F.Y0*

Description

compute F(y|ytmin1) where F is the conditional distribution of untreated potential outcomes for the treated group conditional on ytmin1; This is computed under the copula stability assumption

Usage

```
F.Y0(ytmin1, y.seq, Y0tmin1, Y0tmin2, Y0tqteobj, h = NULL, method = "level")
```

Arguments

ytmin1	the value of ytmin1 to condition on
y.seq	possible values for y to take
Y0tmin1	vector of outcomes for the treated group in period t-1
Y0tmin2	vector of outcomes for the treated group in period t-2
Y0tqteobj	a qte object for obtaining the counterfactual distribution of untreated potential outcomes for the treated group in period t
h	optional bandwidth
method	"level" or "rank" determining whether method should be used conditional on ytmin1 or the rank of ytmin1

Value

distribution F(y|ytmin1)

Examples

```
data(displacements)
ytmin1 <- 10
Y1t <- subset(displacements, year==2011 & treat==1)$learn
Y0tmin1 <- subset(displacements, year==2007 & treat==1)$learn
Y0tmin2 <- subset(displacements, year==2003 & treat==1)$learn
y.seq <- seq(min(c(Y0tmin2,Y0tmin1,Y1t)), max(c(Y0tmin2,Y0tmin1,Y1t)), length.out=100)
cc <- qte::CiC(learn ~ treat,
               t=2011, tmin1=2007, tname="year",
               idname="id", panel=TRUE, data=displacements,
               probs=seq(.05,.95,.01),se=FALSE)
cc$F.treated.tmin2 <- ecdf(subset(displacements, year==2003 & treat==1)$learn)
cc$F.treated.tmin1 <- ecdf(subset(displacements, year==2007 & treat==1)$learn)
F.Y0(ytmin1, y.seq, Y0tmin1, Y0tmin2, cc)
```

F.Y1 *F.Y1*

Description

calculate F(y|ytmin1), the conditional distribution of treated potential outcomes conditional on ytmin1; The order of the variables is due to the way that the function is called later on

Usage

```
F.Y1(ytmin1, y.seq, Y1t, Y0tmin1, h = NULL, method = "level")
```

Arguments

ytmin1	the value of ytmin1 to condition on
y.seq	possible values for y to take
Y1t	vector of outcomes for the treated group in period t
Y0tmin1	vector of outcomes for the treated group in period t-1
h	optional bandwidth
method	"level" or "rank" determining whether method should be used conditional on ytmin1 or the rank of ytmin1

Value

distribution F(y|ytmin1)

Examples

```
data(displacements)
ytmin1 <- 10
Y1t <- subset(displacements, year==2011 & treat==1)$learn
Y0tmin1 <- subset(displacements, year==2007 & treat==1)$learn
y.seq <- seq(min(c(Y0tmin1,Y1t)), max(c(Y0tmin1,Y1t)), length.out=100)
F.Y1(ytmin1, y.seq, Y1t, Y0tmin1)
```

ggattcpo *ggattcpo*

Description

plot the ATT-CPO using ggplot2

Usage

```
ggattcpo(attcpoobj, ylim = NULL)
```

Arguments

attcpoobj	an attcpo object
ylim	optional limits of the plot

ggCSABounds *ggCSABounds*

Description

plot bounds on the quantile of the treatment effect using ggplot2

Usage

```
ggCSABounds(csaboundsobj, tau = seq(0.05, 0.95, 0.05), wdbounds = FALSE,
  otherdist1 = NULL, otherdist2 = NULL)
```

Arguments

csaboundsobj	an object returned from the csa.bounds method
tau	vector of values between 0 and 1 to plot quantiles for
wdbounds	boolean whether or not to also plot Williamson-Downs bounds
otherdist1	optional ecdf of the distribution of the treatment effect under cross sectional rank invariance
otherdist2	optional ecdf of the distribution of the treatment effect under panel rank invariance

Index