Package ‘LARF’

October 12, 2022

Type Package
Title Local Average Response Functions for Instrumental Variable Estimation of Treatment Effects
Version 1.4
Date 2016-07-25
Author Weihua An and Xuefu Wang, Indiana University Bloomington
Maintainer Weihua An <weihuaan@indiana.edu>
Description Provides instrumental variable estimation of treatment effects when both the endogenous treatment and its instrument are binary. Applicable to both binary and continuous outcomes.
Imports Formula
License GPL-3
NeedsCompilation no
Repository CRAN
Date/Publication 2016-07-26 12:04:14

R topics documented:
c401k ................................................................. 2
cvlm ................................................................. 3
Generate.Powers ............................................... 4
larf ................................................................. 4
larf.fit ............................................................ 7
npse ................................................................. 8
predict.larf ....................................................... 9
print.larf ........................................................ 10
summary.larf .................................................... 11
vcov.larf ......................................................... 11

Index 13
Description
Cross-sectional data with 9,275 observations including 11 variables on eligibility for and participation in 401(k) along with income and demographic information.

Usage
data(c401k)

Format
pira  participation in IRA, participation = 1
nettfa net family financial assets in $1000
p401k participation in 401(k), participation = 1
e401k eligibility for 401(k), eligible = 1
inc  income
incsq income square
marr  marital status, married = 1
male  sex, male = 1
age  age
agesq age square
fsize family size

Details
An exemplary data to illustrate the usage of larf. The data includes both a binary outcome (pira) and a continuous outcome (nettfa). The treatment is participation in 401k, p401k. Eligibility for 401(k), e401k, is used as an instrument for p401k.

Source

References

See Also
larf, larf.fit
cvlm

Examples

data(c401k)

cvlm(form.lm, data, m=10, seed = NULL)

Description

Provides cross-validation of a linear regression model

Usage

cvlm(form.lm, data, m=10, seed = NULL)

Arguments

form.lm formula of the regression model.
data data including outcome and covaraites.
m the number of folds to be used in cross-validation.
seed random starting number used to replicate cross-validation.

Details

This function finds the optimal order of the covariates power series through cross-validation.

Value

sumres Sum of residual squares divided by degree of freedom.
df Degree of freedom which equals to the number of valid predictions minus the number of parameters.
m the number of folds to be used in cross-validation.
seed The random seed.

Note

In making the code, we adopted part of the CVlm in DAAG (Maindonald and Braun, 2015).
https://cran.r-project.org/package=DAAG

Author(s)

Weihua An, Departments of Sociology and Statistics, Indiana University Bloomington, <weihuaan@indiana.edu>.
Xuefu Wang, Department of Statistics, Indiana University Bloomington, <wangxuef@umail.iu.edu>.

See Also

larf, npse
Generate.Powers  Generating Powers Series of Variables

Description

Internal function used by npse to generate covariates power series.

Usage

Generate.Powers(X, lambda)

Arguments

X  covariates.
lambda  the maximal order of power series.

Author(s)

Weihua An, Departments of Statistics and Sociology, Indiana University Bloomington, <weihuaan@indiana.edu>. Xuefu Wang, Department of Statistics, Indiana University Bloomington, <wangxuef@u mail.iu.edu>.

See Also

larf, npse

larf  Local Average Response Functions for Instrumental Variable Estimation of Treatment Effects

Description

The function provides instrumental variable estimation of treatment effects when both the endogenous treatment and its instrument are binary. Applicable to both binary and continuous outcomes.

Usage

larf(formula, treatment, instrument, data, method = "LS", AME = FALSE, optimizer = "Nelder-Mead", zProb = NULL)
Arguments

**formula** specification of the outcome model in the form like either \( y \sim x_1 + x_2 \) or \( y \sim X \) where \( X \) is a matrix containing all the covariates excluding the treatment. Also support multi-part formulas (Zeileis and Croissant, 2010). For example, \( y + d \sim x_1 + x_2 \mid z \), where \( d \) represents the treatment and \( z \) the instrument.

**treatment** A vector containing the binary treatment.

**instrument** A vector containing the binary instrument for the endogenous treatment.

**data** an optional data frame. If unspecified, the data will be taken from the working environment.

**method** the estimation method to be used. The default is “LS”, standing for least squares. “ML”, standing for maximum likelihood, is an alternative.

**AME** whether average marginal effects (AME) should be reported. The default is FALSE, in which case marginal effects at the means (MEM) are reported.

**optimizer** the optimization algorithm for the ML method. It should be one of “Nelder-Mead”, “BFGS”, “CG”, “L-BFGS-B”, “SANN”, or “Brent”. See `optim` in R for more detail.

**zProb** a vector containing the probability of receiving the treatment inducement (i.e., instrument = 1) that have been estimated by semiparametrical methods.

Details

`larf` is the high-level interface to the work-horse function `larf.fit`. A set of standard methods (including `print`, `summary`, `coef`, `vcov`, `fitted`, `resid`, `predict`) can be used to extract the corresponding information from a `larf` object.

The function provides instrumental variable estimation of treatment effects when both the endogenous treatment and its instrument (i.e., the treatment inducement) are binary. The method (Abadie, 2003) involves two steps. First, pseudo-weights are constructed from the probability of receiving the treatment inducement. By default the function estimates the probability by a Probit regression. But it also allows users to employ the probability that has been estimated by semiparametric methods. Second, the pseudo-weights are used to estimate the local average response function of the outcome conditional on the treatment and covariates. The function provides both least squares and maximum likelihood estimates of the conditional treatment effects.

Value

**coefficients** Estimated coefficients.

**SE** Standard errors of the estimated coefficients.

**MargEff** Estimated marginal effects, available only for binary outcomes.

**MargStdErr** Standard errors of the estimated marginal effects, available only for binary outcomes.

**vcov** Variance covariance matrix of the estimated coefficients.

**fitted.values** Predicted outcomes based on the estimated model. They are probabilities when the outcome is binary.
Note

We derived part of the code from the Matlab code written by Professor Alberto Abadie, available at http://www.hks.harvard.edu/fs/aabadie/larf.html. We thank Onur Altindag and Behzad Kianian for helpful suggestions on improving the computation.

Author(s)

Weihua An, Departments of Sociology and Statistics, Indiana University Bloomington, <weihuaan@indiana.edu>. Xuefu Wang, Department of Statistics, Indiana University Bloomington, <wangxuef@umail.iu.edu>.

References


See Also

larf.fit, c401k

Examples

data(c401k)
attach(c401k)

## Not run:
# Continuous outcome. Treatment effects of participation in 401(k)
# on net family financial asset
est1 <- larf(nettfa ~ inc + age + agesq + marr + fsize, treatment = p401k, instrument = e401k, data = c401k)
summary(est1)

# Nonparametric estimates of the probability of receiving the treatment inducement
library(mgcv)
firstStep <- gam(e401k ~ s(inc) + s(age) + s(agesq) + marr + s(fsize), data=c401k, family=binomial(link = "probit"))
zProb <- firstStep$fitted
est2 <- larf(nettfa ~ inc + age + agesq + marr + fsize, treatment = p401k, instrument = e401k, data = c401k, zProb = zProb)
summary(est2)

# Binary outcome. Treatment effects of participation in 401(k)
# on participation in IRA
est3 <- larf(pira ~ inc + age + agesq + marr + fsize, treatment = p401k, instrument = e401k, data = c401k)
summary(est3)
larf.fit  

Fitting the Local Average Response Function

Description

It is the work-horse function for its high-level interface larf.

Usage

larf.fit(Y, X, D, Z, method, AME, optimizer, zProb)

Arguments

Y  a vector containing the outcome.
X  a matrix containing the covariates excluding the treatment.
D  a vector containing the binary treatment.
Z  a vector containing the binary instrument for the endogenous treatment.
method the estimation method to be used. The default is “LS”, standing for least squares. “ML”, standing for maximum likelihood, is an alternative.
AME whether average marginal effects (AME) should be reported. The default is FALSE, in which case marginal effects at the means (MEM) are reported.
optimizer the optimization algorithm for the ML method. It should be one of “Nelder-Mead”, “BFGS”, “CG”, “L-BFGS-B”, “SANN”, or “Brent”. See optim in R for more detail.
zProb a vector containing the probability of receiving the treatment inducement (i.e., instrument = 1) that have been estimated by semiparametrical methods.

Author(s)

Weihua An and Xuefu Wang, Departments of Sociology and Statistics, Indiana University Bloomington

See Also

larf, c401k
npse

Nonparametric Power Series Estimation

Description

Use the optimal order of power series of covariates to predict outcome. The optimal order of power series is determined by cross-validation.

Usage

npse(formula, order = 3, m = 10, seed = NULL)

Arguments

formula specification of the outcome model in the form like either z ~ x1 + x2 or z ~ X where X is the covariate matrix.
order the maximal order of power series to be used.
m the number of folds to be used in cross-validation.
seed random starting number used to replicate cross-validation.

Details

This function predicts the outcome based on the optimal order of covariates power series. The optimal order of the power series is determined by cross-validation. For example, it can be used to predict the probability of receiving treatment inducment based on covariates.

Value

fitted Predicted outcomes based on the estimated model. They are probabilities when the outcome is binary.
Lambda The optimal order of power series determined by cross-validation.
Data.opt The data including z and the optimal covariates power series.
CV.Res The residual sum of squares of the cross-validations.
seed The random seed.

Author(s)

Weihua An, Departments of Sociology and Statistics, Indiana University Bloomington, <weihuan@indiana.edu>. Xuefu Wang, Department of Statistics, Indiana University Bloomington, <wangxuef@umail.iu.edu>.

References

predict.larf

Predictions Based on the Estimated LARF

Description

Predict new outcomes based on the model fitted by larf.

Usage

## S3 method for class 'larf'
predict(object, newCov, newTreatment, ...)

Arguments

object      an object of class larf as fitted by larf.
newCov      A matrix containing the new covariates.
newTreatment A vector containing the new binary treatment.
...          currently not used.
Details

Predicted outcomes are based on the estimated coefficients and new covariates and/or new treatment. The predicted outcomes are probabilities when the outcome is binary.

Value

predicted.values

The function returns a vector of the predicted outcomes.

Author(s)

Weihua An and Xuefu Wang, Departments of Statistics and Sociology, Indiana University Bloomington

See Also

larf, larf.fit

Description

Methods to display brief results of a `larf` object.

Usage

```r
## S3 method for class 'larf'
print(x, digits = 4, ...)
```

Arguments

- `x` an object of class "larf" as fitted by `larf`.
- `digits` The number of significant digits to be printed in the reports of the results.
- `...` currently not used.

Author(s)

Weihua An and Xuefu Wang, Departments of Statistics and Sociology, Indiana University Bloomington

See Also

larf, larf.fit
summary.larf

Summary of the Estimated LARF

Description
Summary of an object in the larf class.

Usage
## S3 method for class 'larf'
summary(object, ...)

Arguments
- object: an object of class "larf" as fitted by larf.
- ...: currently not used.

Author(s)
Weihua An and Xuefu Wang, Departments of Statistics and Sociology, Indiana University Bloomington

See Also
larf, larf.fit

vcov.larf

Variance Covariance Matrix of the Parameters in the Estimated LARF

Description
Methods to display the variance covariance matrix of the model parameters estimated by larf.

Usage
## S3 method for class 'larf'
vcov(object, ...)

Arguments
- object: an object of class "larf" as fitted by larf.
- ...: currently not used.
Author(s)

Weihua An and Xuefu Wang, Departments of Statistics and Sociology, Indiana University Bloomington

See Also

larf, larf.fit
Index

c401k, 2, 6, 7
cvlm, 3

Generate.Powers, 4

larf, 2–4, 4, 5, 7, 9–12
larf.fit, 2, 5, 6, 7, 9–12

npse, 3, 4, 8

predict.larf, 9
print.larf, 10
print.summary.larf (summary.larf), 11

summary.larf, 11

vcov.larf, 11