Package ‘genius’

October 24, 2017

Title G-Estimation under No-Interaction with Unmeasured Selection

Date 2017-10-16

Version 1.0

Author BaoLuo Sun and Eric Tchetgen Tchetgen

Maintainer BaoLuo Sun <bluosun@gmail.com>

Description This package implements the MR GENIUS estimator.

Depends R (>= 3.4.1)

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Imports gmm (>= 1.6-1),
BB,
lmtest

Suggests knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

---

bptest *Breusch-Pagan Test*

Description

Performs the Breusch-Pagan test against heteroskedasticity. This function is exported from the "lmtest" package (Achim Zeileis & Torsten Hothorn, 2002). This function provides a way to test the heteroskedasticity identification condition (5) in Lemma 1 of Tchetgen Tchetgen et al. (2017).

Usage

```
bptest(formula, varformula = NULL, studentize = TRUE, data = list())
```

Arguments

`formula`	a symbolic description for the model to be tested (or a fitted `"lm"` object).
`varformula`	a formula describing only the potential explanatory variables for the variance (no dependent variable needed). By default the same explanatory variables are taken as in the main regression model.
`studentize`	logical. If set to `TRUE` Koenker's studentized version of the test statistic will be used.
`data`	an optional data frame containing the variables in the model. By default the variables are taken from the environment which `bptest` is called from.

Details

The Breusch-Pagan test fits a linear regression model to the residuals of a linear regression model (by default the same explanatory variables are taken as in the main regression model) and rejects if too much of the variance is explained by the additional explanatory variables. Under H0 the test statistic of the Breusch-Pagan test follows a chi-squared distribution with `parameter` (the number of regressors without the constant in the model) degrees of freedom.

Value

A list with class `"htest"` containing the following components:

`statistic`	the value of the test statistic.
`p.value`	the p-value of the test.
`parameter`	degrees of freedom.
`method`	a character string indicating what type of test was performed.
`data.name`	a character string giving the name(s) of the data.

References

T.S. Breusch & A.R. Pagan (1979), A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica* **47**, 1287–1294

R. Koenker (1981), A Note on Studentizing a Test for Heteroscedasticity. *Journal of Econometrics* **17**, 107–112.

W. Kr?mer & H. Sonnberger (1986), *The Linear Regression Model under Test*. Heidelberg: Physica

Achim Zeileis & Torsten Hothorn (2002), Diagnostic Checking in Regression Relationships. *R News* **2(3)**, 7-10. `https://CRAN.R-project.org/doc/Rnews/`

Tchetgen Tchetgen, E., Sun, B. and Walter, S. (2017). The GENIUS Approach to Robust Mendelian Randomization Inference. arXiv e-prints.

Examples

```
# the following packages are needed to simulate data
library("msm")
library("MASS")

nIV=10; N=500; beta=1;
phi=rep(-0.5,nIV); gamma=rep(-2,nIV); alpha=rep(-0.5,nIV);
lambda0=1; lambda1=rep(0.5,nIV);
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))

G  = (Gn>0)*1;
U = as.vector(phi%*%t(G))+rnorm(N);
A = as.vector(gamma%*%t(G)) +U + rnorm(N,mean=0,sd=abs(lambda0+as.vector(lambda1%*%t(G))));
Y = as.vector(alpha%*%t(G)) + beta*A + U + rnorm(N);

bptest(A~G);
```

---

genius_addY *MR GENIUS under additive outcome model*

Description

Implements MR GENIUS under an additive outcome model.

Usage

```
genius_addY(Y, A, G, formula = A ~ G, alpha = 0.05, lower = -10,
  upper = 10)
```

Arguments

Y	A numeric vector of outcomes.
A	A numeric vector of exposures (binary values should be coded in 0/1).
G	A numeric matrix of instruments; each column stores values for one instrument (a numeric vector if only a single instrument is available).
formula	An object of class "formula" describing the linear predictor of the model for $E(A\|G)$ (default A G, main effects of all available instruments).
alpha	Significance level for confidence interval (default value=0.05).
lower	The lower end point of the causal effect interval to be searched (default value=-10).
upper	The upper end point of the causal effect interval to be searched (default value=10).

Details

This function implements the estimators given in equations (6) and (12) of Tchetgen Tchetgen et al (2017) for single and multiple instruments, respectively. The term $E(A|G)$ is modelled under the logit and identity links for binary and continuous exposure respectively, with a default linear predictor consisting of the main effects of all available instruments.

Value

A "genius" object containing the following items:

beta.est	The point estimate of the causal effect (on the additive scale) of the exposure on the outcome.
beta.var	The corresponding estimated variance.
ci	The corresponding Wald-type confidence interval at specified significance level.
pval	The p-value for two-sided Wald test of null causal effect (on the additive scale) of the exposure on the outcome.

References

Tchetgen Tchetgen, E., Sun, B. and Walter, S. (2017). The GENIUS Approach to Robust Mendelian Randomization Inference. arXiv e-prints.

Examples

```
# the following packages are needed to simulate data
library("msm")
library("MASS")
expit <- function(x) {
   exp(x)/(1+exp(x))
}

### example with binary exposure, all instruments invalid ###
# true causal effect, beta = 1.0
# Number of instruments, nIV = 10
# Y: vector of outcomes
# A: vector of exposures
# G: matrix of instruments, one column per instrument

nIV=10; N=5000; beta=1;
phi=rep(-0.02,nIV); gamma=rep(-0.15,nIV); alpha=rep(-0.5,nIV);
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))

G  = (Gn>0)*1;
U= as.vector(phi%*%t(G))+ rtnorm(n=N,mean=0.35,lower=0.2,upper=0.5);
A = rbinom(N,1,expit(as.vector(gamma%*%t(G)))+U-0.35-as.vector(phi%*%t(G)));
Y = as.vector(alpha%*%t(G)) + beta*A + U + rnorm(N);

genius_addY(Y,A,G);

### specify a more richly parameterized linear predictor for the model
### of E[A|G] containing all main effects and pairwise interactions of
### instruments

colnames(G)=paste("g",1:10,sep="")

genius_addY(Y,A,G,A~(g1+g2+g3+g4+g5+g6+g7+g8+g9+g10)^2);

### example with continous exposure, all instruments invalid ###

nIV=10; N=500; beta=1;
phi=rep(-0.5,nIV); gamma=rep(-2,nIV); alpha=rep(-0.5,nIV);
lambda0=1; lambda1=rep(0.5,nIV);
```

```
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))

G  = (Gn>0)*1;
U = as.vector(phi%*%t(G))+rnorm(N);
A = as.vector(gamma%*%t(G)) +U + rnorm(N,mean=0,sd=abs(lambda0+as.vector(lambda1%*%t(G))));
Y = as.vector(alpha%*%t(G)) + beta*A + U + rnorm(N);

genius_addY(Y,A,G);
```

`genius_mulA` *MR GENIUS under multiplicative exposure model*

Description

Implements MR GENIUS under a multiplicative exposure model.

Usage

```
genius_mulA(Y, A, G, alpha = 0.05, lower = -10, upper = 10)
```

Arguments

`Y`	A numeric vector of outcomes.
`A`	A numeric vector of exposures (binary values should be coded in 0/1).
`G`	A numeric matrix of instruments; each column stores values for one instrument (a numeric vector if only a single instrument is available).
`alpha`	Significance level for confidence interval (default value=0.05).
`lower`	The lower end point of the causal effect interval to be searched (default value=-10).
`upper`	The upper end point of the causal effect interval to be searched (default value=10).

Details

This function implements the estimator given in Lemma 3 of Tchetgen Tchetgen et al (2017), under a multiplicative exposure model. By default, the log ratio term in equation (9) is modelled as a linear combination of the main effects of all available instruments.

Value

A "genius" object containing the following items:

`beta.est`	The point estimate of the causal effect (on the additive scale) of the exposure on the outcome.
`beta.var`	The corresponding estimated variance.
`ci`	The corresponding Wald-type confidence interval at specified significance level.
`pval`	The p-value for two-sided Wald test of null causal effect (on the additive scale) of the exposure on the outcome.

References

Tchetgen Tchetgen, E., Sun, B. and Walter, S. (2017). The GENIUS Approach to Robust Mendelian Randomization Inference. arXiv e-prints.

Examples

```
#the following package is needed to simulate data
library("MASS")
nIV=10; N=2000; beta=0.5;
gamma=rep(0.5,nIV); alpha=rep(0.5,nIV);phi=rep(0.05,nIV);
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))
G  = (Gn>0)*1;
U = as.vector(phi%*%t(G))+rnorm(N);
#exposure generated from negative binomial distribution
A = rnbinom(N,size=10,mu = exp(as.vector(gamma%*%t(G)) +0.1*U))
Y = as.vector(alpha%*%t(G)) + beta*A + U + rnorm(N);

genius_mulA(Y,A,G);
```

`genius_mulY` *MR GENIUS under multiplicative outcome model*

Description

Implements MR GENIUS under a multiplicative outcome model.

Usage

```
genius_mulY(Y, A, G, formula = A ~ G, alpha = 0.05, lower = -10,
  upper = 10)
```

Arguments

`Y`	A numeric vector of outcomes.
`A`	A numeric vector of exposures (binary values should be coded in 0/1).
`G`	A numeric matrix of instruments; each column stores values for one instrument (a numeric vector if only a single instrument is available).
`formula`	An object of class "formula" describing the linear predictor of the model for $E(A\|G)$ (default A G, main effects of all available instruments).
`alpha`	Significance level for confidence interval (default value=0.05).
`lower`	The lower end point of the causal effect interval to be searched (default value=-10).
`upper`	The upper end point of the causal effect interval to be searched (default value=10).

Details

This function implements MR GENIUS as the solution to the empirical version of equation (14) in Tchetgen Tchetgen et al (2017). The term $E(A|G)$ is modelled under the logit and identity links for binary and continuous exposure respectively, with a default linear predictor consisting of the main effects of all available instruments.

Value

A "genius" object containing the following items:

beta.est	The point estimate of the causal effect (on the multiplicative scale) of the exposure on the outcome.
beta.var	The corresponding estimated variance.
ci	The corresponding Wald-type confidence interval at specified significance level.
pval	The p-value for two-sided Wald test of null causal effect (on the multiplicative scale) of the exposure on the outcome.

References

Tchetgen Tchetgen, E., Sun, B. and Walter, S. (2017). The GENIUS Approach to Robust Mendelian Randomization Inference. arXiv e-prints.

Examples

```
#the following packages are needed to simulate data
library("msm")
library("MASS")

### examples under multiplicative outcome model, all instruments invalid ###
# true causal effect, beta = 1.5
# Number of instruments, nIV = 10
# Y: vector of outcomes
# A: vector of exposures
# G: matrix of instruments, one column per instrument

### binary exposure
nIV=10; N=2000; beta=1.5;
phi=rep(-0.02,nIV); gamma=rep(-0.15,nIV); alpha=rep(-0.5,nIV);
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))
G  = (Gn>0)*1;
U= as.vector(phi%*%t(G))+ rtnorm(n=N,mean=0.35,lower=0.2,upper=0.5);
A = rbinom(N,1,expit(as.vector(gamma%*%t(G)))+U-0.35-as.vector(phi%*%t(G)));
Y = exp(beta*A)*(as.vector(alpha%*%t(G)) + U) + rnorm(N);

genius_mulY(Y,A,G);

### specify a more richly parameterized linear predictor for the model of E[A|G]
### containing all main effects and pairwise interactions of instruments

colnames(G)=paste("g",1:10,sep="")

genius_mulY(Y,A,G,A~(g1+g2+g3+g4+g5+g6+g7+g8+g9+g10)^2);

### continuous exposure
nIV=10; N=2000; beta=0.25;
phi=rep(0.2,nIV); gamma=rep(0.5,nIV); alpha=rep(0.5,nIV);
lambda0=0.5; lambda1=rep(0.5,nIV);
Gn = mvrnorm(N,rep(0,nIV),diag(rep(1,nIV)))
G  = (Gn>0)*1;
U = as.vector(phi%*%t(G))+rnorm(N);
A = as.vector(gamma%*%t(G)) +U + rnorm(N,mean=0,sd=abs(lambda0+as.vector(lambda1%*%t(G))));
Y = exp(beta*A)*(as.vector(alpha%*%t(G)) + U) + rnorm(N);
```

```
genius_mulY(Y,A,G);
```

Index