Package ‘treatSens’

March 9, 2018

Version 2.1.3
Date 2018-03-08
Title Sensitivity Analysis for Causal Inference
Depends R (>= 3.1-0), Rcpp (>= 0.11.0)
Imports methods, stats, graphics, grDevices, mvtnorm, lme4, nlme, 
dbarts (>= 0.9-1),
Suggests doParallel, foreach, testthat (>= 0.9-0)
LinkingTo dbarts (>= 0.9-1), BH (>= 1.60.0), Rcpp (>= 0.12.0), 
RcppEigen
Description Utilities to investigate sensitivity to unmeasured confounding in 
parametric models with either binary or continuous treatment.
NeedsCompilation yes
Biarch yes
License GPL (>= 2)
URL https://github.com/vdorie/treatSens
BugReports https://github.com/vdorie/treatSens/issues
RoxygenNote 5.0.1
Author Nicole Bohme Carnegie [aut, cre], 
Masataka Harada [aut], 
Vincent Dorie [aut], 
Jennifer Hill [ctb]
Maintainer Nicole Bohme Carnegie <nicole.carnegie@montana.edu>
Repository CRAN
Date/Publication 2018-03-08 23:32:57 UTC

R topics documented:

  treatSens-package .................................................. 2
  sensPlot .............................................................. 2
Description

Utilities to investigate sensitivity to unmeasured confounding in causal models of several types.

Details

Run dual-parameter sensitivity analysis for continuous response modeled with linear model using treatSens. Results are available in either tabular form using summary.sensitivity or graphical form using sensPlot.

Author(s)

Authors: Nicole Bohme Carnegie, Masataka Harada, Jennifer Hill
Maintainer: Nicole Bohme Carnegie <carnegin@uwm.edu>

References

Carnegie NB, Hill JH and Harada M. Assessing sensitivity to unmeasured confounding using simulated potential confounders (under review)
Description

sensPlot creates a contour plot that summarizes the results of the sensitivity analysis obtained from GLM.sens. The plot region is defined by the coefficient on $U$ in the outcome model (vertical axis) and that in the treatment model (horizontal axis). Each contour represents the combination of sensitivity parameters for $U$ that lead to the same treatment effect estimate.

Usage

sensPlot(x, contour.levels = NULL, col.zero = "red", lty.zero = 1,
        col.insig = "blue", lty.insig = 1, data.line = TRUE, X.pch = NULL,
        signif.level = 0.05, labcex = 0.75, limit.Xplot = FALSE, txtlab = FALSE,
        which.txtlab = NULL, ...)

Arguments

x an object of class sensitivity.
contour.levels numeric vector of levels at which to draw contour lines. The default is NULL.
col.zero color of the contour representing the combination of zetas that lead to the treatment effect estimate of 0. The default is "red".
lty.zero line type of the contour representing the combination of zetas that lead to the treatment effect estimate of 0. The default is 1 (solid line).
col.insig color of the contour representing the combination of zetas that makes the treatment effect estimate statistically insignificant at a given level. The default is "blue".
lty.insig line type of the contour representing the combination of zetas that makes the treatment effect estimate statistically insignificant at a given level. The default is 1 (solid line)
data.line logical. If TRUE a grey contour corresponding to the treatment effect estimate obtained with sensitivity parameters set equal to the coefficients (across all observed confounders) that are farthest from the origin (0,0). The default is TRUE.
X.pch vector of length 2 giving plotting symbols confounders with a positive association with the outcome and a negative association with the outcome, respectively. The default is c(3,6).
signif.level this option specifies the statistical significance level at which the significance contour is drawn. The default is 0.05.
labcex letter size of the treatment effect estimates on the contours. The default is 0.75.
limit.Xplot logical. If TRUE this option limits the plot region to the minimum and the maximum of the sensitivity parameters and the covariates are plotted on the left or the right end of the figure. If FALSE the contour is extended to the coefficient of the strongest confounders.
txtlab logical. Label plotted covariates with variable name.
which.txtlab numeric vector of covariates to include, i.e. c(1:3) shows labels for first three covariates.
... Other arguments to be passed to all calls to contour
Details

Plots contours of treatment effect estimates under varying combinations of sensitivity parameters, with parameters associated with observed covariates for benchmarking. Options allow highlighting of sensitivity parameters where significance is lost/gained, where treatment effect is reduced to zero, or with treatment effect consistent with most extreme covariate.

Author(s)

Nicole Bohme Carnegie, Masataka Harada, Jennifer Hill

References

Carnegie NB, Hill JH and Harada M. Assessing sensitivity to unmeasured confounding using simulated potential confounders (under review)

See Also

treatSens, plot.default, plot.formula.

Examples

#See the manual for treatSens.

summary.sensitivity.default

Summarizing sensitivity analysis results

Description

Summary method for objects of class "sensitivity".

Usage

## S3 method for class 'sensitivity'
summary(object, ...)
## Default S3 method:
summary.sensitivity(object, digits = 3, signif.level = 0.05, ...)

Arguments

object
  An object of class "sensitivity".
digits
  The number of digits used for rounding.
signif.level
  Function returns a table of combinations of sensitivity parameters where this significance level is lost.
...  Included for consistency with summary generic.
Details
Prints a summary of main features of sensitivity analysis distribution.

Author(s)
Nicole Bohme Carnegie, Masataka Harada, and Jennifer Hill

References
Carnegie NB, Hill JH and Harada M. Assessing sensitivity to unmeasured confounding using simulated potential confounders (under review)

See Also
treatsens

Examples
#See treatsens documentation

treatsens  Performing Dual-Parameter Sensitivity Analysis

Description

treatsens performs dual-parameter sensitivity analysis for unobserved confounding for both continuous and binary treatment variables.

Usage
treatsens(formula, response.covariates = NULL, sensParam = "coef", resp.family = gaussian, trt.family = gaussian, theta = 0.5, grid.dim = c(8,4), standardize = TRUE, nsim = 20, zero.loc = 1/3, verbose = FALSE, buffer = 0.1, weights = NULL, data = NULL, seed = 1234, iter.j = 10, offset = TRUE, core = NULL, spy.range = NULL, spz.range = NULL, trim.wt = 10)

Arguments
formula an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted.
response.covariates an object of class formula (or one that can be coerced to that class): a one-sided formula giving additional predictors to be included in the response model but not the treatment model.
sensParam a string giving the choice of the form of the sensitivity parameters. Supports "coef" for model coefficient and "cor" for partial correlations (with continuous treatment only)
resp.family an assumed family for GLM of the response model. Currently only supports gaussian.

trt.family an assumed family for GLM of the treatment model. The default is gaussian. binomial(link="probit") can be also specified.

theta this option specifies the marginal probability that a binary unobserved confounder takes one (i.e. Pr(U=1)). The default is 0.5.

grid.dim the final dimensions of output grid. GLM.sens draws nsim of unobserved confounders and corresponding treatment effect estimates for each grid cell. The first argument specifies the horizontal (treatment) dimension, and the second argument specifies the vertical (response) dimension. The default is c(9, 5), that is, confounding effects are evaluated at 40 grid points. Note that the dimensions given will be increased by one to an odd number where necessary when zero.loc = "full" or sensitivity parameter ranges are given in order to force inclusion of 0 in the range of sensitivity parameters.

standardize logical. If TRUE all variables except binary ones are standardized to have mean 0 and standard deviation 1. The default is TRUE.

nsim number of simulated Us to generate (and hence to average over) per cell in grid. The default is 20.

zero.loc A parameter used to automatically determine the range of each axis by specifying at what point along the y=x line the "treatment effect=0" contour will cross it. The parameter specifies the fraction of the length of the y=x line that this intersection occurs relative to the entire length of the line segment. The default is 1/3. No need to specify when spy.range and spz.range are specified.

verbose logical. If TRUE, GLM.sens will show detailed messages.

buffer restriction to range of coefficients on U to ensure stability around the edges. The default is 0.1. This option is rarely used.

weights the user can supply a vector of weights or specify one of three causal estimands – "ATE", "ATT", or "ATC" – in which case the algorithm will automatically create appropriate weights. The default is NULL. See Details for information on calculation of weights.

data an optional data frame. If not found in data, the variables are taken from environment(formula), typically the environment from which glm is called. The default is NULL.

seed random seed. The default is 1234.

iter.j number of iterations used to draw simulated U from its conditional distribution in the iterated algorithm required for the binary treatment variable model. The default is 10. This option is used only when trt.family=binomial(link="probit").

offset Models are fit by setting zeta.z*U or zeta.y*U explicitly as an offset (rather than simply including generated U in the model). Only consistent with sensParam = "coef". The default is TRUE.

core number of CPU cores used for parallel processing. The default is NULL, which implies single-threading.

spy.range custom range for the sensitivity parameter associated with Y (the coefficient on U in the outcome model or partial correlation of U with Y given X), e.g. c(0, 2). When this option is specified, zeta.z.range must be also specified and zero.loc will be overridden. The default is NULL.
spz.range  custom range for the sensitivity parameter associated with Z (the coefficient on U in the treatment model or partial correlation of U with Z given X), e.g. c(-2,2). When this option is specified, zetay.range must be also specified and zero.loc will be overridden. The default is NULL.

trim wt the maximum size of weight as a percentage of the sample size of the inferential group for the causal estimand. This option is used only when weights option is specified as "ATE", "ATT" or "ATC". The default is 10. For example, trim.wt=10 in a dataset with 300 treated observations when the estimand has been set to "ATT" gives a maximum weight of 30.

Details
This function performs a dual-parameter sensitivity analysis for unobserved confounding by drawing simulated potential confounders U from the conditional distribution of U given observed response, treatment and covariates.

Weights are calculated for any of three different estimands. The weights for the ATE are calculated as $1/\hat{e}(x)$ for the treatment group observations and $1/(1-\hat{e}(x))$ for the control group observations, where $\hat{e}(x)$ denotes the estimated modification score (that is, the estimate of the probability of being treated conditional on observed covariates, x). The weights for the ATT are 1 for the treatment group observations and $\hat{e}(x)/(1-\hat{e}(x))$ for the control group observations. The weights for the ATC are $(1-\hat{e}(x))/\hat{e}(x)$ for the treatment group observations and 1 for the control group observations. These weights are discussed in Carnegie et al (2014) and similar versions using the propensity score are discussed in a variety of places including Gelman and Hill, 2007.

Value
treatSens returns an object of class "sensitivity".
An object of class "sensitivity" is a list containing the following components:

model.type estimation methods used (i.e. "GLM"). Currently only GLM methods are available, but in future other types of models may be added.
sensParam type of sensitivity parameter used (model coefficient or partial correlation).
tau individual estimated treatment effects corresponding to each grid point (combination of sensitivity parameters) and each drawing of simulated U.
se.tau individual standard error estimates of treatment effects for each grid point and each drawing of simulated U.
zeta.z individual estimated coefficients on U in the treatment model corresponding to each grid point and each drawing of simulated U.
zeta.y individual estimated coefficients on U in the outcome model corresponding to each grid point and each drawing of simulated U.
se.zz individual standard error estimates of the coefficient on U in the treatment model corresponding to each grid point and each drawing of simulated U.
se.zy individual standard error estimates of the coefficient on U in the outcome model corresponding to each grid point and each drawing of simulated U.
Y outcome variable
Z treatment variable
X covariates

sig2.trt individual residual variances for the treatment model corresponding to each grid point and each drawing of simulated U.

sig2.resp individual residual variances for the outcome model corresponding to each grid point and each drawing of simulated U.

tau0 averaged estimated treatment effect assuming no unobserved confounding (i.e. naive model).

se.tau0 averaged standard error estimate of the treatment effect without unobserved confounding (i.e. naive model)

Xcoef the coefficients on the covariates in the treatment model (1st column) and the outcome model (2nd column) without unobserved confounding (i.e. naive model).

Xcoef.plot coefficients as for Xcoef, but after transforming covariates to have positive coefficients in the response model (X = -X, used for plotting purposes).

Note
When weights are specified, treatSens returns clustered robust standard errors.
Matrices and data frames should include only complete cases.

Author(s)
Nicole Bohme Carnegie, Masataka Harada, and Jennifer Hill

References

See Also
sensPlot summary.sensitivity glm

Examples
N = 250 #number of observations

zetay = .5 #coefficient on U in the outcome model
zetaz = .5 #coefficient on U in the treatment model
betaz = c(.75,-.5,.25) #coefficients of X in the treatment model
betay = c(.5,1,-1.5) #coefficients of X in the outcome model
tau = .25 #treatment effect

X = matrix(rnorm(3*N),N,3) #covariates

# Example with continuous treatment
set.seed(836)
U = rnorm(N, 0, 1) #unmeasured confounder
Z = rnorm(N,X %*% betaz + zetaz*U,1) #treatment variable
treatSens.BART

Y = rnorm(N,X %% betay + zetay*U + tau*Z,2) #outcome variable

# sensitivity analysis
# small grid.dim and nsim just for example purposes
out.cont <- treatSens(Y~Z+X, grid.dim = c(3,2), nsim = 5,
standardize = FALSE, verbose = TRUE, zero.loc = "full")

sensPlot(out.cont) # draw contour plot
# You can see when zetay=zetaz=.5, tau is about .25.

# Example with binary treatment
set.seed(725)
U = rbinom(N,1,.5) #unmeasured confounder
ps = pnorm(X*betaz + zetaz*(U-.5)) #propensity score
Z = rbinom(N,1,ps) #treatment variable
epsilon = rnorm(N,0,2) #error term
Y0 = X*betay + zetay*(U-.5) + epsilon #potential outcome(Z=0)
Y1 = X*betay + zetay*(U-.5) + tau + epsilon #potential outcome(Z=1)
Y = Y0*(1-Z) + Y1*Z #realization of potential outcome

# sensitivity analysis
out.bin <- treatSens(Y~Z+X, trt.model = binomial(link="probit"), nsim = 3,
spy.range = c(0,4), spz.range = c(-2,2),grid.dim = c(5,3),
standardize = FALSE, verbose = TRUE)

sensPlot(out.bin) # draw contour plot

treatSens.BART

Dual-Parameter Sensitivity Analysis Using BART

Description

treatSens.BART performs dual-parameter sensitivity analysis for unobserved confounding for binary treatment variables using Bayesian Additive Regression Trees (BART).

Usage

treatSens.BART(formula, 
    trt.model = probitEM(),
    theta = 0.5,
    grid.dim = c(8, 4),
    standardize = TRUE,
    zero.loc = 1 / 3,
    verbose = FALSE,
    buffer = 0.0,
    est.type = "ATE",
    data = NULL,
seed = 1234,  
sim = 200,  
thin = 10,  
nburn = 200,  
threads = NULL,  
spy.range = NULL,  
spz.range = NULL,  
trim.wt = 10,  
benchmarking = "2SD")

**Arguments**

**formula**

an object of class `formula` (or one that can be coerced to that class): a symbolic description of the model to be fitted.

**trt.model**

an assumed fitting technique for the treatment model. The default is `probitEM()`. Alternatives include `probit()` and `bart()`. See the 'Details' section for arguments to these functions.

**theta**

this option specifies the marginal probability that a binary unobserved confounder takes one, i.e. $P(U = 1)$. The default is 0.5.

**grid.dim**

the final dimensions of output grid. `GLM_sens` draws `nsim` of unobserved confounders and corresponding treatment effect estimates for each grid cell. The first argument specifies the horizontal (treatment) dimension, and the second argument specifies the vertical (response) dimension. The default is `c(8, 4)`, that is, confounding effects are evaluated at 32 grid points. Note that the dimensions given will be increased by one to an odd number where necessary when `zero.loc = "full"` or sensitivity parameter ranges are given in order to force inclusion of 0 in the range of sensitivity parameters.

**standardize**

logical. If TRUE all variables except binary ones are standardized to have mean 0 and standard deviation 1.

**zero.loc**

A parameter used to automatically determine the range of each axis by specifying at what point along the y=x line the "treatment effect=0" contour will cross it. The parameter specifies the fraction of the length of the y=x line that this intersection occurs relative to the entire length of the line segment. The default is $1/3$. No need to specify when `spy.range` and `spz.range` are specified.

**verbose**

logical. If TRUE, detailed messages are reported as the algorithm runs.

**buffer**

restriction to range of coefficients on U to ensure stability around the edges. This option is rarely used.

**est.type**

one of three causal estimads – "ATE", "ATT", or "ATC". See Details for information on calculation of weights.

**data**

an optional data frame. If not found in data, the variables are taken from `environment(formula)`, typically the environment from which `bart` is called.

**seed**

random seed, passed to `set.seed`.

**nsim**

number of simulated $U$'s to generate (and hence to average over) per cell in grid.

**nthin**

number of BART tree draws to make per every draw of $U$.

**nburn**

number of draws of $U$ to make and throw away at the start of a run.
nthreads  number of CPU cores used for parallel processing. If NULL, defaults to a guess of the number of "logical" processing units available. This guess returns the correct number on the most commonly used systems.

spy.range  custom range for the sensitivity parameter associated with Y (the coefficient on U in the outcome model or partial correlation of U with Y given X), e.g. c(0, 2). When this option is specified, zeta. range must be also specified and zero.loc will be overridden.

spz.range  custom range for the sensitivity parameter associated with Z (the coefficient on U in the treatment model or partial correlation of U with Z given X), e.g. c(-2, 2). When this option is specified, zeta. range must be also specified and zero.loc will be overridden.

trim.wt  the maximum size of weight as a percentage of the sample size of the inferential group for the causal estimand. For example, trim.wt = 10 in a dataset with 300 treated observations when the estimand has been set to "ATT" gives a maximum weight of 30.

benchmarking  either "2SD" or "1SD". Used to determine the "marginal effects" corresponding to predictors for the plot, where the number of SD units given is the full spread, i.e. distance from below the mean to above.

Details

This function performs a dual-parameter sensitivity analysis for unobserved confounding by drawing simulated potential confounders U from the conditional distribution of U given observed response, treatment and covariates. Bayesian Additive Regression Trees (bart) are used to fit the response surface and potentially the treatment as well. See treatSens for details of the sensitivity analysis algorithm.

Unique options to treatSens.BART are given by the argument trt.model. It is specified as an delayed-evaluation function call of the form:

• probitEM(maxBackstepIterations = 30L) - uses glm to perform stochastic expectation maximization; backstep iterations are an internal parameter to glm.fit
• probit(family = "cauchy", ...) - family can be "cauchy", "t", "normal", or "flat". Additional options are, scale = 4, c(df = 3, scale = 4), scale = 4, and none, respectively.
• bart(k = 2, ntree = 50, keepEvery = 10) - uses binary bart with the given parameters.

Value

treatSens.BART returns an object of class "sensitivity". See treatSens for details.

Author(s)

Vincent Dorie, Nicole Bohme Carnegie, Masataka Harada, and Jennifer Hill

References

See Also
treatSens

Examples

```r
N <- 250  # number of observations
zetay <- 0.5  # coefficient on U in the outcome model
zetaz <- 0.5  # coefficient on U in the treatment model
betaz <- c(0.75, -0.5, 0.25)  # coefficients of X in the treatment model
betay <- c(0.5, 1.0, -1.5)  # coefficients of X in the outcome model
tau <- 0.25  # treatment effect
set.seed(725)
X <- matrix(rnorm(3 * N), N)  # covariates
U <- rbinom(N, 1, 0.5)  # unmeasured confounder
ps <- pnorm(X %*% betaz + zetaz * (U - 0.5))  # propensity score
Z <- rbinom(N, 1, ps)  # treatment variable
epsilon <- rnorm(N, 0.0, 2.0)  # error term
Y0 <- X %*% betay + zetay * (U - 0.5) + epsilon  # potential outcome(Z=0)
Y1 <- X %*% betay + zetay * (U - 0.5) + tau + epsilon  # potential outcome(Z=1)
Y <- Y0 * (1 - Z) + Y1 * Z  # realization of potential outcome

# sensitivity analysis
out.bin <- treatSens.BART(Y ~ Z + X, nsim = 3, nthin = 2, nburn = 0,
                          spy.range = c(0, 4), spz.range = c(-2, 2),
                          grid.dim = c(5, 3),
                          standardize = FALSE, verbose = TRUE)

sensPlot(out.bin)  # draw contour plot
```
Index

*Topic package
  treatSens-package, 2

bart, 11
formula, 5, 10
glm, 8
plot.default, 4
plot.formula, 4
sensPlot, 2, 8
set.seed, 10
summary.sensitivity, 8
summary.sensitivity
  (summary.sensitivity.default), 4
summary.sensitivity.default, 4
treatSens, 4, 5, 5, 11, 12
treatSens-package, 2
treatSens.BART, 9