Package 'rdpower'

August 22, 2018

Type Package

Title Power Calculations for RD Designs

Version 0.3

Author Matias D. Cattaneo, Rocio Titiunik, Gonzalo Vazquez-Bare

Maintainer Gonzalo Vazquez-Bare <gvazquez@umich.edu>

Description

The regression discontinuity (RD) design is a popular quasi-experimental design for causal inference and policy evaluation. The 'rdpower' package provides tools to perform power and sample size calculations in RD designs: rdpower() calculates the power of an RD design and rdsampsi() calculates the required sample size to achieve a desired power. See Cattaneo, Titiunik and Vazquez-Bare (2018) <https://sites.google.com/site/rdpackages/rdpower/Cattaneo-Titiunik-VazquezBare_2018_Stata.pdf> for further methodological details.

Imports rdrobust

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2018-08-21 22:50:02 UTC

R topics documented:

rdpower-package *rdpower: Power and Sample Size Calculations for RD Designs*

Description

The regression discontinuity (RD) design is a popular quasi-experimental design for causal inference and policy evaluation. The 'rdpower' package provides tools to perform power and sample size calculations in RD designs: rdpower() calculates the power of an RD design and rdsampsi() calculates the required sample size to achieve a desired power. This package relies on the rdrobust package. See Calonico, Cattaneo and Titiunik (2014, 2015) and Calonico, Cattaneo, Farrell and Titiunik (2017). For more details, and related Stata and R packages useful for analysis of RD designs, visit https://sites.google.com/site/rdpackages.

Author(s)

Matias Cattaneo, University of Michigan. <cattaneo@umich.edu>

Rocio Titiunik, University of Michigan. <titiunik@umich.edu>

Gonzalo Vazquez-Bare, University of Michigan. <gvazquez@umich.edu>

References

M.D. Cattaneo, R. Titiunik and G. Vazquez-Bare. (2018). Power Calculations for Regression Discontinuity Designs. *Working paper, University of Michigan*.

M.D. Cattaneo, S. Calonico and R. Titiunik. (2014). Robust Data-Driven Inference in the Regression-Discontinuity Design. *Stata Journal 14(4): 909-946*.

M.D. Cattaneo, S. Calonico and R. Titiunik. (2015).rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs. *R Journal 7(1): 38-51*.

M.D. Cattaneo, S. Calonico, M. Farrell and R. Titiunik. (2017).rdrobust: Software for Regression Discontinuity Designs. *Stata Journal 17(2): 372-404*.

rdpower *Power Calculations for RD Designs*

Description

rdpower() performs power calculations for RD designs.

Usage

```
rdpower(data = NULL, cutoff = 0, tau = NULL, alpha = 0.05,
  nsamples = NULL, sampsi = NULL, samph = NULL, all = FALSE,
  bias = NULL, variance = NULL, plot = FALSE, graph.range = NULL,
  covs = NULL, deriv = 0, p = 1, q = NULL, h = NULL, b = NULL,
  rho = NULL, kernel = "triangular", bwselect = "mserd",
  vce = "nn", cluster = NULL, scalepar = 1, scaleregul = 1,
  fuzzy = NULL, level = 90)
```

Arguments

data	a matrix (Y,R) containing the outcome variable and the running variable (as column vectors).
cutoff	the RD cutoff (default is 0).
tau	specifies the treatment effect under the alternative at which the power function is evaluated. The default is half the standard deviation of the outcome for the untreated group.
alpha	specifies the significance level for the power function. Default is 0.05.
nsamples	sets the total sample size to the left, sample size to the left inside the bandwidth, total sample size to the right and sample size to the right of the cutoff inside the bandwidth to calculate the variance when the running variable is not specified. When not specified, the values are calculated using the running variable.
sampsi	sets the sample size at each side of the cutoff for power calculation. The first number is the sample size to the left of the cutoff and the second number is the sample size to the right. Default values are the sample sizes inside the chosen bandwidth.
samph	sets the bandwidths at each side of the cutoff for power calculation. The first number is the bandwidth to the left of the cutoff and the second number is the bandwidth to the right. Default values are the bandwidths used by rdrobust.
all	displays the power using the conventional variance estimator, in addition to the robust bias corrected one.
bias	set bias to the left and right of the cutoff. If not specified, the biases are estimated using rdrobust.
variance	set variance to the left and right of the cutoff. If not specified, the variances are estimated using rdrobust.
plot	plots the power function using the conventional and robust bias corrected standard errors from rdrobust.
graph.range	range of the plot.
covs	option for rdrobust(): specifies additional covariates to be used for estimation and inference.
deriv	option for rdrobust(): specifies the order of the derivative of the regression functions to be estimated.
p	option for rdrobust(): specifies the order of the local-polynomial used to construct the point-estimator.
q	option for rdrobust(): specifies the order of the local-polynomial used to construct the bias-correction.
h	option for rdrobust(): specifies the values of the main bandwidth to be used on the left and on the right of the cutoff, respectively.
b	option for rdrobust(): specifies the values of the bias bandwidth b to be used on the left and on the right of the cutoff, respectively.
rho	option for rdrobust(): specifies the value of rho so that the bias bandwidth b equals b=h/rho.

kernel	option for rdrobust(): kernel function used to construct the local-polynomial estimators.
bwselect	option for rdrobust(): specifies the bandwidth selection procedure to be used.
vce	option for rdrobust(): specifies the procedure used to compute the variance-covariance matrix estimator.
cluster	option for rdrobust(): indicates the cluster ID variable used for the cluster-robust variance estimation with degrees-of-freedom weights.
scalepar	option for rdrobust(): specifies scaling factor for RD parameter of interest.
scaleregul	option for rdrobust(): specifies scaling factor for the regularization terms of bandwidth selectors.
fuzzy	option for rdrobust(): specifies the treatment status variable used to implement fuzzy RD estimation.
level	option for rdrobust(): sets the confidence level for confidence intervals.

Value

power.rbc	power against tau using robust bias corrected standard error
se.rbc	robust bias corrected standard error
sampsi.r	number of observations inside the window to the right of the cutoff
sampsi.l	number of observations inside the window to the left of the cutoff
samph.r	bandwidth to the right of the cutoff
samph.l	bandwidth to the left of the cutoff
alpha	significance level used in power function
tau	treatment effect under alternative hypothesis
bias.r	bias to the right of the cutoff
bias.l	bias to the left of the cutoff
Vr.rb	Robust bias corrected variance to the right of the cutoff
Vl.rb	Robust bias corrected variance to the left of the cutoff
N.r	Total sample size to the right of the cutoff
N.l	Total sample size to the left of the cutoff
power.conv	power against tau using conventional inference
se.conv	conventional standard error

Author(s)

Matias Cattaneo, University of Michigan. <cattaneo@umich.edu>

Rocio Titiunik, University of Michigan. <titiunik@umich.edu>

Gonzalo Vazquez-Bare, University of Michigan. <gvazquez@umich.edu>

References

M.D. Cattaneo, R. Titiunik and G. Vazquez-Bare. (2018). Power Calculations for Regression Discontinuity Designs. *Working paper, University of Michigan*.

Examples

```
# Toy dataset
X <- array(rnorm(2000),dim=c(1000,2))
R <- X[,1] + X[,2] + rnorm(1000)
Y <- 1 + R -.5*R^2 + .3*R^3 + (R>=0) + rnorm(1000)
# Power against tau = 1
tmp <- rdpower(data=cbind(Y,R),tau=1)
# Power against tau = 1 including covariates
tmp <- rdpower(data=cbind(Y,R),tau=1,covs=X)
```

---

rdsampsi *Sample Size Calculations for RD Designs*

Description

`rdsampsi()` performs sample size calculations for RD designs.

Usage

```
rdsampsi(data = NULL, cutoff = 0, tau = NULL, alpha = 0.05,
  beta = 0.8, samph = NULL, nsamples = NULL, all = FALSE,
  bias = NULL, variance = NULL, nratio = NULL, init.cond = NULL,
  plot = FALSE, graph.range = NULL, covs = NULL, deriv = 0,
  p = 1, q = NULL, h = NULL, b = NULL, rho = NULL,
  kernel = "triangular", bwselect = "mserd", vce = "nn",
  cluster = NULL, scalepar = 1, scaleregul = 1, fuzzy = NULL,
  level = 90)
```

Arguments

`data`	a matrix (Y,R) containing the outcome variable and the running variable (as column vectors).
`cutoff`	the RD cutoff (default is 0).
`tau`	specifies the treatment effect under the alternative at which the power function is evaluated. The default is half the standard deviation of the outcome for the untreated group.
`alpha`	specifies the significance level for the power function. Default is 0.05.
`beta`	specifies the desired power. Default is 0.8.
`samph`	sets the bandwidths at each side of the cutoff for power calculation. The first number is the bandwidth to the left of the cutoff and the second number is the bandwidth to the right. Default values are the bandwidths used by `rdrobust`.
`nsamples`	sets the total sample size to the left, sample size to the left inside the bandwidth, total sample size to the right and sample size to the right of the cutoff inside the bandwidth to calculate the variance when the running variable is not specified. When not specified, the values are calculated using the running variable.

all	displays the power using the conventional variance estimator, in addition to the robust bias corrected one.
bias	set bias to the left and right of the cutoff. If not specified, the biases are estimated using rdrobust.
variance	set variance to the left and right of the cutoff. If not specified, the variances are estimated using rdrobust.
nratio	specifies the proportion of treated units in the window. Default is the ratio of the standard deviation of the treated to the sum of the standard deviations for treated and controls.
init.cond	sets the initial condition for the Newton-Raphson algorithm that finds the sample size. Default is the number of observations in the sample with non-missing values of the outcome and running variable.
plot	plots the power function using the conventional and robust bias corrected standard errors from rdrobust.
graph.range	range of the plot.
covs	option for rdrobust(): specifies additional covariates to be used for estimation and inference.
deriv	option for rdrobust(): specifies the order of the derivative of the regression functions to be estimated.
p	option for rdrobust(): specifies the order of the local-polynomial used to construct the point-estimator.
q	option for rdrobust(): specifies the order of the local-polynomial used to construct the bias-correction.
h	option for rdrobust(): specifies the values of the main bandwidth to be used on the left and on the right of the cutoff, respectively.
b	option for rdrobust(): specifies the values of the bias bandwidth b to be used on the left and on the right of the cutoff, respectively.
rho	option for rdrobust(): specifies the value of rho so that the bias bandwidth b equals b=h/rho.
kernel	option for rdrobust(): kernel function used to construct the local-polynomial estimators.
bwselect	option for rdrobust(): specifies the bandwidth selection procedure to be used.
vce	option for rdrobust(): specifies the procedure used to compute the variance-covariance matrix estimator.
cluster	option for rdrobust(): indicates the cluster ID variable used for the cluster-robust variance estimation with degrees-of-freedom weights.
scalepar	option for rdrobust(): specifies scaling factor for RD parameter of interest.
scaleregul	option for rdrobust(): specifies scaling factor for the regularization terms of bandwidth selectors.
fuzzy	option for rdrobust(): specifies the treatment status variable used to implement fuzzy RD estimation.
level	option for rdrobust(): sets the confidence level for confidence intervals.

Value

alpha	significance level
beta	desired power
tau	treatment effect under alternative hypothesis
sampsi.h.tot	total number of observations inside the window
sampsi.h.r	number of observations inside the window to the right of the cutoff
sampsi.h.l	number of observations inside the window to the left of the cutoff
N.r	Total sample size to the right of the cutoff
N.l	Total sample size to the left of the cutoff
samph.r	bandwidth to the right of the cutoff
samph.l	bandwidth to the left of the cutoff
var.r	Robust bias corrected variance to the right of the cutoff
Var.l	Robust bias corrected variance to the left of the cutoff
sampsi.h.tot.cl	implied total number of observations inside the window using conventional s.e.
sampsi.h.r.cl	number of observations inside the window to the right of the cutoff using conventional s.e.
sampsi.h.l.cl	number of observations inside the window to the left of the cutoff using conventional s.e.
no.iter	number of iterations until convergence of the Newton-Raphson algorithm
init.cond	initial condition of the Newton-Raphson algorithm

Author(s)

Matias Cattaneo, University of Michigan. <cattaneo@umich.edu>

Rocio Titiunik, University of Michigan. <titiunik@umich.edu>

Gonzalo Vazquez-Bare, University of Michigan. <gvazquez@umich.edu>

References

M.D. Cattaneo, R. Titiunik and G. Vazquez-Bare. (2018). Power Calculations for Regression Discontinuity Designs. *Working paper, University of Michigan*.

Examples

```
# Toy dataset
X <- array(rnorm(2000),dim=c(1000,2))
R <- X[,1] + X[,2] + rnorm(1000)
Y <- 1 + R -.5*R^2 + .3*R^3 + (R>=0) + rnorm(1000)
# Sample size to achieve power of 0.8 against tau = 1
tmp <- rdsampsi(data=cbind(Y,R),tau=1)
# Sample size against tau = 1 including covariates
tmp <- rdsampsi(data=cbind(Y,R),tau=1,covs=X)
```

Index