Statistics/Data analysis

Title

rdrandinf —— Randomization Inference for RD Designs under Local Randomization.

Syntax

rdrandinf *outvar runvar* [*if*] [*in*] [, **cutoff(#) wl(#) wr(#) statistic(***stat_name***) p(#) evall(#) evalr(#) kernel(***kerneltype***) fuzzy(***fuzzy_var [fuzzy_stat]***) nulltau(#) d(#) dscale(#) ci(***level [tlist]***) interfci(#) bernoulli(***varname***) reps(#) seed(#) covariates(***varlist***) obsmin(#) wmin(#) wobs(#) wstep(#) wmasspoints nwindows(#) rdwstat(***stat_name***) approximate rdwreps(#) level(#) plot graph_options(***graphopts***) obsstep(#) quietly**]

Description

rdrandinf implements randomization inference and related methods for regression discontinuity (RD) designs, employing observations in a specified or data-driven selected window around the cutoff where local randomization is assumed to hold. See Cattaneo, Frandsen and Titiunik (2015) and Cattaneo, Titiunik and Vazquez-Bare (2017) for an introduction to this methodology.

A detailed introduction to this command is given in Cattaneo, Titiunik and Vazquez-Bare (2016).
Companion R functions are also available here.

Companion functions are rdwinselect, rdsensitivity and rdrbounds.

Related Stata and R packages useful for inference in RD designs are described in the following website:

https://rdpackages.github.io/

Options

cutoff(#) specifies the RD cutoff for the running variable *runvar*. Default is **cutoff(0)**.

Window selection

wl(#) specifies the left limit of the window. Default is the minimum value of the running variable.

wr(#) specifies the right limit of the window. Default is the maximum value of the running variable.

Statistic

statistic(*stat_name***)** specifies the statistic to be used. Options are:
diffmeans for difference in means statistic. This is the default option.
ksmirnov for Kolmogorov-Smirnov statistic.
ranksum for Wilcoxon-Mann-Whitney studentized statistic.
all for all three statistics.
The option **ttest** is equivalent to **diffmeans** and included for backward compatibility.

p(#) specifies the order of the polynomial for outcome adjustment model. Default is **p(0)** (constant treatment effect model).

evall(#) specifies the point at the left of the cutoff at which the adjusted outcome is evaluated. Default is the cutoff value.

evalr(#) specifies the point at the right of the cutoff at which the adjusted outcome is evaluated. Default is the cutoff value.

kernel(*kerneltype***)** specifies the type of kernel to use as weighting scheme. Allowed kernel types are **uniform** (uniform kernel), **triangular** (triangular kernel) and **epan** (Epanechnikov kernel). Default is **kernel(uniform)**.

fuzzy(*fuzzy_var [fuzzy_stat]***)** name of the endogenous treatment variable in fuzzy design. Options for statistic in fuzzy designs are:
ar for Anderson-Rubin-type statistic (this is the default option),
tsls for two-stage least squares (TSLS) statistic (only asymptotic approximation).

Inference

nulltau(#) sets the value of the treatment effect under the null hypothesis. Default is **nulltau(0)**.

d(#) effect size for asymptotic power calculation. Default is 0.5 * standard deviation of outcome variable for the control group.

dscale(#) specifies fraction of the standard deviation of the outcome variable for the control group used as alternative hypothesis for asymptotic power calculation. Default is **dscale(.5)**.

ci(*alpha [tlist]***)** calculates a confidence interval for the treatment effect by test inversion, where *alpha* specifies the significance level (typically 0.05 or 0.01) and *tlist* indicates the grid of treatment effects to be evaluated. This option uses **rdsensitivity** to calculate the confidence interval. See rdsensitivity for details. Note: the default tlist can be narrow in some cases, which may truncate the confidence interval. We recommend the user to manually set a large enough tlist.

interfci(#) sets the significance level (alpha) for Rosenbaum's confidence interval under arbitrary interference between units.

bernoulli(*varname***)** specifies that the randomization mechanism is Bernoulli trials (instead of fixed margins randomization). The values of the probability of treatment for each unit must be provided in the variable **varname**.

reps(#) specifies the number of replications. Default is **reps(1000)**.

seed(#) sets the seed for the permutation test. With this option, the user can manually set the desired seed, or can enter the value -1 to use the system seed. Default is **seed(666)**.

Options for rdwinselect

When the window around the cutoff is not specified, **rdrandinf** can select the window automatically using the companion command rdwinselect. The following options are available:

covariates(*varlist***)** specifies the covariates employed by the companion command rdwinselect.

obsmin(#) specifies the minimum number of observations above and below the cutoff in the smallest window employed by the companion command rdwinselect. Default is **obsmin(10)**.

wmin(#) specifies the smallest window to be used (if **minobs(#)** is not specified) by the companion command rdwinselect. Specifying both **wmin(#)** and **obsmin(#)** returns an error.

wobs(#) specifies the number of observations to be added at each side of the cutoff at each step. Default is **wobs(5)**.

wstep(#) specifies the increment in window length (if **obsstep(#)** is not specified) by the companion command rdwinselect. Specifying both **wobs(#)** and **wstep(#)** returns an error.

wmasspoints specifies that the running variable is discrete and each masspoint should be used as a window.

nwindows(#) specifies the number of windows to be used by the companion command rdwinselect. Default is **nwindows(10)**.

rdwstat(#) specifies the statistic to be used by the companion command rdwinselect (see help file for options). Default option is **rdwstat(diffmeans)**.

approximate specifies that covariate balance tests should use a large-sample approximation instead of finite-sample exact randomization inference methods.

rdwreps(#) specifies the number of replications to be used by the companion command rdwinselect. Default is **rdwreps(1000)**.

level(#) specifies the minimum accepted value of the p-value from the covariate balance tests to be used by the companion command rdwinselect. Default is **level(.15)**.

plot draws a scatter plot of the minimum p-value from the covariate balance test against window length implemented by the companion command rdwinselect.

graph_options(*graphopts***)** graph options for plot generated by the companion command rdwinselect.

quietly supress output from the companion command rdwinselect.

obsstep(#) specifies the minimum number of observations to be added on each side of the cutoff by the companion command rdwinselect. This option is deprecated and only included for backward compatibility.

Example: Cattaneo, Frandsen and Titiunik (2015) Incumbency Data

Setup

```
. use rdlocrand_senate.dta, clear
```

Randomization inference with user-specified window and default options

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75)
```

Randomization inference with all statistics

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75) stat(all)
```

Randomization inference with triangular weights

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75) kernel(triangular)
```

Randomization inference on the Kolmogorov-Smirnov statistic with **rdwinselect** window options

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) statistic(ksmirnov)
covariates(dopen population demvoteshlag1) wmin(.5) wstep(.125)
```

Randomization inference with linear adjustment

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75) p(1)
```

Randomization inference under Bernoulli trials with .5 probability of treatment

```
. gen probs=.5
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75) bernoulli(probs)
```

Confidence interval under interference

```
. rdrandinf demvoteshfor2 demmv, cutoff(0) wl(-.75) wr(.75) interfci(.05)
```

Confidence interval for the treatment effect

```
. rdrandinf demvoteshfor2 demmv, wl(-.75) wr(.75) ci(.05 3(1)20)
```

Linear adjustment with effects evaluated at the mean of the running variable

```
. qui sum demmv if abs(demmv)<=.75 & demmv>=0 & demmv!=. & demvoteshfor2!=.
. local mt=r(mean)
. qui sum demmv if abs(demmv)<=.75 & demmv<0 & demmv!=. & demvoteshfor2!=.
. local mc=r(mean)
. rdrandinf demvoteshfor2 demmv, wl(-.75) wr(.75) p(1) evall(`mc') evalr(`mt')
```

Saved results

rdrandinf saves the following in **r()**:

Scalars

r(wl)	left end of window used
r(wr)	right end of window used
r(N)	sample size in used window
r(N_left)	sample size in used window to the left of the cutoff
r(N_right)	sample size in used window to the right of the cutoff
r(p)	order of polynomial in adjusted model
r(obs_stat)	observed statistic
r(randpval)	randomization p-value
r(asy_pval)	asymptotic p-value
r(ci_lb)	lower limit of confidence interval (if **ci** option is specified)
r(ci_ub)	upper limit of confidence interval (if **ci** option is specified)

Locals

r(seed)	seed used in permutations

Matrices

r(obs_stat)	matrix of observed statistics (when **all** is specified)
r(asy_pval)	matrix of asymptotic p-values (when **all** is specified)
r(p_val)	matrix of p-values (when **all** is specified)

References

Cattaneo, M. D., Frandsen, B., and R. Titiunik. 2015. Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate.
Journal of Causal Inference 3(1): 1–24.

Cattaneo, M.D., Titiunik, R. and G. Vazquez-Bare. 2016. Inference in Regression Discontinuity Designs under Local Randomization.
Stata Journal 16(2): 331–367.

Cattaneo, M. D., Titiunik, R. and G. Vazquez-Bare. 2017. Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality.
Journal of Policy Analysis and Management 36(3): 643–681.

Authors

Matias D. Cattaneo, Princeton University, Princeton, NJ. cattaneo@princeton.edu.

Rocio Titiunik, Princeton University, Princeton, NJ. titiunik@princeton.edu.

Gonzalo Vazquez-Bare, UC Santa Barbara, Santa Barbara, CA.
gvazquez@econ.ucsb.edu.