# Package: rddensity (via r-universe)

December 6, 2024

**Type** Package

**Title** Manipulation Testing Based on Density Discontinuity

**Description** Density discontinuity testing (a.k.a. manipulation testing) is commonly employed in regression discontinuity designs and other program evaluation settings to detect perfect self-selection (manipulation) around a cutoff where treatment/policy assignment changes. This package implements manipulation testing procedures using the local polynomial density estimators: rddensity() to construct test statistics and p-values given a prespecified cutoff, rdbwdensity() to perform data-driven bandwidth selection, and rdplotdensity() to construct density plots.

**Version** 2.6

**Author** Matias D. Cattaneo [aut], Michael Jansson [aut], Xinwei Ma [aut, cre]

**Maintainer** Xinwei Ma <x1ma@ucsd.edu>

**Imports** ggplot2, lpdensity (>= 2.2)

**Depends** R (>= 3.1.0)

**License** GPL-2

**Encoding** UTF-8

**RoxygenNote** 7.3.0

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2024-10-06 08:30:02 UTC

## Contents

rddensity-package          *rddensity: Manipulation Testing Based on Density Discontinuity*

### Description

Density discontinuity testing (a.k.a. manipulation testing) is commonly employed in regression discontinuity designs and other program evaluation settings to detect perfect self-selection (manipulation) around a cutoff where treatment/policy assignment changes.

This package implements manipulation testing procedures using the local polynomial density estimators proposed in Cattaneo, Jansson and Ma (2020), and implements graphical procedures with valid confidence bands using the results in Cattaneo, Jansson and Ma (2022, 2023). In addition, this package provides complementary manipulation testing based on finite sample exact binomial testing following the esults in Cattaneo, Frandsen and Titiunik (2015) and Cattaneo, Frandsen and Vazquez-Bare (2017).

A companion Stata package is described in Cattaneo, Jansson and Ma (2018).

Commands: rddensity for manipulation (density discontinuity) testing. rdbwdensity for data-driven bandwidth selection, and rdplotdensity for density plots.

Related Stata and R packages useful for inference in regression discontinuity (RD) designs are described in the website: https://rdpackages.github.io/.

### Author(s)

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Michael Jansson, University of California Berkeley. <mjansson@econ.berkeley.edu>.

Xinwei Ma (maintainer), University of California San Diego. <x1ma@ucsd.edu>.

### References

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2018. On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association* 113(522): 767-779. doi:10.1080/01621459.2017.1285776

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2022. Coverage Error Optimal Confidence Intervals for Local Polynomial Regression. *Bernoulli*, 28(4): 2998-3022. doi:10.3150/21BEJ1445

Cattaneo, M. D., B. Frandsen, and R. Titiunik. 2015. Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate. *Journal of Causal Inference* 3(1): 1-24. doi:10.1515/jci20130010

Cattaneo, M. D., M. Jansson, and X. Ma. 2018. Manipulation Testing based on Density Discontinuity. *Stata Journal* 18(1): 234-261. doi:10.1177/1536867X1801800115

Cattaneo, M. D., M. Jansson, and X. Ma. 2020. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531): 1449-1455. doi:10.1080/01621459.2019.1635480

Cattaneo, M. D., M. Jansson, and X. Ma. 2022. lpdensity: Local Polynomial Density Estimation and Inference. *Journal of Statistical Software*, 101(2): 1–25. doi:10.18637/jss.v101.i02

Cattaneo, M. D., M. Jansson, and X. Ma. 2023. Local Regression Distribution Estimators. *Journal of Econometrics*, 240(2): 105074. doi:10.1016/j.jeconom.2021.01.006

Cattaneo, M. D., R. Titiunik 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. doi:10.1002/pam.21985

McCrary, J. 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics* 142(2): 698-714. doi:10.1016/j.jeconom.2007.05.005

---

| rdbwdensity | *Bandwidth Selection for Manipulation Testing* |
| --- | --- |

---

## Description

rdbwdensity implements several data-driven bandwidth selection methods useful to construct manipulation testing procedures using the local polynomial density estimators proposed in Cattaneo, Jansson and Ma (2020).

A companion Stata package is described in Cattaneo, Jansson and Ma (2018).

Companion command: rddensity for manipulation (density discontinuity) testing.

Related Stata and R packages useful for inference in regression discontinuity (RD) designs are described in the website: https://rdpackages.github.io/.

## Usage

```
rdbwdensity(
  X,
  c = 0,
  p = 2,
  fitselect = "",
  kernel = "",
  vce = "",
  massPoints = TRUE,
  regularize = TRUE,
  nLocalMin = NULL,
  nUniqueMin = NULL
)
```

## Arguments

| | |
| --- | --- |
| X | Numeric vector or one dimensional matrix/data frame, the running variable. |
| c | Numeric, specifies the threshold or cutoff value in the support of X, which determines the two samples (e.g., control and treatment units in RD settings). Default is 0. |
| p | Nonnegative integer, specifies the local polynomial order used to construct the density estimators. Default is 2 (local quadratic approximation). |

| | |
|---|---|
| fitselect | String, specifies the density estimation method. "unrestricted" for density estimation without any restrictions (two-sample, unrestricted inference). This is the default option. "restricted" for density estimation assuming equal distribution function and higher-order derivatives. |
| kernel | String, specifies the kernel function used to construct the local polynomial estimators. "triangular": K(u)=(1-\|u\|)*(\|u\|<=1). This is the default option. "epanechnikov": K(u) = 0.75*(1-u^2)*(\|u\|<=1). "uniform": K(u) = 0.5 * (\|u\|<=1). |
| vce | String, specifies the procedure used to compute the variance-covariance matrix estimator. "plugin" for asymptotic plug-in standard errors. "jackknife" for jackknife standard errors. This is the default option. |
| massPoints | TRUE (default) or FALSE, specifies whether to adjust for mass points in the data. |
| regularize | TRUE (default) or FALSE, specifies whether to conduct local sample size checking. When set to TRUE, the bandwidth is chosen such that the local region includes at least nLocalMin observations and at least nUniqueMin unique observations. |
| nLocalMin | Nonnegative integer, specifies the minimum number of observations in each local neighborhood. This option will be ignored if set to 0, or if regularize=FALSE is used. Default is 20+p+1. |
| nUniqueMin | Nonnegative integer, specifies the minimum number of unique observations in each local neighborhood. This option will be ignored if set to 0, or if regularize=FALSE is used. Default is 20+p+1. |

## Value

| | |
|---|---|
| h | Bandwidths for density discontinuity test, left and right to the cutoff, and asymptotic variance and bias. |
| N | full: full sample size; left/right: sample size to the left/right of the cutoff. |
| opt | Options passed to the function. |
| X_min | Smallest observations to the left and right of the cutoff. |
| X_max | Largest observations to the left and right of the cutoff. |

## Author(s)

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Michael Jansson, University of California Berkeley. <mjansson@econ.berkeley.edu>.

Xinwei Ma (maintainer), University of California San Diego. <x1ma@ucsd.edu>.

## References

Cattaneo, M. D., M. Jansson, and X. Ma. 2018. Manipulation Testing based on Density Discontinuity. *Stata Journal* 18(1): 234-261. doi:10.1177/1536867X1801800115

Cattaneo, M. D., M. Jansson, and X. Ma. 2020. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531): 1449-1455. doi:10.1080/01621459.2019.1635480

## See Also

rddensity

## Examples

```
# Generate a random sample
set.seed(42)
x <- rnorm(2000, mean = -0.5)

# Bandwidth selection
summary(rdbwdensity(X = x, vce="jackknife"))
```

---

rddensity                    *Manipulation Testing Using Local Polynomial Density Estimation*

---

## Description

rddensity implements manipulation testing procedures using the local polynomial density estimators proposed in Cattaneo, Jansson and Ma (2020), and implements graphical procedures with valid confidence bands using the results in Cattaneo, Jansson and Ma (2022, 2023). In addition, the command provides complementary manipulation testing based on finite sample exact binomial testing following the esults in Cattaneo, Frandsen and Titiunik (2015) and Cattaneo, Frandsen and Vazquez-Bare (2017). For an introduction to manipulation testing see McCrary (2008).

A companion Stata package is described in Cattaneo, Jansson and Ma (2018).

Companion commands: rdbwdensity for data-driven bandwidth selection, and rdplotdensity for density plots.

Related Stata and R packages useful for inference in regression discontinuity (RD) designs are described in the website: https://rdpackages.github.io/.

## Usage

```
rddensity(
  X,
  c = 0,
  p = 2,
  q = 0,
  fitselect = "",
  kernel = "",
  vce = "",
  massPoints = TRUE,
  h = c(),
  bwselect = "",
  all = FALSE,
  regularize = TRUE,
  nLocalMin = NULL,
```

```
    nUniqueMin = NULL,
    bino = TRUE,
    binoW = NULL,
    binoN = NULL,
    binoWStep = NULL,
    binoNStep = NULL,
    binoNW = 10,
    binoP = 0.5
)
```

## Arguments

| | |
|---|---|
| X | Numeric vector or one dimensional matrix/data frame, the running variable. |
| c | Numeric, specifies the threshold or cutoff value in the support of X, which determines the two samples (e.g., control and treatment units in RD settings). Default is 0. |
| p | Nonnegative integer, specifies the local polynomial order used to construct the density estimators. Default is 2 (local quadratic approximation). |
| q | Nonnegative integer, specifies the local polynomial order used to construct the bias-corrected density estimators. Default is p+1 (local cubic approximation for default p=2). |
| fitselect | String, specifies the density estimation method. "unrestricted" for density estimation without any restrictions (two-sample, unrestricted inference). This is the default option. "restricted" for density estimation assuming equal distribution function and higher-order derivatives. |
| kernel | String, specifies the kernel function used to construct the local polynomial estimators. "triangular": K(u)=(1−|u|)*(|u|<=1). This is the default option. "epanechnikov": K(u) = 0.75*(1−u^2)*(|u|<=1). "uniform": K(u) = 0.5 * (|u|<=1). |
| vce | String, specifies the procedure used to compute the variance-covariance matrix estimator. "plugin" for asymptotic plug-in standard errors. "jackknife" for jackknife standard errors. This is the default option. |
| massPoints | TRUE (default) or FALSE, specifies whether to adjust for mass points in the data. |
| h | Numeric, specifies the bandwidth used to construct the density estimators on the two sides of the cutoff. If not specified, the bandwidth h is computed by the companion command [rdbwdensity](#) If two bandwidths are specified, the first bandwidth is used for the data below the cutoff and the second bandwidth is used for the data above the cutoff. |
| bwselect | String, specifies the bandwidth selection procedure to be used. "each" based on MSE of each density estimator separately (two distinct bandwidths, hl and hr). "diff" based on MSE of difference of two density estimators (one common bandwidth, hl=hr). "sum" based on MSE of sum of two density estimators (one common bandwidth, hl=hr). "comb" bandwidth is selected as a combination of the alternatives above. This is the default option. For fitselect="unrestricted", it selects median(each,diff,sum). For fitselect = "restricted", it selects min(diff,sum). |

| | |
|---|---|
| all | TRUE or FALSE (default), if specified, will report two testing procedures: conventional test statistic (not valid when using MSE-optimal bandwidth choice) and robust bias-corrected statistic. |
| regularize | TRUE (default) or FALSE, specifies whether to conduct local sample size checking. When set to TRUE, the bandwidth is chosen such that the local region includes at least nLocalMin observations and at least nUniqueMin unique observations. |
| nLocalMin | Nonnegative integer, specifies the minimum number of observations in each local neighborhood. This option will be ignored if set to 0, or if regularize=FALSE is used. Default is 20+p+1. |
| nUniqueMin | Nonnegative integer, specifies the minimum number of unique observations in each local neighborhood. This option will be ignored if set to 0, or if regularize=FALSE is used. Default is 20+p+1. |
| bino | TRUE (default) or FALSE, specifies whether to conduct binomial tests. By default, the initial (smallest) window contains at least 20 observations on each side, and its length is also used as the increment for subsequent windows. This feature is based on the binom.test function. |
| binoW | Numeric, specifies the half length(s) of the initial window. If two values are provided, they will be used for the data below and above the cutoff separately. |
| binoN | Nonnegative integer, specifies the minimum number of observations on each side of the cutoff used for the binomial test. This option will be ignored if binoW is provided. |
| binoWStep | Numeric, specifies the increment in half length(s). |
| binoNStep | Nonnegative integer, specifies the minimum increment in sample size (on each side of the cutoff). This option will be ignored if binoWStep is provided. |
| binoNW | Nonnegative integer, specifies the total number of windows. Default is 10. |
| binoP | Numeric, specifies the null hypothesis of the binomial test. Default is 0.5. |

**Value**

| | |
|---|---|
| hat | left/right: density estimate to the left/right of cutoff; diff: difference in estimated densities on the two sides of cutoff. |
| sd_asy | left/right: standard error for the estimated density to the left/right of the cutoff; diff: standard error for the difference in estimated densities. (Based on asymptotic formula.) |
| sd_jk | left/right: standard error for the estimated density to the left/right of the cutoff; diff: standard error for the difference in estimated densities. (Based on the jackknife method.) |
| test | t_asy/t_jk: t-statistic for the density discontinuity test, with standard error based on asymptotic formula or the jackknife; p_asy/p_jk: p-value for the density discontinuity test, with standard error based on asymptotic formula or the jackknife. |
| hat_p | Same as hat, without bias correction (only available when all=TRUE). |
| sd_asy_p | Same as sd_asy, without bias correction (only available when all=TRUE). |

| sd_jk_p | Same as sd_jk, without bias correction (only available when all=TRUE). |
|---|---|
| test_p | Same as test, without bias correction (only available when all=TRUE). |
| N | full: full sample size; left/right: sample size to the left/right of the cutoff; eff_left/eff_right: effective sample size to the left/right of the cutoff (this depends on the bandwidth). |
| h | left/right: bandwidth used to the left/right of the cutoff. |
| opt | Options passed to the function. |
| bino | Binomial test results. leftWindow/rightWindow: window lengths. leftN/rightN: number of observations. pval: p-values. |
| X_min | left/right: the samllest observation to the left/right of the cutoff. |
| X_max | left/right: the largest observation to the left/right of the cutoff. |

### Author(s)

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Michael Jansson, University of California Berkeley. <mjansson@econ.berkeley.edu>.

Xinwei Ma (maintainer), University of California San Diego. <x1ma@ucsd.edu>.

### References

Cattaneo, M. D., B. Frandsen, and R. Titiunik. 2015. Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate. *Journal of Causal Inference* 3(1): 1-24. doi:10.1515/jci20130010

Cattaneo, M. D., M. Jansson, and X. Ma. 2018. Manipulation Testing based on Density Discontinuity. *Stata Journal* 18(1): 234-261. doi:10.1177/1536867X1801800115

Cattaneo, M. D., M. Jansson, and X. Ma. 2020. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531): 1449-1455. doi:10.1080/01621459.2019.1635480

Cattaneo, M. D., M. Jansson, and X. Ma. 2022. lpdensity: Local Polynomial Density Estimation and Inference. *Journal of Statistical Software*, 101(2): 1–25. doi:10.18637/jss.v101.i02

Cattaneo, M. D., M. Jansson, and X. Ma. 2023. Local Regression Distribution Estimators. *Journal of Econometrics*, 240(2): 105074. doi:10.1016/j.jeconom.2021.01.006

Cattaneo, M. D., R. Titiunik 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. doi:10.1002/pam.21985

McCrary, J. 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics* 142(2): 698-714. doi:10.1016/j.jeconom.2007.05.005

### See Also

rdbwdensity, rdplotdensity

## Examples

```
### Continuous Density
set.seed(42)
x <- rnorm(2000, mean = -0.5)
rdd <- rddensity(X = x, vce = "jackknife")
summary(rdd)

### Bandwidth selection using rdbwdensity()
rddbw <- rdbwdensity(X = x, vce = "jackknife")
summary(rddbw)

### Plotting using rdplotdensity()
# 1. From -2 to 2 with 25 evaluation points at each side
plot1 <- rdplotdensity(rdd, x, plotRange = c(-2, 2), plotN = 25)

# 2. Plotting a uniform confidence band
set.seed(42) # fix the seed for simulating critical values
plot2 <- rdplotdensity(rdd, x, plotRange = c(-2, 2), plotN = 25, CIuniform = TRUE)

### Density discontinuity at 0
x[x > 0] <- x[x > 0] * 2
rdd2 <- rddensity(X = x, vce = "jackknife")
summary(rdd2)
plot3 <- rdplotdensity(rdd2, x, plotRange = c(-2, 2), plotN = 25)
```

---

rddensity_senate          *RD Senate Data*

---

## Description

Extract of the dataset constructed by Cattaneo, Frandsen, and Titiunik (2015), which include measures of incumbency advantage in the U.S. Senate for the period 1914-2010.

## Format

Numeric vector containing 1390 observations:

**margin** Numeric vector. See Cattaneo, Frandsen and Titiunik (2015) regarding details about this dataset.

## Source

Cattaneo, M. D., B. Frandsen, and R. Titiunik. 2015. Randomization Inference in the Regression Discontinuity Design: An Application to the Study of Party Advantages in the U.S. Senate. *Journal of Causal Inference* 3(1): 1-24. doi:10.1515/jci20130010

## rdplotdensity                  *Density Plotting for Manipulation Testing*

### Description

rdplotdensity constructs density plots. It is based on the local polynomial density estimator proposed in Cattaneo, Jansson and Ma (2020, 2023). A companion Stata package is described in Cattaneo, Jansson and Ma (2018).

Companion command: rddensity for manipulation (density discontinuity) testing.

Related Stata and R packages useful for inference in regression discontinuity (RD) designs are described in the website: https://rdpackages.github.io/.

### Usage

```
rdplotdensity(
  rdd,
  X,
  plotRange = NULL,
  plotN = 10,
  plotGrid = c("es", "qs"),
  alpha = 0.05,
  type = NULL,
  lty = NULL,
  lwd = NULL,
  lcol = NULL,
  pty = NULL,
  pwd = NULL,
  pcol = NULL,
  CItype = NULL,
  CIuniform = FALSE,
  CIsimul = 2000,
  CIshade = NULL,
  CIcol = NULL,
  bwselect = NULL,
  hist = TRUE,
  histBreaks = NULL,
  histFillCol = 3,
  histFillShade = 0.2,
  histLineCol = "white",
  title = "",
  xlabel = "",
  ylabel = "",
  legendTitle = NULL,
  legendGroups = NULL,
  noPlot = FALSE
)
```

## Arguments

| | |
|---|---|
| rdd | Object returned by [rddensity](rddensity) |
| X | Numeric vector or one dimensional matrix/data frame, the running variable. |
| plotRange | Numeric, specifies the lower and upper bound of the plotting region. Default is [c-3*hl,c+3*hr] (three bandwidths around the cutoff). |
| plotN | Numeric, specifies the number of grid points used for plotting on the two sides of the cutoff. Default is c(10,10) (i.e., 10 points are used on each side). |
| plotGrid | String, specifies how the grid points are positioned. Options are es (evenly spaced) and qs (quantile spaced). |
| alpha | Numeric scalar between 0 and 1, the significance level for plotting confidence regions. If more than one is provided, they will be applied to the two sides accordingly. |
| type | String, one of "line" (default), "points" or "both", how the point estimates are plotted. If more than one is provided, they will be applied to the two sides accordingly. |
| lty | Line type for point estimates, only effective if type is "line" or "both". 1 for solid line, 2 for dashed line, 3 for dotted line. For other options, see the instructions for ggplot2 or [par](par). If more than one is provided, they will be applied to the two sides accordingly. |
| lwd | Line width for point estimates, only effective if type is "line" or "both". Should be strictly positive. For other options, see the instructions for ggplot2 or [par](par). If more than one is provided, they will be applied to the two sides accordingly. |
| lcol | Line color for point estimates, only effective if type is "line" or "both". 1 for black, 2 for red, 3 for green, 4 for blue. For other options, see the instructions for ggplot2 or [par](par). If more than one is provided, they will be applied to the two sides accordingly. |
| pty | Scatter plot type for point estimates, only effective if type is "points" or "both". For options, see the instructions for ggplot2 or [par](par). If more than one is provided, they will be applied to the two sides accordingly. |
| pwd | Scatter plot size for point estimates, only effective if type is "points" or "both". Should be strictly positive. If more than one is provided, they will be applied to the two sides accordingly. |
| pcol | Scatter plot color for point estimates, only effective if type is "points" or "both". 1 for black, 2 for red, 3 for green, 4 for blue. For other options, see the instructions for ggplot2 or [par](par). If more than one is provided, they will be applied to the two sides accordingly. |
| CItype | String, one of "region" (shaded region, default), "line" (dashed lines), "ebar" (error bars), "all" (all of the previous) or "none" (no confidence region), how the confidence region should be plotted. If more than one is provided, they will be applied to the two sides accordingly. |
| CIuniform | TRUE or FALSE (default), plotting either pointwise confidence intervals (FALSE) or uniform confidence bands (TRUE). |

| | |
|---|---|
| CIsimul | Positive integer, the number of simulations used to construct critical values (default is 2000). This option is ignored if CIuniform=FALSE. |
| CIshade | Numeric, opaqueness of the confidence region, should be between 0 (transparent) and 1. Default is 0.2. If more than one is provided, they will be applied to the two sides accordingly. |
| CIcol | Color of the confidence region. 1 for black, 2 for red, 3 for green, 4 for blue. For other options, see the instructions for ggplot2 or [par](#). If more than one is provided, they will be applied to the two sides accordingly. |
| bwselect | String, the method for data-driven bandwidth selection. Available options are (1) "mse-dpi" (mean squared error-optimal bandwidth selected for each grid point); (2) "imse-dpi" (integrated MSE-optimal bandwidth, common for all grid points); (3) "mse-rot" (rule-of-thumb bandwidth with Gaussian reference model); and (4) "imse-rot" (integrated rule-of-thumb bandwidth with Gaussian reference model). If omitted, bandwidths returned by rddensity will be used. |
| hist | TRUE (default) or FALSE, whether adding a histogram to the background. |
| histBreaks | Numeric vector, giving the breakpoints between histogram cells. |
| histFillCol | Color of the histogram cells. |
| histFillShade | Opaqueness of the histogram cells, should be between 0 (transparent) and 1. Default is 0.2. |
| histLineCol | Color of the histogram lines. |
| title, xlabel, ylabel | |
| | Strings, title of the plot and labels for x- and y-axis. |
| legendTitle | String, title of legend. |
| legendGroups | String Vector, group names used in legend. |
| noPlot | No density plot will be generated if set to TRUE. |

### Details

Bias correction is only used for the construction of confidence intervals/bands, but not for point estimation. The point estimates, denoted by f_p, are constructed using local polynomial estimates of order p, while the centering of the confidence intervals/bands, denoted by f_q, are constructed using local polynomial estimates of order q. The confidence intervals/bands take the form: [f_q - cv * SE(f_q) , f_q + cv * SE(f_q)], where cv denotes the appropriate critical value and SE(f_q) denotes a standard error estimate for the centering of the confidence interval/band. As a result, the confidence intervals/bands may not be centered at the point estimates because they have been bias-corrected. Setting q and p to be equal results on centered at the point estimate confidence intervals/bands, but requires undersmoothing for valid inference (i.e., (I)MSE-optimal bandwidth for the density point estimator cannot be used). Hence the bandwidth would need to be specified manually when q=p, and the point estimates will not be (I)MSE optimal. See Cattaneo, Jansson and Ma (2022, 2023) for details, and also Calonico, Cattaneo, and Farrell (2018, 2022) for robust bias correction methods.

Sometimes the density point estimates may lie outside of the confidence intervals/bands, which can happen if the underlying distribution exhibits high curvature at some evaluation point(s). One possible solution in this case is to increase the polynomial order p or to employ a smaller bandwidth.

**Value**

Estl, Estr    Matrices containing estimation results: (1) grid (grid points), (2) bw (band-widths), (3) nh (number of observations in each local neighborhood), (4) nhu (number of unique observations in each local neighborhood), (5) f_p (point estimates with p-th order local polynomial), (6) f_q (point estimates with q-th order local polynomial, only if option q is nonzero), (7) se_p (standard error corresponding to f_p), and (8) se_q (standard error corresponding to f_q). Variance-covariance matrix corresponding to f_p. Variance-covariance matrix corresponding to f_q. A list containing options passed to the function.

Estplot    A stadnard ggplot object is returned, hence can be used for further customization.

**Author(s)**

Matias D. Cattaneo, Princeton University <cattaneo@princeton.edu>.

Michael Jansson, University of California Berkeley. <mjansson@econ.berkeley.edu>.

Xinwei Ma (maintainer), University of California San Diego. <x1ma@ucsd.edu>.

**References**

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2018. On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association* 113(522): 767-779. doi:10.1080/01621459.2017.1285776

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2022. Coverage Error Optimal Confidence Intervals for Local Polynomial Regression. *Bernoulli*, 28(4): 2998-3022. doi:10.3150/21BEJ1445

Cattaneo, M. D., M. Jansson, and X. Ma. 2018. Manipulation Testing based on Density Discontinuity. *Stata Journal* 18(1): 234-261. doi:10.1177/1536867X1801800115

Cattaneo, M. D., M. Jansson, and X. Ma. 2020. Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531): 1449-1455. doi:10.1080/01621459.2019.1635480

Cattaneo, M. D., M. Jansson, and X. Ma. 2022. lpdensity: Local Polynomial Density Estimation and Inference. *Journal of Statistical Software*, 101(2): 1–25. doi:10.18637/jss.v101.i02

Cattaneo, M. D., M. Jansson, and X. Ma. 2023. Local Regression Distribution Estimators. *Journal of Econometrics*, 240(2): 105074. doi:10.1016/j.jeconom.2021.01.006

**See Also**

rddensity

**Examples**

```
# Generate a random sample with a density discontinuity at 0
set.seed(42)
x <- rnorm(2000, mean = -0.5)
x[x > 0] <- x[x > 0] * 2

# Estimation
rdd <- rddensity(X = x)
```

```
summary(rdd)

# Density plot (from -2 to 2 with 25 evaluation points at each side)
plot1 <- rdplotdensity(rdd, x, plotRange = c(-2, 2), plotN = 25)

# Plotting a uniform confidence band
set.seed(42) # fix the seed for simulating critical values
plot3 <- rdplotdensity(rdd, x, plotRange = c(-2, 2), plotN = 25, CIuniform = TRUE)
```

# Index