Package 'lpdensity'

October 6, 2024

Title Local Polynomial Density Estimation and Inference

Version 2.5

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

Maintainer Xinwei Ma <x1ma@ucsd.edu>

Description Without imposing stringent distributional assumptions or shape restrictions, nonparametric estimation has been popular in economics and other social sciences for counterfactual analysis, program evaluation, and policy recommendations. This package implements a novel density (and derivatives) estimator based on local polynomial regressions, documented in Cattaneo, Jansson and Ma (2022) <doi:10.18637/jss.v101.i02>: lpdensity() to construct local polynomial based density (and derivatives) estimator, and lpbwdensity() to perform data-driven bandwidth selection.

Imports ggplot2, MASS

Depends R (>= 3.1)

License GPL-2

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Repository CRAN

Date/Publication 2024-10-06 06:50:02 UTC

Contents

`lpdensity-package` *lpdensity: Local Polynomial Density Estimation and Inference*

Description

Without imposing stringent distributional assumptions or shape restrictions, nonparametric estimation has been popular in economics and other social sciences for counterfactual analysis, program evaluation, and policy recommendations. This package implements a novel density (and derivatives) estimator based on local polynomial regressions, documented in Cattaneo, Jansson and Ma (2020, 2023).

`lpdensity` implements the local polynomial regression based density (and derivatives) estimator. Robust bias-corrected inference methods, both pointwise (confidence intervals) and uniform (confidence bands), are also implemented. `lpbwdensity` implements the bandwidth selection methods. See Cattaneo, Jansson and Ma (2022) for more implementation details and illustrations.

Related `Stata` and `R` packages useful for nonparametric estimation and inference are available at `https://nppackages.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., 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

`coef.lpbwdensity` *Coef Method for Local Polynomial Density Bandwidth Selection*

Description

The coef method for local polynomial density bandwidth selection objects.

Usage

```
## S3 method for class 'lpbwdensity'
coef(object, ...)
```

Arguments

`object`	Class "lpbwdensity" object, obtained by calling `lpbwdensity`.
`...`	Other arguments.

Value

A matrix containing grid points and selected bandwidths.

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>.

See Also

`lpbwdensity` for data-driven bandwidth selection.

Supported methods: `coef.lpbwdensity`, `print.lpbwdensity`, `summary.lpbwdensity`.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Construct bandwidth
coef(lpbwdensity(X))
```

`coef.lpdensity` *Coef Method for Local Polynomial Density Estimation and Inference*

Description

The coef method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
coef(object, ...)
```

Arguments

`object`	Class "lpdensity" object, obtained by calling `lpdensity`.
`...`	Additional options.

Value

A matrix containing grid points and density estimates using p- and q-th order local polynomials.

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>.

See Also

`lpdensity` for local polynomial density estimation.

Supported methods: `coef.lpdensity`, `confint.lpdensity`, `plot.lpdensity`, `print.lpdensity`, `summary.lpdensity`, `vcov.lpdensity`.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report results
coef(lpdensity(data = X, bwselect = "imse-dpi"))
```

`confint.lpdensity` *Confint Method for Local Polynomial Density Estimation and Inference*

Description

The confint method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
confint(object, parm = NULL, level = NULL, ...)
```

Arguments

`object`	Class "lpdensity" object, obtained by calling `lpdensity`.
`parm`	Integer, indicating which parameters are to be given confidence intervals.
`level`	Numeric scalar between 0 and 1, the significance level for computing confidence intervals
`...`	Additional options, including (i) `grid` specifies a subset of grid points to display the bandwidth; (ii) `gridIndex` specifies the indices of grid points to display the bandwidth (this is the same as `parm`); (iii) `alpha` specifies the significance level (this is `1-level`); (iv) `CIuniform` specifies whether displaying pointwise confidence intervals (`FALSE`, default) or the uniform confidence band (`TRUE`); (v) `CIsimul` specifies the number of simulations used to construct critical values (default is 2000).

Value

A matrix containing grid points and confidence interval end points using p- and q-th order local polynomials.

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>.

See Also

`lpdensity` for local polynomial density estimation.

Supported methods: `coef.lpdensity`, `confint.lpdensity`, `plot.lpdensity`, `print.lpdensity`, `summary.lpdensity`, `vcov.lpdensity`.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report 95% confidence intervals
est1 <- lpdensity(data = X, bwselect = "imse-dpi")
confint(est1)

# Report results for a subset of grid points
confint(est1, parm=est1$Estimate[4:10, "grid"])
confint(est1, grid=est1$Estimate[4:10, "grid"])
confint(est1, gridIndex=4:10)

# Report the 99% uniform confidence band
# Fix the seed for simulating critical values
set.seed(42); confint(est1, level=0.99, CIuniform=TRUE)
set.seed(42); confint(est1, alpha=0.01, CIuniform=TRUE)
```

`lpbwdensity` *Data-driven Bandwidth Selection for Local Polynomial Density Estimators*

Description

`lpbwdensity` implements the bandwidth selection methods for local polynomial based density (and derivatives) estimation proposed and studied in Cattaneo, Jansson and Ma (2020, 2023). See Cattaneo, Jansson and Ma (2022) for more implementation details and illustrations.

Companion command: `lpdensity` for estimation and robust bias-corrected inference.

Related `Stata` and `R` packages useful for nonparametric estimation and inference are available at https://nppackages.github.io/.

Usage

```
lpbwdensity(
  data,
  grid = NULL,
  p = NULL,
  v = NULL,
  kernel = c("triangular", "uniform", "epanechnikov"),
  bwselect = c("mse-dpi", "imse-dpi", "mse-rot", "imse-rot"),
  massPoints = TRUE,
  stdVar = TRUE,
  regularize = TRUE,
  nLocalMin = NULL,
  nUniqueMin = NULL,
  Cweights = NULL,
  Pweights = NULL
)
```

Arguments

data	Numeric vector or one dimensional matrix/data frame, the raw data.
grid	Numeric, specifies the grid of evaluation points. When set to default, grid points will be chosen as 0.05-0.95 percentiles of the data, with a step size of 0.05.
p	Nonnegative integer, specifies the order of the local polynomial used to construct point estimates. (Default is 2.)
v	Nonnegative integer, specifies the derivative of the distribution function to be estimated. 0 for the distribution function, 1 (default) for the density funtion, etc.
kernel	String, specifies the kernel function, should be one of "triangular", "uniform" or "epanechnikov".
bwselect	String, specifies the method for data-driven bandwidth selection. This option will be ignored if bw is provided. Can be (1) "mse-dpi" (default, mean squared error-optimal bandwidth selected for each grid point); or (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).
massPoints	TRUE (default) or FALSE, specifies whether point estimates and standard errors should be adjusted if there are mass points in the data.
stdVar	TRUE (default) or FALSE, specifies whether the data should be standardized for bandwidth selection.
regularize	TRUE (default) or FALSE, specifies whether the bandwidth should be regularized. 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 regularize=FALSE. 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 regularize=FALSE. Default is 20+p+1.
Cweights	Numeric vector, specifies the weights used for counterfactual distribution construction. Should have the same length as the data. This option will be ignored if bwselect is "mse-rot" or "imse-rot".
Pweights	Numeric vector, specifies the weights used in sampling. Should have the same length as the data. This option will be ignored if bwselect is "mse-rot" or "imse-rot".

Value

BW	A matrix containing (1) grid (grid point), (2) bw (bandwidth), (3) nh (number of observations in each local neighborhood), and (4) nhu (number of unique observations in each local neighborhood).
opt	A list containing options passed to the function.

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. 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

Supported methods: `coef.lpbwdensity`, `print.lpbwdensity`, `summary.lpbwdensity`.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Construct bandwidth
bw1 <- lpbwdensity(X)
summary(bw1)

# Display bandwidths for a subset of grid points
summary(bw1, grid=bw1$BW[4:10, "grid"])
summary(bw1, gridIndex=4:10)
```

`lpdensity` *Local Polynomial Density Estimation and Inference*

Description

`lpdensity` implements the local polynomial regression based density (and derivatives) estimator proposed in Cattaneo, Jansson and Ma (2020). Robust bias-corrected inference methods, both pointwise (confidence intervals) and uniform (confidence bands), are also implemented following the results in Cattaneo, Jansson and Ma (2020, 2023). See Cattaneo, Jansson and Ma (2022) for more implementation details and illustrations.

Companion command: `lpbwdensity` for bandwidth selection.

Related `Stata` and `R` packages useful for nonparametric estimation and inference are available at `https://nppackages.github.io/`.

Usage

```
lpdensity(
  data,
  grid = NULL,
  bw = NULL,
  p = NULL,
  q = NULL,
  v = NULL,
  kernel = c("triangular", "uniform", "epanechnikov"),
  scale = NULL,
  massPoints = TRUE,
  bwselect = c("mse-dpi", "imse-dpi", "mse-rot", "imse-rot"),
  stdVar = TRUE,
  regularize = TRUE,
  nLocalMin = NULL,
  nUniqueMin = NULL,
  Cweights = NULL,
  Pweights = NULL
)
```

Arguments

`data`	Numeric vector or one dimensional matrix/data frame, the raw data.
`grid`	Numeric, specifies the grid of evaluation points. When set to default, grid points will be chosen as 0.05-0.95 percentiles of the data, with a step size of 0.05.
`bw`	Numeric, specifies the bandwidth used for estimation. Can be (1) a positive scalar (common bandwidth for all grid points); or (2) a positive numeric vector specifying bandwidths for each grid point (should be the same length as `grid`).
`p`	Nonnegative integer, specifies the order of the local polynomial used to construct point estimates. (Default is `2`.)
`q`	Nonnegative integer, specifies the order of the local polynomial used to construct confidence intervals/bands (a.k.a. the bias correction order). Default is `p+1`. When set to be the same as `p`, no bias correction will be performed. Otherwise it should be strictly larger than `p`.
`v`	Nonnegative integer, specifies the derivative of the distribution function to be estimated. `0` for the distribution function, `1` (default) for the density funtion, etc.
`kernel`	String, specifies the kernel function, should be one of `"triangular"`, `"uniform"`, and `"epanechnikov"`.
`scale`	Numeric, specifies how estimates are scaled. For example, setting this parameter to 0.5 will scale down both the point estimates and standard errors by half. Default is `1`. This parameter is useful if only part of the sample is employed for estimation, and should not be confused with `Cweights` or `Pweights`.
`massPoints`	`TRUE` (default) or `FALSE`, specifies whether point estimates and standard errors should be adjusted if there are mass points in the data.
`bwselect`	String, specifies the method for data-driven bandwidth selection. This option will be ignored if `bw` is provided. Options are (1) `"mse-dpi"` (default, 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).
stdVar	TRUE (default) or FALSE, specifies whether the data should be standardized for bandwidth selection.
regularize	TRUE (default) or FALSE, specifies whether the bandwidth should be regularized. 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 regularize=FALSE. 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 regularize=FALSE. Default is 20+p+1.
Cweights	Numeric, specifies the weights used for counterfactual distribution construction. Should have the same length as the data.
Pweights	Numeric, specifies the weights used in sampling. Should have the same length as the data.

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 an 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 bandwdith 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 (2020, 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

Estimate	A matrix containing (1) grid (grid points), (2) bw (bandwidths), (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).

CovMat_p	The variance-covariance matrix corresponding to f_p.
CovMat_q	The variance-covariance matrix corresponding to f_q.
opt	A list containing options passed to the function.

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. 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

Supported methods: coef.lpdensity, confint.lpdensity, plot.lpdensity, print.lpdensity, summary.lpdensity, vcov.lpdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report results
est1 <- lpdensity(data = X, bwselect = "imse-dpi")
summary(est1)

# Report results for a subset of grid points
summary(est1, grid=est1$Estimate[4:10, "grid"])
summary(est1, gridIndex=4:10)

# Report the 99% uniform confidence band
set.seed(42) # fix the seed for simulating critical values
summary(est1, alpha=0.01, CIuniform=TRUE)

# Plot the estimates and confidence intervals
plot(est1, legendTitle="My Plot", legendGroups=c("X"))
```

```
# Plot the estimates and the 99% uniform confidence band
set.seed(42) # fix the seed for simulating critical values
plot(est1, alpha=0.01, CIuniform=TRUE, legendTitle="My Plot", legendGroups=c("X"))

# Adding a histogram to the background
plot(est1, legendTitle="My Plot", legendGroups=c("X"),
  hist=TRUE, histData=X, histBreaks=seq(-1.5, 1.5, 0.25))
```

`lpdensity.plot` *Plot Method for Local Polynomial Density Estimation and Inference*

Description

This has been replaced by `plot.lpdensity`.

Usage

```
lpdensity.plot(
  ...,
  alpha = NULL,
  type = NULL,
  lty = NULL,
  lwd = NULL,
  lcol = NULL,
  pty = NULL,
  pwd = NULL,
  pcol = NULL,
  grid = NULL,
  CItype = NULL,
  CIuniform = FALSE,
  CIsimul = 2000,
  CIshade = NULL,
  CIcol = NULL,
  hist = FALSE,
  histData = NULL,
  histBreaks = NULL,
  histFillCol = 3,
  histFillShade = 0.2,
  histLineCol = "white",
  title = NULL,
  xlabel = NULL,
  ylabel = NULL,
  legendTitle = NULL,
  legendGroups = NULL
)
```

Arguments

...	Class "lpdensity" object, obtained from calling lpdensity.
alpha	Numeric scalar between 0 and 1, specifies the significance level for plotting confidence intervals/bands. If more than one is provided, they will be applied to each data series accordingly.
type	String, one of "line" (default), "points" and "both", specifies how the point estimates are plotted. If more than one is provided, they will be applied to each data series 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. If more than one is provided, they will be applied to each data series 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. If more than one is provided, they will be applied to each data series 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. If more than one is provided, they will be applied to each data series 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. If more than one is provided, they will be applied to each data series 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 each data series 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. If more than one is provided, they will be applied to each data series accordingly.
grid	Numeric vector, specifies a subset of grid points to plot point estimates. This option is effective only if type is "points" or "both"; or if CItype is "ebar" or "all".
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 each data series accordingly.
CIuniform	TRUE or FALSE (default), plotting either pointwise confidence intervals (FALSE) or uniform confidence bands (TRUE).
CIsimul	Positive integer, specifies the number of simulations used to construct critical values (default is 2000). This option is ignored if CIuniform=FALSE.
CIshade	Numeric, specifies the 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 each data series 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 each data series accordingly.
hist	TRUE or FALSE (default), specifies whether a histogram should be added to the background.
histData	Numeric vector, specifies the data used to construct the histogram plot.
histBreaks	Numeric vector, specifies 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, specifies the title of the plot and labels for the x- and y-axis.
legendTitle	String, specifies the legend title.
legendGroups	String vector, specifies the group names used in legend.

Value

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>.

---

plot.lpdensity *Plot Method for Local Polynomial Density Estimation and Inference*

Description

The plot method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
plot(
  ...,
  alpha = NULL,
  type = NULL,
  lty = NULL,
  lwd = NULL,
  lcol = NULL,
  pty = NULL,
```

```
  pwd = NULL,
  pcol = NULL,
  grid = NULL,
  CItype = NULL,
  CIuniform = FALSE,
  CIsimul = 2000,
  CIshade = NULL,
  CIcol = NULL,
  hist = FALSE,
  histData = NULL,
  histBreaks = NULL,
  histFillCol = 3,
  histFillShade = 0.2,
  histLineCol = "white",
  title = NULL,
  xlabel = NULL,
  ylabel = NULL,
  legendTitle = NULL,
  legendGroups = NULL
)
```

Arguments

`...`	Class "lpdensity" object, obtained from calling `lpdensity`.
`alpha`	Numeric scalar between 0 and 1, specifies the significance level for plotting confidence intervals/bands. If more than one is provided, they will be applied to each data series accordingly.
`type`	String, one of `"line"` (default), `"points"` and `"both"`, specifies how the point estimates are plotted. If more than one is provided, they will be applied to each data series 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`. If more than one is provided, they will be applied to each data series 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`. If more than one is provided, they will be applied to each data series 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`. If more than one is provided, they will be applied to each data series 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`. If more than one is provided, they will be applied to each data series 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

	each data series 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. If more than one is provided, they will be applied to each data series accordingly.
grid	Numeric vector, specifies a subset of grid points to plot point estimates. This option is effective only if type is "points" or "both"; or if CItype is "ebar" or "all".
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 each data series accordingly.
CIuniform	TRUE or FALSE (default), plotting either pointwise confidence intervals (FALSE) or uniform confidence bands (TRUE).
CIsimul	Positive integer, specifies the number of simulations used to construct critical values (default is 2000). This option is ignored if CIuniform=FALSE.
CIshade	Numeric, specifies the 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 each data series 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 each data series accordingly.
hist	TRUE or FALSE (default), specifies whether a histogram should be added to the background.
histData	Numeric vector, specifies the data used to construct the histogram plot.
histBreaks	Numeric vector, specifies 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, specifies the title of the plot and labels for the x- and y-axis.
legendTitle	String, specifies the legend title.
legendGroups	String vector, specifies the group names used in legend.

Value

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>.

See Also

lpdensity for local polynomial density estimation.

Supported methods: coef.lpdensity, confint.lpdensity, plot.lpdensity, print.lpdensity, summary.lpdensity, vcov.lpdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Generate a density discontinuity at 0
X <- X - 0.5; X[X>0] <- X[X>0] * 2

# Density estimation, left of 0 (scaled by the relative sample size)
est1 <- lpdensity(data = X[X<=0], grid = seq(-2.5, 0, 0.05), bwselect = "imse-dpi",
  scale = sum(X<=0)/length(X))
# Density estimation, right of 0 (scaled by the relative sample size)
est2 <- lpdensity(data = X[X>0],  grid = seq(0, 2, 0.05), bwselect = "imse-dpi",
  scale = sum(X>0)/length(X))

# Plot
plot(est1, est2, legendTitle="My Plot", legendGroups=c("Left", "Right"))

# Plot uniform confidence bands
set.seed(42) # fix the seed for simulating critical values
plot(est1, est2, legendTitle="My Plot", legendGroups=c("Left", "Right"), CIuniform=TRUE)

# Adding a histogram to the background
plot(est1, est2, legendTitle="My Plot", legendGroups=c("Left", "Right"),
  hist=TRUE, histBreaks=seq(-2.4, 2, 0.2), histData=X)

# Plot point estimates for a subset of evaluation points
plot(est1, est2, legendTitle="My Plot", legendGroups=c("Left", "Right"),
  type="both", CItype="all", grid=seq(-2, 2, 0.5))
```

---

print.lpbwdensity *Print Method for Local Polynomial Density Bandwidth Selection*

Description

The print method for local polynomial density bandwidth selection objects.

Usage

```
## S3 method for class 'lpbwdensity'
print(x, ...)
```

Arguments

x	Class "lpbwdensity" object, obtained by calling lpbwdensity.
...	Other arguments.

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>.

See Also

lpbwdensity for data-driven bandwidth selection.

Supported methods: coef.lpbwdensity, print.lpbwdensity, summary.lpbwdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Construct bandwidth
print(lpbwdensity(X))
```

---

print.lpdensity *Print Method for Local Polynomial Density Estimation and Inference*

Description

The print method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
print(x, ...)
```

Arguments

x	Class "lpdensity" object, obtained from calling lpdensity.
...	Additional options.

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>.

See Also

lpdensity for local polynomial density estimation.

Supported methods: coef.lpdensity, confint.lpdensity, plot.lpdensity, print.lpdensity, summary.lpdensity, vcov.lpdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report results
print(lpdensity(data = X, bwselect = "imse-dpi"))
```

---

summary.lpbwdensity *Summary Method for Local Polynomial Density Bandwidth Selection*

Description

The summary method for local polynomial density bandwidth selection objects.

Usage

```
## S3 method for class 'lpbwdensity'
summary(object, ...)
```

Arguments

object	Class "lpbwdensity" object, obtained by calling lpbwdensity.
...	Additional options, including (i) grid specifies a subset of grid points to display the bandwidth; (ii) gridIndex specifies the indices of grid points to display the bandwidth.

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>.

See Also

lpbwdensity for data-driven bandwidth selection.

Supported methods: coef.lpbwdensity, print.lpbwdensity, summary.lpbwdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Construct bandwidth
bw1 <- lpbwdensity(X)
summary(bw1)

# Display bandwidths for a subset of grid points
summary(bw1, grid=bw1$BW[4:10, "grid"])
summary(bw1, gridIndex=4:10)
```

`summary.lpdensity` *Summary Method for Local Polynomial Density Estimation and Inference*

Description

The summary method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
summary(object, ...)
```

Arguments

`object`	Class "lpdensity" object, obtained from calling `lpdensity`.
`...`	Additional options, including (i) `grid` specifies a subset of grid points to display results; (ii) `gridIndex` specifies the indices of grid points to display results; (iii) `alpha` specifies the significance level; (iv) `CIuniform` specifies whether displaying pointwise confidence intervals (`FALSE`, default) or the uniform confidence band (`TRUE`); (v) `CIsimul` specifies the number of simulations used to construct critical values (default is `2000`).

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>.

See Also

`lpdensity` for local polynomial density estimation.

Supported methods: `coef.lpdensity`, `confint.lpdensity`, `plot.lpdensity`, `print.lpdensity`, `summary.lpdensity`, `vcov.lpdensity`.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report results
est1 <- lpdensity(data = X, bwselect = "imse-dpi")
summary(est1)

# Report results for a subset of grid points
summary(est1, grid=est1$Estimate[4:10, "grid"])
summary(est1, gridIndex=4:10)

# Report the 99% uniform confidence band
set.seed(42) # fix the seed for simulating critical values
summary(est1, alpha=0.01, CIuniform=TRUE)
```

`vcov.lpdensity` *Vcov Method for Local Polynomial Density Estimation and Inference*

Description

The vcov method for local polynomial density objects.

Usage

```
## S3 method for class 'lpdensity'
vcov(object, ...)
```

Arguments

`object`	Class "lpdensity" object, obtained by calling `lpdensity`.
`...`	Additional options.

Value

`stdErr`	A matrix containing grid points and standard errors using p- and q-th order local polynomials.
`CovMat_p`	The variance-covariance matrix corresponding to `f_p`.
`CovMat_q`	The variance-covariance matrix corresponding to `f_q`.

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>.

See Also

lpdensity for local polynomial density estimation.

Supported methods: coef.lpdensity, confint.lpdensity, plot.lpdensity, print.lpdensity, summary.lpdensity, vcov.lpdensity.

Examples

```
# Generate a random sample
set.seed(42); X <- rnorm(2000)

# Estimate density and report results
vcov(lpdensity(data = X, bwselect = "imse-dpi"))
```

Index