

Title

kdbwselect — Bandwidth Selection Procedures for Kernel Density Estimation and Inference.

Syntax

kdbwselect varname [if] [in] [, eval(gridvar) neval(#) rho(#) kernel(kernelfn)
bwselect(bwmethod) bwcheck(#) imsegrid(#) separator(#)]

Description

- kdbwselect implements bandwidth selectors for kernel density point estimators and inference procedures developed in <u>Calonico</u>, <u>Cattaneo and Farrell (2018)</u>. See also <u>Calonico</u>, <u>Cattaneo and Farrell (2020)</u> for related optimality results. It also implements other bandwidth selectors available in the literature. See Wand and Jones (1995) and Fan and Gijbels (1996) for background references.
- A detailed introduction to this command is given in <u>Calonico</u>, <u>Cattaneo and Farrell</u> (2019).
- Companion command is: kdrobust for point estimation and inference procedures.
- Related Stata and R packages useful for empirical analysis are described in the following website:

https://nppackages.github.io/

Options

- eval(gridvar) specifies the grid of evaluation points for xvar. By default it
 uses 30 equally spaced points over to support of xvar.
- neval(#) specifies the number of evaluation points to estimate the regression
 functions. Default is 30 evaluation points.
- **rho**(#) specifies the value of *rho*, so that the bias bandwidth b equals b=h/rho. Default is **rho**(1) if h is specified but b is not.
- bwselect(bwmethod) bandwidth selection procedure to be used. By default it
 computes both h and b, unless rho is specified, in which case it only computes
 h and sets b=h/rho. Options are:

 - $\textbf{mse-rot} \ \, \texttt{ROT} \ \, \texttt{implementation} \ \, \texttt{of} \ \, \texttt{MSE-optimal} \ \, \texttt{bandwidth.}$
 - imse-dpi second-generation DPI implementation of IMSE-optimal bandwidth.
 - imse-rot ROT implementation of IMSE-optimal bandwidth.
 - ce-dpi second generation DPI implementation of CE-optimal bandwidth.
 - ce-rot ROT implementation of CE-optimal bandwidth.
- Note: MSE = Mean Square Error; IMSE = Integrated Mean Squared Error; CE = Coverage Error; DPI = Direct Plug-in; ROT = Rule-of-Thumb.
 - Default is **bwselect**(*mse-dpi*). For details on implementation see <u>Calonico</u>, <u>Cattaneo and Farrrell (2019)</u>.
- bwcheck(#) specifies an optional positive integer so that the selected bandwidth
 is enlarged to have at least # effective observations available at each
 evaluation point.
- imsegrid(#) number of evaluations points used to compute the IMSE bandwidth selector. Default is 30 points.
- separator(#) draws separator line after every # variables; default is separator(5).

Example:

Setup

. sysuse auto

MSE bandwidth selection procedure

. kdbwselect length

Saved results

kdbwselect saves the following in e():

Scalars

e(N) original number of observations

Macros

e(varname) name of variable

e (bwselect) bandwidth selection choice

e(kernel) kernel choice

Matrices

e(bws) estimation result

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.

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2019. nprobust: Nonparametric Kernel-Based Estimation and Robust Bias-Corrected Inference. Journal of Statistical Software, 91(8): 1-33. doi: 10.18637/jss.v091.i08.

Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2020. <u>Coverage Error Optimal Confidence Intervals for Local Polynomial Regression</u>, working paper.

Fan, J., and Gijbels, I. 1996. Local Polynomial Modelling and Its Applications, London: Chapman and Hall.

Wand, M., and Jones, M. 1995. Kernel Smoothing, Florida: Chapman & Hall/CRC.

Authors

Sebastian Calonico, Columbia University, New York, NY. sebastian.calonico@columbia.edu.

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

Max H. Farrell, University of Chicago, Chicago, IL. max.farrell@chicagobooth.edu.