Univariate kernel density estimation

Ben Jann
ETH Zürich, Switzerland, jann@soz.gess.ethz.ch
August 3, 2007

Abstract. The methods and formulas used by the `kdens` package for Stata 9.2 are discussed in this document. To install `kdens`, type `ssc install kdens` in Stata. Note that the `moremata` package is required (type `ssc install moremata`).

Preliminary note: In addition to the methods presented below, `kdens` also supports confidence interval and variance estimation using bootstrap and jackknife techniques. The implementation of the bootstrap and jackknife for density estimation is straightforward and is therefore not discussed here.

1 Standard kernel density estimation

Let $X_1, \ldots, X_n$ be a sample from X, where X has the probability density function $f(x)$. Furthermore, let $w_1, \ldots, w_n$ be associated weights (set $w_i = 1 \, \forall \, i$ if there are no weights). The density of X can then be estimated as:

$$\widehat{f}_K(x; h) = \frac{1}{W} \sum_{i=1}^{n} \frac{w_i}{h} K\left(\frac{x - X_i}{h}\right) \tag{1}$$

where $W = \sum_{i=1}^{n} w_i$ and $K(z)$ is a kernel function (see Section 9). h is the smoothing parameter (the kernel halfwidth or "bandwidth"). Formula (1) is also used, for example, by official Stata's `kdensity` (see [R] **kdensity**).

2 Adaptive kernel density estimation

The adaptive kernel density estimator is defined as

$$\widehat{f}^a_K(x; h) = \frac{1}{W} \sum_{i=1}^{n} \frac{w_i}{h\lambda_i} K\left(\frac{x - X_i}{h\lambda_i}\right) \tag{2}$$

where λ, the local bandwidth factors, are based on a preliminary fixed bandwidth density estimate. The factors are estimated as

$$\widehat{\lambda}_i = \sqrt{\frac{G\left(\widehat{f}_K(X; h)\right)}{\widehat{f}_K(X_i; h)}}, \quad i = 1, \ldots, n \tag{3}$$

where $G()$ stands for the geometric mean over all i. Note that $G(\lambda) = 1$ and thus $G(h\lambda) = h$. The estimator is based on Abramson (1982). Also see, for example, Silverman (1986, 100–110), Fox (1990, 100–103), Salgado-Ugarte et al. (1993), Salgado-Ugarte and Pérez-Hernández (2003), or Van Kerm (2003).

Technical note: $\widehat{f}_K(X_i;h)$ is determined by linear interpolation if X_i falls between the points at which the preliminary density estimate has been evaluated.

3 Approximate variance estimation

An approximate estimator for the variance of $\widehat{f}_K(x;h)$ at point x is given as

$$\widetilde{V}\{\widehat{f}_K(x;h)\} = \frac{1}{nh}R(K)\widehat{f}_K(x;h) - \frac{1}{n}\widehat{f}_K(x;h)^2 \tag{4}$$

where

$$R(K) = \int_{-\infty}^{\infty} K(z)^2\, dz$$

(for theoretical background see, e.g., Scott 1992, 130). A simple extension of (4) to the adaptive kernel density method is

$$\widetilde{V}\{\widehat{f}^a_K(x;h)\} = \frac{1}{nh\lambda(x)}R(K)\widehat{f}^a_K(x;h) - \frac{1}{n}\widehat{f}^a_K(x;h)^2 \tag{5}$$

where $\lambda(x)$ denotes the bandwidth factor at point x. However, note that (5) understates the true variance since the local bandwidth factors are assumed fixed.

Probability weights can be taken into account by adding a penalty for the amount of variability in the weights distribution. In particular,

$$\widetilde{V}\{\widehat{f}_K(x;h)\} = \frac{\sum_{i=1}^n w_i^2}{W^2}\left[\frac{1}{h}R(K)\widehat{f}_K(x;h) - \widehat{f}_K(x;h)^2\right] \tag{6}$$

where W is the sum of weights as defined above (see Van Kerm 2003 and Burkhauser et al. 1999 for a similar approach). The assumption behind this formula, however, is that the weights w are essentially independent from X, an assertion that may not be appropriate.[1]

Note that (4) differs from the standard variance formula often found in the literature. The standard formula only contains the first term (see, e.g., Van Kerm 2003, Silverman 1986, 40, Härdle et al. 2004) and, although the second term asymptotically disappears, has a quite substantial bias in finite samples. Estimator (4) is more accurate than the standard formula in finite samples.

1. A point could also be made that weights should be omitted from estimation entirely if the assertion of independence is, in fact, true.

4 Exact variance estimation

The variance of $\widehat{f}_K(x;h)$ can be written as

$$\begin{aligned} V\{\widehat{f}_K(x;h)\} &= \frac{1}{n} E\left[K_h(x-X) - E\{K_h(x-X)\}\right]^2 \\ &= \frac{1}{n}\left[E\{K_h(x-X)^2\} - E\{K_h(x-X)\}^2\right] \end{aligned} \tag{7}$$

where $K_h(z) = 1/h\, K(z/h)$ and

$$E\{K_h(x-X)\} = E\{\widehat{f}_K(x;h)\} = \int_{-\infty}^{\infty} K_h(x-y)f(y)\,dy$$

A natural estimators for (7) is

$$\widehat{V}\{\widehat{f}_K(x;h)\} = \frac{1}{n}\left\{\frac{1}{W}\sum_{i=1}^{n}\frac{w_i}{h^2}K\left(\frac{x-X_i}{h}\right)^2 - \widehat{f}_K(x;h)^2\right\} \tag{8}$$

or, in the case of the adaptive method,

$$\widehat{V}\{\widehat{f}^a_K(x;h)\} = \frac{1}{n}\left\{\frac{1}{W}\sum_{i=1}^{n}\frac{w_i}{(h\lambda_i)^2}K\left(\frac{x-X_i}{h\lambda_i}\right)^2 - \widehat{f}^a_K(x;h)^2\right\} \tag{9}$$

where w is assumed to represent frequency weights or analytic weights (for similar formulas see, e.g., Hall 1992 and Fiorio 2004). Again note that (9) will be downward biased because the local bandwidth factors are assumed fixed.

If the weights are sampling weights, an estimator for (7) may be derived as

$$\widehat{V}\{\widehat{f}_K(x;h)\} = \frac{1}{W^2}\sum_{i=1}^{n} w_i^2\left\{\frac{1}{h}K\left(\frac{x-X_i}{h}\right) - \widehat{f}_K(x;h)\right\}^2 \tag{10}$$

In many cases the approximate estimator is quite good and using the exact formula is not worth the extra computational effort. However, the exact formula should be used if the data contain sampling weights. Furthermore, note that both the exact and the approximate variance formulas assume h fixed. Data dependent choice of h, however, may result in additional variability of the density estimate, especially in regions with high curvature.

5 Confidence intervals

Pointwise confidence intervals for $f(x)$ are constructed as

$$\widehat{f}_K(x;h) \pm z_{1-\alpha/2}\sqrt{\widehat{V}\{\widehat{f}_K(x;h)\}} \tag{11}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

Confidence intervals such as (11) may have bad coverage due to the well-known bias in $\widehat{f}_K$. One suggestion to improve coverage is to use an undersmoothed density estimate to construct the confidence intervals (Hall 1992, Fiorio 2004). The implementation of this approach in `kdens` uses

$$h_{\mathrm{us}} = h\frac{n^{1/5}}{n^{\tau}}$$

where τ is the undersmoothing parameter. τ should be larger than $1/5$; $\tau = 1/4$ is the default choice in `kdens` once undersmoothing is requested.

6 Density estimation for bounded variables

Two simple methods to estimate the density of variables with bounded domain are the renormalization method and the reflection method. Both methods produce consistent estimates, but have a bias of order h near the boundaries (compared to the usual bias of order h^2 in the interior). Various more advanced correction techniques with bias of order h^2 at the boundaries have been proposed (see, e.g., Jones and Foster 1996 or Karunamuni and Alberts 2005). One of these methods is the linear combination technique discussed in Jones (1993).

6.1 Renormalization

The most natural way to deal with the boundary problem is to use a standard kernel density estimate and then locally rescale it relative to the amount of local kernel mass that lies within the support of X (see, e.g., Jones 1993, 137). Let L be the lower boundary of the support of X (e.g. $L=0$) and U be the upper boundary (e.g. $U=1$). Furthermore, let

$$a_0(l,u) = \int_l^u K(y)\,dy$$

The renormalization version of the standard estimator then is

$$\widehat{f}^n_K(x;h,L,U) = \frac{1}{a_0\left(\frac{L-x}{h},\frac{U-x}{h}\right)}\widehat{f}_K(x;h) \tag{12}$$

for $x \in [L,U]$. Furthermore, the boundary renormalization analogue to the adaptive estimator in (2) can be written as

$$\widehat{f}^{n,a}_K(x;h,L,U) = \frac{1}{W}\sum_{i=1}^{n}\frac{1}{a_0\left(\frac{L-x}{h\lambda_i},\frac{U-x}{h\lambda_i}\right)}\frac{w_i}{h\lambda_i}K\left(\frac{x-X_i}{h\lambda_i}\right) \tag{13}$$

The approximate variance estimator for $\widehat{f}^n_K$ is

$$\widetilde{V}\{\widehat{f}^n_K(x;h,L,U)\} = \frac{1}{nh}\frac{b\left(\frac{L-x}{h},\frac{U-x}{h}\right)}{a_0\left(\frac{L-x}{h},\frac{U-x}{h}\right)^2}\widehat{f}^n_K(x;h,L,U) - \frac{1}{n}\widehat{f}^n_K(x;h,L,U)^2 \tag{14}$$

where

$$b(l,u) = \int_l^u K(y)^2\, dy$$

(see Jones 1993). The exact variance estimator can be written as

$$\widehat{V}\{\widehat{f}_K^n(x;h,L,U)\} = \frac{1}{a_0\left(\frac{L-x}{h}, \frac{U-x}{h}\right)^2}\widehat{V}\{\widehat{f}_K(x;h)\} \tag{15}$$

in the most simple case. The renormalization variance estimators for the adaptive kernel method and for data containing sampling weights can also be easily derived, but the resulting formulas are more complicated.

6.2 Reflection

The reflection estimator approaches the boundary problem by "reflecting" the data at the boundaries (see, e.g., Silverman 1986, 30, Ćwik and Mielniczuk 1993). In the standard case the reflection estimator is given as

$$\widehat{f}_K^r(x;h,L,U) = \frac{1}{W}\sum_{i=1}^{n}\frac{w_i}{h}K^r(x;X_i,h,L,U) \tag{16}$$

for $x \in [L,U]$, where

$$K^r(x;X,h,L,U) = K\left(\frac{x-X}{h}\right) + K\left(\frac{x+X-2L}{h}\right) + K\left(\frac{x+X-2U}{h}\right)$$

The reflection technique, that is, replacing K with K^r, can be applied analogously to the adaptive estimator and also the exact variance estimators are easily derived. Unfortunately, however, the reflection solution for the approximate variance estimator is more complex and not supported by `kdens`.

6.3 Linear Combination

Let $z = (x-X)/h$, $l = (L-x)/h$, and $u = (U-x)/h$. The linear combination technique then replaces $K(z)$ by

$$K^{lc}(z;l,u) = \frac{a_2(l,u) - a_1(-u,-l)z}{a_2(l,u)a_0(l,u) - (a_1(-u,-l))^2}K(z) \tag{17}$$

where

$$a_1(-u,-l) = \int_{-u}^{-l} yK(y)\, dy, \qquad a_2(l,u) = \int_l^u y^2K(y)\, dy$$

Similar to the reflection technique, exact variance estimates can be obtained by simply plugging K^{lc} into the standard formulas. Approximate variance estimation is not supported.

7 Binned estimation

7.1 Density

Kernel density estimators such as (1) involve lots of computations. One solution to reduce processing time is to estimate the density based on binned data. The binned kernel density estimator is defined as

$$\widetilde{f}_K(g_j; h) = \frac{1}{W} \sum_{\ell=1}^{m} \frac{c_\ell}{h} K\left(\frac{g_j - g_\ell}{h}\right), \quad j = 1, \ldots, m \tag{18}$$

where $g_1, \ldots, g_m$ is a grid of m equally spaced evaluation points and $c_1, \ldots, c_m$ are the associated grid counts with $\sum c_\ell = W$. Given the equal spacing of grid points, equation (18) has a discrete convolution structure and can be calculated using fast Fourier transform (see Wand and Jones 1995, 182–188, for details). This makes the estimator very fast. The grid counts are computed using linear binning as follows. Let g^- and g^+ be the two nearest grid points below and above observation X_i. Then $w_i(g^+ - X_i)/(g^+ - g^-)$ is added to the grid count at g^- and $w_i(X_i - g^-)/(g^+ - g^-)$ is added to the count at g^+. Note that the results from (18) are usually quite accurate even for relatively small m. The rule-of-thumb given by Hall and Wand (1996, 182) is that ,,grid sizes of about 400–500 are adequate for a wide range of practical situations“.

The binned version of the adaptive kernel density estimator is

$$\widetilde{f}^a_K(g_j; h) = \frac{1}{W} \sum_{\ell=1}^{m} \frac{c_\ell}{h\lambda_\ell} K\left(\frac{g_j - g_\ell}{h\lambda_\ell}\right), \quad j = 1, \ldots, m \tag{19}$$

where λ denotes the local bandwidth factors. Unfortunately, the computational shortcut used for (18) is not applicable to (19).

7.2 Variance

Variance estimation is straightforward with binned data. For example,

$$\widetilde{V}\{\widetilde{f}_K(g_j; h)\} = \frac{1}{nh} R(K) \widetilde{f}_K(g_j; h) - \frac{1}{n} \widetilde{f}_K(g_j; h)^2 \tag{20}$$

If sampling weights are applied, a reasonable variance formula for the binned estimator is

$$\widehat{V}\{\widetilde{f}_K(g_j; h)\} = \frac{1}{W^2} \sum_{\ell=1}^{m} c_\ell(w^2) \left\{\frac{1}{h} K\left(\frac{g_j - g_\ell}{h}\right) - \widetilde{f}_K(g_j; h)\right\}^2 \tag{21}$$

with $c(w^2)$ representing linearly binned squared weights.

7.3 Density derivatives and density functionals

Advanced automatic bandwidth selection involves estimating density functionals of the form

$$R^{(r)} = R(f^{(r)}) = \int f^{(r)}(z)^2\,dz = (-1)^r \int f^{(2r)}(z)f(z)\,dz$$

where $f^{(r)}$ denotes the rth derivative of f. A binned approximation estimator employing the gaussian kernel can be written as

$$\widetilde{R}_{\phi}^{(r)}(m,h) = (-1)^r \frac{1}{W} \sum_{j=1}^{m} c_j \widetilde{f}_{\phi}^{(2r)}(g_j;h) \tag{22}$$

where

$$\widetilde{f}_{\phi}^{(r)}(g_j;h) = \frac{1}{W} \sum_{\ell=1}^{m} \frac{c_\ell}{h^{r+1}} \phi^{(r)}\left(\frac{g_j - g_\ell}{h}\right), \quad j = 1,\dots,m \tag{23}$$

and $\phi^{(r)}$ denotes the rth derivative of ϕ, the standard normal density. Equation (23) can be solved as the convolution of fast Fourier transforms.

7.4 Bounded variables

Binned versions of the estimators for bounded variables are usually simple to derive. For example, the binned renormalization density estimator in the standard case is

$$\widetilde{f}_K^{n}(g_j;h,L,U) = \frac{1}{I_K\left(\frac{L-g_j}{h}, \frac{U-g_j}{h}\right)} \widetilde{f}_K(g_j;h), \quad j = 1,\dots,m \tag{24}$$

for $g_j \in [L,U]$.

An exception is the estimation of density derivatives and density functionals where the renormalization and the linear combination methods have no easy solutions. Fortunately, the reflection technique is a valuable alternative in this situation.

8 Data-dependent bandwidth selection

It can be shown from asymptotic theory that the bandwidth

$$h_{\text{opt}} = \left[\frac{R(K)}{\{\sigma_K^2\}^2 R(f'')n}\right]^{\frac{1}{5}} \tag{25}$$

where

$$\sigma_K^2 = \int z^2 K(z)\,dz, \quad R(K) = \int \{K(z)\}^2\,dz$$

is "optimal" in the sense that it minimizes the asymptotic mean integrated squared error (AMISE). Note that σ_K^2, the kernel variance, and $R(K)$, the kernel "roughness", are known properties of the chosen kernel function. However, $R(f'')$, where f'' denotes the second derivative of f, is unknown.

8.1 Quick and simple rules

Normal scale rule

One idea to derive a first guess for h_{opt} is to assume a specific functional form for the density and then solve equation (25). If the density is assumed to be normal, the optimal h can be estimated as

$$\widehat{h}_{\text{N}} = \delta_K \left[\frac{8\sqrt{\pi}}{3}\right]^{\frac{1}{5}} \widehat{\sigma} n^{-\frac{1}{5}} \tag{26}$$

where $\widehat{\sigma}$ is an estimate of scale and δ_K is the "canonical bandwidth" of the chosen kernel function, that is

$$\delta_K = \left(\frac{R(K)}{\{\sigma_K^2\}^2}\right)^{1/5}$$

(see, e.g., Scott 1992, 141–143; Härdle et al. 2004). Usually, $\widehat{\sigma} = \min(s_x\,, IQR_x/1.349)$ is used where s_x is the standard deviation and IQR_x is the inter-quantile range of the observed data.

Oversmoothed bandwidth rule

It can be shown that, given the scale parameter σ, h_{opt} has a simple upper bound (see, e.g., Salgado-Ugarte et al. 1995). This upper bound can be estimated as

$$\widehat{h}_{\text{O}} = \delta_K \left[\frac{243}{35}\right]^{\frac{1}{5}} s_x n^{-\frac{1}{5}} \tag{27}$$

Note that $h_{\text{O}} \simeq 1.08 h_{\text{N}}$. While $\widehat{h}_{\text{O}}$ usually is too large and results in a density estimate that is too smooth, it is a good starting point for subjective choice of bandwidth. In fact, it may be convenient to choose the bandwidth as a fractional of $\widehat{h}_{\text{O}}$, for example $0.8\widehat{h}_{\text{O}}$ or $0.5\widehat{h}_{\text{O}}$.

Optimal of Silverman

Based on simulations studies, Silverman (1986, 45–48) suggested using

$$\widehat{h}_{\text{S}\phi} = 0.9\,\hat{\sigma} n^{-\frac{1}{5}} \tag{28}$$

for the gaussian kernel, which translates to

$$\widehat{h}_{\text{S}} = 1.159\,\delta_K \hat{\sigma} n^{-\frac{1}{5}} \tag{29}$$

in the general case. $\widehat{h}_{\text{S}}$ is used as the default bandwidth estimate in `kdens`. Official Stata's `kdensity` uses $\widehat{h}_{\text{S}\phi}$.

8.2 Sheather-Jones plug-in estimator

The implementation of the Sheather-Jones plug-in estimator $\widehat{h}_{\mathrm{SJPI}}$ closely follows Sheather and Jones (1991) using a gaussian kernel. Bounded variables are taken into account using the reflection technique. Note that `kdens` imposes an upper limit for $\widehat{h}_{\mathrm{SJPI}}$. This limit is $\widehat{h}_{\mathrm{O}}$.

8.3 Direct plug-in estimator

The implementation of the direct plug-in estimator $\widehat{h}_{\mathrm{DPI}}$ in `kdens` closely follows Wand and Jones (1995, 71–74) (also see Wand and Jones 1995, 177–189) using a gaussian kernel. Bounded variables are taken into account using the reflection technique.

8.4 Probability weights

Probability weights inflate the variance of the density estimate and formula (25) is inappropriate. A rough correction is to apply a penalty for the variability of the weights to a standard bandwidth estimate (also see formula 6), i.e.

$$\widehat{h}^w = \left(\frac{n \sum_{i=1}^n w_i^2}{W^2} \right)^{\frac{1}{5}} \cdot \widehat{h}$$

where $\widehat{h}$ is computed by one of the above methods. `kdens` applies this correction if `pweight`s are specified.

9 Kernel functions

Various kernels are supported by `kdens`. Note that the actual kernel functions are not included in the `kdens` package; they are provided by the `moremata` package. Tables 1–6 give an overview of the various kernels and their properties.

10 References

Abramson, I. S. 1982. On Bandwidth Variation in Kernel Estimates. A Square Root Law. *The Annals of Statistics* 10(4): 1217–1223.

Burkhauser, R. V., A. C. Cutts, M. C. Daly, and S. P. Jenkins. 1999. Testing the Significance of Income Distribution Changes over the 1980s Business Cycle: A Cross-National Comparison. *Journal of Applied Econometrics* 14(3): 253–272.

Ćwik, J. and J. Mielniczuk. 1993. Data-dependent bandwidth choice for a grade density kernel estimate. *Statistics & Probability Letters* 16: 397–405.

Fiorio, C. V. 2004. Confidence intervals for kernel density estimation. *The Stata Journal* 4(2): 168–179.

Fox, J. 1990. Describing Univariate Distributions. In *Modern Methods of Data Analysis*, eds. J. Fox and J. S. Long, 58–123. Newbury Park, CA: Sage.

Hall, P. 1992. Effect of Bias Estimation on Coverage Accuracy of Bootstrap Confidence Intervals for a Probability Density. *The Annals of Statistics* 20(2): 675–694.

Hall, P. and M. P. Wand. 1996. On the Accuracy of Binned Kernel Density Estimators. *Journal of Multivariate Analysis* 56: 165–184.

Härdle, W., M. Müller, S. Sperlich, and A. Werwatz. 2004. *Nonparametric and Semiparametric Models. An Introduction*. Berlin: Springer.

Jones, M. C. 1993. Simpel boundary correction for kernel density estimation. *Statistics and Computing* 3: 135–146.

Jones, M. C. and P. J. Foster. 1996. A Simple Nonnegative Boundary Correction Method for Kernel Density Estimation. *Statistica Sinica* 6: 1005–1013.

Karunamuni, R. J. and T. Alberts. 2005. On boundary correction in kernel density estimation. *Statistical Methodology* 2: 191–212.

Salgado-Ugarte, I. H. and M. A. Pérez-Hernández. 2003. Exploring the use of variable bandwidth kernel density estimators. *The Stata Journal* 3(2): 133–147.

Salgado-Ugarte, I. H., M. Shimizu, and T. Taniuchi. 1993. snp6: Exploring the shape of univariate data using kernel density estimators. *Stata Technical Bulletin* 16: 8–19.

—. 1995. snp6.2: Practical rules for bandwidth selection in univariate density estimation. *Stata Technical Bulletin* 26: 23–31.

Scott, D. W. 1992. *Multivariate Density Estimation. Theory, Practice, and Visualization*. New York: Wiley.

Sheather, S. J. and M. C. Jones. 1991. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society. Series B (Methodological)* 53(3): 683–690.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Van Kerm, P. 2003. Adaptive kernel density estimation. *The Stata Journal* 3(2): 148–156.

Wand, M. P. and M. C. Jones. 1995. *Kernel Smoothing*. London: Chapman and Hall.

Table 1: Kernel functions (the kernel functions evaluate to zero if z is outside the indicated support)

Kernel		
Epanechnikov	$K(z) = \frac{3}{4}(1 - \frac{1}{5}z^2)/\sqrt{5}$	if $\|z\| < \sqrt{5}$
Epan2	$K(z) = \frac{3}{4}(1 - z^2)$	if $\|z\| < 1$
Biweight	$K(z) = \frac{15}{16}(1 - z^2)^2$	if $\|z\| < 1$
Triweight	$K(z) = \frac{35}{32}(1 - z^2)^3$	if $\|z\| < 1$
Cosine	$K(z) = 1 + \cos(2\pi z)$	if $\|z\| < \frac{1}{2}$
Gaussian	$K(z) = \phi(z)$	
Parzen	$K(z) = \begin{cases} 8(1-\|z\|)^3/3 \\ \frac{4}{3} - 8z^2 + 8\|z\|^3 \end{cases}$	if $\frac{1}{2} < \|z\| \leq 1$ if $\|z\| \leq \frac{1}{2}$
Rectangular	$K(z) = \frac{1}{2}$	if $\|z\| < 1$
Triangular	$K(z) = 1 - \|z\|$	if $\|z\| < 1$

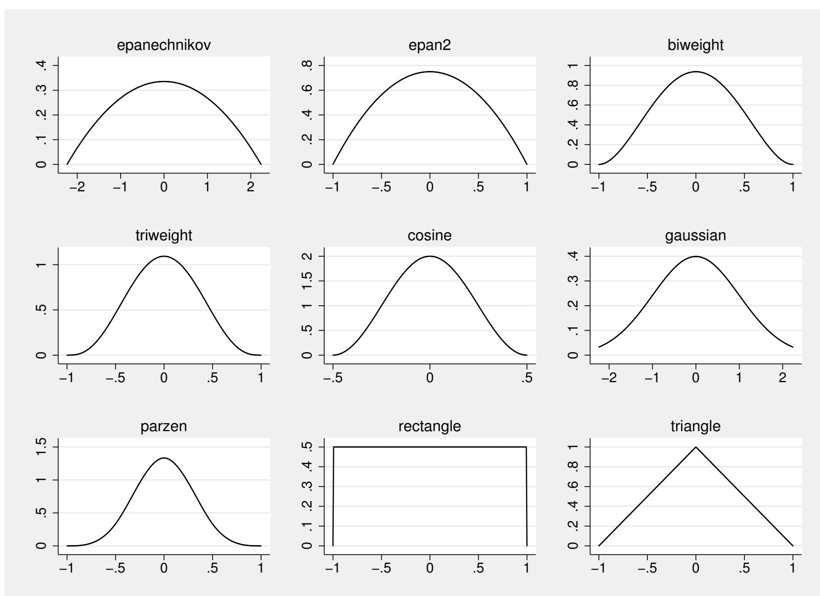

Table 2: Kernel properties

Kernel	$R(K)$	σ_K^2	δ_K	Efficiency
Epanechnikov	$\frac{3}{5\sqrt{5}}$	1	$\left(\frac{3}{5\sqrt{5}}\right)^{\frac{1}{5}}$	1
Epan2	$\frac{3}{5}$	$\frac{1}{5}$	$15^{\frac{1}{5}}$	1
Biweight	$\frac{5}{7}$	$\frac{1}{7}$	$35^{\frac{1}{5}}$	.9939
Triweight	$\frac{350}{429}$	$\frac{1}{9}$	$\left(\frac{9450}{143}\right)^{\frac{1}{5}}$	.9867
Cosine	$\frac{3}{2}$	$\frac{1}{12} - \frac{1}{2\pi^2}$	$\left(\frac{6}{(1/6-1/\pi^2)^2}\right)^{\frac{1}{5}}$	.9897
Gaussian	$\frac{1}{2\sqrt{\pi}}$	1	$\left(\frac{1}{4\pi}\right)^{\frac{1}{10}}$	.9512
Parzen	$\frac{302}{315}$	$\frac{1}{12}$	$2\left(\frac{151}{35}\right)^{\frac{1}{5}}$	.9695
Rectangular	$\frac{1}{2}$	$\frac{1}{3}$	$\left(\frac{9}{2}\right)^{\frac{1}{5}}$	.9295
Triangular	$\frac{2}{3}$	$\frac{1}{6}$	$24^{\frac{1}{5}}$	.9859

Table 3: Kernel integrals (the integrals evaluate to 0 if z is below the kernel support and 1 if above)

Kernel	$\int_{-\infty}^{z} K(y)dy$	
Epanechnikov	$\frac{1}{2} + \frac{3}{4}(z - \frac{1}{15}z^3)/\sqrt{5}$	if $\|z\| < \sqrt{5}$
Epan2	$\frac{1}{2} + \frac{3}{4}(z - \frac{1}{3}z^3)$	if $\|z\| < 1$
Biweight	$\frac{1}{2} + \frac{15}{16}(z - \frac{2}{3}z^3 + \frac{1}{5}z^5)$	if $\|z\| < 1$
Triweight	$\frac{1}{2} + \frac{35}{32}(z - z^3 + \frac{3}{5}z^5 - \frac{1}{7}z^7)$	if $\|z\| < 1$
Cosine	$\frac{1}{2} + z + \frac{\sin(2\pi z)}{2\pi}$	if $\|z\| < \frac{1}{2}$
Gaussian	$\Phi(z)$	
Parzen	$\frac{2}{3} + \frac{8}{3}(z + \frac{3}{2}z^2 + z^3 + \frac{1}{4}z^4)$	if $-1 \leq z < -\frac{1}{2}$
	$\frac{1}{2} + \frac{4}{3}z - \frac{8}{3}z^3 - 2z^4$	if $-\frac{1}{2} \leq z < 0$
	$\frac{1}{2} + \frac{4}{3}z - \frac{8}{3}z^3 + 2z^4$	if $0 \leq z \leq \frac{1}{2}$
	$\frac{1}{3} + \frac{8}{3}(z - \frac{3}{2}z^2 + z^3 - \frac{1}{4}z^4)$	if $\frac{1}{2} < z \leq 1$
Rectangular	$\frac{1}{2} + \frac{1}{2}z$	if $\|z\| < 1$
Triangular	$\frac{1}{2} + z + \frac{1}{2}z^2$	if $-1 < z < 0$
	$\frac{1}{2} + z - \frac{1}{2}z^2$	if $0 \leq z < 1$

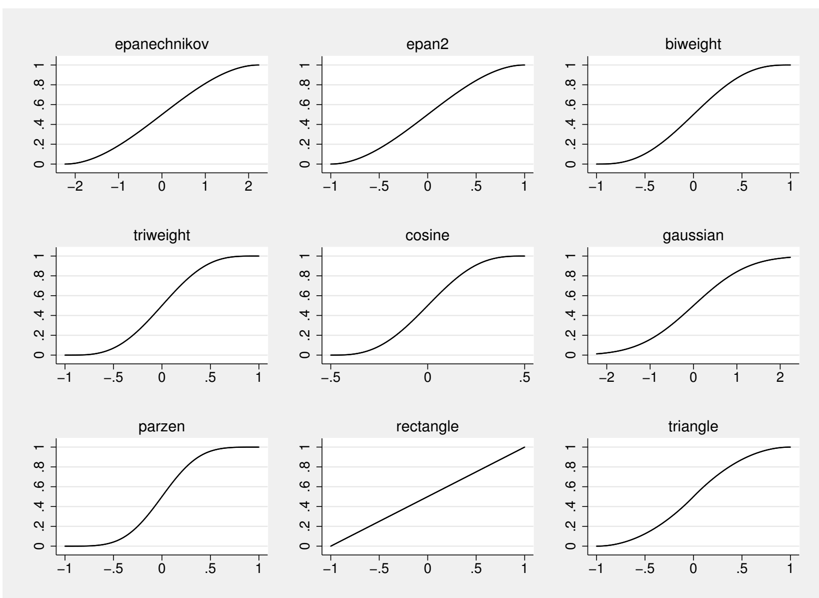

Table 4: Kernel squared integrals (the integrals evaluate to 0 if z is below the kernel support and $R(K)$ if above)

Kernel	$\int_{-\infty}^{z} K(y)^2\,dy$	
Epanechnikov	$\frac{3}{10\sqrt{5}} + \frac{9}{80}(z - \frac{2}{15}z^3 + \frac{1}{125}z^5)$	if $\|z\| < \sqrt{5}$
Epan2	$\frac{3}{10} + \frac{9}{16}(z - \frac{2}{3}z^3 + \frac{1}{5}z^5)$	if $\|z\| < 1$
Biweight	$\frac{5}{14} + \frac{225}{256}(z - \frac{4}{3}z^3 + \frac{6}{5}z^5 - \frac{4}{7}z^7 + \frac{1}{9}z^9)$	if $\|z\| < 1$
Triweight	$\frac{175}{429} + \frac{1225}{1024}(z - 2z^3 + 3z^5 - \frac{20}{7}z^7 + \frac{5}{3}z^9 - \frac{6}{11}z^{11} + \frac{1}{13}z^{13})$	if $\|z\| < 1$
Cosine	$\frac{3}{4} + \frac{3}{2}z + \frac{\sin(2\pi z)}{\pi} + \frac{\cos(2\pi z)\sin(2\pi z)}{4\pi}$	if $\|z\| < \frac{1}{2}$
Gaussian	$\frac{1}{2\sqrt{\pi}}\Phi(\sqrt{2}z)$	
Parzen	$\frac{64}{63} + \frac{64}{9}(z + 3z^2 + 5z^3 + 5z^4 + 3z^5 + z^6 + \frac{1}{7}z^7)$	if $-1 \le z < -\frac{1}{2}$
	$\frac{151}{315} + \frac{16}{9}(z - 4z^3 - 3z^4 + \frac{36}{5}z^5 + 12z^6 + \frac{36}{7}z^7)$	if $-\frac{1}{2} \le z < 0$
	$\frac{151}{315} + \frac{16}{9}(z - 4z^3 + 3z^4 + \frac{36}{5}z^5 - 12z^6 + \frac{36}{7}z^7)$	if $0 \le z \le \frac{1}{2}$
	$-\frac{2}{35} + \frac{64}{9}(z - 3z^2 + 5z^3 - 5z^4 + 3z^5 - z^6 + \frac{1}{7}z^7)$	if $\frac{1}{2} < z \le 1$
Rectangular	$\frac{1}{4} + \frac{1}{4}z$	if $\|z\| < 1$
Triangular	$\frac{1}{3} + z + z^2 + \frac{1}{3}z^3$	if $-1 < z < 0$
	$\frac{1}{3} + z - z^2 + \frac{1}{3}z^3$	if $0 \le z < 1$

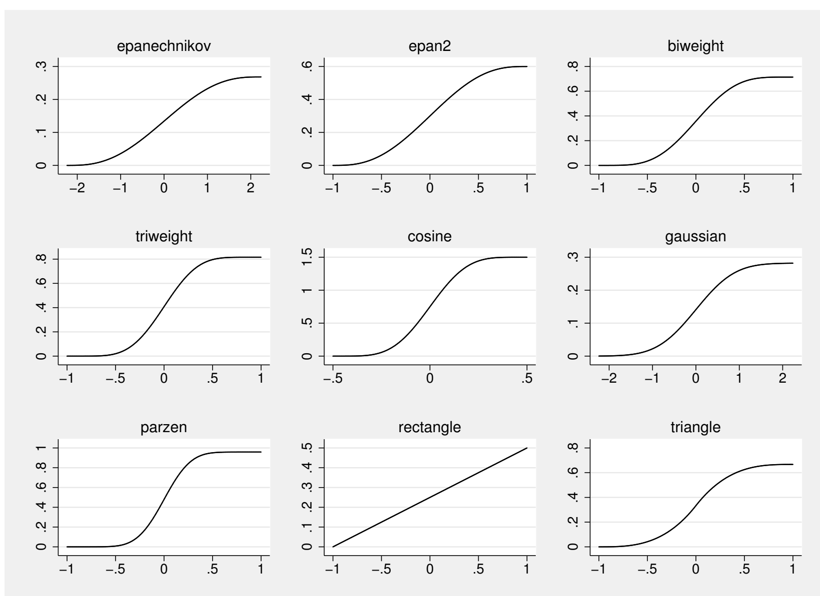

Table 5: Integrals over $yK(y)$ (the integrals evaluate to 0 if z is outside the kernel support)

Kernel	$\int_{-\infty}^{z} yK(y)\,dy$	
Epanechnikov	$-\frac{3\sqrt{5}}{16}+\frac{3}{8\sqrt{5}}z^2-\frac{3}{80\sqrt{5}}z^4$	if $\lvert z\rvert<\sqrt{5}$
Epan2	$-\frac{3}{16}+\frac{3}{8}z^2-\frac{3}{16}z^4$	if $\lvert z\rvert<1$
Biweight	$-\frac{5}{32}+\frac{15}{32}z^2-\frac{15}{32}z^4+\frac{5}{32}z^6$	if $\lvert z\rvert<1$
Triweight	$-\frac{35}{256}+\frac{35}{64}z^2-\frac{105}{128}z^4+\frac{35}{64}z^6-\frac{35}{256}z^8$	if $\lvert z\rvert<1$
Cosine	$-\frac{1}{8}+\frac{1}{2}z^2+\frac{\sin(2\pi z)}{2\pi}z+\frac{\cos(2\pi z)}{4\pi^2}+\frac{1}{4\pi^2}$	if $\lvert z\rvert<\frac{1}{2}$
Gaussian	$-\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}z^2)$	
Parzen	$-\frac{2}{15}+\frac{4}{3}z^2+\frac{8}{3}z^3+2z^4+\frac{8}{15}z^5$	if $-1\le z<-\frac{1}{2}$
	$-\frac{7}{60}+\frac{2}{3}z^2-2z^4-\frac{8}{5}z^5$	if $-\frac{1}{2}\le z<0$
	$-\frac{7}{60}+\frac{2}{3}z^2-2z^4+\frac{8}{5}z^5$	if $0\le z\le\frac{1}{2}$
	$-\frac{2}{15}+\frac{4}{3}z^2-\frac{8}{3}z^3+2z^4-\frac{8}{15}z^5$	if $\frac{1}{2}<z\le 1$
Rectangular	$-\frac{1}{4}+\frac{1}{4}z^2$	if $\lvert z\rvert<1$
Triangular	$-\frac{1}{6}+\frac{1}{2}z^2+\frac{1}{3}z^3$	if $-1<z<0$
	$-\frac{1}{6}+\frac{1}{2}z^2-\frac{1}{3}z^3$	if $0\le z<1$

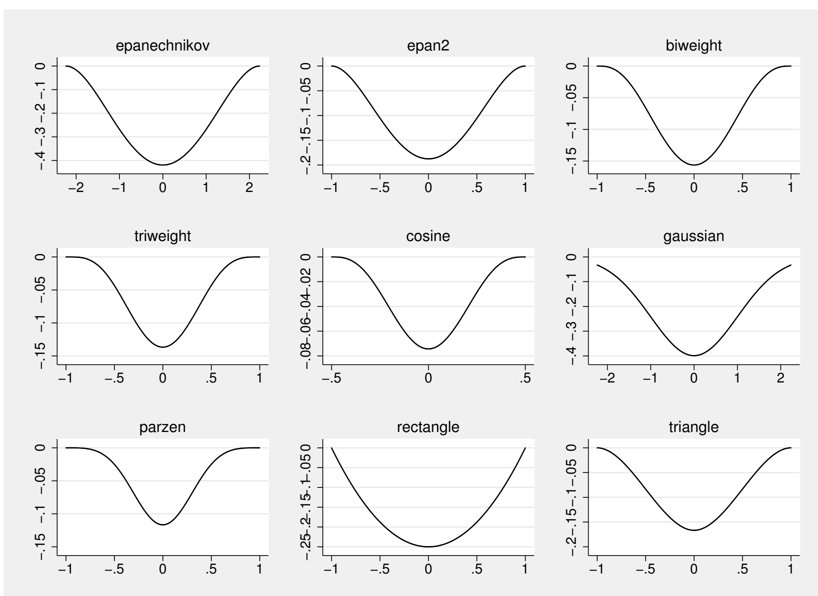

Table 6: Integrals over $y^2K(y)$ (the integrals evaluate to 0 if z is below the kernel support and σ_K^2 if above)

Kernel	$\int_{-\infty}^{z} y^2K(y)\,dy$	
Epanechnikov	$\frac{1}{2}+\frac{1}{4\sqrt{5}}z^3-\frac{3}{100\sqrt{5}}z^5$	if $\lvert z\rvert<\sqrt{5}$
Epan2	$\frac{1}{10}+\frac{1}{4}z^3-\frac{3}{20}z^5$	if $\lvert z\rvert<1$
Biweight	$\frac{1}{14}+\frac{5}{16}z^3-\frac{3}{8}z^5+\frac{15}{112}z^7$	if $\lvert z\rvert<1$
Triweight	$\frac{1}{18}+\frac{35}{96}z^3-\frac{21}{32}z^5+\frac{15}{32}z^7-\frac{35}{288}z^9$	if $\lvert z\rvert<1$
Cosine	$\frac{1}{24}+\frac{1}{3}z^3+\frac{\sin(2\pi z)}{2\pi}z^2+\frac{\cos(2\pi z)}{2\pi^2}z-\frac{\sin(2\pi z)}{4\pi^3}-\frac{1}{4\pi^2}$	if $\lvert z\rvert<\frac{1}{2}$
Gaussian	$-\frac{1}{\sqrt{2\pi}}z\exp(-\frac{1}{2}z^2)+\Phi(z)$	
Parzen	$\frac{2}{45}+\frac{8}{9}z^3+2z^4+\frac{8}{5}z^5+\frac{4}{9}z^6$	if $-1\le z<-\frac{1}{2}$
	$\frac{1}{24}+\frac{4}{9}z^3-\frac{8}{5}z^5-\frac{4}{3}z^6$	if $-\frac{1}{2}\le z<0$
	$\frac{1}{24}+\frac{4}{9}z^3-\frac{8}{5}z^5+\frac{4}{3}z^6$	if $0\le z\le\frac{1}{2}$
	$\frac{7}{180}+\frac{8}{9}z^3-2z^4+\frac{8}{5}z^5-\frac{4}{9}z^6$	if $\frac{1}{2}<z\le 1$
Rectangular	$\frac{1}{6}+\frac{1}{6}z^3$	if $\lvert z\rvert<1$
Triangular	$\frac{1}{12}+\frac{1}{3}z^3+\frac{1}{4}z^4$	if $-1<z<0$
	$\frac{1}{12}+\frac{1}{3}z^3-\frac{1}{4}z^4$	if $0\le z<1$