Economics Department

Central Limit Theorem for Asymmetric Kernel Functionals

MARCELO FERNANDES

ECO No. 2000/1

EUI WORKING PAPERS

EUROPEAN UNIVERSITY INSTITUTE

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 2000/1

Central Limit Theorem for Asymmetric Kernel Functionals

MARCELO FERNANDES

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in March 2000
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

Central Limit Theorem for Asymmetric Kernel Functionals

Marcelo Fernandes

February 2000

Abstract

Asymmetric kernels are quite useful for the estimation of density functions which have bounded support. Gamma kernels are designed to handle density functions whose supports are bounded from one end only, whereas beta kernels are particularly convenient for the estimation of density functions with compact support. These asymmetric kernels are non-negative and free of boundary bias. Moreover, their shape varies according to the location of the data point, thus also changing the amount of smoothing. This paper extends the central limit theorem for degenerate U-statistics in order to compute the limiting distribution of certain asymmetric kernel functionals.

1 Introduction

Fixed kernels are not appropriate to estimate density functions whose supports are bounded in view that they engender boundary bias due to the allocation of weight outside the support in the event that smoothing is applied near the boundary. A proper asymmetric kernel never assigns weight outside the density support and therefore should produce better estimates of the density near the boundary. Indeed, Chen (1999a,b) showed that replacing fixed kernels with asymmetric kernels increases substantially the precision of density estimation close to the boundary. In particular, beta kernels are particularly appropriate to estimate densities with compact support, whereas gamma kernels are more convenient to handle density functions whose supports are bounded from one end only. These asymmetric kernels are non-negative and free of boundary bias. Moreover, their shape varies according to the location of the data point, thus also changing the amount of smoothing.

The aim of this paper is to build on Hall's (1984) central limit theorem for degenerate U-statistics in order to derive asymptotics for asymmetric kernel functionals. The motivation is simple. It is often the case that one must derive the limiting distribution of density functionals such as

$$I = \int_A \varphi(x) \hat{f}^2(x) \mathrm{d}x, \tag{1}$$

where the support A is bounded. Examples abound in econometrics and statistics. Indeed, a central limit theorem for the density functional (1) is useful to study the order of closeness between the integrated square error and the mean integrated squared error in the ambit of non-parametric kernel density estimation. Although there are sharp results for non-parametric density estimation based on fixed kernels (Bickel and Rosenblatt, 1973; Hall, 1984), no results are available for asymmetric kernel density estimation.

Furthermore, goodness-of-fit test statistics are usually driven by second-order asymptotics (e.g. Aït-Sahalia, 1996 and Aït-Sahalia, Bickel and Stoker, 1998), so that density functionals such as (1) arise very naturally in that context. Consider, for instance, one of the goodness-of-fit tests advanced by Fernandes and Grammig (1999) for duration models gauges how large is

$$\Lambda(f,\theta) = \int_0^\infty w(x)[\Gamma_\theta(x) - \Gamma_f(x)]^2 f(x)\mathrm{d}x, \qquad (2)$$

where $w(x)$ is a trimming function and $\Gamma_f(\cdot)$ and $\Gamma_\theta(\cdot)$ denote the non- and parametric hazard rate functions, respectively. It follows from the functional delta method that the asymptotic behaviour of (2) is driven by the leading term of the second functional derivative, namely

$$\int_0^\infty \varphi(x)\left[\hat{f}(x) - f(x)\right]^2 \mathrm{d}x = 2\int_0^\infty w(x)\frac{\Gamma_f(x)}{1-F(x)}\left[\hat{f}(x) - f(x)\right]^2 \mathrm{d}x.$$

Note that duration data are non-negative by definition, hence it is convenient to utilise gamma kernels to avoid boundary bias in the density estimation.

The remainder of this paper is organised as follows. In section 2, I review the properties of beta and gamma kernels. In sections 3 and 4, I apply Hall's (1984) central limit theorem for degenerate U-statistics to derive the limiting distribution of gamma and beta kernel functionals, respectively.

2 Asymmetric kernels

Let $X_1, \ldots, X_T$ be a random sample from an unknown probability density function f defined on a bounded support A. In what follows, I consider that A is either bounded from one end or compact. Without loss of generality, I assume that $A = [0, \infty)$ in the first case, whereas $A = [0, 1]$ in the second context. Finally, assume that the density function f and its second order derivative are bounded and uniformly continuous on the real line.

Instead of the usual non-parametric kernel density estimator

$$\hat{f}(x) = \frac{1}{Th}\sum_{t=1}^{T} K\left(\frac{x - X_t}{h}\right),$$

where K is a fixed kernel function and h is a smoothing bandwidth, consider the asymmetric kernel estimator

$$\hat{f}(x) = \frac{1}{Tb}\sum_{t=1}^{T} K_A(X_t), \tag{3}$$

where $K_A(\cdot)$ corresponds either to the gamma kernel

$$K_{x/b+1,b}(u) = \frac{u^{x/b}\exp(-u/b)}{\Gamma(x/b+1)b^{x/b}} I\{u \in [0,\infty)\} \tag{4}$$

or to the beta kernel

$$K_{x/b+1,(1-x)/b+1}(u) = \frac{u^{x/b}(1-u)^{(1-x)/b}}{B(x/b+1,(1-x)/b+1)} I\{u \in [0,1]\} \tag{5}$$

according to the support under consideration.

Chen (1999a,b) showed that both estimators are boundary bias free in view that the bias is of order $O(b)$ both near the boundaries and in the interior of the support. The absence of boundary bias is due to the fact that asymmetric kernels have the same support of the underlying density, and hence no weight is assigned outside the density support. The trick is that asymmetric kernel functions are flexible enough to vary their shape (and thus the amount of smoothing) according to the location of x within the support.

On the other hand, the asymptotic variance of asymmetric kernels is of higher order $O\left(T^{-1}b^{-1}\right)$ near the boundaries than in the interior, which is of order $O\left(T^{-1}b^{-1/2}\right)$. Nonetheless, this has negligible impact on the integrated variance, thus it does affect the mean integrated square error. Furthermore, it is possible to show that the optimal bandwidth $b_* = O\left(T^{-2/5}\right) = O\left(h_*^2\right)$, where h_* is the optimal bandwidth for fixed kernel estimators. Accordingly, both beta and gamma kernel density estimators achieve the optimal rate of convergence for the mean integrated squared error of non-negative kernels.[1] Lastly, a unique feature for the gamma kernel estimator is that its variance decreases as x increases, though at the expense of an upping in the bias.

3 Gamma kernel functionals

The asymptotic behaviour of gamma kernel functionals of the form (1) is derived using U-statistic theory. For this reason, I utilise a decomposition which forces a degenerate U-statistic to emerge. Let $r_T(x,X) = \varphi^{1/2}(x)K_{x/b+1,b}(X_t)$, $\check{r}_T(x,X) = r_T(x,X) - E_X[r_T(x,X)]$ and $\int_u$ denote the integral over the support of u. Then,

$$\begin{aligned} I &= \int_x \left[\sum_{t=1}^{T} r_T(x,X_t)/T\right]^2 \mathrm{d}x \\ &= \frac{1}{T^2}\sum_{s,t}\int_x r_T(x,X_t)r_T(x,X_s)\mathrm{d}x \\ &= I_1 + I_2 + I_3 + I_4, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \frac{2}{T^2}\sum_{s<t}\int_x \check{r}_T(x,X_t)\check{r}_T(x,X_s)\mathrm{d}x \\ I_2 &= \frac{1}{T^2}\sum_{t}\int_x r_T^2(x,X_t)\mathrm{d}x \\ I_3 &= \frac{T(T-1)}{T^2}\int_x E_X^2[r_T(x,X)]\mathrm{d}x \\ I_4 &= \frac{2T(T-1)}{T^2}\int_x \check{r}_T(x,X_t)E_X[r_T(x,X)]\mathrm{d}x. \end{aligned}$$

The first term stands for a degenerate U-statistic and will contribute with the variance in the limiting distribution. The second term will

[1] Non-negative kernels define the class of second order kernel functions. Higher order kernels may bring about some bias reduction at the expense of assuming negative values (see Müller, 1984, for a list).

contribute with the asymptotic mean, hence it may be interpreted as some sort of asymptotic bias of the functional estimator depending on the context. The third and the fourth terms are, in turn, negligible under a proper choice of the bandwidth b. Suppose the bandwidth b is such that $Tb^{9/4}$ shrinks to zero as sample size T grows. This assumption implies some degree of undersmoothing in view that Chen (1999b) has shown that the optimal bandwidth for gamma kernels is of order $O\left(T^{-2/5}\right)$.

I start by deriving the first two moments of $r_T(x, X)$. Note that

$$E_X[r_T(x,X)] = \varphi^{1/2}(x)\int_X K_{x/b+1,b}(X)f(X)\mathrm{d}X = \varphi^{1/2}(x)E_\zeta[f(\zeta)],$$

where $\zeta \sim \mathcal{G}(x/b+1, b)$. The mean and variance of a $\mathcal{G}(\mu, v)$ are simply μv and μv^2, respectively. Therefore, applying a Taylor expansion yields

$$\begin{aligned} E_\zeta[f(\zeta)] &= f(E_\zeta) + \frac{1}{2}f''(x)V_\zeta + o(b) \\ &= f(x+b) + \frac{1}{2}f''(x)(x+b)b + o(b) \\ &= f(x) + b\left[f'(x) + \frac{1}{2}f''(x)\right] + o(b). \end{aligned}$$

It is noteworthy that the last expression demonstrates that the gamma kernel estimation of the density function has a uniform bias of order $O(b)$. Put differently, the order of magnitude of the bias does not depend on the position of x, that is, whether it is close to the origin or in the interior of the support. To sum up, $E_X[r_T(x,X)] = \varphi^{1/2}(x)f(x) + O(b)$, which implies that $\breve{r}_T(x,X) = O(b)$.

The second moment of $r_T(x, X)$ is computed in similar way. It follows from Chen's (1999b) derivation of the variance of the gamma kernel estimator that

$$\begin{aligned} E_X\left[r_T^2(x,X)\right] &= \varphi(x)\int_X K^2_{x/b+1,b}(X)f(X)\mathrm{d}X \\ &= \varphi(x)B_b(x)E_\eta[f(\eta)], \end{aligned}$$

where

$$B_b(x) = \frac{\Gamma(2x/b+1)/b}{2^{2x/b+1}\Gamma^2(x/b+1)}$$

and $\eta \sim \mathcal{G}(2x/b+1, b)$. Hence applying a Taylor expansion yields

$$\begin{aligned} E_\eta[f(\eta)] &= f(E_\eta) + \frac{1}{2} f''(x) V_\eta + o(b) \\ &= f(2x+b) + \frac{1}{2} f''(x)(2x+b)b + o(b) \\ &= f(x) + f'(x)x + b\left[f'(x) + f''(x)x\right] + o(b) \\ &= f(x) + f'(x)x + O(b). \end{aligned}$$

It follows then that

$$\begin{aligned} E(I_2) &= \frac{1}{T}\int_x E_X\left[r_T^2(x,X)\right] \mathrm{d}x \\ &= \frac{1}{T}\int_x \varphi(x) B_b(x)[f(x) + f'(x)x + O(b)]\mathrm{d}x \\ &= \frac{1}{T}\int_x \varphi(x) B_b(x) f(x)\mathrm{d}x + O(1/T). \end{aligned}$$

For b small enough, Chen (1999b) approximates $B_b(x)$ according to the behaviour of x/b. The motivation stems from the fact that, in the interior of the support, x/b grows without bound as b shrinks to zero, whereas x/b converges to some non-negative constant c in the boundary. The decomposition dictates that

$$B_b(x) \sim \begin{cases} \frac{1}{2\sqrt{\pi}} b^{-1/2} x^{-1/2} & \text{if } x/b \to \infty \\ \\ \frac{\Gamma(2c+1)/b}{2^{2c+1}\Gamma^2(c+1)} & \text{if } x/b \to c, \end{cases}$$

which implies that $B_b(x)$ is higher near the origin. Nonetheless, I show that there is no impact whatsoever in $E(I_2)$.[2]

Let $\delta = b^{1-\epsilon}$, where $0 < \epsilon < 1$. Then,

$$E(I_2) = \frac{1}{T}\int_x \varphi(x) B_b(x) f(x)\mathrm{d}x + O(1/T)$$

[2] This result is analogous to Chen's (1999b) result concerning the variance of the gamma kernel estimator. In particular, the variance mounts as x approaches the boundary, but this increase does not affect the integrated variance of the estimator.

$$
\begin{aligned}
&= \frac{1}{T}\int_0^\delta + \int_\delta^\infty \varphi(x)B_b(x)f(x)\mathrm{d}x + O(1/T) \\
&= \frac{1}{2\sqrt{\pi}T}\int_\delta^\infty b^{-1/2}x^{-1/2}\varphi(x)f(x)\mathrm{d}x + O\left(T^{-1}b^{-\epsilon}\right) \\
&= \frac{b^{-1/2}}{2\sqrt{\pi}T}\int_x \varphi(x)x^{-1/2}f(x)\mathrm{d}x + o\left(T^{-1}b^{-1/2}\right)
\end{aligned}
$$

provided that ϵ is properly chosen and $E\left[\varphi(x)x^{-1/2}\right]$ is finite. Therefore, it ensues that

$$
Tb^{1/4}E(I_2) = \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[x^{-1/2}\varphi(x)\right].
$$

Notice also that

$$
\begin{aligned}
V(I_2) &= \frac{1}{T^3}E\left[\int_x r_T^2(x,X)\mathrm{d}x\right]^2 - \frac{1}{T^3}E^2\left[\int_x r_T^2(x,X)\mathrm{d}x\right] \\
&= \frac{1}{T^3}E\left[\int_x r_T^2(x,X)\mathrm{d}x\right]^2 - \frac{1}{T^3}\left[\int_x E r_T^2(x,X)\mathrm{d}x\right]^2 \\
&= O\left(T^{-3}b^{-1}\right).
\end{aligned}
$$

Thus, $V(Tb^{1/4}I_2) = T^2b^{1/2}V(I_2) = O\left(T^{-1}b^{-1/2}\right)$, which is of order $o(1)$ given the assumption on the bandwidth. Thus, by Chebyshev's inequality,

$$
Tb^{1/4}I_2 - \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[x^{-1/2}\varphi(x)\right] = o_p(1).
$$

The fact that $b = o\left(T^{-4/9}\right)$ also ensures that the third and fourth terms are negligible if properly normalised. Indeed, it follows from

$$
I_3 = \frac{T-1}{T}\int_x E^2[r_T(x,X)]\mathrm{d}x = \frac{T-1}{T}O\left(b^2\right) = O\left(b^2\right)
$$

that $Tb^{1/4}I_3 = O\left(Tb^{9/4}\right)$, which is $o(1)$ by assumption. Furthermore,

$$
I_4 = \frac{2(T-1)}{T^2}\sum_{t=1}^{T}\int_x \breve{r}_T(x,X_t)E_X r_T(x,X)\mathrm{d}x
$$

and hence

$$E(I_4) = \frac{2(T-1)}{T} \int_x E_X[\check{r}_T(x,X)] E_X[r_T(x,X)] \mathrm{d}x = 0$$

given that $\check{r}_T(x,X)$ has zero mean. Besides,

$$E_X \left\{ \int_x \check{r}_T(x,X_t) E_X[r_T(x,X)] \mathrm{d}x \right\}^2 = O\left(b^2\right),$$

which implies that $E\left(I_4^2\right) = O\left(T^{-1}b^2\right)$ and therefore

$$E\left(Tb^{1/4}I_4\right)^2 = T^2 b^{1/2} E\left(I_4^2\right) = O\left(Tb^{5/2}\right) = o(1).$$

Afresh, it stems from Chebyshev's inequality that $Tb^{1/4}I_4 = o_p(1)$.

Finally, recall that $I_1 = \sum_{s<t} H_T(X_t, X_s)$, where

$$H_T(X_t, X_s) = \frac{2}{T^2} \int_x \check{r}_T(x, X_t) \check{r}_T(x, X_s) \mathrm{d}x.$$

Then, I_1 is a degenerate U-statistic in view that $H_T(X_t, X_s)$ is symmetric, centred, and $E\left[H_T(X_t, X_s)|X_s\right] = 0$ almost surely. To see why, note that

$$\begin{aligned} E\left[H_T(X_t, X_s)|X_s\right] &= \frac{2}{T^2} \int_x \check{r}_T(x, X_s) E\left[\check{r}_T(x, X_t)|X_s\right] \mathrm{d}x \\ &= \frac{2}{T^2} \int_x \check{r}_T(x, X_s) E\left[\check{r}_T(x, X_t)\right] \mathrm{d}x \end{aligned}$$

in view of the independence between X_t and X_s. It suffices then to observe that $\check{r}_T(x, X_t)$ has by construction zero mean. Thereby, I apply Hall's (1984) central limit theorem for degenerate U-statistics, which states that if

$$\frac{E_{X_1,X_2}\left\{E^2_{X_1}\left[H_T(X_1,X_1)H_T(X_1,X_2)\right]\right\} + \frac{1}{T} E_{X_1,X_2}\left[H_T^4(X_1,X_2)\right]}{E^2_{X_1,X_2}\left[H_T^2(X_1,X_2)\right]} \to 0 \quad (6)$$

as sample size grows, then

$$I_1 \xrightarrow{d} N\left(0, \frac{T^2}{2} E_{X_1,X_2}\left[H_T^2(X_1,X_2)\right]\right).$$

Tedious algebra shows that (6) holds. Indeed, the two terms of the numerator are of order $O(T^{-12}b^{-2})$ and $O(T^{-9}b^{-3/2})$, respectively, whereas the denominator is of order $O(T^{-8}b^{-1})$. In what follows, I demonstrate the last assertion as a by-product of the derivation of the asymptotic variance above.

Let $V_H = \frac{T^4}{2}E_{X_1,X_2}\left[H_T^2(X_1,X_2)\right]$, then

$$\begin{aligned}
V_H &= 2\int_{X_1,X_2}\left[\int_x \check{r}_T(x,X_1)\check{r}_T(x,X_2)\mathrm{d}x\right]^2 f(X_1,X_2)\mathrm{d}(X_1,X_2) \\
&= 2\int_{x,y}\left[\int_X \check{r}_T(x,X)\check{r}_T(y,X)f(X)\mathrm{d}X\right]^2 \mathrm{d}(x,y) \\
&= 2\int_{x,y}\varphi(x)\varphi(y)E_X^2\left\{\left[K_{x/b+1,b}(X)-E_{K(x,b)}\right]\right. \\
&\qquad\qquad \times\left.\left[K_{y/b+1,b}(X)-E_{K(y,b)}\right]\right\}\mathrm{d}(x,y),
\end{aligned}$$

where $E_{K(u,b)} = E_X\left[K_{u/b+1,b}(X)\right]$. Then, it ensues that

$$V_H = 2\int_{x,y}\varphi(x)\varphi(y)\left[\int_X K_{x/b+1,b}(X)K_{y/b+1,b}(X)\mathrm{d}F(X)\right]^2\mathrm{d}(x,y)+O(b^2)$$

due to the fact that

$$\begin{aligned}
E_{K(x,b)}E_{K(y,b)} &= \int_X K_{x/b+1,b}(X)E_{K(y,b)}\mathrm{d}F(X) \\
&= \int_X E_{K(x,b)}K_{y/b+1,b}(X)\mathrm{d}F(X) \\
&= \int_X E_{K(x,b)}E_{K(y,b)}\mathrm{d}F(X) \\
&= O\left(b^2\right).
\end{aligned}$$

Let $g(X) = f(X)K_{x/b+1,b}(X)$, then

$$V_H = 2\int_{x,y}\varphi(x)\varphi(y)\left[\int_X g(X)\mathrm{d}K_{y/b+1,b}(X)\right]^2\mathrm{d}(x,y)+O(b^2).$$

Applying a Taylor expansion gives

$$\begin{aligned}
\int_X g(X)\mathrm{d}K_{y/b+1,b}(X) &= E_{\mathcal{G}(y/b+1,b)}[g(X)] \\
&= g\left[E_{\mathcal{G}(y/b+1,b)}(X)\right] + \frac{g''(y)V_{\mathcal{G}(y/b+1,b)}(X)}{2} + o(b) \\
&= g(y+b) + \frac{1}{2}g''(y)(y+b)b + o(b) \\
&= g(y) + b\left[g'(y) + \frac{1}{2}g''(y)y\right] + o(b) \\
&= g(y) + O(b).
\end{aligned}$$

This means that

$$\begin{aligned}
V_H &= 2\int_{x,y}\varphi(x)\varphi(y)\left[f(y)K_{x/b+1,b}(y)\right]^2 \mathrm{d}(x,y) + O(b^2) \\
&= 2\int_x \varphi(x)\left[\int_y \varphi(y)f^2(y)K^2_{x/b+1,b}(y)\mathrm{d}y\right]\mathrm{d}x + O(b^2) \\
&= 2\int_x \varphi(x)\left[\int_y h(y)\mathrm{d}K_{x/b+1,b}(y)\right]\mathrm{d}x + O(b^2),
\end{aligned}$$

where $h(y) = \varphi(y)f^2(y)K_{x/b+1,b}(y)$. Afresh, by Taylor expanding, it yields

$$\begin{aligned}
\int_y h(y)\mathrm{d}K_{x/b+1,b}(y) &= E_{\mathcal{G}(x/b+1,b)}[h(y)] \\
&= h\left[E_{\mathcal{G}(x/b+1,b)}(y)\right] + \frac{1}{2}h''(x)V_{\mathcal{G}(x/b+1,b)}(y) + o(b) \\
&= h(x+b) + \frac{1}{2}h''(x)(x+b)b + o(b) \\
&= h(x) + b\left[h'(x) + \frac{1}{2}h''(x)x\right] + o(b) \\
&= h(x) + O(b).
\end{aligned}$$

Therefore,

$$\begin{aligned}
V_H &= 2\int_x \varphi(x)h(x)\mathrm{d}x + O(b) \\
&= 2\int_x \varphi^2(x)f^2(x)K_{x/b+1,b}(x)\mathrm{d}x + O(b) \\
&= 2\int_x \varphi^2(x)f(x)K_{x/b+1,b}(x)\mathrm{d}F(x) + O(b).
\end{aligned}$$

Notice however that using the same technique it is possible to show that

$$\int_X f(X)K^2_{x/b+1,b}(X)\mathrm{d}X = K_{x/b+1,b}(x)f(x) + O(b).$$

Hence, it follows that

$$\begin{aligned} V_H &= 2\int_x \varphi^2(x)f(x)K_{x/b+1,b}(x)\mathrm{d}F(x) + O(b) \\ &= 2\int_x \varphi^2(x)\left[\int_X f(X)K^2_{x/b+1,b}(X)\mathrm{d}X\right]\mathrm{d}F(x) + O(b) \\ &= 2\int_x \varphi^2(x)B_b(x)[f(x) + O(b)]\mathrm{d}F(x) + O(b) \\ &= 2\int_x \varphi^2(x)B_b(x)f(x)\mathrm{d}F(x) + O(b). \end{aligned}$$

By decomposing the integral according to $\delta = b^{1-\epsilon}$, it yields

$$\begin{aligned} V_H &= \int_0^\delta + \int_\delta^\infty 2\varphi^2(x)B_b(x)f(x)\mathrm{d}F(x) + O(b) \\ &= \frac{b^{-1/2}}{\sqrt{\pi}}\int_\delta^\infty \varphi^2(x)x^{-1/2}f(x)\mathrm{d}F(x) + O(b^{-\epsilon}) \\ &= \frac{b^{-1/2}}{\sqrt{\pi}}\int_x \varphi^2(x)x^{-1/2}f(x)\mathrm{d}F(x) + O(b^{-1/2}) \end{aligned}$$

for a properly chosen ϵ and finite $E\left[\varphi^2(x)x^{-1/2}\right]$. Finally, this implies that $E^2_{X_1,X_2}\left[H^2_T(X_1,X_2)\right] = O(T^{-8}b^{-1})$ and that

$$Tb^{1/4}I - \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[x^{-1/2}\varphi(x)\right] \xrightarrow{d} N\left(0, \frac{1}{\sqrt{\pi}}E\left[\varphi^2(x)x^{-1/2}f(x)\right]\right). \tag{7}$$

4 Beta kernel functionals

I derive the asymptotic behaviour of beta kernel functionals using the same approach as before, that is, I consider the decomposition $I = I_1 + I_2 + I_3 + I_4$. The only difference is that $r_T(x, X)$ represents now $\varphi^{1/2}(x)K_{x/b+1,(1-x)/b+1}(X_t)$. Again, the first term stands for a degenerate U-statistic and contributes with the asymptotic variance, whereas the second term provides the asymptotic mean. The third and the fourth terms are, afresh, negligible under proper normalisation provided that the bandwidth b is of order $o\left(T^{-4/9}\right)$. Once more, this assumption implies some degree of undersmoothing in view that Chen (1999a) has shown that the optimal bandwidth for beta kernels is of order $O\left(T^{-2/5}\right)$.

The limiting distribution of beta kernel functionals is perfectly analogous to that derived for gamma kernels. The only distinction stems from the consideration of the upper bound, which engender a correction inversely proportional to the square root of $x(1-x)$ instead of x. More precisely, I show in the sequel that

$$Tb^{1/4}I - \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[\frac{\varphi(x)}{\sqrt{x(1-x)}}\right] \xrightarrow{d} N\left(0, \frac{1}{\sqrt{\pi}}E_x\left[\frac{\varphi^2(x)f(x)}{\sqrt{x(1-x)}}\right]\right). \tag{8}$$

I start by noting that the expectation and variance of a $\mathcal{B}(\mu, v)$ are $v/(\mu+v)$ and $\mu v/[(\mu+v)^2(\mu+v+1)]$, respectively. It is then straightforward to derive the first two moments of $r_T(x, X)$. Indeed,

$$\begin{aligned} E_X[r_T(x,X)] &= \varphi^{1/2}(x)\int_X K_{x/b+1,(1-x)/b+1}(X)f(X)\mathrm{d}X \\ &= \varphi^{1/2}(x)E_\zeta[f(\zeta)], \end{aligned}$$

where $\zeta \sim \mathcal{B}(x/b+1, (1-x)/b+1)$. Therefore, the mean and variance of ζ are

$$\begin{aligned} E_\zeta &= \frac{(1-x)/b+1}{x/b+1+(1-x)/b+1} = \frac{1-x+b}{1+2b} \\ V_\zeta &= \frac{(x/b+1)[(1-x)/b+1]}{(1/b+2)^2(1/b+3)} = x(1-x)b + O\left(b^2\right), \end{aligned}$$

respectively. Applying a Taylor expansion yields

$$\begin{aligned} E_\zeta[f(\zeta)] &= f(E_\zeta) + \frac{1}{2}f''(x)V_\zeta + o(b) \\ &= f\left(\frac{1-x+b}{1+2b}\right) + \frac{1}{2}f''(x)x(1-x)b + o(b) \end{aligned}$$

$$\begin{aligned}
&= f(x) + f'(x)\frac{1-2x+b-2bx}{1+2b} + \frac{1}{2}f''(x)x(1-x)b + o(b) \\
&= f(x) + f'(x)(1-2x)\frac{1+b}{1+2b} + \frac{1}{2}f''(x)x(1-x)b + o(b) \\
&= f(x) + \left[f'(x)(1-2x)\frac{1}{2}f''(x)x(1-x)\right]b + o(b) \\
&= f(x) + O(b),
\end{aligned}$$

which implies that the beta kernel estimation of the density function has a uniform bias of order $O(b)$. To sum up,

$$E_X[r_T(x,X)] = \varphi^{1/2}(x)f(x) + O(b),$$

which implies that $\breve{r}_T(x,X) = O(b)$.

Now I turn to the second moment of $r_T(x,X)$, namely

$$\begin{aligned}
E_X\left[r_T^2(x,X)\right] &= \varphi(x)\int_X K^2_{x/b+1,(1-x)/b+1}(X)f(X)\mathrm{d}X \\
&= \varphi(x)A_b(x)E_\eta[f(\eta)],
\end{aligned}$$

where

$$A_b(x) = \frac{B[2x/b+1, 2(1-x)/b+1]}{B^2[x/b+1,(1-x)/b+1]}$$

and $\eta \sim \mathcal{B}(2x/b+1, 2(1-x)/b+1)$. The mean and variance of η are

$$\begin{aligned}
E_\eta &= \frac{2(1-x)/b+1}{2x/b+1+2(1-x)/b+1} = \frac{2(1-x)+b}{2(1+b)} \\
V_\eta &= \frac{(2x/b+1)[2(1-x)/b+1]}{(2/b+2)^2(2/b+3)} = \frac{1}{2}x(1-x)b + O\left(b^2\right)
\end{aligned}$$

respectively, hence applying a Taylor expansion yields

$$\begin{aligned}
E_\eta[f(\eta)] &= f(E_\eta) + \frac{1}{2}f''(x)V_\eta + o(b) \\
&= f\left(\frac{(2(1-x)+b}{2(1+b)}\right) + \frac{1}{4}f''(x)x(1-x)b + o(b)
\end{aligned}$$

$$
\begin{aligned}
&= f(x) + f'(x)(1-2x)\frac{2+b}{2(1+b)} + \frac{1}{4}f''(x)x(1-x)b + o(b) \\
&= f(x) + \frac{1}{2}\left[f'(x)(1-2x) + \frac{1}{2}f''(x)x(1-x)\right]b + o(b) \\
&= f(x) + O(b).
\end{aligned}
$$

Then, it follows that

$$
\begin{aligned}
E(I_2) &= \frac{1}{T}\int_x E_X\left[r_T^2(x,X)\right]\mathrm{d}x \\
&= \frac{1}{T}\int_x \varphi(x)A_b(x)[f(x)+O(b)]\mathrm{d}x \\
&= \frac{1}{T}\int_x \varphi(x)A_b(x)f(x)]\mathrm{d}x + O(1/T).
\end{aligned}
$$

For b small enough, Chen (1999a) showed that $A_b(x)$ may be approximated according to the location of x within the support. More precisely, x/b and $(1-x)/b$ grows without bound as b shrinks to zero in the interior of the support, whereas either x/b or $(1-x)/b$ converges to some non-negative constant c in the boundaries. The approximation is such that

$$
A_b(x) \sim \begin{cases} \frac{1}{2\sqrt{\pi}}b^{-1/2}[x(1-x)]^{-1/2} & \text{if } x/b \text{ and } (1-x)/b \to \infty \\ \frac{\Gamma(2c+1)/b}{2^{2c+1}\Gamma^2(c+1)} & \text{if } x/b \text{ or } (1-x)/b \to c, \end{cases}
$$

which implies that $A_b(x)$ is of larger order near the boundary. Nonetheless, I show that there is no impact whatsoever in $E(I_2)$.[3]

Let $\delta = b^{1-\epsilon}$, where $0 < \epsilon < 1$. Then,

$$
\begin{aligned}
E(I_2) &= \frac{1}{T}\int_x \varphi(x)A_b(x)f(x)\mathrm{d}x + O(1/T) \\
&= \frac{1}{T}\int_0^\delta + \int_\delta^{1-\delta} + \int_{1-\delta}^1 \varphi(x)A_b(x)f(x)\mathrm{d}x + O(1/T)
\end{aligned}
$$

[3] This result is analogous to Chen's (1999a) result concerning the variance of the beta kernel estimator. In particular, the variance mounts as x approaches the boundary, but this increase does not affect the integrated variance of the estimator.

$$
\begin{aligned}
&= \frac{1}{2\sqrt{\pi}T}\int_{\delta}^{1-\delta} b^{-1/2}[x(1-x)]^{-1/2}\varphi(x)f(x)\mathrm{d}x + O\left(T^{-1}b^{-\epsilon}\right) \\
&= \frac{b^{-1/2}}{2\sqrt{\pi}T}\int_{0}^{1} \varphi(x)[x(1-x)]^{-1/2}f(x)\mathrm{d}x + o\left(T^{-1}b^{-1/2}\right)
\end{aligned}
$$

as long as ϵ is properly chosen and $E\left[\varphi(x)/\sqrt{x(1-x)}\right]$ is finite. Therefore, it ensues that

$$
Tb^{1/4}E(I_2) = \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[\frac{\varphi(x)}{\sqrt{x(1-x)}}\right].
$$

Notice also that

$$
\begin{aligned}
V(I_2) &= \frac{1}{T^3}E\left[\int_x r_T^2(x,X)\mathrm{d}x\right]^2 - \frac{1}{T^3}E^2\left[\int_x r_T^2(x,X)\mathrm{d}x\right] \\
&= \frac{1}{T^3}E\left[\int_x r_T^2(x,X)\mathrm{d}x\right]^2 - \frac{1}{T^3}\left[\int_x Er_T^2(x,X)\mathrm{d}x\right]^2 \\
&= O\left(T^{-3}b^{-1}\right).
\end{aligned}
$$

Thus, $V(Tb^{1/4}I_2) = T^2b^{1/2}V(I_2) = O\left(T^{-1}b^{-1/2}\right)$, which is of order $o(1)$ given the assumption on the bandwidth. Thus, by Chebyshev's inequality,

$$
Tb^{1/4}I_2 - \frac{b^{-1/4}}{2\sqrt{\pi}}E\left[\frac{\varphi(x)}{\sqrt{x(1-x)}}\right] = o_p(1).
$$

Applying exactly the same techniques used in the gamma context, it is straightforward to demonstrate that the third and fourth terms are negligible under proper normalisation. Indeed, the fact that the bandwidth is such that $b = o\left(T^{-4/9}\right)$ suffices to guarantee that $Tb^{1/4}I_3 = o(1)$ and $Tb^{1/4}I_4 = o_p(1)$. Lastly, it is evident given the previous discussion that $I_1 = \sum_{s<t} H_T(X_t, X_s)$, where

$$
H_T(X_t, X_s) = \frac{2}{T^2}\int_x \check{r}_T(x, X_t)\check{r}_T(x, X_s)\mathrm{d}x,
$$

is a degenerate U-statistic. Let $V_H = \frac{T^4}{2} E_{X_1,X_2}\left[H_T^2(X_1,X_2)\right]$, then

$$\begin{aligned} V_H &= 2\int_{X_1,X_2}\left[\int_x \breve{r}_T(x,X_1)\breve{r}_T(x,X_2)\mathrm{d}x\right]^2 f(X_1,X_2)\mathrm{d}(X_1,X_2) \\ &= 2\int_{x,y}\left[\int_X \breve{r}_T(x,X)\breve{r}_T(y,X)f(X)\mathrm{d}X\right]^2 \mathrm{d}(x,y) \\ &= 2\int_{x,y}\varphi(x)\varphi(y)E_X^2\left\{\left[K_{x/b+1,(1-x)/b+1}(X)-E_{K(x,b)}\right]\right. \\ &\qquad\qquad \left.\times\left[K_{y/b+1,(1-y)/b+1}(X)-E_{K(y,b)}\right]\right\}\mathrm{d}(x,y), \end{aligned}$$

where $E_{K(u,b)} = E_X\left[K_{u/b+1,(1-u)/b+1}(X)\right]$. As before, it turns out that

$$V_H \simeq 2\int_{x,y}\varphi(x)\varphi(y)\left[\int_X K_{\frac{x}{b}+1,\frac{1-x}{b}+1}(X)K_{\frac{y}{b}+1,\frac{1-y}{b}+1}(X)\mathrm{d}F(X)\right]^2 \mathrm{d}(x,y)$$

due to the fact that all other terms are of order $O\left(b^2\right)$.

Let $g(X) = f(X)K_{x/b+1,(1-x)/b+1}(X)$ and write

$$V_H \simeq 2\int_{x,y}\varphi(x)\varphi(y)\left[\int_X g(X)\mathrm{d}K_{y/b+1,(1-y)/b+1}(X)\right]^2 \mathrm{d}(x,y).$$

It follows from a Taylor expansion that

$$\begin{aligned} &\int_X g(X)\,\mathrm{d}K_{y/b+1,(1-y)/b+1}(X) \\ &\quad= E_{\mathcal{B}(y/b+1,(1-y)/b+1)}[g(X)] \\ &\quad= g\left[E_{\mathcal{B}(y/b+1,(1-y)/b+1)}(X)\right] + \frac{1}{2}g''(y)V_{\mathcal{B}(y/b+1,(1-y)/b+1)}(X) + o(b) \\ &\quad= g\left(\frac{1-y+b}{1+2b}\right) + \frac{g''(y)y(1-y)b}{2} + o(b) \\ &\quad= g(y) + O(b), \end{aligned}$$

which implies that

$$\begin{aligned} V_H &\simeq 2\int_{x,y}\varphi(x)\varphi(y)\left[f(y)K_{x/b+1,(1-x)/b+1}(y)\right]^2 \mathrm{d}(x,y) \\ &\simeq 2\int_x\varphi(x)\int_y\varphi(y)f^2(y)K^2_{x/b+1,(1-x)/b+1}(y)\mathrm{d}y\mathrm{d}x \\ &\simeq 2\int_x\varphi(x)\int_y h(y)\mathrm{d}K_{x/b+1,(1-x)/b+1}(y)\mathrm{d}x, \end{aligned}$$

where $h(y) = \varphi(y)f^2(y)K_{x/b+1,(1-x)/b+1}(y)$. Applying another Taylor expansion gives forth that

$$\begin{aligned}
&\int_y h(y)\mathrm{d}K_{x/b+1,(1-x)/b+1}(y) \\
&\quad = E_{\mathcal{B}(x/b+1,(1-x)/b+1)}[h(y)] \\
&\quad = h\left[E_{\mathcal{B}(x/b+1,(1-x)/b+1)}(y)\right] + \frac{1}{2}h''(x)V_{\mathcal{B}(x/b+1,(1-x)/b+1)}(y) + o(b) \\
&\quad = h\left(\frac{1-x+b}{1+2b}\right) + \frac{1}{2}h''(x)x(1-x)b + o(b) \\
&\quad = h(x) + O(b).
\end{aligned}$$

Therefore,

$$\begin{aligned}
V_H &\simeq 2\int_x \varphi(x)h(x)\mathrm{d}x \\
&\simeq 2\int_x \varphi^2(x)f^2(x)K_{x/b+1,(1-x)/b+1}(x)\mathrm{d}x \\
&\simeq 2\int_x \varphi^2(x)\left[\int_X f(X)K^2_{x/b+1,(1-x)/b+1}(X)\mathrm{d}X\right]\mathrm{d}F(x) \\
&\simeq 2\int_x \varphi^2(x)A_b(x)[f(x)+O(b)]\mathrm{d}F(x) \\
&\simeq 2\int_x \varphi^2(x)A_b(x)f(x)\mathrm{d}F(x).
\end{aligned}$$

By decomposing the integral according to $\delta = b^{1-\epsilon}$, it yields

$$\begin{aligned}
V_H &\simeq \int_0^\delta + \int_\delta^{1-\delta} + \int_{1-\delta}^1 2\varphi^2(x)A_b(x)f(x)\mathrm{d}F(x) \\
&\simeq \frac{b^{-1/2}}{\sqrt{\pi}}\int_\delta^{1-\delta} \varphi^2(x)[x(1-x)]^{-1/2}f(x)\mathrm{d}F(x) \\
&\simeq \frac{b^{-1/2}}{\sqrt{\pi}}\int_0^1 \varphi^2(x)[x(1-x)]^{-1/2}f(x)\mathrm{d}F(x)
\end{aligned}$$

provided that ϵ is properly chosen and $E\left[\varphi^2(x)[x(1-x)]^{-1/2}\right]$ is finite. Applying Hall's central limit theorem for degenerate U-statistics completes then the proof.

References

Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate, *Review of Financial Studies* **9**: 385–426.

Aït-Sahalia, Y., Bickel, P. J. and Stoker, T. M. (1998). Goodness-of-fit tests for regression using kernel methods, Princeton University, University of California at Berkeley, and Massachusetts Institute of Technology.

Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates, *Annals of Statistics* **1**: 1071–1095.

Chen, S. X. (1999a). Beta kernel estimators for density functions, *Computational Statistics and Data Analysis*. Forthcoming.

Chen, S. X. (1999b). Probability density function estimation using gamma kernels, School of Statistical Science, La Trobe University.

Fernandes, M. and Grammig, J. (1999). Non-parametric specification tests for conditional duration models, European University Institute and University of Frankfurt.

Hall, P. (1984). Central limit theorem for integrated squared error multivariate nonparametric density estimators, *Journal of Multivariate Analysis* **14**: 1–16.

Müller, H. G. (1984). Smooth optimum kernel estimators of densities, regression curves and modes, *Annals of Statistics* **12**: 766–774.

EUI WORKING PAPERS

EUI Working Papers are published and distributed by the European University Institute, Florence

Copies can be obtained free of charge – depending on the availability of stocks – from:

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

To The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI) – Italy
Telefax No: +39/055/4685 636
e-mail: publish@datacomm.iue.it
http://www.iue.it

From Name . . . .
Address. . . . .
. . . .
. . . .
. . . .
. . . .

❑ Please send me a list of EUI Working Papers
❑ Please send me a list of EUI book publications
❑ Please send me the EUI brochure Academic Year 2000/01

Please send me the following EUI Working Paper(s):

No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .
No, Author . . . .
Title: . . . .

Date . . . .

Signature . . . .

Working Papers of the Department of Economics Published since 1999

ECO No. 99/1
Jian-Ming ZHOU
How to Carry Out Land Consolidation - An International Comparison

ECO No. 99/2
Nuala O'DONNELL
Industry Earnings Differentials in Ireland: 1987-1994

ECO No. 99/3
Ray BARRELL/Rebecca RILEY
Equilibrium Unemployment and Labour Market Flows in the UK

ECO No. 99/4
Klaus ADAM
Learning while Searching for the Best Alternative

ECO No. 99/5
Guido ASCARI/Juan Angel GARCIA
Relative Wage Concern and the Keynesian Contract Multiplier

ECO No. 99/6
Guido ASCARI/Juan Angel GARCIA
Price/Wage Staggering and Persistence

ECO No. 99/7
Elena GENNARI
Estimating Money Demand in Italy: 1970-1994

ECO No. 99/8
Marcello D'AMATO/Barbara PISTORESI
Interest Rate Spreads Between Italy and Germany: 1995-1997

ECO No. 99/9
Søren JOHANSEN
A Small Sample Correction for Tests of Hypotheses on the Cointegrating Vectors

ECO No. 99/10
Søren JOHANSEN
A Bartlett Correction Factor for Tests on the Cointegrating Relations

ECO No. 99/11
Monika MERZ/Axel SCHIMMELPFENNIG
Career Choices of German High School Graduates: Evidence from the German Socio-Economic Panel

ECO No. 99/12
Fragiskos ARCHONTAKIS
Jordan Matrices on the Equivalence of the *I* (1) Conditions for VAR Systems

ECO No. 99/13
Étienne BILLETTE de VILLEMEUR
Sequential Decision Processes Make Behavioural Types Endogenous

ECO No. 99/14
Günther REHME
Public Policies and Education, Economic Growth and Income Distribution

ECO No. 99/15
Pierpaolo BATTIGALLI/Marciano SINISCALCHI
Interactive Beliefs and Forward Induction

ECO No. 99/16
Marco FUGAZZA
Search Subsidies vs Hiring Subsidies: A General Equilibrium Analysis of Employment Vouchers

ECO No. 99/17
Pierpaolo BATTIGALLI
Rationalizability in Incomplete Information Games

ECO No. 99/18
Ramon MARIMON/Juan Pablo NICOLINI/Pedro TELES
Competition and Reputation

ECO No. 99/19
Ramon MARIMON/Fabrizio ZILIBOTTI
Employment and Distributional Effects of Restricting Working Time

*out of print

ECO No. 99/20
Leonor COUTINHO
Euro Exchange Rates: What Can Be Expected in Terms of Volatility?

ECO No. 99/21
Bernard FINGLETON
Economic Geography with Spatial Econometrics: A 'Third Way' to Analyse Economic Development and 'Equilibrium', with Application to the EU Regions

ECO 99/22
Mike ARTIS/
Massimiliano MARCELLINO
Fiscal Forecasting: The Track Record of the IMF, OECD and EC

ECO 99/23
Massimo MOTTA/Michele POLO
Leniency Programs and Cartel Prosecution

ECO 99/24
Mike ARTIS/Hans-Martin KROLZIG/
Juan TORO
The European Business Cycle

ECO 99/25
Mathias HOFFMANN
Current Accounts and the Persistence of Global and Country-Specific Shocks: Is Investment Really too Volatile?

ECO 99/26
Mathias HOFFMANN
National Stochastic Trends and International Macroeconomic Fluctuations: The Role of the Current Account

ECO 99/27
Gianmarco I.P. OTTAVIANO/
Jacques-François THISSE
Integration, Agglomeration and the Political Economics of Factor Mobility

ECO 99/28
Gianmarco I.P. OTTAVIANO
Ad usum delphini: A Primer in 'New Economic Geography'

ECO 99/29
Giorgio BASEVI/
Gianmarco I.P. OTTAVIANO
The District Goes Global: Export vs FDI

ECO 99/30
Hans-Martin KROLZIG/Juan TORO
A New Approach to the Analysis of Shocks and the Cycle in a Model of Output and Employment

ECO 99/31
Gianmarco I.P. OTTAVIANO/
Jacques-François THISSE
Monopolistic Competition, Multiproduct Firms and Optimum Product Diversity

ECO 99/32
Stefano MANZOCCHI/
Gianmarco I.P. OTTAVIANO
Outsiders in Economic Integration: The Case of a Transition Economy

ECO 99/33
Roger E.A. FARMER
Two New Keynesian Theories of Sticky Prices

ECO 99/34
Rosalind L. BENNETT/
Roger E.A. FARMER
Indeterminacy with Non-Separable Utility

ECO 99/35
Jess BENHABIB/Roger E.A. FARMER
The Monetary Transmission Mechanism

ECO 99/36
Günther REHME
Distributive Policies and Economic Growth: An Optimal Taxation Approach

ECO 99/37 - Ray BARRELL/
Karen DURY/Ian HURST
Analysing Monetary and Fiscal Policy Regimes Using Deterministic and Stochastic Simulations

ECO 99/38
Chiara FUMAGALLI/Massimo MOTTA
Upstream Mergers, Downstream Mergers, and Secret Vertical Contracts

ECO 99/39
Marcel FRATZSCHER
What Causes Currency Crises: Sunspots, Contagion or Fundamentals?

ECO 99/40
Simone BORGHESI
Intergenerational Altruism and Sustainable Development

***out of print**

ECO 99/41
Jeni KLUGMAN/Alexandre KOLEV
The Role of the Safety Net and the Labour Market on Falling Cash Consumption in Russia: 1994-96. A Quintile-Based Decomposition Analysis

ECO 99/42
Günther REHME
Education, Economic Growth and Personal Income Inequality Across Countries

ECO 99/43
Günther REHME
Why are the Data at Odds with Theory? Growth and (Re-)Distributive Policies in Integrated Economies

ECO 99/44 - Norbert WUTHE
Exchange Rate Volatility: The Impact of Learning Behaviour and the Institutional Framework - A Market Microstructure Approach

* * *

ECO 2000/1
Marcelo FERNANDES
Central Limit Theorem for Asymmetric Kernel Functionals

*out of print