*Econometric Theory*, **25**, 2009, 1716–1733. Printed in the United States of America.
doi:10.1017/S0266466609990302

# INFERENCE ON NONPARAMETRICALLY TRENDING TIME SERIES WITH FRACTIONAL ERRORS

P.M. ROBINSON
*London School of Economics*

The central limit theorem for nonparametric kernel estimates of a smooth trend, with linearly generated errors, indicates asymptotic independence and homoskedasticity across fixed points, irrespective of whether disturbances have short memory, long memory, or antipersistence. However, the asymptotic variance depends on the kernel function in a way that varies across these three circumstances, and in the latter two it involves a double integral that cannot necessarily be evaluated in closed form. For a particular class of kernels, we obtain analytic formulas. We discuss extensions to more general settings, including ones involving possible cross-sectional or spatial dependence.

## 1. INTRODUCTION

Deterministic trend estimation for time series and panel data can be carried out in numerous ways. With macroeconomic time series, use of a polynomial-in-time trend, even a linear trend, is still very common. Classically, polynomial order can be very simply and exactly (in finite samples) tested when the model is expressed in terms of orthogonal polynomials and disturbances are independent, identically distributed, and Gaussian, while, on the other hand, least squares estimates of the coefficients are asymptotically efficient (in the Gauss-Markov sense) even in the presence of general short-range dependence in possibly non-Gaussian stationary disturbances, the technical requirement being that the disturbance spectral density is continuous and positive at zero frequency. Polynomial time trends are also frequently associated with investigation of unit root behavior of the stochastic component. They are also convenient in panel data, especially when the time dimension is not long.

Use of a polynomial, or any parametric function, is inevitably subject to the consequences of misspecification: Typically, large-sample rules of statistical inference are rendered invalid. Moreover, viewed as approximations to general functions of unbounded support, polynomials do not have very good properties. Given

This research was supported by ESRC Grant RES-062-23-0036. I am grateful for the comments of three referees and Rob Taylor. Address correspondence to Peter M Robinson, Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK; e-mail: p.m.robinson@lse.ac.uk.

 © 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

a long time series, an attractive alternative to a high-degree polynomial, and one that is no less ad hoc, is kernel smoothing. Here, the trend is viewed as a smooth, nonparametric function of time.

When the focus is on estimation of the trend, rather than its removal, there is also interest in attaching a measure of variation. If one is prepared to consider estimates at a small number of time points, in the sense of remaining fixed as series length $T$ increases, we need only approximate a finite-dimensional joint distribution. Moreover, the normal approximation to the limit distribution, as $T \to \infty$, can be not only multivariate normal but spherical, in other words, having covariance matrix proportional to the identity. Thus, interval estimation promises to be computationally convenient. Moreover, these kinds of properties can hold not only when disturbances are independent or short-range dependent, but even when they have stationary long memory or antipersistence.

To fix ideas, consider the scalar observable sequence $y_t$, $t = 1, 2, \ldots, T$, described by

$$y_t = r\left(\frac{t}{T}\right) + x_t, \qquad t = 1, \ldots, T, \tag{1.1}$$

where $r$ is a smooth function on the unit interval $(0, 1]$, and $x_t$ is a zero-mean covariance stationary process. Of course $y_t$ also depends on $T$, and so $y_t$, $t = 1, \ldots, T$, forms a triangular array as $T$ increases, but as is common, our notation ignores this. The dependence of the trend $r(t/T)$ on $T$ is required in order to produce a useful asymptotic theory for locally smoothed estimates of $r$.

For a fixed point $\tau$, consider the kernel estimate

$$\hat{r}(\tau) = \frac{1}{Tq} \sum_{t=1}^{T} k\left(\frac{T\tau - t}{Tq}\right) y_t, \tag{1.2}$$

for some $\tau \in (0, 1)$, where $q = q_T$ is a positive, user-chosen, bandwidth sequence, and $k$ is a kernel function satisfying at least

$$\int_{-\infty}^{\infty} k(u)\, du = 1. \tag{1.3}$$

We consider $\hat{r}(\tau_s)$, $s = 1, \ldots, S$, where the $\tau_s$ are distinct values in $(0, 1)$ that remain fixed as $T \to \infty$, as does $S$.

The $\hat{r}(\tau_s)$, after centering at $r(\tau_s)$ and suitable normalization, can be shown to be asymptotically, independently, and normally distributed under regularity conditions that, so far as second-order structure of $x_t$ is concerned, involve only the local-to-zero character of its spectral density, $f(\lambda)$. We assume, following Assumption 1 of Robinson (1997), that it has the form

$$f(\lambda) = g(\lambda)h(\lambda), \qquad -\pi < \lambda \leq \pi, \tag{1.4}$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

where $g$ is an even, nonnegative function that is continuous at $\lambda = 0$, and $h$ is an integrable function such that

$$\delta_j = \int_{-\pi}^{\pi} h(\lambda)\cos(j\lambda)\,d\lambda \tag{1.5}$$

satisfies, for some $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$,

$$\delta_j \sim \phi(d) j^{2d-1}, \quad \text{as } j \to \infty, \quad \text{if } 0 < |d| < \frac{1}{2}, \tag{1.6}$$

$$= 2\pi\Delta_{j0}, \quad \text{if } d = 0, \tag{1.7}$$

where $\Delta_{j0}$ is the Kronecker delta (so $\Delta_{00} = 1$, $\Delta_{j0} = 0$ for $j \neq 0$), and

$$\phi(d) = 2\Gamma(1-2d)\sin(\pi d), \tag{1.8}$$

$$h(0) = 0, \quad \text{if } d \in \left(-\frac{1}{2}, 0\right). \tag{1.9}$$

When $d = 0$ we deduce that $h(\lambda) = 1$, $-\pi < \lambda \leq \pi$, and thus $f(\lambda) = g(\lambda)$, so that $f(\lambda)$ satisfies a mild form of short memory condition. More generally, we can relate the assumptions to the property

$$h(\lambda) \sim \lambda^{-2d}, \quad \text{as } \lambda \to 0+, \tag{1.10}$$

which implies

$$f(\lambda) \sim G\lambda^{-2d}, \quad \text{as } \lambda \to 0+, \tag{1.11}$$

denoting

$$G = g(0). \tag{1.12}$$

When $d \in \left(0, \frac{1}{2}\right)$, (1.10) is equivalent to (1.6) if the $\delta_j$ are quasimonotonically convergent to zero (Yong, 1974, Thm. III-16). Here, $f(\lambda)$ diverges at $\lambda = 0$, and $x_t$ is said to have long memory. For $d \in \left(-\frac{1}{2}, 0\right)$, (1.8) and (1.9) together imply (1.10) (Yong, 1974, Thm. III-31). Here, $f(0) = 0$ and $x_t$ is said to have negative memory, or antipersistence.

These prescriptions do not entail any parametric assumption on the second-order properties of $x_t$, in similar spirit to our nonparametric treatment of trend. Moreover, not only are we likely to be ignorant of the precise value of $d$, but in case it is nonzero we may not even know whether it is positive or negative. Thus there is advantage in an asymptotic distribution theory that assumes only that $d$ lies in the stationary and invertible region $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Robinson (1997) developed such a theory, under the following additional conditions: First, $r(u)$ satisfies a Lipschitz condition of degree $\gamma$, $0 < \gamma \leq 1$, or $r(u)$ is differentiable with derivative satisfying a Lipschitz condition of degree $\gamma - 1$,

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$1 < \gamma \leq 2$. Next, $x_t = \Sigma_{j=-\infty}^{\infty} c_j \varepsilon_{t-j}$, $\Sigma_{j=-\infty}^{\infty} c_j^2 < \infty$, for $t = 0, \pm 1, \ldots$, where the $\varepsilon_t^2$ are uniformly integrable and $\mathrm{E}(\varepsilon_t \,|\mathcal{F}_{t-1}) = 0$, $\mathrm{E}\left(\varepsilon_t^2 \,|\mathcal{F}_{t-1}\right) = 1$, a.s., $t = 0, \pm 1, \ldots$, where $\mathcal{F}_t$ is the $\sigma$-field of events generated by $\{\varepsilon_s, s \leq t\}$. Next, $k(u)$ is even, eventually monotonically nonincreasing, almost everywhere differentiable, and in addition to (1.3), satisfies $k(u) = O((1+u^2)^{-1})$, $k'(u) = O((1+|u|^{1+\eta})^{-1})$ for some $\eta > 0$, $k'(u)$ denoting the derivative of $k(u)$ where it exists. Finally, $(Tq)^{-1} + T^{1/2-d} q^{1/2-d+\gamma} \to 0$, as $T \to 0$.

The linear process assumption on $x_t$ might be replaced by a suitable mixing condition when $d = 0$ (see, e.g., Roussas, Tran, and Ioannides (1992)), while for $d \in (0, \frac{1}{2})$ $x_t$ might instead be represented as a nonlinear function of a Gaussian long memory process (cf. Fox and Taqqu, 1985). However, the linearity assumption conveniently covers simultaneously short memory, long memory, and antipersistence.

Under the above conditions, Theorem 1 of Robinson (1997) states that

$$(Tq)^{1/2-d} \left\{\hat{r}(\tau_s) - r(\tau_s)\right\}, \qquad s = 1, \ldots, S, \tag{1.13}$$

converge in distribution to independent identically distributed $N(0, G\psi(d))$ random variables, where

$$\psi(0) = 2\pi\kappa(0), \tag{1.14}$$

$$\psi(d) = \phi(d)\kappa(d), \qquad 0 < |d| < \frac{1}{2}, \tag{1.15}$$

in which

$$\kappa(d) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(v)\{k(w) - k(v)\}|v - w|^{2d-1}\, dv\, dw, \qquad -\frac{1}{2} < d < 0, \tag{1.16}$$

$$\kappa(0) = \int_{-\infty}^{\infty} k(v)^2 dv, \tag{1.17}$$

$$\kappa(d) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(v)k(w)|v - w|^{2d-1}\, dv\, dw, \qquad 0 < d < \frac{1}{2}. \tag{1.18}$$

The variance $G\psi(d)$ in the limit distribution is desirably constant over the $\tau_s$. This is in contrast to the outcome with kernel estimates of random-design nonparametric regression. Indeed, in that setting, with long memory the desirable independence in the limit distribution may also be lost, the limiting covariance matrix possibly even having only unit rank (see Robinson, 2007). Other asymptotic properties of kernel fixed-design nonparametric regression with long memory disturbances are discussed by Hall and Hart (1990) and Csorgo and Mielniczuk (1995).

Application of the central limit theorem in statistical inference requires estimation of $G$ and $d$. Robinson (1997) justified local Whittle estimates of $G$ and $d$, based on residuals from the nonparametric regression, to the extent that replacing

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$G$ and $d$ by the estimates produces asymptotically valid inferences. This outcome relies also on the continuity of $\psi(d)$, demonstrated in Robinson (1997). The estimate of $d$ is simply substituted for $d$ in $\phi(d)$, while the version (1.16)–(1.18) of $\kappa(.)$ used will automatically reflect the sign of the estimate of $d$.

Calculation of $\kappa(d)$ is not discussed by Robinson (1997). For $d \neq 0$, a double integral is involved, which for general $k(u)$ cannot be obtained analytically. Moreover, some additional care is needed when $d < 0$, in that $k(v)k(w)\,|v-w|^{2d-1}$ and $k(v)^2\,|v-w|^{2d-1}$ are not integrable individually, though their difference is.

The present paper fills a gap by deducing analytic formulas for $\kappa(d)$ for a particular class of $k(u)$, namely, (even) polynomials of arbitrary finite degree on $(-1, 1)$. This class includes the uniform kernel, the "optimal" Epanechnikov kernel, and also bias-reducing higher-order kernels of arbitrary degree. (In an empirical example, Robinson, 1997, employed an alternative specific (cosine-bell) $k(u)$ that is not a member of this class and is not a higher-order kernel, and had to approximate $\kappa(d)$ numerically.) The results are presented in Section 2, with proofs relegated to Section 4, employing some lemmas and a corollary established in Section 5. Section 3 illustrates the results and discusses relevance to more general settings.

## 2. FORMULAS FOR $\kappa(d)$

We focus on $k(u)$ of form

$$k(u) = \sum_{j=0}^{p} a_j u^{2j}, \qquad |u| < 1; \;\; = 0, \qquad |u| \geq 1, \tag{2.1}$$

for prescribed $p \geq 0$. When $p = 0$, $k(u)$ is flat, and (1.3) implies $a_0 = \frac{1}{2}$. When $p = 1$, (1.3) implies $2a_0 + 2a_1/3 = 1$, an important special case being the Epanechnikov mean-squared-error optimal kernel in which $a_0 = \frac{3}{4}$, $a_1 = -\frac{3}{4}$. When $p \geq 1$, higher-order kernels are a possibility, for example, satisfying the constraints

$$\int_{-\infty}^{\infty} u^i k(u)\,du = 0, \qquad 1 \leq i \leq 2p.$$

Thus, under (2.1) we require

$$\sum_{j=0}^{p} a_j \int_{-1}^{1} u^{2j+i}\,du = 0, \qquad 1 \leq i \leq 2p; \tag{2.2}$$

that is,

$$\sum_{j=0}^{p} \frac{a_j}{2j+2\ell+1} = 0, \qquad 1 \leq \ell \leq p.$$

Thus, imposing also (1.3), the vector $a = (a_0, \ldots, a_p)'$ is given by $a = B^{-1}b$, where the $(p+1) \times (p+1)$ matrix $B$ has $(i, j)$th element $B_{ij} = 1/(2(i-1)$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$+2(j-1)+1)$, and $b$ is the $(p+1)\times 1$ vector whose first element is 2 and whose remaining elements are 0. For example, if $p=1$, then $a_0=9/8, a_1=-15/8$. For $p>1$, we can generalize (2.2) by replacing $p$ there by $q\in[1,p]$. Of course, (2.1) excludes a number of popularly used kernels, but except for the issue of kernel order, the actual choice of kernel is often believed to be of secondary importance.

THEOREM 1. *For $d=0$ and $k(u)$ given by (2.1),*

$$\kappa(d)=2\sum_{i=0}^{p}\sum_{j=0}^{p}\frac{a_ia_j}{2i+2j+1}. \tag{2.3}$$

Define, for $b\in(-1,0)$ and even $r,s$,

$$I(r,s;b)=2r!s!\Gamma(b+1)\sum_{\ell=0}^{r}\sum_{m=0}^{s}\frac{(-1)^{\ell+m}2^{\ell+m+b+2}}{(r-\ell)!\,(s-m)!\Gamma(\ell+m+b+3)}. \tag{2.4}$$

THEOREM 2. *For $d\in(0,0.5)$ and $k(u)$ given by (2.1),*

$$\kappa(d)=\sum_{i=0}^{p}\sum_{j=0}^{p}a_ia_jI(2i,2j;2d-1). \tag{2.5}$$

Define, for $b\in(-2,-1)$, even $r$ and even $s\geq 2$,

$$J(r,s;b)=\sum_{n=0}^{s-1}(-1)^nI(r+n,s-1-n;b+1), \tag{2.6}$$

$$K(r;b)=-2r!\Gamma(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!\Gamma(\ell+b+3)}2^{\ell+b+2}. \tag{2.7}$$

THEOREM 3. *For $d\in(-0.5,0)$ and $k(u)$ given by (2.1),*

$$\kappa(d)=\sum_{i=0}^{p}\sum_{j=1}^{p}a_ia_jJ(2i,2j;2d-1)1(p>0)-\sum_{i=0}^{p}\sum_{j=0}^{p}a_ia_jK(2i+2j;2d-1). \tag{2.8}$$

## 3. DISCUSSION

We illustrate the formulas first with $p=0$, whence $a_0=\frac{1}{2}$. From (2.3),

$$\psi(0)=2\pi\kappa(0)=2\pi\times 2a_0^2=\pi.$$

From (2.4),

$$I(0,0;2d-1)=\frac{2^{2d+2}}{2d(2d+1)}.$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

Thus, from (2.5),

$$\psi(d) = 2\Gamma(1-2d)\sin(\pi d)\,a_0^2 I(0,0;2d-1)$$
$$= \frac{1}{2}\Gamma(1-2d)\sin(\pi d)\,\frac{2^{2d+2}}{2d(2d+1)}$$
$$= \Gamma(1-2d)\sin(\pi d)\,\frac{2^{2d}}{d(2d+1)}, \qquad 0 < d < \frac{1}{2},$$

and from (2.8),

$$\psi(d) = -2\Gamma(1-2d)\sin(\pi d)\,a_0^2 K(0;2d-1)$$
$$= 2\Gamma(1-2d)\sin(\pi d)\,a_0^2 I(0,0;2d-1)$$
$$= \Gamma(1-2d)\sin(\pi d)\,\frac{2^{2d}}{d(2d+1)}, \qquad -\frac{1}{2} < d < 0.$$

Notice that as $d \to 0^+$, and $d \to 0^-$, $\psi(d) \to \psi(0)$.

Now take $p = 1$. We have

$$\kappa(0) = 2a_0^2 + \frac{4a_0a_1}{3} + \frac{2a_1^2}{5},$$

so

$$\psi(0) = 2\pi\left(2a_0^2 + \frac{4a_0a_1}{3} + \frac{2a_1^2}{5}\right).$$

Because the formulas become complicated, we finally consider $\psi(d)$ for $0 < d < 1/2$ only. Write

$$I(r,s;b) = 2r!s!\Gamma(b+1)\,M(r,s;b),$$
$$M(r,s;b) = \sum_{\ell=0}^{r}\sum_{m=0}^{s}\frac{(-1)^{\ell+m}2^{b+2+\ell+m}}{(r-\ell)!\,(s-m)!\Gamma(\ell+m+b+3)}.$$

Then

$$M(2,0;b) = \frac{2^{b+1}}{\Gamma(b+3)} - \frac{2^{b+3}}{\Gamma(b+4)} + \frac{2^{b+4}}{\Gamma(b+5)}$$
$$= \frac{2^{b+1}}{\Gamma(b+5)}\{(b+3)(b+4) - 4(b+4) + 8\}$$
$$= \frac{2^{b+1}(b^2+3b+4)}{\Gamma(b+5)},$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

so

$$I(2,0;2d-1)=\frac{2^{2d+2}(2d^2+d+1)}{(2d+3)(2d+2)(2d+1)d}.$$

Next,

$$\begin{aligned}M(2,2;b)&=\frac{M(2,0;b)}{2}+\sum_{\ell=0}^{2}\frac{(-1)^{\ell+1}2^{b+3+\ell}}{(2-\ell)!\Gamma(\ell+b+4)}+\sum_{\ell=0}^{2}\frac{(-1)^{\ell+2}2^{b+4+\ell}}{(2-\ell)!\Gamma(\ell+b+5)}\\&=\frac{2^b}{\Gamma(b+3)}-\frac{2^{b+2}}{\Gamma(b+4)}+\frac{2^{b+3}}{\Gamma(b+5)}-\frac{2^{b+2}}{\Gamma(b+4)}+\frac{2^{b+4}}{\Gamma(b+5)}-\frac{2^{b+5}}{\Gamma(b+6)}\\&\quad+\frac{2^{b+3}}{\Gamma(b+5)}-\frac{2^{b+5}}{\Gamma(b+6)}+\frac{2^{b+6}}{\Gamma(b+7)}\\&=2^b\left\{\frac{1}{\Gamma(b+3)}-\frac{2^3}{\Gamma(b+4)}+\frac{2^5}{\Gamma(b+5)}-\frac{2^6}{\Gamma(b+6)}+\frac{2^6}{\Gamma(b+7)}\right\}.\end{aligned}$$

Thus,

$$\begin{aligned}I(2,2;2d-1)=8\Gamma(2d)\Bigg\{&\frac{2^{2d-1}}{\Gamma(2d+2)}-\frac{2^{2d+2}}{\Gamma(2d+3)}+\frac{2^{2d+4}}{\Gamma(2d+4)}\\&-\frac{2^{2d+5}}{\Gamma(2d+5)}+\frac{2^{2d+5}}{\Gamma(2d+6)}\Bigg\}.\end{aligned}$$

It follows that

$$\begin{aligned}\kappa(d)&=a_0^2I(0,0;2d-1)+2a_0a_1I(2,0;2d-1)+a_1^2I(2,2;2d-1)\\&=a_0^2\frac{2^{2d+2}}{2d(2d+1)}+2a_0a_1\frac{2^{2d+2}(2d^2+d+1)}{(2d+3)(2d+2)(2d+1)d}+a_1^28\Gamma(2d)\\&\quad\times\left\{\frac{2^{2d-1}}{\Gamma(2d+2)}-\frac{2^{2d+2}}{\Gamma(2d+3)}+\frac{2^{2d+4}}{\Gamma(2d+4)}-\frac{2^{2d+5}}{\Gamma(2d+5)}+\frac{2^{2d+5}}{\Gamma(2d+6)}\right\}.\end{aligned}$$

Our results are relevant to more general models than (1.1). For example, we might consider an $N\times 1$ vector of observations $Y_t$ satisfying

$$Y_t=R\left(\frac{t}{T}\right)+X_t,\qquad t=1,\ldots,T, \tag{3.1}$$

where $R$ is an $N\times 1$ vector of smooth functions, and $X_t$ is a vector of zero-mean jointly covariance stationary processes. We can estimate each element of $R(\tau_s)$, $s=1,\ldots,S$, as in (1.2), and deduce from Robinson (1997) $N$ $S$-dimensional

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

central limit theorems. We can also consider a joint, $SN$-dimensional central limit theorem for all the estimates, under suitable conditions on $X_t$. For each $\tau_s$, the asymptotic covariance matrix of the estimate of $R(\tau_s)$ will not be diagonal unless the elements of $X_t$ are incoherent. However, it will have scalar factor $\kappa(d)$, as above.

With panel data, where $N$ might be large, there is interest in a more parsimonious structure than (3.1). One example concerns a panel whose members share the same-shaped nonparametric regression function, but its location and magnitude can vary across individuals. Consider $Y_t = \alpha + \beta r(t/T) + X_t, \quad t = 1, \ldots, T$, where $\alpha$ and $\beta$ are $N \times 1$ vectors, such that $\alpha \neq \beta$ and $\beta \neq 0$, and $X_t$ possibly exhibits spatial or other cross-sectional correlation. Because the location and scale of $r$ are unrestricted, we can, with no loss of generality, take $1'\alpha = 0$, $1'\beta = 1$, where 1 is the $N \times 1$ vector of ones. Thus the cross-sectional aggregates $1'Y_t$ satisfy $1'Y_t = r(t/T) + 1'X_t$, and, taking $y_t = 1'Y_t$, $x_t = 1'X_t$, we revert to model (1.1). With respect to the cross-sectional properties of the elements of $X_t$, these may or may not have identical memory parameters. If they do, then $d$ is the common value, unless certain cointegration is present. If they do not, then generally the largest of them is $d$. When the memory parameters differ, this can influence the "local smoothness parameter" of $f(\lambda)$, but this is not of concern in our central limit theorem for the $\hat{r}(\tau_s)$, though it is relevant to asymptotic theory for estimates of $d$ (see Robinson, 1997).

## 4. PROOFS OF THEOREMS

**Proof of Theorem 1.** This proof is straightforward. ■

**Proof of Theorem 2.** The result is clearly true by substitution if we can show that, for even $r, s$ and for $b = 2d - 1$,

$$\int_{-1}^{1}\int_{-1}^{1} v^r w^s |v - w|^b \, dv \, dw = I(r, s; b).$$

The left-hand side is

$$\begin{aligned}
&\int_0^1\int_0^1 v^r w^s |v-w|^b \, dv \, dw + \int_0^1\int_{-1}^0 v^r w^s |v-w|^b \, dv \, dw \\
&\quad + \int_{-1}^0\int_0^1 v^r w^s |v-w|^b \, dv \, dw + \int_{-1}^0\int_{-1}^0 v^r w^s |v-w|^b \, dv \, dw \\
&\quad = 2\int_0^1\int_0^1 v^r w^s |v-w|^b \, dv \, dw + 2\int_0^1\int_{-1}^0 v^r w^s |v-w|^b \, dv \, dw \\
&\quad = 2\int_0^1 w^s \{F(w; r, b) + G(w; r, b) + H(w; r, b)\} \, dw, \qquad \textbf{(4.1)}
\end{aligned}$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

where

$$F(w;r,b)=\int_0^w v^r\,(w-v)^b\,dv, \qquad G(w;r,b)=\int_w^1 v^r\,(v-w)^b\,dv,$$
$$H(w;r,b)=\int_0^1 v^r\,(w+v)^b\,dv.$$

From Lemmas 1–3 and Corollary 1 in Section 5, (4.1) is

$$\begin{aligned}
&2\int_0^1 w^s\left\{B(r+1,b+1)w^{r+b+1}+r!\sum_{\ell=0}^{r}\frac{1}{(\ell+b+1)\,(r-\ell)!\ell!}w^{r-\ell}(1-w)^{\ell+b+1}\right.\\
&\qquad+r!\Gamma\,(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!\Gamma\,(\ell+b+2)}\,(w+1)^{\ell+b+1}\\
&\qquad\left.-B(r+1,b+1)w^{r+b+1}\right\}dw\\
&=2r!\left\{\sum_{\ell=0}^{r}\frac{1}{(\ell+b+1)\,(r-\ell)!\ell!}\int_0^1 w^{r+s-\ell}(1-w)^{\ell+b+1}dw\right.\\
&\qquad\left.+\Gamma\,(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!\Gamma\,(\ell+b+2)}H(1;s,\ell+b+1)\right\}\\
&=2r!\sum_{\ell=0}^{r}\frac{B(r+s-\ell+1,\ell+b+2)}{(r-\ell)!\,(\ell+b+1)\,\ell!}+2r!\Gamma\,(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!\Gamma\,(\ell+b+2)}\\
&\quad\times\left\{s!\Gamma\,(\ell+b+2)\sum_{m=0}^{s}\frac{(-1)^m}{(s-m)!\Gamma\,(\ell+m+b+3)}2^{\ell+m+b+2}\right.\\
&\qquad\left.-B(s+1,\ell+b+2)\right\}\\
&=\frac{2}{\Gamma\,(r+s+b+3)}\sum_{\ell=0}^{r}\frac{\Gamma(r+s-\ell+1)\Gamma(\ell+b+1)}{(r-\ell)!\ell!}\\
&\quad+2r!s!\Gamma\,(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!}\\
&\quad\times\left\{\sum_{m=0}^{s}\frac{(-1)^m}{(s-m)!\Gamma\,(\ell+m+b+3)}2^{\ell+m+b+2}-\frac{1}{\Gamma\,(\ell+s+b+3)}\right\}
\end{aligned}$$

for even $s$. Using a combinatorial identity (Knuth, 1968, p. 57, form. 18),

$$\sum_{\ell=0}^{r}\frac{(-1)^\ell\Gamma\,(r+s+b+3)}{(r-\ell)!\Gamma\,(\ell+s+b+3)}=\sum_{\ell=0}^{r}\frac{(-1)^\ell\Gamma\,(r+s+b+3)}{\ell!\Gamma\,(r+s-\ell+b+3)}$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$$= (-1)^r \frac{\Gamma(r+s+b+2)}{r!\Gamma(s+b+2)}$$
$$= \frac{\Gamma(r+s+b+2)}{r!\Gamma(s+b+2)},$$

for even $r$. Using another combinatorial identity (Knuth, 1968, p. 58, form. 26, with $t=-1$):

$$\sum_{\ell=0}^{r} \frac{\Gamma(r+s-\ell+1)\Gamma(\ell+b+1)}{(r-\ell)!\ell!} = \frac{s!\Gamma(b+1)\Gamma(r+s+b+2)}{r!\Gamma(s+b+2)}.$$

Then (2.4) follows by substitution. ■

**Proof of Theorem 3.** Clearly, $\kappa(d)$ is

$$\int_{-\infty}^{\infty}\int_{-1}^{1} \sum_{i=0}^{p} a_i v^{2i} \left\{ \sum_{j=0}^{p} a_j (w^{2j} 1(|w|<1) - v^{2j}) \right\} |v-w|^{2d-1} dv\, dw$$
$$= \int_{-1}^{1}\int_{-1}^{1} \sum_{i=0}^{p}\sum_{j=1}^{p} a_i a_j v^{2i} \left(w^{2j} - v^{2j}\right) |v-w|^{2d-1} dv\, dw 1(p>0)$$
$$- 2\int_{1}^{\infty}\int_{-1}^{1} \left( \sum_{j=0}^{p} a_j v^{2j} \right)^2 (w-v)^{2d-1} dv\, dw,$$

with $1(.)$ denoting the indicator function. This can be written as (2.8) with, for $b<-1$, $r\geq 0$, $s\geq 1$,

$$J(r,s;b) = \int_{-1}^{1}\int_{-1}^{1} v^r (w^s - v^s) |v-w|^b dv\, dw,$$
$$K(r;b) = 2\int_{1}^{\infty}\int_{-1}^{1} v^r (w-v)^b dv\, dw,$$

whence it remains to show that definitions (2.6) and (2.7) hold.

For even $r, s$, with $s\geq 2$,

$$J(r,s;b) = \int_{0}^{1}\int_{0}^{1} v^r (w^s - v^s) |v-w|^b dv\, dw + \int_{0}^{1}\int_{-1}^{0} v^r (w^s - v^s) |v-w|^b dv\, dw$$
$$+ \int_{-1}^{0}\int_{0}^{1} v^r (w^s - v^s) |v-w|^b dv\, dw$$
$$+ \int_{-1}^{0}\int_{-1}^{0} v^r (w^s - v^s) |v-w|^b dv\, dw$$
$$= 2\int_{0}^{1}\int_{0}^{1} v^r (w^s - v^s) |v-w|^b dv\, dw$$
$$+ 2\int_{0}^{1}\int_{-1}^{0} v^r (w^s - v^s) (w-v)^b dv\, dw.$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

Now

$$\int_0^1\int_0^1 v^r(w^s-v^s)\,|v-w|^b\,dv\,dw = \int_0^1\int_0^1 v^r\sum_{n=0}^{s-1}v^n w^{s-n-1}(w-v)\,|v-w|^b\,dv\,dw$$

$$= \sum_{n=0}^{s-1}\int_0^1 w^{s-n-1}\{F(w;r+n,b+1)$$

$$-G(w;r+n,b+1)\}\,dw.$$

From Lemma 1 in Section 5,

$$F(w;r+n,b+1) = w^{r+n+b+2}B(r+n+1,b+2).$$

From Lemma 2 in Section 5,

$$G(w;r+n,b+1) = (r+n)!\sum_{\ell=0}^{r+n}\frac{1}{(\ell+b+2)\,(r+n-\ell)!\ell!}w^{r+n-\ell}(1-w)^{\ell+b+2}.$$

Thus,

$$\int_0^1 w^{s-1-n}\{F(w;r+n,b+1)-G(w;r+n,b+1)\}\,dw$$

is

$$B(r+n+1,b+2)\int_0^1 w^{r+s+b+1}dw$$

$$-(r+n)!\sum_{\ell=0}^{r+n}\frac{1}{(\ell+b+2)\,(r+n-\ell)!\ell!}\int_0^1 w^{r+s-\ell-1}(1-w)^{b+\ell+2}dw$$

$$= \frac{B(r+n+1,b+2)}{r+s+b+2} - (r+n)!\sum_{\ell=0}^{r+n}\frac{B(r+s-\ell,\ell+b+3)}{(\ell+b+2)\,(r+n-\ell)!\ell!}$$

$$= \frac{B(r+n+1,b+2)}{r+s+b+2} - \frac{(r+n)!}{\Gamma(r+s+b+3)}\sum_{\ell=0}^{r+n}\frac{\Gamma(r+s-\ell)\Gamma(\ell+b+2)}{(r+n-\ell)!\ell!}. \tag{4.2}$$

Using again Knuth (1968, p. 58, form. 26, with $t=-1$),

$$\sum_{\ell=0}^{r+n}\frac{\Gamma(r+s-\ell)\Gamma(\ell+b+2)}{(r+n-\ell)!\ell!} = \frac{\Gamma(b+2)\Gamma\,(r+s+b+2)\,(s-n-1)!}{(r+n)!\Gamma(s-n+b+2)},$$

so (4.2) is

$$\frac{B(r+n+1,b+2)}{r+s+b+2} - \frac{\Gamma(b+2)(s-n-1)!}{(r+s+b+2)\Gamma(s-n+b+2)}$$

$$= \frac{B(r+n+1,b+2)-B(s-n,b+2)}{r+s+b+2}.$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

Thus,

$$\sum_{n=0}^{s-1}\int_0^1 w^{s-1-n}\{F(w;r+n,b+1)-G(w;r+n,b+1)\}\,dw$$

$$=\frac{1}{r+s+b+2}\sum_{n=0}^{s-1}\{B(r+n+1,b+2)-B(s-n,b+2)\}. \tag{4.3}$$

Next,

$$\int_0^1\int_{-1}^0 v^r(w^s-v^s)(w-v)^b\,dv\,dw$$

$$=\int_0^1\int_0^1 v^r(w^s-v^s)(w+v)^b\,dv\,dw$$

$$=\int_0^1\int_0^1 v^r\sum_{n=0}^{s-1}(-v)^n w^{s-n-1}(w+v)^{b+1}\,dv\,dw$$

$$=\sum_{n=0}^{s-1}(-1)^n\int_0^1\int_0^1 v^{r+n}w^{s-n-1}(w+v)^{b+1}\,dv\,dw$$

$$=\sum_{n=0}^{s-1}(-1)^n\int_0^1 w^{s-n-1}H(w;r+n,b+1)\,dw.$$

From Lemma 3 in Section 5,

$$H(w;r+n,b+1)=(r+n)!\Gamma(b+2)\sum_{\ell=0}^{r+n}\frac{(-1)^\ell}{(r+n-\ell)!\Gamma(\ell+b+3)}(w+1)^{\ell+b+2}$$

$$-(-1)^{r+n}B(r+n+1,b+2)w^{r+n+b+2},$$

so, since $r$ is even, the integral is

$$(r+n)!\Gamma(b+2)\sum_{\ell=0}^{r+n}\frac{(-1)^\ell}{(r+n-\ell)!\Gamma(\ell+b+3)}H(1;s-n-1,\ell+b+2)$$

$$-(-1)^n B(r+n+1,b+2)\int_0^1 w^{r+s+b+1}dw.$$

From Corollary 1 in Section 5,

$$H(1;s-n-1,\ell+b+2)=(s-n-1)!\Gamma(b+3+\ell)$$

$$\times\sum_{m=0}^{s-n-1}\frac{(-1)^m}{(s-n-1)!\Gamma(\ell+m+b+4)}2^{\ell+m+b+3}$$

$$-(-1)^{s-n-1}B(s-n,b+3+\ell).$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

Thus,

$$\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell}}{(r+n-\ell)!\Gamma(\ell+b+3)}H(1;s-n-1,\ell+b+2)$$

$$=(s-n-1)!\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell}}{(r+n-\ell)!}\sum_{m=0}^{s-n-1}\frac{(-1)^{m}2^{\ell+m+b+3}}{(s-n-m-1)!\Gamma(\ell+m+b+4)}$$

$$-\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell+s-n-1}}{(r+n-\ell)!\Gamma(\ell+b+3)}B(s-n,\ell+b+3),$$

and so

$$(r+n)!\Gamma(b+2)\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell}}{(r+n-\ell)!\Gamma(\ell+b+3)}H(1;s-n-1,\ell+b+2)$$

$$=\Gamma(b+2)(r+n)!(s-n-1)!$$

$$\times\sum_{\ell=0}^{r+n}\sum_{m=0}^{s-n-1}\frac{(-1)^{\ell+m}2^{\ell+m+b+3}}{(r+n-\ell)!(s-n-m-1)!\Gamma(\ell+m+b+4)}$$

$$-\Gamma(b+2)(r+n)!\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell+s-n-1}}{(r+n-\ell)!\Gamma(\ell+b+3)}B(s-n,\ell+b+3)$$

$$=\frac{1}{2}I(r+n,s-n-1,b+1)-\Gamma(b+2)(r+n)!$$

$$\times\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell+s-n-1}\Gamma(s-n)}{(r+n-\ell)!\Gamma(\ell+s-n+b+3)}.$$

From Knuth (1968, p. 57, form. 18) again,

$$\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell}\Gamma(r+s+b+3)}{(r+n-\ell)!\Gamma(\ell+s-n+b+3)}=\sum_{\ell=0}^{r+n}\frac{(-1)^{r+n-\ell}\Gamma(r+s+b+3)}{\ell!\Gamma(r+s-\ell+b+3)}$$

$$=\frac{(-1)^{r+n}\Gamma(r+s+b+2)}{(r+n)!\Gamma(s-n+b+2)},$$

so

$$\Gamma(b+2)(r+n)!\Gamma(s-n)\sum_{\ell=0}^{r+n}\frac{(-1)^{\ell+s-n-1}}{(r+n-\ell)!\Gamma(\ell+s-n+b+3)}$$

$$=\Gamma(b+2)\frac{(r+n)!\Gamma(s-n)(-1)^{s-n-1}}{(r+n)!\Gamma(s-n+b+2)}\frac{\Gamma(r+s+b+2)}{\Gamma(r+s+b+3)}$$

$$=-\frac{(-1)^{n}\Gamma(b+2)\Gamma(s-n)}{\Gamma(s-n+b+2)(r+s+b+2)}=-(-1)^{n}\frac{B(s-n,b+2)}{r+s+b+2},$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

since $s$ is even. Also,

$$B(r+n+1,b+2)\int_0^1 w^{r+s+b+1}dw = \frac{B(r+n+1,b+2)}{r+s+b+2}.$$

It follows that

$$\int_0^1 w^{s-n-1}H(w;r+n,b+1)\,dw$$

$$= \frac{1}{2}I(r+n,s-n-1;b+1)$$

$$+(-1)^n\frac{B(s-n,b+2)-B(r+n+1,b+2)}{r+s+b+2}.$$

Then, from (4.3),

$$J(r,s;b) = \frac{2}{r+s+b+2}\sum_{n=0}^{s-1}\{B(r+n+1,b+2)-B(s-n,b+2)\}$$

$$+\sum_{n=0}^{s-1}(-1)^n I(r+n,s-n-1;b+1)$$

$$+\frac{2}{r+s+b+2}\sum_{n=0}^{s-1}(-1)^{2n}\{B(s-n,b+2)-B(r+n+1,b+2)\}$$

$$=\sum_{n=0}^{s-1}(-1)^n I(r+n,s-n-1;b+1),$$

giving (2.6). Next, for even $r$ and $b<-1$,

$$K(r;b) = 2\int_{-1}^1 v^r\left\{\int_1^\infty (w-v)^b dw\right\}dv$$

$$=2\int_{-1}^1 v^r\left[\frac{(w-v)^{b+1}}{b+1}\right]_1^\infty dv$$

$$=-\frac{2}{b+1}\int_{-1}^1 v^r(1-v)^{b+1}\,dv$$

$$=-\frac{2}{b+1}\int_0^1 v^r(1-v)^{b+1}\,dv-\frac{2}{b+1}H(1;r,b+1)$$

$$=-\frac{2}{b+1}B(r+1,b+2)$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$$-\frac{2}{b+1}\left\{r!\Gamma(b+2)\sum_{m=0}^{r}\frac{(-1)^m}{(r-m)!\Gamma(m+b+3)}2^{m+b+2}\right.$$

$$\left.-B(r+1,b+2)\right\}$$

$$=-2r!\Gamma(b+1)\sum_{m=0}^{r}\frac{(-1)^m}{(r-m)!\Gamma(m+b+3)}2^{m+b+2}, \tag{4.4}$$

using Corollary 1 in Section 5, to prove (2.7). ■

## 5. TECHNICAL LEMMAS

We state and prove three lemmas.

LEMMA 1. *For real* $b>-1, r>-1$, *and* $w\in(0,1)$,

$F(w;r,b)=w^{r+b+1}B(r+1,b+1)$.

**Proof.** The proof follows from the definition of the beta function after change of variable. ■

LEMMA 2. *For real* $b\geq 0$, *integer* $r$ *and* $w\in(0,1)$,

$$G(w;r,b)=r!\sum_{\ell=0}^{r}\frac{1}{(\ell+b+1)(r-\ell)!}w^{r-\ell}(1-w)^{\ell+b+1}.$$

**Proof.** We have

$$G(w;r,b)=\int_0^{1-w}(u+w)^r u^b du$$

$$=w^r(1-w)^{b+1}B(1,b+1)_2F_1(-r,b+1;b+2;(w-1)/w),$$

from Gradshteyn and Ryzhik (1994, p. 335, form. 8), the last factor being the hypergeometric series

$${}_2F_1(-r,b+1;b+2;(w-1)/w)$$

$$=\frac{\Gamma(b+2)}{\Gamma(-r)\Gamma(b+1)}\sum_{j=0}^{\infty}\frac{\Gamma(j-r)\Gamma(j+b+1)}{\Gamma(j+b+2)}\frac{\{(w-1)/w\}^j}{j!}$$

$$=1-\frac{r(b+1)}{b+2}\left(\frac{w-1}{w}\right)-\frac{r(1-r)(b+1)(b+2)}{2(b+2)(b+3)}\left(\frac{w-1}{w}\right)^2$$

$$-\frac{r(1-r)(2-r)(b+1)(b+2)(b+3)}{2\times 3(b+2)(b+3)(b+4)}\left(\frac{w-1}{w}\right)^3-\cdots$$

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

$$-\frac{r\,(1-r)\,(2-r)\ldots(r-1-r)\,(b+1)}{2\times 3\times\ldots\times r\,(r+b+1)}\left(\frac{w-1}{w}\right)^r$$

$$=(b+1)r!\sum_{\ell=0}^{r}\frac{1}{(\ell+b+1)\,\ell!\,(r-\ell)!}\left(\frac{1-w}{w}\right)^\ell;$$

see Gradshteyn and Ryzhik (1994, p. 1065), whence the result follows from $B(1,b+1)=(b+1)^{-1}$. ■

LEMMA 3. *For real $b>-1$, integer $r$, and $w\in(0,1)$,*

$$H(w;r,b)=r!\Gamma\,(b+1)\sum_{\ell=0}^{r}\frac{(-1)^\ell}{(r-\ell)!\Gamma\,(\ell+b+2)}\,(w+1)^{\ell+b+1}$$
$$-(-1)^r B(r+1,b+1)w^{r+b+1}.$$

**Proof.** Successively integrating by parts,

$$H(w;r,b)=\left[v^r\frac{(w+v)^{b+1}}{b+1}\right]_0^1-\frac{r}{b+1}\int_0^1 v^{r-1}\,(w+v)^{b+1}\,dv$$

$$=\frac{(w+1)^{b+1}-w^{b+1}1(r=0)}{b+1}-\frac{r}{b+1}\left[v^{r-1}\frac{(w+v)^{b+2}}{b+2}\right]_0^1$$

$$+\frac{r\,(r-1)}{(b+1)\,(b+2)}\int_0^1 v^{r-2}\,(w+v)^{b+2}\,dv$$

$$=\frac{(w+1)^{b+1}-w^{b+1}1(r=0)}{b+1}-\frac{r\left\{(w+1)^{b+2}-w^{b+2}1(r=1)\right\}}{(b+1)\,(b+2)}$$

$$+\frac{r\,(r-1)}{(b+1)\,(b+2)}\left[v^{r-2}\frac{(w+v)^{b+3}}{b+3}\right]_0^1-\frac{r\,(r-1)\,(r-2)}{(b+1)\,(b+2)\,(b+3)}$$

$$\times\int_0^1 v^{r-3}\,(w+v)^{b+3}\,dv$$

$$=\frac{(w+1)^{b+1}-w^{b+1}1(r=0)}{b+1}-\frac{r\left\{(w+1)^{b+2}-w^{b+2}1(r=1)\right\}}{(b+1)\,(b+2)}$$

$$+\frac{r\,(r-1)\left\{(w+1)^{b+3}-w^{b+3}1(r=2)\right\}}{(b+1)\,(b+2)\,(b+3)}-\ldots,$$

giving the result. ■

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press

The following corollary is also used.

COROLLARY 1. *For real* $b > -1$ *and integer* $r$,

$$H(1;r,b) = r!\Gamma(b+1)\sum_{\ell=0}^{r}\frac{(-1)^{\ell}}{(r-\ell)!\Gamma(\ell+b+2)}2^{\ell+b+1} - (-1)^{r}B(r+1,b+1).$$

## REFERENCES

Csorgo, S. & J. Mielniczuk (1995) Nonparametric regression under long-range dependent normal errors. *Annals of Statistics* 23, 1000–1014.

Fox, R. & M.S. Taqqu (1985) Non-central limit theorems for quadratic forms in random variables having long-range dependence. *Annals of Probability* 13, 428–446.

Gradshteyn, I.R. & I.M. Ryzhik (1994) *Table of Integrals, Series and Products.* Academic Press.

Hall P. & J.D. Hart (1990) Nonparametric regression with long-range dependence. *Stochastic Processes and Their Applications* 36, 339–351.

Knuth, D.E. (1968) *The Art of Computer Programming, vol. 1. Fundamental Algorithms.* Addison-Wesley.

Robinson, P.M. (1997) Large-sample inference for nonparametric regression with dependent errors. *Annals of Statistics* 28, 2054–2083.

Robinson, P.M. (2007) Nonparametric regression with spatial data. Working paper, London School of Economics.

Roussas, G., L.T. Tran, & O.A. Ioannides (1992) Fixed design regression for time series: Asymptotic normality. *Journal of Multivariate Analysis* 40, 262–291.

Yong, C. H. (1974) *Asymptotic Behaviour of Trigonometric Series.* Chinese University of Hong Kong.

https://doi.org/10.1017/S0266466609990302 Published online by Cambridge University Press