Varying Kernel Density Estimator for a Positive Time Series

N. Balakrishna and Hira L. Koul
Technical University of Cochin, Kerala and Michigan State University

Abstract

This paper analyzes the large sample of a varying kernel density estimator of the marginal density of a nonnegative stationary and ergodic time series that is also strongly mixing. In particular we obtain an approximation for bias, mean square error and establish asymptotic normality of this density estimator.

1 Introduction

Nonnegative time series often arise in real world applications. A class of nonnegative time series that have seen increased research activity in the last two decades are the so called multiplicative error models. Engle and Russell (1998), Engle (2002), Manganelli (2005), Chou (2005), Engle and Gallo (2006), and Brownlees, Cipollini and Gallo (2012) used these models for analyzing financial durations, trading volume of orders, high-low range of asset prices, absolute value of daily returns, and realized volatility, respectively. Several other applications and properties of these models are discussed in Bauwens and Giot (2001), Bauwens and Veredas (2004), Fernandes and Gramg (2006), Gao, Kim and Saart (2015), among others. Pacarur (2006) and Hautsch (2011) discuss numerous examples of nonegative time series useful in economics and finance, and some of their theoretical properties.

It is of interest to estimate the stationary density of a give nonnegative stationary time series nonparametrically. One way to proceed would be to use a conventional non-parametric kernel estimation method based on a symmetric kernel. There is a vast literature on the asymptotic properties of the nonparametric density estimators based on symmetric kernel for i.i.d. r.v.'s as well as for strongly mixing stationary time series, see, e.g., and Härdle, Litkepohn and Chen (1997), Bosq (1998) and Nze and Doukhan (2004), and references therein.

If the underlying r.v.'s are nonnegative or belong to a finite set, then the use of symmetric kernel for its estimation is not fully justified as it assigns positive mass outside the support set, which leads to the so called the edge effect problem. Such estimators are heavily biased in the tails of the bounded support set. To overcome this problem of boundary bias, several asymmetric kernels have been introduced in the literature in the last two decades. Bagai and Prakasa Rao (1995) proposed kernel type estimators for the density function of nonnegative r.v.'s, where the kernel function is a probability density function on $(0, \infty)$. Chen (1999, 2000) used beta kernel, Chen (2000a) proposed gamma kernel, and Scaillet (2004) introduced inverse gaussian kernel for estimating density functions of non-negative random variables. Chaubey et al. (2012) proposed a density estimator for nonnegative r.v.'s via smoothing of the empirical distribution function using a generalization of Hilles lemma.

Manatsaknov and Sarkasian (2012) (MS) used an inverse gamma type kernel to estimate density of positive i.i.d. r.v.'s. They analyzed its bias and means square error while Koul and Song (2013) (KS) established its asymptotic normality, under the i.i.d. set up. This kernel is given by

$$K_\alpha(y,u) = \frac{1}{u\Gamma(\alpha+1)}\Big(\frac{\alpha y}{u}\Big)^{\alpha+1}\exp\{-(\frac{\alpha y}{u})\}, \qquad \alpha > 0, u > 0, y > 0. \tag{1.1}$$

Several properties of this kernel have been nicely described in the papers of MN and KS. As mentioned in KS, for each x, $K_\alpha(x,\cdot)$ is the density of an Inverse Gamma r.v. with shape parameter $\alpha+1$ and scale parameter αx, having mean x; for each t, $\alpha K_\alpha(x,t)/(\alpha+1)$ is a Gamma density with shape parameter $\alpha+2$ and scale parameter t/α. If we let T_α and X_α be a r.v.'s having density $K_\alpha(x,\cdot)$ and $\alpha K_\alpha(\cdot,t)/(\alpha+1)$, respectively, then

$$\sqrt{\alpha}\Big(T_\alpha/x - 1\Big) \to_d N(0,1), \quad \sqrt{\alpha}\Big(X_\alpha/t - 1\Big) \to_d N(0,1), \quad \text{as } \alpha \to \infty.$$

Here, and in the following, $\to_d$ denotes the convergence in distribution. If we let $h = 1/\sqrt{\alpha}$, then from the above facts it follows that as $\alpha \to \infty$,

$$K_\alpha(x,t) \approx \frac{1}{h}\phi\Big(\frac{x/t-1}{h}\Big) \quad \text{or} \quad K_\alpha(x,t) \approx \frac{1}{h}\phi\Big(\frac{t/x-1}{h}\Big),$$

where ϕ denotes the standard normal density. Therefore, the M-S kernel K_α approximately behaves like the standard normal kernel, while the distance between x and t is not the usual Euclidean distance $|x-t|$, but rather the relative distance $|x-t|/t$ or $|x-t|/x$; for the commonly used kernel function, x and t are symmetric in the sense of difference, while in the kernel $K_\alpha(x,t)$, x and t are asymptotically symmetric in the sense of division; the parameter $1/\sqrt{\alpha}$ plays the role of bandwidth as in commonly used kernel set up.

Now, let $\{Y_i, i \in \mathbb{Z}\}$ be a strictly stationary and ergodic time series taking values in the state-space $\mathbb{R}^{=}\{y : 0 \le y < \infty\}$ with marginal stationary density function f. The proposed density estimator for f based on the kernel specified by (1.1) is

$$\hat{f}_n(y) \;=\; \frac{1}{n}\sum_{i=1}^n K_{\alpha_n}(y, Y_i) = \frac{1}{n}\sum_{i=1}^n \frac{1}{Y_i\Gamma(\alpha_n+1)}\Big(\frac{\alpha_n y}{Y_i}\Big)^{\alpha_n+1} e^{-\alpha_n y/Y_i}. \tag{1.2}$$

In the next section, we study asymptotic behavior of the bias and the mean square error of $\hat{f}_n(y)$, for each $y \ge 0$. We also establish the asymptotic normality of the estimator $\hat{f}_n(y)$. All limits are taken as $n \to \infty$, unless specified otherwise.

2 Main results

In this section we present an approximation to the bias and mean square error of $\hat{f}_n(y)$ for each y fixed, and an asymptotic normality result, under some assumptions, which we shall now state.

Assumption 1 *(C1) The time series* $\{Y_i, i \in \mathbb{Z}\}$ *is a nonnegative, stationary, and ergodic.*

(C2) The joint density $f_i(.,.)$ *of* (Y_0, Y_i) *and the marginal density* f *of* Y_0 *are bounded, for all* $i \in \mathbb{Z}$.

(C3) The densities f_i *and* f *are twice continuously differentiable with the bounded first and second order derivatives.*

(C4) $\alpha_n \to \infty$, $\sqrt{\alpha_n}/n \to 0$.

(C5) $\{Y_i, i \in \mathbb{Z}\}$ *is strongly mixing with mixing coefficients* $\rho(k)$ *such that* $\rho(k) = O(k^{-\beta})$, *for some* $\beta > 3$.

Assumptions (C1)-(C4) are used to analyze bias and means square error, while all assumptions are used to establish the asymptotic normality of $\hat{f}_n(y)$.

The following two lemmas are fundamental for establishing the asymptotic behavior of the bias and mean square error of $\hat{f}_n(y)$. The first lemma below is proved in KS.

Lemma 2.1 *Let* $g(u, p_k, \lambda_k)$ *be a sequence of probability density functions of inverse gamma distributions with shape parameters* p_k *and rate parameters* λ_k, *i.e.,*

$$g(u, p_k, \lambda_k) = \frac{\lambda_k^p}{\Gamma(p_k)} \big(\frac{1}{u}\big)^{p_k+1} \exp\big(-\frac{\lambda_k}{u}\big), \quad u > 0,\ k = 1, 2, ...$$

Define $p_k = k(\alpha_n + 2) - 1$, $\lambda_k = k\alpha_n y, k = 1, 2, ..., y > 0$ *and let* $\ell(u)$ *be a function such that its second order derivative is continuous and bounded on* $(0, \infty)$. *Then for large* α_n *and for all* $y > 0$ *and* $k \geq 1$,

$$\int_0^\infty g(u, p_k, \lambda_k)\ell(u)du = \ell(y) + \frac{(2-2k)y\ell'(y)}{p_k - 1} + \frac{[(2-2k)^2(p_k-2) + k^2\alpha_n^2]y^2\ell''(y)}{2(p_k-1)^2(p_k-2)} + o(\frac{1}{\alpha_n}).$$

In order to analyze the variance and the mean square error of the above density estimator $\hat{f}_n(y)$, we need to be able obtain a useful expression for the $\mathrm{Cov}(K_\alpha(y, Y_0), K_\alpha(y, Y_i))$, which in turn motivates one to obtain an extension of the above lemma to a bivariate setting where ℓ is a function of two variables. The following lemma gives the needed result in this case. For any function $\ell(u, v)$, its partial derivatives, whenever they exist, are denoted as follows: $\ell_u = \frac{\partial}{\partial u}\ell(.,.)$, $\ell_v = \frac{\partial}{\partial v}\ell(.,.)$, $\ell_{uu} = \frac{\partial^2}{\partial u^2}\ell(.,.)$, $\ell_{vv} = \frac{\partial^2}{\partial v^2}\ell(.,.)$, $\ell_{uv} = \frac{\partial^2}{\partial u \partial v}\ell(.,.)$.

Lemma 2.2 *Suppose assumptions (C1) and (C4) hold. Let y, μ_k, λ_k and g be as in Lemma 2.1. Let $\ell(u,v)$ be a nonnegative function that is twice continuously differentiable on $(0,\infty)^2$ with all its derivatives up to the second order assumed to be bounded. Then,*

$$
\begin{aligned}
(2.1) \qquad & \int_0^\infty \int_0^\infty \ell(u,v) g(u,p_k,\lambda_k) g(v,p_k,\lambda_k) du dv \\
&= \ell(y,y) + \frac{(2-2k)y}{p_k-1}\Big[\ell_u(y,y) + \ell_v(y,y)\Big] \\
&\quad + \Big[\frac{(2-2k)^2(p_k-2) + k^2\alpha_n^2}{2(p_k-1)^2(p_k-2)}\Big] y^2[\ell_{uu}(y,y) + \ell_{vv}(y,y)] \\
&\quad + \frac{(2-2k)^2 y^2 \ell_{uv}(y,y)}{(p_k-1)^2} + o\Big(\frac{1}{\alpha_n}\Big).
\end{aligned}
$$

The proof of this lemmas uses bivariate Taylor expansion and arguments similar to those used in the proof of Lemma 2.1. Details are given in the last section below.

To proceed further we need to define

$$
(2.2) \qquad v_n = \frac{f(y)}{2y\sqrt{\pi}}, \qquad w_n = \frac{2}{\sqrt{\alpha_n}} \sum_{i=1}^{n-1} (1 - \frac{i}{n})\Big(f_i(y,y) - f^2(y)\Big).
$$

We are now ready to analyze the bias and MSE of $\hat{f}_n(y)$, as given in the next lemma.

Lemma 2.3 *Assumpe (C1)-(C4) hold. Then*

$$
\begin{aligned}
(2.3) \qquad B_n(y) &:= E\big(\hat{f}_n(y) - f(y)\big) = \frac{1}{\alpha_n - 1}\frac{y^2 f''(y)}{2} + o(\frac{1}{\alpha_n}), \\
(2.4) \qquad Var(\hat{f}_n(y)) &= \frac{\sqrt{\alpha_n}}{n}(v_n + w_n) + o(\frac{\sqrt{\alpha_n}}{n}) + o\Big(\frac{1}{\alpha_n}\Big), \\
(2.5) \qquad MSE(\hat{f}_n(y)) &= \frac{\sqrt{\alpha_n}}{n}(v_n + w_n) + o\Big(\frac{\sqrt{\alpha_n}}{n}\Big) + o\Big(\frac{1}{\alpha_n}\Big).
\end{aligned}
$$

Proof. To begin with we have

$$
\begin{aligned}
(2.6) \qquad E\big(\hat{f}_n(y)\big) &= E\big(K_{\alpha_n}(y, Y_1)\big) \\
&= \int_0^\infty \frac{1}{z\Gamma(\alpha_n+1)} \big(\frac{\alpha_n y}{z}\big)^{\alpha_n+1} e^{-\alpha_n y/z} f(z) dz \\
&= f(y) + \frac{1}{\alpha_n - 1}\big(\frac{y^2 f''(y)}{2}\big) + o(\frac{1}{\alpha_n}).
\end{aligned}
$$

The last equation is obtained by using Lemma 2.1 with k=1. This equation readily yields that the bias $B_n(y)$ of $\hat{f}_n(y)$ satisfies (2.3).

Following the arguments of KS, for all $0 < u < 1$, we have

$$
E\big[\hat{f}_n(\frac{u}{\alpha_n})\big] = \int_0^\infty g(z,p,\lambda) f(z) dz = f\big(\frac{u}{\alpha_n}\big) + O(\frac{1}{\alpha_n}),
$$

and hence $\hat{f}_n(y)$ does not suffer from the boundary effect.

Next, we evaluate the **variance and the mean square error (MSE) of $\hat{f}_n(y)$.** Clearly

$$
\begin{aligned}
(2.7)\quad \mathrm{Var}(\hat{f}_n(y)) &= \frac{1}{n}\mathrm{Var}(K_{\alpha_n}(y,Y_1)) + \frac{2}{n}\sum_{i=1}^{n-1}(1-\frac{i}{n})\mathrm{Cov}\big(K_{\alpha_n}(y,Y_0),K_{\alpha_n}(y,Y_i)\big) \\
&= \tilde{V}_n(y) + C_n(y), \quad \text{say.}
\end{aligned}
$$

To obtain an approximation for $\tilde{V}_n(y) = n^{-1}\mathrm{Var}(K_{\alpha_n}(y,Y_1))$, consider

$$
\begin{aligned}
E\big(K_{\alpha_n}(y,Y_1)\big)^2 &= \int_0^\infty \frac{1}{z^2\Gamma^2(\alpha_n+1)}(\frac{\alpha_n y}{z})^{2(\alpha_n+1)}e^{-2\alpha_n y/z}f(z)dz \\
&= \frac{\Gamma(2\alpha_n+3)}{\Gamma^2(\alpha_n+1)2^{2\alpha_n+2}}\frac{1}{2\alpha_n y}\int_0^\infty \frac{(2\alpha_n y)^{2\alpha_n+3}}{z^{2\alpha_n+4}}e^{-2\alpha_n y/z}f(z)dz \\
&= \frac{\Gamma(2\alpha_n+3)}{\Gamma^2(\alpha_n+1)2^{2\alpha_n+2}}\frac{1}{2\alpha_n y}\int_0^\infty g(z;2\alpha_n+3,2\alpha_n y)f(z)dz.
\end{aligned}
$$

Now applying Lemma 2.1 with $k=2$, the last integral equals to

$$
\begin{aligned}
& f(y) + \frac{(-2)yf'(y)}{2\alpha_n+2} + \frac{[4(\alpha_n+1)+4\alpha_n^2]y^2f''(y)}{2(2\alpha_n+2)^2(2\alpha_n+1)} + o(\frac{1}{\alpha_n}) \\
&= f(y) - \frac{yf'(y)}{\alpha_n+1} + \frac{y^2f''(y)}{2(2\alpha_n+1)} + o(\frac{1}{\alpha_n}).
\end{aligned}
$$

By Stirling approximation,

$$
\begin{aligned}
\frac{\Gamma(2\alpha_n+3)}{\Gamma^2(\alpha_n+1)} &\approx \frac{2^{2\alpha_n+2+1/2}(\alpha_n+1)^{2\alpha_n+2+1/2}e^{-2}}{\sqrt{2\pi}{\alpha_n}^{2\alpha_n+1}} \\
&= \frac{e^{-2}}{2y\sqrt{\pi}}\big(1+\frac{1}{\alpha_n}\big)^{2\alpha_n}\big(1+\frac{1}{\alpha_n}\big)^{5/2}\sqrt{\alpha_n}.
\end{aligned}
$$

Thus

$$
E\big(K_{\alpha_n}(y,Y_1)\big)^2 = \frac{\sqrt{\alpha_n}}{2y\sqrt{\pi}}\big[1+o(1)\big]\Big[f(y) - \frac{yf'(y)}{\alpha_n+1} + \frac{y^2f''(y)}{2(2\alpha_n+1)} + o(\frac{1}{\alpha_n})\Big],
$$

and hence

$$
\frac{1}{n}E\big(K_{\alpha_n}(y,Y_1)\big)^2 = \frac{\sqrt{\alpha_n}f(y)}{2y\sqrt{\pi}n} + o(\frac{\sqrt{\alpha_n}}{n}).
$$

Upon combining this with (2.6), we obtain

$$
(2.8)\qquad \tilde{V}_n(y) = \frac{\sqrt{\alpha_n}f(y)}{2y\sqrt{\pi}n} + o(\frac{\sqrt{\alpha_n}}{n}) = O(\frac{\sqrt{\alpha_n}}{n}).
$$

Next consider the $C_n(y)$ term that involves the covariances. With $f_i(u,v)$ denoting the joint density of (Y_0,Y_i), let $f_{i,uv}$ denote their partial derivatives etc. We have

$$
\mathrm{Cov}\big(K_{\alpha_n}(y,Y_0),K_{\alpha_n}(y,Y_i)\big) = E\big(K_{\alpha_n}(y,Y_0)K_{\alpha_n}(y,Y_i)\big) - \Big(E(K_{\alpha_n}(y,Y_0))\Big)^2.
$$

For the sake of brevity, write $g_k(u) = g(u, p_k, \lambda_k)$. Note that for $k = 1$, from (2.1) of Lemma 2.2, we obtain

$$\int_0^\infty \int_0^\infty \ell(u,v) g_1(u) g_1(v) dudv = \ell(y,y) + \frac{y^2}{2(\alpha_n - 1)}\Big[\ell_{uu}(y,y) + \ell_{vv}(y,y)\Big] + o\Big(\frac{1}{\alpha_n}\Big). \tag{2.9}$$

Now take $\ell(u,v) = f_i(u,v)$ in (2.9) to obtain

$$\begin{aligned}
&\mathrm{E}\big(K_{\alpha_n}(y,Y_0)K_{\alpha_n}(y,Y_i)\big)\\
&= \int_0^\infty \int_0^\infty K_{\alpha_n}(y,u)K_{\alpha_n}(y,v) f_i(u,v) dudv\\
&= \int_0^\infty \int_0^\infty \frac{1}{u\Gamma(\alpha_n+1)}\big(\frac{\alpha_n y}{u}\big)^{\alpha_n+1} e^{-\alpha_n y/u}\\
&\qquad \frac{1}{v\Gamma(\alpha_n+1)}\big(\frac{\alpha_n y}{v}\big)^{\alpha_n+1} e^{-\alpha_n y/v} f_i(u,v) dudv\\
&= \int_0^\infty \int_0^\infty g(u,p_1,\lambda_1) g(v,p_1,\lambda_1) f_i(u,v) dudv\\
&= f_i(y,y) + \frac{y^2}{2(\alpha_n-1)}\Big[f_{i,uu}(y,y) + f_{i,vv}(y,y)\Big] + o\Big(\frac{1}{\alpha_n}\Big).
\end{aligned}$$

Note that we used assumptions (C2)-(C4) here. This fact together with (2.6) yields

$$\begin{aligned}
&\mathrm{Cov}\big(K_{\alpha_n}(y,Y_0), K_{\alpha_n}(y,Y_i)\big)\\
&= f_i(y,y) + \frac{y^2}{2(\alpha_n-1)}\Big[f_{i,uu}(y,y) + f_{i,vv}(y,y)\Big] + o\Big(\frac{1}{\alpha_n}\Big)\\
&\qquad -\Big(f(y) + \frac{1}{\alpha_n-1}\big(\frac{y^2 f''(y)}{2}\big) + o\big(\frac{1}{\alpha_n}\big)\Big)^2.\\
&= f_i(y,y) - f^2(y) + \frac{y^2}{2(\alpha_n-1)}\Big[f_{i,uu}(y,y) + f_{i,vv}(y,y) - 2f(y)f''(y)\Big]\\
&\qquad + o\Big(\frac{1}{\alpha_n}\Big) + o\Big(\frac{1}{\alpha_n^2}\Big).
\end{aligned}$$

Hence,

$$\begin{aligned}
C_n(y) &= \frac{2}{n}\sum_{i=1}^{n-1}(1-\frac{i}{n})\mathrm{Cov}\big(K_{\alpha_n}(y,Y_0), K_{\alpha_n}(y,Y_i)\big)\\
&= \frac{2}{n}\sum_{i=1}^{n-1}(1-\frac{i}{n})\Big(f_i(y,y) - f^2(y)\Big)\\
&\quad + \frac{2}{n}\sum_{i=1}^{n-1}\Big((1-\frac{i}{n})\frac{y^2}{2(\alpha_n-1)}\Big[f_{i,uu}(y,y) + f_{i,vv}(y,y) - 2f(y)f''(y)\Big]\Big)\\
&\quad + o\Big(\frac{1}{\alpha_n}\Big) + o\Big(\frac{1}{\alpha_n^2}\Big)
\end{aligned}$$

$$= \frac{2}{n}\sum_{i=1}^{n-1}(1-\frac{i}{n})\Big(f_i(y,y)-f^2(y)+o\Big(\frac{1}{\alpha_n}\Big)\Big).$$

Again, the last equation is a result of boundedness of the functions and their derivatives guaranteed by the assumptions (C2) and (C3). This result, combined with (2.6), (2.8) and the definition (4.12), readily yields (2.4).

The claim (2.5) about the mean square error now readily follows from (2.3), (2.4) and the fact that $\mathrm{MSE}(\hat{f}_n(y)) := E\big(\hat{f}_n(y)-f(y)\big)^2 = B_n^2(y)+\mathrm{Var}(\hat{f}_n(y))$.

Asymptotic normality of $\hat{f}_n$. Here we shall establish the asymptotic normality of the density estimator $\hat{f}_n(y)$ by applying the central limit theorem for strongly mixing triangular arrays proved by Ekstrom (2014), which is stated here for the sake of completeness as the following lemma.

Lemma 2.4 *Let $\{X_{n,i}, 1\le i\le d_n\}$ be a triangular array of strongly mixing sequences, with $\rho_n(.)$ as the mixing coefficients corresponding to the n-th row. Define $\bar{X}_{n,d_n}=\frac{1}{d_n}\sum_{i=1}^{d_n}X_{n,i}$, and assume the following two conditions. (B1) $E|X_{n,i}-E(X_{n,i})|^{2+\delta}<c$ for some $c>0$ and $\delta>0$, for all n,i*

(B2) $\sum_{k=0}^{\infty}(k+1)^2\rho_n^{\delta/(4+\delta)}(k)<c$ for some $c>0$ and all n.

Then

$$\frac{\bar{X}_{n,d_n}-E(\bar{X}_{n,d_n})}{\sqrt{Var(\bar{X}_{n,d_n})}} \xrightarrow{w} N(0,1).$$

We apply this lemma to obtain the following theorem:

Theorem 2.1 *Suppose the conditions (C1) - (C6) hold. Then,*

$$\Big(n/\sqrt{\alpha_n}\Big)^{1/2}\big(v_n+w_n\big)^{-1/2}\big(\hat{f}_n(y)-f(y)-\frac{y^2f''(y)}{2(\alpha_n-1)}\big)\xrightarrow{w} N(0,1). \tag{2.10}$$

where v_n and w_n are defined in (4.12).

Proof: The strongly mixing property of $\{Y_i\}$ implies that of the triangle array $X_{n,i} := \{K_{\alpha_n}(y,Y_i)\}$. Assumptions $(C5)$ and $(C6)$ imply conditions $(B1)$ and $(B2)$ for these $\{X_{n,i}\}$. Now applying Lemma 2.4 to these X_{ni} with $d_n\equiv n$ and substituting for the expectation and variance, we obtain (2.10).

3 Proof of Lemma 2.2

Proof of Lemma 2.2. The proof of this lemmas uses bivariate Taylor expansion and arguments similar to those used in the proof of Lemma 2.1. Accordingly, for $y>0$ let μ_k be the mean of the inverse gamma distribution with parameters p_k and λ_k and express

$$\mu_k = \frac{\lambda_k}{p_k - 1} = y + \frac{(2-2k)y}{p_k - 1}.$$

Let $\ell(u, v)$ be a bivariate function satisfying the assumed conditions. Expanding $\ell(\mu_k, \mu_k)$ around (y, y) and using Taylor's theorem for bivariate function we rewrite,

$$
\begin{aligned}
(3.1)\quad & \ell(\mu_k, \mu_k) \\
&= \ell(y, y) + \frac{(2-2k)y\ell_u(y,y)}{p_k - 1} + \frac{(2-2k)y\ell_v(y,y)}{p_k - 1} \\
&\quad + \frac{1}{2}\Big[\frac{(2-2k)^2y^2\ell_{uu}(\xi,\xi)}{(p_k-1)^2} + \frac{(2-2k)^2y^2\ell_{vv}(\xi,\xi)}{(p_k-1)^2} + 2\frac{(2-2k)^2y^2\ell_{uv}(\xi,\xi)}{(p_k-1)^2}\Big] \\
&= \ell(y) + A_k(y) + \frac{1}{2}B_k(y), \quad \text{say,}
\end{aligned}
$$

where ξ is a value between $\mu_k = y + (2-2k)y/(p_k - 1)$ and y.

Next, consider the Taylor series expansion of $\ell(u, v)$ around (μ_k, μ_k).

$$
\begin{aligned}
(3.2)\quad & \ell(u, v) \\
&= \ell(\mu_k, \mu_k) + (u - \mu_k)\ell_u(\mu_k, \mu_k) + (v - \mu_k)\ell_v(\mu_k, \mu_k) \\
&\quad + \frac{1}{2}\Big[(u-\mu_k)^2\ell_{uu}(\tilde{u}, \tilde{v}) + (v-\mu_k)^2\ell_{vv}(\tilde{u}, \tilde{v}) + 2(u-\mu_k)(v-\mu_k)\ell_{uv}(\tilde{u}, \tilde{v})\Big], \\
&= \ell(\mu_k, \mu_k) + C_k(u, v) + \frac{1}{2}D_k(u, v), \quad \text{say,}
\end{aligned}
$$

where $\tilde{u}$ and $\tilde{v}$ are the values between u and μ_k and v and μ_k, respectively.

Now rewrite

$$
\begin{aligned}
\ell_{uu}(\tilde{u}, \tilde{v}) &= \ell_{uu}(\mu_k, \mu_k) + [\ell_{uu}(\tilde{u}, \tilde{v}) - \ell_{uu}(\mu_k, \mu_k)], \\
\ell_{uv}(\tilde{u}, \tilde{v}) &= \ell_{uv}(\mu_k, \mu_k) + [\ell_{uv}(\tilde{u}, \tilde{v}) - \ell_{uv}(\mu_k, \mu_k)], \\
\ell_{vv}(\tilde{u}, \tilde{v}) &= \ell_{vv}(\mu_k, \mu_k) + [\ell_{vv}(\tilde{u}, \tilde{v}) - \ell_{vv}(\mu_k, \mu_k)].
\end{aligned}
$$

Then we have the decomposition $D_k = D_{k1} + D_{k2} + D_{k3} + D_{k4} + D_{k5} + D_{k6}$, where

$$
\begin{aligned}
D_{k1}(u, v) &:= (u-\mu_k)^2\ell_{uu}(\mu_k, \mu_k), \quad D_{k2} := (v-\mu_k)^2\ell_{vv}(\mu_k, \mu_k), \\
D_{k3}(u, v) &:= 2(u-\mu_k)(v-\mu_k)\ell_{uv}(\mu_k, \mu_k), \\
D_{k4}(u, v) &:= (u-\mu_k)^2[\ell_{uu}(\tilde{u}, \tilde{v}) - \ell_{uu}(\mu_k, \mu_k)], \\
D_{k5}(u, v) &:= (u-\mu_k)^2[\ell_{vv}(\tilde{u}, \tilde{v}) - \ell_{vv}(\mu_k, \mu_k)], \\
D_{k6}(u, v) &:= 2(u-\mu_k)(v-\mu_k)[\ell_{uv}(\tilde{u}, \tilde{v}) - \ell_{uv}(\mu_k, \mu_k)].
\end{aligned}
$$

For the sake brevity, let $g_k(u) \equiv g(u, p_k, \lambda_k)$. In what follows, the range of integration is over $(0, \infty)$, unless specified otherwise. Note that $\int\int C_k(u, v)g_k(u)g_k(v)dudv = 0$. Thus from (3.2) and the above derivation we obtain

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \ell(u,v)g_k(u)g_k(v)dudv \\
&= \ell(\mu_k,\mu_k) + \frac{1}{2}\int_0^\infty \int_0^\infty D_k(u,v)g_k(u)g_k(v)dudv \\
&= \ell(\mu_k,\mu_k) + \frac{1}{2}\Big(I_1 + I_2 + I_3\Big) + \frac{1}{2}\Big(I_4 + I_5 + I_6\Big), \quad \text{(say)},
\end{aligned}$$

where $I_j := \int\int D_{kj}(u,v)dudv$, $j = 1,\cdots,6$. Because g_k is a density with mean μ_k and variance $\tau^2(k,\alpha_n) = k^2\alpha_n^2 y^2/(p_k-1)^2(p_k-2)$, $I_3 = 0$, and

$$I_1 = \ell_{uu}(\mu_k,\mu_k)\int_0^\infty (u-\mu_k)^2 g_k(u)g_k(v)du = \ell_{uu}(\mu_k,\mu_k)\tau^2(k,\alpha_n).$$

Similarly, $I_2 = \ell_{vv}(\mu_k,\mu_k)\tau^2(k,\alpha_n)$.

Next, arguing as in the proof of Lemma 2.1, see (10), (12), (14) and (15) in KS, one can show that

$$I_j = o(\frac{1}{\alpha_n}), \quad j = 4,5,6.$$

Thus

$$\int\int \ell(u,v)g_k(u)g_k(v)dudv = \ell(\mu_k,\mu_k) + \frac{1}{2}\Big[\ell_{uu}(\mu_k,\mu_k) + \ell_{vv}(\mu_k,\mu_k)\Big]\tau^2(k,\alpha_n) + o\Big(\frac{1}{\alpha_n}\Big).$$

Because ξ and μ_k approach y for large α_n and all the partial derivatives are assumed to be continuous, one can replace $\ell(\mu_k,\mu_k)$ by the expression given in the right hand side of (3.1). Hence, upon plugging in the value of $\tau^2(k,\alpha_n)$, for α_n large, to obtain

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \ell(u,v)g_k(u)g_k(v)dudv \\
&= \ell(y,y) + \frac{(2-2k)y\ell_u(y,y)}{p_k-1} + \frac{(2-2k)y\ell_v(y,y)}{p_k-1} \\
&\quad + \frac{1}{2}\Big[\frac{(2-2k)^2y^2\ell_{uu}(y,y)}{(p_k-1)^2} + \frac{(2-2k)^2y^2\ell_{vv}(y,y)}{(p_k-1)^2} + 2\frac{(2-2k)^2y^2\ell_{uv}(y,y)}{(p_k-1)^2}\Big] \\
&\quad + \frac{1}{2}\Big[\ell_{uu}(y,y) + \ell_{vv}(y,y)\Big]\tau(k,\alpha_n) + o\Big(\frac{1}{\alpha_n}\Big) \\
&= \ell(y,y) + \frac{(2-2k)y}{p_k-1}\Big[\ell_u(y,y) + \ell_v(y,y)\Big] \\
&\quad + \Big[\frac{(2-2k)^2(p_k-2) + k^2\alpha_n^2}{2(p_k-1)^2(p_k-2)}\Big]y^2[\ell_{uu}(y,y) + \ell_{vv}(y,y)] \\
&\quad + \frac{(2-2k)^2y^2\ell_{uv}(y,y)}{(p_k-1)^2} + o\Big(\frac{1}{\alpha_n}\Big).
\end{aligned}$$

This completes the proof of the Lemma 2.2.

4 Estimation of the autoregressive function in MEM model.

We shall now consider the problem of estimating the conditional mean function, given the past, of a Markovian multiplicative error time series model. To describe the multiplicative error model of interest here, let Y_i, $i \in \mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}$, be a discrete time nonnegative stationary process. A Markovian multiplicative error model takes the form

$$Y_i = \tau(Y_{i-1})\varepsilon_i, \quad i \in \mathbb{Z}, \tag{4.1}$$

for some positive measurable function τ defined on $R^+ := [0, \infty)$. Here $\varepsilon_i, i \in \mathbb{Z}$ are independent and identically distributed (i.i.d.) non-negative error random variables (r.v.'s) with $E(\varepsilon_0) = 1$, $E(\varepsilon_0^2) < \infty$. Moreover, ε_i is assumed to be independent of the past information $Y_j, j < i$, for all $i \in \mathbb{Z}$. Thus $\tau(y) = E(Y_i | Y_{i-1} = y)$.

The problem of interest here is to estimate τ. We propose the following estimator for this function based on the kernel K_{α_n}.

$$\hat{\psi}_n(y) = \frac{\sum_{i=1}^n K_{\alpha_n}(y, Y_{i-1})Y_i}{\sum_{i=1}^n K_{\alpha_n}(y, Y_i)} = \frac{\hat{\phi}_n(y)}{\hat{f}_n(y)}, \tag{4.2}$$

where $\hat{\phi}_n(y) = n^{-1}\sum_{i=1}^n K_{\alpha_n}(y, Y_{i-1})Y_i$. Let $\phi(y) = E(K_{\alpha_n}(y, Y_0)Y_1)$, and $\psi(y) := \phi(y)/f(y)$. The following decomposition (Bosq (1998), pp 70) is useful for analyzing the asymptotic properties of the estimator $\hat{\psi}_n$. Consider, suppressing y for simplicity,

$$\hat{\psi}_n - \psi = (\hat{\psi}_n - \psi)\big(\frac{f - \hat{f}_n}{f}\big) + \frac{\psi}{f}(f - \hat{f}_n) + \frac{\hat{\phi}_n - \phi}{f}.$$

Hence

$$E\big(\hat{\psi}_n - \psi\big)^2 = A_n + B_n + C_n, \tag{4.3}$$

where

$$\begin{aligned} A_n &:= \frac{\psi^2}{f^2}E(f - \hat{f}_n)^2 + \frac{1}{f^2}E(\hat{\phi}_n - \phi)^2 + \frac{2\psi}{f^2}E\big((f - \hat{f}_n)(\hat{\phi}_n - \phi))\big), \\ B_n &:= \frac{1}{f^2}E\big((\hat{\psi}_n - \psi)^2(f - \hat{f}_n)^2\big) + \frac{2\psi}{f^2}E\big((\hat{\psi}_n - \psi)(f - \hat{f}_n)^2\big) \\ C_n &:= \frac{2}{f^2}E\big((\hat{\psi}_n - \psi)(f - \hat{f}_n)(\hat{\phi}_n - \phi))\big). \end{aligned}$$

We have already analyzed $E(f - \hat{f}_n)^2$, in (2.5). Now consider the second term of A_n,

$$E(\hat{\phi}_n - \phi)^2 = \text{Bias}^2(\hat{\phi}_n) + \text{Var}\big(\hat{\phi}_n(y)\big). \tag{4.4}$$

Calculations similar to the one used in analyzing the bias of $\hat{f}_n$ yields that

$$\text{Bias}\big(\hat{\phi}_n(y)\big) = E(\hat{\phi}_n(y)) - \phi(y) = E(K_{\alpha_n}(y, Y_0)\psi(Y_0)) - \psi(y)f(y) = O(\frac{1}{\alpha_n}). \tag{4.5}$$

Now consider the

(4.6) $\text{Var}(\hat{\phi}_n(y))$

$$= \frac{1}{n}\text{Var}\big(K_{\alpha_n}(y, Y_0)Y_1\big) + \frac{2}{n}\sum_{i=1}^{n-1}(1 - \frac{i}{n})\text{Cov}\big(K_{\alpha_n}(y, Y_1)Y_2, K_{\alpha_n}(y, Y_{1+i})Y_{i+2}\big).$$

Recall from (4.1) that ε_i are i.i.d. with mean 1 and constant variance, σ^2, and that ε_i is independent of Y_{i-1}, for all $i \in \mathbb{Z}$. Now, consider

$$\begin{aligned}\text{Var}\big(K_{\alpha_n}(y, Y_0)Y_1\big) &= \text{Var}\big(K_{\alpha_n}(y, Y_0)\tau(Y_0)\varepsilon_1\big)\\ &= \sigma^2\, E\big(K^2_{\alpha_n}(y, Y_0)\tau^2(Y_0)\big) - \big[E(K_{\alpha_n}(y, Y_0)\tau(Y_0))\big]^2.\end{aligned}$$

But

$$\begin{aligned}E\big(K^2_{\alpha_n}(y, Y_0)\tau^2(Y_0)\big) &= \int_0^\infty K^2_{\alpha_n}(y, z)\tau^2(z)f(z)dz\\ &= \frac{\Gamma(2\alpha_n + 3)}{\Gamma^2(\alpha_n + 1)2^{2\alpha_n+1}(2\alpha_n y)}\int_0^\infty g(z, 2\alpha_n + 3, 2\alpha_n y)\tau^2(z)f(z)dz\\ &\approx \frac{\sqrt{\alpha_n}}{2\sqrt{\pi y}}f(y)\psi^2(y) + O(\frac{1}{\sqrt{\alpha_n}}).\end{aligned}$$

We used Stirling approximation for gamma functions and Lemma 2.1 for the integral with $\ell(z) = \tau^2(z)f(z)$.
(NB: Assumption should take care of the twice differentiability of F and ψ.)) Hence

$$\text{Var}\big(K_{\alpha_n}(y, Y_0)Y_1\big) \approx \frac{\sqrt{\alpha_n}}{2\sqrt{\pi y}}f(y)\psi^2(y) + O(\frac{1}{\sqrt{\alpha_n}}) + (\psi(y)f(y))^2 + O(\frac{1}{\alpha_n^2}) + O(\frac{1}{\alpha_n}),$$

so that

$$\frac{1}{n}\text{Var}\big(K_{\alpha_n}(y, Y_0)Y_1\big) = O(\frac{\sqrt{\alpha_n}}{n}). \tag{4.7}$$

Next, we shall obtain a bound for the covariance term in (4.6) via Devydov's inequality as before.

$$\begin{aligned}
&\frac{2}{n}\sum_{i=1}^{n-1}(1-\frac{i}{n})\mathrm{Cov}\big(K_{\alpha_n}(y,Y_1)Y_2, K_{\alpha_n}(y,Y_{1+i})Y_{i+2}\big)\\
&\leq\ 4\sum_{i=1}^{n-1}(1-\frac{i}{n})\sqrt{\rho(i)}\big[E(K_{\alpha_n}(y,Y_1)Y_2)^4\big]^{1/2}\\
&\leq\ \frac{8}{(2\pi)^{3/4}y^{3/2}}\big(\frac{\alpha_n^{3/4}}{n}\big)\big[\psi^4(y)f(y)+O(\frac{1}{\alpha_n})\big]^{1/2}\sqrt{2}\sum_{i=1}^{n-1}(1-\frac{i}{n})\sqrt{\rho(i)}.
\end{aligned}$$

Upon combining this with (4.7) and (4.6) we obtain that

$$\begin{aligned}
\mathrm{Var}(\hat{\phi}_n(y)) \ \approx\ & \frac{\sqrt{\alpha_n}}{2n\sqrt{\pi}y}f(y)\psi^2(y)+O\big(\frac{1}{n\sqrt{\alpha_n}}\big)+\frac{1}{n}(\psi(y)f(y))^2+O\big(\frac{1}{n\alpha_n}\big)\\
&+\frac{8}{(2\pi)^{3/4}y^{3/2}}\big(\frac{\alpha_n^{3/4}}{n}\big)\big[\psi^4(y)f(y)+O(\frac{1}{\alpha_n})\big]^{1/2}\sqrt{2}\sum_{i=1}^{n-1}(1-\frac{i}{n})\sqrt{\rho(i)}.
\end{aligned}$$

Hence

$$\mathrm{Var}(\hat{\phi}_n(y)) = O\big(\frac{\alpha_n^{3/4}}{n}\big) = o\big(\frac{\alpha_n}{n}\big).$$

This bound together with (4.5) gives that

$$E(\hat{\phi}_n-\phi)^2 = \mathrm{Var}(\hat{\phi}_n(y))+O\big(\frac{1}{\alpha_n}\big) = O\big(\frac{\alpha_n^{3/4}}{n}\big)+O\big(\frac{1}{\alpha_n}\big). \tag{4.8}$$

The Cauchy-Schwarz inequality yields that the absolute value of the third term of A_n is bounded above as follows.

$$\big|E\big((f-\hat{f}_n)(\hat{\phi}_n-\phi)\big)\big| \ \leq\ \sqrt{E(f-\hat{f}_n)^2E(\hat{\phi}_n-\phi)^2} = O\big(\frac{\alpha_n^{3/4}}{n}\big) = o\big(\frac{\alpha_n}{n}\big).$$

Hence

$$A_n = O\big(\frac{\alpha_n^{3/4}}{n}\big) = o\big(\frac{\alpha_n}{n}\big). \tag{4.9}$$

Next consider B_n and for $\gamma > 0$, we write

$$\begin{aligned}
f^2B_n \ &=\ E\big((\hat{\psi}_n^2-\psi^2)(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|>n^{\gamma})}\big)+E\big((\hat{\psi}_n^2-\psi^2)(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|\leq n^{\gamma})}\big)\\
&=\ B_{1n}+B_{2n}, \mathit{say}
\end{aligned}$$

where

$$\begin{aligned}
|B_{1n}| &= |E\big((\hat{\psi}_n^2-\psi^2)(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|>n^\gamma)}\big)| \\
&\le E\big(\hat{\psi}_n^2(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|>n^\gamma)}\big) \\
&\le E\big((maxY_i)^2(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|>n^\gamma)}\big).
\end{aligned}$$

The last inequality follows from the definition (4.2) of $\hat{\psi}_n$. Now by the condition (**to be aded in the Assumptions**)it follows that B_{1n} is negligible. (NB: Write in terms of o())
Next consider

$$\begin{aligned}
|B_{2n}| &= |E\big((\hat{\psi}_n^2-\psi^2)(\hat{f}_n-f)^2.\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)| \\
&\le (n^\gamma+|\psi|)E\big(|\hat{\psi}_n-\psi)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}(\hat{f}_n-f)^2\big) \\
&\le (n^\gamma+|\psi|)E\big(|\hat{\psi}_n-\psi)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}(\hat{f}_n-f)^2\big[\mathcal{I}_{(|\hat{\psi}_n-\psi|\le n^{-(1+\epsilon)\gamma})}+\mathcal{I}_{(|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma})}\big]\big) \\
&\le 2n^\gamma\big[n^{-(1+\epsilon)\gamma}E\big(\hat{f}_n-f\big)^2+E\big((\hat{f}_n-f)^2|\hat{\psi}_n-\psi|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\mathcal{I}_{(|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma})}\big)\big] \\
&\le 2n^\gamma\big[n^{-(1+\epsilon)\gamma}E\big(\hat{f}_n-f\big)^2\big]+2n^\gamma\big[E(\hat{f}_n-f)^4\big]^{1/2}\big[E\big(|\hat{\psi}_n-\psi|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\mathcal{I}_{(|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma})}\big)^2\big]^{1/2} \\
&\le 2n^{-\epsilon\gamma}E\big(\hat{f}_n-f\big)^2 \\
&+ 2n^\gamma\big[E(\hat{f}_n-f)^4\big]^{1/2}\big[E\big(|\hat{\psi}_n-\psi|^{2v}\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)\big]^{1/2v|}\big[P\big((|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma},|\hat{\psi}_n|\le n^\gamma\big)\big]^{1/2w},
\end{aligned} \tag{4.10}$$

where we have used Schwarz inequality initially and then the Holder inequality with $\frac{1}{v}+\frac{1}{w}=1$
We have already analyzed $E\big(\hat{f}_n-f\big)^2$ and we can write

$$\begin{aligned}
\big[E(\hat{f}_n-f)^4\big]^{1/2} &= \big[E(\hat{f}_n-f)^2.(\hat{f}_n-f)^2\big]^{1/2} \\
&\le \big[E((sup\hat{f}_n(y))^2+f^2).(\hat{f}_n-f)^2\big]^{1/2} \\
&\le \big(\frac{c^2\alpha_n}{y^2}+f^2\big)^{1/2}\big[E(\hat{f}_n-f)^2\big]^{1/2},
\end{aligned}$$

where c is a constant (cf K-S). **(NB: Write in terms of O or o)**.
To simplify the next term consider

$$\big[E\big(|\hat{\psi}_n-\psi|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)^2\big]^{1/2} = \big[E\big((\hat{\psi}_n-\psi)^2\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)\big]^{1/2}$$

From the decomposition (4.3), we can write

$$\begin{aligned}
|\hat{\psi}_n-\psi| &\le \frac{|\hat{\psi}_n|}{f}\big(|\hat{f}_n-E(\hat{f}_n)|+|E(\hat{f}_n)-f|\big)+\frac{1}{f}\big(|\hat{\phi}_n-E(\hat{\phi}_n)|+|E(\hat{\phi}_n)-\phi|\big) \\
&= \frac{|\hat{\psi}_n|}{f}\big(|\hat{f}_n-E(\hat{f}_n)|+Bias(\hat{f}_n)\big)+\frac{1}{f}\big(|\hat{\phi}_n-E(\hat{\phi}_n)|+Bias(\hat{\phi}_n)\big) \\
&= \frac{|\hat{\psi}_n|}{f}\big(|\hat{f}_n-E(\hat{f}_n)|\big)+\big(|\hat{\phi}_n-E(\hat{\phi}_n)|\big)+o(1)
\end{aligned} \tag{4.11}$$

So

$$
\begin{aligned}
E\big((\hat{\psi}_n-\psi)^2\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big) &\le \frac{n^{2\gamma}}{f^2}\big(\mathrm{Var}(\hat{\phi}_n)\big)+Var(\hat{f}_n)+2\frac{n^\gamma}{f}\mathrm{Cov}(\hat{\phi}_n,\hat{f}_n)+o(1)\\
&\le \frac{n^{2\gamma}}{f^2}\big(\mathrm{Var}(\hat{\phi}_n)\big)+Var(\hat{f}_n)+2\frac{n^\gamma}{f}\sqrt{\mathrm{Var}(\hat{\phi}_n)\mathrm{Var}(\hat{f}_n)}\big]+O\big(\frac{1}{\alpha_n}\big)
\end{aligned}
$$

(NB: It should be expressed in terms of O(.), powers of n)
Thus we have

$$
E\big((\hat{\psi}_n-\psi)^2\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)=n^{2\gamma}O\big(\frac{\alpha_n^{3/4}}{n}\big)+O\big(\frac{\alpha_n^{3/4}}{n}\big)+n^\gamma O\big(\frac{\alpha_n^{3/4}}{n}\big)+O\big(\frac{1}{\alpha_n}\big)
$$

Next consider the last term in (4.10)
$P\big((|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma},|\hat{\psi}_n|\le n^\gamma\big)\le P\big((|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma}\big)$.
But from (4.11), we can write
$P\big((|\hat{\psi}_n-\psi|>n^{-(1+\epsilon)\gamma}\big)\le P\big[\frac{|\hat{\psi}_n|}{f}\big(|\hat{f}_n-E(\hat{f}_n)|\big)+\frac{1}{f}\big(|\hat{\phi}_n-E(\hat{\phi}_n)|\big)\ge n^{-(1+\epsilon)\gamma}\big]$
So,
$P\big((|\hat{\psi}_n-\psi|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}>n^{-(1+\epsilon)\gamma}\big)$

$$
\begin{aligned}
&\le P\big[\frac{|\hat{\psi}_n|}{f}\big(|\hat{f}_n-E(\hat{f}_n)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)>\frac{n^{-(1+\epsilon)\gamma}}{2}\big]+P\big[\frac{1}{f}\big(|\hat{\phi}_n-E(\hat{\phi}_n)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)>\frac{n^{-(1+\epsilon)\gamma}}{2}\big]\\
&\le P\big[n^\gamma\big(|\hat{f}_n-E(\hat{f}_n)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)>\frac{f.n^{-(1+\epsilon)\gamma}}{2}\big]+P\big[\big(|\hat{\phi}_n-E(\hat{\phi}_n)|\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\big)>\frac{f.n^{-(1+\epsilon)\gamma}}{2}\big]
\end{aligned}
$$

To simplify further we need a bound for the second term. Let $\delta_n=\frac{f.n^{-(1+\epsilon)\gamma)}}{2}$ and define

$$
\begin{aligned}
V_i&=K_{\alpha_n}(y,Y_{i-1})Y_i\mathcal{I}_{(|\hat{\psi}_n|\le n^\gamma)}\\
W_i&=K_{\alpha_n}(y,Y_{i-1})Y_i\mathcal{I}_{(|\hat{\psi}_n|>n^\gamma)}
\end{aligned}
\tag{4.12}
$$

so that $\hat{\phi}_n=\frac{1}{n}\sum_{i=1}^n(V_i+W_i)$
Now

$$
\begin{aligned}
P(|\hat{\phi}_n-E(\hat{\phi}_n)|>\delta_n) &= P\big[|\sum_{i=1}^n\big((V_i+W_i)-E(V_i+W_i)\big)|>n\delta_n\big]\\
&= P\big[|\sum_{i=1}^n(V_i-E(V_i)|>\frac{n\delta_n}{2}\big]+P\big[|\sum_{i=1}^n(W_i-E(W_i)|>\frac{n\delta_n}{2}\big]
\end{aligned}
$$

(4.13)

If we define

$$\begin{aligned} E_n &= \{|\sum_{i=1}^{n}\big(W_i - E(W_i)\big)| > \frac{n\delta_n}{2}\} \\ &= \{|\sum_{i=1}^{n}\big(K_{\alpha_n}(y, Y_{i-1})Y_i\mathcal{I}_{(Y_i>n^\gamma)} - E(K_{\alpha_n}(y, Y_{i-1})Y_i\mathcal{I}_{(Y_i>n^\gamma)})| > \frac{n\delta_n}{2}\}\big) \end{aligned}$$

then for $a > 0$,

$$\begin{aligned} P(E_n) &\leq \sum_{i=1}^{n} P(|Y_i| > n^\gamma) \\ &\leq ne^{-an^\gamma}E(e^{a|Y_i|}) \end{aligned} \tag{4.14}$$

Next consider V_i in (4.12) and let $X_i = V_i - E(V_i)$ **[NB: to use Theorem 4.1, Bosq, p.31, with** $E(X_i) = 0$**].** Consider

$$\begin{aligned} E(V_i - E(V_i))^2 &\leq E(V_i^2) \\ &\leq E\big(K_{\alpha_n}(y, Y_{i-1})Y_i\mathcal{I}_{(Y_i\leq n^\gamma)}\big)^2 \\ &\leq E\big(K_{\alpha_n}(y, Y_{i-1})\psi(Y_{i-1})\varepsilon_i\big)^2 \\ &= (\sigma^2+1)E\big(K^2_{\alpha_n}(y, Y_{i-1})\psi^2(Y_{i-1})\big) \\ &\approx (\sigma^2+1)\frac{\sqrt{\alpha_n}}{2\sqrt{\pi y}}f(y)\psi^2(y) + O\big(\frac{1}{\sqrt{\alpha_n}}\big) = O\big(\sqrt{\alpha_n}\big) \end{aligned}$$

Consider

$$\begin{aligned} |V_i - E(V_i)| &\leq \big(|K_{\alpha_n}(y, Y_{i-1})Y_i\mathcal{I}_{(Y_i\leq n^\gamma)}| + E|K_{\alpha_n}(y, Y_{i-1})Y_i\mathcal{I}_{(Y_i\leq n^\gamma)}|\big) \\ &\leq 2n^\gamma K^*_{\alpha_n} \end{aligned}$$

where $K^*_{\alpha_n} = max\{K_{\alpha_n}(y, Y_{i-1})\} = \frac{c\sqrt{\alpha_n}}{y}$, and c is a constant.
For $k \geq 3$ we can write

$$\begin{aligned} E(|V_i - E(V_i)|)^k &= E\big(|V_i - E(V_i)|^{k-2}.|V_i - E(V_i)|^2\big) \\ &\leq \big(2n^\gamma K^*_{\alpha_n}\big)^{k-2}.E(V_i - E(V_i))^2 \\ &\leq k!\big(2n^\gamma\frac{c\sqrt{\alpha_n}}{y}\big)^{k-2}.E(V_i - E(V_i))^2 \end{aligned}$$

(NB: after verifying many conditions and checking the required algebra we can apply Lemma 1.4 of Bosq. We also have to fix α_n in terms of n for simplification.) Now applying Lemma 1.4 of Bosq (1998), p. 31, we can show that $B_n = o(n^{1/5})$??. The last term C_n in (4.3) can be simplified (to be done) similarly. Thus the MSE of $\hat{\psi}(y)$ tends to zero.

References

[1] Bagai, Isha and Prakasa Rao, B. L. S. (995). Kernel type density estimates for positive valued random variables. *Sankhya Ser. A*, **57**, 56–67.

[2] Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes, 2nd Edition.* Springer, Berlin.

[3] Bauwens, L. and Giot, P. (2001). *Econometric Modeling of Stock Market Intraday Activity.* Kluwer Academic Publishers, Boston, Dordrecht, London. Advanced Studies in Theoretical and Applied Econometrics, vol. 38.

[4] Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *Journal of Econometrics*, **119**(2), 381–412.

[5] Brownlees, C. T., Cipollini, F., and Gallo, G. M. (2012). Multiplicative error models. In L. Bauwens, C. Hafner, and S. Laurent, editors, *Handbook of Volatility Models and Their Applications*, chapter 9, pages 225–247. John Wiley and Sons, Inc, Hoboken, New Jersey.

[6] Chaubey, Y. P. and Sen, A., and Sen, Pranab K. (2012). A new smooth density estimator for non-negative random variables. *J. of India Statistical Association*, **50**, 83–104.

[7] Chen, Y.-T. and Hsieh, C.-S. (2010). Generalized moment tests for autoregressive conditional duration models. *Journal of Financial Econometrics*, **8**(3), 345–391.

[8] Chen, S.X. (1999), Beta kernel estimators for density functions. *Computational Statistics & Data Analysis*, **31**, 131-145.

[9] Chen, S.X. (2000a). Beta kernel smoothers for regression curves. *Statistica Sinica*, **10**, 73-91.

[10] Chen, S.X. (2000b). Probability density function estimation using gamma kernels. *Ann. of the Institute of Statist. Math.*, **52**, 471-480.

[11] Chen, S.X. (2002). Local linear smoothers using asymmetric kernels. *Ann. of the Institute of Statist. Math.*, **54**, 312-323.

[12] Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. *Journal of Money, Credit and Banking*, **37**(3), 561–82.

[13] Christensen, T., Hurn, A., and Lindsay, K. (2012). Forecasting spikes in electricity prices. *International Journal of Forecasting*, **28**(2), 400–411.

[14] Ekstrum, M. (2014). A general central limit theorem for strong mixing sequences. *Statistics and Probability Letters*, **94**, 236–238.

[15] Engle, R. F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, **17**(5), 425–446.

[16] Engle, R. F. and Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, **131**(1-2), 3–27.

[17] Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, **66**(5), 1127–1162.

[18] Fernandes, M. and Grammig, J. (2006). A family of autoregressive conditional duration models. *Journal of Econometrics*, **130**(1), 1–23.

[19] Hautsch, N. (2011). *Econometrics of Financial High-Frequency Data.* Springer.

[20] Härdle, W., Lutkepohl, H. and Chen, R. (1997). A review of nonparametric time series analysis. *Internationl Statistical Review,* 65, 1, 49 - 72.

[21] Koul, Hira L. and Song, Weixing. (2013). Large sample results for varying kernel regression estimates. *J. Nonparametr. Stat.*, **25**, 829-853.

[22] Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets*, **8**(4), 377–399.

[23] Mnatsakanov, R., and Sarkisian, K. (2012). Varying kernel density estimation on. *Statistics & Probability Letters*, **82**, 1337-1345.

[24] Nze, P. A. and Doukhan, P. (2004). Weak dependence: models and applications to econometrics. *Econometric Theory*, **20**, 995–1045.

[25] Pacurar, M. (2008). Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys*, **22**(4), 711–751.

[26] Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels. *Nonparametric Statistics*, **16**, 217-226.