Short Note

A Note on the Asymptotic Normality of the Kernel Deconvolution Density Estimator with Logarithmic Chi-Square Noise

Yang Zu

Department of Economics, City University London, Northampton Square, EC1V 0HB London, UK;
E-Mail: yang.zu@city.ac.uk; Tel.: +44-0-20-7040-8619; Fax: +44-0-20-7040-8580.

Academic Editor: Kerry Patterson

Abstract: This paper studies the asymptotic normality for the kernel deconvolution estimator when the noise distribution is logarithmic chi-square; both identical and independently distributed observations and strong mixing observations are considered. The dependent case of the result is applied to obtain the pointwise asymptotic distribution of the deconvolution volatility density estimator in discrete-time stochastic volatility models.

Keywords: kernel deconvolution estimator; asymptotic normality; volatility density estimation

JEL classifications: C13, C22, C46, C58

---

1. Introduction

Consider the measurement error model:

$$Y = X + \varepsilon,$$

where X is the signal, while ε is the noise. Assume X is independent of ε; X has density f_X, and ε has density k, so the density of Y, denoted as f_Y, is the convolution of f_X and k:

$$f_Y = f_X * k,$$

where the $*$ denotes convolution.

Assume we observe the realizations $Y_1, \ldots, Y_n$ of Y and that the function k is fully known; one possible estimator for f_X from the noisy observations $Y_1, \ldots, Y_n$ is the kernel deconvolution estimator:

$$\hat{f}_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\mathrm{i}tx} \frac{\phi_K(th)\widehat{\phi_{f_Y}}(t)}{\phi_k(t)} \mathrm{d}t, \tag{1}$$

where:

$$\widehat{\phi_{f_Y}}(t) = \frac{1}{n} \sum_{j=1}^{n} e^{\mathrm{i}tY_j},$$

is the empirical characteristic function of density f_Y, $K(x)$ is a kernel function, ϕ_K and ϕ_k are the Fourier transform of K and k, respectively[1]. The kernel deconvolution estimator was first proposed for the measurement error model by Carroll and Hall [1] and Stefanski and Carroll [2].

Define the kernel deconvolution function as follows:

$$\nu_h(x) := \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\phi_K(t)}{\phi_k(t/h)} e^{-\mathrm{i}tx} \mathrm{d}t;$$

the kernel deconvolution estimator can be written compactly as:

$$\hat{f}_X(x) = \frac{1}{nh} \sum_{j=1}^{n} \nu_h\left(\frac{x - Y_j}{h}\right). \tag{2}$$

In this paper, I show the asymptotic normality for the estimator $\hat{f}_X(x)$ when the distribution of ε is logarithmic chi-square. The asymptotic distribution of the kernel deconvolution estimator has been considered in Fan [3], Fan and Liu [4], Van Es and Uh [5] and Van Es and Uh [6] for identically independently distributed (*i.i.d.*) observations. Masry [7] and Kulik [8] consider various cases for the weakly-dependent observations. However, none of the above research allows the error distribution to be the logarithmic chi-square distribution. I consider both the identical and independently distributed (*i.i.d.*) observations and strong mixing observations in this paper, which complements the above-mentioned literature.

The results obtained in this paper can be applied to obtain the asymptotic distribution of the deconvolution volatility density estimator. The problem of estimating volatility density has been gaining increasing interest in econometrics in recent years; see, e.g., Van Es, Spreij, and Van Zanten [9] and Van Es, Spreij, and Van Zanten [10] for the kernel deconvolution estimator, Comte and Genon-Catalot [11] for the penalized projection estimator and Todorov and Tauchen [12] for the study in the context of high-frequency data. Kernel deconvolution with logarithmic chi-square noise arises naturally when estimating the volatility density in stochastic volatility (SV) models. Existing research (e.g., Van Es, Spreij, and Van Zanten [9] and Van Es, Spreij, and Van Zanten [10]) focuses on the convergence rates of the estimators, and the asymptotic distribution of the estimators is not available.

---

[1] The characteristic function of a random variable with density f is defined as $\phi_f = \int_{\mathbb{R}} e^{\mathrm{i}tx} f(x)\mathrm{d}x$.

In Section 2, I review the probabilistic properties of the logarithmic chi-square distribution; Section 3 presents the asymptotic normality of the estimator, for both *i.i.d.* observations and dependent observations; Section 4 discusses the application of the results to volatility density estimation in SV models; Section 5 concludes the paper.

2. Logarithmic Chi-Square Distribution

The logarithmic chi-square distribution is obtained by taking the logarithm of a chi-square distribution with degrees of freedom of one. The density function of logarithmic chi-square distribution is:

$$k(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}x} e^{-\frac{1}{2}e^x}.$$

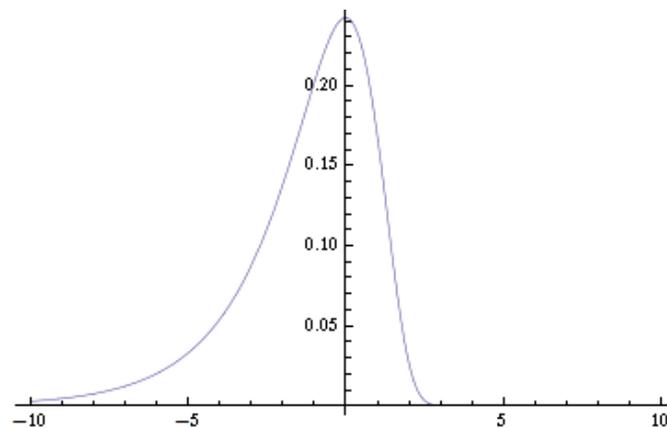The density function of the logarithmic chi-square distribution is asymmetric and is plotted in Figure 1.

Figure 1. Density function of the logarithmic chi-square distribution.

The characteristic function of the logarithmic chi-square distribution is:

$$\phi_k(t) = \frac{1}{\sqrt{\pi}} 2^{\mathrm{i}t} \Gamma\left(\frac{1}{2} + \mathrm{i}t\right),$$

where $\Gamma(.)$ is the gamma function.

Fan [3] studies the quadratic mean convergence rate of the kernel deconvolution estimator; it turns out that the convergence rate of the estimator depends heavily on the type of error distribution. In particular, it is determined by the tail behaviour of the modulus of the characteristic function of the error distribution: the faster the modulus function goes to zero in the tail, the slower the converge rate. The following lemma, which is from Van Es, Spreij, and Van Zanten [10], gives the tail behaviour of $|\phi_k(t)|$.

Lemma 1. *(Lemma 5.1 of Van Es, Spreij, and Van Zanten [10]) For $|t| \to \infty$, we have:*

$$|\phi_k| = \sqrt{2} e^{-\frac{1}{2}\pi|t|}\left(1 + O\left(\frac{1}{|t|}\right)\right), \tag{3}$$

and:

$$\mathrm{Re}\,\phi_k(t) \;=\; |\phi_k|\left(\cos\left[t\log\left(\sqrt{1+4t^2}-t\right)\right]+O\left(\frac{1}{|t|}\right)\right), \tag{4}$$

$$\mathrm{Im}\,\phi_k(t) \;=\; |\phi_k|\left(\sin\left[t\log\left(\sqrt{1+4t^2}-t\right)\right]+O\left(\frac{1}{|t|}\right)\right). \tag{5}$$

From (3), it is known that the modulus of $\phi_k(t)$ decays exponentially fast as $|t| \to \infty$. It thus belongs to the super-smooth density according to the classification in Fan [13]. According to Fan [13], the optimal convergence rate of the estimator is $(\log n)^{-2}$, when $h = (\log n)^{-1}$. Figure 2 plots the modulus function $|\phi_k|$ and its approximation $\sqrt{2}e^{-\frac{1}{2}\pi|t|}$; we notice that the two functions almost coincide at both tails.
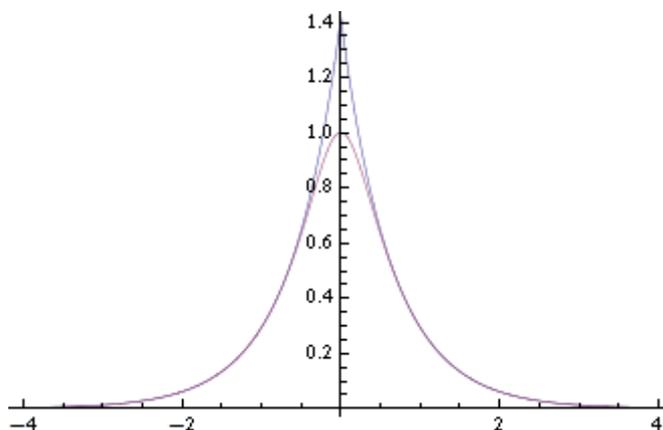

Figure 2. Modulus function of the characteristic function of logarithmic chi-square distribution and its approximation: the higher peak curve is the approximating function $\sqrt{2}e^{-\frac{1}{2}\pi|t|}$.

From (4) and (5), it is known that in both tails, neither the real part nor imaginary part of the characteristic function can dominate the other; this violates the assumptions in the previous works by, e.g., Fan [3] and Masry [7], on studying the asymptotic normality; for super-smooth error distributions, these papers assume either the real part or the imaginary part to be dominant.

3. Asymptotic Normality

In this paper, I consider one particular kernel function, namely the sinc kernel function:

(C1) The sinc kernel function is defined as:

$$K(x) = \frac{\sin(x)}{\pi x},$$

with Fourier transform[2]:

$$\phi_K(t) = I\{|t| \leqslant 1\}.$$

---

[2] In this paper, I follow the convention to define the Fourier transform of a function f to be $\phi_f = \int_{-\infty}^{+\infty} e^{\mathrm{i}tx} f(x)\,\mathrm{d}x$.

The sinc kernel function is favoured in theoretical literature because of the simplicity of its Fourier transform and is thus used here.[3]

3.1. i.i.d. Observations

In this section, I prove the asymptotic normality of the estimator when the observations are *i.i.d.*

Theorem 1. *When the observations are i.i.d. and ε is distributed as logarithmic chi-square, if Assumption (C1) holds, when $\exp(1/h)/n \to 0$ as $n \to \infty$ and $h \to 0$, it holds that,*

$$\frac{\hat{f}_X(x) - K_h * f_X(x)}{\sqrt{\frac{1}{2\pi^2 n} \exp(\pi/h) f_Y(x)}} \to^d N(0,1),$$

where $K_h(x) := (1/h)K(x/h)$.

Proof. Denote:

$$Z_j = \frac{1}{h} \nu_h \left(\frac{x - Y_j}{h} \right),$$

then:

$$\hat{f}(x) = \frac{1}{n} \sum_{j=1}^{n} Z_j.$$

First:

$$
\begin{aligned}
E\hat{f}(x) &= EZ_1 \\
&= E\left[\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{\phi_K(th)\widehat{\phi_{f_Y}}(t)}{\phi_k(t)} dt \right] \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{\phi_K(th)E\left[\widehat{\phi_{f_Y}}(t)\right]}{\phi_k(t)} dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \frac{\phi_K(th)\phi_{f_Y}(t)}{\phi_k(t)} dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \phi_K(th)\phi_{f_X}(t) dt \\
&= K_h * f_X(x),
\end{aligned}
$$

---

[3] Usually, for practical implementations, the following kernels:

$$K_1(x) = \frac{48 \cos x}{\pi x^4} \left(1 - \frac{15}{x^2} \right) - \frac{144 \sin x}{\pi x^5} \left(2 - \frac{5}{x^2} \right),$$

with Fourier transform:

$$\phi_{K_1}(t) = I\{|t| \leqslant 1\} \left(1 - t^2 \right)^3,$$

are used because they have better numerical properties; see Delaigle and Gijbels [14] for the discussions.

Second, I evaluate $\mathrm{Var}\, Z_1$,

$$
\begin{aligned}
\mathrm{Var}\, Z_1 &= \mathrm{Var}\left(\frac{1}{h}\nu_h\left(\frac{x-Y_1}{h}\right)\right) \\
&= \frac{1}{h^2}\left(E\nu_h\left(\frac{x-Y_1}{h}\right)^2 - \left(E\nu_h\left(\frac{x-Y_1}{h}\right)\right)^2\right) \\
&= \frac{1}{h^2}\left(\int \nu_h\left(\frac{x-y}{h}\right)^2 f_Y(y)\mathrm{d}y - (K_h * f_X(x))^2\right) \\
&= \frac{1}{h^2}\left(h\int \nu_h(y)^2\,\mathrm{d}y\, f_Y(x) - (K_h * f_X(x))^2\right) \\
&= \frac{1}{2\pi^2}\exp\left(\frac{\pi}{h}\right)f_Y(x)\left(1 + o(1)\right),
\end{aligned}
\tag{6}
$$

where the last equality is obtained because $K_h * f_X(x) \to f_X(x)$ as $h \to 0$, and $\int |\nu_h(x)|^2\,\mathrm{d}x = \frac{h}{2\pi^2}\exp\left(\frac{\pi}{h}\right)(1 + o(1))$. The latter result is shown as follows,

$$
\begin{aligned}
\int |\nu_h(x)|^2\,\mathrm{d}x &= \frac{1}{2\pi}\int |\phi_{\nu_h}(u)|^2\,\mathrm{d}u \\
&= \frac{1}{2\pi}\int \left|\frac{\phi_K(u)}{\phi_k(u/h)}\right|^2\,\mathrm{d}u \\
&= \frac{h}{\pi}\int_0^{1/h}\left|\frac{1}{\phi_k(u)}\right|^2\,\mathrm{d}u \\
&= \frac{h}{\pi}\left(\int_0^M\left|\frac{1}{\phi_k(u)}\right|^2\,\mathrm{d}u + \int_M^{1/h}\left|\frac{1}{\phi_k(u)}\right|^2\,\mathrm{d}u\right),
\end{aligned}
$$

where M is a very big number. The first term in the brackets is a constant depending on M; the order of the second term can be evaluated as follows,

$$
\begin{aligned}
\int_M^{1/h}\left|\frac{1}{\phi_k(u)}\right|^2\,\mathrm{d}u &= \frac{1}{2\pi}\left(\exp\left(\frac{\pi}{h}\right) - \exp(\pi M)\right) \\
&= \frac{1}{2\pi}\exp\left(\frac{\pi}{h}\right)(1 + o(1)),
\end{aligned}
$$

where I use the fact that when M is big, $|\phi_k(u)|$ can be replaced by its asymptotic approximation. The second term clearly dominates the first term, which is a constant, such that:

$$
\int |\nu_h(x)|^2\,\mathrm{d}x = \frac{h}{2\pi^2}\exp\left(\frac{\pi}{h}\right)(1 + o(1)).
\tag{7}
$$

Here, I use the argument of Butucea [15] to split the integral and show that the tail part of the integral dominates.

A sufficient condition for asymptotic normality is the Lyapunov condition, which reduces to:

$$\frac{E|Z_1 - EZ_1|^{2+\delta}}{n^{\delta/2}[\text{Var}(Z_1)]^{1+\delta/2}} \to 0, \tag{8}$$

for *i.i.d.* data.

For an upper bound for the numerator,

$$
\begin{aligned}
E|Z_1 - EZ_1|^{2+\delta} &\leqslant E|Z_1|^{2+\delta} + |EZ_1|^{2+\delta} \\
&\leqslant 2E|Z_1|^{2+\delta} \\
&= \frac{2}{h^{2+\delta}} \int_{-\infty}^{+\infty} \left| \nu_h \left(\frac{x-y}{h} \right) \right|^{2+\delta} f_Y(y) \mathrm{d}y \\
&\leqslant \frac{C}{h^{2+\delta}} \int_{-\infty}^{+\infty} \left| \nu_h \left(\frac{x-y}{h} \right) \right|^{2+\delta} \mathrm{d}y
\end{aligned} \tag{9}
$$

Now, notice the result from Van Es, Spreij, and Van Zanten [10] and Masry [16] that, for $p > 2$[4],

$$\|\nu_h\|_p \leqslant \|\nu_h\|_\infty^{1-2/p} \|\nu_h\|_2^{2/p}.$$

An upper bound for $\|\nu_h\|_\infty$ is easy to get, as[5]:

$$
\begin{aligned}
\|\nu_h\|_\infty &= \sup_x \left| \frac{1}{2\pi} \int \frac{\phi_K(t)}{\phi_k(t/h)} e^{-\mathrm{i}tx} \mathrm{d}t \right| \\
&\leqslant \frac{1}{2\pi} \int \left| \frac{\phi_K(t)}{\phi_k(t/h)} \right| \mathrm{d}t \\
&\leqslant \frac{\sqrt{2}}{\pi^2} h \exp\left(\frac{\pi}{2h} \right),
\end{aligned}
$$

while $\|\nu_h\|_2^2$ is known from (7), such that:

$$
\begin{aligned}
\int |\nu_h(z)|^p \mathrm{d}z &\leqslant \|\nu_h\|_\infty^{p-2} \|\nu_h\|_2^2 \\
&\leqslant C \times h^{p-2} \exp\left(\frac{\pi(p-2)}{2h} \right) \times h \exp\left(\frac{\pi}{h} \right) \\
&= C \times h^{p-1} \exp\left(\frac{\pi p}{2h} \right),
\end{aligned}
$$

for $p > 2$. Therefore, take $p = 2 + \delta$ and use the result in (9); it then holds that:

$$E|Z_1 - EZ_1|^{2+\delta} \leqslant C \times \exp\left(\frac{\pi(2+\delta)}{2h} \right), \tag{10}$$

this, together with (6), implies that Lyapunov's condition (8) holds, which completes the proof. $\square$

---

[4] This is easy to see by noticing $\int |\nu_h(x)|^p \mathrm{d}x \leqslant \int |\nu_h(x)|^2 |\sup_x \nu_h(x)|^{p-2} \mathrm{d}x$ for $p > 2$.
[5] Here, again, the splitting integral argument as in proving (7) is used; I omit the details for the ease of exposition.

3.2. Strong Mixing Observations

In this section, I consider the model:

$$Y = X + \varepsilon, \tag{11}$$

where X's realizations of $X_1, \cdots, X_n$ are strictly stationary and strong mixing, while the noise realizations $\varepsilon_1, \cdots, \varepsilon_n$ are *i.i.d.* logarithmic chi-square variables, independent of X, such that the observations $Y_1, \cdots, Y_n$ are also strictly stationary and strong mixing.

There are various concepts of dependence; here, I consider the case of α mixing, also called strong mixing, which is the weakest among all of the dependence concepts.

Definition 1. *Let $\{X_t\}$, $t = \cdots, -1, 0, 1, \cdots$ be an infinite sequence of strictly stationary random variables and $\mathcal{F}_i^j$ be the σ-algebra generated by $\{X_t, i \leqslant t \leqslant j\}$; then, the α-mixing coefficient is defined as:*

$$\alpha(k) = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_k^{+\infty}} |P(A)P(B) - P(AB)|.$$

The sequence $\{X_t\}$, $t = \cdots, -1, 0, 1, \cdots$, is called α-mixing if $\alpha(k) \to 0$ as $k \to \infty$.

For the dependent case, a bounded assumption on the joint density of observations is also needed.

(C2) The probability density function of any joint distribution (Y_i, Y_j), $1 \leqslant i < j \leqslant n$, exists and is bounded by a constant.

Now, I give the asymptotic normality theorem. Notice that the mixing assumption here is a litter weaker than that in Masry [7].

Theorem 2. *In model (11), let $X_1, X_2, \cdots, X_n$ be strictly stationary, α-mixing with:*

$$\sum_{k=1}^{\infty} \alpha(k)^{1-2/\delta} < \infty, \tag{12}$$

for some $\delta > 2$; the noises $\varepsilon_1, \cdots, \varepsilon_n$ are i.i.d. logarithmic chi-square variables, independent of X; if (C1) and (C2) hold, when $\exp(1/h)/n \to 0$ as $n \to \infty$ and $h \to 0$, then:

$$\frac{\hat{f}_X(x) - K_h * f_X(x)}{\sqrt{\frac{1}{2\pi^2 n} \exp(\pi/h) f_Y(x)}} \to^d N(0, 1).$$

Proof. First, by strict stationarity and using the ergodic theorem for strong mixing sequences, similarly as in the proof of Theorem 1,

$$E\hat{f}(x) = K_h * f_X(x).$$

Next, the variance of the estimator is evaluated; first:

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{1}{n}\mathrm{Var}(Z_1) + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\mathrm{Cov}(Z_1, Z_{j+1}).$$

Knowing from Theorem 1 that the first term is:

$$\frac{1}{n}\operatorname{Var}(Z_1) = \frac{1}{2\pi^2 n}\exp\left(\frac{\pi}{h}\right)f_Y(x)(1 + o(1)). \tag{13}$$

For the covariance term, first notice:

$$\begin{aligned}
|\operatorname{Cov}(Z_1, Z_{j+1})| &\leqslant |E(Z_1 Z_{j+1})| + (K_h * f_X(x))^2 \\
&\leqslant |E(Z_1 Z_{j+1})| + O(1),
\end{aligned} \tag{14}$$

as $h \to 0$. Now, because:

$$\begin{aligned}
|E(Z_1 Z_{j+1})| &= \frac{1}{h^2}\left|E\left(\nu_h\left(\frac{x - Y_1}{h}\right)\nu_h\left(\frac{x - Y_{j+1}}{h}\right)\right)\right| \\
&= \frac{1}{h^2}\left|E\int\int\frac{\phi_K(t)\phi_K(t')}{\phi_k(t/h)\phi_k(t'/h)}e^{-it\frac{x-Y_1}{h}}e^{-it'\frac{x-Y_{j+1}}{h}}\,dt\,dt'\right| \\
&= \frac{1}{h^2}\left|\int\int\frac{\phi_K(t)\phi_K(t')}{\phi_k(t/h)\phi_k(t'/h)}E\left(E\left(e^{-it\frac{x-X_1-\varepsilon_1}{h}}e^{-it'\frac{x-X_{j+1}-\varepsilon_{j+1}}{h}}\Big|X\right)\right)\,dt\,dt'\right| \\
&= \frac{1}{h^2}\left|\int\int\frac{\phi_K(t)\phi_K(t')}{\phi_k(t/h)\phi_k(t'/h)}\phi_k(t/h)\phi_k(t'/h)E\left(e^{-it\frac{x-X_1}{h}}e^{-it'\frac{x-X_{j+1}}{h}}\right)\,dt\,dt'\right| \\
&= \frac{1}{h^2}\left|\int\int\phi_K(t)\phi_K(t')E\left(e^{-it\frac{x-X_1}{h}}e^{-it'\frac{x-X_{j+1}}{h}}\right)\,dt\,dt'\right| \\
&\leqslant \frac{1}{h^2}\left|\int\int|\phi_K(t)\phi_K(t')|\,E\left(\left|e^{-it\frac{x-X_1}{h}}e^{-it'\frac{x-X_{j+1}}{h}}\right|\right)\,dt\,dt'\right| \\
&\leqslant \frac{1}{h^2}\left|\int\int|\phi_K(t)\phi_K(t')|\,dt\,dt'\right| \\
&\leqslant \frac{C}{h^2},
\end{aligned}$$

where C is a constant; continuing on (14), I get:

$$|\operatorname{Cov}(Z_1, Z_{j+1})| \leqslant C\frac{1}{h^2}(1 + o(1)). \tag{15}$$

On the other hand, using the assumption on the α-mixing coefficients and the covariance inequality for strong mixing sequence in Proposition 2.5 in Fan and Yao [17], for $\delta > 2$,

$$\begin{aligned}
\operatorname{Cov}(Z_1, Z_{j+1}) &\leqslant 8\alpha(j)^{1-2/\delta}\left(E|Z_1|^\delta\right)^{1/\delta}\left(E|Z_{j+1}|^\delta\right)^{1/\delta} \\
&= 8\alpha(j)^{1-2/\delta}\left(E|Z_1|^\delta\right)^{2/\delta} \\
&\leqslant C'\alpha(j)^{1-2/\delta}\exp\left(\frac{\pi}{h}\right).
\end{aligned} \tag{16}$$

Therefore, using (15) and (16),

$$\left| \sum_{j=1}^{n-1} \mathrm{Cov}\left(Z_1, Z_{j+1} \right) \right|$$

$$\leqslant \left| \sum_{j=1}^{m_n} \mathrm{Cov}\left(Z_1, Z_{j+1} \right) \right| + \left| \sum_{j=m_n}^{n-1} \mathrm{Cov}\left(Z_1, Z_{j+1} \right) \right|$$

$$\leqslant C \frac{1}{h^2} m_n + C \exp\left(\frac{\pi}{h} \right) \sum_{j=m_n}^{n-1} \alpha(j)^{1-2/\delta},$$

if one chooses $m_n = \frac{1}{h|\log h|}$, then $m_n \to \infty$ and $m_n h \to 0$, then obviously the first term is $o\left(\exp\left(\frac{\pi}{h} \right) \right)$; the second term is also $o\left(\exp\left(\frac{\pi}{h} \right) \right)$ by noticing the mixing assumption in (12). Then, it is shown that:

$$\left| \sum_{j=1}^{n-1} \mathrm{Cov}\left(Z_1, Z_{j+1} \right) \right| = o\left(\exp\left(\frac{\pi}{h} \right) \right). \tag{17}$$

From (13) and (17), it then follows that:

$$\mathrm{Var}\left(\hat{f}(x) \right) = \frac{1}{2\pi^2 n} \exp\left(\frac{\pi}{h} \right) f_Y(x) \left(1 + o(1) \right).$$

Now, I prove the central limit theorem, for which I use the classical large block-small block argument of proving the central limit theorem for the dependent sequence. First, I make some normalizations, define $\sigma_0 = \left(\frac{1}{2\pi^2} \exp\left(\frac{\pi}{h} \right) f_Y(x) \right)^{1/2}$, and:

$$Z_j' = \frac{Z_j - K_h * f_X(x)}{\sigma_0},$$

then Z_j' has mean zero and unit variance and:

$$\frac{1}{n} \sum_{j=1}^{n} Z_j' = \frac{\hat{f}(x) - K_h * f_X(x)}{\sigma_0};$$

and it will be shown that:

$$\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{n} Z_j' \right) \to^d N(0,1),$$

which is the result that needed to be shown.

First, the set $\{1, \cdots, n\}$ is partitioned into $2k_n + 1$ subsets with large blocks of size l_n and small blocks of size s_n, such that $k_n = \lfloor n/(l_n + s_n) \rfloor$, so the last remaining block has size $n - k_n(l_n + s_n)$. The sizes are such that $l_n \to \infty$, $s_n \to \infty$, $l_n/s_n \to \infty$. Then, we can write:

$$\sum_{j=1}^{n} Z_j' = \sum_{j=1}^{k_n} \xi_j + \sum_{j=1}^{k_n} \eta_j + \zeta,$$

where:

$$\xi_j = \sum_{j'=(j-1)(l_n+s_n)+1}^{(j-1)(l_n+s_n)+l_n} Z'_{j'}$$

$$\eta_j = \sum_{j'=(j-1)(l_n+s_n)+l_n+1}^{j(l_n+s_n)} Z'_{j'}$$

$$\zeta = \sum_{k_n(l_n+s_n)+1}^{n} Z'_j.$$

which are the sum of large blocks, small blocks and the last block, respectively. Then, as a standard procedure for the small block-big block argument, I show the following:

$$\frac{1}{n}E\left(\sum_{j=1}^{k_n}\eta_j\right)^2 = o(1), \tag{18}$$

$$\frac{1}{n}E\zeta^2 = o(1), \tag{19}$$

$$\left|E\exp\left(it\sum_{j=1}^{k_n}\xi_j/\sqrt{n}\right) - \prod_{j=1}^{k_n}E\exp\left(it\xi_j/\sqrt{n}\right)\right| \to 0, \tag{20}$$

$$\frac{1}{n}\sum_{j=1}^{k_n}E\xi_j^2 \to 1, \tag{21}$$

$$\frac{1}{n}\sum_{j=1}^{k_n}E\left[\xi_j^2 I\left(|\xi_j| \geqslant \varepsilon n^{1/2}\right)\right] \to 0, \tag{22}$$

for $\forall \varepsilon > 0$. (18) and (19) say that the small blocks and the last block are of smaller order. (20) says that the large blocks are as if independent in the sense of the characteristic function. Then, (21) and (22) are the Lindeberg-Feller condition for the asymptotic normality for $\sum_{j=1}^{k_n}\xi_j$ under independence.

For (18) and (19), using the moment inequality for the α-mixing sequence in Proposition 2.7 (i) in Fan and Yao [17], it can be shown that:

$$E\left(\sum_{j=1}^{k_n}\eta_j\right)^2 = O(k_n s_n),$$

$$E\zeta^2 = O(n - k_n(l_n + s_n)),$$

notice that the conditions for Proposition 2.7 (i) are satisfied, because by (10), $E\left|Z'_j\right|^\delta < \infty$ for $\delta > 2$; and the mixing assumption (12) implies that $\alpha(j)^{1-2/a} = j^{-b}$ for $b > 1$, which is $\alpha(j) = j^{-ab/(a-2)} = j^{-\frac{1}{2}\times\frac{1}{1/(2b)-1/(ab)}}$; take $\delta = ab$ and $q = 2b$, so the mixing condition is also satisfied.

For (20), using the covariance inequality in Proposition 2.6 in Fan and Yao [17], we have:

$$\left| E \exp \left(\mathrm{it} \sum_{j=1}^{k_n} \xi_j / \sqrt{n} \right) - \prod_{j=1}^{k_n} E \exp \left(\mathrm{it} \xi_j / \sqrt{n} \right) \right|$$
$$\leqslant \quad 16(k_n - 1)\alpha(s_n);$$

this is $o(1)$ by choosing for example $l_n = (nh^{\gamma_1})^{1/2}$, $s_n = (nh^{\gamma_2})^{1/2}$ for $1 < \gamma_1 < \gamma_2$. Then, $k_n = O(n^{1/2} h^{-\gamma_1/2})$, such that for some $q > 1$,

$$k_n \alpha(s_n) \quad = \quad n^{1/2} h^{-\gamma_1/2} \frac{1}{(nh^{\gamma_2})^{q/2}}$$
$$= \quad n^{\frac{(1-q)}{2}} h^{-\frac{(\gamma_2 q + \gamma_1)}{2}};$$

obviously, the above expression is $o(1)$ by the assumption that $\exp(1/h)/n \to 0$, so (20) is proven.

For Feller's condition (21), first use the same strategy as calculating the variance of the estimator; it holds that:

$$E\xi_j^2 = l_n \left(1 + o(1) \right),$$

for any j, because ξ_j is also an infinite sum of the observations. Therefore,

$$\frac{1}{n} \sum_{j=1}^{k_n} E\xi_j^2 = \frac{1}{n} k_n l_n \left(1 + o(1) \right) \to 1.$$

Finally, for Lindeberg's condition (22), first observe that:

$$E \left[\xi_j^2 I \left(|\xi_j| \geqslant \varepsilon n^{1/2} \right) \right] \quad \leqslant \quad \left(E\xi_j^4 \right)^{1/2} P \left(|\xi_j| \geqslant \varepsilon n^{1/2} \right)$$
$$\leqslant \quad \left(E\xi_j^4 \right)^{1/2} \frac{E\xi_j^2}{\left(\varepsilon \sqrt{n} \right)^2},$$

where I first use Holder's inequality and then Markov's inequality. Using again the moment inequality for the strong mixing sequence in Proposition 2.7 in Fan and Yao [17],

$$\left(E\xi_j^4 \right)^{1/2} \frac{E\xi_j^2}{\left(\varepsilon \sqrt{n} \right)^2} \quad \leqslant \quad \left(l_n^2 \right)^{1/2} \times \frac{l_n}{\left(\varepsilon \sqrt{n} \right)^2}$$
$$= \quad \frac{l_n^2}{\varepsilon^2 n},$$

so:

$$\frac{1}{n} \sum_{j=1}^{k_n} E \left[\xi_j^2 I \left(|\xi_j| \geqslant \varepsilon n^{1/2} \right) \right] \quad = \quad O \left(\frac{k_n}{n} \times \frac{l_n^2}{\varepsilon^2 n} \right)$$
$$= \quad \frac{1}{\varepsilon^2} O \left(\frac{l_n}{n} \right) = o(1).$$

Using the Lindeberg-Feller condition and employing the standard argument for the proof of central limit theorems, it can be shown that:

$$\prod_{j=1}^{k_n} E \exp\left(\mathrm{i}t\xi_j/\sqrt{n}\right) \to \exp\left(-\frac{t^2}{2}\right);$$

this together with (18)–(20) entail the stated result. □

4. Application to Density Estimation in the Stochastic Volatility Model

In this section, I consider applying the results of Theorem 2 to obtain the asymptotic distribution of the kernel deconvolution volatility density estimator in SV models. A generic SV model has the following form,

$$y_{t_i} = \sigma_{t_i}\varepsilon_{t_i}, \quad i = 1, \cdots, n, \tag{23}$$

where ε_{t_i}, $i = 1, \cdots, n$ are *i.i.d.* $N(0,1)$; $\{\sigma_{t_i}\}$ is a latent stochastic process called the volatility process; $\{y_{t_i}\}$ is the observed financial returns. The SV model is a popular model used in financial econometrics to describe the evolution of financial returns. Model (23) incorporates popular discrete-time SV models (e.g., Taylor [18]) and the discretized continuous-time SV models, which assume the volatility process to be stationary as special cases (see, e.g., Shephard [19] for a review). Van Es, Spreij, and Van Zanten [9] and Van Es, Spreij, and Van Zanten [10] considered estimating the volatility density using the kernel deconvolution estimator in this model.

Remark 1. *By using the term "stochastic volatility", here, I consider the so-called "genuine stochastic volatility" models, where the volatility process has a separate stochastic driving factor (see, e.g., Shephard and Andersen [20] and Andersen, Bollerslev, Diebold, and Labys [21] for detailed discussions). It thus does not include the ARCH/GARCH class models, where one has explicitly specified one-step-ahead predictive densities. Van Es, Spreij, and Van Zanten [10] considered estimating volatility density in the context of the ARCH/GARCH class of models and had given the rate of convergence for their estimator.*

To apply the general theory derived in Section 3, it is further assumed that the volatility process $\{\sigma_{t_i}\}$ is strictly stationary, and it is independent of ε_{t_i} for $i = 1, \cdots, n$. The independence assumption rules out the leverage effect in stochastic volatility models and, thus, is suitable to apply to, say, exchange rate data, where the leverage effect is rarely observed. Extending the model to allow for the leverage effect is an important, yet challenging task, which is thus left for future research.

The SV model can be written as a measurement error model (11) by taking squares and logarithms on both sides of equation (23),

$$\log y_{t_i}^2 = \log \sigma_{t_i}^2 + \log \varepsilon_{t_i}^2, i = 1, \ldots, n, \tag{24}$$

such that the variable $\log y_i^2$ is the convolution of $\log \eta_i$ with a completely known distribution logarithmic chi-square. Following the notations in the previous sections, write the density functions of $\log y_{t_i}^2$, $\log \sigma_{t_i}^2$ and $\log \varepsilon_{t_i}^2$ to be f_y, f_σ and k, respectively.

If we want to recover the density f_σ of $\log \sigma_{t_i}^2$ from the observations $\{\log y_{t_i}^2\}$, this is a problem of deconvolution with logarithmic chi-square error, and the kernel deconvolution estimator can be used. Van Es, Spreij, and Van Zanten [9] and Van Es, Spreij, and Van Zanten [10] first noticed this connection. Define $Z_j := \log y_j^2$; they use the following estimator to recover $f_\sigma(x)$,

$$\hat{f}_y(x) = \frac{1}{2\pi} \frac{1}{n} \sum_{j=1}^{n} \int_{-\infty}^{+\infty} \frac{\phi_K(th)}{\phi_k(t)} e^{-\mathrm{i}t(x-Z_j)} \mathrm{d}t,$$

where ϕ_K is the Fourier transform of a kernel function K and $\phi_k(t)$ is the characteristic function of the $\log \chi_1^2$ variable. Van Es, Spreij, and Van Zanten [9] and Van Es, Spreij, and Van Zanten [10] derive the convergence rate of the estimator, but a central limit theorem is missing.

If we assume the observed return sequence $\{Z_j\}$, $j = 1, \cdots, n$ is generated by the SV model (23) with a strict stationary, the α-mixing volatility process satisfies (12) and *i.i.d.* errors; a simple application of Theorem 2 will lead to the following corollary.

Corollary 1. *In the stochastic volatility model (23), when the volatility process $\{\sigma_j\}$, $j = 1, \cdots, n$ is α-mixing with (12) satisfied, ε_{t_i}'s are* i.i.d. *$N(0,1)$, independent of the volatility process; when $\exp(1/h)/n \to 0$ as $n \to \infty$ and $h \to 0$, it holds that:*

$$\frac{\hat{f}_\sigma(x) - K_h * f_\sigma(x)}{\sqrt{\frac{1}{2\pi^2 n} \exp(\pi/h) f_y(x)}} \to^d N(0,1).$$

Since the density $f_u(x)$ can be estimated with the observed return sequence $\{\log y_{t_i}^2\}$ consistently using the classical kernel density estimator for any x (see e.g., Fan and Yao [17]), the above result can be used to construct pointwise confidence intervals for the kernel deconvolution density estimator.

5. Conclusions

In this paper, I have proven the asymptotic normality for the kernel deconvolution estimator with logarithmic chi-square noise. I consider both the case of identical and independently distributed observations and strong mixing observations. The results are applied to prove the asymptotic normality of the kernel deconvolution estimator for volatility density in stochastic volatility models.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Carroll, R.; Hall, P. Optimal rates of convergence for deconvolving a density. *J. Am. Stat. Assoc.* **1988**, *83*, 1184–1186.
2. Stefanski, L.; Carroll, R. Deconvoluting kernel density estimators. *Statistics* **1990**, *21*, 169–184.
3. Fan, J. Asymptotic normality for deconvolution kernel density estimators. *Sankhya Indian J. Stat. A* **1991**, *53*, 97–110.
4. Fan, Y.; Liu, Y. A note on asymptotic normality for deconvolution kernel density estimators. *Sankhya Indian J. Stat. A* **1997**, *59*, 138–141.
5. Van Es, B.; Uh, H. Asymptotic normality of nonparametric kernel type deconvolution density estimators: Crossing the Cauchy boundary. *J. Nonparametr. Stat.* **2004**, *16*, 261–277.
6. Van Es, B.; Uh, H. Asymptotic normality of kernel-type deconvolution estimators. *Scand. J. Stat.* **2005**, *32*, 467–483.
7. Masry, E. Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes. *J. Multivar. Anal.* **1993**, *44*, 47–68.
8. Kulik, R. Nonparametric deconvolution problem for dependent sequences. *Electron. J. Stat.* **2008**, *2*, 722–740.
9. Van Es, B.; Spreij, P.; van Zanten, H. Nonparametric volatility density estimation. *Bernoulli* **2003**, *9*, 451–465.
10. Van Es, B.; Spreij, P.; van Zanten, H. Nonparametric volatility density estimation for discrete time models. *Nonparametr. Stat.* **2005**, *17*, 237–251.
11. Comte, F.; Genon-Catalot, V. Penalized projection estimator for volatility density. *Scand. J. Stat.* **2006**, *33*, 875–893.
12. Todorov, V.; Tauchen, G. Inverse realized laplace transforms for nonparametric volatility density estimation in jump-diffusions. *J. Am. Stat. Assoc.* **2012**, *107*, 622–635.
13. Fan, J. On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Stat.* **1991**, *19*, 1257–1272.
14. Delaigle, A.; Gijbels, I. Practical bandwidth selection in deconvolution kernel density estimation. *Comput. Stat. Data Anal.* **2004**, *45*, 249–267.
15. Butucea, C. Asymptotic normality of the integrated square error of a density estimator in the convolution model. *Sort* **2004**, *28*, 9–26.
16. Masry, E. Multivariate probability density deconvolution for stationary random processes. *IEEE Trans. Inf. Theory* **1991**, *37*, 1105–1115.
17. Fan, J.; Yao, Q. *Nonlinear Time Series*; Springer: Berlin, Germany, 2002; Volume 2.
18. Taylor, S.J. Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961–1975. *Time Ser. Anal. Theory Pract.* **1982**, *1*, 203–226.
19. Shephard, N. (Ed) *Stochastic Volatility: Selected Readings*; Oxford University Press: Oxford, UK, 2005.
20. Shephard, N.; Andersen, T.G. Stochastic volatility: Origins and overview. In *Handbook of Financial Time Series*; Andersen, T.G., Davis, R.A., Kreiß, J.P., Mikosch, T., Eds.; Springer: Berlin, Germany, 2009; pp. 233–254.

21. Andersen, T.G.; Bollerslev, T.; Diebold, F.X.; Labys, P. Parametric and nonparametric volatility measurement. *Handb. Financ. Econom.* **2009**, *1*, 67–138.