Deconvolution estimators for invariant densities of stationary processes: Method and Simulation

Thai Phuc Hung[1,*], Nguyen Cao Phong[2]

[1] Faculty of Basics, Soc Trang Community College;
[2] Testing and Quality assurance Department, Mien Tay Construction University;

*Email: tphung@stcc.edu.vn

■Nhận bài: 17/04/2024 ■Sửa bài: 07/06/2024 ■Duyệt đăng: 18/06/2024

ABSTRACT

In this paper, the authors estimate the invariant density function of a stationary process from noisy observed data, using the kernel density estimation technique. We examine observations that have been contaminated by another stationary process with a known invariant density function. The study focuses on stationary processes with strong mixing properties. Both the ordinary smooth and the supersmooth noise density function classes are investigated. The research establishes upper bounds for the mean squared error of the estimator to evaluate the rate of convergence. The theoretical properties of the estimator's convergence are illustrated through simulation studies, in which the authors estimate invariant density functions for two stationary processes from noisy observed data generated using the R language. Additionally, a computational example with data on Duchenne muscular dystrophy is also presented to demonstrate the estimator's effectiveness.

Keywords: *Deconvolution; Kernel density estimation; Mean squared error; Stationary process; Strongly mixing ($\alpha - mixing$).*

1. INTRODUCTION

Let us consider a strongly mixing stationary process X_t with an unknown invariant density function f_X. Suppose we have a sample $Y_{t1},..., Y_{tn}$ drawn from the distribution of $Y_t = X_t + \varepsilon_t$, where $t_1 < t_2 < ... < t_n$. Here, the process ε_t is assumed to be independent of the process X_t. The process Y_t serve as noisy versions of the processes X_t, while ε_t represents a random measurement error process with a known invariant density function g called the error density. Consequently, the invariant density function of the process Y_t is determined by

$$f_Y(y) = \int_{-\infty}^{+\infty} f_X(y-x)\, g(x)\, dx \cdot$$

We assume that the $\alpha - mixing$ stationary processes are observed at discrete time $t_j = j\Delta$ (where Δ is a positive constant). For brevity, we will denote $X_{t_j}, Y_{t_j}, \varepsilon_{t_j}$ by X_j, Y_j, ε_j, respectively. Thus, the model can be expressed as:

$$Y_j = X_j + \varepsilon_j,\ \ j = 1,...,n. \qquad (1)$$

The aim of this paper is to nonparametrically estimate the invariant density function f_X of the stationary process X_t from a sample of noisy observations $Y_1,...,Y_n$.

In the scenario of independent and identically distributed (iid) random variables $X_1,...,X_n \overset{iid}{\sim} X, \varepsilon_1,...,\varepsilon_n \overset{iid}{\sim} \varepsilon$, model (1), investigated by Fan, has provided an optimal rate in mean squared error (MSE) convergence and minimax theory (see Fan [1, 2]). But in reality, many factors are not independent, and assuming they are generated from stationary processes would be more reasonable. Stationary processes are important stochastic processes with many practical applications (such as representing random dynamical systems or the returns of assets, exchange rates, and other economic factors). Therefore, the problem of estimating the invariant density function f_X for a stationary process, either nonparametric or parametric, through collected data is a topic of interest to many researchers in both cases of error-free data (where $\varepsilon_t = 0$, see N'drin and Hili [3, 4]) and noisy data (where $\varepsilon_t \neq 0$, see Masry [5, 6], Trong and Hung [7]). However, in studies estimating f_X, researchers have not yet evaluated the convergence rate of MSE, nor have they conducted simulations and real data calculations.

In this paper, the authors investigate the estimation of f_X for the model (1) under the assumption that both X_t and ε_t are $\alpha - mixing$ stationary processes, suitable for various real-life scenarios. The main research outcomes are the convergence rate of MSE of the estimations presented in *Theorems 3.1* and *Theorem 3.2* under two cases of the error density class. The error density classes are ordinary smooth under assumption $c(1+|p|)^{-\kappa} \le |g^{ft}(p)| \le C(1+|p|)^{-\kappa}$ and supersmooth under assumption

$$c\exp(-\beta |p|^{\gamma}) \le |g^{ft}(p)| \le C\exp(-\beta |p|^{\gamma}),$$

where g^{ft} is the Fourier transform of the error density g and $0 < c \le C, \kappa, \beta, \gamma > 0$ are constants (see Meister [8]). Additionally, the paper conducts simulation and numerical computations with real data to illustrate the achieved results of the theory.

The paper is organized as follows: In Section 2, the proposed estimator is introduced. Section 3 presents the main results on MSE convergence along with outline proofs. Section 4 is dedicated to simulation studies, while Section 5 focuses on empirical applications, specifically analyzing Duchenne muscular dystrophy data. Finally, the paper concludes with a conclusion in Section 6.

2. KERNEL ESTIMATION METHOD

For $\varrho \ge 1$, let denote the set of Lebesgue measurable functions f satisfying

$\| f \|_{\varrho} = \left(\int_{-\infty}^{+\infty} |f(x)|^{\varrho} dx \right)^{1/\varrho} < \infty$. For $f \in L_1$, let $f^{ft}(p) := \int_{-\infty}^{+\infty} e^{ipx} f(x) dx$ denote the Fourier transform of f. In this paper, we address the issue of dependent processes. Various dependence conditions have been extensively studied (see Bradley [9]). Here, we focus on the $\alpha - mixing$ dependence introduced by Rosenblatt. Let $\mathcal{F}_i^k$ denote the $\sigma - algebra$ of events generated by the random variables $\{X_j, 1 \le i \le j \le k\}$. The process $\{X_j\}$ is termed $\alpha - mixing$ (see [9] and Rosenblatt [10]) if

$$\sup_{j \in \mathbb{Z}} \sup_{A \in \mathcal{F}_1^j, B \in \mathcal{F}_{j+k}^{+\infty}} |P[AB] - P[A].P[B]| = \alpha_X(k)$$

$\to 0$ as $k \to +\infty$, where $\alpha_X(k)$ represents the strong mixing coefficient of the process $\{X_j\}$. This coefficient quantifies the degree of dependence. When (X_j) is mutually independent, then $\sup_k \alpha_X(k) = 0$. The condition $\alpha_X(k) \to 0$ implies that X_j and X_{j+k} become "asymptotically independent" as $k \to \infty$. Utilizing these dependence properties, we establish a consistent estimator for deconvolution.

The kernel density estimator used to estimate the unknown invariant density f_X is defined as follows:

$$f_n(x) = \frac{1}{n} \sum_{j=1}^{n} \psi(x) \quad \forall x \in \mathbb{R}, \qquad (2)$$

where $\psi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp[ip(Y_j - x)] \frac{K^{ft}(pb_n)}{g^{ft}(p)} dp,$

b_n is the bandwidth satisfying $b_n > 0$ $\forall n \in \mathbb{N}$, K denotes a known kernel function and K^{ft} represents the Fourier transform of K.

3. MAIN RESULTS

3.1. Some assumptions

(i) *The kernel function* $K \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$, $\int_{-\infty}^{+\infty} K(z) dz = 1$, $\int_{-\infty}^{+\infty} zK(z) dz = 0$, $\int_{-\infty}^{+\infty} z^2 K(z) < \infty$ and the Fourier transform K^{ft} is symmetric function, supported on $[-1;1]$.

(ii) Let $\kappa > 0$. For some $0 < c \le C$, the noise density g belongs to the ordinary smooth class with the condition:

$$c\left(1+|p|\right)^{-\kappa} \le \left|g^{ft}(p)\right| \le C\left(1+|p|\right)^{-\kappa}.$$

(iii) Let $\beta, \gamma > 0$. For some $0 < c \le C$, the noise density g belongs to the suppersmooth class with the condition:

$$c\exp\left(-\beta\,|\,p\,|^{\gamma}\right) \le \left|g^{ft}(p)\right| \le C\exp\left(-\beta\,|\,p\,|^{\gamma}\right).$$

(iv) Let $\delta > 0$. X_t, ε_t have $\alpha - mixing$ coefficients $\alpha_X(k), \alpha_\varepsilon(k) \le O\left(1/k^{2+\delta}\right)$.

(v) The 2-dimensional probability density function $f_{Y_j,Y_k}(u,v)$ exists and is bounded for all $1 \le j,k \le n$.

3.2. Mean square error convergence

3.2.1. Theorem 3.1

Let the assumptions (i), (iii), (iv), (v) hold. If $f\ (x)$ is twice differentiable and $f_X''(x)$ is continuous and bounded, by choosing the bandwidth $b_n = O\left(n^{-1/(5+2\kappa)}\right)$,

$$E\left(f_n(x) - f_X(x)\right)^2 \le O\left(n^{-4/(5+2\kappa)}\right)$$

at continuity points x of $f_Y(x)$.

3.2.2. Theorem 3.2

Let the assumptions (i), (iii), (iv), (v) hold and let $\lambda_{b_n} \in \left(0, 1/(2\beta)\right)$. If $f_X(x)$ is twice differentiable and $f_X''(x)$ is continuous and bounded, by choosing the bandwidth $b_n = \lambda_{b_n}^{-1/\gamma}(\ln n)^{-1/\gamma}$, $E\left(f_n(x) - f_X(x)\right)^2 \le O\left((\ln n)^{-4/\gamma}\right)$ at continuity points x of $f_Y(x)$.

Remark. It is worthwhile to compare the convergence rate of MSE as shown in *Theorem 3.1* with that of iid observations, proven to be optimal in Theorem 2 of [3], and the convergence rate of MSE as shown in *Theorem 3.2* with that of iid observations, proven to be optimal in Theorem 1 of [3]. According to these theorems, if we set $m = 2, \alpha = 0, l = 0$, the convergence rate of MSE remains the same for both scenarios: random variables generated from $\alpha - mixing$ stationary processes and iid random variables. In the iid case, Fan demonstrated that these convergence rates are optimal. This comparison highlights that the convergence rates achieved in *Theorems 3.1* and *3.2* are indeed quite good.

The proofs of the two Theorems are quite lengthy and complex; here, the authors only provide a brief outline.

Proof of Theorem 3.1. We have

$$\text{bias}\left(f_n(x)\right) = E\left(f_n(x)\right) - f_X(x)$$

$$= \int_{-\infty}^{+\infty} K\left(\frac{u}{b_n}\right) f_X(x-u)\,du - f_X(x)$$

$$\overset{v=u/b_n}{=} \int_{-\infty}^{+\infty} K(v)\left[f_X(x - vb_n) - f_X(x)\right]dv$$

$$= \int_{-\infty}^{+\infty} K(v)\Big[f_X'(x)(-vb_n) + \frac{1}{2}f_X''(x)(-vb_n) + o(b_n)^2\Big]dv$$

$$= \frac{b_n^2}{2} f_X''(x) \int_{-\infty}^{+\infty} v^2 K(v)\,dv + o(b_n)^2.$$

Therefore, $\text{bias}\left(f_n(x)\right) \le O\left(b_n^2\right)$.

By Lemma 5.4 of [13], we obtain

$$\text{Var}\left(f_n(x)\right) \le O\left(1/\left(nb_n^{1+2\kappa}\right)\right).$$

By choosing the bandwidth $O\left(n^{-1/(5+2\kappa)}\right)$, we have

$$E\left(f_n(x) - f_X(x)\right)^2$$
$$= \text{Var}\left(f_n(x)\right) + \text{bias}^2\left(f_n(x)\right)$$
$$\le O\left(n^{-4/(5+2\kappa)}\right).$$

Proof of Theorem 3.2. Similarly, we have

$$\text{bias}\left(f_n(x)\right) \le O\left(b_n^2\right).$$

Let $\omega_{b_n} = \exp\left[(-\beta)b_n^{-\gamma}\right]$. A modification to the proof of Lemma 5.4 of **[7]** yields the result

$$\text{Var}\left(f_n(x)\right) \le O\left(1/\left(nb_n\omega_n^2\right)\right).$$

And by choosing the band width $b_n = \lambda_{b_n}^{-1/\gamma}(\ln n)^{-1/\gamma}$, we have

$$E\left(f_n(x) - f_X(x)\right)^2$$
$$= \text{Var}\left(f_n(x)\right) + \text{bias}^2\left(f_n(x)\right)$$
$$\le O\left((\ln n)^{-4/\gamma}\right).$$

4. SIMULATION STUDY

4.1. Cox-Ingersoll-Ross process

The Cox-Ingersoll-Ross (CIR) process is a flexible modeling tool, often used to simulate the fluctuations of variables with a mean-reverting characteristic. In finance, the CIR model is commonly applied to model interest rates and prices, but it also finds applications in other fields such as medicine and data science. For instance, in medicine, it can be used to simulate the dynamics of biological indicators like hormone concentrations, aiding in the assessment of medical tests' performance based on the Receiver Operating Characteristic (ROC) model. In data applications, the CIR model can also be used to simulate the fluctuations of measurement indices in data classification processes, enhancing the performance and reliability of forecasting and classification models.

The CIR process is the solution to the stochastic differential equation

$$dX_t = (\theta_1 - \theta_2 X_t)\,dt + \theta_3 \sqrt{X_t}\,dW_t,$$

where $\theta_1, \theta_2, \theta_3 \in \mathbb{R}_-$.

Under the assumption that $2\theta_1 > \theta_3^2$, there exist a positive τ such that $\alpha_X(k) \le e^{-\tau k}/4$ (see Genon-Catalot, Jeantheau and Lardo [11], Corollary 2.1). Consequently, the Cox-Ingersoll-Ross process is exponentially $\alpha - mixing$ and its invariant distribution follows a gamma density function $f_X(x) = \frac{b^a x^{a-1}}{\Gamma(a)} \exp(-bx)$, where the shape parameter $a = 2\theta_1/\theta_3^2$ and the rate parameter $b = 2\theta_2/\theta_3^2$. Assume that X_t is exponentially $\alpha - mixing$ stationary CIR process with $\theta_1 = 2,\ \theta_2 = 1/2,\ \theta_3 = 1$ (Gamma density with $a = 4, b = 1$) and the noise random variables ε_j are Laplace random variables, which have the density function $g(x) = (1/2)e^{-|x|}$ and the characteristic function $g^{ft}(p) = 1/(1+p^2)$.

We now conduct simulations for the CIR process using the R language. First, we use the 'sde' package to generate data for the process X_t (see Iacus [12]). We also generate data for ε_j and then construct the dataset $\{Y_j = X_j + \varepsilon_j\}_{j=\overline{1,n}}$. Finally, we employ the proposed estimator given in (2) to estimate the invariant density f_X of the process X_t. The authors choose the bandwidth $b_n = n^{-1/9}$ according to *Theorem 3.1* and the kernel function

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

with $K^{ft}(p) = (1-p^2)^3 I_{[-1;1]}(p)$, where

$$I_{[-1;1]}(p) = \begin{cases} 1 & \text{when } p \in [-1,1] \\ 0 & \text{when } p \notin [-1,1] \end{cases}.$$

These simulations encompass three distinct sample sizes: $n = 100$, $n = 200$, and $n = 500$, facilitating an analysis of the sensitivity of the estimation to sample size. Each simulation comprises 100 replications of observations for the CIR process Y_t. The results of the estimator of f_X are presented in Figure 1, and the empirical MSE at various points are shown in Table 1.

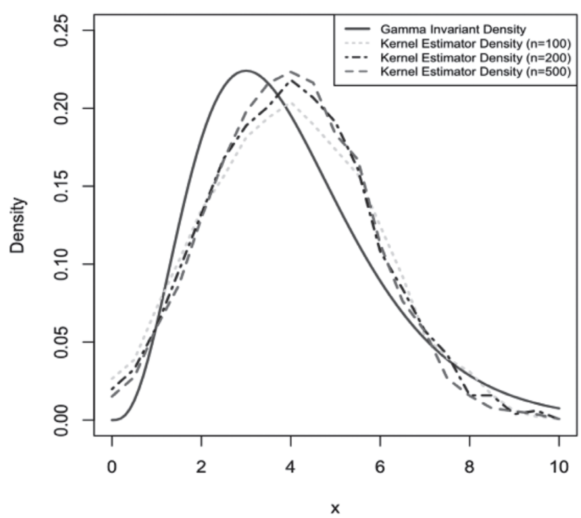

Figure 1. The Gamma invariant density function of the CIR process X_t with $\varepsilon_j \sim Laplace(0,1)$ and its Estimator for different sample sizes.

Table 1. The empirical MSE of the estimator at various points x for the CIR process X_t with $\varepsilon_j \sim Laplace(0,1)$ and different sample sizes.

x	$E\left(f_n(x) - f_X(x)\right)^2$		
	$n = 100$	$n = 200$	$n = 500$
1	0.000108567	0.000033427	0.000073352
2	0.003448514	0.002718298	0.002204957

x			
	$n = 100$	$n = 200$	$n = 500$
3	0.001796847	0.001645794	0.000219733
4	0.000229594	0.000260438	0.000334539
5	0.002749886	0.002513594	0.002970994
6	0.000711336	0.000371551	0.000334867
7	0.000321735	0.000065087	0.000003606
8	0.000047871	0.000099706	0.000060205

4.2. m - Dependent stationary process

An important form of dependence, where distance serves as a measure of dependence, is the $m-dependence$ scenario. A stationary process $\{X_j\}_{j\in\mathbb{Z}}$ *is considered \$m\$-dependent if two sets of random variables* $\{..., X_{k-1}, X_k\}$ and $\{X_h, X_{h+1}, ...\}$ are independent whenever $h-k>m$. It's evident that $\alpha_X(l)=0$ for $l>m$, and the $m-dependence$ stationary process is a specific instance of a strongly mixing stationary process.

The $m-dependent$ stationary process is a powerful tool used to simulate real-world situations, particularly those with strong dependencies among variables over time. One of its primary applications is in the field of finance, where it can model the fluctuations of stock prices, interest rates, and other financial indicators. Additionally, the $m-dependent$ stationary process can be applied in weather forecasting, simulating the variability of weather factors such as temperature, air pressure, and precipitation. Moreover, in data science, the $m-dependent$ stationary process can model dependencies among variables in time series data.

In this simulation, we assume that $\{X_j\}_{j\in\mathbb{Z}}$ is $30-dependent$ stationary process, and its invariant distribution is normal distributions with $\mu=5, \sigma^2=2$. The noise variables ε_j are standar nomal random variables, which have the characteristic function $g^{ft}(p)=\exp(-p^2/2)$.

Now we'll simulate using R. Initially, we generate iid normal random variables $\{U_j\}_{j\in\mathbb{Z}}$. Subsequently, a $30-dependent$ stationary process $\{X_j\}_{j\in\mathbb{Z}}$ is created with $X_j=(1/30)(U_j+U_{j+1}+...+U_{j+29}), j=1,...,n$.

Using the bandwidth $b_n=(\ln n)^{-1/2}$ according to *Theorem 3.2* and the same kernel function detailed in Section 4.1, simulations are executed across three sample sizes: $n=100$, $n=200$, and $n=500$ to examine the impact of sample size on estimation sensitivity. Each simulation is repeated 100 times, and the findings are summarized in in Figure 2 and Table 2.

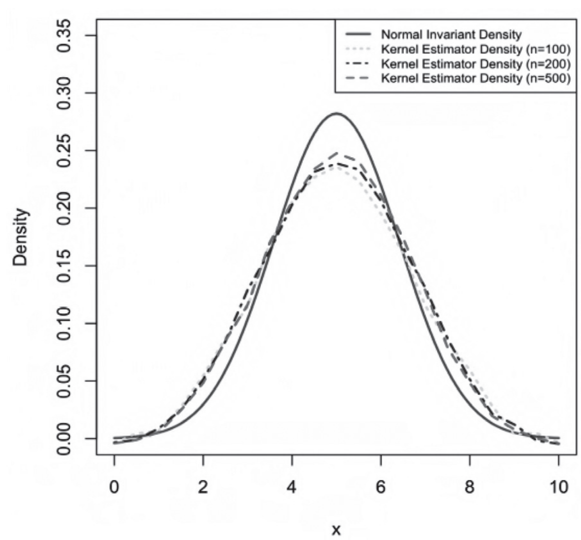

Figure 2. The Normal invariant density function of the $30-dependent$ stationary process X_t with $\varepsilon_j \sim N(0,1)$ and its Estimator for different sample sizes.

Table 2. The empirical MSE of the estimator at various points x for the $30-dependent$ stationary process X_t with $\varepsilon_j \sim N(0,1)$ and different sample sizes.

x	$E(f_n(x)-f_X(x))^2$		
	$n=100$	$n=200$	$n=500$
1	0.000177833	0.000132823	0.000102027
2	0.000950244	0.000703886	0.000513955
3	0.000420404	0.000266311	0.000289713
4	0.001668186	0.000798099	0.000603855
5	0.004383285	0.00349504	0.003039045
6	0.001002422	0.000544171	0.000458958
7	0.000356865	0.000257928	0.000240189
8	0.001177767	0.000582998	0.000534965
9	0.000212997	0.000132662	0.000117591

5. DMD DATA

This dataset pertains to a study on Duchenne muscular dystrophy (DMD), a progressive genetic disorder passed from mother to child (see Reiser [13]). DMD is one of the most severe and common forms of muscular dystrophy in humans, typically manifesting in children before the age of 3 and resulting in fatality by the early 20s. With no effective treatment currently available, screening females as potential carriers of DMD is of paramount importance. The dataset includes blood samples from two groups: normals and DMD carriers. Of particular interest is the serum creatine kinase enzyme. Data is available for 38 carriers and 87 normals, with some subjects providing multiple blood samples treated as replicates. The number of replicates per subject varies from 1 to 7, creating an unbalanced dataset.

Let X and X denote independent random variables representing the distribution of the marker in the populations of diseased and healthy individuals, respectively, with available random samples $X_1,...,X_n$ and $X_1,...,X_m$. According to Reiser, in many situations the variables X_j and X_k are not directly observable but are measured with additive normally distributed random measurement error. Let Y_{jl_j} and Y_{ks_k} denote the l_j th replicate taken on the j th subject in the diseased population and the s_k th replicate taken on the k th subject in the healthy population. Thus

$$Y_{jl_j} = X_j + \varepsilon_{jl_j}, \quad j=1,...,n, l_j=1,...,p_j$$

$$Y_{ks_k} = X_k + \varepsilon_{ks_k}, \quad k=1,...,m, s_k=1,...,q_k,$$

with $\tilde{\varepsilon}_{jl_j} \sim N(0,\sigma_{\tilde{\varepsilon}}^2)$, $\varepsilon_{ks_k} \sim N(0,\sigma_\varepsilon^2)$. In this context, the random variable sequences $Y_{11},...,Y_{1p_1},...,Y_{n1},...,Y_{np_n}$ and $Y_{11},...,Y_{1q_1},...,Y_{m1},...,Y_{mq_m}$ demonstrate $p-dependent$ and $q-dependent$ processes, correspondingly, with $p=\max\{p_1,...,p_n\}$ and $q=\max\{q_1,...,q_m\}$.

Upon examination of the data described in the DMD dataset, it was observed that the marker values exhibited significant skewness. Therefore, a log transformation was applied to improve normality. Under the assumption that $X \sim N(\mu_X,\sigma_X^2)$ and $X \sim N(\mu_X,\sigma_X^2)$, Reiser utilized the maximum likelihood estimation method to estimate the probability density functions f_X of X and f_X of X. Now, with

$$\sigma_\varepsilon^2 = [n(p-1)]^{-1}\sum_j\sum_l\left(Y_{jl_j} - \overline{Y}_j\right)^2$$

and

$$\sigma_\varepsilon^2 = [m(q-1)]^{-1}\sum_k\sum_s\left(Y_{ks_k} - \overline{Y}_k\right)^2,$$

the authors will use proposed estimator given in (2) to estimate f_X and f_X without any distributional assumptions about X and X. We use the kernel function K described in Section 4.1 and choose the bandwidth $b_n = (\ln n)^{-1/2}$ in accordance with *Theorem 3.2*. The results are illustrated in Figure 3 and Figure 4.

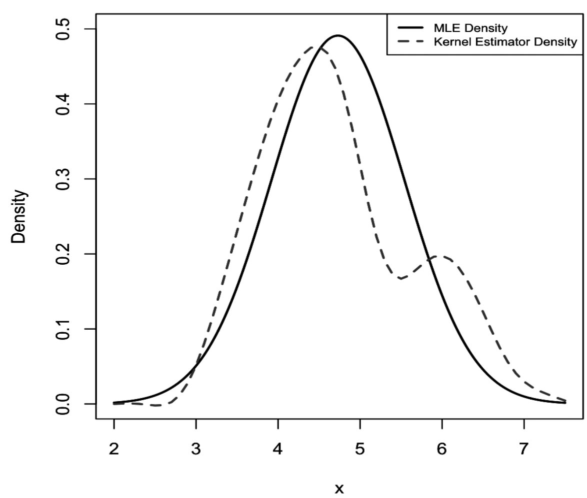

***Figure** 3. The maximum likelihood estimator (MLE) density and kernel estimator density for the marker.*

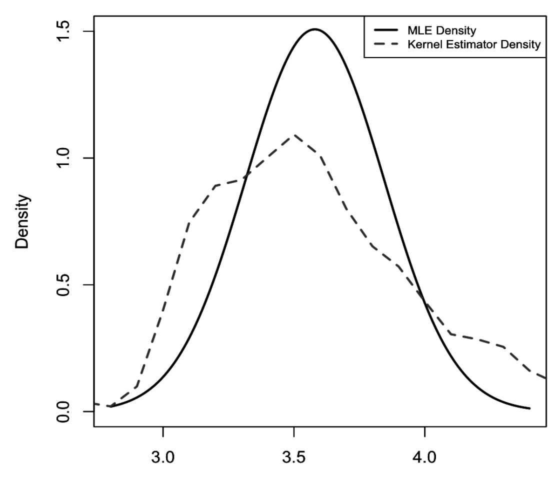

***Figure** 4. The maximum likelihood estimator (MLE) density and kernel estimator density for the healthy.*

6. CONCLUSION

In this study, the authors estimated the invariant density function of a stationary process from noisy observed data using the kernel density estimation method. The results show effective performance, with the convergence rate for the MSE as demonstrated in Theorems and simulations. The convergence rate of MSE is quite good when compared to the iid case. The method was validated with both simulated data, specifically the CIR and stationary processes, and real data on Duchenne muscular dystrophy (DMD).

Key advantages of this study include:

● Reliable Estimations: The kernel density estimation method provided dependable estimates for invariant density functions.

● Practical Applications: Successful application to both simulated and real-world data.

However, the choice of bandwidth was not optimized, which may affect the convergence rate. Future research should focus on this aspect to enhance the method's performance.

Overall, this study contributes valuable insights into the estimation of invariant density functions from noisy data and highlights areas for future improvement.

REFERENCES

[1] J. Fan, "Asymptotic normality for deconvolution kernal density estimators", *Sankhyā: The Indian Journal of Statistics, Series A*, Volume 53, Number 1, 97-110, 1991.

[2] J. Fan, "On the optimal rates of convergence for nonparametric deconvolution problems", *The Annals of Statistics*, Volume 19, Number 3, 1257–1272, 1991.

[3] J.A. N'drin and O. Hili, "Parameter estimation of one-dimensional diffusion process by minimum Hellinger distance method", *Random Operators and Stochastic Equations,* Volume 21, Issue 4, 403-424, 2013. DOI: org/10.1515/rose-2013-0019.

[4] J.A. N'drin and O. Hili, "From nonparametric density estimation to parametric estimation of multidimensional diffusion processes", *Applied mathemathics*, Volume 6, Number 9, 1592-1610, 2015. DOI: 10.4236/am.2015.69142.

[5] E. Masry, "Multivariate probability density deconvolution for stationary random processes", *IEEE Transactions on Information Theory*, Volume 37, Issue 4, 1105-1115, 1991. DOI: 10.1214/07-EJS154.

[6] E. Masry, "Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes", *Journal of multivariate analysis*, Volume 44, Issue 1, 47-68, 1993. DOI: org/10.1006/jmva.1993.1003.

[7] D.D. Trong and T.P. Hung, "Parameter estimation for diffusion process from perturbed discrete observations", *Communications in Statistic – Simulation and computation*, Volume 52, Issue 3, 925-944, 2023. DOI: org/10.1080/03610918.2020. 1871014.

[8] A. Meister, *Deconvolution problems in nonparametric statistics*. Berlin Heidelberg: Springer-Verlag, 2009.

[9] C.R. Bradley, "Basic properties of strong mixing conditions. A survey and some open questions", *Probability Surveys*, Volume 2, 107-144, 2005. DOI: 10.1214/154957805100000104.

[10] M. Rosenblatt, "A central limit theorem and strong mixing conditions", *National Academy of Sciences*, Volume 42, Number 1, 43-47, 1956. DOI: 10.1073/pnas.42.1.43.

[11] V. Genon-Catalot, T. Jeantheau and C. Lardo, "Stochastic volatility models as hidden Markov models and statistical applications", *Bernoulli,* Volume 6, Number 6, 1051-1079, 2000. DOI: org/10.2307/3318471.

[12] S.M. Iacus, *Simulation and inference for stochastic differential equations*. New York: Springer, 2008.

[13] B. Reiser, "Measuring the effectiveness of diagnostic markers in the presence of measurement error through the use of roc curves", *Statistics in Medicine,* Volume 19, Issue 16, 2115-2129, 2000. DOI: 10.1002/1097-0258(20000830)19:16<2115: aid-sim 529>3.0.co;2-m.