ORIGINAL PAPER

A NONPARAMETRIC MODE ESTIMATE UNDER DOUBLY TRUNCATED MODEL

KARIMA ZERFAOUI[1,2], ABIDA ZAHNIT[1], DJABRANE YAHIA[1]

Manuscript received: 14.10.2022; Accepted paper: 01.02.2023;
Published online: 30.03.2023.

Abstract. *The phenomenon of simultaneous left and right double truncation appears in a variety of fields, sush as in medical research. The problem of estimating the mode function for this type of data has not been addressed in the statistical literature. In this paper, we propose a new kernel estimator of the mode in the framework of randomly and doubly truncated model. We establish the strong consistency with a rate for the proposed estimate, and state its asymptotic normality. A simulation study is carried out to illustrate and evaluate the finite sample behavior of the proposed estimator.*

Keywords: *asymptotic normality; kernel mode estimator; left and right truncation; rate of convergence.*

1. INTRODUCTION

Randomly truncated data appear in a variety of fields, including astronomy, medicines, epidemiology and economics. A typical example of random right truncation is the analysis of AIDS data, when the information is restricted to those individuals developing AIDS before some specific date, in such a case, the induction time is said to be right-truncated. In some applications, two-sided (rather than one-sided) random truncation appears. In [1], it is indicated that induction times in AIDS are actually doubly truncated since HIV was unknown prior to 1982 and therefore infected patients would have been incorrectly discarded when developing AIDS before that date. Furthermore, the study [2] investigated quasar luminosities which were doubly truncated by some detection limits.

Nonparametric methods for one-sided (left or right) truncated data were introduced by many authors, see for example [3-5]. However, literature on random double truncation is much scarcer. A possible reason is the absence of closed form estimators, indeed, existing methods for doubly truncated data are iterative and computationally intensive, and these issues make difficult both the theoretical developments and the practical implementations.

The authors of [2] are the first who introduced the NPMLE of the distribution function under double truncation. The literature also contains semiparametric approaches to estimate the distribution function under double truncation. The problem when the distribution of the truncation times is assumed to belong to a given parametric family is investigated by [6-7].

The study [8] introduces two different estimators of kernel density, which are defined as a convolution between a kernel function and an estimator of the cumulative distribution function. Several bandwidth selection procedures for kernel density estimation of randomely double truncated data are introduced and compared; the autors [9] present five bandwidth selection procedures, and give a theoretical justification.

[1] Biskra University, Laboratory of Applied Mathematics, PB145 Biskra, Algeria. E-mail: abida_stat@yahoo.fr; yahia.djabrane@univ-biskra.dz.
[2] Higher School of Management and Digital Economy, Koléa, Algeria. E-mail: karimazarfaoui@gmail.com.

To our knowledge, the problem of estimating the mode function under doubly truncated data has not been addressed in the statistics literature. This is a central object of interest of this paper. In this work we propose a new mode function estimator and establish its almost sure uniform convergence and asymptotic normality. For that perpose, we consider the Vapnik-Cervonenkis (V-C) classes for which uniform exponential inequalities are available. Moreover, functional estimation is based on the kernel method. As an application of the asymptotic normality of our newly estimator, we introduce an asymptotic confidence interval for the mode. Our theoritical results coincide with those obtained in the complete data case.

This paper is organized as follows: in section 2 we define some important and useful results in the random doubly truncation model, then we define the kernel mode estimator under random double truncation. Assumption and main results are given in section 3 with asymptotic normality of the suggested estimator. Section 4 provides a simulation of our estimator. The proofs of the main results are proposed to section 5, finally main conclusion and a final discussion are given in section 6.

2. NOTATIONS AND DEFINITION OF THE ESTIMATOR

We first present some results from the literature for doubly truncated data, and which will be used to define our estimator of the mode. Let Y^* be the random variable of interest with distribution function F, and assume that it is doubly truncated by the random pair (U^*, V^*) with joint distribution function H, where U^* and $V^*(U^*\leq V^*)$are the left and right truncation variables respectively. This means that the triplet (U^*, Y^*, V^*) is observed if and only if $U^* \leq Y^* \leq V^*$. While no information is available when $Y^* < U^*$or $Y^* > V^*$. We assume that Y^* is independent of (U^*, V^*).

Let $(U_i, Y_i, V_i), i = 1, \dots., n,$ denote the sampling information, these are i.i.d. data with the same distribution of (U^*, Y^*, V^*) given $U^* \leq Y^* \leq V^*$. Introduce $\alpha = P(U^* \leq Y^* \leq V^*)$, the probability of no-truncation. It is clear that if α = 0, no data can be observed and therefore, we suppose throughout this paper that α > 0. For any distribution W denote respectively the left and right endpoints of its support by

$$a_w = inf\{t: W(t) > 0\} \text{ and } b_w = inf\{t: W(t) = 1\}.$$

Let $H_1(u) = H(u, \infty)$ and $H_2(v) = H(-\infty, v)$ the marginal distribution functions of U^* and V^* respectively. When $a_{H_1} \leq a_F \leq a_{H_2}$and $b_{H_1} \leq b_F \leq b_{H_2}$, F and H are both identifiable (see [5], for more details). Denote by $f(.)$ the probability density function of Y^* and assume that it has a unique mode defined by

$$\theta = \underset{y \in \mathbb{R}}{\operatorname{argmax}} f(y).$$

As proposed in [8], to define the nonparametric kernel estimator of the density $f(.)$, we first need to introduce the nonparametric maximum likelihood estimator (NPMLE) of the distribution function (df) of Y^*(see [2]). Under the doubly truncated sampling scheme, the NPMLE estimator of the df of Y^* is given by

$$F_n(y) = \alpha_n \int_{-\infty}^{y} \frac{F_n^*(dt)}{G_n(t)}$$

where

$$\alpha_n = \left(\int_{a_F}^{\infty} G_n^{-1}(t) F_n^*(dt) \right)^{-1}$$

is an estimator of α, (See [7]). $F_n^*(y) = n^{-1} \sum_{i=1}^{n} I_{[Y_i \le y]}$ is the ordinary empirical df of the $Y_i's$, and

$$G_n(t) = \int_{\{u \le t \le v\}} H_n(du, dv)$$

is a nonparametric estimator of $G(t) = P(U^* \le t \le V^*)$ which is the probability of sampling a life time $Y^* = t$. Here $H_n(u, v)$ is the NPLME of the joint distribution H of the truncation times, see [8], for more details. Now, our nonparametric estimator of the mode θ is defined as the random θ_n maximizing the estimator of the density $\hat{f}$, that is

$$\hat{f}_n(\hat{\theta}_n) = \sup_{a_F \le y \le b_F} \hat{f}_n(y) \tag{1}$$

where

$$\hat{f}_n(y) = \int K_h(y - t) F_n dt = \frac{\alpha_n}{n h_n} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{y - Y_i}{h_n}\right), \tag{2}$$

K is a probability density function (so-called kernel function) and h_n is a sequence of positive real numbers (so-called bandwidth) which goes to zero as n goes to infinity.

Remark 2.1. Recal that, the estimator $\hat{\theta}_n$ is not necessarily unique and our results are valid for any chosen value satisfying **(1)**. We point out that, our choice can be specified by taking

$$\hat{\theta}_n = inf\{a_F \le t \le b_F \ \ such\ that\ \hat{f}_n(t) = sup_{a_F \le y \le b_F} \hat{f}_n(y)\}.$$

3. ASSUMPTIONS AND MAIN RESULTS

Throughout this paper, we suppose that $a_{H_1} \le a_F \le a_{H_2}$ and $b_{H_1} \le b_F \le b_{H_2}$ and let Ω be a compact set such that $\Omega \subset \Omega_0 = \{y : a_F \le y \le b_F\}$. Consider now the following regularity assumptions:

(A1) The kernel function K is a density function with

$$\int tK(t)dt = 0, \int t^2 K(t)dt < \infty \text{ and } R(K) := \int K^2(t) \quad dt < \infty.$$

(A2) The sequence h_n of bandwidth satisfies

$$h_n \longrightarrow 0 \text{ and } nh_n \longrightarrow \infty\ as\ n \longrightarrow \infty \text{ and } \frac{nh_n}{Logn} \longrightarrow \infty, as\ n \longrightarrow\ \infty.$$

(A3) The kernel K is compactly supported, C^1-probability density, two times differentiable and such that $K, K^{(1)} and\ K^{(2)}$ are integrable, furthermore K, $K^{(2)}$ are lipschitz continous.

(A4) The functions $f(y)$ and $G^{-1}f$ are twice continuously differentiable in y.

(A5) The density $f(.)$ is differentiable up to order 3 and $f^{(2)}(.)$does not vanish.

(A6) The mode θ satisfies the following property: for any $\varepsilon > 0\ and\ t > 0$ there exists

ISSN: 1844 – 9581

$$\beta > 0 \; such \; that \; |\theta - t| \geq \varepsilon \; implies \; |f(\theta) - f(t)| \geq \beta.$$

(A7) h_n satisfies for $n \longrightarrow \infty$, $h_n \longrightarrow 0$; $nh_n^3 \longrightarrow 0 \; and \; nh_n^7 \longrightarrow 0$.

(A8) The sets $\boldsymbol{\mathcal{F}} = \left\{K\left(\frac{x-.}{h}\right): x \in \mathbb{R}, h \in \mathbb{R} - \{0\}\right\}$ is a bounded V-C-class of measurable functions.

3.1. CONSISTENCY

In this section, we show strong consistency with a rate for our proposed estimate.

Proposition 3.1. Under assumptions **(A1)−(A5)** we have

$$\sup_{y \in \Omega} \left|\hat{f}_n(y) - f(y)\right| = O\left(max\left(\left(\frac{logn}{nh_n}\right)^{1/2}, h_n^2\right)\right) \; a.s \; as \; n \longrightarrow \infty.$$

Theorem 3.1. Under the hypotheses of Proposition 3.1. Furthermore, if **(A6)** holds, for *n* large enough, we have

$$\hat{\theta}_n - \theta = O\left(max\left(\left(\frac{logn}{nh_n}\right)^{1/4}, h_n\right)\right) \; a.s \; as \; n \longrightarrow \infty.$$

Remark 3.1. If we choose $h_n = O\left(\left(\frac{logn}{n}\right)^{1/5}\right)$, which is the optimal choice with respect to the almost sure uniform convergence criterion in the density estimation (see [10]), we get

$$\hat{\theta}_n - \theta = O\left(\left(\frac{logn}{n}\right)^{1/5}\right) \; a.s \; as \; n \longrightarrow \infty,$$

which is the optimal rate as that one obtained in the complete data case (see [11]).

3.2. ASYMPTOTIC NORMALITY

Now, suppose that the density function$f(.)$ is unimodal at θ. Under **(A5)** we have

$$f^{(1)}(\theta) = 0 \; and \; f^{(2)}(\theta) < 0.$$

Similarly, we have

$$\hat{f}_n^{(1)}\left(\hat{\theta}_n\right) = 0 \; and \; \hat{f}_n^{(2)}\left(\hat{\theta}_n\right) < 0.$$

Using a Taylor expansion in neighborhood of θ, we get

$$\hat{f}_n^{(1)}\left(\hat{\theta}_n\right) = \hat{f}_n^{(1)}(\theta) + \left(\hat{\theta}_n - \theta\right)\hat{f}_n^{(2)}(\bar{\theta}_n) = 0,$$

where $\bar{\theta}_n$ is between $\hat{\theta}_n$ and θ, which gives

$$\hat{\theta}_n - \theta = -\frac{\hat{f}_n^{(1)}(\theta)}{\hat{f}_n^{(2)}(\bar{\theta}_n)}. \qquad (3)$$

Now, to establish the asymptotic normality, we show that the numerator in **(3)**, is asymptotically normally distributed and the denominator converges in probability to $f^{(2)}(\theta)$. The result is given in the following theorem.

Theorem 3.2. We assume that hypothesis **(A1)−(A6)** hold, then we have

$$\left(\frac{nh_n^3\left(\hat{f}_n^{(2)}(\theta)\right)^2}{\sigma^2}\right)^{1/2}\left(\hat{\theta}_n-\theta\right)\overset{D}{\to}N(0,1)$$

where $\overset{D}{\to}$ means the convergence in distribution, $N(0,1)$ is the standard normal distribution and $\sigma^2=(\alpha f(\theta)/G(\theta))R\left(K^{(1)}\right)$.

Remark 3.2. In complete data case (i.e., $\alpha=G(.)=1$), we have $\sigma^2=f(\theta)R\left(K^{(1)}\right)$, that is what was obtained by [12].

Corollary 3.1. Using a plug-in method by replacing α and f by their estimates, permits us to obtain a convergent estimate σ_n^2 of σ^2. From Theorem 3.2, we get for each fixed $v\in(0,1)$, the following $(1-v)\%$ asymptotic confidence interval for θ, namely,

$$\theta=\hat{\theta}_n\mp\sigma_n\left(nh_n^3\left(\hat{f}_n^{(2)}(\theta)\right)^2\right)^{-1/2}z_{1-v/2},$$

where $z_{1-v/2}$ denotes the $1-v/2$ quantile of the standard normal distribution $N(0,1)$.

4. SIMULATION STUDIES

In this section, we illustrate the finite sample behavior of our estimator $\hat{\theta}_n$ defined in (**1**) and examine its asymptotic normality. Here it is assumed that for double truncation, we consider the case $U^*$and V^* are mutually independent. Results in this section have been obtained with R-Software package DATDA (see, [13]). First, we present two simulated models which permitte to compute the estimator $\hat{\theta}_n$.

Model 1 (exponential decreasing case): The variable Y is distributed as a normal $N(\mu,\sigma^2)$, with density

$$f(y)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2},y\in\mathbb{R},$$

which admits a mode θ equal to the mean µ.

Model 2 (heavy tail case): The variable Y is distributed as a three parameters Weibull(λ, β, γ) with density

$$f(y)=\left(\frac{\beta}{\lambda}\right)\left(\frac{x-\gamma}{\lambda}\right)^{\beta-1}e^{-\left(\frac{x-\gamma}{\lambda}\right)^\beta}1_{(y\geq 0)},$$

where $\gamma\in\mathbb{R}$, β > 0, λ > 0 are the location, shape and scale parameters respectively, of the distribution and which admits a mode θ at the point $\gamma+\lambda\left(1-\frac{1}{\beta}\right)^{1/\beta}$.

We take $U^{*} \sim Exp(b)$ and $V^{*} \sim Exp(c)$, where b and c are chosen in order to obtain the following percentages of truncation (PT): $70\%, 50\%, 30\%$ and 10% corresponding to $(\alpha = 0.3, 0.5, 0.7$ and $0.9)$ respectively.

Note that this choice of truncation percentages is standard in this type of study. It is clear that, lower values of α are of little practical interest, for $\alpha \approx 0$, almost no data can be observed. Using this scheme, $B = 500$ independent samples of size n were generated for each Model. Sample sizes of $n = 50, n = 150$ and $n = 300$ were considered. The truncation occurs when $U^{*} \leq Y^{*} \leq V^{*}$ is violated. This means that, for each trial, the number of simulated data is much larger than n, actually $\alpha \approx \frac{n}{N}$ are needed on average, where recall that α stands for the proportion of no truncation.

The small samples cases (with *n* often somewhere between 20 and 50) is based on the use of robust estimators, which is not the case of this study, and may be considered in our future work.

For each sample, using plug-in estimate σ^2 we estimated the mode and we compute the bias, variance (*Var*) and root mean squared error (*MSE*) of the proposed estimator. Results are displayed in Tables 1 and 2. Recall that in nonparametric estimation, optimality (in the *MSE* sense) is not seriously swayed by the choice of the kernel K but is affected by the choice of the bandwidth h_n.

In this study, the bandwidth h_n is chosen to satisfy the assumptions above, and the kernel *K* is Gaussian.The bandwidth that we used in estimating the mode is that used by [8], based on the minimization of the *AMISE* (f_n) with respect to h_n, which leads to the asymptotically optimal bandwidth:

$$h_{opt} = \left(\frac{\alpha R(K) \int G^{-1} f}{R(f'') \mu_2^2(K)} \right)^{1/5} n^{-1/5} = \left(\frac{4}{3} \alpha \int G^{-1} f \right)^{1/5} \sigma n^{-1/5}.$$

As one can see it in Tables 1 and 2, the quality of the estimator does not seem to be affected by the percentages of truncation, and the *MSE* decreases when the sample size increases. Moreover, to illustrate the behavior of the estimator, we plotted for different values of *n*, the histogram and the corresponding Q−Q-plot against the standard Normal distribution in figures 1 to 4 for model 1 and figures 5 to 8 for model 2. Furthermore, a significance level (p-value) of Shapiro-Wilk normality test is greater than 0.05 in all simulated scenarios. The normality assumption is therefore highly conserved.

Table 1. Average estimated Bias, Variance and MSE, exponential decreasing case.

PT	***N***	***n***	***Bias***	***Var***	***MSE***
70%	167	50	0.01469	0.11657	0.34174
	500	150	0.00097	0.06888	0.26246
	1000	300	0.02786	0.04511	0.21421
50%	100	50	0.03414	0.11857	0.11974
	300	150	0.01619	0.07090	0.07116
	600	300	0.01491	0.04946	0.04968
30%	72	50	-0.00580	0.10191	0.31928
	215	150	-0.00061	0.06139	0.24778
	429	300	-0.00509	0.04429	0.21053
10%	56	50	-0.02383	0.10695	0.32790
	167	150	0.01403	0.06385	0.25307
	334	300	0.00483	0.04243	0.20605

Table 2. Average estimated Bias, Variance and MSE, heavy tail case.

PT	*N*	*n*	*Bias*	*Var*	*MSE*
70%	167	50	-0.00903	0.00818	0.09091
	500	150	0.00273	0.00476	0.06902
	1000	300	-0.00214	0.00456	0.06755
50%	100	50	-0.01539	0.00820	0.09185
	300	150	-0.00770	0.00469	0.06889
	600	300	-0.00558	0.00331	0.05779
30%	72	50	-0.01044	0.00839	0.09221
	215	150	-0.00252	0.00443	0.06663
	429	300	-0.00608	0.00348	0.05929
10%	56	50	0.00480	0.00848	0.09223
	167	150	0.01018	0.00462	0.06874
	334	300	0.00837	0.00337	0.05865

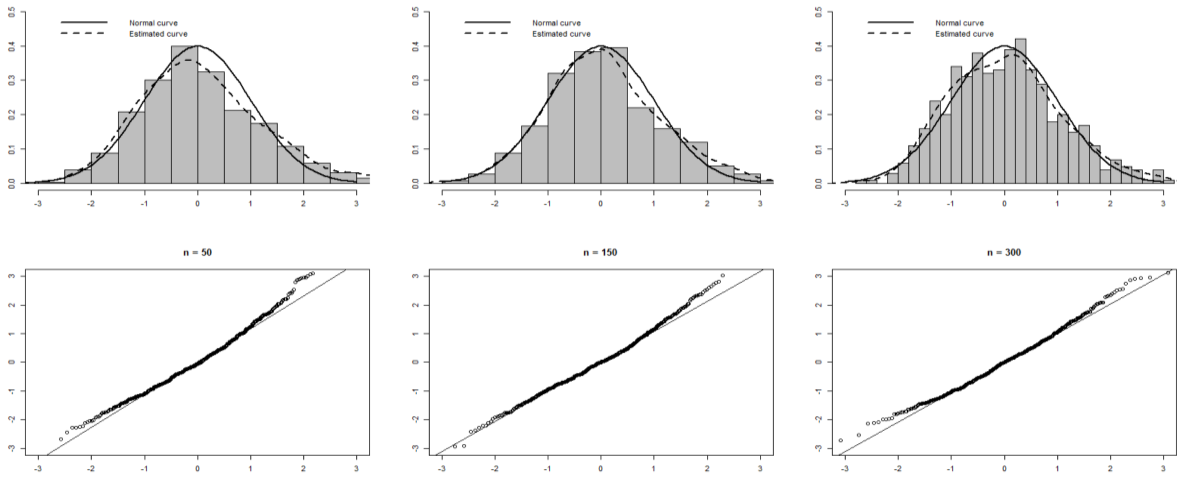

Figure 1. (Model 1): α = .30, B=500, n=50, 150 and 300 respectively.

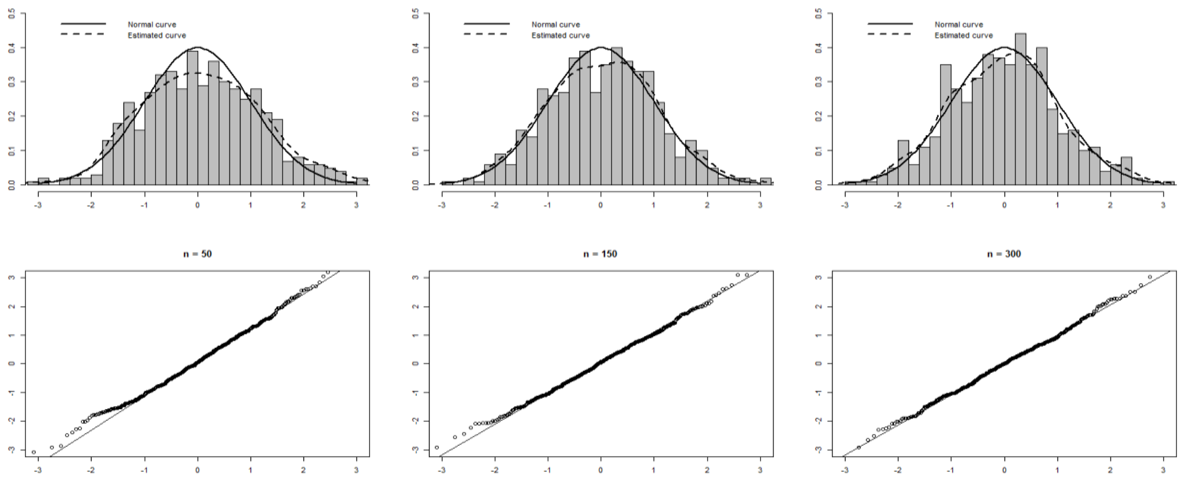

Figure 2. (Model 1): α = .50, B=500, n=50, 150 and 300 respectively.

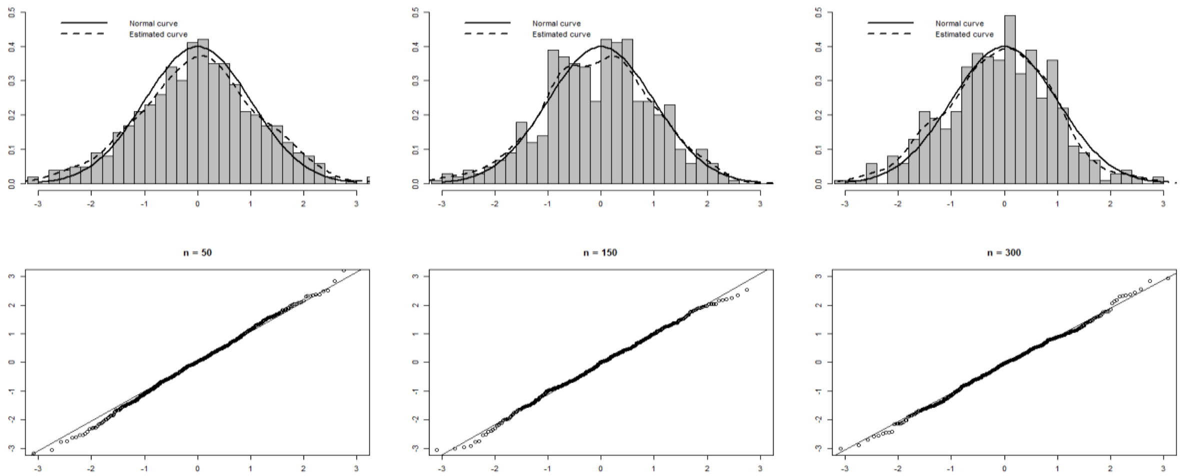

Figure 3. (Model 1): α = .70, B=500, n=50, 150 and 300 respectively.

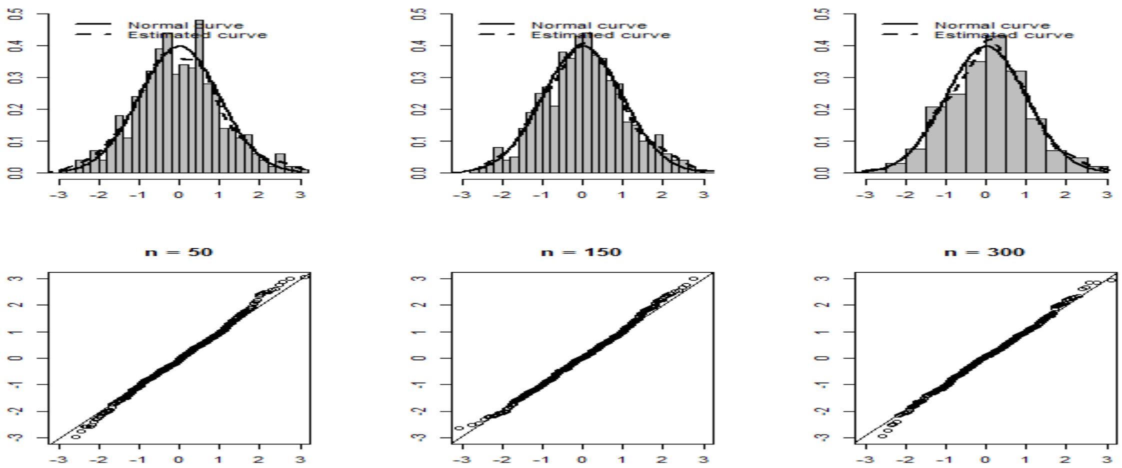

Figure 4. (Model 1): α = .90, B=500, n=50, 150 and 300 respectively.

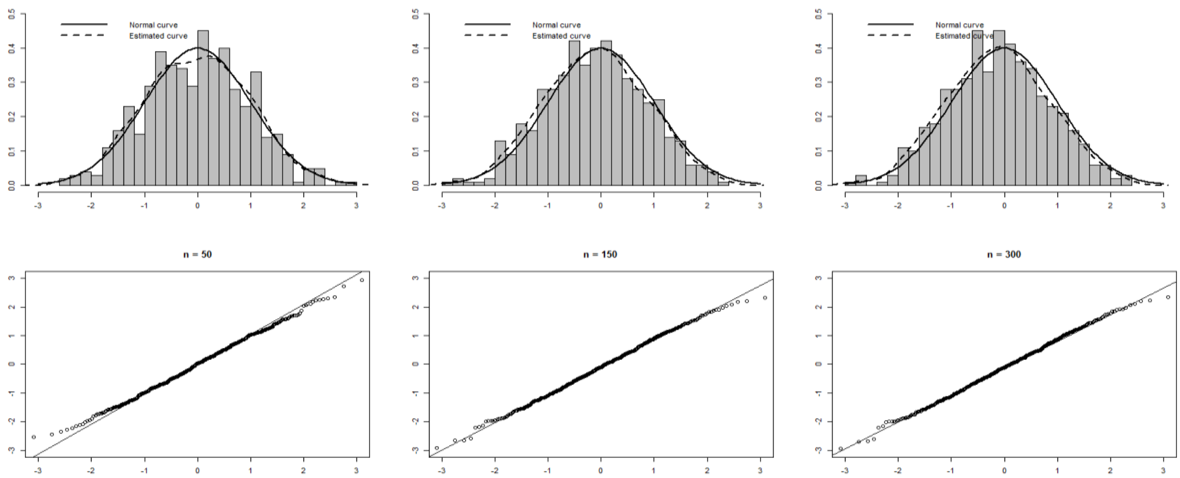

Figure 5. (Model 2): α = .30, B=500, n=50, 150 and 300 respectively.

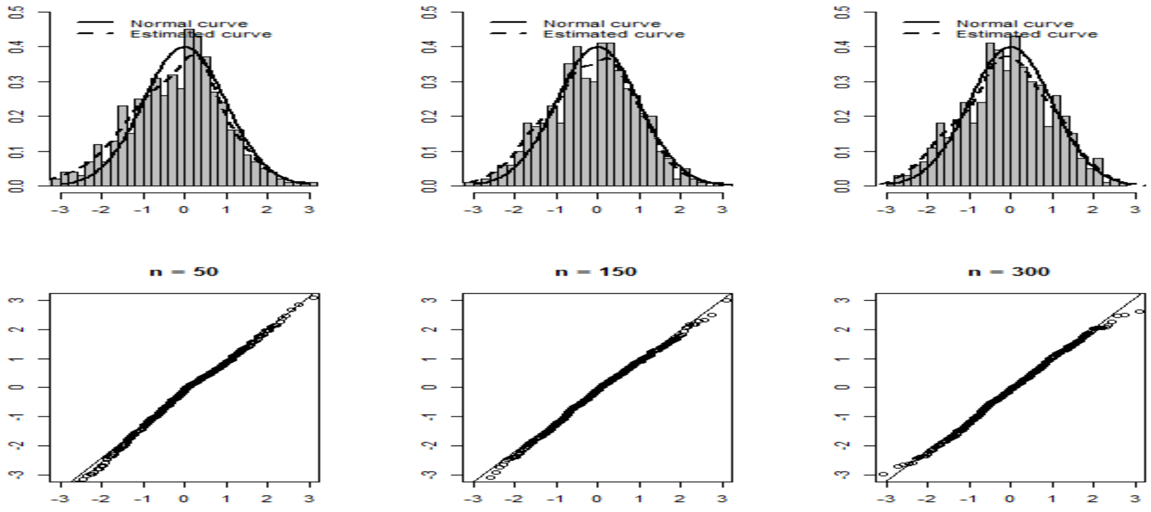

Figure 6. (Model 2): α = .50, B=500, n=50, 150 and 300 respectively.

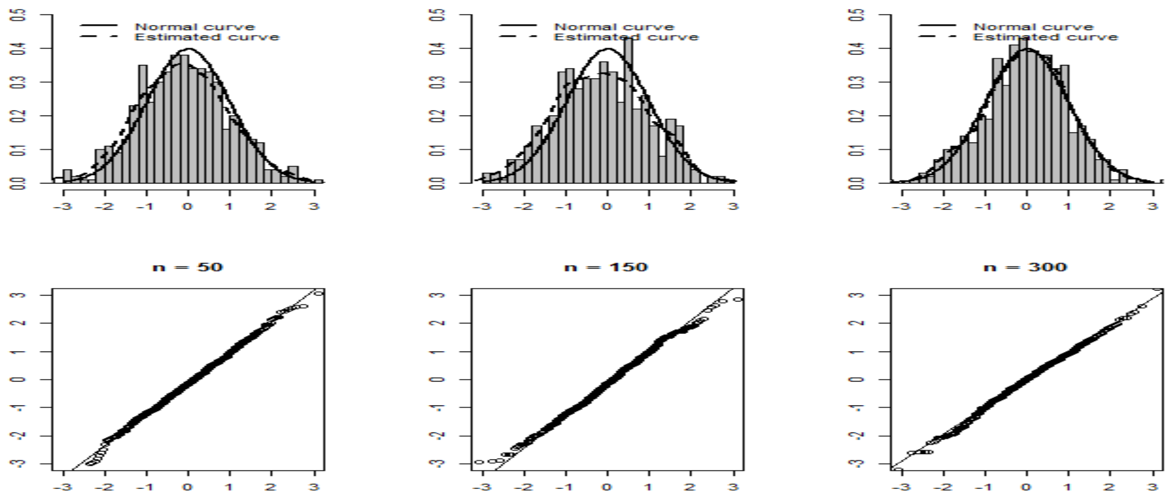

Figure 7. (Model 2): α = .70, B=500, n=50, 150 and 300 respectively.

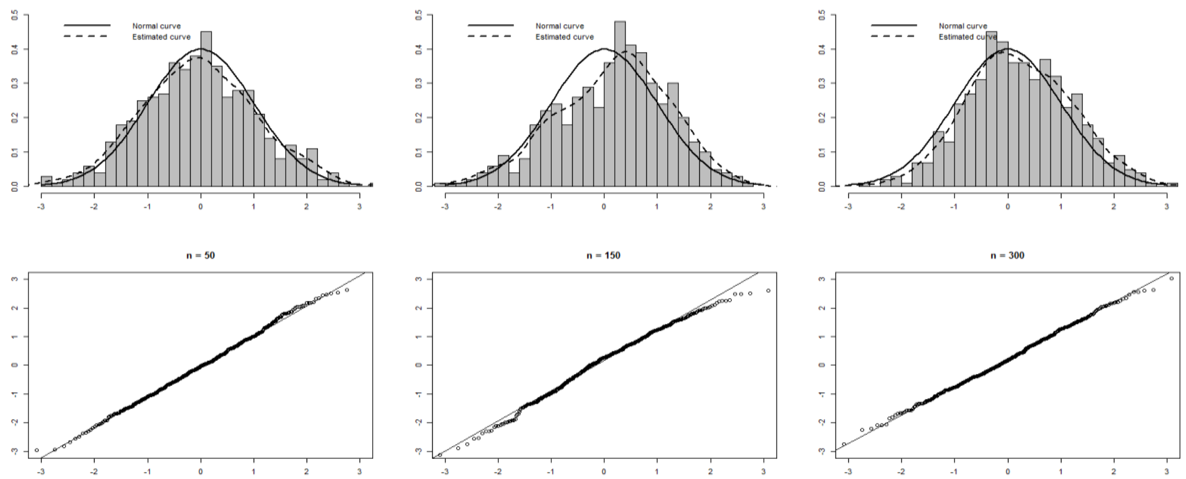

Figure 8. (Model 2): α = .90, B=500, n=50, 150 and 300 respectively.

5. AUXILIARY RESULTS AND PROOFS

For technical reason we need to introduce the pseudo-estimator of the density f, denoted $\tilde{f}_n$, and analogously to **(2)** we define it by

$$\tilde{f}_n(y) = \frac{\alpha}{nh_n}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{y-Y_i}{h_n}\right),$$

and its derivative is given by

$$\tilde{f}_n^{\,(j)}(y) = \frac{\alpha}{nh^{j+1}{}_n}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K^{(j)}\left(\frac{y-Y_i}{h_n}\right).$$

Lemma 5.1. Under assumptions **(A1)**, **(A3)** and **(A5)** we have

$$sup_{y\in\Omega}\left|E\left[\tilde{f}_n(y)\right]-f(y)\right| = O(h_n^2).$$

Proof: Using a change of variable and a Taylor expansion, under assumptions **(A1)**, **(A3)** and **(A5)** we get

$$\begin{aligned}
E\left[\tilde{f}_n(y)\right]-f(y) &= \frac{1}{h_n}\,E\left[\frac{\alpha}{nh_n}\sum_{i=1}^{n}\frac{1}{G(Y_i)}\,K\left(\frac{y-Y_i}{h_n}\right)\right]-f(y)\\
&= \frac{1}{h_{\boldsymbol{n}}}\int\frac{\alpha}{G(t)}K\left(\frac{y-t}{h_n}\right)f(t)dt-f(y)\\
&= \frac{1}{h_{\boldsymbol{n}}}\int K\left(\frac{y-t}{h_n}\right)f(t)dt-f(y)\\
&= \int K(u)\left[f(y)-h_nuf^{(1)}(y)+\frac{h^2u^2}{2}f^{(2)}(y)+O(h_n^2)\right]du-f(y)\\
&= \int K(u)\frac{h_n^2u^2}{2}f^{(2)}(y)du.
\end{aligned}$$

Thus

$$L_1 = \left|E\left[\tilde{f}_n(y)\right]-f(y)\right| \le \frac{h_n^2}{2}\sup_{y\in\Omega}\left|f^{(2)}(y)\right|\int u^2K(u)du = O(h_n^2)$$

the result holds.

Lemma 5.2. Under assumptions **(A2)** and **(A4)**, for n large enough, we have

$$\sup_{y\in\Omega}\left|\hat{f}_n(y)-\tilde{f}_n(y)\right| = O\left((nh_n)^{-1/2}\right).$$

Proof: We have

$$\hat{f}_n(y)-\tilde{f}_n(y) = \frac{\alpha_n}{nh_n}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K\left(\frac{y-Y_i}{h_n}\right)-\frac{\alpha}{nh_n}\sum_{i=1}^{n}\frac{1}{G(Y_i)}\,K\left(\frac{y-Y_i}{h_n}\right)$$

$$\begin{aligned}
L_2 = \left|\hat{f}_n(y)-\tilde{f}_n(y)\right| &\le \frac{1}{nh_n}\left|\frac{\alpha_n}{G_n(y)}-\frac{\alpha}{G(y)}\right|\sum_{i=1}^{n}K\left(\frac{y-Y_i}{h}\right)\\
&\le \frac{1}{nh_n}\sup_{y\in\Omega}\left|\frac{\alpha_n}{G_n(y)}-\frac{\alpha}{G(y)}\right|.
\end{aligned}$$

As pointed in [8], G_n is $\sqrt{n}$-Consistent estimator of G_n. In fact, $\sqrt{n}$-Consistency of G_n is a consequence of that of F_n and H_n^* (see also, [14] for more details). Therefore, under regularity we get the result.

Lemma 5.3. Under assumptions **(A2),(A3),(A5)** and **(A8)**, we have

$$sup_{y\in\Omega}\left|\tilde{f}_n(y) - E\left[\tilde{f}_n(y)\right]\right| = O\left(\sqrt{\frac{Logn}{nh_n}}\right).$$

Proof: Note that

$$L_3 = \sup_{y\in\Omega}\left|\tilde{f}_n(y) - E\left[\tilde{f}_n(y)\right]\right|.$$

Under **(A3), (A8)** the sequence

$$\xi_n = \left\{\Phi_y(u) = \frac{\alpha}{nh_n}\frac{1}{G(y)}K\left(\frac{y-u}{h_n}\right);\ a_F \le y \le b_F\right\}\ n \ge 1$$

is a V-C classes of measurable functions abounded with respective envelope $U_n = \frac{\|K\|_\infty}{nh_n G(y)}$. Moreover, under (**A5)**

$$E\left[\Phi_y^2(Y)\right] \le \frac{1}{nh_n^2}\frac{\|K\|_2^2\|f\|_\infty}{G(y)} \le \frac{\|f\|_\infty}{nh_n^2 G(y)} = \sigma_n^2$$

with $\sigma_n \le U_n$ for n large enough. Applying Talagrand's inequality (see, Proposition 2.2 in [15]) with $t = B_3\sqrt{\frac{logn}{nh_n^2}}$, for a positive constant B_3, we get

$$P\left\{\sup_{\Phi_y\in\,\xi_n}\left|\sum_{i=1}^{n}\left\{\Phi_y(Y_i) - E\left[\Phi_y(Y)\right]\right\}\right| \ge\ B_3\sqrt{\frac{logn}{nh_n^2}}\right\}$$
$$\le B_2\,exp\left\{\frac{-1}{B_2}\frac{B_3\sqrt{\frac{logn}{nh_n^2}}}{\|K\|_\infty}\,nh_nG(y)\,log\left[1+\frac{B_3\sqrt{\frac{logn}{nh_n^2}}\frac{\|K\|_\infty}{nh_nG(y)}}{\left[\sqrt{n}\frac{\sqrt{\|f\|_\infty}}{n\sqrt{h_nG(y)}}+\frac{\|K\|_\infty}{nh_nG(y)}\sqrt{log\frac{VU_n}{\sigma_n}}\right]^2}\right]\right\}$$

Under **(A2)** and using $\log(1+w) \sim w(for\ w \longrightarrow\ 0)$, the last quantity is of order

$$B_2exp\left\{-\frac{1}{B_2}\frac{B_3\sqrt{\frac{logn}{nh_n^2}}}{\|K\|_\infty}\,nh_nG(y)\frac{B_3\sqrt{\frac{logn}{nh_n^2}}\frac{\|K\|_\infty}{nh_nG(y)}}{B_2n\frac{\sqrt{\|f\|_\infty}}{n^2\sqrt{h_nG(y)}}\sigma_n^2}\right\}$$
$$= B_2\,exp\left\{-\frac{1}{B_2^2}\frac{B_3^2\frac{logn}{nh_n^2}}{n\frac{\|f\|_\infty}{n^2h_n^2G(y)}}\right\} = B_2 n^{-\frac{B_3^2\ G(y)}{B_2^2\|f\|_\infty}},$$

which for n large enough and by an appropriate choice of B_3, can be made $O(n^{-3/2})$. The latter being a general term of a summable series, then a direct application of Borel-Cantelli's lemma and the result is

$$L_3 = O\left(\left(\frac{logn}{nh_n^2}\right)^{1/2}\right) \; as \; n \longrightarrow \infty.$$

Proof of Proposition 3.1. Using the triangular inequality, we have

$$\sup_{y\in\Omega}\left|\hat{f}_n(y) - f(y)\right| \le \sup_{y\in\Omega}\left|\hat{f}_n(y) - \tilde{f}_n(y)\right| + \sup_{y\in\Omega}\left|\tilde{f}_n(y) - E\tilde{f}_n(y)\right| + \sup_{y\in\Omega}\left|E\tilde{f}_n(y) - f(y)\right|.$$

Then lemmas **5.1, 5.2** and **5.3** give the result.

Proof of Theorem 3.1. The uniform consistency of $\hat{\theta}_n$ follows from Proposition 3.1, assumption **(A5)** and the following results

$$\begin{aligned}\left|f(\hat{\theta}_n) - f(\theta)\right| &\le \left|f(\hat{\theta}_n) - \hat{f}_n(\hat{\theta}_n)\right| + \left|\hat{f}_n(\hat{\theta}_n) - f(\theta)\right| \\ &\le \sup_{y\in\Omega}\left|\hat{f}_n(y) - f(y)\right| + \left|\hat{f}_n(\hat{\theta}_n) - f(\theta)\right| \le 2\sup_{y\in\Omega}\left|\hat{f}_n(y) - f(y)\right|.\end{aligned} \qquad \textbf{(4)}$$

For the second part, a Taylor expansion of $f(.)$ in neighborhood of θ gives

$$f(\hat{\theta}_n) - f(\theta) = \frac{1}{2}(\hat{\theta}_n - \theta)^2 f''(\bar{\theta})$$

where $\bar{\theta}$ is between $\hat{\theta}_n$ and θ. Thus, (**4**) and assumptions (**A4)** and (**A5)** yield

$$(\hat{\theta}_n - \theta)^2\left|f''(\bar{\theta})\right| \le 4\sup_{y\in\Omega}\left|\hat{f}_n(y) - f(y)\right|$$

Thus,

$$\left|\hat{\theta}_n - \theta\right| \le 2\sqrt{\frac{\sup\left|\hat{f}_n(y) - f(y)\right|}{\left|f''(\bar{\theta})\right|}}.$$

Using **Proposition 3.1,** the proof is complete.

Proof of Theorem 3.2. From **(3)** we have the following decomposition

$$\sqrt{nh_n^3}(\hat{\theta}_n - \theta) = \sqrt{nh_n^3}\frac{\hat{f}_n^{(1)}(\theta) - \tilde{f}_n^{(1)}(\theta)}{\hat{f}_n''(\bar{\theta}_n)} + \sqrt{nh_n^3}\frac{\tilde{f}_n^{(1)}(\theta) - E\left(\tilde{f}_n^{(1)}(\theta)\right)}{\hat{f}_n^{(2)}(\bar{\theta}_n)}$$

$$+\sqrt{nh_n^3}\frac{E(\tilde{f}_n^{(1)}(\theta))}{\hat{f}_n^{(2)}(\bar{\theta}_n)}$$

$$=: \frac{J_1 + J_2 + J_3}{\hat{f}_n^{(2)}(\bar{\theta}_n)}.$$

To prove the results, we establish that J_1 and J_3 are negligible and J_2 is asymptotically normal and $\hat{f}_n^{(2)}(\bar{\theta}_n) \longrightarrow f^{(2)}(\theta)$. For the first term J_1, we have

$$J_1 = \hat{f}_n^{(1)}(\theta) - \tilde{f}_n^{(1)}(\theta) = \frac{\alpha_n}{nh_n^2}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K^{(1)}\left(\frac{\theta - Y_i}{h_n}\right) - \frac{\alpha}{nh_n^2}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K^{(1)}\left(\frac{\theta - Y_i}{h_n}\right).$$

Recall that

$$\sup_y \left|\frac{\alpha_n}{G_n(y)} - \frac{\alpha}{G(y)}\right| = O\left((nh_n)^{-\frac{1}{2}}\right),$$

(see the proof of lemma 5.2.) and since

$$\frac{1}{nh_n^2}\sum_{i=1}^{n} K^{(1)}\left(\frac{\theta - Y_i}{h_n}\right) = (f_n^*)^{(1)}(\theta) \longrightarrow (f^*)^{(1)}(\theta),$$

here, $(f^*)^{(1)}(.)$ denotes the first derivative of the density function $f^*(.)$ of the observed data, which is not necessarily equal to zero at the point θ, which permit us to conclude that $J_1$is negligible. Now, we state the following results for J_3:

$$E\left[\tilde{f}_n^{(1)}(\theta)\right] = \frac{\alpha}{nh_n^2}\int \frac{K^{(1)}\left(\frac{\theta-u}{h_n}\right)}{G(u)} f(u)du = \frac{1}{h_n^2}\int K^{(1)}\left(\frac{\theta-u}{h_n}\right) f(u)du$$
$$= \frac{1}{h_n}\int K^{(1)}(r) f(\theta - rh_n)dr.$$

Integrating by part, we have

$$E\left[\tilde{f}_n^{(1)}(\theta)\right] = \int K(r) f^{(1)}(\theta - rh_n)dr.$$

By Taylor expansion of $f^{(1)}(.)$ around θ, **(A1), (A5)** and the definition of the mode, we get

$$\sqrt{nh_n^3}E\left[\tilde{f}_n^{(1)}(\theta)\right] = \sqrt{nh_n^7}\int r^2 K(r) f^{(3)}(\bar{\theta})dr,$$

where $\bar{\theta}$ is between θ and $\theta - rh_n$, by **(A2), (A5)** and **(A7)**, the negligity of J_3 follows.

Finally, to state theasymptotic normality of J_2, we have to prove that

$$Var[J_2] \longrightarrow \frac{\alpha f(\theta)}{G(\theta)}\int \left[K^{(1)}(r)\right]^2 dr \; as \; n \longrightarrow \infty.$$

Note that

$$Var[J_2] = nh_n^3 Var\left[\frac{\alpha}{h_n^2}\frac{1}{G(y)}K^{(1)}\left(\frac{\theta - y}{h_n}\right)\right]$$

$$= nh_n^3 E\left[\frac{\alpha}{h_n^2}\frac{1}{G(y)}K^{(1)}\left(\frac{\theta - y}{h_n}\right)\right]^2 - nh_n^3\left\{E\left[\frac{\alpha}{h_n^2}\frac{1}{G(y)}K^{(1)}\left(\frac{\theta - y}{h_n}\right)\right]\right\}^2$$

$$=: I_1 + I_2$$

On the one hand, the negligity of J_3 gives as $\mathrm{I}_2 \longrightarrow 0$ as $\mathrm{n} \longrightarrow \infty$.On the other hand, using a change of variable we can write

$$I_1 = \int \frac{\alpha}{G(\theta)}\left[K^{(1)}(r)\right]^2 f(\theta - rh_n)\, dr.$$

Since $\mathrm{G}(.)$ is continous, we have under (A1), (A3), and (A5)

$$I_1 = \frac{\alpha f(0)}{G(\theta)}\int \left[K^{(1)}(r)\right]^2 dr + O(1), as \; n \longrightarrow \infty,$$

which gives a result.

Now, for any $y \in \mathbb{R}$, let us consider the following centered i.i.d random variables

$$R_i(y) = \frac{\alpha}{nh_n^2}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K^{(1)}\left(\frac{\theta - Y_i}{h_n}\right) - E\left[\frac{\alpha}{nh_n^2}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K^{(1)}\left(\frac{\theta - Y_i}{h_n}\right)\right] \; 1 \le i \le n.$$

Then, a simple algebraic calculation gives us

$$\sum_{i=1}^{n} R_i(y) = \sqrt{nh_n^3}\left(\tilde{f}_n^{(1)}(y) - E\left[\tilde{f}_n^{(1)}(y)\right]\right) =: J_2$$

Hence,

$$Var\left(\sum_{i=1}^{n} R_i\right) = nh_n^3 Var\tilde{f}_n(y) \longrightarrow \frac{\alpha f(\theta)}{G(\theta)}\int\left[K^{(1)}(r)\right]^2 dr. \quad \textbf{(5)}$$

We have,

$$R_i^2 \le \frac{2\alpha^2}{nh_n}\frac{1}{G^2(Y_i)}\left(K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right)^2 + \frac{2\alpha^2}{nh_n}E^2\left[\frac{1}{G(Y_i)}K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right]. \quad \textbf{(6)}$$

Note that, the negligity of J_3 implies that the second term in the right hand of equation (**6**) goes to zero as n goes to infinity. By (**5**) there exists $n_0 \in \mathbb{N}^*$, such that $\forall\, n \ge n_0$, we have

$$Var\left(\sum_{i=1}^{n} R_i(y)\right) \ge \frac{\alpha f(\theta)}{2G(\theta)}\int\left[K^{(1)}(r)\right]^2 dr. \quad \textbf{(7)}$$

Now denote by

$$A(Y_i) = \frac{1}{G^2(Y_i)}\left(K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right)^2 + E^2\left[\frac{1}{G(Y_i)}K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right].$$

From (**6**), we have

$$R_i^2(y) \le \frac{2\alpha^2 A(Y_i)}{nh_n}.$$

Using (**7**), for $n \ge n_0$ we have

$$\{R_i^2(y) > \varepsilon^2 Var(\sum_{i=1}^{n} R_i(y))\} \subset \left\{R_i^2(y) > \varepsilon^2\frac{\alpha f(\theta)}{2G(\theta)}\int\left[K^{(1)}(r)\right]^2 dr\right\} = \{R_i^2(y) > 2\varepsilon'\}.$$

Now set $\varepsilon' = \varepsilon^2\frac{\alpha f(\theta)}{4G(\theta)}\int\left[K^{(1)}(r)\right]^2 dr > 0$, then

$$\left\{R_i^2(y) > \varepsilon^2 Var\left(\sum_{i=1}^{n} R_i(y)\right)\right\} \subset \{R_i^2(y) > 2\varepsilon'\} = \left\{\frac{nh_n}{2\alpha^2}R_i^2(y) > \varepsilon' nh_n\right\} \subset \{A(Y_i) > \varepsilon' nh_n\}$$

$$\subset \left\{\frac{1}{G^2(Y_i)}\left(K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right)^2 > \frac{\varepsilon' nh_n}{2}\right\} \cup \left\{E^2\left[\frac{1}{G(Y_i)}K^{(1)}\left(\frac{y-Y_i}{h_n}\right)\right] > \frac{\varepsilon' nh_n}{2}\right\}$$

$$= V_{1n} \cup V_{2n}.$$

By (**6**), for n large enough V_{2n} is empty in the same way, since G is lower bounded and $K^{(1)}$ bounded, by (A7), we have for n large enough that V_{1n} is empty. Therfore, we get $\{R_i^2 > \varepsilon^2 Var(\sum_{i=1}^{n} R_i(y))\} = \emptyset$. We conclude from the Lindeberg's theorem, that J_2 is normally distributed. Now, to complete the proof of Theorem 3.2, it suffices to prove that

$$\hat{f}_n^{(2)}(\bar{\theta}_n) \longrightarrow f^{(2)}(\theta).$$

Observe that one can write the following

$$\hat{f}_n^{(2)}(y) - f^{(2)}(y) = \left\{\frac{1}{h_n^2}\int K^{(2)}\left(\frac{y-t}{h_n}\right)\hat{f}_n(t)dt - \frac{1}{h_n^2}\int K^{(2)}\left(\frac{y-t}{h_n}\right)f(t)dt\right\}$$

$$+\left\{\frac{1}{h_n^2}\int K^{(2)}\left(\frac{y-t}{h_n}\right)f(t)dt - K\left(\frac{y-t}{h_n}\right)f^{(2)}(y)dt\right\}$$

$$=: T_1 + T_2.$$

By Taylor's expansion, integration by part and assumptions (A1), (A2), (A6) and (A7) we have

$$|T_1| \le \frac{1}{h_n}\left|\hat{f}_n(y - rh_n) - f(y - rh_n)\right| \int K^{(2)}(r)dr$$

$$\le \sup_{y\in\Omega}\left|\hat{f}_n(y) - f(y)\right| \int K^{(3)}(r)dr.$$

Then by Proposition 3.1, $|T_1| = O\left[\max\left(\frac{\mathrm{Logn}}{\mathrm{nh_n}}, h_n^2\right)\right]$.

On the other hand, under assumptions (A3), (A5) and integrating by part twice, we have

$$T_2 = \frac{1}{h_n^2}\int K^{(2)}\left(\frac{y-t}{h_n}\right)f(t)dt - \int K\left(\frac{y-t}{h_n}\right)f^{(2)}(t)dt$$

$$= \frac{1}{h_n^2}\int K\left(\frac{y-t}{h_n}\right)f^{(2)}(t)dt - \int K\left(\frac{y-t}{h_n}\right)f^{(2)}(t)dt.$$

Using a Taylor expansion, we have

$$|T_2| \le h_n \int r\ K(r)\left|f^{(3)}(r^*)\right|dr = O(h_n),$$

where r^* is between $y - rh_n$ and y. Finally, we have for n large enough

$$\left|\hat{f}_n^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta)\right| \le \sup_{y\in\Omega}\left|\hat{f}_n^{(2)}(y) - f^{(2)}(y)\right| + \left|f^{(2)}(\bar{\theta}_n) - f^{(2)}(\theta)\right|,$$

which ends the proof.

6. CONCLUSION

The main objective of this paper was to propose a new estimator of the mode function, which is observed under random double truncation. We have presented two simulated models (exponential decreasing case and heavy tail case) and have studied the finite sample behavior of our newly estimator of the mode. It has been shown that, in general the asymptotic normality is highly conserved since the selection of an appropriate bandwidth, which is known to be an important but delicate issue. This paper does not treat the bandwidth selection procedures. It remains a likely topic for future investigations. Also, the small samples cases (with n often somewhere between 20 and 50) is based on the use of robust estimators, which is not the case of this study, and may be considered in our future work. The literature also contains semi-parametric approaches to estimate the density function under double truncation, see for example [2,8]. Thus, in our further research we will study the mode estimation problem when the distribution is assumed to belong to a given parametric family. Finally, the conditional mode estimation function is a natural extension of the present work.

***Acknowledgment**: The authors wish to thank the editor and reviewers for their helpful and constructive comments that have much improved our manuscript.*

REFERENCES

[1] Bilker, W. B., Wang, M. C., *Biomterics*, **52**, 10, 1996.
[2] Efron, B., Petrosian, V., *Journal of the American Statistical Association*, **94**, 824, 1999.
[3] Lynden-Bell, D., *Monthly Notices of the Royal Astronomical Society*, **155**(1), 95, 1971.
[4] Stute, W., *Annals of Statistics*, **21**, 146, 1993.
[5] Woodroofe, M., *Annals of Statistics*, **13**, 163, 1985.
[6] Moreira, C., de Uña-Álvarez, J., *Statistics in Medicine*, **29**(30), 3147, 2010.
[7] Shen, P., *Communications in Statistics-Theory and Methods*, **39**, 3178, 2010.
[8] Moreira, C., de Una-Alvarez, J. *Electronic Journal of Statistics*, **6**, 501, 2012.
[9] Moreira, C., Van Keilegom, I., *Computational Statistics and Data Analysis*, **61**, 107, 2013.
[10] Stone, C.J., *Recent advances in statistics*, Academic Press, Massachusetts, United States, 1983.
[11] Vieu P., *Jornal of Statistics and Probability Letters*, **26**, 297, 1996.
[12] Parzen, E., *Annals of Mathematical Statistics*, **33**(3)**,** 1065, 1962.
[13] Moreira, C., de Uña-Alvarez, J., Crujeiras-Casais, R.M., *Journal of Statistical Software*, **37**(7), 1, 2010.
[14] Shen, P., *Annals of the institute of Statistical Mathematics*, **62**, 853, 2010.
[15] Giné, E., Guillou, A.E., *Annals of Probability*, **27**, 2042, 1999.