Asymmetric kernels for boundary modification in distribution function estimation

Authors: Habib Allah Mombeni
– Statistics Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran (habiballamombeni@gmail.com)

Behzad Mansouri
– Statistics Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran (b.mansouri@scu.ac.ir)

Mohammad Reza Akhoond
– Statistics Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran (mra.biostat@gmail.com)

Abstract:

- Kernel-type estimators are popular in density and distribution function estimation. However, they suffer from boundary effects. In order to modify this drawback, this study has proposed two new kernel estimators for the cumulative distribution function based on two asymmetric kernels including the Birnbaum-Saunders kernel and the Weibull kernel. We show the asymptotic convergence of our proposed estimators in boundary as well as interior design points. We illustrate the performance of our proposed estimators using a numerical study and show that our proposed estimators outperform the other commonly used methods. The illustration of our proposed estimators to a real data set indicates that they provide better estimates than those of the formerly-known methodologies.

Key-Words:

- *Cumulative distribution function; Boundary effects; Kernel-type estimators; Asymmetric kernels.*

AMS Subject Classification:

- 62G05, 62G20.

1. INTRODUCTION

Suppose that $X_1, X_2, \ldots, X_n$ be a set of continuous random variables with unknown cumulative distribution function $F(x)$ which we wish to estimate. The Empirical distribution function provides a uniformly consistent estimate of the cumulative distribution function. However, estimations which are provided by the Empirical distribution are not smooth. Another approach for estimating the cumulative distribution function is to use Kernel-type estimators. Kernel-type estimators for distribution estimation, based on symmetric kernels, have been introduced by authors such as Nadaraya [14] and Watson and Leadbetter [21], and their asymptotic properties have been investigated by Singh et al. [18]. Asymptotical superiority of Kernel-type estimators to the empirical distribution function at a single point in density estimation was shown by Reiss [15] and Falk [5].

Although the symmetric kernels are popular and commonly used in Kernel-type estimators, they are not efficient for those distribution (density) functions which have a compact support due to the boundary bias. This problem is known as *boundary effects* and several approaches have so far been proposed to deal with it in regression and density estimation tasks (Gasser and Muller [6], Rice [16], Gasser et al. [7] and Muller [12]). In a similar manner, Tenreiro [19] proposed some boundary kernels for estimating a cumulative distribution function with a finite interval support. These approaches, hereafter called the Boundary kernel methods or briefly the B-K methods, are based on symmetric kernels.

Asymmetric kernel functions were introduced by Chen [2] as an alternative approach to the boundary correction in kernel density estimation. He proposed the beta kernel density estimator to estimate a density with support on [0, 1]. Chen [3] considered the gamma kernel density estimator to estimate a density with support on $[0, \infty)$. In order to provide a boundary-free estimation for the density function $f(x)$ with support on $[0, \infty)$ by the gamma kernel density estimator, Zhang [22] has shown that having a shoulder at x=0, whose derivative of $f(x)$ is zero at $x = 0$, is a necessary condition. For densities not satisfying this condition, the gamma kernel density estimator suffers from severe boundary problems. This approach was extended for estimating a density with support on $[0, \infty)$ using other asymmetric kernels (Jin and Kawczak [10], Scaillet [17], Hirukawa and Sakudo [8] and Hirukawa and Sakudo [9]).

So far, the boundary effects in density estimation have attracted the attention of many researchers. Accordingly, several methods, using symmetric and asymmetric kernels, have been proposed to solve the problem. However, in the cumulative distribution estimation, the boundary effects have received little if any attention. In this paper, we have focused on estimating those distribution functions with support on $[0, \infty)$ and proposed a new Kernel-type estimator for the cumulative distribution function based on asymmetric kernels. Our estimator at the design point x has the following form:

$$\hat{F}_n(x) = n^{-1} \sum_{i=1}^{n} \bar{K}_{x,b}(X_i), \tag{1.1}$$

where $\bar{K}_{x,b}(t) = \int_t^\infty k_{x,b}(u)du$ and $k(.)$ is an asymmetric kernel function on [0, ∞) with the smoothing parameter b. Thus, the kernel has the same support as the true distribution function. We introduce two estimators by considering two asymmetric kernels including the Birnbaum-Saunders (B-S) kernel and the Weibull kernel. In the next Section, we demonstrate the asymptotic properties of our proposed estimator based on the B-S kernel, hereafter called the B-S kernel estimator. We investigate the rate of convergence of the B-S kernel estimator both in the interior and the boundary points. In Section three, we have run the same study for our second estimator which is based on the Weibull kernel, hereafter called the Weibull kernel estimator. The rest of the paper is organised as follows. Section four is dedicated to illustrating the performance of our proposed estimators. We conducted a comprehensive numerical study and considered various cumulative distribution functions to estimate and compare the performance of our estimators with other existing methods. In section five we have illustrated the performance of our proposed estimators on a real data set. Finally, Section six is devoted to discussions and conclusions.
In this paper, we assume that the cumulative distribution function $F(x)$ is satisfies in the following assumptions:

Assumption 1 The cumulative distribution function $F(x)$ is absolutely continuous with respect to Lebesgue measure on $(0,\infty)$ and has two continuous and bounded derivatives.

Assumption 2 The smoothing parameter $b = b_n > 0$ satisfies $b \to 0$, $as\, n \to \infty$.

Assumption 3 The following integrals

$$\int_0^\infty (xf(x))^2\,dx \quad and \quad \int_0^\infty \left(x^2 f'(x)\right)^2 dx, \tag{1.2}$$

are finite.

Following Hirukawa and Sakudo [9] and 'In order to describe different asymptotic properties of an asymmetric kernel estimator across positions of the design point $x > 0$', we denote by 'interior x' and a sequence of points converging to the boundary or 'boundary x' a design point x that satisfies $x/b \to \infty$ and $x/b \to k$ for some $0 < k < \infty$ as $n \to \infty$, respectively.

2. Asymmetric cumulative distribution function estimation using B-S kernel

In this Section, we aim at demonstrating the asymptotic convergence of our first proposed estimator: Equation (1.1) based on the B-S kernel, i.e. the B-S kernel estimator. To forward this end we will show that the B-S kernel estimator is asymptotically unbiased and consistent. We will obtain an appropriate smoothing parameter for our estimator through minimizing the mean integrated square error. In addition, we will discuss the convergence rate of the B-S kernel estimator in the boundary points.

2.1. Asymptotic properties of the B-S kernel estimator

Consider the Birnbaum–Saunders kernel given by

$$\bar{K}_{B-S}(t;\beta,\alpha) = 1 - \Phi\left(\frac{\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)}{\alpha}\right), \qquad t > 0,\ \alpha > 0,\ \beta > 0, \tag{2.1}$$

where $\Phi(.)$ is the Standard Normal distribution function. Let $\alpha = \sqrt{b}$ and $\beta = x$, where x and b denote the design point and the smoothing parameters, respectively. The B-S kernel estimator for the cumulative distribution function is defined as:

$$\hat{F}_1(x) = n^{-1}\sum_{i=1}^{n} \bar{K}_{B-S}\left(X_i; x, \sqrt{b}\right). \tag{2.2}$$

In what follows we will obtain two approximate expressions for the bias and variance for $\hat{F}_1(x)$ in Lemma 2.1 and Lemma 2.1, respectively. First consider that for the two continuous distribution functions F and G and their corresponding density function f and g, it is easy to show that:

$$E_g\left(F(X)\right) = 1 - E_f\left(G(X)\right). \tag{2.3}$$

where $E_g\left(F(X)\right)$ is the expectation of $F(X)$, when X is a randon variable following the distribution G.

Lemma 2.1. *Suppose that Assumptions 1-3 hold, then we have:*

$$E\left(\hat{F}_1(x)\right) = F(x) + \frac{b}{2}\left(xf(x) + x^2 f'(x)\right) + O(b^2). \tag{2.4}$$

Proof of Lemma 2.1: Since X_i's are identical, we have

$$E_f\left(\hat{F}_1(x)\right) = E_f\left(\bar{K}_{B-S}(T; x, \sqrt{b})\right), \tag{2.5}$$

where T is a random variable following the distribution F. Using equation(2.3) and Taylor expansion, we have:

$$\begin{aligned} E_f\left(\bar{K}_{B-S}(T; x, \sqrt{b})\right) &= E_f\left(1 - K_{B-S}(T; x, \sqrt{b})\right) = E_k\left(F(T)\right) \\ &= F(x) + f(x)E(T - x) + \sum_{i=1}^{\infty} \frac{f^{(j)}(x)}{j!} E(T - x)^{j+1}, \end{aligned} \tag{2.6}$$

where $f^{(j)}(.)$ is the jth derivative of $f(x)$ and now $T \sim k_{x,\sqrt{b}}(t)$ where

$$k_{x,\sqrt{b}}(t) = \frac{t^{-\frac{3}{2}}(t+x)}{\sqrt{2\pi b x}} \exp\{-\tfrac{1}{2b}(\tfrac{t}{x} + \tfrac{x}{t} - 2)\}, \qquad t > 0\ , x > 0\ , b > 0, \tag{2.7}$$

Using the results of Johnson et al. [11], we have:

$$\begin{aligned} E(T-x) &= b\frac{x}{2}, \\ E(T-x)^2 &= \frac{bx^2}{2}(2+3b), \\ E(T-x)^3 &= \frac{9b^2x^3}{2}(3+5b), \end{aligned} \tag{2.8}$$

$$\begin{aligned} \Rightarrow E(F_1(x)) = F(x) + f(x)\left(b\frac{x}{2}\right) + \frac{f'(x)}{2}\left(\frac{bx^2}{2}(2+3b)\right) + \frac{f^{(2)}(x)}{6}\left(\frac{9b^2x^3}{2}(3+5b)\right) \\ + \cdots = F(x) + \frac{b}{2}\left(xf(x) + x^2 f'(x)\right) + O(b^2), \end{aligned}$$

where $f'(.)$ is the first derivative of $f(x)$. □

So, for the interior points, the bias of the B-S kernel estimator is of order $O(b)$. Although this rate of convergence to zero seems disappointing, one should be aware that the smoothing parameter is a function of n. In the remainder of this Section, we will show that by taking this relation into account and considering the rate of convergence based on n, the bias of the B-S kernel estimator is normal (not too bad). We defer a detailed discussion of this matter until later in Section 3 where we provide a comparison between the bias of the B-S kernel estimator and the Weibull kernel estimator. In addition, in the numerical study, we will see that the overall performance of the B-S kernel estimator is not only satisfactory but also better than the other competitors. This achievement is the result of a reduction in the variance of the B-S kernel estimator, as we will see in Lemma 2.2, and what is the so-called trade-off between the variance and the bias.
Now we turn to the variance of the B-S kernel estimator. The following Lemma shows that the variance of $\hat{F}_1(x)$ resembles the variance of the Empirical distribution function to some extent but it involves a negative term which can lead to its superiority over the Empirical distribution function since it has a smaller variance.

Lemma 2.2. *Suppose that Assumptions 1-3 hold, then variance of the B-S kernel estimator can be obtined as:*

$$Var(\hat{F}_1(x)) = n^{-1}F(x)\left(1 - F(x)\right) - n^{-1}b^{\frac{1}{2}}\pi^{-\frac{1}{2}}xf(x) + O(n^{-1}b). \tag{2.9}$$

Proof of Lemma 2.2: First consider that

(2.10)

$$
\begin{aligned}
E\left(\bar{K}^2_{B-S}(T;x,\sqrt{b})\right) &= \int_0^\infty \bar{K}^2_{B-S}(t;x,\sqrt{b})f(t)dt \\
&= \int_0^\infty F(t)\left(2k_{B-S}(t;x,\sqrt{b})\bar{K}_{B-S}(t;x,\sqrt{b})\right)dt \quad \text{(using integral by part)} \\
&= F(x) + f(x)E(Z-x) + \frac{1}{2}f'(x)E(Z-x)^2 + \ldots,
\end{aligned}
$$

where $Z \sim 2k_{B-S}(z;x,\sqrt{b})\bar{K}_{B-S}(z;x,\sqrt{b})$ (a skew probability density function) and

(2.11)

$$
k_{B-S}(z;x,\sqrt{b}) = \frac{z^{-\frac{3}{2}}(z+x)}{\sqrt{2\pi bx}}\exp\{-\frac{1}{2b}(\frac{z}{x}+\frac{x}{z}-2)\}\ , \qquad z>0\ ,x>0\ ,b>0,
$$

By extending the results of Vilca and Leiva [20], we have:

$$
\begin{aligned}
E(Z-x) &= \frac{b^{\frac{1}{2}}x}{2}\left(\omega_1 + b^{\frac{1}{2}}\gamma_2\right), \\
E(Z-x)^2 &= \frac{bx^2}{2}\left(2\gamma_2 + \gamma_4 + b^{\frac{1}{2}}x\omega_3\right),
\end{aligned}
\tag{2.12}
$$

where $\gamma_r = E(W^r)$ and $\omega_r = E(W^r\sqrt{bW^2+4})$. In addition, W is a random variable with a Skewed Normal distribution, i.e. $W \sim SN(0,1,-1)$.
Using the Taylor expansion for $W\sqrt{bW^2+4}$ and $W^3\sqrt{bW^2+4}$, we obtain

$$
W\sqrt{bW^2+4} = 2W + \frac{1}{4}bW^3 - \frac{1}{64}b^2W^5 + O(b^3),
$$

and

$$
W^3\sqrt{bW^2+4} = 2W^3 + \frac{1}{4}bW^5 - \frac{1}{64}b^2W^7 + O(b^3),
$$

Nadarajah and Kotz [13] show that, $E(W) = -\frac{1}{\sqrt{\pi}}$, $E(W^3) = -\sqrt{\frac{5}{4\pi}}$ thus we can deduce that

$$
\gamma_2 = 1, \quad \gamma_4 = 3, \quad \omega_1 \approx -\frac{2}{\sqrt{\pi}}, \quad \omega_3 \approx -\sqrt{\frac{5}{\pi}}.
\tag{2.13}
$$

By substituting γ_2, γ_4, ω_1 and ω_3 in (2.12) and then substituting (2.12) in (2.10), we obtain

$$
E\left(\bar{K}^2_{B-S}(T;x,\sqrt{b})\right) = F(x) - \sqrt{\frac{b}{\pi}}xf(x) + O(b).
$$

Using this result and the result of Lemma 2.1, we have:
(2.14)

$$\begin{aligned}
\mathrm{Var}(\hat{F}_1(x)) &= \mathrm{Var}\left(n^{-1}\sum_{i=1}^{n}\bar{K}_{B-S}(X_i;x,\sqrt{b})\right) = n^{-1}\mathrm{Var}\left(\bar{K}_{B-S}(T;x,\sqrt{b})\right) \\
&= n^{-1}\left\{E\left(\bar{K}^2_{B-S}(T;x,\sqrt{b})\right) - E^2\left(\bar{K}_{B-S}(T;x,\sqrt{b})\right)\right\} \\
&= n^{-1}\left\{F(x) - b^{\frac{1}{2}}\pi^{-\frac{1}{2}}xf(x) + O(b)\right\} \\
&- n^{-1}\left\{\left(F(x) + \frac{b}{2}\left(xf(x) + x^2f'(x)\right) + O(b^2)\right)^2\right\} \\
&= n^{-1}F(x)\left(1 - F(x)\right) - n^{-1}\left(b^{\frac{1}{2}}\pi^{-\frac{1}{2}}xf(x)\right) + O(n^{-1}b).
\end{aligned}$$

□

Using Lemma 2.1 and Lemma 2.2, we can derive an estimate of the mean integrated square error ($MISE$) for the B-S kernel estimator as follows.

(2.15)

$$\begin{aligned}
MISE_{B-S}\left(\hat{F}_1(x)\right) &= \int_0^\infty MSE\left(\hat{F}_1(x)\right)dx \\
&\approx n^{-1}\int_0^\infty F(x)\left(1 - F(x)\right)dx - n^{-1}b^{\frac{1}{2}}\pi^{-\frac{1}{2}}\int_0^\infty xf(x)dx \\
&+ \frac{b^2}{4}\int_0^\infty \left(xf(x) + x^2f'(x)\right)^2 dx.
\end{aligned}$$

This result gives rise to the following proposition.

Proposition 2.1. *The optimal smoothing parameter for the B-S kernel estimator based on minimizing the $MISE$ is*

$$\begin{aligned}
b^{MISE}_{B-S} &= \underbrace{argmin}_{b>0}\left(MISE_{B-S}\left(\hat{F}_1(x)\right)\right) \\
&\approx \left\{\int_0^\infty xf(x)dx\right\}^{\frac{2}{3}}\left\{\pi^{\frac{1}{2}}\int_0^\infty \left(xf(x) + x^2f'(x)\right)^2 dx\right\}^{-\frac{2}{3}} n^{-\frac{2}{3}}
\end{aligned} \tag{2.16}$$

This indicates that the optimal smoothing parameter is of order $O(n^{-2/3})$. By substituting b^{MISE}_{B-S} in (2.15), we have:

$$\begin{aligned}
MISE_{B-S}(\hat{F}_1(x)) &= n^{-1}\int_0^\infty F(x)\left(1 - F(x)\right)dx \\
&- \frac{3}{4}n^{-\frac{4}{3}}\pi^{-\frac{2}{3}}\left\{\int_0^\infty xf(x)dx\right\}^{\frac{4}{3}}\left\{\int_0^\infty \left(xf(x) + x^2f'(x)\right)^2 dx\right\}^{-\frac{1}{3}} \\
&+ O(n^{-\frac{5}{3}})
\end{aligned}$$

$$\Rightarrow MISE_{B-S}(\hat{F}_1(x)) = n^{-1} \int_0^\infty F(x)\,(1 - F(x))\,dx - O(n^{-\frac{4}{3}}).$$

2.2. The performance of the B-S kernel estimator at near boundary points

In order to delve in asymptotic properties of the B-S kernel estimator at the boundary points and compare the rate of its convergence at the boundary points and the interior points, we consider two specific cases for the design point x.

a) In the case where $x = 0$, the B-S kernel is zero, i.e. $\bar{K}_{B-S}\left(t; x, \sqrt{b}\right) = 0$ and, therefore, in this case $\hat{F}_1(0) = 0$ which is remarkable because the ordinary kernel estimator does not satisfy this property.

b) For the case where $x = cb$, where $0 < c < 1$, we have:

$$E\left(\hat{F}_1(x)\right) = F(x) + \frac{cb^2}{2} f(x) + O(b^3), \tag{2.17}$$

and

$$\mathrm{Var}(\hat{F}_1(x)) = n^{-1} F(x)\,(1 - F(x)) - n^{-1} b^{\frac{3}{2}} \pi^{-\frac{1}{2}} f(x) + O(n^{-1} b^2). \tag{2.18}$$

Therefore, we can compute the mean square error (MSE) for the B-S kernel estimator at the boundary points as follows:

$$MSE_{B-S}(\hat{F}_1(x)) \approx n^{-1} F(x)\,(1 - F(x)) - n^{-1} b^{\frac{3}{2}} \pi^{-\frac{1}{2}} f(x) + \frac{c^2 b^4}{4} f^2(x). \tag{2.19}$$

Comparing the bias and variance terms of the B-S kernel estimator at the near boundary and interior points (in equations (2.19) and (2.15), respectively) shows that the bias term is smaller at the near boundary points at the expense of increasing the variance term. Because at the near boundary points, the rate of convergence to zero of the negative portion of variance, which is the gain of smoothing technique over the empirical distribution function, is smaller than that of interior points.
Now it is easy to show that the optimal smoothing parameter which minimizes the MSE is

$$b_{B-S}^{MSE} = O(n^{-\frac{2}{5}}), \tag{2.20}$$

By substituting (2.20) in (2.19), we have:

$$MSE_{B-S}(\hat{F}_1(x)) = n^{-1} F(x)(1 - F(x)) + O(n^{-\frac{8}{5}}). \tag{2.21}$$

3. Asymmetric cumulative distribution function estimation using Weibull kernel

In the previous section, we introduced the B-S kernel estimator and demonstrated its asymptotic consistency. In this section, we will run a similar study and introduce another cumulative distribution function estimator based on the Weibull kernel, i.e. the Weibull kernel estimator.

3.1. Asymptotic properties of the Weibull kernel estimator

Consider the Weibull kernel given by

$$\bar{K}_{wbl}\left(t;\alpha,\beta\right)=\exp\left\{-\left(\frac{t}{\beta}\right)^{\alpha}\right\}\,,\qquad t\geq 0,\alpha>0,\beta>0. \tag{3.1}$$

Since $T\sim$ Weibull(α,β) then we have:

$$E\left(T^{k}\right)=\beta^{k}\Gamma\left(1+\frac{k}{\alpha}\right)\,,\qquad k=1,2,\ldots, \tag{3.2}$$

where $\Gamma\left(1+\frac{k}{\alpha}\right)=1-\frac{k\gamma}{\alpha}+\frac{k^2}{12\alpha^2}\left(\pi^2+6\gamma^2\right)+O\left(\alpha^3\right)$ and $\gamma=0.57721$ is the Euler's constant. Hirukawa and Sakudo [9] proposed an expansion for $\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma^{-2}\left(1+\frac{1}{\alpha}\right)$ as follows:

$$\Gamma\left(1+\frac{2}{\alpha}\right)\Gamma^{-2}\left(1+\frac{1}{\alpha}\right)=1+\frac{\pi^2}{6\alpha^2}+\frac{\gamma\pi^2-3\gamma^3}{2\alpha^3}+O\left(\alpha^{-4}\right). \tag{3.3}$$

Similarly, it is easy to show that

$$\Gamma\left(1+\frac{3}{\alpha}\right)\Gamma^{-3}\left(1+\frac{1}{\alpha}\right)=1+\frac{\pi^2}{2\alpha^2}+2\frac{\gamma\pi^2-3\gamma^3}{\alpha^3}+O\left(\alpha^{-4}\right). \tag{3.4}$$

Let $(\alpha,\beta)=\left(1/b,x/\Gamma\left(1+\alpha^{-1}\right)\right)$ where x and b denote the design point and the smoothing parameters, respectively. Our second asymmetric Kernel-type estimator, i.e. the Weibull kernel estimator, is defined as follows:

$$\hat{F}_2\left(x\right)=n^{-1}\sum_{i=1}^{n}\bar{K}_{wbl}\left(X_i;1/b,x/\Gamma\left(1+b\right)\right). \tag{3.5}$$

The Weibull kernel estimator $\hat{F}_2(x)$ is nonnegative and appropriate to estimate cumulative distribution functions with support on $[0,\infty)$. In what follows, we present the theoretical properties of $\hat{F}_2(x)$ and we will obtain an appropriate smoothing parameter for this estimator through minimizing the mean integrated square error. We will obtain approximate expressions for the bias and variance for $\hat{F}_2(x)$ in Lemma 3.1 and Lemma 3.2, respectively. In addition, we will discuss the convergence rate of the Weibull kernel estimator in the boundary points.

Lemma 3.1. *Suppose that Assumptions 1-3 hold, then the expectation value of $\hat{F}_2(x)$ can be obtained as:*

$$E\left(\hat{F}_2\left(x\right)\right) = F\left(x\right) + b^2\frac{\pi^2 x^2 f'\left(x\right)}{12} + O\left(b^3\right). \tag{3.6}$$

Proof of Lemma 3.1: The proof is analogous with the proof of Lemma 2.1. Using equation(2.3) and Taylor expansion, we have:

$$\begin{aligned} E\left(\hat{F}_2\left(x\right)\right) &= E\left(\bar{K}_{wbl}\left(T; 1/b, x/\Gamma\left(1+b\right)\right)\right) = E_k\left(F\left(T\right)\right) \\ &= F\left(x\right) + f\left(x\right)E\left(T-x\right) + \sum_{j=1}^{\infty}\frac{f^{(j)}\left(x\right)}{j!}E\left(T-x\right)^{j+1}, \end{aligned}$$

where T is a random variable with Weibull $\left(1/b, x/\Gamma\left(1+b\right)\right)$ probability density function. Using equations (3.3) and (3.4), we have:

$$\begin{aligned} &E\left(T-x\right) = 0, \\ &E\left(T-x\right)^2 = \frac{\left(xb\pi\right)^2}{6} + x^2b^3\left(\frac{\gamma\pi^2}{2} - 3\gamma^3\right) + O\left(b^4\right), \\ &E\left(T-x\right)^3 = \left(xb\right)^3\left(\frac{\gamma\pi^2}{2} + 3\gamma^3\right) + O\left(b^4\right). \end{aligned}$$

Now we can conclude that

$$\begin{aligned} E\left(\hat{F}_2(x)\right) &= F(x) + \frac{1}{2}f'(x)\left(\frac{(xb\pi^2)}{6} + x^2b^3\left(\frac{\gamma\pi^2}{2} - 3\gamma^3\right)\right) \\ &\quad + \frac{f^{(2)}(x)}{6}\left((xb)^3\left(\frac{\gamma\pi^2}{2} + 3\gamma^3\right)\right) + \cdots = F(x) + b^2\frac{\pi^2x^2f'(x)}{12} + O(b^3). \end{aligned}$$

□

Note that for the interior points, the bias of the Weibull kernel estimator is of order $O(b^2)$. However, by considering the smoothing parameter as a function of n in Remark 3.1, we will see that in the sense of convergence rate of bias, the Weibull kernel estimator is the same as the B-S kernel estimator. The following lemma provides an approximation for the variance of the Weibull kernel estimator.

Lemma 3.2. *Suppose that Assumptions 1-3 hold, then the variance of $\hat{F}_2(x)$ can be obtained as:*

$$\mathrm{Var}(\hat{F}_2(x)) = n^{-1}F(x)(1-F(x)) - n^{-1}b\ln(2)xf(x) + O(n^{-1}b^2). \tag{3.7}$$

where ln(.) *is the natural logarithm.*

Proof of Lemma 3.2: First note that

(3.8)
$$\begin{aligned}
E\left(\bar{K}_{wbl}^{2}\left(T;1/b,x/\Gamma\left(1+b\right)\right)\right) &= \int_0^\infty \bar{K}_{wbl}^{2}\left(t;1/b,x/\Gamma\left(1+b\right)\right) f\left(t\right) dt \\
&= \int_0^\infty F\left(t\right)\left\{2k_{wbl}\left(t;1/b,x/\Gamma\left(1+b\right)\right)\bar{K}_{wbl}\left(t;1/b,x/\Gamma\left(1+b\right)\right)\right\} dt \\
&= F\left(x\right)+f\left(x\right)E\left(Z-x\right)+\frac{1}{2}f'\left(x\right)E\left(Z-x\right)^2+\ldots,
\end{aligned}$$

where $Z \sim 2k_{wbl}\left(z;1/b,x/\Gamma(1+b)\right)\bar{K}_{wbl}\left(z;1/b,x/\Gamma(1+b)\right), z>0, b>0, x>0$.
It is easy to show that Z is random variable with Weibull($\alpha, \frac{\beta}{2^{\frac{1}{\alpha}}}$)density function.
Since $2^{-\frac{1}{\alpha}} = 1-\frac{\ln(2)}{\alpha}+\frac{\ln^2(2)}{2\alpha^2}+O(\alpha^{-3})$, we have:

(3.9)
$$\begin{aligned}
E(Z-x) &= -bx\ln(2)+\frac{(bx\ln(2))^2}{6}+O(b^3), \\
E(Z-x)^2 &= (xb)^2\left(\ln(2)^2+\frac{\pi^2}{6}\right)+O(b^3).
\end{aligned}$$

By substituting (3.9) in (3.8), we obtain

(3.10)
$$E(\bar{K}_{wbl}^{2}(T;1/b,x/\Gamma(1+b)))) = F(x)-bx\ln(2)f(x)+O(b^2).$$

Using (3.10) and Lemma 3.1, we can deduce that:

(3.11)
$$\begin{aligned}
\mathrm{Var}\left(\hat{F}_2\left(x\right)\right) &= \mathrm{Var}\left(n^{-1}\sum_{i=1}^{n}\bar{K}_{wbl}\left(X_i;1/b,x/\Gamma\left(1+b\right)\right)\right) \\
&= n^{-1}\mathrm{Var}\left(\bar{K}_{wbl}\left(T;1/b,x/\Gamma\left(1+b\right)\right)\right) \\
&= n^{-1}\left\{E\left(\bar{K}_{wbl}^{2}\left(T;1/b,x/\Gamma\left(1+b\right)\right)\right)-E^2\left(\bar{K}_{wbl}\left(T;1/b,x/\Gamma\left(1+b\right)\right)\right)\right\} \\
&= n^{-1}\left\{F\left(x\right)-bx\ln\left(2\right)f\left(x\right)+O\left(b^2\right)-\left(F\left(x\right)+b^2\frac{\pi^2x^2f'\left(x\right)}{12}+O\left(b^3\right)\right)^2\right\} \\
&= n^{-1}F\left(x\right)\left(1-F\left(x\right)\right)-n^{-1}b\ln\left(2\right)xf\left(x\right)+O\left(n^{-1}b^2\right).
\end{aligned}$$

□

Using Lemma (3.1) and Lemma (3.2), we can derive an estimate of the $MISE$ for the Weibull kernel estimator as follows:

(3.12)
$$\begin{aligned}
MISE_{wbl}(\hat{F}_2(x)) \approx n^{-1}\int_0^\infty F(x)(1-F(x))dx - n^{-1}b\ln(2)\int_0^\infty xf(x)dx \\
+ b^4\frac{\pi^4}{144}\int_0^\infty (x^2f'(x))^2dx.
\end{aligned}$$

Now we can select the optimal smoothing parameter based on minimizing the $MISE$.

Proposition 3.1. *The optimal smoothing parameter for the Weibull kernel estimator based on minimizing the MISE of $\hat{F}_2(x)$ in (3.12) is*

$$
\begin{aligned}
b_{wbl}^{MISE} &= \underbrace{argmin}_{b>0}(MISE_{wbl}(\hat{F}_2(x))) \\
&= \left\{36\ln(2)\int_0^\infty xf(x)dx\right\}^{\frac{1}{3}}\left\{\ \pi^4\int_0^\infty (x^2 f'(x))^2 dx\right\}^{-\frac{1}{3}} n^{-\frac{1}{3}}.
\end{aligned}
\tag{3.13}
$$

Note that the optimal smoothing parameter is of order $O(n^{-1/3})$. By substitutingb_{wbl}^{MISE} in (3.12), we have:

(3.14)

$$
\begin{aligned}
MISE_{wbl}(\hat{F}_2(x)) = n^{-1}\int_0^\infty F(x)(1-F(x))dx \\
&- 2.4764(n\pi)^{-\frac{4}{3}}(\ln(2))^{\frac{4}{3}}\left\{\int_0^\infty xf(x)dx\right\}^{\frac{4}{3}}\left\{\int_0^\infty (x^2 f'(x))^2 dx\right\}^{-\frac{1}{3}} \\
&+ O(n^{-\frac{5}{3}})
\end{aligned}
$$

$$
\Rightarrow MISE_{wbl}(\hat{F}_2(x)) = n^{-1}\int_0^\infty F(x)(1-F(x))dx - O(n^{-\frac{4}{3}}).
$$

Remark 3.1. From the two equations (2.16) and (3.13), the optimal smoothing parameter of the B-S kernel estimator and the Weibull kernel estimator are of order $O(n^{-2/3})$ and $O(n^{-1/3})$, respectively. Therefore, in terms of the rate of convergence to zero, we have $b_{B-S}^{MISE} \approx (b_{wbl}^{MISE})^2$. Thus from (2.4) and (3.6), we can conclude that for the interior points, the bias of the B-S kernel estimator has the same rate of convergence to zero as the bias of the Weibull kernel estimator.

3.2. The performance of the Weibull kernel estimator near boundary points

In this sub-section, we run a similar study like what we have done in Section 2.2 in order to investigate the asymptotic properties of the Weibull kernel estimator at the boundary points. This helps us to compare the rate of convergence at the boundary points and the interior points. We consider two specific cases for the design point x.

a) In the case where $x = 0$, we have $\bar{K}_{wbl}(T; 1/b, x/\Gamma(1+b)) = 0$, so in this case, unlike the ordinary kernel estimator, $\hat{F}_2(0) = 0$.

a) For the case where $x = cb$, where $0 < c < 1$, we have:

$$
E(\hat{F}_2(x)) = F(x) + \frac{c^2 b^4}{2} f(x) + O(b^5) \tag{3.15}
$$

and

$$\mathrm{Var}(\hat{F}_2(x)) = n^{-1}F(x)(1-F(x)) - n^{-1}cb^2\ln(2)f(x) + O(n^{-1}b^3). \tag{3.16}$$

So we can compute the MSE for the Weibull kernel estimator at the boundary points as follows:

$$MSE_{wbl}(\hat{F}_2(x)) \approx n^{-1}F(x)(1-F(x)) - n^{-1}cb^2\ln(2)f(x) + \frac{c^4b^8}{4}f^2(x). \tag{3.17}$$

Comparing the two equations (3.17) and (3.14) shows a trade-off between the bias and the variance terms for the Weibull kernel estimator. This is something like what we have seen for the B-S kernel estimator in Section 2. The bias term is again smaller at the near boundary points at the expense of increasing the variance term.
Now it is easy to show that the optimal smoothing parameter which minimizes the above-mentioned MSE is

$$b_{wbl}^{MSE} = O(n^{-\frac{1}{6}}), \tag{3.18}$$

By substituting (3.18) in (3.17), we have:

$$MSE_{wbl}(\hat{F}_2(x)) = n^{-1}F(x)(1-F(x)) + O(n^{-\frac{4}{3}}). \tag{3.19}$$

Remark 3.2. From the two equations (2.20) and (3.18), the optimal smoothing parameter of the B-S kernel estimator and the Weibull kernel estimator are of order $O(n^{-2/5})$ and $O(n^{-1/6})$, respectively. By substituting back these values into the corresponding bias terms of the two estimators, we can deduce that for the near boundary points, the bias of the B-S kernel estimator is of order $O(n^{-4/5})$ while the bias of the Weibull kernel estimator is of order $O(n^{-2/3})$.

4. Numerical study

In this section, we illustrate the performance of the proposed estimators (the B-S kernel estimator and the Weibull kernel estimator) through a simulation study. We compare our proposed estimators with the ordinary kernel method (O-K method), the B-K method and the Empirical distribution method. In both the O-K method and the B-K method, we use the Epanechnikov kernel. In order to select an appropriate bandwidth for the O-K and the B-K methods, we use the optimal bandwidth proposed by Altman and Leger [1] and Tenreiro [19], respectively.
We generated 1000 samples of size n=256 and 1024 from eight various distributions including, 1: Burr (1, 3, 1), 2: Gamma (0.6, 2), 3: Gamma (4, 2), 4: Generalized Pareto (0.4, 1, 0), 5: Halfnormal (0, 1), 6: Lognormal (0, 0.75),

7: Weibull (1.5, 1.5) and 8: Weibull (3, 2). In order to estimate the smoothing parameter for the B-S kernel estimator and the Weibull kernel estimator, we used Gamma density $f(x) = \frac{x^{\alpha-1}\exp\left(-\frac{x}{\beta}\right)}{\beta^{\alpha}\Gamma(\alpha)}$ as a referenced density in equations (2.16) and (3.13), respectively. The parameters (α, β) have been estimated by the method of maximum likelihood estimation.

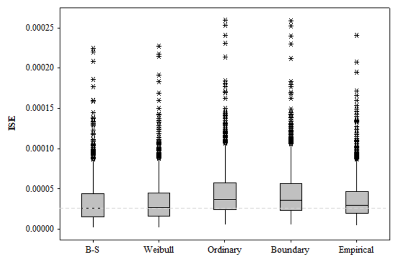
In order to evaluate the performance of our proposed estimators and compare their functionality with other existing methods, we considered the integrated squared error $ISE_i = \int_0^\infty \left(\hat{F}_i(x) - F(x)\right)^2 dx$ as an error metrics, where $\hat{F}_i(x), i = 1, 2, \ldots 5$ stands for the B-S kernel estimator, the Weibull kernel estimator, the O-K method, the B-K method and the Empirical distribution method, respectively. In our setting, we approximated the integral with summation.

Table 1 shows the mean and standard deviation of the ISE for the eight distributions and the two sample sizes over one thousand repetitions. In all cases, the mean and standard deviation of the ISE decreased as the sample size increased. The simulation results show that based on the ISE, regardless of the sample size, our proposed estimators perform better than the other three methods. The only exception is distribution 5: Halfnormal (0, 1) with a sample size of 256 for which the B-K method has a smaller mean of ISE than that of the Weibull kernel estimator. However, even for this case, when the sample size is increased to 1024, both the B-S kernel estimator and the Weibull kernel estimator have a better performance. The comparison between the B-S kernel estimator and the Weibull kernel estimator indicates the superiority of the B-S kernel estimator. This is true, surprisingly, even in estimating two distributions Weibull (1.5, 1.5) and Weibull (3, 2).

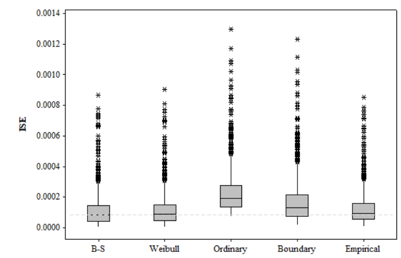
Value($\times 10^{-4}$)		B-S		Weibull		Ordinary		Boundary		Empirical	
N	Example	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
256	1	1.33	1.17	1.37	1.20	1.63	1.27	1.59	1.26	1.53	1.12
	2	4.24	4.17	4.50	4.28	6.37	5.14	5.46	5.14	4.96	4.24
	3	1.37	0.94	1.43	0.98	1.64	1.08	1.64	1.08	1.55	0.91
	4	4.44	4.59	4.70	4.74	6.52	5.46	5.33	5.47	5.14	4.64
	5	2.90	2.77	2.99	2.85	3.60	3.02	2.96	3.03	3.33	2.82
	6	5.89	5.90	6.10	6.05	8.05	6.82	7.59	6.82	6.17	5.51
	7	3.29	3.02	3.37	3.12	4.20	3.34	3.65	3.29	3.74	3.03
	8	2.62	2.40	2.68	2.46	3.09	2.58	3.03	2.57	2.98	2.39
1024	1	0.34	0.28	0.35	0.29	0.46	0.33	0.45	0.32	0.38	0.28
	2	1.18	1.14	1.22	1.17	2.38	1.55	1.78	1.56	1.30	1.14
	3	0.28	0.31	0.28	0.29	0.74	0.55	0.73	0.54	0.28	0.30
	4	1.18	1.21	1.22	1.24	2.18	1.51	1.53	1.54	1.28	1.19
	5	0.74	0.72	0.76	0.73	1.14	0.82	0.81	0.82	0.81	0.71
	6	0.63	0.56	0.64	0.58	1.04	0.70	0.91	0.69	0.67	0.52
	7	0.91	0.81	0.93	0.81	1.30	0.90	1.12	0.89	1.00	0.80
	8	0.69	0.62	0.70	0.62	0.87	0.68	0.86	0.68	0.75	0.62

Table 1: The mean and standard deviation of the ISE in estimating eight distributions via five methods (see the text for explanation) for n=256 and 1024.

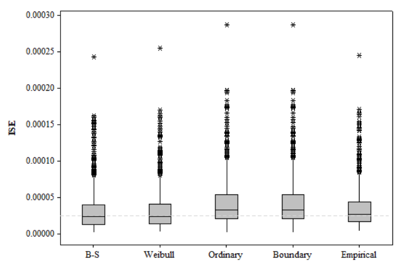
In order to provide a better comparison between the aforementioned methods, we have presented the boxplots of the ISE for the case n=1024 in Figure 1. In this figure, we consider eight boxplots for eight divers' distributions. In the boxplots, the vertical axis shows the ISE and the horizontal axis contains the methods. The dotted line in each of the boxplots shows the lowest median of the ISEs. The overall superiority of the B-S kernel estimator in all cases is obvious. The overall performance of the Weibull kernel estimator is better than the B-K method and the Empirical distribution method. The O-K method shows the worst performance as is expected.



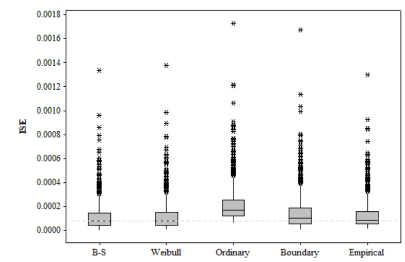
(a) Burr(1,3,1)



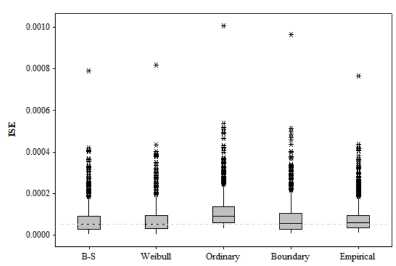
(b) Gamma(0.6,2)



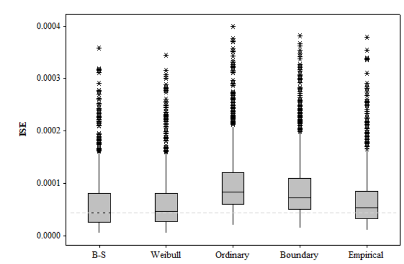
(c) Gamma(4,2)



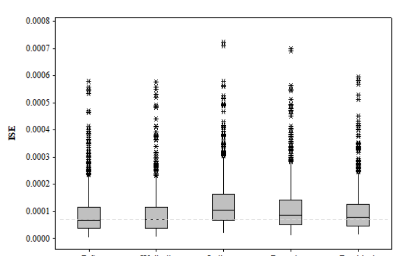
(d) Generalized Pareto(0.4,1,0)



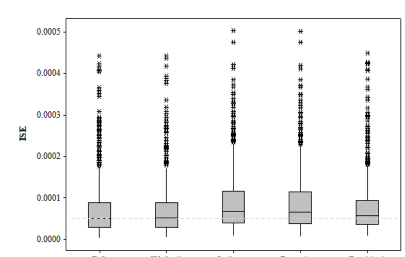
(e) Halfnormal(0,1)



(f) Lognormal(0,0.75)



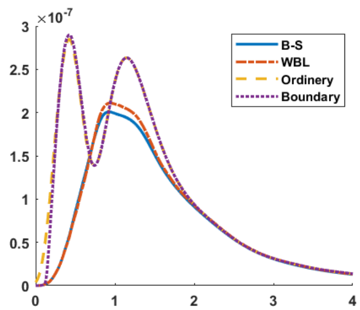
(g) Weibull(1.5,1.5)



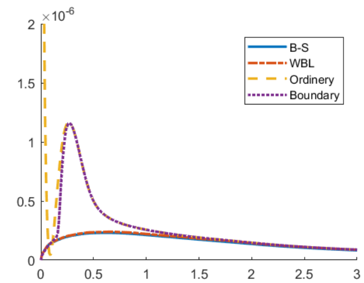
(h) Weibull(3,2)

Figure 1: The boxplots of the ISE in estimating eight distribution functions via five methods in 1000 repetitions (n = 1024) (see text for further explanation).

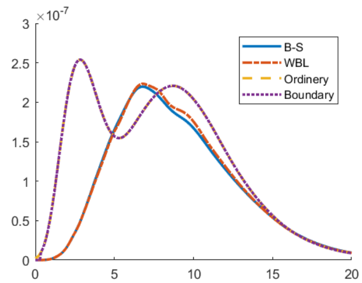
In Figure 2, we provide the results on the mean squared error (MSE) at various points of the support of the considered distributions in 1000 repetitions for the sample size (1024). This helps one to see the performance of the compared methods depending on the point where the distribution function is estimated. To increase the visibility and better compare other kernel-type estimators, we have ignored the Empirical distribution in this Figure. The poor performance of the O-K method at near boundary region is obvious. At the points far from the boundary, the O-K method and the B-K method almost match. Although the amount of MSE is dependent on the design point and the distribution which we want to estimate, the overall performance of the two proposed estimators are better than both the O-K and the B-K methods.Note that, the shape of asymmetric kernels changes with the design point and for the points, those are far enough from the boundary, they become symmetric, and finally all the methods almost match in Figure 2.
Figures **??** to 4 illustrate 30 estimates in blue along with the true distribution in red for the eight different distributions ($n = 256$) via five methods. The density function of these distributions is plotted as well in the top left corner of each image. The boundary bias of the O-K method is obvious. The B-K method remedies this drawback but not completely. In particular, a careful inspection of the figures, especially Gamma (4,2) and Weibull (3,2), for near boundary points, shows that the B-K method suffers from over-estimation. It seems that this problem depends on the shape of the distribution which we wish to estimate. Another striking point is that the Empirical distribution could not provide smooth estimates. In general, the performance of our proposed estimators is satisfying.



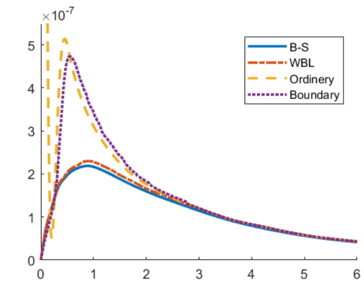
(a) Burr(1,3,1)



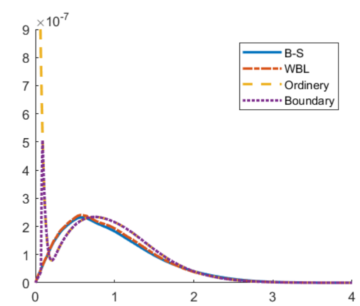
(b) Gamma(0.6,2)



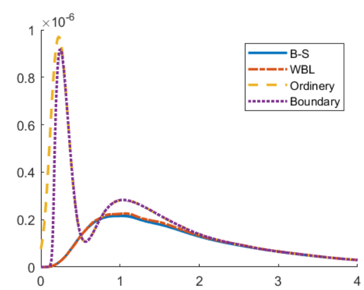
(c) Gamma(4,2)



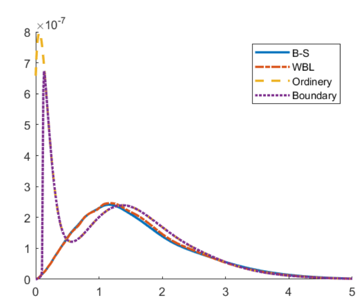
(d) Generalized Pareto(0.4,1,0)



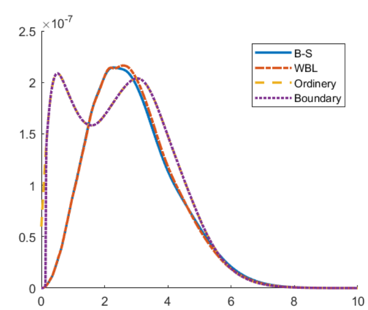
(e) Halfnormal(0,1)



(f) Lognormal(0,0.75)

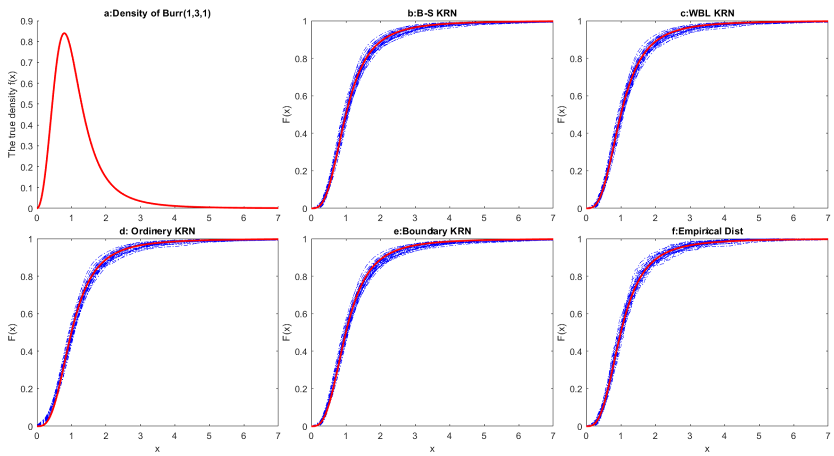


(g) Weibull(1.5,1.5)

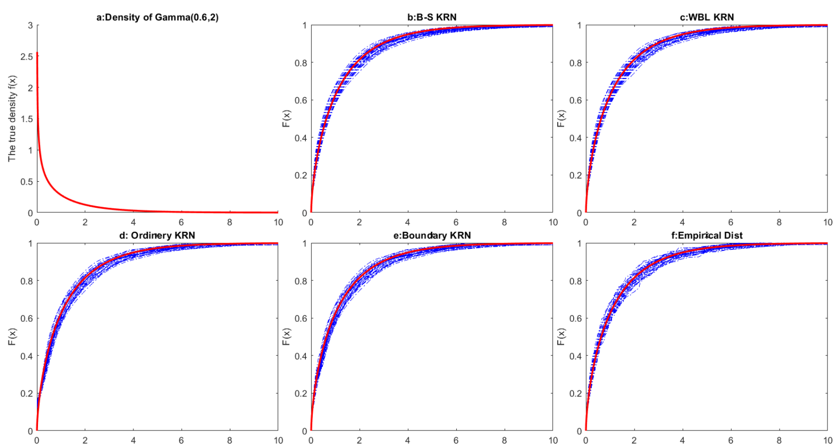


(h) Weibull(3,2)

Figure 2: The Plot of the MSE in estimating eight distribution functions via five methods in 1000 repetitions (n = 1024) (see the text for further explanation).

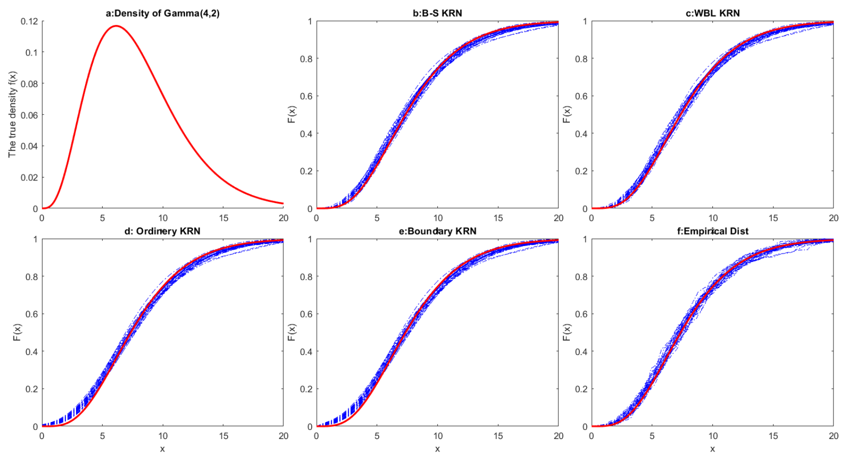


(a) Burr(1,3,1)

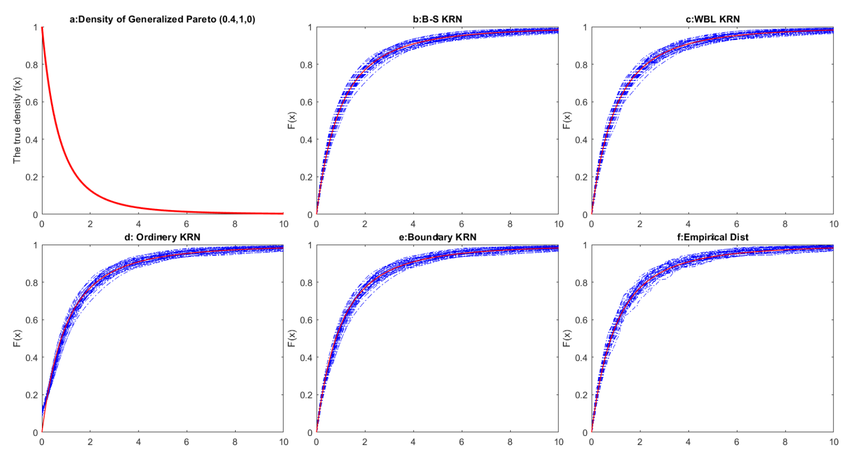


(b) Gamma(0.6,2)

Figure 3: Plots of 30 estimates (in blue) of Burr (1,3,1) and Gamma (0.6,2) via five methods: (b) B-S kernel estimator (top mid), (c) Weibull kernel estimator (top right), (d) O-K method (Bottom left), (e) B-K method (Bottom mid) and (f) Empirical distribution (bottom right). The true distribution is shown in red and sample size n = 256. The top left (a) shows the density function of each distribution.

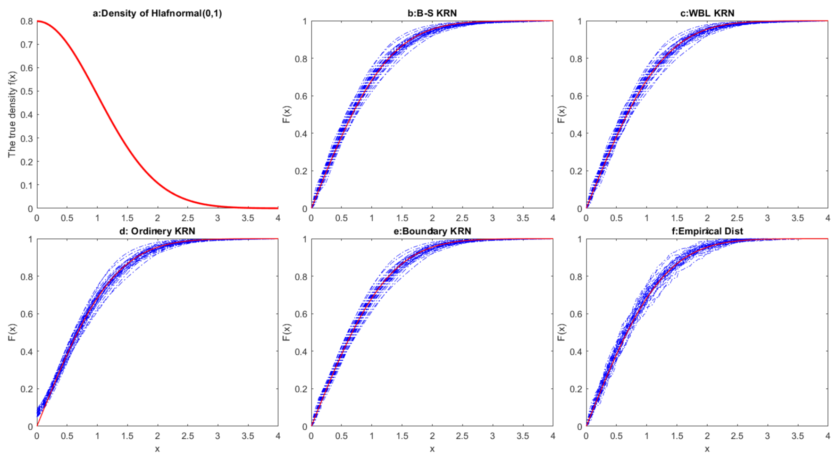


(a) Gamma(4,2)

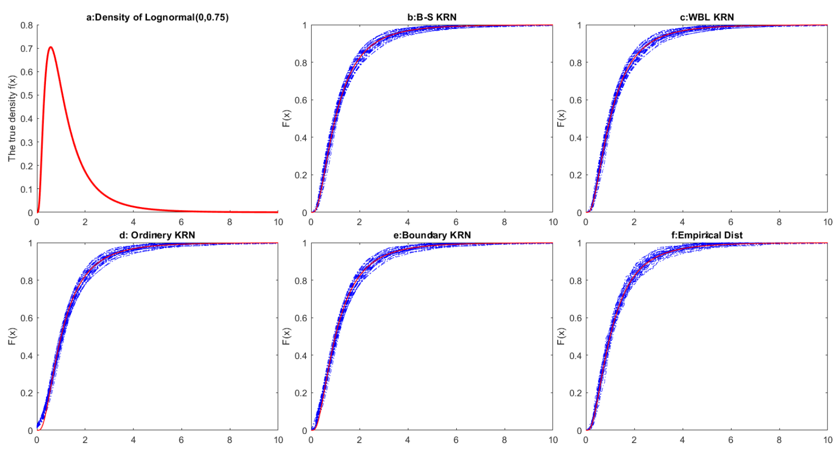


(b) Generalized Pareto(0.4,1,0)

Figure 4: Plots of 30 estimates (in blue) of Gamma (4,2) and Generalized Pareto (0.4, 1, 0) via five methods: (b) B-S kernel estimator (top mid), (c) Weibull kernel estimator (top right), (d) O-K method (Bottom left), (e) B-K method (Bottom mid) and (f) Empirical distribution (bottom right). The true distribution is shown in red and sample size n = 256. The top left (a) shows the density function of each distribution.

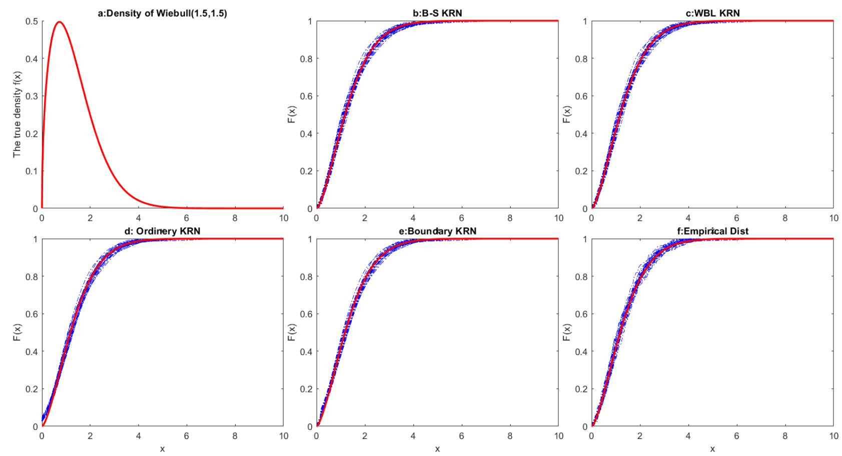


(a) Halfnormal(0,1)

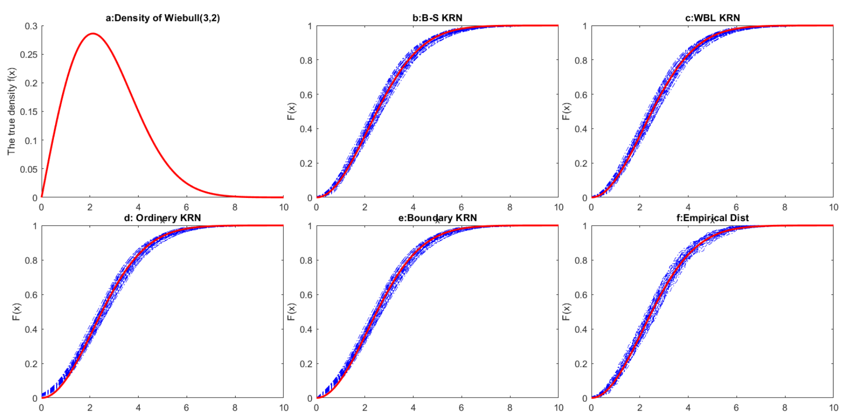


(b) Lognormal(0,0.75)

Figure 5: Plots of 30 estimates (in blue) of Halfnormal (0,1) and Lognormal(0,0.75) via five methods: (b) B-S kernel estimator (top mid), (c) Weibull kernel estimator (top right), (d) O-K method (Bottom left), (e) B-K method (Bottom mid) and (f) Empirical distribution (bottom right). The true distribution is shown in red and sample size n = 256. The top left (a) shows the density function of each distribution.



(a) Weibull(1.5,1.5)



(b) Weibull(3,2)

Figure 6: Plots of 30 estimates (in blue) of Weibull (1.5,1.5) and Weibull (3,2) via five methods: (b) B-S kernel estimator (top mid), (c) Weibull kernel estimator (top right), (d) O-K method (Bottom left), (e) B-K method (Bottom mid) and (f) Empirical distribution (bottom right). The true distribution is shown in red and sample size n = 256. The top left (a) shows the density function of each distribution.

5. Illustration with a real data set

In this Section, we apply our two proposed estimators to a real dataset. The data are the time distance between marriage to the first childbirth. This dataset is a result of a field research performed by Choromzadeh et al. [4] to study the factors that influence childbirth behavioral patterns of women aged 15-49 in a sample of size n=1106 in Ahwaz, Iran. Due to the traditions, many families tend to have children immediately after marriage. Therefore, the data has a natural peak in 9-18 months after marriage. There are rare cases of childbirth in 1-8 months, which are probably the result of pregnancy before marriage. Figure 7 shows the histogram of this dataset. On the other hand, due to the changes in socioeconomic and cultural statuses, there are few families that give birth to their first child in a considerable time after their marriage. Also, there are some families whose delayed first birth is due to sterility problems. Thus, a long tail with sparse data is another considerable feature in the distribution of this dataset. Figure 8 illustrates 5 estimates of the distribution of this data via five methods.

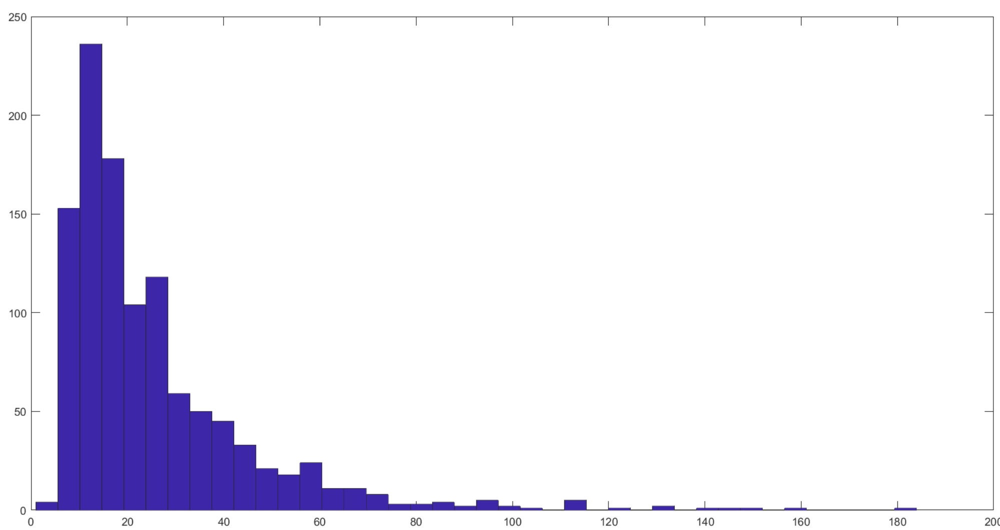


Figure 7: Histogram of the months after marriage before the first child-birth

The methods of choosing the smoothing parameter for various estimators are described in Section 4. Figure 8-a shows that estimates mainly differ at the near origin. In order to provide a better insight, we separately illustrate the estimates in the first 9 months in Figure 8-b. In comparison with the Empirical distribution, the estimates created by the O-K method and the B-K method are similar. It seems they rise too early. In the simulation study, we have seen that these two estimators suffer from over-estimating for near boundary points, especially for those distributions that have the same shape as in Figure 7. The B-S kernel estimate and the Weibull kernel estimate are very close, and the more consistent they are with the Empirical distribution and for this dataset, the more realistic they seem to be.

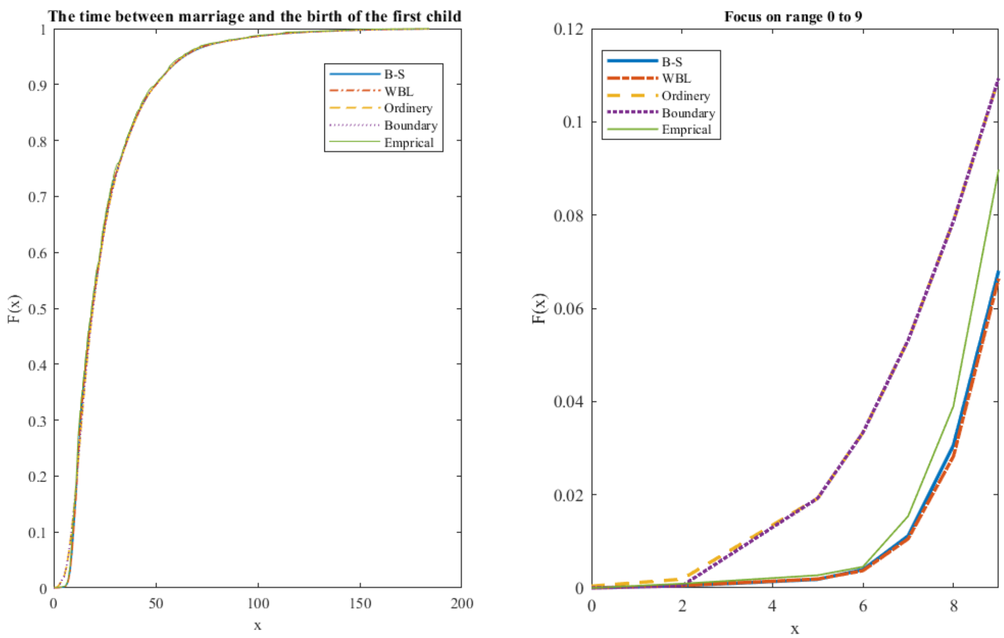


Figure 8: Five estimates of the distribution of first childbirth via five methods: B-S kernel estimator (solid-blue), Weibull kernel estimator (dashed-red), O-K method (dashed-yellow), B-K method (dotted-purple) and Empirical distribution (solid-green)

6. Conclusion and discussion

This paper is devoted to proposing some appropriate estimators for the cumulative distribution functions with non-negative support. To achieve this goal, we proposed a general asymmetric Kernel-type estimator and introduced two asymmetric estimators for the cumulative distribution function. We demonstrated the asymptotic consistency of our proposed estimators and we showed that they are free from boundary effects as well. Comparing our estimators based on the rate of convergence at the boundary points, we found that the B-S kernel estimator was better than the Weibull kernel estimator. In our setting, we estimated the bandwidths of the two estimators based on minimizing the MISE. In order to evaluate the performance of our estimators and compare them with other existing methods, we conducted a numerical study. The results of the numerical study show that both the B-S kernel and the Weibull kernel estimators are superior to the B-K method proposed by Tenreiro [19]. In the numerical study, the B-S kernel estimator achieved the best results and outperformed the Weibull kernel estimator. This is consistent with the good asymptotic properties of the B-S kernel estimator. In this research, we used the B-S kernel and the Weibull kernel as the asymmetric kernels in our general estimator. As a path for future research, one can try other existing asymmetric kernels. Another area for future research can be the estimation of those cumulative distributions with a finite interval support, for instance [a, b]. In addition, application of this type of cumulative distribution estimator in several other fields such as the survival analysis and the copula methods is an interesting topic for future research.

REFERENCES

[1] Altman, N. and Leger,C. (1995). Bandwidth selection for kernel distribution function estimation, *Journal of Statistical Planning and Inference*, **46**, 2, 195–214.

[2] Chen, S. (1999). Beta kernel estimators for density functions, *Computational Statistics and Data Analysis*, **31**, 2, 131–145.

[3] Chen, S. (2000). Probability density function estimation using Gamma kernels, *Annals of the Institute of Statistical Mathematics*, **52**, 3, 471–480.

[4] Choromzadeh, R.; Akhound, M.R. and Rasekh, A. (2015). Factors affecting women's birth intervals: the case of women referred to health centers in Ahwaz, *Jornal of Hayat*, **20**, 4, 35–50.

[5] Falk,M. (1983). Relative efficiency and deficiency of kernel type estimators of smoothdistribution functions, *Statistica Neerlandica*, **37**, 2, 73–83.

[6] Gasser,T. and Muller,H. (1979). Kernels estimation of regresion functions, *Lecture Notes in Mathematics*, **757**, 23–68.

[7] Gasser,T.; Muller,H. and Mammitzsch, V. (1985). Kernels for nonparametriccurve estimation, *Journal of the Royal Statistical Society. Series B*, **47**, 2, 238–252.

[8] Hirukawa, M. and Sakudo, M. (2014). Nonnegative bias reduction methods for density estimation using asymmetric kernels, *Computational Statisticsand and Data Analysis*, **75**, 112–123.

[9] Hirukawa, M. and Sakudo, M. (2015). Family of the generalised gamma kernels: a generator of asymmetric kernels for nonnegative data, *Journal of Nonparametric Statistics*, **27**, 1, 41–63.

[10] Jin, X. and Kawczak, J. (2003). Birnbaum–Saunders and lognormal kernel estimators for modelling durations in high frequency financial data, *Annals of Economics and Finance*, **4**, 1, 103–124.

[11] Johnson, N.; Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distribution*, Volume 2, 2nd Ed. Wiley, New York

[12] Muller, H. (1991). Smooth optimum kernel estimators near endpoints, *Biometrika*, **78**, 3, 521–530.

[13] Nadarajah, S. and Kotz, S. (2003). Skewed distribution generated by normal kernel, *Statistics and Probability Letters*, **65**, 3, 269–277.

[14] Nadaraya, E.A. (1964). Some new estimate for distribtion function, *Theory of Probability and Its Applications*, **9**, 3, 497–500.

[15] Reiss, R.D. (1981). Nonparametric estimation of smooth distribution functions, *Scandinavian Journal of Statistics*, **8**, 2, 116–119.

[16] Rice, J. (1984). Boundary modification for kernel regression, *Communications in Statistics - Theory and Methods*, **13**, 7, 893–900.

[17] Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels, *Journal of Nonparametric Statistics*, **16**, 1, 217–226.

[18] Singh, R.S.; Gasser, T. and Prasad, B. (1983). Nonparametric estimates of distribution functions, *Communications in Statistics - Theory and Methods* , **12**, 18, 2095–2108.

[19] Tenreiro, C. (2013). Boundary kernels for distribution function estimation, *REVSTAT–Statistical Journal*, **11**, 2, 169–190.

[20] Vilca, F. and Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distribtions, *Communications in Statistics - Theory and Methods*, **35**, 2, 229–244.

[21] Watson, G. and Leadbetter, M. (1964). Hazard analysis II *Sankhyā: The Indian Journal of Statistics, Series A*, **21**, 1, 229–244.

[22] Zhang, S. (2010). A note on the performance of the gamma kernel estimators at the boundary, *Statistics and Probablity Letters*, **80**, 7, 548-557.