AIMS Mathematics, 8(1): 1784–1799.
DOI: 10.3934/math.2023092
Received: 11 July 2022
Revised: 19 September 2022
Accepted: 28 September 2022
Published: 24 October 2022

Mathematics

http://www.aimspress.com/journal/Math

Research article

Improved the bias in kernel quantile function estimation

Abdallah Sayah* **and Nassima Almi**

Department of Mathematics and Informatic, Mohamed Khider University, Biskra, Algeria

* **Correspondence:** Email: sayahabdel@yahoo.fr.

Abstract: In this paper, a new estimator for kernel quantile estimation is given to reduce the bias. The asymptotic properties of the proposed estimator was established and it turned out that the bias has been reduced to the fourth power of the bandwidth, while the bias of the estimators considered has the second power of the bandwidth, while the variance remains at the same order. Futhermore, we calculate the optimal bandwidth which minimizes the asymptotic mean squared error. A simulation study and a real data example are carried out to illustrate the performance of the proposed estimator and compared with other existing approaches mentioned.

Keywords: bias reduction; quantile function; kernel estimation; order statistics; bandwidth
Mathematics Subject Classification: 62G05, 62G20

1. Introduction

Percentiles and more generally, quantiles are commonly used when a parametric form for the underlying distribution is not available. The quantile function approach it has been used in exploratory data analysis, applied statistics, reliability and survival analysis by several authors among them Falk [1], Parzen [2] and Azzalini [3] and in the books of Galambos [4] and David [5]. Without loss of generality, for $X_1, ..., X_n$ be independent and identically distributed with an unknown density $f(.)$ and absolutely continuous distribution function $F(.)$, while $X_{(1)}, ..., X_{(n)}$ denote the corresponding order statistics. The quantile function Q is defined to be the left-continuous inverse of F as follows:

$$Q(p) = \inf\{x : F(x) \geq p\} = F^{-1}(p),\ 0 < p < 1.$$

Many nonparametric estimators of the quantile function have been proposed and studied extensively.

Traditionally, the estimator of the distribution function is the empirical function $F_n(x)$, which is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{]-\infty,x]}(X_i),$$

where the indicator function $1_{]-\infty,x]}(X_i) = 1$ if $X_i \le x$ and 0 otherwise.

Theoretical properties of $F_n(x)$ as an estimator of the unknown true distribution function $F(x)$ have been investigated by several authors, see for example Yamato [6], Reiss [7].

A basic estimator of $Q(p)$ is the empirical quantile or the sample quantiles which is given by

$$Q_n(p) = \inf\{x : F_n(x) \ge p\} = X_{([np])},$$

where $[np]$ denotes the integer part of np.

Because of the variability of individual order statistics, the sample quantiles suffer from lack of efficiency. In order to reduce this variability, different approaches of estimating sample quantiles through weighted order statistics have been proposed. We begin by defining the kernel estimators of f and F at x. They are respectively

$$\tilde{f}_n(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right),$$

and

$$\tilde{F}_n(x) = \int_{-\infty}^{x} \tilde{f}_n(t)\,dt = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

where $h = h_n$ is the smoothing parameter (or the bandwidth) since it controls the amount of smoothness in the estimator, and satisfy $h := h_n \to 0$ as $n \to \infty$, $k(.)$ is a kernel function which is a predetermined density function symmetric about 0 and the function K is defined from a kernel k as

$$K(x) = \int_{-\infty}^{x} k(t)\,dt.$$

Let introduce the notation

$$\mu_j = \int_{-\infty}^{\infty} t^j k(t)\,dt, \; j = 1, 2, 3, 4.$$

The corresponding estimator of the quantile function $Q = F^{-1}$ is then defined by

$$Q_n^c(p) = \inf\{x : \tilde{F}_n(x) \ge p\}, \; 0 < p < 1.$$

Nadaraya [8] showed under some assumptions for k, f and h, $Q_n^c(p)$ has an asymptotic standard normal distribution. The almost sure consistency was obtained by Yamato [6] . Ralescu and Sun [9] obtained the necessary and conditions for the asymptotic normality. Azzalini [3] obtains the asymptotic mean squared error of $Q_n^c(p)$:

$$AMSE(Q_n^c(p)) = \frac{h^4}{4}\left(\frac{Q^{(2)}(p)}{(Q^{(1)}(p))^2}\right)^2 \mu_2^2 + \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 - \frac{h}{n} Q^{(1)}(p)\,\psi(k),$$

where $Q^{(1)}$and $Q^{(2)}$ are the first and the second derivative of Q respectively and

$$\psi(k) = 2\int_{-\infty}^{\infty} tk(t)\,K(t)\,dt.$$

It can be seen that the optimal bandwidth for minimizing $AMSE(Q_n^c(p))$ has the form

$$h_{opt}^c = \left(\frac{\left(Q^{(1)}(p)\right)^5 \psi(k)}{n\left(Q^{(2)}(p)\right)^2 \mu_2^2}\right)^{\frac{1}{3}}.$$

Parzen [2] proposed the classical kernel quantile estimator as below:

$$\check{Q}_n(p) = \sum_{i=1}^{n}\left[\int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h}k\left(\frac{p-x}{h}\right)dx\right]X_{(i)}.$$

In practice, Yang [10] proposed the following approximation to $\check{Q}_n(p)$ which is defined by:

$$\check{Q}_n^a(p) = \frac{1}{nh}\sum_{i=1}^{n} X_{(i)}k\left(\frac{p-\frac{i}{n}}{h}\right).$$

Under suitable conditions on F, Falk [1] proposed the following kernel type quantile estimator:

$$\tilde{Q}_n(p) = \frac{1}{h}\int_0^1 Q_n(x)\,k\left(\frac{p-x}{h}\right)dx,$$

this kernel-type quantile estimate can then be approximated by $\check{Q}_n(p)$.

Yang [10] provided the asymptotic normality property and the mean squared consistency of $\check{Q}_n(p)$ and proved that $\check{Q}_n(p)$ and $\check{Q}_n^a(p)$ are asymptotically equivalent in terms of mean squared errors. Falk [1] showed that the asymptotic performance of $\check{Q}_n(p)$ is better than that of the empirical sample quantile $Q_n(p)$ in terms of relative deficiency for appropriately chosen kernels and sufficiently smooth distribution functions. Building on Falk [1], Sheather and Marron [11] gave the asymptotic mean squared error of $\check{Q}_n(p)$. If the second derivative of Q is continuous in a neighborhood of p and if f is not symmetric or f is symmetric but $p \neq \frac{1}{2}$ then the asymptotic mean squared error of $\check{Q}_n(p)$ is

$$AMSE\left(\check{Q}_n(p)\right) = \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 + \frac{h^4}{4}\left(Q^{(2)}(p)\right)^2\mu_2^2 - \frac{h}{n}\left(Q^{(1)}(p)\right)^2\psi(k).$$

The optimal bandwidth for $AMSE\left(\check{Q}_n(p)\right)$ is

$$\check{h}_{opt} = \left(\frac{\left(Q^{(1)}(p)\right)^2 \psi(k)}{n\left(Q^{(2)}(p)\right)^2 \mu_2^2}\right)^{\frac{1}{3}}.$$

When F is symmetric and $p = \frac{1}{2}$ then the asymptotic mean squared error of $\check{Q}_n(p)$ has the form

$$AMSE\left(\check{Q}_n(p)\right) = \frac{1}{n}\left(Q^{(1)}\left(\frac{1}{2}\right)\right)^2\left[0.25 - 0.5h\psi(k) + \left(\frac{1}{nh}\right)\rho(k)\right],$$

where $\rho(k) = \int_{-\infty}^{\infty} k^2(x)\,dx$.

In order to reduce the order of the bias in the case of regression estimation, Choi and Hall. [12] proposed a method, called (skewing method) whose basic idea is to move the center of the adjustment local to the left and right of the point at which one wishes to estimate the curve and take a convex combination of three estimators, the two oppositely shifted estimators and the symmetric estimator and showed that this idea resulted in the reduction of the order of bias to the fourth power of the bandwidth at the expense of a slight increase in variance by a constant factor. Cheng et al. [13] applied the same idea in locally parametric density estimation. And for the same purpose, Kim et al. [14] applied the skewing method to the classical kernel density estimator instead of rather complicated locally parametric density estimators. This method is based on the consideration of the tangent lines of the classical kernel estimator placed at points a little to the left and to the right of the point x where we want to estimate the density, and have shown that by evaluating the two lines tangents and the classical kernel estimator in x and taking a convex combination of these three values can reduce the bias order to the fourth power of the bandwidth.

The aim of this work is to propose a new estimator of the quantile function to reduce the order of the bias and improve its asymptotic properties when p in the interior region, based on convex combination technique of skewed estimators. This idea is inspired by the works of Choi et al. [12], Cheng et al. [13] and Kim et al. [14]. In Section 2, we derive the explicit form of our estimator and its asymptotic properties are illustrated. In Section 3, numerical simulation studies have been carried out to see numerical performance of the proposed estimator, and compared with the existing approaches mentioned previously, real data example is also given. Finally conclusion stated in last section.

2. Main results

In order to reduce the bias in the case of estimating the quantile function, we propose a new estimator $\hat{Q}_n(p)$ of $Q(p)$ using a bias reduction technique in the kernel quantile function estimator based on the technique of convex combination of three estimates $\tilde{Q}_n$, $\bar{Q}_1$ and $\bar{Q}_2$ where for $j = 1, 2$, $\bar{Q}_j$ estimator represents the value at p, of the tangent line which meets $\tilde{Q}_n$ at $p + \varphi h$ and $p - \varphi h$ respectively. Our proposed estimator can be expressed as a linear combination of the kernel quantile function estimator and the kernel quantile density function estimator.

Our proposed estimator at p is

$$\hat{Q}_n(p) = \frac{\tilde{Q}_n(p) + \alpha\left(\bar{Q}_1(p) + \bar{Q}_2(p)\right)}{1 + 2\alpha}, \tag{2.1}$$

where

$$\bar{Q}_1(p) := \tilde{Q}_n(p + \varphi h) - \varphi h \tilde{Q}_n^{(1)}(p + \varphi h),$$
$$\bar{Q}_2(p) := \tilde{Q}_n(p - \varphi h) + \varphi h \tilde{Q}_n^{(1)}(p - \varphi h),$$

$\alpha > 0$ and $\varphi = \varphi(\alpha)$ are constants to be determined.

Our estimator can also be written as

$$\begin{aligned}\hat{Q}_n(p) &= \frac{1}{(1+2\alpha)h}\int_0^1 Q_n(x)\left[k\left(\frac{p-x}{h}\right)+\alpha\left(k\left(\frac{p-x+\varphi h}{h}\right)-\varphi k^{(1)}\left(\frac{p-x+\varphi h}{h}\right)\right)\right.\\ &\quad \left.+\alpha\left(k\left(\frac{p-x-\varphi h}{h}\right)+\varphi k^{(1)}\left(\frac{p-x-\varphi h}{h}\right)\right)\right]dx.\end{aligned}$$

The following theorem shows that the bias of $\hat{Q}_n(p)$ is of order $O(h^4)$, while that of mentioned estimators is $O(h^2)$, also, it shows that the variance of the proposed estimator remains at the same order as the kernel quantile function estimators and it gives the expressions for the bias and the variance of the proposed estimator.

Theorem 2.1. *Assume that Q has four bounded, continuous derivatives in a neighborhood of p and the kernel k is probability density function symmetric about zero, then if $0 < h \to 0$, $nh^2 \to \infty$, we have for all fixed $p \in]\varphi h, 1-\varphi h[$*

$$Bias\left(\hat{Q}_n(p)\right) = \frac{1}{24}h^4 Q^{(4)}(p)\left(\mu_4 - \frac{3(1+6\alpha)}{2\alpha}\mu_2^2\right) + o\left(h^4\right), \tag{2.2}$$

and

$$Var\left(\hat{Q}_n(p)\right) = \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 - \frac{h}{n}\left(Q^{(1)}(p)\right)^2 U(\alpha) + o\left(\frac{h}{n}\right) + o(1), \tag{2.3}$$

where

$$\begin{aligned}U(\alpha) &= \frac{1}{(2\alpha+1)^2}\left[2\alpha^2\left(\psi(k) + 2\varphi^2\int_{-\infty}^{\infty} tk^{(1)}(t)\,k(t)\,dt\right) + 4\alpha\left(\int_{-\infty}^{\infty}(t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)K(t-\varphi)\,dt\right)\right.\\ &+ 4\alpha^2\left(\int_{-\infty}^{\infty}(t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)\left(K(t-2\varphi)+\varphi\int_{-\infty}^{t-2\varphi}k^{(1)}(t)\,dt\right)dt\right) + \psi(k)\\ &+ \left.4\alpha\left(\int_{-\infty}^{\infty} tk(t)\left(K(t+\varphi)-\varphi\int_{-\infty}^{t+\varphi}k^{(1)}(t)\,dt\right)dt\right)\right].\end{aligned}$$

Proof. For $p \in]\varphi h, 1-\varphi h[$ we have

$$\begin{aligned}E\left(\hat{Q}_n(p)\right) &= \frac{1}{1+2\alpha}\left[\alpha\left(\frac{1}{h}\int_0^1 Q(x)\,k\left(\frac{p+\varphi h-x}{h}\right)dx - \frac{\varphi h}{h^2}\int_0^1 Q(x)\,k^{(1)}\left(\frac{p+\varphi h-x}{h}\right)dx\right.\right.\\ &+ \left.\frac{1}{h}\int_0^1 Q(x)\,k\left(\frac{p-\varphi h-x}{h}\right)dx + \frac{\varphi h}{h^2}\int_0^1 Q(x)\,k^{(1)}\left(\frac{p-\varphi h-x}{h}\right)dx\right)\\ &+ \left.\frac{1}{h}\int_0^1 Q(x)\,k\left(\frac{p-x}{h}\right)dx\right]\\ &= \frac{1}{1+2\alpha}\left[\alpha\left(\int_{-\infty}^{\infty}k(t)\,Q(p-h(t-\varphi))\,dt - \varphi h\int_{-\infty}^{\infty}Q^{(1)}(p-h(t-\varphi))\,k(t)\,dt\right)\right.\\ &+ \alpha\left(\int_{-\infty}^{\infty}k(t)\,Q(p-h(t+\varphi))\,dt + \varphi h\int_{-\infty}^{\infty}Q^{(1)}(p-h(t+\varphi))\,k(t)\,dt\right)\\ &+ \left.\int_{-\infty}^{\infty}k(t)\,Q(p-ht)\,dt\right],\end{aligned}$$

using a Taylor expansion of order 4 and of order 3 on the function $Q(p)$ and $Q^{(1)}(p)$ respectively, we have

$$\begin{aligned} E\left(\hat{Q}_n(p)\right) &= Q(p) + \frac{h^2}{2(1+2\alpha)} Q^{(2)}(p)\left((2\alpha+1)\mu_2 - 2\alpha\varphi^2\right) \\ &+ \frac{h^4}{24(1+2\alpha)} Q^{(4)}(p)\left((2\alpha+1)\mu_4 - 12\alpha\varphi^2\mu_2 - 6\alpha\varphi^4\right) + o\left(h^4\right), \end{aligned}$$

therefore, $E\left(\hat{Q}_n(p)\right)$ can be computed by letting the term in h^2 vanish if and only if

$$(2\alpha+1)\mu_2 - 2\alpha\varphi^2 = 0,$$

the equation imply

$$\varphi(\alpha) = \left(\frac{2\alpha+1}{2\alpha}\mu_2\right)^{\frac{1}{2}}.$$

Finally, if we substitute $\varphi(\alpha)$ for $E\left(\hat{Q}_n(p)\right)$, we obtain the bias expression as

$$Bias\left(\hat{Q}_n(p)\right) = \frac{1}{24} h^4 Q^{(4)}(p)\left(\mu_4 - \frac{3(1+6\alpha)}{2\alpha}\mu_2^2\right) + o\left(h^4\right).$$

For the variance term. First, note that

$$\begin{aligned} Var\left(\hat{Q}_n(p)\right) &= \frac{1}{(2\alpha+1)^2}\left[\alpha^2 Var\left(\bar{Q}_1(p)\right) + Var\left(\tilde{Q}_n(p)\right) + \alpha^2 Var\left(\bar{Q}_2(p)\right)\right. \\ &+ 2\alpha^2 Cov\left(\bar{Q}_1(p), \bar{Q}_2(p)\right) \\ &+ \left. 2\alpha Cov\left(\bar{Q}_1(p), \tilde{Q}_n(p)\right) + 2\alpha Cov\left(\bar{Q}_2(p), \tilde{Q}_n(p)\right)\right]. \end{aligned}$$

Now, we will compute each term on the right hand side of $Var\left(\hat{Q}_n(p)\right)$ except $Var\left(\tilde{Q}_n(p)\right)$ which is given by

$$\begin{aligned} Var\left(\tilde{Q}_n(p)\right) &= \frac{1}{n}\left(Q^{(1)}(p)\right)^2\left(-p^2 + 2\int_{-\infty}^{+\infty}(p-ht)\,k(t)\,K(t)\,dt\right) + o\left(\frac{h}{n}\right) \\ &= \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 - \frac{h}{n}\left(Q^{(1)}(p)\right)^2\psi(k) + o\left(\frac{h}{n}\right). \end{aligned}$$

For the first term, we have

$$\begin{aligned} Var\left(\bar{Q}_1(p)\right) &= Var\left(\tilde{Q}_n(p+\varphi h)\right) + (\varphi h)^2 Var\left(\tilde{Q}_n^{(1)}(p+\varphi h)\right) \\ &- 2\varphi h Cov\left(\tilde{Q}_n(p+\varphi h), \tilde{Q}_n^{(1)}(p+\varphi h)\right) \\ &= I_1 + (\varphi h)^2 I_2 - 2\varphi h I_3, \end{aligned}$$

where

$$
\begin{aligned}
I_1 &= E\left[\left(\frac{1}{h}\int_0^1 Q_n(x)\,k\left(\frac{p+\varphi h-x}{h}\right)dx-\frac{1}{h}\int_0^1 Q(x)\,k\left(\frac{p+\varphi h-x}{h}\right)dx\right)^2\right]\\
&= E\left[\left(\int_{\frac{p+\varphi h-1}{h}}^{\frac{p+\varphi h}{h}}\left(Q_n(p+\varphi h-ht)-Q(p+\varphi h-ht)\right)k(t)\,dt\right)^2\right]\\
&= E\left[\left(\int_{-\infty}^{\infty} k(t)\Big((p-h(t-\varphi))-\bar{F}_n(p-h(t-\varphi))\Big)Q^{(1)}(p-h(t-\varphi))\,dt\right)^2\right]\\
&= \frac{\left(Q^{(1)}(p)\right)^2}{n}\int_0^1\left(\int_{-\infty}^{\infty} k(t)\Big((p-h(t-\varphi))-1_{(0,p-h(t-\varphi))}(y)\Big)dt\right)^2dy+o\left(\frac{h}{n}\right)\\
&= \frac{\left(Q^{(1)}(p)\right)^2}{n}\int_0^1\left(-(p+h\varphi)^2+\left(\int_{-\infty}^{\infty}k(t)\,1_{(0,p-h(t-\varphi))}(y)dt\right)^2\right)dy+o\left(\frac{h}{n}\right),
\end{aligned}
$$

where $\bar{F}_n$ is the empirical distribution fonction according to n independent, uniformly on $[0,1]$ distributed random variables and

$$
\begin{aligned}
\left(\int_{-\infty}^{\infty}k(t)\,1_{(0,p-h(t-\varphi))}(y)dt\right)^2 &= \int_0^1\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+h\varphi}{h}}k(t)\,dt\right)^2dy\\
&= \frac{2}{h}\int_0^1 y\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+h\varphi}{h}}k(t)\,dt\right)k\left(\tfrac{p-y+h\varphi}{h}\right)dy\\
&= 2\int_{\frac{p-1+h\varphi}{h}}^{\frac{p+h\varphi}{h}}(p-h(t-\varphi))\,k(t)\left(\int_{\frac{p-1+h\varphi}{h}}^{t}k(t)\,dt\right)dt\\
&= \left(p+h\varphi-2h\int_{-\infty}^{\infty}tk(t)\,K(t)\,dt\right),
\end{aligned}
$$

therefore,

$$
\begin{aligned}
I_1 &= Var\left(\tilde{Q}_n(p+\varphi h)\right)\\
&= \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2-\frac{h}{n}\left(Q^{(1)}(p)\right)^2\psi(k)+o\left(\frac{h}{n}\right)+o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
I_2 &= Var\left(\tilde{Q}_n^{(1)}(p+h\varphi)\right)\\
&= E\left(\frac{1}{h^2}\int_0^1 Q_n(x)\,k^{(1)}\left(\frac{p+\varphi h-x}{h}\right)dx-\frac{1}{h^2}\int_0^1 Q(x)\,k^{(1)}\left(\frac{p+\varphi h-x}{h}\right)dx\right)^2\\
&= \frac{1}{h^2}E\left(\int_{\frac{p+\varphi h-1}{h}}^{\frac{p+\varphi h}{h}}\left(Q_n(p+\varphi h-ht)-Q(p+\varphi h-ht)\right)k^{(1)}(t)\,dt\right)^2\\
&= \frac{1}{nh^2}\int_0^1\left(\int_{-\infty}^{\infty}k^{(1)}(t)\Big((p-h(t-\varphi))-1_{(0,p-h(t-\varphi))}(y)\Big)Q^{(1)}(p-h(t-\varphi))\,dt\right)^2dy\\
&= \frac{1}{nh^2}\left(Q^{(1)}(p)\right)^2\int_0^1\left(\left(\int_{-\infty}^{\infty}k^{(1)}(t)\,1_{(0,p-h(t-\varphi))}(y)dt\right)^2\right)dy+o\left(\frac{1}{nh}\right)+o(1),
\end{aligned}
$$

by similar computations we obtain

$$\int_0^1 \left(\left(\int_{-\infty}^{\infty} k^{(1)}(t)\, 1_{(0,p-h(t-\varphi))}(y) dt \right)^2 \right) dy = 2 \int_{-\infty}^{+\infty} (p - h(t-\varphi))\, k^{(1)}(t) \left(\int_{-\infty}^{t} k^{(1)}(t)\, dt \right) dt,$$

then

$$I_2 = -\frac{2}{nh} \left(Q^{(1)}(p) \right)^2 \int_{-\infty}^{+\infty} t k^{(1)}(t)\, k(t)\, dt + o\left(\frac{1}{nh} \right) + o(1).$$

Also, it can be shown that

$$\begin{aligned}
I_3 &= Cov\left(\tilde{Q}_n(p + \varphi h), \tilde{Q}_n^{(1)}(p + \varphi h) \right) \\
&= E\left[\left(\frac{1}{h} \int_0^1 Q_n(x)\, k\left(\frac{p + h\varphi - x}{h} \right) dx - \frac{1}{h} \int_0^1 Q(x)\, k\left(\frac{p + h\varphi - x}{h} \right) dx \right) \right. \\
&\times \left. \left(\frac{1}{h^2} \int_0^1 Q_n(x)\, k^{(1)}\left(\frac{p + h\varphi - x}{h} \right) dx - \frac{1}{h^2} \int_0^1 Q(x)\, k^{(1)}\left(\frac{p + h\varphi - x}{h} \right) dx \right) \right] \\
&= \frac{\left(Q^{(1)}(p) \right)^2}{nh} \int_0^1 \left(\left(\int_{-\infty}^{\infty} k(t)\,(p - h(t-\varphi)) - 1_{(0,p-h(t-\varphi))}(y) \right) dt \right. \\
&\times \left. \left(\int_{-\infty}^{\infty} k^{(1)}(t) \left((p - h(t-\varphi)) - 1_{(0,p-h(t-\varphi))}(y) \right) dt \right) dy \right) + o\left(\frac{1}{n} \right) \\
&= \frac{\left(Q^{(1)}(p) \right)^2}{nh} \int_0^1 \left(-h(p + h\varphi) + \int_{-\infty}^{\infty} k(t)\, 1_{(0,p-h(t-\varphi))}(y) dt \int_{-\infty}^{\infty} k^{(1)}(t)\, 1_{(0,p-h(t-\varphi))}(y) dt \right) dy + o\left(\frac{1}{n} \right) \\
&= \frac{1}{n} \left(Q^{(1)}(p) \right)^2 \left(\frac{1}{2} - p - h\varphi \right) + o\left(\frac{1}{n} \right) \\
&= o\left(\frac{1}{n} \right) + o(1).
\end{aligned}$$

At last, we combine each terms of $Var\left(\bar{Q}_1(p)\right)$ we have

$$Var\left(\bar{Q}_1(p)\right) = \frac{p(1-p)}{n} \left(Q^{(1)}(p) \right)^2 - \frac{h}{n} \left(Q^{(1)}(p) \right)^2 \left(\psi(k) + 2\varphi^2 \left(\int_{-\infty}^{\infty} t k^{(1)}(t)\, k(t) \right) dt \right) + o\left(\frac{h}{n} \right) + o(1).$$

Similar computations give

$$Var\left(\bar{Q}_2(p)\right) = \frac{p(1-p)}{n} \left(Q^{(1)}(p) \right)^2 - \frac{h}{n} \left(Q^{(1)}(p) \right)^2 \left(\psi(k) + 2\varphi^2 \left(\int_{-\infty}^{\infty} t k^{(1)}(t)\, k(t) \right) dt \right) + o\left(\frac{h}{n} \right) + o(1).$$

Now we will calculate the fourth term on the right hand side of $Var(\hat{Q}_n)$.

We have

$$
\begin{aligned}
Cov\left(\bar{Q}_1(p), \bar{Q}_2(p)\right) &= E\left[\frac{1}{h}\int_0^1 \left(Q_n(x) - Q(x)\right)\left(k\left(\frac{p+h\varphi-x}{h}\right) - \varphi k^{(1)}\left(\frac{p+h\varphi-x}{h}\right)\right)dx\right. \\
&\times \left.\frac{1}{h}\int_0^1 \left(Q_n(x) - Q(x)\right)\left(k\left(\frac{p-h\varphi-x}{h}\right) + \varphi k^{(1)}\left(\frac{p-h\varphi-x}{h}\right)\right)dx\right] \\
&= \frac{-\left(Q^{(1)}(p)\right)^2}{n}p^2 + \frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_0^1\left(\int_{-\infty}^{\infty} k(t)\, 1_{(0,p-h(t-\varphi))}(y)\, dt\right.\right. \\
&\times \left.\int_{-\infty}^{\infty}\left(k(t) + \varphi k^{(1)}(t)\right) 1_{(0,p-h(t+\varphi))}(y)\, dt\right)dy\Bigg) \\
&- \left.\varphi\int_0^1\left(\int_{-\infty}^{\infty} k^{(1)}(t)\, 1_{(0,p-h(t-\varphi))}(y)\, dt \int_{-\infty}^{\infty}\left(k(t) + \varphi k^{(1)}(t)\right) 1_{(0,p-h(t+\varphi))}(y)\, dt\right)dy\right] \\
&+ o\left(\frac{h}{n}\right) \\
&= \frac{-\left(Q^{(1)}(p)\right)^2}{n}p^2 + \frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_0^1\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+h\varphi}{h}} k(t)\, dt \int_{\frac{p-1-h\varphi}{h}}^{\frac{p-y-h\varphi}{h}}\left(k(t) + \varphi k'(t)\right)dt\right)dy\right. \\
&- \left.\varphi\int_0^1\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+\varphi h}{h}} k^{(1)}(t)\, dt \int_{\frac{p-1-h\varphi}{h}}^{\frac{p-y-h\varphi}{h}}\left(k(t) + \varphi k^{(1)}(t)\right)dt\right)dy\right] + o\left(\frac{h}{n}\right).
\end{aligned}
$$

By integration by part, we find

$$
\begin{aligned}
Cov\left(\bar{Q}_1(p), \bar{Q}_2(p)\right) &= \frac{-\left(Q^{(1)}(p)\right)^2}{n}p^2 + \frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_{\frac{p+h\varphi-1}{h}}^{\frac{p+h\varphi}{h}} \left(p - h(t-\varphi)\right)k(t)\left(\int_{\frac{p-1-h\varphi}{h}}^{t-2\varphi} k(t)\, dt\right)dt\right. \\
&+ \int_{\frac{p-1-h\varphi}{h}}^{\frac{p-h\varphi}{h}} \left(p - h(t+\varphi)\right)k(t)\left(\int_{\frac{p-1+h\varphi}{h}}^{t+2\varphi} k(t)\, dt\right)dt + \varphi\left(\int_{\frac{p+h\varphi-1}{h}}^{\frac{p+h\varphi}{h}} \left(p - h(t-\varphi)\right)k(t)\right. \\
&\times \left.\left(\int_{\frac{p-1-h\varphi}{h}}^{t-2\varphi} k^{(1)}(t)\, dt\right)dt + \int_{\frac{p-h\varphi-1}{h}}^{\frac{p-h\varphi}{h}} \left(p - h(t+\varphi)\right)k^{(1)}(t)\left(\int_{\frac{p-1-h\varphi}{h}}^{t+2\varphi} k(t)\, dt\right)dt\right) \\
&- \varphi\left(\int_{\frac{p+h\varphi-1}{h}}^{\frac{p+h\varphi}{h}} \left(p - h(t-\varphi)\right)k'(t)\left(\int_{\frac{p-1-h\varphi}{h}}^{t-2\varphi} k(t)\, dt\right)dt + \int_{\frac{p-h\varphi-1}{h}}^{\frac{p-h\varphi}{h}} \left(p - h(t+\varphi)\right)k(t)\right. \\
&\times \left.\left(\int_{\frac{p-1+h\varphi}{h}}^{t+2\varphi} k^{(1)}(t)\, dt\right)dt\right) - \varphi^2\left(\int_{\frac{p+h\varphi-1}{h}}^{\frac{p+h\varphi}{h}} \left(p - h(t-\varphi)\right)k^{(1)}(t)\left(\int_{\frac{p-1-h\varphi}{h}}^{t-2\varphi} k^{(1)}(t)\, dt\right)dt\right. \\
&+ \left.\left.\int_{\frac{p-h\varphi-1}{h}}^{\frac{p-h\varphi}{h}} \left(p - h(t+\varphi)\right)k^{(1)}(t)\left(\int_{\frac{p-1+h\varphi}{h}}^{t+2\varphi} k^{(1)}(t)\, dt\right)dt\right)\right] + o\left(\frac{h}{n}\right), \\
&= \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 - h\frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_{-\infty}^{\infty} (t+\varphi)\left(k(t) + \varphi k^{(1)}(t)\right)\right. \\
&\times \left(K(t+2\varphi) - \varphi\left(\int_{-\infty}^{t+2\varphi} k^{(1)}(t)\, dt\right)\right)dt + \int_{-\infty}^{\infty} (t-\varphi)\left(k(t) - \varphi k^{(1)}(t)\right) \\
&\times \left.\left(K(t-2\varphi) + \varphi\left(\int_{-\infty}^{t-2\varphi} k^{(1)}(t)\, dt\right)\right)dt\right] + +o\left(\frac{h}{n}\right).
\end{aligned}
$$

$$Cov\left(\bar{Q}_1(p), \bar{Q}_2(p)\right) = \frac{p(1-p)}{n}\left(Q^{(1)}(p)\right)^2 - 2h\frac{\left(Q^{(1)}(p)\right)^2}{n}\int_{-\infty}^{\infty}\left((t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)\right.$$
$$\times \left.\left(K(t-2\varphi)+\varphi\left(\int_{-\infty}^{t-2\varphi}k^{(1)}(t)\,dt\right)\right)\right)dt + +o\left(\frac{h}{n}\right).$$

Since

$$\int_{-\infty}^{\infty}K(t+2\varphi)(t+\varphi)\left(k(t)+\varphi k^{(1)}(t)\right)dt = \int_{-\infty}^{\infty}K(t-2\varphi)(t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)dt,$$
$$-\int_{-\infty}^{\infty}(t+\varphi)\left(k(t)+\varphi k^{(1)}(t)\right)\left(\int_{-\infty}^{t+2\varphi}k^{(1)}(t)\,dt\right)dt = \int_{-\infty}^{\infty}(t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)dt\left(\int_{-\infty}^{t-2\varphi}k^{(1)}(t)\right)dt.$$

For the fifth term on the right hand side of $Var(\hat{Q}_n)$, we have

$$Cov\left(\bar{Q}_1(p), \tilde{Q}_n(p)\right) = E\left[\frac{1}{h}\int_0^1(Q_n(x)-Q(x))\left(k\left(\frac{p+h\varphi-x}{h}\right)+\varphi k^{(1)}\left(\frac{p+h\varphi-x}{h}\right)\right)dx\right.$$
$$\times \left.\frac{1}{h}\int_0^1(Q_n(x)-Q(x))\,k\left(\frac{p-x}{h}\right)dx\right].$$

$$Cov\left(\bar{Q}_1(p), \tilde{Q}_n(p)\right) = \frac{-\left(Q^{(1)}(p)\right)^2}{n}p^2 + \frac{\left(Q^{(1)}(p)\right)^2}{n}\int_0^1\left[\int_{-\infty}^{\infty}k(t)\,1_{(0,p-h(t-\varphi))}(y)dt\int_{-\infty}^{\infty}k(t)\,1_{(0,p-ht)}(y)dt\right.$$
$$+ \left.\int_{-\infty}^{\infty}k^{(1)}(t)\,1_{(0,p-h(t-\varphi))}(y)dt\int_{-\infty}^{\infty}k(t)\,1_{(0,p-ht)}(y)dt\right]dy + o\left(\frac{h}{n}\right)$$
$$= \frac{-\left(Q^{(1)}(p)\right)^2}{n}p^2 + \frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_0^1\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+h\varphi}{h}}k(t)\,dt\int_{\frac{p-1}{h}}^{\frac{p-y}{h}}k(t)\,dt\right)dy\right.$$
$$+ \left.\int_0^1\left(\int_{\frac{p-1+h\varphi}{h}}^{\frac{p-y+h\varphi}{h}}k^{(1)}(t)\,dt\int_{\frac{p-1}{h}}^{\frac{p-y}{h}}k(t)\,dt\right)dy\right] + o\left(\frac{h}{n}\right).$$

By integration by part, we find

$$Cov\left(\bar{Q}_1(p), \tilde{Q}_n(p)\right) = \frac{p(1-p)\left(Q^{(1)}(p)\right)^2}{n}$$
$$- h\frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_{-\infty}^{\infty}(t-\varphi)\left(k(t)-\varphi k^{(1)}(t)\right)K(t-\varphi)\,dt\right.$$
$$+ \left.\int_{-\infty}^{\infty}tk(t)\left(K(t+\varphi)-\varphi\left(\int_{-\infty}^{t+\varphi}k^{(1)}(t)\,dt\right)dt\right)\right] + o\left(\frac{h}{n}\right).$$

Similar computations give the sixth term on the right hand side of $Var(\hat{Q}_n)$,

$$Cov\left(\bar{Q}_2(p), \tilde{Q}_n(p)\right) = \frac{p(1-p)\left(Q^{(1)}(p)\right)^2}{n} - h\frac{\left(Q^{(1)}(p)\right)^2}{n}\left[\int_{-\infty}^{\infty}(t+\varphi)\left(k(t)+\varphi k^{(1)}(t)\right)K(t+\varphi)\,dt\right.$$
$$+ \left.\int_{-\infty}^{\infty}tk(t)\left(K(t-\varphi)+\varphi\left(\int_{-\infty}^{t-\varphi}k^{(1)}(t)\,dt\right)dt\right)\right] + o\left(\frac{h}{n}\right).$$

By adding up all these terms, we have the desired result for the variance.

Corollary 2.1. *The asymptotically optimal bandwidth, in the sense of minimising the asymptotic mean squared error, is given by*

$$h^* = \left(\frac{72\left(Q^{(1)}(p)\right)^2 U(\alpha)}{n\left(Q^{(4)}(p)\left(\mu_4 - \frac{3(1+6\alpha)}{2\alpha}\mu_2^2\right)\right)^2} \right)^{\frac{1}{7}}.$$

The associated AMSE is

$$AMSE_h\left(\hat{Q}_n(p)\right) = \frac{p(1-p)}{n}\left(Q'(p)\right)^2 - \frac{7}{8n}\left(\frac{72\left(\left(Q'(p)\right)^2 U(\alpha)\right)^8}{n\left(Q^{(4)}(p)\left(\mu_4 - \frac{3(1+6\alpha)}{2\alpha}\mu_2^2\right)\right)^2} \right)^{\frac{1}{7}}.$$

□

3. Simulation

There are different ways to compare quantile estimators from finite samples described in the literature and implemented in statistics packages. Particularly, we use two important criterion are the integrated bias (IB) and integrated variance (IV). These can be combined into a single measure, the mean integrated squared error (MISE). Since the variance is not much influenced because the variance of the proposed estimator remains at the same order as existing methods as a results the reduction of the MISE of Q is mainly due to bias.

To measure the performance of our estimator and compared their efficiencies to the existing quantile estimators previously mentioned from different distributions, Weibull distribution with parameters 1 and 1.5, Burr with parameters 1 and 3, Exponential distribution with parameter 1 and the reduced centered normal distribution. Without loss of generality, we randomly select 1000 independent samples of size $n = 100$ and $n = 200$, we have chosen fixed intervals $I_p =]\varphi h, 1 - \varphi h[$ for all distributions used where $\varphi h < 0.5$. Since the properties of kernel estimators do not depend much on which particular kernel is used, if the Epanechnikov kernel is used then the condition $0 < \alpha < 1$ leads that $\varphi > 0.54772$. In this part we choose 6 values of φ summarized in the Table 1. Depending on the values of φ and since the largest value of $\varphi = 2$ and the condition $\varphi h < 0.5$ implies that $h < 0.25$. The value of h used in the simulations was $h = 0.2$, which gave reliably good results.

Table 1. Choice of φ values.

φ	0.75	1.00	1.25	1.50	1.75	2.00
I_p	]0.15, 0.85[	]0.20, 0.80[	]0.25, 0.75[	]0.30, 0.70[	]0.35, 0.65[	]0.40, 0.60[

The numerical results summarized in the following tables: Tables 2–9. Results are re-scaled by the factor 0.01.

Table 2. Values of MISE (IV/ISB) of Weibull distribution for $n = 100$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	9. 188 6 (1.0571/8. 101 5)	9. 162 7 (1.0557/8. 105 0)	9. 166 4 (1.0621/8. 104 3)	8. 761 1 (0.9266/7. 834 5)
1.00	5. 921 9 (0.9142/5.0127)	5.5269 (0.9105/4. 976 6)	5. 584 9 (0.9135/4.9714)	5.4797 (0.8199/4.6597)
1.25	3.7104 (0.7618/2.9494)	3.6610 (0.7588/2.9238)	3.6987 (0.7717/2.9279)	2.3817 (0.7518/1.6286)
1.50	2.2490 (0.6144/1. 644 9)	2. 154 4 (0.5911/1. 627 0)	2.1922 (0.5926/1. 622 1)	1.9710 (0.5816/1. 389 4)
1.75	1.2753 (0.4355/0.8588)	1. 274 1 (0.4247/0.8245)	1. 273 9 (0.4247/0.8231)	1. 062 9 (0.4230/0.6399)
2.00	0.667 2 (0.2815/0.3867)	0.6438 (0.2769/0.3665)	0.6461 (0.2769/0.36811)	0.5010 (0.2754/0.2194)

Table 3. Values of MISE (IV/ISB) of Weibull distribution for $n = 200$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	9.457(0.5658/8. 8974)	8.8694 (0.5531/8. 314 0)	8.8782(0.5537/8. 314 3)	8.5064(0.5426/7. 9643)
1.00	5.82 39 (1. 5803/4.2136)	4. 713 4 (0.4787/4. 172 5)	4. 711 3 (0.5084/4.1724)	4.228 6 (0.4726/3.7641)
1.25	3.7507(0.4103/3. 185 4)	3. 5017 (0.40145/3. 099 7)	3. 5104 (0.4017/3.1087)	3. 262 5 (0.4013/2. 833 5)
1.50	2.2268 (0.354/1. 858 7)	2. 137 6 (0.3 217/1. 758 4)	2. 137 6 (0.32174/1. 782 5)	1. 8590 (0.32080/1. 537 2)
1.75	1.1534(0.2638/0.88164)	1.1090 (0.2506/0.86924)	1.1102(0.2510/0.8689)	0.8708(0.2500/0.63579)
2.00	0.5962 (0.22085/0.3 976)	0.5123 (0.1 1640/0.3 971)	0.5115(0.1200/0.3893)	0.3 411 (0.1121 4/0.2 292 5)

Table 4. Values of MISE (IV/ISB) of Burr distribution for $n = 100$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	10.7340 (1. 170 2/8.8291)	9. 346 2 (1. 677 1/8.2543)	9.3545(1.6809/8. 254 1)	8.6827(1. 248 8/7.4339)
1.00	6.7461(1.4783/5.3466)	6.6901(1.4657/5.2311)	6. 167 8 (1. 463 8/5. 233 0)	5.6428(1. 381 7/4.46109)
1.25	4.1050(1.2393/3. 387 5)	4.0979(1.2318/3.3 14)	4. 098 1 (1.2301/3. 317 7)	3.4330(0.8876/2. 545 4)
1.50	3. 040 9 (1.0361/2.1048)	2.7004(0.9843/2.0503)	2.7 060 (0.9868/2. 057 5)	2. 001 6 (0.6573/1.3442)
1.75	1.9206(0.7264/1.2093)	1.9425 (0.7126/1.1926)	1.9207 (0.7225/1.2205)	1.14989(0.4871/0.6627)
2.00	1. 2135 (0.4427/0.7696)	1.1235(0.4424/0.7569)	1.1457(0.4420/0.6500)	0.70912(0.4629/0.2442)

Table 5. Values of MISE (IV/ISB) of Burr distribution for $n = 200$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	9. 983 8 (1.7280/8.2548)	9.27341 (1. 169 2/7.7088)	9.27284 (1. 159 9/7. 704 3)	8.1973 (1. 047 2/7. 026 2)
1.00	7.1027 (0.8779/5.8809)	7.0958 (0.8749/5.7216)	7.0951 (0.8689/5. 716 3)	5.8918 (0.8337/5.0580)
1.25	4. 5014 (0.6911/3.8128)	4.7006 (0.6723/3.7823)	4.6939 (0.6703/3. 777 3)	3.8024 (0.5920/3.2103)
1.50	3.1951 (0.5241 + 2.4522)	3.0650 (0.5198 + 2.3599)	3.0586 (0.5194/2. 448 5)	2.4480 (0.5060/1.9420)
1.75	1. 897 9 (0.3757 + 1.4422)	1.9070 (0.3742/1.5089)	1.9023 (0.3741/1. 507 0)	1.4713 (0.3743/1.0890)
2.00	0.8724 (0.2316/0.6404) :	0.794 5 (0.2303/0.5642)	0.770 1 (0.2303/0.5398)	0.6900 (0.2305/0.4594)

Table 6. Values of MISE (IV/ISB) of Exponential distribution for $n = 100$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	57. 999 (35.0979/22.9006)	57. 1980 (35.0791 + 22.1193)	24.0065(35.0577/22.1139)	23.8943(5.3292/15.3788)
1.00	18.5647(3.2888/15.9999)	17.2254(3.2865/15.9657)	17.2378(3.2901/15.9475)	12.7930(2.88711/9.9059)
1.25	14.55315(2.8463/11.3721)	13.9715(2.8463/11.3747)	13.2735(2.8182/11.3066)	11.8631(2.7439/9.5296)
1.50	9. 372 1 (1.9991/7.3730)	9. 274 3 (1.9991/7.2752)	8. 810 5 (1. 935 7/6.8748)	6. 778 7 (0.5454/6.2333)
1.75	5. 201 4 (1.0494/4.112)	4. 1309 (1.0518/2.8572)	5. 100 4 (1.0341/4. 047 3)	3. 769 6 (0.9929/2.7767)
2.00	2. 418 5 (0.3700/2.0285)	2. 288 4 (0.3692/1.9192)	2. 276 5 (0.3 692 /1. 907 3)	1. 269 7 (0.3611/0.9086)

Table 7. Values of MISE (IV/ISB) of Exponential distribution for $n = 200$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	22.6039(0.9346/19.4363)	18.3669(0.9538 + 19.6130)	18.387(0.9548/19.6038)	15. 625 (0.2844/15.3405)
1.00	14.9154(0.8959/13.3794)	14.6503(0.8197/13.3492)	14.6683(0.8177/13.3471)	9.6034(0.3030/9.3004)
1.25	10. 018 (0.6958/9.3225)	9.7679(0.6952/9.2020)	9. 985 5 (0.6907/9. 210 9)	5.7404(0.4743/5.2 660)
1.50	6. 8720 (0.5649/6.3071)	6. 885 9 (0.5610/6. 324 9)	5. 768 (0.5648/5.2032)	3. 486 5 (0.4850/3.0015)
1.75	5.1479(0.4520/4.2022)	4.9325(0.4521/4.2634)	3. 956 6 (0.4513/4. 282 6)	1.8050(0.09121/1.7138)
2.00	3. 668 1 (0.3533/3. 314 8)	2.6676(0.3523/2. 746 3)	2.6872(0.3568/2. 762 8)	1.1180(0.0092/1. 108 8)

Table 8. Values of MISE (IV/ISB) of Normal distribution for $n = 100$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	18. 083 (5.3219/12.7606)	18.214(5.3117/12.7738)	17.854(5.1214/12.868)	17. 706 (5.3178/12.3880)
1.00	15. 2618 (2.5817/12.6861)	14.8460(2.5261/12.6420)	14.9096(2.5180/12.6422)	14.0037(1.3735/12.6301)
1.25	9.0551(2.130/7. 050 1)	8.7914(2.1292/6. 970 5)	8. 826 3 (2.1246/6. 970 4)	8.4277(1.4710/6. 956 7)
1.50	4.8945(1.7231/3.4424)	4.8726(1.7182/3.4347)	4.8532(1.7167/3. 435 0)	4.7214(1.5000/3.4213)
1.75	2.9461(1.2814/1. 425 5)	2.48140(1. 278 7/1. 406 9)	2. 489 9 (1. 276 7/1. 408 0)	2.4 29 (1. 032 3/1. 396 7)
2.00	1. 238 7 (0.8183/0.4204)	1. 150 5 (0.8171/0.4124)	1. 155 2 (0.8160/0.4083)	1. 117 3 (0.7090/0.4140)

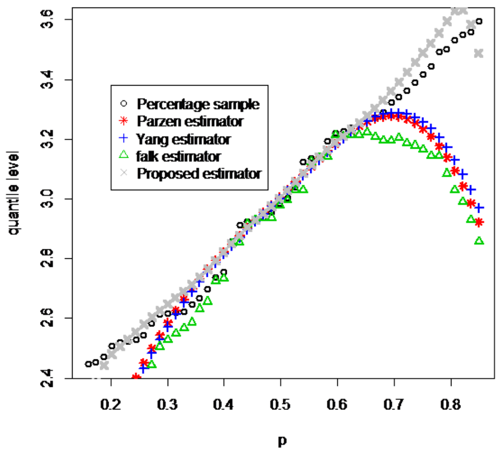
Table 9. Values of MISE (IV/ISB) of Normal distribution for $n = 200$.

φ	$\check{Q}_n^a$	$\check{Q}_n$	$\tilde{Q}_n$	$\hat{Q}_n$
0.75	3. 622 6 (1.4672/2.1574)	3. 5421 (1.40261/2.1494)	3. 539 4 (1. 400 0/2.1394)	2. 543 7 (0.42625/2.1074)
1.00	2.5456 (1.2053/1.3632)	2. 5845 (1.2048/1.2652)	2. 869 (1. 202 5/1.2665)	1. 9132 (0.6603/1.26 52)
1.25	3. 874 9 (1.0148/2. 8601)	3. 476 3 (1.0140/2. 5623)	2. 75 31 (1.0124/1. 962 9)	1. 841 5 (0.3837/1. 457 8)
1.50	1. 505 1 (0.0839/1.4212)	1. 500 2 (0.0824/1.4178)	1. 500 1 (0.0823/1. 417 8)	1. 485 8 (0.0810/1.4048)
1.75	1. 454 4 (0.0667/1. 387 7)	1. 452 1 (0.0646/1. 387 5)	1. 451 9 (0.0644/1. 387 5)	1. 4435 (0.0780/1. 384 4)
2.00	0.565 3 (0.0723/0.4930)	0.562 2 (0.0679/0.4943)	0.5 571 (0.0521/0.3943)	0.4374 (0.0457/0.3902)

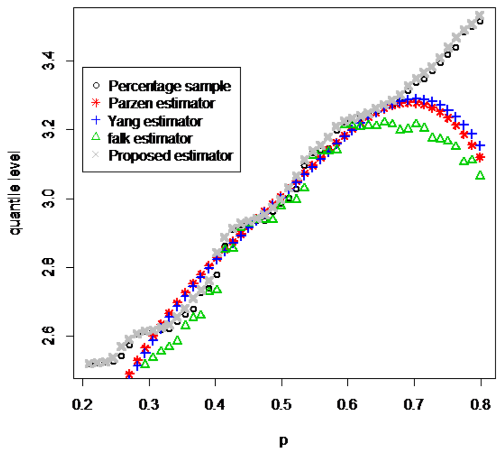
After examining all the tables, we can say that our estimator is well performed when compared with the considering estimators in the sens MISE which imply that is better also in the senses of ISB however the variance is not much influenced but still we can show that the variance of our proposed estimator is smaller than that the compared estimators for all distribution used.

3.1. An example of real data

In this part, we compare the performance of our proposed estimator with the mentioned estimators by using a real data set. The data set consist of 63 observations relates to the strength of carbon fibers tested under tension at gauge lengths of 10 mm. The data has been recently reported and analyzed by Bi and Gui [15] among others. Depending on φ value, and for each quantile interval we have graphed the performance of considering estimators, the results are presented in the following figures.

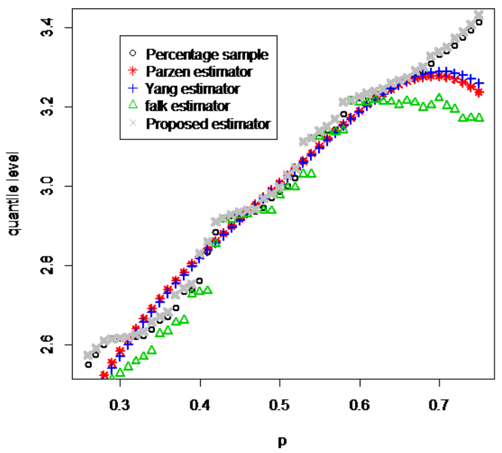


(a) $\varphi = 0.75$

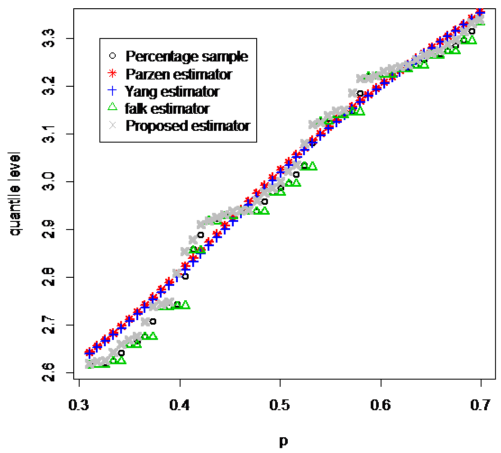


(b) $\varphi = 1.00$

Figure 1. Performance of considering estimators for different values of φ for real data.

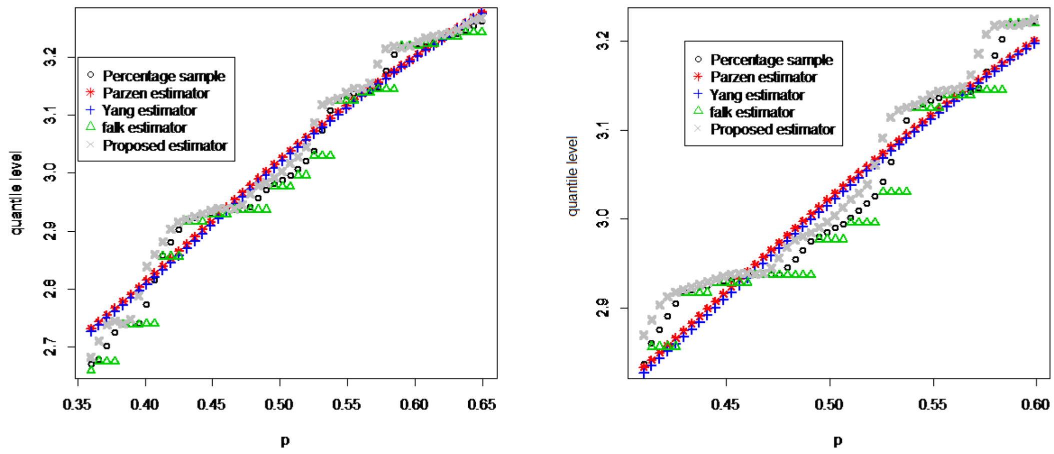


(a) $\varphi = 1.25$



(b) $\varphi = 1.50$

Figure 2. Performance of considering estimators for different values of φ for real data.



(a) $\varphi = 1.75$ (b) $\varphi = 2.00$

Figure 3. Performance of considering estimators for different values of φ for real data.

4. Conclusions

In this paper, we have developed a kernel-smoothed nonparametric quantile estimator by establishing the convex combination technique of skewed estimators in order to improve the bias in kernel quantile function estimator. Depending on the theoretical results it turned out that our proposed estimator reduces the order of bias from $O(h^2)$ to $O(h^4)$, while the variance remains at the same order as the existing estimators. Furthermore, the numerical results it is shown that the MISE of the proposed estimator is smaller than that of the used estimators in all distributions and for each sample size. Note that the reduction of the MISE of $\hat{Q}_n$ is mainly due to the bias, and the variance parts for all estimators are very close. As a results we reveal the superior performance of the proposed estimator, especially if we choose φ big enough. The estimator of the quantile function proposed here is not good at the tails.

Acknowledgments

The authors would like to thank dear reviewers for their constructive and useful comments.

Conflict of interest

All authors declare that there are no conflicts of interest concerning the publication of this paper.

References

1. M. Falk, Relative efficiency and deficiency of kernel type estimators of smooth distribution functions, *Stat. Neerl.*, **37** (1983), 73–83. https://doi.org/10.1111/j.1467-9574.1983.tb00802.x

2. E. Parzen, Nonparametric statistical data modelling, *J. Am. Stat. Assoc.*, **74** (1979), 105–121. https://doi.org/10.1080/01621459.1979.10481621
3. A. Azzalini, A note on the estimation of a distribution function and quantiles by a kernel method, *Biometrika*, **68** (1981), 326–328. https://doi.org/10.1093/biomet/68.1.326
4. J. Galambos, The asymptotic theory of extreme order statistics, In: *The theory and applications of reliability with emphasis on bayesian and nonparametric methods*, New York: Academic Press, 1977, 151–164. https://doi.org/10.1016/B978-0-12-702101-0.50014-7
5. H. A. David, H. N. Nagaraja, *Order statistics*, 3 Eds., New York: Wiley, 2003. https://doi.org/10.1002/0471722162
6. H. Yamato, Uniform convergence of an estimator of a distribution function, *Bull. Math. Stat.*, **15** (1973), 69–78. https://doi.org/10.5109/13073
7. R. D. Reiss, Nonparametric estimation of smooth distribution functions, *Scand. J. Statist.*, **8** (1981), 116–119. https://www.jstor.org/stable/4615820
8. E. A. Nadaraya, Some new estimates for distribution functions, *Theor. Probab. Appl.*, **9** (1964), 497–500. https://doi.org/10.1137/1109069
9. S. S. Ralescu, S. Sun, Necessary and sufficient conditions for the asymptotic normality of perturbed sample quantiles, *J. Stat. Plan. Infer.*, **35** (1993), 55–64. https://doi.org/10.1016/0378-3758(93)90067-G
10. S. S. Yang, A smooth nonparametric estimator of a quantile function, *J. Am. Stat. Assoc.*, **80** (1985), 1004–1011. https://doi.org/10.2307/2288567
11. S. J. Sheater, J. S. Marron, Kernel quatile estimtors, *J. Am. Stat. Assoc.*, **85** (1990), 410–416. https://doi.org/10.1080/01621459.1990.10476214
12. E. Choi, P. Hall, On bias reduction in local linear smoothing, *Biometrika*, **85** (1998), 333–345. https://doi.org/10.1093/biomet/85.2.333
13. M. Y. Cheng, E. Choi, J. Q. Fan, P. Hall, Skewing methods for two-parameter locally parametric density estimation, *Bernoulli*, **6** (2000), 169–182. https://doi.org/10.2307/3318637
14. C. Kim, W. Kim, B. U. Park, Skewing and generalized jackknifing in kernel density estimation, *Commun. Stat. Theor. M.*, **32** (2003), 2153–2162. https://doi.org/10.1081/STA-120024473
15. Q. X. Bi, W. H. Gui, Bayesian and classical estimation of stress-strength reliability for inverse Weibull lifetime models, *Algorithms*, **10** (2017), 71. https://doi.org/10.3390/a10020071