



# Nonparametric estimation of the shape functions and related asymptotic results

M. D. Jiménez-Gamero<sup>1</sup> · J. C. Pardo-Fernández<sup>2</sup>

Received: 27 January 2023 / Revised: 11 April 2023

© The Author(s) 2023

## Abstract

Arriaza et al (Metrika 82:99–124, 2019) introduced the right and left shape functions, which enjoy interesting properties in terms of describing the global form of a distribution. This paper proposes and studies nonparametric estimators of those functions. The estimators involve nonparametric estimation of the quantile and density functions. Pointwise and uniform consistency are proved under general regularity assumptions, as well as the limit in law. Simulations are included to study the practical performance of the proposed estimators. The analysis of a real data set illustrates the methodology.

**Keywords** Shape functions · Nonparametric estimation · Pointwise convergence · Weak convergence

## 1 Introduction

Arriaza et al. (2019) have introduced two functions, the left shape function and the right shape function, which, in some stochastic sense, synthesize the form of the distribution and can be employed to study the behavior of the tails and the symmetry of a random variable. Specifically, let  $X$  be an absolutely continuous random variable with probability density function  $f$  and distribution function  $F$ . For each  $u \in (0, 1)$ , let

$$x_u = F^{-1}(u) = \inf\{x : F(x) \geq u\}.$$

---

✉ J. C. Pardo-Fernández  
juancp@uvigo.gal

M. D. Jiménez-Gamero  
dolores@us.es

<sup>1</sup> Department of Statistics and Operations Research, Universidad de Sevilla, Sevilla, Spain

<sup>2</sup> CITMAga, Universidade de Vigo, Vigo, Spain

The left shape function and the right shape function of  $X$  are defined as

$$L_X(u) = \mathbb{E}\{(X - x_u)^- f(X)\}, \quad u \in (0, 1),$$

and

$$R_X(u) = \mathbb{E}\{(X - x_u)^+ f(X)\}, \quad u \in (0, 1),$$

respectively, provided that the expectations exist, where  $x^- = \max\{0, -x\}$  and  $x^+ = \max\{0, x\}$ ,  $\forall x \in \mathbb{R}$ . Since  $L_X(u) = R_{-X}(1 - u)$ ,  $\forall u \in (0, 1)$  (see Lemma 2.2 in Arriaza et al. 2019), from now on we will restrict our attention to the right shape function,  $R_X$ . In order to simplify notation, the subindex  $X$  will be suppressed from the right shape function, so, from now on, we write  $R(u)$  instead of  $R_X(u)$  when there is no possibility of confusion.

Notice that if  $\mathbb{E}|X| < \infty$  and  $f$  is bounded, then  $R(u)$  is a well-defined quantity for each  $u \in (0, 1)$  since

$$R(u) = \mathbb{E}[\{X - F^{-1}(u)\}^+ f(X)] \leq M \left\{ \mathbb{E}|X| + |F^{-1}(u)| \right\} < \infty, \quad \forall u \in (0, 1),$$

where  $M$  is a positive constant. Moreover,  $R$  is a positive and strictly decreasing function with  $\lim_{u \rightarrow 1^-} R(u) = 0$  (see Remark 2.4 in Arriaza et al. 2019). So we define  $R(1) = 0$  and, if the limit exists,  $R(0) = \lim_{u \rightarrow 0^+} R(u)$ .

The right shape function has several remarkable properties. For example, the limit when  $u$  approaches 1 of the quotient of the right shape functions of two random variables provides useful information on the relative behavior of their residual Rényi entropies of order 2, a measure of interest in reliability and other fields; if  $\mathcal{F}$  is a location-scale family of distribution functions, that is,

$$\mathcal{F} = \{F : F(x) = F_0((x - \mu)/\zeta), \forall x \in \mathbb{R}, \mu \in \mathbb{R}, \zeta > 0\}, \quad (1)$$

for some fixed distribution function  $F_0$ , then for any random variables  $X$  and  $Y$  with distribution function in  $\mathcal{F}$ , we have that  $R_X(u) = R_Y(u)$ ,  $\forall u \in (0, 1)$ , in other words, the right shape function characterizes location-scale families; among many others (see Arriaza et al. 2019). Moreover, if

$$S_X(u) = R_X(u) - R_{-X}(u), \quad u \in (0, 1),$$

then  $S_X(u) = S_Y(u)$ ,  $\forall u \in (0, 1)$ , for all  $X$  and  $Y$  with distribution function in  $\mathcal{F}$ , and  $S_X(u) = 0$ ,  $\forall u \in (0, 1)$ , if and only if the distribution of  $X$  is symmetric. These properties can be used to make inferences. For example, since  $R$  characterizes location-scale families, it may be used to build goodness-of-fit tests of these families. A key step towards the development of statistical procedures based on the right shape function is the study of an estimator of such function. This is just the objective of this paper.

**Remark 1** The definition of the function  $S$  (as before, the subindex  $X$  will be skipped when there is no possibility of confusion) is a bit different from that given in Arriaza et al. (2019), which is  $S^{\text{Arr}}(u) = R_{-X}(1 - u) - R_X(1 - u)$ . Both definitions are closely related:  $S(u) = 0, \forall u \in (0, 1)$ , if and only if  $S^{\text{Arr}}(u) = 0, \forall u \in (0, 1)$ , and  $S(u) \geq 0, \forall u \in (0, 1)$ , if and only if  $S^{\text{Arr}}(u) \leq 0, \forall u \in (0, 1)$ .

Let  $X_1, \dots, X_n$  be a random sample from  $X$ , that is,  $X_1, \dots, X_n$  are independent with the same distribution as  $X$ . Since

$$R(u) = \mathbb{E}\{(X - x_u)^+ f(X)\} = \int \{x - F^{-1}(u)\}^+ f(x) dF(x),$$

where  $\int$  stands for the integral on the whole real line, to estimate  $R(u)$  we propose to replace  $F$  with the empirical distribution function,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\},$$

where  $\mathbf{1}\{\cdot\}$  denotes the indicator function (that is,  $\mathbf{1}\{X_i \leq x\} = 1$  if  $X_i \leq x$  and  $\mathbf{1}\{X_i \leq x\} = 0$  if  $X_i > x$ ), and  $f$  with a kernel estimator

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

where  $\hat{h}$  is the bandwidth and  $K$  is a kernel. We take  $\hat{h} = \hat{\sigma} \times g(n)$ , where  $\hat{\sigma} = \hat{\sigma}(X_1, \dots, X_n)$  is an estimator of  $\sigma = \sigma(X)$ , a spread measure of  $X$ , both of them satisfying  $\sigma(aX + b) = |a|\sigma(X)$  and  $\hat{\sigma}(aX_1 + b, \dots, aX_n + b) = |a|\hat{\sigma}(X_1, \dots, X_n)$ ,  $\forall a, b \in \mathbb{R}$ , and  $g$  is a decreasing function. Further assumptions on  $\hat{h}$  and  $K$  will be specified later. For  $u \in (0, 1)$ , the empirical quantile function,  $F_n^{-1}(u)$ , is defined as follows

$$F_n^{-1}(u) = \inf\{x : F_n(x) \geq u\} = \begin{cases} X_{1:n} & \text{if } u \in I_1 = [0, 1/n], \\ X_{k:n} & \text{if } u \in I_k = ((k - 1)/n, k/n], \quad 2 \leq k \leq n, \end{cases}$$

where  $X_{1:n} \leq \dots \leq X_{n:n}$  denote the order statistics.

Therefore, we consider the following plug-in estimator of  $R(u)$ ,

$$R_n(u) = \frac{1}{n} \sum_{i=1}^n \{X_i - F_n^{-1}(u)\}^+ \hat{f}_n(X_i).$$

Notice that

$$R_n(u) = \begin{cases} \frac{1}{n} \sum_{i=k+1}^n (X_{i:n} - X_{k:n}) \hat{f}_n(X_{i:n}), & u \in I_k, \quad 1 \leq k \leq n - 1, \\ 0 & u \in I_n, \end{cases}$$

and thus,  $R_n$  is a piece-wise constant function. Observe that  $R_n(1) = 0, \forall n$ . The behavior of  $R_n$  at  $u = 1$  is consistent since, as seen before,  $\lim_{u \rightarrow 1^-} R(u) = 0$ . Observe also that, by construction,  $R_n(u) = R_n(u; X_1, \dots, X_n) = R_n(u; aX_1 + b, \dots, aX_n + b), \forall a, b \in \mathbb{R}$ , thus  $R_n(u)$  is location-scale invariant, the same as  $R(u)$ .

Analogously, we consider the following plug-in estimator of  $S(u)$ ,

$$S_n(u) = R_{X,n}(u) - R_{-X,n}(u),$$

where  $R_{X,n}(u)$  stands for  $R_n(u)$ , the estimator of  $R_X(u)$  calculated from the sample  $X_1, \dots, X_n$ , and  $R_{-X,n}(u)$  stands for the estimator of  $R_{-X}(u)$  calculated from the sample  $-X_1, \dots, -X_n$ .

Section 5 of Arriaza et al. (2019) proposes another estimator of  $R(u)$  that consists of replacing both  $f(x)$  and  $F(x)$  with kernel estimators  $f_n(x)$  and  $\hat{F}_n(x) = \int_{-\infty}^x f_n(y)dy$ , respectively. No properties of the resulting estimators of  $R$  and  $S$  were studied there. A main drawback of such estimators is that they do not have an easily computable expression, which must be approximated numerically.

The paper unfolds as follows. Section 2 derives asymptotic properties related to the pointwise and uniform consistency of the proposed estimator of the right shape function. Several results are given under different regularity assumptions. In Sect. 3 results related to the pointwise asymptotic normality and global weak convergence are detailed. In Sect. 4, a simulation study and an application to a real data set illustrate the practical performance of the estimator. This section also contains an application to a goodness-of-fit testing problem for a location-scale family that can be solved by employing the proposed estimator. All computations have been programmed and run in R (R Core Team 2020). Some conclusions and further research possibilities are discussed in Sect. 5. Finally, all proofs are deferred to Sect. 6.

Throughout the paper it will be tacitly assumed that  $X$  is an absolutely continuous random variable with cumulative distribution function  $F$  and bounded probability density function  $f$ ; all limits are taken when  $n \rightarrow \infty$ , where  $n$  denotes the sample size;  $\xrightarrow{\mathcal{L}}$  stands for the convergence in law;  $\xrightarrow{P}$  stands for the convergence in probability;  $\xrightarrow{a.s.}$  stands for the almost sure convergence; for a function  $w : (a, b) \subseteq \mathbb{R} \mapsto \mathbb{R}$  and  $x \in (a, b]$ ,  $w(x-)$  denotes the one-sided limit  $\lim_{y \rightarrow x-} w(y)$ ;  $O_P(1)$  refers to a stochastic sequence bounded in probability and  $o_P(1)$  refers to a stochastic sequence that converges to zero in probability; the kernel function  $K : \mathbb{R} \mapsto \mathbb{R}$  is a probability density function satisfying some of the following assumptions:

### Assumption 1

- (i)  $K$  has compact support and is Lipschitz continuous.
- (ii)  $K$  has bounded variation.
- (iii)  $K$  is symmetric,  $K(x) = K(-x), \forall x \in \mathbb{R}$ .
- (iv) The support of  $K$  is  $[c, d]$ , for some  $\infty < c < d < \infty, K(c) = K(d) = 0, K$  is twice differentiable on  $(c, d)$ , with bounded derivatives  $K'$  and  $K''$ .

The bandwidth  $h = \sigma \times g(n)$  will be assumed to satisfy some of the following assumptions:

**Assumption 2**

- (i)  $h \rightarrow 0$  and  $\sum_{n \geq 1} \exp\{-\varepsilon n h^2\} < \infty, \forall \varepsilon > 0.$
- (ii)  $h \rightarrow 0, nh \rightarrow \infty, nh^4 \rightarrow 0.$
- (iii)  $h \rightarrow 0, nh^2 \rightarrow \infty, nh^4 \rightarrow 0.$

In Assumption 2, notice that (iii) is stronger than (ii). On the other hand, the condition  $\sum_{n \geq 1} \exp\{-\varepsilon n h^2\} < \infty$  in (i) implies  $nh^2 \rightarrow \infty$  in (iii), but does not entail  $nh^4 \rightarrow 0.$

**2 Almost sure limit**

The next theorem gives the almost sure limit of  $R_n(u)$ , for each  $u \in (0, 1).$

**Theorem 1** *Suppose that  $\mathbb{E}|X| < \infty,$  that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0,$  that  $f$  is uniformly continuous, that  $K$  satisfies Assumption 1 (i) and that  $h$  satisfies Assumption 2 (i). Let  $u \in (0, 1).$  Suppose that there is a unique solution in  $x$  of  $F(x-) \leq u \leq F(x).$  Then,  $R_n(u) \xrightarrow{a.s.} R(u).$*

Let  $(a, b)$  denote the support of  $F,$  that is,  $a = \sup\{x : F(x) = 0\}$  and  $b = \inf\{x : F(x) = 1\}.$  A key assumption in Theorem 1 to get the a.s. convergence is the uniform continuity of  $f,$  necessary to get the uniform convergence of  $\hat{f}_n$  to  $f.$  This assumption may not hold, specially if either  $a > -\infty$  or  $b < \infty.$  Nevertheless, if such assumption fails, we still can get the a.s. convergence by using other assumptions, as stated in the next theorem.

**Theorem 2** *Suppose that  $f$  is twice continuously differentiable on  $(a, b),$  that  $\mathbb{E}(X^2) < \infty, \mathbb{E}\{f'(X)^2\} < \infty, \mathbb{E}\{f''(X)^2\} < \infty$  and  $\mathbb{E}\{X^2 f''(X)^2\} < \infty,$  that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0,$  that  $K$  satisfies Assumption 1 (i) and (iii) and that  $h$  satisfies Assumption 2 (i) and (ii). Let  $u \in (0, 1).$  Suppose that there is a unique solution in  $x$  of  $F(x-) \leq u \leq F(x).$  Then,  $R_n(u) \xrightarrow{a.s.} R(u).$*

In general, it is not possible to get the uniform a.s. convergence of  $R_n$  to  $R$  because the convergence of the empirical quantile function to the population quantile function is not uniform, unless  $F$  has finite support. In such a case, the next theorem shows that we also have the uniform convergence of  $R_n$  to  $R.$

**Theorem 3** *Suppose that  $-\infty < a < b < \infty, f$  is continuous in  $(a, b)$  and  $\inf_{0 \leq u \leq 1} f(F^{-1}(u)) > 0.$  Suppose also that  $K$  satisfies Assumption 1 (i), that  $h$  satisfies Assumption 2 (i), and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0.$  Then,*

$$\sup_{0 \leq u \leq 1} |R_n(u) - R(u)| \xrightarrow{a.s.} 0.$$

Let  $w : [0, 1] \mapsto \mathbb{R}$  a measurable positive function and let  $L^2(w)$  denote the separable Hilbert space of (equivalence classes of) measurable functions  $f : [0, 1] \mapsto$

$\mathbb{R}$  satisfying  $\int_0^1 f(u)^2 w(u) du < \infty$ , the scalar product and the resulting norm in  $L^2(w)$  will be denoted by  $\langle f, g \rangle_w = \int_0^1 f(u)g(u)w(u)du$  and  $\|f\|_w = \sqrt{\langle f, f \rangle_w}$ , respectively. If  $w(u) = 1$ ,  $0 \leq u \leq 1$ , then we simply denote  $L^2(w)$ ,  $\langle \cdot, \cdot \rangle_w$  and  $\|\cdot\|_w$  by  $L^2$ ,  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$ , respectively.

As said before, in general, it is not possible to obtain the uniform convergence of  $R_n$  to  $R$ . However, under quite general assumptions, it can be shown that  $R_n$  converges to  $R$  in  $L^2(w)$ . A first result in this sense is given in the following theorem.

**Theorem 4** *Suppose that  $\mathbb{E}(X^2) < \infty$ , that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , that  $f$  is uniformly continuous, that  $K$  satisfies Assumption 1 (i) and that  $h$  satisfies Assumption 2 (i). Then  $R \in L^2$  and  $\|R_n - R\| \xrightarrow{a.s.} 0$ .*

It readily follows that if  $w : [0, 1] \mapsto \mathbb{R}$  is a measurable bounded positive function, then the statement in the previous theorem also holds in  $L^2(w)$ .

**Corollary 1** *Let  $w : [0, 1] \mapsto \mathbb{R}$  be a measurable bounded positive function. Suppose that assumptions in Theorem 4 hold. Then  $R \in L^2(w)$  and  $\|R_n - R\|_w \xrightarrow{a.s.} 0$ .*

The uniform continuity of  $f$  can be replaced with other assumptions.

**Theorem 5** *Let  $w : [0, 1] \mapsto \mathbb{R}$  be a measurable bounded positive function. Suppose that  $f$  is twice continuously differentiable on  $(a, b)$ , that  $\mathbb{E}(X^2) < \infty$ ,  $\mathbb{E}\{f'(X)^2\} < \infty$ ,  $\mathbb{E}\{f''(X)^2\} < \infty$  and  $\mathbb{E}\{X^2 f''(X)^2\} < \infty$ , that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , that  $K$  satisfies Assumption 1 (i) and (iii) and that  $h$  satisfies Assumption 2 (i) and (ii). Then  $R \in L^2(w)$  and  $\|R_n - R\|_w \xrightarrow{a.s.} 0$ .*

**Remark 2** The assumptions in the statements of the previous asymptotic results exclude the optimal rate of the bandwidth for density estimation, to wit,  $n^{-1/5}$ . Notice, however, that the objective here differs from the mere estimation of the density. This situation is not uncommon in nonparametric literature. For instance, when the target is to estimate a distribution function using the kernel method, Azzalini (1981) showed that the optimal bandwidth is of order  $n^{-1/3}$ . Other examples can be found in Pardo-Fernández et al. (2015) and Pardo-Fernández and Jiménez-Gamero (2019).

**Remark 3** All properties studied so far were stated for  $R_n$  as an estimator of  $R$ . Clearly, these properties carry over  $S_n$  as an estimator of  $S$ , which are not given to save space. The finite sample performance  $R_n(u)$  and  $S_n(u)$  as estimators of  $R(u)$  and  $S(u)$ , respectively, will be numerically studied in Sect. 4 for data coming from a uniform distribution.

### 3 Weak limit

We first study the weak limit of  $\sqrt{n} \{R_n(u) - R(u)\}$  at each  $u \in (0, 1)$ .

**Theorem 6** *Suppose that  $f$  is twice continuously differentiable on  $(a, b)$ , that  $\mathbb{E}(X^2) < \infty$ ,  $\mathbb{E}\{f'(X)^2\} < \infty$ ,  $\mathbb{E}\{f''(X)^2\} < \infty$  and  $\mathbb{E}\{X^2 f''(X)^2\} < \infty$ , that  $\sqrt{n}(\hat{\sigma} - \sigma) =$*

$O_P(1)$ , that  $K$  satisfies Assumption 1 (i), (iii) and (iv), and that  $h$  satisfies Assumption 2 (i) and (iii). Let  $u \in (0, 1)$  be such that  $f(F^{-1}(u)) > 0$ , then

$$\sqrt{n} \{R_n(u) - R(u)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(u) + o_P(1),$$

where

$$Y_i(u) = 2[\{X_i - F^{-1}(u)\}^+ f(X_i) - R(u)] + \frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mu(u),$$

$1 \leq i \leq n$ , with

$$\mu(u) = \mathbb{E} \left[ f(X) \mathbf{1}\{F^{-1}(u) < X\} \right],$$

and therefore

$$\sqrt{n} \{R_n(u) - R(u)\} \xrightarrow{\mathcal{L}} Z \sim N(0, \varrho^2(u)),$$

where  $\varrho^2(u) = \mathbb{E}\{Y_1(u)^2\}$ .

Recall that to estimate  $R(u)$  we replaced the population quantile function  $F^{-1}$  with the empirical quantile function  $F_n^{-1}$  and the population density function  $f$  with the kernel estimator  $\hat{f}_n$ . Each of these two replacements have an effect on the asymptotic behavior of  $\sqrt{n} \{R_n(u) - R(u)\}$ : (a) the first replacement is the responsible for the term  $\frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mu(u)$  in the expression of  $Y_i(u)$ ; and (b) the second replacement is the responsible for the coefficient 2 in the first part of  $Y_i(u)$ . Notice that, under the assumptions made, taking the bandwidth data dependent,  $\hat{h} = \hat{\sigma} g(n)$ , has no effect on the asymptotic distribution of  $\sqrt{n} \{R_n(u) - R(u)\}$ .

The result in Theorem 6 can be used to give approximate (in the sense of asymptotic) confidence intervals for  $R(u)$ . Let  $\hat{\varrho}(u)$  denote any consistent estimator of  $\varrho(u)$  (see the explanation below for a candidate). If  $z_v$  is such that  $\Phi(z_v) = v$ , where  $\Phi$  stands for the cumulative distribution function of the standard normal distribution, for a given  $\alpha \in (0, 1)$ , then

$$(R_n(u) - z_{1-\alpha/2} \hat{\varrho}(u) / \sqrt{n}, R_n(u) + z_{1-\alpha/2} \hat{\varrho}(u) / \sqrt{n}) \tag{2}$$

is a random confidence interval for  $R(u)$  with asymptotic confidence level  $1 - \alpha$ . If  $Y_1(u), \dots, Y_n(u)$  were observed, since  $\varrho^2(u) = \mathbb{E}\{Y_1(u)^2\}$ , which also coincides with the variance of  $Y_1(u)$ ,  $\mathbb{V}\{Y_1(u)\}$ , one could consistently estimate  $\varrho^2(u)$  by means of the sample variance of  $Y_1(u), \dots, Y_n(u)$ . The point is that  $Y_1(u), \dots, Y_n(u)$  depend on unknown quantities. Taking into account that  $\mathbb{V}\{Y_1(u)\} = \mathbb{V}\{W_1(u)\}$ , where  $W_i(u) = 2\{X_i - F^{-1}(u)\}^+ f(X_i) + \mathbf{1}\{X_i \leq F^{-1}(u)\} \mu(u) / f(F^{-1}(u))$ ,  $1 \leq i \leq n$ , we propose to replace  $f$  by  $\hat{f}_n$  and  $F$  by  $F_n$  in the expression of  $W_1(u), \dots, W_n(u)$ , giving rise to

$$\hat{W}_i(u) = 2\{X_i - F_n^{-1}(u)\}^+ \hat{f}_n(X_i) + \mathbf{1}\{X_i \leq F_n^{-1}(u)\} \hat{\mu}(u) / \hat{f}_n(F_n^{-1}(u)), \quad 1 \leq i \leq n, \tag{3}$$

where

$$\hat{\mu}(u) = \int \hat{f}_n(x) \mathbf{1}\{F_n^{-1}(u) < x\} dF_n(x) = \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i) \mathbf{1}\{F_n^{-1}(u) < X_i\},$$

and then estimate  $\varrho^2(u)$  by means of the sample variance of  $\hat{W}_1(u), \dots, \hat{W}_n(u)$ , that we denote by  $\hat{\varrho}^2(u)$ . The finite sample performance of the confidence interval in (2) as well as the goodness of  $\hat{\varrho}^2(u)/n$  as an approximation to the variance of  $R_n(u)$  will be examined in Sect. 4 for data coming from a uniform distribution.

Finally, we study the convergence in law of  $n\|R_n - R\|_w^2$ , for some adequate  $w$ . In general,  $n\|R_n - R\|^2$  does not possess a weak limit unless we assume rather strong assumptions on  $F$ . This is because the derivations to study such a limit involve those of  $\sqrt{n}\{F_n^{-1} - F^{-1}\}$  in  $L^2$ , and therefore, it inherits the same limitations (see, for example, del Barrio et al. 2000, 2005). A convenient way to overcome these difficulties is to consider, instead of  $\|\cdot\|^2$ , the norm in  $L^2(w)$ , with  $w(u) = f^2(F^{-1}(u))$ . This weight function is taken for analytical convenience. It may seem a bit odd, because  $f$  (and hence  $F$ ) is unknown in practical applications. Nevertheless, as stated in the Introduction,  $n\|R_n - R\|_w^2$  could be used as a test statistic for testing goodness-of-fit to a location-scale family (1), and in such a case under the null hypothesis  $f(F^{-1}(u)) = \frac{1}{\xi} f_0(F_0^{-1}(u))$ , so we can take  $w(u) = f_0(F_0^{-1}(u))$  which is known in that testing framework. Later on, we will discuss other weight functions

We first show that, under some conditions, the linear approximation for  $\sqrt{n}\{R_n(u) - R(u)\}$  given in Theorem 6 for a fixed  $u \in (0, 1)$ , is valid for all  $u$  in certain intervals, in the  $L^2(w)$  sense.

**Theorem 7** *Suppose that  $f$  is twice continuously differentiable on  $(a, b)$ , that  $\mathbb{E}(X^2) < \infty$ ,  $\mathbb{E}\{f'(X)^2\} < \infty$ ,  $\mathbb{E}\{f''(X)^2\} < \infty$  and  $\mathbb{E}\{X^2 f''(X)^2\} < \infty$ , that  $\sqrt{n}(\hat{\sigma} - \sigma) = O_P(1)$ , that  $K$  satisfies Assumption 1 (i), (iii) and (iv), and that  $h$  satisfies Assumption 2 (i) and (iii). Suppose also that  $f(F^{-1}(u)) > 0$ ,  $u \in (0, 1)$  and*

$$\sup_{0 < u < 1} u(1-u) \frac{|f'(F^{-1}(u))|}{f^2(F^{-1}(u))} \leq \gamma,$$

for some finite  $\gamma > 0$ . Let

$$A = \lim_{x \downarrow a} f(x) < \infty.$$

Suppose also that if  $A = 0$  then  $f$  is nondecreasing on an interval to the right of  $a$ . Let  $w(u) = f^2(F^{-1}(u))$ , then

$$\sqrt{n}\{R_n(u) - R(u)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(u) + r_n(u), \quad u \in \mathbb{I}_n = \left(0, \frac{n-1}{n}\right],$$



$$\int_{\mathbb{I}_n} r_n(u)^2 w(u) du = o_P(1).$$

**Corollary 2** Under assumptions in Theorem 7,  $n\|R_n - R\|_w^2 \xrightarrow{\mathcal{L}} \|Z\|_w^2$ , where  $\{Z(u), 0 \leq u \leq 1\}$  is a zero-mean Gaussian process on  $L^2(w)$  with  $Z(1) = 0$  and covariance function  $cov\{Z(u), Z(s)\} = \mathbb{E}\{Y_1(u) Y_1(s)\}$ ,  $u, s \in (0, 1)$ .

From the proof of Theorem 7 and Theorem 4.6 (i) of del Barrio et al. (2005), the results in Theorem 7 and Corollary 2 keep on being true for any bounded weight function  $w$  satisfying

$$\int_0^1 \frac{u(1-u)}{f^2(F^{-1}(u))} w(u) du < \infty, \quad \lim_{u \uparrow 1} \frac{(1-u) \int_u^1 w(u) du}{f^2(F^{-1}(u))} = 0.$$

Although this result may seem more general than those stated in Theorem 7 and Corollary 2, notice that the choice of an adequate weight function requires a strong knowledge of  $f$ .

As observed after Theorem 6, the replacement of the population quantile function  $F^{-1}$  with the empirical quantile function  $F_n^{-1}$  in the expression of  $R$  to build the estimator  $R_n$  is the responsible for the term  $\frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mu(u)$  in the expression of  $Y_i(u)$ , which makes  $Y_i(u)$  to inherit the properties and difficulties of estimating the quantile function. So, one may wonder if there is any advantage in using the right (left) shape function instead of the quantile function. From a theoretical point of view the answer is yes: in order to use the quantile function one must assume certain conditions on the right tail and on the left tail of the distribution, while if one uses right (left) shape function then only assumptions on the left (right) tail are necessary. This is because both  $R_n(u)$  and  $R(u)$  (respectively,  $L_n(u)$  and  $L(u)$ ) go to 0 as  $u \uparrow 1$  (respectively,  $u \downarrow 0$ ).

**Remark 4** The properties studied in this section for  $R_n - R$  are inherited by  $S_n - S$ . Specifically, let  $u \in (0, 1)$ , then

$$(S_n(u) - z_{1-\alpha/2} \hat{\varrho}_S(u) / \sqrt{n}, S_n(u) + z_{1-\alpha/2} \hat{\varrho}_S(u) / \sqrt{n}) \tag{4}$$

is an approximate (in the sense of asymptotic) confidence interval for  $S(u)$ , where  $\hat{\varrho}_S^2(u)$  is the sample variance of  $V_1(u), \dots, V_n(u)$ , where  $V_i(u) = \hat{W}_{X,i}(u) - \hat{W}_{-X,i}(u)$ ,  $\hat{W}_{X,i}(u)$  are the quantities defined in (3) calculated on the sample  $X_1, \dots, X_n$ , and  $\hat{W}_{-X,i}$  are the quantities defined in (3) calculated on  $-X_1, \dots, -X_n$ ,  $i = 1, \dots, n$ . Notice that if 0 does not belong to the confidence interval (4) one may conclude that the law of  $X$  is not symmetric.

## 4 Some numerical illustrations

### 4.1 Estimation of $R$ and $S$

If  $X$  has a uniform distribution on the interval  $(a, b)$ ,  $X \sim U(a, b)$ , then  $R(u) = 0.5(1 - u)^2$  (see, Table 1 in Arriaza et al. 2019). For several values of  $n$ , we have generated 10,000 random samples with size  $n$  from a distribution  $U(0, 1)$ . For each sample, we have estimated  $R$  using  $R_n$  taking as  $K$  the Epanechnikov kernel (scaled so that it has variance 1) and  $h = sd \times n^{-\tau}$ , with  $\tau = 0.35, 0.40, 0.45, 0.49$ , and  $sd^2$  denoting the sample variance. Notice that for these choices of  $h$  Assumption 2 (i), (ii) and (iii) are met. Tables 1 and 2 show the value of  $R(u)$ , the bias and the standard deviation of the values of  $R_n(u)$ , the mean of the standard deviation estimator  $\hat{\sigma}(u)/\sqrt{(n)}$  (recall that it is based on asymptotic arguments), and the coverage of the confidence interval (2) calculated at the nominal level 95%, for  $u = 0.1, \dots, 0.9$  and  $n = 100, 250, 500, 1000$ . Figure 1 displays the graph of 1000 estimations for  $\tau = 0.45$  in grey joint with the population shape function in black. Looking at these tables and figure we see that the bias and the variance become smaller as  $u$  approaches 1; the standard deviation estimator is, on average, a bit larger than the true standard deviation, specially for smaller sample sizes; the bias depends on the values of  $\tau$  and  $u$ , being negative for smaller values of  $\tau$  and for larger values of  $u$ ; as for the coverage of the confidence interval, it also depends the values of  $\tau$  and  $u$ : it is rather poor for smaller values of  $\tau$  and larger values of  $u$ , this is because in such cases the value of the bias in relation to the standard deviation estimator is non-negligible. As expected from Theorem 3, the differences between  $R_n(u)$  and  $R(u)$  become smaller as  $n$  increases, uniformly in  $u$ .

A similar experiment was carried out for the estimation of  $S$ , whose results are summarized in Tables 3 and 4 and Fig. 2. Looking at these tables and figure we see that the bias is quite small in all cases; that the variance becomes smaller as  $u$  approaches 1 and it also decreases with  $n$ ; the standard deviation estimator is, on average, a bit larger than the true standard deviation, the differences become smaller as  $n$  increases; this fact provokes that the coverage of the confidence intervals is larger than the nominal value, specially for small sample sizes.

### 4.2 Glass fibre breaking strengths

As an illustration, we will analyse a real data set already considered in Arriaza et al. (2019) and which had been previously introduced by Smith and Naylor (1987). The set consists of 63 observations of the breaking strength of glass fibres of 1.5 cm of length collected by the National Physical Laboratory in England (for more details about the data set, see Smith and Naylor 1987). The left panel of Fig. 3 depicts the histogram and the kernel density estimator obtained from the data. As discussed in Arriaza et al. (2019), this explanatory analysis suggests a certain negative skewness in the distribution. The right panel of Fig. 3 displays the estimator  $S_n(u)$ ,  $u \in (0, 1)$ , with  $\hat{h} = sd \times n^{-0.45}$  and taking as  $K$  the Epanechnikov kernel (scaled so that it has variance 1). Other values for  $\hat{h}$  have been investigated and similar results were

**Table 1** Values of  $R(u)$ , bias  $\times 10^3$  (bi) of  $R_n(u)$ , standard deviation  $\times 10^3$  (se) of  $R_n(u)$ , the mean of the standard deviation estimator  $\hat{\sigma}(u)/\sqrt{n} \times 10^3$  (ms), and the coverage of the confidence interval (2) calculated at the nominal level 95% (co), calculated with  $h = sd \times n^{-\tau}$ , for  $u = 0.1, \dots, 0.9$ , and sample sizes  $n = 100, 250$

$n$	$\tau$	R(0.1)	R(0.2)	R(0.3)	R(0.4)	R(0.5)	R(0.6)	R(0.7)	R(0.8)	R(0.9)	
100	0.35	bi	-10.797	-11.692	-11.894	-11.798	-11.023	-9.405	-7.547	-5.116	-2.070
		se	45.235	36.861	29.196	22.188	16.145	10.938	6.182	2.718	0.727
		ms	48.443	40.106	32.244	24.966	18.312	12.403	7.382	3.509	1.158
		co	93.4	92.9	91.9	90.3	87.4	84.0	78.5	66.3	60.0
	0.40	bi	0.268	-3.179	-5.419	-6.639	-6.905	-6.793	-5.865	-4.304	-1.915
		se	47.521	38.489	30.737	23.566	17.266	11.824	6.873	3.040	0.751
		ms	51.228	42.581	34.395	26.791	19.786	13.521	8.125	3.854	1.169
		co	96.0	95.2	94.5	93.0	91.8	89.0	84.7	76.8	66.1
	0.45	bi	9.454	4.778	1.469	-1.284	-2.955	-3.874	-4.128	-3.433	-1.734
		se	48.224	39.738	32.131	24.673	18.085	12.614	7.630	3.490	0.811
		ms	53.980	45.125	36.610	28.738	21.382	14.743	8.951	4.307	1.230
		co	97.0	96.9	96.6	96.3	95.2	93.1	89.0	83.2	72.3
0.49	bi	19.893	13.900	8.250	3.608	0.191	-1.777	-2.742	-2.716	-1.562	
	se	50.227	41.557	33.386	25.966	19.262	13.314	8.132	3.792	0.910	
	ms	56.829	47.602	38.720	30.378	22.719	15.767	9.660	4.705	1.315	
	co	97.8	98.1	97.5	97.3	96.4	95.2	92.4	87.2	76.7	

Table 1 continued

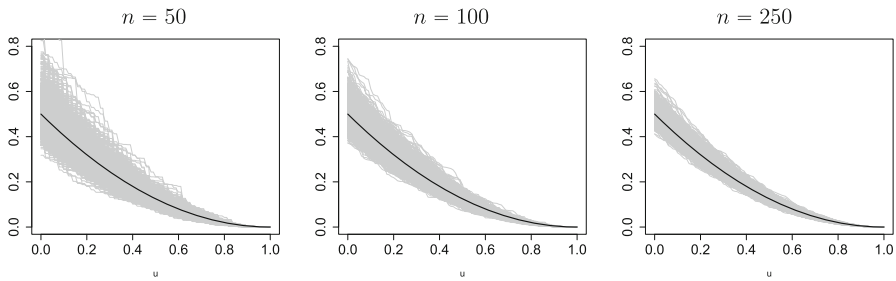
<i>n</i>	$\tau$	R(0.1)	R(0.2)	R(0.3)	R(0.4)	R(0.5)	R(0.6)	R(0.7)	R(0.8)	R(0.9)
		0.405	0.320	0.245	0.180	0.125	0.080	0.045	0.020	0.005
250	bi	-7.600	-8.009	-7.890	-7.637	-7.049	-6.216	-4.947	-3.431	-1.541
	se	29.106	24.170	19.348	14.939	11.092	7.567	4.540	2.081	0.510
	ms	30.583	25.424	20.568	16.033	11.893	8.188	4.981	2.408	0.698
	co	94.1	92.9	92.1	90.9	88.9	85.1	79.1	67.0	32.9
0.40	bi	0.001	-1.561	-2.919	-3.572	-4.096	-3.997	-3.549	-2.635	-1.298
	se	29.866	24.899	20.231	15.817	11.724	8.105	4.898	2.351	0.581
	ms	31.905	26.597	21.588	16.919	12.619	8.759	5.401	2.656	0.755
	co	95.8	95.4	94.9	94.3	92.6	90.3	87.0	79.0	56.2
0.45	bi	7.922	4.943	2.487	0.424	-1.022	-1.899	-2.126	-1.838	-1.080
	se	31.070	25.586	20.971	16.319	12.219	8.494	5.299	2.617	0.663
	ms	33.351	27.877	22.714	17.873	13.409	9.383	5.850	2.936	0.834
	co	96.6	96.9	96.5	96.5	95.4	94.4	91.8	86.7	71.8
0.49	bi	15.937	11.341	7.351	4.072	1.697	0.109	-0.900	-1.198	-0.877
	se	32.235	26.665	21.552	16.901	12.700	8.832	5.563	2.805	0.774
	ms	34.735	29.123	23.769	18.774	14.131	9.952	6.258	3.175	0.922
	co	96.3	96.9	97.1	97.2	97.0	96.4	94.5	91.0	79.2

**Table 2** Values of  $R(u)$ , bias  $\times 10^3$  (bi) of  $R_n(u)$ , standard deviation  $\times 10^3$  (se) of  $R_n(u)$ , the mean of the standard deviation estimator  $\hat{\sigma}(u)/\sqrt{(n)} \times 10^3$  (ms), and the coverage of the confidence interval (2) calculated at the nominal level 95% (co), calculated with  $h = sd \times n^{-\tau}$ , for  $u = 0.1, \dots, 0.9$ , and sample sizes  $n = 500, 1000$

$n$	$\tau$	R(0.1)	R(0.2)	R(0.3)	R(0.4)	R(0.5)	R(0.6)	R(0.7)	R(0.8)	R(0.9)	
500	0.35	bi	-6.737	-6.725	-6.416	-6.103	-5.476	-4.695	-3.810	-2.618	-1.217
		se	21.238	17.547	14.201	11.009	8.149	5.618	3.424	1.666	0.408
		ms	21.580	17.968	14.573	11.410	8.515	5.909	3.650	1.808	0.518
		co	93.1	92.0	91.1	89.9	87.5	84.7	78.4	65.8	29.6
	0.40	bi	-0.508	-1.388	-2.105	-2.587	-2.784	-2.742	-2.518	-1.855	-0.954
		se	21.697	17.955	14.487	11.310	8.452	5.937	3.684	1.824	0.494
		ms	22.368	18.676	15.201	11.958	8.974	6.277	3.924	1.981	0.576
		co	95.4	95.1	94.8	94.3	93.3	91.4	87.7	80.9	58.6
0.45	bi	5.994	4.008	2.149	0.727	-0.386	-1.116	-1.343	-1.215	-0.734	
	se	22.152	18.296	14.768	11.679	8.808	6.186	3.904	1.962	0.568	
	ms	23.186	19.407	15.836	12.495	9.409	6.629	4.186	2.153	0.644	
	co	96.2	96.3	96.7	96.2	95.9	94.7	93.0	88.7	74.4	
	0.49	bi	11.973	8.573	5.886	3.566	1.771	0.419	-0.446	-0.709	-0.547
		se	22.671	18.814	15.214	11.901	9.054	6.328	4.016	2.084	0.636
		ms	23.940	20.070	16.415	12.979	9.813	6.944	4.414	2.291	0.706
		co	95.3	96.1	96.5	96.7	96.8	96.5	95.4	92.8	83.6

Table 2 continued

$n$	$\tau$	R(0.1)	R(0.2)	R(0.3)	R(0.4)	R(0.5)	R(0.6)	R(0.7)	R(0.8)	R(0.9)	
		0.405	0.320	0.245	0.180	0.125	0.080	0.045	0.020	0.005	
1000	0.35	bi se ms co	-5.546 15.069 15.271 92.0	-5.365 12.432 12.728 92.0	-5.207 10.019 10.339 91.3	-4.882 7.787 8.120 89.5	-4.375 5.812 6.087 87.8	-3.710 4.096 4.260 83.8	-2.897 2.574 2.665 77.5	-1.996 1.277 1.346 65.3	-0.965 0.341 0.392 28.2
	0.40	bi se ms co	-0.753 15.119 15.709 95.3	-1.355 12.676 13.128 95.3	-1.820 10.470 10.700 94.5	-2.085 8.261 8.439 93.6	-2.096 6.220 6.358 92.9	-1.986 4.353 4.477 91.5	-1.814 2.722 2.826 88.0	-1.355 1.386 1.456 81.6	-0.721 0.399 0.439 59.7
	0.45	bi se ms	4.484 15.655 16.177	2.965 13.153 13.548	1.578 10.727 11.064	0.547 8.422 8.754	-0.079 6.280 6.619	-0.585 4.437 4.686	-0.804 2.846 2.985	-0.814 1.498 1.560	-0.508 0.453 0.487
	0.49	bi se ms co	8.747 15.793 16.593 94.3	6.480 13.268 13.919 94.8	4.515 10.781 11.394 95.2	2.917 8.533 9.026 95.9	1.551 6.453 6.846 96.1	0.572 4.598 4.862 95.2	-0.064 2.928 3.115 96.2	-0.382 1.552 1.642 93.7	-0.346 0.490 0.527 86.9



**Fig. 1** Graphs of 1000 generated estimations calculated with  $h = sd \times n^{-0.45}$  in grey joint with the population shape function  $R(u)$  in black

obtained. The graph also displays the confidence intervals in (4) for  $S(u)$  calculated at the nominal level 90% for  $u = 0.1, \dots, 0.9$ . The confidence intervals are also detailed in Table 5. Notice that the estimator of the function  $S$  tends to lie over the horizontal axis, which indicates asymmetry in the distribution. Moreover, the confidence intervals for  $S(0.1), S(0.2)$  and  $S(0.4)$  do not contain the zero. This conclusion is in agreement with Arriaza et al. (2019).

### 4.3 An application to testing goodness-of-fit

Now we consider the problem of testing goodness-of-fit to a uniform distribution, that is, we want to test

$$\begin{aligned}
 H_0 : X &\sim U(a, b), \quad \text{for some } a < b, \quad a, b \in \mathbb{R}, \\
 H_1 : X &\not\sim U(a, b), \quad \text{for all } a < b, \quad a, b \in \mathbb{R},
 \end{aligned}$$

on the basis of a sample of size  $n$  from  $X$ . Let  $f_0$  and  $F_0$  denote the probability density function and the cumulative distribution function of the  $U(0, 1)$  law, respectively. Then  $w(u) = f_0(F_0(u)) = \mathbf{1}\{0 \leq u \leq 1\}$ , and thus  $\|\cdot\| = \|\cdot\|_w$ . Let  $R$  denote the right shape function of  $X$  and let  $R_0$  denote the right shape function of a uniform distribution,  $R_0(u) = 0.5(1 - u)^2$ . From Theorem 3, it follows that  $\|R_n - R_0\|$  converges to  $\|R - R_0\|$ , which is equal to 0, under the null, and a positive quantity under alternatives. Thus, it seems reasonable to consider the test that rejects  $H_0$  for large values of  $T_n = n\|R_n - R_0\|^2$ . The critical region is  $T_n \geq t_{n,\alpha}$ , where  $t_{n,\alpha}$  is the  $\alpha$  upper percentile of the null distribution of  $T_n$ , whose value can be calculated by simulation by generating data from a  $U(0, 1)$  law, since the null distribution of  $T_n$  does not depend on the values of  $a$  and  $b$ , but only on  $F_0$ . Notice that  $T_n$  has the readily computable expression

$$T_n = \sum_{k=1}^{n-1} r_k^2 + \frac{n}{20} - \frac{n}{3} \sum_{k=1}^{n-1} c_k r_k,$$

where

$$r_k = \frac{1}{n} \sum_{i=k+1}^n (X_{i:n} - X_{k:n}) \hat{f}_n(X_{i:n}) \quad \text{and} \quad c_k = (1 - (k-1)/n)^3 - (1 - k/n)^3.$$

There are many tests in the statistical literature for testing  $H_0$  against  $H_1$ , and the objective of this section is not to provide an exhaustive list of such tests, but only to suggest a possible application of the results stated in the previous sections. In our view, this as well as other possible applications deserve further separate research, out of the scope of this manuscript. Nevertheless, since  $T_n$  is closely related to the Wasserstein distance between  $F_n$  and  $F_0$ , which is equal to the squared root of the  $L^2$  norm of the difference between the empirical quantile function and the quantile function of  $F_0$ , we carried out a small simulation experiment in order to compare the powers of the newly proposed tests and the one based on the Wasserstein distance. To make the Wasserstein distance invariant with respect to location and scale changes, we consider as test statistic  $W_n = n \|F_n^{-1} - F_0^{-1}\|^2 / \hat{\sigma}^2$ ,  $\hat{\sigma}^2$  denoting the sample variance, and reject the null hypothesis for large values of  $W_n$  (see, del Barrio et al. 2000). The critical region is  $W_n \geq w_{n,\alpha}$ , where  $w_{n,\alpha}$  is the  $\alpha$  upper percentile of the null distribution of  $W_n$ , whose value can be calculated by simulation by generating data from a  $U(0, 1)$  law, since the null distribution of  $W_n$  does not depend on the values of  $a$  and  $b$ , but only on  $F_0$ .

Table 6 displays the values  $t_{n,\alpha}$  and  $w_{n,\alpha}$  for  $n = 30, 50$  and  $\alpha = 0.05$ , calculated by generating 100,000 samples. To calculate  $T_n$  we used the Epanechnikov kernel (scaled so that it has variance 1) and  $\hat{h} = sd \times n^{-\tau}$ ,  $\tau = 0.26, 0.30, 0.35, 0.40, 0.45, 0.49$ . As alternatives, we considered several members of the log-Lindley distribution, a family with support on  $(0, 1)$  with a large variety of shapes (see Gómez-Déniz et al. 2014; Jodrá and Jiménez-Gamero 2016) and probability density function

$$f(x; \kappa, \lambda) = \frac{\kappa^2}{1 + \kappa\lambda} (\lambda - \log(x)) x^{\kappa-1}, \quad x \in (0, 1),$$

for some  $\kappa > 0$  and  $\lambda \geq 0$ . Figure 4 represents the empirical power of the two tests for  $\kappa = 1.5, 2, 2.5$  and  $0.6 \leq \lambda \leq 5$ , calculated by generating 10,000 samples for each combination of the parameter values. Looking at Fig. 4 we see that, although according to Corollary 2 the asymptotic null distribution of  $T_n$  does not depend on the value of  $\hat{h}$ , for finite sample sizes it has an effect on the power of the test, being higher for larger values of  $\hat{h}$ . As expected from the results in Janssen (2000), no test has the largest power against all alternatives. The power increases with the sample size.

## 5 Conclusions and further research

As seen in the Introduction, shape functions have shown to have interesting properties. From a practical point of view, estimators of these functions need to be proposed and studied. This paper has focused on nonparametric estimators of the shape functions. The proposed estimators have been studied both theoretically and numerically. They



**Table 3** Values of  $S(u)$ , bias  $\times 10^3$  (bi) of  $S_n(u)$ , standard deviation  $\times 10^3$  (se) of  $S_n(u)$ , the mean of the standard deviation estimator  $\hat{\rho}_S(u)/\sqrt{(n)} \times 10^3$  (ms), and the coverage of the confidence interval (4) calculated at the nominal level 95% (co), calculated with  $h = sd \times n^{-\tau}$ , for  $u = 0.1, \dots, 0.9$ , and sample sizes  $n = 100, 250$

$n$	$\tau$	S(0.1)	S(0.2)	S(0.3)	S(0.4)	S(0.5)	S(0.6)	S(0.7)	S(0.8)	S(0.9)	
100	0.35	bi	0.088	-0.056	-0.371	0.036	0.182	0.115	-0.068	-0.024	-0.016
		se	86.938	63.030	40.205	23.151	23.957	16.556	9.545	4.175	1.208
		ms	90.263	67.138	44.149	26.466	26.028	17.792	10.571	5.009	1.655
		co	96.2	96.2	96.2	97.4	97.2	97.3	97.7	98.5	99.9
	0.40	bi	0.118	-0.027	-0.353	0.059	0.204	0.133	-0.061	-0.019	-0.013
		se	89.433	64.852	41.914	25.021	25.579	17.963	10.592	4.760	1.271
		ms	94.215	70.097	46.977	29.209	28.298	19.378	11.670	5.532	1.676
		co	96.4	96.7	96.7	97.9	97.4	97.4	97.8	98.3	99.7
	0.45	bi	0.165	0.011	-0.325	0.094	0.237	0.158	-0.05	-0.013	-0.010
		se	91.731	66.673	43.564	26.800	27.142	19.289	11.580	5.372	1.407
		ms	98.327	73.438	50.172	32.254	30.781	21.124	12.882	6.185	1.764
		co	96.8	96.9	97.4	98.4	97.8	97.7	98.0	98.3	99.4
0.49	bi	0.220	0.054	-0.296	0.127	0.268	0.185	-0.035	-0.003	-0.007	
	se	93.567	68.184	44.900	28.186	28.387	20.318	12.339	5.853	1.564	
	ms	101.821	76.480	53.057	34.968	32.937	22.649	13.941	6.790	1.891	
	co	97.1	97.1	97.7	98.6	98.1	97.9	98.1	98.4	99.1	

Table 3 continued

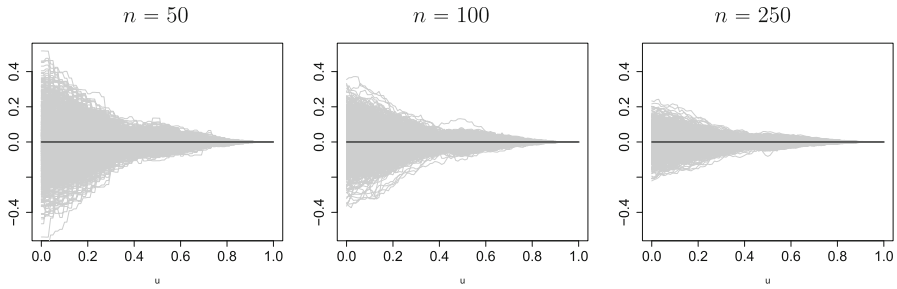
$n$	$\tau$	$S(0.1)$	$S(0.2)$	$S(0.3)$	$S(0.4)$	$S(0.5)$	$S(0.6)$	$S(0.7)$	$S(0.8)$	$S(0.9)$
250	0.35	0.136	-0.046	-0.039	0.148	0.233	0.158	0.063	0.016	-0.006
		55.593	39.989	26.104	15.633	15.961	11.037	6.641	3.137	0.782
		57.026	41.730	27.591	17.031	16.952	11.639	7.096	3.434	0.997
		95.8	95.8	96.0	96.7	96.6	96.3	96.8	97.3	99.1
	0.40	0.145	-0.040	-0.035	0.153	0.240	0.164	0.068	0.020	-0.005
		56.646	40.925	26.987	16.598	16.788	11.761	7.208	3.537	0.897
		58.661	43.167	28.987	18.448	18.129	12.497	7.720	3.801	1.076
		96.0	96.0	96.2	97.0	96.9	96.6	96.8	97.1	98.6
	0.45	0.186	-0.008	-0.009	0.177	0.262	0.183	0.084	0.029	-0.002
		57.690	41.841	27.828	17.462	17.550	12.399	7.704	3.886	1.049
		60.379	44.776	30.549	19.964	19.380	13.388	8.368	4.193	1.198
		96.3	96.3	96.5	97.3	97.3	97.0	97.1	97.1	98.1
	0.49	0.187	-0.012	-0.012	0.179	0.265	0.185	0.087	0.034	-0.001
		58.529	42.554	28.473	18.097	18.156	12.882	8.075	4.139	1.177
		61.940	46.249	31.959	21.309	20.479	14.153	8.923	4.530	1.321
		96.5	96.6	97.0	97.8	97.6	97.2	97.5	97.4	98.0

**Table 4** Values of  $S(u)$ , bias  $\times 10^3$  (bi) of  $S_n(u)$ , standard deviation  $\times 10^3$  (se) of  $S_n(u)$ , the mean of the standard deviation estimator  $\hat{\rho}_S(u)/\sqrt{(n)} \times 10^3$  (ms), and the coverage of the confidence interval (4) calculated at the nominal level 95% (co), calculated with  $h = sd \times n^{-\tau}$ , for  $u = 0.1, \dots, 0.9$ , and sample sizes  $n = 500, 1000$

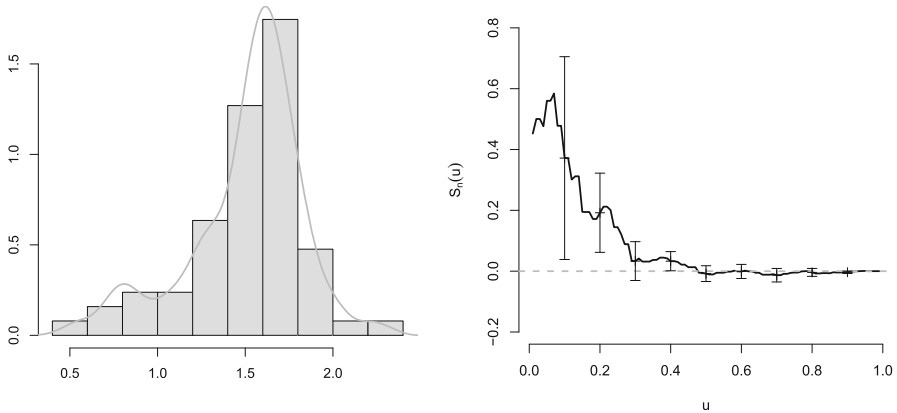
$n$	$\tau$	$S(0.1)$	$S(0.2)$	$S(0.3)$	$S(0.4)$	$S(0.5)$	$S(0.6)$	$S(0.7)$	$S(0.8)$	$S(0.9)$	
500	0.35	bi	0.379	0.330	0.105	0.094	0.043	0.025	-0.027	-0.008	-0.004
		se	39.139	28.375	18.595	11.485	11.66	8.055	4.973	2.395	0.607
		ms	40.121	29.289	19.474	12.156	12.127	8.392	5.196	2.574	0.735
		co	95.7	95.6	95.6	96.0	96.1	96.2	96.2	96.6	98.3
	0.40	bi	0.383	0.334	0.104	0.094	0.041	0.023	-0.029	-0.009	-0.004
		se	39.762	28.938	19.157	12.054	12.151	8.472	5.315	2.649	0.725
		ms	40.976	30.097	20.282	12.977	12.816	8.905	5.581	2.821	0.818
		co	95.8	95.8	95.9	96.5	96.5	96.4	96.4	96.5	97.6
0.45	bi	0.381	0.332	0.099	0.091	0.037	0.018	-0.031	-0.01	-0.003	
	se	40.365	29.454	19.655	12.557	12.598	8.840	5.605	2.855	0.841	
	ms	41.918	31.001	21.169	13.842	13.541	9.424	5.963	3.068	0.919	
	co	96.0	96.0	96.1	96.8	96.8	96.5	96.5	96.7	97.2	
	bi	0.385	0.336	0.099	0.092	0.037	0.015	-0.033	-0.011	-0.004	
	se	40.878	29.873	20.048	12.963	12.951	9.130	5.820	3.000	0.923	
	ms	42.799	31.838	21.986	14.619	14.183	9.875	6.286	3.274	1.007	
	co	96.0	96.2	96.2	97.2	97.0	96.8	96.7	97.1	97.3	

Table 4 continued

$n$	$\tau$	$S(0.1)$	$S(0.2)$	$S(0.3)$	$S(0.4)$	$S(0.5)$	$S(0.6)$	$S(0.7)$	$S(0.8)$	$S(0.9)$	
1000	0.35	0	0	0	0	0	0	0	0	0	
		bi	-0.338	-0.195	-0.134	-0.043	0.044	0.034	0.021	0.029	0.004
		se	27.886	20.216	13.261	8.376	8.419	5.835	3.658	1.852	0.494
		ms	28.274	20.640	13.760	8.681	8.659	6.035	3.779	1.909	0.556
		co	95.6	95.4	95.5	95.7	95.7	95.9	95.9	96.0	97.4
	0.40	bi	-0.336	-0.194	-0.132	-0.039	0.049	0.037	0.025	0.033	0.005
		se	28.223	20.522	13.567	8.707	8.710	6.089	3.864	2.009	0.585
		ms	28.747	21.095	14.220	9.160	9.053	6.335	4.010	2.065	0.624
		co	95.7	95.5	95.8	96.0	95.9	96.1	96.0	96.1	96.7
	0.45	bi	-0.336	-0.192	-0.131	-0.037	0.052	0.040	0.028	0.034	0.006
		se	28.536	20.790	13.825	8.984	8.956	6.301	4.031	2.129	0.660
		ms	29.269	21.596	14.708	9.650	9.457	6.630	4.232	2.213	0.693
		co	95.8	95.6	96.0	96.3	96.2	96.3	96.2	96.3	96.6
	0.49	bi	-0.334	-0.190	-0.129	-0.034	0.056	0.043	0.030	0.036	0.006
		se	28.792	21.004	14.026	9.191	9.150	6.461	4.153	2.214	0.710
		ms	29.762	22.065	15.155	10.084	9.817	6.883	4.418	2.331	0.749
		co	96.0	95.8	96.2	96.6	96.4	96.5	96.4	96.6	96.8



**Fig. 2** Graphs of 1000 generated estimations of calculated with  $h = sd \times n^{-0.45}$  in grey joint with the population function  $S(u)$  in black



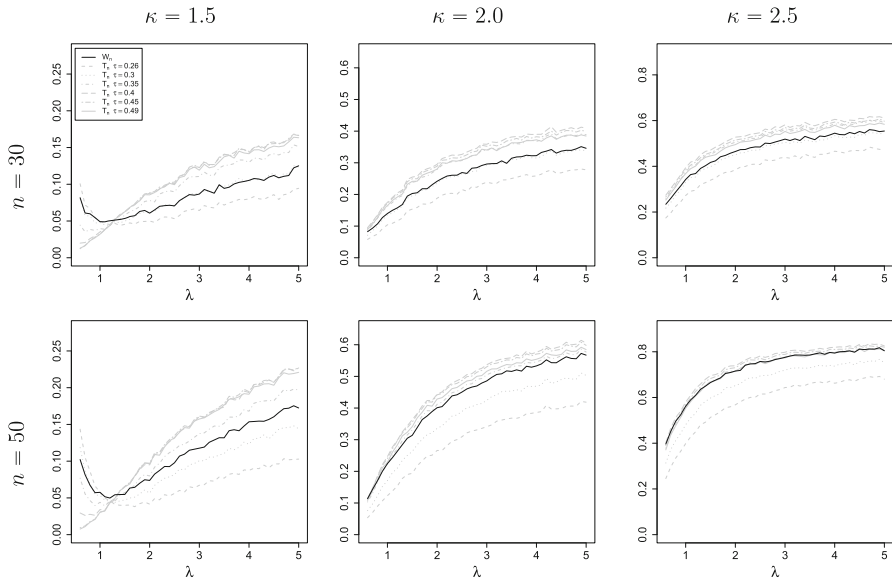
**Fig. 3** Left: histogram and kernel density estimator for Glass fibre breaking strengths data. Right: graph of  $S_n(u), u \in (0, 1)$ , and confidence intervals at the nominal level 90% calculated with  $\hat{h} = sd \times n^{-0.45}$  for  $u = 0.1, \dots, 0.9$

**Table 5** Confidence intervals (4) for  $S(u)$  at nominal level 90%, calculated with  $\hat{h} = sd \times n^{-0.45}$ , for  $u = 0.1, \dots, 0.9$ , for the Glass fibre breaking strengths data

$u$	Lower limit	Upper limit	$u$	Lower limit	Upper limit
0.1	0.0384	0.7052	0.6	-0.0241	0.0223
0.2	0.0619	0.3223	0.7	-0.0355	0.0089
0.3	-0.0306	0.0970	0.8	-0.0170	0.0090
0.4	0.0013	0.0644	0.9	-0.0076	0.0006
0.5	-0.0340	0.0178			

**Table 6** Critical points for  $n = 30, 50$  with  $h = sd \times n^{-\tau}$  and  $\alpha = 0.05$

	$w_{n,\alpha}$	$t_{n,\alpha}$						
		$\tau =$	0.26	0.30	0.35	0.40	0.45	0.49
$n = 30$	2.6839		0.1590	0.1606	0.1736	0.2046	0.2610	0.3213
$n = 50$	2.7199		0.1832	0.1800	0.1876	0.2178	0.2778	0.3438



**Fig. 4** Empirical power based on 10,000 samples with size  $n = 30, 50$  from a log-Lindley distribution

exhibit nice asymptotic properties and the numerical experiments show a reasonable practical behavior. Optimal choice of the smoothing parameter involved in the construction of the estimator has not been dealt with in this piece of research. This issue deserves more investigation and will be considered in future studies.

### 6 Proofs

This section sketches the proofs of the results stated in the previous sections. Along this section,  $M$  is a generic positive constant taking many different values throughout the proofs and  $f_n(x)$  is defined as  $\hat{f}_n(x)$  with  $\hat{h} = \hat{\sigma} \times g(n)$  replaced with  $h = \sigma \times g(n)$ .

**Lemma 1** *Suppose that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , that  $K$  satisfies Assumption 1 (i) and that  $h$  satisfies Assumption 2 (i). Then  $\sup_x |\hat{f}_n(x) - f_n(x)| \xrightarrow{a.s.} 0$ .*

**Proof** We have that

$$\hat{f}_n(x) - f_n(x) = \delta_1(x) + \delta_2(x), \tag{5}$$

with

$$\delta_1(x) = \left( \frac{1}{n\hat{h}} - \frac{1}{nh} \right) \sum_{j=1}^n K \left( \frac{X_j - x}{h} \right), \tag{6}$$

$$\delta_2(x) = \frac{1}{n\hat{h}} \sum_{j=1}^n \left\{ K \left( \frac{X_j - X_i}{\hat{h}} \right) - K \left( \frac{X_j - x}{h} \right) \right\}. \tag{7}$$

Since

$$\delta_1(x) = \frac{\sigma - \hat{\sigma}}{\hat{\sigma}} \frac{1}{nh} \sum_{j=1}^n K \left( \frac{X_j - x}{h} \right) = \frac{\sigma - \hat{\sigma}}{\hat{\sigma}} f_n(x),$$

we can write

$$|\delta_1(x)| \leq \left| \frac{\sigma - \hat{\sigma}}{\hat{\sigma}} \right| \left\{ \sup_x |f_n(x) - \mathbb{E}\{f_n(x)\}| + \sup_x \mathbb{E}\{f_n(x)\} \right\}. \tag{8}$$

Recall that if  $K$  is Lipschitz continuous and has compact support, then it has bounded variation. Under the assumptions made on  $K$  and  $h$  we have that (see, for example, the proof of Theorem 2.1.3 of Prakasa Rao 1983)

$$\sup_x |f_n(x) - \mathbb{E}\{f_n(x)\}| \xrightarrow{a.s.} 0. \tag{9}$$

Since  $f$  is bounded,

$$\mathbb{E}\{f_n(x)\} = \int_{-\infty}^{\infty} K(u) f(x - hu) du \leq M \int_{-\infty}^{\infty} K(u) du = M, \quad \forall x. \tag{10}$$

Using (8), (9), (10) and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , it follows that

$$\sup_x |\delta_1(x)| \xrightarrow{a.s.} 0. \tag{11}$$

Now we study  $\delta_2(x)$ . Since  $K$  is Lipschitz continuous and has compact support,  $\text{Sup}(K)$ , we have that

$$\begin{aligned} D(x) &= \frac{1}{n\hat{h}} \sum_{j=1}^n \left| K \left( \frac{X_j - x}{\hat{h}} \right) - K \left( \frac{X_j - x}{h} \right) \right| \\ &\leq M \frac{\sigma}{\hat{\sigma}} \left| \frac{\sigma}{\hat{\sigma}} - 1 \right| \frac{1}{nh} \sum_{j=1}^n \left| \frac{X_j - x}{h} \right| \mathbf{1}\{(X_j - x)/h \in \text{Sup}(K) \text{ or } (X_j - x)/\hat{h} \in \text{Sup}(K)\}. \end{aligned}$$

Since  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , it follows that for  $n$  large enough,  $\mathbf{1}\{(X_j - x)/h \in \text{Sup}(K) \text{ or } (X_j - x)/\hat{h} \in \text{Sup}(K)\} \subseteq \mathbf{1}\{|X_j - x|/h \leq C\}$ , for certain positive constant  $C$ . Therefore

$$D(x) \leq M \frac{\sigma}{\hat{\sigma}} \left| \frac{\sigma}{\hat{\sigma}} - 1 \right| f_{U,n}(x), \tag{12}$$

where  $f_{U,n}(x)$  is the kernel estimator of  $f$  built by using as kernel the probability density function of the uniform law on the interval  $[-C, C]$ . Proceeding as before, we have that

$$\sup_x f_{U,n}(x) \leq M \text{ a.s.}, \tag{13}$$

Using that  $|\delta_2(x)| \leq D(x)$ , (12), (13) and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , it follows that

$$\sup_x |\delta_2(x)| \xrightarrow{a.s.} 0. \tag{14}$$

The result follows from (5), (11) and (14). □

**Remark 5** From the previous proof, notice that if in the statement of Lemma 1 the assumption  $\hat{\sigma} \xrightarrow{a.s.} \sigma$  is replaced with  $\hat{\sigma} \xrightarrow{P} \sigma$  then it is concluded that  $\sup_x |\hat{f}_n(x) - f_n(x)| \xrightarrow{P} 0$ .

**Proof of Theorem 1** We have that

$$R_n(u) = R_{1n}(u) + R_{2n}(u) + R_{3n}(u) + R_{4n}(u), \tag{15}$$

where

$$\begin{aligned} R_{1n}(u) &= \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ f(X_i), \\ R_{2n}(u) &= \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \{f_n(X_i) - f(X_i)\}, \\ R_{3n}(u) &= \frac{1}{n} \sum_{i=1}^n \left[ \{X_i - F_n^{-1}(u)\}^+ - \{X_i - F^{-1}(u)\}^+ \right] \hat{f}_n(X_i), \\ R_{4n}(u) &= \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \left\{ \hat{f}_n(X_i) - f_n(X_i) \right\}. \end{aligned}$$

From the SLLN, it follows that

$$R_{1n}(u) \xrightarrow{a.s.} R(u). \tag{16}$$

As for  $R_{2n}(u)$ , we have that

$$|R_{2n}(u)| \leq \sup_x |f_n(x) - f(x)| \left\{ \frac{1}{n} \sum_{i=1}^n |X_i| + |F^{-1}(u)| \right\}.$$



Under the assumptions made (see, for example, Theorem 2.1.3 in Prakasa Rao 1983)

$$\sup_x |f_n(x) - f(x)| \xrightarrow{a.s.} 0. \tag{17}$$

Taking into account that  $\frac{1}{n} \sum_{i=1}^n |X_i| \xrightarrow{a.s.} \mathbb{E}|X| < \infty$ , it follows that

$$R_{2n}(u) \xrightarrow{a.s.} 0. \tag{18}$$

In order to study  $R_{3n}(u)$  we first observe that

$$\begin{aligned} & \left\{ X_i - F_n^{-1}(u) \right\}^+ - \left\{ X_i - F^{-1}(u) \right\}^+ \\ &= \begin{cases} 0 & \text{if } X_i \leq \min\{F_n^{-1}(u), F^{-1}(u)\}, \\ X_i - F_n^{-1}(u) & \text{if } F_n^{-1}(u) < X_i \leq F^{-1}(u), \\ -X_i + F^{-1}(u) & \text{if } F^{-1}(u) < X_i \leq F_n^{-1}(u), \\ F^{-1}(u) - F_n^{-1}(u) & \text{if } X_i > \max\{F_n^{-1}(u), F^{-1}(u)\}, \end{cases} \end{aligned} \tag{19}$$

which implies

$$\left| \left\{ X_i - F_n^{-1}(u) \right\}^+ - \left\{ X_i - F^{-1}(u) \right\}^+ \right| \leq \left| F^{-1}(u) - F_n^{-1}(u) \right|, \quad \forall i,$$

and therefore

$$|R_{3n}(u)| \leq |F_n^{-1}(u) - F^{-1}(u)| \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i).$$

Under the assumptions made  $|F_n^{-1}(u) - F^{-1}(u)| \xrightarrow{a.s.} 0$  (see, for example display (1.4.9) in Csörgő 1983). We also have that

$$\frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i) \leq \frac{1}{n} \sum_{i=1}^n f(X_i) + \sup_x |f_n(x) - f(x)| + \sup_x |\hat{f}_n(x) - f_n(x)|. \tag{20}$$

From (17), Lemma 1 and taking into account that  $f$  is bounded, it follows that the right hand-side of inequality (20) is bounded a.s. Therefore,

$$R_{3n}(u) \xrightarrow{a.s.} 0. \tag{21}$$

We have that

$$R_{4n}(u) = R_{41n}(u) + R_{42n}(u), \tag{22}$$

where

$$R_{41n}(u) = \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \delta_1(X_i)$$

and

$$R_{42n}(u) = \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \delta_2(X_i),$$

where  $\delta_1(x)$  and  $\delta_2(x)$  are as defined in (6) and (7), respectively. Since

$$R_{41n}(u) = (\sigma - \hat{\sigma})\{R_{1n}(u) + R_{2n}(u)\}/\hat{\sigma} \tag{23}$$

and  $\hat{\sigma} \xrightarrow{a.s.} \sigma$ , from (16) and (18), it follows that

$$R_{41n}(u) \xrightarrow{a.s.} 0. \tag{24}$$

Using (12), we get that

$$|R_{42n}(u)| \leq M \frac{\sigma}{\hat{\sigma}} \left| \frac{\sigma}{\hat{\sigma}} - 1 \right| R_{U,n}(u), \tag{25}$$

where

$$R_{U,n}(u) = \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ f_{U,n}(X_i).$$

Proceeding as before, it can be seen that  $R_{U,n}(u) \xrightarrow{a.s.} R(u)$ , and hence

$$R_{42n}(u) \xrightarrow{a.s.} 0. \tag{26}$$

The result follows from (15), (16), (18), (21), (22), (24) and (26). □

**Proof of Theorem 2** Let us consider decomposition (15). We have that  $R_{1n}(u) + R_{2n}(u) = T_{1n}(u) + T_{2n}(u)$  with

$$T_{1n}(u) = K(0) \frac{1}{n^2 h} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+,$$

and  $T_{2n}(u) = \frac{n-1}{n} U_n(u)$ , where  $U_n(u) = \frac{1}{n(n-1)} \sum_{i \neq j} H_n(X_i, X_j; u)$  is a degree two U-statistic with symmetric kernel

$$H_n(X_i, X_j; u) = \frac{1}{2} \left[ \{X_i - F^{-1}(u)\}^+ + \{X_j - F^{-1}(u)\}^+ \right] \frac{1}{h} K \left( \frac{X_i - X_j}{h} \right).$$

We first see that

$$T_{1n}(u) \xrightarrow{a.s.} 0. \tag{27}$$

Notice that under the assumptions made,

$$\mathbb{E} \left[ \{X - F^{-1}(u)\}^+ \right] < \infty, \tag{28}$$

therefore, by the SLLN,

$$\frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \xrightarrow{a.s.} \mathbb{E} \left[ \{X - F^{-1}(u)\}^+ \right] < \infty.$$

Finally, since  $K$  is bounded and  $nh \rightarrow \infty$ , (27) follows.

Next we will see that

$$T_{2n}(u) \xrightarrow{a.s.} R(u). \tag{29}$$

With this aim, we first calculate  $\mathbb{E}\{U_n(u)\}$ .

$$\mathbb{E}\{U_n(u)\} = \mathbb{E}\{H_n(X_1, X_2; u)\} = \int \{x - F^{-1}(u)\}^+ \left( \int K(y) f(x - hy) dy \right) f(x) dx.$$

Since  $\int K(y) f(x - hy) dy \rightarrow f(x)$ ,  $\{x - F^{-1}(u)\}^+ \left( \int K(y) f(x - hy) dy \right) f(x) \leq M \{x - F^{-1}(u)\}^+$  and (28), by dominated convergence theorem we have that  $\mathbb{E}\{U_n(u)\} \rightarrow R(u)$ .

Now, let  $A_n(x; u) = \mathbb{E}\{H_n(x, X; u)\}$  and  $\tau = \int x^2 K(x) dx$ . Routine calculations show that

$$\begin{aligned} A_n(x; u) &= A(x; u) + a_n(x; u), \\ A(x; u) &= \{x - F^{-1}(u)\}^+ f(x), \\ a_n(x; u) &= 0.5\tau h^2 f''(x) \{x - F^{-1}(u)\}^+ + 0.5\tau h^2 f'(x) + o(h^2). \end{aligned}$$

Let  $\varepsilon > 0$ . From the stated assumptions  $\mathbb{E}\{a_n(X, u)^2\} = O(h^4)$ , therefore

$$P \left( \frac{1}{n} \sum_{i=1}^n a_n(X_i; u) > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathbb{E} \left[ \left\{ \frac{1}{n} \sum_{i=1}^n a_n(X_i; u) \right\}^2 \right] = O(h^4).$$

Since  $nh^4 \rightarrow 0$ , it follows that

$$\sum_{n \geq 1} P \left( \frac{1}{n} \sum_{i=1}^n a_n(X_i; u) > \varepsilon \right) < \infty,$$

which implies that

$$\frac{1}{n} \sum_{i=1}^n a_n(X_i; u) \xrightarrow{a.s.} 0.$$

Let  $B_n(u) = U_n(u) + \mathbb{E}\{H_n(X_1, X_2; u)\} - \frac{2}{n} \sum_{i=1}^n A_n(X_i; u)$ . It can be seen that  $\mathbb{E}\{B_n(u)^2\} = O(1/n^2 h)$ . Reasoning as before, we get that

$$B_n(u) \xrightarrow{a.s.} 0.$$

Summarizing,

$$T_{2n}(u) = \frac{2}{n} \sum_{i=1}^n A(X_i; u) - R(u) + t_{2n}(u),$$

with  $t_{2n}(u) \xrightarrow{a.s.} 0$ . By the SLLN we have that  $\frac{1}{n} \sum_{i=1}^n A(X_i; u) \xrightarrow{a.s.} R(u)$ , and thus (29) is proven.

Finally, proceeding as in the proof of Theorem 1, one gets that  $R_{in}(u) \xrightarrow{a.s.} 0$ ,  $i = 3, 4$ . This completes the proof.  $\square$

**Lemma 2** *Let  $f$  be a probability density function with finite support,  $(a, b)$ , continuous in  $(a, b)$ . Suppose also that  $K$  satisfies Assumption 1 (ii) and that  $h$  satisfies Assumption 2 (ii). Then  $\sup_{x \in (a,b)} |f_n(x) - f(x)| \xrightarrow{a.s.} 0$ .*

**Proof** From the proof of Theorem 2.1.3 in Prakasa Rao (1983), we have that

$$\sup_x |f_n(x) - \mathbb{E}\{f_n(x)\}| \xrightarrow{a.s.} 0.$$

So, it suffices to see that

$$\sup_{x \in (a,b)} |\mathbb{E}\{f_n(x)\} - f(x)| \rightarrow 0. \quad (30)$$

We have that

$$\mathbb{E}\{f_n(x)\} = \int K(y) f(x - hy) dy.$$

From the continuity of  $f$ , for each fixed  $y \in \mathbb{R}$

$$f(x + hy) \rightarrow f(x), \quad \forall x \in (a, b),$$

and

$$K(y) f(x + hy) \leq MK(y).$$

Thus, (30) holds true by applying dominated convergence theorem. □

**Proof of Theorem 3** First of all, we see that, under the assumptions made,  $R$  is a continuous function on  $[0, 1]$ . Since the function

$$G : \mathbb{R} \times (0, 1) \mapsto \mathbb{R}$$

$$(x, u) \mapsto G(x, u) = \{x - F^{-1}(u)\}^+ f^2(x)$$

is continuous, it follows that  $R(u) = \int G(x, u)dx$  is continuous on  $(0, 1)$ . Recall that  $\lim_{u \rightarrow 1^-} R(u) = 0 = R(1)$  and, if the limit exists,  $R(0) = \lim_{u \rightarrow 0^+} R(u)$ . Thus,  $R$  is continuous at  $u = 1$ . To see that it is also continuous at  $u = 0$  it suffices to see that the limit  $\lim_{u \rightarrow 0^+} R(u)$  exists, which is true since

$$\lim_{u \rightarrow 0^+} \{x - F^{-1}(u)\}^+ f^2(x) = (x - a) f^2(x),$$

$$\{x - F^{-1}(u)\}^+ f^2(x) \leq c(x) = M(x - a) \mathbf{1}\{a \leq x \leq b\}$$

and  $\int c(x)dx < \infty$ , then by dominated convergence theorem

$$\lim_{u \rightarrow 0^+} R(u) = \int \{x - a\}^+ f^2(x)dx < \infty.$$

Next, we consider decomposition (15) and study each term on the right hand-side of such expression. Since  $R$  is a continuous function on  $[0, 1]$ , the point by point convergence of  $R_{1n}(u)$  to  $R(u)$  implies the uniform convergence on  $[0, 1]$ , that is

$$\sup_{0 \leq u \leq 1} |R_{1n}(u) - R(u)| \xrightarrow{a.s.} 0. \tag{31}$$

As for  $R_{2n}(u)$ , taking into account that  $\frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \leq \bar{X} - a$ , with  $\bar{X} = (1/n) \sum_{i=1}^n X_i$ , we have that

$$\sup_{0 \leq u \leq 1} |R_{2n}(u)| \leq \sup_{x \in [a, b]} |f_n(x) - f(x)|(\bar{X} - a).$$

From Lemma 2, taking into account that Assumption 1 (i) implies Assumption 1 (ii) and that  $\bar{X} - a \xrightarrow{a.s.} \mathbb{E}(X) - a < \infty$ , we conclude that

$$\sup_{0 \leq u \leq 1} |R_{2n}(u)| \xrightarrow{a.s.} 0. \tag{32}$$

From (31) and (32), we conclude that

$$\sup_{0 \leq u \leq 1} |R_{1n}(u) + R_{2n}(u) - R(u)| \xrightarrow{a.s.} 0. \tag{33}$$

For  $R_{3n}(u)$  we have that

$$\sup_{0 \leq u \leq 1} |R_{3n}(u)| \leq \sup_{0 \leq u \leq 1} |F_n^{-1}(u) - F^{-1}(u)| \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i).$$

In the proof of Theorem 1 we saw that  $(1/n) \sum_{i=1}^n \hat{f}_n(X_i)$  is bounded a.s. Under the assumptions made (see, for example, p. 6 of Csörgő 1983),

$$\sup_{0 \leq u \leq 1} |F_n^{-1}(u) - F^{-1}(u)| \xrightarrow{a.s.} 0.$$

Therefore

$$\sup_{0 \leq u \leq 1} |R_{3n}(u)| \xrightarrow{a.s.} 0.$$

Finally, taking into account decomposition (22), it suffices to show that  $\sup_{0 \leq u \leq 1} |R_{4in}(u)| \xrightarrow{a.s.} 0$ ,  $i = 1, 2$ . Using (23), (33) and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , one gets that

$$\sup_{0 \leq u \leq 1} |R_{41n}(u)| \xrightarrow{a.s.} 0.$$

Using (25), (33) and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , one similarly gets that

$$\sup_{0 \leq u \leq 1} |R_{42n}(u)| \xrightarrow{a.s.} 0.$$

This concludes the proof. □

**Proof of Theorem 4** First of all, we see that, under the assumptions made,  $R \in L^2$ . We have that

$$\begin{aligned} \int_0^1 R^2(u)du &= \int_0^1 \left( \int_{F^{-1}(u)}^\infty \{x - F^{-1}(u)\} f^2(x)dx \right) \\ &\quad \times \left( \int_{F^{-1}(u)}^\infty \{y - F^{-1}(u)\} f^2(y)dy \right) du = I_1 - 2I_2 + I_3, \end{aligned}$$

with (recall that  $f$  is a bounded function and that  $\int_0^1 F^{-1}(u)^2 du = \mathbb{E}(X^2)$ )

$$I_1 = \int_0^1 \left( \int_{F^{-1}(u)}^\infty x f^2(x)dx \right)^2 du \leq M \mathbb{E}(X^2) < \infty,$$

$$\begin{aligned} I_2 &= \int_0^1 F^{-1}(u) \int_{F^{-1}(u)}^\infty x f^2(x)dx du \leq M \left( \int_0^1 F^{-1}(u)^2 du \int x^2 f(x)dx \right)^{1/2} \\ &= M \mathbb{E}(X^2) < \infty, \end{aligned}$$

$$I_3 = \int_0^1 F^{-1}(u)^2 \left( \int_{F^{-1}(u)}^{\infty} f^2(x) dx \right)^2 du \leq M \int_0^1 F^{-1}(u)^2 du = M \mathbb{E}(X^2) < \infty,$$

and thus  $R \in L^2$ .

Next, we consider decomposition (15) and study each term on the right hand-side of such expression. Since  $R_{1n}$  is an average of integrable i.i.d. random elements whose expectation is  $R(u)$ , applying the SLLN in Hilbert spaces we obtain that (see, for example Theorem 2.4 of Bosq 2000),

$$R_{1n} \xrightarrow{a.s.} R,$$

in  $L^2$ . As for  $R_{2n}(u)$ , we have that

$$|R_{2n}(u)| \leq \sup_x |f_n(x) - f(x)| \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+.$$

Let  $V(u) = \mathbb{E}[\{X_i - F^{-1}(u)\}^+]$ . A parallel reasoning to that used to prove that  $R \in L^2$  shows that  $V \in L^2$ . Now, from the SLLN in Hilbert spaces it follows that  $\frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \xrightarrow{a.s.} V$ , in  $L^2$ . This fact and (17) gives

$$R_{2n} \xrightarrow{a.s.} 0,$$

in  $L^2$ . From (15), (31) and (32), we conclude that

$$\|R_{1n} + R_{2n} - R\| \xrightarrow{a.s.} 0. \tag{34}$$

For  $R_{3n}(u)$  we have that

$$|R_{3n}(u)| \leq |F_n^{-1}(u) - F^{-1}(u)| \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i).$$

As shown in the proof of Theorem 1, the second factor on the right hand-side of the above inequality is a.s. bounded. Since  $\mathbb{E}(X^2) < \infty$ , from Lemma 8.3 in Bickel and Freedman (1981), it follows that  $\|F_n^{-1} - F^{-1}\| \xrightarrow{a.s.} 0$ , and thus  $R_{3n} \xrightarrow{a.s.} 0$ , in  $L^2$ . Finally, taking into account decomposition (22), using (23), (25) (34) and that  $\hat{\sigma} \xrightarrow{a.s.} \sigma > 0$ , one also gets that  $R_{4n} \xrightarrow{a.s.} 0$ , in  $L^2$ . This concludes the proof.  $\square$

**Proof of Theorem 5** Similar developments to those made in the proof of Theorem 2, and sharing the notation used there, show that

$$R_{1n}(u) + R_{2n}(u) = \frac{2}{n} \sum_{i=1}^n A(X_i; u) - R(u) + t_{2n}(u),$$

with  $t_{2n} \xrightarrow{a.s.} 0$  in  $L^2(w)$ . By the SLLN in Hilbert spaces we have that  $\frac{1}{n} \sum_{i=1}^n A(X_i; \cdot) \xrightarrow{a.s.} R$ , in  $L^2(w)$ . Finally, proceeding as in the proof of Theorem 4 we get that  $\|R_{in}\|_w \xrightarrow{a.s.} 0$ ,  $i = 3, 4$ , which completes the proof.  $\square$

**Proof of Theorem 6** Similar developments to those made in the proof of Theorem 2, and sharing the notation used there, show that

$$\sqrt{n} \{R_{1n}(u) + R_{2n}(u) - R(u)\} = \frac{2}{\sqrt{n}} \sum_{i=1}^n [\{X_i - F^{-1}(u)\}^+ f(X_i) - R(u)] + o_P(1). \tag{35}$$

Now, taking into account (19) we can write

$$\sqrt{n}R_{3n}(u) = T_{1n}(u) + T_{2n}(u) + T_{3n}(u), \tag{36}$$

where

$$\begin{aligned} T_{1n}(u) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - F_n^{-1}(u)\} \hat{f}_n(X_i) \mathbf{1}\{F_n^{-1}(u) < X_i \leq F^{-1}(u)\}, \\ T_{2n}(u) &= -\frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - F_n^{-1}(u)\} \hat{f}_n(X_i) \mathbf{1}\{F^{-1}(u) < X_i \leq F_n^{-1}(u)\}, \\ T_{3n}(u) &= \sqrt{n} \{F^{-1}(u) - F_n^{-1}(u)\} \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i) \mathbf{1}\{F^{-1}(u) < X_i\}. \end{aligned}$$

We have that

$$0 \leq T_{1n}(u) \leq \left| \sqrt{n} \{F^{-1}(u) - F_n^{-1}(u)\} \right| \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i) \mathbf{1}\{F_n^{-1}(u) < X_i \leq F^{-1}(u)\}.$$

Under the assumptions made,  $\sqrt{n} \{F_n^{-1}(u) - F^{-1}(u)\} = O_P(1)$  and  $F_n^{-1}(u) \xrightarrow{a.s.} F^{-1}(u)$ , which implies that for each  $\varepsilon > 0$  there exists  $n_0$  such that  $\mathbf{1}\{F_n^{-1}(u) < X_i \leq F^{-1}(u)\} \leq \mathbf{1}\{F^{-1}(u) - \varepsilon < X_i \leq F^{-1}(u)\}$ ,  $\forall n \geq n_0$ , and hence  $\mathbb{E}[\mathbf{1}\{F_n^{-1}(u) < X_i \leq F^{-1}(u)\}] = O(\varepsilon)$ . Proceeding as in the proof of Theorem 1, it can be seen that  $(1/n) \sum_{i=1}^n \hat{f}_n(X_i) = O_P(1)$ . Therefore  $T_{1n}(u) \xrightarrow{P} 0$ . Analogously, it can be seen that  $T_{2n}(u) \xrightarrow{P} 0$  and that  $\frac{1}{n} \sum_{i=1}^n f_n(X_i) \mathbf{1}\{F^{-1}(u) < X_i \leq F_n^{-1}(u)\} = o_P(1)$ . Now, proceeding as in the proof of Theorem 2, it can be seen that

$$\frac{1}{n} \sum_{i=1}^n f_n(X_i) \mathbf{1}\{F^{-1}(u) < X_i\} \xrightarrow{P} \mu(u) = \mathbb{E} \left[ f(X) \mathbf{1}\{F^{-1}(u) < X\} \right]. \tag{37}$$



From Lemma 1,

$$\left| \frac{1}{n} \sum_{i=1}^n \left\{ \hat{f}_n(X_i) - f_n(X_i) \right\} \mathbf{1}\{F^{-1}(u) < X_i\} \right| \leq \sup_x |\hat{f}_n(x) - f_n(x)| \xrightarrow{a.s.} 0.$$

We also have that (see, e.g. Theorem 2.5.2 in Serfling 1980)

$$\sqrt{n} \left\{ F^{-1}(u) - F_n^{-1}(u) \right\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} + o_P(1).$$

Thus, by Slutsky theorem,

$$T_{3n}(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mathbb{E} \left[ f(X) \mathbf{1}\{F^{-1}(u) < X\} \right] + o_P(1).$$

Therefore, it has been shown that

$$\sqrt{n} \{R_{1n}(u) + R_{2n}(u) + R_{3n}(u) - R(u)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(u) + o_P(1). \tag{38}$$

To prove the result, it remains to see that  $\sqrt{n}R_{4n}(u) = o_P(1)$ . Recall decomposition (22). From (23) and (35), it follows that

$$\sqrt{n}R_{41n}(u) = -\sqrt{n}\{\hat{\sigma} - \sigma\}R(u)/\sigma + o_P(1). \tag{39}$$

Now we study  $R_{42n}(u)$ . A Taylor expansion of  $K((X_j - X_i)/\hat{h})$  around  $K((X_j - X_i)/h)$  gives,

$$\begin{aligned} \sqrt{n}R_{42n}(u) &= \frac{\sigma}{\hat{\sigma}^2} \sqrt{n}(\sigma - \hat{\sigma}) \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \\ &\quad \times \frac{1}{nh} \sum_{j=1}^n \frac{X_j - X_i}{h} K' \left( \frac{X_j - X_i}{h} \right) + Q_n(u), \end{aligned}$$

where

$$\begin{aligned} Q_n(u) &= \frac{1}{2\sqrt{n}} \frac{\sigma}{\hat{\sigma}^3} \left\{ \sqrt{n}(\sigma - \hat{\sigma}) \right\}^2 \frac{1}{n} \sum_{i=1}^n \{X_i - F^{-1}(u)\}^+ \\ &\quad \times \frac{1}{nh} \sum_{j=1}^n \left( \frac{X_j - X_i}{h} \right)^2 K'' \left( \frac{X_j - X_i}{\tilde{h}} \right), \end{aligned}$$

where  $\tilde{h} = \alpha h + (1 - \alpha)\hat{h}$ , for some  $\alpha \in (0, 1)$ . The assumptions made imply that  $Q_n(u) = o_P(1)$ . Now proceeding as in the proof of Theorem 2, and taking into account that  $\int u K'(u) f(x + hu) du \rightarrow f(x) \int u K'(u) du = -f(x)$ , we obtain that

$$\sqrt{n}R_{42n}(u) = \sqrt{n}(\hat{\sigma} - \sigma)R(u)/\sigma + o_P(1). \tag{40}$$

Finally, the result follows from (38), (39) and (40). □

**Proof of Theorem 7** Similar developments to those made in the proof of Theorem 2, and sharing the notation used there, show that

$$\begin{aligned} \sqrt{n} \{R_{1n}(u) + R_{2n}(u) - R(u)\} &= \frac{2}{\sqrt{n}} \sum_{i=1}^n \left[ \{X_i - F^{-1}(u)\}^+ f(X_i) - R(u) \right] \\ &\quad + r_{1n}(u), \quad u \in (0, 1), \end{aligned}$$

with  $\int_0^1 r_{1n}(u)^2 w(u) du \xrightarrow{P} 0$ .

As for  $R_{3n}(u)$ , we consider decomposition (36). In the proof of Theorem 6 it was shown that  $T_{1n}(u) \xrightarrow{P} 0$ , for each  $u \in (0, 1)$ . We have that

$$0 \leq T_{1n}(u) \leq \left| \sqrt{n} \{F_n^{-1}(u) - F^{-1}(u)\} \right| \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i).$$

From Theorems 2.1, 3.1.1 and 3.2.1 in Csörgő (1983), it follows that

$$n \int_0^{(n-1)/n} \{F_n^{-1}(u) - F^{-1}(u)\}^2 w(u) du = O_P(1).$$

Under the assumptions made,

$$0 \leq \frac{1}{n} \sum_{i=1}^n \hat{f}_n(X_i) \xrightarrow{P} \mathbb{E}\{f(X)\} < \infty.$$

Thus, by dominated convergence theorem, it follows that

$$\int_0^{(n-1)/n} T_{1n}(u)^2 w(u) du \xrightarrow{P} 0.$$

Analogously it can be seen that

$$\int_0^{(n-1)/n} T_{2n}(u)^2 w(u) du \xrightarrow{P} 0.$$

From Theorems 2.1 and 3.2.1 in Csörgő (1983), it follows that

$$\sup_{0 < u \leq \frac{n-1}{n}} \left| \sqrt{n} \{F_n^{-1}(u) - F^{-1}(u)\} f(F^{-1}(u)) - \frac{1}{\sqrt{n}} \sum_{i=1}^n [\mathbf{1}\{X_i \leq F^{-1}(u)\} - u] \right| \xrightarrow{a.s.} 0.$$

Notice that the convergence in (37) holds for each  $u \in [0, 1]$ . Since  $\mu(u)$  is a continuous function, it follows that such convergence is uniform on the interval  $[0, 1]$ . Therefore,

$$\int_0^{(n-1)/n} \left\{ T_{3n}(u) - \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mu(u) \right\}^2 w(u) du \xrightarrow{P} 0.$$

Summarizing, it has been shown that

$$\begin{aligned} \sqrt{n} \{R_{1n}(u) + R_{2n}(u) + R_{3n}(u) - R(u)\} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i(u) + r_{2n}(u), \\ \int_0^{(n-1)/n} r_{2n}(u)^2 w(u) du &\xrightarrow{P} 0. \end{aligned}$$

The same steps given in the proof of Theorem 6 to show that  $\sqrt{n}R_{4n}(u) = o_P(1)$  can be used to see that  $\int_0^{(n-1)/n} nR_{4n}(u)^2 w(u) du \xrightarrow{P} 0$ . This completes the proof.  $\square$

**Proof of Corollary 2** Define

$$\tilde{Y}_i(u) = \begin{cases} Y_i(u) & \text{if } u \in \mathbb{I}_n = (0, (n-1)/n], \\ 0 & \text{if } u \in ((n-1)/n, 1], \end{cases}$$

$1 \leq i \leq n$ , and  $W_n(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \tilde{Y}_i(u)$ ,  $u \in [0, 1]$ .

From Theorem 7, we have that

$$\begin{aligned} n \int_0^1 \{R_n(u) - R(u)\}^2 w(u) du &= \int_0^1 W_n(u)^2 w(u) du \\ &+ n \int_{(n-1)/n}^1 R(u)^2 w(u) du + o_P(1). \end{aligned} \tag{41}$$

We first see that  $\tilde{Y}_1 \in L^2(w)$ . With this aim we write  $\tilde{Y}_1(u) = \tilde{Y}_{11}(u) - \tilde{Y}_{12}(u) + \tilde{Y}_{13}(u)$ , with

$$\tilde{Y}_{11}(u) = 2\{X_1 - F^{-1}(u)\}^+ f(X_1) \mathbf{1}(u \in \mathbb{I}_n), \quad \tilde{Y}_{12}(u) = 2R(u) \mathbf{1}(u \in \mathbb{I}_n),$$

and

$$\tilde{Y}_{13}(u) = \frac{\mathbf{1}\{X_i \leq F^{-1}(u)\} - u}{f(F^{-1}(u))} \mu(u) \mathbf{1}(u \in \mathbb{I}_n).$$

In the proof of Theorem 4 we saw that  $R, \{X_1 - F^{-1}(u)\}^+ \in L^2$ ; since  $f$  is bounded, we also have that  $\tilde{Y}_{11}, \tilde{Y}_{12} \in L^2(w)$ . As for  $\tilde{Y}_{13}$ ,

$$\int_0^1 \tilde{Y}_{13}^2(u)w(u)du = \int_{\mathbb{I}_n} \left[ \mathbf{1}\{X_1 \leq F^{-1}(u)\} - u \right]^2 \mu^2(u)du \leq \int_0^1 \mu^2(u)du < \infty,$$

because  $\mu(u) \leq M, \forall u \in [0, 1]$ , since  $f$  is bounded. Thus,  $\tilde{Y}_1 \in L^2(w)$ .

From the central limit theorem in Hilbert spaces and the continuous mapping theorem,

$$\int_0^1 W_n(t)^2 w(u)du \xrightarrow{\mathcal{L}} \|Z\|_w^2. \quad (42)$$

Since  $R$  is a decreasing function with  $\lim_{u \uparrow 1} R(u) = 0$ , and  $w$  is bounded

$$0 \leq n \int_{(n-1)/n}^1 R(u)^2 w(u)du \leq Mn \int_{(n-1)/n}^1 R(u)^2 du \leq MR^2((n-1)/n) \rightarrow 0. \quad (43)$$

The result follows from (41)–(43).  $\square$

**Acknowledgements** The authors thank the two anonymous referees for their constructive comments. This work supported by the Grant PID2020-118101GB-I00, Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033).

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arriaza A, Di Crescenzo A, Sordo MA, Suárez-Llorens A (2019) Shape measures based on the convex transform order. *Metrika* 82:99–124
- Azzalini A (1981) A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika* 68(1):326–328
- Bickel PJ, Freedman DA (1981) Some asymptotic theory for the bootstrap. *Ann Stat* 9(6):1196–1217
- Bosq D (2000) Linear processes in function spaces. Theory and applications. *Lecture Notes in Statistics* Vol 149. Springer, New York
- Csörgő M (1983) Quantile processes with statistical applications. Society for Industrial and Applied Mathematics, Philadelphia
- del Barrio E, Cuesta-Albertos JA, Matrán C (2000) Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests. *TEST* 9(1):1–96

- del Barrio E, Giné E, Utzet F (2005) Asymptotics for  $L_2$  functionals of the empirical quantile process, with applications to tests of fit based on weighted wasserstein distances. *Bernoulli* 11(1):131–189
- Gómez-Déniz E, Sordo MA, Calderín-Ojeda E (2014) The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance* 54:49–57
- Janssen A (2000) Global power functions of goodness of fit tests. *Ann Stat* 28(1):239–253
- Jodrá P, Jiménez-Gamero MD (2016) A note on the Log-Lindley distribution. *Insurance* 71:189–194
- Pardo-Fernández JC, Jiménez-Gamero MD (2019) A model specification test for the variance function in nonparametric regression. *AStA Adv Stat Anal* 103:387–410
- Pardo-Fernández JC, Jiménez-Gamero MD, El Ghouh A (2015) A non-parametric ANOVA-type test for regression curves based on characteristic functions. *Scand J Stat* 42(1):197–213
- Prakasa Rao BLS (1983) Nonparametric functional estimation. Academic Press Inc, New York
- R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York
- Smith RL, Naylor JC (1987) A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *J R Stat Soc Ser C* 36:358–369

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.