Ann Inst Stat Math (2011) 63:617–644
DOI 10.1007/s10463-009-0241-5

Edgeworth expansion for the kernel quantile estimator

Yoshihiko Maesono · Spiridon Penev

Received: 28 April 2008 / Revised: 23 March 2009 / Published online: 4 June 2009

Abstract Using the kernel estimator of the pth quantile of a distribution brings about an improvement in comparison to the sample quantile estimator. The size and order of this improvement is revealed when studying the Edgeworth expansion of the kernel estimator. Using one more term beyond the normal approximation significantly improves the accuracy for small to moderate samples. The investigation is nonstandard since the influence function of the resulting L-statistic explicitly depends on the sample size. We obtain the expansion, justify its validity and demonstrate the numerical gains in using it.

Keywords Edgeworth expansion · Kernel quantile estimator · Quantile · Validity

1 Introduction

There has been renewed interest in quantile estimation during the last decade. This is mostly due to the important practical applications of the method in the financial industry and risk assessment. The risk measure VaR is a tantamount quantile. Recent more refined definitions of coherent risk measures also represent suitable transformations of quantiles. The Lorenz curve (Lorenz 1905) as an indicator of income distributions, also represents transformation of quantiles. The paper Ogryczak and Ruszczyński (2002) illustrates the importance of transformations of quantiles in the modern theory

Y. Maesono (✉)
Department of Mathematical Sciences, Faculty of Mathematics, Kyushu University, Fukuoka, Japan
e-mail: maesono@math.kyushu-u.ac.jp

S. Penev
Department of Statistics, School of Mathematics and Statistics, The University of New South Wales, Sydney, Australia
e-mail: S.Penev@unsw.edu.au

of dual stochastic dominance. Estimating these transformations of quantiles requires dedicated approaches and is not discussed further in this paper. However, estimating the quantile itself as accurately as possible is an essential ingredient of the inference; it is important in its own right and will be dealt with in this paper.

For a continuous random variable X with a cumulative distribution function $F(x)$, density function $f(x)$ and $E|X| < \infty$, the pth quantile is defined as $Q(p) = inf\{x : F(x) \geq p\}$. Given a sample $X_1, X_2, \ldots, X_n$, from F, the simplest estimator of $Q(p)$ is the sample quantile estimator ξ_{pn}; that is, the pth quantile of the empirical distribution function $F_n(x) = \frac{1}{n}\sum_{i=1}^n I(X_i \leq x)$. It is a popular estimator that is typically implemented in statistical packages. Under mild conditions, it is asymptotically normal, but its asymptotic variance is $\sigma^2 = \frac{p(1-p)}{f^2(Q(p))}$ which happens to be large particularly in the tails of the distribution where the density $f(x)$ has small values. The tails are exactly the region of interest when evaluating quantiles. Alternative estimators need to be considered and, further terms in the asymptotic expansion need to be taken into account to improve the accuracy for small and for moderate samples.

One obvious choice is the kernel quantile estimator

$$\hat{Q}_{p,h_n} = \frac{1}{h_n}\int_0^1 F_n^{-1}(x)K\left(\frac{x-p}{h_n}\right)\mathrm{d}x, \tag{1}$$

where $F_n^{-1}(x)$ denotes the inverse of the empirical distribution function, and $K(.)$ is a suitably chosen kernel. Intuitively, in (1) we are weighting up different sampling quantiles by using weights that are determined by the kernel instead of just using only one empirical sample quantile. Weighting is chosen carefully to make sure that the observations closer to the empirical pth quantile have higher weight in comparison to the remaining ones. Of course, the bandwidth should be chosen appropriately. Besides the obvious requirement $h_n \to_{n\to\infty} 0$, additional requirements in combination with requirements on the kernel must be placed to ensure consistency, asymptotic normality, asymptotic bias elimination, and higher order accuracy of the estimator $\hat{Q}_{p,h_n}$.

The estimator in (1) has been studied by many researchers in the past (see Falk 1984, 1985; Sheather and Marron 1990; Xiang 1995a,b and the references therein). Clearly, $\hat{Q}_{p,h_n}$ is an L-estimator since it can be written as a weighted sum of the order statistics $X_{(i)}, i = 1, 2, \ldots, n$:

$$\hat{Q}_{p,h_n} = \sum_{i=1}^n v_{i,n}X_{(i)}, \quad v_{i,n} = \frac{1}{h_n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} K\left(\frac{x-p}{h_n}\right)\mathrm{d}x. \tag{2}$$

The difficulty in using standard asymptotic theory about L-statistics when analysing the behaviour of $\hat{Q}_{p,h_n}$ is related to the fact that the weights $v_{i,n}$ in (2) depend explicitly on n in a peculiar way and tend to vanish asymptotically. Standard statements about asymptotic expansions of L-statistics (e.g. Serfling (1980); Helmers (1982)) are related to the case where the L-statistic could be written as $\int_0^1 F_n^{-1}(u)J(u)\mathrm{d}u$ with the *score function* $J(u)$ (which itself is an asymptotic limit) that does *not* involve n. A separate treatment is necessary for (2). The reason is that it is impossible to write its "score function" in such a way since it becomes a delta function in the limit.

The treatment is technically involved and this has slowed down the research on the properties of $\hat{Q}_{p,h_n}$ as an estimator of $Q(p)$. It is particularly interesting to derive higher order expansions for the asymptotic distribution of $\hat{Q}_{p,h_n}$. Indeed, research by Falk (1984) and Xiang (1995b) shows that the kernel quantile estimator is asymptotically equivalent to the empirical quantile estimator up to first order, hence any advantage would be revealed by involving at least one more term in the Edgeworth expansion and for moderate sample sizes. In this paper, we will derive such an expansion, will justify its validity and will illustrate its advantages numerically.

It is interesting to mention an alternative approach towards improving the estimation of a quantile suggested in Kozek (2005). He shows that it is possible to reduce the variance by considering slightly perturbed quantiles that are estimated using M-estimators. The resulting estimators exhibit some asymptotically non-vanishing bias as a payoff for the slightly reduced variance. In our approach, we do not have asymptotic bias. The accuracy for moderate samples is revealed via inclusion of more terms in the Edgeworth expansion.

2 Relations to U-statistics

A standard approach in the study of asymptotic properties of L-statistics is to first decompose them into an U-statistic plus a small-order remainder term and to apply asymptotic theory for the main term, that is, the U-statistic. When the $J(u)$ function depends on n however, such a decomposition leads to an object that is similar to an asymptotic U-statistic (Lai and Wang 1993) yet its defining function $h(\cdot,\cdot)$ explicitly depends on n and standard results about U-statistics can not be used directly. Progress in this direction has been achieved by Xiang (1995a,b). By applying Esseen's smoothing lemma (Feller 1971, Ch. XVI) and exploiting a nice decomposition by Friedrich for the resulting statistic, he has shown that a Berry-Esseen type result of the form

$$P\left(\sqrt{n}|\hat{Q}_{p,h_n} - Q(p)| \le x\sigma_n\right) = 2\Phi(x) - 1 + O(n^{-r}) \tag{3}$$

holds whereby σ_n^2 can be given explicitly in terms of the kernel K and the derivative $Q'(p)$ of $Q(p)$. The rate r depends on the order m of the kernel (for the exact definition of kernel order, see Sect. 3). When $m = 2$ he obtains $O(n^{-1/3})$, which is improved to $O(n^{-5/13})$ for $m = 3$. In his approach, for no kernel order m it is possible to get the reminder order in (3) down to $O(n^{-1/2})$, not to mention $o(n^{-1/2})$. In Xiang and Vos (1997, Lemma 1), the authors have attempted to derive Edgeworth expansion in order to improve (3). However, the result as stated there is incorrect. Both the definition of the expansion function ($G_n(x)$) and the order of the expansion ($o(n^{-1/3})$) do not make sense and the argument used in the derivation is imprecise. Indeed, the authors rely on a technique about U-statistics used by Bickel et al. (1986) but in the latter paper the function $h(\cdot,\cdot)$ in the U-statistic's definition does *not* depend on n.

The purpose of this paper is to suggest a correct Edgeworth expansion for the kernel-based quantile estimator $\hat{Q}_{p,h_n}$ up to order $o(n^{-1/2})$. The derivation is non-trivial. The specific requirement on the bandwidth h_n for the sake of eliminating the asymptotic bias triggers the need to include further contributions from the terms of order $O(n^{-1})$ of the appropriately modified result for U-statistic. These contributions have to be taken into account in order to achieve the desired Edgeworth expansion with remainder of order $o(n^{-1/2})$. Although the general result is involved, the expansion can be simplified significantly in the most typical and practically relevant case of a symmetric compactly supported kernel $K(\cdot)$. We present numerical evidence about the improved accuracy of the Edgeworth expansion in comparison to the normal approximation for moderate samples sizes and some common distributions $F(x)$.

The proofs of our results require novel dedicated approaches. After an extensive search of the literature, we were indeed able to find one recent result (Jing and Wang 2003, Lemma 2.1) concerning the Edgeworth expansion of uniformly integrable transformations of independent identically distributed random variables in the form

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_n(X_i) + \frac{1}{n^{3/2}}\sum_{1\le i<j\le n} W_n(X_i, X_j), \quad n \ge 2, \tag{4}$$

whereby the transformations $V_n(\cdot)$ and $W_n(\cdot,\cdot)$ are explicitly allowed to depend on n. For the sake of completeness, we quote the statement of the Lemma:

Let $W_n(x, y)$ be a symmetric function in its arguments. Assume also that:

a) $EV_n(X_1) = 0$, $EV_n^2(X_1) = 1$, $|V_n(X_1)|^3$ are uniformly integrable and the distribution of $V_n(X_1)$ is non-lattice for all sufficiently large n.
b) $E(W_n(X_1, X_2)|X_1) = 0$, $|W_n(X_1, X_2)|^{5/3}$ are uniformly integrable and $|W_n(X_i, X_j)| \le n^{3/2}$ for all $n \ge 2$ $i \ne j$.

Then, as $n \to \infty$,

$$\sup_x \left| P\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_n(X_i) + \frac{1}{n^{3/2}}\sum_{1\le i<j\le n} W_n(X_i, X_j) \le x\right) - F_n^{(1)}(x)\right| = o(n^{-1/2}), \tag{5}$$

where $F_n^{(1)}(x) = \Phi(x) - \frac{\Phi^{(3)}(x)}{6\sqrt{n}}(EV_n^3(X_1) + 3EV_n(X_1)V_n(X_2)W_n(X_1, X_2))$. Here $\Phi(x)$ denotes the cdf of the standard normal distribution, $\phi(x)$ is its density, and $\Phi^{(3)}(x)$ is the third derivative.

However, as already pointed out, the above Lemma does not quite serve our purposes since we need one more term in the expansion of the U-statistic in order to take into account the effect of the bandwidth h_n.

3 Edgeworth expansion for $\hat{Q}_{p,h_n}$

For simplicity of the exposition, we will be using a compactly supported kernel $K(x)$ on $(-1, 1)$. We will say that the kernel is of order m if

$$K(x) \in L^2(-\infty, \infty), \quad K^{(m)}(x) \in Lip(\alpha) \text{ for some } \alpha > 0,$$
$$\int_{-1}^{1} K(x)\mathrm{d}x = 1, \quad \int_{-1}^{1} x^i K(x)\mathrm{d}x = 0, i = 1, 2, \ldots, m-1, \quad \int_{-1}^{1} x^m K(x)\mathrm{d}x \neq 0.$$

For now, we do not require symmetry of the kernel.

Achieving a fine bias-variance trade-off for the kernel-based estimator is a delicate matter. It will be seen in the proof that the proper choice of the bandwidth is given by

$$h_n = o(n^{-1/4}) \quad \text{and} \quad \lim_{n\to\infty} (n^{1/4} h_n)^{-k} n^{-\beta} = 0 \tag{6}$$

for any $\beta > 0$ and integer k. The typical bandwidth we choose to work with in the numerical implementation is $h_n = n^{-1/4}(\log n)^{-1}$.

Before formulating the main results of the paper, we briefly review the moment evaluations of the H-decomposition. For independent identically distributed random variables $X_1, \ldots, X_n$ and a function $\nu(x_1, \ldots, x_r)$ which is symmetric in its arguments with the property $E[\nu(X_1, \ldots, X_r)] = 0$, we define

$$\rho_1(x_1) = E[\nu(x_1, X_2, \ldots, X_r)],$$
$$\rho_2(x_1, x_2) = E[\nu(x_1, x_2, \ldots, X_r)] - \rho_1(x_1) - \rho_1(x_2), \ldots,$$

and

$$\rho_r(x_1, x_2, \ldots, x_r) = \nu(x_1, x_2, \ldots, x_r) - \sum_{k=1}^{r-1} \sum_{C_{r,k}} \rho_k(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$$

where $\sum_{C_{r,k}}$ indicates that the summation is taken over all integers $i_1, \ldots, i_k$ satisfying $1 \leq i_1 < \cdots < i_k \leq r$. Then it holds

$$E[\rho_k(X_1, \ldots, X_k) | X_1, \ldots, X_{k-1}] = 0 \; a.s. \tag{7}$$

and

$$\sum_{C_{n,r}} \nu(X_{i_1}, \ldots, X_{i_r}) = \sum_{k=1}^{r} \binom{n-k}{r-k} A_k \tag{8}$$

where

$$A_k = \sum_{C_{n,k}} \rho_k(X_{i_1}, \ldots, X_{i_k}). \tag{9}$$

Using Eq. (7) and the moment evaluations of martingales in Dharmadhikari et al. (1968) we have the upper bounds of the absolute moments of A_k. For $q \geq 2$, if $E|\nu(X_1, \ldots, X_r)|^q < \infty$, there exists a positive constant C, which may depend on ν and F but not on n, such that

$$E|A_k|^q \leq Cn^{qk/2} E|\rho_k(X_{i_1}, \ldots, X_{i_k})|^q. \tag{10}$$

Since the empirical distribution $F_n(u)$ is a sum of independent identically distributed (i.i.d.) random variables, we have moment evaluations of $|F_n(u) - F(u)|$. For $q \geq 2$, we get

$$E|F_n(u) - F(u)|^q \leq Cn^{-q/2} F(u)(1 - F(u)) \tag{11}$$

where C is a constant. Hereafter C will denote a generic constant that may change its meaning at different places in the text.

Using the inequalities (10) and (11), we will obtain an asymptotic representation of the standardized quantile estimator with residual $o_L(n^{-1/2})$

$$P\left(|o_L(n^{-1/2})| \geq \gamma_n n^{-1/2}\right) = o(n^{-1/2})$$

where $\gamma_n \to 0$ as $n \to \infty$.

Let $\{Y_i\}_{i=1,\ldots,n}$ be independent random variables uniformly distributed on (0, 1) and define

$$\begin{aligned}
&\bar{Q}(p) = \frac{1}{h_n} \int_0^1 F^{-1}(x) K\left(\frac{x-p}{h_n}\right) \mathrm{d}x, \\
&\hat{I}_x(Y_1) = I(Y_1 \leq p + h_n x) - (p + h_n x), \\
&g_{1n}(Y_1) = -\int_{-1}^1 Q'(p + h_n x) K(x) \hat{I}_x(Y_1) \mathrm{d}x, \\
&\sigma_n^2 = \mathrm{Var}(g_{1n}(Y_1)), \\
&d_{1n} = \sigma_n^{-1} n^{-1/2}, \quad d_{2n} = \sigma_n^{-1} n^{-3/2} h_n^{-1}, \quad d_{3n} = \sigma_n^{-1} n^{-5/2} h_n^{-2}, \\
&g_{2n}(Y_1, Y_2) = -\int_{-1}^1 Q'(p + h_n x) K'(x) \hat{I}_x(Y_1) \hat{I}_x(Y_2) \mathrm{d}x, \\
&g_{3n}(Y_1, Y_2, Y_3) = -\int_{-1}^1 Q'(p + h_n x) K^{(2)}(x) \hat{I}_x(Y_1) \hat{I}_x(Y_2) \hat{I}_x(Y_3) \mathrm{d}x, \\
&\hat{g}_{1n}(Y_1) = -\frac{1}{2} \int_{-1}^1 Q'(p + h_n x) K^{(2)}(x) E[\hat{I}_x^2(Y_2)] \hat{I}_x(Y_1) \mathrm{d}x, \\
&A_{1n} = \sum_{i=1}^n g_{1n}(Y_i), \quad A_{2n} = \sum_{C_{n,2}} g_{2n}(Y_i, Y_j), \quad A_{3n} = \sum_{C_{n,3}} g_{3n}(Y_i, Y_j, Y_k)
\end{aligned}$$

and $\hat{A}_{1n} = \sum_{i=1}^n \hat{g}_{1n}(Y_i)$. Then we have the following lemma.

Lemma 1 *Assume* $\int[F(x)(1-F(x))]^{1/5}\mathrm{d}x < \infty$. *Let* $Q^{(m)}$ *be uniformly bounded in a neighborhood of* p $(0 < p < 1)$ *and* $f(Q(p)) > 0$. *Let* $K(x)$ *be a fourth order kernel (i.e.* $m = 4$*) and* $K^{(4)}(x) \in Lip(\alpha), \alpha > 0$. *Denote* $\delta = Q'(p)(\frac{1}{2} - p) + \frac{1}{2}Q^{(2)}(p)p(1-p)$. *Further, choose* h_n *satisfying* (6). *Then we have*

$$\begin{aligned}&\sigma_n^{-1}\sqrt{n}\big(\hat{Q}_{p,h_n} - \bar{Q}(p)\big)\\&\quad = d_{1n}A_{1n} + d_{2n}A_{2n} + d_{3n}A_{3n} + d_{3n}(n-1)\hat{A}_{1n} + \frac{\delta}{\sigma\sqrt{n}} + o_L(n^{-1/2}).\end{aligned}$$

The lemma shows that to get the Edgeworth expansion of $\sigma_n^{-1}\sqrt{n}(\hat{Q}_{p,h_n} - \bar{Q}(p))$, we can concentrate first on the Edgworth expansion of

$$d_{1n}A_{1n} + d_{2n}A_{2n} + d_{3n}A_{3n} + d_{3n}(n-1)\hat{A}_{1n}$$

with residual term $o_L(n^{-1/2})$. We will also prove the validity of the Edgeworth expansion.

Theorem 1 *Under the assumptions of Lemma* 1 *we have*

$$P\big(\sqrt{n}(\hat{Q}_{p,h_n} - \bar{Q}(p)) \le x\sigma_n\big) = G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) + o(n^{-1/2}). \tag{12}$$

Here

$$\begin{aligned}G_n(x) = \Phi(x) - \phi(x)\Bigg\{&\frac{x^2-1}{6n^{1/2}\sigma_n^3}\left(e_{1n} + \frac{3e_{2n}}{h_n}\right) + \frac{1}{nh_n^2}\\&\times\left(\frac{x}{4\sigma_n^2}\{4e_{5n} + e_{6n}\} + \frac{x^3-3x}{6\sigma_n^4}\{3e_{3n} + e_{4n}\} + \frac{x^5 - 10x^3 + 15x}{8\sigma_n^6}e_{2n}^2\right)\Bigg\}\end{aligned}$$

where

$$\begin{aligned}e_{1n} &= E[g_{1n}^3(Y_1)], \quad e_{2n} = E[g_{1n}(Y_1)g_{1n}(Y_2)g_{2n}(Y_1, Y_2)],\\e_{3n} &= E[g_{1n}(Y_2)g_{1n}(Y_3)g_{2n}(Y_1, Y_2)g_{2n}(Y_1, Y_3)],\\e_{4n} &= E[g_{1n}(Y_1)g_{1n}(Y_2)g_{1n}(Y_3)g_{3n}(Y_1, Y_2, Y_3)],\\e_{5n} &= E[g_{1n}(Y_1)\hat{g}_{1n}(Y_1)] \quad \text{and} \quad e_{6n} = E[g_{2n}^2(Y_1, Y_2)].\end{aligned}$$

Expanding $G_n(x - \frac{\delta}{\sigma\sqrt{n}})$ around x and keeping the $O(\frac{1}{\sqrt{n}})$ terms only we can also write (12) as follows:

$$\begin{aligned}
&P\big(\sqrt{n}(\hat{Q}_{p,h_n} - \bar{Q}(p)) \le x\sigma_n\big) \\
&= \Phi(x) - \phi(x)\left\{\frac{x^2-1}{6n^{1/2}\sigma_n^3}\left(e_{1n} + \frac{3e_{2n}}{h_n}\right) + \frac{1}{nh_n^2}\left(\frac{x}{4\sigma_n^2}\{4e_{5n} + e_{6n}\}\right.\right. \\
&\quad \left.\left. + \frac{x^3-3x}{6\sigma_n^4}\{3e_{3n} + e_{4n}\} + \frac{x^5 - 10x^3 + 15x}{8\sigma_n^6}e_{2n}^2\right)\right\} - \frac{\delta}{\sigma\sqrt{n}}\phi(x) + o(n^{-1/2}).
\end{aligned}$$

The obtained Edgeworth expansion (12) is not for the ultimate quantity of interest $\sqrt{n}(\hat{Q}_{p,h_n} - Q(p))$. Representing

$$\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) = \sqrt{n}(\hat{Q}_{p,h_n} - \bar{Q}(p)) + d_n$$

with $d_n = \sqrt{n}(\bar{Q}(p) - Q(p)) = \sqrt{n}(\frac{1}{h_n}\int_0^1 F^{-1}(x)K(\frac{x-p}{h_n})\mathrm{d}x - Q(p))$ we need to make sure that the asymptotic order of the bias d_n is kept under control. Substituting $x - p = yh_n$ and applying Taylor expansion up to order m of $Q(x) = F^{-1}(x)$ around p, we get

$$d_n = O(\sqrt{n}h_n^m)\int_{-1}^{1} K(y)y^m\mathrm{d}y.$$

Hence

$$\begin{aligned}
&P\big(\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) \le x\sigma_n\big) - G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) \\
&= P\left(\sqrt{n}\left(\hat{Q}_{p,h_n} - \bar{Q}(p)\right) \le \left(x - \frac{d_n}{\sigma_n}\right)\sigma_n\right) - G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) \\
&= G_n\left(x - \frac{d_n}{\sigma_n} - \frac{\delta}{\sigma\sqrt{n}}\right) - G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) + o(n^{-1/2}). \qquad (13)
\end{aligned}$$

Expanding further G_n around the point $x - \frac{\delta}{\sigma\sqrt{n}}$ we finally get

$$\begin{aligned}
&P\left(\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) \le x\sigma_n\right) - G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) \\
&= \phi\left(x - \frac{\delta}{\sigma\sqrt{n}}\right)\left(-\frac{d_n}{\sigma_n}\right) + o(n^{-1/2}). \qquad (14)
\end{aligned}$$

To keep the contribution of d_n under control, we require $d_n = o(n^{-1/2})$, that is, $\sqrt{n}h_n^m = o(n^{-1/2})$. We see now that the requirements $h_n = o(n^{-1/4})$, $m = 4$ that we put for another reason in Theorem 1, also guarantee that $d_n = o(n^{-1/2})$ holds. This is one more argument in favour of the choice $h_n = o(n^{-1/4})$ for the bandwidth.

Moreover, the discussion in Sheather and Marron (1990, p. 411) also demonstrates from another point of view that a reasonable choice of h_n should satisfy the order requirement $h_n = o(n^{-1/4})$. Their argument is that otherwise the dominant term in the MSE expansion of the kernel-quantile estimator may become worse than the one for the empirical quantile estimator, that is, $\frac{p(1-p)}{n}(Q'(p))^2$.

The above discussion together with Theorem 1, leads us to the following

Theorem 2 *Under the assumptions of Theorem* 1, *we have the following Edgeworth expansion with remainder* $o(n^{-1/2})$:

$$P\left(\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) \leq x\sigma_n\right) = G_n\left(x - \frac{\delta}{\sigma\sqrt{n}}\right) + o(n^{-1/2}). \tag{15}$$

Remark 1 Our expansion also leads to a Berry-Esseen bound

$$P\left(\sqrt{n}|\hat{Q}_{p,h_n} - Q(p)| \leq x\sigma_n\right) = 2\Phi(x) - 1 + O(n^{-1/2}).$$

This is an improvement on the result found by Xiang (1995a) whose bound for $m = 4$ only implies $P(\sqrt{n}|\hat{Q}_{p,h_n} - Q(p)| \leq x\sigma_n) = 2\Phi(x) - 1 + O(n^{-7/17})$.

Remark 2 A careful examination of the proof of Theorem 1 shows that for its statement to hold, one does not actually need the moment conditions $\int_{-1}^{1} x^i K(x)\mathrm{d}x = 0, i = 1, 2, \ldots, m-1$ on the kernel $K(x)$ to hold. However, these conditions are required in order to get the order of the bias right so that the ultimate expansion given in Theorem 2 could be obtained.

Remark 3 Finally, we will comment on a crucial special case with important implications about the implementation of the Edgeworth expansion (15). We note that the general statement of Theorem 2 does not require *symmetry* (around zero) of the kernel $K(x)$. However, most of the kernels used in practical applications are usually symmetric. The symmetry assumption is also automatically made in Sheather and Marron (1990) and in other influential papers on kernel quantile estimation. If we do assume that in addition the kernel *is* symmetric around zero (that is, $K(-x) = K(x)$) holds), then obviously $\int_{-1}^{1} K'(x)\mathrm{d}x = 0$ holds. In that case, via Taylor expansion around p, we can see that the terms e_{2n} and e_{6n} in the expansion in Theorems 1 and 2 are of smaller order in comparison to the remaining terms and we get in this case the simpler expression:

$$\begin{aligned} &P\left(\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) \leq x\sigma_n\right) \\ &= \Phi(x) - \phi(x)\left\{\frac{x^2-1}{6n^{1/2}\sigma_n^3}\left(e_{1n} + \frac{3e_{2n}}{h_n}\right) + \frac{1}{nh_n^2}\left(\frac{xe_{5n}}{\sigma_n^2} + \frac{x^3-3x}{6\sigma_n^4}e_{4n}\right)\right\} \\ &\quad - \frac{\delta}{\sigma\sqrt{n}}\phi(x) + o(n^{-1/2}). \end{aligned} \tag{16}$$

For higher-order kernels like the ones we use in this paper it is possible to make them also satisfy the additional condition $\int_{-1}^{1} K''(x)\mathrm{d}x = 0$. If such kernel is chosen then expression (16) simplifies even further and becomes

$$P\left(\sqrt{n}(\hat{Q}_{p,h_n} - Q(p)) \le x\sigma_n\right)$$
$$= \Phi(x) - \phi(x)\frac{x^2-1}{6n^{1/2}\sigma_n^3}\left(e_{1n} + \frac{3e_{2n}}{h_n}\right) - \frac{\delta}{\sigma\sqrt{n}}\phi(x) + o(n^{-1/2}). \quad (17)$$

Remark 4 It is known that in terms of first order performance *with respect to MSE* the kernel quantile estimator can only match the sample quantile (Sheather and Marron 1990). The improvement with respect to the sample quantile can only show up in higher order terms of the MSE approximation (this phenomenon has been called *deficiency*). The works of Falk (1984), Sheather and Marron (1990), Borovkov et al. (1996) show advantages of the kernel quantile estimator with respect to the deficiency criterion. The paper Borovkov et al. (1996) relaxes some of the original requirements of Falk (1984) regarding bandwidth choice. For an appropriate compactly supported in $(-1, 1)$ fourth-order kernel $K(x)$ crucial is the sign of the quantity

$$\psi = \int_{-1}^{1} yK(y)M(y)\mathrm{d}y, \quad M(y) = \int_{-1}^{y} K(x)\mathrm{d}x.$$

When $\psi > 0$, the kernel quantile estimator is better than the sample quantile in terms of deficiency whereas if $\psi < 0$ the sample quantile is better. This underlines the importance of the choice of the analysing kernel for achieving higher order advantage in estimating the quantile. When using such quantile estimator, it is also important to develop higher order expansion for its distribution. The Edgeworth expansion we propose is more accurate than the normal approximation of the kernel quantile and can be used in conjunction of choices where $\psi > 0$ to suggest better estimators than the sample quantile. Relevant practical applications of expansion (17) are in evaluating power of tests of the type $H_0 : Q(p) = q_0$ against one-sided or two-sided alternatives. If the test is based on the statistic $\frac{\sqrt{n}(\hat{Q}_{p,h_n} - q_0)}{\sigma_n}$ then by using (17) one should be able to get better power approximation for such a test in comparison to just using the normal approximation. Another application is in constructing more accurate confidence interval for the quantile $Q(p)$ when the sample size is small to moderate. For a given level α, instead of constructing it in a symmetric way as $\hat{Q}_{p,h_n} \pm z_{\alpha/2}\sigma_n/\sqrt{n}$ one can improve the coverage accuracy by using $(\hat{Q}_{p,h_n} + c_{1-\alpha/2}\sigma_n/\sqrt{n}, \hat{Q}_{p,h_n} + c_{\alpha/2}\sigma_n/\sqrt{n})$ with the quantile values $c_{1-\alpha/2}$ and $c_{\alpha/2}$ obtained by inverting numerically the Edgeworth approximation.

4 Numerical comparisons

Our goal in this section is to demonstrate numerically the effect of the improvement in the approximation of the distribution of the kernel quantile estimator when we move from the normal to the Edgeworth approximation. This effect could be seen for moderate sample sizes such as $n = 15, 30, 40, 50$ (for large n the two approximations become very close). We could choose different kernels satisfying the requirements of Theorem 2 but the effect of the kernel is not that crucial as long as it satisfies the condition $\psi = \int_{-1}^{1} yK(y)M(y)\mathrm{d}y > 0$. Here we only present results obtained with the following symmetric fourth order kernel suggested by H. Müller:

$$K(x) = \frac{315}{512}\left(11x^8 - 36x^6 + 42x^4 - 20x^2 + 3\right)I(|x| \leq 1).$$

It is easily seen that $\psi > 0$ holds. The "asymptotically correct" bandwidth $h_n = n^{-1/4}(\log n)^{-1}$ turned out to be very well adjusted for samples of size $n \geq 50$. For smaller sample sizes it may be necessary to choose a smaller bandwidth, e.g., $0.1n^{-1/4}$, specifically when the value of p is near 0 or 1, to protect against a bias associated with edge effects.

The integral of the second derivative of this kernel is equal to zero hence the approximation (17) can be applied. The improvement effect from using Edgeworth expansion depends on the particular distribution, on the sample size and on the value of p. We include examples where the underlying distribution is normal for the sake of completeness but more spectacular improvement is achieved for skewed distributions, for smaller values of n (e.g., 15, 30) and for p closer to zero or to one. The subsections about the exponential and gamma distribution confirm this.

4.1 Estimation of quantiles of the standard normal distribution

In this case, we can derive easily:

$$Q'(p) = \frac{1}{\phi(\Phi^{-1}(p))}, \quad Q''(p) = \frac{\Phi^{-1}(p)}{\phi^2(\Phi^{-1}(p))}.$$

We compare the numerical values of "true" cumulative distribution function of the standardised random variable $\sqrt{n}(\hat{Q}_{p,h_n} - Q(p))/\sigma_n$ with the cumulative distribution of the standard normal and to the Edgeworth approximation as derived in Theorem 2. Numerical values were compared for two different scenarios: $n = 50$ and $n = 40$, both applied for the same value of $p = 0.1$. The "true" values of the cumulative distribution function were calculated on the basis of the empirical proportions for 2,000,000 repeated simulations. We found that at this number of simulations, the empirical ratios virtually do not change uniformly over the whole range $(-2.5, 2.5)$ of values of the argument up to the fourth decimal place. The values of e_{in} were calculated via averaging the numerical values of the resulting $g_{1n}(Y_1)$, $g_{2n}(Y_1, Y_2)$ values for 500,000 simulated independent uniform (0,1) values Y_1, Y_2 whereas σ_n was calculated as the empirical standard deviation estimator for $\sqrt{\mathrm{Var}(g_{1n}(Y_1))}$ using the 500,000 simulated Y_1 observations. Again, the 500,000 simulations were chosen on the basis of our experimentation showing that the resulting estimated parameter values for e_{in}, and for σ_n were virtually unchanged by further increase of the number of simulations. Tabulated numerical values can be obtained from the authors upon request. In a nutshell, except for a very tiny region of values in the lower tails of the distribution, the Edgeworth expansion approximated much better the true cumulative distribution

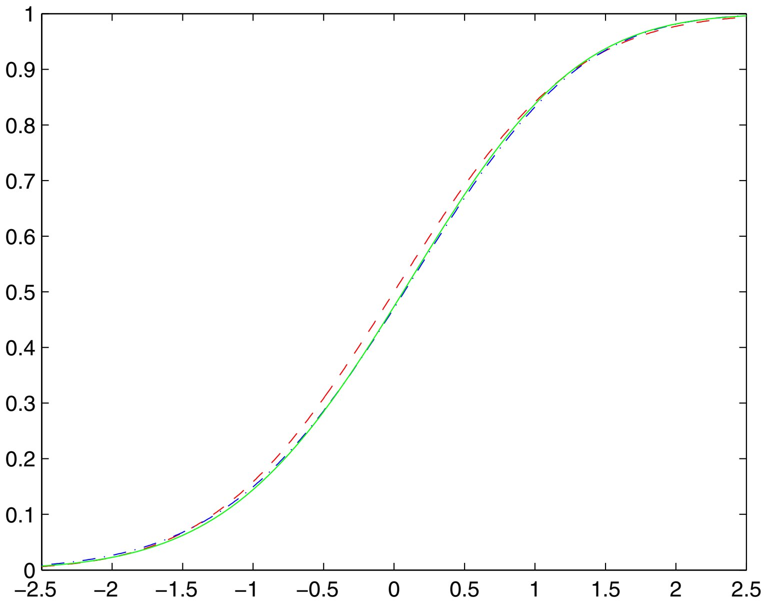

Fig. 1 Standard normal distribution, $p = 0.1, n = 40$. True distribution *continuous line*, Edgeworth approximation: *dot-dashed line*, normal approximation: *dashed line*

function's values. In the small region in the lower tail where the opposite happened, the difference in the two approximations is negligibly small (it is in the third digit after the decimal point only, whereas in the region where the Edgeworth expansion is better, the effect of the improvement is typically observed in the second digit). Since the graphs for $n = 40$ and $n = 50$ are similar, we only present the graph in Fig. 1 corresponding to $n = 40$.

4.2 Estimation of quantiles of the standard exponential distribution

The following calculations are easy to derive in this case:

$$Q(p) = -\ln(1-p), \quad Q'(p) = \frac{1}{1-p}, \quad Q''(p) = \frac{1}{(1-p)^2}.$$

Here, the improvement using the Edgeworth approximation is more significant and can be felt for sample sizes as small as $n = 15$. The improvement is *virtually uniform* all over the whole range of values of the argument (except for a small region in the very left tail where, as is known, the Edgeworth expansion may become negative, in which case it should be set equal to zero). The comparison was done for two scenarios: $n = 30$ and $n = 15$, both applied for the same value of $p = 0.9$. Again, the "true" values of the cumulative distribution were calculated on the basis of the empirical proportions for 2,000,000 repeated simulations and 500,000 auxiliary independent uniform (0,1) pairs of values Y_1, Y_2, were generated. The tabulated numerical values

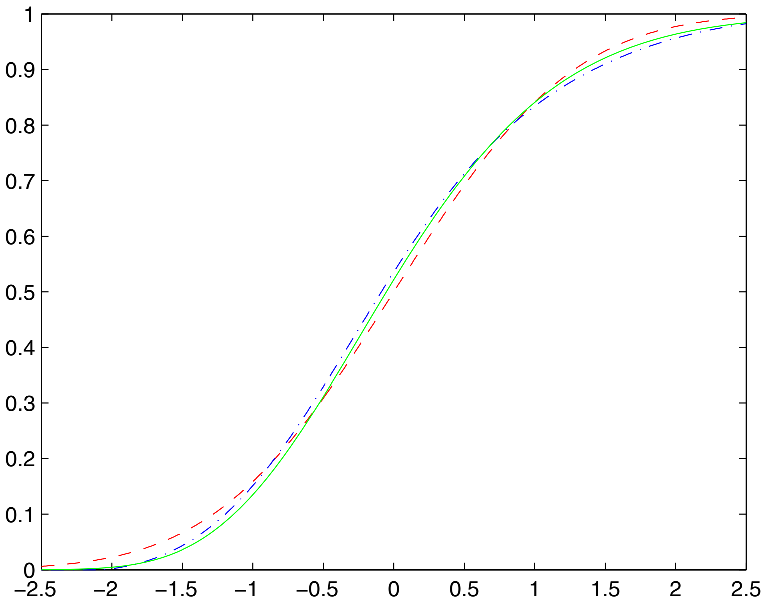

Fig. 2 Standard exponential distribution, $p = 0.9$, $n = 30$. True distribution: *continuous line*, Edgeworth approximation: *dot-dashed line*, normal approximation: *dashed line*

are not included in the paper and are available upon request. A graphical illustration of the same comparison is presented in Fig. 2 (for $n = 30$) and in Fig. 3 (for $n = 15$).

4.3 Estimation of quantiles of the Gamma distribution

There is a reliable algorithm (DiDonato and Morris 1987) for calculating the inverse of the incomplete Gamma function. This alleviates the testing of our approximation for any Gamma distribution. For illustrative purposes, we show some results for the Chi-squared distribution with 4 degrees of freedom. This is a particular Gamma distribution with the density $f(x) = \frac{1}{4}xe^{-x/2}$, $x > 0$. After changing variables, we see that for this distribution, $Q(p)$ is a solution of the equation:

$$p = \int_0^{Q(p)/2} ye^{-y}\mathrm{d}y, \tag{18}$$

which means that $Q(p)$ can be expressed via the inverse of an incomplete Gamma function. The derivatives of $Q(p)$ are then easily obtained by using the $Q(p)$ value. Indeed, the following calculations follow easily by applying integration by parts in (18) and differentiating both sides of the equality with respect to p:

$$Q'(p) = \frac{4 + 2Q(p)}{(1-p)Q(p)}, \quad Q''(p) = \frac{2Q(p)(2 + Q(p)) - 4Q'(p)(1-p)}{(1-p)^2 Q(p)^2}.$$

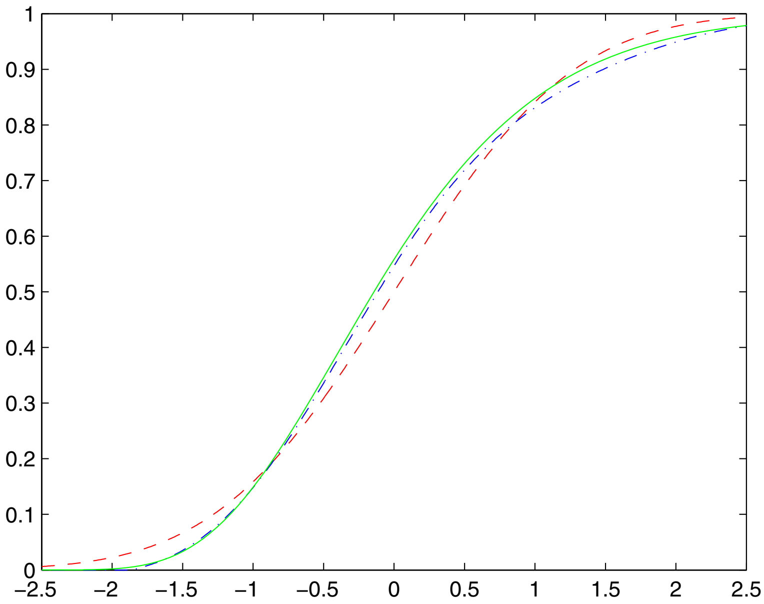

Fig. 3 Standard exponential distribution, $p = 0.9$, $n = 15$. True distribution *continuous line*, Edgeworth approximation: *dot-dashed line*, normal approximation: *dashed line*

The improvement can be felt for sample sizes as small as $n = 15$ and is virtually uniform all over the whole range of values of the argument (except for a small region in the very left tail where the Edgeworth expansion may become negative, in which case it should be equal to zero).

The "true" cumulative distribution function of the standardised random variable $\sqrt{n}(\hat{Q}_{p,h_n} - Q(p))/\sigma_n$ is compared to the cumulative distribution of the standard normal and to the Edgeworth in Table 1. The same number of simulations has been applied as in the previous examples. Numerical values are given for three scenarios: $n = 50$, $n = 30$ and $n = 15$, with $p = 0.9$. However, to keep the paper concise, we only present the graphical results for $n = 15$ (Fig. 4). It can be observed that the gains are impressive.

5 Proofs

The following remark plays an important role in the proofs below.

Remark 5 If $E|R|^c = O(n^{-1/2-c/2-\delta})$ holds for some $c > 0$ and $\delta > 0$, we have

$$P\left\{|R| > (\log n)^{-1} n^{-1/2}\right\} = o(n^{-1/2}).$$

Thus we can ignore R when we discuss asymptotic expansion up to the order $n^{-1/2}$.

Table 1 Chi-square sample

v	Normal	$p = 0.9, n = 50$		$p = 0.9, n = 30$		$p = 0.9, n = 15$	
		Edgeworth	True	Edgeworth	True	Edgeworth	True
−2.5000	0.0062	−0.0002	0.0016	−0.0013	0.0005	−0.0044	0.0002
−2.4000	0.0082	0.0007	0.0026	−0.0005	0.0009	−0.0042	0.0004
−2.3000	0.0107	0.0021	0.0039	0.0007	0.0016	−0.0034	0.0008
−2.2000	0.0139	0.0042	0.0059	0.0026	0.0027	−0.0002	0.0016
−2.1000	0.0179	0.0072	0.0086	0.0054	0.0044	0.0002	0.0030
−2.0000	0.0228	0.0111	0.0124	0.0092	0.0069	0.0036	0.0053
−1.9000	0.0287	0.0164	0.0173	0.0143	0.0105	0.0083	0.0087
−1.8000	0.0359	0.0231	0.0238	0.0209	0.0154	0.0147	0.0138
−1.7000	0.0446	0.0315	0.0319	0.0293	0.0220	0.0229	0.0208
−1.6000	0.0548	0.0419	0.0423	0.0397	0.0306	0.0334	0.0304
−1.5000	0.0668	0.0545	0.0547	0.0523	0.0415	0.0463	0.0430
−1.4000	0.0808	0.0695	0.0698	0.0675	0.0553	0.0620	0.0589
−1.3000	0.0968	0.0871	0.0874	0.0853	0.0718	0.0805	0.0782
−1.2000	0.1151	0.1074	0.1079	0.1059	0.0915	0.1021	0.1008
−1.1000	0.1357	0.1305	0.1313	0.1293	0.1143	0.1267	0.1277
−1.0000	0.1587	0.1563	0.1575	0.1557	0.1406	0.1544	0.1576
−0.9000	0.1841	0.1850	0.1868	0.1848	0.1700	0.1851	0.1910
−0.8000	0.2119	0.2162	0.2185	0.2166	0.2022	0.2185	0.2271
−0.7000	0.2420	0.2499	0.2525	0.2508	0.2371	0.2544	0.2657
−0.6000	0.2743	0.2857	0.2888	0.2871	0.2745	0.2924	0.3061
−0.5000	0.3085	0.3233	0.3265	0.3252	0.3137	0.3321	0.3478
−0.4000	0.3446	0.3622	0.3658	0.3646	0.3545	0.3729	0.3903
−0.3000	0.3821	0.4022	0.4062	0.4050	0.3963	0.4145	0.4328
−0.2000	0.4207	0.4427	0.4470	0.4457	0.4385	0.4561	0.4750
−0.1000	0.4602	0.4833	0.4880	0.4865	0.4808	0.4974	0.5166
0.0000	0.5000	0.5235	0.5281	0.5268	0.5226	0.5379	0.5571
0.1000	0.5398	0.5629	0.5675	0.5662	0.5632	0.5771	0.5961
0.2000	0.5793	0.6012	0.6061	0.6043	0.6025	0.6147	0.6332
0.3000	0.6179	0.6380	0.6431	0.6408	0.6403	0.6503	0.6683
0.4000	0.6554	0.6731	0.6782	0.6755	0.6762	0.6838	0.7010
0.5000	0.6915	0.7062	0.7113	0.7081	0.7100	0.7150	0.7318
0.6000	0.7257	0.7372	0.7427	0.7386	0.7417	0.7439	0.7602
0.7000	0.7580	0.7660	0.7717	0.7668	0.7710	0.7705	0.7864
0.8000	0.7881	0.7925	0.7986	0.7929	0.7981	0.7948	0.8101
0.9000	0.8159	0.8169	0.8232	0.8167	0.8226	0.8170	0.8317
1.0000	0.8413	0.8390	0.8456	0.8384	0.8447	0.8371	0.8513
1.1000	0.8643	0.8591	0.8656	0.8580	0.8648	0.8554	0.8691
1.2000	0.8849	0.8772	0.8836	0.8757	0.8827	0.8719	0.8848

Table 1 continued

v	Normal	$p = 0.9, n = 50$		$p = 0.9, n = 30$		$p = 0.9, n = 15$	
		Edgeworth	True	Edgeworth	True	Edgeworth	True
1.3000	0.9032	0.8935	0.8997	0.8917	0.8986	0.8869	0.8990
1.4000	0.9192	0.9080	0.9139	0.9060	0.9129	0.9005	0.9118
1.5000	0.9332	0.9209	0.9266	0.9187	0.9255	0.9127	0.9230
1.6000	0.9452	0.9323	0.9375	0.9301	0.9364	0.9238	0.9329
1.7000	0.9554	0.9424	0.9472	0.9402	0.9458	0.9338	0.9415
1.8000	0.9641	0.9512	0.9555	0.9490	0.9541	0.9428	0.9493
1.9000	0.9713	0.9589	0.9625	0.9569	0.9611	0.9509	0.9560
2.0000	0.9772	0.9656	0.9687	0.9637	0.9673	0.9581	0.9619
2.1000	0.9821	0.9714	0.9739	0.9697	0.9725	0.9645	0.9671
2.2000	0.9861	0.9764	0.9782	0.9748	0.9770	0.9702	0.9716
2.3000	0.9893	0.9807	0.9820	0.9793	0.9808	0.9751	0.9755
2.4000	0.9918	0.9843	0.9852	0.9831	0.9840	0.9794	0.9790
2.5000	0.9938	0.9873	0.9878	0.9863	0.9867	0.9831	0.9820

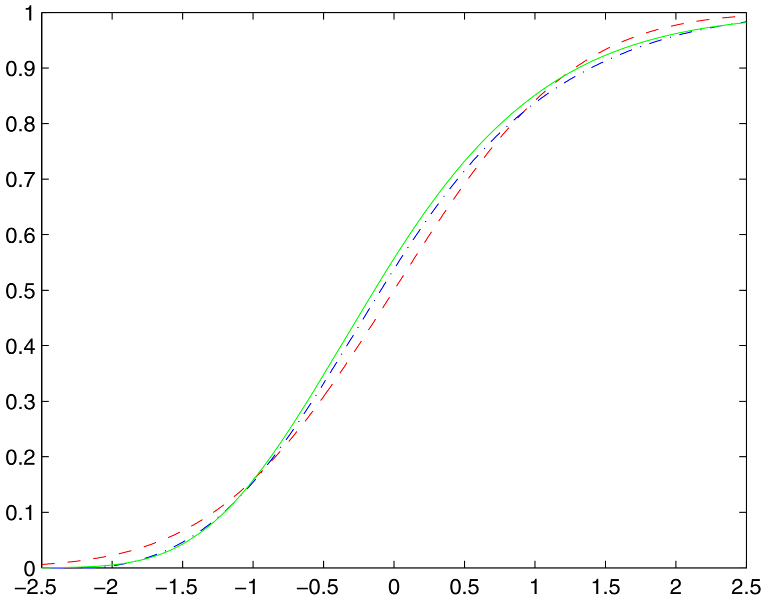

Fig. 4 Chi-square distribution, $df = 4$, $p = 0.9, n = 15$. True distribution *continuous line*, Edgeworth approximation: *dot-dashed line*, normal approximation: *dashed line*

Proof of Lemma 1 Let us denote $S_n(x, p) = \int_0^x K(\frac{s-p}{h_n})\mathrm{d}s$. Let n be sufficiently large. Changing the variables, we get:

$$
\begin{aligned}
\hat{Q}_{p,h_n} - \bar{Q}(p) &= \frac{1}{h_n}\int_{-\infty}^{\infty} x d(S_n(F_n(x), p) - S_n(F(x), p)) \\
&= -\frac{1}{h_n}\int_{-\infty}^{\infty} (S_n(F_n(x), p) - S_n(F(x), p))\mathrm{d}x \\
&= -\int_{-\infty}^{\infty}\left\{\int_{(F(x)-p)/h_n}^{(F_n(x)-p)/h_n} K(s)\mathrm{d}s\right\}\mathrm{d}x \\
&= \sum_{i=1}^{5} J_{in} + R_n
\end{aligned} \tag{19}
$$

where $J_{in} = -\frac{1}{i!}\int_{-\infty}^{\infty} K^{(i-1)}(\frac{F(x)-p}{h_n})\{\frac{F_n(x)-F(x)}{h_n}\}^i\mathrm{d}x$, $i = 1, 2, 3, 4, 5$ and

$$
\begin{aligned}
R_n = -\int_{-\infty}^{\infty}\Bigg\{\int_{(F(x)-p)/h_n}^{(F_n(x)-p)/h_n}&\left[K(u) - \sum_{j=0}^{4}\frac{1}{j!}K^{(j)}\left(\frac{F(x)-p}{h_n}\right)\right. \\
&\left.\times\left(u - \frac{F(x)-p}{h_n}\right)^j\right]\mathrm{d}u\Bigg\}\mathrm{d}x.
\end{aligned}
$$

It follows from the Lipshitz condition of $K^{(4)}(\cdot)$ and $\sigma_n^{-1} = O(1)$ that

$$
\begin{aligned}
&E\left(\sqrt{n}\sigma_n^{-1}R_n\right)^2 \\
&\leq Cnh_n^{-10-2\alpha}\int\int E\left[|F_n(x) - F(x)|^{5+\alpha}|F_n(y) - F(y)|^{5+\alpha}\right]\mathrm{d}x\,\mathrm{d}y \\
&\leq Cnh_n^{-10-2\alpha}\int\int\left\{E|F_n(x) - F(x)|^{10+2\alpha}E|F_n(y) - F(y)|^{10+2\alpha}\right\}^{1/2}\mathrm{d}x\,\mathrm{d}y.
\end{aligned}
$$

From the inequality (11), we have

$$
\begin{aligned}
&\left\{E|F_n(x) - F(x)|^{10+2\alpha}E|F_n(y) - F(y)|^{10+2\alpha}\right\}^{1/2} \\
&\quad\leq Cn^{-5-\alpha}\{F(x)(1 - F(x))\}^{1/2}\{F(y)(1 - F(y))\}^{1/2}.
\end{aligned}
$$

Thus, we can show that

$$
E(\sqrt{n}\sigma_n^{-1}R_n)^2 \leq Cnh_n^{-10-2\alpha}n^{-5-\alpha} = Cn^{-1/2-2/2-\alpha/2}(n^{1/4}h_n)^{-10-2\alpha}
$$

and then, using Remark 5, we obtain $\sqrt{n}\sigma_n^{-1}R_n = o_L(n^{-1/2})$.

From an equation in Shorack and Wellner (1986, p. 103), we have

$$J_{5n} = -\frac{1}{5!n^5h_n^4}\int_{-1}^{1} Q'(p+h_nx)K^{(4)}(x)\left\{\sum_{i=1}^{n}\hat{I}_x(Y_i)\right\}^5 \mathrm{d}x.$$

Utilizing the inequality (11) once more, we can get

$$\begin{aligned}
E\big(\sqrt{n}\sigma_n^{-1}J_{5n}\big)^2 &= O(n^{-9}h_n^{-8})\int_{-1}^{1}\int_{-1}^{1} Q'(p+h_nx)K^{(4)}(x)\\
&\qquad\times E\left[\left|\sum_{i=1}^{n}\hat{I}_x(Y_i)\right|^5\left|\sum_{i=1}^{n}\hat{I}_y(Y_i)\right|^5\right]\mathrm{d}x\,\mathrm{d}y\\
&= O(n^{-4}h_n^{-8}) = O\big(n^{-1/2-2/2}n^{-1/2}(n^{1/4}h_n)^{-8}\big).
\end{aligned}$$

By the assumption on h_n, it holds $n^{-1/2}(n^{1/4}h_n)^{-8} = o(1)$ and therefore we see that $\sqrt{n}\sigma_n^{-1}J_{5n} = o_L(n^{-1/2})$ holds.

Similarly, we have

$$\begin{aligned}
&E\big(\sqrt{n}\sigma_n^{-1}J_{4n}\big)^3\\
&= O\big(n^{-21/2}h_n^{-9}\big)\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} Q'(p+h_nx)K^{(3)}(x)\\
&\quad\times E\left[\left|\sum_{i=1}^{n}\hat{I}_x(Y_i)\right|^4\left|\sum_{i=1}^{n}\hat{I}_y(Y_i)\right|^4\left|\sum_{i=1}^{n}\hat{I}_z(Y_i)\right|^4\right]\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\\
&= O(n^{-9/2}h_n^{-9}) = O\big(n^{-1/2-3/2}n^{-1/4}(n^{1/4}h_n)^{-9}\big).
\end{aligned}$$

Thus we see that $\sqrt{n}\sigma_n^{-1}J_{4n} = o_L(n^{-1/2})$ also holds.

Next we consider J_{3n}. Let us rewrite

$$\begin{aligned}
\sqrt{n}\sigma_n^{-1}J_{3n} &= -\frac{1}{6}d_{3n}\int_{-1}^{1} Q'(p+h_nx)K^{(2)}(x)\left\{\sum_{i=1}^{n}\hat{I}_x(Y_i)\right\}^3\mathrm{d}x\\
&= d_{3n}(J_{3n;1}+J_{3n;2}+A_{3n})
\end{aligned}$$

where

$$\begin{aligned}
J_{3n;1} &= -\frac{1}{6}\sum_{i=1}^{n}\int_{-1}^{1} Q'(p+h_nx)K^{(2)}(x)\hat{I}_x^3(Y_i)\mathrm{d}x,\\
J_{3n;2} &= -\frac{1}{2}\sum_{C_{n,2}}\int_{-1}^{1} Q'(p+h_nx)K^{(2)}(x)\big\{\hat{I}_x^2(Y_i)\hat{I}_x(Y_j)+\hat{I}_x(Y_i)\hat{I}_x^2(Y_j)\big\}\mathrm{d}x.
\end{aligned}$$

Using $\tilde{g}_{1n}(y)$ with $E(\tilde{g}_{1n}(Y_i)) = 0$, we can rewrite $d_{3n} J_{3n;1}$

$$d_{3n} J_{3n;1} = d_{3n} \sum_{i=1}^{n} \tilde{g}_{1n}(Y_i) - d_{3n} E(J_{3n;1}).$$

Using the inequality (10), it is easy to see that $d_{3n} J_{3n;1} = o_L(n^{-1/2})$. Further applying the H-decomposition (8), we have

$$d_{3n} J_{3n;2} = d_{3n}(n-1)\hat{A}_{1n} + o_L(n^{-1/2}).$$

Similarly we can show that

$$\sqrt{n}\sigma_n^{-1} J_{2n} = d_{2n} A_{2n} + \frac{\delta}{\sigma\sqrt{n}} + o_L(n^{-1/2})$$

and

$$\sqrt{n}\sigma_n^{-1} J_{1n} = d_{1n} A_{1n}.$$

Thus we have the desired result. □

Proof of Theorem 1 Let us define

$$\psi_n(t) = E\left[\exp\left\{it\left(\sum_{i=1}^{3} d_{in} A_{in} + d_{3n}(n-1)\hat{A}_{1n}\right)\right\}\right]$$

and

$$\eta_n(t) = E[\exp\{itg_{1n}(Y_1)\}].$$

We observe that

$$E[g_{2n}(Y_1, Y_2)|Y_1] = E[g_{3n}(Y_1, Y_2, Y_3)|Y_1, Y_2] = 0 \quad a.s. \tag{20}$$

and then for $k \geq 2$

$$\begin{gathered} E|A_{1n}|^k \leq Cn^{k/2}, \quad E|\hat{A}_{1n}|^k \leq Cn^{k/2}, \\ E|A_{2n}|^k \leq Cn^k \quad \text{and} \quad E|A_{3n}|^k \leq Cn^{3k/2}. \end{gathered}$$

The main tool used by us in obtaining the desired Edgeworth expansion and in proving the validity is Esseen's smoothing lemma. Let us define

$$\begin{aligned}\hat{\psi}(t) &= \int_{-\infty}^{\infty} \exp\{itx\}\mathrm{d}G_n(x)\\ &= \exp\left\{-\frac{t^2}{2}\right\}\left\{1 + \frac{(it)^2}{4nh_n^2\sigma_n^2}(4e_{5n} + e_{6n}) + \frac{(it)^3}{6n^{1/2}\sigma_n^3}\left(e_{1n} + \frac{3e_{2n}}{h_n}\right)\right.\\ &\quad \left. + \frac{(it)^4}{6nh_n^2\sigma_n^4}(3e_{2n} + e_{4n}) + \frac{(it)^6}{8nh_n^2\sigma_n^6}e_{2n}^2\right\}.\end{aligned}$$

It is easy to see that

$$\int_{n^{1/18}}^{n^{1/2}\log n} \frac{|\hat{\psi}(t)|}{t}\mathrm{d}t = o(n^{-1/2}).$$

Thus, if we can show

$$\text{(I)} = \int_0^{n^{1/18}} \frac{|\psi_n(t) - \hat{\psi}(t)|}{t}\mathrm{d}t = o(n^{-1/2})$$

and

$$\text{(II)} = \int_{n^{1/18}}^{n^{1/2}\log n} \frac{|\psi_n(t)|}{t}\mathrm{d}t = o(n^{-1/2}),$$

we have the desired result. To show that (I) $= o(n^{-1/2})$, we make use of the same evaluation method that has been already applied in the papers Callaert et al. (1980) and Maesono (1987).

Evaluation of (I)

Let us define

$$\psi_{1n}(t) = E\left[\exp\left\{it\sum_{k=1}^{3} d_{kn}A_{kn}\right\}\left(1 + itd_{3n}(n-1)\hat{A}_{1n}\right)\right].$$

Since

$$|\psi_n(t) - \psi_{1n}(t)| \le \frac{t^2}{2}d_{3n}^2(n-1)^2E(\hat{A}_{1n}^2) = O\left(n^{-5}h_n^{-4}(n-1)^2n\right)t^2,$$

we have

$$\int_0^{n^{1/18}} \frac{|\psi_n(t) - \psi_{1n}(t)|}{t}\mathrm{d}t = O\left(n^{-1/2}n^{-7/18}(n^{1/4}h_n)^{-4}\right) = o(n^{-1/2}).$$

Putting

$$\psi_{2n}(t) = E\left[\exp\left\{it\sum_{k=1}^{3} d_{kn}A_{kn}\right\}\right] + E\left[\exp\{itd_{1n}A_{1n}\}itd_{3n}(n-1)\hat{A}_{1n}\right],$$

we have

$$\begin{aligned}|\psi_{1n}(t)-\psi_{2n}(t)| &\le t^2 d_{2n}d_{3n}(n-1)E|A_{2n}\hat{A}_{1n}|+t^2 d_{3n}^2(n-1)E|A_{3n}\hat{A}_{1n}|\\ &= O(n^{-4}h_n^{-3}nn^{3/2})t^2+O(n^{-5}h_n^{-4}nn^2)t^2.\end{aligned}$$

Thus we get

$$\int_0^{n^{1/18}}\frac{|\psi_{1n}(t)-\psi_{2n}(t)|}{t}\mathrm{d}t=o(n^{-1/2}).$$

Similarly we can show that

$$\int_0^{n^{1/18}}\frac{|\psi_{2n}(t)-\psi_{3n}(t)|}{t}\mathrm{d}t=o(n^{-1/2})$$

where

$$\psi_{3n}(t)=E\left[\exp\left\{it\sum_{k=1}^{2}d_{kn}A_{kn}\right\}\right]+E\Big[\exp\{itd_{1n}A_{1n}\}itd_{3n}\{(n-1)\hat{A}_{1n}+A_{3n}\}\Big].$$

Since

$$\begin{aligned}\int_0^{n^{1/18}}\frac{t^3E|d_{2n}^3A_{2n}^3|}{6t}\mathrm{d}t &= O\big(n^{-9/2}h_n^{-3}n^3n^{3/18}\big)\\ &= O\big(n^{-1/2}n^{-1/12}(n^{1/4}h_n)^{-3}\big)=o(n^{-1/2}),\end{aligned}$$

it is sufficient to obtain an approximation of

$$\begin{aligned}\psi_{4n}(t) &= E\left[\exp\{itd_{1n}A_{1n}\}\sum_{j=0}^{2}\frac{(it)^j}{j!}d_{2n}^jA_{2n}^j\right]\\ &\quad +E\Big[\exp\{itd_{1n}A_{1n}\}itd_{3n}\{(n-1)\hat{A}_{1n}+A_{3n}\}\Big].\end{aligned}$$

It is easy to see that

$$\begin{aligned}\psi_{4n}(t) &= \Psi_0^*+itd_{2n}\frac{n(n-1)}{2}\Psi_2^*E_1^*+\frac{(it)^2}{2}d_{2n}^2\frac{n(n-1)}{2}\Psi_2^*E_2^*\\ &\quad +\frac{(it)^2}{2}d_{2n}^2n(n-1)(n-2)\Psi_3^*E_3^*\\ &\quad +\frac{(it)^2}{2}d_{2n}^2\frac{n(n-1)(n-2)(n-3)}{4}\Psi_4^*E_4^*+itd_{3n}(n-1)n\Psi_1^*E_5^*\\ &\quad +itd_{3n}\frac{n(n-1)(n-2)}{6}\Psi_3^*E_6^*\end{aligned}$$

where

$$
\begin{aligned}
\Psi_k^* &= \eta_n^{n-k}(td_{1n}),\\
E_1^* &= E[\exp\{itd_{1n}[g_{1n}(Y_1)+g_{1n}(Y_2)]\}g_{2n}(Y_1,Y_2)],\\
E_2^* &= E[\exp\{itd_{1n}[g_{1n}(Y_1)+g_{1n}(Y_2)]\}g_{2n}^2(Y_1,Y_2)],\\
E_3^* &= E[\exp\{itd_{1n}[g_{1n}(Y_1)+g_{1n}(Y_2)+g_{1n}(Y_3)]\}g_{2n}(Y_1,Y_2)g_{2n}(Y_1,Y_3)],\\
E_4^* &= \left(E[\exp\{itd_{1n}[g_{1n}(Y_1)+g_{1n}(Y_2)]\}g_{2n}(Y_1,Y_2)]\right)^2,\\
E_5^* &= E[\exp\{itd_{1n}[g_{1n}(Y_1)]\}\hat{g}_{1n}(Y_1)]
\end{aligned}
$$

and

$$E_6^* = E[\exp\{itd_{1n}[g_{1n}(Y_1)+g_{1n}(Y_2)+g_{1n}(Y_3)]\}g_{3n}(Y_1,Y_2,Y_3)].$$

Similarly to Callaert et al. (1980), we have that for $0 \le t \le \varepsilon n^{1/2}$

$$|\Psi_k^* - \Psi_k| \le \gamma_n n^{-1} P(t)\exp\{-at^2\}$$

where

$$
\begin{aligned}
\Psi_k = \exp\left\{-\frac{t^2}{2}\right\}\Bigg(1 &- \frac{(it)^2}{2}k + \frac{(it)^3}{6n^{1/2}\sigma_n^3}E[g_{1n}^3(Y_1)]\\
&+ \frac{(it)^4}{24n\sigma_n^4}\left\{E[g_{1n}^4(Y_1)] - 3\sigma_n^4\right\} + \frac{(it)^6}{72n\sigma_n^6}\left\{E[g_{1n}^3(Y_1)]\right\}^2\Bigg),
\end{aligned}
$$

a and ε are positive constants, $\gamma_n \to 0$ as $n \to \infty$, and $P(t)$ is a polynomial in t.

Putting

$$
\begin{aligned}
E_1 = E[g_{2n}(Y_1,Y_2)] &+ itd_{1n}E[\{g_{1n}(Y_1)+g_{1n}(Y_2)\}g_{2n}(Y_1,Y_2)]\\
&+ \frac{(it)^2}{2}d_{1n}^2E[\{g_{1n}(Y_1)+g_{1n}(Y_2)\}^2g_{2n}(Y_1,Y_2)],
\end{aligned}
$$

we have

$$|E_1^* - E_1| \le Cn^{-3/2}|t|^3.$$

It follows from Eq. (20) that

$$E_1 = (it)^2 d_{1n}^2 E[g_{1n}(Y_1)g_{1n}(Y_2)g_{2n}(Y_1,Y_2)].$$

Similarly we can show that

$$
\begin{aligned}
&|E_2^* - E_2| \le Cn^{-1/2}|t|, \quad |E_3^* - E_3| \le Cn^{-3/2}|t|^3,\\
&|E_4^* - E_4| \le Cn^{-3}t^6, \quad |E_5^* - E_5| \le Cn^{-1}t^2 \text{ and } |E_6^* - E_6| \le Cn^{-2}t^4.
\end{aligned}
$$

where

$$
\begin{aligned}
E_2 &= E[g_{2n}^2(Y_1, Y_2)],\\
E_3 &= (it)^2 d_{1n}^2 E[g_{1n}(Y_2) g_{1n}(Y_3) g_{2n}(Y_1, Y_2) g_{2n}(Y_1, Y_3)],\\
E_4 &= \left((it)^2 d_{1n}^2 E[g_{1n}(Y_1) g_{1n}(Y_2) g_{2n}(Y_1, Y_2)]\right)^2,\\
E_5 &= it d_{1n} E[g_{1n}(Y_1) \hat{g}_{1n}(Y_1)]
\end{aligned}
$$

and

$$E_6 = (it)^3 d_{1n}^3 E[g_{1n}(Y_1) g_{1n}(Y_2) g_{1n}(Y_3) g_{3n}(Y_1, Y_2, Y_3)].$$

Similarly to Callaert et al. (1980) using the approximations $\Psi_k, k = 0, \ldots, 3$ and $E_i, i = 1, 2, \ldots, 6$ we can show that

$$\int_0^{n^{1/18}} \frac{|\psi_{4n}(t) - \hat{\psi}(t)|}{t} \mathrm{d}t = o(n^{-1/2})$$

and thus

$$\int_0^{n^{1/18}} \frac{|\psi_n(t) - \hat{\psi}(t)|}{t} \mathrm{d}t = o(n^{-1/2}). \tag{21}$$

Evaluation of (II)
We need an evaluation of $|\psi_n(t)|$ form above. Since the evaluation involves some technical details, we will formulate the final result in the following Lemma 2 below. The proof of the lemma follows the steps outlined in Callaert et al. (1980) and Maesono (1987, p. 192), and makes use of the inequalities (10) and (11).

Lemma 2 *If the conditions of Theorem 1 are satisfied then there exist positive constants $M_s (s = 1, \ldots, 10)$ such that for all t $(-\infty < t < \infty)$, all integers n and m with $8 < m < n - 2$*

$$
\begin{aligned}
|\psi_n(t)| \le |\eta_n(td_{1n})|^{m-8} \Bigg(& M_1 \sum_{s=0}^{4} |t|^s d_{2n}^s m^{s/2} n^{s/2} + M_2 |t| d_{3n} m^{1/2} n^2 \\
& + M_3 t^2 d_{2n} d_{3n} m^2 n^3 + M_4 \sum_{s=0}^{2} |t|^{s+1} d_{2n}^s d_{3n} (n-1) m^{s+1} n^s \Bigg) \\
& + M_5 |t|^5 d_{2n}^5 m^{5/2} n^{5/2} + M_6 |t|^3 d_{2n}^2 d_{3n} m^{3/2} n^2 + M_7 t^2 d_{3n}^2 m n^2 \\
& + M_8 t^4 d_{2n}^3 d_{3n} (n-1) m^2 n^{3/2} + M_9 t^2 d_{3n}^2 (n-1) m n \\
& + M_{10} t^2 d_{3n}^2 (n-1)^2 m.
\end{aligned}
$$

The idea behind the proof of (II) is to decompose into two terms

$$\text{(II)} = \int_{n^{1/18}}^{n^{1/2}\log n} \frac{|\psi_n(t)|}{t}\mathrm{d}t = \int_{n^{1/18}}^{\epsilon n^{1/2}} \frac{|\psi_n(t)|}{t}\mathrm{d}t + \int_{\epsilon n^{1/2}}^{n^{1/2}\log n} \frac{|\psi_n(t)|}{t}\mathrm{d}t$$

for a suitably chosen $\epsilon > 0$ and to evaluate the two terms separately. Let us define for some $\epsilon > 0$

$$\text{(III)} = \int_{n^{1/18}}^{\epsilon n^{1/2}} \frac{|\psi_n(t)|}{t}\mathrm{d}t \quad \text{and} \quad \text{(IV)} = \int_{\epsilon n^{1/2}}^{n^{1/2}\log n} \frac{|\psi_n(t)|}{t}\mathrm{d}t.$$

Similarly to Callaert et al. (1980), we can claim that there exists an $\epsilon > 0$ such that

$$|\eta_n(\theta)| \leq \exp\left(-\frac{1}{3}\theta^2\sigma_n^2\right) \quad \text{for } |\theta| \leq \epsilon/\sigma_n.$$

Thus we have that for $0 \leq t \leq \epsilon n^{1/2}$

$$|\eta_n(td_{1n})| \leq \exp\left(-\frac{t^2}{3n}\right). \tag{22}$$

Using Lemma 2, we will prove that (III) $= o(n^{-1/2})$ holds. Further let us define

$$(\text{III}_1) = \int_{n^{1/18}}^{n^{5/8}h_n^2} \frac{|\psi_n(t)|}{t}\mathrm{d}t, \quad (\text{III}_2) = \int_{n^{5/8}h_n^2}^{nh_n^3} \frac{|\psi_n(t)|}{t}\mathrm{d}t,$$
$$(\text{III}_3) = \int_{nh_n^3}^{n^{3/8}} \frac{|\psi_n(t)|}{t}\mathrm{d}t, \quad (\text{III}_4) = \int_{n^{3/8}}^{\epsilon n^{1/2}} \frac{|\psi_n(t)|}{t}\mathrm{d}t.$$

Choosing $m = [n^{8/9+\mu}]$, $0 < \mu < \frac{1}{45}$, where $[\cdot]$ denotes the largest integer less than $n^{8/9+\mu}$, we can show that $(\text{III}_1) = o(n^{-1/2})$ as follows. Using (22), we have that for $t \in [n^{1/18}, n^{5/8}h_n^2]$

$$|\eta_n(td_{1n})|^{m-8} \leq \exp\left(-\frac{(m-8)n^{1/9}}{3n}\right).$$

By our choice of m, we can show that for any k

$$n^k \exp\left(-\frac{(m-8)n^{1/9}}{3n}\right) = o(n^{-1/2})$$

and

$$\log(n^{5/8}h_n^2)\exp\left(-\frac{(m-8)n^{1/9}}{3n}\right) = o(n^{-1/2}).$$

Thus all terms in Lemma 2 that are multiplied by $|\eta_n(td_{1n})|^{m-8}$, become of order $o(n^{-1/2})$ after their integration within the interval $(n^{1/18}, n^{5/8}h_n^2)$. Moreover, by direct evaluation, we have

$$\int_{n^{1/18}}^{n^{5/8}h_n^2} t^{-1}M_5|t|^5 d_{2n}^5 m^{5/2} n^{5/2} \mathrm{d}t = O(n^{5\mu/2-15/16}(n^{1/4}h_n)^5) = o(n^{-1/2}).$$

Similarly we can show that the remaining terms are all $o(n^{-1/2})$, and this leads us to $(\mathrm{III}_1) = o(n^{-1/2})$.

For (III_2), by choosing $m = [n^{3/4+\mu}]$, $0 < \mu < \frac{1}{20}$, we have immediately $(\mathrm{III}_2) = o(n^{-1/2})$. If we choose $m = [n^{1/2+\mu}](0 < \mu < \frac{1}{20})$ and $m = [n^{\mu}](\frac{1}{4} < \mu < \frac{3}{10})$, we can show that $(\mathrm{III}_3) = o(n^{-1/2})$ and $(\mathrm{III}_4) = o(n^{-1/2})$, respectively.

Now we concentrate on evaluating (IV). Let us define

$$g_{1n}^{*}(Y_i) = -\int_{-1}^{1} Q'(p)K(x)\hat{I}_x(Y_i)\mathrm{d}x, \quad g_{1n}^{**}(Y_i) = -\int_{-1}^{1} Q''(p)xK(x)\hat{I}_x(Y_i)\mathrm{d}x,$$
$$g_{1n}^{***}(Y_i) = -\int_{-1}^{1} Q^{(3)}(p)x^2K(x)\hat{I}_x(Y_i)\mathrm{d}x,$$
$$\hat{g}_{1n}^{*}(Y_i) = \hat{g}_{1n}(Y_i) - \frac{n^2h_n^4}{2(n-1)}g_{1n}^{***}(Y_i), \quad \hat{A}_{1n}^{*} = \sum_{i=1}^{m}\hat{g}_{1n}^{*}(Y_i)$$

and, for a suitably chosen m to be defined below, the characteristic function

$$\psi_n^{*}(t) = E\left[\exp\left\{it\left[d_{1n}\sum_{i=1}^{m}g_{1n}^{*}(Y_i) + d_{1n}h_n\sum_{i=1}^{m}g_{1n}^{**}(Y_i) + d_{3n}(n-1)\hat{A}_{1n}^{*}\right.\right.\right.$$
$$\left.\left.\left. + \sum_{i=m+1}^{n}\{d_{1n}g_{1n}(Y_i) + d_{3n}(n-1)\hat{g}_{1n}(Y_i)\} + d_{2n}A_{2n} + d_{3n}A_{3n}\right]\right\}\right].$$

Using the Taylor expansion and uniform boundedness of $Q^{(4)}(\cdot)$ around p, we can show that for a constant M_{11} the inequality

$$|\psi_n(t) - \psi_n^{*}(t)| \le M_{11}|t|d_{1n}h_n^3m^{1/2}$$

holds. Note that $n^2h_n^4/2(n-1) = o((\log n)^{-4})$. Similarly as in Lemma 2, we can show that

$$\begin{aligned} |\psi_n^{*}(t)| \le\ & |\eta_n^{*}(td_{1n})|^{m-8}P(m,n,t) + M_5|t|^5d_{2n}^5m^{5/2}n^{5/2} \\ & + M_6|t|^3d_{2n}^2d_{3n}m^{3/2}n^2 + M_7t^2d_{3n}^2mn^2 \\ & + M_8t^4d_{2n}^3d_{3n}(n-1)m^2n^{3/2} + M_9t^2d_{3n}^2(n-1)mn \\ & + M_{10}t^2d_{3n}^2(n-1)^2m + M_{11}|t|d_{1n}h_n^3m^{1/2} + M_{12}|t|^5d_{1n}^5h_n^5m^{5/2} \quad (23) \end{aligned}$$

where $\eta_n^*(t) = E[\exp\{itg_{1n}^*(Y_i)\}]$, M_{12} is a constant, and $P(m, n, t)$ is a polynomial of m, n and t.

Now we state why the term containing $|\eta_n^*(td_{1n})|^{m-8}$ can become exponentially dominant. It should be noted that a very careful evaluation is needed to this end. The difficulty arises from the fact that for fixed n, the random variable $g_{1n}^*(Y_1)$ has a non-lattice distribution for fixed n but this distribution becomes lattice in the limit when $n \to \infty$ and a special treatment is needed to show that yet a Cramèr type condition is satisfied. This treatment is demonstrated below. First, we note that $\eta_n^*(t) = e^{itE(Z_1)}E[\exp\{it(Z_1)\}]$ where

$$Z_1 = -\int_{-1}^{1} Q'(p)K(x)I(Y_1 \le p + h_n x)\mathrm{d}x.$$

Then we have $|\eta_n^*(t)| \le |E[e^{itZ_1}]|$. We will evaluate the characteristic function of Z_1 via conditioning on the event $D = \{p - h_n \le Y_1 \le p + h_n\}$. For sufficiently large n, we have

$$\begin{aligned}
|E[e^{itZ_1}]| &= |P(D)E[e^{itZ_1}|D] + P(D^c)E[e^{itZ_1}|D^c]| \\
&= |2h_n E[e^{itZ_1}|D] + (1 - 2h_n)E[e^{itZ_1}|D^c]| \\
&\le 2h_n |E[e^{itZ_1}|D]| + (1 - 2h_n).
\end{aligned}$$

Let us define the random variable $W_1 = (Y_1 - p)/h_n$. Its conditional distribution given D is

$$q(w|D) = \begin{cases} \frac{1}{2}, & -1 \le w \le 1 \\ 0, & \text{others} \end{cases}$$

which *does not depend* on n. Since

$$Z_1 = -Q'(p)\int_{(Y_1-p)/h_n}^{1} K(x)\mathrm{d}x = -Q'(p)\int_{W_1}^{1} K(x)\mathrm{d}x,$$

the conditional distribution of Z_1 also *does not depend* on n and *is non-lattice*. This implies that for $\varepsilon n^{1/2} \le t$ and sufficiently large n, there exist a constant $0 \le c < 1$ such that

$$|\eta_n^*(t)| \le 1 - 2(1 - c)h_n.$$

Note that

$$|\eta_n^*(t)|^{m-8} \le \left\{[1 - 2(1 - c)h_n]^{h_n^{-1}}\right\}^{h_n(m-8)}.$$

Since $\lim_{n\to\infty}[1-2(1-c)h_n]^{h_n^{-1}} = e^{-2(1-c)}$, choosing $m = [n^{\mu}h_n^{-1}](0 < \mu < \frac{1}{30})$ in (23), we can show that (IV)$= o(n^{-1/2})$ because for any constant $k > 0$

$$n^k\left\{[1-2(1-c)h_n]^{h_n^{-1}}\right\}^{h_n(m-8)} = o(n^{-1/2}).$$

Thus $\int_{\varepsilon n^{1/2}}^{n^{1/2}\log n} |\eta_n^*(td_{1n})|^{m-8}P(m,n,t)\mathrm{d}t = o(n^{-1/2})$ holds. We can also evaluate the contributions of the remaining terms in (23) after the integration in $(\epsilon n^{1/2}, n^{1/2}\log n)$ is carried out. Each of these contributions is of order $o(n^{-1/2})$. For instance

$$\begin{aligned}\int_{\varepsilon n^{1/2}}^{n^{1/2}\log n} & t^{-1}M_5t^5d_{2n}^5m^{5/2}n^{5/2}\mathrm{d}t \\ &\leq Cn^{5/2}(\log n)^5n^{-15/2}h_n^{-5}n^{5\mu/2}h_n^{-5/2}n^{5/2} = o(n^{-1/2}).\end{aligned}$$

Thus $\int_{\epsilon n^{1/2}}^{n^{1/2}\log n}\frac{|\psi_n(t)|}{t}\mathrm{d}t = o(n^{-1/2})$ holds and we have the desired result. □

Acknowledgments The authors are very grateful to the two referees for their careful reading of an earlier version of this paper and for their useful comments and suggestions.

References

Bickel, P., Götze, F., van Zwet, W. (1986). The Edgeworth expansion for U-statistics of degree two. *The Annals of Statistics, 14*, 1463–1484.

Borovkov, K., Dehling, H., Pfeifer, D. (1996). On asymptotic behavior of weighted sample quantiles. *Mathematical Methods of Statistics, 5*, 173–186.

Callaert, H., Janssen, P., Veraverbeke, N. (1980). An Edgeworth expansion for U-statistics. *The Annals of Statistics, 8*, 299–312.

Dharmadhikari, S., Fabian, V., Jogdeo, K. (1968). Bounds on the moments of martingales. *The Annals of Mathematical Statistics, 39*, 1719–1723.

DiDonato, A., Morris, A. (1987). ALGORITHM 654: FORTRAN subroutines for computing the incomplete gamma function ratios and their inverse. *Association for Computing Machinery Transactions on Mathematical Software, 13*(3), 318–319.

Falk, M. (1984). Relative deficiency of kernel type estimators of quantiles. *The Annals of Statistics, 12*(1), 261–268.

Falk, M. (1985). Asymptotic normality of the kernel quantile estimator. *The Annals of Statistics, 13*(1), 428–433.

Feller, W. (1971). *An introduction to probability and its applications* (Vol. 2). New York: Wiley.

Helmers, R. (1982). Edgeworth expansions for linear combinations of order statistics. *Mathematical Centre Tracts* (Vol. 105). Amsterdam: Mathematisch Centrum.

Jing, B.-Y., Wang, Q. (2003). Edgeworth expansions for U-statistics under minimal conditions. *The Annals of Statistics, 31*(4), 1376–1391.

Kozek, A. (2005). How to vombine M-estimators to estimate quantiles and a score function. *Sankhya, 67*(2), 277–294.

Lai, T., Wang, J. (1993). Edgeworth expansions for symmetric statistics with applications to bootstrap methods. *Statistica Sinica, 3*, 517–542.

Lorenz, M. (1905). Methods of measuring concentration of wealth. *Journal of the American Statistical Association, 9*, 209–219.

Maesono, Y. (1987). Edgeworth expansion for one-sample U-statistics. *Bulletin of Informatics and Cybernetics, 22*, 189–197.

Ogryczak, W., Ruszczyński, A. (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal of Optimization, 13*(1), 60–78.

Serfling, R. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley.
Sheather, S., Marron, J. (1990). Kernel quantile estimators. *Journal of the American Statistical Association, 85*(410), 410–416.
Shorack, G., Wellner, J. (1986). *Empirical processes with applications to statistics*. New York: Wiley.
Xiang, X. (1995a). A Berry-Esseen theorem for the kernel quantile estimator with application to studying the deficiency of qunatile estimators. *Annals of the Institute of Statistical Mathematics, 47*(2), 237–251.
Xiang, X. (1995b). Defficiency of the sample quantile estimator with respect to kernel quantile estimators for censored data. *The Annals of Statistics, 23*(3), 836–854.
Xiang, X., Vos, P. (1997). Quantile estimators and covering probabilities. *Journal of Nonparametric Statistics, 7*, 349–363.