The Annals of Statistics
1984, Vol. 12, No. 1, 261–268

RELATIVE DEFICIENCY OF KERNEL TYPE ESTIMATORS OF QUANTILES

BY MICHAEL FALK

Universität-Gesamthochschule Siegen

In this paper the asymptotic relative deficiency of the sample q-quantile with respect to kernel type estimators of the q-quantile is evaluated. The comparison is based on the mean square errors of the estimators. The result suggests a purely analytic measure of performance within the class of kernels. It is notable that a similar situation occurs when kernel estimators of a distribution function are studied.

1. Introduction and main result. Let P be a probability measure on the real line with distribution function F. The empirical estimator of the q-quantile, say $q(F)$, is given by the sample q-quantile $q_n := q(F_n) = Z_{r_n:n}$, where $r_n = \min\{j \in \{1, \cdots, n\}: j/n \geq q\}$, $Z_{i:n}$ denotes the ith order statistic in a sample of n independent random variables identically distributed according to P, and F_n is the accompanying empirical distribution function.

Sample quantiles have been extensively studied in the statistical literature; references can be found in the books by David (1981) and Galambos (1978).

For obvious reasons one might hope that averaging over order statistics close to the sample q-quantile leads to estimators of better performance. This idea was carried out by Reiss (1980a) who proved that the asymptotic relative deficiency of the sample q-quantile with respect to a linear combination of finitely many order statistics quickly tends to infinity as the sample size increases.

Averaging over all order statistics leads to kernel type estimators

$$\hat{q}_n(F_n) := \int_0^1 F_n^{-1}(x)\alpha_n^{-1}k\left(\frac{q-x}{\alpha_n}\right)dx$$

for an appropriate kernel k and a bandwidth $\alpha_n > 0$ where hereafter G^{-1} denotes the generalized inverse of a distribution function G, i.e. $G^{-1}(p) := \inf\{t \in \mathbb{R}: G(t) \geq p\}$, $p \in (0, 1)$.

Estimators of this form are extensively studied in the literature of nonparametric density estimation (see, for example, Scott et al., 1977, and Wertz, 1978). The kernel estimator of the q-quantile is mentioned in Parzen (1979), page 113, and Reiss (1980b), and a "discrete" version was also used in Reiss (1980b) for testing the hypothesis $q(F) < r$ against the alternative $q(F) > r + C_n n^{-1/2}$.

In the present paper we investigate the mean square errors (MSE) of q_n and $\hat{q}_n$, i.e. $E((q_n - q(F))^2)$ and $E((\hat{q}_n - q(F))^2)$, respectively, and establish an asymptotic representation of the relative deficiency $i(n) - n$ of q_n with respect

Received March 1983; revised August 1983.

AMS 1980 *subject classifications.* Primary 62G05; secondary 62G20, 62G30.

Key words and phrases. Sample q-quantile, kernel estimator, mean square error, deficiency, Bahadur approximation.

to $\hat{q}_n$, where $i(n)$ is defined by

$$i(n) := \min\{j \in \mathbb{N}: \mathrm{MSE}(q_j) \le \mathrm{MSE}(\hat{q}_n)\}.$$

Our main result is the following.

THEOREM. *Assume that* $\lim_{t\to\infty} t^{\delta}(1 - F(t) + F(-t)) = 0$ *for some* $\delta > 0$ *and that* F^{-1}*is* $(m + 1)$*-times differentiable in a neighborhood of* $q \in (0, 1)$, $m \ge 2$, *with bounded* $(m + 1)$th *derivative and* $(F^{-1})'(q) > 0$. *Assume further that the kernel* k *has finite support* $[-c, c]$ *and fulfills* $\int k(x)\,dx = 1$, $\int x^i k(x)\,dx = 0$, $i = 1, \cdots, m$. *Then, if* $\alpha_n n^{1/4} \to_{n\in\mathbb{N}} \infty$ *and* $\alpha_n n^{1/(2m+1)} \to_{n\in\mathbb{N}} 0$, $\mathrm{MSE}(q_n)$ *and* $\mathrm{MSE}(\hat{q}_n)$ *are finite if* n *is large and*

$$\lim_{n\in\mathbb{N}}\left(\frac{i(n) - n}{n\alpha_n}\right) = 2\int xk(x)K(x)\,dx/(q(1 - q))$$

where $K(x) := \int_{-c}^{x} k(y)\,dy$.

Notice that this result remains true if the sample q-quantile is replaced by $Z_{r_n:n}$, where $r_n \in \{1, \cdots, n\}$, $n \in \mathbb{N}$, fulfills $|q - r_n/n| = O(n^{-5/8})$.

The number $\psi(k) := 2\int xk(x)K(x)\,dx$ can obviously be regarded as a measure of the asymptotic performance within the class of kernels and its sign determines whether one does better with the sample q-quantile, i.e. if $\psi(k) < 0$, or with the corresponding kernel estimator, i.e. if $\psi(k) > 0$. In either case $i(n) - n$ is of order $n\alpha_n$ if $\psi(k)$ does not vanish.

The functional ψ occurs also as a measure of performance when kernel estimators of a smooth distribution function are considered (see Reiss, 1981, and Falk, 1983). A discrete analogue to ψ is given in Reiss (1980b), formula (3.7).

Denote by K_m the class of kernels with support $[-1, 1]$ which fulfill $\int k(x)\,dx = 1$, $\int x^i k(x)\,dx = 0$, $i = 1, \cdots, m$, $\int k^2(x)\,dx < \infty$.

We know from Falk (1983), Theorem 2, where the functional ψ is extensively studied that, if p_m denotes the unique polynomial of degree not greater than m in K_m

$$0 < \psi(p_m) \sim (\pi m)^{-1}.$$

Furthermore, we know that there is no kernel in K_m which maximizes ψ over K_m, i.e. there is no optimal kernel in K_m. However, it was shown by Mammitzsch (1983) that

$$\sup_{k\in K_m}\psi(k) = \binom{2[m/2]}{[m/2]}^2 2^{-4[m/2]-1}, \quad m \in \mathbb{N},$$

where $[x]$ denotes the integral part of $x \in \mathbb{N}$, and thus

$$\psi(p_m)/\sup_{k\in K_m}\psi(k) \to_{m\in\mathbb{N}} 1.$$

This entails that

$$p_m(x) = \sum_{j=0}^{[m/2]}\left(-\frac{1}{4}\right)^j \frac{4j + 1}{2}\binom{2j}{j} l_{2j}(x), \quad m \in \mathbb{N},$$

where l_j denotes the Legendre-polynomial of degree j on $[-1, 1]$, are nearly optimal kernels within a certain class for constructing kernel estimators of a distribution function as well as of quantiles.

2. Auxiliary results and proofs. When dealing with the mean square error we are concerned with the problem of computing moments of q_n and $\hat{q}_n$. To this end we establish the following two auxiliary results which are of interest in their own. For further results on moments of order statistics we refer to Sections 3 and 4 in David (1981), Bickel (1967) and Hall (1978).

LEMMA 1. *Assume that* $\lim_{t\to\infty} t^\delta(1 - F(t) + F(-t)) = 0$ *for some* $\delta > 0$. *Assume further that* F^{-1} *is twice differentiable in a neighborhood of* $q \in (0, 1)$ *with bounded second derivative. If* $r_n/n \to_{n\in\mathbb{N}} q$, *where* $1 \le r_n \le n$, *then the mean and variance of* $Z_{r_n:n}$ *exist if n is large and*

$$E((Z_{r_n:n} - q(F))^2)$$
$$= (n+2)^{-1}\frac{r_n}{n+1}\left(1 - \frac{r_n}{n+1}\right)(F^{-1})'^2\left(\frac{r_n}{n+1}\right) + O\left(\left(\frac{r_n}{n} - q\right)^2\right) + O(n^{-3/2}).$$

LEMMA 2. *Assume that* $\lim_{t\to\infty} t^\delta(1 - F(t) + F(-t)) = 0$ *for some* $\delta > 0$ *and that* F^{-1} *is* $(m+1)$*-times differentiable in a neighborhood of* $q \in (0, 1)$ *with continuous second derivative if* $m = 1$ *and with bounded* $(m+1)$th *derivative if* $m \ge 2$. *Assume further that the kernel k has finite support* $[-c, c]$ *and fulfills* $\int k(x)\,dx = 1$, $\int x^i k(x)\,dx = 0$, $i = 1, \cdots, m$. *Then* $\mathrm{MSE}(\hat{q}_n)$ *is finite if n is large and*

$$n\,\mathrm{MSE}(\hat{q}_n) = (F^{-1})'^2(q)\left\{q(1-q) - 2\alpha_n \int xk(x)K(x)\,dx\right\}$$
$$+ O(\alpha_n^{2m+2}n) + O(n^{-1/4}) + o(\alpha_n).$$

PROOF OF LEMMA 1. Denote by $X_{i:n}$ the ith order statistic in a sample of n independent and uniformly on (0, 1) distributed random variables on some probability space $(\Omega, \mathscr{A}, \tilde{P})$. Then

$$E(Z_{r_n:n}^2) = E((F^{-1}(X_{r_n:n}))^2)$$
$$= E((F^{-1}(X_{r_n:n}))^2 1_{M_n}) + E((F^{-1}(X_{r_n:n}))^2 1_{M_n^c}) =: I_n + II_n,$$

where $M_n := \{|X_{r_n:n} - r_n/(n+1)| \le \varepsilon\}$, ε sufficiently small, and 1_A denotes the indicator function of an event A.

Hölder's inequality together with Lemma 1 in Wellner (1977) implies

$$II_n \le E((F^{-1}(X_{r_n:n}))^4)^{1/2}\tilde{P}(M_n)^{1/2} = E((F^{-1}(X_{r_n:n}))^4)^{1/2}O(\exp(-n)).$$

Next we show that $E((F^{-1}(X_{r_n:n}))^4)$ is finite and uniformly bounded if n is large.

Fubini's Theorem implies

$$E((F^{-1}(X_{r_n:n}))^4) \le \int_0^\infty \tilde{P}\{X_{r_n:n} \ge F(t^{1/4})\}\,dt + \int_0^\infty \tilde{P}\{X_{r_n:n} \le F(-t^{1/4})\}\,dt$$

$$=: A_n + B_n.$$

We show that A_n, $n \in \mathbb{N}$, is uniformly bounded if n is large. Similar arguments yield that this is also true for B_n, $n \in \mathbb{N}$. From formula (2.1.6) in David (1981) we know

$$A_n = \frac{n!}{\{(r_n-1)!(n-r_n)!\}} \int_0^\infty \int_{F(t^{1/4})}^1 x^{r_n-1}(1-x)^{n-r_n}\,dx\,dt.$$

For $\eta > 0$ there exists $C_1 > 0$ such that if $t \ge C_1$ then $1 - F(t^{1/4}) \le \eta t^{-\delta/4}$. Thus,

$$A_n \le C_1 + \frac{n!}{\{(r_n-1)!(n-r_n)!\}} \eta^{n-r_n} \int_{C_1}^\infty t^{-(n-r_n)\delta/4}\,dt.$$

Obviously it suffices to show that for appropriately chosen η the sequence $(n!/\{(r_n-1)!(n-r_n)!\})\eta^{n-r_n}$, $n \in \mathbb{N}$, tends to zero.

Stirling's formula implies

$$n!/\{(r_n-1)!(n-r_n)!\} = O(n^{1/2}(r_n/n)^{-r_n}(1-r_n/n)^{-(n-r_n)}).$$

Choose $\rho > 0$ and $C_2 > 1$ such that $q^{-1} - 1 - \rho > 0$, $C_2^{q^{-1}-1-\rho} \ge 4q^{-1}$. Then, for $\eta := (1-q)/(2C_2)$ and n sufficiently large $\eta \le (1 - r_n/n)/C_2$ and thus

$$\eta^{n-r_n}(r_n/n)^{-r_n}(1-r_n/n)^{-(n-r_n)} \le (r_n/n)^{-r_n} C_2^{-(n-r_n)}$$

$$= \{(r_n/n)C_2^{(n/r_n)-1}\}^{-r_n} \le \{(q/2)C_2^{q^{-1}-\rho-1}\}^{-r_n} \le 2^{-r_n}.$$

This implies that A_n, $n \in \mathbb{N}$, is uniformly bounded if n is large.

Finally we treat I_n. Taylor's formula together with Lemmata 1 and 2 in Wellner (1977) implies if n is large

$$I_n = E(\{F^{-1}(r_n/(n+1)) + (F^{-1})'(r_n/(n+1))(X_{r_n:n} - r_n/(n+1))$$

$$+ 2^{-1}(F^{-1})''(\theta)(X_{r_n:n} - r_n/(n+1))^2\}^2 \cdot 1_{M_n})$$

$$= F^{-1}(r_n/(n+1))^2 + (F^{-1})'^2(r_n/(n+1))E((X_{r_n:n} - r_n/(n+1))^2)$$

$$+ F^{-1}(r_n/(n+1))E(\{(F^{-1})''(\theta)(X_{r_n:n} - r_n/(n+1))^2\} \cdot 1_{M_n}) + O(n^{-3/2}).$$

An analogous expansion of $E(Z_{r_n:n})$ together with example 3.1.1 and formula 3.1.6 in David (1981) and elementary computations complete the proof.

PROOF OF LEMMA 2. The approximate variance of the kernel estimator is given by

$$E\left(\left\{\int_0^1 (F_n^{-1}(x) - F^{-1}(x))\alpha_n^{-1}k\left(\frac{q-x}{\alpha_n}\right) dx\right\}^2\right)$$

$$= E\left(\left\{\int_{(q-1)/\alpha_n}^{q/\alpha_n} (F_n^{-1}(q-\alpha_n x) - F^{-1}(q-\alpha_n x))k(x)\, dx\right\}^2\right)$$

$$= E\left(\left\{\int_{-c}^{c} (F^{-1}(\bar{F}_n^{-1}(q-\alpha_n x)) - F^{-1}(q-\alpha_n x))k(x)\, dx\right\}^2\right)$$

for α_n small enough, where $\bar{F}_n$ denotes the empirical distribution function according to n independent, uniformly on (0, 1) distributed random variables on some probability space $(\Omega, \mathscr{A}, \tilde{P})$. In order to apply Taylor's formula the above integral is split into two terms

$$E\left(\left\{\int (F^{-1}(\bar{F}_n^{-1}(q-\alpha_n x)) - F^{-1}(q-\alpha_n x))k(x)\, dx\right\}^2 1_{M_n}\right)$$

$$+ E\left(\left\{\int (F^{-1}(\bar{F}_n^{-1}(q-\alpha_n x)) - F^{-1}(q-\alpha_n x))k(x)\, dx\right\}^2 1_{M_n^c}\right) =: A_n + B_n,$$

where $M_n := \{\sup_{x\in[-c,c]} |\bar{F}_n^{-1}(q-\alpha_n x) - q + \alpha_n x| \le \varepsilon\}$ for ε being sufficiently small.

B_n is up to an additive constant bounded by

$$E(\max\{(F^{-1}(X_{[(q-\rho)n]:n}))^2, (F^{-1}(X_{[(q+\rho)n]:n}))^2\} \cdot 1_{M_n^c}) = O(\exp(-n)), \tag{1}$$

for ρ being sufficiently small, which is immediate from the proof of Lemma 1 and the inequality by Dvoretzky et al. (1956) since $\sup_{p\in(0,1)} |\bar{F}_n^{-1}(p) - p| = \sup_{t\in[0,1]} |\bar{F}_n(t) - t|$.

Applying Taylor's formula to A_n we derive

$$\begin{aligned} A_n = {} & E\left(\left\{\int k(x)(F_n^{-1}(q-\alpha_n x) - (q-\alpha_n x))(F^{-1})'(q-\alpha_n x)\, dx\right\}^3\right) \\ & + O(E(\sup_{t\in[0,1]} |\bar{F}_n(t) - t|^3)) + O(\exp(-n)). \end{aligned} \tag{2}$$

Furthermore,

$$n^{3/2}E(\sup_{t\in[0,1]} |\bar{F}_n(t) - t|^3) = \int_0^\infty \tilde{P}\{\sup_{t\in[0,1]} |\bar{F}_n(t) - t| \ge n^{-1/2}t^{1/3}\}\, dt$$

$$\le C \int_0^\infty \exp(-2t^{2/3})\, dt < \infty,$$

where C denotes the constant occurring in the inequality by Dvoretzky et al.

(1956). Thus, $E(\sup_{t\in[0,1]}|\bar{F}_n(t) - t|) = O(n^{-3/2})$ and therefore

$$A_n = E\left(\left\{\int k(x)(\bar{F}_n^{-1}(q - \alpha_n x) - (q - \alpha_n x))(F^{-1})'(q - \alpha_n x)\,dx\right\}^2\right) + O(n^{-3/2})$$

$$= E\left(\left\{\int k(x)(q - \alpha_n x - \bar{F}_n(q - \alpha_n x))(F^{-1})'(q - \alpha_n x)\,dx + \int k(x)(\bar{F}_n^{-1}(q - \alpha_n x) - (q - \alpha_n x) + \bar{F}_n(q - \alpha_n x) - (q - \alpha_n x))(F^{-1})'(q - \alpha_n x)\,dx\right\}^2\right) + O(n^{-3/2}). \tag{3}$$

Since the second term above is the remainder term of the first Bahadur quantile-approximation (Bahadur, 1966), Theorem 1 in Duttweiler (1973) implies

$$A_n = E\left(\left\{\int k(x)(q - \alpha_n x - \bar{F}_n(q - \alpha_n x))(F^{-1})'(q - \alpha_n x)\,dx\right\}^2\right) + O(n^{-5/4}).$$

Furthermore,

$$E\left(\left\{\int k(x)(q - \alpha_n x - \bar{F}_n(q - \alpha_n x))(F^{-1})'(q - \alpha_n x)\,dx\right\}^2\right)$$

$$= n^{-1}\int_0^1\left\{\int k(x)(q - \alpha_n x - 1_{(0,q-\alpha_n x]}(y))(F^{-1})'(q - \alpha_n x)\,dx\right\}^2 dy$$

$$= n^{-1}\int_0^1\left\{\int k(x)(q - \alpha_n x - 1_{(0,q-\alpha_n x]}(y))(F^{-1})'(q)\,dx\right\}^2 dy + o(n^{-1}\alpha_n) \tag{4}$$

$$= n^{-1}(F^{-1})'^2(q)\left\{q(1 - q) - 2\alpha_n\int xk(x)K(x)\,dx\right\} + o(n^{-1}\alpha_n)$$

which follows from elementary computations.

Combining (1) – (4) we get

$$E\left(\left\{\int (F_n^{-1}(x) - F^{-1}(x))\alpha_n^{-1}k\left(\frac{q - x}{\alpha_n}\right)dx\right\}^2\right)$$

$$= n^{-1}(F^{-1})'^2(q)\left\{q(1 - q) - 2\alpha_n\int xk(x)K(x)\,dx\right\} + o(n^{-1}\alpha_n) + O(n^{-5/4}).$$

Since for sufficiently small α_n the approximate bias equals $\int k(x)\{F^{-1}(q -$

$\alpha_n x) - F^{-1}(q)\}\, dx = O(\alpha_n^{m+1})$ the assertion of Lemma 2 is immediate from standard arguments.

We remark that it is possible to utilize a differentiation argument instead of the Bahadur approximation in the preceding proof (see Chapter 8 in Serfling, 1980, for details). However, this would require restrictive conditions on k.

PROOF OF THE THEOREM. Let $r_n \in \{1, \cdots, n\}$, $n \in \mathbb{N}$, be such that $|q - r_n/n| = O(n^{-5/8})$. Obviously, $i(n) = \min\{j \in \mathbb{N}: \mathrm{MSE}(Z_{r_j:j}) \le \mathrm{MSE}(\hat{q}_n)\}$ tends to infinity as n increases. Therefore, from Lemmata 1 and 2 and the definition of $i(n)$

$$\mathrm{MSE}(\hat{q}_n) = n^{-1}\{(F^{-1})'^2(q)q(1-q) + o(1)\} \ge \mathrm{MSE}(Z_{r_{i(n)}:i(n)})$$

$$= i(n)^{-1}\{(F^{-1})'^2(q)q(1-q) + O(i(n)^{-1/4})\}$$

which implies $\limsup_{n \in \mathbb{N}} i(n)/n \ge 1$ yielding

$$i(n) \ge (1 + O(n^{-1/4}))(F^{-1})'^2(q)q(1-q)/\mathrm{MSE}(\hat{q}_n).$$

A similar argument yields

$$i(n) \le (1 + O(n^{-1/4}))(F^{-1})'^2(q)q(1-q)/\mathrm{MSE}(\hat{q}_n).$$

The assertion now follows from Lemma 2 and elementary computations.

REFERENCES

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

BICKEL, P. J. (1967). Some contributions to the theory of order statistics. *Proc. 5th Berkeley Symp.* **I** 575–591.

DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

DUTTWEILER, D. L. (1973). The mean-square error of Bahadur's order statistic approximation. *Ann. Statist.* **1** 446–453.

DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

FALK, M. (1983). Relative efficiency and deficiency of kernel type estimators of smooth distribution functions. *Statist. Neerlandica* **37** 73–83.

GALAMBOS, J. (1978). *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York.

HALL, P. (1978). Some asymptotic expansions of moments of order statistics. *Stochastic Process. Appl.* **7** 265–275.

MAMMITZSCH, V. (1983). On the asymptotically optimal solution within a certain class of kernel type estimators. *Statist. Decisions*, to appear.

PARZEN, E. (1979). Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.* **74** 105–121.

REISS, R.-D. (1980a). Estimation of quantiles in certain nonparametric models. *Ann. Statist.* **8** 87–105.

REISS, R.-D. (1980b). One-sided tests for quantiles in certain nonparametric models. In: *Nonparametric Statistical Inference* (Colloq. Math. Soc. János Bolyai, Budapest 1980) 759–772.

REISS, R.-D. (1981). Nonparametric estimation of smooth distribution functions. *Scand. J. Statist.* **8** 116–119.

SCOTT, D. W., TAPIA, R. A. and THOMPSON, J. R. (1977). Kernel density estimation revisited. *Nonlinear Anal.* **1** 339–372.

SERFLING, R. J. (1980) *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

WELLNER, J. A. (1977). A law of the iterated logarithm for functions of order statistics. *Ann. Statist.* **5** 481–494.

WERTZ, W. (1978). *Statistical Density Estimation. A survey. Vandenhoeck and Ruprecht, Göttingen.*

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SIEGEN
HÖLDERLINSTR. 3
5900 SIEGEN 21
WEST GERMANY