Sankhyā : The Indian Journal of Statistics
Special Issue on Quantile Regression and Related Methods
2005, Volume 67, Part 2, pp 277-294

How to Combine M-estimators to Estimate Quantiles and a Score Function

Andrzej S. Kozek
Macquarie University, Sydney, Australia

Abstract

In Kozek (2003) it has been shown that proper linear combinations of some M-estimators provide efficient and robust estimators of quantiles of near normal probability distributions. In the present paper we show that this approach can be extended in a natural way to a general case, not restricted to a vicinity of a specified probability distribution. The new class of nonparametric quantile estimators obtained this way can also be viewed as a special class of linear combinations of kernel-smoothed quantile estimators with a varying window width. The new estimators are consistent and can be made more efficient than the popular quantile estimators based on kernel smoothing with a single bandwidth choice, like those considered in Nadaraya (1964), Azzalini (1981), Falk (1984) and Falk (1985). The present approach also yields simple and efficient nonparametric estimators of a score function $J(p) = -\frac{f'(Q(p))}{f(Q(p))}$, where $f = F'$ and $Q(p)$ is the quantile function, $Q(p) = F^{-1}(p)$.

AMS (2000) *subject classification.* Primary 62G99; secondary 62G35, 62G30.
Keywords and phrases. Asymptotic properties, kernel estimators, M-estimators, quantiles, score function, smoothing.

1 Introduction

Let Y be a random variable with a cumulative distribution function (cdf) F defined on the real line. Quantile function

$$Q(p) = F^{-1}(p) = \inf\{y : F(y) \geq p, p \in (0,1)\}$$

is frequently used in many areas ranging from statistical data analysis to risk measurement in Econometrics, cf. Parzen (1979), Engle and Manganelli (2000) and Koenker and Xiao (2004). If $Y_1, Y_2, \ldots, Y_n$ is a sample from F and $\hat{F}_n$ is the corresponding empirical distribution function then the sample quantile function

$$\hat{q}_n(p) = \hat{F}_n^{-1}(p) \tag{1}$$

is used frequently as an estimator of the population quantiles.

Empirical quantiles are known to be deficient (Falk (1984)) with respect to some nonparametric, smoothed quantile estimators. It is also known, even in a more general context, that for small sample sizes smoothing reduces the variance of the sample functionals, cf. Fernholz (1993) and Fernholz (1997). However, since the window width converges to zero the asymptotic variance of the smoothed estimator is identical with that of the non-smoothed sample functionals.

Falk (1984), Kaigh and Sorto (1993) and Cheng and Parzen (1997) report in detail about classes of distribution functions where the empirical quantiles are superseded in finite samples by some of their nonparametric competitors. The class of kernel smoothed empirical quantiles has been the most extensively studied in the literature, cf. Parzen (1979), Falk (1984), Falk (1985), Cheng and Parzen (1997), Falk (1985) and Falk and Reiss (1989). Other classes include Bernstein polynomial type estimators studied in Muñoz Pérez and Fernández Palacín (1987), Kaigh and Sorto (1993) and Cheng (1995) and the perturbed sample quantile estimators, coinciding with the quantile of a kernel estimator of the cumulative distribution function studied in Nadaraya (1964), Azzalini (1981), Yamato (1972/73), Mack (1987), Ralescu (1992), Ralescu and Sun (1993) and Ralescu (1996).

In the present paper we introduce and explore properties of a new class of approximations to quantiles based on M-functionals and derive asymptotic properties of the corresponding sample estimators. Our approach is similar to the method of perturbed sample quantile estimators introduced in Nadaraya (1964), however, by contrast with the traditional smoothing techniques, we consider here the effects of using a constant window width. We show that in this way one can get excellent approximations to quantile functionals based on M-estimators. Our estimator $\hat{Q}_n(p)$ equals the intercept of a particular polynomial regression of degree 3, fitted to values of several perturbed sample quantile estimators, each with a different window width h_i, $i = 1, 2, \ldots, m$. The regression polynomial of variable h is particular because it's linear term is not present. Moreover, we show that the regression coefficient at h^2 can be used to estimate the score function $J(p) = -\frac{f'(Q(p))}{f(Q(p))}$.

Our estimators are asymptotically unbiased and, for a broad class of cumulative probability distribution functions, they have asymptotic variance lower than the variance of the corresponding empirical quantiles. Hence they can easily compete with empirical quantiles in applications to regression quantiles introduced in Bassett and Koenker (1978). We refer to Green and Kozek (2003) for some applications of M-regression quantiles to weather modelling.

The paper is organized as follows. In Section 2 we present links between M-estimation, kernel estimation and perturbed quantiles. In Section 3 we derive properties of M-functionals corresponding to the perturbed quantiles. In Section 4 we show that the asymptotic variance of the empirical perturbed quantiles is decreasing for h in a vicinity of zero. In Section 5 we present our combined estimators of quantiles and score function and report some simulations providing further insight into their asymptotic properties. In Section 6 we present the asymptotic theory of our estimators.

2 M-Estimators, Kernel Estimators and Perturbed Quantiles

Let Y be a random variable with a cdf F and let K be another cdf, which can be chosen by an analyst. Let Z be a random variable independent of X with a cdf K. Let $h > 0$ and $X = Y - hZ$. Then

$$H_h(x) = P(X \leq x) = \int_{-\infty}^{\infty} \left[1 - K\left(\frac{y-x}{h}\right)\right] F(dy) = 1 - E_F K\left(\frac{Y-x}{h}\right). \tag{2}$$

If H_h is continuous and increasing then for every $p \in (0,1)$ equation

$$H_h(x) - p = 1 - E_F K\left(\frac{Y-x}{h}\right) - p = 0 \tag{3}$$

has a unique solution, say $Q_{p,h}(F)$, a p-quantile of H_h. Equation (3) may be considered as an estimating equation of the p-quantile of H_h. The resulting estimator $\tilde{q}_p(h) = Q_{p,h}\left(\hat{F}_n\right)$ solving equation

$$1 - \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{Y_i - x}{h}\right) - p = 0 \tag{4}$$

is identical, in the case of $K_1(z) = 1 - K(-z)$, with the kernel quantile estimator introduced in Nadaraya (1964), except that we are not changing h with the sample size n.

The cdf H_h given by (2) can be considered as a cdf of a perturbed probability distribution corresponding to F. Consequently, the functional $Q_{p,h}(F)$ is referred to as *a perturbed p-quantile of F*, cf. Ralescu (1992) and Ralescu and Sun (1993). Let us note that

$$\begin{aligned}
\gamma(\theta) &= 2\int_0^\theta (H_h(x) - p)\, dx \\
&= 2E_F \int_0^\theta \left(1 - K\left(\frac{Y-x}{h}\right) - p\right) dx \\
&= -E_F \int_0^\theta M_p'\left(\frac{Y-x}{h}\right) dx \\
&= E_F\left[M_p\left(\frac{Y-\theta}{h}\right) - M_p\left(\frac{Y}{h}\right)\right], \qquad (5)
\end{aligned}$$

where

$$M_p(y) = \int_0^y (2K(u) - 1)\, du + (2p-1)\, y. \qquad (6)$$

Since H_h is non-decreasing the functions $\gamma(\theta)$ and $M_p(y)$ are convex. If H_h is strictly increasing on its support interval then $\gamma(\theta)$ is strictly convex, $Q_{p,h}(F)$ is the unique minimizer of $\gamma(\theta)$ and hence it can be consistently estimated by an M-estimator $Q_{p,h}\left(\hat{F}_n\right)$, where $\hat{F}_n$ is an empirical cdf based on a sample from F. If K is also continuous and increasing on its support interval, then $Q_{p,h}\left(\hat{F}_n\right)$ solves as well equation (4).

Let us note that by kernel smoothing interpretation of the estimating equation (4) the parameter h can be called a window width or a smoothing parameter. If, however, equation (4) is considered as defining an M-estimator (or Z-estimator, cf, van der Vaart and Wellner, 1996), then h can be referred to as a scale parameter.

The link of the estimating equation (2) with the M-functional (5) plays an important role in the present paper. By keeping h fixed we can readily use the asymptotic theory of M-estimators to explore properties of both quantile functional $Q_{p,h}(F)$ and of its sample estimator $Q_{p,h}\left(\hat{F}_n\right)$.

3 Perturbed Quantile Functionals for Small h

We shall assume that Z has a cdf K with a compact support and two moments κ_k, $k = 1, 2$ such that $\kappa_1 = 0$ and $\kappa_2 > 0$. By (5), the functional $Q_{p,h}(F)$, called a perturbed p-quantile of F, coincides with a p -quantile of $X = Y - hZ$. $Q_{p,h}(F)$ also minimizes

$$\mathbf{M}_{p,h}(\theta) = E_F M_p\left(\frac{Y-\theta}{h}\right), \qquad (7)$$

where $M_p(y)$ is given by (6). Assuming that Y and Z have probability density functions $f(y)$ and $k(z)$, respectively, the probability density function of X is given by

$$g_h(x) = \int_{-\infty}^{\infty} \frac{1}{h} f(x+z) k\left(\frac{z}{h}\right) dz.$$

The following lemma provides further justification of the name *perturbed p-quantile of F* for $Q_{p,h}(F)$. It shows that $Q_{p,h}(F)$ differs from $Q(p)$, the quantile of F, by at most by a term of order h.

LEMMA 3.1 *If Z has a compact support with a density function positive in a neighbourhood of zero and the density function $f(y)$ is positive in a neighbourhood of $Q(p)$ then, for small h, we have*

$$|Q_{p,h}(F) - Q(p)| \leq c_K h,$$

where the constant c_K depends on the compact support of K.

PROOF OF LEMMA 3.1. Let c_K be such that $|Z| \leq c_K$ with probability 1. Then, with probability 1 we have

$$\begin{aligned} \{Y - hZ \leq Q(p) - hc_K\} &= \{Y \leq Q(p) + h(-c_K + Z)\} \subset \{Y \leq Q(p)\} \\ &\subset \{Y \leq Q(p) + h(c_K + Z)\} = \{Y - hZ \leq Q(p) + hc_K\}. \end{aligned}$$

Since $P(Y \leq Q(p)) = p$ and our assumptions guarantee the uniqueness of quantiles $Q(p)$ and $Q_{p,h}(F)$ we conclude that

$$Q(p) - hc_K \leq Q_{p,h}(F) \leq Q(p) + hc_K.$$

□

We need Lemma 3.1 to prove Theorem 3.1 below. Theorem 3.1 shows that, under some additional week assumptions on F and K, $Q_{p,h}(F)$ differs from $Q(p)$ by a term $\frac{\kappa_2}{2} J(p) h^2 + o(h^2)$ where $J(p)$ is a score function

$$J(p) = -\frac{f'(Q(p))}{f(Q(p))}, \tag{8}$$

and where $f = F'$. The score function $J(p)$ plays an important role in nonparametric statistics, cf. Hájek and Šidák (1967), Parzen (1979) and Behnen and Neuhaus (1989). In Section 5, apart from new estimators of quantiles, we also will discuss new estimators of $J(p)$ suggested by Theorem 3.1.

Theorem 3.1 *Assume that F has three continuous derivatives vanishing at $-\infty$ and the M-function is given by (6), where K is a cdf with a compact support, the first moment $\kappa_1 = 0$ and the second moment $\kappa_2 > 0$. If f is positive at $Q(p)$ then the perturbed quantile $Q_{p,h}(F)$ has the following Taylor expansion*

$$Q_{p,h}(F) = Q(p) + \frac{\kappa_2}{2} J(p) h^2 + o\left(h^2\right), \tag{9}$$

where $Q(p)$ is the p-quantile of F.

Proof of Theorem 3.1. Let q be a p-quantile of U, q_o a p-quantile of Y and let $c = q - q_o$. By the definition of quantile, Lemma 3.1 and by Taylor expansion we have the following.

$$\begin{aligned}
p = \int_{-\infty}^{q} g_h(u)\, du &= \int_{-\infty}^{q}\int_{-\infty}^{\infty} \frac{1}{h} f(u+z) k\left(\frac{z}{h}\right) dz du \\
&= \int_{-\infty}^{q} \left(f(u) + \frac{\kappa_2}{2} f''(u) h^2\right) du + o\left(h^2\right) \\
&= F(q_0 + c) + \frac{\kappa_2}{2} f'(q) h^2 + o\left(h^2\right) \\
&= p + f(q_0) c + \frac{1}{2} f'(q_0) c^2 + \frac{\kappa_2}{2} f'(q) h^2 + o\left(h^2\right) \\
&= p + f(q_0) c + \frac{1}{2} f'(q_0) c^2 \\
&\qquad + \frac{\kappa_2}{2}\left(f'(q_0) + c f''(q_0)\right) h^2 + o\left(h^2\right).
\end{aligned}$$

Hence, we get

$$\left(\frac{\kappa_2}{2} f'(q_0) h^2 + o\left(h^2\right)\right) + \left(f(q_0) + \frac{\kappa_2}{2} f''(q_0) h^2\right) c + \frac{1}{2} f'(q_0) c^2 = 0. \tag{10}$$

Now, applying Lemma 3.1 we choose the proper root $c(h)$ of (10) and with some algebra we get

$$\begin{aligned}
c(h) = \frac{1}{2f'(q_0)}\Bigg(&-\kappa_2 f''(q_0) h^2 - 2f(q_0) + \\
&\sqrt{4f(q_0)^2 + \kappa_2^2 f''(q_0)^2 h^4 + 4\kappa_2 f''(q_0) h^2 f(q_0) - 4\kappa_2 f'(q_0)^2 h^2 + o\left(h^2\right)}\Bigg).
\end{aligned} \tag{11}$$

Next, by Taylor expansion

$$\sqrt{1 + x^2} = 1 + \frac{1}{2} x + o(x)$$

applied to the square root at (11) we get, again with some algebra,

$$c(h) = -\frac{\kappa_2}{2}\frac{f'(q_0)}{f(q_0)}h^2 + o\left(h^2\right).$$

Consequently, we obtain expansion (9)

$$q = q_o + c = q_o - \frac{\kappa_2}{2}\frac{f'(q_0)}{f(q_0)}h^2 + o\left(h^2\right). \tag{12}$$

□

4 Variances of Perturbed Sample Quantiles for Small h.

To estimate the perturbed quantiles $Q_{p,h}(F)$ one can use either a solution of the sample version (4) of estimating equations (3) or, equivalently, by minimizing a sample version

$$\widehat{\mathbf{M}}_{p,h}(\theta) = \frac{1}{n}\sum_{i=1}^{n} M_p\left(\frac{Y_i - \theta}{h}\right) \tag{13}$$

of the convex functional $\mathbf{M}_{p,h}$ given by (7). The estimator, to be denoted by $\hat{Q}_{p,h} = Q_{p,h}\left(\hat{F}_n\right)$, is asymptotically normal $AN\left(Q_{p,h}(F), \frac{\sigma^2(h)}{n}\right)$ (cf. Huber, 1981, p.50, Corollary 2.5) with the asymptotic variance $\sigma^2(h)$ given by

$$\sigma^2(h) = AVar\left(\hat{Q}_{p,h}\right) = \frac{\int \psi^2(y, Q_{p,h}(F))dF(y)}{\left(\int \psi_t(y, Q_{p,h}(F))dF(y)\right)^2}, \tag{14}$$

where

$$\psi(y,t) = 2K\left(\frac{y-t}{h}\right) - 2(1-p)$$

and

$$\psi_t(y,t) = \frac{\partial}{\partial t}\psi(y,t) = -\frac{2}{h}k\left(\frac{y-t}{h}\right).$$

Let $f = F'$. Notice that for $q = Q_{p,h}(F)$ and $q_0 = Q(p)$ we have by Theorem 3.1 that $q - q_0 = \frac{\kappa_2}{2}J(p)h^2 + o(h^2)$. With this notation, (14) can be written as follows

$$\sigma^2(h) = \frac{\int \left[2K\left(\frac{y-q}{h}\right) - 2(1-p)\right]^2 dF(y)}{\left(\frac{2}{h}\int k\left(\frac{y-q}{h}\right)F(dy)\right)^2} = \frac{J_1(h)}{\left(J_2(h)\right)^2}. \tag{15}$$

To find the behaviour of the asymptotic variance $\sigma^2(h)$ for small h we need to find Taylor expansion in h for both $J_1(h)$ and $J_2(h)$. We assume here that the probability distribution corresponding to the cdf K is symmetric, non-degenerate and concentrated on interval $(-1,1)$. We have

$$\begin{aligned}
J_1(h) &= \int_{-\infty}^{\infty}\left[2K\left(\frac{y-q}{h}\right)-2(1-p)\right]^2 dF(y)\\
&= 4(1-p)^2F(q-h)+4p^2\left(1-F\left(q+h\right)\right)\\
&\quad +4\int_{q-h}^{q+h}\left[K\left(\frac{y-q}{h}\right)-(1-p)\right]^2 dF(y)\\
&= 4p\left(1-p\right)+4p^2\left(F\left(q_0\right)-F\left(q+h\right)\right)+4(1-p)^2\big(F(q-h)\\
&\quad -F\left(q_0\right)\big)+4\int_{q-h}^{q+h}\left[K\left(\frac{y-q}{h}\right)-(1-p)\right]^2 dF(y)\\
&= 4p\left(1-p\right)+4p^2f\left(q_0\right)\left(q_0-q-h\right)+4(1-p)^2f\left(q_0\right)\left(q-h-q_0\right)+o\left(h\right)\\
&\quad +4\int_{q-h}^{q+h}\left[K\left(\frac{y-q}{h}\right)-(1-p)\right]^2 dF(y)\\
&= 4p\left(1-p\right)-4f\left(q_0\right)\left(p^2+(1-p)^2\right)h\\
&\quad +4f\left(q_0\right)\int_{q-h}^{q+h}\left[K\left(\frac{y-q}{h}\right)-(1-p)\right]^2 dy+o\left(h\right)\\
&= 4p\left(1-p\right)-8Af\left(q_0\right)h+o\left(h\right)
\end{aligned}\tag{16}$$

where

$$A=\int_0^1 K\left(u\right)\left(1-K\left(u\right)\right)du>0 \tag{17}$$

for $K\left(u\right)\neq 1_{[0,\infty)}\left(u\right)$. The last equality in (16) follows from the following derivation, valid for symmetric cdf $K\left(y\right)=1-K\left(-y\right)$ concentrated on $[-1,1]$.

$$\begin{aligned}
\int_{q-h}^{q+h}\left[K\left(\frac{y-q}{h}\right)-(1-p)\right]^2 dy &= h\int_{-1}^{1}\left[K(u)-(1-p)\right]^2 du\\
&= h\int_0^1\left((p-K(u))^2+(K(u)-(1-p))^2\right)du\\
&= h\left(p^2+(1-p)^2+2\int_0^1 K\left(u\right)\left(K\left(u\right)-1\right)du\right).
\end{aligned}$$

Next, we shall find the Taylor expansion for $J_2(h)$.

$$\begin{aligned} J_2(h) &= \int_{-\infty}^{\infty} \left(\frac{2}{h} k\left(\frac{y-q}{h}\right)\right) F(dy) \\ &= \frac{2}{h}\int_{-\infty}^{\infty} k\left(\frac{y-q}{h}\right)\Big(f\left(q\right)+f'\left(q\right)\left(y-q\right) \\ &\qquad\qquad + \frac{1}{2}f''(q)\left(y-q\right)^2 + o\left(h^2\right)\Big)dy \\ &= 2f\left(q\right)+\kappa_2 h^2 f''(q) + o\left(h^2\right) \\ &= 2f\left(q_0\right)+\kappa_2 h^2\left(f''(q_0)+f'\left(q_0\right)J\left(p\right)\right)+o\left(h^2\right) \end{aligned} \tag{18}$$

By (15)–(18) we get the right derivative of the ratio

$$\frac{d}{dh}\sigma^2(h)\left|_{h=0+}\right. = \left(J_1' J_2^2 - 2J_2' J_2 J_1\right)/J_2^4\Big|_{h=0+} = -2\frac{A}{f\left(q_0\right)}.$$

Hence, the variance $\sigma^2(h)$ is decreasing for sufficiently small h. We summarize this result in the following theorem.

THEOREM 4.1 *Assume that the probability distribution corresponding to the cdf K is symmetric, non-degenerated, concentrated on interval $(-1,1)$ and with a derivative $k=K'$ on $(-1,1)$ such that $k(0)>0$. If F is four times continuously differentiable with $F'=f$ then the asymptotic variance $\sigma^2(h)$ of $\hat{Q}_{p,h}$ is decreasing in a vicinity of $h=0$ and*

$$\sigma^2(h) = \frac{p\left(1-p\right)}{f^2(Q(p))} - 2\frac{A}{f\left(Q(p)\right)}h + o(h), \tag{19}$$

where $A>0$ is given by (17).

Let us note that (19) implies that the limit of asymptotic variances $\sigma^2(h)$ equals the asymptotic variance of the sample quantile (cf. Mosteller, 1946)

$$\lim_{h\to 0}\sigma^2(h) = \frac{p\left(1-p\right)}{f^2(Q(p))}.$$

We refer to Green (2002) for simulations in the case of the Uniform $U(0,1)$, Exponential $E(\beta=2)$ and the standard normal $N\left(0,1\right)$ distributions. The results of simulations show very good agreement with the theoretical results. In Figures 1 – 3 we show in the left-hand side graphs the dependence on h of $Q_{ph}\left(F\right)$, the perturbed p-quantiles of F and in the

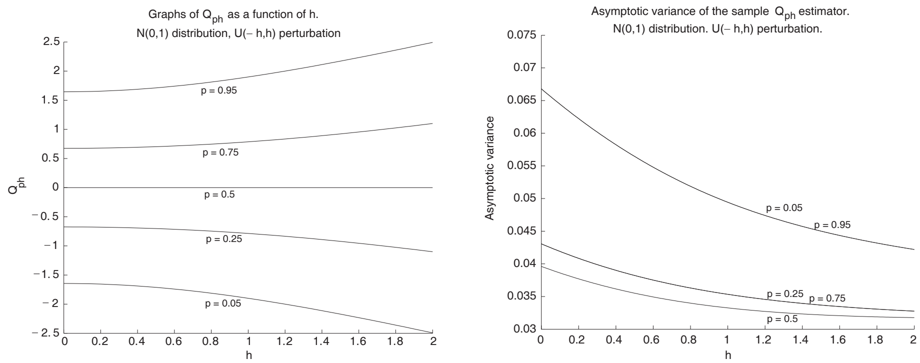

Figure 1: Perturbed quantile functionals $Q_{p,h}(F)$ (left) and asymptotic variances of sample perturbed quantiles $\hat{Q}_{p,h}$ (right) in the case of normal $N(0,1)$ probability distribution and uniform perturbation $U(-h,h)$.

right-hand side graphs, the corresponding asymptotic variances of the sample functionals $Q_{ph}\left(\hat{F}_n\right)$.

Figures 1 – 3 show how the neighbourhood of $h = 0$, over which the asymptotic variance is decreasing, depends both on the distribution function F and on the value of $p \in (0,1)$. Let us note the discontinuity of the first derivative of the variance of $Q_{ph}\left(\hat{F}_n\right)$ in Figures 2 and 3. These points correspond to distances between the corresponding quantiles and the boundary of the support of the probability distribution of Y. Though smoothness of the functional $Q_{ph}(F)$ is not affected at these points yet, the estimator of the score function $J(p)$ breaks down at $p = 0.05$ as we can see in Figure 8. This provides a practical tip on that the window widths h exceeding distance from quantiles to the boundary of the support should not be included into the window design. These examples also remind that conclusions of Theorems 3.1 and 4.1 are valid only in the regions where the Taylor expansion provides an adequate approximation.

5 Combined Perturbed Quantiles

Theorems 3.1 and 4.1 suggest that, similarly as in Kozek (2003), by perturbing the sample distribution we may produce robust estimators with good statistical properties. The perturbed quantile functionals $Q_{p,h}(F)$ differ from the quantile $Q(p)$ of F only by $\frac{\kappa_2}{2} J(p)h^2 + o\left(h^2\right)$ while the corresponding M-estimators $Q_{p,h}\left(\hat{F}_n\right)$ have smaller asymptotic variance than

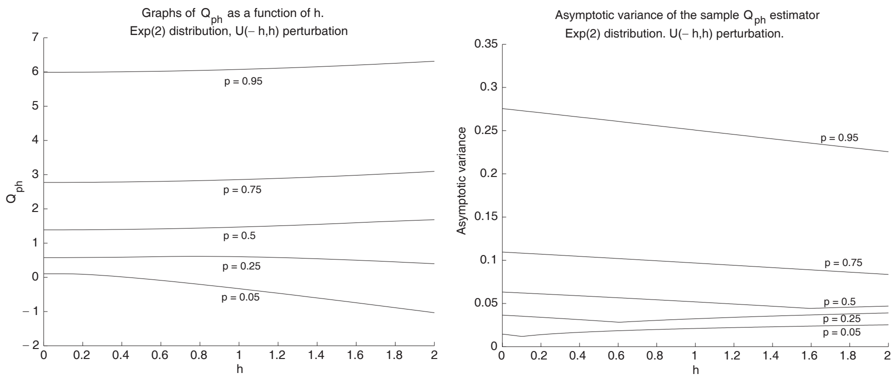

Figure 2: Perturbed quantile functionals $Q_{p,h}(F)$ (left) and asymptotic variances of sample perturbed quantiles $\hat{Q}_{p,h}$ (right) in the case of exponential $Exp(2)$ probability distribution and uniform perturbation $U(-h, h)$.

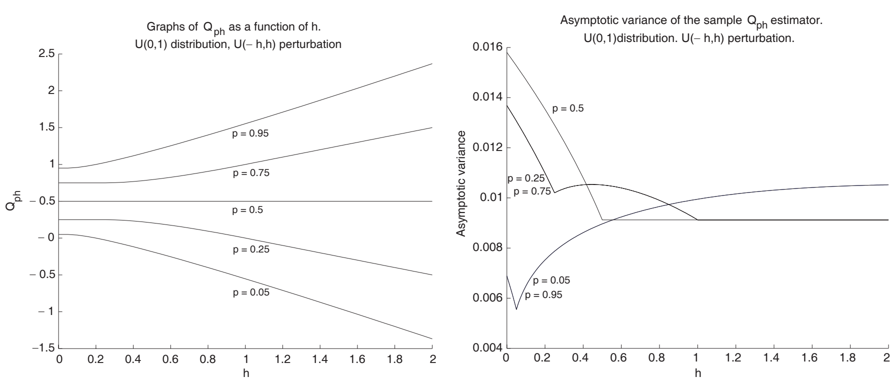

Figure 3: Perturbed quantile functionals $Q_{p,h}(F)$ (left) and asymptotic variances of sample perturbed quantiles $\hat{Q}_{p,h}$ (right) in the case of uniform $U(0,1)$ probability distribution and uniform perturbation $U(-h, h)$.

that of the sample quantile. In an attempt to combine these features we consider the following strategy to estimate simultaneously quantiles and the score function.

1. Let $Y_1, \ldots, Y_n$ be a sample from F. Choose $p \in (0,1)$ and select a window design, ie. a set of values
$$0 \le h_1 < h_2 < \cdots < h_m, \text{ with } m > 3. \tag{20}$$
2. Calculate estimators $Q_{p,h_i}\left(\hat{F}_n\right)$ for $i = 1, 2, \ldots, m$.
3. Find the least squares method approximation to values $\left(h_i, Q_{p,h_i}(\hat{F}_n)\right)$, $i = 1, 2, \ldots, m$, by a polynomial
$$q(h) = \beta_0 + \beta_1 h^2 + \beta_2 h^3. \tag{21}$$
4. Denote by $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_2$ the coefficients of the fitted polynomial.
5. Set $\hat{Q}_n(p) = \hat{\beta}_0$ as an estimator of the quantile $Q(p)$.
6. Set $\hat{J}_n(p) = 2\hat{\beta}_1/\kappa^2$ as an estimator of the score function $J(p) = -\frac{f'(Q(p))}{f(Q(p))}$.

We have the following heuristic motivation for estimators $\hat{Q}_n(p)$ and $\hat{J}_n(p)$. Theorem 3.1 implies that in the Taylor expansion of $Q_{p,h}(F)$ the linear term in h vanishes, so, a polynomial $q(h)$ approximating $Q_{p,h}$ should have zero as a linear term. We suggest to take for $q(h)$ a polynomial of degree 3, with the quadratic term estimating the score function and, as the Taylor expansion is valid only locally, the cube term should, hopefully, reduce the bias of $\hat{J}_n$. The window design $\{h_1, h_2, \cdots, h_m\}$ allows for information about the quantile to be inferred from M-estimators $Q_{p,h_i}\left(\hat{F}_n\right)$ having lower variance than the sample quantile. The intercept $\hat{\beta}_0$ estimates the quantile, up to $o(h_m)$, in an unbiased way and has lower variance than the empirical quantile. The window design plays an important role in the variance reduction, however the choice of an optimal window design is beyond the scope of the present paper.

In Figures 4–6 we show a comparison of asymptotic standard deviations of the sample quantiles and of our estimators $\hat{Q}_n(p)$.

Figures 7 and 8 show the score function $J(p)$ and the averaged values of estimators $\hat{J}_n(p)$.

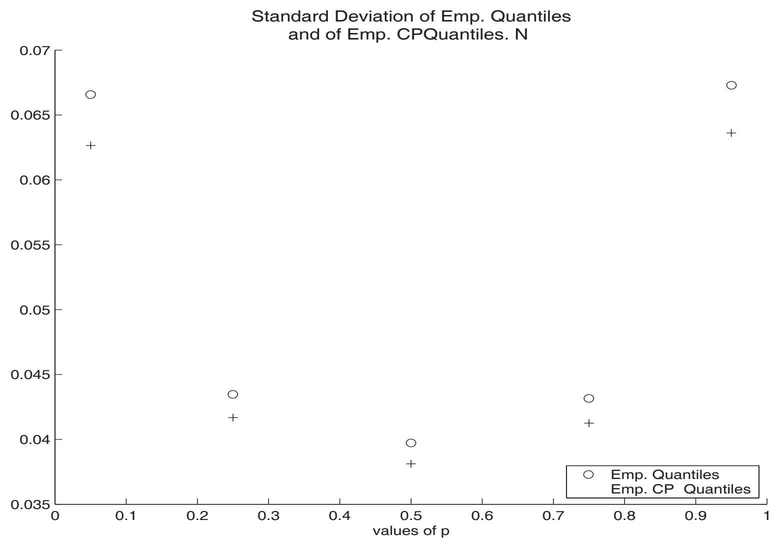

Figure 4: Asymptotic variances of empirical quantiles and of sample combined perturbed quantiles for normal distribution $N(0,1)$, uniform $U(-h,h)$ perturbation and the window design $(0, 0.1, 0.5, 1, 2)$.

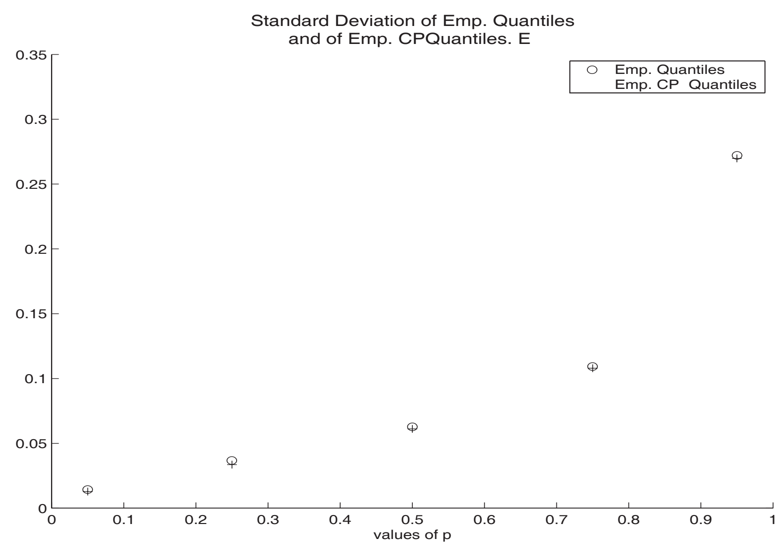

Figure 5: Asymptotic variances of empirical quantiles and of sample combined perturbed quantiles for exponential distribution $Exp(2)$, uniform $U(-h,h)$ perturbation and the window design $(0, 0.1, 0.5, 1, 2)$.

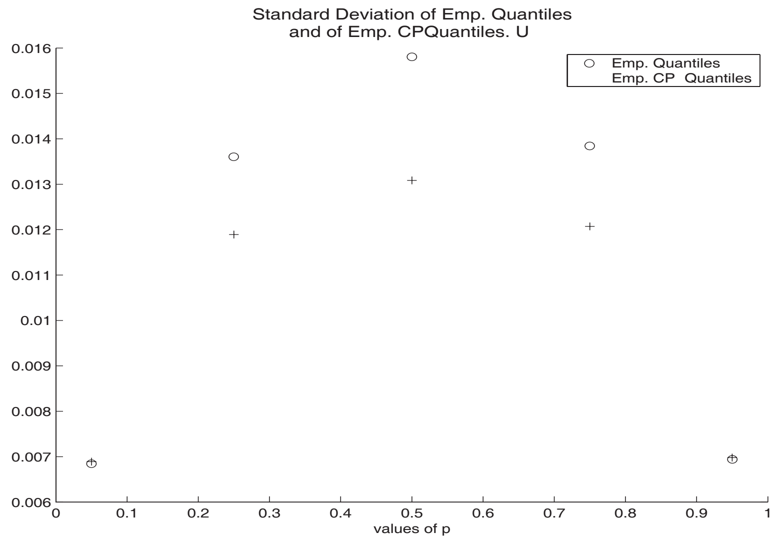

Figure 6: Asymptotic variances of empirical quantiles and of sample combined perturbed quantiles for uniform distribution $U(0,1)$, uniform $U(-h,h)$ perturbation and the window design $(0,0.1,0.5,1,2)$.

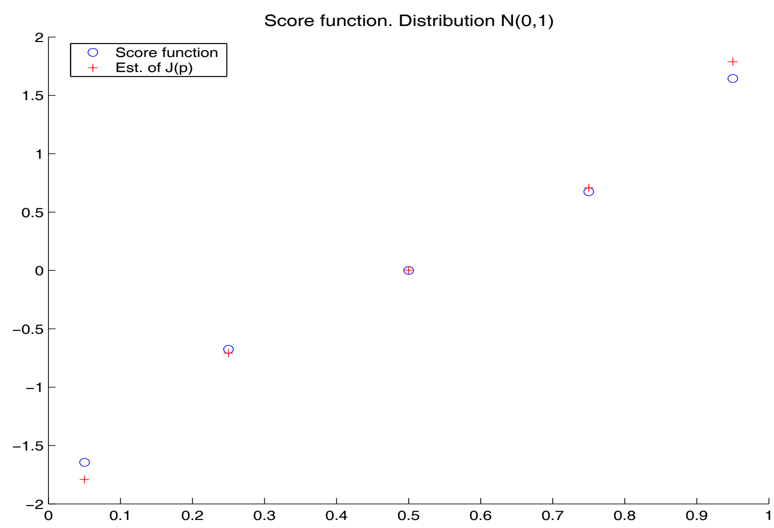

Figure 7: Score function $J(p)$ and the mean of estimator $\hat{J}_n(p)$ values. Normal distribution $N(0,1)$, uniform $U(-h,h)$ perturbation and the window design $(0,0.1,0.5,1,2)$.

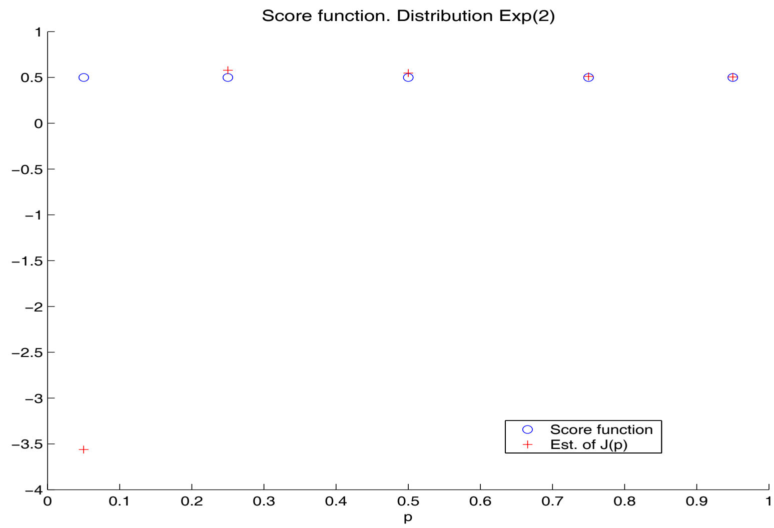

Figure 8: Score function $J(p)$ and the mean of estimator $\hat{J}_n(p)$ values. Exponential distribution $Exp(2)$, uniform $U(-h,h)$ perturbation and the window design $(0, 0.1, 0.5, 1, 2)$. Notice inconsistency of $\hat{J}_n$ at $p = 0.05$, where the window design ranges far beyond the support of the exponential distribution.

6 Asymptotic Theory

In the present section we derive the joint asymptotic distribution of estimators $\hat{Q}_n(p)$ and $\hat{J}_n(p)$. Let

$$\widehat{\mathbf{Q}}_{n,p,h_1,\ldots,h_m}\left(\hat{F}_n\right) = \left[Q_{p,h_1}\left(\hat{F}_n\right), \ldots, Q_{p,h_m}\left(\hat{F}_n\right)\right]^T$$

denote a vector of estimators $Q_{p,h_j}\left(\hat{F}_n\right), j = 1, 2, \ldots, m$, based on sample $Y_1, Y_2, \ldots, Y_n$ and minimizing (13) with $h = h_1, h_2, \ldots, h_m$, respectively. By Corollary 3.2, p. 133 of Huber (1981) we get asymptotic distribution of this vector of M-estimators

$$\widehat{\mathbf{Q}}_{n,p,h_1,\ldots,h_m}\left(\hat{F}_n\right) \sim AN\left(\mathbf{Q}_{p,h_1,\ldots,h_n}(F), \frac{1}{n}\mathbf{\Sigma}\right), \tag{22}$$

where

$$\mathbf{Q}_{p,h_1,\ldots,h_n}(F) = \left[Q_{p,h_1}(F), \ldots, Q_{p,h_m}(F)\right]^T \tag{23}$$

and

$$\mathbf{\Sigma} = \left(\frac{\int \psi(y,Q_{p,h_i}(F))\psi(y,Q_{p,h_j}(F))dF(y)}{\int \psi_t(y,Q_{p,h_i}(F))\psi_t(y,Q_{p,h_j}(F))dF(y)} \right)_{\substack{i=1,\dots,m \\ j=1,\dots,m}}. \qquad (24)$$

The regression design matrix corresponding to the windows design $\{h_1, \dots, h_m\}$ is given by

$$\mathbf{D} = \begin{bmatrix} 1 & h_1^2 & h_1^3 \\ 1 & h_2^2 & h_2^3 \\ \vdots & \vdots & \vdots \\ 1 & h_m^2 & h_m^3 \end{bmatrix}$$

and hence the coefficients of the the least squares polynomial approximation to values of $\widehat{\mathbf{Q}}_{n,p,h_1,\dots,h_m}\left(\hat{F}_n\right)$ are given by

$$\left[\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right]^T = \mathbf{B}\widehat{\mathbf{Q}}_{n,p,h_1,\dots,h_m}\left(\hat{F}_n\right) = \left(\mathbf{D}^T\mathbf{D}\right)^{-1}\mathbf{D}^T\widehat{\mathbf{Q}}_{n,p,h_1,\dots,h_m}\left(\hat{F}_n\right). \qquad (25)$$

Let $\mathbf{C} = [\mathbf{B}_{ij}]_{\substack{i=1,2 \\ j=1,\dots,m}}$ be the upper, of size $(2,m)$, sub-matrix of $\mathbf{B} = \left(\mathbf{D}^T\mathbf{D}\right)^{-1}\mathbf{D}^T$. With some algebra we get

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \frac{\left(S_4S_6 - S_5^2\right)T_0 + (S_3S_5 - S_2S_6)T_2 + (S_2S_5 - S_3S_4)T_3}{\left(S_4S_6 - S_5^2\right)S_0 + (S_3S_5 - S_2S_6)S_2 + (S_2S_5 - S_3S_4)S_3} \\ \frac{(S_5S_3 - S_2S_6)T_0 + \left(S_0S_6 - S_3^2\right)T_2 + (S_3S_2 - S_0S_5)T_3}{(S_5S_3 - S_2S_6)S_2 + \left(S_0S_6 - S_3^2\right)S_4 + (S_3S_2 - S_0S_5)S_5} \end{bmatrix}, \qquad (26)$$

where

$$S_i = \sum_{j=1}^{m} h_j^i \qquad \text{and} \qquad T_i = \sum_{j=1}^{m} h_j^i Q_{ph_j}\left(\hat{F}_n\right). \qquad (27)$$

The joint asymptotic distribution of $\hat{Q}_n(p)$ and $\hat{J}_n(p)$ is given by

$$\begin{bmatrix} \hat{Q}_n(p) \\ \hat{J}_n(p) \end{bmatrix} \sim AN\left(\mathbf{C}\mathbf{Q}_{p,h_1,\dots,h_n}(F), \frac{1}{n}\mathbf{C}\mathbf{\Sigma}\mathbf{C}^T\right). \qquad (28)$$

References

AZZALINI, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, **68**, 326-328.

BASSETT, G. and KOENKER, R. (1978). Asymptotic theory of least absolute error regression. *J. Amer. Statist. Assoc.*, **73**, 618-622.

BEHNEN, K. and NEUHAUS, G. (1989). *Rank Tests with Estimated Scores and their Application.* Teubner Skripten zur Mathematischen Stochastik (Teubner Texts on Mathematical Stochastics), B.G. Teubner, Stuttgart.

CHENG, C. (1995). The Bernstein polynomial estimator of a smooth quantile function. *Statist. Probab. Lett.*, **24**, 321-330.

CHENG, C. and PARZEN, E. (1997). Unified estimators of smooth quantile and quantile density functions. *J. Statist. Plann. Inference*, **59**, 291-307.

ENGLE, R. and MANGANELLI, S. (2000), Caviar: Conditional autoregressive value at risk by regression quantiles. In *Econometric Society World Congress 2000, 0841.*

FALK, M. (1984). Relative deficiency of kernel type estimators of quantiles. *Ann. Statist.*, **12**, 261-268.

FALK, M. (1985). Asymptotic normality of the kernel quantile estimator. *Ann. Statist.*, **13**, 428-433.

FALK, M. and REISS, R.-D. (1989). Weak convergence of smoothed and nonsmoothed bootstrap quantile estimates. *Ann. Probab.*, **17**, 362-371.

FERNHOLZ, L.T. (1993). Smoothed versions of statistical functionals. In *New Directions in Statistical Data Analysis and Robustness (Ascona, 1992)*, Monte Veritá, Birkhäuser, Basel, 61-72.

FERNHOLZ, L.T. (1997). Reducing the variance by smoothing. *J. Statist. Plann. Inference*, **57** (Special issue on Robust Statistics and Data Analysis I), 29-38.

GREEN, H. (2002). *Modelling Weather Data by Approximate Regression Quantiles.* Master Project, Department of Statistics, Macquarie University.

GREEN, H. M. and KOZEK, A.S. (2003). Modelling Weather Data by Approximate Regression Quantiles. *Australian and New Zealand Industrial and Applied Mathematics Journal (E)*, **44**, C229-C248.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

HUBER, P.J. (1981). *Robust Statistics.* Wiley, New York.

KAIGH, W.D. and SORTO, M.A. (1993). Subsampling quantile estimator majorization inequalities. *Statist. Probab. Lett.*, **18**, 373-379.

KOENKER, R. and XIAO, Z. (2004). Unit root quantile autoregression inference. *J. Amer. Statist. Assoc.*, **99**, 775-787.

KOZEK, A.S. (2003). On M-estimators and normal quantiles. *Ann. Statist.*, **31**, 1170-1185.

MACK, Y.P. (1987). Bahadur's representation of sample quantiles based on smoothed estimates of a distribution function. *Probab. Math. Statist.*, **8**, 183-189.

MOSTELLER, F. (1946). On some useful "inefficient" statistics. *Ann. Math. Statist.*, **17**, 377-408.

MUÑOZ PÉREZ, J. and FERNÁNDEZ PALACÍN, A. (1987). Estimating the quantile function by Bernstein polynomials. *Comput. Statist. Data Anal.*, **5**, 391-397.

NADARAYA, É. (1964). Some new estimates for distribution function. *Theory Probab. Appl.*, **9**, 497-500.

PARZEN, E. (1979). Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.*, **74**, 105-131 (with comments by J.W. Tukey, R.E. Welsch, W.F. Eddy, D.V. Lindley, M.E. Tarter and E.L. Crow, and a rejoinder by the author).

RALESCU, S.S. (1992). Asymptotic deviations between perturbed empirical and quantile processes. *J. Statist. Plann. Inference*, **32**, 243-258.

Ralescu, S.S. (1996). A Bahadur-Kiefer law for the Nadaraya empiric-quantile processes. *Teor. Veroyatnost. i Primenen.*, **41**, 380-392.

Ralescu, S.S. and Sun, S. (1993). Necessary and sufficient conditions for the asymptotic normality of perturbed sample quantiles. *J. Statist. Plann. Inference*, **35**, 55-64.

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes, with Applications to Statistics*. Springer-Verlag, New York.

Yamato, H. (1972-73). Uniform convergence of an estimator of a distribution function. *Bull. Math. Statist.*, **15**, 69-78.

Andrzej S. Kozek
Department of Statistics
C5C, Macquarie University
Sydney, NSW 2109, Australia
E-mail: akozek@efs.mq.edu.au

Paper received: July 2004; revised March 2005.