How to combine M-estimators to estimate
quantiles and a score function.

Andrzej S. Kozek

Department of Statistics, C5C, Macquarie University
Sydney, NSW 2109, Australia

October 29, 2004

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 estima-
tors 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, $Q(p) = F^{-1}(p)$. 
Running Title: M-estimators of quantiles


Key words 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 & Manganelli (2000) and Franke & Mwita (2003). 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

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

is used frequently as an estimator of the population quantiles.

Empirical quantiles are known to be deficient 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 & Sorto (1993) and Cheng & 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 quan-

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. Referring to the similarity with the kernel estimation, 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 & Koenker (1978). We refer to
Green & 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

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

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

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

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 $\tilde{q}_p(h) = Q_{p,h}(\hat{F}_n)$
solving equation

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

(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 & Sun (1993). Let us note that

\[
\gamma (\theta) = 2 \int_{0}^{\theta} (H_h (x) - p) \, dx \\
= 2E_F \int_{0}^{\theta} \left( 1 - K \left( \frac{Y - x}{h} \right) - p \right) \, dx \\
= -E_F \int_{0}^{\theta} M' \left( \frac{Y - x}{h} \right) \, dx \\
= E_F \left[ M_p \left( \frac{Y}{h} \right) - M_p \left( \frac{Y - \theta}{h} \right) \right],
\]

(5)

where

\[
M_p (y) = \int_{0}^{y} (2K (u) - 1) \, du + (2p - 1) \, y.
\]

(6)

For \(H_h\) increasing on its support interval the functions \(\gamma (\theta)\) and \(M_p (y)\) are convex. Clearly, \(Q_{p,h} (F)\) is then the unique minimizer of \(\gamma (\theta)\) and hence it can be 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 \(F\). Let us note that, whenever \(K\) 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 the 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 & 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 symmetric distribution with a cdf $K$. By (5), the functional $Q_{p,h}(F)$, called a perturbed p-quantile of $F$, is coinciding with a $p$-quantile of $X = Y - hZ$. $Q_{p,h}(F)$ is also minimizing

$$M_{p,h}(\theta) = E_{F}M_{p}\left(\frac{Y - \theta}{h}\right),$$

(7)

where $M_{p}(y)$ is given by (6). By the symmetry of $Z$ we have $X \sim Y - hZ \sim Y + hZ$ and, 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 theorem 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 $F$, only by a term of order $h^2$ multiplied by a score function

$$J(p) = - \frac{f'(Q(p))}{f(Q(p))}.$$
where \( f = F' \). The score function \( J(p) \) plays an important role in nonparametric statistics, cf. Hájek & Šidák (1967), Parzen (1979) and Behnen & Neuhaus (1989) and in the following we will discuss new estimators of \( J(p) \) suggested by Theorem 1.

**Theorem 1** Assume that \( F \) has three continuous derivatives vanishing at \(-\infty\) and the \( M\)-function is given by (6), where \( K \) is a cdf of a symmetric distribution with a compact support and the second moment \( \kappa_2 \). If \( f(Q_p) > 0 \) then the perturbed quantile \( Q_{p,h}(F) \) has the following Taylor expansion

\[
Q_{p,h}(F) = Q(p) + \frac{\kappa^2}{2} J(p) h^2 + o(h^2),
\]

(9)

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

**Proof of Theorem 1.**

Let \( q \) be a \( p\)-quantile of \( U \), \( q_o \) a \( p\)-quantile of \( Y \) and let \( c = q - q_o \). Denote also by \( \kappa_m \) the \( m\)-th moment of \( Z \). By the definition of quantile and by Taylor expansion we have the following.

\[
\begin{align*}
 p &= \int_{-\infty}^{q} g_h(u) \, du = \int_{-\infty}^{q} \int_{-\infty}^{\infty} \frac{1}{h} f(u - z) k\left(\frac{z}{h}\right) \, dz \, du \\
 &= \int_{-\infty}^{q} \left( f(u) + \frac{\kappa^2}{2} f''(u) h^2 \right) \, du + o(h^2) \\
 &= F(q_0 + c) + \frac{\kappa^2}{2} f''(q) h^2 + o(h^2) \\
 &= p + f(q_0) c + \frac{1}{2} f'(q_0) c^2 + \frac{\kappa^2}{2} f'(q) h^2 + o(h^2) \\
 &= p + f(q_0) c + \frac{1}{2} f'(q_0) c^2 + \frac{\kappa^2}{2} (f'(q_0) + cf''(q_0)) h^2 + o(h^2).
\end{align*}
\]

Hence, by skipping the \( o(h^2) \) contribution we get

\[
f(q_0) c + \frac{1}{2} f'(q_0) c^2 + \frac{\kappa^2}{2} (f'(q_0) + cf''(q_0)) h^2 = 0.
\]

8
By taking Taylor expansion of order 2 of the solution to the last equation we get

\[ c(h) = -\frac{\kappa_2 f'(q_0)}{2} \frac{1}{f(q_0)} h^2 + o(h^2) \]

and, consequently, we obtain expansion (9)

\[ q = q_o + c = q_o - \frac{\kappa_2 f'(q_0)}{2} \frac{1}{f(q_0)} h^2 + o(h^2) . \quad (10) \]

\[ \diamond \]

4 Variances of perturbed sample quantiles for small \( h \).

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

\[ \widehat{M}_{p,h}(\theta) = \frac{1}{n} \sum_{i=1}^{n} M_p \left( \frac{Y_i - \theta}{h} \right) \quad (11) \]

of the convex functional \( M_{p,h} \) given by (7). The estimator, to be denoted by \( \hat{Q}_{p,h} = Q_{p,h}(\hat{F}_n) \), 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

\[ \sigma^2(h) = Var \left( \hat{Q}_{p,h} \right) = \frac{\int \psi^2(y, Q_{p,h}(F)) dF(y)}{\left( \int \psi(y, Q_{p,h}(F)) dF(y) \right)^2}, \quad (12) \]

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 \( t = Q_{p,h} (F) \) we have

\[
\sigma^2(h) = \frac{\int \left[ 2K\left( \frac{y-t}{h} \right) - 2(1-p) \right]^2 dF(y)}{\left( \frac{2}{h} \int k\left( \frac{y-t}{h} \right) F(dy) \right)^2} \tag{13}
\]

and hence

\[
\lim_{h \to 0} \sigma^2(h) = \frac{\int \left[ 2 \times 1_{[Q_p(F),\infty)}(y) - 2(1-p) \right]^2 dF(y)}{(2f(Q_p(F)))^2} = \frac{\int \left[ 4 \times 1_{[Q_p(F),\infty)}(y) + 4(1-p)^2 - 8 \times 1_{[Q_p(F),\infty)}(y)(1-p) \right] dF(y)}{(2f(Q_p(F)))^2} = \frac{4(1-p) + 4(1-p)^2 - 8(1-p)^2}{(2f(Q_p(F)))^2} = \frac{p(1-p)}{f^2(Q_p(F))}. \tag{14}
\]

This shows that the limit of asymptotic variances \( \sigma^2(h) \) equals the asymptotic variance of the sample quantile (1), cf. Mosteller (1946).

To find the behaviour of the asymptotic variance \( \sigma^2(h) \) we need to find Taylor expansion for both numerator and denominator of (13). We assume here that the probability distribution corresponding to the cdf \( K \) is symmetric, concentrated on interval \((-1,1)\) and that the derivative \( k = K' \) is positive on \((-1,1)\). For \( t = Q_{p,h} (F) \) we have

\[
\sigma^2(h) = \frac{\int \left[ 2K\left( \frac{y-t}{h} \right) - 2(1-p) \right]^2 f(y)dy}{\left( \frac{2}{h} \int k\left( \frac{y-t}{h} \right) \left( f(t) + f'(t)(y-t) + \frac{1}{2} f''(t)(y-t)^2 + o(h^2) \right) dy \right)^2} \nonumber
\]

\[
= \frac{\int \left[ 2K\left( \frac{y-t}{h} \right) - 2(1-p) \right]^2 f(y)dy}{4 \left( \int f(t) + \frac{1}{2} h^2 K_2 f''(t) + o(h^2) \right)^2} \nonumber
\]

\[
= \frac{\int \left[ 2 \times 1_{[t,\infty)}(y) - 2(1-p) + 2 \left( K\left( \frac{y-t}{h} \right) - 1_{[t,\infty)}(y) \right) \right]^2 f(y)dy}{4 \left( \int f(Q_p(F)) + \frac{1}{2} h^2 K_2 (f''(Q_p(F)) - f'(Q_p(F)) + o(h^2)) \right)^2} \nonumber
\]

and with some algebra we get

\[
= \frac{4p(1-p) + 4 \int \left( K\left( \frac{y-t}{h} \right) - 1_{[t,\infty)}(y) \right)^2 f(y)dy + 8 \left( p \int_t^{t+h} K\left( \frac{y-t}{h} \right) f(y)dy - (1-p) \int_{t-h}^{t} K\left( \frac{y-t}{h} \right) f(y)dy \right)}{4 \left( \int f(Q_p(F)) + \frac{1}{2} h^2 K_2 (f''(Q_p(F)) - f'(Q_p(F)) + o(h^2)) \right)^2} \nonumber
\]

\[
= \frac{4p(1-p) + Ah^2 - Bh + o(h^2)}{4 \left( \int f(Q_p(F)) + \frac{1}{2} h^2 K_2 (f''(Q_p(F)) - f'(Q_p(F)) + o(h^2)) \right)^2} \tag{15}
\]
for some positive $A > 0$ and $B \geq 0$, where

$$0 < 4 \int \left( K\left( \frac{y-t}{h} \right) - 1_{[t,\infty)}(y) \right)^2 f(y)dy = Ah^2 + o(h^2)$$

and

$$0 > p \int_t^{t+h} \left( K\left( \frac{y-t}{h} \right) - 1 \right) f(y)dy - (1-p) \int_{t-h}^{t} K\left( \frac{y-t}{h} \right) f(y)dy = -Bh + o(h).$$

Hence, the variance $\sigma^2(h)$ is non-increasing for sufficiently small $h$ and $\frac{d}{dh}\sigma^2(h)|_{h=0} \leq 0$. We summarize this result in the following theorem.

**Theorem 2** Assume that the probability distribution corresponding to the cdf $K$ is symmetric, concentrated on interval $(-1,1)$ and that the derivative $k = K'$ is positive on $(-1,1)$. If $F$ is three times continuously differentiable with $F'(t) = f(t)$ then

$$\lim_{h \to 0} \sigma^2(h) = \frac{p(1-p)}{f^2(Q_p(F))},$$

where the limit equals the variance of the empirical quantile and

$$\sigma^2(h) = \frac{4p(1-p) - Ah + Bh^2 + o(h^3)}{4 \left( \left. \frac{f'(Q_p(F))}{f(Q_p(F))} + \frac{h^2}{2} \right) \left. \left( f''(Q_p(F)) - f'(Q_p(F)) \right) + o(h^2) \right|_{h=0} \right)^2,$$

where constants $A > 0$ and $B \geq 0$. This implies that the variance $\sigma^2(h)$ is decreasing in the vicinity of 0.

We refer to Green (2002) for simulations in the case of the Uniform $U(0,1)$, Exponential $E(\beta = 2)$ and the standard normal 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 $Q_{ph}(F)$, the perturbed p-quantiles of $F$ and in the right-hand side graphs, the corresponding asymptotic variances of the sample functionals $Q_{ph}\left( \hat{F}_n \right)$. 

11
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 $U(-h,h)$.

Figures 1 – 3 show how the neighborhood 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 $Q_{ph}(\hat{F}_n)$ 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 1 and 2 are valid only in the regions where the Taylor expansion provides an adequate approximation.
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 $U(-h, h)$.
5 Combined perturbed quantiles.

Theorems 1 and 2 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 $\frac{\kappa^2}{2} J(p) h^2 + o(h^2)$ while the corresponding M-estimators $Q_{p,h}(\hat{F}_n)$ have smaller asymptotic variance than 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, i.e. a set of values

$$0 \leq h_1 < h_2 < \cdots < h_m, \text{ with } m > 3.$$  \hspace{1cm} (16)

2. Calculate estimators $Q_{p,h_i}(\hat{F}_n)$ 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

$$q(h) = \beta_0 + \beta_1 h^2 + \beta_2 h^3.$$  \hspace{1cm} (17)

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 $Q(p)$ and

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 1 implies that in the Taylor expansion of $Q_{p,h}(F)$ the linear term in $h$ vanishes, so, a polynomial
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 $(0,0.1,0.5,1,2)$.

$q(h)$ 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.

Below, 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 $\hat{J}_n(p)$. 
Standard Deviation of Emp. Quantiles and of Emp. CP−Quantiles.

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

Standard Deviation of Emp. Quantiles and of Emp. CP−Quantiles.

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 $(0, 0.1, 0.5, 1, 2)$. 

16
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 $\hat{J}_n(p)$. First, by Corollary 3.2, p. 133 of Huber (1981) we get the asymptotic distribution of a vector of M-estimators.

$$\hat{Q}_{n,p,h_1,\ldots,h_m}(\hat{F}_n) = \begin{bmatrix} Q_{p,h_1}(\hat{F}_n) & \ldots & Q_{p,h_m}(\hat{F}_n) \end{bmatrix}^T$$

$$\sim AN\left(Q_{p,h_1,\ldots,h_m}(F), \frac{1}{n}\Sigma\right),$$

where

$$\Sigma = \left( \int \psi(y,Q_{p,h_j}(F))\psi(y,Q_{p,h_i}(F))dF(y) \right)_{i=1,\ldots,m} \left( \int \psi^t(y,Q_{p,h_j}(F))\psi^t(y,Q_{p,h_i}(F))dF(y) \right)_{j=1,\ldots,m}.$$  \hspace{1cm} (20)

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

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

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

$$\begin{bmatrix} \hat{\beta}_0 \hat{\beta}_1 \hat{\beta}_2 \end{bmatrix}^T = B\hat{Q}_{n,p,h_1,\ldots,h_m}(\hat{F}_n) = (D^T D)^{-1} D^T \hat{Q}_{n,p,h_1,\ldots,h_m}(\hat{F}_n).$$

Let $C = [B_{ij}]_{i=1,2,\ldots,m}$ be the left-upper sub-matrix of $B$ of size $(2, m)$. With some algebra we get

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \frac{(S_4 S_6 - S_1^2)}{S_4} T_0 + (S_1 S_5 - S_2 S_4) T_2 + (S_2 S_5 - S_4 S_3) T_3 \\ \frac{(S_4 S_6 - S_1^2)}{S_4} S_0 + (S_1 S_5 - S_2 S_4) S_2 + (S_2 S_5 - S_4 S_3) S_3 \\ (S_5 S_3 - S_2 S_6) T_0 + (S_0 S_6 - S_2^2) T_2 + (S_4 S_2 - S_6 S_5) T_3 \\ (S_5 S_3 - S_2 S_6) S_2 + (S_0 S_6 - S_2^2) S_4 + (S_4 S_2 - S_6 S_5) S_5 \end{bmatrix}. $$

(22)
where
\[
S_i = \sum_{j=1}^{m} h_{ij}^j \quad \text{and} \quad T_i = \sum_{j=1}^{m} h_{ij}^j Q_{pj} \left( \hat{F}_n \right).
\]
(23)

The joint asymptotic distribution of \( \hat{Q}_n(p) \) and \( \hat{J}_n(p) \) is given by
\[
\left[ \begin{array}{c}
\hat{Q}_n(p) \\
\hat{J}_n(p)
\end{array} \right] \sim AN \left( CQ_{p,h_1,...,h_n}(F), \frac{1}{n} C\Sigma C^T \right).
\]
(24)

References


Springer-Verlag, New York. With applications to statistics.

Yamato, H. (1972/73), ‘Uniform convergence of an estimator of a distribution function’,