Ann. Inst. Statist. Math.
Vol. 47, No. 2, 371–384 (1995)

RELATING QUANTILES AND EXPECTILES UNDER WEIGHTED-SYMMETRY*

BELKACEM ABDOUS AND BRUNO REMILLARD

Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières, Québec G9A 5H7, Canada

(Received May 10, 1993; revised October 5, 1994)

Abstract. Recently, quantiles and expectiles of a regression function have been investigated by several authors. In this work, we give a sufficient condition under which a quantile and an expectile coincide. We extend some classical results known for mean, median and symmetry to expectiles, quantiles and weighted-symmetry. We also study split-models and sample estimators of expectiles.

Key words and phrases: Quantiles, expectiles, weighted-symmetry, split-models.

1. Introduction

Given an explanatory variable X and a response variable Y, a regression model attempts to describe the relationship between X and Y. Typical regression models have the form

$$Y = m(X) + \varepsilon, \tag{1.1}$$

where the function m is usually referred to as the regression function or regression mean, and ε is a known random variable. These kind of models have been widely studied. However other features of Y such as its extreme behavior did not receive as much attention. Recently, in the economics literature, several authors investigated conditional percentiles. These quantities are found to be useful descriptors of regression data sets. It started with a paper of Koenker and Basset (1978) where they defined conditional quantiles (or regression quantiles) via an asymmetric absolute loss. In the context of linear models (i.e. $m(x) = a + bx$), if (x_i, y_i), $i = 1, \dots, n$ is a cloud of points in $\mathbb{R}^2$, and α is fixed in $(0, 1)$, then a quantile line of order α is defined by $y = \hat{a}_\alpha + \hat{b}_\alpha x$ where $\hat{a}_\alpha$ and $\hat{b}_\alpha$ minimize

$$\sum_{i=1}^{n} \rho_\alpha(y_i - a - bx_i)$$

* Work supported by the Natural Science and Engineering Council of Canada and by the Université du Québec à Trois-Rivières.

with

$$\rho_\alpha(t) = |\alpha - 1_{\{t\le 0\}}||t| = \begin{cases} \alpha|t| & \text{if } t > 0 \\ (1-\alpha)|t| & \text{if } t \le 0. \end{cases} \tag{1.2}$$

Later Newey and Powell (1987) criticized the use of regression quantiles in linear models and instead of (1.2), they proposed an asymmetric quadratic loss:

$$\psi_\alpha(t) = |\alpha - 1_{\{t\le 0\}}|t^2 = \begin{cases} \alpha t^2 & \text{if } t > 0 \\ (1-\alpha)t^2 & \text{if } t \le 0. \end{cases} \tag{1.3}$$

These authors called the resulting curves regression expectiles. More details and results on regression expectiles are given by Efron (1991) who used an equivalent criteria to (1.3) namely

$$\psi_\omega(t) = (\omega 1_{\{t>0\}} + 1_{\{t\le 0\}})t^2,$$

where ω is a positive constant. Finally Breckling and Chambers (1988) embedded both quantiles and expectiles in the general class of M-estimators by proposing asymmetric M-estimators. Some non-parametric estimates of regression expectiles were proposed: smoothing splines (Wang (1992)), kernel estimates (Abdous (1992)).

We will see that several models of distributions used in practice indeed possess what we call weighted-symmetry, a notion that generalizes the classical notion of symmetry. Under the hypothesis of symmetry, it is well-known that the center of symmetry, the mean (when it exists) and the median coincide. As pointed out by Rosenberger and Gasko (1983), this fact provides threefold heuristic basis for defining the location parameter; it also enriches the meaning of location parameter for symmetric distributions and provides a starting point in the search for robust estimators. Motivated by this useful property, we show that in general, under the hypothesis of weighted-symmetry one expectile and one quantile also coincide with the center of weighted-symmetry, thus yielding two estimators of that important parameter of location. In addition, this provides a class of asymmetric distributions which have a "natural" location parameter with the same threefold interpretation as the center of symmetry of a symmetric distribution.

This work is organized as follows: Section 2 gives some properties of expectiles and introduces the notion of weighted-symmetry; Section 3 deals with an extension of weighted-symmetry and studies split-models; Section 4 characterizes weighted-symmetry and Section 5 gives some asymptotic results on sample estimators of quantiles and expectiles.

2. Properties of expectiles

From now on, we will consider only the case of a single random variable, since the regression case (1.1) can be obtained by studying the random variable ε.

DEFINITION 1. Let X be a random variable with distribution function F. Let $\alpha \in (0,1)$ be fixed. Suppose that X has a finite second moment. We say that μ_α is an α-th expectile of F if

$$\mu_\alpha = \operatorname{argmin}_\theta E[|\alpha - 1_{\{X \leq \theta\}}|(X-\theta)^2]. \tag{2.1}$$

Remark that when X has only a finite absolute moment of order one, then instead of (2.1), one may define μ_α by

$$\mu_\alpha = \operatorname{argmin}_\theta E[\psi_\alpha(X-\theta) - \psi_\alpha(X)]$$

where ψ_α is given by (1.3).

Quantiles and expectiles both characterize a distribution function although they are different in nature. As an illustration, Fig. 1 plots curves of quantiles and expectiles of the standard normal $N(0,1)$.

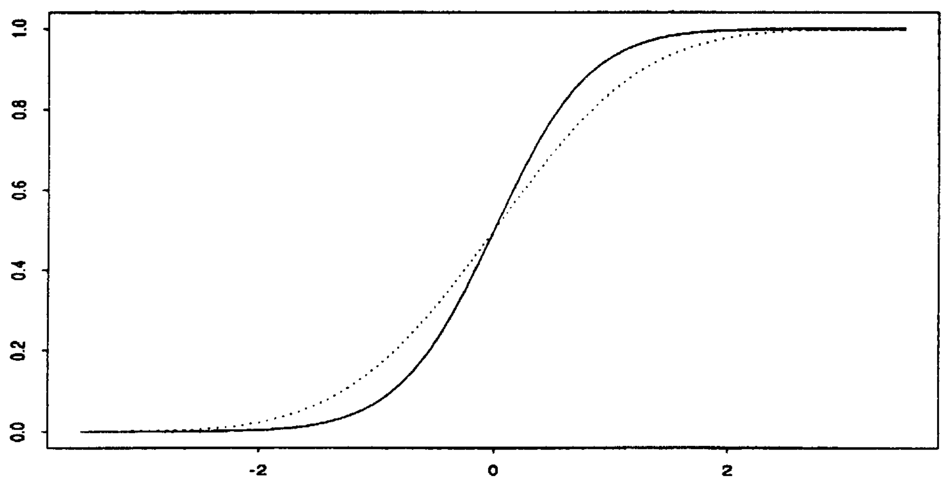

Fig. 1. Expectiles (solid curve) and quantiles (broken curve) of a normal $N(0,1)$.

Since $E[\psi_\alpha(X-\theta) - \psi_\alpha(X)]$ is differentiable in θ, μ_α satisfies

$$\alpha E[|X - \mu_\alpha|] = E[1_{\{X \leq \mu_\alpha\}}|X - \mu_\alpha|],$$

which is equivalent to saying that $G(\mu_\alpha) = \alpha$, if the law of X is not concentrated at one point, where

$$G : t \mapsto \frac{\int_{-\infty}^{t} |t-x| dF(x)}{\int_{-\infty}^{+\infty} |t-x| dF(x)}.$$

Hence μ_α may be expressed as the α-quantile of the distribution function G.

The following theorem due to Newey and Powell (1987) gives some properties of expectiles.

THEOREM 2.1. *Suppose that* $m = E(X)$ *exists, and put* $I_F = \{x \mid 0 < F(x) < 1\}$. *Then for each* $0 < \alpha < 1$, *a unique solution* μ_α *of* (1.2) *exists and has the following properties*:

(i) *As a function* $\mu_\alpha : (0,1) \mapsto \mathbb{R}$, μ_α *is strictly monotonic increasing.*

(ii) *The range of* μ_α *is* I_F *and* μ_α *maps* $(0,1)$ *onto* I_F.

(iii) *For* $\tilde{X} = a + bX$, *the* α*-th expectile* $\tilde{\mu}_\alpha$ *of* $\tilde{X}$ *satisfies*:

$$\tilde{\mu}_\alpha = \begin{cases} a + b\mu_\alpha & \text{if } b > 0 \\ a + b\mu_{1-\alpha} & \text{if } b < 0. \end{cases}$$

PROOF. See Newey and Powell ((1987), Theorem 1).

If the mean exists and if the median is uniquely defined, we have $\mu_{1/2}$ = mean and $\xi_{1/2}$ = median; if in addition F is symmetric about θ, i.e.

$$1 - F(\theta + x) - F(\theta - x - 0) = 0, \qquad \forall x \in \mathbb{R}, \tag{2.2}$$

where $F(t - 0) = \lim_{h \uparrow 0} F(t + h)$, then

$$\xi_{1/2} = \mu_{1/2} = \theta.$$

Our aim is to extend a similar relationship to any α in $(0,1)$. To this end, observe that when the mean $\mu_{1/2}$ exists, it satisfies the following equation:

$$\int_0^\infty [1 - F(\mu_{1/2} + z) - F(\mu_{1/2} - z)]dz = 0. \tag{2.3}$$

Moreover, the integrand in the previous equation corresponds to the term on left-hand-side of (2.2). Thus, for an α-th expectile, a similar equation to (2.3) should give us a generalization of the classical symmetry. Indeed, for $\alpha \in (0,1)$ and any distribution function F such that μ_α exists, μ_α satisfies

$$\int_0^\infty [1 - F(\mu_\alpha + z) - \omega F(\mu_\alpha - z)]dz = 0$$

where $\omega = (1-\alpha)/\alpha$. Therefore one possible generalization of the classical symmetry is given by the following definition.

DEFINITION 2. A random variable X or its distribution function F is said to be weighted-symmetric about θ if and only if F is such that

$$1 - F(\theta + |x|) - \omega F(\theta - |x| - 0) = 0, \qquad \forall x \in \mathbb{R}, \tag{2.4}$$

where $\omega = (1 - F(\theta))/F(\theta - 0)$.

Remark that equation (2.4) in the last definition is equivalent to the following: for any Borel subset A in $(0, \infty)$,

$$\omega P(X - \theta \in -A) = P(X - \theta \in A).$$

This concept of weighted-symmetry is not new. It has been used by Parent (1965) in order to construct sequential signed rank tests for the one-sample location problem. Wolfe (1974) gave a characterization of weighted-symmetry and some related results. Our motivation in studying weighted-symmetry came from the reading of the nice papers of Aki (1987) and Nabeya (1987) who considered weighted-symmetry as an extension of the problem of testing symmetry about zero. In fact, weighted-symmetry is a special case of alternatives to symmetry proposed by Lehmann (1953). In the next theorem, we explore the relationship between the center of weighted-symmetry, and certain quantile and expectile.

From now on, we will suppose that the distribution function F is such that the quantiles are uniquely defined, i.e. the support of the underlying random variable is a closed finite interval or an interval of the form $[a, +\infty)$, $(-\infty, b]$ or $(-\infty, +\infty)$.

THEOREM 2.2. (i) *Assume that there exists $\omega > 0$ such that for all $x \geq 0$, the following condition is satisfied*:

$$1 - F(\xi_\alpha + x) - \omega F(\xi_\alpha - x - 0) \geq 0,$$

where $\alpha = (1+\omega)^{-1}$. Then

$$\xi_\alpha \leq \mu_\alpha.$$

(ii) *Suppose that F is weighted-symmetric about θ and $(1-F(\theta))/F(\theta-0) = \omega$. If $\alpha = (1+\omega)^{-1}$, then*

$$\xi_\alpha = \mu_\alpha = \theta.$$

Moreover, for any $u \in (0, \alpha)$,

$$\theta = \frac{1}{2}\xi_u + \frac{1}{2}\xi_{1-u\omega}.$$

PROOF. (i) Put $\Gamma(t) = \int_0^\infty [(1 - F(t+z)) - \omega F(t-z)]dz$. By definition of μ_α, $\Gamma(\mu_\alpha) = 0$. Hence

$$\begin{aligned}\Gamma(\xi_\alpha) &= \Gamma(\xi_\alpha) - \Gamma(\mu_\alpha)\\ &= \int_0^\infty \{[F(\mu_\alpha + z) - F(\xi_\alpha + z)] + \omega[F(\mu_\alpha - z) - F(\xi_\alpha - z)]\}dz.\end{aligned}$$

On the other hand, if we assume that

$$[(1 - F(\xi_\alpha + z)) - \omega F(\xi_\alpha - z - 0)] \geq 0, \qquad \forall z \geq 0,$$

then $\Gamma(\xi_\alpha) \geq 0$. This implies that there exists at least one $z_0 \geq 0$ such that

$$[F(\mu_\alpha + z_0) - F(\xi_\alpha + z_0)] + \omega[F(\mu_\alpha - z_0) - F(\xi_\alpha - z_0)] \geq 0.$$

Hence $\xi_\alpha \leq \mu_\alpha$.

(ii) Weighted-symmetry about θ entails that

$$\omega F(\theta - 0) = 1 - F(\theta) \leq \min(\omega F(\theta), 1 - F(\theta - 0)).$$

Consequently

$$F(\theta - 0) \leq \frac{1}{1+\omega} = \alpha \leq F(\theta),$$

i.e. $\theta = \xi_\alpha$. As for the assertion $\theta = \mu_\alpha$, the hypothesis of weighted-symmetry gives

$$\int_0^\infty [(1 - F(\theta + z)) - \omega F(\theta - z - 0)]dz = 0.$$

An integration by parts enables to write

$$\int_{\mathbb{R}} (1_{\{x>\theta\}} + \omega 1_{\{x\leq\theta\}})(x - \theta)dF(x) = 0,$$

or equivalently

$$\int_{\mathbb{R}} |1 - \alpha 1_{\{x\leq\theta\}}|(x - \theta)dF(x) = 0,$$

i.e. $\theta = \mu_\alpha$. Next, let $u \in (0, \alpha)$ be fixed; weighted-symmetry implies that $\theta = \xi_\alpha$. Set one $z_0 = \theta - \xi_u$. If $u_0 = F(\theta - 0) < u_1 = F(\theta)$, then $\xi_u = \theta$ for all $u \in [u_0, u_1]$. Using weighted-symmetry, we get $u_1 = 1 - \omega u_0$. Therefore, for any $u \in [u_0, \alpha]$, $1 - \omega u \in [\alpha, u_1]$; hence $\xi_u = \theta = \xi_{1-\omega u}$. Finally, if $z_0 > 0$, then weighted-symmetry implies that

$$1 - F(\theta + z_0) = \omega F(\theta - z_0 - 0) = \omega F(\xi_u - 0) \leq u\omega$$

and

$$1 - F(\theta + z_0 - 0) = \omega F(\theta - z_0) = \omega F(\xi_u) \geq u\omega.$$

Combining the last two inequalities, we obtain

$$F(\theta + z_0 - 0) \leq 1 - u\omega \leq F(\theta + z_0),$$

proving that $\xi_{1-u\omega} = \theta + z_0 = 2\theta - \xi_u$. This completes the proof of the theorem.

Remark. Part (i) of Theorem 2.2 is an extension of a sufficient condition for the mean, median, mode inequality given by Van Zwet (1979).

3. Extension of weighted-symmetry and split-models

Weighted-symmetry, as introduced in Definition 2, is somewhat restrictive. It compares the distribution function at two symmetrical points $\theta + |x|$ and $\theta - |x|$. However, in many practical cases, one needs to compare F at $\theta + |x|$ and $\theta - |x|/\omega'$ with $\omega' > 0$. Therefore, we propose the following generalization of weighted-symmetry, which will replace Definition 2 in the sequel.

DEFINITION 3. A random variable X or its distribution function F is said to be weighted-symmetric about θ if and only if there exist two positive constants ω_1 and ω_2 such that

$$1 - F(\theta + |x|) - \omega_1 F(\theta - |x|/\omega_2 - 0) = 0, \qquad \forall x \in \mathbb{R}. \tag{3.1}$$

Observe that ω_1 must satisfy: $\omega_1 = (1 - F(\theta))/F(\theta - 0)$. In addition, equation (3.1) in Definition 3 is equivalent to the following useful property: for any Borel subset A of $(0, \infty)$, we have

$$\omega_1 P(X - \theta \in -A) = P(X - \theta \in \omega_2 A).$$

Weighted-symmetry is not a mathematical construction just for the fun of it. We now give a brief review of some weighted-symmetric models used in practice.

One such model is the two-piece normal, also called the joined half-Gaussian distribution or the split-normal distribution in the literature.

DEFINITION 4. (i) A random variable X has the continuous two-piece normal (CTPN) distribution with parameters θ, $\sigma_1 > 0$ and $\sigma_2 > 0$ if its probability density function is

$$f(x) = \begin{cases} A \exp(-(x-\theta)^2/(2\sigma_1^2)) & \text{if } x \le \theta \\ A \exp(-(x-\theta)^2/(2\sigma_2^2)) & \text{if } x > \theta \end{cases}$$

where $A = 2/(\sqrt{2\pi}(\sigma_1 + \sigma_2))$.

(ii) A random variable X has the discontinuous two-piece normal (DTPN) distribution with parameters θ, $\sigma_1 > 0$ and $\sigma_2 > 0$ if its probability density function is

$$f(x) = \begin{cases} A_1 \exp(-(x-\theta)^2/(2\sigma_1^2)) & \text{if } x \le \theta \\ A_2 \exp(-(x-\theta)^2/(2\sigma_2^2)) & \text{if } x > \theta \end{cases}$$

with $A_1 = 2\sigma_2/(\sqrt{2\pi}\sigma_1(\sigma_1 + \sigma_2))$ and $A_2 = 2\sigma_1/(\sqrt{2\pi}\sigma_2(\sigma_1 + \sigma_2))$.

We can see easily that both the CTPN and the DTPN distributions are weighted-symmetric with $\omega_1 = \omega_2 = \sigma_2/\sigma_1$ and $\omega_1 = 1/\omega_2 = \sigma_1/\sigma_2$ respectively.

Split-normal models were used and studied by several authors, perhaps the first one being Fechner (1897) (see Runnenburg (1978)). Later, Gibbons and Mylroie (1973) used these models to fit impurity profiles data in ion-implementation research. Since then, they have been used in applied physics, see Gibbons *et* *al.* (1975). An application in economics may be found in Aigner *et al.* (1976) who used split-normal models in the estimation of production frontiers. Properties of split-normal models have been investigated by John (1982) and Kimber (1985). Finally Lefrançois (1989) applied split-normal models to the estimation of future values in a forecasting process.

More general models than the DTPN models have been considered by Gupta (1967). The corresponding probability density may be written:

$$f(x) = g(x)1(x \leq 0) + g(x/\tau)\tau^{-1}1(x > 0),$$

where $\tau > 0$ and g is symmetric about 0. These models are weighted-symmetric with

$$\omega_1 = 1 \qquad \text{and} \qquad \omega_2 = \frac{1}{\tau}.$$

A very simple and curious weighted-symmetric model is the uniform distribution on (a, b). Indeed, for any $c \in (a, b)$, the uniform distribution is weighted-symmetric about c with

$$\omega_1 = \omega_2 = \frac{b-c}{c-a}.$$

Finally, let X_1 and X_2 be two independent exponential random variables with parameters β_1 and β_2 respectively. Then $Z = X_2 - X_1$ is weighted-symmetric with

$$\omega_1 = \omega_2 = \frac{\beta_2}{\beta_1}.$$

More generally, any symmetric model may be split in order to obtain a weighted-symmetric model. Indeed, let Z be a random variable with distribution function G symmetric about 0. Let $\theta \in \mathbb{R}$, $\alpha \in (0,1)$, $\sigma_1 > 0$ and $\sigma_2 > 0$. Then

$$X = \begin{cases} \theta - \sigma_1|Z| & \text{with probability } \alpha \\ \theta + \sigma_2|Z| & \text{with probability } (1-\alpha) \end{cases}$$

is a random variable whose distribution function is

$$F(x) = \begin{cases} 2\alpha G\left(\dfrac{x-\theta}{\sigma_1}\right) & \text{if } x \leq \theta \\ (2\alpha - 1) + 2(1-\alpha)G\left(\dfrac{x-\theta}{\sigma_2}\right) & \text{if } x > \theta \end{cases}$$

and X is weighted-symmetric with

$$\omega_1 = \frac{1-\alpha}{\alpha} \qquad \text{and} \qquad \omega_2 = \frac{\sigma_2}{\sigma_1}.$$

For particular values of α, σ_1 and σ_2, we retrieve CTPN, DTPN models, and Gupta's alternatives.

Some characteristics of F may be easily derived from those of G. We summarize some of them in the following proposition. The proof of this proposition is omitted since these results are obtained by standard manipulations.

PROPOSITION 3.1. (a) *If G is unimodal then F also does.*

(b) *Let β be fixed in $(0,1)$. Put $\xi_\beta^{(F)}$ and $\xi_\beta^{(G)}$ for β-th quantiles of F and G respectively. Then*

$$\xi_\beta^{(F)} = \begin{cases} \theta + \sigma_1 \xi^{(G)}_{\beta/(2\alpha)} & \text{if } \beta \le \alpha \\ \theta - \sigma_2 \xi^{(G)}_{(1-\beta)/(2(1-\alpha))} & \text{if } \beta > \alpha. \end{cases}$$

(c) *Let $r > 0$, if $\nu^+_{r,G} = E(Z^r 1_{\{Z \ge 0\}})$ is finite, then*

$$E(X-\theta)^r = \nu^+_{r,G}[2(1-\alpha)\sigma_2^r + 2\alpha(-\sigma_1)^r],$$

and

$$E|X-\theta|^r = \nu^+_{r,G}[2(1-\alpha)\sigma_2^r + 2\alpha\sigma_1^r].$$

Finally, let us mention that the relationships between quantiles, expectiles and weighted-symmetry given in Theorem 2.2 still hold for the weighted-symmetry given in Definition 3.

THEOREM 3.1. *Let $\omega_1 > 0$ and $\omega_2 > 0$ be fixed. Set $\alpha = (1+\omega_1)^{-1}$ and $\alpha' = (1+\omega_1\omega_2)^{-1}$.*

(i) *Assume that for all $x \ge 0$, the following condition is satisfied:*

$$1 - F(\xi_\alpha + x) - \omega_1 F(\xi_\alpha - x/\omega_2 - 0) \ge 0.$$

Then

$$\xi_\alpha \le \mu_{\alpha'}.$$

(ii) *Suppose that F is weighted-symmetric about θ relatively to (ω_1, ω_2). Then*

$$\xi_\alpha = \mu_{\alpha'} = \theta.$$

Moreover, for any $u \in (0,\alpha)$,

$$\theta = \frac{\xi_{1-\omega_1 u} + \omega_2 \xi_u}{1+\omega_2}.$$

PROOF. The proof is similar to the proof of Theorem 2.2.

4. Characterization of weighted-symmetry

When signed ranks are used for continuous random variables, the theory of distribution-free tests for the center of symmetry is essentially based on the following property.

"If a random variable X is symmetric about θ, then
$|X-\theta|$ and $1_{\{X \ge \theta\}}$ are stochastically independent."

Wolfe (1974) extended this result to the weighted-symmetry given in Definition 2. In the following theorem, we establish the same result for the weighted-symmetry introduced in Definition 3.

THEOREM 4.1. *Let ω_1 and ω_2 be two positive constants. Let X be a random variable with distribution function F and such that $F(\theta - 0) = F(\theta) = (1+\omega_1)^{-1}$. Put*

$$|X-\theta|_{\omega_2} = |X-\theta|\{1_{\{X>\theta\}} + \omega_2 1_{\{X\le\theta\}}\},$$

and

$$\operatorname{sign}(X-\theta) = \begin{cases} 1 & \text{if } X>\theta \\ 0 & \text{if } X=\theta \\ -1 & \text{if } X<\theta. \end{cases}$$

Then $|X-\theta|_{\omega_2}$ and $\operatorname{sign}(X-\theta)$ are independent if and only if X is weighted-symmetric about θ relatively to (ω_1, ω_2).

PROOF. First, we prove sufficiency. We have to show that if F is weighted-symmetric about θ, then, for any $x>0$ and $y=-1,0,1$

$$\Pr(|X-\theta|_{\omega_2} \le x, \operatorname{sign}(X-\theta) = y) = \Pr(|X-\theta|_{\omega_2} \le x) \Pr(\operatorname{sign}(X-\theta) = y). \tag{4.1}$$

Indeed, when $y=1$,

$$\Pr(|X-\theta|_{\omega_2} \le x, \operatorname{sign}(X-\theta)=1) = \Pr(\theta < X \le x+\theta) = F(x+\theta) - F(\theta),$$
$$\Pr(|X-\theta|_{\omega_2} \le x) = F(\theta+x) - F\left(\theta - \frac{x}{\omega_2} - 0\right) = \frac{1+\omega_1}{\omega_1} F(\theta+x) - \frac{1}{\omega_1},$$
$$\Pr(\operatorname{sign}(X-\theta)=1) = \frac{\omega_1}{1+\omega_1}.$$

This gives (4.1). Cases $y=0$ and $y=-1$ may be proved similarly.

Next, it remains to show that if the two random variables $|X-\theta|_{\omega_2}$ and $\operatorname{sign}(X-\theta)$ are stochastically independent, then F is weighted-symmetric about θ. Indeed, the independence of $|X-\theta|_{\omega_2}$ and $\operatorname{sign}(X-\theta)$ entails that for any $x>0$

$$\begin{aligned}\Pr(|X-\theta|_{\omega_2} \le x, \operatorname{sign}(X-\theta)=1) &= F(x+\theta) - F(\theta) \\ &= \Pr(|X-\theta|_{\omega_2} \le x)\Pr(\operatorname{sign}(X-\theta)=1) \\ &= \left[F(\theta+x) - F\left(\theta - \frac{x}{\omega_2} - 0\right)\right] \frac{\omega_1}{1+\omega_1},\end{aligned}$$

which in turn implies that F is weighted-symmetric about θ relatively to (ω_1, ω_2).

Remark. The characterization given in Theorem 4.1 may be used to extend many results on rank tests to weighted-symmetry's tests. For example, Wilcoxon's signed rank test can be generalized in the following way. Suppose you have a sample $X_1, \ldots, X_n$ and the null hypothesis is: this sample comes from a weighted-symmetric distribution about θ relatively to (ω_1, ω_2). To test that hypothesis, define the statistic

$$AR_n = \sum_{i=1}^{n} \operatorname{sign}(X_i - \theta) \operatorname{Rank}(|X_i - \theta|_{\omega_2}).$$

Note that we obtain Wilcoxon's signed rank statistic when $\omega_2 = 1$. Then, under the null hypothesis, AR_n has the same distribution as $\sum_{i=1}^{n} i\varepsilon_i$, where the ε_i are i.i.d. random variables such that $P(\varepsilon_i = 1) = \omega_1/(1+\omega_1)$ and $P(\varepsilon_i = -1) = (1+\omega_1)^{-1}$.

5. Asymptotic behavior of sample expectiles and quantiles

In this section we deal with sample expectiles and quantiles. Results below are quite standard and are easily obtained by means of M-estimation theory. Note that some of these results were obtained by Breckling and Chambers (1988) for M-quantiles.

Consider a sample $X_1, \ldots, X_n$ of size n, from F. Fix α in $(0,1)$, then an α-th sample quantile of F is given by:

$$\xi_{n,\alpha} = \operatorname{argmin}_{\theta} \sum_{i=1}^{n} |\alpha - 1_{\{(X_i - \theta) \leq 0\}}||X_i - \theta|.$$

In other words, if the sample arises form the density

$$f_1(x, \theta, \alpha) = \alpha(1-\alpha) \exp\left(-|\alpha - 1_{\{(x-\theta)\leq 0\}}||x - \theta|\right), \qquad x \in \mathbb{R},$$

then $\xi_{n\alpha}$ is the maximum likelihood estimator of the location parameter θ. Thus among all estimators of θ the α-th sample quantile has the minimum asymptotic variance at $f_1(\cdot, \theta, \alpha)$.

Similarly, $\mu_{n,\alpha'}$ the sample α-th expectile of F satisfies

$$\mu_{n,\alpha'} = \operatorname{argmin}_{\theta} \sum_{i=1}^{n} [1 + (\omega - 1) 1_{\{(X_i-\theta)\leq 0\}}](X_i - \theta)^2,$$

$$\text{(with } \alpha' = (1+\omega)^{-1}\text{)}.$$

The iteratively reweighted least squares algorithm may be used to evaluate $\mu_{n,\alpha'}$. Observe that $\mu_{n,\alpha'}$ is the maximum likelihood estimator of θ when the sample comes from the probability density

$$f_2(x, \theta, \alpha') = \frac{2\sqrt{\omega}}{\sqrt{2\pi}(1+\sqrt{\omega})} \cdot \exp\left(-[1 + (\omega - 1) 1_{\{(x-\theta)\leq 0\}}](x - \theta)^2/2\right), \qquad x \in \mathbb{R}.$$

In order to determine how $\mu_{n,\alpha'}$ and $\xi_{n,\alpha}$ asymptotically behave, it suffices to make use of the useful concepts of the influence curve discussed in Hampel (1974) and

M-estimation theory (see, e.g., Serfling (1980)). Indeed, standard manipulations show that the gross-error-sensitivity (see Hampel (1968, 1974)) of $\mu_{\alpha'}$ is infinite, indicating the non-robustness of $\mu_{\alpha'}$ and its extreme sensitivity to the influence of wild observations. Whereas the α-th-quantile ξ_α has a bounded gross-error-sensitivity which means that ξ_α is not affected by contamination of the data by gross errors. The expectile $\mu_{\alpha'}$ is resistant to systematic rounding and grouping since its local-shift-sensitivity is bounded. The α-th quantile has an infinite local-shift-sensitivity since its influence curve has a jump at ξ_α. This means that ξ_α is affected by rounding and grouping. Finally, both ξ_α and $\mu_{\alpha'}$ are not protected against sufficiently large outliers because their rejection points are infinite.

Next, it is well known that if ξ_α is unique then $\xi_{n,\alpha} \to \xi_\alpha$ a.s. as $n \to \infty$, and if $f(\xi_\alpha) > 0$ then $\xi_{n,\alpha}$ is asymptotically normal with mean ξ_α and variance $\sigma_1^2 = \alpha(1-\alpha)/(nf^2(\xi_\alpha))$ (see Serfling (1980), p. 77). Similar results hold for $\mu_{n,\alpha'}$. Standard results in M-estimation theory enable us to show that under mild conditions, $\mu_{n,\alpha'} \to \mu_{\alpha'}$ a.s. as $n \to \infty$, and $\mu_{n,\alpha'}$ is asymptotically normal with mean $\mu_{\alpha'}$ and variance σ_2^2

$$\sigma_2^2 = \frac{\int [1 + (\omega - 1)1_{\{(x-\mu_{\alpha'})\leq 0\}}]^2 (x - \mu_{\alpha'})^2 dF(x)}{n[1 + F(\mu_{\alpha'})(\omega - 1)]^2}.$$

When ξ_α and $\mu_{\alpha'}$ coincide i.e. F is weighted-symmetric about ξ_α (or $\mu_{\alpha'}$), then the comparison of $\xi_{n,\alpha}$ and $\mu_{n,\alpha'}$ may be made by means of the asymptotic relative efficiency $e(\xi_{n,\alpha}, \mu_{n,\alpha'}) = \sigma_2^2/\sigma_1^2$. For the particular case $\alpha = 1/2$ and symmetry, Lehmann ((1983), p. 359) showed that the asymptotic relative efficiency of the sample median to the sample mean is bounded below by $1/3$. A similar result holds for $e(\xi_{n,\alpha}, \mu_{n,\alpha'})$ and unimodal weighted-symmetric densities. Without loss of generality, we will assume that the center of weighted-symmetry is zero.

THEOREM 5.1. *If the underlying density f is unimodal and weighted-symmetric about zero or more generally f is weighted-symmetric about zero and satisfies*

$$f(x) \leq f(0) \quad \forall x,$$

then, the asymptotic relative efficiency of $\xi_{n,\alpha}$ to $\mu_{n,\alpha'}$ is such that

$$e(\xi_{n,\alpha}, \mu_{n,\alpha'}) \geq \frac{1}{3}.$$

The lower bound being attained for uniform distributions and no others.

PROOF. A mimic of the proof given by Lehmann ((1983), p. 359) enables to prove that the distribution function minimizing $e(\xi_{n,\alpha}, \mu_{n,\alpha'})$ is the uniform distribution on $(-(1+\sqrt{\omega})^{-1}, \sqrt{\omega}(1+\sqrt{\omega})^{-1})$. This concludes the proof.

Open questions. Many results on median, mean and symmetry are to be extended to quantiles, expectiles and weighted-symmetry. Some of them are

- $\mu_{n,\alpha'}$ should be more appropriate for estimating the center of weighted-symmetric short-tailed distributions, and $\xi_{n,\alpha}$ should be better for weighted-symmetric long-tailed distributions.

• When we estimate the center of a weighted-symmetric distribution, estimators based on criteria similar to that of Breckling and Chambers (1988) are more robust than $\xi_{n,\alpha}$ and $\mu_{n,\alpha'}$.

• When ω the parameter of weighted-symmetry is unknown, how it would be estimated? etc.

Acknowledgements

The authors wish to thank the two anonymous referees for helpful comments and suggestions on an earlier version of the paper.

References

Abdous, B. (1992). Kernel estimators of regression expectiles, Tech. Report, Université du Québec à Trois-Rivières.

Aigner, D., Amemiya, T. and Poirier, D. (1976). On the estimation of production frontiers: maximum likelihood estimation of the parameters of a discontinous density function, *Internat. Econom. Rev.*, **17**, 372–396.

Aki, S. (1987). On nonparametric tests for symmetry, *Ann. Inst. Statist. Math.*, **39**, 457–472.

Breckling, J. and Chambers, R. (1988). M-quantiles, *Biometrika*, **75**, 761–771.

Efron, B. (1991). Regression percentiles using asymmetric squared error loss, *Statistica Sinica*, **1**, 93–125.

Fechner, G. Th. (1897). *Kollektivmasslehre*, Leipzig, Engelman.

Gibbons, J. F. and Mylroie, S. (1973). Estimation of impurity profiles in ion-implanted amorphous targets using joined half-Gaussian distributions, *Applied Physical Letters*, **22**, 568–569.

Gibbons, J. F., Johnson, W. S. and Mylroie, S. (1975). *Projected Range Statistics*, 2nd ed., Wiley, New York.

Gupta, M. K. (1967). An asymptotically nonparametric test of symmetry, *Ann. Math. Statist.*, 849–866.

Hampel, F. R. (1974). The influence curve and its role in robust estimation, *J. Amer. Statist. Assoc.*, **69**, 383–393.

John, S. (1982). The three-parameter two-piece normal family of distributions and its fitting, *Comm. Statist. Theory Methods*, **11**, 879–885.

Kimber, A. C. (1985). Methods for the two-piece normal distribution, *Comm. Statist. Theory Methods*, **14**, 235–245.

Koenker, R. and Basset, G. (1978). Regression quantiles, *Econometrica*, **46**, 33–50.

Lefrançois, P. (1989). Allowing for asymmetry in forecast errors: Results from a Monte-Carlo study, *International Journal of Forcasting*, **5**, 99–110.

Lehmann, E. L. (1953). The power of rank tests, *Ann. Math. Statist.*, **24**, 23–43.

Lehmann, E. L. (1983). *Theory of Point Estimation*, Wiley, New York.

Nabeya, S. (1987) On Aki's nonparametric test for symmetry, *Ann. Inst. Statist. Math.*, **39**, 473–482.

Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing, *Econometrica*, **55**, 819–847.

Parent, E. A. (1965). Sequential ranking procedures, Tech. Report, No. 80, Department of Statistics, Stanford University, California.

Rosenberger, J. L. and Gasko, M. (1983). Comparing location estimators: trimmed means, medians, and trimean, *Understanding Robust and Exploratory Data Analysis* (eds. D. C. Hoaglin, F. Mosteller and J. W. Tukey), 297–338, Wiley, New York.

Runnenburg, J. Th. (1978). Mean, median, mode, *Statistica Neerlandica*, **32**, 73–79.

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

van Zwet, W. R. (1979). Mean, median, mode II, *Statistica Neerlandica*, **33**, 1–5.

Wang, Y. (1992). Smoothing splines for non-parametric regression percentiles, Tech. Report, No. 9207, Department of Statistics, University of Toronto.

Wolfe, D. A. (1974). A characterization of population weighted-symmetry and related results, *J. Amer. Statist. Assoc.*, **69** 819–822.