giusto da guardare, è sulla convergenza dei quantili nella Quantile regression

Econometric Theory, **2**, 1986, 191–201. Printed in the United States of America.

STRONG CONSISTENCY OF REGRESSION QUANTILES AND RELATED EMPIRICAL PROCESSES

GILBERT W. BASSETT

University of Illinois at Chicago

ROGER W. KOENKER

University of Illinois at Champaign–Urbana

The strong consistency of regression quantile statistics (Koenker and Bassett [4]) in linear models with iid errors is established. Mild regularity conditions on the regression design sequence and the error distribution are required. Strong consistency of the associated empirical quantile process (introduced in Bassett and Koenker [1]) is also established under analogous conditions. However, for the proposed estimate of the conditional distribution function of **Y**, *no regularity conditions on the error distribution are required for uniform strong convergence*, thus establishing a Glivenko-Cantelli-type theorem for this estimator.

1. INTRODUCTION

In several recent papers, Koenker and Bassett [1, 5] and Bassett and Koenker [1], we have explored the problem of estimating linear models for conditional quantile functions of related random variables. This approach complements classical least-squares methods for linear models as well as recent robust methods which focus exclusively on estimation of conditional central tendency.

The authors wish to express their appreciation to the editor and readers for an extremely careful and constructive review. The support of the NSF through grant SES-8408567 is also gratefully acknowledged.

We will consider the linear model

$$\mathbf{Y}_i = x_i\beta + u_i \qquad i = 1, 2, \ldots, n \tag{1.1}$$

where $x_i \in \mathbf{R}^p$ is a nonstochastic row vector, and β is an unknown p-dimensional parameter. The initial element of x_i will be taken to be 1 for all i, so the first element of β may be interpreted as an intercept parameter. Given an observed sample $y_i, \ldots, y_n$, we let $\hat{B}(\theta)$ denote the set of solutions to the problem

$$\min_{b \in \mathbf{R}^p} \sum_{i=1}^{n} \rho_\theta(y_i - x_i b) \tag{1.2}$$

where $\theta \in [0, 1]$ and $\rho_\theta(\cdot)$ is the "check function",

$$\rho_\theta(u) = \begin{cases} \theta u & u \geqslant 0 \\ (\theta - 1)u & u < 0 \end{cases}. \tag{1.3}$$

Elements of $\hat{B}(\theta)$ are denoted by $\hat{\beta}_\theta$. In Koenker and Bassett [4] we showed that with $\{u_i\}$ iid F, the sequence of solutions $\{\hat{\beta}_\theta\}$ had the property that, under mild conditions on the sequence of designs and the assumption that F had a positive density in a neighborhood of the θth quantile, $\mathbf{Q}(\theta) = \mathbf{F}^{-1}(\theta)$, $\sqrt{n}(\hat{\beta}_\theta - \beta - \xi_\theta)$ converged in law to a p-variate Gaussian distribution with mean vector 0, $\xi_\theta = (\mathbf{Q}(\theta), 0, \ldots, 0)' \in \mathbf{R}^p$. Thus, under the foregoing conditions $\hat{\beta}_\theta$ is weakly consistent for $\beta_\theta = \beta + \xi_\theta$. In Koenker and Bassett [5] we showed that similar asymptotic behavior prevailed in sequences of linear models with heteroscedasticity of order $O(1/\sqrt{n})$. In this paper we maintain the hypothesis of iid errors while relaxing our previous smoothness conditions on the error distribution F. We begin by treating the behavior of the p-dimensional regression quantiles and conclude by treating the associated empirical processes introduced in Bassett and Koenker [1].

The estimate of the conditional quantile function of $\mathbf{Y}$ we propose is

$$\hat{\mathbf{Q}}_\mathbf{Y}(\theta \mid x) = \inf\{xb \mid b \in \hat{\mathbf{B}}(\theta)\}. \tag{1.4}$$

It was previously shown that at $x = \bar{x} = n^{-1}\sum x_i$, the sample paths of the random function

$$\hat{\mathbf{Q}}_\mathbf{Y}(\theta) \equiv \hat{\mathbf{Q}}_\mathbf{Y}(\theta \mid \bar{x}) \tag{1.5}$$

are non-decreasing, left continuous jump functions on (0, 1). However, unlike the ordinary empirical quantile function to which $\hat{\mathbf{Q}}_\mathbf{Y}(\theta)$ specializes when $\mathbf{X}_n = 1_n$, an n-vector of ones, $\hat{\mathbf{Q}}_\mathbf{Y}(\theta)$ jumps at irregularly spaced points on

(0, 1). Similarly, one may show that

$$\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta + 0) \equiv \lim_{\varepsilon \downarrow 0} \hat{\mathbf{Q}}_{\mathbf{Y}}(\theta + \varepsilon) \tag{1.6}$$
$$= \sup \{\bar{x}b \mid b \in \hat{\mathbf{B}}(\theta)\}$$

is a nondecreasing, right-continuous jump function on (0, 1). It was also shown that properly normalized versions of these processes have finite dimensional distributions which converge to those of the Brownian Bridge process. This estimate of the quantile function may be "inverted" to obtain an empirical distribution function for the linear model

$$\hat{\mathbf{F}}_{\mathbf{Y}}(y) = \sup \{\theta \mid \hat{\mathbf{Q}}_{\mathbf{Y}}(\theta) \leqslant y\} \tag{1.7}$$

Portnoy [7] has recently shown that this estimator converges weakly to the Brownian Bridge process.

Almost sure convergence results are established below under mild regularity conditions on the design and the distribution function of the errors for $\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta)$. In the case of our proposed estimate of the error distribution $\mathbf{F}_{\mathbf{Y}}(y)$, *no regularity conditions are required on* **F**, thus providing a natural extension of the Glivenko-Cantelli theorem to the realm of linear models. The latter result provides an intriguing alternative to methods based on residuals for assessing distributional features of linear models with iid errors.

It has been long recognized that least squares residuals provide a rather unsatisfactory basis for assessing departures from the Gaussian error hypothesis in linear models. See, for example, the comments by Weisberg [9]. It is well known that the empirical distribution function of least-squares residuals is biased toward the Gaussian shape. Similar bias may be demonstrated for residuals based on other preliminary estimates, see Bassett and Koenker [1], and the references cited there. Thus it is of obvious interest to investigate methods for estimating the shape of the error distribution which do not rely on an empirical process based upon residuals. Our proposed methods offer a natural generalization of unidimensional empirical process to the linear regression model while avoiding some of the problems inherent in choosing a preliminary estimator from which to compute residuals.

2. STRONG CONSISTENCY OF REGRESSION QUANTILES

We will assume throughout that we have data generated from the model,

$$\mathbf{Y}_i = x_i\beta + u_i, \tag{2.1}$$

The errors u_i are assumed to be independently and identically distributed with distribution function **F**. About **F** we will assume that it is a proper, right-continuous distribution function. Its "inverse" will be denoted by

$$\mathbf{Q}(\theta) = \inf\{u \mid \mathbf{F}(u) \geqslant \theta\} \tag{2.2}$$

so $\mathbf{Q}$ is left continuous on $[0, 1]$. The parameter β is an unknown p-vector. The sequence $\{x_i\}$ of design vectors is assumed to "contain an intercept," that is, $x_{i1} = 1$ for all i and to satisfy, the following regularity conditions:

$$d_n \equiv \inf_{\|\omega\|=1} n^{-1} \sum_{i=1}^{n} |x_i\omega| \geqslant d > 0 \qquad n > n_0 \tag{D1}$$

$$\mathbf{D}_n \equiv \sup_{\|\omega\|=1} n^{-1} \sum_{i=1}^{n} (x_i\omega)^2 \leqslant \mathbf{D} < \infty \qquad n > n_0 \tag{D2}$$

(D1) is essentially an identifiability condition while (D2) bounds the rate of growth of design rows facilitating certain uniform continuity arguments.

We may now state:

THEOREM 2.1. *If* (D1) *and* (D2) *hold and* $\mathbf{F}$ *has a unique* θ*th quantile, that is,* $\mathbf{Q}(\theta) = \mathbf{Q}(\theta + 0)$, *then any sequence of solutions* $\{\hat{\beta}_n(\theta)\}$ *to problem (1.2) satisfies* $\hat{\beta}_n(\theta) \to \beta + \xi_\theta$, *almost surely.*

Proof. Consider,

$$\mathbf{R}_n(\delta) \equiv n^{-1} \sum_{i=1}^{n} r_i(\delta)$$

$$\equiv n^{-1} \sum_{i=1}^{n} [\rho_\theta(u_i - x_i\delta - \mathbf{Q}(\theta)) - \rho_\theta(u_i - \mathbf{Q}(\theta))] \tag{2.3}$$

Since $\mathbf{R}_n(0) = 0$, $\mathbf{R}_n(\delta) \leqslant 0$ on its minimum set. And $\mathbf{R}_n(\delta)$ is a sum of convex functions and therefore convex, and convexity assures that $\mathbf{R}_n(\delta)$ is strictly positive outside the Δ-ball, if for any $\Delta > 0$,

$$\liminf_{n\to\infty} \inf_{\|\delta\|=\Delta} \mathbf{R}_n(\delta) > 0 \tag{2.4}$$

Thus it suffices to show that (2.4) holds almost surely.

We begin by establishing that at each δ,

$$\mathbf{R}_n(\delta) - \mathbf{E}(\mathbf{R}_n(\delta)) \to 0 \quad \text{a.s.} \tag{2.5}$$

Since $|r_i(\delta)| < |x_i\delta|$, we have,

$$\sum_{i=1}^{\infty} \operatorname{Var}(r_i(\delta))/i^2 \leqslant \sum (x_i\delta)^2/i^2 \tag{2.6}$$

and using (D2),

$$\leqslant \|\delta\|^4 \mathbf{D}^2 \sum (x_i\delta)^2 / (\sum (x_i\delta)^2)^2$$

which is convergent (see Hardy, Littlewood, and Polya [3, p. 120]) so Kolmogorov's criterion is satisfied and (2.5) follows. This may be strengthened to uniform convergence on compacts by noting that for any $\delta_0 \in \mathbf{R}^p$, we have, for $n \geqslant n_o$,

$$\begin{aligned} \sup_{\|\delta-\delta_0\|\leqslant\varepsilon} \{|R_n(\delta) - R_n(\delta_0)|\} &\leqslant \sup_{\|\delta-\delta_0\|\leqslant\varepsilon} \{n^{-1}\sum |x_i(\delta-\delta_0)|\} \\ &\leqslant \varepsilon \sup_{\|\omega\|=1} n^{-1}\sum |x_i\omega| \\ &\leqslant \varepsilon \mathbf{D}_n^{1/2} \\ &\leqslant \varepsilon \mathbf{D}^{1/2}. \end{aligned} \tag{2.7}$$

Thus to establish (2.4) it remains to show that $\mathbf{E}R_n(\delta)$ is bounded away from 0, for any $\|\delta\| > 0$. For $\alpha > 0$, let

$$\begin{aligned} g(\alpha) &= \mathbf{E}[\rho_\theta(u - \alpha - \mathbf{Q}(\theta)) - \rho_\theta(u - \mathbf{Q}(\theta)] \\ &= \int_{-\infty}^{Q+\alpha} (1-\theta)\alpha\, d\mathbf{F}(u) - \int_{Q+\alpha}^{\infty} \alpha\theta\, d\mathbf{F}(u) - \int_{Q}^{Q+\alpha} (u - \mathbf{Q})\, d\mathbf{F}(u) \end{aligned} \tag{2.8}$$

and integrating by parts gives,

$$g(\alpha) = \int_{Q}^{Q+\alpha} (\mathbf{F}(u) - \theta)\, du. \tag{2.9}$$

For $\alpha < 0$, the sign and limits of integration are simply reversed. The function g is convex, $g(\alpha) \geqslant 0$, and $g(\alpha) = 0$ only for $0 \leqslant \alpha \leqslant \mathbf{Q}(\theta + 0) - \mathbf{Q}(\theta)$. Thus, when θ is unique $g(\alpha) = 0$ only at 0. Now, let $h(\alpha)$ denote the convex hull of $g(\alpha)$ and $g(-\alpha)$, that is, let h be the greatest convex function such that $h(\alpha) \leqslant g(\alpha)$ and $h(\alpha) \leqslant g(-\alpha)$ for all $\alpha \in \mathbf{R}$; see e.g., Rockefeller [8 p. 37]. Then we have,

$$\begin{aligned} \mathbf{E}R_n(\delta) &\geqslant n^{-1} \sum_{i=1}^{n} h(x_i\delta) \\ &= n^{-1}\sum h(|x_i\delta|) \\ &\geqslant h(n^{-1}\sum |x_i\delta|) \\ &\geqslant h(d\|\delta\|) \end{aligned} \tag{2.10}$$

by the symmetry of h, Jensen's inequality and (D1) respectively. The function $h(\alpha)$ is positive for $\alpha \neq 0$ by the uniqueness of the θth quantile, thus completing the proof. ■

Reviewing the preceding argument it is clear that uniqueness of the θth quantile is needed only to argue that $\mathbf{ER}_n(\delta)$ has a unique minimum at the origin. When the θth quantile is not unique then $\mathbf{ER}_n(\delta)$ has a larger minimum set, but it is straightforward to show that solutions to (1.2) converge almost surely to elements of this set. (See Koenker and Bassett [6] for an example of weak but not strong convergence in this context.) We can therefore establish a slightly more general form of the previous result.

THEOREM 2.2. *Let* $\mathbf{\Lambda}_n(\mathbf{\Delta}) = \{\delta \in \mathbf{R}^p \mid \mathbf{ER}_n(\delta) \leqslant \mathbf{\Delta}\}$, *the* $\mathbf{\Delta}$*-level set of* $\mathbf{ER}_n(\delta)$. *If* (D1) *and* (D2) *hold and* $\mathbf{F}$ *is any proper, right-continuous function then any sequence of solutions to (1.2),* $\{\hat{\beta}_n(\theta)\}$, *satisfies*

$$\hat{\beta}_n(\theta) - \beta - \xi_\theta \in \Lambda_n(\Delta) \quad \text{a.s.} \tag{2.11}$$

for all $\mathbf{\Delta} > 0$.

Proof. Let $\bar{\mathbf{\Lambda}}_n(\mathbf{\Delta})$ denote the boundary of $\mathbf{\Lambda}_n(\mathbf{\Delta})$. By the convexity argument of the preceding proof it suffices to show

$$\liminf_{n \to \infty} \inf_{\delta \in \bar{\Lambda}_n(\Delta)} R_n(\delta) > 0 \quad \text{a.s.} \tag{2.12}$$

This can be established by showing that $\bar{\mathbf{\Lambda}}_n(\mathbf{\Delta})$ is bounded, since we can then appeal to (D2), and use $\mathbf{ER}_n(\delta)$ to approximate $\mathbf{R}_n(\delta)$ on the compact set $\bar{\mathbf{\Lambda}}_n(\mathbf{\Delta})$ as in the proof of Theorem 2.1.

The sequence of steps in (2.10) remains valid here and implies

$$\mathbf{\Lambda}_n(\mathbf{\Delta}) \subset \{\delta \in \mathbf{R}^p \mid h(\|\delta\| d) \leqslant \mathbf{\Delta}\} \tag{2.13}$$

and the larger set is bounded because (i.) $d > 0$, and (ii.) the convex function h is zero only on the *bounded* interval with the endpoints $\pm[\mathbf{Q}(\theta + 0) - \mathbf{Q}(\theta)]$ and this completes the proof ■

To complete this section we consider the case in which $\mathbf{F}$ has positive mass at $\mathbf{Q}(\theta)$. Under this condition we can strengthen Theorem 2.1. The following result plays a crucial role in the next section on empirical processes.

THEOREM 2.3. *Fix* $\theta \in (0, 1)$, *and assume,* $\zeta(\theta) = \mathbf{F}(\mathbf{Q}(\theta)) - \mathbf{F}(\mathbf{Q}(\theta) - 0) > 0$, *so* $\mathbf{F}$ *has strictly positive mass at the* θ*th quantile. Then for any* $\varepsilon > 0$, *there is an* n_o *such that*

$$\mathbf{P}\{\hat{\mathbf{B}}_n(\theta) = \beta + \xi_\theta, n > n_o\} \geqslant 1 - \varepsilon. \tag{2.14}$$

Proof. Consider the directional derivative

$$\begin{aligned}\mathbf{R}'_n(\delta, w) &= n^{-1}\sum r'_i(\delta, w)\\ &= n^{-1}\sum\left[\tfrac{1}{2} - \theta - \tfrac{1}{2}sgn^*(u_i - x_i\delta, x_i w)\right]x_i w\end{aligned} \tag{2.15}$$

where $sgn^*(u, v) = sgn(u)$ for $u \neq 0$ and $sgn(v)$ otherwise. We must establish,

$$\liminf_{n\to\infty}\ \inf_{||w||=1}\ \mathbf{R}'_n(0, w) > 0 \quad \text{a.s.} \tag{2.16}$$

Following the approach used in the proof of Theorem 2.1 we begin by showing that for fixed w,

$$\mathbf{R}'_n(0, w) - \mathbf{ER}'_n(0, w) \to 0 \quad \text{a.s.} \tag{2.17}$$

Since $|r'_i(0, w)| \leqslant |x_i w|$

$$\sum_{i=1}^{\infty} Var(r'_i(0, w))/i^2 \leqslant \sum_{i=1}^{\infty} (x_i w)^2/i^2 < \infty \tag{2.18}$$

by (D2). So Kolmogorov's criterion implies (2.18). This may be strengthened to uniform convergence on the sphere $||w|| = 1$ using (D2) and the continuity of $\mathbf{R}'_n(0, w)$ in w, as in the argument following (2.6). Thus it remains to verify that $\mathbf{ER}'_n(0, w)$ is bounded away from zero.

Now,

$$\mathbf{E}r'_i(0, w) = \begin{cases}[\mathbf{F}(\mathbf{Q}(\theta)) - \theta]x_i w & \text{if } x_i w \geqslant 0\\ [\mathbf{F}(\mathbf{Q}(\theta) - 0) - \theta]x_i w & \text{otherwise}\end{cases} \tag{2.19}$$

so setting $m(\theta) = \min\{\mathbf{F}(\mathbf{Q}(\theta)) - \theta, \theta - \mathbf{F}(\mathbf{Q}(\theta) - 0)\}$ we have,

$$\mathbf{E}r'_i(0, w) \geqslant m(\theta)|x_i w|, \tag{2.20}$$

hence, using (D1), and the fact that $\zeta(\theta) > 0$,

$$\liminf_{n\to\infty}\ \inf_{||w||=1}\ \mathbf{ER}'_n(0, w) \geqslant m(\theta)d > 0, \tag{2.21}$$

which completes the proof ∎

3. STRONG CONVERGENCE OF EMPIRICAL PROCESSES BASED ON REGRESSION QUANTILES

Here we wish to investigate the strong convergence of $\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta)$ and $\hat{\mathbf{F}}_{\mathbf{Y}}(\theta)$ using results from the previous section. We may begin by noting that given the iid error assumption of model (1.2), the θth conditional quantile function of $\mathbf{Y}$ given x may be written as,

$$\mathbf{Q}_{\mathbf{Y}}(\theta \mid x) = x\beta + \mathbf{Q}(\theta) \tag{3.1}$$

We will restrict attention as previously to

$$\mathbf{Q}_{\mathbf{Y}}(\theta) = \mathbf{Q}_{\mathbf{Y}}(\theta \mid \bar{x}) = \bar{x}\beta + \mathbf{Q}(\theta). \tag{3.2}$$

We may now state:

THEOREM 3.1. *If* (D1) *and* (D2) *hold and* $\mathbf{Q}(\theta)$ *is continuous on a closed interval* $\boldsymbol{\Theta} \subset (0, 1)$ *then*

$$\sup_{\theta \in \boldsymbol{\Theta}} |\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta) - \mathbf{Q}_{\mathbf{Y}}(\theta)| \to 0 \quad \text{a.s.} \tag{3.3}$$

Proof. From (2.1) we have pointwise convergence of $\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta)$ and $\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta + 0)$ to $\mathbf{Q}_{\mathbf{Y}}(\theta)$ and using the monotonicity and continuity of $\mathbf{Q}_{\mathbf{Y}}(\theta)$ this may be strengthened to uniform almost sure convergence on $\boldsymbol{\Theta}$. See, for example, the argument in Billingsley [2, p. 233] for the Glivenko-Cantelli theorem ∎

The standard argument for Glivenko-Cantelli also shows that the uniform, almost sure convergence of $\hat{\mathbf{F}}$ follows from the pointwise convergence results:

$$\hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta)) - \mathbf{F}(\mathbf{Q}_{\mathbf{Y}}(\theta)) \to 0 \quad \text{a.s.} \tag{3.4}$$

$$\hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta) - 0) - \mathbf{F}(\mathbf{Q}_{\mathbf{Y}}(\theta) - 0) \to 0 \quad \text{a.s.} \tag{3.5}$$

We will establish (3.4) and (3.5) without any smoothness conditions on $\mathbf{F}$ using two lemmas.

LEMMA 3.1. *If* (D1) *and* (D2) *hold then at each* $\theta \in (0, 1)$,

$$\limsup_{n \to \infty} \hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta)) \leqslant \mathbf{F}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta)) \quad \text{a.s.} \tag{3.6}$$

and

$$\liminf_{n \to \infty} \hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta) - 0) \geqslant \mathbf{F}(\mathbf{Q}_{\mathbf{Y}}(\theta) - 0) \quad \text{a.s.} \tag{3.7}$$

Proof. It suffices to establish the inequalities:

$$\liminf_{n \to \infty} \hat{\mathbf{Q}}_{\mathbf{Y}}(\theta) \geqslant \mathbf{Q}_{\mathbf{Y}}(\theta) \quad \text{a.s.} \tag{3.8}$$

$$\limsup_{n \to \infty} \hat{\mathbf{Q}}_{\mathbf{Y}}(\theta + 0) \leqslant \mathbf{Q}_{\mathbf{Y}}(\theta + 0) \quad \text{a.s.} \tag{3.9}$$

From Theorem 2.2 we have for any $\mathbf{\Delta}$

$$\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta) \geqslant \mathbf{Q}_{\mathbf{Y}}(\theta) + \inf\{\bar{x}\delta \,|\, \delta \in \mathbf{\Lambda}_n(\mathbf{\Delta})\} \quad \text{a.s.} \tag{3.10}$$

and

$$\hat{\mathbf{Q}}_{\mathbf{Y}}(\theta + 0) \leqslant \mathbf{Q}_{\mathbf{Y}}(\theta) + \sup\{\bar{x}\delta \,|\, \delta \in \mathbf{\Lambda}_n(\mathbf{\Delta})\} \quad \text{a.s.} \tag{3.11}$$

By Jensen's inequality, and using (D1),

$$\begin{aligned} \{\bar{x}\delta \,|\, \lambda \in \mathbf{\Lambda}_n(\mathbf{\Delta})\} &= \{\bar{x}\delta \,|\, n^{-1} \sum_{i=1}^{n} g(x_i\delta) < \mathbf{\Delta}\} \\ &\subset \{\bar{x}\delta \,|\, g(\bar{x}\delta) < \mathbf{\Delta}\}. \end{aligned} \tag{3.12}$$

Since $g(\cdot)$ is convex and zero only on $[0, \mathbf{Q}(\theta + 0) - \mathbf{Q}(\theta)]$, recall (2.9) above, we have

$$\lim_{\Delta \downarrow 0} \inf\{\alpha \,|\, g(\alpha) < \Delta\} = 0 \tag{3.13}$$

and

$$\lim_{\Delta \downarrow 0} \sup\{\alpha \,|\, g(\alpha) < \Delta\} = \mathbf{Q}(\theta + 0) - \mathbf{Q}(\theta)$$

which imply the conclusion of the lemma ∎

Remarks

1. For continuous $\mathbf{F}$, (3.6-7) implies

$$\liminf_{n \to \infty} \hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta) - 0) = \limsup_{n \to \infty} \hat{\mathbf{F}}_{\mathbf{Y}}(\mathbf{Q}_{\mathbf{Y}}(\theta)) = \theta. \tag{3.14}$$

And by standard arguments for the Glivenko-Cantelli theorem, again see Billingsley [2, p. 233], this implies

$$\sup_{u \in \mathbf{R}} |\hat{\mathbf{F}}_{\mathbf{Y}}(u) - \mathbf{F}_{\mathbf{Y}}(u)| \to 0 \quad \text{a.s.} \tag{3.15}$$

2. Lemma 3.1 does not imply (3.15) when $\mathbf{F}$ has discontinuities. This is illustrated by the following example in which $\hat{\mathbf{Q}}$ converges to $\mathbf{Q}$ uniformly but the corresponding $\hat{\mathbf{F}}$ fails to converge to $\mathbf{F}$. Let $\mathbf{Q}(\theta) = 0$ for $\theta \in (0, 1]$ so that

the associated *df* is

$$\mathbf{F}(u) = \begin{cases} 0 & u < 0 \\ 1 & u \geqslant 0 \end{cases}. \tag{3.16}$$

Consider,

$$\hat{\mathbf{Q}}(\theta) = \begin{cases} -n^{-1} & \theta \in (0, \frac{1}{2}] \\ n^{-1} & \theta \in (\frac{1}{2}, 1] \end{cases} \tag{3.17}$$

and

$$\hat{\mathbf{F}}(u) = \begin{cases} 0 & u < -n^{-1} \\ \frac{1}{2} & -n^{-1} \leqslant u \leqslant n^{-1}. \\ 1 & n^{-1} \leqslant u \end{cases} \tag{3.18}$$

Then $\hat{\mathbf{Q}}(\cdot)$ satisfies (3.8-9), $\hat{\mathbf{Q}}$ actually converges to $\mathbf{Q}$ uniformly, but

$$\lim_{n \to \infty} \hat{\mathbf{F}}(\mathbf{Q}(\tfrac{1}{2})) = \hat{\mathbf{F}}(0) = \tfrac{1}{2} \neq 1 = \mathbf{F}(0) \tag{3.19}$$

so $\hat{\mathbf{F}}$ fails to converge to $\mathbf{F}$. Note that since the pointwise limit of $\hat{\mathbf{F}}$ is neither right or left continuous at 0 the limit fails to be a proper distribution function.

An immediate consequence, however, of the Theorem 2.3 is that under the same conditions,

$$\mathbf{P}\{\mathbf{Q_Y}(\theta) = \hat{\mathbf{Q}}_Y(\theta) = \hat{\mathbf{Q}}_\mathbf{Y}(\theta + 0) = \mathbf{Q_Y}(\theta + 0),\ n > n_o\} > 1 - \varepsilon \tag{3.20}$$

which implies,

LEMMA 3.2. *For any θ such that* $\mathbf{F}(\mathbf{Q}(\theta)) > \mathbf{F}(\mathbf{Q}(\theta) - 0)$,

$$\limsup_{n \to \infty} \hat{\mathbf{F}}_\mathbf{Y}(\hat{\mathbf{Q}}_\mathbf{Y}(\theta)) \geqslant \mathbf{F_Y}(\mathbf{Q_Y}(\theta)) \quad \text{a.s.} \tag{3.21}$$

$$\liminf_{n \to \infty} \hat{\mathbf{F}}_\mathbf{Y}(\hat{\mathbf{Q}}(\theta) - 0) \leqslant \mathbf{F_Y}(\mathbf{Q}(\theta) - 0) \quad \text{a.s.} \tag{3.22}$$

Lemmas 3.1 and 3.2 imply (3.4) and (3.5) which in turn establish our main result:

THEOREM 3.2. *Under* (D1) *and* (D2): $\sup_{u \in \mathbf{R}} |\hat{\mathbf{F}}_\mathbf{Y}(u) - \mathbf{F_Y}(u)| \to 0 \quad \text{a.s.}$

This provides a reasonably satisfactory extension of the Glivenko-Cantelli Theorem for linear models with p fixed. It remains to be seen what can be done when $p \to \infty$ with n.

REFERENCES

1. Bassett, G. W. & R. Koenker. An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association* 77, (1982): 407–415.
2. Billingsley, P. *Probability and Measure*. New York: Wiley, 1979.
3. Hardy, G. H., J. E. Littlewood, & G. Polya. *Inequalities*. Cambridge: Cambridge University Press, 1967.
4. Koenker, R. & G. W. Bassett. Regression Quantiles. *Econometrica* 46, (1978): 33–50.
5. Koenker, R. & G. W. Bassett. Robust Tests for Heteroscedasticity Based on Regression Quantiles. *Econometrica* 50, (1982): 43–61.
6. Koenker, R. W. & G. W. Bassett. Four (Pathological) Examples in Asymptotic Statistics. *The American Statistician* 38, (1984): 201–212.
7. Portnoy, S. Tightness of the sequence of empirical cdf processes defined from regression fractiles. In J. Franke, W. Hardle, and D. Martin, (eds.) *Robust and Non-linear Time-Series Analysis*. New York: Springer-Verlag, 1983.
8. Rockafeller, R. T. *Convex Analysis*. Princeton: Princeton University Press, (1970).
9. Weisberg, S. Comment on a Paper of McDonald and White. *Journal of the American Statistical Association* 75, (1980): 28–31.