Open access • Journal Article • DOI:10.1017/S0266466600011488

Strong consistency of regression quantiles and related empirical processes

— Source link

Gilbert W. Bassett, Roger Koenker

Institutions: University of Illinois at Chicago, University of Illinois at Urbana–Champaign

Published on: 01 Aug 1986 - Econometric Theory (Cambridge University Press)

Topics: Strong consistency, Quantile, Estimator, Conditional probability distribution and Linear model

BEBR

FACULTY WORKING
PAPER NO. 1086

Strong Consistency of Regression Quantiles and Related Empirical Processes

Gilbert W. Bassett
Roger W. Koenker

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1086

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

November, 1984

Strong Consistency of Regression Quantiles and Related Empirical Processes

Gilbert W. Bassett
University of Illinois at Chicago

Roger W. Koenker, Professor
Department of Economics

The support of the NSF is gratefully acknowledged.

Abstract

The strong consistency of regression quantile statistics (Koenker and Bassett (1978)) 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 (1982)) 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 (1978, 1982) and Bassett and Koenker (1982), 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.

We might hypothesize that the θ^{th} conditional quantile function of $Y_1, \ldots, Y_n$ is a linear function of a p-vector of exogenous variables x, i.e.,

$$Q_Y(\theta \mid x) = x\beta_\theta \tag{1.1}$$

Given this hypothesis, one may ask under what conditions can we consistently estimate the parameter vector β_θ? In Koenker and Bassett (1978) we showed that for the linear model with iid errors any sequence of solutions $\{\hat{\beta}_\theta\}$ to the problem

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

where $\rho(\cdot)$ is the check function,

$$\rho_\theta(u) = \begin{cases} \theta u & u \geq 0 \\ (\theta - 1)u & u < 0 \end{cases}$$

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 $Q(\theta) = F^{-1}(\theta)$, $\sqrt{n}\,(\hat{\beta}_\theta - \beta)$ converged in law to a k-variate Gaussian distribution with mean vector $\xi_\theta = (Q(\theta), 0, ..., 0)' \in \mathbf{R}^p$. Thus, under the foregoing conditions $\hat{\beta}_\theta$ is weakly consistent for $\beta_\theta = \beta + \xi_\theta$. In Koenker and Bassett (1982) we showed that similar asymptotic behavior prevailed in sequences of linear models with heteroscedasticity of order $0(1/\sqrt{n})$.

Here 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 (1982). Almost sure convergence results are established under mild regularity conditions on the design and the distribution function of the errors. In the case of our proposed estimate of the error distribution *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. Applications to bootstrapping and other model diagnostics immediately suggest themselves.

2. Strong Consistency of Regression Quantiles

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

$$Y_i = x_i \beta + u_i , \tag{2.1}$$

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

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

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

$$\liminf_{n \to \infty} d_n = \liminf \inf_{\|\omega\| = 1} n^{-1} \Sigma \, |x_i \omega| = d > 0 \tag{D1}$$

$$\liminf_{n \to \infty} D_n = \liminf \sup_{\|\omega\| = 1} n^{-1} \Sigma (x_i \omega)^2 = D < \infty \tag{D2}$$

We may now state:

Theorem 1. If D1 and D2 hold and F has a unique θ^{th} quantile then any sequence of solutions $\{\hat{\beta}_n(\theta)\}$ to problem (1.3) satisfies $\hat{\beta}_n(\theta) \to \beta + \xi_\theta$ almost surely.

Proof. Fix θ and consider,

$$R_n(\delta) \equiv n^{-1} \sum_{i=1}^{n} r_i(\delta) \tag{2.3}$$

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

Since $R_n(0) = 0$, and $R_n(\delta)$ is a sum of convex functions and therefore convex, it suffices to show that for any $\Delta > 0$

$$\liminf_{n \to \infty} \inf_{||\delta|| = \Delta} R_n(\delta) > 0 \quad \text{a.s.} \tag{2.4}$$

We begin by establishing that

$$R_n(\delta) \to E\, R_n(\delta) \quad \text{a.s.} \tag{2.5}$$

Since $r_i(\delta) \leq |x_i \delta|$, Kolmogorov's criterion yields

$$\sum_{i=1}^{\infty} \mathrm{Var}\,(r_i(\delta))/i^2 \leq \Sigma(x_i \delta)^2/i^2 \tag{2.6}$$

$$\leq ||\delta||^4 D^2 \Sigma(x_i \delta)^2/(\Sigma(x_i \delta)^2)^2$$

which is convergent so (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,

$$\sup_{||\delta - \delta_0|| \leq \epsilon} |R_n(\delta) - R_n(\delta_0)| \leq \sup\{n^{-1}\Sigma|x_i(\delta - \delta_0)|\} \tag{2.7}$$

$$\leq \epsilon \sup_{||\omega|| = 1} n^{-1}\Sigma|x_i \omega|$$

$$\leq \epsilon D_n^{1/2}$$

$$\leq \epsilon D^{1/2} .$$

So it remains to show that $ER_n(\delta)$ is bounded away from 0, for any

$||\delta|| > 0$. For $\alpha > 0$,

$$g(\alpha) = E[\rho_\theta(u - \alpha - Q(\theta)) - \rho_\theta(u - Q(\theta))] \qquad (2.8)$$

$$= \int_{-\infty}^{Q+\alpha} (1-\theta)\alpha dF(u) - \int_{Q+\alpha}^{\infty} \alpha\theta dF(u) - \int_{Q}^{Q+\alpha} (u - Q)dF(u)$$

and integrating by parts gives,

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

For $\alpha < 0$, the sign and limits of integration are simply reversed. The function g is convex, $g(0) = 0$, and $g(\alpha) > 0$ for any $\alpha \neq 0$ by the uniqueness of the θ^{th} quantile. Now, let $h(\alpha)$ denote the convex hull of $g(\alpha)$ and $g(-\alpha)$ then we have,

$$ER_n(\delta) \geq n^{-1} \sum_{i=1}^{n} h(x_i \delta) \qquad (2.10)$$

$$\geq n^{-1} \Sigma h(|x_i \delta|)$$

$$\geq h(n^{-1} \Sigma |x_i \delta|)$$

$$\geq h(d \, ||\delta||)$$

by the symmetry of h, Jensen's inequality and (D2) respectively. The function $h(\alpha)$ is bounded away from zero for $\alpha \neq 0$ by the uniqueness of the θ^{th} quantile, thus completing the proof.

Reviewing the preceding argument it is clear that uniqueness of the θ^{th} quantile is needed only to argue that $ER_n(\delta)$ has a unique minimum at the origin. When the θ^{th} quantile is not unique then $ER_n(\delta)$ has a larger minimum set, but it is straightforward to show that solutions to (1.3) converge almost surely to elements of this set. See Koenker and Bassett(1984) for an example of weak but not strong convergence in this framework. The following result will be used in subsequent sections.

Theorem 2.2. Suppose F is any proper, right continuous distribution function and D1-D2 prevail. Let $\Lambda_n(\Delta) = \{\delta \in \mathbf{R}^p \mid ER_n(\delta) \leq \Delta\}$, the Δ-level set of $ER_n(\delta)$. Then any sequence of solutions to (1.3), $\{\hat{\beta}_n(\theta)\}$, satisfies

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

for all $\Delta > 0$.

Proof. When $Q(\theta + 0) = Q(\theta)$ this follows immediately from Theorem 1 since trivially, $0 \in \Lambda_n(\Delta)$ for all Δ. When,

$$\zeta(\theta) = Q(\theta + 0) - Q(\theta) > 0 \tag{2.12}$$

the function $g(\alpha)$ is identically zero on $[0, \zeta(\theta)]$ and therefore $ER_n(\delta)$ will attain a minimum of zero on a *set* containing the origin. It suffices to verify that the (necessarily convex) sets $\Lambda_n(\Delta)$ are bounded since we may then, by the prior argument, finitely cover the boundary of any such set and by the uniform convergence of $R_n(\delta)$ to $ER_n(\delta)$ on compacts we may conclude that any solution to (1.3) lies inside this boundary.

For any $\lambda \in \Lambda_n(\Delta)$ we have $h(d\,||\lambda||) \leq \Delta$ by the argument of (2.10). Since $h(\cdot)$ is convex and zero only on the bounded interval $[-\zeta(\theta), \zeta(\theta)]$, for any $\Delta > 0$, there exists an $M < \infty$ such that $h(\alpha) \leq \Delta$ implies $|\alpha| < M$. Thus $||\lambda|| < M/d$ for any $\lambda \in \Lambda_n(\Delta)$, and this completes the proof.

3. Strong Convergence of Empirical Processes Based on Regression Quantiles

In Bassett and Koenker (1982) we proposed an estimate of the conditional quantile function of Y given x,

$$\hat{Q}_Y(\theta \mid x) = \inf\{xb \mid b \in \hat{B}(\theta)\} \tag{3.1}$$

where as above

$$\hat{B}(\theta) = \{b \in \mathbf{R}^p \mid \Sigma \rho_\theta(y_i - x_i b) = \min!\}. \tag{3.2}$$

There it is shown that at $x = \bar{x} = n^{-1} \Sigma x_i$, the sample paths of the random function

$$\hat{Q}_Y(\theta) \equiv \hat{Q}_Y(\theta \mid \bar{x}) \tag{3.3}$$

are non-decreasing, left continuous jump functions on (0, 1). However, unlike the ordinary empirical quantile function to which $\hat{Q}_Y(\theta)$ specializes when $X_n = 1_n$; $\hat{Q}_Y(\theta)$, jumps at irregularly spaced points on (0, 1). Similarly, one may show that

$$\hat{Q}_Y(\theta + 0) \equiv \lim_{\epsilon \to 0} \hat{Q}_Y(\theta + \epsilon) \tag{3.4}$$

$$= \sup \{\bar{x}b \mid b \in \hat{B}(\theta)\}$$

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. Portnoy (1983) has recently strengthened these results to establish weak convergence of the estimate

$$\hat{F}_Y(y) = \sup\{\theta \mid \hat{Q}_Y(\theta) \le y\} \tag{3.5}$$

to the Brownian bridge.

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

$$Q_Y(\theta \mid x) = x\beta + Q(\theta) \tag{3.6}$$

We will restrict attention as previously to

$$Q_Y(\theta) = Q_Y(\theta \mid \bar{x}) = \beta\bar{x} + Q(\theta). \tag{3.7}$$

We may now state:

Theorem 3.1. If D1 and D2 hold and $Q(\theta)$ is continuous on a closed interval $\Theta \subset (0, 1)$ then

$$\sup_{\theta \in \Theta} \left| \hat{Q}_Y(\theta) - Q_Y(\theta) \right| \to 0 \quad \text{a.s.} \tag{3.8}$$

Proof. From Theorem 2.1 we have pointwise convergence of $\hat{Q}_Y(\theta)$ and $\hat{Q}_Y(\theta + 0)$ to $Q_Y(\theta)$ and using the monotonicity and continuity of $Q_Y(\theta)$ this may be strengthened to uniform almost sure convergence on Θ. See, for example, the argument in Billingsley (1979, p. 233) for the Glivenko-Cantelli theorem.

When there are jumps in $Q_Y(\cdot)$ some new difficulties arise, since $\hat{Q}_Y(\theta)$ may oscillate between $Q_Y(\theta)$ and $Q_Y(\theta + 0)$, but based on Theorem 2.2 we may establish the following result.

Theorem 3.2. If D1 and D2 hold then

$$\sup_{u \in \mathbf{R}} \left| \hat{F}_Y(u) - F_Y(u) \right| \to 0 \quad \text{a.s.} \tag{3.9}$$

Proof. We will consider the case of continuous F first; discontinuities in F are treated in the Appendix. We will begin by fixing θ and establishing the inequalities:

$$\liminf_{n \to \infty} \hat{Q}_Y(\theta) \geq Q_Y(\theta) \quad \text{a.s.} \tag{3.10}$$

$$\limsup_{n \to \infty} \hat{Q}_Y(\theta + 0) \leq Q_Y(\theta + 0) \quad \text{a.s.} \tag{3.11}$$

From Theorem 2.2 we have for any

$$\hat{Q}_Y(\theta) \geq Q_Y(\theta) + \inf \{\lambda \bar{x} \mid \lambda \in \Lambda_n(\Delta)\} \tag{3.12}$$

and

$$\hat{Q}_Y(\theta + 0) \leq Q_Y(\theta) + \sup \{\lambda \bar{x} \mid \lambda \in \Lambda_n(\Delta)\} \tag{3.13}$$

so, by Jensen's inequality, using D1,

$$\{\lambda \bar{x} \mid \lambda \in \Lambda_n(\Delta)\} = \{\lambda \bar{x} \mid n^{-1} \sum_{i=1}^{n} g(x_i \lambda) < \Delta\} \tag{3.14}$$

$$\subset \{\lambda \bar{x} \mid g(\lambda \bar{x}) < \Delta\} .$$

Since $g(\cdot)$ is convex and zero only on $[0, Q(\theta + 0) - Q(\theta)]$ we have

$$-o(\Delta) \leq \inf \{\alpha \mid g(\alpha) < \Delta\} \tag{3.15}$$

$$\leq \sup \{\alpha \mid g(\alpha) < \Delta\}$$

$$\leq Q(\theta + 0) - Q(\theta) + o(\Delta)$$

so the inequalities follow letting $\Delta \to 0$. It is straightforward to verify that (3.10-11) imply

$$\limsup_{n \to \infty} \hat{F}_Y(Q_Y(\theta)) \leq F_Y(Q_Y(\theta)) \tag{3.16}$$

and

$$\liminf_{n \to \infty} \hat{F}_Y(Q_Y(\theta) - 0) \geq F(Q_Y(\theta) - 0) . \tag{3.17}$$

So by continuity of F we have

$$\liminf_{n \to \infty} \hat{F}_Y(Q_Y(\theta) - 0) = \limsup_{n \to \infty} \hat{F}_Y(Q_Y(\theta)) = \theta . \tag{3.18}$$

And by standard arguments for the Glivenko-Cantelli theorem, again see, Billingsley (1979, p. 233), this implies (3.9).

Appendix

When F has discontinuities some new difficulties arise which are taken up in this appendix. Principally, we must verify that $\hat{F}$ converges to a proper distribution function. See the remark following the the proof of the lemma. To complete the proof of Theorem 3.2 we will need:

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

$$P\{\hat{B}_n(\theta) = \beta + \xi_\theta, \quad n > n_o\} \geq 1 - \lambda. \tag{A.1}$$

Proof. As in (2.3) consider,

$$R_n(\delta) = n^{-1} \sum r_i(\delta) = n^{-1} \sum \rho_\theta(u_i - x_i \delta - Q(\theta)) - \rho_\theta(u_i - Q(\theta)) \tag{A.2}$$

and corresponding directional derivitives,

$$R_n'(\delta, w) = n^{-1} \sum r_i'(\delta, w) = n^{-1} \sum [\tfrac{1}{2} - \theta - \tfrac{1}{2} sgn^*(u_i, x_i w)] x_i w \tag{A.3}$$

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

$$\liminf_{n \to \infty} \inf_{||w||=1} R'_n(0, w) > 0 \quad a.s. \tag{A.4}$$

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

$$R'_n(0,w) - ER'_n(0,w) \to 0 \quad a.s. \tag{A.5}$$

Since $|r'_i(0,w)| \leq |x_i w|$ application of Kolmogorov's criterion yields,

$$\sum_{i=1}^{\infty} Var(r'_i(0,w))/i^2 \leq \sum_{i=1}^{\infty} (x_i w)^2/i^2 < \infty \tag{A.6}$$

by (D2). This may be strengthened to uniform convergence on the sphere $||w|| = 1$ using (D2) and the continuity of $R'_n(0,w)$ in w. Hence, we have,

$$\liminf_{n \to \infty} \inf_{||w||=1} R_n'(0,w) \to \liminf_{n \to \infty} \inf_{||w||=1} ER_n'(0,w) \quad a.s. \tag{A.7}$$

Now,

$$Er_i'(0,w) = \begin{cases} [F(Q(\theta))-\theta]x_i w & \text{if } x_i w \geq 0 \\ [F(Q(\theta)-0)-\theta]x_i w & \text{otherwise} \end{cases} \tag{A.8}$$

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

$$Er_i'(0,w) \geq m(\theta)|x_i w|, \tag{A.9}$$

hence, using (D1),

$$\liminf_{n \to \infty} \inf_{||w||=1} ER_n'(0,w) \geq m(\theta)d > 0, \tag{A.10}$$

which completes the proof.

An immediate consequence of the lemma is that under the same conditions,

$$P\{Q_Y(\theta) = \hat{Q}_Y(\theta) = \hat{Q}_Y(\theta+0) = Q_Y(\theta+0),\ n > n_o\} > 1 - \lambda \tag{A.11}$$

which implies,

$$\limsup_{n \to \infty} \hat{F}_Y(\hat{Q}_Y(\theta)) \geq F_Y(Q_Y(\theta)) \tag{A.12}$$

$$\liminf_{n \to \infty} \hat{F}_Y(\hat{Q}_Y(\theta-0)) \leq F_Y(Q_Y(\theta-0)) \tag{A.13}$$

Which together with (3.16-17) imply that at points of discontinuity in F we have the required convergence. Since there are only countably many such points this completes the proof.

Remark. The following example illustrates the necessity of the lemma. Let $Q(\theta) = 0$ for $\theta \in (0,1]$ so that the associated df is

$$F(u) = \begin{cases} 0 & u < 0 \\ 1 & u \geq 0 \end{cases}. \tag{A.14}$$

Consider,

$$\hat{Q}(\theta) = \begin{cases} -n^{-1} & \theta \in (0,½] \\ n^{-1} & \theta \in (½,1] \end{cases} \tag{A.15}$$

and

$$\hat{F}(u) = \begin{cases} 0 & u < -n^{-1} \\ ½ & -n^{-1} \leq u \leq n^{-1} \\ 1 & n^{-1} \leq u \end{cases}. \tag{A.16}$$

Then $\hat{Q}(\cdot)$ satisfies (3.10-11) but not (A.11). $\hat{Q}$ actually converges to Q uniformly, but

$$\lim_{n \to \infty} \hat{F}(Q(½)) = \hat{F}(0) = ½ \neq 1 = F(0) \tag{A.17}$$

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

References

Bassett, G. W. and Koenker R. (1982). An empirical quantile function for linear models with iid errors, *Journal of the American Statistical Association*, 407-415.

Billingsley, P. (1979) *Probability and Measure*, (New York: Wiley).

Koenker, R. and Bassett, G. W. (1978) Regression Quantiles, *Econometrica*, 46, 33-50.

Koenker, R. and Bassett, G. W. (1982) Robust Tests for Heteroscedasticity based on Regression Quantiles, *Econometrica*, 50, 43-61.

Koenker, R.W. and Bassett, G.W. (1984) Four (Pathological) Examples in Asymptotic Statistics, *The American Statistician*, 38, 201-212.

Portnoy, S. (1983) Tightness of the sequence of empiric cdf processes defined from regression fractiles, Proceedings of the Conference on Robust and Non-linear Methods in Time-Series Analysis, Heidelburg, FRG.