Lecture 2(c). Introduction to Quantile Regression Models and Methods, Continued

Victor Chernozhukov

May 2009.

Outline

1. Monotonization of Quantile Curves

2. Outline of Asymptotics*

1. Monotonization of Quantile Curves

Crossing/ Non-Monotonicity Problem for Quantile Regression

- The material of this part of the lecture is based on the paper "Quantile and Probability Curves without Crossing" by Chernozhukov, Fernandez-Val, and Galichon, 2010, Econometrica.

- A related paper is "Improving Point and Interval Estimates of Monotone Functions by Rearrangement", by Chernozhukov, Fernandez-Val, and Galichon, 2010, Biometrika.

- estimated quantile curve

$$u \mapsto x'\widehat{\beta}(u)$$

need not be monotone in the probability index u

- curves are fitted pointwise

$$\widehat{\beta}(u) = \arg\min_{\beta} \sum \rho_u(Y_i - X_i'\beta)$$

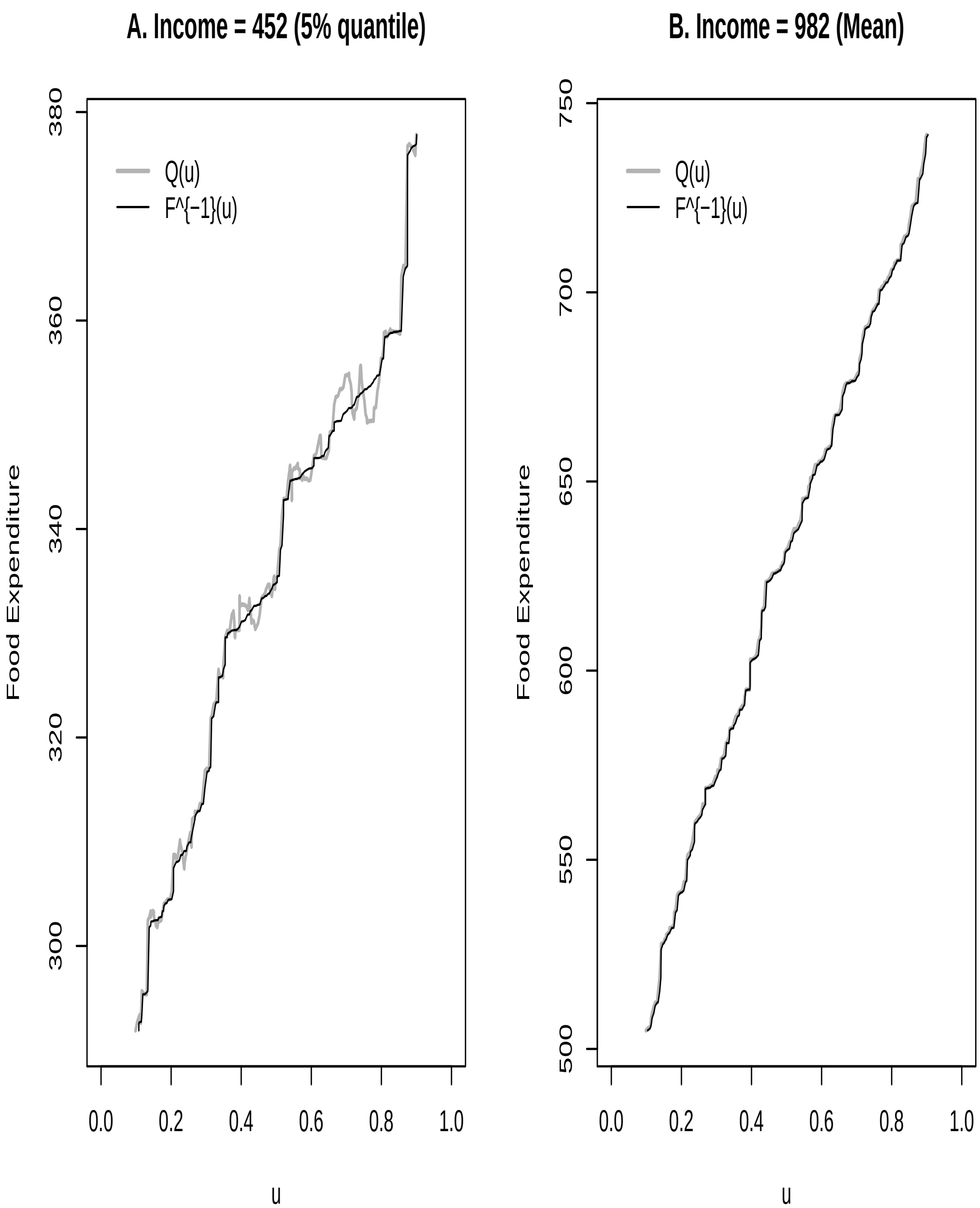
where

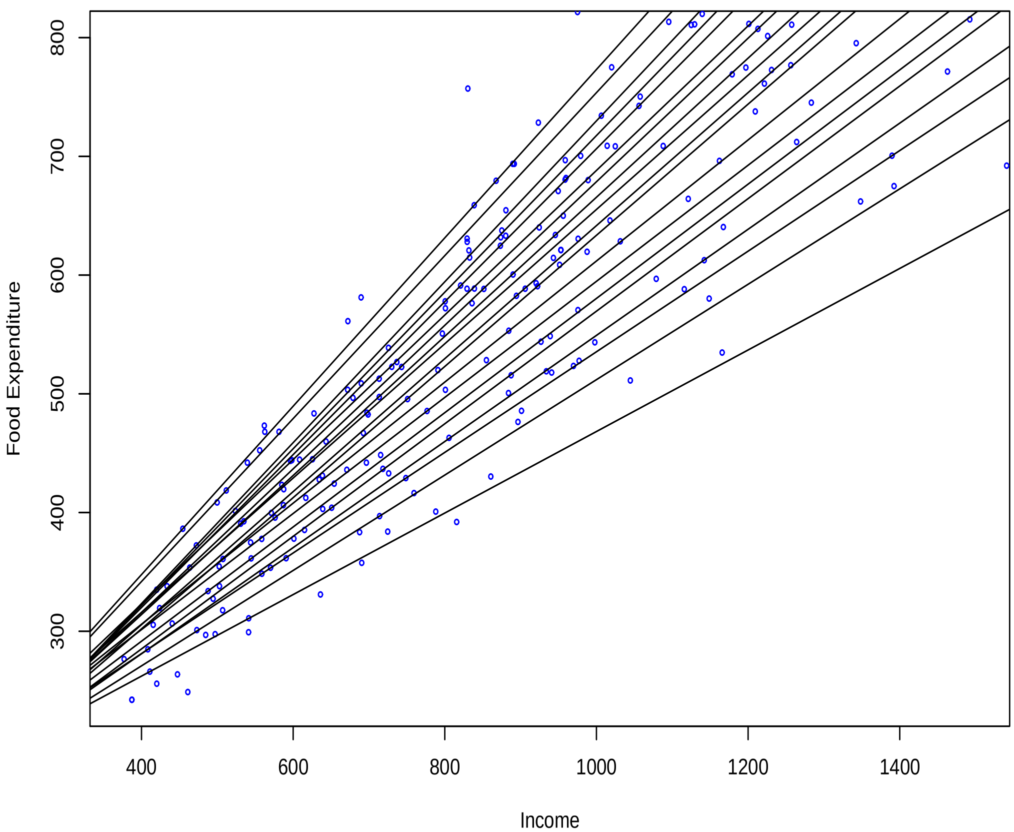
$$\rho_u(e) = ue^{+} + (1-u)e^{-}$$

without imposing monotonicity restrictions

- Example: Engel curves

Example: Engel Curves





Food Expenditure
300
400
500
600
700
800
400
600
800
1000
1200
1400
Income

Two reasons cause crossing:

- estimation error

- approximation error

We focus the discussion on the first (see CFG, 2010, for discussion of the second.)

Rearrangement

- define

$$Y_x := x'\widehat{\beta}(U) \text{ where } U \sim U(0,1)$$

- induces proper cdf

$$\widehat{F}(y|x) = \int_0^1 1\{x'\widehat{\beta}(u) \leq y\}du$$

- now go one step further and invert the cdf to get

$$\widehat{F}^{-1}(u|x) = \inf\{y : \widehat{F}(y|x) \geq u\}$$

- this delivers a proper (monotone) quantile function

Rearrangement

- this operation is known in variational analysis and in operations research as the rearrangement operator (Hardy, Polya, Littlewood, 1940s)

- CFG 2010 establish the statistical properties of the curves

$$u \mapsto \widehat{F}^{-1}(u|x) \quad \text{and} \quad u \mapsto \widehat{F}(u|x)$$

Large Sample Statistical Properties

- State the main results in terms of the canonical linear setup

- Assume that

$$\sqrt{n}x'(\widehat{\beta}(u) - \beta(u)) \Rightarrow x'G(u) \qquad (1)$$

in the metric space of bounded functions $\ell^{\infty}(0,1) \times \mathcal{X}$, as discussed earlier. Here $\mathcal{X}$ is a compact region.

- this assumption is general – holds under general sampling conditions and does not depend on the nature of the underlying model (e.g. ordinary or endogenous). Also can exclude the tail quantiles, by considering interval $[\epsilon, 1-\epsilon]$ instead of $(0, 1)$.

Theorem. Under the stated and other mild assumptions, CFG (2007) show that

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'[G(u)]$$

in $\ell^\infty((0,1)\times\mathcal{X})$, which coincides with the asymptotics of the original curve $u \mapsto x'\beta(u)$.

This has a convenient practical implication:

The empirical non-monotone curve can be rearranged to be monotonic *without* affecting its (first order) asymptotic properties.

This result applies when the population curves are monotone, in particular, under correct specification.

Finite Sample Estimation Property

Theorem: The rearranged quantile function is *closer* to the true quantile function $Q_0(u,x)$, for all $p \in [1,\infty]$:

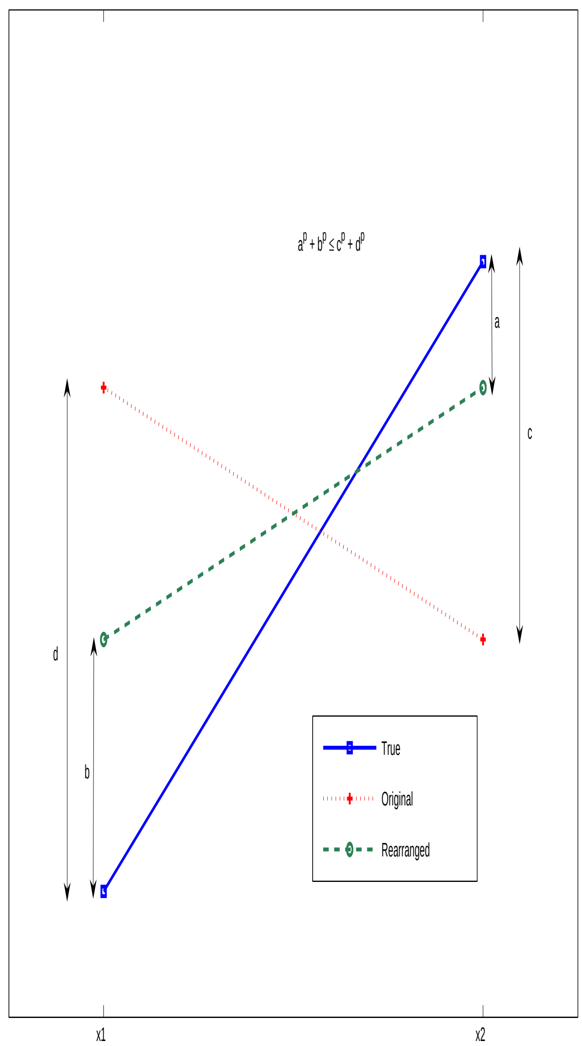
$$\begin{aligned}&\left(\int_{\mathcal{U}} |Q_0(u,x) - \widehat{F}^{-1}(u|x)|^p du\right)^{1/p} \\ &\leq \left(\int_{\mathcal{U}} |Q_0(u,x) - x'\widehat{\beta}(u)|^p du\right)^{1/p}\end{aligned} \tag{2}$$

where inequality is strict for $p \in (1,\infty)$ whenever $u \mapsto x'\widehat{\beta}(u)$ is decreasing on a subset of $\mathcal{U}$ of positive Lebesgue measure and the true quantile function $u \mapsto Q_0(u,x)$ is increasing.

- Rearranged quantile curves have a smaller estimation error than the original curves whenever the latter are not monotone.

- This property is generic and is independent of the sample size.

Proof is based on:



2. Asymptotics for Quantile Regression Processes

I outline the main ideas only.

Sample Quantiles. Asymptotic theory for quantile regression to be presented can be understood by looking at the sample quantiles as the example. Let T be a compact subset of $(0, 1)$ and Y have a distribution function F with a continuous density f. The u-quantiles $Q(u)$ solve the equation

$$F(Q(u)) = u.$$

We assume that $f(Q(u)) > 0$ on $u \in T$. Let $\widehat{F}$ be the empirical distribution function and $\widehat{Q}$ be the empirical quantile function, which satisfies

$$|\widehat{F}(\widehat{Q}(u)) - u| \leq 1/n.$$

Thus, we have that

$$\sqrt{n}[\widehat{F}[\widehat{Q}(u)] - u] = O_p(1/\sqrt{n}).$$

which is the same as

$$\underbrace{\sqrt{n}[\widehat{F}[\widehat{Q}(u)] - F[\widehat{Q}(u)]]}_{a(u)} + \underbrace{\sqrt{n}[F[\widehat{Q}(u)] - F[Q(u)]]}_{b(u)}$$
$$= O_p(1/\sqrt{n}). \tag{3}$$

Recall that we have that

$$\mathbb{G}_n = \sqrt{n}(\widehat{F} - F) \to_d \mathbb{G}_F,$$

which implies $\|a\|_T = O_p(1)$, so that $\|b\|_T = O_p(1)$, which can be shown to imply that

$$\|\widehat{Q} - Q\|_T \to_p 0.$$

The latter fact and the stochastic equicontinuity of $\mathbb{G}_n$ in turn imply that uniformly on $u \in T$

$$a(u) = \underbrace{\sqrt{n}[\widehat{F}[Q(u)] - F[Q(u)]]}_{c(u)} + o_p(1).$$

Indeed, by the stochastic equicontinuity of $\mathbb{G}_n$, which is a part of the Donsker theorem, uniformly on $u \in T$

$$|a(u) - c(u)| \leq \sup_{|t-t'|\leq\delta_n} |\mathbb{G}_n(t) - \mathbb{G}_n(t')| \to_p 0,$$

for any $\delta_n \searrow 0$.

By Taylor expansion and consistency, uniformly on $u \in T$,

$$b(u) = (f(Q(u)) + o_p(1))\sqrt{n}(\hat{Q}(u) - Q(u)).$$

Therefore, uniformly on T,

$$\sqrt{n}(\hat{Q}(u)-Q(u)) = (f(Q(u))^{-1}+o_p(1))c(u)+o_p(1).$$

Note that

$$c(\cdot) \to_d \mathbb{G}_F \circ Q = \mathbb{B},$$

where $\mathbb{B}$ is the standard Brownian bridge (i.e. has covariance function of the form $min(u, u') - u \cdot u'$). Conclude that

$$\sqrt{n}(\hat{Q}(\cdot) - Q(\cdot)) \to_d f(Q(\cdot))^{-1}\mathbb{B}(\cdot).$$

Remark.*[Stochastic equicontinuity of $\mathbb{G}_n$]. Recall that process $\mathbb{G}_n$ is stochastically equicontinuous with respect to $L_2(P)$ metric, $\rho(t,t') = [[F(t) - F(t')]^2]^{1/2}$. Indeed,

$$\sup_{\rho(t,t')\leq\delta} |\mathbb{G}_n(t) - \mathbb{G}_n(t')| \to_d \sup_{\rho(t,t')\leq\delta} |\mathbb{G}(t) - \mathbb{G}(t')|,$$

so that tail probabilities of the process on the left are asymptotically controlled by the tail probabilities of the stochastically equicontintuous process on the right. In particular, since the right side converges to zero as $\delta \searrow 0$, we can conclude that

$$\sup_{\rho(t,t')\leq o(1)} |\mathbb{G}_n(t) - \mathbb{G}_n(t')| \to_p 0.$$

Finally the metric ρ is continuous with respect to the metric $|t-t'|$, yielding the result, so that

$$\sup_{|t-t'|\leq o(1)} |\mathbb{G}_n(t) - \mathbb{G}_n(t')| \to_p 0.$$

To summarize, in the non-regression quantile case, the two key ingredients are:

1. a functional central limit theorem for the empirical distribution function,

$$\mathbb{G}_n = \sqrt{n}(\widehat{F} - F) \to_d \mathbb{G}_F,$$

where $\mathbb{G}_F$ is a F-Brownian bridge with continuous paths a.s. For this all we need to know is that the class of cells $\{1\{Y \leq y\}, y \in \mathbb{R}\}$ is P-Donsker and for continuity of paths we need F to be continuous.

2. continuity of $f = F'$ and its uniform boundedness away from zero at u-quantiles $Q(u)$ of interest:

$$f(Q(u)) > 0$$

uniformly on $u \in T = [\epsilon, 1 - \epsilon]$.

Discussion.* In the quantile regression case, we similarly need

(a) P-Donkerness of the function class

$$\mathcal{G} = \{g(Z, \beta, u) = (u - 1\{Y \leq X'\beta\})X_j, \\ \beta \in R^k, u \in T, j = 1, ..., d\} \quad (4)$$

This can be demonstrated using bracketing arguments, provided $E\|X\|^2 < \infty$. P-Donskerness gives us a functional central limit theorem for

$$\mathbb{G}_n(\beta, u) := n^{-1/2} \sum (g(Z, \beta, u) - Eg(Z, \beta, u)),$$

$$\mathbb{G}_n(\beta, u) \to_d \mathbb{G}(\beta, u) \text{ in } \ell^\infty(\mathbb{R}^d \times T)$$

where $\mathbb{G}$ is a Gaussian process with continuous paths. For continuity of paths we need only that $g(Z, \beta, u)$ has a square integrable envelope and is pointwise continuous in (β, u) a.s.

(b) Continuity of

$$\frac{\partial}{\partial \beta} Eg(Z, \beta, u)$$

as well as uniform invertibility and continuity of

$$J(u) = \frac{\partial}{\partial \beta} Eg(Z, \beta, u)|_{\beta=\beta(u)}$$

over $u \in T$

Then, arguing similarly to the sample quantile case, one can conclude that

$$\sqrt{n}(\widehat{\beta}(\cdot) - \beta(\cdot)) \to_d J(\cdot)^{-1}\mathbb{G}(\beta(\cdot), \cdot) \text{ in } \ell^{\infty}(T)^k.$$