

## Quantile and Expectile Regression Models

A. Charpentier (Université de Rennes 1)

with A.D. Barry & K. Oualkacha (UQàM)

ESC Rennes, December 2016.

<http://freakonometrics.hypotheses.org>



## Mediane/Quantiles ( $\ell_1$ -norm)

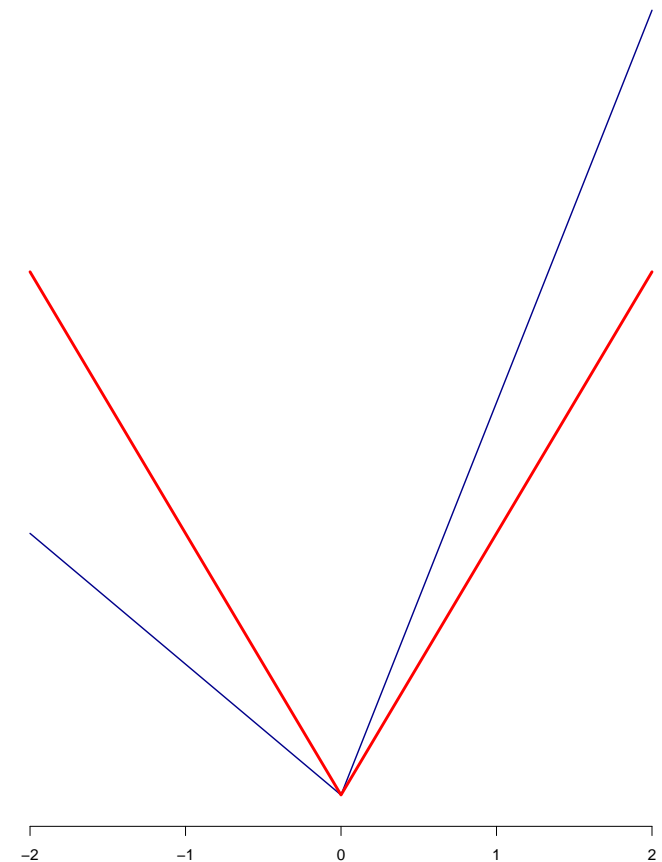
Empirical median  $\hat{m}(\mathbf{y})$  is solution of

$$\hat{m}(\mathbf{y}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n \underbrace{\frac{1}{2} |y_i - \theta|}_{=r_{1/2}^Q(y_i - \theta)} \right\}.$$

Empirical quantile  $\hat{q}(\alpha, \mathbf{y})$  is solution of

$$\hat{q}(\alpha, \mathbf{y}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\alpha}^Q(y_i - \theta) \right\},$$

with  $r_{\alpha}^Q(u) = |\alpha - \mathbf{1}(u \leq 0)| \cdot |u|$ .



## Quantiles

Consider  $Y \sim F$ , and a level  $\alpha \in (0, 1)$ , then  $q(\alpha, Y) = \inf\{y; F_Y(y) \geq \alpha\}$ .

Equivalently

$$q(\alpha, Y) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \mathbb{E} \left[ r_{\alpha}^Q(Y - \theta) \right] \right\}, \text{ with } r_{\alpha}^Q(u) = |\alpha - \mathbf{1}(u \leq 0)| \cdot |u|$$

The empirical version, with a sample  $\mathbf{y} = \{y_1, \dots, y_n\}$ , is

$$\hat{q}(\alpha, \mathbf{y}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\alpha}^Q(y_i - \theta) \right\}.$$

The conditional  $\alpha$ -quantile of  $Y|\mathbf{x}$  is  $q(\alpha, Y, \mathbf{x}) = \inf\{y; F_{Y|\mathbf{x}}(y) \geq \alpha\}$ . Assuming that  $F_{Y|\mathbf{x}}^{-1}(\alpha) = \mathbf{x}_i^{\top} \boldsymbol{\beta}^Q(\alpha)$ , quantile regression parameters are obtained from sample  $(\mathbf{y}, \mathbf{X}) = \{(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)\}$  as

$$\hat{\boldsymbol{\beta}}^Q(\alpha, \mathbf{y}, \mathbf{X}) = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^p} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\alpha}^Q(y_i - \mathbf{x}_i^{\top} \boldsymbol{\beta}^Q(\alpha)) \right\}.$$

## Expected value/Expectiles ( $\ell_2$ -norm)

Empirical mean  $\bar{y}$  is solution of

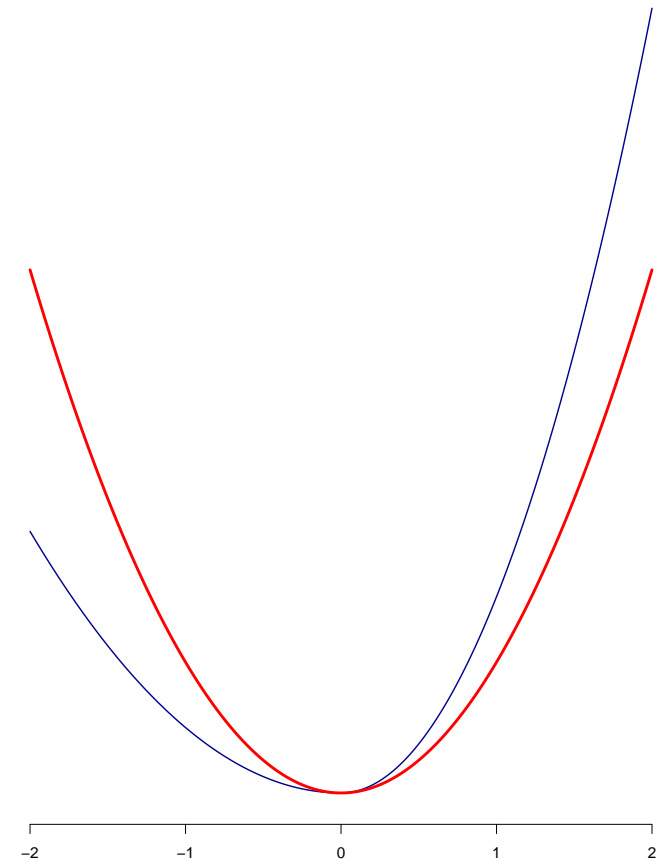
$$\bar{y} = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n \underbrace{\frac{1}{2} [y_i - \theta]^2}_{=r_{1/2}^E(y_i - \theta)} \right\}.$$

Empirical expectile  $\hat{\mu}(\tau, \mathbf{y})$  is solution of

$$\hat{\mu}(\tau, \mathbf{y}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\tau}^E(y_i - \theta) \right\},$$

with  $r_{\tau}^E(u) = |\tau - \mathbf{1}(u \leq 0)| \cdot u^2$ .

See



## Expectiles

Consider  $Y \sim F$ , and a level  $\tau \in (0, 1)$ ,

$$\mu(\tau, Y) = \operatorname{argmin}_{\theta \in \mathbb{R}} \mathbb{E}\{r_{\tau}^E(Y - \theta)\} \text{ with } r_{\tau}^E(u) = |\tau - \mathbf{1}(u \leq 0)| \cdot u^2.$$

The empirical version, with a sample  $\mathbf{y} = \{y_1, \dots, y_n\}$  is

$$\hat{\mu}(\tau, \mathbf{y}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\tau}^E(y_i - \theta) \right\}.$$

The conditional  $\tau$ -expectile of  $Y|\mathbf{x}$  is

$$\mu(\tau, Y, \mathbf{x}) = \operatorname{argmin}_{\theta \in \mathbb{R}} \mathbb{E}\{r_{\tau}^E(Y - \theta)|\mathbf{x}\},$$

and assuming that  $\mu(\tau, \mathbf{x}) = \mathbf{x} \top \boldsymbol{\beta}^E(\tau)$ , parameters of the expectile regression are

$$\hat{\boldsymbol{\beta}}^E(\tau, \mathbf{y}, \mathbf{X}) = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^p} \left\{ \frac{1}{n} \sum_{i=1}^n r_{\tau}^E(y_i - \mathbf{x}_i \top \boldsymbol{\beta}^E(\tau)) \right\}.$$

## Quantiles and Expectiles

Observe that  $q(\alpha, Y)$  is solution of

$$\alpha = F(q(\alpha, Y)) = \mathbb{E}[\mathbf{1}(Y < q(\alpha, Y))]$$

while  $\mu(\tau, Y)$  is solution of

$$\tau = \frac{\mathbb{E}[|Y - \mu(\tau, Y)| \cdot \mathbf{1}\{Y < \mu(\tau, Y)\}]}{\mathbb{E}[|Y - \mu(\tau, Y)|]}$$

## Quantile Regression with Fixed Effects (QRFE)

In a panel linear regression model,  $y_{i,t} = \mathbf{x}_{i,t}^\top \boldsymbol{\beta} + u_i + \varepsilon_{i,t}$ , where  $u$  is an unobserved individual specific effect.

In a fixed effects models,  $u$  is treated as a parameter. Quantile Regression is

$$\min_{\boldsymbol{\beta}, u} \left\{ \sum_{i,t} r_{\alpha}^Q(y_{i,t} - [\mathbf{x}_{i,t}^\top \boldsymbol{\beta} + u_i]) \right\}$$

Consider Penalized QRFE, as in [Koenker & Biliias \(2001\)](#),

$$\min_{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k, u} \left\{ \sum_{k,i,t} \omega_k r_{\alpha_k}^Q(y_{i,t} - [\mathbf{x}_{i,t}^\top \boldsymbol{\beta}_k + u_i]) + \lambda \sum_i |u_i| \right\}$$

where  $\omega_k$  is a relative weight associated with quantile of level  $\alpha_k$ .

## Quantile Regression with Random Effects (QRRE)

Assume here that  $y_{i,t} = \mathbf{x}_{i,t}^\top \boldsymbol{\beta} + \underbrace{u_i + \varepsilon_{i,t}}_{=\eta_{i,t}}$ .

Quantile Regression Random Effect (QRRE) yields solving

$$\min_{\boldsymbol{\beta}} \left\{ \sum_{i,t} r_{\alpha}^Q(y_{i,t} - \mathbf{x}_{i,t}^\top \boldsymbol{\beta}) \right\}$$

which is a weighted asymmetric least square deviation estimator.

Let  $\Sigma = [\sigma_{s,t}(\alpha)]$  denote the matrix

$$\sigma_{ts}(\alpha) = \begin{cases} \alpha(1 - \alpha) & \text{if } t = s \\ \mathbb{E}[\mathbf{1}\{\varepsilon_{it}(\alpha) < 0, \varepsilon_{is}(\alpha) < 0\}] - \alpha^2 & \text{if } t \neq s \end{cases}$$

If  $(nT)^{-1} \mathbf{X}^\top \{\mathbb{I}_n \otimes \Sigma_{T \times T}(\alpha)\} \mathbf{X} \rightarrow \mathbf{D}_0$  as  $n \rightarrow \infty$  and  $(nT)^{-1} \mathbf{X}^\top \Omega_f \mathbf{X} = \mathbf{D}_1$ , then

$$\sqrt{nT} \left( \hat{\boldsymbol{\beta}}^Q(\alpha) - \boldsymbol{\beta}^Q(\alpha) \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbf{D}_1^{-1} \mathbf{D}_0 \mathbf{D}_1^{-1}\right).$$



## Expectile Regression with Random Effects (ERRE)

Quantile Regression Random Effect (QRRE) yields solving

$$\min_{\beta} \left\{ \sum_{i,t} r_{\alpha}^E(y_{i,t} - \mathbf{x}_{i,t}^{\top} \beta) \right\}$$

One can prove that

$$\hat{\beta}^E(\tau) = \left( \sum_{i=1}^n \sum_{t=1}^T \hat{\omega}_{i,t}(\tau) \mathbf{x}_{it} \mathbf{x}_{it}^{\top} \right)^{-1} \left( \sum_{i=1}^n \sum_{t=1}^T \hat{\omega}_{i,t}(\tau) \mathbf{x}_{it} y_{it} \right),$$

where  $\hat{\omega}_{it}(\tau) = |\tau - \mathbf{1}(y_{it} < \mathbf{x}_{it}^{\top} \hat{\beta}^E(\tau))|$ .

## Expectile Regression with Random Effects (ERRE)

If  $W = \text{diag}(\omega_{11}(\tau), \dots, \omega_{nT}(\tau))$ , set

$$\bar{W} = \mathbb{E}(W), H = \mathbf{X}^T \bar{W} \mathbf{X} \text{ and } \Sigma = \mathbf{X}^T \mathbb{E}(W \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T W) \mathbf{X}.$$

and then

$$\sqrt{nT} \{ \hat{\boldsymbol{\beta}}^E(\tau) - \boldsymbol{\beta}^E(\tau) \} \xrightarrow{\mathcal{L}} \mathcal{N}(0, H^{-1} \Sigma H^{-1}).$$

# Application to Real Data

