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

$$\hat{m}(y) = \arg\min_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} |y_i - \theta| \right\}.$$  

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

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

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

@freakonometrics
Quantiles

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

$$q(\alpha, Y) = \arg\min_{\theta \in \mathbb{R}} \left\{ \mathbb{E} \bigcap \alpha^Q (Y - \theta) \right\}, \text{ with } r^Q_\alpha(u) = |\alpha - 1(u \leq 0)| \cdot |u|$$

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

$$\hat{q}(\alpha, y) = \arg\min_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^{n} r^Q_\alpha(y_i - \theta) \right\}.$$

The conditional $\alpha$-quantile of $Y|x$ is $q(\alpha, Y, x) = \inf \{ y; F_{Y|x}(y) \geq \alpha \}$. Assuming that $F_{Y|x}^{-1}(\alpha) = x_i^T \beta^Q(\alpha)$, quantile regression parameters are obtained from sample $(y, X) = \{(y_1, x_1), \ldots, (y_n, x_n)\}$ as

$$\hat{\beta}^Q(\alpha, y, X) = \arg\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{n} \sum_{i=1}^{n} r^Q_\alpha(y_i - x_i^T \beta^Q(\alpha)) \right\}.$$
Expected value/Expectiles ($\ell_2$-norm)

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

$$\bar{y} = \arg\min_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2}[y_i - \theta]^2 \right\}.$$

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

$$\hat{\mu}(\tau, y) = \arg\min_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^{n} r^E_{\tau}(y_i - \theta) \right\},$$

with $r^E_{\tau}(u) = |\tau - 1(u \leq 0)| \cdot u^2$.

See
**Expectiles**

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

$$\mu(\tau, Y) = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}\{r^E_\tau(Y - \theta)\} \text{ with } r^E_\tau(u) = |\tau - 1(u \leq 0)| \cdot u^2.$$

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

$$\hat{\mu}(\tau, y) = \arg\min_{\theta \in \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^{n} r^E_\tau(y_i - \theta) \right\}.$$

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

$$\mu(\tau, Y, x) = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}\{r^E_\tau(Y - \theta) | x\},$$

and assuming that $\mu(\tau, x) = x^T \beta^E(\tau)$, parameters of the expectile regression are

$$\hat{\beta}^E(\tau, y, X) = \arg\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{n} \sum_{i=1}^{n} r^E_\tau(y_i - x_i^T \beta^E(\tau)) \right\}.$$
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, \( y_{i,t} = x_{i,t}^T \beta + u_i + \varepsilon_{i,t} \),
where \( u \) is an unobserved individual specific effect.

In a fixed effects models, \( u \) is treated as a parameter. Quantile Regression is

\[
\min_{\beta,u} \left\{ \sum_{i,t} r^Q_{\alpha}(y_{i,t} - [x_{i,t}^T \beta + u_i]) \right\}
\]

Consider Penalized QRFE, as in Koenker & Bilias (2001),

\[
\min_{\beta_1, \ldots, \beta_\kappa, u} \left\{ \sum_{k,i,t} \omega_k r^Q_{\alpha_k}(y_{i,t} - [x_{i,t}^T \beta_k + u_i]) + \lambda \sum_{i} |u_i| \right\}
\]

where \( \omega_k \) is a relative weight associated with quantile of level \( \alpha_k \).
Quantile Regression with Random Effects (QRRE)

Assume here that \( y_{i,t} = x_{i,t}^T \beta + u_i + \varepsilon_{i,t} \).

Quantile Regression Random Effect (QRRE) yields solving

\[
\min_\beta \left\{ \sum_{i,t} r_Q^{\alpha}(y_{i,t} - x_{i,t}^T \beta) \right\}
\]

which is a weighted asymmetric least square deviation estimator.

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

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

If \((nT)^{-1} X^T \{I_n \otimes \Sigma_{T \times T}(\alpha)\} X \to D_0\) as \(n \to \infty\) and \((nT)^{-1} X^T \Omega_f X = D_1\), then

\[
\sqrt{nT} \left( \hat{\beta}^Q(\alpha) - \beta^Q(\alpha) \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left( 0, D_1^{-1} D_0 D_1^{-1} \right).
\]
**Expectile Regression with Random Effects (ERRE)**

Quantile Regression Random Effect (QRRE) yields solving

$$\min_{\beta} \left\{ \sum_{i,t} r^E_{\alpha}(y_{i,t} - x_{i,t}^T\beta) \right\}$$

One can prove that

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

where \( \hat{\omega}_{it}(\tau) = |\tau - 1(y_{it} \leq x_{it}^T\hat{\beta}^E(\tau))| \).
Expectile Regression with Random Effects (ERRE)

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

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

and then

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

![Graphs showing ERRE and QRRE for various variables including Intercept, Education, Experience, Experience^2, Time, Ability, Mother ed., Father ed., Broken home, and Siblings.](image-url)