Random Effects Quantile Regression

Adam Maidman

December 8, 2014
Suppose we have $n$ clusters with cluster size $m_i$ where $i = 1, \ldots, n$.

Let $\{(y_{ij}, x_{ij})\}$, $j = 1, \ldots, m_i$ denote observations from the $i^{th}$ cluster drawn from the model:

$$Q_{i, \tau}\{y \mid x, b_i(\tau)\} = x^T b_i(\tau)$$

We assume that
(1) $b_i(\tau)$ are $q$ dimensional random variables
(2) with common mean $\beta(\tau)$,
(3) covariance matrix $\Sigma_i(\tau)$,
(4) and densities $g_i(b_i|\theta)$. 

Random Effects Quantile Regression Model
Random Effects Quantile Regression Model

To compare the model with the conditional mean model, we write

$$y_{ij} = x_{ij}^T \beta + x_{ij}^T b_i^* + e_{ij}$$

where $b_i^*$ are mean-0 random effects.
Estimation

We define a semiparametric likelihood-like criterion function for the random coefficients $b_i$,

$$\hat{b}_i = \text{arg min}_{b \in \mathbb{R}^q} \sum_{j=1}^{m_i} \rho_\tau(y_{ij} - x_{ij}^T b)$$

The following empirical likelihood for the coefficient $b_i$ follows from this and other technical details

$$L_{m_i}(b) = \max \left\{ \prod_{j=1}^{m_i} p_j \left| \sum_{j=1}^{m_i} p_j \varphi_\tau(y_{ij} - x_{ij}^T b)x_{ij} = 0, \right.\right.$$

$$\left. \sum_{j=1}^{m_i} p_j = 1, \ 0 \leq p_j \leq 1 \right\}$$
Estimation

We now define a semiparametric likelihood-like criterion for $b_i$ by

$$
\tilde{L}_{m_i}(b_i|\theta) = L_{m_i}(b)g_i(b|\theta)
$$

and for $\theta$ by

$$
\prod_{i=1}^{n} \int_{b_i} \tilde{L}_{m_i}(b_i|\theta) \, db_i
$$
Bayesian Framework

We let $\theta = (\beta, \Sigma)$ and

$$
\begin{align*}
  b_i | \theta & \sim N_q(\beta, \Sigma) \\
  \pi(\theta) & \propto |\Sigma|^{-(q+1)/2},
\end{align*}
$$

giving the posteriors

$$
\begin{align*}
  \Sigma | b_1, \ldots, b_n & \sim IW_{n-1}(S_n) \\
  \beta | \Sigma, b_1, \ldots, b_n & \sim N_q \left( n^{-1} \sum_{i=1}^n b_i, n^{-1} \Sigma \right)
\end{align*}
$$

$$
S_n = \sum_{i=1}^n (b_i - \bar{b}_i)(b_i - \bar{b}_i)^T
$$
Bayesian Framework

\( b_i \)'s not directly observable, so draw from

\[
\tilde{\omega}_n(b_i | \theta) \propto L_{m_i}(b_i) g_i(b_i | \theta)
\]

\[
\Sigma | b_1, \ldots, b_n \sim IW_{n-1}(S_n)
\]

\[
\beta | \Sigma, b_1, \ldots, b_n \sim N_q \left(n^{-1} \sum_{i=1}^{n} b_i, n^{-1} \Sigma\right)
\]
Results for $b_{4,1}$

Model: $Q_{i,\tau}\{y \mid \mathbf{x}, \mathbf{b}_i(\tau)\} = \mathbf{x}^T \mathbf{b}_i(\tau), \quad \mathbf{b}_i \in \mathbb{R}^2, \quad n = 4$

Authors ran the chain for 4,000 steps and threw out first 1,000 steps.
I ran the chain for 35,000 steps and kept everything.
Results for $b_{4,1}$

![ACF plot](image)

<table>
<thead>
<tr>
<th>Steps</th>
<th>Estimates</th>
<th>MCSE</th>
<th>Estimates ± 2SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000-4,000</td>
<td>1.326726</td>
<td>0.06241951</td>
<td>(1.202, 1.451)</td>
</tr>
<tr>
<td>1-35,000</td>
<td>1.247893</td>
<td>0.02210512</td>
<td>(1.204, 1.292)</td>
</tr>
</tbody>
</table>

**Table:** Estimates and Monte Carlo Standard Errors
Conclusions

- MCMC can be used to obtain estimates for the Random Effects Quantile Regression Model
- Consider running the chain longer to obtain reliable results (obvious!)