University of Alberta
Department of Mathematical and Statistical Sciences
Technical Report S139

A CENTRAL LIMIT THEOREM FOR NONLINEAR QUANTILE REGRESSION

Selvakkadunko Selvaratnam[2]

April 2, 2019

Abstract This technical report contains unpublished material, relevant to the article *Model-Robust Designs for Nonlinear Quantile Regression* (S. Selvaratnam, L. Kong, D.P. Wiens). Equation numbers and bibliographic items refer to those in the article.

Proof of Theorem 1*:* The experimenter aims to minimize the loss function

$$\mathcal{L}_E(\boldsymbol{\theta}) = \sum_{i=1}^{n} \rho_\tau[y_i - F(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)].$$

An equivalent loss function is (Oberhofer and Haupt 2016, Asymptotic theory for nonlinear quantile regression under weak dependence, *Econometric Theory* 32: 686-713)

$$\mathcal{L}_H(\boldsymbol{\gamma}) = \sum_{i=1}^{n} \{\rho_\tau[u_i - h_i(\boldsymbol{\gamma})] - \rho_\tau[u_i]\}, \tag{A.1}$$

where $h_i(\boldsymbol{\gamma}) = F(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau + (\boldsymbol{\gamma}/\sqrt{n})) - F(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau)$, $\boldsymbol{\gamma} = \sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_\tau)$, and $u_i = \delta(\boldsymbol{x}_i) + \sigma(\boldsymbol{x}_i)\varepsilon_i$. We expand the function $F(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau + (\boldsymbol{\gamma}/\sqrt{n}))$ by applying Taylor's expansion

$$F(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau + (\boldsymbol{\gamma}/\sqrt{n})) = F(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau) + \boldsymbol{f}'(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau)\frac{\boldsymbol{\gamma}}{\sqrt{n}} + o(1).$$

The loss function in (A.1) is equivalent to the following objective function (Yang et al. 2018, Quantile regression for robust inference on varying coefficient partially linear models, *Journal of the Korean Statistical Society* 47: 172-184):

$$\mathcal{L}_T(\boldsymbol{\gamma}) = \sum_{i=1}^{n} \{\rho_\tau[u_i - \boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma}/\sqrt{n}] - \rho_\tau[u_i]\}. \tag{A.2}$$

Note the identity of Knight (Knight [17]):

$$\rho_\tau(r - s) - \rho_\tau(r) = -s[\tau - I(r \le 0)] + \int_0^s [I(r \le t) - I(r \le 0)]dt.$$

[2]Department of Mathematical and Statistical Sciences; University of Alberta, Edmonton, Alberta; Canada T6G 2G1. e-mail: `selvarat@ualberta.ca`

We apply Knight's identity to (A.2). Thus, we have

$$\begin{aligned}\mathcal{L}_T(\boldsymbol{\gamma}) &= -\sum_{i=1}^{n} s_{ni}\psi_\tau(u_i) + \sum_{i=1}^{n}\int_0^{s_{ni}}[I(u_i \le t) - I(u_i \le 0)]dt \\ &= Z_{1n}(\boldsymbol{\gamma}) + Z_{2n}(\boldsymbol{\gamma}),\ \text{where } s_{ni} = \boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\frac{\boldsymbol{\gamma}}{\sqrt{n}},\end{aligned}$$

where

$$\begin{aligned}Z_{1n}(\boldsymbol{\gamma}) &= -\sum_{i=1}^{n} s_{ni}\psi_\tau(u_i) \text{ and } Z_{2n}(\boldsymbol{\gamma}) = \sum_{i=1}^{n} Z_{2ni}(\boldsymbol{\gamma}) \text{ for} \\ Z_{2ni}(\boldsymbol{\gamma}) &= \int_0^{s_{ni}}[I(u_i \le t) - I(u_i \le 0)]dt.\end{aligned}$$

Let us consider

$$\begin{aligned}E[Z_{1n}(\boldsymbol{\gamma})] &= -\frac{\boldsymbol{\gamma}'}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)E[\psi_\tau(u_i)] \\ &= -\frac{\boldsymbol{\gamma}'}{\sqrt{n}}\sum_{i=1}^{n}\{\tau - G[-\delta^*(\boldsymbol{x}_{(i)})]\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau) \\ &= -\boldsymbol{\gamma}'\frac{\sqrt{n}}{n}\sum_{i=1}^{n}\{g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)}) + o(1)\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau) \\ &\to -\boldsymbol{\gamma}'\left\{\frac{1}{n}\sum_{i=1}^{n}\delta^*(\boldsymbol{x}_{(i)})\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\right\}g_\varepsilon(0)\sqrt{n} \\ &= -\boldsymbol{\gamma}'\left\{\sum_{i=1}^{N}\xi_i\delta^*(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i, \boldsymbol{\theta}_\tau)\right\}g_\varepsilon(0)\sqrt{n} \\ &= -\boldsymbol{\gamma}'\boldsymbol{\mu}g_\varepsilon(0)\sqrt{n}.\end{aligned}$$

Also, we have

$$\begin{aligned}\mathrm{Var}[Z_{1n}(\boldsymbol{\gamma})] &= \sum_{i=1}^{n} s_{ni}^2\mathrm{Var}[\psi_\tau(u_i)] \\ &= \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n} G[-\delta^*(\boldsymbol{x}_{(i)})]\{1 - G[-\delta^*(\boldsymbol{x}_{(i)})]\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma} \\ &= \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\{\tau - g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)}) + o(1)\}\{1 - \tau + g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)}) + o(1)\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma} \\ &\to \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\{\tau - g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)})\}\{1 - \tau + g_\varepsilon(\tau)\delta^*(\boldsymbol{x}_{(i)})\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma} \\ &= \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\left\{\begin{array}{c}\tau(1-\tau) \\ +(2\tau - 1)g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)}) - g_\varepsilon(0)^2\delta^*(\boldsymbol{x}_{(i)})^2\end{array}\right\}\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma}. \quad \text{(A.3)}\end{aligned}$$

Let us consider the second term in braces in (A.3),

$$\begin{aligned}
& \boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\{(2\tau-1)g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)})\}\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{\gamma} \\
= \ & (2\tau-1)g_\varepsilon(0)\boldsymbol{\gamma}'\left\{\frac{1}{n}\sum_{i=1}^{n}\delta^*(\boldsymbol{x}_{(i)})\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\
= \ & (2\tau-1)g_\varepsilon(0)\boldsymbol{\gamma}'\left\{\sum_{i=1}^{N}\xi_i\delta^*(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\
= \ & N(2\tau-1)g_\varepsilon(0)\boldsymbol{\gamma}'\left\{\frac{1}{N}\sum_{i=1}^{N}\xi_i\delta^*(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\
\leq \ & N\boldsymbol{\gamma}'[(2\tau-1)g_\varepsilon(0)]\max_{1\leq i\leq N}\{\parallel \boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\parallel \xi_i\}\parallel\boldsymbol{\gamma}\parallel\frac{1}{N}\sum_{i=1}^{N}\delta^*(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau) \\
\to \ & \mathbf{0} \qquad [\text{by } (5)].
\end{aligned}$$

Thus, we can conclude

$$\boldsymbol{\gamma}'\frac{1}{n}\sum_{i=1}^{n}\{(2\tau-1)g_\varepsilon(0)\delta^*(\boldsymbol{x}_{(i)})\}\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{\gamma}\to\mathbf{0}. \tag{A.4}$$

The third term in braces in (A.3) is

$$\begin{aligned}
& g_\varepsilon(0)^2\boldsymbol{\gamma}'\left\{\frac{1}{n}\sum_{i=1}^{n}\delta^*(\boldsymbol{x}_{(i)})^2\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\
= \ & Ng_\varepsilon(0)^2\boldsymbol{\gamma}'\left\{\frac{1}{N}\sum_{i=1}^{N}\xi_i\delta^*(\boldsymbol{x}_i)^2\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\
= \ & Ng_\varepsilon(0)^2\boldsymbol{\gamma}'\frac{1}{N}\sum_{i=1}^{N}\xi_i\delta^*(\boldsymbol{x}_i)^2\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{f}'(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{\gamma} \\
\leq \ & Ng_\varepsilon(0)^2\boldsymbol{\gamma}'\max_{1\leq i\leq N}\{\xi_i\parallel\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\parallel|\delta^*(\boldsymbol{x}_i)|\}\parallel\boldsymbol{\gamma}\parallel\frac{1}{N}\sum_{i=1}^{N}\delta^*(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau) \\
\to \ & \mathbf{0} \qquad [\text{by } (5)].
\end{aligned}$$

So, we have

$$g_\varepsilon(0)^2\frac{1}{n}\sum_{i=1}^{n}\delta^*(\boldsymbol{x}_{(i)})^2\boldsymbol{\gamma}'\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{\gamma}\to\mathbf{0}. \tag{A.5}$$

By using (A.4) (A.5), and(A.3), we obtain

$$\begin{aligned}\mathrm{Var}[Z_{1n}(\boldsymbol{\gamma})] &\rightarrow \boldsymbol{\gamma}'\tau(1-\tau)\left\{\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\ &\rightarrow \boldsymbol{\gamma}'\tau(1-\tau)E[\boldsymbol{f}(\boldsymbol{x},\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x},\boldsymbol{\theta}_\tau)]\boldsymbol{\gamma} \\ &= \boldsymbol{\gamma}'\tau(1-\tau)\boldsymbol{P}_0\boldsymbol{\gamma}.\end{aligned}$$

Therefore, we have

$$Z_{1n}(\boldsymbol{\gamma}) \xrightarrow{\mathcal{D}} -\boldsymbol{\gamma}'\boldsymbol{w} \text{ where } \boldsymbol{w} \sim N(\boldsymbol{\mu} g_\varepsilon(0)\sqrt{n}, \tau(1-\tau)\boldsymbol{P}_0). \tag{A.6}$$

Next, we consider the component $Z_{2n}(\boldsymbol{\gamma})$:

$$Z_{2n}(\boldsymbol{\gamma}) = \sum_{i=1}^{n} E[Z_{2ni}(\boldsymbol{\gamma})] + \sum_{i=1}^{n}\{Z_{2ni}(\boldsymbol{\gamma}) - E[Z_{2ni}(\boldsymbol{\gamma})]\}.$$

We have

$$\begin{aligned}E[Z_{2n}(\boldsymbol{\gamma})] &= \sum_{i=1}^{n} E[Z_{2ni}(\boldsymbol{\gamma})] \\ &= \sum_{i=1}^{n}\int_0^{s_{ni}}\left\{G\left[-\delta_n^*(\boldsymbol{x}_{(i)}) + \frac{t}{\sigma(\boldsymbol{x}_{(i)})}\right] - G[-\delta_n^*(\boldsymbol{x}_{(i)})]\right\}dt \\ &= \sum_{i=1}^{n}\int_0^{s_{ni}} g_\varepsilon[-\delta_n^*(\boldsymbol{x}_{(i)})]\frac{t}{\sigma(\boldsymbol{x}_{(i)})}dt + o(1) \\ &= \frac{1}{2n}\sum_{i=1}^{n}\frac{g_\varepsilon[-\delta_n^*(\boldsymbol{x}_{(i)})]}{\sigma(\boldsymbol{x}_{(i)})}\boldsymbol{\gamma}'\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{\gamma} + o(1) \\ &\rightarrow \frac{1}{2}g_\varepsilon(0)\boldsymbol{\gamma}'\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma(\boldsymbol{x}_{(i)})}\boldsymbol{f}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x}_{(i)},\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\ &= \frac{1}{2}g_\varepsilon(0)\boldsymbol{\gamma}'\left\{\sum_{i=1}^{N}\xi_i\frac{1}{\sigma(\boldsymbol{x}_i)}\boldsymbol{f}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\boldsymbol{f}^{'}(\boldsymbol{x}_i,\boldsymbol{\theta}_\tau)\right\}\boldsymbol{\gamma} \\ &= \frac{1}{2}\boldsymbol{\gamma}'g_\varepsilon(0)\boldsymbol{P}_1\boldsymbol{\gamma}.\end{aligned}$$

Moreover, we have the bound

$$
\begin{aligned}
\mathrm{Var}[Z_{2n}(\boldsymbol{\gamma})] &\le \sum_{i=1}^{n} E\left\{\int_0^{s_{ni}} [I(u_i \le t) - I(u_i \le 0)]dt\right\}^2 \\
&\le \sum_{i=1}^{n} E\left\{\int_0^{s_{ni}} dt \int_0^{s_{ni}} [I(u_i \le t) - I(u_i \le 0)]dt\right\} \\
&= \sum_{i=1}^{n} E\left\{\boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\frac{\boldsymbol{\gamma}}{\sqrt{n}} \int_0^{s_{ni}} [I(u_i \le t) - I(u_i \le 0)]dt\right\} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{f}'(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\boldsymbol{\gamma} E\left\{\int_0^{s_{ni}} [I(u_i \le t) - I(u_i \le 0)]dt\right\} \\
&\le \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \|\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\| \|\boldsymbol{\gamma}\| E\left\{\int_0^{s_{ni}} [I(u_i \le t) - I(u_i \le 0)]dt\right\} \\
&\le \left\{\frac{1}{\sqrt{n}} \max_{1 \le i \le n} \|\boldsymbol{f}(\boldsymbol{x}_{(i)}, \boldsymbol{\theta}_\tau)\|\right\} \|\boldsymbol{\gamma}\| E[Z_{2n}(\boldsymbol{\gamma})].
\end{aligned}
$$

We have $\mathrm{Var}[Z_{2n}(\boldsymbol{\gamma})] \to 0$. Using $E[Z_{2n}(\boldsymbol{\gamma})] \to \frac{1}{2}\boldsymbol{\gamma}' g(0)\boldsymbol{P}_1\boldsymbol{\gamma}$ and $\mathrm{Var}[Z_{2n}(\boldsymbol{\gamma})] \to 0$, we can obtain

$$E\left[Z_{2n}(\boldsymbol{\gamma}) - \frac{1}{2}\boldsymbol{\gamma}' g_\varepsilon(0)\boldsymbol{P}_1\boldsymbol{\gamma}\right]^2 \to 0.$$

Therefore, we have

$$Z_{2n}(\boldsymbol{\gamma}) \to \frac{1}{2}\boldsymbol{\gamma}' g_\varepsilon(0)\boldsymbol{P}_1\boldsymbol{\gamma}. \tag{A.7}$$

Because of (A.6) and (A.7), we have

$$Z_n(\boldsymbol{\gamma}) \xrightarrow{\mathcal{D}} Z(\boldsymbol{\gamma}),$$

where $Z(\boldsymbol{\gamma}) = -\boldsymbol{\gamma}'\boldsymbol{w} + \frac{1}{2}\boldsymbol{\gamma}' g_\varepsilon(0)\boldsymbol{P}_1\boldsymbol{\gamma}$ and $\boldsymbol{w} \sim N(\boldsymbol{\mu} g_\varepsilon(0)\sqrt{n}, \tau(1-\tau)\boldsymbol{P}_0)$. The convexity of the limiting objective function $Z(\boldsymbol{\gamma})$ ensures the uniqueness of the minimizer $\hat{\boldsymbol{\gamma}} = \frac{1}{g_\varepsilon(0)}\boldsymbol{P}_1^{-1}\boldsymbol{w}$. Therefore, we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{n\tau} - \boldsymbol{\theta}_\tau) = \hat{\boldsymbol{\gamma}}_n = \arg\min Z_n(\boldsymbol{\gamma}) \xrightarrow{\mathcal{D}} \hat{\boldsymbol{\gamma}} = \arg\min Z(\boldsymbol{\gamma}).$$

Thus, we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{n\tau} - \boldsymbol{\theta}_\tau - \boldsymbol{P}_1^{-1}\boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N\left(\mathbf{0}, \frac{\tau(1-\tau)}{g_\varepsilon(0)^2}\boldsymbol{P}_1^{-1}\boldsymbol{P}_0\boldsymbol{P}_1^{-1}\right),$$

as required. ■