Confidence Corridors for Multivariate Generalized Quantile Regression

Shih-Kang Chao, Katharina Proksch
Wolfgang Karl Härdle
Holger Dette

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Chair of Stochastic
Ruhr-Universität Bochum
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://www.ruhr-uni-bochum.de/mathematik3
Treatment effect

- Treatments (program, policy, intervention) affect distributions (income, age)
- $Y_1(Y_0)$: the treatment (control) group; $D$: dummy variable, 
  - Mean: the average treatment effect $\Delta_m = \mathbb{E}[Y_1 - Y_0]$
  - Quantile: $\Delta_\tau = \hat{F}_{1,n}^{-1}(\tau) - \hat{F}_{0,n}^{-1}(\tau)$
- If the experiment is randomized:
  $\mathbb{E}[Y_1 - Y_0] = \mathbb{E}[Y_1|D = 1] - \mathbb{E}[Y_0|D = 0]$
- Measure $\Delta_m$ through a dummy-variable regression:
  \[ Y_i = \alpha + D_i \gamma + X_i^\top \beta + e_i, \quad \text{(Location shift)} \]
  \[ Y_i = \alpha + X_i^\top (\beta + D_i \gamma) + e_i, \quad \text{(scaling)} \]
Quantile treatment effect (QTE)

Doksum (1974): if we define $\Delta(y)$ as the "horizontal distance" between $F_0$ and $F_1$ at $y$ so that

$$F_1(y) = F_0\{y + \Delta(y)\},$$

then $\Delta(y)$ can be expressed as

$$\Delta(y) = F_0^{-1}\{F_1(y)\} - y,$$

changing variable with $\tau = F_1(y)$, one gets the quantile treatment effect:

$$\Delta_\tau = \Delta\{F_1^{-1}(\tau)\} \overset{\text{def}}{=} F_0^{-1}(\tau) - F_1^{-1}(\tau).$$
Figure 1: Heterogeneous horizontal shifts in distribution.
Stochastic dominance (SD)

Conditional stochastic dominance (CSD): Given state variables \( X \), \( Y_1 \) conditionally stochastically dominates \( Y_0 \) if:

\[
F_{1|X}(y|x) \leq F_{0|X}(y|x) \quad \text{a.s. for all } y, x, \quad (1)
\]

Take \( \tau = F_{0|X}(y|x) \), so \( y = F_{0|X}^{-1}(\tau|x) \). Apply \( F_{1|X}^{-1} \) to (??):

\[
F_{0|X}^{-1}(\tau|x) \leq F_{1|X}^{-1}\left\{ F_{0|X}(y|x)|x \right\} = F_{1|X}^{-1}(\tau|x) \quad \forall x, \tau
\]

it preserves the inequality
Which one helps more?

**Type I: Risk reduction CSD**

- Earnings growth
- \( \tau \)
- \( F_0 \)
- \( F_1 \)

**Type II: Potential enhancement CSD**

- Earnings growth
- \( \tau \)
- \( F_0 \)
- \( F_1 \)
Confidence corridors (CC)

(a) $d = 1$

(b) $d = 2$
Distribution comparison & model diagnosis

- It is a common procedure to compare distributions or perform goodness-of-fit test in econometrics.
- Parametric inference: requires prior knowledge on the correct stochastic specification.
- Nonparametric inference gives more flexibility.
- **The big question**: How to statistically compare the nonparametric curve/surfaces?
Confidence corridors: a history

- Bickel & Rosenblatt (1973b) density estimation: \( d = 1 \)
- Rosenblatt (1976) density estimation: \( d \geq 1 \)
- Johnston (1982) function estimation: \( d = 1 \)
- Hardle (1989) \( M \)-estimator: \( d = 1 \)
- Guo & Hardle (2011) ER estimator: \( d = 1 \)
- Hardle & Song (2010) QR estimator: \( d = 1 \)
- Bickel & Rosenblatt (1973a) Max Dev. Dist. Random Field
- Smirnov (1950) Histogram Estimator
- CPHD GQR estimator: \( d \geq 1 \)
Some recent developments

- Claeskens and van Keilegom (2003): local polynomial mean estimator
- Gené and Nickl (2010): adaptive density estimation with wavelets and kernel
- Liu and Wu (2010): long memory, strictly stationary time series density estimation
- Fan and Liu (2013): one dimensional, generic (semi)parametric quantile estimator, avoid estimating conditional density
Outline

1. Motivation ✓
2. Method and Theoretical Results
3. Bootstrap
4. Simulation
5. Application to National Supported Work (NSW) Demonstration data
Additive error model

- Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sequence of i.i.d. random vectors in $\mathbb{R}^{d+1}$ and consider the nonparametric regression model

$$Y_i = \theta(X_i) + \varepsilon_i, \quad i = 1, \ldots, n,$$

where $\theta$ is an aspect of $Y$ conditional on $X$ such as the $\tau$-quantile, the $\tau$-expectile regression curve, $\varepsilon_i$ i.i.d. with $\tau$-quantile/expectile 0.

- Heterogeneity: $\varepsilon_i$ is allowed to be correlated with $X$
Confidence intervals

1 \( - \alpha \)-confidence interval

\[
P \left( \hat{\theta}_n(x) - B_n(x) \leq \theta_0(x) \leq \hat{\theta}_n(x) + B_n(x) \right) = 1 - \alpha
\]

- Confidence statement for one fixed \( x \).
- Only pointwise information!
- Cannot be used to check for global statements without a correction
Confidence corridors

Uniform $1 - \alpha$-confidence corridor on a compact set $D$

$$\mathbb{P}\left(\hat{\theta}_n(x) - \Phi_n(x) \leq \theta_0(x) \leq \hat{\theta}_n(x) + \Phi_n(x) \forall x \in D\right) = 1 - \alpha$$

- True values of $\theta_0(x)$ covered for all $x \in D$ simultaneously by the band with probability $1 - \alpha$.

- Global information about $\theta_0$ on $D$. 
Distribution of the maximal deviation

\[ \mathbb{P} \left( \hat{\theta}(x) - \Phi_n(x) \leq \theta_0(x) \leq \hat{\theta}(x) + \Phi_n(x) \forall x \in \mathcal{D} \right) = 1 - \alpha \]

Goal: Find \( \Phi_n \) such that the equality holds approximately.

Suppose: \( \sup_{x \in \mathcal{D}} w_n(x) |\hat{\theta}(x) - \theta_0(x)| \leq \varphi_n \)

with probability \( 1 - \alpha \) and \( n \to \infty \), which implies

\[ |\hat{\theta}(x) - \theta_0(x)| \leq \frac{\varphi_n}{w_n(x)} \overset{\text{def}}{=} \Phi_n(x) \text{ for all } x \in \mathcal{D}. \]

Approach:

Approximation of the distribution of \( \sup_{x \in \mathcal{D}} w_n(x) |\hat{\theta}(x) - \theta_0(x)|. \)
Consider the local constant estimator

\[ \hat{\theta}_n(x) \overset{\text{def}}{=} \arg \min_{\theta} n^{-1} \sum_{i=1}^{n} K_h(x - X_i) \rho_\tau(Y_i - \theta) \]

with a kernel \( K \) and loss function \( \rho_\tau \).

Uniform nonparametric Bahadur representation:

\[
\sup_{x \in \mathcal{D}} \left| \hat{\theta}_n(x) - \theta_0(x) - \frac{1}{nS_{n,0,0}(x)} \sum_{i=1}^{n} K_h(x - X_i) \psi_\tau \{ Y_i - \theta_0(x) \} \right| = O \left\{ (\log n/nh^d)^{3/4} \right\}, \quad \text{a.s.}[P]
\]
Bahadur representation

\[ S_{n,0,0}(x)(\hat{\theta}_n(x) - \theta_0(x)) \approx \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) \psi_{\tau}\{ Y_i - \theta_0(x) \} \]

\[ \psi_{\tau}(u) = \begin{cases} 
1(u \leq 0) - \tau, & \text{Quantile;} \\
2\{1(u \leq 0) - \tau\}|u|, & \text{Expectile.}
\end{cases} \]

\[ S_{n,0,0}(x) = \begin{cases} 
f_{Y\mid X}(\theta_0(x)\mid x)f_X(x) + O(h^s), & \text{Q;} \\
2[\tau - F_{Y\mid X}(\theta_0(x)\mid x)(2\tau - 1)]f_X(x) + O(h^s), & \text{E.}
\end{cases} \]
Approximating empirical process

\[ V_n^{-1/2} S_{n,0,0}(x) \left\{ \hat{\theta}_n(x) - \theta_0(x) \right\} \]

\[ \approx V_n^{-1/2} \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) \psi \{ Y_i - \theta_0(x) \} \]

\[ \approx \frac{1}{\sqrt{h^d f_X(x) \sigma^2(x)}} \int \int K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} dZ_n(y, u) \]

\[ Y_n(X) \]

\[ \square \text{ with the centered empirical process} \]

\[ Z_n(y, u) \overset{\text{def}}{=} n^{1/2} \{ F_n(y, u) - F(y, u) \}. \]
The empirical processes of QR

\[ Y_n(x) \xrightarrow{\mathcal{L}} Y_0,n(x) \xrightarrow{\mathcal{O}_p\{(\log n)^{-1}\}} Y_1,n(x) \xrightarrow{\text{strong approx.}} \text{Gaussian} \]

\[ Y_4,n(x) \xrightarrow{\mathcal{L}} Y_3,n(x) \xrightarrow{\mathcal{O}_p(h^{1/2-\delta})} Y_2,n(x) \]

\[ Y_5,n(x) \xrightarrow{\mathcal{O}_p(h^{1-\delta})} \text{Gaussian & Stationary} \]

Rosenblatt (1976): \( \sup_x Y_5,n(x) \xrightarrow{\mathcal{L}} \text{Gumbel} \)
The empirical process of ER

\[ Y_n(x) \overset{\mathcal{O}_p\{(\log n)^{-1}\}}{\longrightarrow} Y_{0,n}(x) \overset{\text{strong approx.}}{\longrightarrow} Y_{1,n}(x) \overset{\mathcal{O}_p\{(\log n)^{-1}\}}{\longrightarrow} \text{Gaussian} \]

\[ Y_{4,n}(x) \overset{\mathcal{L}}{\longrightarrow} Y_{3,n}(x) \overset{\mathcal{O}_p(h^{1-\delta})}{\longrightarrow} Y_{2,n}(x) \]

\[ \mathcal{O}_p(h^{d/2}) \]

\[ Y_{5,n}(x) \overset{\mathcal{O}_p(h^{1-\delta})}{\longrightarrow} \text{Gaussian & Stationary} \]

Rosenblatt (1976): \( \sup_x Y_{5,n}(x) \overset{\mathcal{L}}{\longrightarrow} \text{Gumbel} \)
Step 1: Support truncation

\[ Y_0(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2(x)}} \int \int K \left( \frac{x-u}{h} \right) \psi \{ y - \theta_0(x) \} dZ_n(y,u), \]

- \( \Gamma_n = \{ y : |y| \leq a_n \} \)
- \( \sigma^2_n(x) = E[\psi^2(Y - \theta_0(x))1(Y_i \leq a_n)|X = x] \)
- Claim: \( \| Y_0 - Y_{n,0} \| = O_P((\log n)^{-1/2}) \)
Step 1: Support truncation

\[ Y_{0,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} \, dZ_n(y, u), \]

- \( \Gamma_n = \{ y : |y| \leq a_n \} \)
- \( \sigma_n^2(x) = E \left[ \psi^2 \{ Y - \theta_0(x) \} \mathbf{1}(Y_i \leq a_n) | X = x \right] \)
- Claim: \( \| Y_0 - Y_{n,0} \| = \mathcal{O}_P \{ (\log n)^{-1/2} \} \)
Step 1: Support truncation

- Show \( (Y_{n,0} - Y_{n,0})(x) \xrightarrow{P} 0 \) for each \( x \) and tightness.

  - **Tightness Lemma**

- Necessary to control the decay of the tail of distribution of \( Y \)

- Watch out for difference in quantile and expectile regression:
  - Quantile: very weak assumption (A2)
  - Expectile: exploding boundary deteriorates the strong approximation rate \( \rightarrow \) requiring at least finite forth conditional moment (EA2)
Step 2: Strong approximation

\[ Y_{0,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x - u}{h} \right) \psi \{y - \theta_0(x)\} \, dZ_n(y, u), \]

where

\[ T(y, u) = \{ F_{X_1 | Y}(u_1 | y), F_{X_2 | Y}(u_2 | u_1, y), \ldots, \]
\[ F_{X_d | x_{d-1}, \ldots, x_1, Y}(u_d | u_{d-1}, \ldots, u_1, y), F_Y(y) \} \]

is the Rosenblatt transformation and

\[ B_n(T(y, u)) = W_n(T(y, u)) - F(y, u)W(1, \ldots, 1) \]

a multivariate Brownian bridge.

Claim: \[ \|Y_{0,n} - Y_{1,n}\| = O_p\left\{ (\log n)^{-1} \right\}, \text{ a.s.} \]
Step 2: Strong approximation

\[ Y_{1,n}(x) = \frac{1}{\sqrt{h^d f_X(x)\sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x-u}{h} \right) \psi \{ y - \theta_0(x) \} \ dB_n(T(y,u)) \]

where

\[ T(y,u) = \{ F_{X_1|Y}(u_1|y), F_{X_2|Y}(u_2|u_1,y), \ldots, \]
\[ \quad F_{X_d|X_{d-1},\ldots,X_1,Y}(u_d|u_{d-1},\ldots,u_1,y), F_Y(y) \} \]

is the Rosenblatt transformation and

\[ B_n(T(y,u)) = W_n(T(y,u)) - F(y,u)W(1,\ldots,1) \]

a multivariate Brownian bridge.

Claim: \( \| Y_{0,n} - Y_{1,n} \| = O_p\{ (\log n)^{-1} \}, \text{ a.s.} \)
Step 3: Brownian bridge → Wiener sheet

\[ Y_{2,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} \, dB_n \left( T(y, u) \right). \]

Claim: \( \| Y_{1,n} - Y_{2,n} \| = O_p \left( h^{d/2} \right) \)

- by integration by parts
- since \( |W(1, ..., 1)| \leq O_P(1) \).
Step 3: Brownian bridge \( \rightarrow \) Wiener sheet

\[
Y_{2,n}(x) = \frac{1}{\sqrt{h^df_X(x)\sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x-u}{h} \right) \psi \{ y - \theta_0(x) \} \, dW_n(T(y,u)).
\]

Claim: \( \| Y_{1,n} - Y_{2,n} \| = O_p\left( h^{d/2} \right) \)

- by integration by parts
- since \( |W(1, \ldots, 1)| \leq O_P(1) \).
Step 4: Stationarise the process

\[ Y_{2,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x-u}{h} \right) \psi \{ y - \theta_0(x) \} \, dW_n \left( T(y, u) \right) \]

Claim: \[ \| Y_{2,n} - Y_{3,n} \| = O_P(h^{1-\delta}), \text{ for any } \delta > 0 \]

A supremum concentration inequality for Gaussian field is applied.

Meerschaert et al. (2013)
Step 4: Stationarise the process

\[ Y_{3,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \int_{\Gamma_n} K \left( \frac{x - u}{h} \right) \psi \{y - \theta_0(u)\} \, dW_n(T(y, u)) \]

Claim: \[ \|Y_{2,n} - Y_{3,n}\| = O_P(h^{1-\delta}), \text{ for any } \delta > 0 \]

A supremum concentration inequality for Gaussian field is applied.  

Meerschaert et al. (2013)
Step 5: Equally distributed

\[ Y_{3,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int \int_{\Gamma_n} K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(u) \} \, dW_n(T(y, u)). \]

Claim: \( Y_{3,n} \overset{d}{=} Y_{4,n} \)

A computation of the covariance functions gives the result.
Step 5: Equally distributed

\[ Y_{4,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int K \left( \frac{x - u}{h} \right) \sqrt{\sigma_n^2(u) f_X(u)} dW(u). \]

Claim: \( Y_{3,n} \overset{d}{=} Y_{4,n} \)

A computation of the covariance functions gives the result.
Step 6: Final stationarisation

\[ Y_{4,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma_n^2(x)}} \int K \left( \frac{x - u}{h} \right) \sqrt{\sigma_n^2(u) f_X(u)} dW(u). \]

\[ Y_{5,n}(x) = \frac{1}{\sqrt{h^d}} \int K \left( \frac{x - u}{h} \right) dW(u). \]

Claim: \( \| Y_{4,n} - Y_{5,n} \| = O_p(h^{1-\delta}) \) for \( \delta > 0 \).

Supremum concentration inequality for Gaussian field is again applied.

Meerschaert et al. (2013)
Maximal deviation for nonparametric QR

Theorem (1)

Under regularity conditions, \( \text{vol}(\mathcal{D}) = 1 \),

\[
\begin{align*}
\mathbb{P} \left\{ (2d \kappa \log n)^{1/2} \left( \sup_{x \in \mathcal{D}} \left[ r(x) |\hat{\theta}_n(x) - \theta_0(x)| \right] / \|K\|_2 - d_n \right) < a \right\} \\
\rightarrow \exp \left\{ -2 \exp(-a) \right\},
\end{align*}
\]

as \( n \to \infty \), where \( \hat{\theta}_n(x) \) and \( \theta_0(x) \) are the local constant quantile estimator and the true quantile function.
Corollary (RQ-CC)

Under the assumptions of Theorem ??, an approximate $(1 - \alpha) \times 100\%$ confidence corridor over $\alpha \in (0, 1)$ is

$$\hat{\theta}_n(t) \pm (nh^d)^{-1/2} \{\tau (1 - \tau) \|K\|_2 / \hat{f}_X(t)\}^{1/2} \hat{f}_\varepsilon |X \{0|t\}^{-1} \left\{d_n + c(\alpha) (2d \kappa \log n)^{-1/2}\right\},$$

where $c(\alpha) = \log 2 - \log |\log (1 - \alpha)|$ and $\hat{f}_X(t)$, $\hat{f}_\varepsilon |X \{0|t\}$ are consistent estimates for $f_X(t)$, $f_\varepsilon |X \{0|t\}$.
Theorem (2)

Under regularity conditions, \( \text{vol}(\mathcal{D}) = 1, \)

\[
P \left\{ (2d \kappa \log n)^{1/2} \left( \sup_{x \in \mathcal{D}} \left[ r(x) |\hat{\theta}_n(x) - \theta_0(x)| \right] / \| K \|_2 - d_n \right) < a \right\} 
\]

\[
\rightarrow \exp \left\{ -2 \exp(-a) \right\},
\]

as \( n \to \infty, \) where \( \hat{\theta}_n(x) \) and \( \theta_0(x) \) are the local constant expectile estimator and the true expectile function.
Corollary (RE-CC)

*Under the assumptions of Theorem ??*, an approximate \((1 - \alpha) \times 100\%\) confidence corridor over \(\alpha \in (0, 1)\) is

\[
\hat{\theta}_n(t) \pm (nh^d)^{-1/2} \left\{ \hat{\sigma}^2(x) \| K \|_2 \hat{f}_X(t) \right\}^{1/2} 0.5 \left[ \tau - \hat{F}_{\epsilon|X}(0|x)(2\tau - 1) \right]^{-1} \left\{ d_n + c(\alpha)(2d\kappa \log n)^{-1/2} \right\},
\]

where \(c(\alpha) = \log 2 - \log |\log(1 - \alpha)|\) and \(\hat{f}_X(t), \hat{\sigma}^2(x)\) and \(\hat{F}_{\epsilon|X}(0|x)\) are consistent estimates for \(f_X(t), \sigma^2(x)\) and \(F_{\epsilon|X}(0|x)\).
Estimating scaling factors

we propose to estimate $F_{\varepsilon|X}$, $f_{\varepsilon|X}$ and $\sigma^2(x)$ based on residuals $\hat{\varepsilon}_i = Y_i - \hat{\theta}_n(x_i)$:

$$\hat{F}_{\varepsilon|X}(v|x) = n^{-1} \sum_{i=1}^{n} G \left( \frac{v - \hat{\varepsilon}_i}{h_0} \right) L_{\bar{h}}(x - X_i)/\hat{f}_X(x)$$ (3)

$$\hat{f}_{\varepsilon|X}(v|x) = n^{-1} \sum_{i=1}^{n} g_{h_0} (v - \hat{\varepsilon}_i) L_{\bar{h}}(x - X_i)/\hat{f}_X(x)$$ (4)

$$\hat{\sigma}^2(x) = n^{-1} \sum_{i=1}^{n} \psi^2(\hat{\varepsilon}_i) L_{\bar{h}}(x - X_i)/\hat{f}_X(x)$$ (5)

where $G$ is a CDF, $g$ and $L$ are a kernel functions, $h_0, \bar{h} \to 0$ and $nh_0 \bar{h}^d \to \infty$
Lemma

Under regularity conditions, we have

1. \[ \sup_{v \in I} \sup_{x \in D} |\hat{F}_\varepsilon(v|x) - F_\varepsilon(v|x)| = O_p(n^{-\lambda}) \]
2. \[ \sup_{v \in I} \sup_{x \in D} |\hat{f}_\varepsilon(v|x) - f_\varepsilon(v|x)| = O_p(n^{-\lambda}) \]
3. \[ \sup_{x \in D} |\hat{\sigma}^2(x) - \sigma^2(x)| = O_p(n^{-\lambda_1}) \]

where \( n^{-\lambda} = h_0^2 + h^s + \bar{h}^2 + (nh_0\bar{h}^d)^{-1/2} \log n + (nh^d)^{-1/2} \log n \),
and \( n^{-\lambda_1} = h^s + \bar{h}^2 + (n\bar{h}^d)^{-1/2} \log n + (nh^d)^{-1/2} \log n \).
Bootstrap

Smooth bootstrap:

\[ \hat{f}_{\varepsilon,X}(\nu, x) = \frac{1}{n} \sum_{i=1}^{n} g_{h_0}(\nu - \hat{\varepsilon}_i)L_{\bar{h}}(x - X_i), \quad (6) \]

where \( g \) and \( L \) are kernels and \( h_0, \bar{h} \rightarrow 0, \, nh_0\bar{h}^d \rightarrow \infty \)

Define

\[ \hat{\theta}^*(x) - \hat{\theta}_n(x) \]

\[ = \frac{1}{n \hat{S}_{n,0,0}(x)} \sum_{i=1}^{n} K_h(x - X_i^*)\psi(\varepsilon_i^*) - E^* \left[ K_h(x - X_i^*)\psi(\varepsilon_i^*) \right] \]

Remove the bias

\[ \hat{S}_{n,0,0}(x) = \left\{ \begin{array}{ll} \hat{f}_{\varepsilon|x}(0|x)\hat{f}_X(x), & \text{quantile case;} \\ 2[\tau - \hat{F}_{\varepsilon|x}(0|x)(2\tau - 1)]\hat{f}_X(x), & \text{exptile case.} \end{array} \right. \]
**Theorem (Bootstrap)**

*Under regularity conditions, let*

\[
  r^*(x) = \sqrt{\frac{nh^d}{\hat{f}_X(x) \sigma_*^2(x)}} \hat{S}_{n,0,0}(x),
\]

*Then as* \( n \to \infty \),

\[
  \mathbb{P}^* \left\{ (2d \kappa \log n)^{1/2} \left( \sup_{x \in \mathcal{D}} \left[ r^*(x) |\hat{\theta}^*(x) - \hat{\theta}_n(x) | \right] / \| K \|_2 - d_n \right) < a \right\} \to \exp \left[ -2 \exp(-a) \right], \ a.s.
\]

**Lemma**

*Under regularity conditions,*

\[
  \| \sigma_*^2(x) - \hat{\sigma}^2(x) \| = o_p^* \left( (\log n)^{-1/2} \right), \ a.s.
\]

Corollary

Under the regularity conditions, the bootstrap confidence set is defined by

$$\left\{ \theta : \sup_{x \in D} \left| \frac{\hat{S}_{n,0,0}(x)}{\sqrt{\hat{f}_X(x)\hat{\sigma}^2(x)}} \left[ \hat{\theta}_n(x) - \theta(x) \right] \right| \leq \xi_\alpha \right\},$$

where $\xi_\alpha$ satisfies

$$P^* \left( \sup_{x \in D} \left| \frac{\hat{S}_{n,0,0}(x)}{\sqrt{\hat{f}_X(x)\hat{\sigma}^2(x)}} \left[ \hat{\theta}^*(x) - \hat{\theta}_n(x) \right] \right| \leq \xi_\alpha \right) = 1 - \alpha,$$

where $\alpha$ is the level of the test and $\hat{S}_{n,0,0}$ is defined as in (??).
Implementation problem for QR: The CC (??) for QR tends to be too narrow.

Figure 2: Confidence corridors: regression quantiles $\tau = 50\%$. Green: Asymptotic confidence band. Blue: Bootstrap confidence band.
Bootstrap CC for QR

Observation:

\[
\hat{f}_{\varepsilon|x}(0|x) = n^{-1} \sum_{i=1}^{n} g_{h_0}(\hat{\varepsilon}_i) L_{\bar{h}}(x - X_i) / \hat{f}_X(x) \\
\hat{f}_{Y|x}(\hat{\theta}_n(x)|x) = n^{-1} \sum_{i=1}^{n} g_{h_1} \left( Y_i - \hat{\theta}_n(x) \right) L_{\bar{h}}(x - X_i) / \hat{f}_X(x),
\]

(10)  

(11)

are NOT equivalent in finite sample, and \( \hat{f}_{Y|x}(\hat{\theta}_n(x)|x) \) accounts more for the bias

Bootstrap CC for QR

Hence, we propose to construct CC for QR by

\[
\left\{ \theta : \sup_{x \in D} \left| \sqrt{\hat{f}_X(x)} \hat{f}_Y \{ \hat{\theta}_n(x) \} x \right| \left[ \hat{\theta}_n(x) - \theta(x) \right] \right| \leq \xi_{\alpha}^\dagger \right\},
\]

where \( \xi_{\alpha}^\dagger \) satisfies

\[
P^* \left( \sup_{x \in D} \left| \hat{f}_X(x)^{-1/2} \frac{\hat{f}_Y \{ \hat{\theta}_n(x) \} x}{\hat{f}_X \{ \hat{\theta}_n(x) \} x} \left[ A^*_n(x) - E^* A^*_n(x) \right] \right| \leq \xi_{\alpha}^\dagger \right) \approx 1 - \alpha.
\]
Simulated coverage probabilities

Generating process: $d = 2$

$$Y_i = f(X_{1,i}, X_{2,i}) + \sigma(X_{1,i}, X_{2,i})\varepsilon_i,$$

- $f(x_1, x_2) = \sin(2\pi x_1) + x_2.$
- $(X_1, X_2)$ supported on $[0, 1]^2$ with corr. $= 0.2876$.
- $\varepsilon_i \sim N(0, 1)$ i.i.d.
- Specification for $\sigma(X_1, X_2)$:
  - Homogeneity: $\sigma(X_1, X_2) = \sigma_0$, for $\sigma_0 = 0.2, 0.5, 0.7$
  - Heterogeneity:
    $$\sigma(X_1, X_2) = \sigma_0 + 0.8X_1(1 - X_1)X_2(1 - X_2)$$
    for $\sigma_0 = 0.2, 0.5, 0.7$
Simulated coverage probabilities

- Quantile regression bandwidth choice:
  - Rule-of-thumb for conditional density in R package np
  - Yu and Jones (1998) quantile regression adjustment (not applied to expectile)
  - Undersmoothed by $n^{-0.05}$

- Expectile bandwidth choice: Rule-of-thumb for conditional density and undersmoothed by $n^{-0.05}$

- $n = 100, 300, 500$.
  2000 simulation runs are carried out.
Table 1: Nonparametric quantile model asymptotic coverage probability. Nominal coverage is 95%. The digit in the parentheses is the volume.

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous</th>
<th></th>
<th>Heterogeneous</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau = 0.5$</td>
<td>$\tau = 0.2$</td>
<td>$\tau = 0.8$</td>
<td>$\tau = 0.5$</td>
</tr>
<tr>
<td>$\sigma_0 = 0.2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.000(0.366)</td>
<td>0.109(0.720)</td>
<td>0.104(0.718)</td>
<td>0.000(0.403)</td>
</tr>
<tr>
<td>300</td>
<td>0.000(0.304)</td>
<td>0.130(0.518)</td>
<td>0.133(0.519)</td>
<td>0.002(0.349)</td>
</tr>
<tr>
<td>500</td>
<td>0.000(0.262)</td>
<td>0.117(0.437)</td>
<td>0.142(0.437)</td>
<td>0.008(0.296)</td>
</tr>
<tr>
<td>$\sigma_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.070(0.890)</td>
<td>0.269(1.155)</td>
<td>0.281(1.155)</td>
<td>0.078(0.932)</td>
</tr>
<tr>
<td>300</td>
<td>0.276(0.735)</td>
<td>0.369(0.837)</td>
<td>0.361(0.835)</td>
<td>0.325(0.782)</td>
</tr>
<tr>
<td>500</td>
<td>0.364(0.636)</td>
<td>0.392(0.711)</td>
<td>0.412(0.712)</td>
<td>0.381(0.669)</td>
</tr>
<tr>
<td>$\sigma_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.160(1.260)</td>
<td>0.381(1.522)</td>
<td>0.373(1.519)</td>
<td>0.155(1.295)</td>
</tr>
<tr>
<td>300</td>
<td>0.438(1.026)</td>
<td>0.450(1.109)</td>
<td>0.448(1.110)</td>
<td>0.481(1.073)</td>
</tr>
<tr>
<td>500</td>
<td>0.533(0.888)</td>
<td>0.470(0.950)</td>
<td>0.480(0.949)</td>
<td>0.564(0.924)</td>
</tr>
</tbody>
</table>
Table 2: Nonparametric quantile model bootstrap coverage probability. Nominal coverage is 95%. The digit in the parentheses is the volume.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.325(0.676)</td>
<td>.784(0.954)</td>
<td>.783(0.954)</td>
<td>.409(0.717)</td>
<td>.779(0.983)</td>
<td>.778(0.985)</td>
</tr>
<tr>
<td>300</td>
<td>.442(0.457)</td>
<td>.896(0.609)</td>
<td>.894(0.610)</td>
<td>.580(0.504)</td>
<td>.929(0.650)</td>
<td>.922(0.649)</td>
</tr>
<tr>
<td>500</td>
<td>.743(0.411)</td>
<td>.922(0.502)</td>
<td>.921(0.502)</td>
<td>.839(0.451)</td>
<td>.950(0.535)</td>
<td>.952(0.536)</td>
</tr>
</tbody>
</table>

For $\sigma_0 = 0.2$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.929(1.341)</td>
<td>.804(1.591)</td>
<td>.818(1.589)</td>
<td>.938(1.387)</td>
<td>.799(1.645)</td>
<td>.773(1.640)</td>
</tr>
<tr>
<td>300</td>
<td>.950(0.920)</td>
<td>.918(1.093)</td>
<td>.923(1.091)</td>
<td>.958(0.973)</td>
<td>.919(1.155)</td>
<td>.923(1.153)</td>
</tr>
<tr>
<td>500</td>
<td>.988(0.861)</td>
<td>.968(0.943)</td>
<td>.962(0.942)</td>
<td>.990(0.902)</td>
<td>.962(0.986)</td>
<td>.969(0.987)</td>
</tr>
</tbody>
</table>

For $\sigma_0 = 0.5$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.976(1.811)</td>
<td>.817(2.112)</td>
<td>.808(2.116)</td>
<td>.981(1.866)</td>
<td>.826(2.178)</td>
<td>.809(2.176)</td>
</tr>
<tr>
<td>300</td>
<td>.986(1.253)</td>
<td>.919(1.478)</td>
<td>.934(1.474)</td>
<td>.983(1.308)</td>
<td>.930(1.537)</td>
<td>.920(1.535)</td>
</tr>
<tr>
<td>500</td>
<td>.996(1.181)</td>
<td>.973(1.280)</td>
<td>.968(1.278)</td>
<td>.997(1.225)</td>
<td>.969(1.325)</td>
<td>.962(1.325)</td>
</tr>
</tbody>
</table>

For $\sigma_0 = 0.7$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.976(1.811)</td>
<td>.817(2.112)</td>
<td>.808(2.116)</td>
<td>.981(1.866)</td>
<td>.826(2.178)</td>
<td>.809(2.176)</td>
</tr>
<tr>
<td>300</td>
<td>.986(1.253)</td>
<td>.919(1.478)</td>
<td>.934(1.474)</td>
<td>.983(1.308)</td>
<td>.930(1.537)</td>
<td>.920(1.535)</td>
</tr>
<tr>
<td>500</td>
<td>.996(1.181)</td>
<td>.973(1.280)</td>
<td>.968(1.278)</td>
<td>.997(1.225)</td>
<td>.969(1.325)</td>
<td>.962(1.325)</td>
</tr>
</tbody>
</table>
Table 3: Nonparametric expectile model asymptotic coverage probability. Nominal coverage is 95%. The digit in the parentheses is the volume.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
<th>$\tau = 0.5$</th>
<th>$\tau = 0.2$</th>
<th>$\tau = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_0 = 0.2$</th>
<th></th>
<th>$\sigma_0 = 0.5$</th>
<th></th>
<th>$\sigma_0 = 0.7$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>.000(0.428)</td>
<td>.000(0.333)</td>
<td>.000(0.333)</td>
<td>.000(0.463)</td>
<td>.000(0.362)</td>
<td>.000(0.361)</td>
</tr>
<tr>
<td>300</td>
<td>.049(0.341)</td>
<td>.000(0.273)</td>
<td>.000(0.273)</td>
<td>.079(0.389)</td>
<td>.001(0.316)</td>
<td>.002(0.316)</td>
</tr>
<tr>
<td>500</td>
<td>.168(0.297)</td>
<td>.000(0.243)</td>
<td>.000(0.243)</td>
<td>.238(0.336)</td>
<td>.003(0.278)</td>
<td>.002(0.278)</td>
</tr>
<tr>
<td>100</td>
<td>.007(0.953)</td>
<td>.000(0.776)</td>
<td>.000(0.781)</td>
<td>.007(0.997)</td>
<td>.000(0.818)</td>
<td>.000(0.818)</td>
</tr>
<tr>
<td>300</td>
<td>.341(0.814)</td>
<td>.019(0.708)</td>
<td>.017(0.709)</td>
<td>.355(0.862)</td>
<td>.017(0.755)</td>
<td>.018(0.754)</td>
</tr>
<tr>
<td>500</td>
<td>.647(0.721)</td>
<td>.067(0.645)</td>
<td>.065(0.647)</td>
<td>.654(0.759)</td>
<td>.061(0.684)</td>
<td>.068(0.684)</td>
</tr>
<tr>
<td>100</td>
<td>.012(1.324)</td>
<td>.000(1.107)</td>
<td>.000(1.107)</td>
<td>.010(1.367)</td>
<td>.000(1.145)</td>
<td>.000(1.145)</td>
</tr>
<tr>
<td>300</td>
<td>.445(1.134)</td>
<td>.021(1.013)</td>
<td>.013(1.016)</td>
<td>.445(1.182)</td>
<td>.017(1.062)</td>
<td>.016(1.060)</td>
</tr>
<tr>
<td>500</td>
<td>.730(1.006)</td>
<td>.062(0.928)</td>
<td>.078(0.929)</td>
<td>.728(1.045)</td>
<td>.068(0.966)</td>
<td>.066(0.968)</td>
</tr>
</tbody>
</table>

Table 4: Nonparametric expectile model bootstrap coverage probability. Nominal coverage is 95%. The digit in the parentheses is the volume.

<table>
<thead>
<tr>
<th></th>
<th>Homogeneous</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$\tau = 0.5$</td>
<td>$\tau = 0.2$</td>
<td>$\tau = 0.8$</td>
<td>$\tau = 0.5$</td>
<td>$\tau = 0.2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_0 = 0.2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.686(2.191)</td>
<td>.781(2.608)</td>
<td>.787(2.546)</td>
<td>.706(2.513)</td>
<td>.810(2.986)</td>
<td>.801(2.943)</td>
</tr>
<tr>
<td>300</td>
<td>.762(0.584)</td>
<td>.860(0.716)</td>
<td>.876(0.722)</td>
<td>.788(0.654)</td>
<td>.877(0.807)</td>
<td>.887(0.805)</td>
</tr>
<tr>
<td>500</td>
<td>.771(0.430)</td>
<td>.870(0.533)</td>
<td>.875(0.531)</td>
<td>.825(0.516)</td>
<td>.907(0.609)</td>
<td>.904(0.615)</td>
</tr>
<tr>
<td>$\sigma_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.886(5.666)</td>
<td>.906(6.425)</td>
<td>.915(6.722)</td>
<td>.899(5.882)</td>
<td>.927(6.667)</td>
<td>.913(6.571)</td>
</tr>
<tr>
<td>300</td>
<td>.956(1.508)</td>
<td>.958(1.847)</td>
<td>.967(1.913)</td>
<td>.965(1.512)</td>
<td>.962(1.866)</td>
<td>.969(1.877)</td>
</tr>
<tr>
<td>500</td>
<td>.968(1.063)</td>
<td>.972(1.322)</td>
<td>.972(1.332)</td>
<td>.972(1.115)</td>
<td>.971(1.397)</td>
<td>.974(1.391)</td>
</tr>
<tr>
<td>$\sigma_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.913(7.629)</td>
<td>.922(8.846)</td>
<td>.935(8.643)</td>
<td>.929(8.039)</td>
<td>.935(9.057)</td>
<td>.932(9.152)</td>
</tr>
<tr>
<td>300</td>
<td>.969(2.095)</td>
<td>.969(2.589)</td>
<td>.971(2.612)</td>
<td>.974(2.061)</td>
<td>.972(2.566)</td>
<td>.979(2.604)</td>
</tr>
<tr>
<td>500</td>
<td>.978(1.525)</td>
<td>.976(1.881)</td>
<td>.967(1.937)</td>
<td>.981(1.654)</td>
<td>.978(1.979)</td>
<td>.974(2.089)</td>
</tr>
</tbody>
</table>

Application to NSW demonstration data

- National Supported Work (NSW): a randomized, temporary employment program carried out in the US in 1970s to help the disadvantaged workers
- 297 obs. treatment group; 425 obs. control group, all male
- Delgado and Escancian (2013): heterogeneity effect in age; nonnegative treatment effect
- $X_1$: Age; $X_2$: schooling in years; $Y$: Earning difference 78-75 (in thousand $)
- Bootstrap: 10,000 repetition
Unconditional kernel densities. Magenta: treatment group. Blue: control group. $h_{tr} = 1.652$. $h_{co} = 1.231$. 

Figure 3: Confidence corridors for treatment (net surfaces) and control group (solid surfaces).

(a) $\tau = 10\%$

(b) $\tau = 20\%$
(a) $\tau = 80\%$

(b) $\tau = 90\%$

Figure 4: Confidence corridors for treatment (net surfaces) and control group (solid surfaces).
Figure 5: Quantile estimates for treatment (solid surfaces) and confidence corridors control group (net surfaces).
Figure 6: Quantile estimates for treatment (solid surfaces) and confidence corridors control group (net surfaces).

(a) $\tau = 80\%$

(b) $\tau = 90\%$
Summary

- The nonnegative CSD is not rejected, confirming the findings of Delgado and Escanciano (2013)
- Heterogeneous effect in age and schooling in years: individuals who are older and spend more time in the school benefit more from the treatment
- We show: treatment raises the potential for realizing higher earnings growth, but does little in reducing the risk of realizing low earnings growth
Confidence Corridors for Multivariate Generalized Quantile Regression

Shih-Kang Chao, Katharina Proksch
Wolfgang Karl Härdle
Holger Dette

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Chair of Stochastic
Ruhr-Universität Bochum

http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://www.ruhr-uni-bochum.de/mathematik3
Assumptions

(A1) $K$ is of order $s - 1$ (see (A3)) has bounded support $[-A, A]^d$, continuously differentiable up to order $d$ (and are bounded); i.e. $\partial^\alpha K \in L^1(\mathbb{R}^d)$ exists and is continuous for all multi-indices $\alpha \in \{0, 1\}^d$.

(A2) The increasing sequence $\{a_n\}_{n=1}^\infty$ satisfies

$$\left(\log n\right) h^{-3d} \int_{|y| > a_n} f_Y(y) dy = O(1)$$

(12)

and

$$\left(\log n\right) h^{-d} \int_{|y| > a_n} f_{Y|X}(y|x) dy = O(1), \text{ for all } x \in \mathcal{D}$$

as $n \to \infty$ hold.

(A3) The true function $\theta_0(x)$ is continuously differentiable and is in Hölder class with order $s > d$. 
Assumptions

(A4) $f_X(x)$ is continuously differentiable and its gradient is uniformly bounded. In particular, $\inf_{x \in D} f_X(x) > 0$.

(A5) The joint probability density function $f(y, u)$ is positive and continuously differentiable up to $s$th order (needed for Rosenblatt transform), and the conditional density $f_Y|X(y|X = x)$ is continuously differentiable with respect to $x$.

(A6) $h$ satisfies $\sqrt{nh^d} h^s \sqrt{\log n} \to 0$ (undersmoothing), and $nh^{3d} \to \infty$ as $n \to \infty$
Assumptions

(EA2) $\sup_{x \in D} \left| \int v^{b_1} f_{\varepsilon}(v|x) dv \right| < \infty$, where $b_1$ satisfies

$$n^{-1/6} h^{-d/2-3d/(b_1-2)} = O(n^{-\nu}), \quad \nu > 0.$$

e.g. when $h = n^{-1/(2s+d)}$, then $b_1 > (4s + 14d)/(2s + d - 3)$. 

▶ Thm RE-Band
▶ Emp. process ER
Assumptions

(B1) $L$ is a Lipschitz, bounded, symmetric kernel. $G$ is Lipschitz continuous cdf, and $g$ is the derivative of $G$ and is also a density, which is Lipschitz continuous, bounded, symmetric and five times continuously differentiable kernel.

(B2) $F_{\epsilon|X}(v|x)$ is in $s'+1$ order Hölder class with respect to $v$ and continuous in $x$, $s'>\max\{2,d\}$. $f_X(x)$ is in second order Hölder class with respect to $x$ and $v$. $E[\psi^2(\epsilon_i)|x]$ is second order continuously differentiable with respect to $x \in \mathcal{D}$.

(B3) $nh_0\bar{h}^d \to \infty$, $h_0, \bar{h} = \mathcal{O}(n^{-\nu})$, where $\nu > 0$. 

Scaling factors
Assumptions

(C1) There exist an increasing sequence \( c_n \), \( c_n \to \infty \) as \( n \to \infty \) such that

\[
(\log n)^3 (nh^{6d})^{-1} \int_{|v|>c_n/2} f_\varepsilon(v)dv = O(1), \quad (13)
\]

as \( n \to \infty \).

(EC1) \( \sup_{x \in \mathcal{D}} \left| \int v^b f_\varepsilon(x(v|x))dv \right| < \infty \), where \( b \) satisfies

\[
n^{-\frac{1}{6} + \frac{4}{b^2} - \frac{1}{b} h - \frac{d}{2} - \frac{6d}{b}} = O(n^{-\nu}), \quad \nu > 0, \text{ (Thm. ??)}
\]

and

\[
b > 2(2s' + d + 1)/(2s' + 3). \text{ (Lemma ??)}
\]
Quantile regression notations

\[ h = n^{-\kappa}, \quad \rho_\tau(u) = |\tau - 1(u < 0)||u|, \quad \psi(u) = 1(u \leq 0) - \tau \]

\[ d_n = (2d\kappa \log n)^{1/2} + (2d\kappa(\log n))^{-1/2} \left[ \frac{1}{2}(d - 1) \log \log n^\kappa \right. \]

\[ + \log \left\{ (2\pi)^{-1/2}H_2(2d)^{(d-1)/2} \right\}, \]

\[ H_2 = (2\pi\|K\|_2^2)^{-d/2} \det(\Sigma)^{1/2}, \quad \Sigma_{ij} = \int \frac{\partial K}{\partial u_i} \frac{\partial K}{\partial u_j} d\mathbf{u}, \]

\[ r(\mathbf{x}) = \sqrt{\frac{nh^d f_X(\mathbf{x})}{\tau(1 - \tau)}} f_\varepsilon|\mathbf{x}(0|\mathbf{x}), \]
Expectile regression notations

\[ h = n^{-\kappa}, \quad \rho_\tau(u) = |\tau - 1(u < 0)|u^2, \quad \varphi(u) = -2\{\tau - 1(u < 0)\}|u| \]

\[ d_n = (2d\kappa \log n)^{1/2} + (2d\kappa (\log n))^{-1/2} \left[ \frac{1}{2}(d - 1) \log \log n^\kappa \right. \]

\[ \left. + \log \left\{ (2\pi)^{-1/2} H_2 (2d)^{(d-1)/2} \right\} \right], \]

\[ H_2 = (2\pi \|K\|_2^2)^{-d/2} \det(\Sigma)^{1/2}, \quad \Sigma_{ij} = \int \frac{\partial K}{\partial u_i} \frac{\partial K}{\partial u_j} du, \]

\[ r(x) = \sqrt{\frac{nh^d f_X(x)}{\sigma^2(x)}} 2\left[\tau - F_{\varepsilon|x}(0|x)(2\tau - 1)\right], \]

\[ \sigma^2(x) = \mathbb{E}[\varphi^2(Y - \theta_0(x))|X = x]. \]
Approximations

\[ Y_n(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2(x)}} \int \int K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} dZ_n(y, u) \]

\[ Y_{0,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \int \Gamma_n K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} dZ_n(y, u) \]

\[ Y_{1,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \int \Gamma_n K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} dB_n(T(y, u)) \]

\[ Y_{2,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \int \Gamma_n K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(x) \} dW_n(T(y, u)) \]

\[ Y_{3,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \int \Gamma_n K \left( \frac{x - u}{h} \right) \psi \{ y - \theta_0(u) \} dW_n(T(y, u)) \]
Approximations

\[ Y_{4,n}(x) = \frac{1}{\sqrt{h^d f_X(x) \sigma^2_n(x)}} \int \sqrt{\sigma^2_n(u)f_X(u)K\left(\frac{x-u}{h}\right)} \, dW(u) \]

\[ Y_{5,n}(x) = \frac{1}{\sqrt{h^d}} \int K\left(\frac{x-u}{h}\right) \, dW(u) \]
Lemma (Bickel and Wichura (1971))

If \( \{X_n\}_{n=1}^{\infty} \) is a sequence in \( D[0,1]^d \), \( P(X \in [0,1]^d) = 1 \). For \( B, C \) neighboring blocks in \([0,1]^d\), constants \( \lambda_1 + \lambda_2 > 1, \gamma_1 + \gamma_2 > 0 \), \( \{X_n\}_{n=1}^{\infty} \) is tight if

\[
E[|X_n(B)|^{\gamma_1}|X_n(C)|^{\gamma_2}] \leq \mu(B)^{\lambda_1}\mu(C)^{\lambda_2},
\]

where \( \mu(\cdot) \) is a finite nonnegative measure on \([0,1]^d\) (for example, Lebesgue measure), and the increment of \( X_n \) on the block \( B \), denoted \( X_n(B) \), is defined by

\[
X_n(B) = \sum_{\alpha \in \{0,1\}^d} (-1)^{d-|\alpha|} X_n(s + \odot(t-s))
\]
Neighboring Blocks

Definition
A block $B \subset D$ is a subset of $D$ of the form $B = \Pi_i(s_i, t_i]$ with $s$ and $t$ in $D$; the $p$th-face of $B$ is $\Pi_{i \neq p}(s_i, t_i]$. Disjoint blocks $B$ and $C$ are $p$-neighbors if they abut and have the same $p$th face; they are neighbors if they are $p$-neighbors for some $p$ (for example, when $d = 3$, the blocks $(s, t] \times (a, b] \times (c, d]$ and $(t, u] \times (a, b] \times (c, d]$ are 1-neighbors for $s \leq t \leq u$).
Examples

- $d = 1$: $B = (s, t]$, $X_n(B) = X_n(t) - X_n(s)$;
- $d = 2$: $B = (s_1, t_1] \times (s_2, t_2]$. $X_n(B) = X_n(t_1, t_2) - X_n(t_1, s_2) + X_n(s_1, s_2) - X_n(s_1, t_2)$;
- For general $d$, $B = \prod_{i=1}^d (s_i, t_i]$, let $s = (s_1, \ldots, s_d)^\top$, $t = (t_1, \ldots, t_d)^\top$, then where $\odot$ denotes the vector of componentwise products.

Bickel & Wichura (1971)
Lemma (Meerschaert et al. (2013))

Suppose that $Y = \{Y(t), t \in \mathbb{R}^d\}$ is a centered Gaussian random field with values in $\mathbb{R}$, and denote

$$d(s, t) \overset{\text{def}}{=} d_Y(s, t) = (\mathbb{E}|Y(t) - Y(s)|^2)^{1/2}, \quad s, t \in \mathbb{R}^d.$$  

Let $D$ be a compact set contained in a cube with length $r$ in $\mathbb{R}^d$ and let $\sigma^2 = \sup_{t \in D} \mathbb{E}[Y(t)^2]$. For any $m > 0$, $\epsilon > 0$, define

$$\gamma(\epsilon) = \sup_{s, t \in D, \|s-t\| \leq \epsilon} d(s, t), \quad Q(m) = (2 + \sqrt{2}) \int_1^\infty \gamma(m2^{-y^2})dy.$$  

Then for all $a > 0$ which satisfy $a \geq (1 + 4d \log 2)^{1/2}(\sigma + a^{-1})$,

$$P\left\{\sup_{t \in S} |Y(t)| > a\right\} \leq 2^{2d+2} \left(\frac{r}{Q^{-1}(1/a) + 1}\right)^d \frac{\sigma + a^{-1}}{a} \exp\left\{-\frac{a^2}{2(\sigma + a^{-1})^2}\right\}$$

where $Q^{-1}(a) = \sup\{m : Q(m) \leq a\}$.  

Step 4  
Step 6
Generate Bivariate Uniform Samples

The bivariate samples \((X_1, X_2)\) are generated as follows:

1. Generate \(n\) pairs of bivariate normal variables \((Z_1, Z_2)\) with correlation \(\rho_N\) and variance 1
2. Transform the normal r.v.: \((X_1, X_2) = (\Phi(Z_1), \Phi(Z_2))\), where \(\Phi(\cdot)\) is the standard normal distribution function
3. Let \(\rho_U\) be the correlation of \((X_1, X_2)\), the following relation is true:

\[
\rho_U = \frac{6}{\pi} \arcsin \frac{\rho_N}{2}.
\]

Details: Falk (1999)
References

Bickel, P. J. & Rosenblatt, M.
On Some Global Measures of the Deviations of Density Function Estimates,

Bickel, P. J. & Rosenblatt, M.
Two-dimensional random fields,

Dedecker, J.; Merlevéde, F. & Rio, E.
Strong Approx. of the empi. distri. function for abs. regular seq. in $\mathbb{R}^d$
Delgado, M. A. & Escanciano, J. C.
Conditional Stochastic Dominance Testing

Efron, B.
Regression Percentiles using Asymmetric squared error loss,

Falk, M.
A simple approach to the generation of uniformly distributed random variables with prescribed correlations.

Guo, M. & Härdle, W.
Simultaneous confidence bands for expectile functions,
References

Härdle, W.
Asymptotic maximal deviation of $M$-smoothers

Härdle, W. & Song, S.
Confidence Bands in Quantile Regression,

Johnston, G. J.
Prob. of max. dev. for nonp. regression function estimates
J. Multivariate Anal, 1982, 12, 402-414

Kong, E.; Linton, O. Xia, Y. Uniform Bahadur representation for
local polynomial estimates of $M$-reg. and its application to the additive model
Econometric Theory, 2010, 26, 1529-1564

Li, Q., Lin, J. & Racine, J. S.
Optimal Bandwidth Selection for Nonp. Cond. Dist. and Quantile Functions,
References

Mas-Colell, A., Whinston, M. D.  Green, J. R.
Microeconomic Theory

Meerschaert, M. M., Wang, W. and Xiao, Y.
Fernique-Type Inequalities and Moduli of Continuity for Anisotropic Gaussian Random Fields
Transactions of the American Mathematical Society, 2013, 365, 1081-1107

Owen, A. B.
Multidim. Variation for quasi-Monte Carlo
References

Rosenblatt, M.
On the Max. Dev. of $k$-Dimensional Density Estimates,

Smirnov, N. V.
On the construction of confidence regions for the dens. of distri. of random variables,
Doklady Akad. Nauk SSSR, 1950, 74, 189-191 (Russian)