

Bootstrap-Assisted Inference for Generalized Grenander-type Estimators

Matias D. Cattaneo¹, Michael Jansson², and Kenichi Nagasawa³

April 2023

¹Princeton University.

²UC-Berkeley.

³University of Warwick.

Setup: Grenander-Type Estimators [Westling & Carone, 2020]

Monotone function estimators $\widehat{\theta}_n(x)$ at interior point x exhibit:

$$n^{q/(1+2q)}(\widehat{\theta}_n(x) - \theta_0(x)) \rightsquigarrow \frac{1}{\partial\Phi_0(x)}\partial_{-}\text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_x^q)(0),$$

$$\mathcal{M}_x^q(v) = \frac{\partial^q\theta_0(x)\partial\Phi_0(x)}{(1+q)!}v^{1+q}, \quad q = \min\{j \in \mathbb{N} : \partial^j\theta_0(x) \neq 0\}.$$

as $n \rightarrow \infty$, where:

- ▶ $\theta_0(\cdot)$ is density, regression or hazard function, among other possibilities.
- ▶ \mathcal{G}_x is zero-mean Gaussian process (nonstationary, rough cov kernel).
- ▶ \mathcal{M}_x^q is non-random drift function (usually quadratic, but not always).
- ▶ Φ_0 unknown non-decreasing and càdlàg function.
- ▶ $\text{GCM}_I(f)$ is greatest convex minorant of function f on interval I .

Motivation: Conducting inference can be challenging.

Setup: Grenander-Type Estimators [Westling & Carone, 2020]

$$n^{q/(1+2q)}(\widehat{\theta}_n(x) - \theta_0(x)) \rightsquigarrow \frac{1}{\partial\Phi_0(x)} \partial_{-} \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_x^q)(0),$$

$$\mathcal{M}_x^q(v) = \frac{\partial^q \theta_0(x) \partial \Phi_0(x)}{(1+q)!} v^{1+q}, \quad q = \min \{j \in \mathbb{N} : \partial^j \theta_0(x) \neq 0\}.$$

Examples:

- ▶ Isotonic density [Grenander, 1956] and extensions (censoring, covariates).
- ▶ Isotonic regression [Brunk, 1970] and extensions (censoring, covariates).
- ▶ Monotone Hazard [Rao, 1970] and extensions (censoring, covariates).
- ▶ Current Status [Ayer et al. 1955] and extensions (censoring, covariates).

Problem: bootstrap inconsistent [Kosorok, 2008; Sen, Banerjee & Woodroffe, 2010].

- ▶ Restore bootstrap validity: modifying the distribution used when resampling (subsampling, m -out-of- n bootstrap, smooth bootstrap).
- ▶ This paper: Modifying (“reshaping”) functional form estimator [CJN, 2020].

Leading Example (Today): Isotonic Density Estimation

Model:

- ▶ X_1, \dots, X_n i.i.d. with support $[0, 1]$.
- ▶ $F(x) = \mathbb{P}[X_i \leq x]$ absolutely continuous.
- ▶ $\partial F(x) = f(x)$ monotone (e.g., non-decreasing).

Estimand: $\theta_0(x) = f(x)$ for interior point x .

Estimator: for \mathcal{F} the class of non-decreasing densities supported on $[0, 1]$,

$$\hat{\theta}_n(\cdot) = \arg \max_{f \in \mathcal{F}} \sum_{i=1}^n \log f(X_i)$$

$$\implies \quad \hat{\theta}_n(x) = \partial_- \text{GCM}_{[0,1]}(\hat{\Gamma}_n)(x), \quad \hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$$

Asy Dist: under regularity, $q = 1$ and $\Phi_0(x) = x$,

$$\sqrt[3]{n}(\hat{\theta}_n(x) - \theta_0(x)) \rightsquigarrow \partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_{x,1})(0) \sim \sqrt[3]{4f(x)\partial f(x)} \underset{v \in \mathbb{R}}{\operatorname{argmin}} \{ \mathcal{G}(v) + v^2 \}$$

Second Example: Isotonic Regression Estimation

Model:

- ▶ $(Y_1, X_1), \dots, (Y_n, X_n)$ i.i.d. with X_i on support I , and $F(x) = \mathbb{P}[X_i \leq x]$.
- ▶ $\mu(x) = \mathbb{E}[Y_i | X_i = x]$ differentiable and monotone (e.g., non-decreasing).

Estimand: $\theta_0(x) = \mu(x)$ for interior point x .

Estimator: for \mathcal{F} the class of non-decreasing functions supported on I ,

$$\begin{aligned}\hat{\theta}_n(x) &= \operatorname{argmin}_{f \in \mathcal{F}} \sum_{i=1}^n (Y_{(i)} - f(X_{(i)}))^2 \\ \implies \hat{\theta}_n(x) &= \left(\partial_- \text{GCM}_{[0,1]}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-) \right) \circ \hat{\Phi}_n(x), \\ \hat{\Gamma}_n(x) &= \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x) Y_i, \quad \hat{\Phi}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x).\end{aligned}$$

Asy Dist: under regularity, $q = 1$ and $\Phi_0(x) = F(x)$,

$$\sqrt[3]{n}(\hat{\theta}_n(x) - \theta_0(x)) \rightsquigarrow \frac{1}{\partial \Phi_0(x)} \partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_{x,1})(0) \sim \sqrt[3]{\frac{4\sigma^2(x) \partial \mu(x)}{f(x)}} \operatorname{argmin}_{v \in \mathbb{R}} \{\mathcal{G}(v) + v^2\}$$

Third Example: Covariate-Adjusted Isotonic Regression

Model:

- ▶ $(Y_1, X_1, \mathbf{Z}'_1), \dots, (Y_n, X_n, \mathbf{Z}'_n)$ i.i.d. with X_i on support I , and $F(x) = \mathbb{P}[X_i \leq x]$.
- ▶ $\mu(x) = \mathbb{E}[\mathbb{E}[Y_i|X_i = x, \mathbf{Z}_i]]$ differentiable and monotone (e.g., non-decreasing).

Estimand: $\theta_0(\mathbf{x}) = \mathbb{E}[\mathbb{E}[Y_i|X_i = \mathbf{x}, \mathbf{Z}_i]]$.

Estimator:

$$\hat{\theta}_n(\mathbf{x}) = \left(\partial_- \text{GCM}_{[0,1]}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-) \right) \circ \hat{\Phi}_n(\mathbf{x})$$

$$\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x) \left[\frac{Y_i - \hat{\mu}(X_i, \mathbf{Z}_i)}{\hat{f}(\mathbf{Z}_i|X_i)} + \frac{1}{n} \sum_{j=1}^n \hat{\mu}(X_i, \mathbf{Z}_j) \right], \quad \hat{\Phi}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x).$$

Asy Dist: under regularity, $q = 1$ and $\Phi_0(x) = F(x)$,

$$\sqrt[3]{n}(\hat{\theta}_n(\mathbf{x}) - \theta_0(\mathbf{x})) \rightsquigarrow \frac{1}{\partial \Phi_0(\mathbf{x})} \partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_{\mathbf{x}} + \mathcal{M}_{\mathbf{x},1})(0) \sim c_0(\mathbf{x}) \underset{v \in \mathbb{R}}{\operatorname{argmin}} \{ \mathcal{G}(v) + v^2 \}$$

- ▶ Causal inference and program evaluation: Westling, Gilbert & Carone (2020, JRSSB).

Generalized Grenander-Type Estimators: Framework

Estimand: $\theta_0(\mathbf{x})$ at interior point \mathbf{x} , a monotone function.

Estimator:

$$\widehat{\theta}_n(\mathbf{x}) = [\partial_- \text{GCM}_{J_n}(\widehat{\Gamma}_n \circ \widehat{\Phi}_n^-)] \circ \widehat{\Phi}_n(\mathbf{x}),$$

for some (application specific) $\widehat{\Gamma}_n \rightarrow_{\mathbb{P}} \Gamma_0$ and $\widehat{\Phi}_n \rightarrow_{\mathbb{P}} \Phi_0$.

Asy Dist:

$$n^{\mathfrak{q}/(1+2\mathfrak{q})}(\widehat{\theta}_n(\mathbf{x}) - \theta_0(\mathbf{x})) \rightsquigarrow \frac{1}{\partial \Phi_0(\mathbf{x})} \partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_{\mathbf{x}} + \mathcal{M}_{\mathbf{x}}^{\mathfrak{q}})(0),$$

$$\mathcal{M}_{\mathbf{x}}^{\mathfrak{q}}(v) = \frac{\partial^{\mathfrak{q}} \theta_0(\mathbf{x}) \partial \Phi_0(x)}{(1 + \mathfrak{q})!} v^{1+\mathfrak{q}}, \quad \mathfrak{q} = \min \{j \in \mathbb{N} : \partial^j \theta_0(\mathbf{x}) \neq 0\}.$$

Our Goals:

- ▶ Develop valid bootstrap-assisted & automatic distributional approximations.
- ▶ Develop valid inference procedures: e.g., confidence intervals for $\theta_0(\mathbf{x})$.

Isotonic Density Estimation: Distribution Theory

- ▶ $\theta_0(x) = f(x)$, a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.

Switching Relation:

$$\mathbb{P} \left[\sqrt[3]{n} (\hat{\theta}_n(x) - \theta_0(x)) > t \right] = \mathbb{P} \left[\arg \max_{v \in S_n} \left\{ \hat{G}_{x,n}(v) + t \hat{L}_{x,n}(v) + M_{x,n}(v) \right\} < 0 \right],$$

where

$$\hat{G}_{x,n}(v) = -n^{2/3} \left[\hat{\Gamma}_n(x + vn^{-1/3}) - \Gamma_0(x + vn^{-1/3}) - \hat{\Gamma}_n(x) - \Gamma_0(x) \right]$$

$$\hat{L}_{x,n}(v) = v$$

$$M_{x,n}(v) = -n^{2/3} \left[\Gamma_0(x + vn^{-1/3}) - \Gamma_0(x) - \theta_0(x)vn^{-1/3} \right]$$

and

$$S_n = [-xn^{1/3}, \infty) \quad \text{with} \quad \mathbb{1}(v \in S_n) \rightarrow \mathbb{1}(v \in I)$$

Isotonic Density Estimation: Distribution Theory

- ▶ $\theta_0(x) = f(x)$, a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.

Switching Relation:

$$\mathbb{P} \left[\sqrt[3]{n}(\hat{\theta}_n(x) - \theta_0(x)) > t \right] = \mathbb{P} \left[\arg \max_{v \in S_n} \left\{ \hat{G}_{x,n}(v) + t\hat{L}_{x,n}(v) + M_{x,n}(v) \right\} < 0 \right],$$

where

$$\hat{G}_{x,n}(v) = -n^{2/3} \left[\hat{\Gamma}_n(x + vn^{-1/3}) - \Gamma_0(x + vn^{-1/3}) - \hat{\Gamma}_n(x) - \Gamma_0(x) \right] \rightsquigarrow \mathcal{G}_x(v)$$

$$\hat{L}_{x,n}(v) = v$$

$$M_{x,n}(v) = -n^{2/3} \left[\Gamma_0(x + vn^{-1/3}) - \Gamma_0(x) - \theta_0(x)vn^{-1/3} \right] \rightarrow -\mathcal{M}_x^q(v) = -\frac{\partial f(x)}{2}v^2$$

which implies

$$\arg \max_{v \in S_n} \left\{ \hat{G}_{x,n}(v) + t\hat{L}_{x,n}(v) + M_{x,n}(v) \right\} \rightsquigarrow \arg \max_{v \in I} \left\{ \mathcal{G}_x(v) + tv - \mathcal{M}_x^q(v) \right\}$$

Isotonic Density Estimation: Distribution Theory

$$\begin{aligned}\mathbb{P}\left[\sqrt[3]{n}(\widehat{\theta}_n(x) - \theta_0(x)) > t\right] &= \mathbb{P}\left[\arg\max_{v \in S_n} \left\{\widehat{G}_{x,n}(v) + t\widehat{L}_{x,n}(v) + M_{x,n}(v)\right\} < 0\right] \\ &\rightarrow \mathbb{P}\left[\arg\max_{v \in I} \left\{\mathcal{G}_x(v) + tv - \mathcal{M}_x^q(v)\right\} < 0\right] \\ &= \mathbb{P}\left[\partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_x^q)(0) > t\right]\end{aligned}$$

where

$$\mathcal{M}_x^q = \frac{\partial f(x)}{2} v^2$$

► In this stylized example,

$$q = 1, \quad \Phi_0(x) = x, \quad \mathcal{G}_x(v) = \sqrt{f(x)}\mathcal{W}(v),$$

with $\mathcal{W}(\cdot)$ two-sided Wiener process.

Next: let's investigate what happens when we apply the nonparametric bootstrap...

Isotonic Density Estimation: Bootstrapping

- ▶ $\theta_0(x) = f(x)$ a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.
- ▶ $\hat{\theta}_n^*(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n^*)(x)$ with $\hat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i^* \leq x)$.

Switching Relation:

$$\mathbb{P}^* \left[\sqrt[3]{n}(\hat{\theta}_n^*(x) - \hat{\theta}_n(x)) > t \right] = \mathbb{P}^* \left[\arg \max_{v \in S_n^*} \left\{ \hat{G}_{x,n}^*(v) + t \hat{L}_{x,n}^*(v) + \hat{M}_{x,n}^*(v) \right\} < 0 \right],$$

where

$$\hat{G}_{x,n}^*(v) = -n^{2/3} \left[\hat{\Gamma}_n^*(x + vn^{-1/3}) - \hat{\Gamma}_n(x + vn^{-1/3}) - \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x) \right]$$

$$\hat{L}_{x,n}^*(v) = v$$

$$\hat{M}_{x,n}^*(v) = -n^{2/3} \left[\hat{\Gamma}_n(x + vn^{-1/3}) - \hat{\Gamma}_n(x) - \hat{\theta}_n(x)vn^{-1/3} \right]$$

and

$$S_n^* = [-xn^{1/3}, \infty) \quad \text{with} \quad \mathbb{1}(v \in S_n^*) \rightarrow \mathbb{1}(v \in I)$$

Isotonic Density Estimation: Bootstrapping

- ▶ $\theta_0(x) = f(x)$ a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.
- ▶ $\hat{\theta}_n^*(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n^*)(x)$ with $\hat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i^* \leq x)$.

NP Bootstrap Failure:

$$\arg \max_{v \in S_n^*} \left\{ \hat{G}_{x,n}^*(v) + t \hat{L}_{x,n}^*(v) + M_{x,n}^*(v) \right\} \not\rightsquigarrow_{\mathbb{P}} \arg \max_{v \in I} \{ \mathcal{G}_{x,\alpha}(v) + tv + \mathcal{M}_{x,\alpha}(v) \}$$

because

$$\hat{G}_{x,n}^*(v) = -n^{2/3} [\hat{\Gamma}_n^*(x + vn^{-1/3}) - \hat{\Gamma}_n(x + vn^{-1/3}) - \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x)] \rightsquigarrow_{\mathbb{P}} \mathcal{G}_x(v)$$

$$\hat{L}_{x,n}^*(v) = v$$

$$M_{x,n}^*(v) = -n^{2/3} [\hat{\Gamma}_n(x + vn^{-1/3}) - \hat{\Gamma}_n(x) - \hat{\theta}_n(x)vn^{-1/3}] \not\rightsquigarrow_{\mathbb{P}} -\mathcal{M}_{x,n}(v) = -\frac{\partial f(x)}{2}v^2$$

- ▶ **Recall:** for Asy Dist, we instead had

$$M_{x,n}(v) = -n^{2/3} [\Gamma_0(x + vn^{-1/3}) - \Gamma_0(x) - \theta_0(x)vn^{-1/3}] \rightarrow -\mathcal{M}_{x,n}(v) = -\frac{\partial f(x)}{2}v^2$$

Isotonic Density Estimation: Recap and Intuition

- ▶ $\theta_0(x) = f(x)$ a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.

Asymptotic Distribution:

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}[\sqrt[3]{n}(\hat{\theta}_n(x) - \theta_0(x)) \leq t] - \mathbb{P}[\partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_x^q)(0) \leq t] \right| \rightarrow 0$$

NP bootstrap: $\hat{\theta}_n^*(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n^*)(x)$ with $\hat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i^* \leq x)$ is invalid,

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}^*[\sqrt[3]{n}(\hat{\theta}_n^*(x) - \hat{\theta}_n(x)) \leq t] - \mathbb{P}[\sqrt[3]{n}(\hat{\theta}_n(x) - \theta_0(x)) \leq t] \right| \not\rightarrow_{\mathbb{P}} 0$$

This paper: consistency can be achieved by reshaping $\hat{\Gamma}_n^*$.

- ▶ Intuition:
 - ▶ around x , $\hat{\Gamma}_n(x)$ has mean $\Gamma_0(x) \approx \Gamma_0(x) + f(x)(x - x) + \frac{1}{2} \partial f(x)(x - x)^2$
 - ▶ whereas the mean of $\hat{\Gamma}_n^*(x)$ under the bootstrap distribution is given by $\hat{\Gamma}_n(x)$.
- ▶ Reshaping: Let $\partial \tilde{f}_n(x)$ denote a consistent estimator of $\partial f(x)$, then

$$\tilde{\Gamma}_n^*(x) = \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x) + \hat{\theta}_n(x)(x - x) + \frac{1}{2} \partial \tilde{f}_n(x)(x - x)^2$$

Isotonic Density Estimation: Bootstrap Consistency

- ▶ $\theta_0(x) = f(x)$ a monotone density with $\Gamma_0(x) = \mathbb{P}(X_i \leq x)$.
- ▶ $\hat{\theta}_n(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n)(x)$ with $\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$.
- ▶ $\hat{\theta}_n^*(x) = \partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n^*)(x)$ with $\hat{\Gamma}_n^*(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i^* \leq x)$.
- ▶ $\tilde{\theta}_n^*(x) = \partial_- \text{GCM}_{J_n}(\tilde{\Gamma}_n^*)(x)$ with

$$\tilde{\Gamma}_n^*(x) = \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x) + \hat{\theta}_n(x)(x - x) + \frac{1}{2} \partial \tilde{f}_n(x)(x - x)^2.$$

Bootstrap-Assisted Validity:

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}^*[\tilde{\theta}_n^*(x) - \hat{\theta}_n(x) \leq t] - \mathbb{P}[\hat{\theta}_n(x) - \theta_0(x) \leq t] \right| \xrightarrow{\mathbb{P}} 0$$

provided that

$$\partial \tilde{f}_n(x) \rightarrow_{\mathbb{P}} \partial f(x).$$

- ▶ The consistency requirement is mild and easy to achieve automatically.
- ▶ Confidence intervals based on kernel estimator perform well in simulations.

Generalized Grenander-Type Estimators: General Case

- ▶ $\theta_0(x)$ at interior point x , a monotone function.
- ▶ $\hat{\theta}_n(x) = [\partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)] \circ \hat{\Phi}_n(x)$, with $(\hat{\Gamma}_n, \hat{\Phi}_n) \rightarrow_{\mathbb{P}} (\Gamma_0, \Phi_0)$.

Asymptotic Distribution: using their Generalized Switching Relation,

$$n^{q/(1+2q)} (\hat{\theta}_n(x) - \theta_0(x)) \rightsquigarrow \frac{1}{\partial \Phi_0(x)} \partial_- \text{GCM}_{\mathbb{R}}(\mathcal{G}_x + \mathcal{M}_x^q)(0),$$

$$\mathcal{M}_x^q(v) = \frac{\partial^q \theta_0(x) \partial \Phi_0(x)}{(1+q)!} v^{1+q}, \quad q = \min \{j \in \mathbb{N} : \partial^j \theta_0(x) \neq 0\}.$$

Bootstrap-assisted Inference using reshaped estimator:

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}^*[\tilde{\theta}_n^*(x) - \hat{\theta}_n(x) \leq t] - \mathbb{P}[\hat{\theta}_n(x) - \theta_0(x) \leq t] \right| \xrightarrow{\mathbb{P}} 0$$

where

$$\tilde{\theta}_n^*(x) = [\partial_- \text{GCM}_{J_n}(\tilde{\Gamma}_n^* \circ (\hat{\Phi}_n^*)^-)] \circ \hat{\Phi}_n^*(x)$$

with

$$\tilde{\Gamma}_n^*(x) = \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x) + \hat{\theta}_n(x) \hat{\Phi}_n(x - x) + \tilde{\mathcal{M}}_{x,q}(x - x),$$

and provided that

$$\tilde{\mathcal{M}}_{x,q}(v) \rightarrow_{\mathbb{P}} \mathcal{M}_x^q(v).$$

Grenander-Type Estimators: Bootstrap-assisted Inference

Valid percentile CI for monotone function $\theta_0(x)$ at interior point x :

$$I_{1-a}^*(x) = \left[\hat{\theta}_n(x) - q_{1-a/2}^* \quad , \quad \hat{\theta}_n(x) - q_{a/2}^* \right]$$

where

$$\blacktriangleright \hat{\theta}_n(x) = [\partial_- \text{GCM}_{J_n}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)] \circ \hat{\Phi}_n(x), \quad (\hat{\Gamma}_n, \hat{\Phi}_n) \rightarrow_{\mathbb{P}} (\Gamma_0, \Phi_0).$$

$$\blacktriangleright q_a^* = \inf\{t \in \mathbb{R} : \mathbb{P}^*[\tilde{\theta}_n^*(x) - \hat{\theta}_n(x) \leq t] \geq a\},$$

$$\tilde{\theta}_n^*(x) = [\partial_- \text{GCM}_{J_n}(\tilde{\Gamma}_n^* \circ (\hat{\Phi}_n^*)^-)] \circ \hat{\Phi}_n^*(x),$$

$$\tilde{\Gamma}_n^*(x) = \hat{\Gamma}_n^*(x) - \hat{\Gamma}_n(x) + \hat{\theta}_n(x)\hat{\Phi}_n(x - x) + \tilde{\mathcal{M}}_{x,q}(x - x),$$

$$\tilde{\mathcal{M}}_{x,q}(v) \rightarrow_{\mathbb{P}} \mathcal{M}_x^q(v).$$

Key outstanding issue: How to construct $\tilde{\mathcal{M}}_{x,q}(v)$?

Bootstrap-assisted Inference: Drift Estimation

Recall: $\widehat{\theta}_n(x) = [\partial_- \text{GCM}_{J_n}(\widehat{\Gamma}_n \circ \widehat{\Phi}_n^-)] \circ \widehat{\Phi}_n(x)$, $(\widehat{\Gamma}_n, \widehat{\Phi}_n) \rightarrow_{\mathbb{P}} (\Gamma_0, \Phi_0)$,

$$\mathcal{M}_x^q(v) = \frac{\partial^q \theta_0(x) \partial \Phi_0(x)}{(1+q)!} v^{1+q}, \quad q = \min \{j \in \mathbb{N} : \partial^j \theta_0(x) \neq 0\}.$$

- ▶ Sometimes, $\partial^q \theta_0(x)$ and $\partial \Phi_0(x)$ easy to characterize (if q is known!) and estimate.
 - ▶ For example, $\pi_0(x) = \partial f(x)$ and $\partial \Phi_0(x) = 1$ in Isotonic Density Estimation.
- ▶ In general, if q is known, $\tilde{\mathcal{M}}_{x,q}(v)$ based on numerical derivative estimator:

$$\frac{\partial^q \theta_n(\widehat{x}) \partial \widehat{\Phi}_n(x)}{(q+1)!} = \epsilon_n^{-(q+1)} \sum_{k=1}^{q+1} (-1)^{k+q+1} \binom{q+1}{k} [\widehat{\Upsilon}_n(x + k\epsilon_n) - \widehat{\Upsilon}_n(x)]$$

where

- ▶ $\widehat{\Upsilon}_n(v) = \widehat{\Gamma}_n(v) - \widehat{\theta}_n(x) \widehat{\Phi}_n(v)$ and $\epsilon_n > 0$ is a (small) tuning parameter.
- ▶ $\tilde{\mathcal{M}}_{x,q}(v) \rightarrow_{\mathbb{P}} \mathcal{M}_x^q(v)$ requires $\epsilon_n \rightarrow 0$ and $n\epsilon_n^{1+2q} \rightarrow \infty$.
- ▶ Under additional conditions, MSE-optimal ϵ_n can be obtained.
- ▶ Possible to develop estimator $\tilde{\mathcal{M}}_{x,q}(v)$ adaptive to unknown $q \leq \bar{q}$.

	DGP 1				DGP 2				DGP 3			
	$\hat{\mathcal{D}}_{1,n}$	$\hat{\mathcal{D}}_{3,n}$	Coverage	Length	$\hat{\mathcal{D}}_{1,n}$	$\hat{\mathcal{D}}_{3,n}$	Coverage	Length	$\hat{\mathcal{D}}_{1,n}$	$\hat{\mathcal{D}}_{3,n}$	Coverage	Length
Standard			0.832	0.370			0.835	0.516			0.904	0.028
m-out-of-n												
$m = \lceil n^{1/2} \rceil$			0.900	0.413			0.909	0.583			0.940	0.031
$m = \lceil n^{2/3} \rceil$			0.872	0.399			0.879	0.556			0.921	0.029
$m = \lceil n^{4/5} \rceil$			0.856	0.391			0.862	0.544			0.913	0.029
Reshaped												
Oracle	1.000	0.000	0.942	0.393	1.00	0.000	0.943	0.549	0.000	1.000	0.943	0.029
ND known q	1.045	0.000	0.950	0.396	1.04	0.000	0.944	0.543	0.000	1.012	0.935	0.028
ND robust	1.045	0.633	0.951	0.398	1.04	0.981	0.953	0.556	0.014	1.012	0.959	0.030

Discussion and Conclusion

- ▶ Nonparametric bootstrap fails for Grenander-Type Estimators.
- ▶ Other valid resampling methods available change the bootstrap distribution.
- ▶ This paper:
 - ▶ Employs standard nonparametric bootstrap.
 - ▶ Reshapes estimator to deal with bootstrap inconsistency.
- ▶ Our method applies to many problems in econ, stats, biostats, ML and beyond:
 - ▶ Isotonic density [Grenander, 1956] and extensions (censoring, covariates).
 - ▶ Isotonic regression [Brunk, 1970] and extensions (censoring, covariates).
 - ▶ Monotone Hazard [Rao, 1970] and extensions (censoring, covariates).
 - ▶ Current Status [Ayer et al. 1955] and extensions (censoring, covariates).
- ▶ **Coming soon:** Smoothed pairwise maximum rank correlation and related problems.
 - ▶ \sqrt{n} -consistent U-process optimizer but with Chernoff-type distribution