Bootstrap-Assisted Inference for Generalized Grenander-type Estimators

Matias D. Cattaneo, Princeton University

Michael Jansson, UC-Berkeley and CREATES

Kenichi Nagasawa, University of Warwick

June 2021

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

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

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

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

as $n \to \infty$, where:

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

Motivation: Conducting inference can be challenging.

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

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

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

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: NP bootstrap inconsistent [Kosorok, 2008; Sen, Banerjee & Woodroofe, 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:

- $X_1, \ldots, X_n$ 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: $\theta_0(\mathsf{x}) = f(\mathsf{x})$ for interior point x.

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

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

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

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

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

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: $\theta_0(\mathsf{x}) = \mu(\mathsf{x})$ for interior point x.

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

$$\hat{\theta}_n(x) = \operatorname*{argmin}_{f \in \mathcal{F}} \sum_{i=1}^{n} (Y_{(i)} - f(X_{(i)}))^2$$

$$\implies \qquad \hat{\theta}_n(\mathsf{x}) = \Big(\partial_- \mathrm{GCM}_{[0,1]}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)\Big) \circ \hat{\Phi}_n(\mathsf{x}),$$

$$\hat{\Gamma}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i \leq x) Y_i, \qquad \hat{\Phi}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i \leq x).$$

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

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

Third Example: Covariate-Adjusted Isotonic Regression

Model:

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

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

Estimator:

$$\hat{\theta}_n(\mathsf{x}) = \Big(\partial_- \mathrm{GCM}_{[0,1]}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)\Big) \circ \hat{\Phi}_n(\mathsf{x})$$

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

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

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

Generalized Grenander-Type Estimators: Framework

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

Estimator:

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

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

Asy Dist:

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

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

Our Goals:

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

Isotonic Density Estimation: Distribution Theory

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

Switching Lemma:

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

where

$$\begin{aligned}
\hat{G}_{\mathsf{x},n}(v) &= -n^{2/3}[\hat{\Gamma}_n(\mathsf{x} + vn^{-1/3}) - \Gamma_0(\mathsf{x} + vn^{-1/3}) - \hat{\Gamma}_n(\mathsf{x}) - \Gamma_0(\mathsf{x})] \\
\hat{L}_{\mathsf{x},n}(v) &= v \\
M_{\mathsf{x},n}(v) &= -n^{2/3}[\Gamma_0(\mathsf{x} + vn^{-1/3}) - \Gamma_0(\mathsf{x}) - \theta_0(\mathsf{x})vn^{-1/3}]
\end{aligned}$$

and

$$S_n = [-\mathsf{x}n^{1/3}, \infty) \qquad \text{with} \qquad \mathbb{1}(v \in S_n) \to \mathbb{1}(v \in I)$$

Isotonic Density Estimation: Distribution Theory

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

Switching Lemma:

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

where

$$\begin{aligned}
\hat{G}_{\mathsf{x},n}(v) &= -n^{2/3}[\hat{\Gamma}_n(\mathsf{x}+vn^{-1/3}) - \Gamma_0(\mathsf{x}+vn^{-1/3}) - \hat{\Gamma}_n(\mathsf{x}) - \Gamma_0(\mathsf{x})] \rightsquigarrow \mathcal{G}_{\mathsf{x}}(v) \\
\hat{L}_{\mathsf{x},n}(v) &= v \\
M_{\mathsf{x},n}(v) &= -n^{2/3}[\Gamma_0(\mathsf{x}+vn^{-1/3}) - \Gamma_0(\mathsf{x}) - \theta_0(\mathsf{x})vn^{-1/3}] \to -\mathcal{M}_{\mathsf{x},\mathfrak{q}}(v) = -\frac{\partial f(\mathsf{x})}{2}v^2
\end{aligned}$$

which implies

$$\underset{v \in S_n}{\operatorname{argmax}}\left\{\hat{G}_{\mathsf{x},n}(v) + t\hat{L}_{\mathsf{x},n}(v) + M_{\mathsf{x},n}(v)\right\} \rightsquigarrow \underset{v \in I}{\operatorname{argmax}}\left\{\mathcal{G}_{\mathsf{x}}(v) + tv - \mathcal{M}_{\mathsf{x},\mathfrak{q}}(v)\right\}$$

Isotonic Density Estimation: Distribution Theory

$$\begin{aligned}
\mathbb{P}\left[\sqrt[3]{n}(\hat{\theta}_n(\mathsf{x}) - \theta_0(\mathsf{x})) > t\right] &= \mathbb{P}\left[\operatorname*{argmax}_{v \in S_n} \left\{\hat{G}_{\mathsf{x},n}(v) + t\hat{L}_{\mathsf{x},n}(v) + M_{\mathsf{x},n}(v)\right\} < 0\right] \\
&\to \mathbb{P}\left[\operatorname*{argmax}_{v \in I} \{\mathcal{G}_{\mathsf{x}}(v) + tv - \mathcal{M}_{\mathsf{x},\mathfrak{q}}(v)\} < 0\right] \\
&= \mathbb{P}\left[\partial_- \mathrm{GCM}_{\mathbb{R}}(\mathcal{G}_{\mathsf{x}} + \mathcal{M}_{\mathsf{x},\mathfrak{q}})(0) > t\right]
\end{aligned}$$

where

$$\mathcal{M}_{\mathsf{x},\mathfrak{q}} = \frac{\partial f(\mathsf{x})}{2} v^2$$

- In this stylized example,

  $$\mathfrak{q} = 1, \qquad \Phi_0(x) = x, \qquad \mathcal{G}_{\mathsf{x}}(v) = \sqrt{f(x)}\mathcal{W}(v),$$

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

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

Isotonic Density Estimation: Bootstrapping

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

Switching Lemma:

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

where

$$\begin{aligned}
\hat{G}_{\mathsf{x},n}^*(v) &= -n^{2/3}[\hat{\Gamma}_n^*(\mathsf{x}+vn^{-1/3}) - \hat{\Gamma}_n(\mathsf{x}+vn^{-1/3}) - \hat{\Gamma}_n^*(\mathsf{x}) - \hat{\Gamma}_n(\mathsf{x})] \\
\hat{L}_{\mathsf{x},n}^*(v) &= v \\
M_{\mathsf{x},n}^*(v) &= -n^{2/3}[\hat{\Gamma}_n(\mathsf{x}+vn^{-1/3}) - \hat{\Gamma}_n(\mathsf{x}) - \hat{\theta}_n(\mathsf{x})vn^{-1/3}]
\end{aligned}$$

and

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

Isotonic Density Estimation: Bootstrapping

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

NP Bootsrtap Failure:

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

because

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

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

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

Isotonic Density Estimation: Recap and Intuition

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

Asymptotic Distribution:

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

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

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

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

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

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

Isotonic Density Estimation: Bootstrap Consistency

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

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

Bootstrap-Assisted Validity:

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

provided that

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

- 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

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

Asymptotic Distribution: using their Generalized Switching Lemma,

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

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

Bootstrap-assisted Inference using reshaped estimator:

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

where

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

with

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

and provided that

$$\tilde{\mathcal{M}}_{\mathsf{x},\mathfrak{q}}(v) \to_{\mathbb{P}} \mathcal{M}_{\mathsf{x},\mathfrak{q}}(v).$$

Grenander-Type Estimators: Bootstrap-assisted Inference

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

$$\mathrm{I}^*_{1-a}(\mathsf{x}) = \left[\hat{\theta}_n(\mathsf{x}) - q^*_{1-a/2} \quad , \quad \hat{\theta}_n(\mathsf{x}) - q^*_{a/2}\right]$$

where

- $\hat{\theta}_n(\mathsf{x}) = [\partial_- \mathrm{GCM}_{J_n}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)] \circ \hat{\Phi}_n(\mathsf{x}), \qquad (\hat{\Gamma}_n, \hat{\Phi}_n) \to_{\mathbb{P}} (\Gamma_0, \Phi_0).$

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

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

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

$$\tilde{\mathcal{M}}_{\mathsf{x},\mathfrak{q}}(v) \to_{\mathbb{P}} \mathcal{M}_{\mathsf{x},\mathfrak{q}}(v).$$

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

Bootstrap-assisted Inference: Drift Estimation

Recall: $\hat{\theta}_n(\mathsf{x}) = [\partial_- \mathrm{GCM}_{J_n}(\hat{\Gamma}_n \circ \hat{\Phi}_n^-)] \circ \hat{\Phi}_n(\mathsf{x})$, $\qquad (\hat{\Gamma}_n, \hat{\Phi}_n) \to_{\mathbb{P}} (\Gamma_0, \Phi_0)$,

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

- Sometimes, $\partial^{\mathfrak{q}}\theta_0(\mathsf{x})$ and $\partial\Phi_0(\mathsf{x})$ easy to caracterize (if $\mathfrak{q}$ is known!) and estimate.
  - For example, $\pi_0(\mathsf{x}) = \partial f(\mathsf{x})$ and $\partial\Phi_0(\mathsf{x}) = 1$ in Isotonic Density Estimation.

- In general, if $\mathfrak{q}$ is known, $\tilde{\mathcal{M}}_{x,\mathfrak{q}}(v)$ based on numerical derivative estimator:

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

  where

  - $\hat{\Upsilon}_n(v) = \hat{\Gamma}_n(v) - \hat{\theta}_n(\mathsf{x})\hat{\Phi}_n(v)$ and $\epsilon_n > 0$ is a (small) tuning parameter.
  - $\tilde{\mathcal{M}}_{\mathsf{x},\mathfrak{q}}(v) \to_{\mathbb{P}} \mathcal{M}_{\mathsf{x},\mathfrak{q}}(v)$ requires $\epsilon_n \to 0$ and $n\epsilon_n^{1+2\mathfrak{q}} \to \infty$.
  - Under additional conditions, MSE-optimal ϵ_n can be obtained.

- Possible to develop estimator $\tilde{\mathcal{M}}_{\mathsf{x},\mathfrak{q}}(v)$ adaptive to unknown $\mathfrak{q} \leq \bar{\mathfrak{q}}$.

	DGP 1			DGP 2			DGP 3		
	h, ϵ	Coverage	Length	h, ϵ	Coverage	Length	h, ϵ	Coverage	Length
Standard									
		0.828	0.146		0.808	0.172		0.821	0.155
m-out-of-n									
$m = \lceil n^{1/2} \rceil$		1.000	0.438		0.995	0.495		0.998	0.452
$m = \lceil n^{2/3} \rceil$		0.989	0.314		0.979	0.360		0.989	0.328
$m = \lceil n^{4/5} \rceil$		0.953	0.235		0.937	0.274		0.948	0.248
PI: $\tilde{M}_n^{\mathtt{PI}}$									
$0.7 \cdot h_{\mathtt{MSE}}$	0.264	0.955	0.157	0.202	0.947	0.183	0.209	0.957	0.165
$0.8 \cdot h_{\mathtt{MSE}}$	0.302	0.954	0.157	0.231	0.946	0.182	0.239	0.952	0.165
$0.9 \cdot h_{\mathtt{MSE}}$	0.339	0.951	0.156	0.260	0.944	0.181	0.269	0.949	0.164
$1.0 \cdot h_{\mathtt{MSE}}$	0.377	0.949	0.154	0.289	0.941	0.180	0.299	0.948	0.163
$1.1 \cdot h_{\mathtt{MSE}}$	0.415	0.940	0.151	0.318	0.938	0.178	0.329	0.944	0.161
$1.2 \cdot h_{\mathtt{MSE}}$	0.452	0.934	0.147	0.347	0.934	0.176	0.359	0.939	0.158
$1.3 \cdot h_{\mathtt{MSE}}$	0.490	0.922	0.142	0.376	0.928	0.173	0.389	0.935	0.155
$h_{\mathtt{AMSE}}$	0.380	0.949	0.154	0.300	0.940	0.180	0.333	0.943	0.161
$\hat{h}_{\mathtt{AMSE}}$	0.364	0.950	0.155	0.290	0.941	0.180	0.401	0.930	0.154

	DGP 1			DGP 2			DGP 3		
	h, ϵ	Coverage	Length	h, ϵ	Coverage	Length	h, ϵ	Coverage	Length
Standard									
		0.828	0.146		0.808	0.172		0.821	0.155
m-out-of-n									
$m = \lceil n^{1/2} \rceil$		1.000	0.438		0.995	0.495		0.998	0.452
$m = \lceil n^{2/3} \rceil$		0.989	0.314		0.979	0.360		0.989	0.328
$m = \lceil n^{4/5} \rceil$		0.953	0.235		0.937	0.274		0.948	0.248
ND: $\tilde{M}_n^{\mathtt{ND}}$									
$0.7 \cdot \epsilon_{\mathtt{MSE}}$	0.726	0.954	0.158	0.527	0.947	0.183	0.554	0.952	0.165
$0.8 \cdot \epsilon_{\mathtt{MSE}}$	0.830	0.956	0.159	0.602	0.947	0.182	0.633	0.950	0.164
$0.9 \cdot \epsilon_{\mathtt{MSE}}$	0.933	0.956	0.160	0.678	0.944	0.181	0.712	0.949	0.163
$1.0 \cdot \epsilon_{\mathtt{MSE}}$	1.037	0.956	0.159	0.753	0.942	0.180	0.791	0.948	0.162
$1.1 \cdot \epsilon_{\mathtt{MSE}}$	1.141	0.955	0.159	0.828	0.940	0.179	0.870	0.946	0.161
$1.2 \cdot \epsilon_{\mathtt{MSE}}$	1.244	0.956	0.160	0.904	0.936	0.177	0.949	0.943	0.160
$1.3 \cdot \epsilon_{\mathtt{MSE}}$	1.348	0.960	0.163	0.979	0.935	0.176	1.028	0.940	0.159
$\epsilon_{\mathtt{AMSE}}$	0.927	0.956	0.160	0.731	0.943	0.180	0.812	0.948	0.162
$\hat{\epsilon}_{\mathtt{AMSE}}$	0.888	0.956	0.159	0.708	0.943	0.181	0.978	0.942	0.159

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