Robust Inference for Partially Observed Functional Response Data

Yeonjoo Park[1], Xiaohui Chen[2], and Douglas G. Simpson[2]

[1] *University of Texas at San Antonio and*
[2] *University of Illinois at Urbana-Champaign*

Supplementary Material

This supplement contains two sections. Section S1 includes technical proofs of propositions, lemmas, and theorems, and Section S2 presents figures and detailed results from simulation studies.

S1 Appendix: Proofs

Proof of Proposition 1. Under M1, M4, A1-A3, we write the pdf of the marginal distribution of Y at $t \in \mathrm{C}$ as $f(y)$ and it is assumed to be symmetric (or even function) about $\alpha(t)$. Then

$$E_{P_Y}[\psi\big(Y(t)-\alpha(t)\big)] = \int_{-\infty}^{\infty} \psi\big(Y(t)-\alpha(t)\big) f\big(Y(t)-\alpha(t)\big) dy = 0, \quad t \in \mathrm{C},$$

under the assumption of odd function $\psi(\cdot)$. Thus, $\theta(t) = \alpha(t)$. $\square$

Proof of Proposition 2. Under A1-A2, equation (2.3) implies that $E_{P_Y}\big[\rho(Y(t)-\alpha(t)+\alpha(t)-\theta(t))\big]$ equals specific value at each $t \in \mathrm{C}$, say $c_1(t)$. Under B1, the marginal distribution of $Y(t)-\alpha(t)$, $t \in \mathrm{C}$, does not depend on t with the probability measure P_Z. Then we can equivalently write

$$E_{P_Z}\big[\rho(Z - \{\alpha(t)-\theta(t)\})\big] = c_1,$$

and $\{\alpha(t)-\theta(t)\} = c_1 + c_2$, where constant c_2 is determined by P_Z. Let $c = c_1 + c_2$ then we can write $\theta(t) = \alpha(t) + c$. $\square$

Proof of Theorem 1. Denote $T(P_\varepsilon)(t)$ by $\theta_\varepsilon(t)$. Under D2, for any $\upsilon > 0$, there exists $\delta > 0$,

$$\begin{aligned}
P(\sup_{t\in\mathrm{C}} |\theta_\varepsilon(t)-\theta(t)| > \upsilon) &\leqslant P(\sup_{t\in\mathrm{C}} \big[M(t,\theta_\varepsilon,P) - M(t,\theta,P)\big] > \delta) \\
&\leqslant P(\sup_{t\in\mathrm{C}} \big[M(t,\theta_\varepsilon,P) - M(t,\theta_\varepsilon,P_\varepsilon) + M(t,\theta,P_\varepsilon) - M(t,\theta,P)\big] > \delta) \\
&\leqslant P(\sup_{t\in\mathrm{C}} |M(t,\theta_\varepsilon,P) - M(t,\theta_\varepsilon,P_\varepsilon)| > \delta/2) \\
&\qquad + P(\sup_{t\in\mathrm{C}} |M(t,\theta,P_\varepsilon) - M(t,\theta,P)| > \delta/2).
\end{aligned}$$

By D1, $T(P_\varepsilon)(t)$ is uniformly continuous as $\varepsilon \to 0$. $\square$

Proof of Theorem 2. By the estimating equation of (2.4),

$$\begin{aligned}
0 &= (1-\varepsilon)E_P\big[\delta(t)\psi(Y(t),\theta_\varepsilon(t))\big] + \varepsilon\delta^*(t)\psi(Y^*(t),\theta_\varepsilon(t)) \\
&= (1-\varepsilon)E_P\big[\delta(t)\{\psi(Y(t),\theta_\varepsilon(t)) - \psi(Y(t),\theta(t))\}\big] + \varepsilon\delta^*(t)\psi(Y^*(t),\theta_\varepsilon(t)) \\
&= (1-\varepsilon)E_P\big[\delta(t)\frac{\psi(Y(t),\theta_\varepsilon(t)) - \psi(Y(t),\theta(t))}{\varepsilon}\big] + \delta^*(t)\psi(Y^*(t),\theta_\varepsilon(t))
\end{aligned}$$

Let $\varepsilon \to 0$, then

$$0 = E_P\big[\delta(t)\dot{\psi}(Y(t),\theta(t))\big]\dot{\theta}(t) + \delta^*(t)\psi(Y^*(t),\theta(t)).$$

Thus,

$$\dot{\theta}(t) = IF_T(Y^*,\delta^*)(t) = \frac{\delta^*(t)\psi(Y^*(t)-\theta(t))}{-E_P\big[\delta(t)\dot{\psi}(Y(t),\theta(t))\big]},$$

and the bounded $\psi(\cdot)$ implies $\gamma_T^\infty < \infty$. $\square$

Proof of Lemma 1. Under the sampling scheme condition M2, we can define the empirical process

$$\mathbb{G}_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^n [h(t,V_i) - \mathbb{E}\, h(t,V_i)], \quad t \in \mathrm{C},$$

where $V_1,\ldots,V_n$ are i.i.d. random variables in $\mathcal{V}$ with common distribution f. Alternatively, we may write

$$\mathbb{G}_n(g) = \frac{1}{\sqrt{n}}\sum_{i=1}^n [g(V_i) - \mathbb{E}\, g(V_i)], \quad g \in \mathcal{G}$$

with the identification of g by h_t for a given missing scheme h. Then

$$W_n = \frac{1}{\sqrt{n}}\sup_{t\in \mathrm{C}} \mathbb{G}_n(t) = \frac{1}{\sqrt{n}}\sup_{g\in\mathcal{G}} \mathbb{G}_n(g).$$

Recall that $H : \mathcal{V} \to \{0,1\}$ is a measurable envelope for $\mathcal{G}$. Set $M = \max_{1\leqslant i\leqslant n} H(V_i)$. By the local maximal inequality Chernozhukov, Chetverikov, and Kato (2014) with the locality parameter $\delta = 1$, there is a universal constant $C > 0$ such that

$$E[\sup_{g\in\mathcal{G}} \mathbb{G}_n(g)] \leqslant C\Big\{J(1,\mathcal{G},H)\|H\|_{f,2} + \frac{\|M\|_2 J^2(1,\mathcal{G},H)}{\sqrt{n}}\Big\}.$$

Since $|H| \leqslant 1$ and $\|M\|_2 \leqslant 1$, we get

$$E[\sup_{g\in\mathcal{G}} \mathbb{G}_n(g)] \leqslant C\Big\{J(1,\mathcal{G},H) + \frac{J^2(1,\mathcal{G},H)}{\sqrt{n}}\Big\}.$$

Then it is immediate that

$$E[W_n] \leqslant C\Big\{\frac{J(1,\mathcal{G},H)}{\sqrt{n}} + \frac{J^2(1,\mathcal{G},H)}{n}\Big\} \leqslant C\frac{J(1,\mathcal{G},H)}{\sqrt{n}} \max\Big\{1, \frac{J(1,\mathcal{G},H)}{\sqrt{n}}\Big\}.$$

$\square$

Proof of Theorem 3. For $t \in \mathrm{C}$, if $|\hat{\theta}_n(t) - \theta(t)| > \epsilon$, then $M(t,\hat{\theta}_n,P) - M(t,\theta,P) > \delta_t$ by D2, and $\sup_t \big[M(t,\hat{\theta}_n,P) - M(t,\theta,P)\big] > \delta$, where $\delta = \sup_t \delta_t$. Then

$$
\begin{aligned}
P(\sup_{t\in \mathrm{C}} |\hat\theta_n(t) - \theta(t)| > \epsilon) &\leqslant P(\sup_{t\in \mathrm{C}} \big[M(t,\hat\theta_n,P) - M(t,\theta,P)\big] > \delta) \\
&= P(\sup_{t\in \mathrm{C}} \big[M(t,\hat\theta_n,P) - M(t,\hat\theta_n,P_n) + M(t,\hat\theta_n,P_n) - M(t,\theta,P_n) \\
&\qquad + M(t,\theta,P_n) - M(t,\theta,P)\big]) > \delta) \\
&\leqslant P(\sup_{t\in \mathrm{C}} \big[M(t,\hat\theta_n,P) - M(t,\hat\theta_n,P_n) + M(t,\theta,P_n) - M(t,\theta,P)\big] > \delta) \\
&\leqslant P(\sup_{t\in \mathrm{C}} |M(t,\hat\theta_n,P) - M(t,\hat\theta_n,P_n)| > \delta/2) \\
&\qquad + p(\sup_{t\in \mathrm{C}} |M(t,\theta,P_n) - M(t,\theta,P)| > \delta/2)
\end{aligned}
$$

By D1, $\hat\theta_n(t)$ uniformly converges to $\theta(t)$ over C as $n \to \infty$. □

Proof of Theorem 4. Let $\widetilde{Z}_n(t) = n^{-1/2}\sum_{i=1}^n \delta_i(t)V_i(t)/b(t)$. For any $t_1,\dots,t_K \in \mathrm{C}$, denote $\widetilde{\boldsymbol{Z}}_n = (\widetilde{Z}_n(t_1),\dots,\widetilde{Z}_n(t_K))^T$. By the multivariate CLT and the independence between δ_i and V_i, we have

$$\widetilde{\boldsymbol{Z}}_n \xrightarrow{d} N(0,\Xi),$$

where $\Xi = \{\vartheta_{jk}\}_{j,k=1}^K$ is the $K\times K$ covariance matrix with $\vartheta_{jk} = v(t_j,t_k)\gamma(t_j,t_k)/[b(t_j)b(t_k)]$. By Theorem 7.4.2 in Hsing and Eubank (2015), the process $\{\widetilde{Z}_n(t) : t \in \mathrm{C}\}$ is a random element in the Hilbert space $\mathbb{H} = L^2(\mathrm{C},\mathcal{B}(\mathrm{C}),\mu)$, where μ is a finite measure on C. Then it follows from Theorem 7.7.6 in Hsing and Eubank (2015) for i.i.d. Hilbert space valued random variables that

$$\{\widetilde{Z}_n(t) : t \in \mathrm{C}\} \rightsquigarrow \mathrm{GP}(0,\vartheta),$$

where the finite-dimensional restrictions of ϑ is given by the covariance matrix Ξ. Note that

$$\sup_{t\in \mathrm{C}} \left|\widetilde{Z}_n(t) - Z_n(t)\right| \le \sup_{t\in \mathrm{C}} |\widetilde{Z}_n(t)| \cdot \sup_{t\in \mathrm{C}} \left|1 - \frac{b(t)}{\overline{\delta}_n(t)}\right|,$$

where $\overline{\delta}_n(t) = n^{-1}\sum_{i=1}^n \delta_i(t)$. Note that

$$|\overline{\delta}_n(t)| \ge b(t) - |\overline{\delta}_n(t) - b(t)| \ge \inf_{t\in \mathrm{C}} b(t) - W_n,$$

where

$$W_n = \sup_{t\in \mathrm{C}} |n^{-1}\sum_{i=1}^n [\delta_i(t) - b(t)]|.$$

By Lemma 3.1, $E[W_n] = O(n^{-1/2})$. Since $\sup_{t\in \mathrm{C}} |\widetilde{Z}_n(t)| = O_P(1)$, we have

$$\sup_{t\in \mathrm{C}} \left|\widetilde{Z}_n(t) - Z_n(t)\right| = O_P(n^{-1/2}).$$

Then Theorem 4 is an immediate consequence of Slutsky's lemma. □

Proof of Theorem 5. The estimating equation (2.4) can be equivalently written as,

$$\frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\psi(Y_i(t), \hat{\theta}_n(t)) = 0.$$

By mean value theorem,

$$\frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\psi(Y_i(t), \theta(t)) + \frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\dot{\psi}(Y_i(t), \tilde{\theta}_n(t))\big(\hat{\theta}_n(t) - \theta(t)\big) = 0,$$

where $\theta(t) \leqslant \tilde{\theta}_n(t) \leqslant \hat{\theta}_n(t)$, $t \in \mathrm{C}$. Rearranging terms, we get

$$\sqrt{n}\big(\hat{\theta}_n(t) - \theta(t)\big) = -\underbrace{\Big[\frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\dot{\psi}(Y_i(t), \tilde{\theta}_n(t))\Big]^{-1}}_{(1)} \underbrace{\frac{1}{\sum_{j=1}^n \delta_j(t)} \sqrt{n} \sum_{i=1}^n \delta_i(t)\psi(Y_i(t), \theta(t))}_{(2)},$$

where

$$\begin{aligned}
(1) &= \frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\Big[\dot{\psi}(Y_i(t), \tilde{\theta}_n(t)) - \dot{\psi}(Y_i(t), \theta(t))\Big] + \frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\dot{\psi}(Y_i(t), \theta(t)) \\
&= \frac{1}{\sum_{j=1}^n \delta_j(t)/n} \sum_{i=1}^n \Big[\frac{\delta_i(t)}{n} - \frac{b(t)}{n}\Big]\Big[\dot{\psi}(Y_i(t), \tilde{\theta}_n(t)) - \dot{\psi}(Y_i(t), \theta(t))\Big] \\
&\qquad + \frac{b(t)}{\sum_{j=1}^n \delta_j(t)/n}\Big[\frac{1}{n} \sum_{i=1}^n \dot{\psi}(Y_i(t), \tilde{\theta}_n(t)) - \dot{\psi}(Y_i(t), \theta(t))\Big] \\
&\qquad + \frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\dot{\psi}(Y_i(t), \theta(t)) \\
&\leqslant \sup_{t\in \mathrm{C}}\Big\{\frac{1}{\sum_{j=1}^n \delta_j(t)/n}\Big\} \sup_{t\in \mathrm{C}}\Big|\sum_{i=1}^n \Big[\frac{\delta_i(t)}{n} - \frac{b(t)}{n}\Big]\Big[\dot{\psi}(Y_i(t), \tilde{\theta}_n(t)) - \dot{\psi}(Y_i(t), \theta(t))\Big]\Big| \\
&\qquad\qquad + \sup_{t\in \mathrm{C}}\Big\{\frac{b(t)}{\sum_{j=1}^n \delta_j(t)/n}\Big\} \sup_{t\in \mathrm{C}}\Big|\frac{1}{n}\sum_{i=1}^n \dot{\psi}(Y_i(t), \tilde{\theta}_n(t)) - \dot{\psi}(Y_i(t), \theta(t))\Big| \\
&\qquad\qquad + \frac{1}{\sum_{j=1}^n \delta_j(t)} \sum_{i=1}^n \delta_i(t)\dot{\psi}(Y_i(t), \theta(t)).
\end{aligned}$$

As $n \to \infty$, under given conditions and Corollary 1, it is $O_P(n^{-1/2}) + o_P(1) + E_{P_Y}\dot{\psi}(Y(t), \theta(t))$. By Theorem 4, term (2) converges to Gaussian Process with mean zero and covariance function $\varphi(s,t) = \mathrm{Cov}\left\{\psi(Y(t), \theta(t)), \psi(Y(s), \theta(s))\right\}\frac{v(s,t)}{b(s)b(t)}$, where $v(s,t) = E_{P_\delta}[\delta(s)\delta(t)]$. Then it an immediate consequence of Slutsky's lemma. $\square$

Proof of Corollary 2. The convergence of numerator of T_n follows about the same lines as those in the proof of Theorem 1 of Shen and Faraway (2004) under functional limit theorem for robust M-estimator under partial sampling structure. The denominator of T_n, $trace(\hat{\xi}(s,t))$, converges to $trace(\xi(s,t))$ with consistent estimator $\hat{\xi}$. By Slutsky's theorem, proof is done. $\square$

Proof of Corollary 3. By Karhuneun-Loéve theorem, $\xi(s,t) = \sum_{r=1}^{k} \kappa_r e_r(t) e_r(s)$ and we have $\sqrt{n}(\hat{\theta}_n(t) - \theta(t)) = \sum_{r=1}^{k} \eta_r e_r(t)$, where

$$\eta_r = \int_{\mathrm{C}} \sqrt{n}\big(\hat{\theta}_n(t) - \theta(t)\big) e_r(t) dt \sim AN(0, \kappa_r), \;\; r = 1, ..., k.$$

Then we can write

$$\begin{aligned}\sqrt{n}\langle(\hat{\theta}_n(\cdot) - \theta(\cdot)), \phi(\cdot)\rangle = \sqrt{n}\Big(\langle\hat{\theta}_n(t), \phi(\cdot)\rangle - c\Big) &= \int \big(\sum_{r=1}^{k} \eta_r e_r(t)\big)\phi(t) dt \\ &= \sum_{r=1}^{k} \eta_r \int e_r(t)\phi(t) dt = \sum_{r=1}^{k} \eta_r \langle e_r(\cdot), \phi(\cdot)\rangle,\end{aligned}$$

Under the assumption of $tr(\xi) < \infty$, $\sum_{r=1}^{k} \eta_r$ converges in distribution especially to normal distribution. Thus, $\sum_{r=1}^{k} \eta_r \langle e_r(\cdot), \phi(\cdot)\rangle$ also converges to normal distribution under $\langle e_r(t), \phi(t)\rangle < ||e_r(t)|| \cdot ||\phi(t)|| = c < \infty$. The asymptotic variance is derived as $\tau^2 = \mathrm{Var}\Big\{ \sum_{r=1}^{k} \eta_r \langle e_r(\cdot), \phi(\cdot)\rangle \Big\} = \sum_{r=1}^{k} k_r^2 \kappa_r$, where $k_r^2 = \langle e_r(\cdot), \phi(\cdot)\rangle^2 = \int\int_{\mathrm{C}} \phi(s)\phi(t) e_r(s) e_r(t) ds dt$.

□

S2 Appendix: Additional Figure and Results from Simulation Studies

S2.1 Simulated heavy-tailed Data

We present simulated data in Figure 1 under six heavy-tailed or contaminated scenarios considered in Section 5.

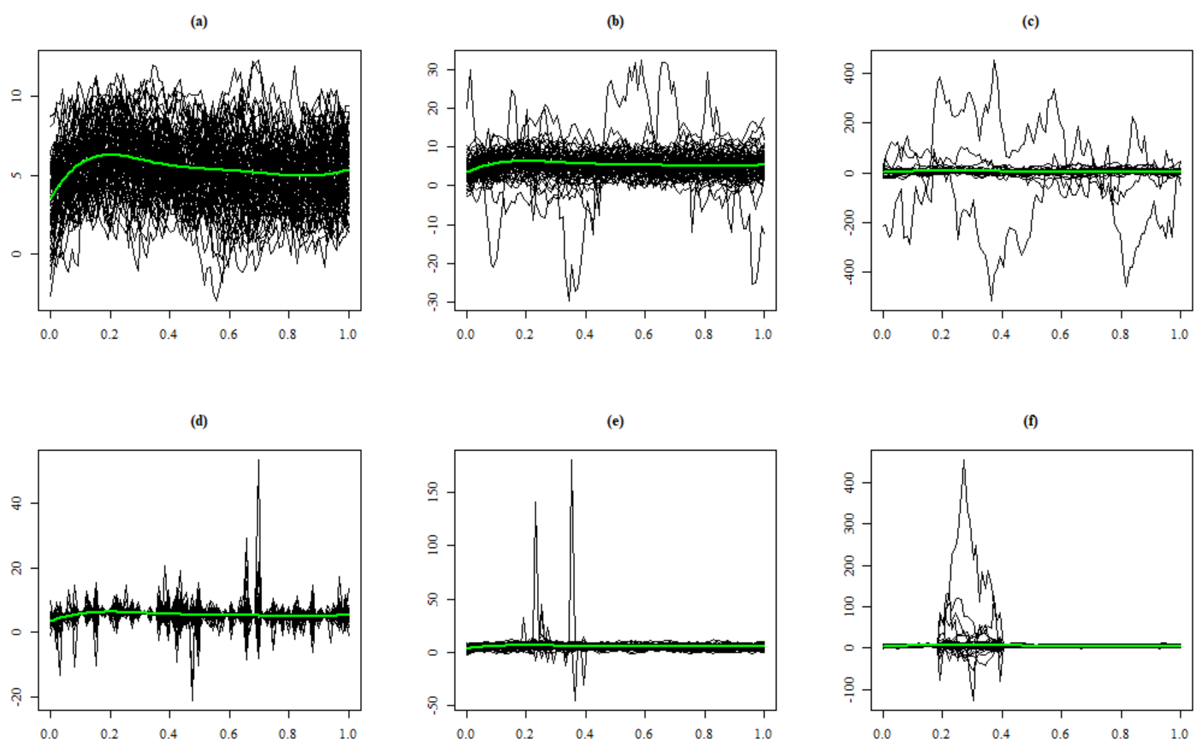

Figure 1: Simulated data from the scenario of (a) Gaussian, (b) t_3, (c) Cauchy, (d) white-noise t_3 with random scales, Gaussian partially contaminated by (e) Cauchy white-noise, and by (f) Cauchy processes. Smooth central line indicates location function.

S2.2 Additional Results for Simulation Studies

Figure 2 displays the estimation performances of robust estimators under partially observed functional data, especially incomplete curves observed at randomly selected interval among a fixed number of pre-specified intervals. Boxplots show similar behaviors that we observe from the results under random interval structure in Section 5.1.2. Our proposed marginal M-estimator achieves superior estimation accuracies compared two competitors under Cauchy data and Gaussian data with Cauchy contaminated noise. Although slightly lower errors are observed from two competing functional estimators under (a) Gaussian and (b) $t(3)$ distributed data compared to ours, differences do not seem significant and almost similar.

Tables 1, 2, and 3 provide detailed results for simulations on bootstrapped functional trend test in Section 5.2 with coverage probabilities and the median length of bootstrapped confidence intervals for projected coefficients to quadratic, linear, and constant basis functions, respectively

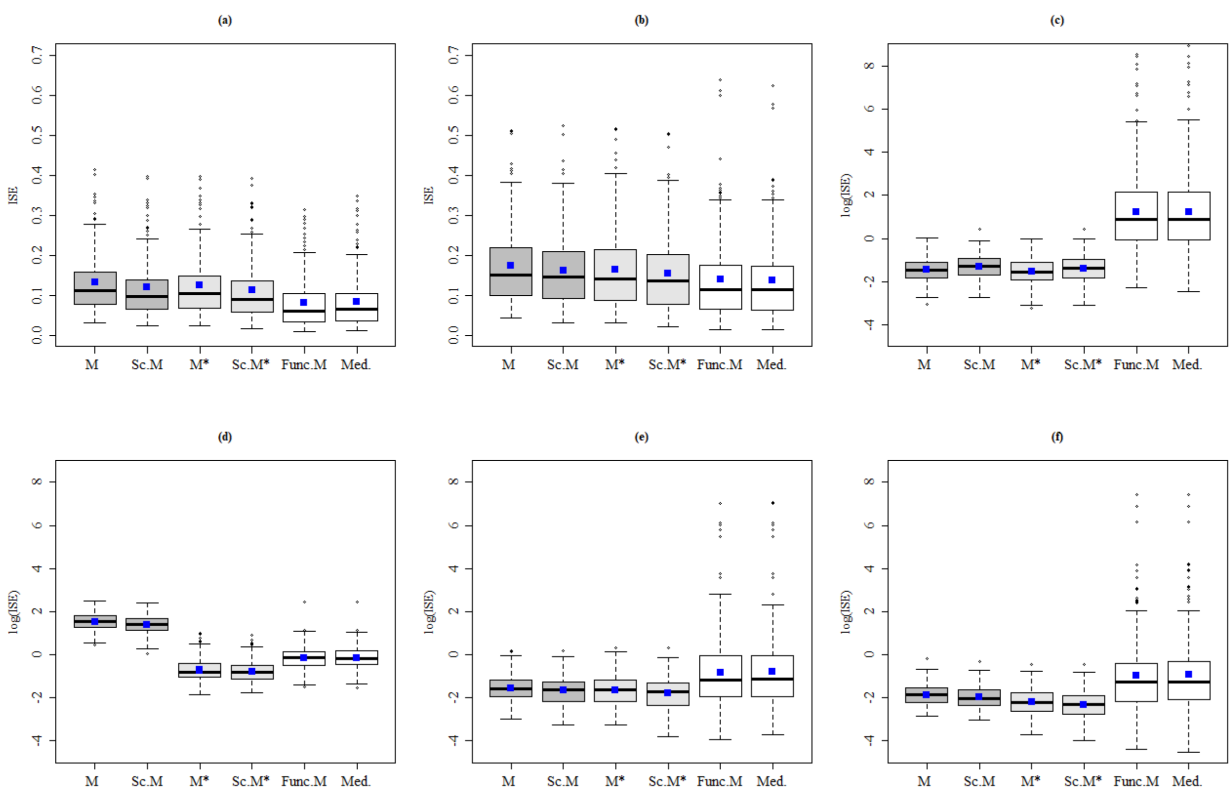

Figure 2: Boxplots of ISE or log transformed ISE over 500 replications from the marginal M-estimator (M), marginal scaled M-estimator (Sc.M), marginal M-estimator under pre-smoothed curves (M*), marginal scaled M-estimator under pre-smoothed curves (Sc.M*), functional M-estimator (Func.M), and functional Median (Med.) under partially observed data at randomly selected interval among a fixed number of pre-specified intervals from (a) Gaussian, (b) t_3, (c) Cauchy, (d) white-noise t_3 with random scales, Gaussian partially contaminated by (e) Cauchy white-noise, and by (f) Cauchy processes. Square dots represent mean values.

Table 1: Coverage probabilities and the median length of bootstrapped confidence intervals (in parenthesis) of projection coefficient to Quadratic basis function from M-estimator (M), scaled M-estimator (Sc.M), and mean over 500 repetitions

	Regular			Irregular1			Irregular2			Dense snippets		
	Mean	Mt	Sc.M	Mean	M	Sc.Mt	Mean	M	Sc.M	Mean	M	Sc.M
Gaussian	0.94	0.942	0.942	0.936	0.954	0.954	0.942	0.946	0.952	0.910	0.915	0.915
	(0.14)	(0.15)	(0.15)	(0.23)	(0.24)	(0.24)	(0.22)	(0.24)	(0.24)	(0.17)	(0.19)	(0.19)
t(3)	0.46	0.948	0.95	0.94	0.95	0.95	0.938	0.944	0.948	0.930	0.920	0.920
	(0.23)	(0.18)	(0.18)	(0.35)	(0.29)	(0.29)	(0.35)	(0.28)	(0.28)	(0.27)	(0.22)	(0.22)
Cauchy	0.938	0.946	0.954	0.912	0.950	0.956	0.902	0.960	0.966	0.960	0.880	0.885
	(1.55)	(0.22)	(0.24)	(2.00)	(0.37)	(0.42)	(2.41)	(0.37)	(0.40)	(4.10)	(0.29)	(0.31)
Cont.1	0.934	0.952	0.954	0.910	0.944	0.946	0.926	0.948	0.954	0.920	0.925	0.935
	(0.83)	(0.21)	(0.22)	(0.95)	(0.34)	(0.35)	(0.86)	(0.34)	(0.35)	(1.19)	(0.28)	(0.28)
Cont.2	0.880	0.940	0.940	0.888	0.938	0.940	0.928	0.950	0.954	0.925	0.935	0.920
	(0.83)	(0.21)	(0.22)	(0.94)	(0.34)	(0.36)	(0.99)	(0.35)	(0.36)	(1.17)	(0.28)	(0.28)

Table 2: Coverage probabilities and the median length of bootstrapped confidence intervals (in parenthesis) of projection coefficient to Linear basis function from M-estimator (M), scaled M-estimator (Sc.M), and mean over 500 repetitions

	Regular			Irregular1			Irregular2			Dense snippets		
	Mean	Mt	Sc.M	Mean	M	Sc.Mt	Mean	M	Sc.M	Mean	M	Sc.M
Gaussian	0.940	0.946	0.948	0.940	0.940	0.950	0.944	0.958	0.960	0.825	0.710	0.730
	(0.19)	(0.20)	(0.20)	(0.28)	(0.30)	(0.30)	(0.30)	(0.32)	(0.32)	(0.32)	(0.36)	(0.36)
t(3)	0.928	0.934	0.934	0.946	0.958	0.966	0.940	0.946	0.950	0.820	0.790	0.815
	(0.30)	(0.23)	(0.23)	(0.44)	(0.35)	(0.35)	(0.47)	(0.36)	(0.37)	(0.44)	(0.40)	(0.41)
Cauchy	0.908	0.948	0.942	0.896	0.946	0.95	0.902	0.952	0.940	0.910	0.680	0.70
	(2.06)	(0.29)	(0.31)	(2.56)	(0.46)	(0.51)	(3.23)	(0.48)	(0.54)	(3.31)	(0.57)	(0.62)
Cont.1	0.892	0.940	0.938	0.928	0.956	0.954	0.926	0.942	0.950	0.855	0.785	0.805
	(1.42)	(0.31)	(0.32)	(1.53)	(0.47)	(0.49)	(1.39)	(0.52)	(0.54)	(1.65)	(0.56)	(0.58)
Cont.2	0.918	0.942	0.944	0.888	0.930	0.934	0.936	0.948	0.952	0.870	0.715	0.725
	(1.42)	(0.31)	(0.32)	(1.55)	(0.47)	(0.49)	(1.55)	(0.47)	(0.48)	(1.40)	(0.53)	(0.55)

Table 3: Coverage probabilities and the median length of bootstrapped confidence intervals (in parenthesis) of projection coefficient to Constant basis function from M-estimator (M), scaled M-estimator (Sc.M), and mean over 500 repetitions

	Regular			Irregular1			Irregular2			Dense snippets		
	Mean	Mt	Sc.M	Mean	M	Sc.Mt	Mean	M	Sc.M	Mean	M	Sc.M
Gaussian	0.950	0.956	0.950	0.944	0.952	0.952	0.954	0.956	0.954	0.690	0.640	0.665
	(0.26)	(0.27)	(0.27)	(0.35)	(0.36)	(0.36)	(0.34)	(0.35)	(0.35)	(0.32)	(0.39)	(0.40)
t(3)	0.908	0.954	0.954	0.916	0.934	0.936	0.940	0.958	0.960	0.735	0.760	0.785
	(0.41)	(0.30)	(0.31)	(0.52)	(0.43)	(0.43)	(0.52)	(0.40)	(0.41)	(0.43)	(0.42)	(0.43)
Cauchy	0.902	0.936	0.934	0.904	0.956	0.960	0.926	0.954	0.952	0.915	0.765	0.785
	(2.92)	(0.38)	(0.42)	(3.04)	(0.55)	(0.61)	(3.51)	(0.51)	(0.57)	(3.25)	(0.58)	(0.60)
Cont.1	0.918	0.964	0.966	0.912	0.950	0.960	0.922	0.948	0.952	0.755	0.690	0.710
	(1.18)	(0.34)	(0.34)	(1.25)	(0.47)	(0.48)	(1.17)	(0.49)	(0.51)	(0.99)	(0.54)	(0.55)
Cont.2	0.918	0.960	0.962	0.874	0.930	0.930	0.914	0.942	0.940	0.735	0.645	0.680
	(1.17)	(0.34)	(0.34)	(1.27)	(0.47)	(0.48)	(1.28)	(0.44)	(0.45)	(0.84)	(0.49)	(0.50)

S2.3 Validity of Robust Inference for Partially Observed Functional Data under Sparse Design Points

To evaluate the numerical feasibility and performance of robust inference under partially observed functional data recorded at sparse points, we apply sparse sampling scheme to sets

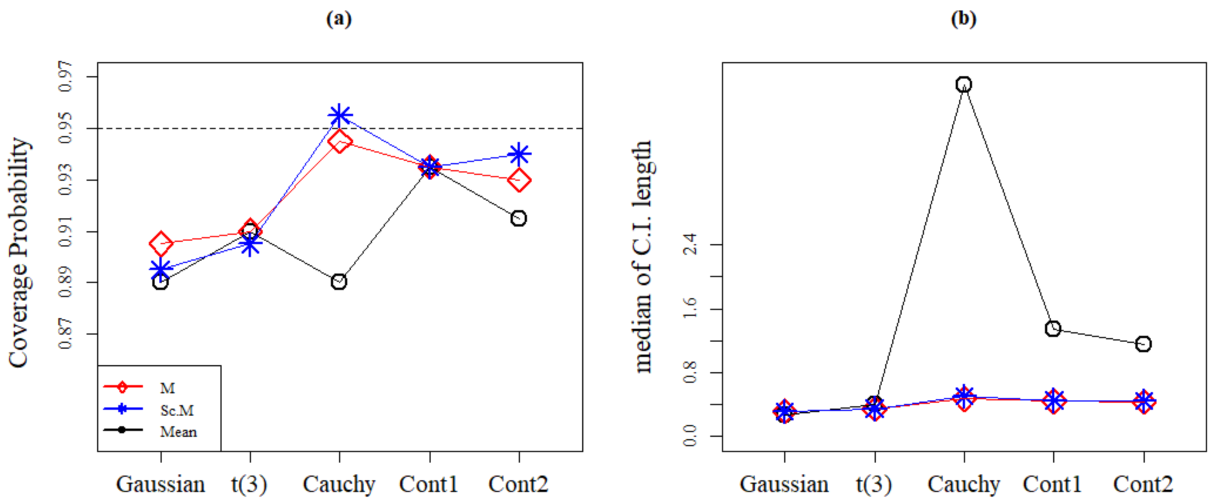

Figure 3: (a) Coverage probabilities of bootstrapped confidence intervals of projection coefficients to quadratic function under Gaussian, t_3, Cauchy, and two contaminated data from M-estimator (M), scaled M-estimator (Sc.M), and Mean functions over 500 repetitions under sparse data. (b) Median length of bootstrapped confidence intervals of projection coefficient

of curves generated under five distributional assumptions considered in Section 5.2.

Specifically, we first define ϵ-equispaced grid points, $t_0, t_1, ..., t_{1/\epsilon}$, over $[0, 1]$, for sufficiently small $\epsilon > 0$, then generate $l_i = \min(v_{i1}, v_{i2})$ and $u_i = \max(v_{i1}, v_{i2})$ from v_{ij}, $j = 1, 2$, i.i.d. from a discrete uniform random variable V on $\{t_0, t_1, ..., t_{1/\epsilon}\}$ to set the lower and upper bounds of the subinterval of each curve. Let t_{ij}, $j = 1, ..., n_i$ denote grid points within each individual random subinterval and $t_{i1} = l_i$, $t_{in_i} = u_i$ by definition. We then assume Bernoulli distribution to draw binary indicator, $\delta(t_{ij}) \overset{i.i.d.}{\sim} Bernoulli(p)$, where p controls the sparsity of the data. In our simulation, we set $p = 0.4$ and the sample size as $n = 200$.

As the generated data have regular sparse design points, the marginal M-estimator is applied to each ϵ-equispaced grid point and Figure 3 displays coverage probabilities of 95% bootstrapped confidence intervals for quadratic coefficient of location parameter and median length of bootstrapped confidence intervals. First it can be seen that robust inferential test performs well even under sparse design by detecting true quadratic trend with $90 - 95\%$ coverage probabilities under various distributional settings. We also observe stable behaviors of confidence intervals with almost constant length of confidence intervals even with existence of heavy-tailed curves or contaminations. On the other hand, unstable performance is observed from inferential test based on functional means.

Via simulation studies, we illustrate the numerical feasibility and validity of robust inferential method for fragmented data observed at sparse grid points. Although further extension on theory, especially for conditions on partial sampling process, is required to fit spares structure to our proposed framework, promising numerical results shine a light on the generalization of our approach even to sparsely observed data.

References

Chernozhukov, V., D. Chetverikov, and K. Kato (2014). Gaussian approximation of suprema of empirical processes. *Annals of Statistics 42*(4), 1564–1597.

Hsing, T. and R. Eubank (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. John Wiley & Sons.

Shen, Q. and J. Faraway (2004). An F type test for linear models with functional responses. *Statistica Sinica 14*, 1239–1257.