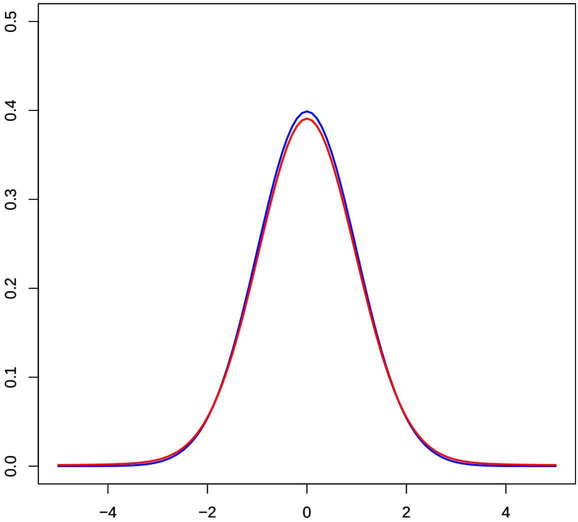
ERCIM 2013 14. December 2013

Qualitatively Robust Estimation for Stochastic Processes

joint work with **Katharina Strohriegl**

Robert Hable
Department of Mathematics
University of Bayreuth

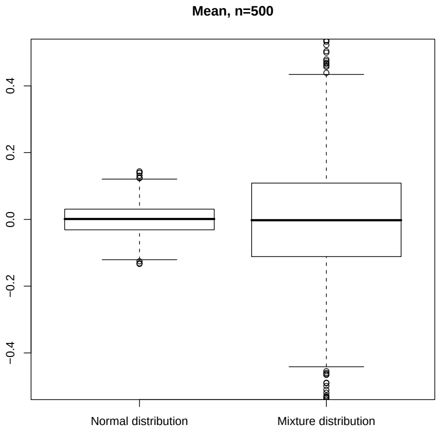
Robustness – Example



Robustness – Example

"mean" applied in 1000 runs

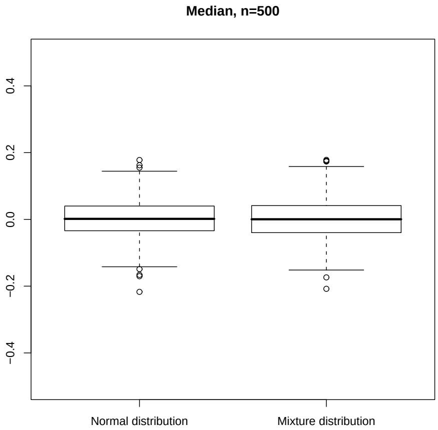
each run consists of a sample with 500 data points



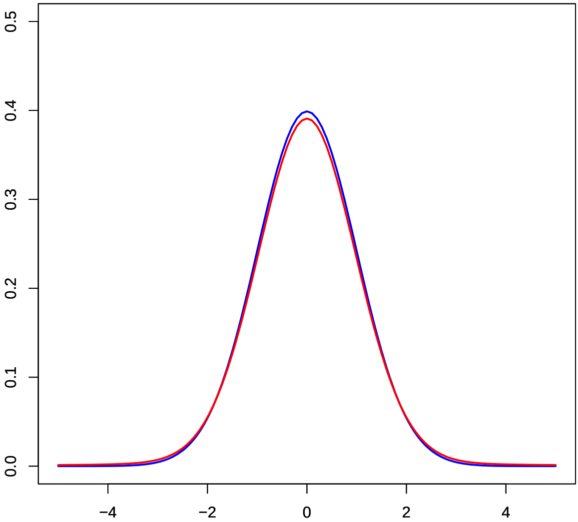
Robustness – Example

"median" applied in 1000 runs

each run consists of a sample with 500 data points



Robustness – Example



Qualitative Robustness: i.i.d.

Model (parametric or nonparametric)

$$\mathcal{P}$$

and

$$Z_i \sim P \quad \text{i.i.d.}\,, \qquad P \in \mathcal{P}$$

Estimator

$$S_n \;:\; \mathcal{Z}^n \to H\,, \qquad (z_1,\ldots,z_n) \mapsto S_n(z_1,\ldots,z_n)$$

Qualitative Robustness: i.i.d.

Model (parametric or nonparametric)

$$\mathcal{P}$$

and

$$Z_i \sim P \quad \text{i.i.d.}\,, \qquad P \in \mathcal{P}$$

Estimator

$$S_n \ : \ \mathcal{Z}^n \to H\,, \qquad (z_1, \ldots, z_n) \mapsto S_n(z_1, \ldots, z_n)$$

Robustness: different notions

small model violations
small errors in the data
} should not change the result too much

Qualitative Robustness: i.i.d.

Estimator

$$S_n \ : \ \mathcal{Z}^n \to \Theta\,, \qquad (z_1,\ldots,z_n) \mapsto S_n(z_1,\ldots,z_n)$$

Qualitative Robustness: Hampel (1968)

- Small errors in the data
  - Small errors in many of the data points (rounding etc.)
  - Large errors in a few data points (gross errors, outliers)

Qualitative Robustness: i.i.d.

Estimator

$$S_n \ : \ \mathcal{Z}^n \to \Theta\,, \qquad (z_1,\ldots,z_n) \mapsto S_n(z_1,\ldots,z_n)$$

Qualitative Robustness: Hampel (1968)

- Small errors in the data
  - Small errors in many of the data points (rounding etc.)
  - Large errors in a few data points (gross errors, outliers)
- "should not change the result too much"

  P : ideal distribution ($P \in \mathcal{P}$)

  Q : real distribution (maybe $Q \notin \mathcal{P}$)

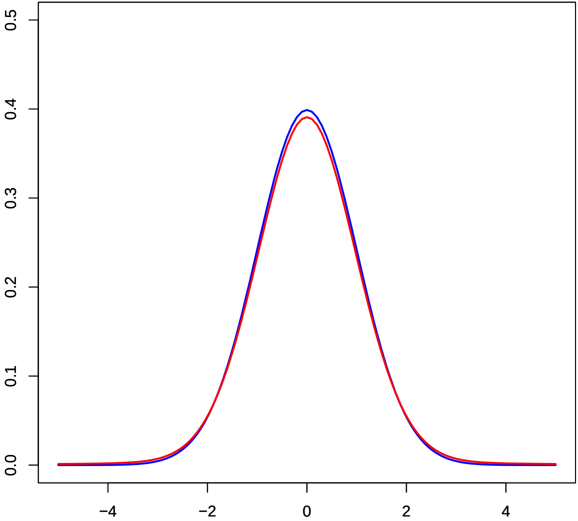
$$Q \text{ close to } P \qquad \longrightarrow \qquad S_n(Q^m) \text{ close to } S_n(P^m)$$

$S_n(P^m)$: ideal distribution of the estimator

$S_n(Q^m)$: real distribution of the estimator

(distribution of the estimator $=$ performance of the estimator)

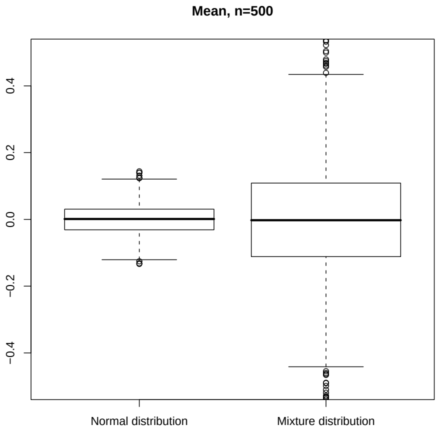
Robustness – Example



Robustness – Example

"mean" applied in 1000 runs

each run consists of a sample with 500 data points



Qualitative Robustness: i.i.d.

Estimator

$$S_n \;:\; \mathcal{Z}^n \to \Theta\,, \qquad (z_1,\ldots,z_n) \mapsto S_n(z_1,\ldots,z_n)$$

Qualitative Robustness: Hampel (1968)

The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is qualitatively robust if

$$\forall\, P \;\; \forall\, \epsilon > 0 \;\; \exists\, \delta > 0 \;\text{ such that }\; \forall\, Q \;\text{ with }\; d_{\text{Pro}}(Q,P) < \delta$$

$$\sup_{n\in\mathbb{N}} d_{\text{Pro}}\big(S_n(Q^n), S_n(P^n)\big) < \varepsilon$$

Qualitative Robustness: i.i.d.

Estimator

$$S_n \;:\; \mathcal{Z}^n \to \Theta\,, \qquad (z_1,\dots,z_n) \mapsto S_n(z_1,\dots,z_n)$$

Qualitative Robustness: Hampel (1968)

The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is qualitatively robust if

$$\forall\, P \;\; \forall\, \epsilon > 0 \;\; \exists\, \delta > 0 \;\text{ such that }\; \forall\, Q \;\text{ with }\; d_{\mathrm{Pro}}(Q,P) < \delta$$

$$\sup_{n\in\mathbb{N}} d_{\mathrm{Pro}}\big(S_n(Q^n), S_n(P^n)\big) < \varepsilon$$

Prokhorov distance between probability measures Q and P:

$$d_{\mathrm{Pro}}(Q,P) \;=\; \inf\left\{\delta > 0 \,\middle|\, Q(A) \le P(A^\delta) + \delta \;\; \forall\, A\right\}$$

$$A^\delta \;=\; \left\{z \,\middle|\, \inf_{z'\in A} d(z,z') < \delta\right\}$$

Qualitative Robustness: i.i.d.

Estimator

$$S_n \;:\; \mathcal{Z}^n \to \Theta\,, \qquad (z_1,\dots,z_n) \mapsto S_n(z_1,\dots,z_n)$$

Qualitative Robustness: Hampel (1968)

The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is qualitatively robust if

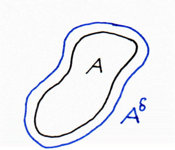
$$\forall\, P \;\; \forall\, \epsilon > 0 \;\; \exists\, \delta > 0 \;\text{ such that }\; \forall\, Q \;\text{ with }\; d_{\text{Pro}}(Q,P) < \delta$$

$$\sup_{n\in\mathbb{N}} d_{\text{Pro}}\big(S_n(Q^n), S_n(P^n)\big) < \varepsilon$$

Prokhorov distance between probability measures Q and P:

$$d_{\text{Pro}}(Q,P) = \inf\left\{\delta > 0 \,\middle|\, Q(A) \le P(A^\delta) + \delta \;\; \forall\, A\right\}$$

$$A^\delta = \left\{z \,\middle|\, \inf_{z'\in A} d(z,z') < \delta\right\}$$



Qualitative Robustness: Hampel's Theorem

Hampel (1968, 1971) (generalized by Cuevas, 1988):
If a sequence of estimators

$$S_n \; : \quad \mathcal{Z}^n \;\to\; H\,, \qquad n \in \mathbb{N}\,,$$

can be represented by a functional

$$S \;:\; \mathcal{M}_1(\mathcal{Z}) \;\to\; H\,, \qquad P \;\mapsto\; S(P)\,,$$

Qualitative Robustness: Hampel's Theorem

Hampel (1968, 1971) (generalized by Cuevas, 1988):
If a sequence of estimators

$$S_n \ : \quad \mathcal{Z}^n \ \to \ H\,, \qquad n \in \mathbb{N}\,,$$

can be represented by a functional

$$S \ : \ \mathcal{M}_1(\mathcal{Z}) \ \to \ H\,, \qquad P \ \mapsto \ S(P)\,,$$

i.e.

$$S_n\big((x_1,y_1),\ldots,(x_n,y_n)\big) \ = \ S\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{(x_i,y_i)}\right) \quad \forall\,(x_i,y_i) \in \mathcal{X}\times\mathcal{Y}$$

for every $n \in \mathbb{N}$

Qualitative Robustness: Hampel's Theorem

Hampel (1968, 1971) (generalized by Cuevas, 1988):
If a sequence of estimators

$$S_n \ : \quad \mathcal{Z}^n \ \to \ H \, , \qquad n \in \mathbb{N} \, ,$$

can be represented by a functional

$$S \ : \ \mathcal{M}_1(\mathcal{Z}) \ \to \ H \, , \qquad P \ \mapsto \ S(P) \, ,$$

Qualitative Robustness: Hampel's Theorem

Hampel (1968, 1971) (generalized by Cuevas, 1988):
If a sequence of estimators

$$S_n \ : \quad \mathcal{Z}^n \ \to \ H\,, \qquad n \in \mathbb{N}\,,$$

can be represented by a functional

$$S \ : \ \mathcal{M}_1(\mathcal{Z}) \ \to \ H\,, \qquad P \ \mapsto \ S(P)\,,$$

and if

the functional S is continuous (w.r.t. the Prokhorov metric),

Qualitative Robustness: Hampel's Theorem

Hampel (1968, 1971) (generalized by Cuevas, 1988):
If a sequence of estimators

$$S_n \; : \quad \mathcal{Z}^n \; \to \; H \, , \qquad n \in \mathbb{N} \, ,$$

can be represented by a functional

$$S \; : \; \mathcal{M}_1(\mathcal{Z}) \; \to \; H \, , \qquad P \; \mapsto \; S(P) \, ,$$

and if

the functional S is continuous (w.r.t. the Prokhorov metric),

then

the sequence of estimators $(S_n)_{n \in \mathbb{N}}$ is qualitatively robust.

so far: **i.i.d. case**

now: **non-i.i.d. case**

Definition of Qualitative Robustness

i.i.d.-case:

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q \in \mathcal{M}_1(\mathcal{Z})$,

$$d_{\mathrm{Pro}}(Q, P) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q^n), S_n(P^n)\big) < \varepsilon \, .$$

Definition of Qualitative Robustness

i.i.d.-case:

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q \in \mathcal{M}_1(\mathcal{Z})$,

$$d_{\mathrm{Pro}}(Q, P) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q^n), S_n(P^n)\big) < \varepsilon \, .$$

or, alternatively,

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called Π-robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1\big(\mathcal{Z}^n\big)$,

$$d_{\mathrm{Pro}}(Q_n, P^n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P^n)\big) < \varepsilon \, .$$

Definition of Qualitative Robustness

i.i.d.-case:

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called Π-robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1\big(\mathcal{Z}^n\big)$,

$$d_{\mathrm{Pro}}(Q_n, P^n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P^n)\big) < \varepsilon \, .$$

Definition of Qualitative Robustness

i.i.d.-case:

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called Π-robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P^n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P^n)\big) < \varepsilon \, .$$

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called Π-robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon \, .$$

Note: Process Z_i, $i \in \mathbb{N}$; $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.
Then, $P_n = \mathcal{L}(\mathbf{D}_n)$

Definition of Qualitative Robustness

i.i.d.-case:

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called Π-robust in P if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P^n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P^n)\big) < \varepsilon \,.$$

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon \,.$$

Note: Process Z_i, $i \in \mathbb{N}$; $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.
Then, $P_n = \mathcal{L}(\mathbf{D}_n)$

Note on Metrics

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon \, .$$

Note on Metrics

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon \; .$$

But: Some care is needed . . .

The

$$\text{distance} \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) \quad \text{on} \quad \mathcal{M}_1(\mathcal{Z}^n)$$

depends on

$$\text{metric} \quad d_n \quad \text{on} \quad \mathcal{Z}^n.$$

Note on Metrics

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon \,.$$

But: Some care is needed . . .

The

$$\text{distance} \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) \quad \text{on} \quad \mathcal{M}_1\big(\mathcal{Z}^n\big)$$

depends on

$$\text{metric} \quad d_n \quad \text{on} \quad \mathcal{Z}^n.$$

$\longrightarrow$ Which metric d_n to choose?

Note on Metrics

- Obvious choice: **product metric**

$$d_{n,2}(z, z') \;=\; \sqrt{\sum_{i=1}^{n} d(z_i, z_i')^2}$$

Problem: sample mean would be qualitatively robust

Note on Metrics

non-i.i.d.-case: Bustos (1981)

- The sequence of estimators $(S_n)_{n\in\mathbb{N}}$ is called qualitatively robust in $P_n \in \mathcal{M}_1(\mathcal{Z}^n)$, $n \in \mathbb{N}$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $Q_n \in \mathcal{M}_1(\mathcal{Z}^n)$,

$$d_{\mathrm{Pro}}(Q_n, P_n) < \delta \quad \Rightarrow \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) < \varepsilon\ .$$

But: Some care is needed . . .

The

$$\text{distance} \quad d_{\mathrm{Pro}}\big(S_n(Q_n), S_n(P_n)\big) \quad \text{on} \quad \mathcal{M}_1(\mathcal{Z}^n)$$

depends on

$$\text{metric} \quad d_n \quad \text{on} \quad \mathcal{Z}^n.$$

$\longrightarrow$ Which metric d_n to choose?

Note on Metrics

- Obvious choice: **product metric**

$$d_{n,2}(z, z') \;=\; \sqrt{\sum_{i=1}^{n} d(z_i, z_i')^2}$$

Problem: sample mean would be qualitatively robust

Note on Metrics

- Obvious choice: **product metric**

$$d_{n,2}(z, z') = \sqrt{\sum_{i=1}^{n} d(z_i, z_i')^2}$$

Problem: sample mean would be qualitatively robust

- Following Boente et al. (1987), we use

$$d_n(z, z') = \inf\left\{\varepsilon > 0 : \sharp\{i : d(z_i, z_i') \geq \varepsilon\}/n \leq \varepsilon\right\}.$$

  - captures the intuitive meaning of qualitative robustness
  - sample mean never qualitatively robust

Hampel's Theorem

i.i.d.

- Estimator S_n represented by a functional S on $\mathcal{M}_1(\mathcal{Z})$:

$$S_n(D_n) \;=\; S(\mathbb{P}_{D_n})\,, \qquad D_n \in (\mathcal{Z})^n\,.$$

- S is continuous.
- Then,

$$S_n, \;\; n \in \mathbb{N}, \;\; \text{is qualitatively robust.}$$

(Hampel (1971) and Cuevas (1988))

Hampel's Theorem

i.i.d.

- Estimator S_n represented by a functional S on $\mathcal{M}_1(\mathcal{Z})$:

$$S_n(D_n) \;=\; S(\mathbb{P}_{D_n})\,, \qquad D_n \in (\mathcal{Z})^n\;.$$

- S is continuous.
- Then,

$$S_n, \;\; n \in \mathbb{N}, \;\; \text{is qualitatively robust.}$$

(Hampel (1971) and Cuevas (1988))

non-i.i.d.:

Is there something similar?
Which kind of processes should be considered?

Varadarajan Property

How is the i.i.d. assumption used in the proof of

$$S \text{ is continuous} \qquad \Rightarrow \qquad S_n \text{ is } \Pi\text{-robust} \qquad ?$$

Varadarajan Property

How is the i.i.d. assumption used in the proof of

$$S \text{ is continuous} \qquad \Rightarrow \qquad S_n \text{ is } \Pi\text{-robust} \qquad ?$$

Varadarajan's Theorem: *If $Z_i \sim P$, $i \in \mathbb{N}$,* **i.i.d.**, *then, for almost every $\omega \in \Omega$,*

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \text{weakly}$$

or, in other words,

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad a.s.$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Varadarajan Property

How is the i.i.d. assumption used in the proof of

$$S \text{ is continuous} \qquad \Rightarrow \qquad S_n \text{ is } \Pi\text{-robust} \qquad ?$$

Varadarajan's Theorem: *If $Z_i \sim P$, $i \in \mathbb{N}$,* **i.i.d.**, *then, for almost every $\omega \in \Omega$,*

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \text{weakly}$$

or, in other words,

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \text{a.s.}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Definition: *Z_n, $n \in \mathbb{N}$, is a Varadarajan-process if there is a probability measure P such that*

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \text{a.s.}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Varadarajan Property

How is the i.i.d. assumption used in the proof of

$$S \text{ is continuous} \qquad \Rightarrow \qquad S_n \text{ is } \Pi\text{-robust} \qquad ?$$

Varadarajan's Theorem: *If $Z_i \sim P$, $i \in \mathbb{N}$,* **i.i.d.**, *then, for almost every $\omega \in \Omega$,*

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

or, in other words,

$$d_{\mathrm{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{a.s.}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Definition: *Z_n, $n \in \mathbb{N}$, is a strong Varadarajan-process if there is a probability measure P such that*

$$d_{\mathrm{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{a.s.}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Varadarajan Property

How is the i.i.d. assumption used in the proof of

$$S \text{ is continuous} \qquad \Rightarrow \qquad S_n \text{ is } \Pi\text{-robust} \qquad ?$$

Varadarajan's Theorem: *If $Z_i \sim P$, $i \in \mathbb{N}$,* **i.i.d.**, *then, for almost every $\omega \in \Omega$,*

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

or, in other words,

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{a.s.}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Generalization of Hampel's Theorem: non-i.i.d.

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\mathrm{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \;\longrightarrow\; 0 \qquad \text{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Generalization of Hampel's Theorem: non-i.i.d.

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \;\longrightarrow\; 0 \qquad \textit{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Theorem: *Let the sequence of estimators S_n, $n \in \mathbb{N}$, be represented by a functional S. If S is continuous and Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process,*
then S_n, $n \in \mathbb{N}$, is qualitatively robust in

$$P_n \,=\, \mathcal{L}\big((Z_1, \ldots, Z_n)\big) \,=\, \mathcal{L}\big(\mathbf{D}_n\big)\,, \quad n \in \mathbb{N}\,.$$

Proof: Essentially, by following the lines of the proofs in Hampel (1971) and Cuevas (1988).

Varadarajan-Processes

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\mathrm{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \;\longrightarrow\; 0 \qquad \textit{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Varadarajan-Processes

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

$\longrightarrow$ a minimal condition when working with empirical measures (?)

Varadarajan-Processes

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\text{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \textit{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

$\longrightarrow$ a minimal condition when working with empirical measures (?)

Examples based on laws of large numbers:

Definition: Steinwart et al. (2009)

Process $(Z_n)_{n\in\mathbb{N}}$ has a **weak law of large numbers** for events:

$$\frac{1}{n}\sum_{i=1}^{n} 1_B \circ Z_i \xrightarrow[n\to\infty]{} c_B \quad \text{in probability} \qquad \forall B \in \mathcal{B}$$

Varadarajan-Processes

Definition: *Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process if there is a probability measure P such that*

$$d_{\mathrm{Pro}}\big(\mathbb{P}_{\mathbf{D}_n}, P\big) \longrightarrow 0 \qquad \text{in probability}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

$\longrightarrow$ a minimal condition when working with empirical measures (?)

Examples based on laws of large numbers:

Definition: Steinwart et al. (2009)

Process $(Z_n)_{n\in\mathbb{N}}$ has a **strong law of large numbers** for events:

$$\frac{1}{n}\sum_{i=1}^{n} 1_B \circ Z_i \xrightarrow[n\to\infty]{} c_B \quad \text{a.s.} \qquad \forall B \in \mathcal{B}$$

Varadarajan-Processes

Definition: Steinwart et al. (2009)

Process $(Z_n)_{n\in\mathbb{N}}$ has a **weak law of large numbers** for events:

$$\frac{1}{n}\sum_{i=1}^{n} 1_B \circ Z_i \xrightarrow[n\to\infty]{} c_B \quad \text{in probability} \qquad \forall B \in \mathcal{B}$$

Theorem: Strohriegl & Hable (2013)

weak law of large numbers $\Rightarrow$ weak Varadarajan

Varadarajan-Processes

Definition: Steinwart et al. (2009)

Process $(Z_n)_{n\in\mathbb{N}}$ has a **strong law of large numbers** for events:

$$\frac{1}{n}\sum_{i=1}^{n} 1_B \circ Z_i \xrightarrow[n\to\infty]{} c_B \quad \text{a.s.} \qquad \forall B \in \mathcal{B}$$

Theorem: Strohriegl & Hable (2013)

strong law of large numbers $\Rightarrow$ strong Varadarajan

Examples for Varadarajan-Processes

- **Examples:** (simply taken from the literature)
  - the i.i.d.-case, of course
  - stationary ergodic processes
  - certain Markov-chains
  - mixing processes

Examples for Varadarajan-Processes

- **Examples:** (simply taken from the literature)
  - the i.i.d.-case, of course
  - stationary ergodic processes
  - certain Markov-chains
  - mixing processes

- **Example**: Weakly Dependent Processes (Dedecker et al., 2007)

  Theorem: Strohriegl & Hable (2013)

  $(Z_n)_{n\in\mathbb{N}}$ is a weak Varadarajan-process if one of the following weak dependence conditions holds:
  - κ-weak dependence with $\kappa(r) = \mathcal{O}(r^{-\kappa})$, $\kappa > 2$,
  - ζ-weak dependence with $\zeta(r) = \mathcal{O}(r^{-\kappa})$, $\kappa > 1$,
  - λ-weak dependence with $\lambda(r) = \mathcal{O}(r^{-\lambda})$, $\lambda > 4$.

Side-Product: Generalizations of Varadarajan's Theorem

Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **i.i.d.**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Side-Product: Generalizations of Varadarajan's Theorem

Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **i.i.d.**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Generalization of Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **pairwise independent**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Side-Product: Generalizations of Varadarajan's Theorem

Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **i.i.d.**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Side-Product: Generalizations of Varadarajan's Theorem

Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **i.i.d.**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Generalization of Varadarajan's Theorem:

If $Z_i \sim P$, $i \in \mathbb{N}$, **stationary ergodic**,
then, for almost every $\omega \in \Omega$,

$$\mathbb{P}_{\mathbf{D}_n(\omega)} \longrightarrow P \qquad \textit{weakly}$$

where $\mathbf{D}_n = (Z_1, \ldots, Z_n)$.

Violation of Varadarajan Property

Recall:

Theorem: *Let the sequence of estimators S_n, $n \in \mathbb{N}$, be represented by a functional S. If S is continuous and Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process, then S_n, $n \in \mathbb{N}$, is qualitatively robust in*

$$P_n = \mathcal{L}\big((Z_1, \ldots, Z_n)\big) = \mathcal{L}\big(\mathbf{D}_n\big), \quad n \in \mathbb{N}.$$

Violation of Varadarajan Property

Recall:

Theorem: *Let the sequence of estimators S_n, $n \in \mathbb{N}$, be represented by a functional S. If S is continuous and Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process, then S_n, $n \in \mathbb{N}$, is qualitatively robust in*

$$P_n = \mathcal{L}\big((Z_1, \ldots, Z_n)\big) = \mathcal{L}\big(\mathbf{D}_n\big)\,, \quad n \in \mathbb{N}\,.$$

- ideal data would be from $P_n = \mathcal{L}\big((Z_1, \ldots, Z_n)\big)$
- observed data are from $Q_n = \mathcal{L}\big((\tilde{Z}_1, \ldots, \tilde{Z}_n)\big)$

Note:

Observed process does not have to be Varadarajan!

Violation of Varadarajan Property

Recall:

Theorem: *Let the sequence of estimators S_n, $n \in \mathbb{N}$, be represented by a functional S. If S is continuous and Z_n, $n \in \mathbb{N}$, is a weak Varadarajan-process, then S_n, $n \in \mathbb{N}$, is qualitatively robust in*

$$P_n = \mathcal{L}\big((Z_1, \ldots, Z_n)\big) = \mathcal{L}\big(\mathbf{D}_n\big), \quad n \in \mathbb{N}.$$

- ideal data would be from $P_n = \mathcal{L}\big((Z_1, \ldots, Z_n)\big)$
- observed data are from $Q_n = \mathcal{L}\big((\tilde{Z}_1, \ldots, \tilde{Z}_n)\big)$

Note:

Observed process does not have to be Varadarajan!

$\longrightarrow$ **Small violations of ergodicity, stationarity, etc. covered!**

Related Work

Other generalizations of Hampel's Theorem: non-i.i.d. case

- special case of stationary ergodic processes, different definition of qualitative robustness
  - Cox (1981)
  - Boente et al. (1982)
- identically distributed observations (special definition of qualitative robustness only for this case)
  - Zähle (2012)
  - Zähle (2013)

Example: M-Estimator

M-estimators for location

$$\sum_{i=1}^{n} \psi\big(z_i - S_n(z_1, \ldots, z_n)\big) \overset{!}{=} 0$$

are represented by a functional S defined by

$$\int \psi(z - S(P))\, P(dz) \overset{!}{=} 0\,.$$

Example: M-Estimator

M-estimators for location

$$\sum_{i=1}^{n} \psi\big(z_i - S_n(z_1, \ldots, z_n)\big) \overset{!}{=} 0$$

are represented by a functional S defined by

$$\int \psi(z - S(P))\, P(dz) \overset{!}{=} 0\,.$$

Classical result: i.i.d. case

M-estimator is qualitatively robust in P if

- ψ is monotone increasing
- ψ is bounded
- $S(P)$ is unique

Example: M-Estimator

M-estimators for location

$$\sum_{i=1}^{n} \psi\big(z_i - S_n(z_1, \ldots, z_n)\big) \overset{!}{=} 0$$

are represented by a functional S defined by

$$\int \psi(z - S(P))\, P(dz) \overset{!}{=} 0\,.$$

Classical result: i.i.d. case

M-estimator is qualitatively robust in P if

- ψ is monotone increasing
- ψ is bounded
- $S(P)$ is unique

$\longrightarrow$ This result also holds for all Varadarajan processes!

References

- **K. Strohriegl and R. Hable (2013)**: Qualitatively Robust Estimation for Stochastic Processes. *Submitted*.
- **I. Steinwart, D. Hush, and C. Scovel (2009)**: Learning from dependent observations. *Journal of Multivariate Analysis*, 100:175–194.
- **F. R. Hampel (1971)**: A general qualitative definition of robustness. *Annals of Mathematical Statistics*, 42:1887–1896.
- **G. Boente, R. Fraiman, and V. J. Yohai (1987)**: Qualitative robustness for stochastic processes. *The Annals of Statistics*, 15:1293–1312.
- **H. Zähle (2013)**: Qualitative robustness of von Mises statistics based on strongly mixing data. *Statistical Papers*, to appear.

Supported by **DFG** grant CH 291/2-1