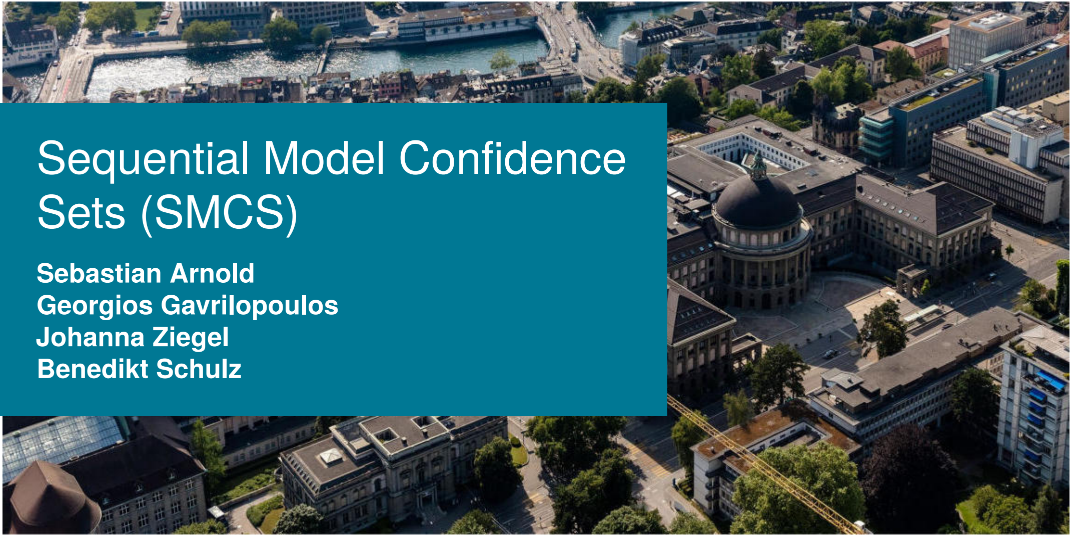
ETH zürich

Seminar for Statistics
ETH Zürich



Sequential Model Confidence Sets (SMCS)

Sebastian Arnold
Georgios Gavrilopoulos
Johanna Ziegel
Benedikt Schulz

Outline

1. Model Confidence Sets (MCS)
2. Sequential Model Confidence Sets
3. Mathematical Tools
4. Implementation
5. Weakly Superior Objects

Uncertainty Quantification

Hypothesis Testing

Parameter Space
Continuous
Confidence Intervals

Overview

Uncertainty Quantification

Hypothesis Testing	***Forecast Evaluation***
Parameter Space	Space of Models
Continuous	Discrete
Confidence Intervals	Model Confidence Sets

Overview

Uncertainty Quantification

Hypothesis Testing	***Forecast Evaluation***
Parameter Space	Space of Models
Continuous	Discrete
Confidence Intervals	Model Confidence Sets

Model Confidence Set (MCS): A subset of the model space containing the *optimal* model.

Model Optimality

Set of forecasts $\mathcal{M}_0 := \{1, \ldots, m\}$.

Model Optimality

Set of forecasts $\mathcal{M}_0 := \{1, \dots, m\}$.

Target: Conditional risk difference

$$\mathbb{E}\left[L_{i,t} - L_{j,t} \,\middle|\, \mathcal{F}_{t-1}\right] =: \mu_{ij,t}$$

where $L_{i,t} = L(f_{i,t}, Y_t)$.

Model Optimality

Set of forecasts $\mathcal{M}_0 := \{1, \dots, m\}$.

Target: Conditional risk difference

$$\mathbb{E}\left[L_{i,t} - L_{j,t} \,\middle|\, \mathcal{F}_{t-1}\right] =: \mu_{ij,t}$$

where $L_{i,t} = L(f_{i,t}, Y_t)$.

Set of Optimal Models

If $\mu_{ij,t}$ does not depend on t,

$$\mathcal{M}^{\star} := \left\{ i \in \mathcal{M}_0 : \mu_{ij} \leq 0 \quad \text{for all } j \in \mathcal{M}_0 \right\}.$$

Model Confidence Set (MCS)

Definition of a $(1-\alpha)$-MCS

A subset of $\mathcal{M}_0$ such that

$$\mathbb{P}\left(\mathcal{M}^{\star} \subset \widehat{\mathcal{M}}_{1-\alpha}\right) \geq 1-\alpha.$$

Originally proposed by Hansen et al. (2011).

Model Confidence Set (MCS)

Definition of a $(1-\alpha)$-MCS

A subset of $\mathcal{M}_0$ such that

$$\mathbb{P}\left(\mathcal{M}^\star \subset \widehat{\mathcal{M}}_{1-\alpha}\right) \geq 1-\alpha.$$

Originally proposed by Hansen et al. (2011).

Assumptions: strict stationarity and α-mixing conditions for $d_{ij,t} = L_{i,t} - L_{j,t}$.

Methods: CLT $\Rightarrow$ asymptotic test for $H_0 : \mu_{ij} = 0$ for all $i, j \in \mathcal{M}_0$.

Outline

Strongly Superior Models (inspired by Henzi and Ziegel (2022))

$$\boxed{\mu_{ij,t} = \mathbb{E}\left[L_{i,t} - L_{j,t} \,\middle|\, \mathcal{F}_{t-1}\right]}$$

Strongly superior models

$$\mathcal{M}^{\mathsf{s},\star} := \left\{ i \in \mathcal{M}_0 \mid \mu_{ij,t} \leq 0 \text{ for all } t \in \mathbb{N} \text{ and all } j \in \mathcal{M}_0 \right\}.$$

Strongly Superior Models (inspired by Henzi and Ziegel (2022))

$$\boxed{\mu_{ij,t} = \mathbb{E}\left[L_{i,t} - L_{j,t} \,\middle|\, \mathcal{F}_{t-1}\right]}$$

Strongly superior models

$$\mathcal{M}^{\mathsf{s},\star} := \left\{ i \in \mathcal{M}_0 \mid \mu_{ij,t} \leq 0 \text{ for all } t \in \mathbb{N} \text{ and all } j \in \mathcal{M}_0 \right\}.$$

Toy Example

$y_1, y_2, \ldots \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and, for $i = 1, \ldots, m$, the forecasting model i outputs

$$x_{i,1}, x_{i,2}, \ldots \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, i).$$

Then, for the CRPS loss, $\mu_{1j,t} \leq 0$ for all $t \geq 1$ and all $j \in \mathcal{M}_0$.

Weakly Superior Models (inspired by Choe and Ramdas (2021))

$$\Delta_{ij,t} = \frac{1}{t} \sum_{s=1}^{t} \mathbb{E}\left[L_{i,s} - L_{j,s} \,\middle|\, \mathcal{F}_{s-1}\right]$$

(Uniformly) Weakly superior

$$\mathcal{M}^{\text{uw},\star} := \{i \in \mathcal{M}_0 \mid \Delta_{ij,t} \leq 0 \text{ for all } t \in \mathbb{N} \text{ and all } j \in \mathcal{M}_0\}$$

Toy example

$y_1, y_2, \ldots \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and the models $i = 2, \ldots, m$ are defined as before. Model $i = 1$ outputs

$$x_{1,1}, x_{1,2}, \ldots \sim \begin{cases} \mathcal{N}(0,1), & t \notin 7\mathbb{N} \\ \mathcal{N}(0,3), & t \in 7\mathbb{N} \end{cases}$$

Sequential Model Confidence Sets

Definition of a Sequential MCS

A sequence $(\widehat{\mathcal{M}}_{1-\alpha,t})_{t\in\mathbb{N}}$ of subsets of $\mathcal{M}_0$ such that

$$\mathbb{P}\left(\mathcal{M}^\star \subset \widehat{\mathcal{M}}_{1-\alpha,t} \quad \text{for all } t \in \mathbb{N}\right) \geq 1-\alpha.$$

Sequential Model Confidence Sets

Definition of a Sequential MCS

A sequence $(\widehat{\mathcal{M}}_{1-\alpha,t})_{t\in\mathbb{N}}$ of subsets of $\mathcal{M}_0$ such that

$$\mathbb{P}\left(\mathcal{M}^\star \subset \widehat{\mathcal{M}}_{1-\alpha,t} \quad \text{for all } t \in \mathbb{N}\right) \geq 1-\alpha.$$

Goals:

- Non-Asymptotic.
- No stationarity assumptions on the loss differences.
- Time-Uniform Coverage.

Outline

E-processes as Test Statistics

For all models $i, j \in \mathcal{M}_0$, we test the hypothesis

$$\mathcal{H}_{ij} : f_{i,t} \text{ is better than } f_{j,t} \text{ for all } t \in \mathbb{N}.$$

E-processes as Test Statistics

For all models $i, j \in \mathcal{M}_0$, we test the hypothesis

$$\mathcal{H}_{ij} : f_{i,t} \text{ is better than } f_{j,t} \text{ for all } t \in \mathbb{N}.$$

- We have an **e-process**, i.e. a process $(E_{ij,t})_{t=1}^{\infty}$ such that $\mathbb{P}_{\mathcal{H}_{ij}}\left(\sup_{t\in\mathbb{N}} E_{ij,t} > \frac{1}{\alpha}\right) \leq \alpha$.
- We compute the average, which is an e-process for $\mathcal{H}_{i\cdot} := \bigcap_{j\in\mathcal{M}_0} \mathcal{H}_{ij}$:

$$E_{i\cdot,t} := \frac{1}{m-1} \sum_{j\neq i} E_{ij,t}$$

- Multiple testing adjustment generalizing Vovk and Wang (2021):

$$E^{\star}_{i\cdot,t} = \min_{I\subset\{1,\ldots,m\}:i\in I} \frac{1}{|I|} \sum_{j\in I} E_{j\cdot,t},$$

E-processes as Test Statistics

For all models $i, j \in \mathcal{M}_0$, we test the hypothesis

$$\mathcal{H}_{ij} : f_{i,t} \text{ is better than } f_{j,t} \text{ for all } t \in \mathbb{N}.$$

- We have an **e-process**, i.e. a process $(E_{ij,t})_{t=1}^{\infty}$ such that $\mathbb{P}_{\mathcal{H}_{ij}}\left(\sup_{t\in\mathbb{N}} E_{ij,t} > \frac{1}{\alpha}\right) \leq \alpha$.
- We compute the average, which is an e-process for $\mathcal{H}_{i\cdot} := \bigcap_{j\in\mathcal{M}_0} \mathcal{H}_{ij}$:

$$E_{i\cdot,t} := \frac{1}{m-1}\sum_{j\neq i} E_{ij,t}$$

- Multiple testing adjustment generalizing Vovk and Wang (2021):

$$E^{\star}_{i\cdot,t} = \min_{I\subset\{1,\dots,m\}:i\in I} \frac{1}{|I|}\sum_{j\in I} E_{j\cdot,t},$$

Construction of the SMCS

$$\widehat{\mathcal{M}}_t = \left\{i \in \mathcal{M}_0 \mid E^{\star}_{i\cdot,t} < 1/\alpha\right\}, \quad t \in \mathbb{N}.$$

Outline

1. Model Confidence Sets (MCS)
2. Sequential Model Confidence Sets
3. Mathematical Tools
4. Implementation
5. Weakly Superior Objects

Strongly superior objects

If $d_{ij,t} := L_{i,t} - L_{j,t}$ is (conditionally) bounded, a **suitable e-process** (Henzi and Ziegel, 2022) is

$$E_{ij,t} := \prod_{s=1}^{t} \Big[1 + \lambda_s (L_{i,s} - L_{j,s})\Big].$$

Strongly superior objects

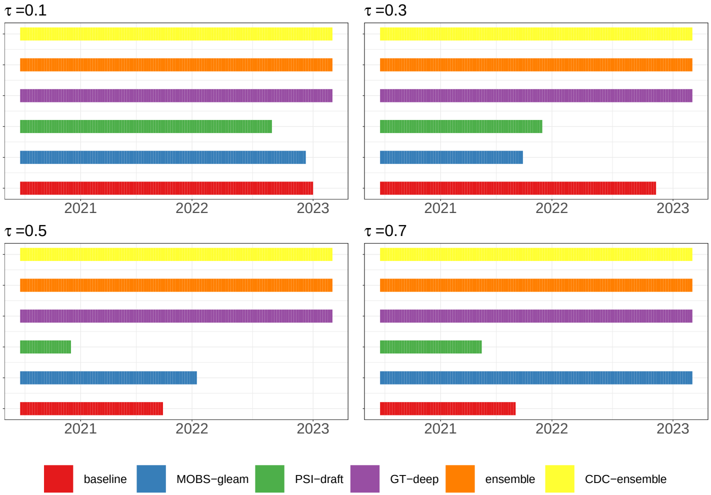
If $d_{ij,t} := L_{i,t} - L_{j,t}$ is (conditionally) bounded, a **suitable e-process** (Henzi and Ziegel, 2022) is

$$E_{ij,t} := \prod_{s=1}^{t} \Big[1 + \lambda_s (L_{i,s} - L_{j,s})\Big].$$

COVID-19 Case Study

- 1-week-ahead quantile forecasts for *COVID-19-related deaths in the USA*.
- Evaluation with the GPL score $S_\tau(x, y) = (\mathbb{1}\{x \geq y\} - \tau)\,(\log x - \log y)$.

Case Study Results



Uniformly weakly superior objects

A **suitable e-process** (Choe and Ramdas, 2021) is

$$E_{ij,t} := \exp\left\{\lambda_{ij}\sum_{s=1}^{t} d_{ij,s} - \psi_{E,c_{ij}}(\lambda_{ij})V_{ij,t}\right\}.$$

Uniformly weakly superior objects

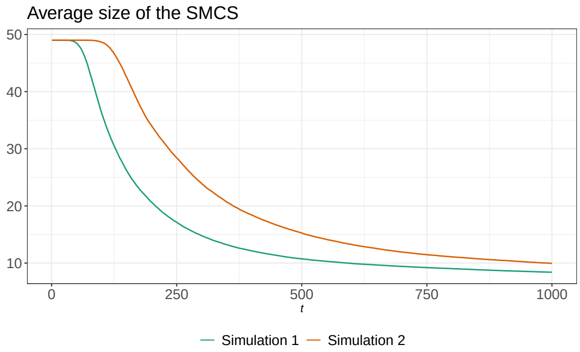
A **suitable e-process** (Choe and Ramdas, 2021) is

$$E_{ij,t} := \exp\left\{\lambda_{ij}\sum_{s=1}^{t} d_{ij,s} - \psi_{E,c_{ij}}(\lambda_{ij})V_{ij,t}\right\}.$$

Simulation Design

- Sample $(Y_t)_{t=1}^N$, where $Y_t \sim \mathcal{N}(Y_{t-1}, 1)$.
- Consider forecasting models $f_{i,t} = \mathcal{N}(Y_{t-1} + \varepsilon_i, 1 + \delta_i)$ for $(\varepsilon_i, \delta_i) \in \{0, \pm 0.2, \pm 0.4, \pm 0.6\}^2$.
- Evaluate them with the CRPS.
- The one corresponding to $\varepsilon_{i_0} = \delta_{i_0} = 0$ is the only optimal one.

Simulation Results



Outline

1. Model Confidence Sets (MCS)
2. Sequential Model Confidence Sets
3. Mathematical Tools
4. Implementation
5. Weakly Superior Objects

Weakly superior objects

$\Delta_{ij,t} = \frac{1}{t} \sum_{s=1}^{t} \mathbb{E}[L_{i,s} - L_{j,s} \,|\, \mathcal{F}_{s-1}]$

- Before: $\mathcal{M}^{\text{uw},\star} := \left\{ i \in \mathcal{M}_0 \,\middle|\, \Delta_{ij,t} \leq 0 \text{ for all } t \in \mathbb{N} \text{ and all } j \in \mathcal{M}_0 \right\}$.
- Now: $\mathcal{M}_t^{\text{w},\star} := \left\{ i \in \mathcal{M}_0 \,\middle|\, \Delta_{ij,t} \leq 0 \text{ for all } j \in \mathcal{M}_0 \right\}$.

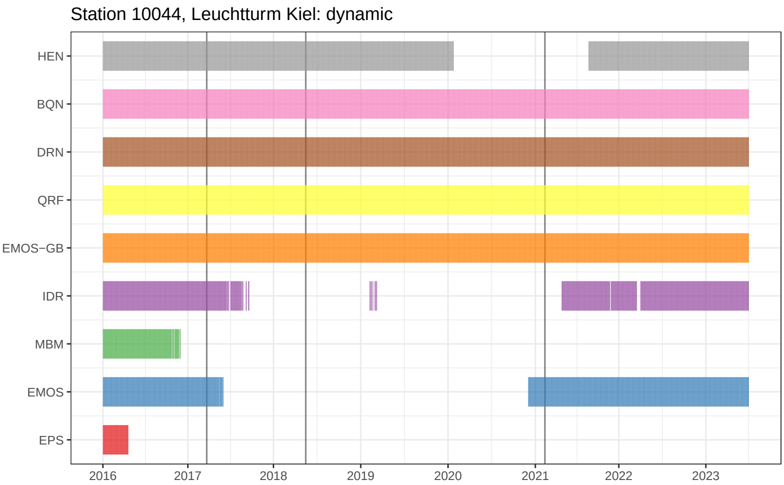
Weakly superior objects

$\Delta_{ij,t} = \frac{1}{t}\sum_{s=1}^{t} \mathbb{E}[L_{i,s} - L_{j,s} \,|\, \mathcal{F}_{s-1}]$

- Before: $\mathcal{M}^{\text{uw},\star} := \left\{ i \in \mathcal{M}_0 \,\middle|\, \Delta_{ij,t} \leq 0 \text{ for all } t \in \mathbb{N} \text{ and all } j \in \mathcal{M}_0 \right\}$.
- Now: $\mathcal{M}_t^{\text{w},\star} := \left\{ i \in \mathcal{M}_0 \,\middle|\, \Delta_{ij,t} \leq 0 \text{ for all } j \in \mathcal{M}_0 \right\}$.

A model can reappear after it has been eliminated.

Case study on wind gusts



Bibliography

Y. J. Choe and A. Ramdas. Comparing sequential forecasters. *Operations Research*, 2021.

P. R. Hansen, A. Lunde, and J. M. Nason. The model confidence set. *Econometrica*, 79(2):453–497, 2011.

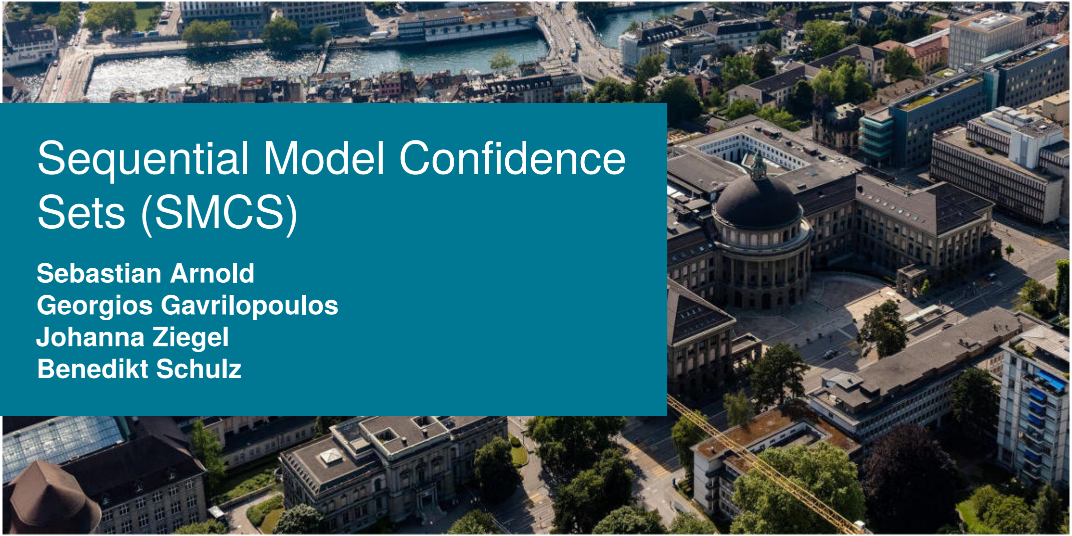
A. Henzi and J. F. Ziegel. Valid sequential inference on probability forecast performance. *Biometrika*, 109(3):647–663, 2022.

V. Vovk and R. Wang. E-values: Calibration, combination and applications. *Ann. Stat.*, 49(3): 1736–1754, 2021.

ETH zürich

Thank you!

Seminar for Statistics
ETH Zürich



Sequential Model Confidence Sets (SMCS)

Sebastian Arnold
Georgios Gavrilopoulos
Johanna Ziegel
Benedikt Schulz

Outline

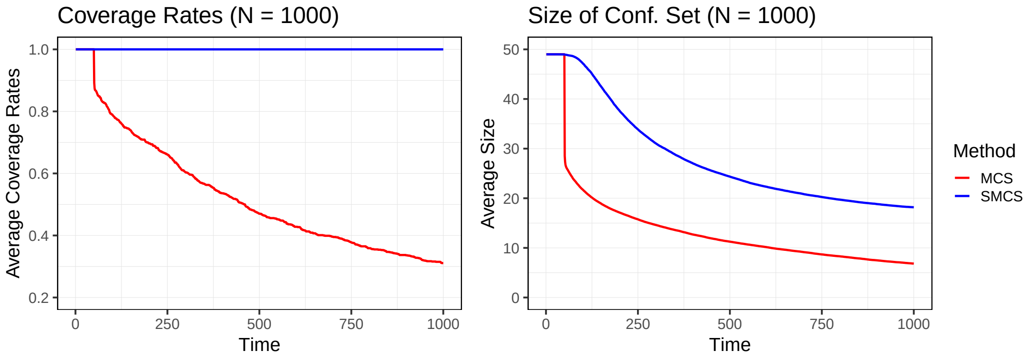
6. Further Empirical Results

7. Implementation of the SMCS for Weak Superior Objects

8. Further questions

9. Technical Details

Further Empirical Results



Outline

Implementation

Assume that $|d_{ij,t}| \leq c_{ij}/2$.

- Consider $M_{ij,t}(x) = \exp\left\{\lambda_{ij} \sum_{s=1}^{t} d_{ij,t} - \lambda_{ij} t x - \psi_{E,c_{ij}}(\lambda_{ij}) V_{ij,t}\right\}$.

Implementation

Assume that $|d_{ij,t}| \leq c_{ij}/2$.

- Consider $M_{ij,t}(x) = \exp\left\{\lambda_{ij} \sum_{s=1}^{t} d_{ij,t} - \lambda_{ij} t x - \psi_{E,c_{ij}}(\lambda_{ij}) V_{ij,t}\right\}$.
- Average out:

$$M_t(\boldsymbol{X}) = \frac{1}{m(m-1)} \sum_{\substack{i,j=1,\dots,m \\ i \neq j}} M_{ij,t}(x_{ij}).$$

Implementation

Assume that $|d_{ij,t}| \leq c_{ij}/2$.

- Consider $M_{ij,t}(x) = \exp\left\{\lambda_{ij} \sum_{s=1}^{t} d_{ij,t} - \lambda_{ij} t x - \psi_{E,c_{ij}}(\lambda_{ij}) V_{ij,t}\right\}$.
- Average out:

$$M_t(\boldsymbol{X}) = \frac{1}{m(m-1)} \sum_{\substack{i,j=1,\ldots,m \\ i \neq j}} M_{ij,t}(x_{ij}).$$

- Invert: $C_t = \left\{\boldsymbol{X} \in \mathbb{R}_0^{m\times m} \mid M_t(\boldsymbol{X}) \leq 1/\alpha\right\}$.

Implementation

Assume that $|d_{ij,t}| \leq c_{ij}/2$.

- Consider $M_{ij,t}(x) = \exp\left\{\lambda_{ij} \sum_{s=1}^{t} d_{ij,t} - \lambda_{ij} t x - \psi_{E,c_{ij}}(\lambda_{ij}) V_{ij,t}\right\}$.
- Average out:

$$M_t(\boldsymbol{X}) = \frac{1}{m(m-1)} \sum_{\substack{i,j=1,\ldots,m \\ i \neq j}} M_{ij,t}(x_{ij}).$$

- Invert: $C_t = \left\{\boldsymbol{X} \in \mathbb{R}_0^{m \times m} \mid M_t(\boldsymbol{X}) \leq 1/\alpha\right\}$.
- Exclude $i \in \mathcal{M}_0$ from the SMCS if $\exists j \neq i$ such that

$$C_t \cap \{\boldsymbol{X} \in \mathbb{R}_0^{m \times m} \mid x_{ij} \leq 0\} = \emptyset.$$

Implementation

Assume that $|d_{ij,t}| \leq c_{ij}/2$.

- Consider $M_{ij,t}(x) = \exp\left\{\lambda_{ij} \sum_{s=1}^{t} d_{ij,t} - \lambda_{ij} t x - \psi_{E,c_{ij}}(\lambda_{ij}) V_{ij,t}\right\}$.
- Average out:

$$M_t(\boldsymbol{X}) = \frac{1}{m(m-1)} \sum_{\substack{i,j=1,\dots,m \\ i \neq j}} M_{ij,t}(x_{ij}).$$

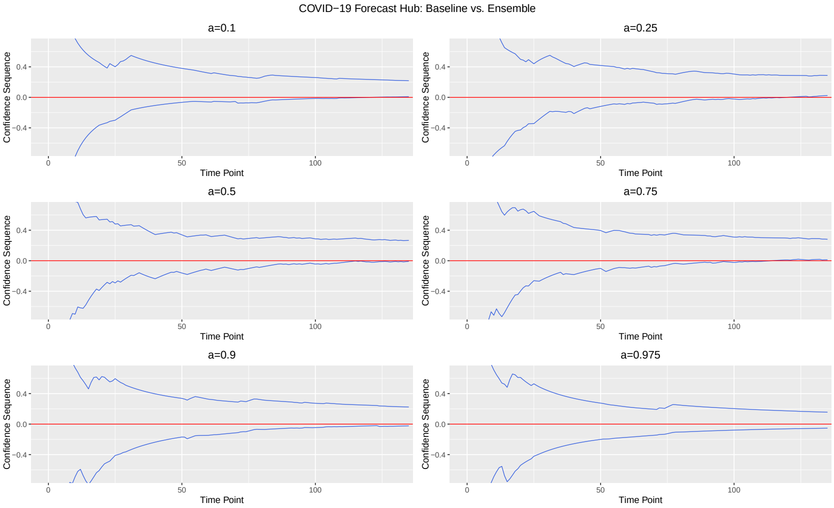
- Invert: $C_t = \left\{\boldsymbol{X} \in \mathbb{R}_0^{m \times m} \mid M_t(\boldsymbol{X}) \leq 1/\alpha\right\}$.
- Exclude $i \in \mathcal{M}_0$ from the SMCS if $\exists j \neq i$ such that

$$C_t \cap \{\boldsymbol{X} \in \mathbb{R}_0^{m \times m} \mid x_{ij} \leq 0\} = \emptyset.$$

- Output

$$\widehat{M}_t = \left\{i \in \mathcal{M}_0 \,\middle|\, C_t \cap \{\boldsymbol{X} \in \mathbb{R}_0^{m \times m} \mid x_{ij} \leq 0\} \neq \emptyset\right\}.$$

Examples of univariate confidence sequences



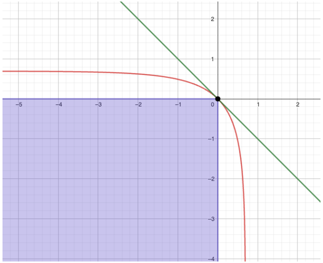
Outline

Open questions

1. Predictable weights for multiple testing?

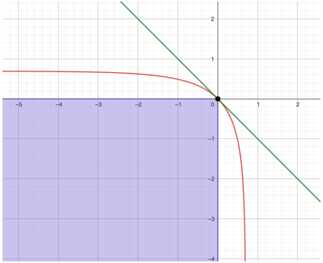
Open questions

1. Predictable weights for multiple testing?
2. Better approximation to the maximum than the average.



Open questions

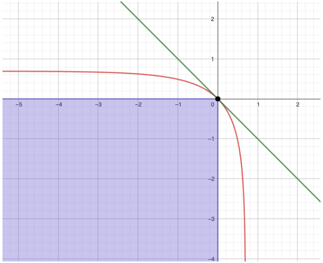
1. Predictable weights for multiple testing?
2. Better approximation to the maximum than the average.



3. Sequential Model Confidence Sets for Model Selection (with AIC/BIC).

Open questions

1. Predictable weights for multiple testing?
2. Better approximation to the maximum than the average.



3. Sequential Model Confidence Sets for Model Selection (with AIC/BIC).
4. More general loss functions.

Outline

Why not e-processes for Weakly Superior objects?

The process

$$E_{ij,t} := \prod_{s=1}^{t} \left[1 + \lambda_s d_{ij,s}\right]$$

is an e-process **only** under $\mathcal{H}^{\mathsf{s}}_{ij} = \{\mathbb{Q} \in \mathfrak{B}(\Omega) : \mu_{ij,t} \leq 0$ for all $t \in \mathbb{N}\}$, **but not** under $\mathcal{H}^{\mathsf{w}}_{ij,t} := \{\mathbb{Q} \in \mathfrak{B}(\Omega) : \mu_{ij,t} \leq 0\}$.