Bayesian inference for multivariate extreme value distributions

Sebastian Engelke

Clément Dombry, Marco Oesting

Toronto, *Fields Institute*, May 4th, 2016

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

SWISS NATIONAL SCIENCE FOUNDATION

Main motivation

- For a parametric model $\mathbf{Z} \sim F_\theta$ of multivariate **max-stable distributions**, the full LLHs are usually not available.

Main motivation

- For a parametric model $\mathbf{Z} \sim F_\theta$ of multivariate **max-stable distributions**, the full LLHs are usually not available.
- **Maximum likelihood estimation** is infeasible.

Main motivation

- For a parametric model $\mathbf{Z} \sim F_\theta$ of multivariate **max-stable distributions**, the full LLHs are usually not available.
- **Maximum likelihood estimation** is infeasible.
- **Goal:** Use full LLHs in Bayesian setup to
  - **improve frequentist efficiency**,

Main motivation

- For a parametric model $\mathbf{Z} \sim F_\theta$ of multivariate **max-stable distributions**, the full LLHs are usually not available.
- **Maximum likelihood estimation** is infeasible.
- **Goal:** Use full LLHs in Bayesian setup to
  - **improve frequentist efficiency**,
  - allow for **Bayesian methods in multivariate extremes**.

Max-stable distributions and partitions

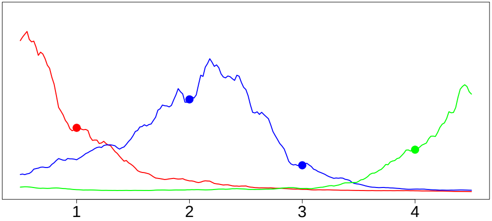
- Let $\mathbf{Z}$ be a d-dimensional max-stable distribution and $\mathbf{X}_1, \ldots, \mathbf{X}_n$ (with std. Fréchet margins) in its MDA.

Max-stable distributions and partitions

- Let $\mathbf{Z}$ be a d-dimensional max-stable distribution and $\mathbf{X}_1, \ldots, \mathbf{X}_n$ (with std. Fréchet margins) in its MDA.
- Componentwise maxima $\mathbf{M}_n = \max_{i=1}^n \mathbf{X}_i / n \overset{d}{\approx} \mathbf{Z}$.

Max-stable distributions and partitions

- Let $\mathbf{Z}$ be a d-dimensional max-stable distribution and $\mathbf{X}_1, \ldots, \mathbf{X}_n$ (with std. Fréchet margins) in its MDA.
- Componentwise maxima $\mathbf{M}_n = \max_{i=1}^n \mathbf{X}_i / n \overset{d}{\approx} \mathbf{Z}$.
- Partition Π_n of the set $\{1, \ldots, d\}$ of **occurrence times** of maxima, i.e., j and k in the same set if $M_{n,j}$ and $M_{n,k}$ come from same $\mathbf{X}_i$.



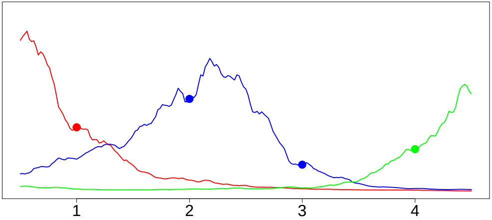
Example: $d = 4$, $n = 3$, $\Pi_n = \{\{1\}, \{2, 3\}, \{4\}\}$.

Max-stable distributions and partitions

- We have the weak limit as $n \to \infty$

$$(\mathbf{M}_n, \Pi_n) \xrightarrow{d} (\mathbf{Z}, \Pi),$$

where Π is the **limit partition** of occurrence times of $\mathbf{Z}$.



Example: $d = 4$, $n = 3$, $\Pi_n = \{\{1\}, \{2, 3\}, \{4\}\}$.

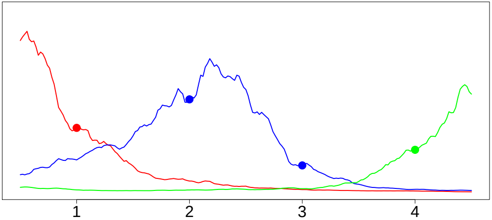
Max-stable distributions and partitions

- We have the weak limit as $n \to \infty$

$$(\mathbf{M}_n, \Pi_n) \xrightarrow{d} (\mathbf{Z}, \Pi),$$

where Π is the **limit partition** of occurrence times of $\mathbf{Z}$.

- Π_n and Π are distributions on the set $\mathcal{P}_\mathbf{d}$ **of all partitions** of $\{1, \ldots, d\}$.



Example: $d = 4$, $n = 3$, $\Pi_n = \{\{1\}, \{2, 3\}, \{4\}\}$.

Max-stable distributions and likelihoods

- $\mathbb{P}(\mathbf{Z} \leq \mathbf{z}) = e^{-V(\mathbf{z})}$, where V is the **exponent measure** of $\mathbf{Z}$.

Max-stable distributions and likelihoods

- $\mathbb{P}(\mathbf{Z} \leq \mathbf{z}) = e^{-V(\mathbf{z})}$, where V is the **exponent measure** of $\mathbf{Z}$.
- Likelihood of $\mathbf{Z}$ is

$$L_{\mathbf{Z}}(\mathbf{z}) = \sum_{\pi \in \mathcal{P}_d} L_{\mathbf{Z},\Pi}(\mathbf{z}, \pi),$$

where the Stephenson–Tawn LLH (2005, *Biometrika*) is **joint LLH of $\mathbf{Z}$ and Π** and has the explicit form

$$L_{\mathbf{Z},\Pi}(\mathbf{z}, \pi) = e^{-V(\mathbf{z})} \prod_{j=1}^{|\pi|} \left\{-V_{\pi^{(j)}}(\mathbf{z})\right\},$$

where $V_{\pi^{(j)}}$ are **partial derivatives** of V in directions $\pi^{(j)}$, and the partition $\pi = (\pi^{(1)}, \ldots, \pi^{(|\pi|)})$ consists of $|\pi|$ sets.

Max-stable distributions and likelihoods

- $\mathbb{P}(\mathbf{Z} \leq \mathbf{z}) = e^{-V(\mathbf{z})}$, where V is the **exponent measure** of $\mathbf{Z}$.
- Likelihood of $\mathbf{Z}$ is
$$L_{\mathbf{Z}}(\mathbf{z}) = \sum_{\pi \in \mathcal{P}_d} L_{\mathbf{Z},\Pi}(\mathbf{z}, \pi),$$
where the Stephenson–Tawn LLH (2005, *Biometrika*) is **joint LLH of $\mathbf{Z}$ and Π** and has the explicit form
$$L_{\mathbf{Z},\Pi}(\mathbf{z}, \pi) = e^{-V(\mathbf{z})} \prod_{j=1}^{|\pi|} \left\{-V_{\pi^{(j)}}(\mathbf{z})\right\},$$
where $V_{\pi^{(j)}}$ are **partial derivatives** of V in directions $\pi^{(j)}$, and the partition $\pi = (\pi^{(1)}, \ldots, \pi^{(|\pi|)})$ consists of $|\pi|$ sets.
- Since $|\mathcal{P}_d|$ is the **dth Bell number** (very large), the use of $L_{\mathbf{Z}}(z)$ is infeasible.

Parametric inference methods for $\mathbf{Z} \sim F_\theta$

With information on partition (occurrence times observed):

- Stephenson and Tawn (2005, *Biometrika*) use partition information and the joint LLH $\mathbf{L}_{\mathbf{Z},\boldsymbol{\Pi}}(\mathbf{z},\boldsymbol{\pi})$.
  But: **Possible bias** if $\Pi_n \neq \Pi$.

Parametric inference methods for $\mathbf{Z} \sim F_\theta$

With information on partition (occurrence times observed):

- Stephenson and Tawn (2005, *Biometrika*) use partition information and the joint LLH $\mathbf{L}_{\mathbf{Z},\mathbf{\Pi}}(\mathbf{z}, \boldsymbol{\pi})$.
  But: **Possible bias** if $\Pi_n \neq \Pi$.

- Wadsworth (2015, *Biometrika*) proposes **bias correction** by using $\mathbf{L}_{\mathbf{Z},\mathbf{\Pi_n}}(\mathbf{z}, \boldsymbol{\pi})$.

Parametric inference methods for $\mathbf{Z} \sim F_\theta$

With information on partition (occurrence times observed):

- Stephenson and Tawn (2005, *Biometrika*) use partition information and the joint LLH $\mathbf{L_{Z,\Pi}(z,\pi)}$.
  But: **Possible bias** if $\Pi_n \neq \Pi$.
- Wadsworth (2015, *Biometrika*) proposes **bias correction** by using $\mathbf{L_{Z,\Pi_n}(z,\pi)}$.

Without information on partition (occurrence times unobserved):

- Composite (pairwise) likelihood Padoan et al. (2010, *JASA*).
  But: **Losses in efficiency**.

Parametric inference methods for $\mathbf{Z} \sim F_\theta$

With information on partition (occurrence times observed):

- Stephenson and Tawn (2005, *Biometrika*) use partition information and the joint LLH $\mathbf{L}_{\mathbf{Z},\boldsymbol{\Pi}}(\mathbf{z}, \boldsymbol{\pi})$.
  But: **Possible bias** if $\Pi_n \neq \Pi$.
- Wadsworth (2015, *Biometrika*) proposes **bias correction** by using $\mathbf{L}_{\mathbf{Z},\boldsymbol{\Pi}_n}(\mathbf{z}, \boldsymbol{\pi})$.

Without information on partition (occurrence times unobserved):

- Composite (pairwise) likelihood Padoan et al. (2010, *JASA*).
  But: **Losses in efficiency**.
- Dombry et al. (2013, *Biometrika*) introduces **Gibbs sampler** to obtain conditional partitions of $\mathbf{L}_{\boldsymbol{\Pi}|\mathbf{Z}}(\boldsymbol{\pi}|\mathbf{z})$.

Parametric inference methods for $\mathbf{Z} \sim F_\theta$

With information on partition (occurrence times observed):

- Stephenson and Tawn (2005, *Biometrika*) use partition information and the joint LLH $\mathbf{L_{Z,\Pi}(z, \pi)}$.
  But: **Possible bias** if $\Pi_n \neq \Pi$.
- Wadsworth (2015, *Biometrika*) proposes **bias correction** by using $\mathbf{L_{Z,\Pi_n}(z, \pi)}$.

Without information on partition (occurrence times unobserved):

- Composite (pairwise) likelihood Padoan et al. (2010, *JASA*).
  But: **Losses in efficiency**.
- Dombry et al. (2013, *Biometrika*) introduces **Gibbs sampler** to obtain conditional partitions of $\mathbf{L_{\Pi|Z}(\pi|z)}$.
- Thibaud et al. (2015, `http://arxiv.org/abs/1506.07836`) use this Gibbs sampler to choose Π automatically and obtain posterior $\mathbf{L(\theta|z)}$ in a **Bayesian framework** for Brown–Resnick processes.

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.
- Establish Bayesian framework which uses full LLHs for **general max-stable distributions** (without observed partition information).

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.
- Establish Bayesian framework which uses full LLHs for **general max-stable distributions** (without observed partition information).
- **Improve (frequentist) efficiency** of estimates compared to existing methods (pairwise LLH,...)

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.
- Establish Bayesian framework which uses full LLHs for **general max-stable distributions** (without observed partition information).
- **Improve (frequentist) efficiency** of estimates compared to existing methods (pairwise LLH,...)
  - Without partition information.

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.
- Establish Bayesian framework which uses full LLHs for **general max-stable distributions** (without observed partition information).
- **Improve (frequentist) efficiency** of estimates compared to existing methods (pairwise LLH,...)
  - Without partition information.
  - With partition information: ignore partition information.

Goals

- **Recall**: MLE with $L_{\mathbf{Z}}(\mathbf{z})$ is infeasible.
- Establish Bayesian framework which uses full LLHs for **general max-stable distributions** (without observed partition information).
- **Improve (frequentist) efficiency** of estimates compared to existing methods (pairwise LLH,...)
  - Without partition information.
  - With partition information: ignore partition information.
- Explore applications of **Bayesian methodology in multivariate extremes**.

Bayesian framework

Parametric model $\mathbf{Z} \sim F_\theta$ with $\theta \in \Theta$.

Observations from $\mathbf{Z}$:	$\mathbf{z_1, \ldots, z_n} \in \mathbb{R}^{\mathbf{d}}$
Unobserved partitions:	$\boldsymbol{\pi}_\mathbf{1}, \ldots, \boldsymbol{\pi}_\mathbf{n} \in \boldsymbol{\mathcal{P}}_\mathbf{d}$
Model parameters:	$\boldsymbol{\theta} \in \boldsymbol{\Theta}$

Bayesian framework

Parametric model $\mathbf{Z} \sim F_\theta$ with $\theta \in \Theta$.

Observations from $\mathbf{Z}$:		$\mathbf{z_1, \ldots, z_n} \in \mathbb{R}^{\mathbf{d}}$
Unobserved partitions: (latent variables)		$\boldsymbol{\pi}_{\mathbf{1}}, \ldots, \boldsymbol{\pi}_{\mathbf{n}} \in \boldsymbol{\mathcal{P}}_{\mathbf{d}}$
Model parameters:	$\gamma \sim$	$\boldsymbol{\theta} \in \boldsymbol{\Theta}$

- **Bayesian setup:** Introduce prior γ on $\boldsymbol{\Theta}$ and try to obtain the posterior distribution $\mathbf{L}(\boldsymbol{\theta}, \boldsymbol{\pi}_{\mathbf{1}}, \ldots, \boldsymbol{\pi}_{\mathbf{n}} | \mathbf{z_1}, \ldots \mathbf{z_n})$.

Markov Chain Monte Carlo

- The posterior has to be evaluated by **MCMC**.

Markov Chain Monte Carlo

- The posterior has to be evaluated by **MCMC**.
- **Gibbs sampler** for partitions and **Metropolis–Hastings** for θ, using

$$
\begin{aligned}
L(\theta, \pi_1, \ldots, \pi_n | \mathbf{z}_1, \ldots \mathbf{z}_n) &\propto \gamma(\theta) \prod_{i=1}^{n} L(\mathbf{z}_i, \pi_i) \\
&\propto \gamma(\theta) \prod_{i=1}^{n} \prod_{j=1}^{|\pi_i|} \left\{ -V_{\pi_i^{(j)}}(\mathbf{z}_i) \right\}
\end{aligned}
$$

Markov Chain Monte Carlo

- The posterior has to be evaluated by **MCMC**.
- **Gibbs sampler** for partitions and **Metropolis–Hastings** for θ, using

$$L(\theta, \pi_1, \ldots, \pi_n | \mathbf{z}_1, \ldots \mathbf{z}_n) \propto \gamma(\theta) \prod_{i=1}^{n} L(\mathbf{z}_i, \pi_i)$$

$$\propto \gamma(\theta) \prod_{i=1}^{n} \prod_{j=1}^{|\pi_i|} \left\{ -V_{\pi_i^{(j)}}(\mathbf{z}_i) \right\}$$

- We need **derivatives of the exponent measure** $V_{\pi^{(j)}}(\mathbf{z})$.

Examples

- **Logistic model** with $\theta \in \Theta = (0,1)$:

$$V_{\pi^{(j)}}(\mathbf{z}) = \theta^{1-|\pi^{(j)}|} \frac{\Gamma(|\pi^{(j)}| - \theta)}{\Gamma(1-\theta)} \left(\sum_{k=1}^{d} z_k^{-1/\theta} \right)^{\theta - |\pi^{(j)}|} \prod_{k \in \pi^{(j)}} z_k^{-1-1/\theta}$$

Examples

- **Logistic model** with $\theta \in \Theta = (0,1)$:

$$V_{\pi^{(j)}}(\mathbf{z}) = \theta^{1-|\pi^{(j)}|} \frac{\Gamma(|\pi^{(j)}| - \theta)}{\Gamma(1-\theta)} \left(\sum_{k=1}^{d} z_k^{-1/\theta} \right)^{\theta - |\pi^{(j)}|} \prod_{k \in \pi^{(j)}} z_k^{-1-1/\theta}$$

- **Many other models:**
  - Brown–Resnick; cf. Thibaud et al. (2015, *Biometrika*)
  - Extremal-t
  - Reich–Shaby
  - Dirichlet
  - ...

Extremal dependence: max-stable data

- Simulate $n = 100$ samples $\mathbf{z}_1, \dots \mathbf{z}_n$ from the d-dim. **max-stable logistic model** for $\mathbf{Z}$ with parameter $\theta_0 \in \{0.1, 0.7, 0.9\}$; see Dombry et al. (2016, *Biometrika*).

Extremal dependence: max-stable data

- Simulate $n = 100$ samples $\mathbf{z}_1, \ldots \mathbf{z}_n$ from the d-dim. **max-stable logistic model** for $\mathbf{Z}$ with parameter $\theta_0 \in \{0.1, 0.7, 0.9\}$; see Dombry et al. (2016, *Biometrika*).
- Run MC with **uniform prior** γ on $(0, 1)$, and take **empirical median** of posterior $L(\theta|\mathbf{z}_1, \ldots \mathbf{z}_n)$ as point estimate $\hat{\theta}_{\mathrm{Bayes}}$.

Extremal dependence: max-stable data

- Simulate $n = 100$ samples $\mathbf{z}_1, \ldots \mathbf{z}_n$ from the d-dim. **max-stable logistic model** for $\mathbf{Z}$ with parameter $\theta_0 \in \{0.1, 0.7, 0.9\}$; see Dombry et al. (2016, *Biometrika*).
- Run MC with **uniform prior** γ on $(0, 1)$, and take **empirical median** of posterior $L(\theta|\mathbf{z}_1, \ldots \mathbf{z}_n)$ as point estimate $\hat{\theta}_{\mathrm{Bayes}}$.

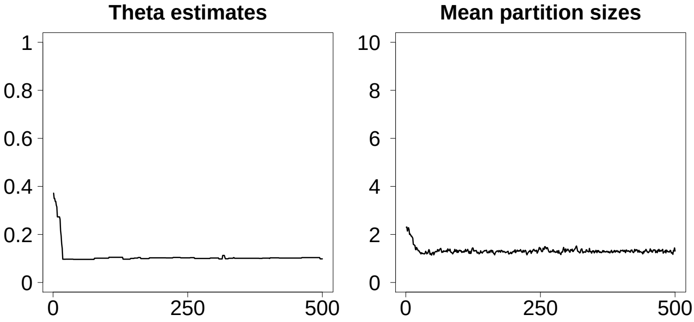


Figure : Markov chains whose stationary distributions are the posterior of θ_0 (left) and the mean partition size (right); $\boldsymbol{\theta_0 = 0.1}$, $d = 10$.

Extremal dependence: max-stable data

- Simulate $n = 100$ samples $\mathbf{z}_1, \ldots \mathbf{z}_n$ from the d-dim. **max-stable logistic model** for $\mathbf{Z}$ with parameter $\theta_0 \in \{0.1, 0.7, 0.9\}$; see Dombry et al. (2016, *Biometrika*).
- Run MC with **uniform prior** γ on $(0, 1)$, and take **empirical median** of posterior $L(\theta|\mathbf{z}_1, \ldots \mathbf{z}_n)$ as point estimate $\hat{\theta}_{\mathrm{Bayes}}$.

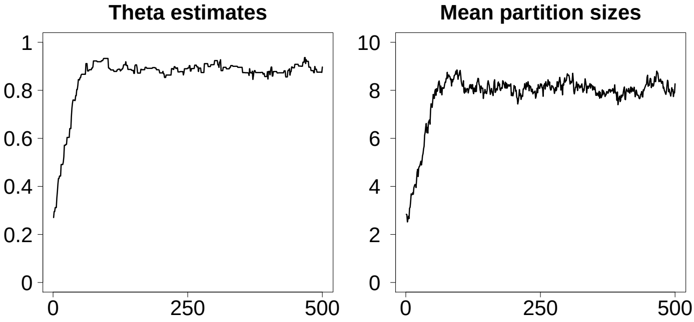


Figure : Markov chains whose stationary distributions are the posterior of θ_0 (left) and the mean partition size (right); $\boldsymbol{\theta_0 = 0.9}$, $d = 10$.

Extremal dependence: max-stable data

- Compare with pairwise composite LLH estimator $\hat{\theta}_{\mathrm{PL}}$.

Extremal dependence: max-stable data

- Compare with pairwise composite LLH estimator $\hat{\theta}_{\mathrm{PL}}$.

	$\theta_0 = 0.1$			$\theta_0 = 0.7$			$\theta_0 = 0.9$		
d	6	10	50	6	10	50	6	10	50
$\mathrm{Bias}(\hat{\theta}_{\mathrm{Bayes}})$	2	2	2	10	6	1	-6	-3	2
$s(\hat{\theta}_{\mathrm{Bayes}})$	36	27	12	240	179	79	239	182	84
$\mathrm{Bias}(\hat{\theta}_{\mathrm{PL}})$	1	0	2	13	12	16	26	31	41
$s(\hat{\theta}_{\mathrm{PL}})$	40	30	13	275	237	173	313	273	246

Table : Sample bias and standard deviation of $\hat{\theta}_{\mathrm{Bayes}}$ and $\hat{\theta}_{\mathrm{PL}}$, estimated from 1500 estimates; figures multiplied by 10000.

Observations:

- Posterior median $\hat{\theta}_{\mathrm{Bayes}}$ is **unbiased**.
- Substantially **reduced std. deviations** by using full LLHs.

Extremal dependence: max-stable data

dimension d	$\theta_0 = 0.1$ 6	10	50	$\theta_0 = 0.7$ 6	10	50	$\theta_0 = 0.9$ 6	10	50
$\frac{MSE(\hat{\theta}_{\text{Bayes}})}{MSE(\hat{\theta}_{\text{PL}})}$	82	83	78	76	57	21	58	44	11

Table : Relative efficiencies (**%**) of $\hat{\theta}_{\text{Bayes}}$ compared to $\hat{\theta}_{\text{PL}}$, estimated from 1500 estimates.

Observations:

- Increasing efficiency gain for **lower dependence** and **higher dimensions**; see also Huser et al. (2015, *Extremes*)).

Extremal dependence: max-domain of attraction

- For $i = 1, \ldots, n$, $n = 100$, simulate $b = 50$ samples $\mathbf{x}_{1,i}, \ldots \mathbf{x}_{b,i}$ from **outer power Clayton copula** in MDA of logistic model with parameter θ_0.

Extremal dependence: max-domain of attraction

- For $i = 1, \ldots, n$, $n = 100$, simulate $b = 50$ samples $\mathbf{x}_{1,i}, \ldots \mathbf{x}_{b,i}$ from **outer power Clayton copula** in MDA of logistic model with parameter θ_0.
- Take block maxima $\tilde{\mathbf{z}}_i = \max_{j=1,\ldots,b} \mathbf{x}_{j,i}$ as input data.

Extremal dependence: max-domain of attraction

- For $i = 1, \ldots, n$, $n = 100$, simulate $b = 50$ samples $\mathbf{x}_{1,i}, \ldots \mathbf{x}_{b,i}$ from **outer power Clayton copula** in MDA of logistic model with parameter θ_0.
- Take block maxima $\tilde{\mathbf{z}}_i = \max_{j=1,\ldots,b} \mathbf{x}_{j,i}$ as input data.
- Compare $\hat{\theta}_{\mathrm{Bayes}}$, $\hat{\theta}_{\mathrm{PL}}$ based on $\{\tilde{\mathbf{z}}_i\}_i$, and ST estimator $\hat{\theta}_{\mathrm{ST}}$ and its bias-corrected version $\hat{\theta}_{\mathrm{W}}$ based on $\{(\tilde{\mathbf{z}}_i, \pi_i)\}_i$.

Extremal dependence: max-domain of attraction

- For $i = 1, \ldots, n$, $n = 100$, simulate $b = 50$ samples $\mathbf{x}_{1,i}, \ldots \mathbf{x}_{b,i}$ from **outer power Clayton copula** in MDA of logistic model with parameter θ_0.
- Take block maxima $\tilde{\mathbf{z}}_i = \max_{j=1,\ldots,b} \mathbf{x}_{j,i}$ as input data.
- Compare $\hat{\theta}_{\text{Bayes}}$, $\hat{\theta}_{\text{PL}}$ based on $\{\tilde{\mathbf{z}}_i\}_i$, and ST estimator $\hat{\theta}_{\text{ST}}$ and its bias-corrected version $\hat{\theta}_{\text{W}}$ based on $\{(\tilde{\mathbf{z}}_i, \pi_i)\}_i$.

	$\theta_0 = 0.1$		$\theta_0 = 0.7$		$\theta_0 = 0.9$	
dimension d	6	10	6	10	6	10
$\hat{\theta}_{\text{Bayes}}$	100	100	100	100	100	100
$\hat{\theta}_{\text{PL}}$	83	81	71	60	62	51
$\hat{\theta}_{\text{ST}}$	90	97	47	23	16	7
$\hat{\theta}_{\text{W}}$	90	97	110	70	70	26

Table : Relative efficiencies as $\text{MSE}(\hat{\theta}_{\text{Bayes}})/\text{MSE}(\hat{\theta})$ (in **%**) for different estimators $\hat{\theta}$; estimated from 1500 estimates.

Observations:

- **Efficiency gain** by using full information remains large.
- Automatic choice of partition in $\hat{\theta}_{\text{Bayes}}$ **more robust** than ST- likelihood with fixed partition; **no bias correction** needed.

More results and work in progress

Estimation of marginal parameters:

- **Simultaneous estimation** of marginals and dependence parameters possible.
- Astonishing observation: This substantially **reduces MSE for shape estimation** (weak improvement for location and scale).

More results and work in progress

Estimation of marginal parameters:

- **Simultaneous estimation** of marginals and dependence parameters possible.
- Astonishing observation: This substantially **reduces MSE for shape estimation** (weak improvement for location and scale).

Bayesian hypothesis testing:

- Ex.: Testing asymptotic **independence against dependence**

$$H_0 : \theta_0 = 1 \text{ against } H_1 = \theta_0 \in (0,1).$$

- Compare **posterior probabilities** $\mathbb{P}(H_0|\mathbf{z}_1, \ldots, \mathbf{z}_n)$ and $\mathbb{P}(H_1|\mathbf{z}_1, \ldots, \mathbf{z}_n)$.
- Many other tests are possible (symmetry, regression coefficients,...).

More results and work in progress

Estimation of marginal parameters:

- **Simultaneous estimation** of marginals and dependence parameters possible.
- Astonishing observation: This substantially **reduces MSE for shape estimation** (weak improvement for location and scale).

Bayesian hypothesis testing:

- Ex.: Testing asymptotic **independence against dependence**

$$H_0 : \theta_0 = 1 \text{ against } H_1 = \theta_0 \in (0, 1).$$

- Compare **posterior probabilities** $\mathbb{P}(H_0 | \mathbf{z}_1, \ldots, \mathbf{z}_n)$ and $\mathbb{P}(H_1 | \mathbf{z}_1, \ldots, \mathbf{z}_n)$.
- Many other tests are possible (symmetry, regression coefficients,...).

Asymptotic limit results for posterior median.

References

C. Dombry, S. Engelke, and M. Oesting.
Exact simulation of max-stable processes.
Biometrika, 2016.
To appear.

C. Dombry, F. Eyi-Minko, and M. Ribatet.
Conditional simulation of max-stable processes.
Biometrika, 100(1):111–124, 2013.

R. Huser, A. C. Davison, and M. G. Genton.
Likelihood estimators for multivariate extremes.
Extremes, 19:79–103, 2016.

S. A. Padoan, M. Ribatet, and S. A. Sisson.
Likelihood-based inference for max-stable processes.
J. Amer. Statist. Assoc., 105:263–277, 2010.

A. Stephenson and J. A. Tawn.
Exploiting occurrence times in likelihood inference for componentwise maxima.
Biometrika, 92(1):213–227, 2005.

E. Thibaud, J. Aalto, D. S. Cooley, A. C. Davison, and J. Heikkinen.
Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures.
Available from `http://arxiv.org/abs/1506.07836`, 2015.

J. L. Wadsworth.
On the occurrence times of componentwise maxima and bias in likelihood inference for multivariate max-stable distributions.
Biometrika, 102:705–711, 2015.