Rank-based Estimation under Asymptotic Dependence and Independence, with Applications to Spatial Extremes

Michaël Lalancette[1] Sebastian Engelke[2] Stanislav Volgushev[1]

[1]Department of Statistical Sciences, University of Toronto

[2]Research Center for Statistics, University of Geneva

SSC Annual Meeting, 08/06/2021

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Given 50 years of rainfall data, EVT finds

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Given 50 years of rainfall data, EVT finds
  - the 1-in-100-years rainfall (or even 1-in-10 000 years)

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Given 50 years of rainfall data, EVT finds
  - the 1-in-100-years rainfall (or even 1-in-10 000 years)
  - the probability of at least $x$mm rainfall (x larger than anything observed)

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Given 50 years of rainfall data, EVT finds
  - the 1-in-100-years rainfall (or even 1-in-10 000 years)
  - the probability of at least $x$mm rainfall (x larger than anything observed)
- Mathematically, EVT uses largest data points to estimate the whole tail of one (or multiple) random variable(s)

Tail dependence

- Typical problem: $(X, Y) =$ (rainfall at location 1, rainfall at location 2)

Tail dependence

- Typical problem: $(X, Y) =$ (rainfall at location 1, rainfall at location 2)
- Estimate

$$\mathbb{P}(X \geq u, Y \geq v),$$

for large (unobserved) values u, v

Tail dependence

- Typical problem: $(X, Y) =$ (rainfall at location 1, rainfall at location 2)
- Estimate

$$\mathbb{P}(X \geq u, Y \geq v),$$

  for large (unobserved) values u, v
- Univariate methods $\Rightarrow$ estimation of marginal tails of X and Y

Tail dependence

- Typical problem: $(X, Y) =$ (rainfall at location 1, rainfall at location 2)
- Estimate

$$\mathbb{P}(X \geq u, Y \geq v),$$

  for large (unobserved) values u, v
- Univariate methods $\Rightarrow$ estimation of marginal tails of X and Y
- Dependence structure in tail regions

Tail dependence

- Copula-based approach: represent the tail dependence structure by

$$R(x, y) := \lim_{t \downarrow 0} t^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty),$$

F_1 and F_2 are marginal d.f. of X, Y

Tail dependence

- Copula-based approach: represent the tail dependence structure by

$$R(x, y) := \lim_{t \downarrow 0} t^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty),$$

  F_1 and F_2 are marginal d.f. of X, Y

- Extrapolation: limit allows to write

$$\mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty) \approx tR(x, y)$$

Tail dependence

- Copula-based approach: represent the tail dependence structure by

$$R(x,y) := \lim_{t \downarrow 0} t^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty),$$

  F_1 and F_2 are marginal d.f. of X, Y

- Extrapolation: limit allows to write

$$\mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty) \approx tR(x,y)$$

- So estimation of $R \Rightarrow$ estimation of far tail probabilities

Asymptotic independence

- What if X and Y are independent? Then $R(x, y) = 0$

Asymptotic independence

- What if X and Y are independent? Then $R(x, y) = 0$
- Many distributions satisfy $R(x, y) = 0$ (*asymptotically independent*)

Asymptotic independence

- What if X and Y are independent? Then $R(x, y) = 0$
- Many distributions satisfy $R(x, y) = 0$ (*asymptotically independent*)
- Example: (X, Y) is jointly Normal with correlation $\rho \in (-1, 1)$

Asymptotic independence

- What if X and Y are independent? Then $R(x, y) = 0$
- Many distributions satisfy $R(x, y) = 0$ (*asymptotically independent*)
- Example: (X, Y) is jointly Normal with correlation $\rho \in (-1, 1)$
- Modelling/inference based on $R \Rightarrow$ cannot distinguish different asympt. independent distributions

Alternative tail representation: the function c

- Instead of R, we consider the function

$$c(x,y) := \lim_{t\downarrow 0} q(t)^{-1}\mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

Alternative tail representation: the function c

- Instead of R, we consider the function

$$c(x,y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- q is chosen so that the limit is non-trivial

Alternative tail representation: the function c

- Instead of R, we consider the function

$$c(x,y) := \lim_{t\downarrow 0} q(t)^{-1}\mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- q is chosen so that the limit is non-trivial
- As opposed to R, c has non-trivial form whether X, Y are asympt. independent or dependent

Alternative tail representation: the function c

- Instead of R, we consider the function

$$c(x,y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- q is chosen so that the limit is non-trivial
- As opposed to R, c has non-trivial form whether X, Y are asympt. independent or dependent
- Examples:
  - X and Y are independent: $c(x,y) = xy$

Alternative tail representation: the function c

- Instead of R, we consider the function

$$c(x,y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- q is chosen so that the limit is non-trivial
- As opposed to R, c has non-trivial form whether X, Y are asympt. independent or dependent
- Examples:
  - X and Y are independent: $c(x,y) = xy$
  - X and Y are Normal with correlation ρ: $c(x,y) = (xy)^{1/(1+\rho)}$

Non-parametric estimation (unfeasible)

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$

Non-parametric estimation (unfeasible)

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$
- Wish to estimate

$$c(x, y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

Non-parametric estimation (unfeasible)

- (iid) observations $(X_1, Y_1), \dots, (X_n, Y_n)$
- Wish to estimate

$$c(x,y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- For a small $t_n > 0$,

$$\widehat{c}_n(x,y) := q(t_n)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left\{ \widehat{F}_1(X_i) \geq 1 - t_n x, \widehat{F}_2(Y_i) \geq 1 - t_n y \right\},$$

where $\widehat{F}_j$ are the empirical d.f.

Non-parametric estimation (unfeasible)

- (iid) observations $(X_1, Y_1), \dots, (X_n, Y_n)$
- Wish to estimate

$$c(x, y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- For a small $t_n > 0$,

$$\widehat{c}_n(x, y) := q(t_n)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left\{\widehat{F}_1(X_i) \geq 1 - t_n x, \widehat{F}_2(Y_i) \geq 1 - t_n y\right\},$$

where $\widehat{F}_j$ are the empirical d.f.

Parametric method

- Suppose $c = c_{\theta_0} \in \{c_\theta : \theta \in \Theta\}$

Parametric method

- Suppose $c = c_{\theta_0} \in \{c_\theta : \theta \in \Theta\}$
- Estimate θ_0 by

$$\widehat{\theta} := \arg\min_{\theta} \left\| \int g\, c_\theta - \int g\, \widehat{c}_n \right\|,$$

for a weight function $g : [0, \infty)^2 \to \mathbb{R}^p$

Parametric method

- Suppose $c = c_{\theta_0} \in \{c_\theta : \theta \in \Theta\}$
- Estimate θ_0 by

$$(\widehat{\theta}, \widehat{\sigma}) := \arg\min_{\theta, \sigma} \left\| \sigma \int g\, c_\theta - q(t_n) \int g\, \widehat{c}_n \right\|,$$

for a weight function $g : [0, \infty)^2 \to \mathbb{R}^p$

Parametric method

- Suppose $c = c_{\theta_0} \in \{c_\theta : \theta \in \Theta\}$
- Estimate θ_0 by

$$(\widehat{\theta}, \widehat{\sigma}) := \arg\min_{\theta, \sigma} \left\| \sigma \int g\, c_\theta - q(t_n) \int g\, \widehat{c}_n \right\|,$$

  for a weight function $g : [0, \infty)^2 \to \mathbb{R}^p$

- If $\widehat{c}_n \approx c$, hopefully $\widehat{\theta} \approx \theta_0$ and $\widehat{\sigma} \approx q(t_n)$

Spatial tail dependence

- Typical problem: $\{X(s) : s \in \mathbb{R}^2\}$ = rainfall over a spatial domain

Spatial tail dependence

- Typical problem: $\{X(s) : s \in \mathbb{R}^2\}$ = rainfall over a spatial domain
- Simultaneous estimation of marginal tails of $X(s)$

Spatial tail dependence

- Typical problem: $\{X(s) : s \in \mathbb{R}^2\}$ = rainfall over a spatial domain
- Simultaneous estimation of marginal tails of $X(s)$
- Given observations at d locations, estimate

$$\mathbb{P}(F_s(X(s)) \geq 1 - tx(s), \text{for each } s \in \mathcal{S}), \qquad (*)$$

for small t and finite collection of locations $\mathcal{S}$

Spatial tail dependence

- Typical problem: $\{X(s) : s \in \mathbb{R}^2\}$ = rainfall over a spatial domain
- Simultaneous estimation of marginal tails of $X(s)$
- Given observations at d locations, estimate

$$\mathbb{P}(F_s(X(s)) \geq 1 - tx(s), \text{for each } s \in \mathcal{S}), \qquad (*)$$

  for small t and finite collection of locations $\mathcal{S}$
- Locations $\mathcal{S}$ potentially unobserved

Spatial tail dependence

- Typical problem: $\{X(s) : s \in \mathbb{R}^2\}$ = rainfall over a spatial domain
- Simultaneous estimation of marginal tails of $X(s)$
- Given observations at d locations, estimate

$$\mathbb{P}(F_s(X(s)) \geq 1 - tx(s), \text{for each } s \in \mathcal{S}), \tag{$*$}$$

  for small t and finite collection of locations $\mathcal{S}$
- Locations $\mathcal{S}$ potentially unobserved
- Spatial tail dependence model: probabilities $(*)$ for all collections $\mathcal{S}$

Spatial tail dependence

- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

Spatial tail dependence

- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

- Spatial model (θ_{spatial}) $\Rightarrow$ Bivariate model for each pair ($\theta_{\text{pair}}(\theta_{\text{spatial}})$)

Spatial tail dependence

- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

- Spatial model (θ_{spatial}) $\Rightarrow$ Bivariate model for each pair ($\theta_{\text{pair}}(\theta_{\text{spatial}})$)
- Bivariate methodology $\Rightarrow$ Estimator $\widehat{\theta}_{\text{pair}}$ for each pair

Spatial tail dependence

- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

- Spatial model (θ_{spatial}) $\Rightarrow$ Bivariate model for each pair ($\theta_{\text{pair}}(\theta_{\text{spatial}})$)
- Bivariate methodology $\Rightarrow$ Estimator $\widehat{\theta}_{\text{pair}}$ for each pair
- Fit spatial model via least squares:

$$\left\| \theta_{\text{pair}}(\theta_{\text{spatial}}) - \widehat{\theta}_{\text{pair}} \right\|^2$$

Spatial tail dependence

- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

- Spatial model (θ_{spatial}) $\Rightarrow$ Bivariate model for each pair ($\theta_{\text{pair}}(\theta_{\text{spatial}})$)
- Bivariate methodology $\Rightarrow$ Estimator $\widehat{\theta}_{\text{pair}}$ for each pair
- Fit spatial model via least squares:

$$\sum_{\text{pair} \in \text{observed pairs}} \left\| \theta_{\text{pair}}(\theta_{\text{spatial}}) - \widehat{\theta}_{\text{pair}} \right\|^2$$

Spatial tail dependence

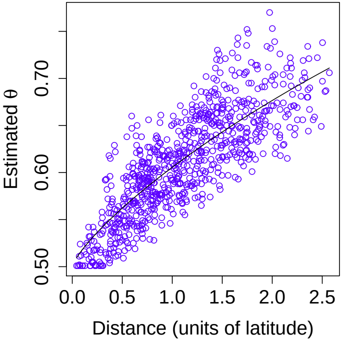
- Key: popular models are identifiable by finite number of pairwise prob.

$$\mathbb{P}(F_{s_1}(X(s_1)) \geq 1 - tx(s_1), F_{s_2}(X(s_2)) \geq 1 - tx(s_2))$$

- Spatial model (θ_{spatial}) $\Rightarrow$ Bivariate model for each pair ($\theta_{\text{pair}}(\theta_{\text{spatial}})$)
- Bivariate methodology $\Rightarrow$ Estimator $\widehat{\theta}_{\text{pair}}$ for each pair
- Fit spatial model via least squares:

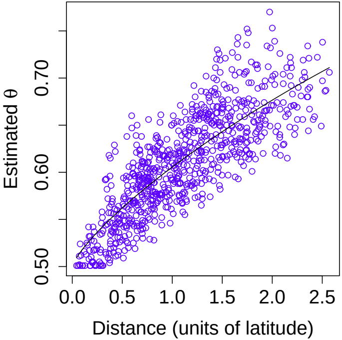
$$\widehat{\theta}_{\text{spatial}} := \arg\min_{\theta_{\text{spatial}}} \sum_{\text{pair} \in \text{observed pairs}} \left\| \theta_{\text{pair}}(\theta_{\text{spatial}}) - \widehat{\theta}_{\text{pair}} \right\|^2$$

Rainfall over Victoria region (AUS), 1967-2017



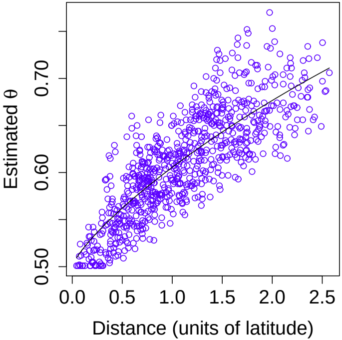
- 40 locations $\Rightarrow$ 780 pairs

Rainfall over Victoria region (AUS), 1967-2017



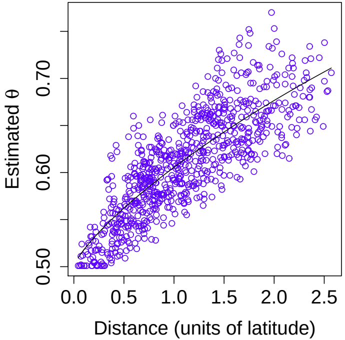
- 40 locations $\Rightarrow$ 780 pairs
- Fractal inverted Brown–Resnick process

Rainfall over Victoria region (AUS), 1967-2017



- 40 locations $\Rightarrow$ 780 pairs
- Fractal inverted Brown–Resnick process
- Pairwise parameter depends on Euclidean distance

Rainfall over Victoria region (AUS), 1967-2017



- 40 locations $\Rightarrow$ 780 pairs
- Fractal inverted Brown–Resnick process
- Pairwise parameter depends on Euclidean distance
- Blue dots $= \widehat{\theta}_{\text{pair}}$, black curve $=$ fitted spatial model

Thanks for your attention! Questions?

Lalancette, M., S. Engelke, and S. Volgushev (2021). Rank-based estimation under asymptotic dependence and independence, with applications to spatial extremes. *Ann. Stat.*, to appear.

Other contributions in the paper:

- Novel examples of (bivariate and spatial) tail models
- Asymptotic theory for all (bivariate and spatial) estimators
- Simulation studies

mic-lalancette.github.io