Modelling the extremal dependence with Bernstein polynomials

Simone Padoan
Bocconi University of Milan, Italy
simone.padoan@unibocconi.it

UCA-MERE in NCAR, Boulder, Colorado, US
25-28 April, 2016

Thanks to: Giulia Marcon, Isadora Antoniano.
1. Motivation;

2. Introduction and background;

3. Polynomial representation of the extremal dependence in Bernstein form;

4. Bayesian modelling;

5. Simulation results and real data application;
Motivations

- Our interest is on estimating the probability that simultaneous extreme events occur, in different locations of a region.

**Figure**: French map with 49 weather stations. Weekly precipitation maxima in the period 1993–2011.
Literature

• Theoretical background: de Haan L. and A. Ferreira (2006), Beirlant, J. et al. (2001);


Other papers on the same area

• modeling the extremal dependence: Boldi and Davison (2007), Sabourin and Naveau (2014), Guillotte et al. (2011), de Haan L. and A. Ferreira (2006) Ch 7, to name a few....
Preliminary: bivariate max-stable distributions

1) Let $X \sim F(x)$. If $F$ is in the max-domain of attraction of an extreme value distribution $G$, then

$$F^n(a_n x + b_n) \rightarrow G(x), \quad n \rightarrow \infty \quad x \in \mathbb{R}^2.$$  

The marginal distributions are members of the GEV (de Haan L. and A. Ferreira, 2006).

2) If we assume to work with common unit Fréchet margins, then

$$V_0(y) = 2 \int_{[0,1]} \max \left( \frac{w}{y_1}, \frac{1-w}{y_2} \right) H(dw), \quad y \in \mathbb{R}_+^2, \quad (1)$$

where $V_0(y) := -\log G_0(y)$ is the exponent function and $H$ is a probability distribution function on $[0, 1]$, named angular measure.

3) We can express the dependence in two ways, working with: $H$ or $V_0$. 

5/23
1st representation: angular measure

(C1) $H$ is satisfying the mean constraints

$$\int_{[0,1]} w \, H(dw) = 1/2 = \int_{[0,1]} (1 - w) \, H(dw) \quad (2)$$

- We assume that $H$ is such that

\[
\begin{align*}
    p_0 &= H(\{0\}) \quad \text{if} \quad w = 0, \\
    \hat{h}(w) &= H(\{1\}) \quad \text{if} \quad 0 < w < 1, \\
    p_1 &= H(\{1\}) \quad \text{if} \quad w = 1,
\end{align*}
\]
2nd representation: Pickands dependence function

- Take \( r = y_1 + y_2 \) and \( t = y_1 / r \), then for \( t \in [0, 1] \) we have

\[
V_0(y) = \left( \frac{1}{y_1} + \frac{1}{y_2} \right)^2 \int_{[0,1]} \max \{ w(1-t), (1-w)t \} H(dw) =: A(t),
\]

- The function \( A : [0, 1] \mapsto [1/2, 1] \) satisfies the conditions:

(C2) is convex on \([0, 1]\);

(C3) \( 1/2 \leq \max(t, 1-t) \leq A(t) \leq 1; \)
Relationship between $A$ and $H$

- It can be shown that (Beirlant, J. et. al, 2004)

$$A(t) = 1 + 2 \int_0^t H([0, w])dw - t, \quad t \in [0, 1],$$

and

$$A'(w) = 2H([0, w]) - 1 \quad (3)$$

see Beirlant, J. et. al (2004) for details.

- Rephrasing the densities we have

$$\begin{cases}
    p_0 = \frac{A'(0) + 1}{2} & \text{if } w = 0, \\
    \dot{h}(w) = \frac{A''(w)}{2} & \text{if } 0 < w < 1, \\
    p_1 = \frac{1 - A'(1)}{2} & \text{if } w = 1,
\end{cases}$$

where $A'(1) = \sup_{w \in [0,1]} A'(w)$. 
Bernstein Polynomial representation of $A$

$$A_k(t) = \sum_{j \leq k} \beta_j b_j(t; k), \quad t \in [0, 1], \quad j = 0, \ldots, k, \quad (4)$$

where $\beta_j$ are coefficients and

$$b_j(t; k) = \binom{k}{j} t^j (1 - t)^{k-j}, \quad \binom{k}{j} = \frac{k!}{j!(k - j)!}, \quad j = 0, \ldots, k.$$

are Bernstein polynomial bases.

$A_k$ is a valid Pickands function, satisfying conditions (C2)–(C3), if

(R1) $\beta_0 = \beta_k = 1 \geq \beta_j$ for $j = 1, \ldots, k - 1$;

(R2) $\beta_1 = \frac{k-1+2p_0}{k}$ and $\beta_{k-1} = \frac{k-1+2p_1}{k}$;

(R3) $\beta_{j+2} - 2\beta_{j+1} + \beta_j \geq 0$ for $j = 0, \ldots, k - 2$. 


Link with the angular measure

- Formula (3) suggests

\[
H_k(w) = \frac{1}{2} \left\{ \sum_{j \leq k-1} k(\beta_{j+1} - \beta_j) b_j(w; k - 1) + 1 \right\}
\]

\[
= \cdots = \sum_{j \leq k-1} \left( \frac{k(\beta_{j+1} - \beta_j) + 1}{2} \right) \cdot \eta_j b_j(w; k - 1),
\]

where \( b_j(w; k - 1) \) satisfies

\[
b_j(w; k - 1) = w^{j-1} (1-w)^{k-1-j},
\]

\[
\int_0^1 w h_k(w) \, dw + p_1 = 1/2 = p_0 + \int_0^1 (1 - w) h_k(w) \, dw,
\]

where \( h_k(w) = H'_k(w) \) and so \( H_k(w) \) is a valid angular measure!!

- It can be checked that the \( \eta \) coefficients satisfy the restrictions:

\[
(R4) \quad 0 \leq p_0 = \eta_0 \leq \eta_1 \leq \cdots \leq \eta_{k-2} \leq \eta_{k-1} = 1 - p_1 \leq 1;
\]

\[
(R5) \quad \sum_{j \leq k-1} \eta_j = k/2;
\]

implying

\[
\int_0^1 w h_k(w) \, dw + p_1 = 1/2 = p_0 + \int_0^1 (1 - w) h_k(w) \, dw,
\]
Summing up

**Result:** A Bernstein polynomial $A_k$ satisfying restrictions (R1)–(R3) implies that the Bernstein polynomial $H_k$, defined by the coefficients

$$
\eta_j = \frac{1}{2} + \frac{k}{2} (\beta_{j+1} - \beta_j), \quad j = 0, \ldots, k - 1,
$$

is a **valid angular measure** satisfying the condition (C1).

Vice-versa, a Bernstein polynomial $H_k$ satisfying restrictions (R4) and (R5), implies that the Bernstein polynomial $A_k$, defined by the coefficients

$$
\beta_{j+1} = \frac{1}{k} \left\{ 2 \sum_{i=0}^{j} \eta_i + k - j - 1 \right\}, \quad j = 0, \ldots, k - 1,
$$

is a **valid Pickands dependence function** satisfying conditions (C2)-(C3).
Bayesian non-parametric model

1. We define the **angular measure** with a Bernstein polynomial $H_k$.

2. We construct a **prior distribution** on $\mathcal{H}$, i.e. the space of probability measures satisfying (C1).

3. A prior on $\mathcal{H}$ is given by the joint distribution

   \[ \Pi(\theta) = \Pi(\eta_k \mid k) \Pi(k), \quad \eta_k = (\eta_0, \ldots, \eta_{k-1}), \]

   defined on $\theta := (k, \eta_k)$, such that (R4)-(R5) are satisfied.

4. We define the **likelihood function** using the Pickands, i.e.

   \[ L(y \mid \theta) = G''_0(y), \quad V_0(y) = (y_1^{-1} + y_2^{-1})A_k(t). \]

5. The **posterior distribution** is proportional to

   \[ \Pi(\theta \mid y) \propto L(y \mid \theta) \cdot \Pi(\theta). \]
Prior distribution

- The complete prior is:

\[ \Pi(\theta) = \Pi(\eta_1, \ldots, \eta_{k-2} | p_1, p_0, k) \Pi(p_1 | p_0) \Pi(p_0) \Pi(k), \]

where

A) \( \Pi(k) := \text{Poi}(k + 3 | \kappa), \) for \( k = 1, 2 \ldots, \) and \( \kappa > 0; \)

B) \( \Pi(p_0) := \text{Unif}(0, a), \) with \( a < 1/2; \)

C) \( \Pi(p_1 | p_0) := \text{Unif}(0, a - p_0); \)

D) \( \Pi(\eta_k | p_1, p_0, k) := \text{Unif}(\mathcal{I}_1) \prod_{j=2}^{k-2} \text{Unif}(\mathcal{I}_j | \eta_1, \ldots, \eta_{j-1}), \)

where

\[ \mathcal{I}_j = \left\{ \max \left[ \eta_{j-1}, \frac{k}{2} - \sum_{i=0}^{j-1} \eta_i + (k - i)(p_1 - 1) \right] ; \right\} \]

\[ \frac{1}{k-j} \left( \frac{k}{2} - \sum_{i=0}^{j-1} \eta_i - 1 + p_1 \right). \]
The log-likelihood function is given by

\[
\ell(y_{1:n} | \theta) = - \sum_{i=1}^{n} \left( \frac{1}{y_{1,i}} + \frac{1}{y_{2,i}} \right) A_k(t_i)
\]

\[
+ \sum_{i=1}^{n} \log \left[ \frac{\{A_k(t) - tA'_k(t)\}\{A_k(t) + (1 - t)A'_k(t)\}}{(y_{1,i}y_{2,i})^2} \right]
\]

\[
+ \frac{A''_k(t)}{(y_{1,i} + y_{2,i})^3}
\],

where

\[
A_k(t) = \sum_{j \leq k} \beta_j b_j(t; k), \quad t \in [0, 1], \quad j = 0, \ldots, k.
\]
Inference I

- The model inference is based on a **trans-dimensional** MCMC posterior simulation scheme. That is:

1) We extend $\Pi(\eta_k, k)$ to

$$
\Pi(\eta_\infty, k) = \Pi(\eta_k | k) \Pi(k) \prod_{j \geq k} \Pi(\eta_j), \quad \eta_\infty = (\eta_0, \eta_1, \ldots),
$$

where $\Pi(\eta_j)$ is any fully known distribution.

2) At the Markov chain state $s$, we update $(\eta_k^{(s)}, k^{(s)})$ by a **Metropolis-Hasting** step with proposal distribution

$$
q(\eta_\infty, k | \eta_\infty^{(s)}, k^{(s)}) = q_\eta(\eta_k | k) q_k(k | k^{(s)}) \prod_{j \geq k} q_\eta(\eta_j),
$$

where $q_\eta(\eta_k | k)$ and $q_\eta(\eta_j)$ coincides with the prior distributions.
Inference II

- and

\[ q_k \left( k = k^{(s)} + 1 \mid k^{(s)} \right) = \begin{cases} 
1 & \text{if } k^{(s)} = 3 \\
1/2 & \text{if } k^{(s)} > 3
\end{cases} \]

- and

\[ q_k \left( k = k^{(s)} - 1 \mid k^{(s)} \right) = \begin{cases} 
0 & \text{if } k^{(s)} = 3 \\
1/2 & \text{if } k^{(s)} > 3.
\end{cases} \]

3) Given the current state \( s \), the acceptance probability for the proposal, indexed by \( s + 1 \), is, for any \( k^{(s)} > 3 \), equal to

\[
p \left( k^{(s+1)}, \eta_{k^{(s+1)}}, k^{(s)}, \eta_{k^{(s)}} \right) = \frac{(k^{(s)} - 3)! \ k^{(s+1) - k^{(s)}}}{(k^{(s+1)} - 3)!} \ \frac{L(y_{1:n} \mid k^{(s+1)}, \eta_{k^{(s+1)}})}{L(y_{1:n} \mid k^{(s)}, \eta_{k^{(s)}})}
\]

and for \( k^{(s)} = 3 \) we have the multiplying factor \( 1/2 \).
Monthly maxima of daily rainfall

Monthly maxima of hourly rainfall

Aéroport de Paris–Orly

Aérodrome de Brétigny

Aéroport d'Aubenas

Montelimar
Estimated extremal dependence

1) Aéroport de Paris versus Aérodrome de Brétigny

2) Aéroport d’Aubenas versus Montelimar
An important goal is to **predict** the probability of future simultaneous extremes. We can compute e.g.

\[
P(Y_1 > y_1^*, Y_2 > y_2^* | Y_{1:n}) = \int_{\theta \in \Theta} P(Y_1 > y_1^*, Y_2 > y_2^* | \theta) \Pi^n(\theta | Y_{1:n}) d\theta
\]

where \( y_1^*, y_2^* > 0 \) are **unobserved high thresholds**.

With our framework, for high thresholds we have

\[
P(Y_1 > y_1^*, Y_2 > y_2^* | \theta) \approx \frac{1}{k} \sum_{j=0}^{k-2} (\eta_{j+1} - \eta_j) \left\{ \frac{j + 1}{y_1} B(v|j + 2, k - j - 1) + \frac{k - j}{y_2} B(1 - v|k - j, j + 1) \right\},
\]

where \( v = y_1^* / (y_1^* + y_2^*) \) and \( B(\cdot|a, b) \) is the Beta distribution.

Therefore, (6) can be approximated by a **Monte Carlo estimate**.
We consider some unobserved high thresholds: 18 cm for the first pair (left) and 80 cm for the second (right). We compute

- $P(X_1 > 18 | X_2 > 18) \approx 0.176$ and $P(X_2 > 18 | X_1 > 18) \approx 0.557$.
- $P(X_1 > 80 | X_2 > 80) \approx 0.453$ and $P(X_2 > 80 | X_1 > 80) \approx 0.303$. 
Thank you!
Some References


