"Statistics for Near Independence in Multivariate Extremes"

A.W. Ledford and J.A. Tawn

Discussion by Dan Cooley
NCAR Extremes Reading Group
April 12, 2006

Forward

These two Ledford and Tawn articles are a precursor to the Heffernan and Tawn paper we looked at several weeks ago. The first(?) alternative to traditional multivariate extreme value theory.

- Review of classical multivariate extreme value theory
- Ledford and Tawn article 1
- A few notes on article 2

Results from Univariate EVT

- $M_t = \max_{i=1,\ldots,t} X_i$.
- A dist. G is *max stable* if $\mathbb{P}\left(\frac{M_t - b_t}{a_t} \leq x\right) = \mathbb{P}(X \leq x) \Leftrightarrow G^t(a_t x + b_t) = G(x)$.
- $F^t(a_t x + b_t) \to G(x), \Rightarrow G$ is max stable
- All univariate max stable distributions are GEV.

Results from Univariate EVT

- $M_t = \max_{i=1,\dots,t} X_i$.
- A dist. G is *max stable* if $\mathbb{P}\left(\frac{M_t - b_t}{a_t} \le x\right) = \mathbb{P}(X \le x) \Leftrightarrow G^t(a_t x + b_t) = G(x)$.
- $F^t(a_t x + b_t) \to G(x), \Rightarrow G$ is max stable
- All univariate max stable distributions are GEV.

Max Stable Example: Unit Fréchet

$$\mathbb{P}\{X_i \le x\} = F(x) = exp\{-x^{-1}\}$$

Normalizing constants: $a_t = t$, $b_t = 0$

$$\mathbb{P}\{M_t/t \le x\} = F^t(tx) = \left[exp\left\{-\frac{1}{tx}\right\}\right]^t = exp\left\{-\frac{t}{tx}\right\} = exp\{-x^{-1}\}$$

Regular Variation

"Regularly varying functions are those functions which behave asymptotically like power functions".

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{-\alpha}$$

Heavy-tailed random variables have regularly varying tails.
Example: Unit Fréchet

$$\begin{aligned} \lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} &= \lim_{t\to\infty} \frac{1-exp\{-(tx)^{-1}\}}{1-exp\{-(t)^{-1}\}} \\ &= \lim_{t\to\infty} \frac{(tx)^{-1}+O(t^{-2})}{t^{-1}+O(t^{-2})} \\ &= \frac{1}{x} \quad \text{(e.g. } 1-F(x) \in RV_1\text{)} \end{aligned}$$

If $U \in RV_\alpha$ then $U(x) = \mathcal{L}(x)x^{-\alpha}$.

Multivariate EVT:

$\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ iid F, $\boldsymbol{X}_i = [X_{i,1}, \ldots, X_{i,d}]^T$

- $\boldsymbol{M}_t = [\max_{i=1,\ldots,t} X_{i,1}, \ldots, \max_{i=1,\ldots,t} X_{i,d}]$.
- A dist. G is max stable if $\mathbb{P}\left(\frac{\boldsymbol{M}_t - \boldsymbol{b}_t}{\boldsymbol{a}_t} \leq \boldsymbol{x}\right) = \mathbb{P}(\boldsymbol{X} \leq \boldsymbol{x}) \Leftrightarrow G^t(\boldsymbol{a}_t\boldsymbol{x} + \boldsymbol{b}_t) = G(\boldsymbol{x})$.
- $F^t(\boldsymbol{a}_t\boldsymbol{x} + \boldsymbol{b}_t) \to G(\boldsymbol{x}) \Rightarrow G$ is max stable [deHaan/Resnick, 1977].
- Q: Can the multivariate EVD be characterized?
  A: Yes, but not easily.

Multivariate EVD's: Characterization 1

WLOG, assume G is a d-dimensional EVD with unit Fréchet margins [Resnick, 1987]. Let $E = [0, \infty]^d \setminus \{\mathbf{0}\}$. Then:

$$G(\boldsymbol{x}) = \exp\left(-\left[\int_0^1 \max\left(\frac{w_1}{x_1}, \ldots, \frac{w_d}{x_d}\right) dH(w)\right]\right),$$

where

$$\sum_{j=1}^{d} w_j = 1,$$

and H is a finite measure on $B = \{\boldsymbol{x} \in E : ||\boldsymbol{x}|| = 1\}$ such that

$$\int_0^1 w dH(w) = 1.$$

Multivariate EVD's: Characterization 2

Again, assume G is a d-dimensional EVD with unit Fréchet margins. Let $E = [0, \infty]^d \setminus \{\mathbf{0}\}$. Then, there exists a non-homogeneous Poisson process on E with intensity measure

$$\mu \left\{ \boldsymbol{x} \in E : ||\boldsymbol{x}|| > r, \frac{\boldsymbol{x}}{||\boldsymbol{x}||} \in A \right\} = \frac{H(A)}{r},$$

such that

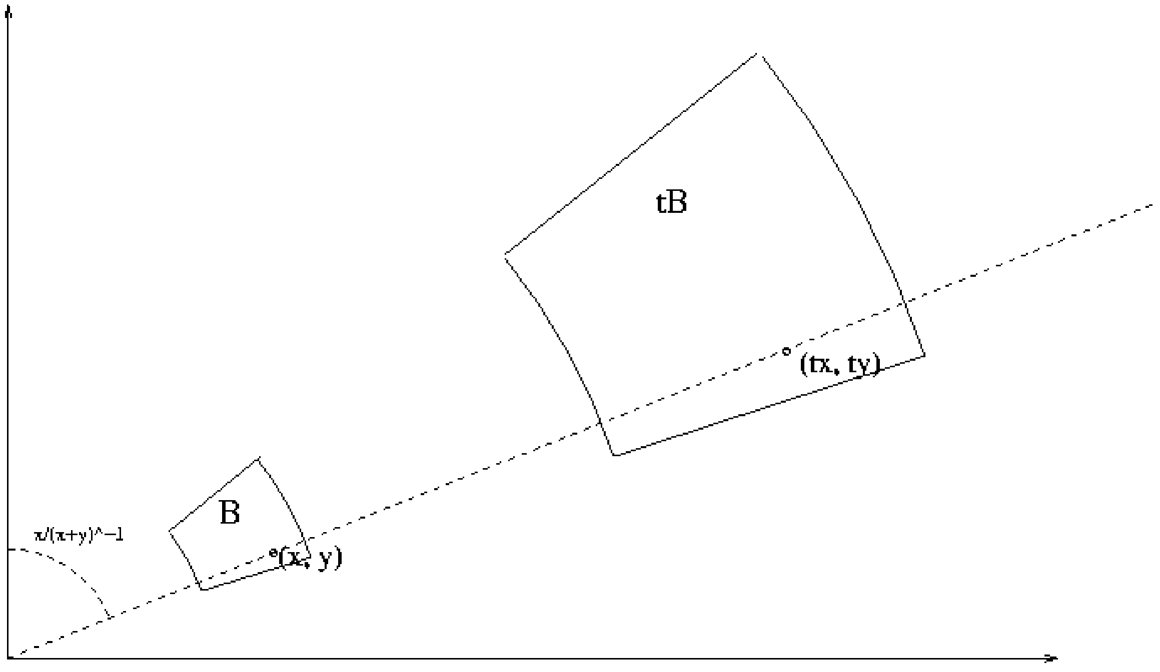
$$G(\boldsymbol{x}) = \exp\left(-\mu\{(\mathbf{0}, \boldsymbol{x}]^c\}\right).$$

Idea: The value of $G(\boldsymbol{x})$ depends only on the "radius" ($||\boldsymbol{x}||$) and the "angular" ($\frac{\boldsymbol{x}}{||\boldsymbol{x}||}$).

Where does characterization 2 come from?

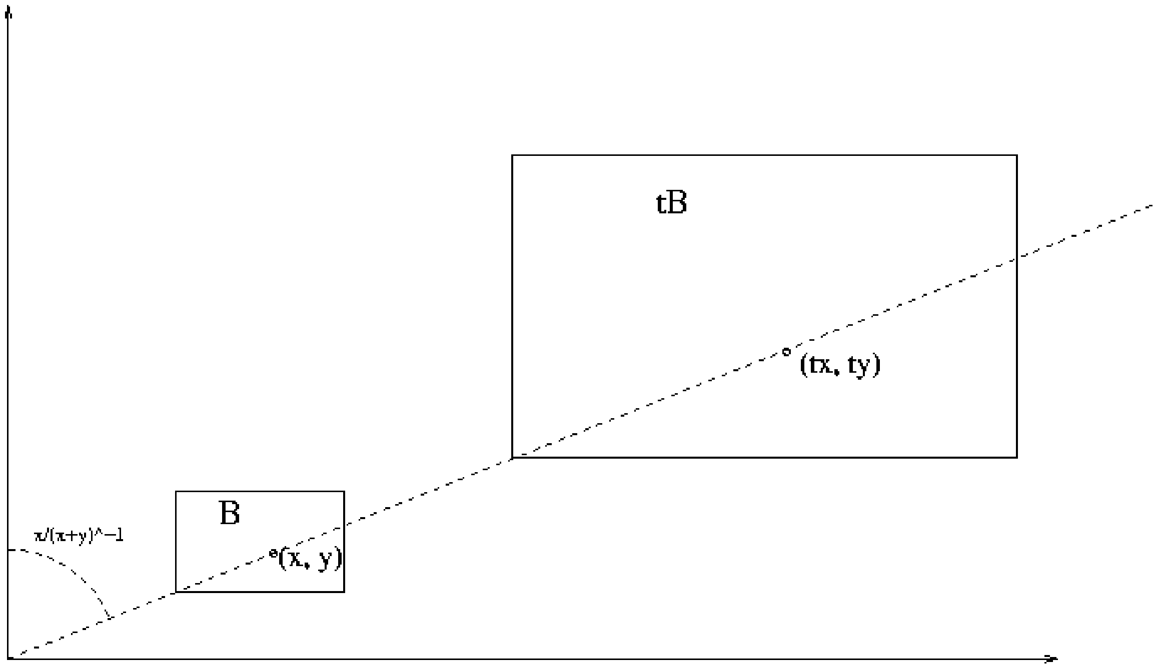
$$\mu\left\{\boldsymbol{x} \in E : ||\boldsymbol{x}|| > r, \frac{\boldsymbol{x}}{||\boldsymbol{x}||} \in A\right\} = \frac{H(A)}{r},$$

Given a set B such that $B = \left\{\boldsymbol{x} : \frac{\boldsymbol{x}}{||\boldsymbol{x}||} \in A, ||\boldsymbol{x}|| \in (a, b)\right\}$, measure is such that $\mu\{B\} = t\mu\{tB\}$.



Where does characterization 2 come from?

Measure theory: Works for all "shapes" B.



With such a measure, we get max stability:

$$G^t(t\boldsymbol{x}) = [\exp(-\mu\{(\mathbf{0}, t\boldsymbol{x}]^c\})]^t = \exp(-t\mu\{(\mathbf{0}, \boldsymbol{t}x]^c\}) = G(x)$$

Multivariate regular variation

...♪ Kinda goes something like this ♪...

$$\mu_x\{tA\} = \frac{\mathbb{P}(x^{-1}\boldsymbol{X} \in tA)}{\mathbb{P}(||\boldsymbol{X}|| > tx)} \frac{\mathbb{P}(||\boldsymbol{X}|| > tx)}{\mathbb{P}(||\boldsymbol{X}|| > x)} \to \mu(A)t^{-\alpha}$$

- Still has angular idea
- Fréchet case: $\alpha = 1$

Asymptotic Independence

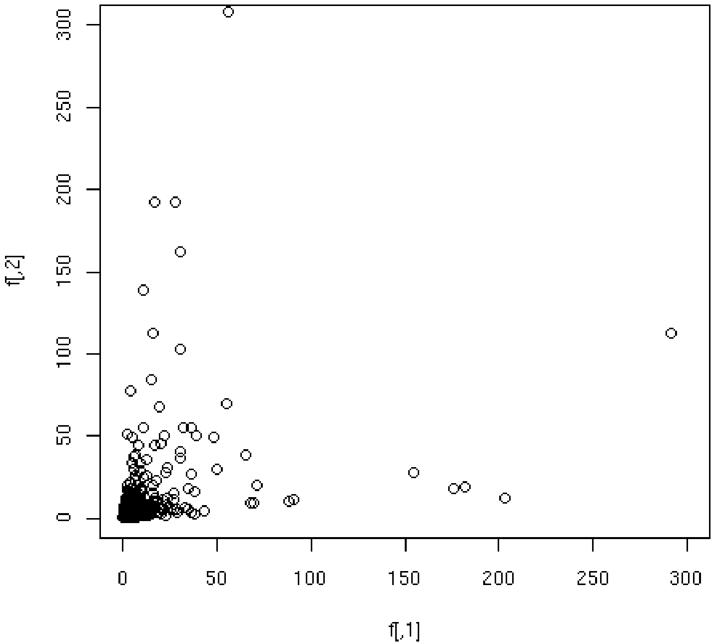
Let (X_1, X_2) be a bivariate random variable with common marginals. Then X_1 and X_2 are *asymptotically independent* if

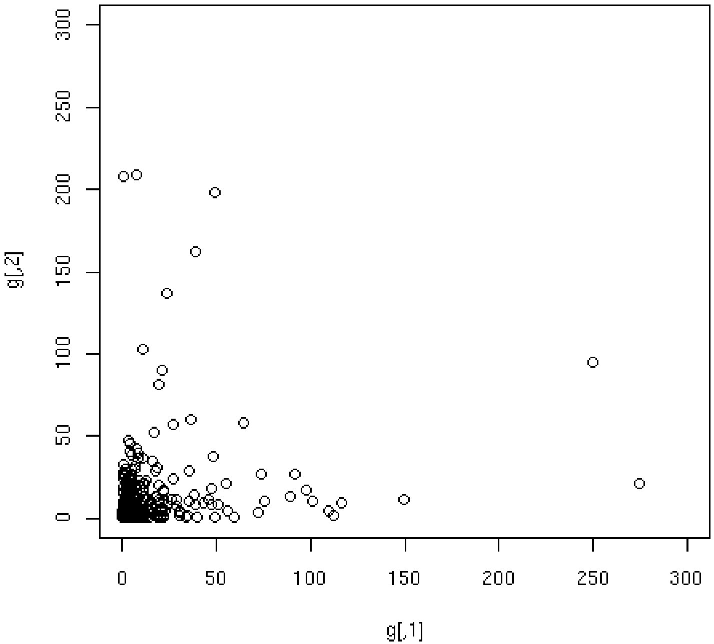
$$\lim_{u \to \infty} \mathbb{P}\{X_1 > u | X_2 > u\} = 0.$$

Example: (X_1, X_2) bivariate normal with $\rho < 1$

Q: How best to handle the middle ground between asymptotically dependent and complete independence?

Example





L&T: 1. Introduction

- Univariate review of threshold exceedance models (GPD and PP).
- Multivariate threshold models:
  - point process model of de Haan (1985)
  - model of Smith, Tawn, and Coles (1997)
- Asymptotic independence
- Problems with above approaches when components are (asymptotically) independent

L&T: 2. Multivariate Threshold Model

- Equations 2.2 and 2.3 ??
- Equations 2.5 and 2.8
- Asymptotic independence case

L&T: 3. Inference for the Model

- Develop censored likelihood (3.2)
- Assume a logistic dependence structure
- Develop score statistic under assumption of independence (3.4)
- Proposition 1

L&T: 4. Variable Thresholds

Similar result holds

L&T: 5. Inference for Asymptotic Independence

- Proposition 2
- η introduced
- Examples
- Estimation for η

L&T: 6. Applications

- Simulations
- Two Examples (wave/surge and wind/rain)

Dan's simulated data

100 points of largest radius used, $X_{min} = min(X_1, X_2)$, GEV fit to X_{min}, $\widehat{\eta} = \widehat{\xi}$

	Normal ($\rho = .8$, $Fr(1)$ Marg.)	Logistic ($\alpha = .7$)
η	.9	1
$\widehat{\eta}$	0.684	0.889
SE	0.106	0.148

L&T II

- From: $\mathbb{P}(Z_1 > r, Z_2 > r) \sim \mathcal{L}(r) r^{-1/\eta}$
  To: $\mathbb{P}(Z_1 > z_1, Z_2 > z_2) \sim \mathcal{L}_1(z_1, z_2) z_1^{-c_1} z_2^{-c_2} + \ldots$
  where $c_1 + c_2 = 1/\eta$
- Make some assumptions, develop a likelihood
- Develop techniques for fitting a model (pretty cool results)
- Point process representation for asymptotic independence

References

de Haan, L. (1985). Extremes in higher dimensions: the model and some statistics. *Proceedings 45th Session of the International Statistical Institute*.

de Haan, L. and Resnick, S. (1977). Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitstheorie*, 4:317–337.

Fougères, A. (2004). Multivariate extremes. In B. Finkenstadt and H. Rootzen, editor, *Extreme Values in Finance, Telecommunications and the Environment*, pages 373–388. Chapman and Hall CRC Press, London.

Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions*. Imperial College Press, London.

Mikosch, T. (2004). How to model multivariate extremes if one must? http://www.gloriamundi.org/picsresources/tm.pdf.

Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag, New York.

Smith, R., Tawn, J., and Coles, S. (1997). Markov chain models for threshold exceedances. *Biometrika*, 84:249–268.