TWO DEPENDENCE MEASURES FOR MULTIVARIATE EXTREME VALUE DISTRIBUTIONS

WEISSMAN, ISHAY

Technion - Israel Institute of Technology, Israel

ieriw01@ie.technion.ac.il

Outline:

- Introduction

- Dependence measures τ_1, τ_2

- Examples

- Relations between τ_1, τ_2

- Combining two models

- Conclusions

1. Introduction

$$\mathbf{X} = (X_1, X_2, \cdots, X_d) \sim G(\mathbf{x}), \quad \mathbf{x} \in \mathbf{R}^d,$$

where G is a multivariate extreme value distribution function. WOLOG with Fréchet margins: for all j

$$G_j(x) = \exp\{-1/x\} \quad (x > 0)$$

and exponent function

$$\lambda(\mathbf{x}) = -\log G(\mathbf{x}),$$

$$G^t(t\mathbf{x}) = G(\mathbf{x}) \quad \Rightarrow \quad t\lambda(t\mathbf{x}) = \lambda(\mathbf{x}), \ (t > 0).$$

Since for MEVD

$$\min_{1 \le j \le d} G_j(x_j) \ge G(\mathbf{x}) \ge \prod_{j=1}^{d} G_j(x_j),$$

in our case

$$\max \frac{1}{x_j} \le \lambda(\mathbf{x}) \le \sum \frac{1}{x_j}.$$

(complete dependence) (total independence)

The homogeneity of λ implies that $\lambda(tx)/\Sigma(tx_j)^{-1}$ does not depend on t. Define the (generalized) Pickands dependence function

$$A(\mathbf{v}) = \lambda(v_1^{-1}, v_2^{-1}, \cdots, v_d^{-1}) \quad \mathbf{v} \in \Omega,$$

$$\text{where} \qquad \Omega = \{\mathbf{v} : v_j \geq 0, \ \Sigma v_j = 1\}$$

is the unit-simplex. It follows that

$$\frac{1}{d} \leq A_0(\mathbf{v}) =: \max v_j \leq A(\mathbf{v}) \leq 1,$$

$$\lambda(\mathbf{x}) = A(\mathbf{v})\Sigma x_j^{-1},$$

where $v_j = x_j^{-1}/\Sigma x_i^{-1}$.

$$\eta = A\left(\frac{1}{d}, \cdots, \frac{1}{d}\right)$$

has an interesting interpretation:

$$P\left\{\max_{1 \leq j \leq d} X_j \leq z\right\} = \exp\{-\lambda(z, z, \cdots, z)\}$$

$$= \exp\{-d\eta/z\} = \{\exp\{-1/z\}\}^{d\eta}.$$

Hence, $\theta = d\eta$ is the *extremal coefficient* of $(X_1, X_2, \cdots, X_d)$.

$$\theta = 1 \Leftrightarrow \text{complete dependence}$$

$$\theta = d \Leftrightarrow \text{total independence}$$

Schlather and Tawn (2002) analyse $\theta_B = |B|\eta_B$ for all 2^d possible subsets B of $\{1, 2, \cdots, d\}$.
From de Haan and Resnick (1977) and Pickands (1981)

$$A(\mathbf{v}) = \int_\Omega \max v_j a_j U(d\mathbf{a})$$

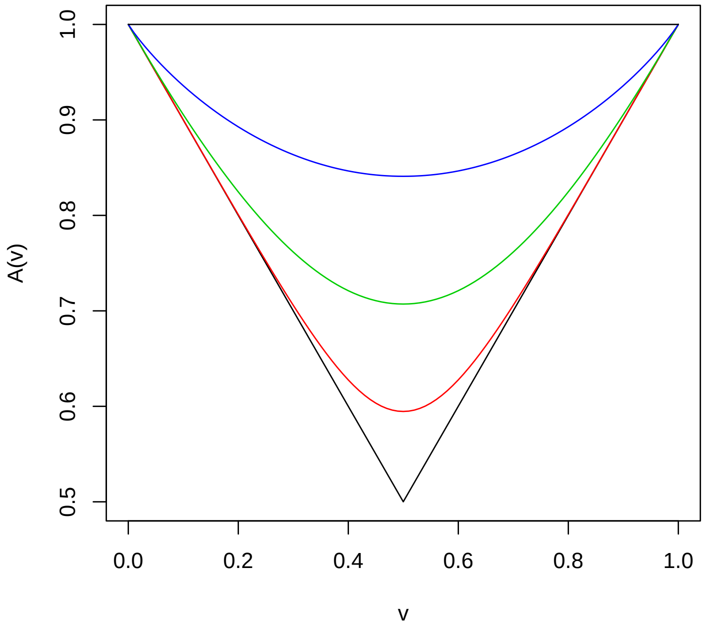
for a finite positive measure U,

$$U(\Omega) = d \text{ and } \int_\Omega a_j U(d\mathbf{a}) = 1 \text{ for all } j.$$

The function A is convex because for $0 \leq \alpha \leq 1$,

$$\max\{(\alpha v_j + (1-\alpha) w_j) a_j\}$$

$$\leq \alpha \max v_j a_j + (1-\alpha) \max w_j a_j.$$



Pickands dependence function for the Logistic Model $A(v) = (v^{1/\alpha} + (1-v)^{1/\alpha})^{\alpha}$ with $\alpha = 0$, .25, .50, .75, 1.

2. Measures of Dependence: Rescaling η, a natural measure of dependence is

$$\tau_1 = \frac{1 - A\left(\frac{1}{d}, \cdots, \frac{1}{d}\right)}{\max_A \left\{1 - A\left(\frac{1}{d}, \cdots, \frac{1}{d}\right)\right\}}$$

$$= \frac{d-\theta}{d-1} = \frac{d}{d-1}(1-\eta)$$

An alternative measure is

$$\tau_2 = \frac{\int_\Omega (1-A(\mathbf{v}))d\mathbf{v}}{\max_A \int_\Omega (1-A(\mathbf{v}))d\mathbf{v}}$$

$$= \frac{\int_\Omega (1-A(\mathbf{v}))d\mathbf{v}}{\int_\Omega (1-A_0(\mathbf{v}))d\mathbf{v}} =: \frac{S_d(A)}{S_d(A_0)}.$$

Which one is preferred?

Similar question: mode vs. mean.

Expect from dependence measure that for

$$A = \alpha A_0 + (1-\alpha)\cdot 1$$

$$\Rightarrow \qquad \tau = \alpha.$$

Indeed, for this **mixture model**

$$\tau_1 = \tau_2 = \alpha.$$

To compute τ_2 we need a formula for S_{A_0}, the volume above A_0:

d	S_{A_0}
2	$1/4 = .2500$
3	$7/36 = .19444$
4	.07986
5	.02264

Until very recently the challenge was to find a formula for $S_d(A_0)$. My colleague **Shmuel Onn** derived and proved

$$\boxed{S_d(A_0) = \tfrac{1}{(d-1)!} - \tfrac{B_d}{d!}}$$

where

$$B_d = \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{d}\right)$$

is the harmonic sum.

Other (bivariate) measures of dependence:

In the literature (Beirlant et al, 2004) we encounter

$$\tau_K = \text{Kendall's tau} = 4EC(G_1(X_1), G_2(X_2)) - 1,$$

$$\rho_S = \text{Spearman's rho} = corr(G_1(X_1), G_2(X_2)),$$

$$\rho = corr(\log G_1(X_1), \log G_2(X_2)).$$

Tawn (1988) mentioned τ_1 for $d = 2$. I have not seen τ_2. These are all marginal-free and for mixture distributions (not mixture exponents):

$$(X_1, X_2) = \left\{ \begin{array}{ccc} (U, V) & \text{w.p.} & 1 - \alpha \\ (U, U) & \text{w.p.} & \alpha \end{array} \right\}$$

U, V independent,

$$\tau_K = \rho_S = \rho = \alpha.$$

3. Examples.

Let $V_1, V_2, \cdots$ be i.i.d. unit-Fréchet.

Mixture model: For $0 \le \alpha \le 1$

$$\lambda(x, y) = \alpha \max(x^{-1}, y^{-1}) + (1 - \alpha)(x^{-1} + y^{-1}).$$

That is

$$X = \max(\alpha V_1, (1 - \alpha) V_2)$$

$$Y = \max(\alpha V_1, (1 - \alpha) V_3).$$

$$A(v) = \alpha \max(v, 1 - v) + (1 - \alpha) \cdot 1 \quad (v \in [0, 1]).$$

$$\tau_1 = \tau_2 = \alpha.$$

$$\tau_K = \rho = \frac{\alpha}{2 - \alpha} \le \rho_S = \frac{3\alpha}{4 - \alpha} \le \alpha.$$

$\alpha = \tau_1 = \tau_2$	ρ_S	$\tau_K = \rho$
0	0	0
1/4	1/5	1/7
1/2	3/7	1/3
3/4	9/13	3/5
1	1	1

Mixed model:

$$\lambda(x,y) = \frac{1}{x} + \frac{1}{y} - \frac{\alpha}{x+y}$$

$$A(v) = 1 - \alpha(1-v)v$$

$$\tau_1 = \frac{\alpha}{2}, \quad \tau_2 = \frac{2}{3}\alpha$$

$$\tau_K = \frac{8tan^{-1}(\alpha/(4-\alpha))^{1/2}}{\alpha^{1/2}(4-\alpha)^{1/2}} - 2$$

$$\rho = \frac{8tan^{-1}(\alpha/(4-\alpha))^{1/2}}{\alpha^{1/2}(4-\alpha)^{3/2}} - \frac{2-\alpha}{4-\alpha}$$

$$\rho_S = 12\left\{\frac{8tan^{-1}(\alpha/(8-\alpha))^{1/2}}{\alpha^{1/2}(8-\alpha)^{3/2}} + \frac{1}{8-\alpha}\right\} - 3$$

α	τ_K	ρ	τ_1	ρ_S	τ_2
0	0	0	0	0	0
.25	.0877	.0901	.1250	.1299	.1667
.50	.1853	.1958	.2500	.2702	.3333
.75	.2947	.3215	.3750	.4222	.5000
1	.4184	.4728	.5000	.5874	.6667

de Haan - Resnick model:

$$\lambda(x,y,z) = \frac{1}{2}\{\max(x^{-1},y^{-1}) + \max(x^{-1},z^{-1})$$

$$+\max(y^{-1},z^{-1})\}$$

$$X_1 = \max(V_1,V_2)/2$$

$$X_2 = \max(V_1,V_3)/2$$

$$X_3 = \max(V_2,V_3)/2$$

$$A(\mathbf{v}) = \frac{1}{2}\{\max(v_1,v_2) + \max(v_1,v_3) + \max(v_2,v_3)\}$$

$$\eta = A(1/3,1/3,1/3) = 1/2, \;\; \tau_1 = (3/2)(1-\eta) = 3/4$$

$$\tau_2 = \frac{36}{7}\cdot\frac{1}{8} = \frac{9}{14} = .642857$$

$$\tau_1(1,2) = \tau_2(1,2) = 1/2$$

(Introducing X_3 to the system increases the dependence)

Non-symmetric model:

$$X_1 = \max(V_1/2, V_2/4, V_3/4)$$

$$X_2 = \max(2V_1/3, V_2/3)$$

$$X_3 = V_3$$

$$A(\mathbf{v}) = \max(.75v_1, v_2) + \max(.25v_1, v_3)$$

$$(v_1 + v_2 + v_3 = 1)$$

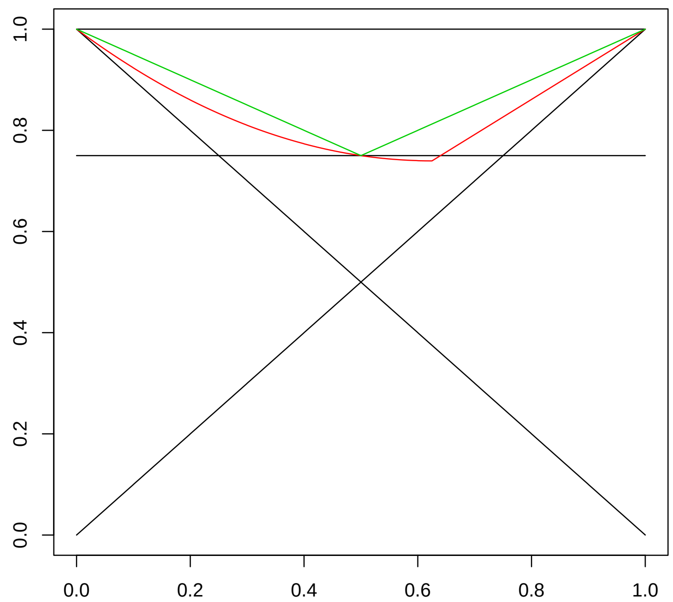
$$\eta = 2/3, \;\; \tau_1 = 1/2, \;\; \tau_2 = (36/7).104762 = .53876$$

$$\tau_1(1,2) = 3/4 = .75 \quad \tau_2(1,2) = 6/7 = .85714$$

$$\tau_1(1,3) = 1/4 = .25 \quad \tau_2(1,3) = 4/10 = .4$$

4. Relations between τ_1 and τ_2.

Theorem. For $d = 2$, $\tau_1 \leq \tau_2$.



Proof. Denote $1 - h = A(1/2, 1/2))$,

$$\Rightarrow \qquad \tau_1 = 2h.$$

Define the mixture model (green graph)

$$A^*(\mathbf{v}) = \tau_1 A_0(\mathbf{v}) + 1 - \tau_1,$$

$$\Rightarrow \qquad \tau_1^* = \tau_1 = \tau_2^*.$$

Since A is convex, $A \leq A^*(=$ at $(1/2, 1/2))$,

$$\int_\Omega (1 - A) \geq \int_\Omega (1 - A^*) = \tau_1 \int_\Omega (1 - A_0)$$

$$\tau_2 = \frac{\int_\Omega (1 - A)}{\int_\Omega (1 - A_0)} \geq \tau_1.$$

This is a perfect proof for $d = 2$. For $d \geq 3$, the picture is misleading, namely, $A \leq A^*$ is not necessarily true. Here is a counter example: de Haan-Resnick model.

For $v_1 \geq v_2 \geq v_3$, $v_1 + v_2 + v_3 = 1$,

$$A(\mathbf{v}) = v_1 + \frac{v_2}{2}, \;\; A^*(\mathbf{v}) = \frac{3}{4} v_1 + \frac{1}{4}.$$

Since $v_2 \geq v_3 \Leftrightarrow v_2 \geq (1 - v_1)/2$,

$$A(\mathbf{v}) - A^*(\mathbf{v}) = \frac{1}{4}v_1 + \frac{1}{2}v_2 - \frac{1}{4} \geq 0,$$

with equality when $v_1 \geq 1/3,\ v_2 = v_3 = (1 - v_1)/2$.

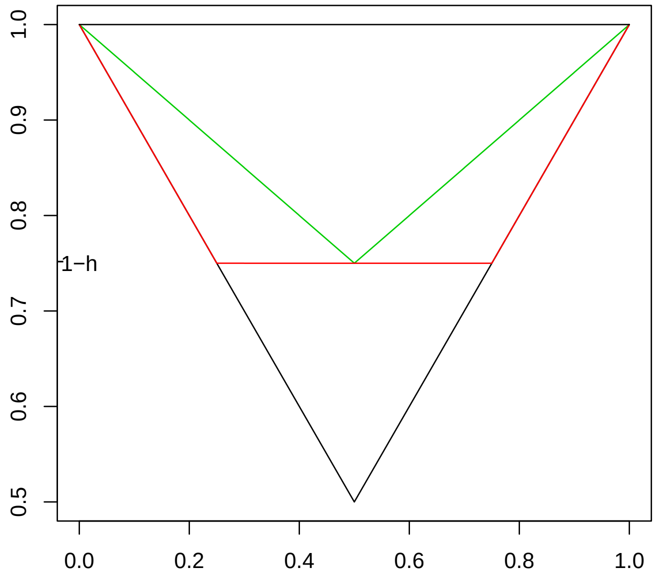
For the logistic model

$$A(\mathbf{v}) = (v_1^{1/\alpha} + v_2^{1/\alpha} + v_3^{1/\alpha})^\alpha,$$

$$A(1/3, 1/3, 1/3) = 3^{\alpha-1}.$$

α	$\tau_1 = (3 - 3^\alpha)/2$	$\tau_2 = \int_\Omega (1 - A) 36/7$
0	1	1
1/4	.8420	.9457
1/2	.6340	.7670
3/4	.3602	.4559
1	0	0

For $d = 2$, how big can the difference $\tau_2 - \tau_1$ be?



Consider all (symmetric, $d = 2$) models for which $A(1/2) = 1 - h$ so that $\tau_1 = 2h$.

All the A functions must be bounded between the green graph and the red one. The green graph corresponds to a mixture model with $\alpha = 2h = \tau_1 = \tau_2$:

$$X_1 = \max(2hV_1, (1-2h)V_2)$$

$$X_2 = \max(2hV_1, (1-2h)V_3).$$

The red A corresponds to "cross over" model:

$$X_1 = \max(hV_1, (1-h)V_2)$$

$$X_2 = \max((1-h)V_1, hV_2)$$

for which

$$\tau_1 = 2h, \quad \tau_2 = 4h(1-h) = 1-(1-\tau_1)^2.$$

$$\max_h(\tau_2 - \tau_1) = \frac{1}{4}$$

occurs at $h = 1/4,\ \ \tau_1 = 1/2,\ \ \tau_2 = 3/4.$

To be fair, one could hold the area (volume) constant (i.e. τ_2) and let τ_1 vary. For instance, all triangles with height h have

$$\tau_2 = 2h, \ \ (0 \leq h \leq 1/2)$$

but

$$\frac{h}{1-h} \leq \tau_1 \leq 2h = \tau_2.$$

$$h = 1/4, \quad 1/3 \leq \tau_1 \leq 1/2 = \tau_2.$$

Combining two models.

$$\mathbf{X} = (X_1, \cdots, X_k), \ \ \mathbf{Y} = (Y_1, \cdots, Y_m)$$

are combined into

$$\mathbf{Z} = (X_1, \cdots, X_k, Y_1, \cdots, Y_m), \quad k + m = d.$$

To study the dependence measures of $\mathbf{Z}$ we must know the dependence between $\mathbf{X}$ and $\mathbf{Y}$. If they are independent we can compute τ_1 and τ_2 :

$$A(\mathbf{v}) = tA_1(\mathbf{u}) + (1-t)A_2(\mathbf{w}) \ \ (\mathbf{v} \in \Omega_d),$$

$$t = v_1 + \cdots + v_k, \ \ \mathbf{u} \in \Omega_k, \ \ \mathbf{w} \in \Omega_m,$$

$$u_i = \frac{v_i}{t}, \ 1 \le i \le k; \quad w_i = \frac{v_{k+i}}{(1-t)}, \ 1 \le i \le m.$$

The Jacobian of the transformation

$$(v_1, \cdots, v_{d-1}) \mapsto (t, u_1, \cdots, u_{k-1}, w_1, \cdots, w_{m-1})$$

is $$J = t^{k-1}(1-t)^{m-1}.$$

$$1 - A = t(1 - A_1) + (1-t)(1 - A_2)$$

$$S(A) = \int_{\Omega_d} (1 - A(\mathbf{v}))d\mathbf{v} =$$

$$= \int_0^1 \int_{\Omega_k} \int_{\Omega_m} (1 - A) d\mathbf{u} d\mathbf{w} t^{k-1}(1-t)^{m-1} dt$$

$$= \frac{1}{(m-1)!} \int_0^1 t^k (1-t)^{m-1} dt \int_{\Omega_k} (1 - A_1(\mathbf{u}))d\mathbf{u}$$

$$+ \frac{1}{(k-1)!} \int_0^1 t^{k-1}(1-t)^m dt \int_{\Omega_m} (1 - A_2(\mathbf{w}))d\mathbf{w}$$

$$= \frac{k!}{d!} S_k(A_0)\tau_{2,1} + \frac{m!}{d!} S_m(A_0)\tau_{2,2}$$

$$\boxed{\tau_2 = \tfrac{k-B_k}{d-B_d}\tau_{2,1} + \tfrac{m-B_m}{d-B_d}\tau_{2,2}}$$

where

$$B_k = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}.$$

For $k = m = 2$,

$$\tau_2 = \frac{6}{23}(\tau_{2,1} + \tau_{2,2}).$$

Similar, even simpler, is the treatment of τ_1:

$$A\left(\frac{1}{d}, \cdots, \frac{1}{d}\right) = \frac{k}{d} A_1 \left(\frac{1}{k}, \cdots, \frac{1}{k}\right) + \frac{m}{d} A_2 \left(\frac{1}{m}, \cdots, \frac{1}{m}\right)$$

$$= \frac{k}{d}\left(1 - \frac{k-1}{k}\tau_{1,1}\right) + \frac{m}{d}\left(1 - \frac{m-1}{m}\tau_{1,2}\right).$$

$$\boxed{\tau_1 = \tfrac{k-1}{d-1}\tau_{1,1} + \tfrac{m-1}{d-1}\tau_{1,2}}$$

For $k = m = 2$

$$\tau_1 = \frac{1}{3}\left(\tau_{1,1} + \tau_{1,2}\right)$$

Note, the sums of the weights are not equal to 1 but tend to 1 as k, m both tend to ∞.

The results here are lower bounds for τ_1, τ_2 when the two models are dependent.

Conclusions

- Conventional correlation coefficients measure pair-wise dependence, while τ_1, τ_2 are reasonable dependence measures for $d \geq 2$.

- For the mixture model τ_1, τ_2 are equal to what we desire.

- The results for combining independent models can serve as lower bounds in case the two models are dependent.

References

[1] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., de Wall, D. and Ferro, C. (2004) *Statistic of Extremes: Theory and Applications*. Wiley.

[2] de Haan, L., and Resnick, S.I. (1977) Limit theory for multivariate sample extremes. *Z. Wahr. verw. Gebeiete,* **40**,317 - 337.

[3] Pickands, J. III (1981) Multivariate extreme value distributions. *Proceedings, 43rd Session of the ISI.* **Book 2**, 859 - 858.

[4] Schlather, M. and Tawn, J. (2002) Inequalities for the extremal coefficients of multivariate extreme value distributions. *Extremes,* **5**, 87 - 102.

[5] Tawn, J.A. (1988) Bivariate extreme value theory: models and estimation. *Biometrika* **75**, 397 - 415.