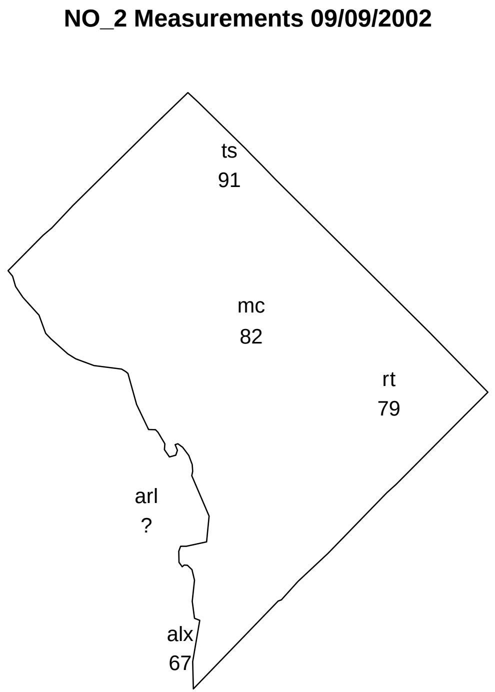
Applications of Tail Dependence I: Interpolating Extreme Air Pollution Levels

Dan Cooley
Department of Statistics
Colorado State University

Joint work with:
Richard Davis, Columbia University
Philippe Naveau, LSCE

Washington DC Air Pollution Measurements

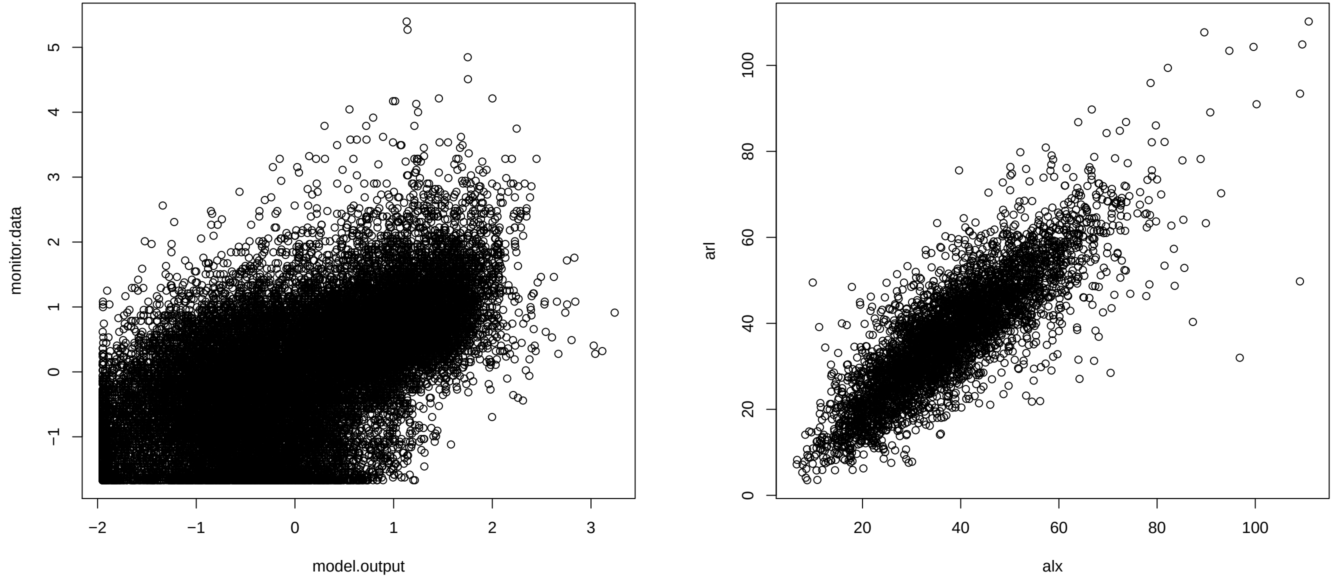


- values are high; each exceeds the 0.97 empirical quantile.
- aim: use observed values to predict/interpolate at unobserved locations.

Outline

- Part A: Background on multivariate extremes.
  (Statistical Application Point-of-View)
  - What is meant by tail dependence?
  - Asymptotic dependence and *measuring* tail dependence.
  - *Modeling* tail dependence.
    - Marginal and dependence effects.
    - Multivariate regular variation and angular measure.
  - Illustration of an extreme value analysis: estimating probability of falling in a risk region.
- Part B: Approximating the conditional density via the angular measure.
  - A Model for the Angular Measure
  - Approximating the Conditional Density when Observed are Large.
  - Washington DC pollution application.

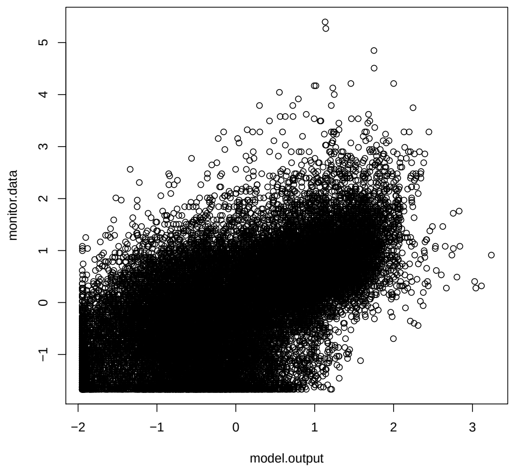
Tail Dependence

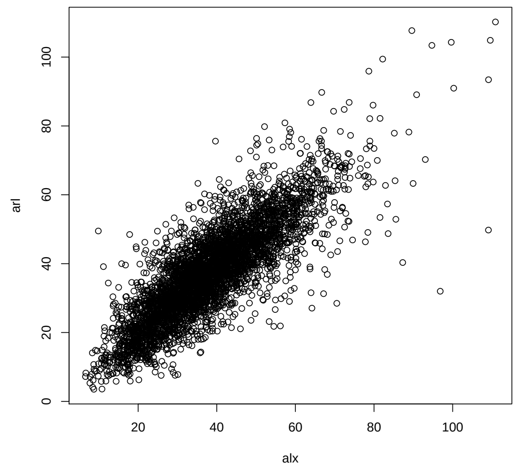


A central aim of multivariate extremes is trying to find an appropriate structure to describe *tail dependence*.

NOT Tail Dependence: Correlation

$$\rho = \frac{E[(X - \mu_x)(Y - \mu_Y)]}{\sqrt{E[(X - \mu_x)^2]E[(Y - \mu_y)^2]}}$$





$\hat{\rho} = 0.59$ $\hat{\rho} = 0.83$

Correlation measures "spread from center", does not focus on extremes.

A Start: Asymptotic Dependence/Independence

A random vector (X, Y) *with common marginals* is termed asymptotically independent if

$$\lim_{u \to x^+} P(X > u \mid Y > u) = 0.$$

Or if X has cdf F_X and Y has cdf F_Y, then

$$\lim_{u \to 1} P(F_X(X) > u \mid F_Y(Y) > u) = 0.$$

If limits is > 0, then X and Y are *asymptotically dependent*.

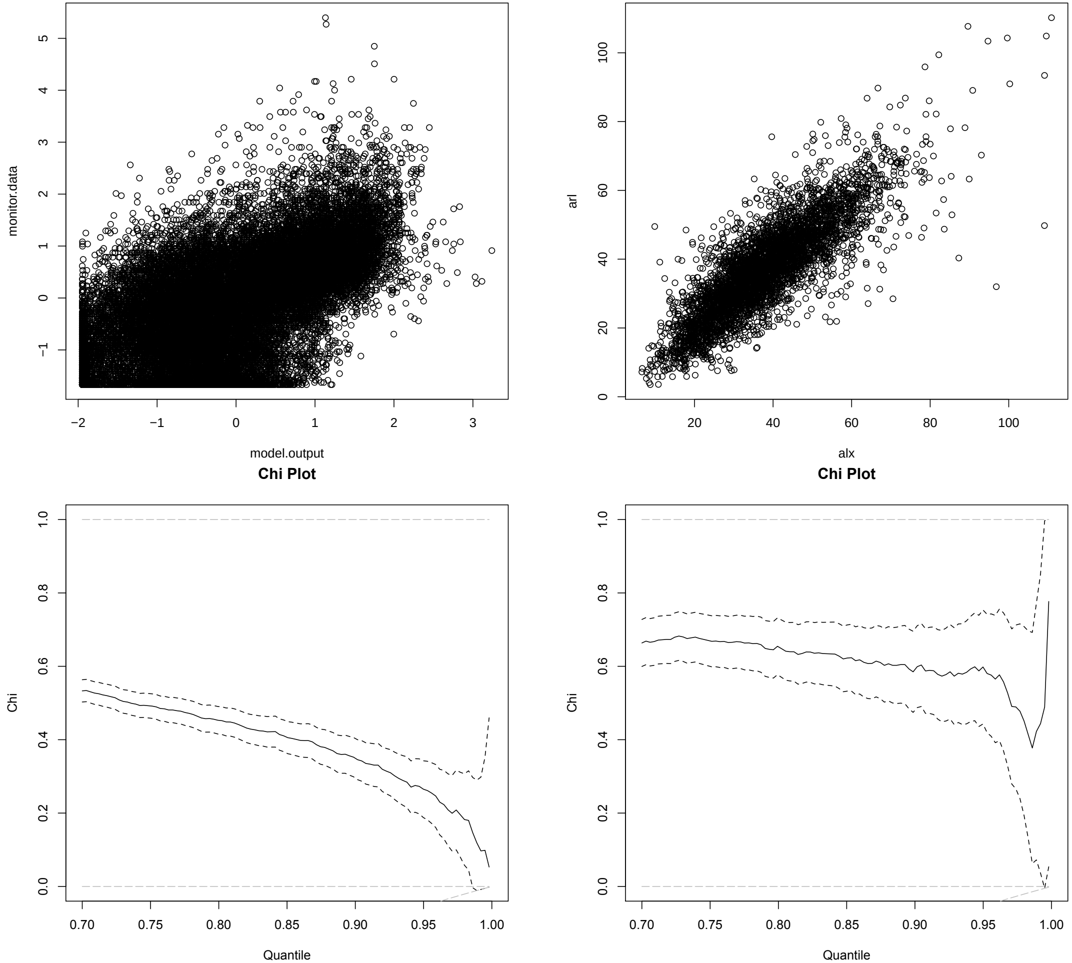
To talk about tail dependence, we need to know something about what it means to be in the tail of each component:

- have a common marginal,
- or account for different marginals.

Asymptotic dependence/independence is a way to *begin* to talk about tail dependence, but doesn't yield whole picture.

Tail Dependence of Examples

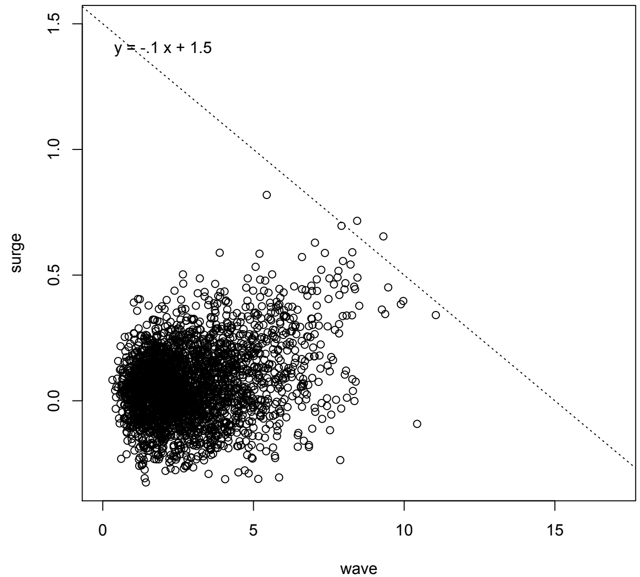
$\widehat{\chi}$ is an empirical measure of asymptotic dependence.

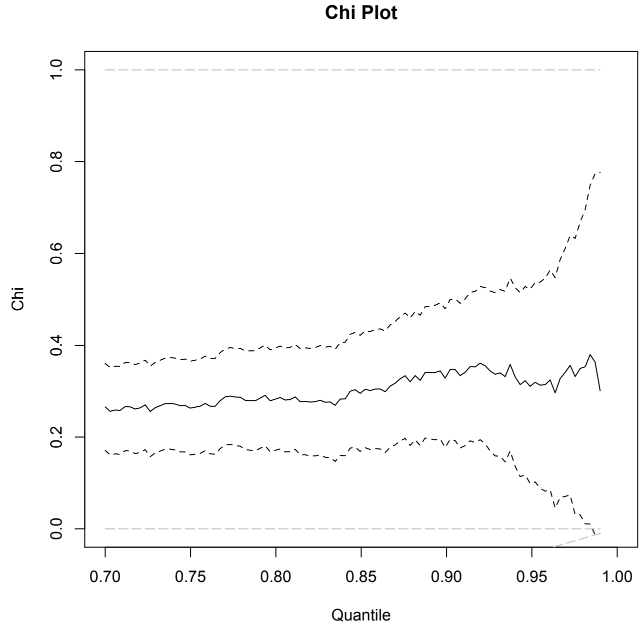


Application of MV Extremes

Typical Goal: estimate the probability of landing in the risk region.

Wave height and storm surge data (Coles, 2001).





Data appear tail dependent, but risk estimate requires more than just a summary measure of tail dependence.

Multivariate Regular Variation

Idea: Joint tail behavior like a power function.

So What? Because it is defined in terms of tail behavior, it provides a framework for describing the joint tail.

Let $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^T \geq \boldsymbol{0}$ be a random vector, define $\mathcal{C}$ to be the set $[\boldsymbol{0}, \infty] \setminus \boldsymbol{0}$ and let $\{b_n\}$ be such that $P(\|\boldsymbol{Z}\| > b_n) \sim n^{-1}$.

Then $\boldsymbol{Z}$ is regularly varying if:

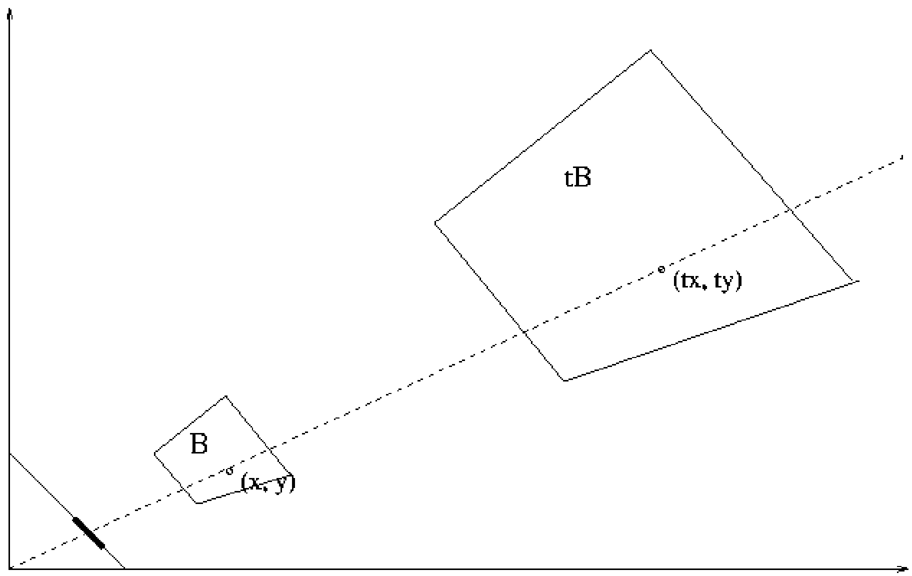
$$nP\left(\frac{\boldsymbol{Z}}{b_n} \in \cdot\right) \xrightarrow{v} \nu(\cdot),$$

where ν is a positive measure, v denotes vague convergence (Resnick, 2007), and $\|\cdot\|$ is any norm.

It can be shown that:

$$\nu(tB) = t^{-\alpha}\nu(B).$$

Scaling Property in a Picture



$$\nu(tB) = t^{-\alpha}\nu(B).$$

- What's ν? A measure, but not a probability measure.
- Nice sets aren't easily described by Cartesian coordinates.
- Scaling property suggests a (pseudo-)polar coordinate transformation.

Regular Variation and the Angular Measure

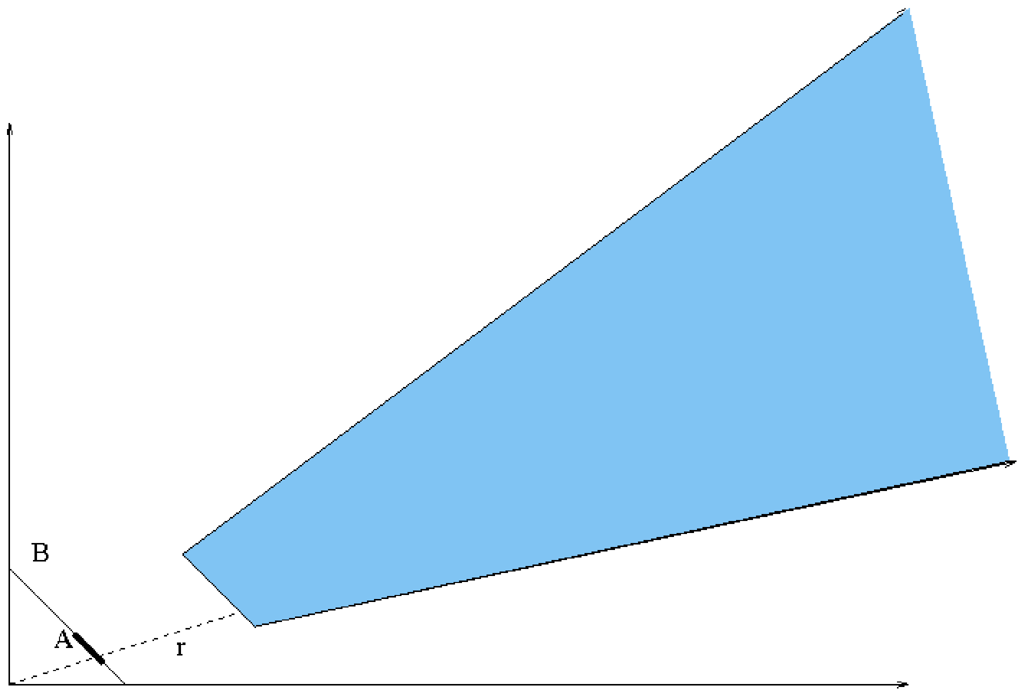
Another Definition: Let $R = \|\boldsymbol{Z}\|$ and $\boldsymbol{W} = \|\boldsymbol{Z}\|^{-1}\boldsymbol{Z}$. $\boldsymbol{Z}$ is regular varying if there exists a normalizing sequence $\{b_n\}$ where $P(b_n^{-1}\|\boldsymbol{Z}\| > r) \sim 1/n$, such that

$$nP\left(b_n^{-1}R > r, \boldsymbol{W} \in A\right) \xrightarrow{v} r^{-\alpha}H(A)$$

where d is the dimension of $\boldsymbol{Z}$, and where H is some probability measure on the unit 'ball' $S_d = \{\boldsymbol{z} \in \mathbb{R}^d \mid \|\boldsymbol{z}\| = 1\}$.

- measure on right is a product measure.
- so...
  - LHS: "as points get big (radial component)"
  - RHS: "radial and angular comps. become independent"
- angular measure H describes distribution of directions – completely describes dependence.
- note: definition requires a common tail behavior (often not true: wave and surge data).

Polar Decomposition in a Picture



$$nP\left(b_n^{-1}R > r, \boldsymbol{W} \in A\right) \xrightarrow{v} r^{-\alpha}H(A)$$

To obtain the result, we looked at a convenient set.
Nice sets are pie-shaped regions.

Regular Variation and Point Processes

$$nP\left(\frac{\boldsymbol{Z}}{b_n} \in \cdot\right) \xrightarrow{v} \nu(\cdot); \quad nP\left(b_n^{-1}R > r, \boldsymbol{W} \in A\right) \xrightarrow{v} r^{-\alpha}H(A)$$

$\{\boldsymbol{Z}_i\}, i = 1, 2, \ldots$ iid copies of $\boldsymbol{Z}$,

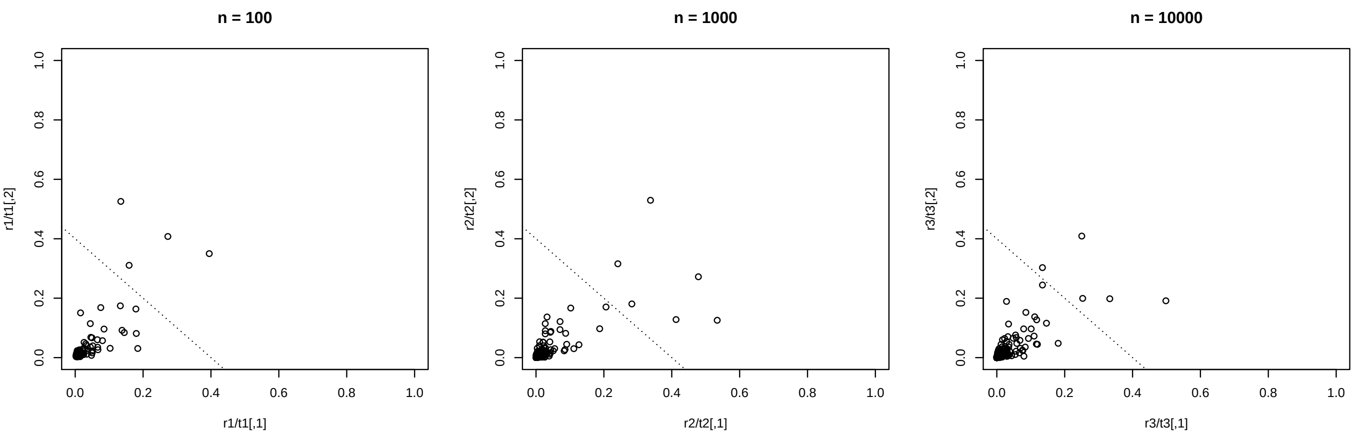
$$\sum_{i=1}^{n} \epsilon_{\boldsymbol{Z}_i/b_n} \xrightarrow{d} PRM(\nu),$$

where $\nu(dr \times d\boldsymbol{w}) = r^{-(\alpha+1)} dr H(d\boldsymbol{w})$.

If H continuously differentiable, then h is the angular density.

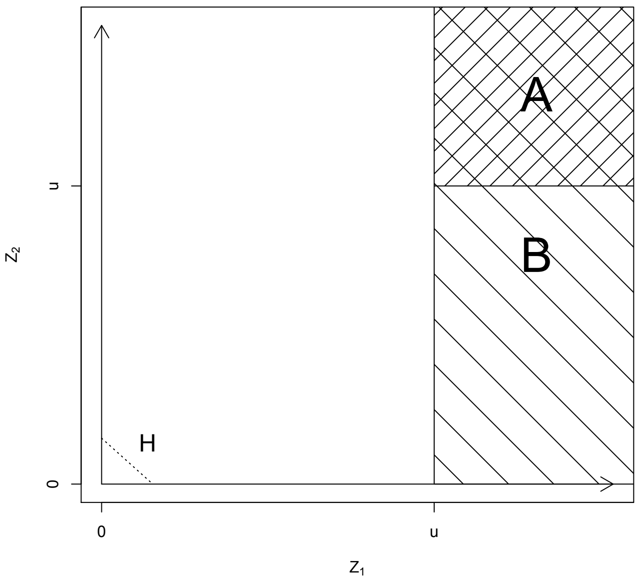
R Demo

Point Processes in a Picture



What's $\nu(B)$? It's the expected number of (normalized) points in set B.

Measuring Tail Dependence, Revisited



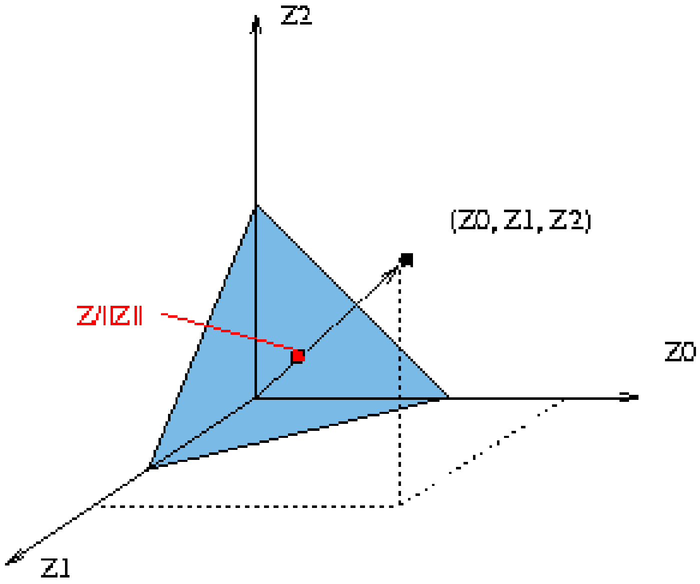
$$\chi = \frac{\nu\Big([u,\infty]\times[u,\infty]\Big)}{\nu\Big([u,\infty]\times[0,\infty]\Big)} = \frac{\int_{S_d} \min\left(\frac{w_1}{z_1},\frac{w_2}{z_2}\right) H(d\boldsymbol{w})}{\int_{S_d} \frac{w_1}{z_1} H(d\boldsymbol{w})}$$

last equality assumes L_1 norm and $\alpha = 1$

- Several other dependence metrics out there.
- Most measure bivariate dependence.

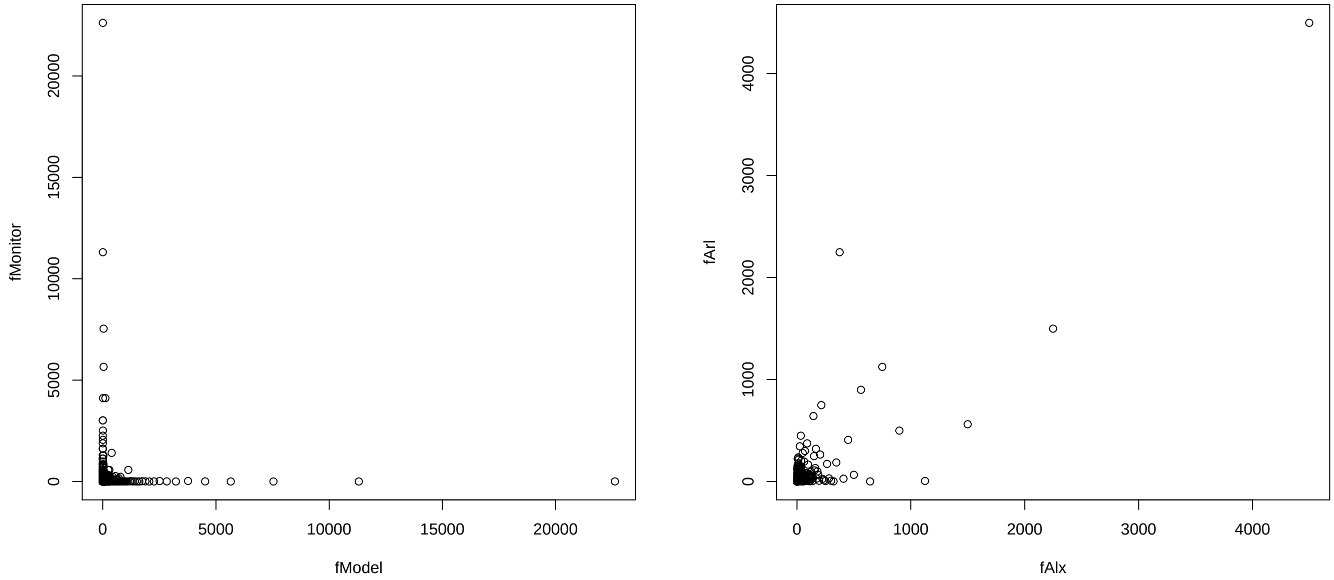
Statistical Practice utilizing MV Regular Variation

- convert marginals to a common and convenient heavy-tailed distribution.
- similar *in approach* to copula methods, *models differ*.
- goal is to model the angular (or spectral) measure H.

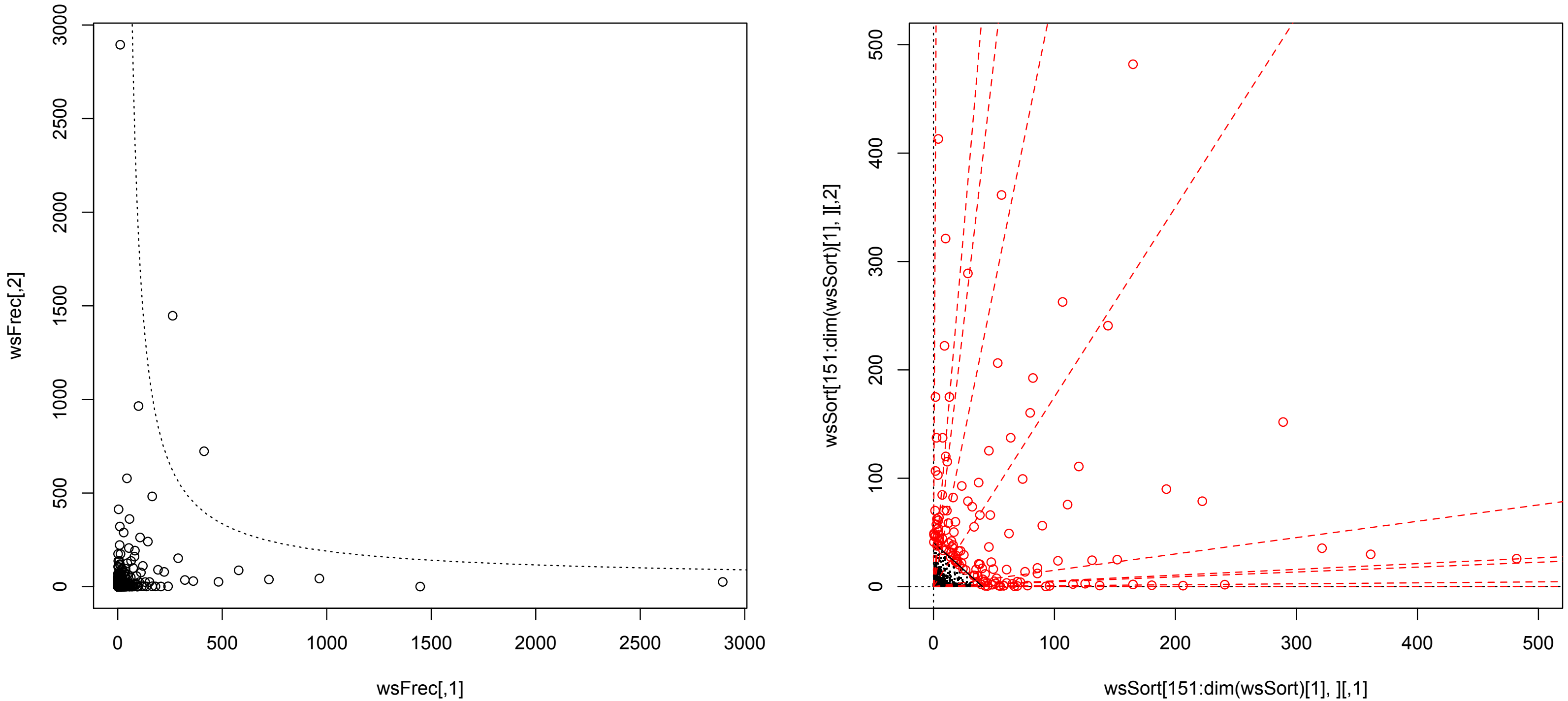


Transformed Data: Air Pollution Datasets

We choose $\alpha = 1$, accentuates large values, will also use L_1 norm.



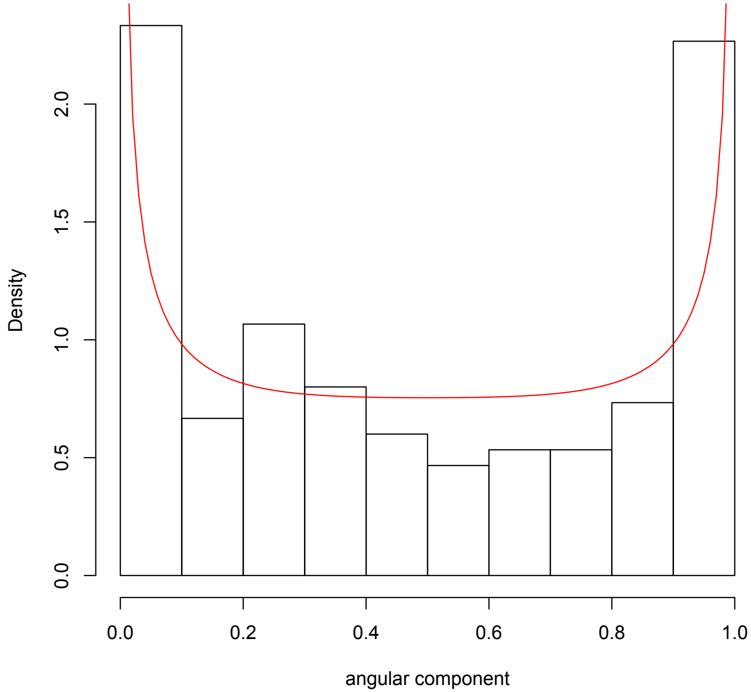
Transformed Wave/Surge Data



Largest 150 observations shown in red; approx 0.95 empirical quantile or radius of 40.6.

Goal: To estimate risk we need to estimate the dependence structure in the tail.

Estimating the Angular Measure



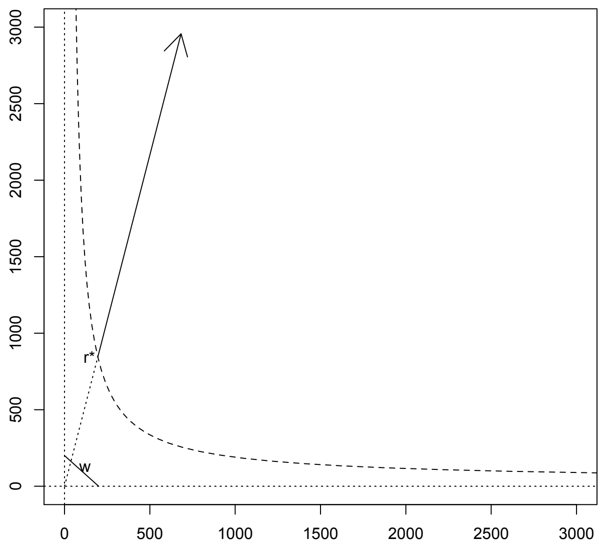
Logistic Model:

$$h(w) = \frac{1}{2}(1/\beta - 1)\Big(w(1-w)\Big)^{-1-1/\beta}\Big(w^{-1/\beta} + (1-w)^{-1/\beta}\Big)^{\beta-2}$$

ML estimate: $\widehat{\beta} = .680(.018)$.

Probability assoc. with Risk Region (1)

$$\nu(A^*) = \int_0^1 \int_{r^*}^{\infty} r^{-2} h(w) dr dw$$



Using fitted logistic model:

$$\widehat{\nu}(A^*) = 0.00079$$

Probability assoc. with Risk Region (2)

$$
\begin{aligned}
nP\left(\frac{\boldsymbol{Z}}{2n} \in A\right) &\approx \nu(A) \\
nP(\boldsymbol{Z} \in 2nA) &\approx \nu(A) \\
\Rightarrow nP(\boldsymbol{Z} \in A^*) &\approx \nu\left(\frac{A^*}{2n}\right) = 2n\nu(A^*) \\
\Rightarrow P(\boldsymbol{Z} \in A^*) &\approx 2\nu(A^*) \stackrel{\text{est}}{=} 0.00158.
\end{aligned}
$$

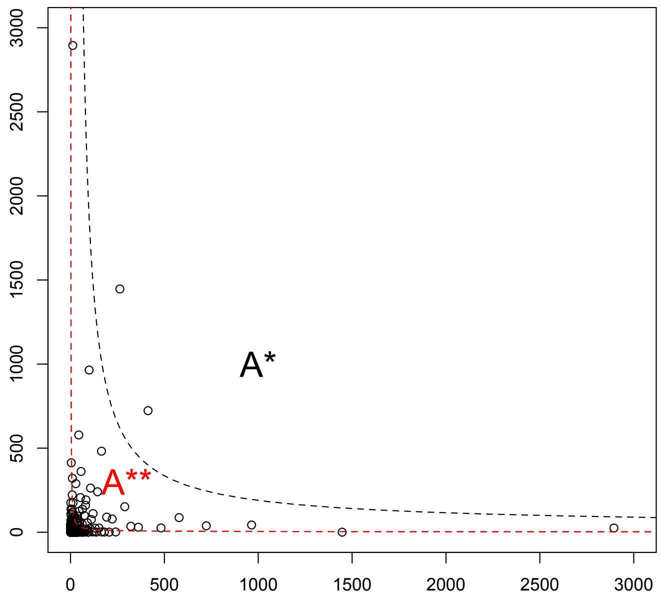
Empirical probability:

$$2/2894 = 0.000691$$

If $p = 0.00158$, probability of two exceedances is

$$\binom{2894}{2}(.00158)^2(1 - .00158)^{2892} = 0.11$$

Expanded Set Estimate (Nonparametric)



$$A^{**} = A^*/10$$
$$\widehat{P}(\boldsymbol{Z} \in A^{**}) = 44/2894 = 0.0152$$
$$\Rightarrow \widehat{P}(\boldsymbol{Z} \in A^*) = 0.00152$$

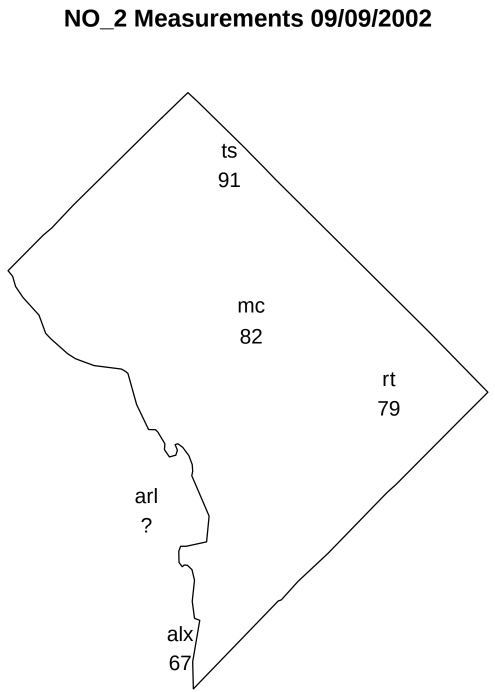
Take-away Messages for Part A

- Tail dependence is different than what we usually think of as dependence.
- In regular variation framework, tail dependence is completely described by the angular measure.
- Regular variation provides a mathematical framework for describing tail behavior—leads to a polar decomposition.
- Current statistical practice often separately handles marginal effects and tail dependence (although the two-step approach illustrated is not always used).
- Extreme value analyses often try to assess the probability associated with a risk region.

Outline

- Part A: Background on multivariate extremes.
  - What is meant by tail dependence?
  - Asymptotic dependence and *measuring* tail dependence.
  - *Modeling* tail dependence.
    - Marginal and dependence effects.
    - Multivariate regular variation and angular measure.
  - Illustration of an extreme value analysis: estimating probability of falling in a risk region.
- Part B: Approximating the conditional density via the angular measure.
  - A Model for the Angular Measure
  - Approximating the Conditional Density when Observed are Large.
  - Washington DC pollution application.

Washington DC Air Pollution Measurements

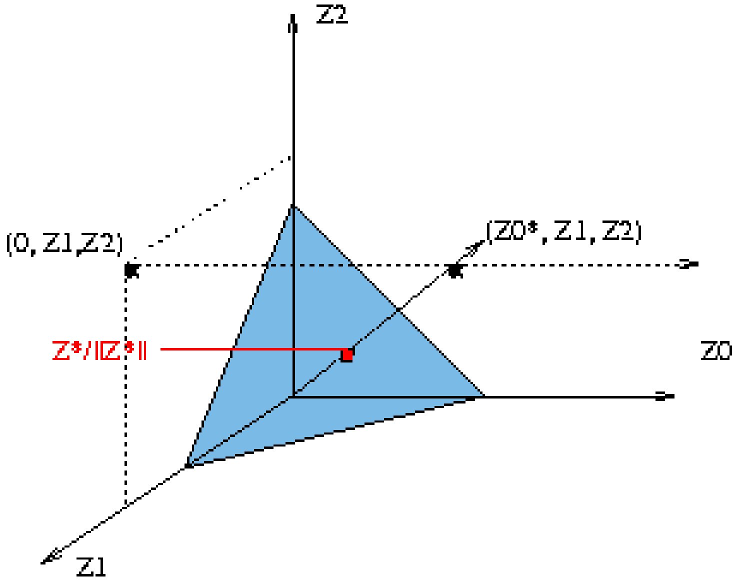


- values are high; each exceeds the 0.97 empirical quantile.
- aim: use observed values to predict/interpolate at unobserved locations.

Motivation

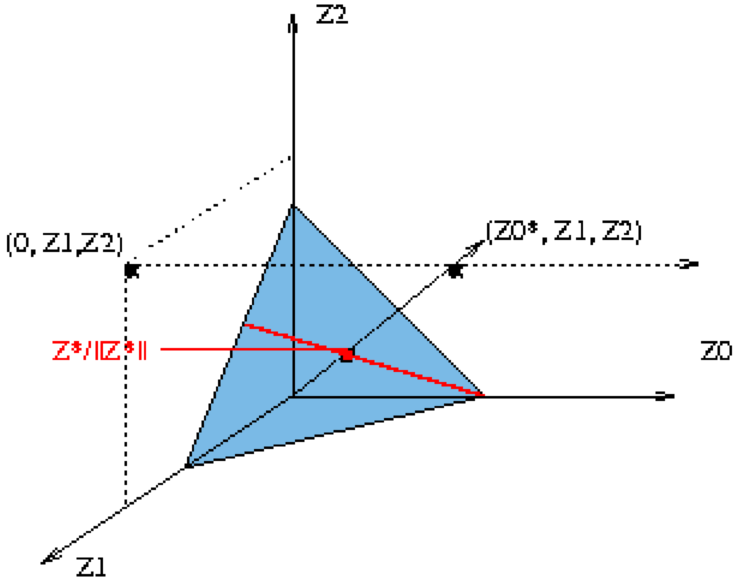
- Air pollution (and other variables) are of greatest interest when values are large.
- Linear prediction methods (e.g., Kriging) are well-suited for center of the distribution.
  - use second-moment properties – based on covariances or correlations.
  - *almost* a Gaussian assumption.
- Utilize extreme value theory to describe *tail dependence*.
- Point prediction may not be very useful; instead try to approximate the *conditional density*.
  - What is probability amount exceeds a specified level?
  - What is a probabilistic upper bound on the pollution level?
- An atypical application of multivariate extremes.

Approximating the Conditional Density when Observations are Large



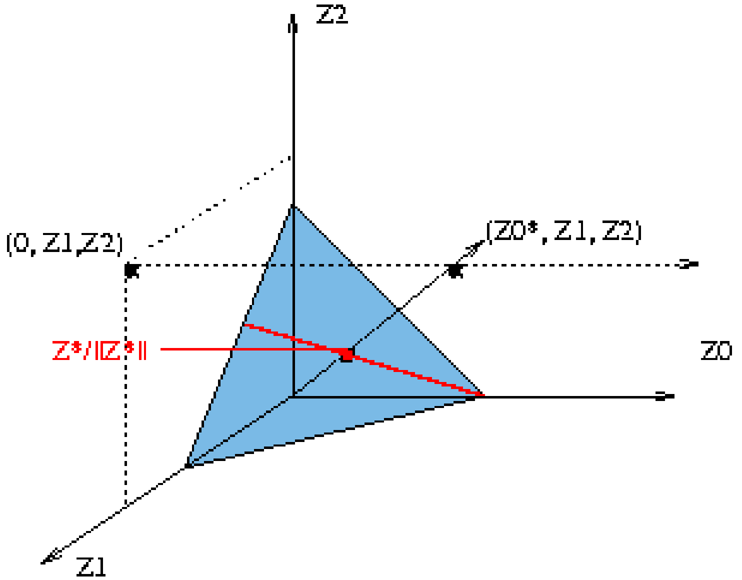
Assume Z_1, Z_2 are observed and large and Z_0 is unobserved. Any predictor Z_0^* will yield a point $\boldsymbol{Z}^* = (Z_0^*, Z_1, Z_2)$ which can be mapped back to S_{p-1} as $\frac{\boldsymbol{Z}^*}{\|\boldsymbol{Z}^*\|_1}$.

Approximating the Condtional Density when Observations are Large



Given the radius is large, by knowing the values of the angular density at $\frac{\boldsymbol{Z}^*}{\|\boldsymbol{Z}^*\|_1}$ and the value of the "radius" $\|\boldsymbol{Z}^*\|_1$, we aim to approximate the values of the joint "density" and in turn the *conditional "density"*.

Approximating the Condtional Density when Observations are Large



We need:

1. A model for the angular measure.
2. To clarify what we mean by “density”.

Moment Conditions for the Angular Measure

In general, H can be *any* probability measure.

However, if we assume that $Z_i, i = 1, \dots p$ have a common marginal distribution with $\alpha = 1$. Then for the ith marginal component,

$$
\begin{aligned}
nP\left(\frac{Z_i}{a_n} > z\right) &\to \nu\{\boldsymbol{x} \in \mathcal{C} : x_i > z\} \\
&= \int_{S_{p-1}} \int_{\frac{z}{w_i}}^{\infty} r^{-2} dr dH(\boldsymbol{w}) \\
&= \frac{1}{z} \int_{S_{p-1}} w_i dH(\boldsymbol{w}).
\end{aligned}
$$

Since we have assumed a common marginal, this implies that

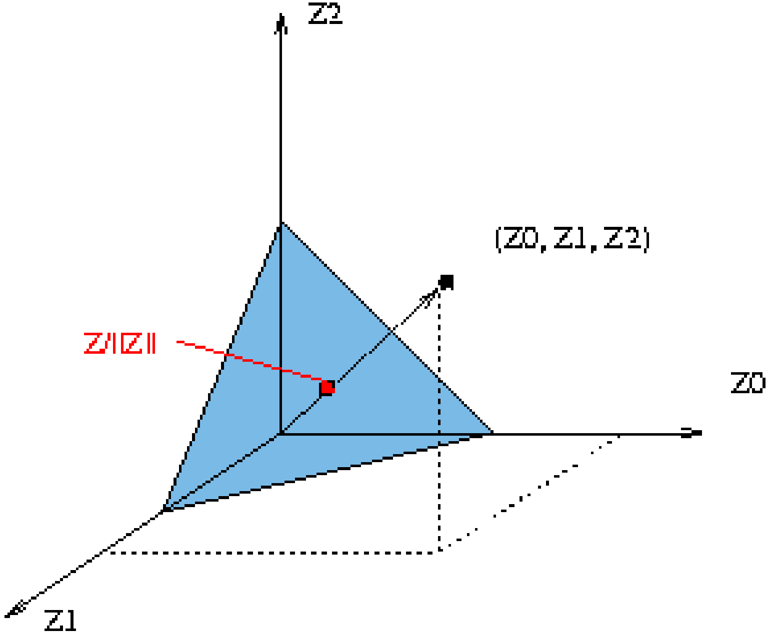
$$\int_{S_{p-1}} w_1 dH(\boldsymbol{w}) = \int_{S_{p-1}} w_j dH(\boldsymbol{w})$$

for all $j = 2, \dots, p$.

Center of Mass Condition

If $\alpha = 1$, it is useful to choose the L_1 norm: $\|\boldsymbol{z}\| = z_1 + \ldots + z_p$.

With this norm, S_{p-1} is unit simplex and $\int_{S_{p-1}} w_i dH(\boldsymbol{w}) = p^{-1}$.



Parametric Models for MV Extremes

Parametric models have been suggested for the exponent measure function $V(\boldsymbol{z})$ or angular density $h(\boldsymbol{w})$.

$$V(\boldsymbol{z}) = \int_{S_d} \max_i \frac{w_i}{z_i} H(d\boldsymbol{w})$$

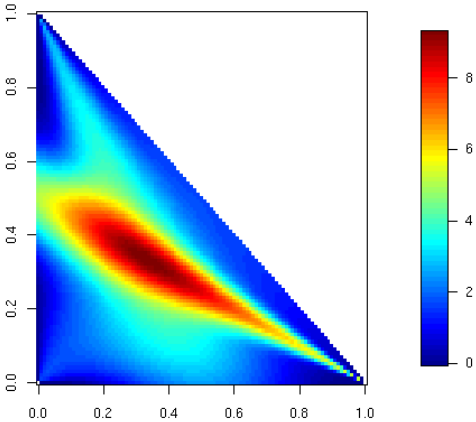
Exponent measure function $V(\boldsymbol{z})$	Angular density $h(\boldsymbol{w})$
• Logistic	• Dirichlet (Coles and Tawn, 1991)
• Asymmetric Logistic (Tawn, 1988)	• *Pairwise Beta* (Cooley et al., 2010)
• Negative Logistic (Joe, 1990)	• Geometric Approach (Ballani and Schlather, 2011)

Pairwise Beta Angular Measure

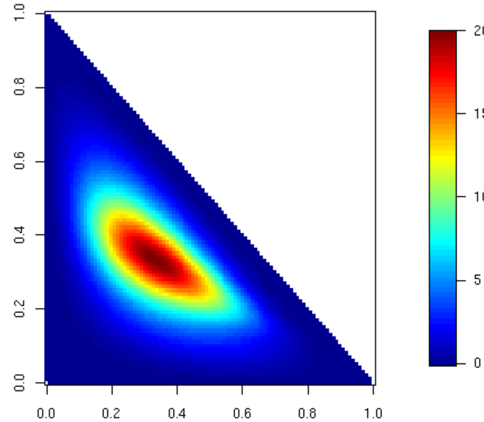
$$\begin{aligned} h(\boldsymbol{w};\alpha,\boldsymbol{\beta}) &= K_p(\alpha)\sum_{1\le i<j\le p} h_{i,j}(\boldsymbol{w};\alpha,\beta_{i,j}), \\ \text{where}\quad h_{i,j}(\boldsymbol{w};\alpha,\beta_{i,j}) &= (w_i+w_j)^{2\alpha-1}(1-(w_i+w_j))^{\alpha(p-2)-p+2} \\ &\quad \times \frac{\Gamma(2\beta_{i,j})}{(\Gamma(\beta_{i,j}))^2}\left(\frac{w_i}{w_i+w_j}\right)^{\beta_{i,j}-1}\left(\frac{w_j}{w_i+w_j}\right)^{\beta_{i,j}-1}, \\ \text{and}\quad K_p(\alpha) &= \frac{2(p-3)!}{p(p-1)\sqrt{p}}\frac{\Gamma(\alpha p+2)}{\Gamma(2\alpha+1)\Gamma(\alpha(p-2))(\alpha p+1)} \end{aligned}$$

Advantages:

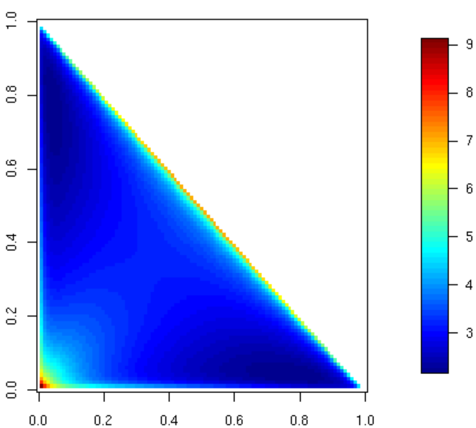
- no adjustment necessary to get center of mass condition
- parameters have some interpretation: α controls overall dependence, $\beta_{i,j}$'s control pairwise dependence
- largely specified by pairwise parameters



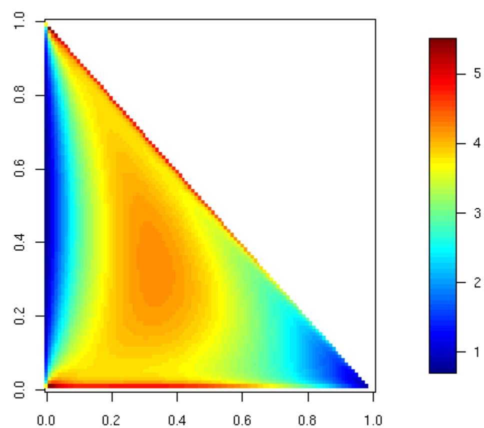
$\alpha = 1, \beta = (2, 4, 15)$



$\alpha = 4, \beta = (2, 4, 15)$



$\alpha = 1, \beta = (2, .5, .5)$

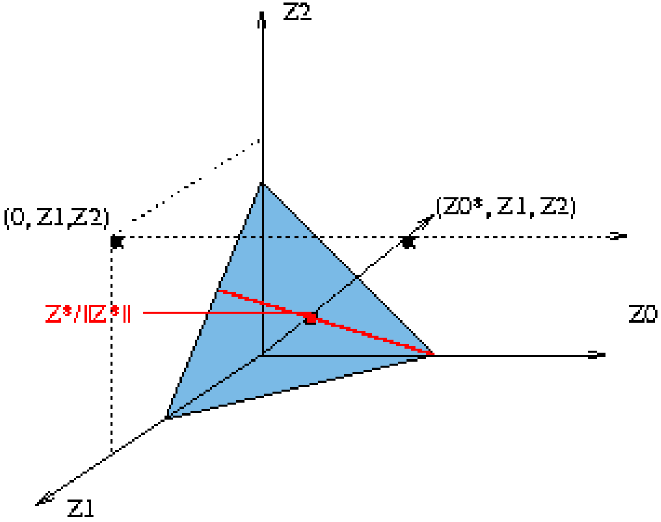


$\alpha = 1, \beta = (2, 2, .5)$

Idea of a Conditional Density

Assume $\alpha = 1$. In the application, we will make a marginal transformation so that this holds.

We need to work in Cartesian coordinates.



A change of variables argument yields the Cartesian point process intensity function:

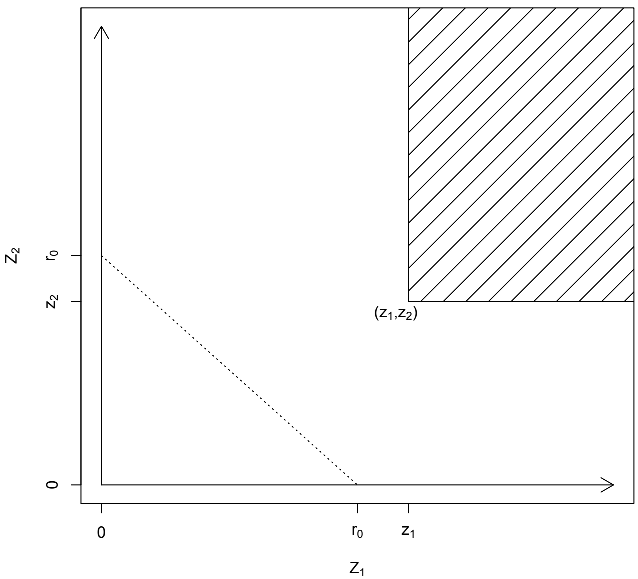
$$\nu(dr \times d\boldsymbol{w}) = r^{-2} dr h(\boldsymbol{w}) d\boldsymbol{w} \;\Rightarrow\; \nu(d\boldsymbol{z}) = \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}) d\boldsymbol{z}.$$

Obtaining a Conditional Density

We have to work a little to obtain a “density”.

Define the conditional survival function

$$F_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0) = P\left(\frac{\boldsymbol{Z}}{a_n} \in [\boldsymbol{z}, \infty) \;\middle|\; \frac{\|\boldsymbol{Z}\|}{a_n} > r_0\right).$$



Toward a Limiting Density for Large Values

$$\begin{aligned}
F_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0) &= P\left(\frac{\boldsymbol{Z}}{a_n} \in [\boldsymbol{z}, \infty) \;\middle|\; \frac{\|\boldsymbol{Z}\|}{a_n} > r_0\right) \\
&= \frac{nP\left(\frac{\boldsymbol{Z}}{a_n} \in [\boldsymbol{z}, \infty)\right)}{nP\left(\frac{\|\boldsymbol{Z}\|}{a_n} > r_0\right)} \\
&\to \frac{\nu([\boldsymbol{z}, \infty))}{\nu(\{\boldsymbol{z} \mid \|\boldsymbol{z}\| > r_0\})} \\
&= r_0 \nu([\boldsymbol{z}, \infty)), \text{ because } \textstyle\int_{r > r_0} r^{-2} dr = r_0^{-1} \\
&= r_0 \int_{[\boldsymbol{z}, \infty)} \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}) d\boldsymbol{z}
\end{aligned}$$

We wish to speak of $f_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0)$, a limiting joint density of $\boldsymbol{Z}/a_n$ given $\|\boldsymbol{Z}\|/a_n > r_0$. We will assume that

$$f_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0) \to r_0 \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}); \text{ for } \|\boldsymbol{z}\| > r_0$$

as $n \to \infty$. True if $\frac{d}{d\boldsymbol{z}} F_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0) \overset{unif}{\longrightarrow} r_0 \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1})$.

Example: Bivariate Logistic

$$P(Z_1 \le z_1, Z_2 \le z_2) = \exp[-(z_1^{-1/\beta} + z_2^{-1/\beta})^\beta] \text{ for } \beta \in (0,1].$$

$$h(\boldsymbol{w}) = \frac{1}{2}\left(\frac{1}{\beta} - 1\right)(w_1 w_2)^{-1/\beta-1}\left(w_1^{-1/\beta} + w_2^{-1/\beta}\right)^{\beta-2}.$$

Let $a_n = 2n$. Then,

$$\begin{aligned} P\left(\frac{\boldsymbol{Z}}{2n} \in [\boldsymbol{z}, \infty)\right) &= (2nz_1)^{-1} + (2nz_2)^{-1} \\ &\quad - \left((2nz_1)^{-1/\beta} + (2nz_2)^{-1/\beta}\right)^\beta + o(n^{-1}) \\ \Rightarrow F_{\boldsymbol{Z}/2n}(\boldsymbol{z}, r_0) &\to \frac{1}{2} r_0 \left(z_1^{-1} + z_2^{-1} - (z_1^{-1/\beta} + z_2^{-1/\beta})^\beta\right). \end{aligned}$$

Differentiating, we obtain:

$$\begin{aligned} f_{\boldsymbol{Z}/2n}(\boldsymbol{z}, r_0) &\to \frac{1}{2} r_0 \left(\beta^{-1} - 1\right)\left(z_1^{-1/\beta} + z_2^{-1/\beta}\right)^{\beta-2} z_1^{-1/\beta-1} z_2^{-1/\beta-1} \\ &= r_0 \|\boldsymbol{z}\|^{-3} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}). \end{aligned} \tag{1}$$

Approximate Conditional Density for Large Values

Assume n is fixed, but large enough such that

$$f_{\boldsymbol{Z}/a_n}(\boldsymbol{z}, r_0) \approx r_0 \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}).$$

We wish to approximate $f_{\boldsymbol{Z}}(\boldsymbol{z}, r_*)$, the density of $\boldsymbol{Z}$ given that $\|\boldsymbol{Z}\| > r_*$ where r_* is large.

$$\begin{aligned} f_{\boldsymbol{Z}}(\boldsymbol{z}, r_*) &\approx r_0 \|\boldsymbol{z}/a_n\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}) a_n^{-d} \\ &= r_* \|\boldsymbol{z}\|^{-(d+1)} h(\boldsymbol{z}\|\boldsymbol{z}\|^{-1}), \end{aligned}$$

where $r_* = a_n r_0$, and thus is large.

Thus, the conditional distribution of $[Z_d \mid \boldsymbol{Z}_{-d} = \boldsymbol{z}_{-d}]$ when $\|\boldsymbol{z}_{-d}\| > r_*$

$$f_{Z_d \mid \boldsymbol{Z}_{-d}}(z_d \mid \boldsymbol{z}_{-d}) \approx \frac{\|\boldsymbol{z}\|^{-(d+1)} h\left(\frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\right)}{\int_0^\infty \|\boldsymbol{z}(t)\|^{-(d+1)} h\left(\frac{\boldsymbol{z}(t)}{\|\boldsymbol{z}(t)\|}\right) dt}.$$

Approximation Example: Trivariate Logistic

The trivariate logistic is a regularly varying random vector with distribution

$$P(Z_1 \le z_1, Z_2 \le z_2, Z_3 \le z_3) = \exp[-(z_1^{-1/\beta} + z_2^{-1/\beta} + z_3^{-1/\beta})^{\beta})]$$

for $\beta \in (0, 1]$.

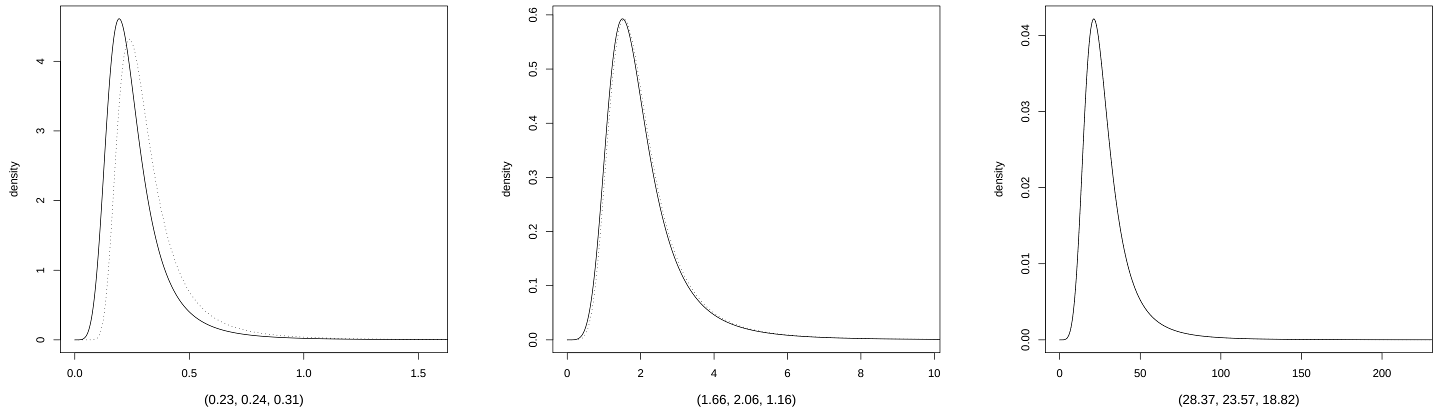
The angular measure of the trivariate logistic is given by

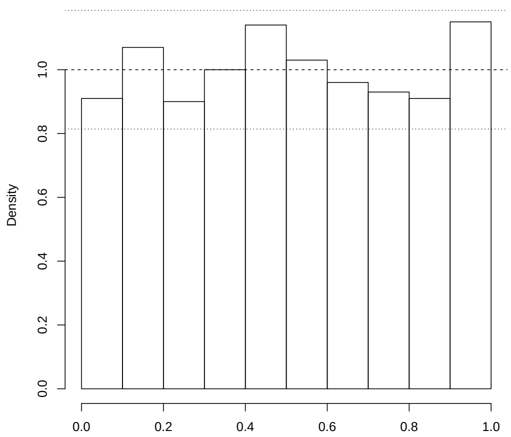
$$h(\boldsymbol{w}) = \frac{1}{3}\left(\frac{1}{\beta} - 1\right)\left(\frac{2}{\beta} - 1\right)(w_1 w_2 w_3)^{-1/\beta-1}\left(w_1^{-1/\beta} + w_2^{-1/\beta} + w_3^{-1/\beta}\right)^{\beta-3}.$$

We wish to find $[Z_3 \mid Z_1 = z_1, Z_2 = z_2]$.

True conditional density is known.
Our approximation should improve as size of $|(z_1, z_2)|$ increases.

Approximation Example





PIT histogram of the largest 1000 of 5000 total simulations.

Washington DC Data

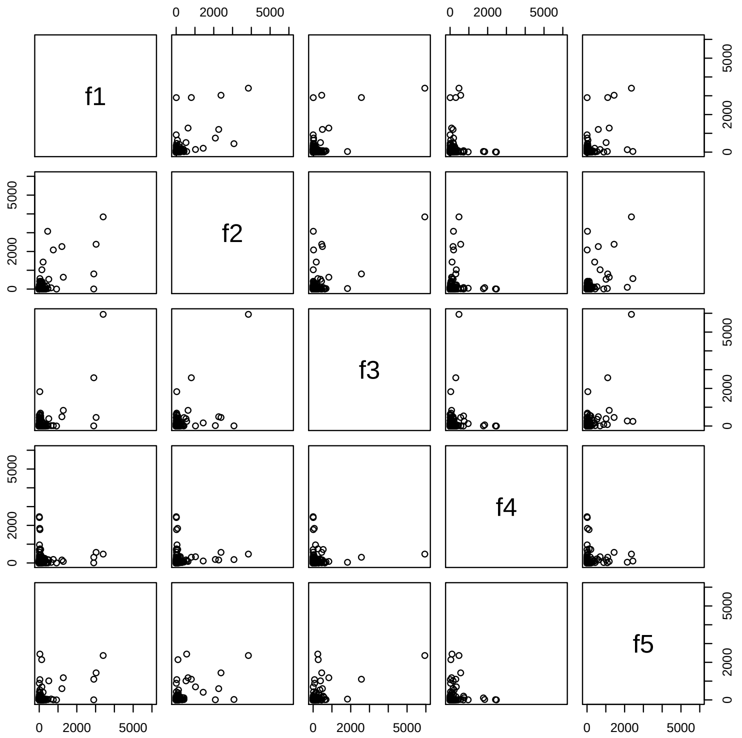
4497 daily observations between January 1, 1995 and January 31, 2010.

Divided into a training set of 2998 observations and a test set of 1499 observations.

Ignore temporal dependence, assume stationarity of tail dependence structure.

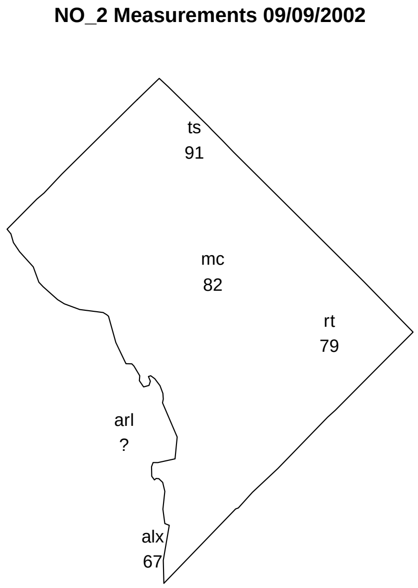
Angular measure model fit at 0.93 quantile.

Transformed Washington DC Data



Pollution at different sites exhibits tail dependence, some pairs stronger than others. Need a flexible angular measure model.

Fitted Pairwise Beta Model



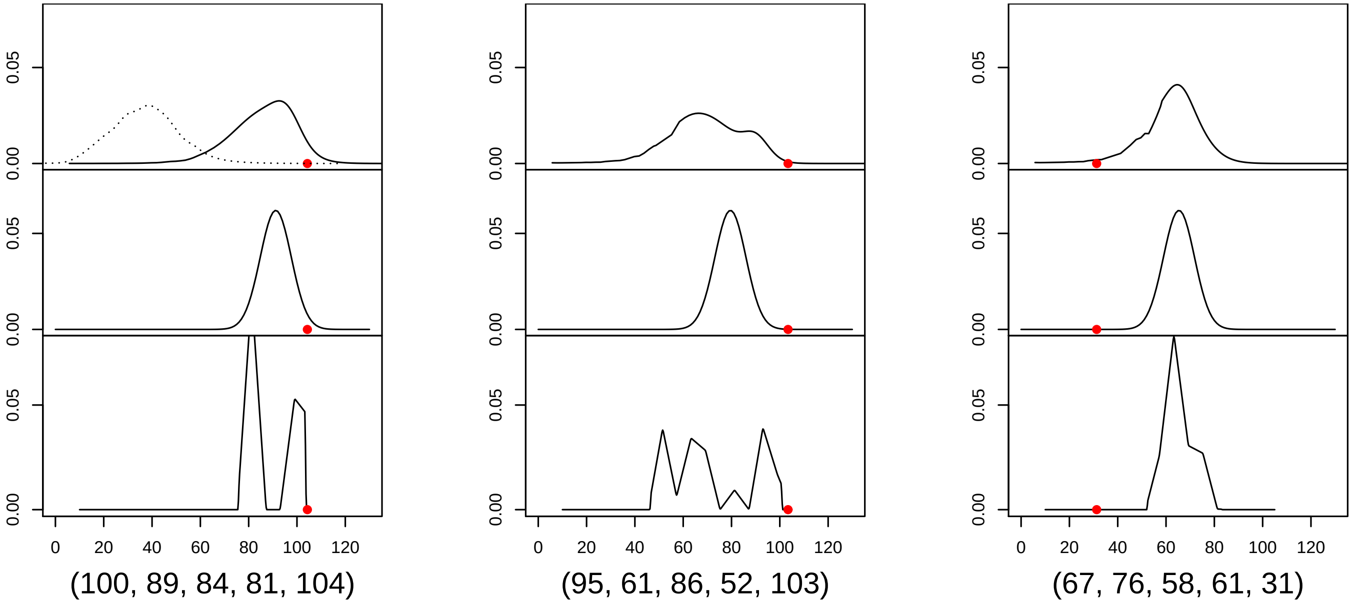
1: alx
2: mc
3: rt
4: ts
5: arl

$\widehat{\gamma}$	$\widehat{\beta}_{1,2}$	$\widehat{\beta}_{1,3}$	$\widehat{\beta}_{1,4}$	$\widehat{\beta}_{1,5}$	$\widehat{\beta}_{2,3}$	$\widehat{\beta}_{2,4}$	$\widehat{\beta}_{2,5}$	$\widehat{\beta}_{3,4}$	$\widehat{\beta}_{3,5}$	$\widehat{\beta}_{4,5}$
0.37	0.51	0.64	0.56	6.11	0.76	1.64	0.96	0.56	0.98	1.01
(0.03)	(0.18)	(0.28)	(0.19)	(2.59)	(0.44)	(1.08)	(0.51)	(0.20)	(0.51)	(0.61)

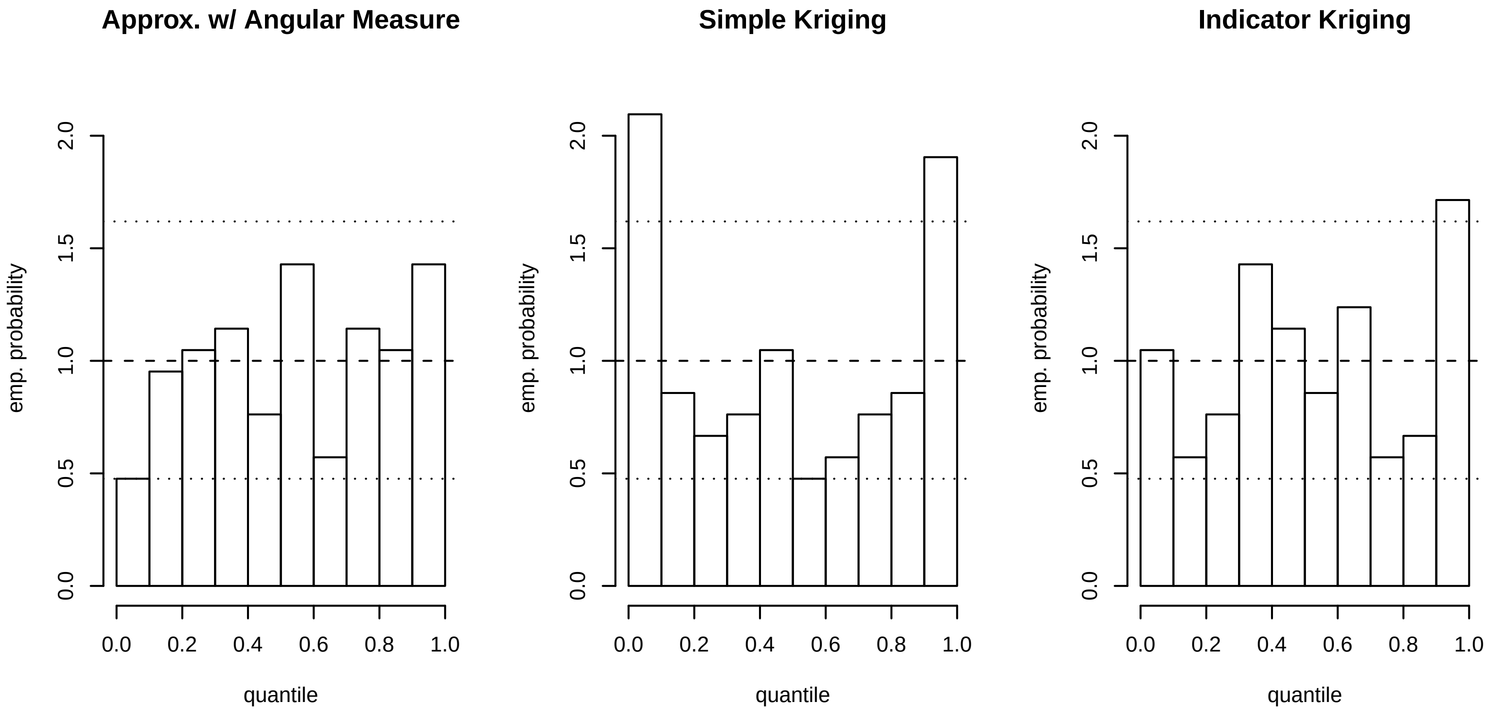
At this point, we assume to have a model which captures the tail dependence of the measurements for these five locations.

Approximating Conditional Density at Arlington

We compare our method to kriging and indicator kriging.



PIT Histograms



Evaluating Quantile Scores

How well does each method predict a high quantile?

Interpret as a conditional upper bound.

Quantile	0.99		0.95		0.90		0.75	
	Cvg	QVS	Cvg	QVS	Cvg	QVS	Cvg	QVS
Angular Measure	0.97	40.97	0.93	134.77	0.88	225.68	0.70	398.97
Simple Kriging	0.92	65.80	0.83	170.04	0.81	246.26	0.65	378.27
Indicator Kriging	0.90	67.80	0.86	153.41	0.83	238.63	0.73	377.20
Sampling Error	(0.01)	–	(0.02)	–	(0.03)	–	(0.04)	–

Summary for Part B

- Our interest lies in cases when observations are large . . .
- . . . so we model *only* the tail of the distribution and use angular measure to approximate conditional density.
- Approach allows us to answer any related question (e.g. 95% quantile of predicted distribution, probability of exceeding a level of interest).
- Seems to outperform methods devised for entire distribution (and it should!)
- An interesting application of *multivariate* extremes.
- Was not really spatial as we divided into training and test sets.

References

Ballani, F. and Schlather, M. (2011). A construction principle for multivariate extreme value distributions. *Biometrika*, 98:633–645.

Coles, S. and Tawn, J. (1991). Modeling multivariate extreme events. *Journal of the Royal Statistical Society, Series B*, 53:377–92.

Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer-Verlag London Ltd., London.

Cooley, D., Davis, R. A., and Naveau, P. (2010). The pairwise beta: A flexible parametric multivariate model for extremes. *Journal of Multivariate Analysis*, 101:2103–2117.

Cooley, D., Davis, R. A., and Naveau, P. (2012). Approximating the conditional density given large observed values via a multivariate extremes framework, with application to environmental data. *Annals of Applied Statistics*, 6:1406–1429.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistics and Probability Letters*, 9:75–81.

Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

Tawn, J. (1988). Bivariate extreme value theory: models and estimation. *Biometrika*, 75:397–415.