Multivariate Scenario Sets in the Non-Gaussian World

Discussion

Alexander J. McNeil[1]

[1] Heriot-Watt University Edinburgh

BIRS, Banff, 23 May 2013

a.j.mcneil@hw.ac.uk

Statement of Problem

You are given a multivariate (non-Gaussian) distribution (either theoretical or empirical) and asked to construct a set of multivariate scenarios that contains plausible scenarios but excludes the "most extreme" scenarios.

I aim to generate some discussion of this question including:

1. Why is this problem interesting?
2. How would you go about it? Using density or depth (half-space, simplex, other)?
3. Is the computation of the set feasible?

Financial Risk

- The random vector $\boldsymbol{X}$ represents a set of financial risk factors that effect the profitability of a portfolio or the solvency of a company.
- Which possible values of $\boldsymbol{X}$ should we worry about? We can't worry about all of them (particularly in high dimensions) and have to specify the set S of plausible scenarios.
- Among the plausible scenarios $\boldsymbol{x} \in S$ we might want to examine the worst possible impact $\ell(\boldsymbol{x})$ for some function ℓ. For a portfolio of assets this might simply be a linear function.
  (LSLE - least solvent likely event.)
- Related problem. Among a particular set of ruin scenarios $\{\boldsymbol{x} : \boldsymbol{x} \in R\}$ what is the most plausible way of being ruined?
  (MLRE - most likely ruin event.)

Notation

- For any point $\mathbf{y} \in \mathbb{R}^d$ and any directional vector $\mathbf{u} \in \mathbb{R}^d \setminus \{0\}$, consider the closed half space

$$H_{\mathbf{y},\mathbf{u}} = \left\{\mathbf{x} \in \mathbb{R}^d \ : \ \mathbf{u}'\mathbf{x} \leq \mathbf{u}'\mathbf{y}\right\},$$

  bounded by the hyperplane through $\mathbf{y}$ with normal vector $\mathbf{u}$.
- The probability of the half-space is written

$$P_{\mathbf{X}}(H_{\mathbf{y},\mathbf{u}}) = P(\mathbf{u}'\mathbf{X} \leq \mathbf{u}'\mathbf{y}).$$

- We define an α-quantile function on $\mathbb{R}^d \setminus \{0\}$ by writing $q_\alpha(\mathbf{u})$ for the α-quantile of the random variable $\mathbf{u}'\mathbf{X}$.

Quantile Depth Set

Let $\alpha > 0.5$ be fixed. We write our scenario set in two ways:

1. $$Q_\alpha = \bigcap \{H_{\mathbf{y},\mathbf{u}} \; : \; P_{\mathbf{X}}(H_{\mathbf{y},\mathbf{u}}) \geq \alpha\} \, ,$$

   the intersection of all closed half spaces with probability at least α;

2. $$Q_\alpha = \{\mathbf{x} \; : \; \mathbf{u}'\mathbf{x} \leq q_\alpha(\mathbf{u}) \, , \forall \mathbf{u}\} \; , \qquad (1)$$

   the set of points for which linear combinations are no larger than the quantile function.

Relation to Usual Depth Set

Our set differs very slightly from the usual depth set.
Depth at a point $\boldsymbol{x}$ is usally defined to be

$$\text{depth}(\boldsymbol{x}) = \inf_{\boldsymbol{u}:\boldsymbol{u}\neq\boldsymbol{0}} P_{\boldsymbol{X}}(H_{\boldsymbol{x},\boldsymbol{u}}) \,,$$

and the depth set to be

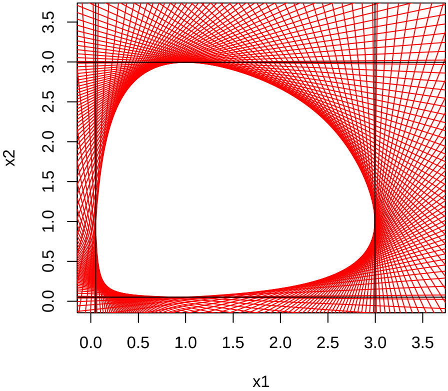
$$D_\alpha = \left\{\boldsymbol{x} \in \mathbb{R}^d \;:\; \text{depth}(\boldsymbol{x}) \geq 1 - \alpha\right\} \,,$$

i.e. points which are at least $1 - \alpha$ deep into the distribution. It may be shown [Rousseeuw and Ruts, 1999] that

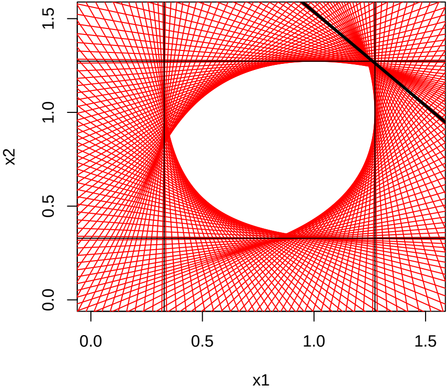
$$D_\alpha = \bigcap \{H_{\mathbf{y},\mathbf{u}} \;:\; P_{\mathbf{X}}(H_{\mathbf{y},\mathbf{u}}) > \alpha\} \;.$$

Q_α and D_α coincide when $\boldsymbol{X}$ has a density.

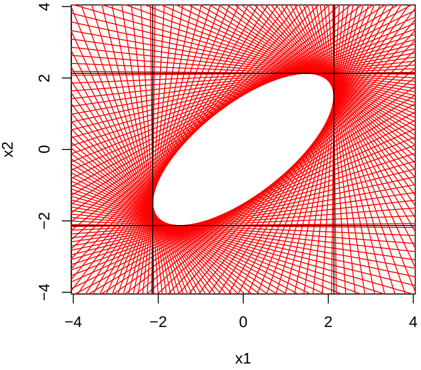
Two Independent Exponentials, $Q_{0.95}$



Two Independent Exponentials, $Q_{0.75}$



A bivariate Student distribution, $Q_{0.95}$



$\nu = 4$, $\rho = 0.7$

Commentary on examples

- Note how the depth set in the exponential case has a smooth boundary for $\alpha = 0.95$. (Supporting hyperplanes in every direction.)
- Note how the depth set in the exponential case has a sharp corners for $\alpha = 0.75$. (No supporting hyperplanes in some directions.)
- The depth set for an elliptical distribution is an ellipsoid.
- For elliptical distributions both the contours of equal depth and the contours of equal density are ellipsoidal.

Literature

- Origins of concepts: data depth [Tukey, 1975]; multivariate analogues of quantiles [Eddy, 1984].
- Multivariate trimming: [Nolan, 1992, Massé and Theodorescu, 1994].
- Depth function for population distributions: [Rousseeuw and Ruts, 1999].
- Estimation: [Ruts and Rousseeuw, 1996, Rousseeuw and Strufy, 1998, Rousseeuw et al., 1999]
- Other concepts of depth (such as simplex): [Liu et al., 1999, Zuo and Serfling, 2000].
- Use of concepts in risk analysis: [McNeil and Smith, 2012].

Coherent Risk Measures, Linear Portfolios

A risk measure $\varrho : \mathcal{M} \mapsto \mathbb{R}$ is said to be coherent on a set of random variables $\mathcal{M}$ if it satisfies the following axioms for random variables $L \in \mathcal{M}$ (representing financial losses).

Monotonicity. $L_1 \leq L_2 \Rightarrow \varrho(L_1) \leq \varrho(L_2)$.

Translation invariance. For $m \in \mathbb{R}$, $\varrho(L + m) = \varrho(L) + m$.

Subadditivity. For $L_1, L_2 \in \mathcal{M}$, $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$.

Positive homogeneity. For $\lambda \geq 0$, $\varrho(\lambda x) = \lambda \varrho(x)$.

Let **X** be a fixed random vector and define the linear portfolio set:

$$\mathcal{M} = \left\{ L : L = m + \boldsymbol{\lambda}' \mathbf{X},\ m \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^d \right\}.$$

Key Result

Theorem

A risk measure ϱ on the linear portfolio set $\mathcal{M}$ is coherent if and only if it has the representation

$$\varrho(L) = \varrho(m + \boldsymbol{\lambda}'\mathbf{X}) = \sup\{m + \boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in S_\varrho\} \tag{2}$$

where S_ϱ is the scenario set

$$S_\varrho = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{u}'\mathbf{x} \leq \varrho(\mathbf{u}'\mathbf{X}), \forall \mathbf{u} \in \mathbb{R}^d\}\,.$$

The scenario set is a closed convex set and we may conclude that, for given $\boldsymbol{\lambda}$, there is a worst case scenario (obtainable by convex optimization)

$$\mathbf{x}_{LSLE} = \arg\max\{\boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in S_\varrho\}\,.$$

The Case of VaR

Let us suppose the risk measure $\varrho = \text{VaR}_\alpha$ for some value $\alpha > 0.5$. Then the scenario set S_ϱ is as given in (1), i.e.

$$\{\mathbf{x} \in \mathbb{R}^d : \mathbf{u}'\mathbf{x} \leq q_\alpha(\mathbf{u}), \forall \mathbf{u} \in \mathbb{R}^d\} = Q_\alpha \, .$$

- However VaR_α is not a coherent risk measure in general.
- It is a coherent risk measure for linear portfolios of elliptically-distributed risks.
- In other cases the relationship (2) must break down.

The Case of VaR for Elliptical Distributions

Theorem

Suppose that $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ *(an elliptical distribution centred at* $\boldsymbol{\mu}$ *with dispersion matrix* Σ *and type* ψ*) and let* $\mathcal{M}$ *be the space of linear portfolios. Then* VaR_α *is coherent on* $\mathcal{M}$ *for* $\alpha > 0.5$.

In the elliptical case the scenario set is

$$Q_\alpha = \{\mathbf{x} : (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \leq k_\alpha^2\}$$

where $k_\alpha = \mathrm{VaR}_\alpha(Y)$ and $Y \sim E_1(0, 1, \psi)$.
The worst case scenario for a given portfolio is easily computed (Lagrange multipliers) to be

$$\mathbf{x}_{\mathrm{LSLE}} = \boldsymbol{\mu} + \frac{\Sigma\boldsymbol{\lambda}}{\sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}} k_\alpha,$$

and the corresponding loss is

$$\mathrm{VaR}_\alpha(m + \boldsymbol{\lambda}'\mathbf{X}) = m + \boldsymbol{\lambda}'\mathbf{x}_{\mathrm{LSLE}} = m + \boldsymbol{\lambda}'\boldsymbol{\mu} + \sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}} k_\alpha.$$

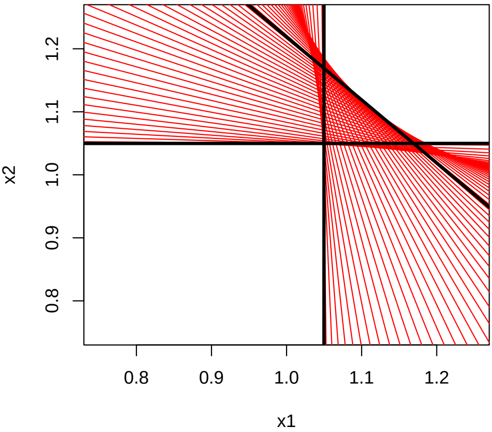
The Case of VaR for Non-Elliptical Distributions

- In the non-elliptical case it may happen that VaR_α is not coherent on $\mathcal{M}$ for some value of α. In such situations we may find portfolio weights $\boldsymbol{\lambda}$ such that

$$\text{VaR}_\alpha(L) = \text{VaR}_\alpha(m + \boldsymbol{\lambda}'\mathbf{X}) > \sup\left\{m + \boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in Q_\alpha\right\} .$$

- Such a situation was shown earlier. It occurs when some lines bounding half-spaces with probablity α are not supporting hyperplanes for the set Q_α, i.e. they do not touch it.
- In such situations we can construct explicit examples to show that VaR_α violates the property of subadditivity.

Two Independent Exponentials, $Q_{0.65}$



Demonstration of Super-Additivity

- In previous slide we set $\alpha = 0.65$ and consider loss $L = X_1 + X_2$.
- Diagonal line is $x_1 + x_2 = q_\alpha(X_1 + X_2)$ which obviously intersects axes at $(0, q_\alpha(X_1 + X_2))$ and $(q_\alpha(X_1 + X_2), 0)$.
- Horizontal (vertical) lines are at $q_\alpha(X_1)$.
- We infer
  1. $x_1 + x_2 < q_\alpha(X_1 + X_2)$ in the depth set;
  2. $\sup \{x_1 + x_2 : \mathbf{x} \in Q_\alpha\}$ is a poor lower bound
  3. $q_\alpha(X_1 + X_2) > q_\alpha(X_1) + q_\alpha(X_2)$ (non-subadditivity of quantile risk measure)

Remark: An Upper Bound for VaR

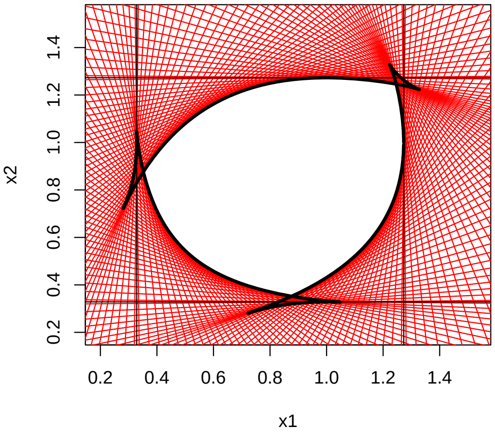
- Assume differentiability of quantile function q_α and define an outer scenario set as

$$O_\alpha = \{\mathbf{x} \ : \ \mathbf{x} = \nabla q_\alpha(\mathbf{u}), \mathbf{u} \neq \mathbf{0}\}$$

- Let $\psi(\mathbf{u}) = \sup\{\mathbf{u}'\mathbf{x} \ : \ \mathbf{x} \in O_\alpha\}$ be the worst scenario in this set.
- It can be shown that $q_\alpha(\mathbf{u}) \leq \psi(\mathbf{u})$ with equality for all $\mathbf{u} \in \mathbb{R}^d$ if and only if q_α is sub-additive.
- This gives the upper bound for VaR:

$$\mathrm{VaR}_\alpha(L) = \mathrm{VaR}_\alpha(m + \boldsymbol{\lambda}'\mathbf{X}) \leq \sup\left\{m + \boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in O_\alpha\right\} .$$

Outer Set for $Q_{0.72}$



The Case of Expected Shortfall

Consider the expected shortfall risk measure $\varrho = \text{ES}_\alpha$, which is known to be a coherent risk measure given by

$$\text{ES}_\alpha(L) = \frac{\int_\alpha^1 \text{VaR}_\theta(L)d\theta}{1-\alpha}, \quad \alpha \in (0.5, 1),$$

and write $e_\alpha(\mathbf{u}) := \text{ES}_\alpha(\mathbf{u}'\mathbf{X})$.
Since expected shortfall is a coherent risk measure (irrespective of $\mathbf{X}$) it must have the stress test representation

$$\text{ES}_\alpha(L) = \varrho(m + \boldsymbol{\lambda}'\mathbf{X}) = \sup\{m + \boldsymbol{\lambda}'\mathbf{x} : \mathbf{x} \in E_\alpha\}$$

where

$$E_\alpha := \{\mathbf{x} \ : \ \mathbf{u}'\mathbf{x} \le e_\alpha(\mathbf{u}), \forall \mathbf{u}\}.$$

The Case of ES for Elliptical Distributions

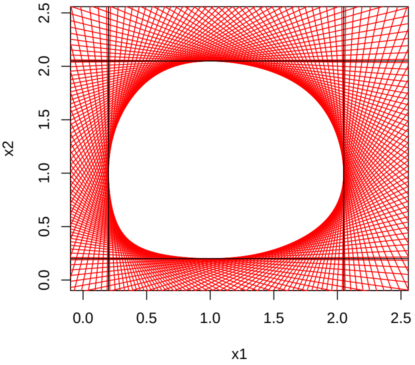
If $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \psi)$ is elliptically distributed then the scenario set is simply the ellipsoidal set

$$E_\alpha = \{\mathbf{x} \; : \; (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \leq l_\alpha^2\},$$

where $l_\alpha = \mathrm{ES}_\alpha(Y)$ and $Y \sim E_1(0, 1, \psi)$.
The worst case scenario is given by

$$\mathbf{x}_{\mathrm{LSLE}} = \boldsymbol{\mu} + \frac{\Sigma\boldsymbol{\lambda}}{\sqrt{\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}}} \, l_\alpha \; .$$

The Case of ES for Non-Elliptical Distributions



The set $E_{0.65}$. Recall that $Q_{0.65}$ did not have smooth boundary.

When Can Depth Sets be Computed?

- For a given random vector $\boldsymbol{X}$ in $\mathbb{R}^d$ (assumed to have a density) we would like to be able to say whether a point $\boldsymbol{x}$ is in the depth set Q_α. Is it plausible or not?
- Equivalently, is it true that

$$\text{depth}(\boldsymbol{x}) = \inf_{\boldsymbol{u} \neq \boldsymbol{0}} P_{\boldsymbol{X}}(H_{\boldsymbol{x},\boldsymbol{u}}) \geq 1 - \alpha \quad ?$$

- It is particularly nice if we can get a parametric equation for Q_α.
- For elliptical distributions we get ellipsoids.
- What about copulas? It seems less easy to compute Q_α. An exception is the independence copula for $d = 2$ where

$$Q_\alpha = \left\{ \boldsymbol{x} \in \mathbb{R}^2 \ : \ 2\min(x_1, 1 - x_1)\min(x_2, 1 - x_2) \geq 1 - \alpha \right\} .$$

  [Rousseeuw and Ruts, 1999]
- It seems to be possible to compute the sets for skew-t distributions (Giorgi 2013).
- Generalized hyperbolic distributions?

Normal Inverse-Gaussian Model

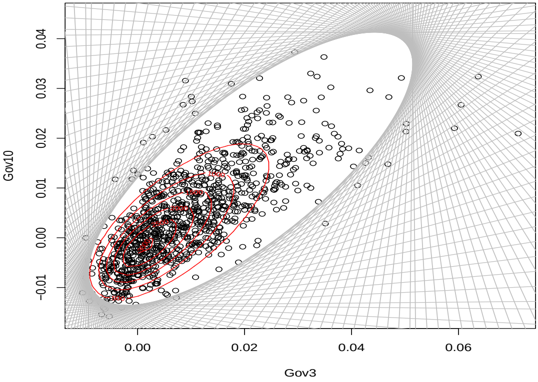


Figure: Points are changes in yields for 3-year and 10-year government bonds. A NIG distribution has been fitted and scenario sets calculated.

Independence Copula, $Q_{0.75}$

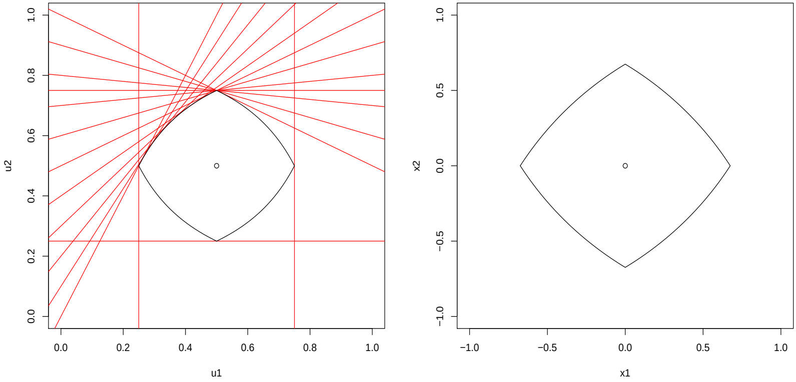


Figure: Left: $Q_{0.75}$ for the independence copula; note all hyperplanes supporting. Right: set transformed to Gaussian scale; note - not a circle!

Independence Copula, $Q_{0.99}$

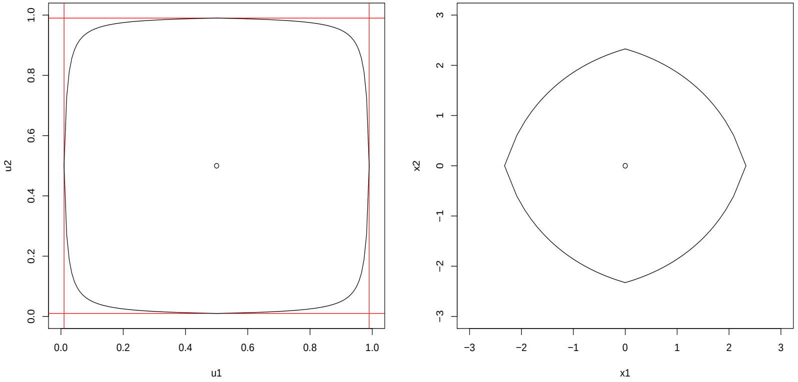


Figure: Left: $Q_{0.99}$ for the independence copula. Right: set transformed to Gaussian scale; note - still not a circle!

Empirical Estimates of Depth and Depth Contours

- Recall that $\text{depth}(\boldsymbol{x}) = \inf_{\boldsymbol{u} \neq \boldsymbol{0}} P(\boldsymbol{u}'\boldsymbol{X} \leq \boldsymbol{u}'\boldsymbol{x})$.
- Given data vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ we compute

$$\widehat{\text{depth}}(\boldsymbol{x}) = \inf_{\boldsymbol{u} \neq \boldsymbol{0}} \frac{1}{n} \sum_{i=1}^{n} I(\boldsymbol{u}'\boldsymbol{X}_i \leq \boldsymbol{u}'\boldsymbol{x}) \,.$$

- Exact computation for $d = 2$ and $d = 3$ possible. Approximate algorithms for $d > 3$ and/or n large [Ruts and Rousseeuw, 1996, Rousseeuw and Strufy, 1998].
- Plot of depth contours often called a bagplot [Rousseeuw et al., 1999].
- R package *depth* available including function *isodepth*.
- Literature on other empirical depth measures [Liu et al., 1999].

Non-Parametric Estimation

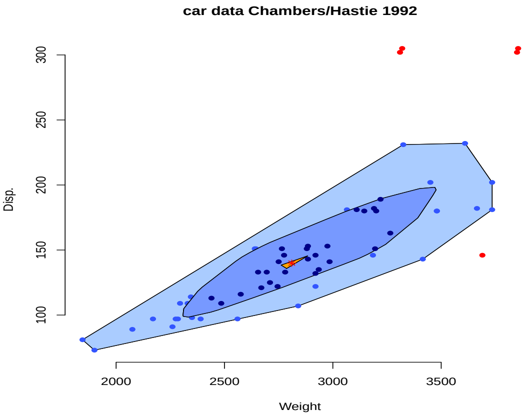
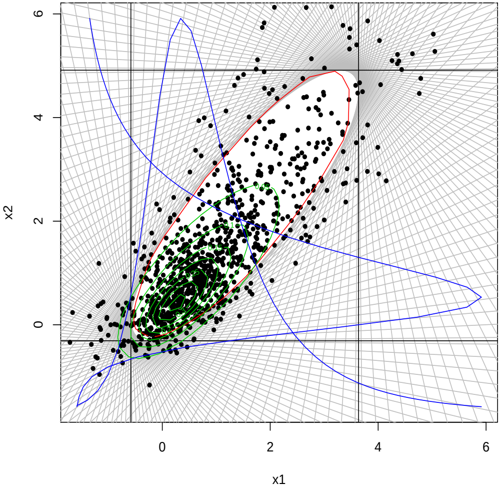


Figure: A so-called bagplot.

NIG Example, $Q_{0.95}$



References

Eddy, W. (1984).
Set-valued orderings for bivariate data.
In Ambartzumian, R. and Weil, W., editors, *Stochastic Geometry, Geometric Statistics, Stereology (Proceedings of an Oberwohlfach Conference in 1983)*, pages 79–90. Teubner-Texte für Mathematik 56, Leipzig.

Liu, R., Parelius, J., and Singh, K. (1999).
Multivariate analysis by data depth: descriptive statistics, graphics and inference (with discussion).
Annals of Statistics, 27(3):783–858.

Massé and Theodorescu (1994).
Halfplane trimming for bivariate distributions.
Journal of Multivariate Analysis, 48:188–202.

References (cont.)

McNeil, A. and Smith, A. (2012).
Multivariate stress scenarios and solvency.
Insurance: Mathematics and Economics, 50(3):299–308.

Nolan, D. (1992).
Asymptotics for multivariate trimming.
Stochastic Processes and their Applications, 42:157–169.

Rousseeuw, P. and Ruts, I. (1999).
The depth function of a population distribution.
Metrika, 49:213–244.

Rousseeuw, P., Ruts, I., and Tukey, J. (1999).
The bagplot: a bivariate boxplot.
The American Statistician, 53(4):382–387.

References (cont.)

Rousseeuw, P. and Strufy, A. (1998).
Computing location depth and regression depth in higher dimensions.
Statistics and Computing, 8:193–203.

Ruts, I. and Rousseeuw, P. (1996).
Computing depth contours of bivariate clouds.
Computational Statistics and Data Analysis, 23:153–168.

Tukey, J. (1975).
Mathematics and the picturing of data.
In *Proceedings of the International Congress of Mathematicians, Vancouver*, volume 2, pages 523–531.

Zuo, Y. and Serfling, R. (2000).
General notions of statistical depth function.
Annals of Statistics, 28(2):461–482.