LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
INSTITUT FÜR STATISTIK

Georg Schollmeyer

Lower quantiles for complete lattices

Technical Report Number 207, 2017
Department of Statistics
University of Munich

http://www.stat.uni-muenchen.de

Lower quantiles for complete lattices

Georg Schollmeyer

Abstract

In this technical report a notion of (lower) quantiles for data or random variables with values in a complete lattice is developed. We list a number of desirable properties a reasonable notion of quantiles should have and analyze for different proposals of quantiles, which of these properties they fulfill. It turns out that one of the constructions has acceptable properties and can thus be used for analyzing lattice-valued data.

Keywords: lower quantile, complete lattice, level function, representation invariance, data depth

1 Introduction

In this technical report we develop a generalization of the known concept of univariate (lower) quantiles to the case of a data set or a random variable with values in a complete lattice $\mathbb{L}$. Such random variables will be called $\mathbb{L}$-valued random variables in the sequel.

For the case of multivariate data in $\mathbb{R}^d$, there already exists a broad literature (cf., e.g., [Mosler, 2013]) generalizing the concept of a quantile and especially the median as a measure of location to the multivariate situation. One basic tool for establishing such generalizations is the notion of *data depth*. A depth-function measures how deep a certain data point lies w.r.t. a data cloud or w.r.t. a probability distribution. The most popular depth-function is Tukey's half-space depth (Tukey [1975]). Other proposals are for example simplicial depth ([Liu et al., 1990]), Oja depth ([Oja, 1983]), zonoid depth [Koshevoy and Mosler, 1997] or L^p depth ([Zuo and Serfling, 2000]), to name just a few. A more recent proposal for data depth is the Monge-Kantorovich depth ([Chernozhukov et al., 2017]) based on the Monge-Kantorovich theory of measure transportation.

However, the above treatments[1] of data depth are basically devoted to the Euclidean d - space as a linear space and put focus on affine invariance[2] as a desirable property[3]. This property seems natural especially if one has a geometrical understanding of $\mathbb{R}^d$ (cf. also [Chaudhuri, 1996]). However, in some situations of application like for example in multidimensional poverty/inequality analysis (cf., e.g., Alkire et al. [2015]) or in the analysis of multidimensional psychological latent constructs like for example "autoritarism" (cf., e.g., Duckitt et al. [2010]), a geometrical understanding is not suggesting itself. An order theoretic underpinning looking at the point-wise ordering, where for example a person could be termed as less poor than another person if she is less poor w.r.t. all considered dimensions of poverty, seems to be more natural in such cases.

For real-valued random variables one common definition of an α-quantile of a random variable X is the following: A value q with the property $\mathbb{P}(X \leq q) \geq \alpha$ and $\mathbb{P}(X \geq q) \geq 1 - \alpha$ is called an α-quantile of X. Analogously, for a data set $(x_1, \ldots, x_n)$ a value q is called an α-quantile if it is an α-quantile for a random variable X that has the same distribution as the empirical distribution of the sample $(x_1, \ldots, x_n)$. This would translate into: q is an α-quantile for the data set $(x_1, \ldots, x_n)$ iff there are at least $\lceil \alpha \cdot n \rceil$ data points that are greater than or equal to q and there are at least $\lceil (1-\alpha) \cdot n \rceil$ values lower than or equal to q. Especially a 0.5-quantile (also called median) is every value between the $\lfloor \frac{n+1}{2} \rfloor$-th and the $\lceil \frac{n+1}{2} \rceil$-th value of the ordered data. So generally, α-quantiles are not unique. Sometimes, this non-uniqueness is avoided by taking for example the average of the $\lfloor \frac{n+1}{2} \rfloor$-th and the $\lceil \frac{n+1}{2} \rceil$-th value of the ordered data for the 0.5-quantile, but this implicitly uses the addition and multiplication in $\mathbb{R}$ and not only the order structure in $\mathbb{R}$. If only the order structure of the data represented by real numbers has meaning, then this way to proceed is somehow arbitrary (but not necessarily unreasonable). Furthermore, if not all values in $\mathbb{R}$ have a meaning in the sense that only values in a proper subset $M \subset \mathbb{R}$ are possible data-values, than not every value q in $\left[x_{[\lfloor \frac{n+1}{2} \rfloor]}, x_{[\lceil \frac{n+1}{2} \rceil]} \right]$ is a member of M and so it is reasonable to consider only values $q \in M$ as quantiles. Another way to cope with the non-uniqueness of quantiles is an asymmetric definition of a lower (and an upper) quantile:

[1]For a treatment within the lattice-valued case, which has actually a very similar starting point like that of the present paper, see Cardin [2012].

[2]Affine invariance means that the depth-function is invariant under a simultaneous affine transformation of both the considered data point and the data cloud/probability distribution, cf., [Mosler, 2013, p. 3].

[3]An axiomatic approach to data depth can be found in Liu et al. [1990] and Zuo and Serfling [2000], who propose four properties that could generally be considered desirable for a statistical depth function: a) affine invariance, b) maximality at the center, c) linear monotonicity relative to the deepest point and d) vanishing at infinity.

Definition 1. Let X be a real-valued random variable with distribution function F_X. A value $q \in \mathbb{R}$ is lower α-quantile for $\alpha \in \mathrm{Im}(F_X)$ if it satisfies one of the following equivalent conditions:

1. q is a minimal element (w.r.t. the ordering in $\mathbb{R}$) in $F_X^{-1}(\alpha)$.

2. q is a minimal element in $F_X^{-1}(\uparrow \alpha)$.

3. q is the smallest element in $F_X^{-1}(\alpha)$.

4. q is the smallest element in $F_X^{-1}(\uparrow \alpha)$.

Here, $\uparrow a := \{b \mid b \geq a\}$ (and similarly $\downarrow a := \{b \mid b \leq a\}$). Analogously, q is an upper α-quantile if it is a lower $(1-\alpha)$-quantile w.r.t. the dual order $(\mathbb{R}, \geq)$. For a data set $(x_1, \ldots, x_n)$ a value q is a lower (upper) α-quantile if it is a lower (upper) α-quantile for a random variable X with the same distribution as the empirical distribution of the data.

The statement of the different equivalent conditions in the above definition is mentioned to indicate that for generalizing the notion of a quantile, there are many thinkable ways. introduced because for a generalization of the concept of lower quantiles for $\mathbb{L}$-valued random variables the different conditions lead to different generalizations of lower quantiles. In the next sections, we will try to find a reasonable generalization for the notion of a quantile for lattice-valued data or random variables.

The paper is structured as follows: In Section 2 we firstly list some desirable properties a reasonable generalization of a quantile concept should have and formalize a notion of representation invariance which is one of these desirable properties. In Section 3 we give some simple proposals of candidates for quantiles (called prequantiles) that satisfy at least a subset of the desired properties. In Section 4 we further modify one of the more promising prequantile constructions from Section 3 to finally satisfy all but one of the listed desirable properties. (Only a "richness"-desire of quantiles cannot be guaranteed to be fulfilled.) In Section 5 we define and briefly analyze a qualitative and a quantitative "measuring-mapping" based on the developed quantile construction, which can be seen as the order theoretic counterpart to depth-contours and depth-functions, respectively, that are known from classical multivariate data depth. Finally Section 6 concludes by briefly indicating possible areas of application.

2 Basic conceptualizations

Before developing different possible generalizations of the concept of quantiles to lattice-valued data or random variables, we list different conditions that a reasonable generalization of a lower quantile should satisfy:

i) It should be independent of the concrete representation of the underlying lattice and the random variable (representation invariance).

ii) It should be non-arbitrary. (This is only an informal statement that is exemplified shortly in the sequel.)

iii) It should be defined for every arbitrary random variable with values in a given lattice. In particular, it should not be restricted to empirical distributions.

If we think of generalized lower quantiles as a descriptive tool for summarizing a data set, then the structure of lower quantiles should be not too complex and at least not more complex than the raw data itself. Thus, especially the following further properties would be desirable:

iv) For a given level α, there should be maximal one corresponding lower quantile.

v) Lower quantiles for different levels should be comparable.

vi) Lower quantiles should be minimal in the sense that for a quantile q with level α there should be no other element $q' \leq q$ with the same level α.

vii) If possible, the system of quantiles should be "rich enough" to allow for a fine-graded data analysis. (This is of course also only an informal statemnt.)

For $\mathbb{L}$-valued random variables, events of the form $\{\omega \in \Omega \mid X(\omega) \leq q\}$ need not to be measurable, so for generalized quantiles we use the inner measure m_* of the image measure m of X, here. Concretely, in the sequel, we will work with what we will call a belief structure[4] in the sequel:

Definition 2. A belief structure is a quadrupel $(\mathbb{L}, m, \mathcal{F}, B_m)$ where $\mathbb{L} = (L, \leq)$ is a complete lattice, $\mathcal{F}$ is a σ-algebra on L , $m : \mathcal{F} \longrightarrow [0,1]$ is a probability measure on the measurable space $(L, \mathcal{F})$ and $B_m : L \longrightarrow [0,1] : x \mapsto m_*(\downarrow x) = m_*(\{y \in L \mid y \leq x\})$ is the associated belief function associated to m, where $m_* : 2^L \longrightarrow [0,1] : A \mapsto \sup_{B \in \mathcal{F}, B \subseteq A} m(B)$ is the inner measure associated to m.

Remark 1. Note that for a belief structure $(\mathbb{L}, m, \mathcal{F}, B_m)$, the inner measure m_* is continuous from above and supermodular, see [Denneberg, 1994, p. 22]. Furthermore, the continuity from above and the supermodularity also translate to the belief function B_m.

[4]The choice of the name belief structure is due to the similarity to the notion of a belief function in the Dempster Shafer theory of evidence, cf., Shafer [1976]. Compared to classical Dempster Shafer theory which deals with the power set as the underlying lattice, we are here concerned with arbitrary lattices, for generalizations of the Dempster Shafer approach to lattices, see e.g., [Grabisch, 2009, Zhou, 2013].

Now, for given $\alpha \in \mathrm{Im}(B_m)$ there need not to exist a minimal element in $B_m^{-1}(\alpha)$ at all.[5] Furthermore, often there is more than one minimal element in $B_m^{-1}(\alpha)$. If one wishes to overcome this non-uniqueness issue, one could choose one element among all minimal elements in $B_m^{-1}(\alpha)$. One could choose arbitrarily the “first element that comes across”, but this would be arbitrary and also generally not representation invariant because such a choice is possibly dependent on the exact representation of the underlying lattice. If one does the “choice”[6] only with the help of the lattice operations and the belief structure than one could expect that the choice is representation invariant. To make this more precise, we state the following definition:

Definition 3. A mapping $f : 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}}$ is called representation invariant if for every order-automorphism Φ on $\mathbb{L}$ and for every $A \in 2^{\mathbb{L}}$ we have

$$f(A) = \tilde{\Phi}(f(\Phi^{-1}(A))$$

where $\tilde{\Phi} : 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}} : A \mapsto \{\Phi(a) \mid a \in A\}$ is the set-valued version of Φ. Analogously, a mapping $f : 2^{\mathbb{L}} \longrightarrow \mathbb{L}$ is called representation invariant if the associated mapping $\tilde{f} : 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}} : A \mapsto \{f(A)\}$ is representation invariant.

Furthermore, a mapping $f : \mathscr{P}(\mathbb{L}) \longrightarrow 2^{\mathbb{L}}$, where $\mathscr{P}(\mathbb{L})$ is the space of all probability measures on an arbitrary measurable space $(\mathbb{L}, \mathcal{E})$, is called representation invariant if for every order-automorphism Φ on $\mathbb{L}$ and for every $m \in \mathscr{P}(\mathbb{L})$ we have

$$f(m) = \tilde{\Phi}(f(m \circ \tilde{\Phi})).$$

Note that for $m : \mathcal{F} \longrightarrow [0,1]$, by abuse of notation with $(m \circ \tilde{\Phi})$ we mean the mapping $(m \circ \tilde{\Phi}) : \Phi^{-1}(\mathcal{F}) \longrightarrow [0,1] : A \mapsto m(\tilde{\Phi}(A))$ with restricted domain $\Phi^{-1}(\mathcal{F})$.

Analogously, a mapping $f : \mathscr{P}(\mathbb{L}) \longrightarrow \mathbb{L}$ is called representation invariant if the associated mapping $\tilde{f} : \mathscr{P}(\mathbb{L}) \longrightarrow 2^{\mathbb{L}} : A \mapsto \{f(A)\}$ is representation invariant.

With this definition, we can show that the $\bigwedge$-operation, the $\bigvee$-operation and the mapping that chooses all elements in the preimage $B_m^{-1}(S)$ for some fixed set

[5] For real-valued random variables there always exists exactly one such minimal element, namely $\bigwedge B_m^{-1}(\alpha)$, because the underlying probability measure m and thus also m_* is continuous from above (see, e.g., [Denneberg, 1994, p. 22]) and $\bigwedge B_m^{-1}(\alpha)$ can be obtained as the infimum $\bigwedge_{n=1}^{\infty} q_i$ of a decreasing sequence $(q_n)_{n \in \mathbb{N}}$ in $B_m^{-1}(\alpha)$ and thus we have $B_m(\bigwedge B_m^{-1}(\alpha)) = m_*(\downarrow \bigwedge B_m^{-1}(\alpha)) = m_*(\downarrow \bigwedge_{n=1}^{\infty} q_i) = m_*(\bigcap_{n=1}^{\infty} \downarrow q_n) = \lim_{n\to\infty} m_*(\downarrow q_n) = \alpha$.

[6] Actually, due to reasons of symmetry, a true choice from the $B_m^{-1}(\alpha)$ is not possible without violating the non-arbitrariness demand.

$S \subseteq [0,1]$ are representation invariant:

Lemma 1. Let $(\mathbb{L}, m, \mathcal{F}, B_m)$ be a belief structure and let $S \subseteq [0,1]$. The mappings

$$\begin{aligned} f :& 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}} : A \mapsto \{\bigwedge A\}, \\ f' :& 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}} : A \mapsto \{\bigvee A\}, \\ g :& \mathscr{P}(\mathbb{L}) \longrightarrow 2^{\mathbb{L}} : m \mapsto B_m^{-1}(S) \end{aligned}$$

are representation invariant. Furthermore, for $f_1 : \mathscr{P}(\mathbb{L}) \longrightarrow 2^{\mathbb{L}}$ and $f_2 : 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}}$ the composition $f_2 \circ f_1$ is representation invariant if f_2 and f_1 are representation invariant.

Proof: For every order-automorphism Φ on $\mathbb{L}$ and for every A in $2^{\mathbb{L}}$ we have $f(A) = \{\bigwedge A\} = \tilde{\Phi}(\Phi^{-1}(\bigwedge A)) = \tilde{\Phi}(\{\bigwedge \Phi^{-1}(A)\}) = \tilde{\Phi}(f(\Phi^{-1}(A))$. Analogously one proofs that f' is representation invariant.

For $m \in \mathscr{P}(\mathbb{L})$ and $S \subseteq [0,1]$ we have

$$\begin{aligned} g(m) &= B_m^{-1}(S) \\ &= \{x \in \mathbb{L} \mid B_m(x) \in S\} \\ &= \{\Phi(y) \mid y \in \mathbb{L}, m_*(\downarrow \Phi(y)) \in S\} \\ &= \{\Phi(y) \mid y \in \mathbb{L}, m_*(\tilde{\Phi}(\downarrow y)) \in S\} \\ &= \{\Phi(y) \mid y \in \mathbb{L}, (m_* \circ \tilde{\Phi})(\downarrow y) \in S\} \\ &\overset{*}{=} \{\Phi(y) \mid y \in \mathbb{L}, (m \circ \tilde{\Phi})_*(\downarrow y) \in S\} \\ &= \tilde{\Phi}(g(m \circ \tilde{\Phi})). \end{aligned}$$

The equality $*$ is valid because for arbitrary $A \in 2^{\mathbb{L}}$ we have

$$\begin{aligned} (m \circ \tilde{\Phi})_*(A) &= \sup_{\substack{B \in \Phi^{-1}(\mathcal{F}), \\ B \subseteq A}} (m \circ \tilde{\Phi})(B) \\ &= \sup_{\substack{B \in \Phi^{-1}(\mathcal{F}), \\ B \subseteq A}} m(\tilde{\Phi}(B)) \\ &= \sup_{\substack{B \in \mathcal{F}, \\ B \subseteq \tilde{\Phi}(A)}} m(B) \\ &= m_*(\tilde{\Phi}(A)) = (m_* \circ \tilde{\Phi})(A). \end{aligned}$$

Finally, for representation invariant maps $f_1 : \mathscr{P}(\mathbb{L}) \longrightarrow 2^{\mathbb{L}}$ and $f_2 : 2^{\mathbb{L}} \longrightarrow 2^{\mathbb{L}}$ we have

$$\begin{aligned}(f_2 \circ f_1)(m) &= f_2(f_1(m)) \\ &= \tilde{\Phi}(f_2(\Phi^{-1}(f_1(m)))) \\ &= \tilde{\Phi}(f_2(\Phi^{-1}(\tilde{\Phi}(f_1(m \circ \tilde{\Phi})) \\ &= \tilde{\Phi}(f_2(f_1(m \circ \tilde{\Phi}))) \\ &= \tilde{\Phi}((f_2 \circ f_1)(m \circ \tilde{\Phi})),\end{aligned}$$

which shows that in this case $f_2 \circ f_1$ is representation invariant, too. ∎

With this result it is possible to choose a set of candidates for a quantile of given level α for example as $B_m^{-1}(\alpha)$ (or alternatively as the set $\min B_m^{-1}(\alpha)$ of minimal elements in $B_m^{-1}(\alpha)$) if this set is not empty. But if one wants only one quantile for a given level α, the choice of one element of the set $B_m^{-1}(\alpha)$ (or $\min B_m^{-1}(\alpha)$) can still be arbitrary to some extent. For example if one chooses that $q \in B_m^{-1}(\alpha)$ with the smallest value $m_*(\uparrow q)$ then (if there is only one such element q, which is generally not the case) the choice is unique and representation invariant, but this choice seems somehow arbitrary: If the quantile q is intended to be understood as some kind of "representative" element under which a proportion α of all the probability mass lies and that q 'represents' this fraction of probability mass then there is no reason to choose that element for which additionally there lies the least probability mass above.

Another way to overcome the non-uniqueness issue is to not choose among the elements in $B_m^{-1}(\alpha)$ but to 'aggregate' all elements as $q = \bigvee B_m^{-1}(\alpha)$ or $\tilde{q} = \bigwedge B_m^{-1}(\alpha)$ to obtain a unique element q or $\tilde{q}$ but with the problem that q and $\tilde{q}$ are generally not elements of $B_m(\alpha)$ anymore, so with this construction one 'looses the α-level', but if one declares $B_m(q)$ (or $B_m(\tilde{q})$) as the new, "corrected" level β, then one has still unique elements q for some $\beta \in \mathrm{Im}(B_m)$. Since for q we still have $q \in B_m^{-1}(\uparrow \alpha)$ the proposal q still holds the original level α in a conservative sense and we therefore start the analysis of further properties with the proposal q. (But as it will turn out later, the proposal $\tilde{q}$ seems to be more natural, note also that for real-valued random variables the proposal $\tilde{q}$ (and not the proposal q) corresponds to the above definition of lower quantiles for this special situation.) In the sequel, we will call the level α we started with the prelevel and we will call the final level β the actual level or shortly the level.

3 Construction of prequantiles

Definition 4. Let $(\mathbb{L}, m, \mathcal{F}, B_m)$ be a belief structure. Every element $q \in \mathbb{L}$ with $B_m(q) = \alpha$ is called a (lower) α-prequantile. A (lower) prequantile q that is minimal in $B_m^{-1}(\alpha)$ is called a (lower) α-quantile.

Remark 2. In the sequel, for simplicity, we will call lower (pre)quantiles simply (pre)quantiles.

Now we can investigate the properties of different proposals for the construction of prequantiles. Later, we will derive quantiles from this prequantiles that particularly satisfy all properties $i) - vi)$ at least for lattices that are linearly order co-Lindelöf. We say here that a complete lattice $\mathbb{L}$ is linearly order co-Lindelöf if for every chain T in $\mathbb{L}$ there exists a countable subcain $S \subseteq T$ with $\bigwedge S = \bigwedge T$. Generally, it seems that one cannot guarantee property $vii)$ "richness", the richness of the final system of quantiles is very dependent on the data situation for our construction.

Construction A

Let $(\mathbb{L}, m, \mathcal{F}, B_m)$ be a belief structure. For $\alpha \in \mathrm{Im}(B_m)$ define $\beta(\alpha) := B_m(\bigvee B_m^{-1}(\alpha))$ and

$$q_{\beta(\alpha)} := \bigvee B_m^{-1}(\alpha)$$

as a prequantile with level β. This simple construction leads to a (representation invariant) set

$$\mathfrak{Q}_A := \{q_{\beta(\alpha)} \mid \alpha \in \mathrm{Im}(B_m)\}$$

of prequantiles. But the set $\mathfrak{Q}_A$ generally does not satisfy properties $iv) - vi)$. To see this, look at Figure 1. This figure shows a simple complete lattice given by its Hasse graph. The probability measure m is indicated by the numbers that are written at the elements. The probability for one element of the lattice is proportional to the given numbers. To make the graphs more readable, we have omitted values that are one. Additionally, we also indicated with a label q, which elements are quantiles. A multiple of the actual level is given in the index at the label q and a multiple of the prelevel is given in brackets. The constants we multiplied all the numbers is simply the smallest integer such that one has only integers and no fraction in the figure. This has only been made to make the figure better readable. Here, for example for $\alpha = 2/8$ and $\alpha = 3/8$ we have $B_m^{-1}(2/8) = \{a, b\}$ and $B_m^{-1}(3/8) = \{c, d\}$ and thus the prequantiles e and f induced by the prelevels $\alpha = 2/8$ and $\alpha = 3/8$ are two incomparable prequantiles of the same level $\beta = 5/8$. This shows that properties $iv)$ and $v)$ are not satisfied. Furthermore, the greatest element $\top$ is a 1-prequantile that is not minimal because the element g is also a 1-prequantile that is lower than $\top$.

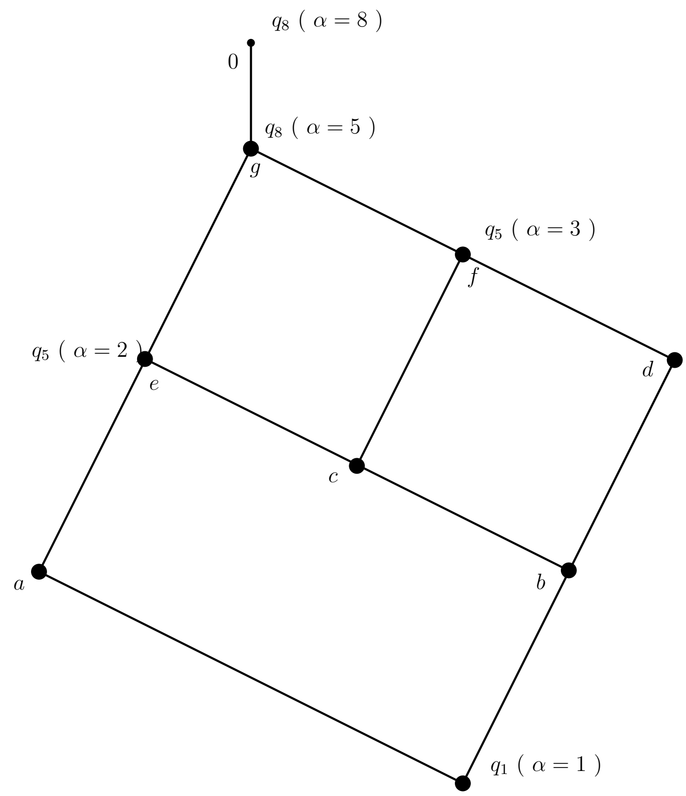

Figure 1: Construction of lower quantiles A.

Construction A'

Let $(\mathbb{L}, m, \mathcal{F}, B_m)$ be a belief structure. For $\alpha \in \mathrm{Im}(B_m)$ define $\beta(\alpha) := B_m(\bigvee\{q \mid q \text{ minimal in } B_m^{-1}(\alpha))\}$ and

$$q_{\beta(\alpha)} := \bigvee\{q \mid q \text{ minimal in } B_m^{-1}(\alpha)\}$$

as a prequantile with level β and

$$\mathfrak{Q}_{A'} := \{q_{\beta(\alpha)} \mid \alpha \in \mathrm{Im}(B_m)\}$$

as the set of all corresponding prequantiles. This construction seems to overcome the non-minimality disadvantage of construction A (at least for lattices that are linearly order co-Lindelöf) but it still has the same other disadvantages of construction A, namely in not satisfying property iv) uniqueness and v) comparability, which can also be seen in Figure 1. Here, the only difference between construction A and A' is that for the construction A' the top element is not a prequantile in $\mathfrak{Q}_{A'}$.

Construction B

This construction looks at the order structure of the underlying lattice $\mathbb{L}$ as obtained as the intersection of linear orders. We consider the set

$$\mathrm{lin}(\mathbb{L}) := \{(L, R) \mid R \supseteq \leq \quad \& \quad R \text{ linear order on } L\}$$

of all linear orders that extend the order $\leq$ of the given lattice $\mathbb{L} = (L, \leq)$. Then, for $\alpha \in \mathrm{Im}(B_m)$ and $(L, R) \in \mathrm{lin}(\mathbb{L})$ look at the induced belief structure $((L, R), m, \mathcal{F}, B_m^R)$ where $B_m^R(x) = m_*(\{y \mid yRx\})$. Take q_α^R as the minimal element (w.r.t. R) in the set $(B_m^R)^{-1}(\uparrow \alpha)$ (Assume that there exists one such minimal element and note that there is always at most one such element because R is a linear order.) Finally define $\beta(\alpha) := B_m(\bigvee\{q_\alpha^R \mid (L, R) \in \mathrm{lin}(\mathbb{L})\})$ and as a β-prequantile the element

$$q_{\beta(\alpha)} := \bigvee\{q_\alpha^R \mid (L, R) \in \mathrm{lin}(\mathbb{L})\}$$

. If such a construction is possible for every $\alpha \in \mathrm{Im}(B_m)$, the obtained set

$$\mathfrak{Q}_B := \{q_{\beta(\alpha)} \mid \alpha \in \mathrm{Im}(B_m)\}$$

satisfies the properties $i) - iii)$. If we restrict the focus on empirical[7] belief structures, then also property vi) seems to be fulfilled. If the empirical belief structure

[7] With an empirical belief structure we mean a belief structure where the underlying probability measure m is an empirical measure induced by a data sample.

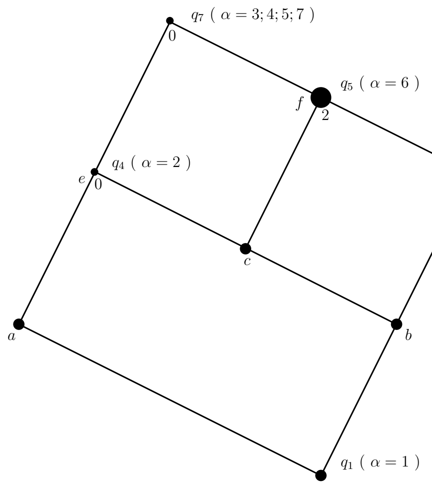

Figure 2: Construction of lower quantiles B.

is furthermore non-degenerate (meaning that the data sample inducing the probability measure m does not does not have ties), then also properties *iv*) and *v*) seem to be satisfied.

Remark 3. The assumption that the empirical belief structure is non-degenerate is actually needed as one can see in Figure 2. The prequantiles q_4 and q_5 are in fact incomparable. Note that this problem can be circumvented if one replaces every element with mass k/n by a chain of k elements with mass $1/n$ respectively.

Construction C

In the Construction B the supremum in the expression $q_{\beta(\alpha)} := \bigvee\{q_\alpha^R \mid (L,R) \in \text{lin}(\mathbb{L})\}$ is the supremum in $\mathbb{L}$. If one takes into account that the supremum of a set can be represented as the infimum of the upper bounds of this set, then there is a possibility to modify construction B a little bit by considering the infimum of all elements that are greater than or equal to every q_α^R, but not w.r.t. the order $\leq$ but w.r.t. the order R. Then one gets

$$q_{\beta(\alpha)} = \bigwedge\{q \mid \forall (L,R) \in \text{lin}(\mathbb{L}) : q_\alpha^R R q\}. \tag{1}$$

Furthermore we have the equivalence

$$q_\alpha^R R q \iff B_m^R(q) \geq \alpha \iff q \in \left(B_m^R\right)^{-1}(\uparrow \alpha)$$

and because of

$$B_m(q) \geq \alpha \iff \forall R \in \text{lin}(\mathbb{L}) : B_m^R(q) \geq \alpha$$

the above Construction is equivalent to the more simple definition

$$q_{\beta(\alpha)} := \bigwedge\{q \mid q \in B_m^{-1}(\uparrow \alpha)\} \tag{2}$$

with $\beta(\alpha) := B_m(\bigwedge\{q \mid q \in B_m^{-1}(\uparrow \alpha)\})$. Thus, construction C leading to the set

$$\mathfrak{Q}_C := \{q_{\alpha(\beta)} \mid \alpha \in \mathrm{Im}(B_m)\}$$

of prequantiles can be motivated by two different views on the relation $\leq$: Firstly, the relation $\leq$ can be understood as induced by all its linear extensions and secondly, the relation $\leq$ can be understood as the primitive notion, and for both understandings, the naturally arising constructions (1) and (2) lead to the same result. Note also that for this construction $B_m^{-1}(\alpha)$ needs not to have minimal elements.

Lemma 2. The set $\mathfrak{Q}_C$ of quantiles from construction C satisfies the properties $i) - iii)$ and property $v)$.

Proof: Property $i)$ is valid because of Lemma 1. Property $ii)$ is satisfied in the sense that only the property of being in $B_m^{-1}(\alpha)$ and the infimum operation in $\mathbb{L}$ is used. For $\alpha \leq \alpha'$ we have $\{q \mid q \in B_m^{-1}(\uparrow \alpha)\} \supseteq \{q \mid q \in B_m^{-1}(\uparrow \alpha')\}$ and thus $\bigwedge\{q \mid q \in B_m^{-1}(\uparrow \alpha)\} \leq \bigwedge\{q \mid q \in B_m^{-1}(\uparrow \alpha')\}$, so $q_{\beta(\alpha)} \leq q_{\beta(\alpha')}$ (and also $\beta(\alpha) \leq \beta(\alpha')$). This shows that property $v)$ is satisfied.

■

Remark 4. The quantiles in $\mathfrak{Q}_C$ need not to satisfy properties $iv)$ and $vi)$: In Figure 3 for $\alpha = 2/4$ and $\alpha = 3/4$ one gets the different prequantiles c and b with the same level $2/4$. Furthermore, the prequantile c is not minimal, because the prequantile b (with the same level $2/4$) is lower than c.

4 Derivation of quantiles from prequantiles

To finally construct quantiles that additionally satisfy properties $iv)$ and $vi)$ we use a result that is known to the author by Stefan E. Schmidt (personal communication, c.f. also Kwuida and Schmidt [2011]):

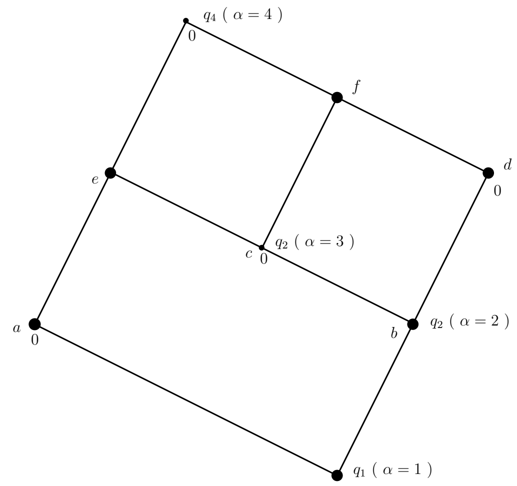

Figure 3: Construction of lower quantiles C.

Theorem 1. Let $f : \mathbb{L} \longrightarrow \mathbb{M}$ be an isotone mapping from a complete lattice $\mathbb{L}$ into a partially ordered, quasi cancellative monoid $\mathbb{M}$.[8]

1. If f is supermodular[9] then $\mathcal{K}_f := \{x \in \mathbb{L} \mid x \text{ is minimal in } f^{-1}(f(x))\}$ is a kernel system[10] on L.
2. If f is submodular[11] then $\mathcal{H}_f := \{x \in \mathbb{L} \mid x \text{ is maximal in } f^{-1}(f(x))\}$ is a closure system[12] on L.

Proof: We show only the first part. The second part is analogously provable. Denote with $\perp$ the smallest element of the lattice $\mathbb{L}$. Then firstly $\perp \in \mathcal{K}_f$ since $x \geq \perp$ for all $x \in \mathbb{L}$. Let now I be an arbitrary (non-empty) index set and $x_i \in \mathcal{K}_f$ for all $i \in I$. If $z := \bigvee_{i \in I} x_i$ is not minimal in $f^{-1}(f(z))$ then we would have some $y < z$ with $f(y) = f(z)$ and some x_i with $x_i \nleq y$ and thus $x_i \wedge y < x_i$.

[8]A monoid $\mathbb{M} = (M, +, e)$ consists of a set M that is equipped with an operation $+ : M \times M \longrightarrow M$ that is associative and has a neutral element e. A partially ordered monoid is a monoid $(M, +, e)$ that is additionally equipped with an ordering $\leq$ on M that is compatible with the operation $+$ in the sense that we have $x \leq y \Longrightarrow x + z \leq y + z \;\&\; z + x \leq z + y$ for all $x, y, z \in M$. A partially orderd monid is called quasi cancellative if we have $x + z \leq y + z \Longrightarrow x \leq y$ and $z + x \leq z + y \Longrightarrow x \leq y$ for all $x, y, z \in M$.

[9]A mapping $f : \mathbb{L} \longrightarrow \mathbb{L}$ is called supermodular if $f(x) + f(y) \leq f(x \wedge y) + f(x \vee y)$ for all $x, y \in \mathbb{L}$.

[10]A kernel system on a lattice $\mathbb{L}$ is a subset of $\mathbb{L}$ that contains the smallest element $\perp$ and that is furthermore closed under arbitrary suprema.

[11]A mapping $f : \mathbb{L} \longrightarrow \mathbb{L}$ is called submodular if $f(x) + f(y) \geq f(x \wedge y) + f(x \vee y)$ for all $x, y \in \mathbb{L}$.

[12]A closure system on a lattice $\mathbb{L}$ is a subset of $\mathbb{L}$ that contains the greatest element $\top$ and that is furthermore closed under arbitrary infima.

With the supermodularity of f and $x_i \vee y \leq z$ we get

$$\begin{aligned} f(x_i) + f(z) &= f(x_i) + f(y) \\ &\leq f(x_i \wedge y) + f(x_i \vee y) \\ &\leq f(x_i \wedge y) + f(z) \end{aligned}$$

which shows $f(x_i) \leq f(x_i \wedge y)$ and thus $f(x_i) = f(x_i \wedge y)$, which is a contradiction to the minimality of x_i in $f^{-1}(f(x_i))$. ■

Corollary 1. Let $(\mathbb{L}, \mathcal{F}, m, B_m)$ be a belief structure. The set

$$\mathfrak{Q} := \{x \in \mathbb{L} \mid x \text{ is minimal in } B_m^{-1}(B_m(x))\}$$

is a kernel system.

Definition 5. If for a mapping $f : \mathbb{L} \longrightarrow \mathbb{M}$ the system $\mathcal{K}_f := \{x \in \mathbb{L} \mid x \text{ is minimal in } f^{-1}(f(x))\}$ is a kernel system, then the corresponding kernel operator $k_f : \mathbb{L} \longrightarrow \mathbb{L}$ is defined as

$$k_f(x) = \bigvee_{y \in \mathcal{K}_f, y \leq x} y.$$

Analogously if $\mathcal{H}_f := \{x \in \mathbb{L} \mid x \text{ is maximal in } f^{-1}(f(x))\}$ is a closure system, then we denote with h_f the corresponding closure operator defined by

$$h_f(x) = \bigwedge_{y \in \mathcal{H}_f, y \geq x} y.$$

Remark 5. Note that generally $f^{-1}(f(x))$ needs not to have minimal elements at all.

Lemma 3. Let $(\mathbb{L}, \mathcal{F}, m, B_m)$ be a belief structure on a lattice $\mathbb{L}$. Then if furthermore $\mathbb{L}$ is linearly order co-Lindelöf (meaning that for every chain T in $\mathbb{L}$ there exists a countable subchain $S \subseteq T$ with $\bigwedge S = \bigwedge T$) we have

$$\forall x \in \mathbb{L} : \quad B_m(k_{B_m}(x)) = B_m(x)$$

and thus all ${B_m}^{-1}(B_m(x))$ have minimal elements.

Proof: For $x \in \mathbb{L}$ let $c := B_m(x)$ and $M = \{y \in \mathbb{L} \mid B_m(y) = c\}$. We show that all chains in M have a lower bound in M and thus because of Zorn's lemma every $a \in M$ lies above a minimal element of M. Then from $x \in M$ it follows $B_m(k_{B_m}(x)) = B_m(x)$. To show that all chains in M have a lower bound in M take some arbitrary chain $C \subseteq M$. Then the lower bound $\bigwedge C$ is in M because there is a countable subchain

$S = \{s_1 \geq s_2 \geq \ldots\} \subseteq C$ with $\bigwedge S = \bigwedge C$ and because the inner measure m_* is continuous from above we have $B_m(\bigwedge C) = B_m(\bigwedge S) = m_* \left(\downarrow \bigwedge_{i=1}^{\infty} s_i \right) = m_* \left(\bigcap_{i=1}^{\infty} \downarrow s_i \right) = \lim_{n\to\infty} m_* (\downarrow s_i) = c$. ∎

The next example shows that one generally cannot drop the co-Lindelöf assumption:

Example 1. $\mathbb{L} = (2^{[0,1]}, \subseteq)$ is not linearly order co-Lindelöf: Take the set of all co-countable subsets of $[0,1]$ and choose a maximal chain T of this set (this is possible because of Zorn's lemma). Then $\bigwedge T = \emptyset$ because if there was an element $x \in \bigwedge T$, the chain $T \cup \{\bigwedge T \backslash \{x\}\}$ would be a strict superchain of T which is in contradiction with the maximality of T. For a countable subchain $S \subseteq T$ we have $(\bigwedge S)^c = \bigvee \{s^c \mid s \in S\}$ which is a countable union of countable sets, thus countable. Therefore $\bigwedge S$ must be uncountable and thus nonempty. This shows that there does not exist a countable subchain S with $\bigwedge S = \bigwedge T$.
Furthermore, there exists a probability measure m on $\mathbb{L}$ such that for example $B_m^{-1}(1)$ has no minimal elements: First define for $A \in 2^{\mathbb{L}}$ the set $S_A := \bigcup \{B \in A \mid \#B = 1\}$ of all singletons of A. Then define $m : \{A \in 2^{\mathbb{L}} \mid S_A \in \mathcal{B}([0,1])\} \longrightarrow [0,1] : A \mapsto \lambda(S_A)$, where λ is the Lebesgue measure and $\mathcal{B}[(0,1)]$ is the Borel σ-algebra on $[0,1]$.

It is clear that every set $T \in B_m^{-1}(1)$ has uncountable many singletons and for an arbitrary singleton $t \in T$ we have $T\backslash\{t\} \in B_m^{-1}(1)$ which means that there cannot be minimal elements in $B_m^{-1}(1)$.

Lemma 4. Let $f : \mathscr{P}(\mathbb{L}) \longrightarrow \mathbb{L}$ be a representation invariant map. Then the mapping

$$f_k : \mathscr{P}_n(\mathbb{L}) \longrightarrow \mathbb{L} : m \mapsto k_{B_m}(f(m)) = \bigvee \{a \mid a \in \mathcal{K}_{B_m}, a \leq f(m)\}$$

is a representation invariant, lower quantile-valued mapping.

Proof: Let $m \in \mathscr{P}(\mathbb{L})$ be given. First we mention that $\mathcal{K}_{B_{m\circ\tilde{\Phi}}} = \Phi^{-1}(\mathcal{K}_{B_m})$. With the representation invariance of f we get

$$\begin{aligned}
\Phi(f_k(m \circ \tilde{\Phi})) &= \Phi\left(\bigvee \{x \mid x \in \mathcal{K}_{B_{m\circ\tilde{\Phi}}}, x \leq f(m \circ \tilde{\Phi})\}\right) \\
&= \Phi\left(\bigvee \{x \mid x \in \Phi^{-1}(\mathcal{K}_{B_m}), x \leq \Phi^{-1}(f(m))\}\right) \\
&= \bigvee \{\Phi(x) \mid x \in \Phi^{-1}(\mathcal{K}_{B_m}), x \leq \Phi^{-1}(f(m))\} \\
&= \bigvee \{x \mid x \in \mathcal{K}_{B_m}, x \leq f(m)\} \\
&= f_k(m).
\end{aligned}$$

This proves the representation invariance of f_k. The fact that $f_k(m)$ is a lower quantile follows from Corollary 1. ∎

5 Quantitative and qualitative data analysis

From now on we focus on the set $\mathfrak{Q}_C$ of construction C and recall that the set of quantiles obtained by construction C and the application of the kernel operator k_{B_m} of Section 4 is given as

$$\mathfrak{Q} := \{k_{B_m}\left(\bigwedge B_m^{-1}(\uparrow \alpha)\right) \mid \alpha \in \mathrm{Im}(B_m)\}.$$

Define furthermore the lower order-completion of $\mathfrak{Q}$ as $\overline{\mathfrak{Q}} = \{\bigwedge M \mid M \subseteq \mathfrak{Q}\}$. The elements of $\overline{\mathfrak{Q}}$ are called quasi-quantiles, here. Note that $\overline{\mathfrak{Q}}$ is still a chain. With this set of quantiles or quasi-quantiles one can do descriptive data analysis for lattice-valued data. Since we have not only the quantile system, but also the underlying belief structure, we can introduce both a qualitative and a quantitative "measurement mapping" for data analysis:

Definition 6. Let $(\mathbb{L}, m, \mathcal{F}, B_m)$ be a belief structure. The mapping

$$\Phi_m : \mathbb{L} \longrightarrow \mathbb{L} : x \mapsto \bigwedge\{q \in \mathfrak{Q} \mid q \geq x\}$$

is called the quantile mapping. It maps every element x to the smallest quasi-quantile $q \in \overline{\mathfrak{Q}}$ that is still greater than or equal to x. Note that $\Phi(x)$ does not need to be a quantile, but at least it is an infimum of quantiles, i.e., a quasi-quantile. The quantile mapping can be used for qualitative data analysis in the sense that every x is mapped to that quasi-quantile q that is in some sense that representative element of the easier to understand chain $\overline{\mathfrak{Q}}$ that is closest to x and still lies above x.

The mapping

$$\lambda_m : \mathbb{L} \longrightarrow [0, 1] : x \mapsto B_m(\Phi_m(x))$$

is called the level function. The level function can be understood as a quantitative "measurement" mapping that quantifies for a given $x \in \mathbb{L}$, how high the smallest quasi-quantile q representing x (in the sense of still lying above x) is in terms of the amount of probability mass lying below q.

Theorem 2. Let $(\mathbb{L}, \mathcal{F}, m, B_m)$ be a belief structure and let X be a $\mathbb{L}$-valued random variable that is distributed according to the image measure m. If $\mathbb{L}$ is linearly order Lindelöf (meaning that for every chain T in $\mathbb{L}$ there exists a countable subchain $S \subseteq T$ with $\bigvee S = \bigvee T$) and if every quasiquantile is measurable (meaning that $\forall q \in \overline{\mathfrak{Q}} : \downarrow q \in \mathcal{F}$) then the level function λ_m has the following pivot property:

$$\forall z \in \mathbb{L} : m(\{x \in \mathbb{L} \mid \lambda_m(x) \leq \lambda_m(z)\}) = \lambda_m(z).$$

Proof: Define $\alpha := \lambda_m(z)$ and $w := \bigvee T$ with $T := \{\Phi_m(x) \mid \lambda_m(x) \leq \alpha\}$. Then, because all quasiquantiles are measurable and because $w = \bigvee S$ for some countable subchain $S \subseteq T$, we have $\lambda_m(w) \leq \alpha$ because m is continuous from below. Because $\Phi_m(z) \in T$ we actually have $\lambda_m(w) = \alpha$. Now, we proof that

$$\lambda_m(x) \leq \alpha \iff x \leq w :$$

if: From $x \leq w$ it immediately follows $\lambda_m(x) \leq \lambda_m(w) = \alpha$.
only if: Let $\lambda_m(x) \leq \alpha$. Then $\Phi(x) \in T$ and thus $x \leq \Phi(x) \leq w$. Finally note that $w = \bigvee T = \bigvee S$ is also measurable and we have

$$\alpha = \lambda_m(w) = m(\downarrow w) = m(\{x \in \mathbb{L} \mid \lambda_m(x) \leq \alpha\}).$$

■

6 Conclusion

In this paper, we introduced a notion of quantiles for data sets or random variables with values in a complete lattice. We showed that acceptable properties for an asymmetric notion of lower quantiles can be reached. The next step would be to apply this notion of lower quantiles. Actually, there is a broad range of applications thinkable, here, because one only needs the structure of a complete lattice. The presentation of the theoretical results in this paper was of course very technical and the paper is mainly only devoted to showing that a very simple idea of lower quantiles actually works in some acceptable way. To give a short example of application, and to come back to the the introduction that mentioned multivariate quantile concepts in Euclidean d-space, consider the complete lattice of all convex sets in $\mathbb{R}^d$, ordered by set inclusion. Then, if one identifies points in $x \in \mathbb{R}^d$ with singletons $\{x\}$, the quantile notion developed herein, especially the level function turns out to be essentially a monotone transformation of Tukey's half-space depth, with the nice additional feature of having the pivot property defined in Theorem 2. Additionally, since one works in the space of all convex sets of $\mathbb{R}^d$ one cannot only deal with points as singletons, but one can also deal with convex set-valued data or random variables.

Interesting areas of applications are all sorts of application, where one has the structure of a complete lattice, but not much more structure like that of Euclidean d-space. One natural example is the case of ranking data, where one has a data set of (maybe only partial) orderings on a basic set C and the space of all orderings on C, treated as relations (i.e., subsets of $C \times C$) is naturally equipped with the set inclusion as an underlying ordering. Compared to this, in classical statistics one usually firstly has to introduce some metric in the space (which may seem a little bit unnatural) to do a quantitative data analysis based on methods for data in a metric space.

References

S. Alkire, J. Foster, S. Seth, M. E. Santos, J. M. Roche, and P. Ballon. *Multidimensional Poverty Measurement and Analysis.* Oxford University Press, 2015.

M. Cardin. A quantile approach to integration with respect to non-additive measures. In V. Torra, Y. Narukawa, B. López, and M. Villaret, editors, *International Conference on Modeling Decisions for Artificial Intelligence*, pages 139–148. Springer, 2012.

P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434):862–872, 1996.

V. Chernozhukov, A. Galichon, M. Hallin, and M. Henry. Monge–Kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics*, 45(1):223–256, 2017.

D. Denneberg. *Non-additive Measure and Integral.* Kluwer Academic Publishers, 1994.

J. Duckitt, B. Bizumic, S. W. Krauss, and E. Heled. A tripartite approach to right-wing authoritarianism: The authoritarianism-conservatism-traditionalism model. *Political Psychology*, 31(5):685–715, 2010.

M. Grabisch. Belief functions on lattices. *International Journal of Intelligent Systems*, 24(1):76–95, 2009.

G. Koshevoy and K. Mosler. Zonoid trimming for multivariate distributions. *The Annals of Statistics*, 25(5):1998–2017, 1997.

L. Kwuida and S. E. Schmidt. Valuations and closure operators on finite lattices. *Discrete Applied Mathematics*, 159(10):990–1001, 2011.

R. Y. Liu et al. On a notion of data depth based on random simplices. *The Annals of Statistics*, 18(1):405–414, 1990.

K. Mosler. Depth statistics. In C. Becker, R. Fried, and S. Kuhnt, editors, *Robustness and Complex Data Structures: Festschrift in Honour of Ursula Gather*, pages 17–34. Springer, 2013.

H. Oja. Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, 1(6):327–332, 1983.

G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

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

C. Zhou. Belief functions on distributive lattices. *Artificial Intelligence*, 201:1 – 31, 2013.

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