Nonparametric Multivariate Descriptive Measures Based on Spatial Quantiles

Robert Serfling[1]
University of Texas at Dallas

March 2003

Final version for *Journal of Statistical Planning and Inference*, to appear 2003

[1]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA. Email: serfling@utdallas.edu. Website: www.utdallas.edu/∼serfling. Support under National Science Foundation Grant DMS-0103698 is gratefully acknowledged.

Abstract

An appealing way of working with probability distributions, especially in nonparametric inference, is through "descriptive measures" that characterize features of particular interest. One attractive approach is to base the measures on quantiles. Here we consider the multivariate context and utilize the "spatial quantiles", a recent vector extension of univariate quantiles that is becoming increasingly popular. In terms of these quantiles, we introduce and study nonparametric measures of multivariate location, spread, skewness and kurtosis. In particular, we define a useful "location" functional which augments the well-known "spatial" median and a "volume" functional which plotted as a "spatial scale curve" yields a convenient two-dimensional characterization of the spread of a multivariate distribution of any dimension. These spatial location and volume functionals also play roles in the formulation of "spatial" skewness and kurtosis functionals which reduce to known versions in the univariate case. We also define corresponding spatial "asymmetry" and "kurtosis" curves which are new devices even in the univariate case. Tailweight and peakedness measures, as distinct from kurtosis, are also discussed. To aid better understanding of the spatial quantiles as a foundation for nonparametric multivariate inference and analysis, we also provide some basic perspective on them: their interpretations, properties, strengths and weaknesses.

AMS 1991 Subject Classification: Primary 62H05 Secondary 62G20.

Key words and phrases: Nonparametric; Multivariate distributions; Spatial quantiles; Location; Scale; Skewness; Kurtosis; Peakedness; Tailweight.

1 Introduction

An intuitively appealing way of working with probability distributions is through "descriptive measures" that characterize features of particular interest. This approach is especially useful when the distributions are otherwise rather unspecified, as in exploratory and nonparametric inference. In the multivariate case, which is the setting of this paper, the formulation of such measures involves interesting conceptual and technical challenges.

In the univariate case, many notions of descriptive measures are *quantile-based*, exploiting the natural order of the real line. For extension to the multivariate case, one must select a particular version of multivariate quantiles (see Serfling, 2002b, for a partial review of various *ad hoc* notions). A compelling choice – which we adopt here – is the *spatial quantiles*, which were introduced by Chaudhuri (1996) and Koltchinskii (1997) as a certain form of generalization of the univariate case based on the L_1 norm. The spatial quantiles have induced a variety of useful new nonparametric multivariate methods and are receiving increasing interest. To go beyond these developments, it is timely to carry out a study of basic descriptive measures defined in terms of the spatial quantiles. In this spirit, and utilizing the spatial quantiles as a basis, the present paper treats measures of *location*, *spread*, *skewness* and *kurtosis*.

Location in this sense already has a representative measure: the well-known *spatial median*, which dates back at least to Hayford (1902) and has been much studied. Indeed, it has enjoyed considerable success among competing notions of multivariate location (see Small, 1990, for an overview of multidimensional medians). Almost as important as location is spread, for which spatial versions are already noted in Chaudhury (1996). Here we provide further development of these two measures. While location and spread occupy central roles and have the broadest application, next most important are skewness and kurtosis, which serve, for example, to characterize the way in which a distribution deviates from normality. Building on our treatment of spatial location and spread, we formulate spatial measures of skewness and kurtosis.

Besides supporting the conceptual foundations for nonparametric inference about multivariate populations, our study of "spatial" descriptive measures and their basic properties also provides a foundation for development of interesting new methodological tools, for example diagnostics. These aspects are beyond the present scope, however, and will be explored elsewhere. A number of issues and directions for further investigation are opened up by our treatment, some of these left implicit and some noted in Remarks 3.1 and Section 3.3. Although the focus here is on the multivariate setting, some of the specializations of our formulations to the univariate case are novel as well and merit further investigation.

For background on nonparametric descriptive measures in general, we cite for the univariate case a landmark treatment of location and spread by Bickel and Lehmann (1975a,b, 1976, 1979) and a unified development also encompassing skewness and kurtosis by Oja (1981). Diverse extensions to the multivariate case have been proposed (see Oja, 1983, for general contributions and discussion, and Kotz, Balakrishnan and Johnson, 2000, for an overview on skewness and kurtosis). In particular, some of the multivariate approaches of the present paper parallel those of Liu, Parelius and Singh (1999) concerning the use of central regions based on depth functions and those of Avérous and Meste (1990, 1997a,b) concerning the construction of functionals for location and skewness. Further background including other recent developments will be referenced in the sequel.

In making application of the spatial quantiles, it is important to fully understand their features. We provide, therefore, a carefully drawn perspective on the spatial quantiles: their interpretations, properties, strengths and weaknesses. This is presented in Section 2, where we also treat location and spread, defining a useful spatial "location functional" augmenting the spatial median, and defining a spatial "volume" functional given by the volume of spatial "central regions" of increasing size. One role of the volume functional is to provide, through a plot as a "spatial scale curve", a convenient two-dimensional characterization of the spread of a multivariate distribution of any dimension. This is the "spatial" analogue of the scale curve introduced in the context of central regions based on statistical depth functions by Liu, Parelius and Singh (1999). We endorse their emphasis on the appeal and importance of visualizing features of multivariate distributions by *one-dimensional curves*: "the very simplicity of such objects ... makes them powerful as a general tool for the practicing statistician". Matrix-valued dispersion measures are more informative, however, on the shape and orientation of the underlying distribution, and spatial versions of these are also discussed in Section 2.

Besides having direct appeal in their own rights, the spatial location and volume functionals are utilized in formulating a spatial "skewness functional" and corresponding measures of asymmetry (Section 3.1) and a spatial "kurtosis functional" and related tailweight and peakedness measures (Section 3.2). Some authors treat kurtosis, tailweight and peakedness as equivalently measuring the same feature (or its inverse), but here we follow those authors who distinguish these as separate although interrelated entities. In this regard, we interpret kurtosis as a measure of the degree of shift of probability mass from the "shoulders" of a distribution toward the center and/or the tails. For multivariate distributions, this conceptualization appears to be new. It is illustrated in Figure 3.1.

2 Spatial quantiles and related location and spread measures

2.1 The spatial quantiles: formulation and overview

For univariate Z with $E|Z| < \infty$, and for $0 < p < 1$, the L_1-based definition of univariate quantiles characterizes the pth quantile as any value θ minimizing

$$E\{|Z-\theta| + (2p-1)(Z-\theta)\} \tag{1}$$

(Ferguson, 1967, p. 51). As an extension to $\mathbb{R}^d$, "spatial" or "geometric" quantiles were introduced by Chaudhuri (1996) as follows. First (1) is rewritten as

$$E\{|Z-\theta| + u(Z-\theta)\}, \tag{2}$$

where $u = 2p-1$, thus re-indexing the univariate pth quantiles for $p \in (0,1)$ by u in the open interval $(-1,1)$. Then d-dimensional "quantiles" are formulated by extending this index set to the open unit ball $\mathbb{B}^{d-1}(0)$ and minimizing a generalized form of (2),

$$E\{\Phi(u, X-\theta) - \Phi(u, X)\}, \tag{3}$$

where X and θ are $\mathbb{R}^d$-valued and $\Phi(u,t) = \|t\| + \langle u, t\rangle$ with $\|\cdot\|$ the usual Euclidean norm and $\langle\cdot,\cdot\rangle$ the usual Euclidean inner product. (Subtraction of $\Phi(u, X)$ in (3) eliminates the need of a moment assumption.) This yields, corresponding to the underlying distribution function F for X on $\mathbb{R}^d$, and for $u \in \mathbb{B}^{d-1}(0)$, a "$u$th quantile" $Q_F(u)$ having both direction and magnitude. In particular, the well-known spatial median is given by $Q_F(0)$, which we shall also denote by M_F.

It is easily checked that, for each $u \in \mathbb{B}^{d-1}(0)$, the quantile $Q_F(u)$ may be represented as the solution x of

$$-E\{(X-x)/\|X-x\|\} = u. \tag{4}$$

An important inference from (4) is that we may attach to each point x in $\mathbb{R}^d$ a spatial quantile interpretation: namely, it is that spatial quantile $Q_F(u_x)$ indexed by the average unit vector u_x

pointing to x from a random point having distribution F. Since u_x is uniquely determined by (4) and satisfies $x = Q_F(u_x)$, we interpret u_x as the inverse at x of the spatial quantile function Q_F and denote it by $Q_F^{-1}(x)$. When the solution x of (4) is not unique, as illustrated for the univariate case in Section 2.4 below, multiple points x can have a common value of $Q_F^{-1}(x)$.

From (4) it also follows that "central" and "extreme" quantiles $Q_F(u)$ correspond to $\|u\|$ being close to 0 and 1, respectively. Thus we may think of the quantiles $Q_F(u)$ as indexed by a directional "outlyingness" parameter u whose magnitude measures outlyingness quantitatively, and thus we may measure the outlyingness of any point x quantitatively by the corresponding magnitude $\|u_x\|$ $= \|Q_F^{-1}(x)\|$.

Another immediate consequence of (4) is that $Q_F(u)$ is obtained by inverting the map

$$t \to -E\{(X-t)/\|X-t\|\}, \tag{5}$$

from which it is seen that spatial quantiles are a special case of the "M-quantiles" introduced by Breckling and Chambers (1988) and also treated by Koltchinskii (1997), Breckling, Kokic and Lübke (2001), and Kokic, Breckling and Lübke (2002).

The function $Q_F^{-1}(x) = -E\{(X-x)/\|X-x\|\}$ appearing in (4) and (5) is called by Möttönen and Oja (1995) the "spatial rank function", as it generalizes the univariate centered rank function, $2F(x)-1$, and similarly indicates the average direction and distance of an observation from the median. The spatial quantile function and the spatial rank function are simply inverses of each other. In the setting of the multivariate location model $F(x-\theta)$, the sample analogue rank function evaluated at a point θ_0 provides a "spatial sign test" statistic for the hypothesis $H_0 : \theta = \theta_0$.

Further, (4) yields the following useful property of the spatial quantile function. For the case that F is *centrally symmetric* about M_F, that is, $X - M_F$ and $M_F - X$ are identically distributed, the corresponding median-centered spatial quantile function Q_F is *skew-symmetric*:

$$Q_F(-u) - M_F = -(Q_F(u) - M_F),\ u \in \mathbb{B}^{d-1}(0). \tag{6}$$

This is easily derived (or see Koltchinskii, 1997, p. 448).

As a final application of (4), it is readily derived that the spatial quantiles are *equivariant* with respect to *shift*, *orthogonal*, and *homogeneous scale transformations*. That is, if the distribution F is transformed by $x \mapsto Ax + b$, with A proportional to an orthogonal matrix and b an arbitrary vector, then the same mapping applied to the original quantile function at u yields the quantile function of the transformed distribution, subject to the reindexing $u \mapsto u' = (\|u\|/\|Au\|)Au$:

$$Q_{AX+b}((\|u\|/\|Au\|)Au) = A\,Q_X(u) + b,\ u \in \mathbb{B}^{d-1}(0), \tag{7}$$

where for convenience we denote Q_G also by Q_Y for Y having distribution G. In particular, the spatial median of the transformed distribution is given by the same mapping applied to the spatial median of the original distribution: $M_{AX+b} = A\,M_X + b$. Note that the quantity $\|u\|$ is preserved under the reindexing, that is, $\|u'\| = \|u\|$, yielding the interpretation that the outlyingness measure associated with a given point x is *invariant* under the given linear transformation, that is, $\|Q_{AX+b}^{-1}(x)\| = \|Q_X^{-1}(x)\|$ for each $x \in \mathbb{R}^d$. In terms of a data cloud in $\mathbb{R}^d$, the sample spatial quantile function changes in the manner prescribed by (7) if the cloud of observations becomes translated, or homogeneously rescaled, or rotated about the origin, or reflected about a $(d-1)$-dimensional hyperplane through the origin. In view of the singular value decomposition of matrices, equivariance with respect to an *arbitrary affine* transformation $x \mapsto Ax + b$ fails only in the case that the action by A includes *heterogeneous* scale transformations of the coordinate variables.

Besides the above features which are of general utility, we note a number of compelling strong points possessed by the spatial quantile function:

- As discussed in Chaudhuri (1996), the solution $Q_F(u)$ to (4) always exists for any u, and it is unique if $d \geq 2$ and F is not supported on a straight line.

- It characterizes the associated distribution, in the sense that $Q_F = Q_G$ implies $F = G$ (see Koltchinskii, 1997, Cor. 2.9).

- It serves effectively as a basis for a variety of useful methodological techniques. For example, the extension of the *regression quantiles* of Koenker and Bassett (1978) for univariate response problems to the case of *multiresponse regresssion* is discussed in Chaudhuri (1996) and Koltchinskii (1997). As an analogue of procedures widely used in univariate data analysis, Marden (1998) illustrates the use of *bivariate QQ-plots* based on spatial quantiles, along with some related devices, and Chakraborty (2001) develops similar methods based on a modified type of sample spatial quantile (discussed below). Also, as noted above, notions of *multivariate ranks* may be based on spatial quantiles — see Jan and Randles (1994), Möttönen and Oja (1995), Chaudhuri (1996), Choi and Marden (1997), and Möttönen, Oja and Tienari (1997). This suggests the possibility of *spatial rank-rank plots*. Finally, in the present paper we see that the spatial quantile function may serve effectively as a basis for some appealing *nonparametric multivariate descriptive measures.*

- *Computation* of the sample spatial quantile function for a data set $X_1, \ldots, X_n$ via

$$-\frac{1}{n}\sum_{i=1}^{n} \frac{X_i - x}{\|X_i - x\|} = u \tag{8}$$

  is straightforward (see Chaudhuri, 1996), whereas, for example, many of the depth-based notions of multivariate quantiles are computationally intensive. (We note that the left-hand side of (8) is the *sample* version of the centered rank function discussed above. Likewise,

$$-\frac{1}{2n}\left[\sum_{i=1}^{n} \frac{X_i - x}{\|X_i - x\|} + \sum_{i=1}^{n} \frac{-X_i - x}{\| - X_i - x\|}\right]$$

  gives the sample *spatial signed-rank function* of Möttönen and Oja, 1995).

- We note from (8) a *robustness property* of $Q_n(u)$: its value remains unchanged if the points X_i are moved outward along the rays joining them with $Q_n(u)$. Moreover, it has *favorable breakdown point* (50% for the median – see Kemperman, 1987, and Lopuhaä and Rousseeuw, 1991) and *bounded influence function* (Koltchinskii, 1997, p. 459).

- While the formulation of spatial quantiles as a solution of an L_1 optimization problem is quite different from that of multivariate quantiles defined in terms of *statistical depth functions* as boundary points of depth-based central regions of specified probability, in Serfling (2002c) it is seen that the spatial quantiles indeed possess a useful depth-based representation, in terms of a new "spatial depth function" which is quite natural: $D(x, F) = 1 - \|Q_F^{-1}(x)\|$. See also Vardi and Zhang (2000).

- It is relatively straightforward to extend spatial quantiles to the setting of *Banach spaces*, as discussed in Kemperman (1987) for the spatial median and by Chaudhuri (1996) and Chakraborty (2001) for the general case.

- *Asymptotic theory* for sample spatial quantiles has been developed. See Chaudhuri (1996) and Koltchinskii (1994, 1997) for results covering weak convergence, Bahadur representations, and Bahadur-Kiefer approximations.

On the other hand, the spatial quantile function also has some drawbacks, for which, however, various remedies have been proposed (although they are somewhat problematic):

- While in the univariate case pth quantiles possess important probabilistic interpretations, as points demarking tail regions of specified probability, no such interpretation of $Q_F(u)$ holds in the higher dimensional case. That is, for $d \geq 2$ the index u has no direct probabilistic interpretation. On the other hand, probabilistic interpretations can be attached indirectly, through an appropriate reparameterization based on the probability weight of the central regions, as discussed in Section 2.3 below. It is not easy, however, to characterize the relevant mapping.

- As noted earlier, the equivariance (7) and the invariance of outlyingness fail to hold under *heterogeneous* rescaling of coordinate variables. This can be of practical concern in applications involving coordinates with differing measurement scales. As pointed out by Chakraborty (2001, p. 391), we would like the outlyingness measure of a data point not to depend on the choice of coordinate system. This is especially important in the case of data points which are potential "outliers". For applications with coordinates measured in a common unit, however, the above equivariance is sufficient. In this regard, Marden (1998) comments that in some cases it may be satisfactory to transform variables to have similar scales at the outset of data analysis. Likewise, as pointed out by Van Keilegom and Hettmansperger (2002), when the variables of interest have a special physical interpretation, there is no interest in affinely transforming them.

  One approach toward resolution of equivariance issues is to suitably modify the estimators. Thus, to replace the sample spatial median by a fully affine equivariant version, Chakraborty, Chaudhuri and Oja (1998) apply a "transformation-retransformation" (TR) approach due to Chaudhuri and Sengupta (1993), whereby the data are reexpressed in a selected "data-driven coordinate system", in terms of which the spatial median is computed and transformed back to the original coordinate system. With proper formulation of the data-based coordinate system, the modified sample spatial median becomes fully affine equivariant. Extension of this approach to the sample spatial quantiles in general is carried out by Chakraborty (2001), who establishes that if the given data are transformed by $x \mapsto Ax+b$, for any $d \times d$ nonsingular matrix A and arbitrary vector b, then the sample TR quantile function at u and the sample TR quantile function of the affinely transformed data satisfy the equivariance relation (7). It should be noted, however, that the TR approach is equivalent to modification of the objective function (3) in a complicated way (see Chakraborty (2001, p. 385) that yields as minimizer a *random* quantile function $Q_{F,n}^{(\mathrm{TR})}$ depending not only upon F but also upon on a subset of $d+1$ observations arbitrarily selected from the sample from F. The relationship of $Q_{F,n}^{(\mathrm{TR})}$ to the spatial quantile function Q_F is not clear, nor is geometric interpretation straightforward. It is this random quantile function that the sample TR quantile function "estimates" (consistently in a conditional sense), and, therefore, the judicious use of this approach depends upon the particular purposes of application.

  Another way to modify the objective function (3) to produce affine equivariance is to replace $\|X - \theta\|$ in the definition of Φ by $\sqrt{(X-\theta)'\Sigma^{-1}(X-\theta)}$, where Σ is the covariance matrix of X, as suggested by Isogai (1985) and Rao (1988) in treating the spatial median. Of course, one may question whether the resulting coordinate system produced by so standardizing the coordinate variables has appeal from a geometric standpoint. In any case, practical implementation requires use of a consistent and affine equivariant estimator of Σ, preferably one that is also robust. (See Section 3.3 below for related discussion.)

- For some data sets the sample spatial quantile contours (more generally, sample M-quantiles) can lie well beyond the convex hull of the data. See Breckling, Kokic and Lübke (2001), and Kokic, Breckling and Lübke (2002) for illustration with the $\|u\| = 0.9$ contour for a "cigar-shaped" data set and for proposed solutions consisting of modifications of the objective function (3) in a way equivalent to introducing a weight function of form $w(u, X - x)$ into the expectation in (4). The properties of such modifications remain to be explored.

Taking an overall view of the strengths and shortcomings of the spatial quantile function, we regard it as a very attractive type of multivariate quantile that merits continued investigation and development.

2.2 Spatial location measures

As discussed in Section 1, the standard "spatial" location measure is the well-known *spatial median* given by $Q_F(0) = M_F$. Additional forms of location measure are generated by quantile-based "L-functionals" (e.g., Serfling, 1980, in the univariate case), which in the present context are given by (vector-valued) weighted averages of the spatial quantile function, $\int_{\mathbb{B}^{d-1}(0)} Q_F(u)\, \mu(du)$, with respect to signed measures $\mu(du)$ on the index set $\mathbb{B}^{d-1}(0)$. See Chaudhuri (1996) and Chakraborty (2001) for some discussion. Here we specialize to a particular class of *location measures*, defined by

$$\ell_F(r) = \int_{\mathbb{S}_r^{d-1}(0)} Q_F(u)\, m(du), \ \ 0 \leq r < 1,$$

where $\mathbb{S}_r^{d-1}(0)$ is the sphere (the surface of the ball) of radius r centered at the origin 0, and $m(du)$ is the uniform measure on this sphere. Note that $\ell_F(0)$ is just M_F. Moreover, in the case of centrally symmetric F, it follows readily from (6) that $\ell_F(r) \equiv M_F$. Considered as a function of r, we call $\ell_F(\cdot)$ the *location functional* corresponding to F through its associated spatial quantile function. It is easily seen that $\ell_F(r)$ is equivariant with respect to shift, orthogonal and homogeneous scale transformations.

Various applications are supported by the spatial location functional. For example, as pointed out by Chaudhuri (1996), a spatial version of multivariate trimmed mean is given by the integral of $Q_F(u)$ with respect to the uniform measure on a subset of $\mathbb{B}^{d-1}(0)$ of form $\{u : \|u\| \leq \beta\}$. In terms of the location functional, this is just $\int_0^\beta \ell_F(r)\, dr$. Further, this location functional plays a direct role in defining a spatial skewness measure in Section 3.1.

Of course, there are other notions of location functional that may be associated with the spatial median. For example, Avérous and Meste (1997b) extend the univariate interquantile intervals to multivariate "median balls" indexed by their radii, as a family of "central regions" which provide *optimal* summaries in a certain L_1 sense, and a corresponding location functional is defined by the centers of the balls. This location functional also is identically M_F in the case of a centrally symmetric F.

2.3 Spatial central regions and spread measures

Corresponding to the spatial quantile function Q_F, we call

$$C_F(r) = \{Q_F(u) : \|u\| \leq r\}$$

the rth *central region*. When F is centrally symmetric, the skew-symmetry of $Q_F - M_F$ given by (6) yields that the regions $C_F(r)$ have the nice property of being *symmetric* sets, in the sense that for each point x in $C_F(r)$ its reflection about M_F is also in $C_F(r)$. From the discussion in Section 2.1, it is clear that the central regions $C_F(r)$ are equivariant under shift, orthogonal and homogeneous scale transformations.

The (real-valued) *volume functional* corresponding to Q_F is defined by

$$v_F(r) = \text{volume}\,(C_F(r)),\ 0 \le r < 1.$$

For each r, $v_F(r)$ provides a dispersion measure, as noted by Chaudhuri (1996). It is invariant under shift and orthogonal transformations, and $v_F(r)^{1/d}$ is equivariant under homogeneous scale transformations. As an increasing function of the variable r, $v_F(r)$ characterizes the dispersion of F in terms of expansion of the central regions $C_F(r)$.

Analogous to the scale curve introduced by Liu, Parelius and Singh (1999) in connection with *depth-based* central regions indexed by their *probability weight*, the spatial volume functional may likewise be plotted as a "scale curve" over $0 \le r < 1$, thus providing a convenient two-dimensional device for the viewing or comparing of multivariate distributions of any dimension. Illustrations of the depth-based scale curves are included in Liu, Parelius and Singh (1999) and, for elliptical distributions along with detailed elucidation and inference approaches, in Hettmansperger, Oja and Visuri (1999). The latter suggest and illustrate in the bivariate case a *PP*-plot of the empirical cdf's of the elliptical areas determined by the data in each sample. These ideas may be exploited for the spatial scale curve as well. Alternatively, two multivariate distributions F and G may be compared via a *spread-spread plot*, the graph of $v_G\, v_F^{-1}$, as introduced for the univariate case in Balanda and MacGillivray (1990). Besides having intrinsic appeal as just described, the volume functional plays key roles in defining skewness and kurtosis measures in Section 3.

Since the central regions are ordered and increase with respect to the spatial "outlyingness" parameter r that describes their boundaries, i.e., $r < r'$ implies $C_F(r) \subset C_F(r')$, their probability weights p increase with r. Thus the central regions and associated volume functional and scale curve can equivalently be indexed by the probability weight of the central region. This relationship may be described by a mapping $\psi_F : r \mapsto p_r \in [0, 1)$, with inverse $\psi_F^{-1} : p \mapsto r_p$ (thus $p_r = \psi_F(r)$ and $r_p = \psi_F^{-1}(p)$), but characterization of this mapping is complicated.

An alternative notion of spatial dispersion function based on the median balls discussed in Section 2.2 is developed by Avérous and Meste (1997b). Under regularity conditions on F, the probability weight of a median ball is a nondecreasing function of its radius, even in cases when the balls are not ordered by inclusion. This yields an analogue of the scale curve described above.

Matrix-valued dispersion measures. As an analogue of the usual covariance matrix, one can also consider *matrix-valued* dispersion measures based on the spatial quantiles, e.g.,

$$S(F) = \int_{\mathbb{B}^{d-1}(0)} (Q_F(u) - M_F)(Q_F(u) - M_F)'\lambda(du),$$

for measures $\lambda(du)$ on $\mathbb{B}^{d-1}(0)$. For example, a suitable choice of $\lambda(\cdot)$ yields a *trimmed dispersion measure*. Such scatter matrices contain information on the shape and orientation of the probability distribution as well as on the variations and mutual dependence of the coordinate variables. Real-valued "generalized variance" measures are provided by the corresponding determinants. See Serfling (2002b) for related discussion. We note that the spatial version of $S(F)$ satisfies the "covariance equivariance"

$$S(F_{AX+b}) = AS(F_X)A'$$

for all $d \times d$ (proportionally) *orthogonal* A and all $b \in \mathbb{R}^d$.

2.4 A simple illustration: the univariate case

To illustrate the above definitions in familiar terms, note that for $d = 1$ and univariate F, we have $\mathbb{B}^0(0) = (-1, +1)$, $\mathbb{S}^0_r(0) = \{-r, r\}$, $M_F = F^{-1}(\frac{1}{2})$, and $Q_F(u) = F^{-1}(\frac{1}{2} + \frac{u}{2})$, $-1 < u < 1$. It is readily seen that $Q_F^{-1}(x) = 2F(x) - 1$, the usual univariate centered rank function, and, accordingly, $|2F(x) - 1|$ serves as a measure of the outlyingness of x relative to the distribution F on $\mathbb{R}$. Note that if F is constant over an interval $[x_1, x_2]$, then $F(x)$ and thus also $Q_F^{-1}(x)$ are constant over this interval.

The location functional corresponding to Q_F is

$$\ell_F(r) = \tfrac{1}{2}[F^{-1}(\tfrac{1}{2} - \tfrac{r}{2}) + F^{-1}(\tfrac{1}{2} + \tfrac{r}{2})],\ 0 \le r < 1,$$

which is comprised of the midpoints of the (nested) "interquantile intervals"

$$[F^{-1}(\tfrac{1}{2} - \tfrac{r}{2}),\ F^{-1}(\tfrac{1}{2} + \tfrac{r}{2})],$$

which in fact are the rth central regions $C_F(r)$ that shrink to M_F as $r \to 0$. This location functional is the same one suggested by Avérous and Meste (1990) as providing through its graph an L_1-sense location "parameter" more informative than any typical real-valued parameter.

The volume functional is given by the widths of these intervals,

$$v_F(r) = F^{-1}(\tfrac{1}{2} + \tfrac{r}{2}) - F^{-1}(\tfrac{1}{2} - \tfrac{r}{2}),\ 0 \le r < 1, \tag{9}$$

which increase with r. This is recognized to be a classical nonparametric spread measure arising in many treatments of skewness and kurtosis in the univariate case (see, for example, Avérous and Meste, 1990, and Balanda and MacGillivray, 1990, and also Section 3 below).

3 Spatial skewness and kurtosis measures

3.1 A spatial skewness functional

In general, a skewness measure should be location- and scale-free and, in the case of a "symmetric" distribution, equal 0. Classical *univariate* quantitative skewness measures thus have the form of a difference of two location measures divided by a scale measure, whereby skewness then is characterized by a sign indicating *direction* and a magnitude measuring *asymmetry*. Along with such measures, associated notions of the *ordering* of distributions according to their skewness have been developed. See van Zwet (1964), Doksum (1975), Oja (1981), MacGillivray (1986), and Benjamini and Krieger (1996) for background and extensive discussion.

Extension of the above notion of a skewness measure to the *multivariate* case should in principle yield a *vector*, thus again characterizing skewness by both a direction and an asymmetry measure. Here, of course, one must specify a notion of multivariate symmetry relative to which skewness represents a deviation. In the present paper we require a quantitative skewness measure to reduce to the null vector in the case of *central symmetry*, as defined in Section 2.1.

Despite the natural appeal of a *vector* notion of multivariate skewness, the classical treatment of the multivariate case has tended to focus upon asymmetry, developing a rich variety of real-valued measures that generalize the univariate case, but leaving largely unattended directional measures of skewness and the ordering of distributions by skewness. For a brief overview, see Kotz, Balakrishnan and Johnson (2000, Section 44.20). Recently, however, Avérous and Meste (1997a) open up a broader treatment by introducing two *vector-valued skewness functionals* oriented to the spatial median, along with corresponding definitions of quantitative skewness, directional qualitative skewness, and directional ordering of multivariate distributions. In particular, one of their skewness functionals is given by the difference of the "median balls" location functional (discussed above in Section 2.2) and the spatial median M_F, divided by a fixed real-valued scale parameter, the inverse of the density of F evaluated at the spatial median. In the same vein, but utilizing instead the *spatial* location and volume functionals, we formulate a *spatial skewness functional*:

$$s_F(r) = 2\,\frac{\ell_F(r) - M_F}{v_F(r)^{1/d}},\ \ 0 < r < 1, \tag{10}$$

which in the case of centrally symmetric F reduces appropriately to the null vector, each r. Note that the scale factor in the denominator is allowed to depend on r. The power $1/d$ for $v_F(r)$ makes $s_F(r)$ invariant under any homogeneous scale transformation.

For each $r = r_0$, $s_F(r_0)$ represents a *quantitative* vector-valued skewness measure, indicating an overall direction of skewness. More generally, such a measure is given by any weighted average, $\beta_\mu(F) = \int_0^1 s_F(r)\,\mu(dr)$, taken with respect to a probability measure $\mu(dr)$ on $[0, 1)$ not depending on F. Further, we obtain quantitative *real-valued* measures of the skewness of F in any particular direction h, taken from the median M_F, by taking scalar products with the vector measures: $\langle s_F(r), h\rangle$, $0 < r < 1$, and $\langle \beta_\mu(F), h\rangle$.

Of course, one also may take $\langle s_F(r), h\rangle$, $0 < r < 1$, as a *functional* real-valued measure of skewness in the direction h. This provides a basis for straightforward *qualitative* notions of skewness. The following definitions and proposition parallel for "spatial skewness" the treatment of Avérous and Meste (1997a) for "median balls" skewness. The distribution F is called *weakly skew* in the direction h if $\langle s_F(r), h\rangle$ is nonnegative for each r, *strongly skew* if $\langle s_F(r), h\rangle$ is increasing in r. A related *ordering* of distributions, "F is less weakly skew than G in the direction h", is defined by

$$F \prec_h G \quad \Leftrightarrow \quad \langle s_G(r) - s_F(r), h\rangle \geq 0 \text{ for each } r.$$

Let $\overline{F}$ denote the distribution induced from F by the mapping $x - M_F \mapsto -(x - M_F)$. Then we have

PROPOSITION 3.1 *For any distribution F and any direction h, F is weakly skew in the direction h if and only if $\overline{F} \prec_h F$. Moreover, if F is weakly skew in the direction h, then $\langle \beta_\mu(F), h\rangle \geq 0$ for any probability measure μ.*

The proof is straightforward. In the univariate case, the interpretation of the first statement of Proposition 3.1 is that F is skew to the right if and only if $\overline{F}$ is less skew to the right than F.

Asymmetry measures. Our vector-valued spatial skewness functional yields a corresponding real-valued *asymmetry functional*, which we express in the form

$$\|s_F(r)\| = 2\,\frac{\|\int_{\mathbb{S}_r^{d-1}(0)} Q_F(u)\,m(du) - M_F\|}{v_F(r)^{1/d}},\ \ 0 < r < 1,$$

from which is obtained a real-valued *index of asymmetry* $A_F = \sup_{0<r<1}\|s_F(r)\|$. The latter index extends a measure of asymmetry for the univariate case suggested by MacGillivray (1986) and likewise may be used to order distributions: "F is less asymmetric than G", written $F \prec_A G$, if $A_F \leq A_G$. We might also consider $\int_0^1 \|s_F(r)\|\,dr$ as an asymmetry measure.

An alternative asymmetry functional, differing from $\|s_F(r)\|$ in the numerator, is proposed by Chaudhuri (1996):

$$\frac{\sup_{\|u\|=r}\|Q_F(u) + Q_F(-u) - 2M_F\|}{v_F(r)^{1/d}},\ \ 0 < r < 1. \tag{11}$$

Like $\|s_F(r)\|$, by (6) it equals 0 in the case of centrally symmetric F. As seen below, it coincides with $\|s_F(r)\|$ in the univariate case. Also, its supremum over r yields an alternative asymmetry measure.

See Oja (1983) and Avérous and Meste (1997a) for extensive discussion of other asymmetry measures, including, for example, those of form $(\mu_1 - \mu_2)'\Sigma^{-1}(\mu_1 - \mu_2)$, where μ_1 and μ_2 are two different location measures for F, such as the mean and the spatial median, or the Wilks generalized mean and the Oja generalized median, and Σ is the covariance matrix of F. These extend the univariate Pearson-type measures (see below) and may be compared to the asymmetry measure of Mardia (1970, 1974), $E\{(X - \mu_F)'\Sigma^{-1}(Y - \mu_F)\}^3$, where X and Y are independent random vectors having distribution F and μ_F is the mean of F. □

Asymmetry curves. Analogous to the scale curve discussed Section 2.3, a plot of the asymmetry functional $\|s_F(r)\|$, $0 < r < 1$, as a "spatial skewness curve" provides a convenient two-dimensional summary of the skewness of a multivariate distribution. Likewise we may plot a *directional* version $\langle s_F(r), h\rangle$, $0 < r < 1$, for any selected direction h.

An alternative summary, related to (11), is given by a plot of

$$\sup_{\|u\|=r} \frac{\|Q_F(u) - M_F\|}{\|Q_F(-u) - M_F\|},\ 0 < r < 1.$$

By (6), we see that in the case of centrally symmetric F, this curve follows the constant level 1. In the univariate case it is equivalent to a plot of $F^{-1}(1-p) - F^{-1}(\frac{1}{2})$ versus $F^{-1}(\frac{1}{2}) - F^{-1}(p)$, which is discussed in Gilchrist (2000).

Another type of asymmetry curve is obtained by adapting one given by Liu, Parelius and Singh (1999) in the context of depth-based central regions. For each r, let $I_F(r)$ denote the intersection of the central region $C_F(r)$ and its reflection about M_F, and let $w(r)$ denote the ratio of the volume of $I_F(r)$ to that of $C_F(r)$, over $0 < r < 1$. Since, as seen in Section 2.3, for centrally symmetric F the intersection $I_F(r)$ coincides with $C_F(r)$ and thus $w(r) \equiv 1$, a departure of F from central symmetry about M_F is indicated by the degree to which the curve $w(r)$ lies below the constant level 1.

For measuring departure from other types of symmetry, namely spherical, elliptical, and angular, Liu, Parelius and Singh (1999) construct similar summary curves using central regions in somewhat different ways. See their paper for illustrations. These approaches too can be adapted to the context of the spatial central regions. □

Further remarks on the univariate case. Utilizing Section 2.4, in the case $d = 1$ we have

$$\begin{aligned} s_F(r) &= \frac{F^{-1}(\frac{1}{2} - \frac{r}{2}) + F^{-1}(\frac{1}{2} + \frac{r}{2}) - 2\,M_F}{F^{-1}(\frac{1}{2} + \frac{r}{2}) - F^{-1}(\frac{1}{2} - \frac{r}{2})} \\ &= b_2(\tfrac{1}{2} - \tfrac{r}{2}),\ 0 < r < 1, \end{aligned} \tag{12}$$

where

$$b_2(\alpha) = \frac{F^{-1}(\alpha) + F^{-1}(1-\alpha) - 2\,M_F}{F^{-1}(1-\alpha) - F^{-1}(\alpha)},\ 0 < \alpha < \tfrac{1}{2},$$

which is a general skewness functional formulated by Oja (1981) and shown to be compatible with the skewness ordering of van Zwet (1964), and which is further discussed by Groeneveld and Meeden (1984), Avérous and Meste (1997a), and Gilchrist (2000). Indeed, Avérous and Meste (1997a) suggest positivity of the numerator appearing in (12) as a notion of "qualitative weak skewness to the right". The sample analogue form was introduced by David and Johnson (1956), and its practical role as a diagnostic tool is discussed by Parzen (1979) and Benjamini and Krieger

(1996). The special case involving the quartiles, $b_2(\frac{1}{4}) = s_F(\frac{1}{2}) = (Q_1 + Q_3 - 2Q_2)/(Q_3 - Q_1)$, has a long history dating to Galton (1875). See Brys, Hubert and Struyf (2003) for a study of the robustness of $b_2(\frac{1}{4})$ and $b_2(\frac{1}{8})$, along with some other skewness measures they propose. As noted by Groeneveld and Meeden (1984), the ratio of integrals of the numerator and denominator of $b_2(\cdot)$ yields another skewness measure,

$$\frac{\mu_F - M_F}{E|X - M_F|},$$

which compares with the early skewness measure $(\mu_F - M_F)/\sigma$ and with perhaps the earliest measure of all, $(\mu_F - \widetilde{M}_F)/\sigma$, where $\widetilde{M}_F$ denotes the mode of F. The latter measure is due to Pearson (1895). See also MacGillivray (1986) for useful discussion.

Some of the asymmetry curves discussed above are new even in the univariate case. A close variant, however, is given by Benjamini and Krieger (1996), who plot the numerator of (12) versus its denominator, as a "skewness-spread" plot to detect departures from symmetry. □

Remarks 3.1 (i) Through the mapping $p \mapsto r_p \in [0, 1)$ discussed in Section 2.3, one can define for any $\alpha \leq 1/2$ an index of asymmetry for the central (in the spatial sense) $100(1 - 2\alpha)\%$ part of the distribution F: $\sup_{\alpha \leq p \leq 1/2} \|s_F(r_p)\|$. This extends a univariate treatment by MacGillivray (1986).

(ii) By varying in (10) the choice of "spatial" location functional differenced with M_F and the choice of scaling in the denominator, competing "spatial skewness functionals" can be generated. It is of interest to study these possibilities comparatively with respect to invariance properties, notions of ordering, and other criteria.

(iii) The asymptotic behavior of sample versions of the vector-valued skewness measures has not been pursued, although some partial results are available. Let us discuss $s_F(\cdot)$, for example, with sample versions of Q_F, M_F, ℓ_F, v_F and s_F denoted by Q_n $(= Q_{F_n})$, M_n, ℓ_n, v_n and s_n. For any fixed u, the asymptotic normality of $Q_n(u)$ along with a Bahadur representation is derived in Chaudhuri (1996). More generally, as a special case of results for M-quantiles in general, Koltchinskii (1997, Theorems 4.2 and 5.7) establishes weak convergence of the standardized function $Q_n(\cdot)$ (along with a related Bahadur representation) and of corresponding L-statistics, yielding asymptotic normality of M_n, $\ell_n(r)$, and the difference $\ell_n(r) - M_n$, thus taking care of the numerator of $s_n(r)$ for any fixed r. For the denominator, partial but incomplete results on asymptotic normality of $v_n(r)$ are available in Serfling (2002c). In order to obtain asymptotic normality of $s_n(r)$ itself, however, it remains to develop a unified and complete treatment that yields *joint* asymptotic normality of $\ell_n(r) - M_n$ and $v_n(r)$. This is deferred to a future study. □

3.2 A spatial kurtosis functional

As seen above, although many differing notions of skewness have been formulated, there is general agreement on the intuitive meaning of skewness and on the kind of distributional feature it is intended to capture. With kurtosis, however, not only are there many versions, but also there is variation in views about what feature the term should denote. According as the standardized fourth central moment $\sigma^{-4}E(X - \mu)^4$ is greater than or less than 3 (the value for normal distributions), a distribution is classified as "leptokurtic" (peaked) or "platykurtic" (flat). In this way the term "kurtosis" has become associated with measurement of "peakedness" or its inverse, "tailweight", but it has been found not to be a perfect discriminator in this sense. Finecan (1964) suggested that this quantity actually measures the shift of probability mass away from the points $\mu \pm \sigma$ (called the "shoulders") and toward either the middle or the tails, and Moors (1986) showed how kurtosis may be interpreted as measuring the dispersion about these points. See Balanda and MacGillivray (1988) for useful discussion.

To avoid the complication that kurtosis becomes somewhat entangled with skewness in describing the shape of an asymmetric distribution, many authors have restricted the treatment of kurtosis to symmetric distributions, or to symmetrized versions of asymmetric distributions. In particular, for the case of univariate *symmetric* distributions, a *quantile-based* kurtosis functional was introduced by Groeneveld and Meeden (1984) as the (univariate) skewness measure (12) applied to the folded random variable $|X - M_F|$. The resulting functional may be expressed as

$$\frac{F^{-1}(\frac{3}{4}-\frac{p}{4}) + F^{-1}(\frac{3}{4}+\frac{p}{4}) - 2\,F^{-1}(\frac{3}{4})}{F^{-1}(\frac{3}{4}+\frac{p}{4}) - F^{-1}(\frac{3}{4}-\frac{p}{4})},\ 0 < p < 1. \tag{13}$$

It explicitly manifests the "shoulders" of the symmetric distribution F not as $\mu \pm \sigma$ but rather as the 1st and 3rd quartiles. See Balanda and MacGillivray (1988) and Groeneveld (1998) for detailed discussion of this functional, as well as MacGillivray and Balanda (1988) and Balanda and MacGillivray (1990) for discussion of issues surrounding the asymmetric case. See Gilchrist (2000) for some "upper" and "lower" kurtosis measures which are variants of (13). □

Turning now to the multivariate case, and drawing upon the above discussion, we consider kurtosis to characterize the relative degree, in a location- and scale-free sense, to which probability mass of a distribution is diminished in the "shoulders" and heavier in the either the center or tails or both. We thus distinguish *peakedness*, *kurtosis* and *tailweight* as distinct, although very much interrelated, features of a distribution. As an analogue of (13) based on the spatial volume functional, we introduce a *spatial kurtosis functional*:

$$k_F(r) = \frac{v_F(\frac{1}{2}-\frac{r}{2}) + v_F(\frac{1}{2}+\frac{r}{2}) - 2\,v_F(\frac{1}{2})}{v_F(\frac{1}{2}+\frac{r}{2}) - v_F(\frac{1}{2}-\frac{r}{2})},\ 0 < r < 1. \tag{14}$$

This functional is invariant under shift, orthogonal and homogeneous scale transformations.

Interpretation of $k_F(r)$ may thus be based on consideration of the boundary of the central region $C_F(\frac{1}{2})$ as representing the "shoulders" of the multivariate distribution, separating a "central part" from a complementary "tail part". The quantity $k_F(r)$ measures the relative volumetric difference between regions within and without the shoulders, which are defined by shifting the "outlyingness" parameter $\frac{1}{2}$ of the shoulders by equal amounts $\frac{r}{2}$ toward the center and toward the tails. The nature of $k_F(r)$ is easily understood via Figure 3.1, which displays this measure as the difference between the volumes of specified regions A and B, divided by their sum:

$$\frac{\text{volume}(B) - \text{volume}(A)}{\text{volume}(B) + \text{volume}(A)}.$$

An important point to note is that $k_F(r)$ has straightforward meaning and appeal without regard to considerations of symmetry. Conceptualizing kurtosis in a general way for multivariate distributions even yields helpful clarifications for the univariate case, for which $k_F(r)$ yields a natural extension of (13) to asymmetric distributions and reduces to (13) for symmetric F.

Remarks 3.2 As emphasized above, we interpret kurtosis as measuring a feature which is interrelated with peakedness and tailweight but not to be equated with either of these. Here we comment on peakedness and tailweight as separate from kurtosis.

(i) *Tailweight.* A family of tailweight measures based on the spatial quantiles is given by

$$t_F(r,s) = \frac{v_F(r)}{v_F(s)},\ 0 < r < s < 1, \tag{15}$$

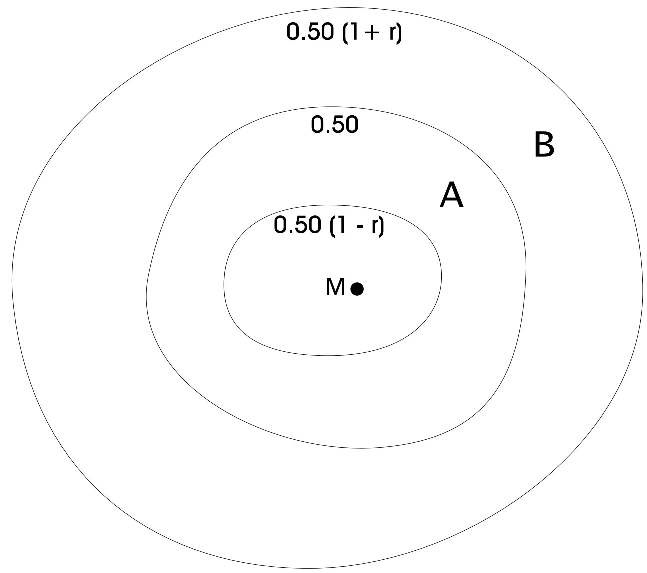

FIGURE 3.1. Median M and central regions $C_F(\frac{1}{2} - \frac{r}{2})$, $C_F(\frac{1}{2})$, $C_F(\frac{1}{2} + \frac{r}{2})$, with $A = C_F(\frac{1}{2}) - C_F(\frac{1}{2} - \frac{r}{2})$, and $B = C_F(\frac{1}{2} + \frac{r}{2}) - C_F(\frac{1}{2})$.

which reduces in the univariate case to ratios of the spread functional (9) evaluated at different points (treated by Balanda and MacGillivray, 1990). Using the term "kurtosis" for tailweight, a similar multivariate extension using *depth-based* central regions is given by Liu, Parelius and Singh (1999), who introduce a "fan plot" exhibiting the curves $t_F(r, s)$ for a fixed choice of r and selected choices of s. They also introduce other forms of tailweight measures, i.e, a Lorenz curve and a "shrinkage plot", which likewise may be formulated analogously in terms of the spatial quantile function. Several multivariate distributions or data sets may be compared with respect to tailweight on the basis of their respective (either spatial or depth-based) fan plots, Lorenz curves, or shrinkage plots. Asymptotics for sample versions of the kurtosis functional $k_F(\cdot)$ and these other transforms of the volume functional may be derived from the asymptotics for the scale curve as discussed in Section 3.3 below.

Bickel and Lehmann (1975a) suggest that a measure of "kurtosis" (meaning tailweight) is given by any suitable ratio of two scale measures. Typical tailweight measures indeed are of this form, but such a restriction is too restrictive for the more refined notion of kurtosis as distinct from tailweight. The numerator of (13), for example, is not a scale measure (see MacGillivray and Balanda, 1988, and Balanda and MacGillivray, 1990, for discussion).

(ii) *Peakedness.* The term "peakedness" is traditionally used synonymously with "concentration" or inversely with "dispersion" or "scatter". For key definitions and developments, in the univariate case see Brown and Tukey (1946), Birnbaum (1948), and Bickel and Lehmann (1976) and in the multivariate case Sherman (1955), Eaton (1982), Oja (1983), Olkin and Tong (1988), and Zuo and Serfling (2000). In particular, the latter authors provide a depth-based notion for ordering distributions by "more scattered": relative to a depth function $D(x, \cdot)$, the distribution F on $\mathbb{R}^d$ is more scattered than the distribution G if the D-based volume functional for F lies above that of G. As an appropriate analogue in terms of the spatial quantile functional, we thus define:

F is more scattered (less peaked) than G if $v_F(r) \geq v_G(r)$, $0 < r < 1$. This provides an alternative to the notions of Oja (1983) and Zuo and Serfling (2000), and in the univariate case it essentially reduces, via (9), to the notion of Brown and Tukey (1946). □

3.3 Concluding Remarks

A desirable property of any skewness, kurtosis, or other shape measure is that it satisfy "reverse implications", that is, that distributions can be ordered by the given measure (see MacGillivray and Balanda, 1988, p. 320). An important further investigation is to explore possible orderings associated with the spatial skewness and kurtosis functionals.

In the univariate case, detailed investigation of the interrelationships among spread, peakedness, skewness, kurtosis, and tailweight measures and associated orderings has been carried out by Oja (1981), MacGillivray and Balanda (1988), Balanda and MacGillivray (1990), and MacGillivray (1992), among others. A similar study of the (multivariate) spatial versions of these measures would be of interest.

One way to compare kurtosis or tailweight across two samples is via the corresponding sample kurtosis functionals. Another method, suggested by Hettmansperger, Oja and Visuri (1999) and illustrated in the bivariate case, is a *standardized* form of *PP* scale plot (as discussed in Section 2.3 above), whereby an S-shaped curve indicates difference in kurtosis.

The kurtosis functional $k_F(r)$ given by (14), and the tailweight functional $t_F(r, s)$ given by (15) for fixed s, are Hadamard differentiable transforms of the volume functional $v_F(r)$. Thus the weak convergence of suitably defined empirical spatial kurtosis and tailweight processes may be obtained routinely, using standard results on weak convergence of transformed random elements (e.g., van der Vaart, 1998), from results on weak convergence of the spatial empirical volume process. At this point, however, only incomplete results (Serfling, 2002c) are available for empirical volume processes.

Acknowledgment

Very useful comments by J. Avérous, V. Koltchinkskii, H. Oja, Xin Dang and two anonymous referees have led to substantive improvements in the paper and are greatly appreciated. Also, the author gratefully acknowledges support under National Science Foundation Grant DMS-0103698.

References

[1] Avérous, J. and Meste, M. (1990). Location, skewness and tailweight in L_s-sense: a coherent approach. *Mathematische Operationsforshung und Statistik, Series Statistics* **21** 57–74.

[2] Avérous, J. and Meste, M. (1997a). Skewness for multivariate distributions: two approaches. *Annals of Statistics* **25** 1984–1997.

[3] Avérous, J. and Meste, M. (1997b). Median balls: an extension of the interquantile intervals to multivariate distributions. *Journal of Multivariate Analysis* **63** 222–241.

[4] Balanda, K. P. and MacGillivray, H. L. (1988). Kurtosis: a critical review. *The American Statistician* **42** 111–119.

[5] Balanda, K. P. and MacGillivray, H. L. (1990). Kurtosis and spread. *The Canadian Journal of Statistics* **18** 17–30.

[6] Benjamini, Y. and Krieger, A. M. (1996). Concepts and measures for skewness with data-analytic implications. *Canadian Journal of Statistics* **24** 131–140.

[7] Bickel, P. J. and Lehmann, E. L. (1975a). Descriptive statistics for nonparametric models. I. Introduction. *Annals of Statistics* **3** 1038–1044.

[8] Bickel, P. J., and Lehmann, E. L. (1975b). Descriptive statistics for nonparametric models. II. Location. *Annals of Statistics* **4** 1045–1069.

[9] Bickel, P. J., and Lehmann, E. L. (1976). Descriptive statistics for nonparametric models. III. Dispersion. *Annals of Statistics* **4** 1139–1158.

[10] Bickel, P. J., and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models. IV. Spread. *Contributions to Statistics: Jaroslav Hájek Memorial Volume* (J. Jurečková, ed.), pp. 33–40. Academia, Prague.

[11] Birnbaum, Z. W. (1948). On random variables with comparable peakedness. *Annals of Mathematical Statistics* **19** 76–81.

[12] Breckling, J. and Chambers, R. (1988). M-quantiles. *Biometrika* **75** 761–771.

[13] Breckling, J., Kokic, P. and Lübke, O. (2001). A note on multivariate M-quantiles. *Statistics & Probability Letters* **55** 39–44.

[14] Brown, G., and Tukey, J. W. (1946). Some distributions of sample means. *Annals of Mathematical Statistics* **7** 1–12.

[15] Brys, G., Hubert, M. and Struyf, A. (2003). A comparison of some new measures of skewness. In *Developments in Robust Statistics (ICORS 2001)*, pp. 98–113.

[16] Chakraborty, B. (2001). On affine equivariant multivariate quantiles. *Annals of the Institute of Statistical Mathematics* **53** 380–403.

[17] Chakraborty, B., Chaudhuri, P., and Oja, H. (1998). Operating transformation and retransformation on spatial median and angle test. *Statista Sinica* **8** 767–784.

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

[19] Chaudhuri, P. and Sengupta, D. (1993). Sign tests in multidimension: Inference based on the geometry of the data cloud. *Journal of the American Statistical Association* **88** 1363–1370.

[20] Choi, K. and Marden, J. (1997). An approach to multivariate rank tests in multivariate analysis of variance. *Journal of the American Statistical Association* **92** 1581–1590.

[21] David, F. N. and Johnson, N. L. (1956). Some tests of significance with ordered variables. *Journal of the Royal Statistical Society, Series B* **18** 1–31.

[22] Doksum, K. A. (1975). Measures of location and asymmetry. *Scandinavian Journal of Statistics* **2** 11–22.

[23] Eaton, M. L. (1982). A review of selected topics in multivariate probability inequalities. *Annals of Statistics* **10** 11–43.

[24] Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York.

[25] Finecan, H. M. (1964). A note on kurtosis. *Journal of the Royal Statistical Society, Series B* **26** 111–112.

[26] Galton, F. (1875). Statistics by intercomparison: with remarks on the Law of Frequency of Error. *Philosophical Magazine*, 4th series, **49** 33–46.

[27] Gilchrist, W. G. (2000). *Statistical Modelling with Quantile Functions*. Chapman & Hall.

[28] Groeneveld, R. A. (1998). A class of quantile measures for kurtosis. *The American Statistician* **51** 325–329.

[29] Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *The Statistician* **33** 391–399.

[30] Hayford, J. F. (1902). What is the center of an area, or the center of a population? *Journal of the American Statistical Association* **8** 47–58.

[31] Hettmansperger, T. P., Oja, H. and Visuri, S. (1999). Discussion of Liu, Parelius and Singh (1999). *Annals of Statistics* **27** 845–854.

[32] Isogai, T. (1985). Some extension of Haldane's multivariate median and its application. *Annals of Institute of Statistical Mathematics* **37** 289–301.

[33] Jan, S. L. and Randles, R. H. (1994). A multivariate signed sum test for the one sample location problem. *Journal of Nonparametric Statistics* **4** 49–63.

[34] Kemperman, J. H. B. (1987). The median of a finite measure on a Banach space. In *Statistical Data Analysis Based On the L_1-Norm and Related Methods* (Y. Dodge, ed.), pp. 217–230, North-Holland.

[35] Koenker, R. and Bassett, G. (1978). Regression quantiles. *Journal of the American Statistical Association* **82** 851–857.

[36] Kokic, P., Breckling, J., and Lübke, O. (2002). A new definition of multivariate M-quantiles. In *Statistical Data Analysis Based On the L_1-Norm and Related Methods* (Y. Dodge, ed.), pp. 15–24. Birkhaüser, Basel.

[37] Koltchinskii, V. (1994). Bahadur-Kiefer approximation for spatial quantiles. In *Probability in Banach Spaces* (J. Hoffman-Jørgensen, J. Kuelbs and M. Marcus, eds.), pp. 394–408. Birkhaüser, Basel.

[38] Koltchinskii, V. (1997). M-estimation, convexity and quantiles. *Annals of Statistics* **25** 435–477.

[39] Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions, Volume 1: Models and Applications*, 2nd edition. Wiley, New York.

[40] Liu, R. Y., Parelius, J. M. and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion). *Annals of Statistics* **27** 783–858.

[41] Lopuhaä, H. P. and Rousseeuw, J. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Annals of Statistics* **19** 229-248.

[42] MacGillivray, H. L. (1986). Skewness and asymmetry: measures and orderings. *Annals of Statistics* **14** 994–1011.

[43] MacGillivray, H. L. (1992). Shape properties of the g-and-h and Johnson families. *Communications in Statistics – Theory and Methods* **21** 1233–1250.

[44] MacGillivray, H. L. and Balanda, K. P. (1988). The relationships between skewness and kurtosis. *Australian Journal of Statistics* **30** 319–337.

[45] Marden, J. I. (1998). Bivariate QQ-plots and spider web plots. *Statistica Sinica* **8** 813–826. *Sankhya B* **36** 115–128.

[46] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57** 519–530.

[47] Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya B* **36** 115–128.

[48] Moors, J. J. A. (1986). The meaning of kurtosis: Darlington reexamined. *The American Statistician* **40** 283–284.

[49] Möttönen, J. and Oja, H. (1995). Multivariate spatial sign and rank methods. *Journal of Nonparametric Statistics* **5** 201–213.

[50] Möttönen, J., Oja, H., and Tienari, J. (1997). On the efficiency of multivariate spatial sign and rank tests. *Annals of Statistics* **25** 542–552.

[51] Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics* **8** 154–168.

[52] Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics and Probability Letters* **1** 327–333.

[53] Olkin, I. and Tong, Y. L. (1988). Peakedness in multivariate distributions. In *Statistical Decision Theory and Related Topics IV*, Vol. 2 (S.S. Gupta and J. O. Berger, eds.), pp. 373–384, Springer.

[54] Parzen, E. (1979). Nonparametric statistical data modeling (with discussion). *Journal of the American Statistical Association* **74** 105–131.

[55] Pearson, K. (1895). Contributions to the mathematical theory of evolution II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Statistical Society, London, Series A* **186** 343.

[56] Rao, C. R. (1988). Methodology based on the L_1 norm in statistical inference. *Sankhyā Series A* **50** 289–313.

[57] Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

[58] Serfling, R. (2002a). Generalized quantile processes based on multivariate depth functions, with applications in nonparametric multivariate analysis. *Journal of Multivariate Analysis* **83** 232–247.

[59] Serfling, R. (2002b). Quantile functions for multivariate analysis: approaches and applications. *Statistica Neerlandica* **56** 214–232.

[60] Serfling, R. (2002c). A depth function and a scale curve based on spatial quantiles. In *Statistical Data Analysis Based On the L_1-Norm and Related Methods* (Y. Dodge, ed.), pp. 25-38. Birkhaüser, Basel. (modified postprint available at `www.utdallas.edu/~serfling`)

[61] Sherman, S. (1955). A theorem on convex sets with applications. *Annals of Mathematical Statistics* **26** 763–767.

[62] Small, C. G. (1990). A survey of multidimensional medians. *International Statistical Institute Review* **58** 263–277.

[63] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

[64] Van Keilegom, I. and Hettmansperger, T. (2002). Inference based on multivariate M-estimators based on bivariate censored data. *Journal of the American Statistical Association* **97** 328–336.

[65] van Zwet, W. R. (1964). *Convex Transformations of Random Variables*. Math. Centrum, Amsterdam.

[66] Vardi, Y. and Zhang, C.-H. (2000). The multivariate L_1-median and associated data depth. *Proceedings of National Academy of Science USA* **97** 1423–1426.

[67] Zuo, Y. and Serfling, R. (2000). Nonparametric notions of multivariate "scatter measure" and "more scattered" based on statistical depth functions. *Journal of Multivariate Analysis* **75** 62–78.