Distribution shape and size, equivariance properties of the influence function, and transformation groups

Frank Critchley
Department of Statistics, Open University
Guobing Lu
Department of Social Medicine, University of Bristol
Richard A. Atkinson
School of Mathematics and Statistics, University of Birmingham
and Dong Qian Wang
School of Mathematical and Computing Sciences,
Victoria University of Wellington

December 2003

SUMMARY

Statistical analysis of the shape and size of empirical distributions on $\mathbb{E}^k$ is well-established. Filtered invariants and scale-isovariants are studied as first steps towards corresponding analysis of more general multivariate distributions. A complete account of dispersion shape and size indices is obtained, along with general invariance and equivariance properties of the influence function. A corresponding diagnostic methodology based on complementary use of shape and size measures is outlined. A transformation group approach unifies presentation.

Some key words: Diagnostics; Equivariance; Influence function; Invariance; Isovariance; Location transformation; Orthogonal transformation; Scale transformation; Shape; Size; Transformation group.

1 INTRODUCTION

1.1 Distribution shape and size

The present paper has three related but distinct objectives, as follows.

Shape and size are complementary, varying independently of each other. For example, there are no constraints between the height : diameter ratio of a cylinder and its volume.

A great deal of progress has been made recently in the statistical analysis of the shape and size of finite Euclidean point sets. Equivalently, given sample size, of empirical distributions on $\mathbb{E}^k$ $(k \geq 1)$. Transformation groups a central rôle. The key to analysing *all* the information about shape in a multivariate data set is an explicit *complete* (or 'maximal') invariant under the relevant group. For size analysis, the key idea is a form of equivariance under scale transformation.

Overviews of this field are available in, for example, Dryden and Mardia (1998), Kendall, Barden, Carne and Le (1999) or Small (1996).

Shape and size problems for general multivariate distributions are more difficult, if not impossible, hindering both asymptotic and influence analysis. A first objective is to provide some tools with which, at least, to approach these problems.

A fundamental difficulty for shape analysis in this wider context is the lack, in general, of an explicit complete invariant under the relevant group. Progress can, however, be made by restricting attention to invariants which depend only on a specified feature of interest of a distribution. An explicit complete account of all such *filtered* invariants may then be available. This provides valuable insight in itself, while a collection of such accounts – one for each of several features of interest – provides a fuller, structured, account of all shape invariants. For example, as developed here, we may study all shape invariants filtered by (central) moments of given low order $l = 1, 2, \ldots$.

Size analysis in this wider context requires extension of the twin notions of scale transformation and cone beyond their usual real vector space setting. To this end, a (quite general) account of *scale-isovariants* is given.

A second objective is to understand *why* two propositions are true. Their statement involves the following preliminaries.

1.2 Preliminaries

Throughout, $\subseteq$ denotes inclusion, $\subset$ strict inclusion, $\mathbb{A}$ a nonempty set, $\mathrm{id}_{\mathbb{A}}$ the identity transformation on $\mathbb{A}$ and $\mathrm{dom}(\mathbb{A})$ the set of all functions having domain $\mathbb{A}$, while $t(\mathbb{A}') := \{t(a) : a \in \mathbb{A}'\}$ ($t \in \mathrm{dom}(\mathbb{A})$, $\mathbb{A}' \subseteq \mathbb{A}$). Again, $\circ$ denotes function composition as defined by

$$\forall\, t_1 \in \mathrm{dom}(\mathbb{A}), \forall\, t_2 \in \mathrm{dom}(t_1(\mathbb{A})), (t_2 \circ t_1)(a) := t_2(t_1(a)). \tag{1}$$

For any $\xi, \phi \in \mathrm{dom}(\mathbb{A})$, we say that ξ is *ϕ-filtered* if $\xi(a)$ depends on a only via $\phi(a)$. That is, if the diagram

$$\begin{array}{ccc} & a & \\ \swarrow & & \searrow \\ \phi(a) & \dashrightarrow & \xi(a) \end{array}$$

can be completed by a (necessarily unique) function $\zeta : \phi(\mathbb{A}) \to \xi(\mathbb{A})$ such that $\xi = \zeta \circ \phi$.

In a convenient convention, brackets around (sets of) distributions occurring in arguments of functions are always square. The L_p norm is denoted $\|\cdot\|_p$ ($p \geq 1$). Four sets, and mappings between them, are central to the study of distribution shape and size:

- $\mathbb{F}_k := \{$all distributions F on $\mathbb{E}^k : \mu[F]$ and $\Sigma[F]$ exist$\}$, where $\mu[F] := \int x\, dF(x)$ and $\Sigma[F] := \int (x - \mu[F])(x - \mu[F])^T dF(x)$,
- $\mathbb{D}_k := \Sigma[\mathbb{F}_k] = \{$all $k \times k$ dispersion matrices $\Sigma = (\sigma_{rs})\}$,

- $\mathbb{L}_k := \lambda(\mathbb{D}_k) = \{\lambda = (\lambda_r) \in \mathbb{E}^k : \lambda_1 \geq ... \geq \lambda_k \geq 0\}$, where $\lambda(\Sigma)$ contains the eigenvalues of Σ in nonincreasing order, and
- $\mathbb{W}_k := w(\mathbb{L}_k) = \{0_{\mathbb{E}^k}\} \cup \{\text{all probability vectors in } \mathbb{L}_k\}$, where

$$w(\lambda) := \begin{cases} \lambda / \|\lambda\|_1 & (\lambda \neq 0_{\mathbb{E}^k}). \\ 0_{\mathbb{E}^k} & (\lambda = 0_{\mathbb{E}^k}). \end{cases}$$

Exploiting the square bracket convention, there is no confusion in, for example, abbreviating '$\mathbb{W}_k = (w \circ \lambda \circ \Sigma)[\mathbb{F}_k]$' to '$\mathbb{W}_k = w[\mathbb{F}_k]$'.

In the present general multivariate context, *distribution shape* is concerned with features of distributions invariant under location, scale and orthogonal transformation. In the analysis of finite Euclidean point sets, 'reflection shape' is used with the same meaning to emphasise that the orthogonal transformations involved are not restricted to be rotations. Accordingly we work here with $\mathbb{O}_k$ rather than $\mathbb{SO}_k$ (in the usual notation), and let $\mathbb{G}$ denote the set of all transformations

$$g = (\kappa, Q, x_\circ) : x \to \kappa Q(x - x_\circ) \text{ on } \mathbb{E}^k, (\kappa > 0, Q \in \mathbb{O}_k, x_\circ \in \mathbb{E}^k),$$

with subsets $\mathbb{G}_l := \{g_l = (1, I_k, x_\circ) : x_\circ \in \mathbb{E}^k\}$, $\mathbb{G}_s := \{g_s = (\kappa, I_k, 0_{\mathbb{E}^k}) : \kappa > 0\}$ and $\mathbb{G}_o := \{g_o = (1, Q, 0_{\mathbb{E}^k}) : Q \in \mathbb{O}_k\}$.

Of central importance here is the fact that $\mathbb{G}$ induces a corresponding set $\widetilde{\mathbb{G}}$ of transformations $\widetilde{g}$ on $\mathbb{F}_k$ via $g \to \widetilde{g}$, where $\widetilde{g}[F]$ denotes the distribution of $g(X)$ induced by $X \sim F$.

1.3 Two propositions

Consider the following two propositions.

Proposition I: If $\iota[F] = \zeta(w[F])$ for some $\zeta \in \text{dom}(\mathbb{W}_k)$, then $\iota[F]$ is (clearly) unchanged by $F \to \widetilde{g}[F]$ $(g \in \mathbb{G})$. In fact, among all Σ-filtered functions ι in $\text{dom}[\mathbb{F}_k]$, *only* the w-filtered ones have this invariance property.

Proposition II: The functional $\text{trace}(\Sigma[F])$ is invariant under both location and orthogonal transformation, responding quadratically to scale changes. Its influence function, $\|x - \mu[F]\|_2^2 - \text{trace}(\Sigma[F])$, enjoys these *same* properties.

A transformation group approach meets our second objective in the sense that it provides general results of which these two propositions are special cases.

These results also meet a third objective. Namely, the provision of a framework within which an (outlier) diagnostic methodology based on complementary use of dispersion shape and size measures can be developed.

1.4 Organisation

The paper is organised as follows.

Complementary accounts of transformation groups are available in, for example, Eaton (1988), Kawakubo (1991) or Lehmann (1991). Both terminology and notation vary substantially between authors, depending on orientation. To avoid confusion, we establish ours in Section 2. Elements of both transformation groups and distribution shape are covered here, the former topic being illustrated with examples from the latter. The rôle of filtered functions is emphasised, and equivariance properties of distribution shape maps established.

Filtered invariants lead to a complete account of dispersion-filtered measures of distribution shape (Section 3). Scale-isovariants, a special form of equivariant central to distribution size analysis, are studied in Section 4 (and, in greater generality, at Appendix C). A complete account of dispersion-filtered measures of distribution size is then given in Section 5. General invariance and equivariance properties of the influence function are described in Section 6, where a corresponding (outlier) diagnostic methodology based on complementary use of dispersion shape and size measures is outlined.

A few well-known results are stated without proof in Section 2. For brevity, straightforward proofs are omitted throughout.

The further preliminaries given below are used.

1.5 Further preliminaries

Let $t \in \mathrm{dom}(\mathbb{A})$, $\mathbb{A}' \subseteq \mathbb{A}$ and $\mathbb{B} \subseteq t(\mathbb{A})$. Then $\{a \in \mathbb{A} : t(a) \in \mathbb{B}\}$, the *largest* subset of $\mathbb{A}$ that t maps onto $\mathbb{B}$, is called the inverse image of $\mathbb{B}$ under t, denoted $t^{-1}(\mathbb{B})$. If, further, t is $1-1$ with inverse function t^{-1}, the inverse image of $\mathbb{B}$ under t is the same as the image of $\mathbb{B}$ under t^{-1} and there is no ambiguity in the notation $t^{-1}(\mathbb{B})$, which is then the *unique* subset of $\mathbb{A}$ that t maps onto $\mathbb{B}$. Thus, in general,

$$t(t^{-1}(\mathbb{B})) = \mathbb{B} \text{ but } t^{-1}(t(\mathbb{A}')) \supseteq \mathbb{A}'. \tag{2}$$

Again, $\forall\, t_1 \in \mathrm{dom}(\mathbb{A}), \forall\, t_2 \in \mathrm{dom}(t_1(\mathbb{A})), \forall\, \mathbb{B} \subseteq (t_2 \circ t_1)\mathbb{A}$,

$$(t_2 \circ t_1)^{-1}\mathbb{B} = t_1^{-1}(t_2^{-1}(\mathbb{B})). \tag{3}$$

A quadruple $(\mathbb{F}, \mathbb{D}, \mathbb{L}, \mathbb{W})$ of nonempty subsets of, respectively, $\mathbb{F}_k$, $\mathbb{D}_k$, $\mathbb{L}_k$ and $\mathbb{W}_k$ is said to *correspond*, or to be *generated by* $\mathbb{W}$, if $\mathbb{L} = w^{-1}(\mathbb{W})$, $\mathbb{D} = \lambda^{-1}(\mathbb{L})$ and $\mathbb{F} = \Sigma^{-1}(\mathbb{D})$. By definition, $(\mathbb{F}_k, \mathbb{D}_k, \mathbb{L}_k, \mathbb{W}_k)$ itself is a corresponding quadruple while, using (2) and (3), we have

Proposition 1 *(Corresponding quadruples)*

(a) Let $\mathbb{F}$, $\mathbb{D}$, $\mathbb{L}$ and $\mathbb{W}$ be nonempty subsets of $\mathbb{F}_k$, $\mathbb{D}_k$, $\mathbb{L}_k$ and $\mathbb{W}_k$ respectively. Then the following three statements are equivalent:
[1] $(\mathbb{F}, \mathbb{D}, \mathbb{L}, \mathbb{W})$ *corresponds.*
[2] $\mathbb{L} = w^{-1}(\mathbb{W}), \mathbb{D} = \lambda^{-1}(\mathbb{L})$ *and* $\mathbb{F} = (\lambda \circ \Sigma)^{-1}(\mathbb{L})$.
[3] $\mathbb{L} = w^{-1}(\mathbb{W}), \mathbb{D} = (w \circ \lambda)^{-1}(\mathbb{W})$ *and* $\mathbb{F} = (w \circ \lambda \circ \Sigma)^{-1}(\mathbb{W})$.
(b) Let $(\mathbb{F}_i, \mathbb{D}_i, \mathbb{L}_i, \mathbb{W}_i)$ $(i \in \mathbf{I})$ *correspond. Then:*
(i) $(\cup_{i\in\mathbf{I}}\mathbb{F}_i, \cup_{i\in\mathbf{I}}\mathbb{D}_i, \cup_{i\in\mathbf{I}}\mathbb{L}_i, \cup_{i\in\mathbf{I}}\mathbb{W}_i)$ *corresponds.*
(ii) $(\cap_{i\in\mathbf{I}}\mathbb{F}_i, \cap_{i\in\mathbf{I}}\mathbb{D}_i, \cap_{i\in\mathbf{I}}\mathbb{L}_i, \cap_{i\in\mathbf{I}}\mathbb{W}_i)$ *corresponds whenever* $\cap_{i\in\mathbf{I}}\mathbb{W}_i \neq \varnothing$.
(c) If $(\mathbb{F}, \mathbb{D}, \mathbb{L}, \mathbb{W})$ *and* $(\mathbb{F}', \mathbb{D}', \mathbb{L}', \mathbb{W}')$ *correspond, then:*
$\mathbb{F} \subseteq \mathbb{F}' \Leftrightarrow \mathbb{D} \subseteq \mathbb{D}' \Leftrightarrow \mathbb{L} \subseteq \mathbb{L}' \Leftrightarrow \mathbb{W} \subseteq \mathbb{W}'$.

Table 1 identifies five corresponding quadruples in each of which $\mathbb{W}$ comprises those members of $\mathbb{W}_k$ satisfying the condition stated. Let $\mathbb{M}_k := \{$all $k \times k$ real symmetric matrices$\}$. Then, for example, $\mathbb{W}_k^{\deg} = \mathbb{L}_k^{\deg} = \{0_{\mathbb{E}^k}\}$ while $\mathbb{D}_k^{\deg} = \{0_{\mathbb{M}_k}\}$, so that $\mathbb{F}_k^{\deg} = \{\widehat{F}_x : x \in \mathbb{E}^k\}$ where $\widehat{F}_x$ denotes the distribution *degenerate* at $x \in \mathbb{E}^k$. Again, $\mathbb{D}_k^{+}$ comprises all positive definite members of $\mathbb{D}_k$, and $\mathbb{D}_k^{>}$ all members of $\mathbb{D}_k$ each of whose eigenvalues is simple.

Quadruple	$(\mathbb{F},$	$\mathbb{D},$	$\mathbb{L},$	$\mathbb{W})$	**Condition**
Q1	$(\mathbb{F}_k^{\deg},$	$\mathbb{D}_k^{\deg},$	$\mathbb{L}_k^{\deg},$	$\mathbb{W}_k^{\deg})$	$w_1 = 0$
Q2	$(\mathbb{F}_k^{(0)},$	$\mathbb{D}_k^{(0)},$	$\mathbb{L}_k^{(0)},$	$\mathbb{W}_k^{(0)})$	$w_1 > 0$
Q3	$(\mathbb{F}_k^{+},$	$\mathbb{D}_k^{+},$	$\mathbb{L}_k^{+},$	$\mathbb{W}_k^{+})$	$w_k > 0$
Q4	$(\mathbb{F}_k^{>},$	$\mathbb{D}_k^{>},$	$\mathbb{L}_k^{>},$	$\mathbb{W}_k^{>})$	$w_1 > ... > w_k$
Q5	$(\mathbb{F}_k^{+>},$	$\mathbb{D}_k^{+>},$	$\mathbb{L}_k^{+>},$	$\mathbb{W}_k^{+>})$	$w_1 > ... > w_k > 0$

Table 1: Five corresponding quadruples.

Note: Q2 and Q3 coincide for $k = 1$; Q4 and Q5 are only defined for $k > 1$.

Of course, $\mathbb{W}_k^{\deg}$ and $\mathbb{W}_k^{(0)}$ partition $\mathbb{W}_k$ while, $\forall\, k > 1$, $\mathbb{W}_k^{+>} = \mathbb{W}_k^{+} \cap \mathbb{W}_k^{>}$ satisfies $\mathbb{W}_k^{+>} \subset \mathbb{W}_k^{>} \subset \mathbb{W}_k^{(0)}$ and $\mathbb{W}_k^{+>} \subset \mathbb{W}_k^{+} \subset \mathbb{W}_k^{(0)}$, exactly similar relations holding amongst the corresponding subsets of $\mathbb{F}_k$, $\mathbb{D}_k$ and $\mathbb{L}_k$ by Proposition 1.

Finally, note that the empirical distribution $\widehat{F} := \sum_{i=1}^{n} n^{-1} \widehat{F}_{x_i}$ of a random sample $\{x_i\}_{i=1}^{n}$ from $F \in \mathbb{F}_k$ belongs itself to $\mathbb{F}_k$. Accordingly, replacing F by $\widehat{F}$ as needed, the following development provides a unified treatment of population and sample analysis, affording both insight and economy of presentation.

2 TRANSFORMATION GROUPS AND DISTRIBUTION SHAPE

2.1 Groups of transformations

A nonempty collection $\mathbb{T}$ of transformations $t : \mathbb{A} \to \mathbb{A}$ is called a *group of transformations on* $\mathbb{A}$, written $\mathbb{T} \rightsquigarrow \mathbb{A}$, if it is closed under both composition and inversion. In particular, each $t \in \mathbb{T}$ is, then, a $1-1$ transformation of $\mathbb{A}$ onto itself. For example, $\mathbb{G} \rightsquigarrow \mathbb{E}^k$ with

$$(\kappa_2, Q_2, x_2) \circ (\kappa_1, Q_1, x_1) = (\kappa_2\kappa_1, Q_2Q_1, x_1 + \kappa_1^{-1} Q_1^T x_2)$$

and $(\kappa, Q, x)^{-1} = (\kappa^{-1}, Q^T, -\kappa Q x)$. Again, $\widetilde{\mathbb{G}} \rightsquigarrow \mathbb{F}_k$.

A transformation group is, thus, an abstract group in which function composition, defined at (1), plays the rôle of group multiplication. Otherwise said, $\mathbb{T} \rightsquigarrow \mathbb{A}$ means that 'the group $\mathbb{T}$ acts on the left of $\mathbb{A}$ with action $a \to t(a)$'.

2.2 Subgroups

A subset $\mathbb{T}_\circ$ of an abstract group $\mathbb{T}$ is called a *subgroup* of $\mathbb{T}$ if $\mathbb{T}_\circ$ is itself an abstract group with respect to group multiplication on $\mathbb{T}$. Thus, if $\mathbb{T} \rightsquigarrow \mathbb{A}$ and $\mathbb{T}_\circ \subseteq \mathbb{T}$, $\mathbb{T}_\circ$ is a subgroup of $\mathbb{T}$ if and only if $\mathbb{T}_\circ \rightsquigarrow \mathbb{A}$. For example, $\mathbb{G}_l$, $\mathbb{G}_s$ and $\mathbb{G}_o$ are subgroups of $\mathbb{G}$.

2.3 Group homomorphism and isomorphism

A map from one abstract group to another is called a *group homomorphism* if it is onto and preserves group multiplication (so that the identity maps to the identity, and inverses to inverses), and a *group isomorphism* if, additionally, it is $1-1$. In particular, every group is homomorphic to its *trivial subgroup*, comprising just the identity.

Whereas homomorphic groups are similar, isomorphic groups are identical, differing only in the labelling of their elements. Accordingly, isomorphism is an equivalence relation (to be denoted $\cong$) on the set of all groups.

When no confusion is possible, we may identify isomorphic groups. For example, we may identify $\mathbb{G}$ and $\widetilde{\mathbb{G}}$ via the isomorphism $g \to \widetilde{g}$.

2.4 Orbits

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$. Then $\mathbb{T}$ induces a natural equivalence relation, denoted $\overset{\mathbb{T}}{\sim}$, on $\mathbb{A}$ via $a_1 \overset{\mathbb{T}}{\sim} a_2 \iff \exists t \in \mathbb{T}$ such that $a_2 = t(a_1)$. For each $a \in \mathbb{A}$, its $\mathbb{T}$-*orbit* $\langle a \rangle_{\mathbb{T}} := \{t(a) : t \in \mathbb{T}\}$ is the unique $\overset{\mathbb{T}}{\sim}$ class containing a. In particular, the $\mathbb{T}$-orbits $\{\langle a \rangle_{\mathbb{T}} : a \in \mathbb{A}\}$ partition $\mathbb{A}$. The group $\mathbb{T}$ is called *transitive* (*on* $\mathbb{A}$) if $\mathbb{A}$ comprises a single $\mathbb{T}$-orbit.

For example, $\mathbb{G}_l$ is transitive on $\mathbb{E}^k$, while $N_k(\mu_1, \Sigma_1)$ and $N_k(\mu_2, \Sigma_2)$ in $\mathbb{N}_k := \{$all multivariate normal distributions on $\mathbb{E}^k\}$ are equivalent under $\mathbb{G}_l$ if and only if $\Sigma_2 = \Sigma_1$. These same distributions are equivalent under the full distribution shape group $\mathbb{G}$ if and only if $\Sigma_2 = \kappa^2 Q \Sigma_1 Q^T$ for some $\kappa > 0$ and $Q \in \mathbb{O}_k$. That is, if and only if $(w \circ \lambda)(\Sigma_2) = (w \circ \lambda)(\Sigma_1)$.

2.5 Transformation groups on nonempty subsets

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$ and $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$. Then $\mathbb{A}_{\mathbb{T}}$ is called $\mathbb{T}$-*closed* – or *closed under* $\mathbb{T}$ – if $\langle a_{\mathbb{T}} \rangle_{\mathbb{T}} \subseteq \mathbb{A}_{\mathbb{T}}$ for all $a_{\mathbb{T}} \in \mathbb{A}_{\mathbb{T}}$. For example, each of $\mathbb{F}_k^{\deg}$, $\mathbb{F}_k^{(0)}$, $\mathbb{F}_k^{+}$, $\mathbb{F}_k^{>}$, $\mathbb{F}_k^{+>}$ and $\mathbb{N}_k$ is $\mathbb{G}$-closed.

The $\mathbb{T}$-*closure* of $\mathbb{A}_{\mathbb{T}}$, denoted $\overline{\mathbb{A}_{\mathbb{T}}}$, is defined as the smallest $\mathbb{T}$-closed set containing $\mathbb{A}_{\mathbb{T}}$ – that is, $\overline{\mathbb{A}_{\mathbb{T}}} := \cup\{\langle a_{\mathbb{T}} \rangle_{\mathbb{T}} : a_{\mathbb{T}} \in \mathbb{A}_{\mathbb{T}}\}$ – so that:

$$\mathbb{A}_{\mathbb{T}} \text{ is } \mathbb{T}\text{-closed } \Leftrightarrow \mathbb{A}_{\mathbb{T}} = \overline{\mathbb{A}_{\mathbb{T}}} \Leftrightarrow \mathbb{A}_{\mathbb{T}} \text{ is a union of } \mathbb{T}\text{-orbits.}$$

In what follows, whenever $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$ is considered under $\mathbb{T}$ it is assumed implicitly, if not explicitly, that $\mathbb{A}_{\mathbb{T}}$ is $\mathbb{T}$-closed.

Note that the map $t \to t|_{\mathbb{A}_{\mathbb{T}}}$ need not be $1-1$, even if $\mathbb{A}_{\mathbb{T}}$ is $\mathbb{T}$-closed.

Example 2 *(A simple counterexample)*

Let $\mathbb{A} = \{a_1, a_2, a_3\}$ *and* $\mathbb{T} = \{\mathrm{id}_{\mathbb{A}}, t\}$ *where* $t(a_1) = a_1$, $t(a_2) = a_3$ *and* $t(a_3) = a_2$. *Then* $\{a_1\}$ *is* $\mathbb{T}$*-closed, but* $\mathrm{id}_{\mathbb{A}}|_{\{a_1\}} = \mathrm{id}_{\{a_1\}} = t|_{\{a_1\}}$.

A $\mathbb{T}$-closed set $\mathbb{A}_{\mathbb{T}}$ is called ($\mathbb{T}$-)*regular* if, in fact, $t \to t|_{\mathbb{A}_{\mathbb{T}}}$ *is* $1-1$. Regularity is inherited by subgroups and supersets in the obvious way.

Proposition 3 *(Inheritance of regularity)*

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$ *and* $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$ *be* $\mathbb{T}$*-regular. Then:*

(a) $\mathbb{A}_{\mathbb{T}}$ *is* $\mathbb{T}_\circ$*-regular for every subgroup* $\mathbb{T}_\circ$ *of* $\mathbb{T}$.

(b) $\widetilde{\mathbb{A}}_{\mathbb{T}}$ *is* $\mathbb{T}$*-regular for every* $\mathbb{T}$*-closed* $\widetilde{\mathbb{A}}_{\mathbb{T}}$ *with* $\mathbb{A}_{\mathbb{T}} \subseteq \widetilde{\mathbb{A}}_{\mathbb{T}} \subseteq \mathbb{A}$.

Now, defining $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}} := \{t|_{\mathbb{A}_{\mathbb{T}}} : t \in \mathbb{T}\}$, we have

Proposition 4 *(Homomorphism of* $\mathbb{T}$ *to* $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$*)*

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$ *and* $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$ *be* $\mathbb{T}$*-closed. Then* $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}} \rightsquigarrow \mathbb{A}_{\mathbb{T}}$, $t \to t|_{\mathbb{A}_{\mathbb{T}}}$ *being a group homomorphism. In particular,* $t \to t|_{\mathbb{A}_{\mathbb{T}}}$ *is a group isomorphism if and only if* $\mathbb{A}_{\mathbb{T}}$ *is* $\mathbb{T}$*-regular.*

In the regular case, we may abbreviate $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}} \rightsquigarrow \mathbb{A}_{\mathbb{T}}$ to $\mathbb{T} \rightsquigarrow \mathbb{A}_{\mathbb{T}}$ without confusion.

It turns out that every $\mathbb{G}$-closed subset of $\mathbb{F}_k$ is $\mathbb{G}$-regular. We have

Proposition 5 *(Every* $\mathbb{G}$*-closed subset of* $\mathbb{F}_k$ *is* $\mathbb{G}$*-regular)*

Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ *be* $\mathbb{G}$*-closed. Then* $\mathbb{F}$ *is* $\mathbb{G}$*-regular, and so* $\mathbb{G} \rightsquigarrow \mathbb{F}$.

Proof. Let $g_1 = (\kappa_1, Q_1, x_1), g_2 = (\kappa_2, Q_2, x_2) \in \mathbb{G}$ with $g_1[F] = g_2[F]$, $\forall\, F \in \mathbb{F}$. By transitivity of $\mathbb{G}_l$, $\mu[\mathbb{F}] = \mathbb{E}^k$ so that $\kappa_1 Q_1(\mu - x_1) = \kappa_2 Q_2(\mu - x_2)$ for all $\mu \in \mathbb{E}^k$. Putting $\mu = x_1$ gives $x_1 = x_2$ whence, as required, $g_1 = g_2$. ∎

Corollary 6 *(Examples of* $\mathbb{G}$*-regular sets)*

$\mathbb{F}_k^{\deg}$, $\mathbb{F}_k^{(0)}$, $\mathbb{F}_k^{+}$, $\mathbb{F}_k^{>}$, $\mathbb{F}_k^{+>}$ *and* $\mathbb{N}_k$ *are* $\mathbb{G}$*-regular subsets of* $\mathbb{F}_k$.

2.6 Canonical forms

A subset $\mathbb{A}_\circ$ of $\mathbb{A}$ is called *a set of (*$\mathbb{T}$*-)canonical forms* (*of* $\mathbb{A}$) if it comprises exactly one member of each $\mathbb{T}$-orbit. That is, if for each $a \in \mathbb{A}$ there is a unique $a_\circ \in \mathbb{A}_\circ$ such that $a \overset{\mathbb{T}}{\sim} a_\circ$, in which case $a_\circ$ is called *the* ($\mathbb{T}$-)*canonical form* of a (*in* $\mathbb{A}_\circ$).

For example, $\mathbb{N}_{k:\circ} := \{N_k(0_{\mathbb{E}^k}, \mathrm{diag}(w)) : w \in \mathbb{W}_k\}$ is a set of $\mathbb{G}$-canonical forms of $\mathbb{N}_k$. However, $\{F \in \mathbb{F}_k : \mu[F] = 0_{\mathbb{E}^k}, \Sigma[F] = \mathrm{diag}(w), w \in \mathbb{W}_k\}$ is *not* a set of $\mathbb{G}$-canonical forms of $\mathbb{F}_k$ since, for each $w \in \mathbb{W}_k$, there are infinitely many $F \in \mathbb{F}_k$ with $w[F] = w$.

2.7 Invariants and complete invariants

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$. Then $\iota \in \mathrm{dom}(\mathbb{A})$ is called *a* ($\mathbb{T}$-)*invariant* (*on* $\mathbb{A}$) if, for all $a \in \mathbb{A}$ and $t \in \mathbb{T}$, $\iota(t(a)) = \iota(a)$. That is, if ι is constant on $\mathbb{T}$-orbits. A constant function on $\mathbb{A}$ is called a *trivial* invariant.

Note that there is no constraint on the image space of an invariant, which is usually chosen with convenience in mind. We call a real-valued invariant an invariant *index*.

A $\mathbb{T}$-invariant ι on $\mathbb{A}$ is called *complete* if $\iota(a_1) = \iota(a_2) \Rightarrow \langle a_1 \rangle_{\mathbb{T}} = \langle a_2 \rangle_{\mathbb{T}}$. That is, if ι takes different values on different $\mathbb{T}$-orbits. Thus, the value $\iota(a)$ of a complete $\mathbb{T}$-invariant completely specifies which $\mathbb{T}$-orbit a belongs to. All complete invariants are equivalent in the sense that their sets of constancy coincide.

For example, $\Sigma[\cdot]$ is a $\mathbb{G}_l$-invariant on $\mathbb{F}_k$, while $\mathrm{rank}(\Sigma[\cdot])$ is a $\mathbb{G}$-invariant there, but neither is complete. Again, $w[\cdot]$ is a complete $\mathbb{G}$-invariant on $\mathbb{N}_k$, but not on $\mathbb{F}_k$.

A complete invariant always exists, albeit implicitly. For, by the Axiom of Choice, a set of $\mathbb{T}$-canonical forms of $\mathbb{A}$ always exists while, for any such set $\mathbb{A}_\circ$, $\iota_{\mathbb{A}_\circ} : a \to a_\circ$ is called the corresponding *canonical* complete invariant. However, an *explicit* canonical complete invariant is not, in general, available.

For example, whereas $\iota_{\mathbb{N}_{k:\circ}}$ is explicit for $\mathbb{G}$ on $\mathbb{N}_k$, we lack an explicit complete $\mathbb{G}$-invariant on $\mathbb{F}_k$. It is precisely this lack which motivates the study of filtered invariants below.

2.8 Inheritance of invariance properties by closed subsets

A closed subset inherits invariance properties as follows.

Proposition 7 *(Transformation groups on subsets)*
Let $\mathbb{T} \rightsquigarrow \mathbb{A}$, $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$ be $\mathbb{T}$-closed, ι be a $\mathbb{T}$-invariant on $\mathbb{A}$ and $\mathbb{A}_\circ$ be a set of $\mathbb{T}$-canonical forms of $\mathbb{A}$. Then:
(a) $\iota|_{\mathbb{A}_{\mathbb{T}}}$ is a $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$-invariant on $\mathbb{A}_{\mathbb{T}}$, being complete whenever ι is.
(b) $\mathbb{A}_\circ \cap \mathbb{A}_{\mathbb{T}}$ is a set of $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$-canonical forms of $\mathbb{A}_{\mathbb{T}}$.

For example, $\mathrm{rank}(\Sigma[\cdot])$ is a $\mathbb{G}$-invariant on each of the $\mathbb{G}$-closed subsets $\mathbb{F}_k^{\mathrm{deg}}$, $\mathbb{F}_k^{(0)}$, $\mathbb{F}_k^{+}$, $\mathbb{F}_k^{>}$ and $\mathbb{F}_k^{+>}$ of $\mathbb{F}_k$. Again, the singleton $\{N_k(0_{\mathbb{E}^k}, k^{-1}I_k)\}$ is a set of $\mathbb{G}$-canonical forms of $\mathbb{N}_k^I := \{N_k(\mu, \sigma^2 I_k) : \mu \in \mathbb{E}^k, \sigma^2 > 0\}$, the restriction of $\iota_{\mathbb{N}_{k:\circ}}$ to $\mathbb{N}_k^I$ being the corresponding canonical complete invariant.

2.9 Inheritance of invariance properties by subgroups

Not all invariance properties are inherited by subgroups.

Lemma 8 *(Orbits under groups and subgroups)*
Let $\mathbb{T} \rightsquigarrow \mathbb{A}$ and $\mathbb{T}_\circ$ be a subgroup of $\mathbb{T}$. Then, every $\mathbb{T}$-orbit is a union of $\mathbb{T}_\circ$-orbits. That is, $a_2 \in \langle a_1 \rangle_{\mathbb{T}} \Rightarrow \langle a_2 \rangle_{\mathbb{T}_\circ} \subseteq \langle a_1 \rangle_{\mathbb{T}}$.

Proposition 9 *(Transformation subgroups)*

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$*,* $\mathbb{T}_\circ$ *be a subgroup of* $\mathbb{T}$ *and* $\mathbb{A}_\circ$ *be a subset of* $\mathbb{A}$*. Then:*

(a) ι *is a* $\mathbb{T}$*-invariant on* $\mathbb{A} \Rightarrow \iota$ *is a* $\mathbb{T}_\circ$*-invariant on* $\mathbb{A}$*.*

However:

(b) $\mathbb{A}_\circ$ *is a set of* $\mathbb{T}$*-canonical forms* $\not\Rightarrow \mathbb{A}_\circ$ *is a set of* $\mathbb{T}_\circ$*-canonical forms.*

(c) ι *is a complete* $\mathbb{T}$*-invariant on* $\mathbb{A} \not\Rightarrow \iota$ *is a complete* $\mathbb{T}_\circ$*-invariant on* $\mathbb{A}$*.*

Proof. Lemma 8 implies (a) at once. For (b) and (c), let $\mathbb{T} \rightsquigarrow \mathbb{A}$ as in Example 2, $\mathbb{T}_\circ = \mathbb{T}_{\mathrm{id}_{\mathbb{A}}}$, $\mathbb{A}_\circ = \{a_1, a_2\}$ and $\iota = \iota_{\mathbb{A}_\circ}$. ■

For example, $\iota_{\mathbb{N}_{k:\circ}}$ is a $\mathbb{G}_l$-invariant – but not a complete $\mathbb{G}_l$-invariant – on $\mathbb{N}_k$. Again, $\mathbb{N}_{k:\circ}$ is not a set of $\mathbb{G}_l$-canonical forms of $\mathbb{N}_k$.

2.10 Characterisation of invariants

Invariants can be characterised in terms of filtered functions, as follows.

Lemma 10 *(Necessary and sufficient condition for* ξ *to be* ϕ*-filtered)*

Let $\xi, \phi \in \mathrm{dom}(\mathbb{A})$*. The following statements are equivalent:*

[1] ξ *is* ϕ*-filtered.*

[2] $\phi(a_1) = \phi(a_2) \Rightarrow \xi(a_1) = \xi(a_2)$*.*

Clearly, any function of an invariant is an invariant. In particular, any function of a complete invariant is an invariant. In fact, *every* invariant arises in this latter way. Lemma 10 gives at once the well-known

Theorem 11 *(Characterisation of invariants)*

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$*,* ι_c *be a complete* $\mathbb{T}$*-invariant on* $\mathbb{A}$ *and* $\iota \in \mathrm{dom}(\mathbb{A})$*. Then:*

$$\iota \text{ is a } \mathbb{T}\text{-invariant on } \mathbb{A} \Leftrightarrow \iota \text{ is } \iota_c\text{-filtered.}$$

For example, $\iota \in \mathrm{dom}(\mathbb{N}_k)$ is a $\mathbb{G}$-invariant on $\mathbb{N}_k \Leftrightarrow \iota$ is w-filtered.

Thus, the relationship between invariants and complete invariants mirrors that between sufficient statistics and minimal sufficient statistics.

2.11 Equivariants

Equivariants are defined in terms of filtered functions. If $\mathbb{T} \rightsquigarrow \mathbb{A}$, $\phi \in \mathrm{dom}(\mathbb{A})$ and $t \in \mathbb{T}$, we say that ϕ *is a* t*-equivariant (on* $\mathbb{A}$*)* if $\phi \circ t$ is ϕ-filtered. That is, if the diagram

$$\begin{array}{ccc} a & \rightarrow & t(a) \\ \downarrow & & \downarrow \\ \phi(a) & \dashrightarrow & \phi(t(a)) \end{array}$$

can be completed by a (necessarily unique) function ${}^{\phi}t : \phi(\mathbb{A}) \rightarrow \phi(\mathbb{A})$ such that $\phi \circ t = {}^{\phi}t \circ \phi$. Lemma 10 gives at once

Proposition 12 *(Characterisation of* t*-equivariants)*

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$*,* $\phi \in \mathrm{dom}(\mathbb{A})$ *and* $t \in \mathbb{T}$*. The following statements are equivalent:*

[1] ϕ *is a* t*-equivariant.*

[2] $\phi(a_1) = \phi(a_2) \Rightarrow \phi(t(a_1)) = \phi(t(a_2))$*.*

We call ϕ *a* $\mathbb{T}$*-equivariant (on* $\mathbb{A}$*)* if ϕ is a t-equivariant (on $\mathbb{A}$) for each $t \in \mathbb{T}$.

Equivariance is inherited by subgroups and subsets in the obvious way.

Proposition 13 *(Inheritance of equivariance)*

Let ϕ *be a* $\mathbb{T}$*-equivariant on* $\mathbb{A}$*. Then:*

(a) For any subgroup $\mathbb{T}_\circ$ *of* $\mathbb{T}$*,* ϕ *is a* $\mathbb{T}_\circ$*-equivariant on* $\mathbb{A}$*.*

(b) For any $\mathbb{T}$*-closed* $\varnothing \subset \mathbb{A}_\mathbb{T} \subseteq \mathbb{A}$*,* $\phi|_{\mathbb{A}_\mathbb{T}}$ *is a* $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$*-equivariant on* $\mathbb{A}_\mathbb{T}$*, with*

$$^{\phi|_{\mathbb{A}_\mathbb{T}}}(t|_{\mathbb{A}_\mathbb{T}}) = (^{\phi}t)|_{\phi(\mathbb{A}_\mathbb{T})}. \tag{4}$$

If $\mathbb{A}_\mathbb{T}$ is $\mathbb{T}$-regular, a $\mathbb{T}$-equivariant ϕ on $\mathbb{A}$ may also, without confusion, be called a $\mathbb{T}$-equivariant on $\mathbb{A}_\mathbb{T}$, and either side of (4) abbreviated to $^{\phi}t$.

For any $\mathbb{T}$-equivariant ϕ on $\mathbb{A}$, $^{\phi}\mathbb{T} := \{^{\phi}t : t \in \mathbb{T}\}$ has the following properties.

Theorem 14 *(Properties of* $^{\phi}\mathbb{T}$*)*

Let ϕ *be a* $\mathbb{T}$*-equivariant on* $\mathbb{A}$*. Then:*

(a) $^{\phi}\mathbb{T} \rightsquigarrow \phi(\mathbb{A})$*,* $t \to {}^{\phi}t$ *being a group homomorphism of* $\mathbb{T}$ *to* $^{\phi}\mathbb{T}$*.*

(b) For any $\psi \in \mathrm{dom}(\phi(\mathbb{A}))$*:*

$$\psi \textit{ is a } {}^{\phi}\mathbb{T}\textit{-equivariant on } \phi(\mathbb{A}) \Leftrightarrow \psi \circ \phi \textit{ is a } \mathbb{T}\textit{-equivariant on } \mathbb{A}$$

in which case, for each $t \in \mathbb{T}$*,* $^{\psi}(^{\phi}t) = {}^{\psi\circ\phi}t$*, so that* $^{\psi}(^{\phi}\mathbb{T}) = {}^{\psi\circ\phi}\mathbb{T}$ *and the alternative commutative completions shown in the following diagram are equivalent:*

$$\begin{array}{ccccccc}
\longleftarrow & & \mathbb{A} & \overset{t}{\to} & \mathbb{A} & & \longrightarrow \\
| & & \phi\downarrow & & \phi\downarrow & & | \\
| & & \phi(\mathbb{A}) & \overset{^{\phi}t}{\to} & \phi(\mathbb{A}) & & | \\
\psi\circ\phi\downarrow & & \psi\downarrow & & \psi\downarrow & & \downarrow\psi\circ\phi \\
| & & \psi(\phi(\mathbb{A})) & \overset{^{\psi}(^{\phi}t)}{\dashrightarrow} & \psi(\phi(\mathbb{A})) & & | \\
| & & ||| & & ||| & & | \\
\longrightarrow & & (\psi\circ\phi)(\mathbb{A}) & \overset{^{\psi\circ\phi}t}{\dashrightarrow} & (\psi\circ\phi)(\mathbb{A}) & & \longleftarrow
\end{array}$$

Finally, note that when $^{\phi}\mathbb{T}$ is the trivial group, $\mathbb{T}$-equivariance reduces to $\mathbb{T}$-invariance. In this sense, equivariance is the more general concept.

The next section summarises key equivariance properties of the distribution shape maps $\mu[\cdot], \Sigma[\cdot], \lambda(\cdot)$ and $w(\cdot)$, illustrating Proposition 13 and Theorem 14.

2.12 Equivariance and distribution shape

Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ and recall (Proposition 5) that $\mathbb{G} \rightsquigarrow \mathbb{F}$ if and only if $\mathbb{F}$ is $\mathbb{G}$-closed, so that restricting attention now to such subsets of $\mathbb{F}_k$ is without loss.

By linearity of expectation, we have at once

Proposition 15 *(Equivariance properties of $\mu[\cdot]$)*
Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ be $\mathbb{G}$-closed. Then, $\mu[g[F]] = g[\mu[F]]$, $(g \in \mathbb{G}, F \in \mathbb{F})$ so that $\mu[\cdot]$ is a $\mathbb{G}$-equivariant on $\mathbb{F}$ and ${}^{\mu}\mathbb{G} \rightsquigarrow \mathbb{E}^k$, $g \to {}^{\mu}g$ being a group isomorphism of $\mathbb{G}$ to ${}^{\mu}\mathbb{G}$.

Again, abbreviating ${}^{\lambda\circ\Sigma}g$ to ${}^{\lambda}g$ and ${}^{w\circ\lambda\circ\Sigma}g$ to ${}^{w}g$ without confusion, we have

Proposition 16 *(Equivariance properties of distribution shape maps)*
Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ be $\mathbb{G}$-closed and $g = (\kappa, Q, x_\circ) \in \mathbb{G}$. Then:

(a) (i) $\Sigma[\cdot]$ is a $\mathbb{G}$-equivariant on $\mathbb{F}$, with ${}^{\Sigma}g : \Sigma \to \kappa^2 Q\Sigma Q^T$.
(ii) ${}^{\Sigma}\mathbb{G} := \{{}^{\Sigma}g : g \in \mathbb{G}\} \rightsquigarrow \mathbb{D} := \Sigma[\mathbb{F}]$,
$g \to {}^{\Sigma}g$ being a group homomorphism of $\mathbb{G}$ to ${}^{\Sigma}\mathbb{G}$.

(b) (i) $\lambda(\cdot)$ is a ${}^{\Sigma}\mathbb{G}$-equivariant on $\mathbb{D}$, with ${}^{\lambda}g : \lambda \to \kappa^2\lambda$.
(ii) ${}^{\lambda}\mathbb{G} := \{{}^{\lambda}g : g \in \mathbb{G}\} \rightsquigarrow \mathbb{L} := \lambda(\mathbb{D})$,
$g \to {}^{\lambda}g$ being a group homomorphism of $\mathbb{G}$ to ${}^{\lambda}\mathbb{G}$.

(c) (i) $w(\cdot)$ is a ${}^{\lambda}\mathbb{G}$-equivariant on $\mathbb{L}$, with ${}^{w}g = \mathrm{id}_{\mathbb{W}}$.
(ii) ${}^{w}\mathbb{G} := \{\mathrm{id}_{\mathbb{W}}\} \rightsquigarrow \mathbb{W} := w(\mathbb{L})$,
$g \to {}^{w}g$ being a group homomorphism of $\mathbb{G}$ to ${}^{w}\mathbb{G}$.

Using Proposition 13, we may replace $\mathbb{G}$ in Proposition 16 by any of its subgroups, making any induced changes to ${}^{\Sigma}\mathbb{G}$ or ${}^{\lambda}\mathbb{G}$. For example,

$$\begin{array}{lcl} \text{if } g_o & = & (1, Q, 0_{\mathbb{E}^k}) \in \mathbb{G}_o,\ {}^{\Sigma}g_o : \Sigma \to Q\Sigma Q^T \text{ and } {}^{\lambda}g_o = \mathrm{id}_{\mathbb{L}} \text{ while,} \\ \text{if } g_s & = & (\kappa, I_k, 0_{\mathbb{E}^k}) \in \mathbb{G}_s,\ {}^{\Sigma}g_s : \Sigma \to \kappa^2\Sigma,\ {}^{\lambda}g_s = {}^{\lambda}g \text{ remaining unchanged.} \end{array}$$

Overall, we have the commutative diagram:

$$\begin{array}{ccc} \mathbb{F} & \stackrel{g}{\to} & \mathbb{F} \\ \Sigma\downarrow & & \Sigma\downarrow \\ \mathbb{D} & \stackrel{{}^{\Sigma}g}{\to} & \mathbb{D} \\ \lambda\downarrow & & \lambda\downarrow \\ \mathbb{L} & \stackrel{{}^{\lambda}g}{\to} & \mathbb{L} \\ & \stackrel{w}{\searrow} \quad \stackrel{w}{\swarrow} & \\ & \mathbb{W} & \end{array}$$

None of the maps $g \to {}^{\Sigma}g$, $g \to {}^{\lambda}g$ and $g \to {}^{w}g$ are *iso*morphisms – their images being independent of $x_\circ$, of Q and of κ respectively. Accordingly, $\mathbb{G}$, ${}^{\Sigma}\mathbb{G}$, ${}^{\lambda}\mathbb{G}$ and ${}^{w}\mathbb{G}$ are progressively smaller groups, corresponding to the fact that $\Sigma[\cdot]$ takes out location (and overall sign: $\text{cov}(X) = \text{cov}(-X)$), $\lambda(\cdot)$ takes out (all other) orthogonal transformations and $w(\cdot)$ takes out scale. That the final induced group ${}^{w}\mathbb{G}$ is trivial mirrors the fact that each $g = (\kappa, Q, x_\circ) \in \mathbb{G}$ has a unique decomposition as

$$g = g_s \circ g_o \circ g_l, (g_s \in \mathbb{G}_s, g_o \in \mathbb{G}_o, g_l \in \mathbb{G}_l) \tag{5}$$

given by $g_l = (1, I_k, x_\circ)$, $g_o = (1, Q, 0_{\mathbb{E}^k})$ and $g_s = (\kappa, I_k, 0_{\mathbb{E}^k})$.

Again, illustrating the reduction of equivariance to invariance noted above, the fact that ${}^{w}\mathbb{G}$ acts trivially on $\mathbb{W}$ corresponds to w being a ${}^{\lambda}\mathbb{G}$-invariant on $\mathbb{L}$. Indeed, we have:

Proposition 17 *(Complete invariants for $\mathbb{D}$, $\mathbb{L}$ and $\mathbb{W}$)*
Let $\mathbb{F}$, $\mathbb{D}$, $\mathbb{L}$ and $\mathbb{W}$ be as in Proposition 16. Then:

(a) $w \circ \lambda$ is a complete ${}^{\Sigma}\mathbb{G}$-invariant on $\mathbb{D}$ so that

$$\iota \in \text{dom}(\mathbb{D}) \textit{ is a } {}^{\Sigma}\mathbb{G}\textit{-invariant on } \mathbb{D} \Leftrightarrow \iota \textit{ is } (w \circ \lambda)\textit{-filtered.}$$

(b) w is a complete ${}^{\lambda}\mathbb{G}$-invariant on $\mathbb{L}$ so that

$$\iota \in \text{dom}(\mathbb{L}) \textit{ is a } {}^{\lambda}\mathbb{G}\textit{-invariant on } \mathbb{L} \Leftrightarrow \iota \textit{ is } w\textit{-filtered.}$$

(c) $\text{id}_{\mathbb{W}}$ is a complete ${}^{w}\mathbb{G}$-invariant on $\mathbb{W}$ so that

$$\textit{every } \iota \in \text{dom}(\mathbb{W}) \textit{ is a } {}^{w}\mathbb{G}\textit{-invariant on } \mathbb{W}.$$

3 FILTERED INVARIANTS AND DISPERSION SHAPE

3.1 Distribution shape indices

Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ be $\mathbb{G}$-closed, so that $\mathbb{G} \rightsquigarrow \mathbb{F}$. *A distribution shape measure* on $\mathbb{F}$ is defined as a $\mathbb{G}$-invariant on $\mathbb{F}$, and a distribution shape *index* as a real-valued distribution shape measure.

For particular subsets $\mathbb{F}$, an explicit complete $\mathbb{G}$-invariant exists. In this case Theorem 11 describes *all* distribution shape measures on $\mathbb{F}$. For example, the distribution shape measures on $\mathbb{N}_k$ *are* the w-filtered members of $\text{dom}(\mathbb{N}_k)$. However, as noted above, such explicit complete $\mathbb{G}$-invariants do not exist in general, and we therefore study appropriately filtered $\mathbb{G}$-invariants. These arise as the special case $\mathbb{T} = \mathbb{G}$ and $\mathbb{A} = \mathbb{F}$ of the following theorem.

3.2 Filtered invariants

Let $\mathbb{T} \rightsquigarrow \mathbb{A}$. Then ϕ-filtered $\mathbb{T}$-invariants exist for every $\phi \in \mathrm{dom}(\mathbb{A})$, trivial invariants being obvious examples. In considering ϕ-filtered $\mathbb{T}$-invariants, it is natural to restrict attention to functions ϕ which are $\mathbb{T}$-equivariants. These filtered invariants can be characterised as follows.

Theorem 18 *(Characterisation of ϕ-filtered $\mathbb{T}$-invariants)*
Let ϕ be a $\mathbb{T}$-equivariant on $\mathbb{A}$, $\iota \in \mathrm{dom}(\mathbb{A})$ and ζ_c be a complete ${}^{\phi}\mathbb{T}$-invariant on $\phi(\mathbb{A})$. Then:
ι is a ϕ-filtered $\mathbb{T}$-invariant on $\mathbb{A}$
$\Leftrightarrow$ $\iota = \zeta \circ \phi$ for some ${}^{\phi}\mathbb{T}$-invariant ζ on $\phi(\mathbb{A})$
$\Leftrightarrow$ ι is $(\zeta_c \circ \phi)$-filtered,
in which case ζ is complete whenever ι is.

Proof. By definition,
ι is a ϕ-filtered $\mathbb{T}$-invariant on $\mathbb{A}$
$\Leftrightarrow$ $[\iota = \zeta \circ \phi$ for some $\zeta \in \mathrm{dom}(\phi(\mathbb{A}))$ and $\forall t \in \mathbb{T}, \zeta \circ \phi = \zeta \circ \phi \circ t]$,
from which the stated equivalences follow, the second using Theorem 11.

Suppose now that $\iota = \zeta \circ \phi$ is a complete $\mathbb{T}$-invariant on $\mathbb{A}$ and let $\zeta(\phi(a_1)) = \zeta(\phi(a_2))$. Then $a_2 = t(a_1)$ for some $t \in \mathbb{T}$ so that, as required, $\phi(a_2) = {}^{\phi}t(\phi(a_1))$ for some ${}^{\phi}t \in {}^{\phi}\mathbb{T}$. ■

Theorem 18 reduces the study of ϕ-filtered $\mathbb{T}$-invariants to that of ${}^{\phi}\mathbb{T}$-invariants, describing all of them whenever a complete ${}^{\phi}\mathbb{T}$-invariant can be made explicit. This is usually easier than making a complete $\mathbb{T}$-invariant explicit, precisely because $t \to {}^{\phi}t$ is in general many-to-one, so that ${}^{\phi}\mathbb{T}$ is correspondingly smaller than $\mathbb{T}$ (as, for example, in Proposition 16).

3.3 Dispersion shape

Distribution shape measures can depend on many different features of F, low order moments being of natural interest. In particular, $\mu[\cdot]$ and $\Sigma[\cdot]$ are both widely used distributional summaries, with the right equivariance property.

However,

Proposition 19 *(μ-filtered distribution shape measures are constant)*
There are no nontrivial μ-filtered distribution shape measures.

Proof. Let $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$ be $\mathbb{G}$-closed, $\iota := \zeta \circ \mu$ be a μ-filtered $\mathbb{G}$-invariant on $\mathbb{F}$ and $F \in \mathbb{F}$. Then:
$\iota[F]$
$= \zeta(\mu[g[F]]), (\forall g \in \mathbb{G})$, as ι is a $\mathbb{G}$-invariant
$= \zeta(g(\mu[F])), (\forall g \in \mathbb{G})$, by linearity of expectation
$= \zeta(0_{\mathbb{E}^k})$, by transitivity of $\mathbb{G}_l$. ■

In view of Proposition 19, we focus now on *dispersion shape measures* defined as Σ-filtered distribution shape measures. Again, dispersion shape *indices* are just real-valued dispersion shape measures.

Dispersion shape measures can be conveniently characterised as follows. Theorem 18 together with Propositions 16(a) and 17(a), give at once (Proposition I above):

Theorem 20 *(Characterisation of dispersion shape measures)*
Let $\mathbb{G} \rightsquigarrow \mathbb{F}$ and $\iota \in \mathrm{dom}(\mathbb{F})$. Then the following two statements are equivalent:
[1] ι is a dispersion shape measure on $\mathbb{F}$.
[2] $\iota = \zeta \circ w$ for some $\zeta \in \mathrm{dom}(w[\mathbb{F}])$.
In particular, a nontrivial Σ-filtered distribution shape measure on $\mathbb{F}$ exists if and only if $k > 1$.

Indices of dispersion shape therefore include the following familiar quantities.

Corollary 21 *(Examples of dispersion shape indices)*
(a) For each $r = 1, ..., k$, the proportion $w_r[F]$ of total variance explained by the r^{th} principal component is a dispersion shape index on $\mathbb{F}_k^>$.
(b) For each $r = 1, ..., k$, the proportion $w_1[F]+...+w_r[F]$ (respectively, $w_{k-r+1}[F]+ ... + w_k[F]$) of total variance explained by the first (respectively, last) r principal components is a dispersion shape index on $\mathbb{F}_k^>$.
(c) Defining $\iota_\circ \in \mathrm{dom}(\mathbb{F}_k^+)$ by $2\log \iota_\circ[F] := nk \log \left(\frac{k^{-1}\,\mathrm{trace}\,\Sigma[F]}{\sqrt[k]{\det \Sigma[F]}} \right)$, $\iota_\circ$ is a dispersion shape index on $\mathbb{F}_k^+$, while $\iota_\circ[\widehat{F}]$ is the likelihood ratio test statistic for dispersion sphericity based on a multivariate normal n-sample, ($\widehat{F} \in \mathbb{F}_k^+$ requiring $n > k$).

By definition, every dispersion shape measure on $\mathbb{F}$ is a distribution shape measure on $\mathbb{F}$. Whereas the converse is false in general, conditions under which it does hold are studied in Appendix A. In particular, every distribution shape measure on $\mathbb{N}_k$ is shown there to be a dispersion shape measure on $\mathbb{N}_k$.

4 SCALE-ISOVARIANTS AND DISTRIBUTION SIZE

4.1 Isovariants

A $\mathbb{T}$-equivariant ϕ (on $\mathbb{A}$) is called *a $\mathbb{T}$-isovariant (on $\mathbb{A}$)* if $t \to {}^{\phi}t$ is a group isomorphism of $\mathbb{T}$ to ${}^{\phi}\mathbb{T}$. Equivalently, if $t \to {}^{\phi}t$ is $1-1$. For example, $\mu[\cdot]$ is a $\mathbb{G}$-isovariant, but the $\mathbb{G}$-equivariant $\Sigma[\cdot]$ is not.

To signal this distinction between $\mathbb{T}$-equivariants in general and $\mathbb{T}$-isovariants in particular, we denote the latter by η rather than ϕ, wherever convenient.

Isovariance is inherited directly by subgroups and, under appropriate conditions, by subsets (Proposition 39, Appendix B). When $\mathbb{A}_\mathbb{T}$ is $\mathbb{T}$-regular, so that $\mathbb{T} \rightsquigarrow \mathbb{A}_\mathbb{T}$, there is no confusion in abbreviating '$\eta|_{\mathbb{A}_\mathbb{T}}$ is a $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$-isovariant on $\mathbb{A}_\mathbb{T}$' to 'η is a $\mathbb{T}$-isovariant on $\mathbb{A}_\mathbb{T}$'.

Isovariants play a key rôle here, *scale-isovariants* being a special case of intrinsic interest in size analysis arising in connection with *scale transformation* (multiplication by a positive constant) between *cones* (nonempty sets closed under such transformations). Their defining characteristic is intuitive: changing scale by $\kappa > 0$ induces their multiplicative response by a factor κ^l which, as $l > 0$, increases with κ. Positive-valued scale-isovariants are of particular interest.

There are natural ways to make these ideas precise in nontrivial real vector spaces, and more widely, as follows.

4.2 Scale-isovariants on nontrivial cones in real vector spaces

Let $\mathbb{S}$ denote the group of all scale transformations $a \to \kappa a$, $(a > 0,\ \kappa > 0)$ on $(0, \infty)$. Using the obvious isomorphism, we identify $\mathbb{S}$ with the abstract group of positive reals under scalar multiplication, and denote its general member by $\kappa > 0$.

Again, let $\mathbb{V}$ denote a real vector space with zero element $0_{\mathbb{V}}$. We call $\{0_{\mathbb{V}}\}$ the *trivial* nonempty subset of $\mathbb{V}$, and any other nonempty subset of $\mathbb{V}$ *nontrivial*.

Supposing now that $\mathbb{V}$ itself is nontrivial, and putting $\kappa_{\mathbb{V}} : v \to \kappa v$ $(v \in \mathbb{V})$, $\kappa \to \kappa_{\mathbb{V}}$ is a group isomorphism of $\mathbb{S}$ to $\mathbb{S}_{\mathbb{V}} := \{\kappa_{\mathbb{V}} : \kappa > 0\}$ under which $\mathbb{S} \rightsquigarrow \mathbb{V}$. Defining a cone in $\mathbb{V}$ as a nonempty $\mathbb{S}$-closed subset of $\mathbb{V}$, we have at once

Proposition 22 *(Characterisation of $\mathbb{S}$-regular cones in $\mathbb{V}$)*
A cone $\mathbb{V}_{\mathbb{S}}$ in $\mathbb{V}$ is $\mathbb{S}$-regular if and only if it is nontrivial, in which case $\mathbb{S} \rightsquigarrow \mathbb{V}_{\mathbb{S}}$.

Let now $l > 0$ and $\mathbb{V}$ and $\mathbb{V}_*$ be nontrivial real vector spaces, with $\mathbb{V}_{\mathbb{S}}$ a nontrivial cone in $\mathbb{V}$. Then, $\eta : \mathbb{V}_{\mathbb{S}} \to \mathbb{V}_*$ is called *an* (l^{th} *order*) *scale-isovariant* (*from* $\mathbb{V}_{\mathbb{S}}$ *to* $\mathbb{V}_*$) if η is an $\mathbb{S}$-isovariant on $\mathbb{V}_{\mathbb{S}}$ with ${}^{\eta}\kappa : v_* \to \kappa^l v_*$ and we have

Proposition 23 *(Scale-isovariants on nontrivial cones in real vector spaces)*
Let $l > 0$, $\mathbb{V}$ and $\mathbb{V}_*$ be nontrivial real vector spaces, $\mathbb{V}_{\mathbb{S}}$ be a nontrivial cone in $\mathbb{V}$, and $\eta : \mathbb{V}_{\mathbb{S}} \to \mathbb{V}_*$. Again, let $\mathbb{V}'_{\mathbb{S}}$ be a nontrivial subcone of $\mathbb{V}_{\mathbb{S}}$. Then:

$$\begin{aligned} & \eta \text{ is an } l^{th} \text{ order scale-isovariant} \\ \Leftrightarrow\ & \eta(\mathbb{V}_{\mathbb{S}}) \text{ is a nontrivial cone in } \mathbb{V}_*, \text{ with } \eta(\kappa v) = \kappa^l \eta(v), \end{aligned}$$

in which case:
(a) $\mathbb{S} \rightsquigarrow \eta(\mathbb{V}_{\mathbb{S}})$.
(b) The order of η is well-defined
(that is, $\left[\eta \text{ is a scale-isovariant of order } l \text{ and of order } \widetilde{l}\right] \Rightarrow l = \widetilde{l}$).
(c) $\eta|_{\mathbb{V}'_{\mathbb{S}}}$ is an l^{th} order scale-isovariant $\Leftrightarrow \eta(\mathbb{V}'_{\mathbb{S}})$ is a nontrivial cone in $\mathbb{V}_*$.

4.3 Distribution size measures on nontrivial cones in $\mathbb{F}_k$

More widely, $\kappa \to (\kappa, I_k, 0_{\mathbb{E}^k})$ is an isomorphism of $\mathbb{S}$ to $\mathbb{G}_s$, under which $\mathbb{S} \rightsquigarrow \mathbb{F}_k$ and we may put $\kappa F := (\kappa, I_k, 0_{\mathbb{E}^k})[F]$ without confusion. Defining a cone in $\mathbb{F}_k$ to be a nonempty $\mathbb{S}$-closed subset of $\mathbb{F}_k$, calling $\{\widehat{F}_{0_{\mathbb{E}^k}}\}$ the *trivial* nonempty subset of $\mathbb{F}_k$ and any other nonempty subset of $\mathbb{F}_k$ *nontrivial*, we have

Proposition 24 *(Characterisation of $\mathbb{S}$-regular cones in $\mathbb{F}_k$)*
Let $\mathbb{F}_{\mathbb{S}}$ be a cone in $\mathbb{F}_k$. Then:

$$\begin{array}{llll} & \mathbb{F}_{\mathbb{S}} \textit{ is } \mathbb{S}\textit{-regular} & \Leftrightarrow & \mathbb{F}_{\mathbb{S}} \textit{ is nontrivial} \\ \Leftrightarrow & [\mu[\mathbb{F}_{\mathbb{S}}] \textit{ or } \Sigma[\mathbb{F}_{\mathbb{S}}] \textit{ is nontrivial}] & \Leftrightarrow & [\mu[\mathbb{F}_{\mathbb{S}}] \neq \{0_{\mathbb{E}^k}\} \textit{ or } \mathbb{F}_{\mathbb{S}} \not\subseteq \mathbb{F}_k^{\text{deg}}], \end{array}$$

in which case $\mathbb{S} \rightsquigarrow \mathbb{F}_{\mathbb{S}}$.

Let now $l > 0$, $\mathbb{F}_{\mathbb{S}}$ be a nontrivial cone in $\mathbb{F}_k$, and $\mathbb{V}$ be a nontrivial real vector space. Then, $\eta : \mathbb{F}_{\mathbb{S}} \to \mathbb{V}$ is called *a distribution size measure (on $\mathbb{F}_{\mathbb{S}}$)* – interchangeably, *a scale-isovariant* (*from* $\mathbb{F}_{\mathbb{S}}$ *to* $\mathbb{V}$) – (*of order* l) if η is an $\mathbb{S}$-isovariant on $\mathbb{F}_{\mathbb{S}}$, with ${}^{\eta}\kappa : v \to \kappa^l v$.

The parallel between Proposition 23 and Proposition 25 below reflects the fact that both are special cases of a single, more general result (Theorem 42, Appendix C).

Proposition 25 *(Distribution size measures on nontrivial cones in $\mathbb{F}_k$)*
Let $l > 0$, $\mathbb{F}_{\mathbb{S}}$ be a nontrivial cone in $\mathbb{F}_k$, $\mathbb{V}$ be a nontrivial real vector space, and $\eta : \mathbb{F}_{\mathbb{S}} \to \mathbb{V}$. Again, let $\mathbb{F}'_{\mathbb{S}}$ be a nontrivial subcone of $\mathbb{F}_{\mathbb{S}}$. Then:

$$\begin{array}{ll} & \eta \textit{ is an } l^{th}\textit{-order distribution size measure} \\ \Leftrightarrow & \eta[\mathbb{F}_{\mathbb{S}}] \textit{ is a nontrivial cone in } \mathbb{V}, \textit{ with } \eta[\kappa F] = \kappa^l \eta[F], \end{array}$$

in which case:
(a) $\mathbb{S} \rightsquigarrow \eta[\mathbb{F}_{\mathbb{S}}]$.
(b) The order of η is well-defined.
(c) $\eta|_{\mathbb{F}'_{\mathbb{S}}}$ *is an l^{th}-order distribution size measure*
$\Leftrightarrow \eta[\mathbb{F}'_{\mathbb{S}}]$ *is a nontrivial cone in $\mathbb{V}$.*

For example, with $\mathbb{F}_{\mathbb{S}}$ denoting a nontrivial cone in $\mathbb{F}_k$, $\mu[\cdot]$ is a first-order distribution size measure on $\mathbb{F}_{\mathbb{S}}$ iff $\mu[\mathbb{F}_{\mathbb{S}}]$ is nontrivial, while $\Sigma[\cdot]$ is a second-order distribution size measure there iff $\mathbb{F}_{\mathbb{S}} \not\subseteq \mathbb{F}_k^{\text{deg}}$. More generally, any l^{th} (central) moment η defined throughout $\mathbb{F}_{\mathbb{S}}$ is an l^{th}-order distribution size measure there iff $\eta[\mathbb{F}_{\mathbb{S}}]$ is nontrivial.

Corresponding quadruples provide natural examples of cones, as follows. With $\text{diag}(x) := (x_r \delta_{rs}) \in \mathbb{M}_k$ for any $x = (x_r) \in \mathbb{E}^k$, we have

Proposition 26 *(Cone structure of corresponding quadruples)*
Let $(\mathbb{F}, \mathbb{D}, \mathbb{L}, \mathbb{W})$ correspond. Then, $\mathbb{L} = \{\kappa w : \kappa > 0, w \in \mathbb{W}\}$ and $\mathbb{D} = \{\kappa Q \,\text{diag}(w) Q^T : \kappa > 0, Q \in \mathbb{O}_k, w \in \mathbb{W}\}$ are cones in $\mathbb{E}^k$ and in $\mathbb{M}_k$ respectively, both being nontrivial if and only if $\mathbb{W}$ is so, while $\mathbb{F}$ is always a nontrivial cone in $\mathbb{F}_k$.

4.4 Distribution size indices on nontrivial cones in $\mathbb{F}_k$

Attention focuses now on distribution size *indices*, defined as distribution size measures taking values in $(0, \infty)$. Indeed, being a nontrivial cone (Proposition 25), $\eta[\mathbb{F}_{\mathbb{S}}]$ *is* $(0, \infty)$ for any distribution size index η on any nontrivial cone $\mathbb{F}_{\mathbb{S}}$.

Being $\mathbb{S}$-invariant, it is natural to compare two distribution size indices of the same order by their ratio (equivalently, by the difference in their logarithms). The requirement of positivity, rather than mere nonnegativity, ensures that this is always possible. At the same time note that, for any distribution size index η on any nontrivial cone $\mathbb{F}_\mathbb{S}$, and for any $F \in \mathbb{F}_\mathbb{S}$, $\eta[\kappa F] \to 0_+$ as $\kappa \to 0_+$, as is natural since $\kappa F \to \widehat{F}_{0_{\mathbb{E}^k}}$ in the same limit.

5 DISPERSION SIZE

5.1 Filtered distribution size indices

There seems to be no general way to describe the set of all distribution size indices on a given nontrivial cone $\mathbb{F}_\mathbb{S}$. However, as in shape analysis, progress can be made by considering appropriately *filtered* indices, low order (central) moments again being natural candidate filters.

Let η_0 denote a given distribution size measure on $\mathbb{F}_\mathbb{S}$ of order l_0, taking values in a nontrivial real vector space $\mathbb{V}$. Then $\Delta(\mathbb{F}_\mathbb{S};\eta_0)$, the set of all distribution size indices on $\mathbb{F}_\mathbb{S}$ filtered by η_0, can be analysed as follows. It has natural decomposition $\Delta(\mathbb{F}_\mathbb{S};\eta_0) = \cup\{\Delta(\mathbb{F}_\mathbb{S};\eta_0;l) : l > 0\}$ whose general member

$$\Delta(\mathbb{F}_\mathbb{S};\eta_0;l) := \{\eta \in \Delta(\mathbb{F}_\mathbb{S};\eta_0) : \eta \text{ has order } l\}$$

is characterised, in the next result, in terms of scale-isovariants of given order on $\eta_0[\mathbb{F}_\mathbb{S}]$. These, in turn, have been characterised in Proposition 23.

There is natural interest in the subset $\widetilde{\Delta}(\mathbb{F}_\mathbb{S};\eta_0) := \Delta(\mathbb{F}_\mathbb{S};\eta_0;l_0)$ comprising all η_0-filtered distribution size indices of the same order as η_0. We call such indices *homogeneous.* Again, a first-order scale-isovariant defined on a nontrivial cone in a real vector space is called *positively homogeneous*.

Theorem 45 (Appendix C) specialises to

Theorem 27 *(Characterisation of $\Delta(\mathbb{F}_\mathbb{S};\eta_0;l)$ and $\widetilde{\Delta}(\mathbb{F}_\mathbb{S};\eta_0)$)*

Let $\mathbb{V}$ be a nontrivial real vector space, η_0 a distribution size measure of order l_0 from a nontrivial cone $\mathbb{F}_\mathbb{S}$ in $\mathbb{F}_k$ to $\mathbb{V}$, and $\eta : \mathbb{F}_\mathbb{S} \to (0,\infty)$. Then:

(a) $\eta \in \Delta(\mathbb{F}_\mathbb{S};\eta_0;l)$

$\Leftrightarrow \eta = \zeta \circ \eta_0$ *for some scale-isovariant* $\zeta : \eta_0[\mathbb{F}_\mathbb{S}] \to (0,\infty)$ *of order* l/l_0*.*

(b) $\eta \in \widetilde{\Delta}(\mathbb{F}_\mathbb{S};\eta_0)$

$\Leftrightarrow \eta = \widetilde{\zeta} \circ \eta_0$ *for some positively homogeneous* $\widetilde{\zeta} : \eta_0[\mathbb{F}_\mathbb{S}] \to (0,\infty)$*.*

It is easy to see from Theorem 27 that there are no η_0-filtered distribution size indices, of any order, on $\mathbb{F}_\mathbb{S}$ if $0_\mathbb{V} \in \eta_0[\mathbb{F}_\mathbb{S}]$, but that in all other cases a homogeneous such index exists provided only that $\mathbb{V}$ admits a norm, denoted generically by $\|\cdot\|_\mathbb{V}$. We have

Corollary 28 *(Existence of an η_0-filtered distribution size index)*

Let $\mathbb{V}$ be a nontrivial real vector space admitting a norm $\|\cdot\|_{\mathbb{V}}$, and η_0 a distribution size measure of order l_0 from a nontrivial cone $\mathbb{F}_{\mathbb{S}}$ in $\mathbb{F}_k$ to $\mathbb{V}$. Then:

(a) $\Delta(\mathbb{F}_{\mathbb{S}};\eta_0) \neq \varnothing \Rightarrow 0_{\mathbb{V}} \notin \eta_0[\mathbb{F}_{\mathbb{S}}]$.

(b) $\|\eta_0[\cdot]\|_{\mathbb{V}} \in \widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\eta_0)$ *whenever* $0_{\mathbb{V}} \notin \eta_0[\mathbb{F}_{\mathbb{S}}]$.

(c) Accordingly, $\widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\eta_0) \neq \varnothing \Leftrightarrow 0_{\mathbb{V}} \notin \eta_0[\mathbb{F}_{\mathbb{S}}]$.

In particular, for any cone $\mathbb{F}_{\mathbb{S}}$ in $\mathbb{F}_k$, $\|\mu[\cdot]\|_{\mathbb{E}^k} \in \widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\mu) \Leftrightarrow 0_{\mathbb{E}^k} \notin \mu[\mathbb{F}_{\mathbb{S}}]$, while $\|\Sigma[\cdot]\|_{\mathbb{M}_k} \in \widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\Sigma) \Leftrightarrow \mathbb{F}_{\mathbb{S}} \subseteq \mathbb{F}_k^{(0)}$.

5.2 Dispersion size indices

Let $\mathbb{F}_{\mathbb{S}}$ be a nontrivial cone in $\mathbb{F}_k$. As in the shape case, we focus now on the filter $\eta_0 = \Sigma$, so that $l_0 = 2$. Homogeneous indices being of natural interest, we focus further on $\widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\Sigma)$ whose members we call the *dispersion size indices* on $\mathbb{F}_{\mathbb{S}}$, requiring $\mathbb{F}_{\mathbb{S}} \subseteq \mathbb{F}_k^{(0)}$ for $\widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\Sigma) \neq \varnothing$.

Again, if $\mathbb{F}_{\mathbb{S}}$ is $\mathbb{G}_o$-closed, a dispersion size index on $\mathbb{F}_{\mathbb{S}}$ is called *orthogonally invariant* if it is also a $\mathbb{G}_o$-invariant on $\mathbb{F}_{\mathbb{S}}$, the set of all such being denoted $\widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$.

These indices have convenient characterisations. Part (a) of the following result is immediate from Theorem 27. Part (b) rests on the fact that an orthogonally invariant function of a real symmetric matrix depends only on its eigenvalues. We have:

Theorem 29 *(Characterisation of (orthogonally invariant) dispersion size indices)*

Let $\mathbb{F}_{\mathbb{S}} \subseteq \mathbb{F}_k^{(0)}$ be a nontrivial cone in $\mathbb{F}_k$ and $\eta : \mathbb{F}_{\mathbb{S}} \to (0,\infty)$. Then:

(a) $\quad \eta \in \widetilde{\Delta}(\mathbb{F}_{\mathbb{S}};\Sigma)$

$\Leftrightarrow \eta = \widetilde{\zeta} \circ \Sigma$ *for some positively homogeneous* $\widetilde{\zeta} : \Sigma[\mathbb{F}_{\mathbb{S}}] \to (0,\infty)$.

(b) If $\mathbb{F}_{\mathbb{S}}$ is also closed under $\mathbb{G}_o$, then:

$\eta \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$

$\Leftrightarrow \eta = \widetilde{\zeta}_* \circ \lambda$ *for some positively homogeneous* $\widetilde{\zeta}_* : \lambda[\mathbb{F}_{\mathbb{S}}] \to (0,\infty)$.

5.3 Connections between dispersion shape and size indices

Orthogonally invariant dispersion size indices are of particular interest because of their close connections with dispersion shape indices.

Using Theorem 29, examples include the following.

Corollary 30 *(Examples of orthogonally invariant dispersion size indices)*

(a) Let $\mathbb{F}_{\mathbb{S}}$ be any nontrivial subcone of $\mathbb{F}_k^{(0)}$. Then:

(i) $\|\lambda[\cdot]\|_{\mathbb{E}^k} \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$.

(ii) For each $r = 1, ..., k$, $\lambda_1[\cdot] + ... + \lambda_r[\cdot] \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$.

In particular, $\mathrm{trace}(\Sigma[\cdot]) \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$*.*

(b) Let now $\mathbb{F}_{\mathbb{S}}$ *be a nontrivial subcone of* $\mathbb{F}_k^+ \subset \mathbb{F}_k^{(0)}$*. Then:*

(i) For each $r = 1, ..., k$*,* $\lambda_r[\cdot] \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$*.*

(ii) The spectral geometric mean $\mathrm{GM}(\lambda[F]) := (\det \Sigma[F])^{1/k} \in \widetilde{\Delta}_o(\mathbb{F}_{\mathbb{S}};\Sigma)$*.*

Now, any $\eta \in \Delta(\mathbb{F}_{\mathbb{S}};\Sigma)$ inherits from Σ the property of being a $\mathbb{G}_l$-invariant on every $\mathbb{G}_l$-closed $\mathbb{F}_{\mathbb{S}}$. Accordingly, recalling (5), we have

Proposition 31 *(Ratios of orthogonally invariant dispersion size indices)*
For any $\mathbb{G}$*-closed* $\mathbb{F} \subseteq \mathbb{F}_k^{(0)}$*, (any real-valued function of) the ratio of any two members of* $\widetilde{\Delta}_o(\mathbb{F};\Sigma)$ *is a* $\mathbb{G}$*-invariant on* $\mathbb{F}$*.*

Restricting attention now to $\mathbb{F}_k^+$, a second general connection is established by recalling (Theorem 20) that:

$\iota : \mathbb{F}_k^+ \to \mathbb{R}$ is a dispersion shape index if and only if it is w-filtered

while, by Corollary 30, for each $r = 1, ..., k$,

$w_r[\cdot]$ is the ratio of two orthogonally invariant dispersion size indices on $\mathbb{F}_k^+$.

6 EQUIVARIANCE PROPERTIES OF THE INFLUENCE FUNCTION

Influence function measures are widely used in both robust statistics (Hampel et al., 1986) and diagnostics (Cook and Weisberg, 1982).

These measures enjoy the general equivariance properties established below, leading to a framework for an enhanced form of influence analysis.

Our definition of the influence function differs simply, but fundamentally, from the usual case.

6.1 Definition

Let $\mathbb{V}_\circ$ denote now a nontrivial real vector space endowed with a given norm $\|\cdot\|_\circ$ and T a function defined on $\varnothing \subset \mathbb{F}_T \subseteq \mathbb{F}_k$ and taking values in $\mathbb{V}_\circ$. A pair $[x;F] \in \mathbb{I}_k := \mathbb{E}^k \times \mathbb{F}_k$ is called T-compatible if $[(1-\varepsilon)F + \varepsilon\widehat{F}_x] \in \mathbb{F}_T$ for all small enough $\varepsilon \geq 0$. Let $\mathbb{I}_T$ denote the set of all pairs $[x;F] \in \mathbb{I}_k$ such that (a) $[x;F]$ is T-compatible and (b) $\lim_{\varepsilon\to 0_+}\left(\frac{T[(1-\varepsilon)F+\varepsilon\widehat{F}_x]-T[F]}{\varepsilon}\right)$, denoted $I(x;T,F)$, exists. By definition, $\mathbb{I}_T \subseteq \mathbb{E}^k \times \mathbb{F}_T$. We call T *normal* when, as is usual, $\mathbb{I}_T = \mathbb{E}^k \times \mathbb{F}_T$. The limit here is taken in the sense of norm convergence on $\mathbb{V}_\circ$. That is,

$$\left\| \left(\frac{T[(1-\varepsilon)F + \varepsilon\widehat{F}_x] - T[F]}{\varepsilon}\right) - I(x;T,F)\right\|_\circ \to 0 \text{ as } \varepsilon \to 0_+.$$

In particular, $I(x;T,F)$ itself belongs to $\mathbb{V}_\circ$.

In the literature, $I(x;T,F)$ is usually thought of as a function of x for fixed T and F. In the current context, it is more insightful to fix T only and to define its influence function $I_T : \mathbb{I}_T \to \mathbb{V}_\circ$ by $I_T[x;F] := I(x;T,F)$. This simple distinction is fundamental to the development below, which takes place in $\mathbb{I}_k$.

The notation $T > 0$ denotes that $(\mathbb{V}_\circ, \|\cdot\|_\circ) = (\mathbb{E}, \|\cdot\|_2)$ with $T[\mathbb{F}_T] \subseteq (0,\infty)$. In this case, we have

$$I_{\log T}[x;F] = I_T[x;F]/T[F] \text{ on } \mathbb{I}_T = \mathbb{I}_{\log T}. \tag{6}$$

6.2 An extended isomorphism

The isomorphism $g \to \widetilde{g}$ of $\mathbb{G}$ to $\widetilde{\mathbb{G}}$ extends naturally, in two senses, as follows.

First, let $\mathbb{H}$ denote the set of all nonsingular affine transformations h on $\mathbb{E}^k$, so that $\mathbb{G}$ is a subgroup of $\mathbb{H}$, while $\mathbb{H} \rightsquigarrow \mathbb{E}^k$. Then, with $\widetilde{h}[F]$ denoting the distribution of $h(X)$ induced by $X \sim F$, $h \to \widetilde{h}$ is an isomorphism of $\mathbb{H}$ to $\widetilde{\mathbb{H}} := \{\widetilde{h} : h \in \mathbb{H}\} \rightsquigarrow \mathbb{F}_k$, which we may use to identify $\mathbb{H}$ with $\widetilde{\mathbb{H}}$.

Second, we may extend $\mathbb{H} \rightsquigarrow \mathbb{F}_k$ to $\mathbb{H} \rightsquigarrow \mathbb{I}_k$ via $h[x;F] := [h(x);h(F)]$. Of course, if $\mathbb{H}_\circ$ is a subgroup of $\mathbb{H}$ and $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_k$, $\mathbb{I}_k[\mathbb{F}] := \mathbb{E}^k \times \mathbb{F}$ is $\mathbb{H}_\circ$-closed if and only if $\mathbb{F}$ is so. In particular, identifying $\mathbb{H}_\circ = \mathbb{G}_s$ with $\mathbb{S}$ as above, $\varnothing \subset \mathbb{I}_k[\mathbb{F}] \subseteq \mathbb{I}_k$ is called a *cone* in $\mathbb{I}_k$ iff $\mathbb{F}$ is a cone in $\mathbb{F}_k$, $\mathbb{S}$-regularity of $\mathbb{E}^k$ implying at once

Proposition 32 *(Regularity of $\mathbb{I}_k[\mathbb{F}]$)*
Every cone $\mathbb{I}_k[\mathbb{F}]$ in $\mathbb{I}_k$ is $\mathbb{S}$-regular.

Now, $\mathbb{I}_k$ has a natural convex structure defined, for $0 \le \varepsilon \le 1$, by

$$(1-\varepsilon)[x;F] + \varepsilon[x_*;F_*] := [(1-\varepsilon)x + \varepsilon x_*; (1-\varepsilon)F + \varepsilon F_*]$$

which is preserved by $\mathbb{H}$ in the sense that

$$h[(1-\varepsilon)[x;F] + \varepsilon[x_*;F_*]] = (1-\varepsilon)h[x;F] + \varepsilon h[x_*;F_*].$$

In particular, as $h[\widehat{F}_x] = \widehat{F}_{h(x)}$,

$$h[(1-\varepsilon)F + \varepsilon\widehat{F}_x] = (1-\varepsilon)h[F] + \varepsilon\widehat{F}_{h(x)}. \tag{7}$$

Any subgroup $\mathbb{H}_\circ \subseteq \mathbb{H}$ automatically preserves convexity.

6.3 Inheritance by I_T of equivariance properties of T

Recall that both invariance (Section 2.11) and (scale-) isovariance (Section 4) are special cases of equivariance.

Now, for any normal T, $\mathbb{I}_k[\mathbb{F}] \subseteq \mathbb{I}_T$ whenever $\mathbb{F} \subseteq \mathbb{F}_T$. In this case, I_T inherits on $\mathbb{I}_k[\mathbb{F}]$ equivariance properties of T on $\mathbb{F}$, as follows.

Theorem 33 *(Inheritance by I_T of equivariance of T)*
Let $\mathbb{V}_\circ$ be a nontrivial real vector space with norm $\|\cdot\|_\circ$ and $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_T \subseteq \mathbb{F}_k$, where $T : \mathbb{F}_T \to \mathbb{V}_\circ$ is normal. Again, let $\mathbb{H}_\circ$ be a subgroup of $\mathbb{H}$ and suppose that $\mathbb{F}$ is $\mathbb{H}_\circ$-closed.
Consider now the following two statements made for all $h \in \mathbb{H}_\circ|_\mathbb{F}$, with $c(h) \in \mathbb{R}$:
[1] *T is an $\mathbb{H}_\circ|_\mathbb{F}$-equivariant on $\mathbb{F}$ with ${}^T h : T[F] \to c(h)T[F]$.*
[2] *I_T is an $\mathbb{H}_\circ|_\mathbb{F}$-equivariant on $\mathbb{I}_k[\mathbb{F}]$, with ${}^{I_T} h : I_T[x; F] \to c(h)I_T[x; F]$.*
Then:
(a) [1] $\Rightarrow$ [2].
(b) If also $T > 0$, [1] $\Rightarrow$ *$I_{\log T}$ is an $\mathbb{H}_\circ|_\mathbb{F}$-invariant on $\mathbb{I}_k[\mathbb{F}]$.*

Proof. Using equation (7) and [1], (a) follows from the definition of $I_T[h[x; F]]$, implying (b) via (6). ■

Specialising to the case $c(h) \equiv 1$, gives at once:

Theorem 34 *(Inheritance by I_T of invariance of T)*
Let $\mathbb{V}_\circ$ be a nontrivial real vector space with norm $\|\cdot\|_\circ$ and $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_T \subseteq \mathbb{F}_k$, where $T : \mathbb{F}_T \to \mathbb{V}_\circ$ is normal. Again, let $\mathbb{H}_\circ$ be a subgroup of $\mathbb{H}$ and suppose that $\mathbb{F}$ is $\mathbb{H}_\circ$-closed.
Consider now the following two statements made for all $h \in \mathbb{H}_\circ|_\mathbb{F}$:
[1] *T is an $\mathbb{H}_\circ|_\mathbb{F}$-invariant on $\mathbb{F}$.*
[2] *I_T is an $\mathbb{H}_\circ|_\mathbb{F}$-invariant on $\mathbb{I}_k[\mathbb{F}]$.*
Then:
(a) [1] $\Rightarrow$ [2].
(b) If also $T > 0$, [1] $\Rightarrow$ *$I_{\log T}$ is an $\mathbb{H}_\circ|_\mathbb{F}$-invariant on $\mathbb{I}_k[\mathbb{F}]$.*

Again, specialising Theorem 33 to the case where $\mathbb{H}_\circ = \mathbb{S}$, $\mathbb{F}$ is $\mathbb{H}_\circ$-regular and $c(\kappa) = \kappa^l$ $(\kappa > 0,\ l > 0)$ gives

Theorem 35 *(Inheritance by I_T of scale-isovariance of T)*
Let $\mathbb{V}_\circ$ be a nontrivial real vector space with norm $\|\cdot\|_\circ$ and $\varnothing \subset \mathbb{F} \subseteq \mathbb{F}_T \subseteq \mathbb{F}_k$, where $T : \mathbb{F}_T \to \mathbb{V}_\circ$ is normal. Again, let $l > 0$ and $\mathbb{F}$ be a nontrivial cone.
Consider now the following two statements:
[1] *T is an l^{th}-order distribution size measure on $\mathbb{F}$.*
[2] *I_T is an l^{th}-order scale-isovariant on $\mathbb{I}_k[\mathbb{F}]$.*
Then:
(a) [1] $\Rightarrow$ [2].
(b) If also $T > 0$, [1] $\Rightarrow$ *$I_{\log T}$ is an $\mathbb{S}$-invariant on $\mathbb{I}_k[\mathbb{F}]$.*

These three theorems have a variety of corollaries. We focus here on their special cases where T is either a dispersion shape index or an orthogonally invariant dispersion size index. Theorem 34 gives at once

Corollary 36 *(Equivariance properties of I_T for dispersion shape indices)*
Let ι be a dispersion shape index on a $\mathbb{G}$-closed subset $\mathbb{F}$ of $\mathbb{F}_k$, so that ι is a $\mathbb{G}$-invariant on $\mathbb{F}$. Then, I_ι inherits the same property on $\mathbb{I}_k[\mathbb{F}]$.

Again, recalling Proposition 3(a) and Corollary 28, Theorems 34 and 35 gives

Corollary 37 *(Equivariance properties of I_T for orthogonally invariant dispersion size indices)*
Let η be an orthogonally invariant dispersion size index on a $\mathbb{G}$-closed subset $\mathbb{F}$ of $\mathbb{F}_k^{(0)}$, so that η is a second-order scale-isovariant, and a $\mathbb{G}_o$- and $\mathbb{G}_l$- invariant, on $\mathbb{F}$. Then, I_η inherits the same properties on $\mathbb{I}_k[\mathbb{F}]$. Moreover, $I_{\log \eta}$ is a $\mathbb{G}$-invariant on $\mathbb{I}_k[\mathbb{F}]$.

By Corollary 30, Corollary 37 has Proposition II as a special case, adding that $I_{\log \operatorname{trace} \Sigma}$ is a $\mathbb{G}$-invariant on $\mathbb{G}$-closed subsets $\mathbb{F}$ of $\mathbb{F}_k^{(0)}$.

The same analysis applied to $\eta = \mathrm{GM}(\lambda[\cdot])$ shows that $I_{\log \det \Sigma}$ is a $\mathbb{G}$-invariant on $\mathbb{G}$-closed subsets $\mathbb{F}$ of $\mathbb{F}_k^+$. However, as $\det \Sigma[h[F]] = (\det A)^2 \det \Sigma[F]$ for any $h : x \to A(x - x_\circ)$ in $\mathbb{H}$, we may use Theorem 33 to find the stronger result that $I_{\log \det \Sigma}$ is an $\mathbb{H}$-invariant there. In fact:

$$I_{\log \det \Sigma}[x; F] = (x - \mu[F])^T \Sigma^{-1}[F](x - \mu[F]) - k.$$

6.4 Complementary influence analysis

The above results provide a framework in which to develop a complementary shape-and-size form of influence analysis that is appropriately invariant. This is described briefly here, further developments being indicated.

Distribution shape and size varying independently of each other, a given perturbation may affect either, neither, or both. An enhanced influence analysis is therefore obtained by appropriate, complementary use of pairs of distribution shape and size measures.

Now, the group under which it is appropriate to require the results of an influence analysis to be invariant varies according to context. For example, the group $\mathbb{G}$ may be appropriate in a principal component analysis context in which all the variables measured are commensurable. In such a case, Corollaries 36 and 37 ensure $\mathbb{G}$-invariance of an analysis based on $(I_\iota, I_{\log \eta})$ for any dispersion shape index ι and orthogonally invariant dispersion size index η. Complete accounts of both types of index are given in Theorems 20 and 29 respectively, $k > 1$ being required for a nontrivial ι (when $k = 1$, appropriate skewness or kurtosis measures of shape may, in practice, be used instead). Such complementary influence analysis is further developed and applied in a related paper.

Further research is required when $\mathbb{H}$ is the appropriate invariance group for an influence analysis. Whereas, as we have seen, we may use $I_{\log \det \Sigma}$, the notion of dispersion shape is no longer helpful under $\mathbb{H}$, any two dispersion matrices of the same rank (in particular, any two nonsingular dispersion matrices) being affinely equivalent. However, we may consider other $\mathbb{H}$-invariant measures of distribution shape, appropriate skewness and kurtosis measures being natural candidates provided the necessary higher-order moments exist.

Again, making the necessary changes (including explicit treatment of regularity when first moments do not exist), we may consider extensions of the above framework from $\mathbb{H} \rightsquigarrow \mathbb{I}_k$ to $\mathbb{H} \rightsquigarrow \overline{\mathbb{I}}_k$, where

$$\overline{\mathbb{I}}_k := \mathbb{E}^k \times \overline{\mathbb{F}}_k \text{ with } \overline{\mathbb{F}}_k := \{\textit{all} \text{ distributions on } \mathbb{E}^k\},$$

and from the influence function to other aspects of robust statistics such as the maxbias curve. Finally, it will be of interest to extend the above treatment of single-case perturbations to include multiple case effects. The salience-based approach of Critchley et al. (2001) will be relevant here.

ACKNOWLEDGEMENTS

We are grateful to Laurie Davies and Joe Eaton for helpful comments, and to the EPSRC for support under research grant GR/K08246.

7 APPENDIX A Relations between distribution and dispersion shape measures

Let $\mathbb{G} \rightsquigarrow \mathbb{F}$ and recall that, whereas every dispersion shape measure on $\mathbb{F}$ is (by definition) a distribution shape measure on $\mathbb{F}$, the converse is false in general.

Let now $\mathbb{D} := \Sigma[\mathbb{F}]$ so that ${}^{\Sigma}\mathbb{G} \rightsquigarrow \mathbb{D}$. Then $\overset{\Sigma\mathbb{G}}{\sim}$ on $\mathbb{D}$ induces an equivalence relation, $\overset{\mathbb{G}_{\mathbb{D}}}{\sim}$ say, on $\mathbb{F}$ via $F_1 \overset{\mathbb{G}_{\mathbb{D}}}{\sim} F_2 \Leftrightarrow \Sigma[F_1] \overset{\Sigma\mathbb{G}}{\sim} \Sigma[F_2]$ and we have

$$F_1 \overset{\mathbb{G}}{\sim} F_2 \Rightarrow \Sigma[F_1] \overset{\Sigma\mathbb{G}}{\sim} \Sigma[F_2], \tag{8}$$

so that each $\overset{\mathbb{G}_{\mathbb{D}}}{\sim}$ class is a union of $\overset{\mathbb{G}}{\sim}$ classes. However, these classes do not in general coincide. That is, the converse of (8) is also false in general.

Rather, as we now show, one of these converses holds if and only if the other does, both obtaining under appropriate conditions on $\mathbb{F}$. For example, a nonempty proper subset $\mathbb{F}$ of $\mathbb{F}_k$ is said to be a collection of distributions determined by their first two moments if $F \to (\mu[F], \Sigma[F])$ $(F \in \mathbb{F})$ is $1-1$ and we have

Proposition 38 *(Conditions for equivalence of distribution and dispersion shape measures)*

Let $\mathbb{G} \rightsquigarrow \mathbb{F}$, $\mathbb{D} := \Sigma[\mathbb{F}]$ and $\iota \in \mathrm{dom}(\mathbb{F})$. Consider the following three statements:

[1] $\mathbb{F}$ is a collection of distributions determined by their first two moments.

[2] $F_1 \overset{\mathbb{G}}{\sim} F_2 \Leftrightarrow \Sigma[F_1] \overset{\Sigma\mathbb{G}}{\sim} \Sigma[F_2]$.

[3] ι is a dispersion shape measure $\Leftrightarrow$ ι is a distribution shape measure.

Then,

$$[1] \ \Rightarrow \ [2] \ \Leftrightarrow \ [3].$$

Proof. Recall, first, that the forwards implication in both [2] and [3] always holds.

[1] $\Rightarrow$ [2]: Let $F_1, F_2 \in \mathbb{F}$ with $\Sigma[F_2] = \kappa^2 Q \Sigma[F_1] Q^T$ for some $\kappa > 0$ and $Q \in \mathbb{O}_k$. Then $(\kappa, Q, \mu[F_1] - \kappa^{-1} Q^T \mu[F_2]) F_1$ belongs to $\mathbb{F}$ and has the same two first moments as F_2. It therefore *is* F_2.

[2] $\Rightarrow$ [3]: Let ι be a distribution shape measure on $\mathbb{F}$ and let $F_1, F_2 \in \mathbb{F}$ with $\Sigma[F_1] = \Sigma[F_2]$. Then, $F_1 \overset{\mathbb{G}}{\sim} F_2$ by [2], and so $\iota[F_1] = \iota[F_2]$. Thus, by Lemma 10, ι is Σ-filtered.

[3] $\Rightarrow$ [2]: The proof is by negation. Suppose then that [2] fails, noting that this is equivalent to the existence of $F_1, F_2 \in \mathbb{F}$ with $\Sigma[F_1] = \Sigma[F_2]$ but $F_1 \overset{\mathbb{G}}{\nsim} F_2$. Let $\mathbb{F}_*$ be a set of $\mathbb{G}$-canonical forms of $\mathbb{F}$ and let ι_* be the corresponding complete $\mathbb{G}$-invariant on $\mathbb{F}$. Then, ι_* is a distribution shape measure on $\mathbb{F}$ but not a dispersion shape measure, since $\iota_*[F_1] \neq \iota_*[F_2]$. ■

In particular, $\mathbb{N}_k$ satisfies both $\mathbb{G} \rightsquigarrow \mathbb{N}_k$ and [1] of Proposition 38, so that distribution and dispersion shape measures coincide for multivariate normal distributions.

8 APPENDIX B Inheritance of isovariance

Proposition 39 *(Inheritance of isovariance)*

Let η be a $\mathbb{T}$-isovariant on $\mathbb{A}$, $\mathbb{T}_\circ$ be a subgroup of $\mathbb{T}$, and $\varnothing \subset \mathbb{A}_\mathbb{T} \subseteq \mathbb{A}$ be $\mathbb{T}$-closed. Then:
(a) η is a $\mathbb{T}_\circ$-isovariant on $\mathbb{A}$.
(b) If η is $1-1$, $\eta|_{\mathbb{A}_\mathbb{T}}$ is a $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$-isovariant on $\mathbb{A}_\mathbb{T}$.
(c) Consider the following three statements:
[1] *$\eta(\mathbb{A}_\mathbb{T})$ is ${}^\eta\mathbb{T}$-regular.*
[2] *$t|_{\mathbb{A}_\mathbb{T}} \to {}^\eta t|_{\eta(\mathbb{A}_\mathbb{T})}$ is $1-1$.*
[3] *$\eta|_{\mathbb{A}_\mathbb{T}}$ is a $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$-isovariant on $\mathbb{A}_\mathbb{T}$.*
Then, [1] $\Rightarrow$ [2] $\Leftrightarrow$ [3].
(d) If, further, $\mathbb{A}_\mathbb{T}$ is $\mathbb{T}$-regular, all three statements in (c) are equivalent and, when they hold, $\eta|_{\widetilde{\mathbb{A}}_\mathbb{T}}$ is a $\mathbb{T}|_{\widetilde{\mathbb{A}}_\mathbb{T}}$-isovariant on $\widetilde{\mathbb{A}}_\mathbb{T}$ for every $\mathbb{T}$-closed set $\widetilde{\mathbb{A}}_\mathbb{T}$ with $\mathbb{A}_\mathbb{T} \subseteq \widetilde{\mathbb{A}}_\mathbb{T} \subseteq \mathbb{A}$.

Proof. (a) Immediate.

(b) Equivariance of $\eta|_{\mathbb{A}_\mathbb{T}}$ is immediate from Proposition 13(b), while:

$$[\eta \text{ is } 1-1] \Rightarrow \left[{}^\eta t|_{\eta(\mathbb{A}_\mathbb{T})} = {}^\eta t_*|_{\eta(\mathbb{A}_\mathbb{T})} \Rightarrow t|_{\mathbb{A}_\mathbb{T}} = t_*|_{\mathbb{A}_\mathbb{T}}\right].$$

(c) [1] $\Rightarrow$ [2] : Using first ${}^\eta\mathbb{T}$-regularity of $\eta(\mathbb{A}_\mathbb{T})$, and then $\mathbb{T}$-isovariance of η, we have: ${}^\eta t|_{\eta(\mathbb{A}_\mathbb{T})} = {}^\eta t_*|_{\eta(\mathbb{A}_\mathbb{T})} \Rightarrow {}^\eta t = {}^\eta t_* \Rightarrow t = t_*$.
[2] $\Leftrightarrow$ [3] : Using Proposition 13(b) as before, this is now immediate.

(d) Suppose now that $\mathbb{A}_\mathbb{T}$ is $\mathbb{T}$-regular and that its superset $\widetilde{\mathbb{A}}_\mathbb{T}$ in $\mathbb{A}$ is $\mathbb{T}$-closed. Using Proposition 13(b) again, it suffices to observe that, under $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$-isovariance of $\eta|_{\mathbb{A}_\mathbb{T}}$,

$${}^\eta t|_{\eta(\widetilde{\mathbb{A}}_\mathbb{T})} = {}^\eta t_*|_{\eta(\widetilde{\mathbb{A}}_\mathbb{T})} \Rightarrow {}^\eta t|_{\eta(\mathbb{A}_\mathbb{T})} = {}^\eta t_*|_{\eta(\mathbb{A}_\mathbb{T})} \Rightarrow t|_{\mathbb{A}_\mathbb{T}} = t_*|_{\mathbb{A}_\mathbb{T}} \Rightarrow t = t_*.$$

■

Proposition 39(d) is 'best possible' in the sense that putting $\mathbb{A}_\mathbb{T} = \{a_1\}$ and $\eta = \mathrm{id}_\mathbb{A}$ in Example 2 shows that $\mathbb{T}$-regularity of $\mathbb{A}_\mathbb{T}$ is indispensable to [2] $\Rightarrow$ [1], while Example 40 below shows that it is possible to have:

η a $\mathbb{T}$-isovariant on $\mathbb{A}$,

$\varnothing \subset \mathbb{A}_\mathbb{T} \subset \widetilde{\mathbb{A}}_\mathbb{T} \subset \mathbb{A}$, with $\mathbb{A}_\mathbb{T}$ and $\widetilde{\mathbb{A}}_\mathbb{T}$ respectively $\mathbb{T}$-closed and $\mathbb{T}$-regular,

$\eta|_{\mathbb{A}_\mathbb{T}}$ a $\mathbb{T}|_{\mathbb{A}_\mathbb{T}}$-isovariant on $\mathbb{A}_\mathbb{T}$

and yet:

$\eta|_{\widetilde{\mathbb{A}}_\mathbb{T}}$ *not* a $\mathbb{T}|_{\widetilde{\mathbb{A}}_\mathbb{T}}$-isovariant on $\widetilde{\mathbb{A}}_\mathbb{T}$.

Example 40 *(An isovariance counterexample)*
It is straightforward to show that the following example has the properties claimed:
$\mathbb{A}_\mathbb{T} = \{0\} \subset \widetilde{\mathbb{A}}_\mathbb{T} = \{0, \pm 1\} \subset \mathbb{A} = \{0, \pm 1, \pm 2\}$,
$\mathbb{T} = \{\mathrm{id}_\mathbb{A}, t\}$, *where* $t : a \to -a$,
and $\eta(a) = |a|$ $(a \in \widetilde{\mathbb{A}}_\mathbb{T})$, $\eta(a) = a$ *else.*

9 APPENDIX C Scale-isovariance revisited

Propositions 22 and 24 establish, respectively, that Propositions 23 and 25 are special cases of Theorem 42 below. Again, Theorem 45 below specialises to Theorem 27 of Section 5.

9.1 Scale transformation and cones

Let $\mathbb{S}$ be as defined above (Section 4.2) and $\mathbb{S} \overset{\alpha}{\cong} \mathbb{T}$ denote that α is an isomorphism of $\mathbb{S}$ to $\mathbb{T}$. If $\mathbb{S} \overset{\alpha}{\cong} \mathbb{T}$ and $\mathbb{T} \rightsquigarrow \mathbb{A}$, we call $\mathbb{T}$ *the α-isomorphic scale transformation group on* $\mathbb{A}$ and write $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$. We have at once

Proposition 41 *(Characterisation of α-isomorphic scale transformation groups) Let $\alpha \in \mathrm{dom}(\mathbb{S})$ and $\mathbb{T} \rightsquigarrow \mathbb{A}$. Then:*

$$\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A} \Leftrightarrow [\alpha \text{ is } 1-1, \mathbb{T} = \{\alpha(\kappa) : \kappa > 0\} \text{ and } \alpha(\kappa_2 \circ \kappa_1) = \alpha(\kappa_2) \circ \alpha(\kappa_1)]$$

in which case $\mathbb{T}$ is Abelian, $\alpha(1) = \mathrm{id}_{\mathbb{A}}$ and $(\alpha(\kappa))^{-1} = \alpha(\kappa^{-1})$ $(\kappa > 0)$.

Suppose now that $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$ and $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$. Then we call $\mathbb{A}_{\mathbb{T}}$ an *α-cone* (*in* $\mathbb{A}$) if it is closed under $\mathbb{T} = \alpha(\mathbb{S})$. In particular, $\mathbb{A}$ itself is an α-cone.

For any such α-cone $\mathbb{A}_{\mathbb{T}}$, $\mathbb{T}$ and $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$ are homomorphic but not, in general, isomorphic via $t \to t|_{\mathbb{A}_{\mathbb{T}}}$ (Proposition 4). However, provided $\mathbb{A}_{\mathbb{T}}$ is regular, $\mathbb{S} \overset{\alpha}{\cong} \mathbb{T} \cong \mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$ and $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}} \rightsquigarrow \mathbb{A}_{\mathbb{T}}$. In this case, without confusion, we also call $\mathbb{T}$ *the α-isomorphic scale transformation group on* $\mathbb{A}_{\mathbb{T}}$ and write $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}_{\mathbb{T}}$.

9.2 Scale-isovariants

Recall that the definition of a function $\eta(\cdot)$ as a $\mathbb{T}$-isovariant on $\mathbb{A}$ requires no pre-existing structure on its image space. Rather, such an isovariant induces a transformation group structure there, namely ${}^{\eta}\mathbb{T} \rightsquigarrow \eta(\mathbb{A})$. In particular, if $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$ and η is a $\mathbb{T}$-isovariant on $\mathbb{A}$,

$${}^{\eta}\mathbb{T} \overset{{}^{\eta}\alpha}{\rightsquigarrow} \eta(\mathbb{A}) \text{ where } {}^{\eta}\alpha : \kappa \to {}^{\eta}(\alpha(\kappa)) \text{ is a group isomorphism of } \mathbb{S} \text{ to } {}^{\eta}\mathbb{T}. \quad (9)$$

This lack of pre-existing structure contrasts sharply with the definition of a scale-isovariant, for the following reason.

There is never a unique α such that $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$. In particular, for any $l > 0$, the power transformation $\pi^{(l)} : \kappa \to \kappa^l$ is an isomorphism of $\mathbb{S}$ to itself, so that $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A} \Leftrightarrow \mathbb{T} \overset{\alpha \circ \pi^{(l)}}{\rightsquigarrow} \mathbb{A}$. Accordingly, a well-defined notion of the order of a scale-isovariant between sets must be *relative* to prescribed scale transformation groups on those sets – albeit that there is often a natural choice of such groups, as with the special types of scale-isovariant defined in Section 4.

Suppose then that $\eta : \mathbb{A} \to \mathbb{A}_*$, where $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$ and $\mathbb{T}_* \overset{\alpha_*}{\rightsquigarrow} \mathbb{A}_*$. Then, for any $l > 0$, we call η *an l^{th}-order (α, α_*)-scale-isovariant* (*from* $\mathbb{A}$ *to* $\mathbb{A}_*$) if η is a

$\mathbb{T}$-isovariant on $\mathbb{A}$ with

$$^{\eta}\alpha = \alpha_{*,\eta(\mathbb{A})} \circ \pi^{(l)}, \tag{10}$$

where $\alpha_{*,\eta(\mathbb{A})} : \kappa \to \alpha_*(\kappa)|_{\eta(\mathbb{A})}$, in which case

$$\alpha_{*,\eta(\mathbb{A})} = {}^{\eta}\alpha \circ \pi^{(1/l)} \text{ is a group isomorphism of } \mathbb{S} \text{ to } {}^{\eta}\mathbb{T} = \mathbb{T}_*|_{\eta(\mathbb{A})}. \tag{11}$$

When, as in Section 4, the isomorphisms (α, α_*) are clear from the context, we may omit them and refer only to scale-isovariants. A first-order scale-isovariant is called *positively homogeneous.*

We have

Theorem 42 *(Characterisation and properties of scale-isovariants)*
Let $l > 0$, $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$, $\mathbb{T}_* \overset{\alpha_*}{\rightsquigarrow} \mathbb{A}_*$ *and* $\eta : \mathbb{A} \to \mathbb{A}_*$. *Again, let* $\varnothing \subset \mathbb{A}_{\mathbb{T}} \subseteq \mathbb{A}$ *be* $\mathbb{T}$*-regular, so that* $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}_{\mathbb{T}}$. *Then:*
η *is an* l^{th}*-order* (α, α_*)*-scale-isovariant from* $\mathbb{A}$ *to* $\mathbb{A}_*$
$\Leftrightarrow$ $\eta(\mathbb{A})$ *is a regular* α_**-cone, with*

$$\eta \circ \alpha(\kappa) = \alpha_*(\kappa^l)|_{\eta(\mathbb{A})} \circ \eta, (\kappa > 0), \tag{12}$$

in which case:
(a) $\mathbb{T}_* \overset{\alpha_*}{\rightsquigarrow} \eta(\mathbb{A})$.
(b) The order of η *is well-defined.*
(c) [1] $\eta|_{\mathbb{A}_{\mathbb{T}}}$ *is an* l^{th}*-order* (α, α_*)*-scale-isovariant from* $\mathbb{A}_{\mathbb{T}}$ *to* $\mathbb{A}_*$
$\Leftrightarrow$ [2] $\eta(\mathbb{A}_{\mathbb{T}})$ *is a regular* α_**-cone.*

Proof. *Necessity*: By definition, $\eta(\mathbb{A})$ is $\mathbb{T}_*$-closed and (12) holds. Moreover, as $\alpha_{*,\eta(\mathbb{A})}$ is $1-1$ by (11), $\eta(\mathbb{A})$ is $\mathbb{T}_*$-regular.

Sufficiency: By hypothesis, η is here a $\mathbb{T}$-equivariant satisfying (10), while

$$\alpha_*(\kappa_1^l)|_{\eta(\mathbb{A})} = \alpha_*(\kappa_2^l)|_{\eta(\mathbb{A})} \Rightarrow \alpha_*(\kappa_1^l) = \alpha_*(\kappa_2^l) \Rightarrow \kappa_1 = \kappa_2$$

using, in turn, regularity of $\eta(\mathbb{A})$ and the fact that α_* and $\pi^{(l)}$ are $1-1$.

Suppose now that $\eta : \mathbb{A} \to \mathbb{A}_*$ is an (α, α_*)-scale-isovariant of order l.
(a) This is immediate from $\mathbb{T}_*$-regularity of $\eta(\mathbb{A})$ and Proposition 4.
(b) Suppose, further, that η is also an (α, α_*)-scale-isovariant of order $\widetilde{l}$. Then

$$\alpha_{*,\eta(\mathbb{A})} \circ \pi^{(l)} = \alpha_{*,\eta(\mathbb{A})} \circ \pi^{(\widetilde{l})}$$

and it suffices to note that $\alpha_{*,\eta(\mathbb{A})}$ is invertible and that $l \to \pi^{(l)}$ is $1-1$.
(c) Using the definition of an (α, α_*)-scale-isovariant, Proposition 39(d) and (11), we have:

[1] $\Leftrightarrow$ [$\eta|_{\mathbb{A}_{\mathbb{T}}}$ is a $\mathbb{T}|_{\mathbb{A}_{\mathbb{T}}}$-isovariant on $\mathbb{A}_{\mathbb{T}}$] $\Leftrightarrow$ [$\eta(\mathbb{A}_{\mathbb{T}})$ is ${}^{\eta}\mathbb{T}$-regular] $\Leftrightarrow$ [2]. ■

9.3 Filtered scale-isovariants

Scale-isovariants filtered by a given $\mathbb{T}$-isovariant or scale-isovariant can be characterised as follows.

Lemma 43 *(Characterisation of isovariant-filtered isovariants)*
Let $\mathbb{T} \rightsquigarrow \mathbb{A}$, η_0 be a $\mathbb{T}$-isovariant on $\mathbb{A}$ and $\zeta \in \mathrm{dom}(\eta_0(\mathbb{A}))$. Then:

$$\zeta \circ \eta_0 \textit{ is a } \mathbb{T}\textit{-isovariant on } \mathbb{A} \Leftrightarrow \zeta \textit{ is a } {}^{\eta_0}\mathbb{T}\textit{-isovariant on } \eta_0(\mathbb{A})$$

in which case, for each $t \in \mathbb{T}$, ${}^{(\zeta\circ\eta_0)}t = {}^{\zeta}({}^{\eta_0}t)$, so that ${}^{(\zeta\circ\eta_0)}\mathbb{T} = {}^{\zeta}({}^{\eta_0}\mathbb{T})$.

Proof. This is an immediate corollary to Theorem 14(b), noting that:
η_0 is a $\mathbb{T}$-isovariant
$\Rightarrow t \to {}^{\eta_0}t$ is $1-1$
$\Rightarrow [t \to {}^{(\zeta\circ\eta_0)}t$ is $1-1 \Leftrightarrow {}^{\eta_0}t \to {}^{(\zeta\circ\eta_0)}t$ is $1-1]$. ∎

Theorem 44 *(Characterisation of isovariant-filtered scale-isovariants)*
Let $l > 0$, $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$, $\mathbb{T}_* \overset{\alpha_*}{\rightsquigarrow} \mathbb{A}_*$ and $\eta : \mathbb{A} \to \mathbb{A}_*$. Further, let η_0 be a $\mathbb{T}$-isovariant on $\mathbb{A}$. Then the following two statements are equivalent:
[1] *η is an η_0-filtered l^{th}-order (α, α_*)-scale-isovariant from $\mathbb{A}$ to $\mathbb{A}_*$.*
[2] *$\eta = \zeta \circ \eta_0$ for some l^{th}-order $({}^{\eta_0}\alpha, \alpha_*)$-scale-isovariant ζ from $\eta_0(\mathbb{A})$ to $\mathbb{A}_*$.*

Proof. Consider the following two statements:
[1′] For some $\zeta \in \mathrm{dom}(\eta_0(\mathbb{A}))$, $\eta = \zeta \circ \eta_0$ is a $\mathbb{T}$-isovariant on $\mathbb{A}$ with

$${}^{(\zeta\circ\eta_0)}(\alpha(\kappa)) = \alpha_*(\kappa^l)|_{(\zeta\circ\eta_0)(\mathbb{A})}, (\kappa > 0).$$

[2′] $\eta = \zeta \circ \eta_0$ for some ${}^{\eta_0}\mathbb{T}$-isovariant ζ on $\eta_0(\mathbb{A})$ with

$${}^{\zeta}({}^{\eta_0}\alpha(\kappa)) = \alpha_*(\kappa^l)|_{\zeta(\eta_0(\mathbb{A}))}, (\kappa > 0).$$

Then [1] ⇔ [1′] and [2] ⇔ [2′] by definition, while [1′] ⇔ [2′] is immediate from Lemma 43 and the definition (9) of ${}^{\eta_0}\alpha$. ∎

Finally, using Theorem 44 with (10), we have

Theorem 45 *(Characterisation of scale-isovariant-filtered scale-isovariants)*
Let $\mathbb{T} \overset{\alpha}{\rightsquigarrow} \mathbb{A}$, $\mathbb{T}_i \overset{\alpha_i}{\rightsquigarrow} \mathbb{A}_i$, $(i = 0, 1)$, η_0 be an (α, α_0)-scale-isovariant from $\mathbb{A}$ to $\mathbb{A}_0$ of order l_0 and $\eta : \mathbb{A} \to \mathbb{A}_1$. Then the following two statements are equivalent:

[1] *η is an η_0-filtered (α, α_1)-scale-isovariant of order $l_1 l_0$*
[2] *$\eta = \eta_1 \circ \eta_0$ for some (α_0, α_1)-scale-isovariant $\eta_1 : \eta_0(\mathbb{A}) \to \mathbb{A}_1$ of order l_1*

so that, when they hold, for $i = 0, 1$:
η_i and η have the same order $\Leftrightarrow \eta_{1-i}$ is positively homogeneous.

10 REFERENCES

Cook, R. D. and Weisberg, S. (1982) *Residuals and Influence in Regression.* Chapman and Hall: London.

Critchley, F., Atkinson, R. A., Lu, G. and Biazi, E. (2001) Influence analysis based on the case sensitivity function. *J. Royal Statistical Society*, B, $63 : 307 - 323$.

Dryden, I. L. and Mardia, K. V. (1998) *Statistical Shape Analysis.* Wiley: New York.

Eaton, M. L. (1988) *Group Invariance Applications in Statistics.* Regional Conference Series in Probability and Statistics, Volume 1, Hayward, California: Institute of Mathematical Statistics.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986) *Robust Statistics: The Approach Based on Influence Functions.* Wiley: New York.

Kawakubo, K. (1991) *The Theory of Transformation Groups*, Oxford University Press.

Kendall, D. G., Barden D., Carne T. K. and Le H. (1999). *Shape and Shape Theory.* Wiley: New York.

Lehmann, E. L. (1991) *Theory of Point Estimation.* California: Wadsworth & Brooks/Cole.

Small, C. G. (1996) *The Statistical Theory of Shape.* Springer: New York.