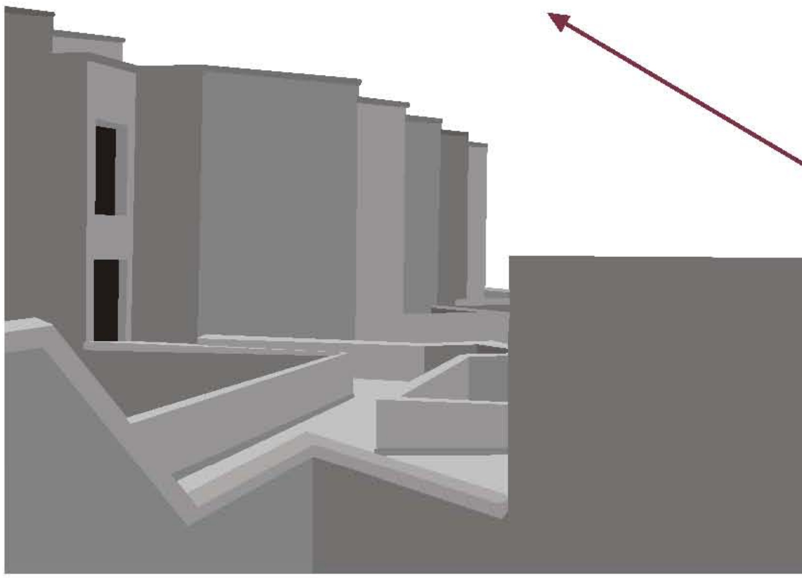

Comunicaciones del CIMAT

FINITE ELLIPTICAL CONFIGURATION DISTRIBUTIONS: INFERENCE AND APPLICATIONS

Francisco J. Caro-Lopera, José A. Díaz-García and Graciela González-Farías

Comunicación del CIMAT No I-08-15/25-09-2008
(PE/CIMAT)

CIMAT

Finite Elliptical Configuration Distributions: Inference and Applications

Francisco J. Caro-Lopera

Department of Basic Mathematics
Centro de Investigación en Matemáticas
Callejón de Jalisco s/n, Mineral de Valenciana
36240 Guanajuato, Guanajuato
México

José A. Díaz-García

Universidad Autónoma Agraria Antonio Narro
Department of Statistics and Computation
25315 Buenavista, Saltillo
Coahuila, México

Graciela González-Farías

Department of Statistics
Centro de Investigación en Matemáticas
Callejón de Jalisco s/n, Mineral de Valenciana
36240 Guanajuato, Guanajuato
México

Abstract

Based on the infinite noncentral elliptical configuration density and the matrix generalised Kummer relation, the finite configuration distribution is derived. It avoids the use of approximations, asymptotic distributions and infinite series of zonal polynomials. The general expression serves in the configuration context for transforming the corresponding infinite series density of zonal polynomials into a polynomial of low degree. Applications in Biology, Medicine and image analysis are studied under a non Gaussian distribution a priori ratified by a model dimension criterion.

Key words: Noncentral elliptical configuration density, matrix generalised Kummer relation, zonal polynomials.
PACS: 62E15, 60E05, 62H99.

Email addresses: fjcaro@cimat.mx (Francisco J. Caro-Lopera),

1 Introduction

Recently, Caro-Lopera *et al* (2008a) provided the necessary mathematical tools in integration and partition theory, for deriving the noncentral configuration density of an elliptical model; in particular, exact expressions for the classical elliptical families were obtained: the Kotz, the Pearson VII type, the Bessel and the Logistic configuration densities. Also, they proposed the general procedure for performing inference of any elliptical configuration model and it was set in such manner that some modifications to the published efficient numerical algorithms for confluent infinite series type involving zonal polynomials (Koev and Edelman (2006)), can be used.

Then, Caro-Lopera *et al* (2008b) proposed a further simplification of the closed computational problem: the study of finite configuration densities. They derived a subfamily of finite configurations based on a Kotz type distribution and as a simple example of their use, exact inference for testing configuration mean differences in some applied problems were provided. Thus, by using formulae of low degree zonal polynomials, some two dimensional applications of the shape literature were studied.

Finally, the main result for proving the conjecture about the finite configuration follows from Díaz-García and Caro-Lopera (2008) which provides an extension of the known matrix Kummer relation of Herz (1955); the so called generalised matrix Kummer relation is based on a function which admits a Taylor expansion in zonal polynomials.

Section 2 provides a survey of the infinite noncentral elliptical configuration density and the matrix generalised Kummer relation. Then, the main result of this paper is proposed in Section 3 and a non Gaussian finite density is derived as corollary, it will support the applications. In Section 4 the inference procedure is outlined and applied in Section 5 with a non gaussian model ratified by the Schwarz criterion, it also studies two dimensional applications based on exact formulae for zonal polynomials of Caro-Lopera *et al* (2007). The applications include: mouse vertebra, gorilla skulls, girl and boy craniofacial studies, brain MR scans of schizophrenic patients and postcode recognition.

2 Preliminary results

First, we give a summary of the matrix generalised Kummer relation.

---

jadiaz@uaaan.mx (José A. Díaz-García), farias@cimat.mx (Graciela González-Farías).

Definition 1 *Let* $\mathbf{X} > 0$ *be an* $m \times m$ *positive definite matrix. The hypergeometric generalised function* ${}_1P_1$ *of matrix argument is defined by*

$$ {}_1P_1(f(t, \operatorname{tr}(\mathbf{X})) : a; c; \mathbf{X}) = \sum_{t=0}^{\infty} \frac{f(t, \operatorname{tr}(\mathbf{X}))}{t!} \sum_{\tau} \frac{(a)_\tau}{(c)_\tau} C_\tau(\mathbf{X}), \qquad (1)$$

where $\sum_\tau$ *denotes the summation over all partitions* τ, $\tau = (t_1, \cdots, t_m)$, $t_1 \geq t_2 \cdots \geq t_m > 0$, *of* t, $C_\tau(X)$ *is the zonal polynomial of* X *corresponding to* τ, *the function* $f(t, \operatorname{tr}(\mathbf{X}))$ *is independent of* τ *and the generalised hypergeometric coefficient* $(b)_\tau$ *is given by*

$$ (\beta)_\tau = \prod_{i=1}^{m} \left(\beta - \frac{1}{2}(i-1) \right)_{t_i}, $$

where

$$ (b)_t = b(b+1) \cdots (b+t-1), \quad (b)_0 = 1. $$

Here $\mathbf{X}$, the argument of the function, is a complex symmetric $m \times m$ matrix and the parameters a, c are arbitrary complex numbers. The parameter c can not be zero or an integer or a half-integer $\leq (m-1)/2$. If the parameter a is a negative integer, say, $a = -l$, then the function (1) is a polynomial of degree ml, because for $t \geq ml + 1$, $(a)_\tau = (-l)_\tau = 0$, see Muirhead (1982, p. 258). In particular note that, ${}_1P_1(1 : a; c; \mathbf{X}) = {}_1F_1(a; c; \mathbf{X})$.

So, using this notation we see that the Kummer relation of Herz (1955), ${}_1F_1(a; c; \mathbf{X}) = \operatorname{etr}(\mathbf{X})\, {}_1F_1(c-a; c; -\mathbf{X})$, is a particular case of a general type of expressions with the following form

$$ {}_1P_1\left(f^{(t)}(0) : a; c; \mathbf{X}\right) = {}_1P_1\left(f^{(t)}(\operatorname{tr}(\mathbf{X})) : c-a; c; -\mathbf{X}\right), \qquad (2)$$

where $f^{(t)}(y)$ denotes the t-th derivative of the function $f(y)$ which admits a Taylor expansion in zonal polynomials.

Now, the present work is based on the following result (Díaz-García and Caro-Lopera (2008)):

Theorem 2 *If* $f(y)$ *admits a Taylor expansion in zonal polynomials, then the generalised Kummer relation is given by*

$$ {}_1P_1\left(f^{(t)}(0) : a; c; \mathbf{X}\right) = {}_1P_1\left(f^{(t)}(\operatorname{tr}(\mathbf{X})) : c-a; c; -\mathbf{X}\right), \qquad (3)$$

where $\mathbf{X} > 0$, $\Re(c) > (m-1)/2$ *and* a *is arbitrary (or at least* $\Re(a) > (m-1)/2$, *if the integral representation of* ${}_1F_0$ *is used, see Herz (1955, p. 485) or Muirhead (1982, Corollary 7.3.5)).*

We end this section given the general expression of the infinite noncentral configuration density.

First we recall the basic definitions of elliptical distributions and configurations (see Gupta and Varga (1993) and Goodall and Mardia (1993), respectively).

We say that $\mathbf{X} : N \times K$ has a matrix variate elliptically contoured distribution if its density respect to the Lebesgue measure is given by:

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{|\mathbf{\Sigma}|^{K/2}|\mathbf{\Theta}|^{N/2}} h(\mathrm{tr}((\mathbf{X}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{X}-\mu)\mathbf{\Theta}^{-1})),$$

where $\mu : N \times K$, $\mathbf{\Sigma} : N \times N$, $\mathbf{\Theta} : K \times K$, $\mathbf{\Sigma}$ positive definite ($\mathbf{\Sigma} > \mathbf{0}$), $\mathbf{\Theta} > \mathbf{0}$. Such a distribution is denoted by $\mathbf{X} \sim \mathcal{E}_{N\times K}(\mu, \mathbf{\Sigma}, \mathbf{\Theta}, h)$.

Definition 3 *Two figures $\mathbf{X} : N \times K$ and $\mathbf{X}_1 : N \times K$ have the same configuration, or affine shape, if $\mathbf{X}_1 = \mathbf{X}\mathbf{E} + \mathbf{1}_N e'$, for some translation $e : K \times 1$ and a nonsingular $\mathbf{E} : K \times K$.*

The configuration coordinates are constructed in two steps summarized in the expression

$$\mathbf{L}\mathbf{X} = \mathbf{Y} = \mathbf{U}\mathbf{E}. \tag{4}$$

The matrix $\mathbf{U} : N-1 \times K$ contains configuration coordinates of $\mathbf{X}$. Let $\mathbf{Y}_1 : K \times K$ be nonsingular and $\mathbf{Y}_2 : q = N - K - 1 \geq 1 \times K$, such that $\mathbf{Y} = (\mathbf{Y}_1' \mid \mathbf{Y}_2')'$. Define also $\mathbf{U} = (\mathbf{I} \mid \mathbf{V}')'$, then $\mathbf{V} = \mathbf{Y}_2\mathbf{Y}_1^{-1}$ and $\mathbf{E} = \mathbf{Y}_1$.

Where $\mathbf{L}$ is an $N-1 \times N$ Helmert sub-matrix.

Then, the general case of the configuration density under a non-isotropic non-central elliptical model is the following (Caro-Lopera *et al* (2008a)):

Theorem 4 *If $\mathbf{Y} \sim \mathcal{E}_{N-1\times K}(\mu_{N-1\times K}, \mathbf{\Sigma}_{N-1\times N-1} \otimes \mathbf{I}_K, h)$, for $\mathbf{\Sigma}$ positive definite ($\mathbf{\Sigma} > \mathbf{0}$), $\mu \neq 0_{N-1\times K}$, then the configuration density is given by*

$$\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\mathbf{\Sigma}|^{\frac{K}{2}}|\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U}|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \sum_{t=0}^{\infty} \frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)} \sum_{r=0}^{\infty} \frac{1}{r!}\left[\mathrm{tr}\left(\mu'\mathbf{\Sigma}^{-1}\mu\right)\right]^r$$
$$\times \sum_{\tau} \frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau} C_\tau(\mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1})S, \tag{5}$$

where

$$S = \int_0^\infty h^{(2t+r)}(y) y^{\frac{K(N-1)}{2}+t-1} dy < \infty. \tag{6}$$

Now, we have the tools for deriving the main result of this paper.

3 Finite noncentral elliptical configuration density

Our motivation for studying finite shape densities, comes from the elliptical extension of the finite gaussian configuration density and the open problem for computations of confluent series type involved in the infinite configuration densities. It is known, that the zonal polynomials are computable very fast by Koev and Edelman (2006), but the problem now resides in the convergence and the truncation of the series of zonal polynomials. In fact, in the same cited reference we read:

> *"Several problems remain open, among them automatic detection of convergence .... and it is unclear how to tell when convergence sets in. Another open problem is to determine the best way to truncate the series. "*

Thus the implicit numerical difficulties for truncation of any configuration density motivate two areas of investigation: one, continue the numerical approach started by (Koev and Edelman (2006)) with the confluent hypergeometric functions and extend it to the case of some configuration series type Kotz, Pearson VII, Bessel, Logistic; or second, find an alternative way, generalise the gaussian configuration density, which become finite after the Kummer relation, to a possible finite elliptical configuration density based on the corresponding generalised Kummer relation.

In this section, we provide the above mentioned second approach.

First let us denote the elliptical configuration density of Theorem 4 by

$$A_1P_1(g(t,\mathbf{X}):a;c;\mathbf{X}), \tag{7}$$

where

$$ {}_1P_1(g(t,\mathbf{X}):a;c;\mathbf{X})=\sum_{t=0}^{\infty}\frac{g(t,\mathbf{X})}{t!}\sum_{\tau}\frac{(a)_\tau}{(c)_\tau}C_\tau(\mathbf{X}),$$

see (1),

$$A=\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\mathbf{\Sigma}|^{\frac{K}{2}}|\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U}|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \tag{8}$$

$$g(t,\mathbf{X})=\sum_{r=0}^{\infty}\frac{\left[\operatorname{tr}\left(\mu'\mathbf{\Sigma}^{-1}\mu\right)\right]^r}{r!\Gamma\left(\frac{K(N-1)}{2}+t\right)}\int_0^\infty h^{(2t+r)}(y)y^{\frac{K(N-1)}{2}+t-1}dy, \tag{9}$$

$$\mathbf{X}=\mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1},\ a=\frac{N-1}{2},\ c=\frac{K}{2}.$$

Unfortunately, the configuration density $A_1P_1(g(t,\mathbf{X}):a;c;\mathbf{X})$ is an infinite series, given that a and c are positive, (recall that N is the number of landmarks, K is the dimension and $N-K-1\geq 1$). So a truncation is needed if we want to perform inference with the modified confluent hypergeometric type series algorithms.

Now, the finiteness of the general configuration density (7) follows from theorem 2.

For any $g(\cdot)$ in (9), associated to an elliptical model $h(\cdot)$ in (5), exists a unique $f(\cdot)$ which admits a Taylor expansion, such that

$$g(t, \mathbf{X}) = f^{(t)}(0). \tag{10}$$

In other words, $g(t)$ is the coefficient of $\frac{y^t}{t!}$ in the Taylor expansion

$$f(y) = \sum_{t=0}^{\infty} \frac{f^{(t)}(0)}{t!} y^t = \sum_{t=0}^{\infty} \frac{g(t)}{t!} y^t.$$

Then with (10) and Theorem 2, we have that:

Theorem 5 *If* $\mathbf{Y} \sim E_{N-1\times K}(\mu_{N-1\times K}, \mathbf{\Sigma}_{N-1\times N-1} \otimes \mathbf{I}_K, h)$, $\mathbf{\Sigma} > 0$, K *is even (odd) and* N *is odd (even), respectively, then the finite non-isotropic noncentral configuration density is given by*

$$A_1P_1\left(f^{(t)}(\mathrm{tr}(\mathbf{X})) : -\left(\frac{N-1}{2} - \frac{K}{2}\right); \frac{K}{2}; -\mathbf{X}\right), \tag{11}$$

and it is a polynomial of degree $K\left(\frac{N-1}{2} - \frac{K}{2}\right)$ *in the latent roots of* $\mathbf{X} = \mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1}$. *Where*

$$A = \frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\mathbf{\Sigma}|^{\frac{K}{2}}|\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U}|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)},$$

$f(y)$ *admits a Taylor expansion and it is uniquely defined by*

$$f^{(t)}(0) = g(t, \mathbf{X}),$$

via

$$g(t, \mathbf{X}) = \sum_{r=0}^{\infty} \frac{\left[\mathrm{tr}\left(\mu'\mathbf{\Sigma}^{-1}\mu\right)\right]^r}{r!\Gamma\left(\frac{K(N-1)}{2} + t\right)} \int_0^{\infty} h^{(2t+r)}(y) y^{\frac{K(N-1)}{2}+t-1} dy.$$

Proof. The result follows by taking $a = \frac{N-1}{2}$ and $c = \frac{K}{2}$ in Theorem 2 and noting that for $N - K - 1 \geq 1$, even (odd) K and odd (even) N, respectively, then $c - a$ is a negative integer and by Caro-Lopera *et al* (2008b), this implies that ${}_1P_1\left(f^{(t)}(\mathrm{tr}(\mathbf{X})) : -\left(\frac{N-1}{2} - \frac{K}{2}\right); \frac{K}{2}; -\mathbf{X}\right)$ is a polynomial of degree

$K\left(\frac{N-1}{2}-\frac{K}{2}\right)$ in the latent roots of $\mathbf{X}$. Finally the relation between g and f in (10) expresses the finite configuration density in terms of f. $\square$

Alternatively, Theorem 5 can be stated as follows:

Theorem 6 *Let* $\mathbf{Y} \sim E_{N-1\times K}(\mu_{N-1\times K}, \mathbf{\Sigma}_{N-1\times N-1} \otimes \mathbf{I}_K, h)$*,* $\mathbf{\Sigma} > 0$*,* K *is even (odd) and* N *is odd (even), respectively, then the noncentral configuration density is finite invariant under the elliptical family and is given by (11).*

We must highlight that every noncentral elliptically contoured density h deserves a deeply study in order to obtain the corresponding function f of Theorem 5.

Finally, we derive a particular finite configuration density based on a Kotz model, which will support the applications in the final section.

Corollary 7 *If* $\mathbf{Y} \sim \mathcal{E}_{N-1\times K}(\mu_{N-1\times K}, \mathbf{\Sigma}_{N-1\times N-1} \otimes \mathbf{I}_K, h)$*, with* $\mathbf{\Sigma} > \mathbf{0}$*, even (odd)* K*, odd (even)* N*, respectively, then the finite non-isotropic noncentral Kotz I (*$T=3$*) configuration density is given by*

$$
\begin{aligned}
&A(\pi^p p)^{-1} \exp(\operatorname{tr}\mathbf{X} + w) \\
&\times {}_1P_1\left(\left[(p - w - \operatorname{tr}\mathbf{X})^2 + p\right] - \left[2(p - w - \operatorname{tr}\mathbf{X}) + 1\right] t + t^2 : c-a; c; -\mathbf{X}\right),
\end{aligned}
$$

and it is a polynomial of degree $K\left(\frac{N-1}{2}-\frac{K}{2}\right)$ *in the latent roots of* $\mathbf{X} = \mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1}$*,*
where $p = \frac{K(N-1)}{2}$*,* $w = \operatorname{tr}(-R\mu'\mathbf{\Sigma}^{-1}\mu)$*,* $a = \frac{N-1}{2}$*,* $c = \frac{K}{2}$ *and* A *is given in (8).*

Proof. From Caro-Lopera *et al* (2008a), if $\mathbf{Y} \sim \mathcal{E}_{N-1\times K}(\mu_{N-1\times K}, \mathbf{\Sigma}_{N-1\times N-1} \otimes \mathbf{I}_K, h)$, with $\mathbf{\Sigma} > \mathbf{0}$, then the Kotz type $s = 1$ non-isotropic noncentral configuration density is given by

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\mathbf{\Sigma}|^{\frac{K}{2}}|\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U}|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(T-1+\frac{K(N-1)}{2}\right)}$$
$$\times\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(-R\mu'\mathbf{\Sigma}^{-1}\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(R\mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1}\right)$$
$$\times\left\{\Gamma\left(T-1+\frac{K(N-1)}{2}+t\right)\right.$$
$$\left.+\sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right](-1)^m\Gamma\left(T-1-m+\frac{K(N-1)}{2}+t\right)\right\}.$$

Taking $T=3$, we have that

$$g(t,\mathbf{X})=(\pi^p p)^{-1}\exp(w)\left[\left((p-w)^2+p\right)-(2(p-w)+1)\,t+t^2\right]=f^{(t)}(0).$$

Which implies that

$$f(y)=(\pi^p p)^{-1}\exp(y+w)\left[\left((p-w)^2+p\right)-2(p-w)y+y^2\right]=\sum_{t=0}^{\infty}\frac{f^{(t)}(0)}{t!}y^t,$$

then

$$f^{(t)}(y)=(\pi^p p)^{-1}\exp(y+w)\left[\left((p-w-y)^2+p\right)-(2(p-w-y)+1)\,t+t^2\right],$$

and the required result follows. □

4 Exact inference for finite elliptical configuration models

By using the previous section, the full group of elliptical densities becomes finite, and so the addressed inference procedure can be improved in such way that it can be performed with the existing formulae for zonal polynomials available since the 60's.

The proposal is to use the elliptically contoured distribution to model population configurations (11) for some particular cases. For this, consider a random sample of n independent and identically distributed observations $\mathbf{U}_1,\ldots,\mathbf{U}_n$ obtained from

$$\mathbf{Y}_i\sim E_{N-1\times K}(\mu_{N-1\times K},\sigma^2\mathbf{I}_{N-1}\otimes\mathbf{I}_K,h),\quad i=1,\ldots,n,$$

by mean of (4).

Let $CD(\mathbf{U};\mathcal{U},\sigma^2)$ be the exact configuration density, where $\mathcal{U}$ is the location parameter matrix of the configuration population and σ^2 is the population scale parameter. Both $\mathcal{U}$ and σ^2 are the parameters to estimate. More exactly, let $\mu \neq \mathbf{0}_{N-1\times K}$ be the parameter matrix of the elliptical density $\mathbf{Y}$ considered in Theorem 4; if we write it as $\mu = (\mu_1' \mid \mu_2')'$, where $\mu_1 : K\times K$ (nonsingular) and $\mu_2 : q = N-K-1 \geq 1\times K$, then, according to (4), we can define the configuration location parameter matrix $\mathcal{U} : N-1\times K$ as follows: $\mathcal{U} = (\mathbf{I}_K \mid \mathcal{V}')'$ where $\mathcal{V} = \mu_2\mu_1^{-1}$; and $\mathcal{V} : q = N-K-1 \geq 1\times K$ contains $q\times K$ configuration location parameters to estimate. Then, the maximum likelihood estimators for location and scale parameters are

$$(\widetilde{\mathcal{V}}, \widetilde{\sigma^2}) = \arg\sup_{\mathcal{V},\sigma^2} \log L(\mathbf{U}_1,\ldots,\mathbf{U}_n;\mathcal{V},\sigma^2). \tag{12}$$

Given that the likelihood function is a polynomial of low degree, the numerical optimization can be performed easily. The initial point for the routines can be set as follows, consider the Helmertized landmark data $\mathbf{Y}_i \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2\mathbf{I}_{N-1}\otimes\mathbf{I}_K, h)$ $i=1,\ldots,n$ (see (4)) and let $\widetilde{\mu} = (\widetilde{\mu}_1' \mid \widetilde{\mu}_2')'$ and $\tilde{\sigma}^2$ be the maximum likelihood estimators of the location parameter matrix $\mu_{N-1\times K}$ and the scale parameter σ^2 of the elliptical distribution under consideration, so, given that

$$\mathbf{U}_i'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}_i(\mathbf{U}_i'\mathbf{\Sigma}^{-1}\mathbf{U}_i)^{-1} = \mathbf{Y}_i'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{Y}_i(\mathbf{Y}_i'\mathbf{\Sigma}^{-1}\mathbf{Y}_i)^{-1},$$

then an initial point can be $x_0 = (vec'(\mathcal{V}_{\mathbf{0}}'), \sigma_0^2)$, where $\mathcal{V}_0 = \widetilde{\mu}_2\widetilde{\mu}_1^{-1}$ and $\sigma_0^2 = \tilde{\sigma}^2$.

So, the exact inference procedure can be outlined in the next few steps:

(1) Available distributions, families of finite elliptical configuration densities: Consider a list of finite configuration densities, some of them, including the finite Kotz type for a positive integer T, can be derived as in corollary 7, but a more enriched list including other subfamilies of Kotz, Bessel and Logistic can be established via Theorem 5. We must note that any elliptical function $h(\cdot)$ which satisfies the conditions of theorems 4 and 5 is appropriate. Most of the applications in statistical theory of shape reside on the isotropic model (see Dryden and Mardia (1998)) $\mathbf{\Sigma} = \sigma^2\mathbf{I}_{N-1}$. Also, we can consider a more enriched structure, for example $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N-1}^2)$, and similar diagonal structures. In this paper we will perform inference with the isotropic version of the finite non Gaussian distribution of corollary 7; but we must highlight that the present procedure can be studied with any finite configuration density.

(2) Choosing the elliptical configuration density: Consider k elliptical models, then perform the maximization of the likelihood function separately for each model $j=1,\ldots,k$, obtaining say, $M_j(\mathbf{Y}_1,\ldots,\mathbf{Y}_n)$, then Schwarz's

criterion for a large-sample is given by (Schwarz (1978))

$$\textit{Choose the model for which} \quad \log M_j(\mathbf{Y}_1, \ldots, \mathbf{Y}_n) - \frac{1}{2}k_j \log n \quad \textit{is largest},$$

where k_j is the dimension (number of parameters) of the model j.

(3) Mean Configuration: Once the elliptical model is selected, we find the estimators of location and scale parameters of configuration by mean of (12). The crucial point here is the computation of the configuration density; but as we proved in theorem 5, any configuration density is a polynomial of degree $K\left(\frac{N-1}{2} - \frac{K}{2}\right)$ in the latent roots of $\mathbf{X} = \mathbf{U}'\mathbf{\Sigma}^{-1}\mu\mu'\mathbf{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\mathbf{\Sigma}^{-1}\mathbf{U})^{-1}$, when the number of landmarks N of the figure is selected odd (even) according to the even (odd) dimension K, respectively. For the published shape theory applications, see for example Dryden and Mardia (1998), the maximum number of landmarks considered is 21 in two dimensions, which supposes a configuration density reduced to a polynomial of degree 18 in the two eigenvalues of the corresponding matrix, and recall that these zonal polynomials of lower degree are known since 60's, in fact for two dimensional applications there are exact formulae easily computable for these cases, see James (1968), Caro-Lopera *et al* (2007).

(4) Hypothesis testing: Finally, given that the likelihood can be evaluated and optimized, then a sort of likelihood ratio tests can be performed for testing a particular configuration for a population, or testing for differences in configuration between two populations, or testing one-dimensional uniform shear of two populations, etc. The large sample standard likelihood ratio tests are the most frequently used, see for example Dryden and Mardia (1998), by mean of Wilk's theorem.

Thus, the whole inference procedure of the above four steps can be carried out for a particular landmark data (for example from Dryden and Mardia (1998), Bookstein (1991)), and we can consider the inference problem in configuration densities solved.

5 Applications

In this section we consider planar classical applications in the statistical shape analysis. The following situations are sufficiently studied by shape based on euclidian transformations and asymptotic formulae. They are also studied under the finite gaussian configuration density in Caro-Lopera *et al* (2008b). We will study them with a non Gaussian finite configuration density which is a better model according to Schwarz (1978). Then, exact formulae from James (1968), or Caro-Lopera *et al* (2007) are used for computing the corresponding low polynomials.

We will test configuration differences under the exact finite configuration density of Corollary 7, and the applications include mouse vertebra, gorilla skulls, girl and boy craniofacial studies, brain MR scans of schizophrenic patients and postcode recognition; these problems were full detailed by Dryden and Mardia (1998), Bookstein (1991) in the Euclidean transformation case and by Caro-Lopera *et al* (2008b) in the gaussian configuration case, and we refer the reader for problems definitions and comparison, here we just provide the Schwarz criterion, the maximum likelihood estimators and the hypothesis testing.

In each problem we give the Schwarz criterion. It ratifies in all the cases that the non gaussian case, the Kotz model for $s = 1$ and $T = 3$ (denoted by K) is better than the gaussian model (G).

5.1 Mouse vertebra

Inference is based on a confluent polynomial of two degree in the two eigenvalues of the zonal polynomial argument, then the corresponding Schwarz criterion, and the maximum likelihood estimators for location and scale parameters of the three groups are given in Table 5.1.

Group	Schwarz's criterion $\begin{smallmatrix} G \\ K \end{smallmatrix}$	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\sigma^2}$
Control	193.11 241.21	-0.10706	0.15531	0.0005196	-0.96918	0.001522
Large	133.88 168.75	-0.084809	0.12754	0.00046022	-1.0849	0.0025147
Small	138.13 170.94	-0.09291	0.21516	0.00049112	-1.0249	0.0014573

Table 5.1. Mouse vertebra.

Tests for scale parameters show differences between C-L and L-S, but equality in the C-S case. The likelihood ratios (based on $-2\log\Lambda \approx \chi^2_4$) for the tests $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$, provide the p-values: 7E-10 for C-L; 5.5E-9 for L-S and 1.52E-5 for C-S. So, we can say that the three groups have different mean configurations.

5.2 *Gorilla skulls*

The Schwarz's criterion and the estimators of the configuration location and scale parameters for the two groups are given in Table 5.2.

Group	Schwarz's criterion $\frac{G}{K}$	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\sigma^2}$
Female	247.97 303	-0.28044	0.31254	-0.4241	-0.59859	0.27498	-1.477	0.73698	-1.2682	0.0056265
Male	225.09 276.9	-0.33364	0.42289	-0.43669	-0.57459	0.30191	-1.2947	0.73113	-1.0522	0.0052423

Table 5.2. Gorilla skulls.

A test for scale parameters shows significant difference between the two sexes, thus the likelihood ratio (based on $-2\log\Lambda \approx \chi^2_8$) for $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranial difference between the sexes of the apes, provides a p-value of 9E-23. It explains strong evidence for differences in configuration locations in both sexes.

5.3 *The university school study subsample*

In this case, the Schwarz's criterion and the estimators of the configuration location and scale parameters for the two groups are given in Table 5.3.

Group	Schwarz's criterion $\frac{G}{K}$	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\sigma^2}$
Male	318.32 384.55	-1.243	2.2021	0.4629	-1.3781	-0.91416	0.66198	0.15916	-0.071978	0.0042093
Female	240.01 289.53	-1.2512	2.2322	0.43771	-1.3843	-0.92988	0.70325	0.16349	-0.078691	0.0040228

Table 5.3. The university school study subsample.

A test for scale parameters between the two populations show important differences, so the likelihood ratio (based on $-2\log\Lambda \approx \chi^2_8$) for $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranialfacial difference between the boys and girls, gives a p-value of 4.73E-1. Thus there is not a significant difference between the configuration locations of both populations. It is in agreement with Bookstein (1991) in a different shape context.

5.4 Brain MR scans of schizophrenic patients

The Schwarz's criterion and the estimators of the configuration location and scale parameters for the two groups are given in Table 5.4.

Group	Schwarz's criterion $\begin{smallmatrix} G \\ K \end{smallmatrix}$	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$
Normal	151.91 190.98	-0.63748	2.6977	-1.2779	-2.8394	-0.42348	-1.0037
Squizo	155.21 193.84	-0.68378	2.393	-1.1457	-2.8457	-0.37395	-1.0762

$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\mathcal{V}_{51}}$	$\widetilde{\mathcal{V}_{52}}$	$\widetilde{\mathcal{V}_{61}}$	$\widetilde{\mathcal{V}_{62}}$
-0.31385	-2.313	-0.30722	-3.5306	0.35869	-0.90155
-0.23316	-2.1941	-0.20335	-3.3257	0.38057	-0.84656

$\widetilde{\mathcal{V}_{71}}$	$\widetilde{\mathcal{V}_{72}}$	$\widetilde{\mathcal{V}_{81}}$	$\widetilde{\mathcal{V}_{82}}$	$\widetilde{\mathcal{V}_{91}}$	$\widetilde{\mathcal{V}_{92}}$
0.15615	-2.2229	0.851	-0.75601	1.8709	0.86781
0.2013	-2.1099	0.84561	-0.56803	1.7927	0.88583

$\widetilde{\mathcal{V}_{10,1}}$	$\widetilde{\mathcal{V}_{10,2}}$	$\widetilde{\sigma^2}$
-0.14238	0.20834	0.015538
-0.079099	0.13773	0.012046

Table 5.4. Brain MR scans of schizophrenic patients.

A test for scale parameters shows significant differences in the two populations. The corresponding test for mean configuration differences, based on $-2\log\Lambda \approx \chi^2_{20}$, gives a p-value of 1.2E-2. Dryden and Mardia (1998) conclude a mean shape difference, but in our case the configuration difference is definitely insignificant. This suggest a comparison of both methods by using a dimension model criterion based on a small sample study, however, the configuration conclusion ratifies some studies about the classification of schizophrenia by MR scans.

5.5 Postcode recognition

Table 5.5 shows the Schwarz criterion, the configuration location and scale parameter estimates, and the configuration coordinates of a template number 3 digit, with two equal sized arcs, and 13 landmarks (two coincident) lying on two regular octagons see Dryden and Mardia (1998), p.153.

Group	Schwarz's criterion $\frac{G}{K}$	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$
Digit 3	-856.06 -750.14	-0.80231	1.9375	-2.13	1.5841	-2.7378	0.81656
Template		-2.0908	2.2071	-4.0409	2.8051	-4.5904	2.2904

$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\mathcal{V}_{51}}$	$\widetilde{\mathcal{V}_{52}}$	$\widetilde{\mathcal{V}_{61}}$	$\widetilde{\mathcal{V}_{62}}$
-2.8289	-0.065051	-2.5941	0.71421	-2.7185	1.2938
-4.2069	1.3688	-3.3126	1.7582	-3.5881	2.7053

$\widetilde{\mathcal{V}_{71}}$	$\widetilde{\mathcal{V}_{72}}$	$\widetilde{\mathcal{V}_{81}}$	$\widetilde{\mathcal{V}_{82}}$	$\widetilde{\mathcal{V}_{91}}$	$\widetilde{\mathcal{V}_{92}}$
-3.1824	1.6758	-3.8353	1.3376	-4.0863	0.33044
-5.4996	4.0629	-7.5557	4.8428	-8.2514	4.4208

$\widetilde{\mathcal{V}_{10,1}}$	$\widetilde{\mathcal{V}_{10,2}}$	$\widetilde{\sigma^2}$
-3.7925	-0.65763	0.1023
-6.9108	2.8899	

Table 5.5. Postcode recognition.

The important difference is proved by a test based on $-2\log\Lambda \approx \chi^2_{20}$, with approximately zero p-value. So there is strong evidence that the configuration location does not have the configuration of the ideal template for digit 3.

Some other studies are suggested, the comparison of the best model of transformation (Euclidean vs Configuration) via Schwarz criterion, however all the shape densities (based on Euclidean transformation) lack of the finite property, and the inference involving infinite series of zonal polynomials again is problematic.

Acknowledgment

This research work was partially supported by CONACYT-México, research grant no. 81512 and IDI-Spain, grant MTM2005-09209. This paper was written during J. A. Díaz- García's stay as a visiting professor at the Department of Statistics and O. R. of the University of Granada, Spain.

References

F. L. Bookstein. (1991). *Morphometric tools for landmark data.* Cambridge, England: Cambridge University Press.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2007). A formula for Jack polynomials of the second order. *Appl. Math.* (Warsaw) 34, 113–119.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2008a). Noncentral elliptical configuration density. Submitted.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías.(2008b). Inference in statistical shape theory: Elliptical configuration densities. Submitted.

J. A. Díaz-García and F. J. Caro-Lopera. (2008). Matrix generalised Kummer relation. Submitted.

I. L. Dryden and K.V. Mardia. (1998). *Statistical shape analysis.* John Wiley and Sons, Chichester.

C. R. Goodall, and K.V. Mardia. (1993). Multivariate Aspects of Shape Theory. *Ann. Statist.*, 21, 848–866.

A. K. Gupta, and T. Varga. (1993). *Elliptically Contoured Models in Statistics*, Kluwer Academic Publishers, Dordrecht.

C. S. Herz. (1955). Bessel Functions of Matrix Argument. *Ann. Math.*, 61, 474–523.

A. T. James. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist.*, 39, 1711–1718.

P. Koev and A. Edelman. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comp.* 75, 833–846.

R. J. Muirhead. (1982). *Aspects of multivariate statistical theory.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc.

G. Schwarz. (1978). Estimating the dimension of a model. *Ann. Statist.*, 6 (2), 461–464.