Sankhyā : The Indian Journal of Statistics
2016, Volume 78-A, Part 1, pp. 19-42

Geodesic Hypothesis Testing for Comparing Location Parameters in Elliptical Populations

Patricia Giménez, Jorge N. López and Lucas Guarracino
Universidad Nacional de Mar del Plata, Mar del Plata, Argentina

Abstract

In this paper we study the geometry of the differentiable manifold associated with two samples of symmetric distributions in the real line equipped with the Fisher information as Riemannian metric. Expressions for the entries of the information matrix are obtained under different assumptions. The geodesic or Rao distance induced by this geometry is used to construct asymptotic parametrization-invariant testing procedures for comparing location parameters. As special cases, we obtain new asymptotic tests for the two sample Behrens-Fisher and Fieller-Creasy problems. Testing equality of several location parameters is also considered. It is shown that when scale parameters are equal, the geodesic test statistic is a strictly monotone increasing function of the Wald statistic. Empirical results for the Student-t distribution provide evidence that the geodesic test statistic has good sampling properties in terms of level and power.

AMS (2000) *subject classification.* Primary 62F03; Secondary 53C99.
Keywords and phrases. Statistical manifolds, Fisher information, Geodesic distance, Hypothesis testing, Elliptical distributions, Comparing location parameters

1 Introduction

The classical statistics for testing hypothesis in parametric models use a variety of ways of measuring how far the unrestricted maximum likelihood estimator is from the null model. Particularly, the Wald test does so at the unrestricted maximum likelihood estimator in terms of an arbitrary measure of distance on the manifold which is determined by the way the restrictions are formulated. This distance is not an intrinsic quantity, meaning that it depends on the parametrization used. Gregory and Veall (1985) showed that this non-invariance can lead to drastically different conclusions, even in a very simple model.

Differential geometrical arguments can be exploited to propose parametrization-invariant testing procedures, considering quantities that can be defined in purely geometrical terms. Bhattacharyya (1943) introduced the information Riemannian geometry approach on a family of probability distributions. Rao (1945) defined the concept of geodesic distance (or Rao distance) between parametric probability distributions of a particular family, which have a manifold structure. The geodesic distance on the manifold is defined by introducing the Fisher information as a Riemannian metric, together with the Levi-Civita connection whose geodesics are curves of minimum length. The geodesic distance between two points, e.g. two distributions, is the same regardless of the parametrizations and it is this distance which might be a good candidate for useful test statistics. For broader accounts of the application of differential geometry to statistics see the review chapters or monographs by Amari (1985), Barndorff-Nielsen, Cox and Reid (1986), Kass (1989), Kass and Vos (1997) and Murray and Rice (1993).

The idea of the geodesic testing procedure is to trace out a geodesic from the unrestricted maximum likelihood estimator to the set defined by the null hypothesis, which under appropriate regularity conditions determines a submanifold. That is, to seek a curve of minimum length in the manifold, starting from the unrestricted maximum likelihood estimator and ending somewhere on the null model. The geodesic test statistic is then the square of the minimized arc length, and under standard regularity conditions, it has asymptotically a central chi-squared distribution under the null hypothesis (Kass and Vos, 1997).

The geodesic test based on geodesic distances can be seen, in a sense, as pointed out by Critchley, Marriott and Salmon (1996), as a parametrization-invariant version of the Wald test. Some other alternative geometrical solutions to the non-invariance problem of the Wald statistics include those of Le Cam (1990), which are based on the Hellinger distance and Larsen and Jupp (2003), based on some differential geometrical ideas of the yoke geometry. A $C(\alpha)$ type test, not completely invariant under reparametrization but very nearly so, is discussed by Davidson (1992).

Geodesic distances can be used directly to define a geodesic test, but in general they need to be worked out specifically. A serious practical problem for geodesic tests in higher dimensions is that they require the solution of a system of second order non-linear differential equations for the geodesic curves. Both the solution of the differential equations and the evaluation of the distance can be difficult. Moreover, the point in the null model, which is closest to the unrestricted maximum likelihood estimator must be found by dropping a geodesic which cuts the submanifold orthogonally. The geodesic is unique stating the additional assumption that the manifold is simply connected. Geodesics have been determined for a limited number of distributions. Rao (1987) summarizes a number of these geodesic distances in statistical manifold. Atkinson and Mitchell (1981) derived geodesic estimation for some classes of distributions. Mitchell (1988) investigated the manifold of univariate elliptical distributions. Relevant contributions in this field also include Burbea (1986), Burbea and Rao (1982, 1984), Choi and Kiefer (2011), Efron (1975), Micchelli and Noakes (2005) and Oller and Corcuera (1995) among many others. Some hypothesis testing based on geodesic distances are considered in Skovgaard (1984), for normal distributions, Villarroya and Oller (1993), for the inverse gaussian distribution and Berkane, Oden and Bentler (1997), for multivariate elliptical distributions with equal location. Cubedo and Oller (2002) also proposed using geodesic distances to examine classical hypothesis testing problems from the point of view of model selection.

The aim of the present paper is to study the basic geometry of two samples of univariate elliptical distributions, which as a matter of fact coincides with the class of symmetric distributions on the real line, and derive geodesic test statistics for a general linear hypothesis about location parameters. The cases of equal and unequal scale parameters and different specifications for the probability density function, including independent or pairwise uncorrelated observations, are considered. These usual assumptions lead to manifolds with constant negative sectional curvature or product of them. This fact, although not frequent, as can be seen by results given in Patrangenaru (1994), allows us to evaluate geodesic distances based on the Poincaré hyperbolic space. In particular, for any two points of this manifold there is one and only one geodesic line that joins them (see Hicks, 1965 and Spivak, 1979). Relationship with Wald statistics is explored. As special cases, we have a new asymptotic procedure, which is parametrization invariant, for the well-known Behrens-Fisher and Fieller-Creasy problems.

In Section 2, we introduce the necessary geometric and statistical background in order to define a geodesic testing procedure. In Section 3 we analyse the basic geometry for two samples of univariate elliptical distributions. We obtain the information matrices and geodesic distances on some manifolds, under different assumptions. Geodesic test statistics for a linear hypothesis of location parameters are derived in Section 4. Also, we extend some results to the problem of testing equality of several location parameters. Section 5 presents Monte Carlo simulation results to illustrate the behaviour of the geodesic testing procedure. Concluding remarks are made in Section 6.

2 Statistical Manifolds and Geodesic Tests

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^T$ be a set of p ($p \geq 1$) real, continuous parameters with parameter space Θ, an open subset of $\mathbb{R}^p$.

We consider a general statistical model $\mathcal{M} = \{p(\mathbf{y}; \boldsymbol{\theta}) : \ \boldsymbol{\theta} \in \Theta\}$, as a set of probability density functions, indexed by $\boldsymbol{\theta}$, given the observed data $\mathbf{y} = (y_1, \ldots, y_n)^T$ on a random sample. The family $\mathcal{M}$ is regarded as a manifold, with the parameter $\boldsymbol{\theta}$ playing the role of a coordinate system on it. These coordinates can be changed by admissible (smooth) transformations being three times continuously differentiable and having a non-singular functional determinant. In this way, $\mathcal{M}$ is equipped with a differentiable structure and $\mathcal{M}$ is called a differentiable manifold.

Let

$$\ell(\boldsymbol{\theta}; \mathbf{y}) = \ln p(\mathbf{y}; \boldsymbol{\theta})$$

denote the corresponding log-likelihood function and

$$\partial_i \ell(\boldsymbol{\theta}; \mathbf{y}) = \frac{\partial}{\partial \theta_i} \ln p(\mathbf{y}; \boldsymbol{\theta}), \quad i = 1, \ldots, p,$$

the components of the score function. We assume that standard regularity conditions hold; see, for instance, Amari (1985, Section 2.1).

We can identify the tangent space $T_{\boldsymbol{\theta}}\mathcal{M}$ at each fixed $p(\mathbf{y}; \boldsymbol{\theta}) \in \mathcal{M}$ as the vector space of random variables spanned by $\{\partial_i \ell(\boldsymbol{\theta}; \mathbf{y}), \ \ i = 1, \ldots, p\}$. Under the regularity conditions, this vector space has dimension p, the dimension of $\mathcal{M}$. Then, an inner product

$$g_{ij}(\boldsymbol{\theta}) = \langle \partial_i \ell(\boldsymbol{\theta}; \mathbf{y}), \partial_j \ell(\boldsymbol{\theta}; \mathbf{y}) \rangle = \mathrm{E}_{\boldsymbol{\theta}}[\partial_i \ell(\boldsymbol{\theta}; \mathbf{y}) \partial_j \ell(\boldsymbol{\theta}; \mathbf{y})]$$

is defined from the basis of $T_{\boldsymbol{\theta}}\mathcal{M}$, where $\mathrm{E}_{\boldsymbol{\theta}}$ denotes expectation with respect to $p(\mathbf{y}; \boldsymbol{\theta})$. Of course, the metric tensor $G(\boldsymbol{\theta}) = (g_{ij}(\boldsymbol{\theta}))$ is the Fisher information matrix.

A connection allows us to determine which curves in the manifold shall be called "geodesic" or "straight". Generalizing familiar Euclidean ideas, these are defined to be those curves along which the tangent vector does not change. A metric tensor induces in a natural way an associated connection called the Levi-Civita or metric connection. The Levi-Civita connection has the property that its geodesics are locally length-minimizing curves joining their endpoints. This property is not a general one for geodesics corresponding to non-metric connections as, for example, the one-parameter family of α-connections introduced by Amari (1982).

Let $\boldsymbol{\theta}(t) = (\theta_1(t), \ldots, \theta_p(t))$ denote a curve in Θ joining $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$, where t is the parameter. Suppose that t_1 and t_2, with $t_1 < t_2$, are the values of t such that $\boldsymbol{\theta}(t_1) = \boldsymbol{\theta}^1$ and $\boldsymbol{\theta}(t_2) = \boldsymbol{\theta}^2$. The geodesic distance or Riemannian distance between $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$ is the minimum arc-length of all these curves, that is

$$d\left(\boldsymbol{\theta}^1, \boldsymbol{\theta}^2\right) = \min_{\boldsymbol{\theta}(t)} \int_{t_1}^{t_2} \sqrt{\sum_{i=1}^{p}\sum_{j=1}^{p} g_{ij}(\boldsymbol{\theta}(t)) \frac{d\theta_i(t)}{dt}\frac{d\theta_j(t)}{dt}} dt. \tag{2.1}$$

Such a curve $\boldsymbol{\theta}(t)$ of minimum length is given as the solution of the following second-order non-linear differential equation

$$\frac{d^2\theta_k}{dt^2} + \sum_{i=1}^{p}\sum_{j=1}^{p} \Gamma_{ij}^k \frac{d\theta_i}{dt}\frac{d\theta_j}{dt} = 0, \quad k = 1, \ldots, p,$$

where Γ_{ij}^k is the Christoffel symbol of the first kind for the Levi-Civita connection and is defined by

$$\Gamma_{ij}^k = \frac{1}{2}\sum_{m=1}^{p}\left(\frac{\partial g_{im}(\boldsymbol{\theta})}{\partial\theta_j} + \frac{\partial g_{jm}(\boldsymbol{\theta})}{\partial\theta_i} - \frac{\partial g_{ij}(\boldsymbol{\theta})}{\partial\theta_m}\right) g^{mk}(\boldsymbol{\theta}), \quad i, j, k = 1, \ldots, p,$$

where $(g^{mk}(\boldsymbol{\theta}))$ is the inverse of the matrix $(g_{ij}(\boldsymbol{\theta}))$.

Closed form of geodesic distances have been determined for a limited number of distributions. Rao (1987) summarizes a number of these geodesic distances in statistical manifolds. In particular, Mitchell (1988) obtains the geodesic distance between two univariate elliptical distributions of the same class. The geodesic distance can be used directly to define a geodesic test as follows.

For a hypothesis testing problem

$$H_0\colon \mathbf{h}(\boldsymbol{\theta}) = \mathbf{0} \quad \text{versus} \quad H_1\colon \mathbf{h}(\boldsymbol{\theta}) \neq \mathbf{0},$$

where $\mathbf{h}\colon \Theta \to \mathbb{R}^r$ $(r < p)$, is a vector value function such that the $p \times r$ matrix $H(\boldsymbol{\theta}) = (\partial/\partial\boldsymbol{\theta})\mathbf{h}(\boldsymbol{\theta})$ exists and is continuous in $\boldsymbol{\theta}$ and $\operatorname{rank}(H(\boldsymbol{\theta})) = r$, the null hypothesis selects a subset of the parameter space $\Theta_0 = \{\boldsymbol{\theta} \in \Theta\colon \mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}\}$ and defines a submanifold

$$\mathcal{N} = \{p(\mathbf{y}; \boldsymbol{\theta})\colon \boldsymbol{\theta} \in \Theta_0\},$$

of $\mathcal{M}$ of dimension $p - r$.

The distance from the unrestricted maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ to the submanifold $\mathcal{N}$ is defined as

$$\rho(\hat{\boldsymbol{\theta}}, \mathcal{N}) = \inf\{d(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) \colon \boldsymbol{\theta} \in \Theta_0\}, \tag{2.2}$$

where d is the distance defined in (2.1). Note that $\rho(\hat{\boldsymbol{\theta}}, \mathcal{N})$ is well-defined, since ρ is bounded from below.

The geodesic test statistic of the hypothesis $\mathcal{N}$, based on $\mathbf{y}$ is defined as

$$T = \rho^2(\hat{\boldsymbol{\theta}}, \mathcal{N}).$$

In Section 7.6.1 of Kass and Vos (1997), it is shown that T has asymptotically a central chi-squared distribution with $\dim \mathcal{M}$-$\dim \mathcal{N} = r$ degrees of freedom, under H_0.

Minimization in (2.2) and derivation of the test statistics may be difficult and they need to be worked out specifically, but the idea behind the geodesic test is extremely simple. The null hypothesis is rejected if the distance between the estimated distribution and the distribution under H_0 is too big. We can obtain the test statistic by directly minimizing the distance function or by calculating the geodesic curve from $\hat{\boldsymbol{\theta}}$ which is orthogonal to $\mathcal{N}$.

For calculating the geodesic distances required in this paper, simplifications are possible. The simplifications involve reducing the metric to one of a hyperbolic geometry and using the properties of that geometry. In particular, for any two points of these manifolds, there exists only one geodesic line joining them. The method proceeds as follows.

If we have a p-dimensional manifold with coordinates $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^T$ where the metric tensor is given by

$$G(\boldsymbol{\theta}) = \theta_p^{-2} \operatorname{diag}(A_1, \ldots, A_{p-1}, B), \tag{2.3}$$

with $A_i,\ i = 1, \ldots, p$ and B constants, then we consider the linear transformation $\boldsymbol{\eta} = M\boldsymbol{\theta}$, where $M = \operatorname{diag}(\sqrt{A_1}, \ldots, \sqrt{A_{p-1}}, 1)$. In terms of the new parameter we obtain

$$\tilde{G}(\boldsymbol{\eta}) = \theta_p^{-2} \operatorname{diag}(1, \ldots, 1, B),$$

which is the basic metric of the hyperbolic half-space $\mathbb{H}^p_{\theta_p > 0}$ with constant sectional curvature $K = -1/B$ (do Carmo, 1992; Ratcliffe, 1994). Then, it is not difficult to show that the distance between two points $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots . \hat{\theta}_p)^T$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^T$ can be expressed as

$$d(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sqrt{B} \operatorname{acosh} \left(\frac{\sum_{i=1}^{p-1} A_i B^{-1} (\hat{\theta}_i - \theta_i)^2 + \hat{\theta}_p^2 + \theta_p^2}{2\hat{\theta}_p \theta_p} \right). \tag{2.4}$$

Two sample test problems considered in this paper lead, under the usual statistical assumptions, to manifolds with constant negative sectional curvature or product of them. Many standard statistical distributions also lead to manifolds with this property (Jensen, 1995). However, it is worth noting that this is rarely the case. Indeed, we consider the two dimensional case, where the metric tensor can be expressed as

$$G(\mu, \phi) = \begin{pmatrix} A(\phi)^2 & 0 \\ 0 & B(\phi)^2 \end{pmatrix}$$

and use the Cartan triple method, in Patrangenaru (1994), from the structural Maurer-Cartan equations:

$$d\omega^1 + \omega \wedge \omega^2 = 0, \quad d\omega^2 - \omega \wedge \omega^1 = 0, \quad d\omega = K\, \omega^1 \wedge \omega^2, \tag{2.5}$$

where ω is the connection form, ω^i is the fundamental form and $K = K(\mu, \phi)$ is the curvature. In this case, $\omega^1 = A(\phi)d\mu$ and $\omega^2 = B(\phi)d\phi$.

We can write $\omega = X(\mu, \phi)d\mu + Y(\mu, \phi)d\phi$, then from the two first equations in (2.5), we have $X = A'(\phi)/B(\phi)$ and $Y = 0$, where "$'$" denotes derivative with respect to ϕ. From the third equation in (2.5) we have

$$\begin{aligned} d((A'(\phi)/B(\phi))d\mu) &= K\, \omega^1 \wedge \omega^2 \\ -\frac{\partial(A'(\phi)/B(\phi))}{\partial\phi} d\mu \wedge d\phi &= K A(\phi)B(\phi) d\mu \wedge d\phi. \end{aligned}$$

Therefore, we see that the curvature of this statistical manifold is constant if and only if there is a constant K such that

$$K A(\phi)B(\phi) + (A'(\phi)/B(\phi))' = 0. \tag{2.6}$$

Statistical problems considered in this paper lead to the particular case of $A(\phi) = \sqrt{A_1}/\phi$ and $B(\phi) = \sqrt{B_1}/\phi$, with A_1 and B_1 constants. Then, solving (2.6) we obtain $K = -1/B_1$.

3 Basic Geometry for Two Samples of Univariate Elliptical Distributions

In this section we analyse some statistical manifolds associated to two samples from different univariate elliptical distributions. We derive simple, general expressions for the entries of the information matrices and the geodesic distances in the manifolds, under the assumptions of equal and unequal scale parameters and different specifications for the probability density function.

Elliptical probability distributions, introduced by Kelker (1970) and further discussed by Cambanis, Huang and Simons (1981) and Fang, Kotz and Ng (1990), have received greater interest in the literature as effective tools for multivariate modelling. This class of distributions contains the normal distribution, Student t, contaminated normal, Cauchy and power exponential, among others. Some results on the geometry of elliptical distributions and geodesic estimation in special cases have been given by Berkane et al. (1997) and Mitchell (1988, 1989).

A n-dimensional random vector $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$ is said to have an elliptical distribution with location parameter $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ and scale matrix Ψ, a $n \times n$ positive definite matrix, if its density is of the form

$$p(\mathbf{y}; \boldsymbol{\mu}, \Psi) = |\Psi|^{-1/2} g((\mathbf{y} - \boldsymbol{\mu})^T \Psi^{-1} (\mathbf{y} - \boldsymbol{\mu})), \ \mathbf{y} \in \mathbb{R}^n,$$

where the function $g\colon \mathbb{R} \to [0, \infty)$ is such that $\int_0^\infty u^{n-1} g(u^2)\, du < \infty$. The function g is known as the density generator. We use the notation $\mathbf{Y} \sim \mathrm{EL}_n^g(\boldsymbol{\mu}, \Psi)$. Its characteristic function is

$$\varphi(\mathbf{t}) = \exp\left(i\mathbf{t}^T \boldsymbol{\mu}\right) h\left(\mathbf{t}^T \Psi \mathbf{t}\right),$$

for some function h, where $i = \sqrt{-1}$. Provided they exist, $\mathrm{E}(\mathbf{Y}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\mathbf{Y}) = c_g \Psi$, where $c_g = -2[dh(u)/du]_{u=0}$ is a positive constant. An elliptical distribution with $\boldsymbol{\mu} = \mathbf{0}$ and $\Psi = I_n$, the $n \times n$ identity matrix, is called a spherical density. If $\mathbf{Y} \sim \mathrm{EL}_n^g(\boldsymbol{\mu}, \Psi)$, then $\mathbf{Z} = \Psi^{-1/2}(\mathbf{Y} - \boldsymbol{\mu}) \sim \mathrm{EL}_n^g(\mathbf{0}, I_n)$.

Let $\mathbf{Y} = (\mathbf{Y}_1^T, \mathbf{Y}_2^T)^T$ represent two samples from different univariate elliptical distributions $\mathrm{EL}_1^g(\mu_i, \phi_i^2),\ i = 1, 2$. That is $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})^T$ corresponds to the i-th sample with $Y_{ij} \sim \mathrm{EL}_1^g(\mu_i, \phi_i^2),\ i = 1, 2,\ j = 1, \ldots, n_i$. Let be $\boldsymbol{\theta} = (\mu_1, \phi_1, \mu_2, \phi_2)^T$. Three different specifications for the probability density function (pdf) of $\mathbf{Y}$, that preserve these marginals, can be assumed.

Case 1. *Observations of the combined sample $\mathbf{Y}$ of size $N = n_1 + n_2$ are assumed independent and therefore the pdf of $\mathbf{Y}$ is given by*

$$p(\mathbf{y}, \boldsymbol{\theta}) = \phi_1^{-n_1} \phi_2^{-n_2} \prod_{j=1}^{n_1} g\left(\left(\frac{y_{1j} - \mu_1}{\phi_1}\right)^2\right) \prod_{j=1}^{n_2} g\left(\left(\frac{y_{2j} - \mu_2}{\phi_2}\right)^2\right). \tag{3.1}$$

Case 2. *Samples $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are independent but observations within each sample $Y_{i1}, \ldots, Y_{in_i},\ i = 1, 2.$ are pairwise uncorrelated. That is*

$\mathbf{Y}_i \sim EL^g_{n_i}(\mu_i \mathbf{1}_{n_i}, \phi_i^2 I_{n_i})$, *where* $\mathbf{1}_{n_i}$ *is the column vector of* n_i *ones,* $i = 1, 2$. *Then, the pdf of* $\mathbf{Y}$ *is given by*

$$p(\mathbf{y}, \boldsymbol{\theta}) = \phi_1^{-n_1} \phi_2^{-n_2} g\left(\sum_{j=1}^{n_1} \left(\frac{y_{1j} - \mu_1}{\phi_1}\right)^2\right) g\left(\sum_{j=1}^{n_2} \left(\frac{y_{2j} - \mu_2}{\phi_2}\right)^2\right). \quad (3.2)$$

CASE 3. *Observations of the combined sample* $\mathbf{Y}$ *are pairwise uncorrelated, but not necessarily independent. That is,* $\mathbf{Y} \sim EL^g_N(\boldsymbol{\mu}, \Psi)$, *where* $\boldsymbol{\mu} = (\mu_1 \mathbf{1}^T_{n_1}, \mu_2 \mathbf{1}^T_{n_2})^T$ *and* $\Psi = diag(\phi_1^2 \mathbf{1}^T_{n_1}, \phi_2^2 \mathbf{1}^T_{n_2})$. *Then, the pdf of* $\mathbf{Y}$ *is given by*

$$p(\mathbf{y}, \boldsymbol{\theta}) = \phi_1^{-n_1} \phi_2^{-n_2} g\left(\sum_{j=1}^{n_1} \left(\frac{y_{1j} - \mu_1}{\phi_1}\right)^2 + \sum_{j=1}^{n_2} \left(\frac{y_{2j} - \mu_2}{\phi_2}\right)^2\right). \quad (3.3)$$

For simplicity we consider the same density generator function g for both elliptical distributions in Case 1 and 2, but all results in the paper can be easily extended for two different elliptical distributions with density generators g_1 and g_2, in these cases.

In next subsections we consider the statistical manifold

$$\mathcal{M} = \{p(\mathbf{y}; \boldsymbol{\theta})\colon \boldsymbol{\theta} \in \Theta\}, \quad (3.4)$$

with $p(\mathbf{y}; \boldsymbol{\theta})$ given by (3.1), (3.2) or (3.3), under the assumptions of equal and unequal scale parameters.

Henceforth, the following notation is used

$$Z \sim \mathrm{EL}^g_1(0, 1), \quad \mathbf{Z}_i \sim \mathrm{EL}^g_{n_i}(\mathbf{0}, I_{n_i}),\ i = 1, 2 \quad \text{and} \quad \mathbf{Z} \sim \mathrm{EL}^g_N(\mathbf{0}, I_N).$$

Related to Z, $\mathbf{Z}_i,\ i = 1, 2$ and $\mathbf{Z}$ we define

$$W_0 = \frac{d \log g(Z^2)}{d(Z^2)}, \quad W_i = \frac{d \log g(\|\mathbf{Z}_i\|^2)}{d(\|\mathbf{Z}_i\|^2)},\ i = 1, 2 \quad \text{and} \quad W = \frac{d \log g(\|\mathbf{Z}\|^2)}{d(\|\mathbf{Z}\|^2)}.$$

3.1. Equal scale parameters. If we assume that the two elliptical distributions have the same scale parameter $\phi_1 = \phi_2 = \phi$, the statistical manifold (3.4) has dimension three, with $\boldsymbol{\theta} = (\mu_1, \mu_2, \phi)^T$, $\Theta = \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{>0}$ and $p(\mathbf{y}; \boldsymbol{\theta})$ given by (3.1), (3.2) or (3.3), with $\phi_1 = \phi_2 = \phi$.

Assuming that all expectations encountered exist, after some calculations, it is verified that the information matrix or metric tensor has the simple form

$$G(\boldsymbol{\theta}) = \phi^{-2}\text{diag}(A_1, A_2, B) \tag{3.5}$$

where $A_i,\ i = 1, 2$ and B are constants. Expressions for the constants are summarized in Table 1

Closed forms for constants $A_i,\ i = 1, 2$ and B can be obtained for example for the normal, Student-t, Cauchy and Pearson VII distributions. For the normal distribution, densities (3.1), (3.2) and (3.3) coincide, $A_i = n_i,\ i = 1, 2$ and $B = 2N$ in matrix (3.5). Table 2 summarizes entries of the metric tensor for the Student-t distribution with ν degrees of freedom.

We see that the metric tensor in (3.5) has the form (2.3), then, using (2.4), the geodesic distance between $\hat{\boldsymbol{\theta}} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\phi})$ and $\boldsymbol{\theta} = (\mu_1, \mu_2, \phi)$ is given by

$$d(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sqrt{B}\,\text{acosh}\left(\frac{\sum_{i=1}^{2} A_i B^{-1}(\hat{\mu}_i - \mu_i)^2 + \hat{\phi}^2 + \phi^2}{2\hat{\phi}\phi}\right). \tag{3.6}$$

3.2. Unequal Scale Parameters. If we assume unequal scale parameters, the manifold (3.4) has dimension four, with $\boldsymbol{\theta} = (\mu_1, \phi_1, \mu_2, \phi_2)^T,\ \Theta = \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R}_{>0}$ and $p(\mathbf{y}; \boldsymbol{\theta})$ given by (3.1), (3.2) or (3.3). In this case, the Fisher information matrix is given by

$$G(\boldsymbol{\theta}) = \begin{pmatrix} G_{11}(\boldsymbol{\theta}) & G_{12}(\boldsymbol{\theta}) \\ G_{12}(\boldsymbol{\theta}) & G_{22}(\boldsymbol{\theta}) \end{pmatrix}, \tag{3.7}$$

with

$$G_{ii}(\boldsymbol{\theta}) = \phi_i^{-2}\,\text{diag}(A_i, B_i),\ i = 1, 2 \quad \text{and} \quad G_{12}(\boldsymbol{\theta}) = \phi_1^{-1}\phi_2^{-1}\,\text{diag}(0, C),$$

where $A_i,\ B_i,\ i = 1, 2$ and C are constants given in Table 3.

Table 1: Entries of the information matrix (3.5)

	Case 1	Case 2	Case 3
A_i $i = 1, 2$	$4n_i\text{E}\left[W_0^2 Z^2\right]$	$4\text{E}\left[W_i^2\lVert\mathbf{Z}_i\rVert^2\right]$	$\frac{4n_i}{N}\text{E}\left[W^2\lVert\mathbf{Z}\rVert^2\right]$
B	$N\left(4\text{E}\left[W_0^2 Z^4\right] - 1\right)$	$\sum_{i=1}^{2}\left(4\text{E}\left[W_i^2\lVert\mathbf{Z}_i\rVert^4\right] - n_i^2\right)$	$4\text{E}\left[W^2\lVert\mathbf{Z}\rVert^4\right] - N^2$

Table 2: Entries of the information matrix (3.5), for Student-t distribution with ν degrees of freedom

	Case 1	Case 2	Case 3
A_i $i=1,2$	$\frac{n_i(\nu+1)}{\nu+3}$	$\frac{n_i(\nu+n_i)}{\nu+n_i+2}$	$\frac{n_i(\nu+N)}{\nu+N+2}$
B	$\frac{2N\nu}{\nu+3}$	$\sum_{i=1}^{2}\left(\frac{n_i(n_i+2)(\nu+n_i)}{\nu+n_i+2}-n_i^2\right)$	$\frac{2N\nu}{\nu+N+2}$

We see that, in general, when we have unequal scale parameters, the information matrix (3.7) for Case 3 is not diagonal and the geometry differs from that in Cases 1 and 2.

For the normal distribution, $A_i = n_i,\ B_i = 2n_i$ and $C = 0$ in matrix (3.7). Closed forms for constants $A_i,\ B_i$ and C for the Student-t distribution with ν degrees of freedom are presented in Table 4.

If $C = 0$ (Cases 1 and 2), the Riemannian space is the product of the two Riemannian spaces with metric tensors $G_{11}(\boldsymbol{\theta})$ and $G_{22}(\boldsymbol{\theta})$ and the squared geodesic distance between $\hat{\boldsymbol{\theta}} = (\hat{\mu}_1, \hat{\phi}_1, \hat{\mu}_2, \hat{\phi}_2)$ and $\boldsymbol{\theta} = (\mu_1, \phi_1, \mu_2, \phi_2)$ is the sum of the squared distances of each space, that is

$$d^2(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sum_{i=1}^{2} d_i^2((\hat{\mu}_i, \hat{\phi}_i), (\mu_i, \phi_i)), \tag{3.8}$$

where

$$d_i^2((\hat{\mu}_i, \hat{\phi}_i), (\mu_i, \phi_i)) = B_i \operatorname{acosh}^2\left(\frac{A_i B_i^{-1}(\hat{\mu}_i - \mu_i)^2 + \hat{\phi}_i^2 + \phi_i^2}{2\hat{\phi}_i \phi_i}\right),\ i = 1, 2,$$

is obtained as before, using (2.4) with the metric tensor $G_{ii}(\boldsymbol{\theta})$.

Table 3: Entries of the information matrix (3.7)

	Case 1	Case 2	Case 3
A_i $i=1,2$	$4n_i\mathrm{E}[W_0^2 Z^2]$	$4\mathrm{E}[W_i^2\lVert\mathbf{Z}_i\rVert^2]$	$\frac{4n_i}{N}\mathrm{E}[W^2\lVert\mathbf{Z}\rVert^2]$
B_i $i=1,2$	$n_i(4\mathrm{E}[W_0^2 Z^4]-1)$	$4\mathrm{E}[W_i^2\lVert\mathbf{Z}_i\rVert^4]-n_i^2$	$\frac{4n_i(n_i+2)}{N(N+2)}\mathrm{E}[W^2\lVert\mathbf{Z}\rVert^4]-n_i^2$
C	0	0	$\frac{4n_1n_2}{N(N+2)}\mathrm{E}[W^2\lVert\mathbf{Z}\rVert^4]-n_1n_2$

Table 4: Entries of the information matrix (3.7) for the Student-t distribution with ν degrees of freedom

	Case 1	Case 2	Case 3
A_i, $i=1,2$	$\dfrac{n_i(\nu+1)}{\nu+3}$	$\dfrac{n_i(\nu+n_i)}{\nu+n_i+2}$	$\dfrac{n_i(\nu+N)}{\nu+N+2}$
B_i, $i=1,2$	$\dfrac{2n_i\nu}{\nu+3}$	$\dfrac{2n_i\nu}{\nu+n_i+2}$	$\dfrac{2n_i(\nu+N-n_i)}{\nu+N+2}$
C	0	0	$-\dfrac{2n_1n_2}{\nu+N+2}$

It is noted that when we consider only uncorrelated observations (Case 3), the geometry is different. In this case the metric is not a product of metrics, neither has constant sectional curvature, then to calculate geodesic distance is somewhat troublesome. In this paper we restrict to Cases 1 and 2 for the hypothesis testing problem.

In deriving the entries of the metric tensors (3.5) and (3.7) we use the fact that if $\mathbf{Z} \sim \mathrm{EL}_n^g(\mathbf{0}, I_n)$ and $W = d\log(\|\mathbf{Z}\|^2)/d(\|\mathbf{Z}\|^2)$, then $\mathrm{E}(W\|\mathbf{Z}\|^2) = -n/2$. On the other hand, several calculations of the off-diagonal entries of the matrix involve expectations of odd functions of $\mathbf{Z}$ having a symmetric density, and then these are equal to zero.

4 Geodesic Testing Procedures for Comparing Location Parameters

We consider two samples $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of sizes n_1 and n_2, where the pdf of the combined sample $\mathbf{Y} = (\mathbf{Y}_1^T, \mathbf{Y}_2^T)^T$ can be given by (3.1), (3.2) or (3.3), with equal or unequal scale parameters. We first focus on the problem of testing $H_0\colon a\mu_1 + b\mu_2 = c$ with a, b and c fixed values in $\mathbb{R}$, against the alternative hypothesis $H_1\colon a\mu_1 + b\mu_2 \neq c$, and derive the geodesic test statistic. As particular cases we obtain the statistics for the difference and ratio of locations. In this way, we propose a new asymptotic testing procedure, with χ_1^2 distribution, which is parametrization-invariant, for the well-known Behrens-Fisher and Fieller-Creasy problems in elliptical populations. Also, we extend some results to the problem of testing equality of several location parameters.

4.1. The Two Sample Problem with Equal Scale Parameters. Consider that the pdf of the combined sample $\mathbf{Y} = (\mathbf{Y}_1^T, \mathbf{Y}_2^T)^T$ is given by

(3.1), (3.2) or (3.3), with $\phi_1 = \phi_2 = \phi$. The null hypothesis defines a two-dimensional submanifold of the three-dimensional manifold $\mathcal{M}$,

$$\mathcal{N} = \{p(\mathbf{y};\boldsymbol{\theta})\colon \boldsymbol{\theta} \in \Theta \text{ and } a\mu_1 + b\mu_2 = c\},$$

where $\boldsymbol{\theta} = (\mu_1, \mu_2, \phi)^T$ and $\Theta = \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{>0}$.

The distance from the unrestricted maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ to the submanifold $\mathcal{N}$ is obtained by minimizing the expression (3.6) for $\boldsymbol{\theta} = \boldsymbol{\theta}^0 = (\mu_1^0, \mu_2^0, \phi^0)^T \in \mathcal{N}$. These minimizing values are

$$\mu_1^0 = \frac{A_1 b^2 \hat{\mu}_1 - A_2 ab\hat{\mu}_2 + A_2 ac}{A_1 b^2 + A_2 a^2}, \quad \mu_2^0 = \frac{A_2 a^2 \hat{\mu}_2 - A_1 ab\hat{\mu}_1 + A_1 bc}{A_1 b^2 + A_2 a^2}$$

and

$$\phi^0 = \sqrt{\hat{\phi}^2 + \frac{A_1 A_2}{(A_1 b^2 + A_2 a^2)B}(a\hat{\mu}_1 + b\hat{\mu}_2 - c)^2}.$$

Replacing this value of $\boldsymbol{\theta}^0$ in (3.6) we get the desired distance $\rho^2(\hat{\boldsymbol{\theta}}, \mathcal{N})$, and the statistic of the geodesic test is given by

$$T = B\,\mathrm{acosh}^2(\sqrt{B^{-1}W + 1}),$$

where

$$W = \frac{(a\hat{\mu}_1 + b\hat{\mu}_2 - c)^2}{\hat{\phi}^2\left(\frac{a^2}{A_1} + \frac{b^2}{A_2}\right)}$$

is the Wald statistic.

Then, the geodesic test is a strictly monotone increasing function of the Wald statistic. By the other hand, through a second-order Taylor expansion, we have that $T = W + o_p(1)$.

Setting $a = 1,\ b = -1$ and $c = 0$, we obtain the geodesic test statistic for testing the equality of location parameters, $H_0\colon \mu_1 = \mu_2$.

Particularly, in the normal case, the geodesic test statistic, is given by

$$T = 2N\,\mathrm{acosh}^2\left(\sqrt{\frac{W}{2N} + 1}\right),$$

where W is the Wald statistic, which differs only slightly from t^2, being t the Student-t statistic, with $n_1 + n_2 - 2$ degrees of freedom, usually used for this problem. Then, the geodesic test is equivalent to the one usually used for the problem. Choosing $a = 1,\ b = -\kappa$ and $c = 0$, we obtain the geodesic test statistic for the Fieller-Creasy problem about testing the ratio of locations, $H_0\colon \mu_1/\mu_2 = \kappa$. Again, the geodesic test statistic is asymptotically equivalent to the Student-t statistic usually used in the normal case.

4.2. The Two Sample Problem with Unequal Scale Parameters. Here we shall consider the problem of testing a linear hypothesis of location parameters when the scale parameters are not equal. We derive a procedure for testing H_0 in Cases 1 and 2. In particular, let $\mathbf{Y}_1$ and $\mathbf{Y}_2$ be two samples of sizes n_1 and n_2, respectively, whose joint pdf is given by (3.1) or (3.2). That is,

$$Y_{ij} \sim \mathrm{EL}_1^g(\mu_i, \phi_i), \quad i = 1, 2; \quad j = 1, 2, \ldots, n_i.$$

The null hypothesis defines a three-dimensional submanifold of the four-dimensional manifold $\mathcal{M}$,

$$\mathcal{N} = \{p(\mathbf{y}; \boldsymbol{\theta}) \colon \boldsymbol{\theta} \in \Theta \text{ and } a\mu_1 + b\mu_2 = c\},$$

where $\boldsymbol{\theta} = (\mu_1, \phi_1, \mu_2, \phi_2)^T$ and $\Theta = \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R}_{>0}$.

The distance from the unrestricted maximum likelihood estimator $\hat{\boldsymbol{\theta}} = (\hat{\mu_1}, \hat{\phi_1}, \hat{\mu_2}, \hat{\phi_2})^T$ to the submanifold $\mathcal{N}$ is obtained by minimizing expression (3.8) for $\boldsymbol{\theta} = \boldsymbol{\theta}^0 = (\mu_1^0, \phi_1^0, \mu_2^0, \phi_2^0)^T \in \mathcal{N}$. Since $\boldsymbol{\theta}^0 \in \mathcal{N}$, we have that

$$\mu_1^0 = \frac{-b\mu_2^0 + c}{a}, \quad \phi_1^0 = \sqrt{\frac{A_1}{B_1 a^2}\left(a\hat{\mu}_1 + b\mu_2^0 - c\right)^2 + \hat{\phi_1}^2} \text{ and } \phi_2^0 = \sqrt{\frac{A_2}{B_2}\left(\mu_2^0 - \hat{\mu}_2\right)^2 + \hat{\phi_2}^2}.$$

Substituting these values in (3.8) the geodesic statistic test is given by

$$T = \min_{\tilde{\mu}} d^2(\tilde{\mu}),$$

where

$$\begin{aligned} d^2(\tilde{\mu}) &= B_1 \operatorname{acosh}^2\left[\sqrt{\frac{A_1}{a^2 B_1}\left(\frac{a\hat{\mu}_1 + b\tilde{\mu} - c}{\hat{\phi}_1}\right)^2 + 1}\right] \\ &\quad + B_2 \operatorname{acosh}^2\left[\sqrt{\frac{A_2}{B_2}\left(\frac{\hat{\mu}_2 - \tilde{\mu}}{\hat{\phi}_2}\right)^2 + 1}\right]. \end{aligned}$$

A closed form for the statistic T is not obtained in this case.

Particularly, new asymptotic and parametrization-invariant test statistics are derived for the Behrens-Fisher problem $(a = 1,\ b = -1,\ c = 0)$ and for the Fieller-Creasy problem $(a = 1,\ b = -\kappa,\ c = 0)$, with unequal scale parameters.

4.3. Testing Equality of Several Location Parameters. While we focus our attention on the two-sample problem, the above results can actually be extended to derive a geodesic statistic for testing equality of several location parameters. Consider that $\mathbf{Y} = (\mathbf{Y_1}, \dots, \mathbf{Y}_k)^T$ represent k samples under consideration, $k \geq 2$, where $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^T$, with $Y_{ij} \sim \mathrm{EL}_1^g(\mu_i, \phi_i^2)$, $i = 1, \dots k$, $j = 1 \dots, n_i$. Analogous to Cases 1, 2 and 3, we can consider that observations of the combined sample are independent, the samples are independent but with uncorrelated observations within each sample or the observations are pairwise uncorrelated but not necessarily independent.

It is desired to test the hypothesis

$$H_0\colon \mu_1 = \mu_2 = \cdots = \mu_k.$$

If we assume that the k populations have the same scale parameter ϕ, the manifold $\mathcal{M}$ has dimension $k{+}1$, with coordinates $\boldsymbol{\theta} = (\mu_1, \mu_2, \dots, \mu_k, \phi)^T$, $\boldsymbol{\theta} \in \Theta = \mathbb{R}^k \times \mathbb{R}_{>0}$ and metric tensor

$$G(\boldsymbol{\theta}) = \phi^{-2}\mathrm{diag}(A_1, A_2, \dots, A_k, B),$$

where the constants A_i, $i = 1, \dots, k$ and B are analogous to those in Table 1, with i varying from 1 to k and $N = \sum_{i=1}^k n_i$. The submanifold $\mathcal{N}$ associated to H_0 has dimension 2.

Then, analogous to Section 4.1, it can be shown that the geodesic test statistic is given by

$$T = B\,\mathrm{acosh}^2\left[\sqrt{\frac{\sum_{i=1}^k A_i B^{-1}(\hat{\mu}_i - \mu^0)^2 + \hat{\phi}^2 + {\phi^0}^2}{2\hat{\phi}\phi^0}}\right],$$

where

$$\mu^0 = \frac{\sum_{i=1}^k \hat{\mu}_i}{\sum_{i=1}^k A_i} \quad \text{and} \quad \phi^0 = \sqrt{\hat{\phi}^2 + \frac{\sum_{i=1}^k (\hat{\mu}_i - \mu^0)^2}{B}},$$

with $\hat{\mu}_i$, $i = 1, \dots k$ and $\hat{\phi}$ the maximum likelihood estimators of μ_i and ϕ, respectively. After some algebra, we have again that

$$T = B\ \mathrm{acosh}^2(\sqrt{B^{-1}W + 1}),$$

where

$$W = \frac{1}{\sum_{i=1}^k A_i} \sum_{i<j} A_i A_j \left(\frac{\hat{\mu}_i - \hat{\mu}_j}{\hat{\phi}}\right)^2$$

is the Wald statistic and then $T = W + o_p(1)$. In the normal case, W is equivalent, when $N \to \infty$, to the usual F-ratio test with a Fisher distribution with $k-1$ and $N-k$ degrees of freedom. So, we have that the geodesic test is equivalent to the one usually used.

On the other hand, if we assume unequal scale parameters, the manifold $\mathcal{M}$ has dimension $2k$, with coordinates $\boldsymbol{\theta} = (\mu_1, \phi_1, \ldots, \mu_k, \phi_k)^T$, $\boldsymbol{\theta} \in \Theta = (\mathbb{R} \times \mathbb{R}_{>0})^k$ and metric tensor for Cases 1 and 2 given by

$$G(\boldsymbol{\theta}) = \mathrm{diag}(G_{11}(\boldsymbol{\theta}), G_{22}(\boldsymbol{\theta}), \ldots, G_{kk}(\boldsymbol{\theta})),$$

with $G_{ii}(\boldsymbol{\theta}) = \phi_i^{-2}\mathrm{diag}(A_i, B_i)$, where the constants A_i and B_i, $i = 1, \ldots, k$ are analogous to those in Table 3, with i varying from 1 to k and $N = \sum_{i=1}^{k} n_i$.

In this case the submanifold $\mathcal{N}$ associated to H_0 has dimension $k+1$ and the geodesic test statistic is given by

$$T = \min_{\tilde{\mu}} d^2(\tilde{\mu}),$$

where

$$d^2(\tilde{\mu}) = \sum_{i=1}^{k} B_i \,\mathrm{acosh}^2 \sqrt{\frac{A_i}{B_i}\left(\frac{\tilde{\mu} - \hat{\mu}_i}{\hat{\phi}_i}\right)^2 + 1},$$

with $\hat{\mu}_i$ and $\hat{\phi}_i$, the maximum likelihood estimators of μ_i and ϕ_i, $i = 1, \ldots, k$.

Under H_0, T is asymptotically distributed as a chi-squared with $k-1$ degrees of freedom.

5 Monte Carlo Results

As an illustration, this section presents Monte Carlo simulation results to analyse the behaviour of the geodesic testing procedure for equality of location parameters, to see how the sizes and powers of the test may depend on sample sizes, on alternatives from the null and on the ratios of scale parameters. Comparisons with the Wald test are included.

We consider the special case of the Student-t distribution in Case 1. In this situation we have closed forms for constants in the test statistics and it leads itself to EM algorithms for maximum likelihood estimation. Following Lange, Little and Taylor (1989) we obtain the equations that define the M-step of the algorithm as the corresponding to the maximum likelihood in the normal case, but weighted by the unknown "missing" values, which are estimated at the E-step of the algorithm.

All simulations were performed using the statistical package R (R Development Core Team, 2013); they are all based on 10 000 independently generated data sets. Two samples of Student-t distributions with $\nu = 5$ degrees of freedom, with parameters (μ_1, ϕ_1^2) and (μ_2, ϕ_2^2), are generated under specifications in Case 1. The sample size pairs considered are $(n_1, n_2) = (4, 8)$, $(5, 15)$, $(10, 10)$, $(15, 45)$, $(30, 30)$ and $(25, 75)$. For the Behrens-Fisher problem we consider $\rho^2 = \phi_1^2/\phi_2^2$ and the true values of the parameters are fixed in $\mu_1 = \mu_2 = 0$, $\phi_2^2 = 1$. When $\rho^2 = 1$, the two scale parameters are equal, when ρ^2 is farther apart from one, the two scale parameters are further different from each other where the Behrens-Fisher problem becomes severe. For the equal scale parameter case the true values of the parameters are fixed in $\mu_1 = \mu_2 = 0$, $\phi_1^2 = \phi_2^2 = \phi^2 = 1$. In conducting power comparisons, the noncentrality parameters considered were

$$\tau = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\phi_1^2}{n_1} + \frac{\phi_2^2}{n_2}}} \quad \text{and} \quad \tau = \frac{\mu_1 - \mu_2}{\phi\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

for the Behrens-Fisher and equal scale parameter cases, respectively. The null hypothesis is true when $\tau = 0$. Alternative hypothesis with $\tau = 1, 2, 3, 4, 5$ were considered. The nominal level was fixed in $\alpha = 0.05$. An adjustment was performed using the empirical distribution of the statistics generated from the simulations such that the empirical sizes is corrected to be 0.05. We include performance of the Wald test, which is given by

$$W = \frac{(\hat{\mu}_1 - \hat{\mu}_2)^2}{\hat{\phi}^2\left(\frac{1}{A_1} + \frac{1}{A_2}\right)} \quad \text{and} \quad W = \frac{(\hat{\mu}_1 - \hat{\mu}_2)^2}{\frac{\phi_1^2}{A_1} + \frac{\phi_2^2}{A_2}},$$

for the cases of equal and unequal location parameters, respectively.

The empirical results provide the evidence that the geodesic test statistic has good sampling properties in terms of level and power. Tables 5 and 6 give the empirical levels of the geodesic test and of those corresponding from the Wald test, for equal and unequal scale parameters, respectively. As the sample sizes increase, the empirical levels are close to the nominal level. When we have unequal scale parameters, the effect of the ratio of location parameters becomes less important with large sample sizes. The empirical levels of the geodesic test are closer to the nominal level than those of the Wald test. The difference between the empirical levels of both tests is most significant for small sample sizes and when the ratio of location parameters increases.

Figures 1 and 2 plot the relative quantile discrepancies versus the corresponding asymptotic quantiles of the null distribution of geodesic and Wald

Table 5: Empirical levels for Geodesic (T) and Wald (W) tests of equality of location parameters with equal scale parameters, for a Student-t distribution with 5 degrees of freedom. Nominal level: α

α		(n_1, n_2)					
		(4,8)	(5,15)	(10,10)	(15,45)	(30,30)	(25,75)
0.1	T	0.1515	0.1364	0.1320	0.1134	0.1089	0.1065
	W	0.1633	0.1420	0.1392	0.1158	0.1107	0.1079
0.05	T	0.0918	0.0779	0.0758	0.0578	0.0572	0.0551
	W	0.1042	0.0852	0.0815	0.0609	0.0592	0.0560
0.01	T	0.0317	0.0209	0.0226	0.0116	0.0127	0.0112
	W	0.0425	0.0271	0.0284	0.0129	0.0134	0.0117

Table 6: Empirical levels for Geodesic (T) and Wald (W) tests of equality of location parameters with unequal scale parameters, for a Student-t distribution with 5 degrees of freedom. Nominal level: α.

α	ρ^2		(n_1, n_2)					
			(4,8)	(5,15)	(10,10)	(15,45)	(30,30)	(25,75)
0.1	0.1	T	0.1346	0.1258	0.1342	0.1127	0.1082	0.1041
		W	0.1465	0.1330	0.1440	0.1148	0.1121	0.1062
	1	T	0.1484	0.1455	0.1217	0.1162	0.1064	0.1095
		W	0.1653	0.1600	0.1298	0.1203	0.1079	0.1127
	10	T	0.1830	0.1703	0.1285	0.1202	0.1083	0.1143
		W	0.2109	0.1952	0.1389	0.1278	0.1119	0.1189
0.05	0.1	T	0.0766	0.0664	0.0750	0.0565	0.0571	0.0522
		W	0.0899	0.0750	0.0871	0.0587	0.0602	0.0536
	1	T	0.0888	0.0832	0.0680	0.0625	0.0545	0.0565
		W	0.1059	0.0981	0.0757	0.0678	0.0565	0.0600
	10	T	0.1167	0.1033	0.0700	0.0651	0.0567	0.0592
		W	0.1494	0.1296	0.0827	0.0739	0.0600	0.0638
0.01	0.1	T	0.0257	0.0158	0.0240	0.0131	0.0137	0.0104
		W	0.0349	0.0214	0.0310	0.0144	0.0171	0.0112
	1	T	0.0279	0.0250	0.0190	0.0144	0.0114	0.0129
		W	0.0423	0.0367	0.0252	0.0171	0.0127	0.0153
	10	T	0.0472	0.0378	0.0217	0.0166	0.0128	0.0149
		W	0.0756	0.0572	0.0284	0.0222	0.0151	0.0182

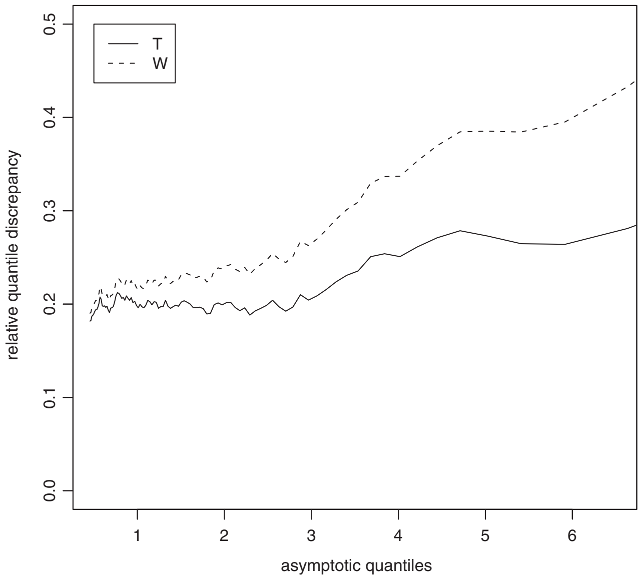

Figure 1: Relative quantile discrepancies plot of Geodesic (T) and Wald (W) statistics, for testing equality of location parameters with equal scales parameters, for a Student-t with 5 degrees of freedom. $n_1 = n_2 = 10$ and $\alpha = 0.05$

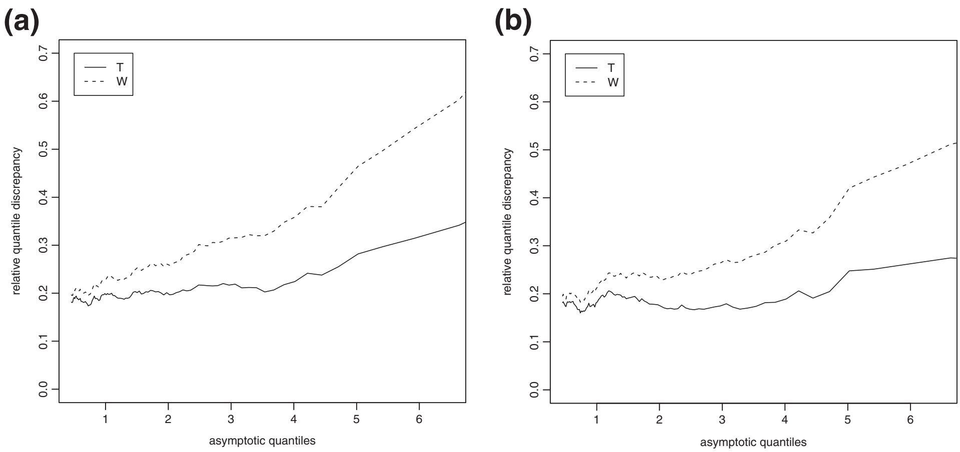

Figure 2: Relative quantile discrepancies plot of Geodesic (T) and Wald (W) statistics, for testing equality of location parameters with unequal scales parameters, for a Student-t with 5 degrees of freedom. $n_1 = n_2 = 10$, $\alpha = 0.05$, (a) $\rho^2 = 0.1$, (b) $\rho^2 = 10$

Table 7: Empirical powers for Geodesic test of equality of location parameters with equal scale parameters, for a Student-t distribution with 5 degrees of freedom. $\alpha = 0.05$

	(n_1, n_2)					
τ	(4,8)	(5,15)	(10,10)	(15,45)	(30,30)	(25,75)
1	0.1241	0.1286	0.1285	0.1344	0.1390	0.1405
2	0.3511	0.3715	0.2618	0.4076	0.4020	0.4047
3	0.6410	0.6745	0.6669	0.7261	0.7206	0.7312
4	0.8484	0.8780	0.8796	0.9188	0.9212	0.9278
5	0.9504	0.9694	0.9671	0.9864	0.9868	0.9898

test statistics. Relative quantile discrepancy is defined as the difference between quantiles estimated by simulation and asymptotic quantiles divided by the latter. The closer to zero this discrepancy, the better the approximation of the exact null distribution of the test statistic by the limiting chi-squared distribution. We see that the relative discrepancy curve based on geodesic statistic is closer to the limiting χ^2_1 distribution than those of the Wald statistic.

Table 8: Empirical powers for Geodesic test of equality of location parameters with unequal scale parameters, for a Student-t distribution with 5 degrees of freedom. $\alpha = 0.05$

		(n_1, n_2)					
τ	ρ^2	(4,8)	(5,15)	(10,10)	(15,45)	(30,30)	(25,75)
1	0.1	0.1268	0.1333	0.1326	0.1357	0.1365	0.1416
	1	0.1270	0.1337	0.1320	0.1343	0.1411	0.1388
	10	0.1203	0.1288	0.1380	0.1341	0.1330	0.1398
2	0.1	0.3437	0.3723	0.3603	0.4005	0.3977	0.4093
	1	0.3505	0.3631	0.3702	0.3896	0.4062	0.3986
	10	0.2994	0.3346	0.3754	0.3820	0.3893	0.3899
3	0.1	0.6391	0.6747	0.6502	0.7204	0.7110	0.7341
	1	0.6196	0.6332	0.6746	0.7043	0.7227	0.7184
	10	0.5207	0.5773	0.6661	0.6903	0.7082	0.7067
4	0.1	0.8441	0.8813	0.8593	0.9155	0.9123	0.9287
	1	0.8183	0.8355	0.8819	0.8977	0.9230	0.9170
	10	0.7112	0.7708	0.8683	0.8877	0.9155	0.9118
5	0.1	0.9463	0.9703	0.9530	0.9831	0.9834	0.9895
	1	0.9262	0.9346	0.9677	0.9764	0.9871	0.9860
	10	0.8424	0.8947	0.9614	0.9701	0.9832	0.9825

Tables 7 and 8 provide power calculations of the geodesic test. In terms of powers, geodesic and Wald tests perform similarly, with sometimes one, and sometimes the other, being slightly superior, so we do not include powers of Wald. The power of the geodesic test increases as τ and sample sizes increase and decreases as the ratio of location parameters increases. These results are typical and comparable with that of asymptotically optimal tests provided by Wald, likelihood ratio and score statistics for normal populations, reported, for example by Best and Rayner (1987).

Simulation experiments for Cases 2 and 3, when scale parameters are equal and for Case 2, when scale parameters differ, were also computed. The analysis showed that the geodesic statistic perform similarly to that of Case 1. Also, other null hypothesis were considered with a similar performance.

6 Conclusions

In this paper we have studied the geometry of the differentiable manifold associated with two samples of symmetric distributions in the real line equipped with the Fisher information matrix as Riemannian metric. We have obtained general expressions for the entries of the information matrices and geodesic distances in manifolds defined under usual statistical assumptions, including equal or unequal scale parameters and independent or pairwise uncorrelated observations. For calculating geodesic distances simplifications were possible by reducing metrics to one of a hyperbolic geometry, except for the case of unequal scale parameters and uncorrelated observations of the combined sample. Calculating geodesic distances in this case merits further investigation.

Based on geodesic distances we propose an asymptotic and parametric-invariant testing procedure for a general linear hypothesis about location parameters from any elliptical distribution. As particular cases, the statistics for the difference and ratio of locations can be obtained. In this way, new asymptotic procedures for the well know Behrens-Fisher and Fieller-Creasy problems are proposed. Geometry can be easily extended for several elliptical populations. In particular, an extension for testing equality of several location parameters is derived, generalizing the two sample problem. When the scale parameters are equal the geodesic statistic is a strictly monotone increasing function of the Wald statistic. From a second-order Taylor expansion of the geodesic statistic we see that the Wald statistic emerges as the leading term.

Empirical results based on a small simulation study for testing equality of location parameters in Student-t distributions, provide evidence that the geodesic test statistic has good sampling properties, outperforming the Wald test in terms of level and with a similar performance in terms of power. Also, the null distribution of the geodesic statistic is closer to the limiting chi-squared distribution, than those of the Wald statistic.

The geometry derived in this paper can also be used to calculate, for example, geodesic statistics for testing equality of scale parameters and simultaneous equality of location and scale parameters.

Acknowledgements. The authors are very grateful to a referee for many helpful comments and suggestions.

References

AMARI, S.-I. (1982). Geometrical theory of asymptotic ancillarity and conditional inference. *Biometrika*, **69**, 1–17.

AMARI, S.-I (1985). *Differential-geometrical methods in statistics.*, Lecture Notes in Statistics 28. Springer, Berlin.

ATKINSON, C. and MITCHELL, A.F.S. (1981). Rao's distance measure. *Sankhya*, **43 Ser A**, 345–365.

BARNDORFF-NIELSEN, O.E., COX, D.R. and REID, N. (1986). The role of differential geometry in statistical theory. *Int. Stat. Rev.*, **54**, 83–96.

BERKANE, M., ODEN, K. and BENTLER, P.M. (1997). Geodesic estimation in elliptical distributions. *J. Multivariate Anal.*, **63**, 35–46.

BEST, D.J. and RAYNER, J.C.W. (1987). Welch's approximate solution for the Behrens-Fisher problem. *Technometrics*, **29**, 205–210.

BHATTACHARYYA, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, **35**, 99–109.

BURBEA, J. (1986). Informative geometry of probability spaces. *Expo. Math.*, **4**, 347–378.

BURBEA, J. and RAO, C.R. (1982). Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. *J. Multivariate Anal.*, **12**, 575–596.

BURBEA, J. and RAO, C.R. (1984). Differential metrics in probability spaces. *Probab. Math. Stat.*, **3**, 241–258.

CAMBANIS, S., HUANG, S. and SIMONS, G. (1981). On the theory of elliptically contoured distributions. *J. Multivariate Anal.*, **11**, 368–385.

CHOI, H. and KIEFER, N. (2011). Geometry of the log-likelihood ratio statistic in misspecified models. *J. Statist. Plann. Inference*, **141**, 2091–2099.

CRITCHLEY, F., MARRIOTT, P. and SALMON, M. (1996). On the differential geometry of the Wald test with nonlinear restrictions. *Econometrica*, **64**, 1213–1222.

CUBEDO, M. and OLLER, J.M. (2002). Hypothesis testing: a model selection approach. *J. Statist. Plann. Inference*, **108**, 3–21.

DAVIDSON, R. (1992). *The Geometry of the Wald test*. Queen's University, Mimeo.

DO CARMO, M.P. (1992). *Riemannian geometry*. Springer, Berlin.

EFRON, B. (1975). Defining the curvature of a statistical problem (with applications to second-order efficiency) (with discussion). *Ann. Statist.*, **3**, 1189–1242.

FANG, K.T., KOTZ, S. and NG, K.W. (1990). *Symmetric multivariate and related distributions*. Chapman and Hall, London-New York.

GREGORY, A. and VEALL, M.R. (1985). Formulating Wald tests of nonlinear restrictions. *Econometrica*, **50**, 549–563.

HICKS, N.J. (1965). *Notes on differential geometry*. Van Nostran, Princeton.

JENSEN, U. (1995). A Review of the Derivation and Calculation of Rao Distances with an Application to Portfolio Theory. In *Advances in Econometrics and Quantitative Economics: Essays in Honor of C.R. Rao*, (P. Maddala, G.S. Phillips and T. Srinivasan, eds.). Blackwell, Cambridge, pp. 413–462.

KASS, R.E. (1989). The geometry of asymptotic inference (with discussion). *Stat. Sci.*, **4**, 188–234.

KASS, R.E. and VOS, P.W. (1997). *Geometrical foundations of asymptotic inference*. Wiley, New York.

KELKER, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya*, **32**, 419–430.

LANGE, K.L., LITTLE, J.A. and TAYLOR, M.G. (1989). Robust statistical modelling using the t-distribution. *J. Amer. Statist. Assoc.*, **84**, 881–896.

LARSEN, P.V. and JUPP, P.E. (2003). Parametrization-invariant Wald tests. *Bernoulli*, **9**, 167–182.

LE CAM, L. (1990). On the standard asymptotic confidence ellipsoids of Wald. *ISI Rev.*, **58**, 129–152.

MICCHELLI, C.A. and NOAKES, L. (2005). Rao distances. *J. Multivariate Anal.*, **92**, 97–115.

MITCHELL, A.F.S. (1988). Statistical manifolds of univariate elliptic distributions. *Int. Stat. Rev.*, **56**, 1–16.

MITCHELL, A.F.S. (1989). The information matrix, skewness tensor and α-connections for the general multivariate elliptical distribution. *Ann. Inst. Statist. Math.*, **41**, 289–304.

MURRAY, M.K. and RICE, J.W. (1993). *Differential geometry and statistics*. Chapman and Hall, London.

OLLER, J.M. and CORCUERA, J.M. (1995). Intrinsic analysis of statistical estimation. *Ann. Statist.*, **5**, 1562–1581.

PATRANGENARU, V. (1994). Locally homogeneous riemannian manifolds and Cartan triples. *Geom. Dedicata*, **50**, 143–164.

R DEVELOPMENT CORE TEAM (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISSN 3-900051-07-0, URL http://www.R-project.org.

RAO, C.R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, **37**, 81–91.

RAO, C.R. (1987). Differential Metrics in Probability Spaces. In *Differential Geometry in Statistical Inference*. IMS, Hayward, pp. 217–237.

RATCLIFFE, J.G. (1994). *Foundations of hyperbolic manifolds*. Springer, Berlin.

SKOVGAARD, L.T. (1984). A Riemannian geometry of the multivariate normal model. *Scand. J. Stat.*, **11**, 211–223.

SPIVAK, M. (1979). *A comprehensive introduction to differential geometry*. Publish or Perish, Inc., Berkeley.

VILLARROYA, A. and OLLER, J.M. (1993). Statistical test for the inverse gaussian distribution based on Rao distance. *Sankhya A*, **55**, 80–103.

Patricia Giménez
Jorge N. López
Lucas Guarracino
Departamento de Matemática, FCEYN
Universidad Nacional de Mar del
Plata, Funes 3350, CP 7600,
Mar del Plata, Argentina
E-mail: pcgimene@mdp.edu.ar

Paper received: 3 February 2014.