



Published in final edited form as:

Aust N Z J Stat. 2017 March 1; 59(1): 119–135. doi:10.1111/anzs.12183.

An Improved Test of Equality of Mean Directions for the Langevin-von Mises-Fisher Distribution

Pavlina Rumcheva¹ and Brett Presnell²

U.S. Centers for Disease Control and Prevention and University of Florida

Abstract

A multi-sample test for equality of mean directions is developed for populations having Langevin-von Mises-Fisher distributions with a common unknown concentration. The proposed test statistic is a monotone transformation of the likelihood ratio. The high-concentration asymptotic null distribution of the test statistic is derived. In contrast to previously suggested high-concentration tests, the high-concentration asymptotic approximation to the null distribution of the proposed test statistic is also valid for large sample sizes with any fixed nonzero concentration parameter. Simulations of size and power show that the proposed test outperforms competing tests. An example with three-dimensional data from an anthropological study illustrates the practical application of the testing procedure.

Keywords

directional data; likelihood ratio; mean resultant length; high-concentration asymptotic distribution

1. Introduction

Directional data are observed in many scientific fields and are especially common in the earth sciences. Directions are measured in various dimensions and can be represented as unit vectors or combinations of angles. The analysis of such data requires statistical methods that properly account for the structure of the sample space. In this article we represent directions as unit vectors, and thus the sample space will be the unit circle in two dimensions and the unit sphere in three dimensions. In general, the sample space of a d -dimensional direction is the unit hypersphere in \mathbb{R}^d denoted by $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = 1\}$, where $\|\cdot\|$ is the usual Euclidean norm for vectors. We derive all results in this paper for the general d -dimensional case.

Let \mathbf{U} be a random direction in \mathbb{R}^d , that is, $\mathbf{U} \in \mathbb{S}^{d-1}$. Two important characteristics of the distribution of \mathbf{U} are its mean direction $\boldsymbol{\mu} = \mathbf{E}(\mathbf{U})/\|\mathbf{E}(\mathbf{U})\|$ and its mean resultant length $\rho = \|\mathbf{E}(\mathbf{U})\|$, which measure the location and concentration of the distribution, respectively (here $\mathbf{E}(\mathbf{U})$ represents the usual componentwise expectation for vectors). The mean resultant

Correspondence to: Pavlina Rumcheva.

¹National Center for Health Statistics, U.S. Centers for Disease Control and Prevention, 3311 Toledo Road, Hyattsville, MD 20782, USA prumcheva@cdc.gov

²Department of Statistics, University of Florida, 102 Griffin-Floyd Hall, Gainesville, FL 32611, USA

length takes values in the interval $[0, 1]$, with larger values of ρ indicating higher concentration: if $\rho = 1$, then \mathbf{U} is a constant; and if $\rho = 0$, then the mean direction is not defined, as when \mathbf{U} is uniformly distributed on the sphere.

Models for directional data often employ the Langevin-von Mises-Fisher distribution, $L_d(\boldsymbol{\mu}, \kappa)$, whose density with respect to the uniform distribution on the unit hypersphere is

$$f(\mathbf{u}; \boldsymbol{\mu}, \kappa) = \frac{(\kappa/2)^{d/2-1}}{\Gamma(d/2)I_{d/2-1}(\kappa)} \exp(\kappa \boldsymbol{\mu}^\top \mathbf{u}), \quad \mathbf{u} \in \mathbb{S}^{d-1}, \quad (1)$$

where $\boldsymbol{\mu} \in \mathbb{S}^{d-1}$, $\kappa > 0$, Γ denotes the gamma function, and I_ν is the modified Bessel function of the first kind and order ν (see Abramowitz & Stegun 1970, page 374). This distribution is unimodal and rotationally symmetric about its mean direction $\boldsymbol{\mu}$, with mean resultant length (see Watson 1983, page 201)

$$\rho = A_d(\kappa) = \frac{I_{d/2}(\kappa)}{I_{d/2-1}(\kappa)}. \quad (2)$$

The mean resultant length ρ is a strictly increasing function of the concentration parameter κ , with $\rho \downarrow 0$ as $\kappa \downarrow 0$ and $\rho \uparrow 1$ as $\kappa \uparrow \infty$. Thus, larger values of κ correspond to more concentrated distributions, with $\kappa = 0$ corresponding to the uniform distribution on the sphere and $\kappa = \infty$ corresponding to the point mass at $\boldsymbol{\mu}$.

Given several independent random samples of directions of equal concentration, one often wishes to test for equality of the corresponding population mean directions, the directional analogue of the classical one-way analysis of variance problem. Let $\mathbf{U}_{i1}, \dots, \mathbf{U}_{in_i}$, $i = 1, \dots, k$, be k independent random samples of sizes n_i from $L_d(\boldsymbol{\mu}_i, \kappa)$, $i = 1, \dots, k$, where $\boldsymbol{\mu}_i$ is the mean direction of the i th sample and κ is the common unknown concentration. We wish to test the null hypothesis that the k mean directions are equal against the alternative that at least two mean directions differ; that is, we wish to test

$$H_0: \boldsymbol{\mu}_1 = \dots = \boldsymbol{\mu}_k \quad \text{versus} \quad H_a: \boldsymbol{\mu}_i \neq \boldsymbol{\mu}_j \text{ for at least one pair } (i, j). \quad (3)$$

Before discussing how this can be accomplished with existing tests, we introduce some additional definitions and notation. For the i th sample, the *resultant vector* is the sum $\sum_{j=1}^{n_i} \mathbf{U}_{ij}$ and the *sample mean vector* is the average $\bar{\mathbf{U}}_i = n_i^{-1} \sum_{j=1}^{n_i} \mathbf{U}_{ij}$. The *resultant length* is denoted by $R_i = \|\sum_{j=1}^{n_i} \mathbf{U}_{ij}\|$ and the *sample mean resultant length* by $\bar{R}_i = n_i^{-1} R_i$. The *sample mean direction* is then defined as the unit vector $\bar{R}_i^{-1} \bar{\mathbf{U}}_i$. For the combined sample

with a total sample size $n = n_1 + \dots + n_k$, we will denote the sample mean vector, resultant length, and sample mean resultant length by \bar{X} , R , and \bar{R} , respectively.

Watson & Williams (1956) proposed a high-concentration F-test of (3), applicable to any number of samples in any dimension, generalizing a test previously introduced by Watson (1956) for the spherical (3-dimensional) case. The Watson & Williams test is based on a decomposition of the total variation represented by $2\kappa(n - R)$ into within-sample and between-sample components, as expressed by the formula

$$2\kappa(n - R) = 2\kappa\left(n - \sum_{i=1}^k R_i\right) + 2\kappa\left(\sum_{i=1}^k R_i - R\right).$$

Following Watson (1956), Watson & Williams noted that, under H_0 ,

$$2\kappa\left(\frac{n - \sum_{i=1}^k R_i}{\sum_{i=1}^k R_i - R}\right) \xrightarrow{d} \left(\frac{\chi_{(n-k)(d-1)}^2}{\chi_{(k-1)(d-1)}^2}\right) \text{ as } \kappa \rightarrow \infty, \quad (4)$$

where $\chi_{(n-k)(d-1)}^2$ and $\chi_{(k-1)(d-1)}^2$ are independent chi-squared distributed random variables with $(n - k)(d - 1)$ and $(k - 1)(d - 1)$ degrees of freedom, respectively, and \xrightarrow{d} denotes convergence in distribution. The test statistic suggested by Watson & Williams,

$$W = \frac{(n - k)\left(\sum_{i=1}^k R_i - R\right)}{(k - 1)\left(n - \sum_{i=1}^k R_i\right)}, \quad (5)$$

is the ratio of the between-sample and within-sample variabilities, divided by the appropriate degrees of freedom. The null hypothesis of equal mean directions is rejected for large values of W , whose high-concentration asymptotic distribution under the null hypothesis is F with $(k - 1)(d - 1)$ numerator and $(n - k)(d - 1)$ denominator degrees of freedom.

In the circular (2-dimensional) case, Stephens (1972) incorporated a multiplicative factor into the Watson & Williams statistic (5) in order to improve the F-approximation. Mardia & Jupp (2000, page 191) suggested a similar correction for the spherical (3-dimensional) case, and Stephens (1992) provided further details for the general d -dimensional case.

Watson & Williams (1956) also developed exact tests of (3) based on the conditional distribution of $R_1 + \dots + R_k$ given R , which does not depend on the common unknown κ . Intuitively, if the value of $R_1 + \dots + R_k$ is large relative to R , then a departure from the null hypothesis of equal mean directions is indicated (see Figure 7.1 of Stephens 1992). In the circular case with two samples, Watson & Williams derived the joint density of R_1 and R_2

given R , and Stephens (1972) provided tables with critical values for the conditional test based on the test statistic $R_1 + R_2$ given R . For the two-sample problem in three dimensions, Stephens (1969) derived the conditional density of R_1 and R_2 given R using the joint distribution of R_1 , R_2 , and R derived earlier by Fisher (1953), and provided tables with critical values for the significance test under certain restrictions. As far as we are aware, critical values have not been calculated for higher dimensions or for comparing more than two groups (also see Mardia & Jupp 2000, page 222).

Harrison, Kanji & Gadsden (1986) suggested a multi-sample test of (3) for circular data, which was later generalized for any dimension in Mardia & Jupp (2000, page 225). This test is based on the ANOVA decomposition

$$\sum_{i=1}^k \sum_{j=1}^{n_i} \|u_{ij} - \bar{u}\|^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} \|u_{ij} - \bar{u}_i\|^2 + \sum_{i=1}^k n_i \|\bar{u}_i - \bar{u}\|^2,$$

which can be rewritten as

$$n(1 - \bar{R}^2) = \left(n - \sum_{i=1}^k n_i \bar{R}_i^2 \right) + \left(\sum_{i=1}^k n_i \bar{R}_i^2 - n \bar{R}^2 \right),$$

where the terms represent, respectively, the total variation, variation within samples and variation between samples. This suggests the test statistic

$$A = \frac{(n - k) \left(\sum_{i=1}^k n_i \bar{R}_i^2 - n \bar{R}^2 \right)}{(k - 1) \left(n - \sum_{i=1}^k n_i \bar{R}_i^2 \right)}, \quad (6)$$

which compares between-sample and within-sample variability. The null hypothesis of equal mean directions is rejected for large values of A , which is referred to its high-concentration null distribution, which is again F with $(k - 1)(d - 1)$ and $(n - k)(d - 1)$ degrees of freedom.

Of course the hypotheses in (3) can also be tested using the likelihood-ratio test statistic (see Mardia & Jupp 2000, page 224)

$$G = 2 \left(\hat{\kappa}_a \sum_{i=1}^k R_i - \hat{\kappa}_0 R + na_d(\hat{\kappa}_a) - na_d(\hat{\kappa}_0) \right), \quad (7)$$

where $\hat{\kappa}_0$ and $\hat{\kappa}_a$ are the maximum likelihood estimators (MLEs) of κ under the null and alternative hypotheses, respectively, and $a_d(\kappa) = \log(\kappa^{d/2-1}/I_{d/2-1}(\kappa))$. The null hypothesis is rejected for large values of G , whose large-sample asymptotic null distribution is chi-

squared with $(k - 1)(d - 1)$ degrees of freedom. For small values of κ , Mardia & Jupp (2000, page 225) suggested incorporating a multiplicative factor into G to improve the chi-squared approximation. In the circular case, Upton (1976) retained the chi-squared reference distribution, but used an approximation of the likelihood-ratio test statistic obtained by applying the approximations and substitutions described in Upton (1973). Note however that Upton's (1976) primary motivation was to develop a test that could be calculated easily, and specifically, one that did not require evaluation of Bessel functions, considerations which are not particularly relevant in today's computing environment.

In this paper we present an improved version of the multi-sample likelihood-ratio test for equality of mean directions in the general d -dimensional case, assuming a common, unknown concentration. Our procedure is based on a simple, monotone transformation of the likelihood-ratio statistic which can be applied with any number of samples in any dimension. We show that, under the null hypothesis, the high-concentration asymptotic distribution of our modified likelihood-ratio statistic is F with $(k - 1)(d - 1)$ and $(n - k)(d - 1)$ degrees of freedom. We demonstrate by simulation that the null distribution of the test statistic is well approximated by this F distribution even with small sample sizes and moderate concentrations. Our simulations of size and power show that our test is superior to competing tests. We also illustrate the practical application of our test by analyzing three-dimensional directions of primate vertebral facets.

2. The High-Concentration Likelihood-Ratio Test

Let $U_c = (U_{11}, \dots, U_{1n_1}, \dots, U_{k1}, \dots, U_{kn_k})$ denote the $d \times n$ matrix of the combined k independent random samples of sizes n_i from $L_d(\mu_i, \kappa)$, $i = 1, \dots, k$. The likelihood function, using (1), is

$$L(\mu_1, \dots, \mu_k; \kappa; U_c) = \left(\frac{(\kappa/2)^{d/2 - 1}}{\Gamma(d/2) I_{d/2 - 1}(\kappa)} \right)^n \exp \left(\kappa \sum_{i=1}^k \sum_{j=1}^{n_i} \mu_i^\top U_{ij} \right).$$

Under the null hypothesis in (3), the MLEs of the common mean direction and the common concentration κ are $\hat{\mu}_0 = \sqrt{R} \bar{R}$ and $\hat{\kappa}_0 = A_d^{-1}(\bar{R})$ respectively (see Mardia & Jupp 2000, page 224), where $A_d^{-1}(\cdot)$ denotes the inverse of the function $A_d(\cdot)$ defined in (2). Under the alternative hypothesis, the MLEs of μ_i and κ are $\hat{\mu}_{ai} = \sqrt{R_i} \bar{R}_i$ and $\hat{\kappa}_a = A_d^{-1}(\tilde{R})$, where $\tilde{R} = n^{-1} \sum_{i=1}^k R_i$. Therefore, the likelihood ratio is

$$\Lambda = \frac{L(\hat{\mu}_0, \hat{\kappa}_0; U_c)}{L(\hat{\mu}_{a1}, \dots, \hat{\mu}_{ak}, \hat{\kappa}_a; U_c)} = \left(\frac{(\hat{\kappa}_0^{d/2 - 1} / I_{d/2 - 1}(\hat{\kappa}_0)) \exp(\hat{\kappa}_0 A_d(\hat{\kappa}_0))}{(\hat{\kappa}_a^{d/2 - 1} / I_{d/2 - 1}(\hat{\kappa}_a)) \exp(\hat{\kappa}_a A_d(\hat{\kappa}_a))} \right)^n. \quad (8)$$

For large κ , the $L_d(\boldsymbol{\mu}, \kappa)$ distribution can be approximated by a spherical normal distribution in the hyperplane tangent to \mathbb{S}^{d-1} at $\boldsymbol{\mu}$ (see Mardia & Jupp 2000, page 172). For testing (3), this suggests the following power transformation of (8), which is analogous to the transformation relating the likelihood ratio statistic and the usual “ratio of sums of squares” F-test statistic in a classical one-way analysis of variance (see Scheffe 1961, page 36):

$$P = \frac{n-k}{k-1} \left(\Lambda^{-2/(n(d-1))} - 1 \right). \quad (9)$$

We will show that the high-concentration asymptotic null distribution of P is F with $(k-1)(d-1)$ and $(n-k)(d-1)$ degrees of freedom. Of course, for fixed positive concentration, the null distribution of the likelihood-ratio test statistic $G = -2 \log \Lambda$ in (7) is approximately $\chi^2_{(k-1)(d-1)}$ when the group sizes n_j are large. Using the Maclaurin series expansion of the exponential function, we see that under the null hypothesis, for large n ,

$$P = \frac{n-k}{k-1} \left(\exp \left(\frac{G}{n(d-1)} \right) - 1 \right) = \frac{G}{(k-1)(d-1)} + O_p(n^{-1}),$$

and therefore P is approximately distributed as $\chi^2_{(k-1)(d-1)}$ scaled by $((k-1)(d-1))^{-1}$.

This scaled chi-squared distribution is also the large- n limit of the F reference distribution suggested by the high-concentration theory, and therefore the F approximation is applicable in both high-concentration and large-sample settings. Note that this is not true for the tests based on W in (5) and A in (6).

In the remainder of this section, we will be concerned with establishing the high-concentration null distribution of P . We adopt the standard notation O and O_p for quantities that are bounded and bounded in probability, respectively. For brevity, we will omit $\kappa \rightarrow \infty$ in expressions involving $\overset{d}{\rightarrow}$, O or O_p , taking this as implied. We will also take it as implied that statements involving O_p hold for the null hypothesis.

For large κ , the modified Bessel function $I_\nu(\kappa)$ can be expanded as (see Abramowitz & Stegun 1970, 9.7.1)

$$I_\nu(\kappa) = \frac{\exp(\kappa)}{(2\pi\kappa)^{1/2}} \left(1 - \frac{4\nu^2 - 1}{8\kappa} + \frac{(4\nu^2 - 1)(4\nu^2 - 9)}{128\kappa^2} + O(\kappa^{-3}) \right). \quad (10)$$

Using this, Schou (1978) determined that the mean resultant length has expansion

$$A_d(\kappa) = 1 - \frac{d-1}{2\kappa} + \frac{(d-1)(d-3)}{8\kappa^2} + O(\kappa^{-3}). \quad (11)$$

Letting $y = y(\kappa) = 1 - A_d(\kappa)$, or equivalently, $\kappa = A_d^{-1}(1 - y)$, from (11) it follows that, as $y \downarrow 0$,

$$\frac{1}{\kappa} = \frac{2}{d-1}y + \frac{(d-3)}{(d-1)^2}y^2 + O(y^3). \quad (12)$$

We will apply this expansion to each of the random variables $\hat{\kappa}_0$ and $\hat{\kappa}_a$.

Under H_0 , from (4), it follows that

$$2\kappa n(1 - \bar{R}) \xrightarrow{d} \chi_{(n-1)(d-1)}^2,$$

and thus

$$\begin{aligned} 1 - \bar{R} &= O_p(\kappa^{-1}) \\ 1 - \tilde{R} &= O_p(\kappa^{-1}). \end{aligned} \quad (13)$$

Since the MLEs of κ under the null and alternative hypotheses are $\hat{\kappa}_0 = A_d^{-1}(\bar{R})$ and $\hat{\kappa}_a = A_d^{-1}(\tilde{R})$ respectively, we separately take $y = 1 - \bar{R}$ and $y = 1 - \tilde{R}$ in the asymptotic expansion (12) for κ^{-1} , and use (13), to obtain

$$\begin{aligned} \frac{1}{\hat{\kappa}_0} &= \frac{2(1 - \bar{R})}{d-1} + O_p(\kappa^{-2}) \\ \frac{1}{\hat{\kappa}_a} &= \frac{2(1 - \tilde{R})}{d-1} + O_p(\kappa^{-2}). \end{aligned} \quad (14)$$

Now, from the asymptotic expansions for $I_{(d-2)/2}(\kappa)$ and $A_d(\kappa)$ in (10) and (11), and the Maclaurin series expansions of the exponential function $\exp(x)$ and geometric series $(1 - x)^{-1}$, we have

$$\begin{aligned} \frac{\kappa^{d/2-1}}{\Gamma_{d/2-1}(\kappa)} \exp(\kappa A_d(\kappa)) &= \kappa^{(d-1)/2} (2\pi)^{1/2} \exp(-\kappa(1 - A_d(\kappa))) (1 + O(\kappa^{-1}))^{-1} \\ &= \kappa^{(d-1)/2} (2\pi)^{1/2} \exp(-2^{-1}(d-1) + O(\kappa^{-1})) (1 + O(\kappa^{-1}))^{-1} \\ &= \kappa^{(d-1)/2} (2\pi)^{1/2} \exp(-2^{-1}(d-1)) (1 + O(\kappa^{-1})). \end{aligned}$$

Substituting this expression in (8), and using the expansions in (14) for $\hat{\kappa}_0^{-1}$ and $\hat{\kappa}_a^{-1}$, yields

$$\Lambda = \left(\frac{1 - \tilde{R} + O_p(\kappa^{-2})}{1 - \bar{R} + O_p(\kappa^{-2})} \right)^{(d-1)/2} \left(\frac{1 + O_p(\kappa^{-1})}{1 + O_p(\kappa^{-1})} \right)^n. \quad (15)$$

Now, using (15) in (9), together with the binomial series for $(1+x)^a$, $a \in \mathbb{R}^1$, and the result $(1 - \tilde{R})^{-1} = O_p(\kappa)$ implied by (4), we have

$$\begin{aligned} P &= \frac{n-k}{k-1} \left(\frac{1 - \bar{R} + O_p(\kappa^{-2})}{(1 - \tilde{R})(1 + O_p(\kappa^{-1}))} \right) \left(1 + O_p(\kappa^{-1}) \right) - 1 \\ &= \frac{n-k}{k-1} \left(\frac{1 - \bar{R}}{(1 - \tilde{R})} + O_p(\kappa) O_p(\kappa^{-2}) \right) \left(1 + O_p(\kappa^{-1}) \right) - 1 \\ &= \frac{n-k}{k-1} \left(\frac{\tilde{R} - \bar{R}}{1 - \tilde{R}} + O_p(\kappa^{-1}) \right) = W + O_p(\kappa^{-1}), \end{aligned}$$

where W is the Watson & Williams test statistic (5). Therefore, under the null hypothesis, the test statistic P is equal to W up to a first-order approximation, and from this it follows that the high-concentration asymptotic null distribution of P is F with $(k-1)(d-1)$ and $(n-k)(d-1)$ degrees of freedom, as stated earlier.

Using additional terms in the asymptotic expansions, one can show that

$$P = W + \frac{n-k}{k-1} \left(\frac{1 - (d-2)^2}{2(d-1)^2} \right) \frac{(\tilde{R} - \bar{R})(1 - \bar{R})}{(1 - \tilde{R})} + O_p(\kappa^{-2}).$$

In the second term on the right-hand side of this expression, the factor $(\tilde{R} - \bar{R})(1 - \bar{R})(1 - \tilde{R})$ is of order κ^{-1} in probability and is always non-negative. The factor $(1 - (d-2)^2)/(2(d-1)^2)$ is zero for $d=3$, and thus, as $\kappa \rightarrow \infty$, the difference between the test statistics P and W diminishes faster for $d=3$ than for other dimensions. Also, notice that for $d=2$, this factor is positive (0.5), and therefore the test statistic P will be greater than W for κ sufficiently large; for $d=4$, this coefficient is negative, and thus P will be smaller than W for large values of κ . These facts are reflected in the simulations in the next section.

It can also be shown, using the equality $1 - \bar{R}^2 = 2(1 - \bar{R})(1 - (1 - \bar{R})/2)$, that $A = W + O_p(\kappa^{-1})$. However, the simulation results of the next section will demonstrate that the test based on P outperforms the tests based on W , A , and G , as the null distribution of P is approximated better by the F distribution in both the high-concentration and large-sample settings.

3. Simulations and Comparisons

Simulations of size and power were carried out for the proposed test with statistic P , and for the tests based on the statistics W , G , and A described in the Introduction. We also included the modified version of W given by the formula $M = (1 + 3/(8\hat{\kappa}_0))W$ for $d = 2$ (Stephens 1972), and $M = (1 - 1/(5\hat{\kappa}_0^2))W$ for $d = 3$ (Mardia & Jupp 2000, page 191), where $\hat{\kappa}_0$ is the MLE of κ under H_0 . Since all tests under consideration are approximate, we first examine how well each of them maintains its nominal size (following Upton 1976, we refer to this as the *accuracy* of the test). We then compare the powers of the tests. Upton's (1976) test for the circular case was not considered in our simulations, because simulations in Upton's (1976) paper showed that this test is similar in power but slightly less accurate compared to Stephens's (1972) test using the statistic M .

The performance of the tests was investigated for dimensions $d = 2$ and $d = 3$ with common mean resultant lengths $\rho = 0.10, 0.20, 0.30, 0.40, 0.45, 0.50, 0.60, 0.75, 0.85, 0.95$, and sample size combinations $(n_1, n_2) = (10, 10), (10, 20), (20, 20), (20, 40), (40, 40)$ for the two-sample problem ($k = 2$) and $(n_1, n_2, n_3) = (10, 10, 10), (10, 20, 20), (20, 20, 40), (20, 30, 40), (40, 40, 40)$ for the three-sample problem ($k = 3$). Limited simulations were also performed for dimensions $d = 4, d = 10$, and $d = 100$. We focus here on simulation results for values of $\rho \geq 0.45$, but note that for $\rho < 0.45$, only the test based on P maintained a size reasonably close to its nominal size.

The generation of a random direction from the Langevin-von Mises-Fisher distribution can be accomplished with the procedures described in Best & Fisher (1979) for $d = 2$ and Fisher, Lewis & Willcox (1981) for $d = 3$. We used R (R Core Team 2015) for all computations. The R package *circular* (Agostinelli & Lund 2013) was used to generate random directions for the two-dimensional case, and the R package *movMF* (Hornik & Grün 2014) for the cases $d = 4, d = 10$, and $d = 100$.

All tests considered are invariant to rotation of the circle/sphere, so in all cases one of the population mean directions can be set to the angle 0 for $d = 2$, and to the direction with colatitude and longitude both equal to 0 for $d = 3$; we will refer to these as the reference directions. In the simulations for power, for $k = 2$, the second population mean direction was chosen to be separated from the reference by the angle $\delta = 5, 10, 15, 20, 25, 30, 35, 40$ degrees. For $k = 3$, two scenarios were used: in the first, two groups shared a common population mean direction defined as the reference, and the third group had a mean direction δ degrees away; in the second, the three groups had coplanar mean directions, two of which formed an angle of 2δ degrees and the third was the bisector of this angle. All test statistics were simulated 100 000 times for each combination of sample size, mean resultant length ρ , and separation angle δ .

Under the null hypotheses, p-values should be uniformly distributed on the interval $(0, 1)$. In Figures 1 and 2, Q-Q plots are shown comparing the p-values obtained by simulation under the null hypothesis to the target uniform distribution for $d = 2, k = 2, (n_1, n_2) = (10, 10)$, and for $d = 3, k = 3, (n_1, n_2, n_3) = (20, 30, 40)$. In these plots, the ordered observed p-values (vertical axes) obtained under the null hypothesis are plotted against the corresponding

quantiles (horizontal axes) of the uniform distribution. The axes are truncated to 0 to 0.10 to provide more detail in the range of most interest in hypothesis testing. A test maintains its nominal size if the simulated curve coincides with the (dashed) diagonal line; the test is conservative if the curve lies above the diagonal line, and anti-conservative if the curve lies below the diagonal line.

Table 1 provides further details by giving the actual sizes of the tests for nominal sizes $\alpha = 0.01, 0.05, 0.10$. Tables 2 and 3 compare the powers of the tests at nominal size $\alpha = 0.05$ as the mean resultant length ρ increases from 0.45 to 0.85. Similar figures and tables for the other scenarios considered are provided in the supplementary materials.

Based on these results, we make the following observations:

- With the exception of the standard likelihood-ratio test based on G (using the chi-squared approximation), the accuracy of each test improves with increasing ρ . However, the modified likelihood-ratio test based on P is consistently the best at maintaining its size, especially for values of ρ less than 0.75.
- As expected from the discussion at the end of Section 2, when $\rho = 0.45$, the Watson & Williams test based on W is conservative for $d = 2$, but anti-conservative for $d = 3$. When $\rho < 0.45$, this test changes to being anti-conservative for $d = 2$, but remains anti-conservative for $d = 3$.
- The test based on M is usually better than W , but our simulations show that this is not always the case. In particular, M seems to be less accurate than W in the two-dimensional case when $\rho = 0.50$.
- For $d = 2$ and $d = 3$, the test based on A is conservative even for very large values of ρ , and its power never exceeded the power of any of the other tests when $\rho = 0.45$ (see supplementary figures and tables).
- The accuracy of the likelihood-ratio test statistic G is poor, although it does improve with increasing sample size. Among the other tests, only the accuracy of the modified likelihood-ratio test based on P seems to improve with increasing sample size, at least over the range of sample sizes considered here (see supplementary figures and tables).
- As expected (see Tables 2 and 3), the power of each test increases with ρ , and, not surprisingly, the anti-conservative tests based on M and G (and on W when $d = 3$) have an artificial power advantage which is reflected in these tables. However, this advantage appears to be entirely explained by the discrepancy in the actual size of the tests. For example, if we use our simulation results to calibrate the tests based on P and W to have equal size, then the powers of the tests become essentially equal (see Tables 5 and 6 in the supplementary materials).
- Finally, the test based on P appears to maintain its advantage over the tests based on W , M , and G as the dimension increases to $d = 4$, $d = 10$, and $d = 100$. For $d = 4$ and $d = 10$, the test with P again seems to perform better than the test with A . For $d = 100$, P and A appear to maintain their nominal size even for

concentrations down to $\rho = 0.10$, and they also seem to have similar power. Note that for $d = 4$, $d = 10$ and $d = 100$, the multiplicative factor in M was calculated using only the first three terms of the expansion provided for the correction in Stephens (1992); the performance of M may improve if more terms of this expansion are used.

In conclusion, these simulation results show that the test based on P maintains its size and power across the range of concentrations usually encountered, much better than any of the other tests considered. This test does not require any special adjustments that depend on dimension or (estimated) concentration, nor does it require the use of the bootstrap or other computationally intensive techniques. There seems to be no reason that this version of the likelihood-ratio test should not be adopted in preference to the other tests in this comparison.

4. Orientations of Primate Vertebral Facets

Anthropologists have long been interested in the relationship between form and function. One specific area of research is the relationship between skeletal form and locomotion, including in particular the spine morphology and locomotor forces transmitted in the lumbar vertebrae in orthograde primates (see Johnson & Shapiro 1998).

In this example we consider the three-dimensional orientation of the last lumbar inferior facets in three species of primates: chimpanzees, gorillas, and humans. The data were collected at the Division of Anthropology and the Division of Vertebrate Zoology at the American Museum of Natural History in New York City, by Dorion A. Keifer, as part of Keifer's (2005) master's thesis research conducted at the Department of Anthropology at the University of Florida. Details on the methods and procedures for specimen selection and for measuring different vertebral elements can be found in Keifer's (2005) thesis.

To illustrate the application of the proposed testing procedure, we consider the normal vector to the last lumbar right inferior facet, as this vector is considered important for the development of biomechanical models on the force direction in the facet. Table 4 contains the three-dimensional directions of the normal vectors (rounded to three decimal places) for the three samples of primates, with the sample mean directions and sample mean resultant lengths provided in the last two rows. One outlying observation in the humans sample and two in the gorillas sample were excluded from the original data set. Note that our numerical results were obtained using the original (full precision) data. Figure 3 shows the Lambert's azimuthal equal-area projection (see Fisher, Lewis & Embleton 1987) of the data onto the plane tangent to the sphere at the point of the overall sample mean direction of the three samples combined, after applying the rotation (i) of Fisher & Best (1984).

To test the goodness-of-fit of the Langevin-von Mises-Fisher three-dimensional distribution to the data, we performed the tests on colatitude, uniformity, and normality described in Fisher & Best (1984) for each of the three samples, and we found no evidence against the null hypothesis of Langevin distribution at the 0.05 level. Before testing for equality of the mean directions, Bartlett's test of homogeneity was performed to verify the assumption of equal concentrations (see Mardia & Jupp 2000, page 226). The test was expected to perform satisfactorily when $\bar{R} = 0.67$, which is the case in this example as $\bar{R} = 0.97$. The data in the

three samples are highly concentrated, and the p-value from Bartlett's test is 0.127 indicating no evidence of unequal concentrations. We then tested for equality of the three mean directions using the test statistic P and found a statistically significant difference (p-value < 0.0001). Pairwise tests on the mean directions were then performed using P , finding statistically significant differences for all three pairs of primate species at the 0.05 level after accounting for the multiple testing with Holm's (1979) simultaneous testing procedure: Human & Gorilla (p-value < 0.0001), Human & Chimpanzee (p-value < 0.0001), and Gorilla & Chimpanzee (p-value = 0.008).

5. Conclusion

In this paper we propose an improved version of the multi-sample likelihood-ratio test of equality of mean directions for populations having Langevin-von Mises-Fisher distributions with a common unknown concentration in the general d -dimensional case. Our test statistic is a monotone transformation of the likelihood ratio. We show that the high-concentration asymptotic null distribution of the test statistic is the F distribution. We also demonstrate that the F approximation to the null distribution is applicable in the large-sample setting. Our simulations of size and power show that the proposed test outperforms competing tests. The proposed test performs well even with small sample sizes and moderate concentrations across all dimensions considered. Further examination of the behavior of the test under departures from the assumptions of Langevin distribution or equal concentrations would provide beneficial information on its robustness. Additional simulations on the size and power of the test in dimensions higher than three would extend its practical applicability.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The first author thanks the Department of Statistics at the University of Florida, USA, the School of Public Health at the University of Sydney, Australia, and the National Center for Health Statistics at the Centers for Disease Control and Prevention, USA, for their support while the author worked on the results and preparation of this paper. The authors also thank the reviewers for their helpful comments, which significantly improved the paper.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official views of the National Center for Health Statistics, U.S. Centers for Disease Control and Prevention.

References

- Abramowitz, M., Stegun, IA. Handbook of Mathematical Functions. New York: Dover Publications; 1970.
- Agostinelli, C., Lund, U. R package circular: Circular Statistics (version 0.4-7). CA: Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University; Venice, Italy. UL: Department of Statistics, California Polytechnic State University, San Luis Obispo, California, USA; 2013. URL <https://r-forge.r-project.org/projects/circular/>
- Best DJ, Fisher NI. Efficient simulation of the von Mises distribution. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1979; 28:152–157.
- Fisher NI, Best DJ. Goodness-of-fit tests for Fisher's distribution on the sphere. Australian & New Zealand Journal of Statistics. 1984; 26:142–150.

- Fisher, NI., Lewis, T., Embleton, JJ. *Statistical Analysis of Spherical Data*. Cambridge: Cambridge University Press; 1987.
- Fisher NI, Lewis T, Willcox ME. Tests of discordancy for samples from Fisher's distribution on the sphere. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 1981; 30:230–237.
- Fisher RA. Dispersion on a sphere. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*. 1953; 217:295–305.
- Harrison D, Kanji GK, Gadsden RJ. Analysis of variance for circular data. *Journal of Applied Statistics*. 1986; 13:123–138.
- Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979; 6:65–70.
- Hornik K, Grün B. movMF: An R package for fitting mixtures of von Mises-Fisher distributions. *Journal of Statistical Software*. 2014; 58:1–31. URL <http://www.jstatsoft.org/v58/i10/>.
- Johnson SE, Shapiro LJ. Positional behavior and vertebral morphology in atelines and cebines. *American Journal of Physical Anthropology*. 1998; 105:333–354. [PubMed: 9545076]
- Keifer, DA. Master's thesis. University of Florida; Gainesville, FL, USA: 2005. Last lumbar facet and pedicle orientation in orthograde primates.
- Mardia, KV., Jupp, PE. *Directional Statistics*. Chichester: Wiley; 2000.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; Vienna, Austria: 2015. URL <https://www.R-project.org>
- Scheffé, H. *The Analysis of Variance*. New York: Wiley; 1961.
- Schou G. Estimation of the concentration parameter in von Mises-Fisher distributions. *Biometrika*. 1978; 65:369–377.
- Stephens MA. Multi-sample tests for the Fisher distribution for directions. *Biometrika*. 1969; 56:169–181.
- Stephens MA. Multisample tests for the von Mises distribution. *Journal of the American Statistical Association*. 1972; 67:456–461.
- Stephens, MA. On Watson's ANOVA for Directions. In: Mardia, KV., editor. *The Art of Statistical Science: A Tribute to G. S. Watson*. Chichester: Wiley; 1992. p. 75-85.
- Upton GJG. Single-sample tests for the von Mises distribution. *Biometrika*. 1973; 60:87–99.
- Upton GJG. More multisample tests for the von Mises distribution. *Journal of the American Statistical Association*. 1976; 71:675–678.
- Watson GS. Analysis of dispersion on a sphere. *Monthly Notices of the Royal Astronomical Society, Geophysical Supplement*. 1956; 7:153–159.
- Watson, GS. *Statistics on Spheres*. New York: Wiley; 1983.
- Watson GS, Williams EJ. On the construction of significance tests on the circle and the sphere. *Biometrika*. 1956; 43:344–352.

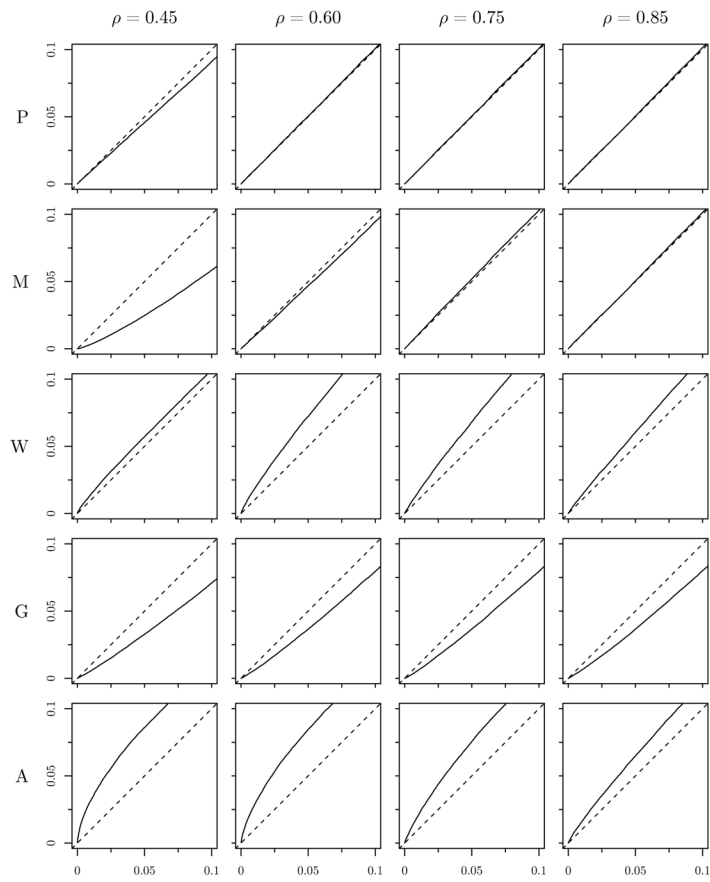


Figure 1. Size of the tests with statistics P , M , W , G , and A (100 000 simulated values each), examined by plotting the ordered p-values (vertical axis) versus uniform quantiles (horizontal axis), for $d = 2$, $k = 2$, $(n_1, n_2) = (10, 10)$, and $\rho = 0.45, 0.60, 0.75, 0.85$.

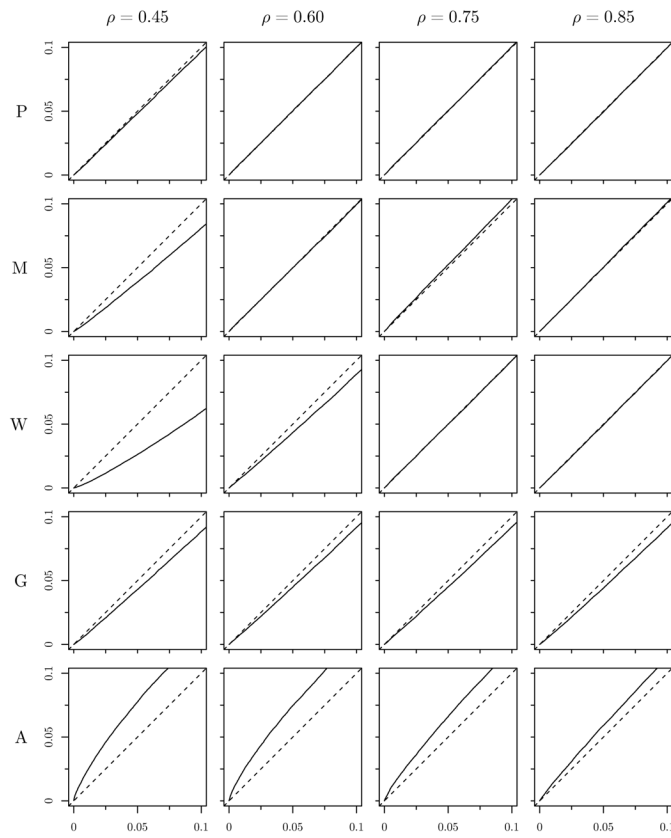


Figure 2. Size of the tests with statistics P , M , W , G , and A (100 000 simulated values each), examined by plotting the ordered p-values (vertical axis) versus uniform quantiles (horizontal axis), for $d = 3$, $k = 3$, $(n_1, n_2, n_3) = (20, 30, 40)$, and $\rho = 0.45, 0.60, 0.75, 0.85$.

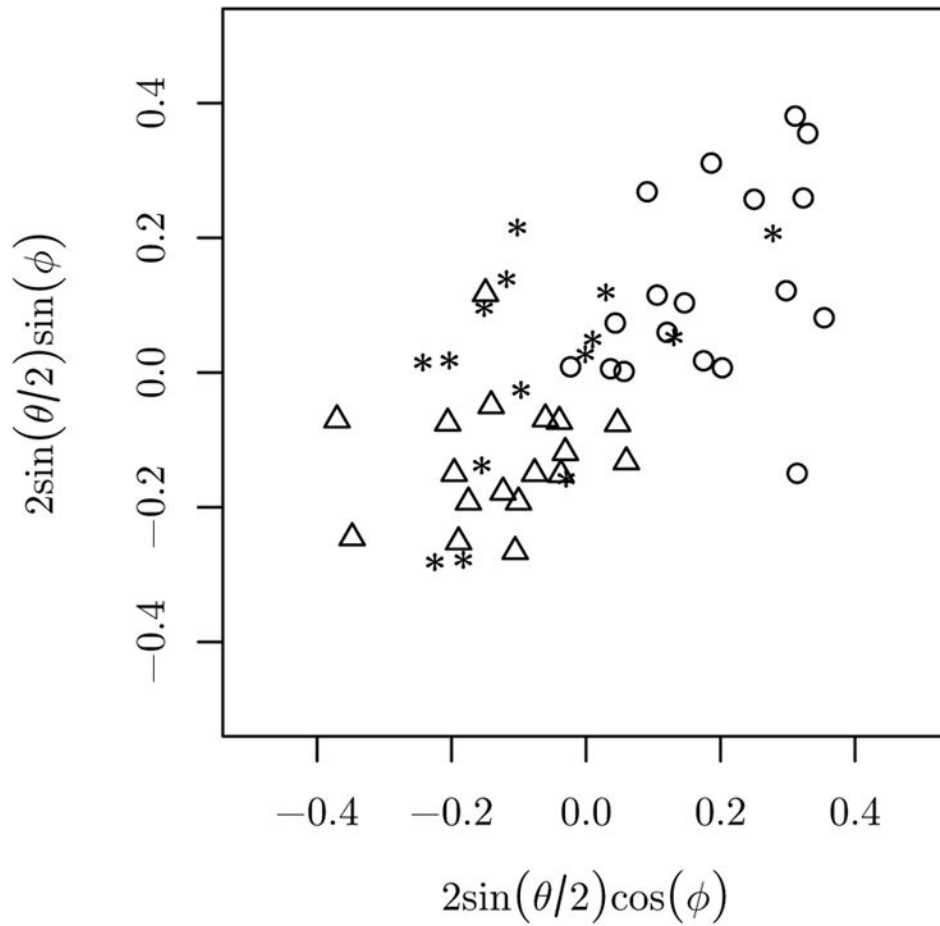


Figure 3. Three-dimensional directions of the normal vectors to the last lumbar right inferior facets in the samples of humans (circles), gorillas (asterisks), and chimpanzees (triangles) (Table 4), projected onto the plane tangent to the sphere at the point of the combined sample mean direction using Lambert’s azimuthal equal-area projection. The angles θ and ϕ denote the colatitude and longitude of the direction respectively, after applying the rotation (i) of Fisher & Best (1984).

Table 1

Actual size of the tests based on the statistics $P, M, W, G,$ and A (100 000 simulated values each), at nominal size $\alpha = 0.01, 0.05, 0.10,$ for $d = 2, k = 2,$ $(n_1, n_2) = (10, 10),$ and $d = 3, k = 3, (n_1, n_2, n_3) = (20, 30, 40),$ and $\rho = 0.45, 0.60, 0.75, 0.85.$

$d = 2, k = 2, (n_1, n_2) = (10, 10)$						
$\rho =$	0.45			0.60		
$\alpha =$	0.01	0.05	0.10	0.01	0.05	0.10
P	0.011	0.055	0.110	0.010	0.050	0.099
M	0.024	0.088	0.154	0.011	0.053	0.106
W	0.007	0.043	0.093	0.005	0.033	0.072
G	0.017	0.073	0.135	0.016	0.066	0.122
A	0.002	0.021	0.064	0.002	0.023	0.064
$\rho =$	0.75			0.85		
$\alpha =$	0.01	0.05	0.10	0.01	0.05	0.10
P	0.010	0.050	0.099	0.010	0.050	0.099
M	0.010	0.048	0.097	0.010	0.050	0.099
W	0.006	0.035	0.077	0.007	0.041	0.085
G	0.016	0.066	0.123	0.016	0.066	0.122
A	0.004	0.030	0.072	0.006	0.037	0.081
$d = 3, k = 3, (n_1, n_2, n_3) = (20, 30, 40)$						
$\rho =$	0.45			0.60		
$\alpha =$	0.01	0.05	0.10	0.01	0.05	0.10
P	0.011	0.052	0.103	0.010	0.050	0.100
M	0.014	0.064	0.122	0.010	0.050	0.100
W	0.022	0.087	0.155	0.012	0.057	0.112
G	0.012	0.058	0.112	0.012	0.056	0.109
A	0.003	0.028	0.070	0.004	0.030	0.073
$\rho =$	0.75			0.85		

$p = 3, k = 3, (n_1, n_2, n_3) = (20, 30, 40)$

$\alpha =$	$\beta = 0.45$			$\beta = 0.60$		
	0.01	0.05	0.10	0.01	0.05	0.10
$\alpha =$	0.009	0.049	0.099	0.010	0.051	0.100
P	0.009	0.049	0.099	0.010	0.051	0.100
M	0.009	0.048	0.097	0.010	0.050	0.099
W	0.010	0.050	0.100	0.010	0.051	0.100
G	0.011	0.055	0.108	0.012	0.056	0.109
A	0.005	0.036	0.081	0.007	0.042	0.088

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Power of the tests with statistics P, M, W , and G (100 000 simulated values each) at nominal size $\alpha = 0.05$, for $d = 2, k = 2, (n_1, n_2) = (10, 10)$, mean directions separated by $\delta = 5, 10, 15, 20, 25, 30, 35, 40$ degrees, and $\rho = 0.45, 0.60, 0.75, 0.85$.

δ	$\rho = 0.45$				$\rho = 0.60$			
	P	M	W	G	P	M	W	G
5	0.06	0.09	0.05	0.08	0.05	0.06	0.04	0.07
10	0.06	0.10	0.05	0.08	0.06	0.07	0.04	0.08
15	0.07	0.11	0.06	0.09	0.08	0.09	0.06	0.10
20	0.09	0.13	0.07	0.11	0.11	0.12	0.08	0.13
25	0.10	0.15	0.08	0.13	0.14	0.15	0.11	0.17
30	0.12	0.18	0.10	0.15	0.18	0.20	0.14	0.22
35	0.15	0.20	0.12	0.18	0.23	0.25	0.18	0.28
40	0.17	0.24	0.15	0.21	0.29	0.31	0.23	0.33

δ	$\rho = 0.75$				$\rho = 0.85$			
	P	M	W	G	P	M	W	G
5	0.06	0.05	0.04	0.07	0.06	0.06	0.05	0.08
10	0.08	0.08	0.06	0.10	0.10	0.10	0.09	0.13
15	0.12	0.11	0.09	0.14	0.17	0.17	0.14	0.20
20	0.17	0.16	0.13	0.20	0.26	0.26	0.23	0.31
25	0.24	0.23	0.19	0.28	0.38	0.38	0.34	0.43
30	0.32	0.31	0.26	0.37	0.50	0.50	0.46	0.56
35	0.41	0.40	0.34	0.46	0.63	0.63	0.59	0.68
40	0.50	0.49	0.43	0.55	0.74	0.74	0.70	0.79

Table 3

Power of the tests with statistics P, M, W , and G (100 000 simulated values each) at nominal size $\alpha = 0.05$, for $d = 3, k = 3, (n_1, n_2, n_3) = (20, 30, 40)$, two equal mean directions and one separated by $\delta = 5, 10, 15, 20, 25, 30, 35, 40$ degrees, and $\rho = 0.45, 0.60, 0.75, 0.85$.

δ	$\rho = 0.45$				$\rho = 0.60$			
	P	M	W	G	P	M	W	G
5	0.06	0.07	0.09	0.06	0.06	0.06	0.07	0.07
10	0.07	0.09	0.12	0.08	0.10	0.10	0.11	0.11
15	0.10	0.12	0.16	0.11	0.17	0.17	0.18	0.18
20	0.15	0.18	0.22	0.16	0.27	0.28	0.30	0.29
25	0.22	0.25	0.30	0.23	0.42	0.42	0.45	0.44
30	0.30	0.34	0.39	0.32	0.58	0.58	0.61	0.60
35	0.40	0.44	0.50	0.42	0.72	0.73	0.75	0.74
40	0.50	0.54	0.61	0.52	0.85	0.85	0.87	0.86

δ	$\rho = 0.75$				$\rho = 0.85$			
	P	M	W	G	P	M	W	G
5	0.07	0.07	0.07	0.08	0.10	0.09	0.10	0.10
10	0.16	0.15	0.16	0.17	0.27	0.27	0.27	0.29
15	0.32	0.31	0.32	0.34	0.58	0.58	0.58	0.60
20	0.55	0.54	0.55	0.57	0.85	0.85	0.85	0.86
25	0.76	0.76	0.76	0.78	0.97	0.97	0.97	0.97
30	0.91	0.90	0.91	0.91	1.00	1.00	1.00	1.00
35	0.97	0.97	0.97	0.98	1.00	1.00	1.00	1.00
40	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00

Three-dimensional directions of the normal vectors to the last lumbar right inferior facets in samples of three species of primates (humans, gorillas, and chimpanzees) rounded to three decimal places. The mean direction \bar{R}^{-1} and the mean resultant length \bar{R} are calculated for each sample separately.

Table 4

	Humans			Gorillas			Chimpanzees		
	(u_1, u_2, u_3)			(u_1, u_2, u_3)			(u_1, u_2, u_3)		
1	0.402	-0.623	-0.671	0.551	-0.794	-0.257	0.204	-0.962	-0.181
2	0.621	-0.472	-0.626	0.456	-0.872	-0.177	0.311	-0.887	-0.341
3	0.445	-0.756	-0.481	0.439	-0.818	-0.371	0.302	-0.936	-0.181
4	0.477	-0.722	-0.502	0.178	-0.974	-0.143	0.288	-0.910	-0.297
5	0.414	-0.815	-0.406	0.513	-0.761	-0.397	0.259	-0.945	-0.199
6	0.407	-0.809	-0.425	0.526	-0.598	-0.605	0.251	-0.929	-0.271
7	0.554	-0.540	-0.634	0.177	-0.967	-0.185	0.390	-0.890	-0.238
8	0.207	-0.743	-0.637	0.406	-0.869	-0.281	0.355	-0.873	-0.335
9	0.452	-0.635	-0.626	0.436	-0.755	-0.490	0.338	-0.845	-0.415
10	0.631	-0.577	-0.518	0.457	-0.804	-0.381	0.180	-0.948	-0.261
11	0.495	-0.733	-0.466	0.519	-0.824	-0.227	0.372	-0.912	-0.174
12	0.574	-0.581	-0.577	0.274	-0.899	-0.340	0.362	-0.877	-0.315
13	0.621	-0.648	-0.441	0.456	-0.879	-0.138	0.268	-0.930	-0.250
14	0.380	-0.742	-0.552	0.309	-0.925	-0.222	0.282	-0.860	-0.424
15	0.395	-0.751	-0.529	0.611	-0.745	-0.268	0.384	-0.923	-0.009
16	0.472	-0.779	-0.412				0.537	-0.812	-0.228
17	0.427	-0.833	-0.351				0.220	-0.975	-0.024
18	0.645	-0.464	-0.607				0.281	-0.900	-0.334
\bar{R}^{-1}	0.487	-0.691	-0.535	0.430	-0.850	-0.305	0.314	-0.916	-0.251
\bar{R}	0.98			0.98			0.99		

Data source: Keifer (2005).