

Statistical inference of Langevin distribution for directional data

Yoko WATAMORI

(Received September 20, 1994)

ABSTRACT. Comparison of some estimators and multi-sample tests about mean directions for the Langevin distribution have been studied. Before displaying main results, the background of directional statistics is briefly considered.

We derive the expectations and MSEs (mean square errors) of the MLEs (Maximum Likelihood estimators) of concentration parameter, κ and mean direction, μ in the forms of asymptotic expansions. We also compare the marginal MLE of κ with the MLE. It is shown that the estimators so modified as to satisfy a higher order asymptotic unbiasedness are the same in a higher order asymptotic sense. Further, it is shown that those estimators have smaller MSEs than the original ones when κ is not so small, but for small κ the MLE is preferable.

Next we consider some multi-sample tests for mean directions, μ_i 's. Two cases are studied in detail. Namely, all μ_i 's are on the same but unknown axis and μ_i 's are in the given subspace.

1. Introduction

This paper is concerned primarily with the Langevin distribution for directional data. In general, there are three basic approaches to directional statistics, which are called embedding, wrapping and intrinsic approaches. For a discussion of these approaches, see Jupp and Mardia [1989]. They are usually used in different area, depending on their merits. In inferential problems, the embedding approach is commonly used, *i.e.*, we consider to embed $(p - 1)$ -dimensional sphere \mathcal{S}^{p-1} into p -dimensional Euclidean space \mathcal{R}^p . We also discuss a little more about this topic in the last section.

In section 2, we summarize some backgrounds of directional statistics, especially related to the Langevin distribution with a little refinement or improvement. Most of them are found in Mardia [1972] and Watson [1983a]. Therefore we give only an outline without strict proofs. Those are used in subsequent sections and will be helpful in understanding the rest of this paper. Although we do not put importance on applications here, Fisher et al. [1987] and Fisher [1993] have given a lot of examples.

1991 *Mathematics Subject Classification.* 62H11, 62F12, 62H15.

Key words and phrases. Asymptotic expansion, Embedding, Wrapping, Central limit theorem, Langevin distribution, Marginal maximum likelihood estimator, Maximum likelihood estimator.

A random vector \mathbf{x} in \mathcal{R}^p of its length $\|\mathbf{x}\|$ unity is said to have a p -variate Langevin distribution $M_p(\boldsymbol{\mu}, \kappa)$ if its probability density function is given by

$$\{a_p(\kappa)\}^{-1} \exp(\kappa \boldsymbol{\mu}' \mathbf{x}) \quad (1.1)$$

on the $(p-1)$ -dimensional unit sphere $\mathcal{S}^{p-1} = \{\mathbf{x} \mid \mathbf{x} \in \mathcal{R}^p, \|\mathbf{x}\| = (\mathbf{x}'\mathbf{x})^{1/2} = 1\}$, where C' denotes a transpose of a matrix C , $\|\boldsymbol{\mu}\| = 1$ and $\kappa > 0$. The normalizing constant is given by

$$a_p(\kappa) = (2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa) \kappa^{-\frac{p}{2}+1},$$

where $I_\nu(\kappa)$ is the modified Bessel function of the first kind of order ν . The parameters $\boldsymbol{\mu}$ and κ are called the mean direction vector and the concentration parameter, respectively. This form of distributions was first introduced by Watson and Williams [1956], but the distributions in the special cases $p=2$ and $p=3$ were first derived by von Mises [1918] and Fisher [1953], respectively.

The Langevin distribution plays an important role among distributions on a sphere. It is a member of (curved-) exponential family, whose shape is symmetric and unimodal. So it is one of the most common distributions for directional data. For the special case $p=2$, there is another common distribution that is called wrapped normal distribution. These two distributions have quite similar properties, but there exist some differences between them as discussed by Collett and Lewis [1981]. Though the normalizing constant of the density (1.1) is complicated, it is relatively easy to carry out asymptotic approximations or expansions. Therefore asymptotic theories have been investigated.

The problems of estimating $\boldsymbol{\mu}$ and κ have been investigated by several authors. For example, Best and Fisher [1981] noted that the MLE of κ is quite biased by a simulation study and they proposed a new estimator in the case of von Mises or Fisher distributions, *i.e.*, $p=2, 3$. Ducharme and Milasevic [1987, 1990] proposed to estimate κ as well as $\boldsymbol{\mu}$ based on the spatial median. Fisher and Lewis [1983] considered estimating common mean directions for a circle or a sphere. Watson [1986] discussed the optimality of estimations about modal directions for the class of rotationally symmetric distributions. Ronchetti [1992] has given optimal robust estimators for the concentration parameter and proposed a general class of M-estimators for directional data. Moreover, Bartels [1984] considered the case of bidirectional mixture of von Mises distributions and Spurr and Koutbeiy [1991] compared various methods for estimating parameters in this case.

In section 3, we consider the MLE $\hat{\kappa}$ and the marginal MLE $\tilde{\kappa}$ of κ . Then

we derive the modified estimators of $\hat{\kappa}$ and $\tilde{\kappa}$ which satisfy higher order asymptotic unbiasedness in the case when the sample size n is large. It is shown that these two modified estimators are the same as in a higher order asymptotic sense. Further, it is shown that with relatively large n those estimators have smaller MSEs than the original ones when κ is not so small. On the other hand, MLE is preferable for small κ . We also consider the case when κ is large. The asymptotic distributions of these estimators are different, according to the cases when n is large or κ is large.

For various testing problems on mean directions and concentration parameters, Watson [1983a, 1983b, 1984] obtained asymptotic distributions of some test statistics. Watson [1983c] also extended large sample theory related to these tests to a more general class of distributions with rotational symmetries. Stephens [1969] studied multi-sample tests for the Fisher distribution, *i.e.*, $p = 3$ and derived a two-sample exact test for the equality of modal vectors or of concentration parameters. Stephens [1972] also investigated multi-sample tests for the von Mises distribution, *i.e.*, $p = 2$. Higher order asymptotic theories have been considered recently. Chou [1986] derived asymptotic expansions of some statistics proposed by Watson [1983a] about a mean direction and Hayakawa [1990] also derived those of some other test statistics. Fujikoshi and Watamori [1992], Watamori [1992] studied some related testing problems, too.

In sections 4 and 5, we consider two multi-sample hypotheses about mean directions. One is that all mean directions are contained in a given subspace of \mathcal{S}^{p-1} . The other is that all mean directions are on the same but unknown line. The latter one is an extension of the equality of mean directions studied by Hayakawa [1992]. He also considered the equality of concentration parameters. We derive asymptotic null and nonnull distributions of some test statistics and compare these. Section 4 is concerned with the large sample situation. On the other hand, section 5 is concerned with the highly concentrated situation.

In section 6, we discuss more fundamental notions of directional statistics. We proceed discussion on different stance from that of previous sections. First we point out some weak points of embedding approaches. Next some possibility of overcoming them is considered, based on a recent work due to Watamori and Kakimizu [1994]. Some related topics and recent trend are also discussed briefly.

2. Backgrounds

In directional statistics, we consider the case where data are on a unit sphere. Some of them are directions and others are axial data. We

distinguish directions from axial data by taking account of their signs. Because spheres are compact and have non-zero curvature, we should proceed statistical theories with special care. Distributions on a sphere can be regarded as a kind of singular ones by embedding them into a Euclidean space. However, the singular nature does not mean that all distributions on a sphere have singular dispersion matrices, but that any distribution on a sphere has the following moment relationship,

$$\text{tr } \Sigma + \|\boldsymbol{\mu}\|^2 = 1, \quad (2.1)$$

where $\text{tr } \Sigma$ is the trace of the dispersion matrix and $\boldsymbol{\mu}$ is the mean vector.

As one of the motivations why we study the distributions on a sphere, it may be noted that there are three data sets mainly quoted. One is 'the declinations and dips of remnant magnetization in nine specimens of Icelandic lava flows of 1947-48' given in Fisher [1953]. The others are 'the vanishing angles of 209 homing pigeons in a clock resetting experiment' (Mardia [1972, p. 123]) and 'the direction of perihelia of 448 long period comets' (Mardia [1975a]). We may extend the underlying space to a more general manifold, however, there is no practical and natural data set suitable to other manifold just like wind data corresponding to a circle. None the less, statistical theories on a sphere have not got enough study and it seems that they used to have been paid little attention by many statisticians.

2.1. Rotational symmetry

Let \mathbf{x} be a random vector which takes values on \mathcal{S}^{p-1} . It is said that \mathbf{x} has rotationally symmetric distribution about a given unit vector $\boldsymbol{\mu} \in \mathcal{S}^{p-1}$ if its density depends only on $\boldsymbol{\mu}'\mathbf{x}$. When we decompose \mathbf{x} as

$$\mathbf{x} = \boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{x} + (I_p - \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{x},$$

its density does not depend on the orthogonal complement of $\boldsymbol{\mu}$. More generally, we call it rotational symmetry that the distribution is invariant under any rotation in an s -dimensional subspace V . The Langevin distribution satisfies this condition. Another example of this class is the Scheiddegger-Watson distribution with density proportional to $\exp\{\kappa(\boldsymbol{\mu}'\mathbf{x})^2\}$. Let $p \geq 3$, $t = \boldsymbol{\mu}'\mathbf{x}$ and $d\omega_p$ denotes the area element on \mathcal{S}^{p-1} . Then

$$d\omega_p = (1 - t^2)^{\frac{1}{2}(p-3)} dt d\omega_{p-1}.$$

Thus if $\mathbf{x} \sim M_p(\boldsymbol{\mu}, \kappa)$ then its probability element is decomposed as follows:

$$\begin{aligned} \{a_p(\kappa)\}^{-1} \exp \kappa \boldsymbol{\mu}'\mathbf{x} \cdot d\omega_p &= \{a_p(\kappa)\}^{-1} e^{\kappa t} (1 - t^2)^{\frac{1}{2}(p-3)} dt d\omega_{p-1} \\ &= \omega_{p-1} \{a_p(\kappa)\}^{-1} e^{\kappa t} (1 - t^2)^{\frac{1}{2}(p-3)} dt \{\omega_{p-1}\}^{-1} d\omega_{p-1}. \end{aligned} \quad (2.2)$$

Here, $\omega_p^{-1} = \{2\pi^{p/2}/\Gamma(p/2)\}^{-1}$ is the density of a uniform distribution on \mathcal{S}^{p-1} . From (2.2) we obtain the density of t as

$$\frac{\omega_{p-1}}{a_p(\kappa)} e^{\kappa t} (1-t^2)^{\frac{1}{2}(p-3)}, \quad -1 \leq t \leq 1. \quad (2.3)$$

Integrating (2.3) about t from -1 to 1 , we get

$$\int_{-1}^1 \frac{\omega_{p-1}}{a_p(\kappa)} e^{\kappa t} (1-t^2)^{\frac{1}{2}(p-3)} dt = 1.$$

Thus

$$\frac{a_p(\kappa)}{\omega_{p-1}} = \int_{-1}^1 e^{\kappa t} (1-t^2)^{\frac{1}{2}(p-3)} dt. \quad (2.4)$$

As another decomposition of \mathbf{x} , consider

$$\mathbf{x} = \mu t + (1-t^2)^{\frac{1}{2}} \boldsymbol{\xi}, \quad \|\boldsymbol{\xi}\| = 1, \quad \boldsymbol{\xi} \perp \mu.$$

From the decomposition (2.2) of the probability element, we know that $\boldsymbol{\xi}$ has a uniform distribution on \mathcal{S}^{p-2} , t and $\boldsymbol{\xi}$ are independent and

$$E(\boldsymbol{\xi}) = \mathbf{0}, \quad D(\boldsymbol{\xi}) = E(\boldsymbol{\xi}\boldsymbol{\xi}') = (I_p - \mu\mu')/(p-1),$$

where E and D denote the expectation and the dispersion, respectively. Now let

$$A_p(\kappa) = \frac{a'_p(\kappa)}{a_p(\kappa)} = \frac{d}{d\kappa} \{\log a_p(\kappa)\}. \quad (2.5)$$

By differentiating both sides of (2.4) about κ , it is shown that $(d'/d\kappa)A_p(\kappa) = A_p^{(r)}(\kappa)$ is the $(r+1)$ th cumulant of t . Thus

$$\begin{aligned} E(t) &= A_p(\kappa), & \text{Var}(t) &= A_p'(\kappa), \\ E(t^2) &= A_p'(\kappa) + A_p(\kappa)^2. \end{aligned}$$

Further, $A_p(\kappa)$ is non-decreasing and convex on $(0, \infty)$, and takes its minimum at $\kappa = 0$ and maximum as $\kappa \rightarrow \infty$ in the range $[0, 1]$, i.e.,

$$0 \leq A_p(\kappa) \leq 1.$$

It is also known that $A_p''(\kappa) \leq 0$, so $A_p'(\kappa)$ is non-increasing and takes its minimum as $\kappa \rightarrow \infty$ and maximum at $\kappa = 0$ in the range $\left[0, \frac{1}{p}\right]$, i.e.,

$$0 \leq A'_p(\kappa) \leq \frac{1}{p}.$$

These properties of $A_p(\kappa)$ come from properties of Bessel functions. For a general theory of Bessel functions, see, *e.g.*, Watson [1980]. From the above results we obtain

$$\begin{aligned} E(\mathbf{x}) &= \boldsymbol{\mu}E(t) + E(1-t^2)^{\frac{1}{2}}E(\boldsymbol{\xi}) \\ &= \boldsymbol{\mu}E(t) \\ &= A_p(\kappa)\boldsymbol{\mu}, \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} D(\mathbf{x}) &= E(\mathbf{x}\mathbf{x}') - \{E(t)\}^2 \boldsymbol{\mu}\boldsymbol{\mu}' \\ &= \boldsymbol{\mu}\boldsymbol{\mu}'E(t^2) + E(1-t^2)E(\boldsymbol{\xi}\boldsymbol{\xi}') - \{E(t)\}^2 \boldsymbol{\mu}\boldsymbol{\mu}' \\ &= \text{Var}(t)\boldsymbol{\mu}\boldsymbol{\mu}' + \frac{E(1-t^2)}{p-1}(I_p - \boldsymbol{\mu}\boldsymbol{\mu}') \end{aligned} \tag{2.7}$$

$$= A'_p(\kappa)\boldsymbol{\mu}\boldsymbol{\mu}' + \frac{A_p(\kappa)}{\kappa}(I_p - \boldsymbol{\mu}\boldsymbol{\mu}') = \Sigma, \quad \text{say.} \tag{2.8}$$

Here we note that the equations (2.6) and (2.7) hold for any rotationally symmetric distribution about $\boldsymbol{\mu}$. Further, the equation (2.8) is obtained from the Riccati equation, which is shown in section 2.3. Moreover, from the equations (2.6) and (2.7) we have

$$\begin{aligned} \text{tr } D(\mathbf{x}) + \|E(\mathbf{x})\|^2 &= \text{Var}(t) + E(1-t^2) + E(t)^2 \\ &= 1 + \text{Var}(t) - \text{Var}(t) \\ &= 1, \end{aligned}$$

which shows the moment relationship (2.1) as we mentioned before. The moment relationship can be also proved by taking the expectations of both sides of $\mathbf{x}'\mathbf{x} = 1$.

2.2. Distribution theory

Let $\mathbf{x} \in \mathcal{S}^{p-1}$ and $\mathbf{x} = \mathbf{u}(\boldsymbol{\theta})$ denote the polar coordinate transformation of \mathbf{x} , where

$$u_j(\boldsymbol{\theta}) = \cos \theta_j \prod_{k=0}^{j-1} \sin \theta_k, \quad j = 1, \dots, p, \quad \sin \theta_0 = \cos \theta_p = 1,$$

$$0 \leq \theta_j \leq \pi, \quad j = 1, \dots, p-2, \quad 0 \leq \theta_{p-1} \leq 2\pi.$$

The subscript denotes the j th coordinate. Then the Jacobian J_p is given by

$$J_p(\boldsymbol{\theta}) = \prod_{j=2}^{p-1} \sin^{p-j} \theta_{j-1}, \quad J_2(\boldsymbol{\theta}) = 1.$$

Assume that each \mathbf{x}_i has identical and independent Langevin distribution (i.i.d.), i.e.,

$$\mathbf{x}_1, \dots, \mathbf{x}_n \sim \text{i.i.d. } M_p(\boldsymbol{\mu}, \kappa).$$

Let

$$\mathbf{x}_. = \sum \mathbf{x}_i, \quad R = \|\mathbf{x}_.\|, \quad \bar{\mathbf{x}} = \frac{1}{n} \sum \mathbf{x}_i = \frac{1}{n} \mathbf{x}_.,$$

and $\bar{\boldsymbol{\theta}}$ denote the polar coordinates of $\mathbf{x}_./\|\mathbf{x}_.\| = \mathbf{x}_./R = \bar{\mathbf{x}}/\|\bar{\mathbf{x}}\|$, i.e.,

$$\mathbf{x}_. = R \cdot \mathbf{u}(\bar{\boldsymbol{\theta}}).$$

Then it is shown that

$$\mathbf{u}(\bar{\boldsymbol{\theta}}) | R \sim M_p(\boldsymbol{\mu}, \kappa R), \quad (2.9)$$

(see e.g., Mardia *et al.* [1979, p. 473] but by their notations, $a_p(\kappa)$ and c_p are corresponding to $a_p^{-1}(\kappa)$ and ω_p^{-1} here, respectively). We note that the distribution of $\mathbf{u}(\bar{\boldsymbol{\theta}})$ obviously depends on R . The exact distribution of R was derived by Mardia [1975]. In particular, for $p = 2$ or 3 , Mardia [1972] and Fisher [1953] gave the density functions in more reduced forms. But these distributions are too complicated and it seems that further reductions have not been done yet.

When $\kappa \rightarrow 0$, it is shown that

$$\lim_{\kappa \rightarrow 0} a_p(\kappa) = \omega_p,$$

and hence \mathbf{x} has a uniform distribution on \mathcal{S}^{p-1} . Applying this to (2.9), we have $\mathbf{u}(\bar{\boldsymbol{\theta}})$ and R are independent when κ tends to zero. Conversely, if $\mathbf{u}(\bar{\boldsymbol{\theta}})$ and R are independent, then \mathbf{x}_i has a uniform distribution on \mathcal{S}^{p-1} under the assumption that \mathbf{x}_i has a density. This gives a characterization of a uniform distribution. For other characterizations, see Mardia [1975b].

2.3. Maximum likelihood estimation and the normalizing constant

The maximum likelihood estimator (MLE) $\hat{\boldsymbol{\mu}}$ of $\boldsymbol{\mu}$ is given by

$$\hat{\boldsymbol{\mu}} = \frac{\mathbf{x}_.}{\|\mathbf{x}_.\|} = \frac{\bar{\mathbf{x}}}{\|\bar{\mathbf{x}}\|}, \quad (2.10)$$

and the MLE $\hat{\kappa}$ of κ satisfies the following equation,

$$A_p(\hat{\kappa}) = \|\bar{\mathbf{x}}\|, \quad (2.11)$$

(Mardia *et al.* [1979]). Both are consistent estimators. The Langevin distribution is the only one that the MLE $\hat{\mu}$ is given by (2.10) among all rotationally symmetric distributions about μ with densities of the form $f(\mu'x)$. This shows somewhat analogue of the normal distribution, *i.e.*, it is the only one that the sample mean is always the MLE of population mean among all continuous distributions on \mathcal{R}^p . The proof is given in Watson [1983a, p. 89]. His proof needs the differentiability of $f(t)$ about t and measurability of f'/f , but Bingham and Mardia [1975] proved a more strong result by assuming only that $f(t)$ is lower semi-continuous at $t = 1$.

In order to study asymptotic behaviors of $\hat{\mu}$ and $\hat{\kappa}$, it is fundamental to obtain asymptotic approximations or refinements of $a_p(\kappa)$, $A_p(\kappa)$ and the inverse function of $A_p(\kappa)$ when $\kappa \rightarrow \infty$ or $\kappa \rightarrow 0$. First we consider $a_p(\kappa)$. From the argument about the normalizing constant of t in section 2.1, we know that it satisfies the integral equation (2.4). By differentiating both sides of (2.4), we have

$$\kappa a_p''(\kappa) + (p-1)a_p'(\kappa) - \kappa a_p(\kappa) = 0. \quad (2.12)$$

Using this, it is shown (Fujikoshi and Watamori [1992]) that for $\kappa \rightarrow \infty$,

$$\begin{aligned} a_p(\kappa) &= (2\pi)^{\frac{1}{2}(p-1)} \kappa^{-\frac{1}{2}(p-1)} e^\kappa \left\{ 1 - \frac{1}{8\kappa} (p-1)(p-3) \right. \\ &\quad \left. + \frac{1}{128\kappa^2} (p+1)(p-1)(p-3)(p-5) + O(\kappa^{-3}) \right\}. \end{aligned} \quad (2.13)$$

Thus by taking logarithm of (2.13), we get

$$\begin{aligned} \log a_p(\kappa) &= \frac{1}{2}(p-1) \log 2\pi + \kappa - \frac{1}{2}(p-1) \log \kappa \\ &\quad - \frac{1}{8}(p-1)(p-3) \left(\frac{1}{\kappa} + \frac{1}{2\kappa^2} \right) + O(\kappa^{-3}). \end{aligned} \quad (2.14)$$

For $\kappa \rightarrow 0$, since

$$\frac{d^r}{d\kappa^r} a_p(\kappa) \Big|_{\kappa=0} = \begin{cases} 0 & r: \text{ odd,} \\ \prod_{j=0}^{r-1} \frac{2j+1}{(p+2j)} \omega_p & r: \text{ even,} \end{cases}$$

$a_p(\kappa)$ is expanded in Taylor series as

$$a_p(\kappa) = \omega_p \left\{ 1 + \frac{1}{2p} \kappa^2 + \frac{1}{8p(p+2)} \kappa^4 + \frac{1}{48p(p+2)(p+4)} \kappa^6 \right\} + O(\kappa^8).$$

Next we consider $A_p(\kappa)$ defined by (2.5). Differentiating both sides of (2.5), we have

$$A_p'(\kappa) = \frac{a_p''(\kappa)}{a_p(\kappa)} - \left\{ \frac{a_p'(\kappa)}{a_p(\kappa)} \right\}^2.$$

Combining this with (2.12), it is shown that $A_p(\kappa)$ satisfies the Ricatti equation

$$A_p'(\kappa) = 1 - A_p(\kappa)^2 - \frac{p-1}{\kappa} A_p(\kappa). \quad (2.15)$$

For $\kappa \rightarrow \infty$, we obtain that

$$A_p(\kappa) = 1 - \frac{p-1}{2} \kappa^{-1} + \frac{(p-1)(p-3)}{8} \kappa^{-2} + \frac{(p-1)(p-3)}{8} \kappa^{-3} + O(\kappa^{-4}),$$

and this is coincident with the result obtained by differentiating both sides of (2.14). For $\kappa \rightarrow 0$, we obtain that

$$A_p(\kappa) = \frac{1}{p} \kappa + \frac{1}{p^2(p+2)} \kappa^3 + \frac{1}{p^3(p+2)(p+4)} \kappa^5 + O(\kappa^7).$$

In studying properties of $\hat{\kappa}$, we need asymptotic approximations of the inverse function of $A_p(\kappa)$, $\kappa = A_p^{-1}(y)$ for all y in $[0, 1]$. Using the equation (2.15), we can find the asymptotic expansion of κ for small y ,

$$\kappa = py + \frac{p^2}{(p+2)} y^3 + \frac{p^3(p+8)}{(p+2)(p+4)} y^5 + O(y^7),$$

and when y is almost unity,

$$\frac{1}{\kappa} = \frac{2}{p-1} (1-y) + \frac{p-3}{p-1} (1-y)^2 + \frac{p-3}{(p-1)^2} (1-y)^3 + O((1-y)^4).$$

For more details about $A_p(\kappa)$, see Watson [1983a, Appendix A.2. pp. 190–195].

Using the above limiting results we may approximate the MLE $\hat{\kappa}$ in (2.11) for the cases of $\kappa \rightarrow \infty$ and $\kappa \rightarrow 0$ as

$$\hat{\kappa} \doteq \begin{cases} \frac{p-1}{2(1-\|\bar{\mathbf{x}}\|)} = \frac{n(p-1)}{2(n-\|\mathbf{x}\|)} & \text{as } \kappa \rightarrow \infty, \\ p\|\bar{\mathbf{x}}\| = \frac{p}{n}\|\mathbf{x}\| & \text{as } \kappa \rightarrow 0. \end{cases}$$

Let L and η denote the likelihood function and the polar coordinates of μ , respectively. Since

$$\begin{aligned}\log L &= -n \log a_p(\kappa) + \kappa \mu' \mathbf{x} \\ &= \text{const.} - n \log a_p(\kappa) + R \kappa \mathbf{u}'(\eta) \mathbf{u}(\bar{\theta}),\end{aligned}$$

we have

$$\begin{aligned}E \left[-\frac{\partial^2 \log L}{\partial \kappa^2} \right] &= n A'_p(\kappa), \\ E \left[-\frac{\partial^2 \log L}{\partial \kappa \partial \eta_i} \right] &= 0, & i = 1, \dots, p-1, \\ E \left[-\frac{\partial^2 \log L}{\partial \eta_i \partial \eta_j} \right] &= 0, & i \neq j, \\ E \left[-\frac{\partial^2 \log L}{(\partial \eta_i)^2} \right] &= n \kappa A_p(\kappa) \prod_{j=0}^{i-1} \sin^2 \eta_j & j = 1, \dots, p-1.\end{aligned}$$

Hence, for large n , we can conclude that $\hat{\kappa}, \bar{\theta}_1, \dots, \bar{\theta}_{p-1}$ are asymptotically independent and normally distributed with means $\kappa, \eta_1, \dots, \eta_{p-1}$ and

$$\begin{aligned}\text{var}(\hat{\kappa}) &= [n A'_p(\kappa)]^{-1}, \\ \text{var}(\bar{\theta}_i) &= \frac{1}{n} \{ \kappa A_p(\kappa) \prod_{j=0}^{i-1} \sin^2 \eta_j \}^{-1}, & i = 1, \dots, p-1,\end{aligned}$$

(Mardia *et al.* [1979, p. 439]).

2.4. Large sample theory ($n \rightarrow \infty$)

When $\mathbf{x} \sim M_p(\mu, \kappa)$, the dispersion matrix of \mathbf{x} is given by

$$\Sigma = A'_p(\kappa) \mu \mu' + \frac{A_p(\kappa)}{\kappa} (I_p - \mu \mu'),$$

see, (2.8). Then the determinant of Σ , $\det \Sigma$ is

$$\det \Sigma = A'_p(\kappa) \cdot \left(\frac{A_p(\kappa)}{\kappa} \right)^{(p-1)}.$$

From the discussion in section 2.1 we know that both $A_p(\kappa)$ and $A'_p(\kappa)$ are positive for non-zero κ , so $\det \Sigma > 0$. Applying the central limit theorem, for $n \rightarrow \infty$,

$$\begin{aligned}\sqrt{n}(\bar{\mathbf{x}} - A_p(\kappa)\mu) &\rightarrow N_p(\mathbf{0}, \Sigma), \\ \therefore \sqrt{n}\Sigma^{-\frac{1}{2}}(\bar{\mathbf{x}} - A_p(\kappa)\mu) &\rightarrow N_p, & (2.16)\end{aligned}$$

where N_p denotes the p -dimensional normal distribution with mean $\mathbf{0}$ and unit dispersion matrix I_p , i.e., $N_p(\mathbf{0}, I_p)$. Multiplying $\boldsymbol{\mu}'$ from the left and noting that $\boldsymbol{\mu}'\Sigma^{-1/2} = (1/\sqrt{A'_p(\kappa)})\boldsymbol{\mu}'$, we have

$$\frac{\sqrt{n}}{\sqrt{A'_p(\kappa)}}(\boldsymbol{\mu}'\bar{\mathbf{x}} - A_p(\kappa)) \rightarrow N_1.$$

Let us choose a $p \times (p-1)$ matrix B_1 that satisfies the following conditions

$$B_1 B_1' = I_p - \boldsymbol{\mu}\boldsymbol{\mu}', \quad B_1' B_1 = I_{p-1},$$

i.e., a $p \times p$ matrix $(\boldsymbol{\mu} B_1)$ is an element of $\mathbf{O}(p)$, the orthogonal group of order p . Since $B_1' \boldsymbol{\mu} = 0$ and $B_1' \Sigma^{-1/2} = \sqrt{\kappa/A_p(\kappa)} B_1'$, we obtain

$$\begin{aligned} \sqrt{n} \sqrt{\frac{\kappa}{A_p(\kappa)}} B_1' \bar{\mathbf{x}} &\rightarrow N_{p-1}, \\ \therefore n \frac{\kappa}{A_p(\kappa)} \|B_1' \bar{\mathbf{x}}\|^2 &\rightarrow \chi_{p-1}^2, \end{aligned}$$

where χ_p^2 denotes a central chi-square distribution with p degrees of freedom. For the unique solution $\hat{\kappa}$ of (2.11), it holds that

$$\begin{aligned} \sqrt{n} A'_p(\hat{\kappa})(\hat{\kappa} - \kappa) &\rightarrow N_1(0, A'_p(\kappa)), \\ \therefore \sqrt{n A'_p(\hat{\kappa})}(\hat{\kappa} - \kappa) &\rightarrow N_1. \end{aligned}$$

It is also shown that

$$2n A_p(\hat{\kappa}) \kappa (1 - \boldsymbol{\mu}' \hat{\boldsymbol{\mu}}) \rightarrow \chi_{p-1}^2.$$

We note that the above limiting results can be elaborated by deriving their asymptotic expansions. Some of them are given in the rest of this paper. It is fundamental to obtain a refinement of the result (2.16), which is discussed in section 4.1.

Next we consider the testing problem that $\boldsymbol{\mu}$ is in a given subspace V . Without loss of generality, we may express V as

$$V = \{\boldsymbol{\mu} \mid \boldsymbol{\mu} = B_0 \boldsymbol{\zeta}, \boldsymbol{\zeta}' \boldsymbol{\zeta} = 1\}, \quad (2.17)$$

where B_0 is a given $p \times s$ matrix such that $B_0' B_0 = I_s$. In a special case $B_0 = \boldsymbol{\mu}_0$, the hypothesis becomes $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ or $\boldsymbol{\mu} = -\boldsymbol{\mu}_0$, and so slightly different from the hypothesis $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, but we may think that these two hypotheses are essentially the same. In fact the results of Watamori [1992] show Watson statistics and the likelihood ratio criterion for the two hypotheses have the same asymptotic expansions. In order to compare powers of these tests, we

consider a sequence of the alternatives,

$$\begin{aligned}\boldsymbol{\mu} &= (\boldsymbol{\mu}_0 + n^{-\frac{1}{2}}\boldsymbol{\delta}) \|\boldsymbol{\mu}_0 + n^{-\frac{1}{2}}\boldsymbol{\delta}\|^{-1} \\ &= (\boldsymbol{\mu}_0 + n^{-\frac{1}{2}}\boldsymbol{\delta})(1 + 2n^{-1}\lambda)^{-\frac{1}{2}},\end{aligned}\quad (2.18)$$

where $\boldsymbol{\mu}_0 = B_0\boldsymbol{\zeta}$, $B_0'\boldsymbol{\delta} = 0$ and $\lambda = \boldsymbol{\delta}'\boldsymbol{\delta}/2$. Let B_s be a $p \times (p-s)$ matrix such that $(B_0B_s) \in \mathbf{O}(p)$. Then

$$\frac{n\kappa}{A_p(\kappa)} \|B_s'\bar{\mathbf{x}}\| \rightarrow \chi_{p-s}^2(A_p(\kappa)\kappa\lambda), \quad (2.19)$$

where $\chi_p^2(\gamma)$ denotes a noncentral chi-square distribution with p degrees of freedom and noncentrality parameter γ .

Most of the results in this section are found in Watson [1983a, §4.2, §4.4, 1983b] and Watamori [1992] gives an extended result of (2.19).

2.5. Approximations in highly concentrated case ($\kappa \rightarrow \infty$)

In this section we give an outline of asymptotic approximations for basic distributions in a highly concentrated case ($\kappa \rightarrow \infty$). The limiting results are found in Watson [1983a, §4.5, 1984]. Some refinements of these and other distributions are obtained in subsequent sections.

Let $U = 2\kappa(1-t)$, where $t = \boldsymbol{\mu}'\mathbf{x}$. From the expansion (2.13) of $a_p(\kappa)$ we can approximate $a_p(\kappa)$ for large κ as

$$a_p(\kappa) \doteq (2\pi)^{\frac{1}{2}(p-1)} \kappa^{-\frac{1}{2}(p-1)} e^\kappa.$$

Substituting this approximation to (2.3), we obtain the density of U as

$$f(u) \doteq \frac{1}{2^{\frac{p-1}{2}} \Gamma(\frac{p-1}{2})} e^{-\frac{u}{2}} u^{\frac{p-1}{2}-1}.$$

Since this is the density function of χ_{p-1}^2 ,

$$U \rightarrow \chi_{p-1}^2.$$

Here we use the fact that

$$t = \boldsymbol{\mu}'\mathbf{x} \rightarrow 1 \quad (\text{in prob.}).$$

If we choose $p \times (p-1)$ matrix B_1 as in the previous section, then the projection of \mathbf{x} onto the orthogonal complement of the space spanned by $\boldsymbol{\mu}_0$ is given by $B_1B_1'\mathbf{x} = (1-t^2)^{\frac{1}{2}}\boldsymbol{\xi}$. So,

$$\sqrt{\kappa}B_1B_1'\mathbf{x} = \{\kappa(1+t)(1-t)\}^{\frac{1}{2}}\boldsymbol{\xi} \rightarrow U^{\frac{1}{2}}\boldsymbol{\xi},$$

where U and ξ are independent. By a characterization of the normal distribution, it follows that

$$\sqrt{\kappa} B_1 B_1' \mathbf{x} \rightarrow N_p(\mathbf{0}, B_1 B_1').$$

Multiplying B_1' from the left and noting $B_1' B_1 = I_{p-1}$, we obtain

$$\sqrt{\kappa} B_1' \mathbf{x} \rightarrow N_{p-1}.$$

Applying this to $\mu' \mathbf{x} = \mu' \sum \mathbf{x}_i = \sum \mu' \mathbf{x}_i$ and $B_1' \mathbf{x} = B_1' \sum \mathbf{x}_i = \sum B_1' \mathbf{x}_i$, it follows that

$$\begin{aligned} 2\kappa(n - \mu' \mathbf{x}) &\rightarrow \chi_{n(p-1)}^2, \\ \sqrt{\kappa} B_1' \mathbf{x} &\rightarrow N_{p-1}, \\ \frac{\kappa}{n} \|B_1' \mathbf{x}\|^2 &\rightarrow \chi_{p-1}^2. \end{aligned}$$

Next we consider to test the hypothesis that μ is in a given subspace V when κ is given, where V is given by (2.17). As a sequence of the alternatives, we consider now the one obtained from (2.18) by replacing n with κ , *i.e.*,

$$\mu = (\mu_0 + \kappa^{-\frac{1}{2}} \delta)(1 + 2\kappa^{-1} \lambda)^{-\frac{1}{2}},$$

where μ_0 , δ and λ are the same as in (2.18). Then

$$\begin{aligned} \sqrt{\kappa} B_s B_s' \mathbf{x} &\rightarrow N_p(\delta, B_s B_s'), \\ \sqrt{\kappa} B_s' \mathbf{x} &\rightarrow N_{p-s}(\delta, I_{p-s}), \\ \frac{\kappa}{n} \|B_s' \mathbf{x}\|^2 &\rightarrow \chi_{p-s}^2(n\lambda), \end{aligned} \tag{2.20}$$

where B_s is the same as in (2.19). It is easily seen that $\|\mathbf{x}\|^2 = \|B_0' \mathbf{x}\|^2 + \|B_s' \mathbf{x}\|^2$ and $\|\mathbf{x}\|$, $\|B_0' \mathbf{x}\| \rightarrow n$ as $\kappa \rightarrow \infty$. Therefore (2.20) is rewritten as

$$2\kappa(\|\mathbf{x}\| - \|B_0' \mathbf{x}\|) \rightarrow \chi_{p-s}^2(n\lambda).$$

Further, when κ is unknown, it is shown that under the hypothesis $\mu = \mu_0$,

$$\frac{(\|\mathbf{x}\| - \mu' \mathbf{x}) / (p-1)}{(n - \|\mathbf{x}\|) / (n-1)(p-1)} \rightarrow F_{p-1, (p-1)(n-1)}.$$

2.6. Explanatory notes

In the backgrounds mentioned above we collect mainly the results related to the subsequent sections, and we restrict topics on the Langevin distribution. In particular, sections 2.4 and 2.5 are the bases of our main

results. Some of the characterization problems are given in sections 2.2 and 2.3. For other ones, Mardia [1975c] has proposed the characterizations by maximum entropy and others in 2 or 3 dimension.

Though some higher dimensional applications have been considered (see, *e.g.*, Stephens [1982]), the main applications have been restricted to 2 or 3 dimension in practice. This may be probably due to that \mathcal{S}^1 and \mathcal{S}^2 are so familiar with us and hence we cannot be free from them on the contrary. On the other hand, most of the results for directional distributions can be extended to higher dimensions. From this point of view, we hope that useful higher dimensional applications will be appeared in the near future.

3. Parameter estimation

3.1. Primaries

We consider the problem of estimating μ and κ , based on a random sample x_1, \dots, x_n of size n from $M_p(\mu, \kappa)$. Let

$$\bar{x} = \frac{1}{n} \sum x_j, \quad \text{and} \quad R = n \|\bar{x}\|.$$

Here R is the resultant length of sample vectors.

The MLE $\hat{\kappa}$ and the marginal MLE $\tilde{\kappa}$ proposed by Schou [1978] can be given as the solutions of the following equations:

$$A_p(\hat{\kappa}) = \|\bar{x}\|,$$

and

$$\begin{aligned} \tilde{\kappa} &= 0, & \text{if } R \leq n^{\frac{1}{2}}, \\ nA_p(\tilde{\kappa}) &= RA_p(\tilde{\kappa}R), & \text{if } R > n^{\frac{1}{2}}, \end{aligned}$$

respectively. Schou [1978] has shown that $\Pr(\tilde{\kappa} = 0) \rightarrow 0$, $\sqrt{n}(\hat{\kappa} - \kappa)$ and $\sqrt{n}(\tilde{\kappa} - \kappa)$ are asymptotically $N(0, 1/A')$ as $n \rightarrow \infty$.

The MLE of μ is given as follows:

$$\hat{\mu} = \frac{1}{\|\bar{x}\|} \bar{x} = \frac{n}{R} \bar{x}.$$

For simplicity we denote $A_p(\kappa)$ and its derivatives about κ as $A, A', A'', A^{(3)}, \dots$, respectively.

3.2. Large sample case

Watamori [1992] derived the asymptotic expansions of distributions of

some test statistics for testing mean directions. The method is based on expanding test statistics in the terms of a normalized statistic $\mathbf{y} = \sqrt{n} \Sigma^{-1/2}(\bar{\mathbf{x}} - A\boldsymbol{\mu})$ and evaluating their characteristic functions. We will discuss the details under a general setting in sections 4 and 5. The same method can be applied for obtaining large sample approximations of $\hat{\kappa}$ and $\tilde{\kappa}$.

The estimators $\hat{\kappa}$ and $\tilde{\kappa}$ are expanded in terms of \mathbf{y} as

$$\begin{aligned} \sqrt{n}(\hat{\kappa} - \kappa) &= \frac{1}{\sqrt{A'}} \boldsymbol{\mu}' \mathbf{y} + \frac{1}{\sqrt{n}} \left\{ \frac{1}{2A'\kappa} \mathbf{y}'(I_p - \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{y} - \frac{A''}{2A'^2} (\boldsymbol{\mu}' \mathbf{y})^2 \right\} \\ &\quad + \frac{1}{n} \left\{ \left(-\frac{1}{2A'\sqrt{A'\kappa}} + \frac{A''}{2A'^2\sqrt{A'\kappa}} \right) \boldsymbol{\mu}' \mathbf{y} \cdot \mathbf{y}'(I_p - \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{y} \right. \\ &\quad \left. + \left(\frac{A''^2}{2A'^3\sqrt{A'}} - \frac{A^{(3)}}{6A'^2\sqrt{A'}} \right) (\boldsymbol{\mu}' \mathbf{y})^3 \right\} + O_p(n^{-\frac{3}{2}}), \\ \sqrt{n}(\tilde{\kappa} - \kappa) &= \frac{1}{\sqrt{A'}} \boldsymbol{\mu}' \mathbf{y} + \frac{1}{\sqrt{n}} \left\{ \frac{1}{2A'\kappa} \mathbf{y}'(I_p - \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{y} - \frac{A''}{2A'^2} (\boldsymbol{\mu}' \mathbf{y})^2 - \frac{p-1}{2A'\kappa} \right\} \\ &\quad + \frac{1}{n} \left\{ \left(-\frac{1}{2A'\sqrt{A'\kappa}} + \frac{A''}{2A'^2\sqrt{A'\kappa}} \right) \boldsymbol{\mu}' \mathbf{y} \cdot \mathbf{y}'(I_p - \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{y} \right. \\ &\quad \left. + \left(\frac{A''^2}{2A'^3\sqrt{A'}} - \frac{A^{(3)}}{6A'^2\sqrt{A'}} \right) (\boldsymbol{\mu}' \mathbf{y})^3 \right. \\ &\quad \left. + \left(\frac{p-1}{2A'\sqrt{A'\kappa^2}} + \frac{(p-1)A''}{2A'^2\sqrt{A'\kappa}} \right) \right\} + O_p(n^{-\frac{3}{2}}). \end{aligned}$$

The characteristic functions of $\sqrt{n}(\hat{\kappa} - \kappa)$ and $\sqrt{n}(\tilde{\kappa} - \kappa)$ can be evaluated in expanded forms, based on asymptotic expansions of the distribution of \mathbf{y} , which is given in section 4.1. Inverting the resultant characteristic function, we have the following theorem.

THEOREM 3.1 *Let $f_{\hat{\kappa}}(x_1)$ and $f_{\tilde{\kappa}}(x_2)$ be the density functions of $x_1 = \sqrt{n}(\hat{\kappa} - \kappa)$ and $x_2 = \sqrt{n}(\tilde{\kappa} - \kappa)$, respectively. Then $f_{\hat{\kappa}}(x_1)$ and $f_{\tilde{\kappa}}(x_2)$ can be expanded for large n as follows:*

$$\begin{aligned} f_{\hat{\kappa}}(x_1) &= \sqrt{A'} \phi(\sqrt{A'}x_1) \left[1 + \frac{1}{\sqrt{n}} \left\{ \left(\frac{p-1}{2\kappa} + \frac{A''}{2A'} \right) x_1 - \frac{A''}{3} x_1^3 \right\} \right. \\ &\quad \left. + \frac{1}{n} \left\{ \left(-\frac{(p-1)(p-3)}{8A'\kappa^2} + \frac{(p-1)A''}{4A'^2\kappa} - \frac{5A''^2}{24A'^3} + \frac{A^{(3)}}{8A'^2} \right) \right. \right. \\ &\quad \left. \left. + \left(\frac{(p-1)(p-3)}{8\kappa^2} + \frac{(p-1)A''}{4A'\kappa} - \frac{A''^2}{8A'^2} + \frac{A^{(3)}}{4A'} \right) x_1^2 \right. \right. \end{aligned}$$

$$\begin{aligned}
& - \left(\frac{(p-1)A''}{6\kappa} + \frac{A''^2}{6A'} + \frac{A^{(3)}}{8} \right) x_1^4 + \frac{A''^2}{18} x_1^6 \Big\} + O(n^{-\frac{3}{2}}) \Big], \\
f_{\tilde{\kappa}}(x_2) = & \sqrt{A'} \phi(\sqrt{A'} x_2) \left[1 + \frac{1}{\sqrt{n}} \left(\frac{A''}{2A'} x_2 - \frac{A''}{3} x_2^3 \right) \right. \\
& + \frac{1}{n} \left\{ \left(-\frac{(p-1)}{4A'\kappa^2} - \frac{5A''^2}{24A'^3} + \frac{A^{(3)}}{8A'^2} \right) \right. \\
& + \left(\frac{p-1}{4\kappa^2} - \frac{A''^2}{8A'^2} + \frac{A^{(3)}}{4A'} \right) x_2^2 \\
& \left. \left. - \left(\frac{A''^2}{6A'} + \frac{A^{(3)}}{8} \right) x_2^4 + \frac{A''^2}{18} x_2^6 \right\} + O(n^{-\frac{3}{2}}) \right],
\end{aligned}$$

where $\phi(x) = (1/\sqrt{2\pi}) \exp(-x^2/2)$.

From Theorem 3.1 we can obtain expectations, MSEs and concentration probabilities of $\hat{\kappa}$ and $\tilde{\kappa}$. The expectations of $\hat{\kappa}$ and $\tilde{\kappa}$ are expanded as

$$\begin{aligned}
E[\hat{\kappa}] &= \kappa + \frac{1}{n} \left(\frac{p-1}{2A'\kappa} - \frac{A''}{2A'^2} \right) + O(n^{-2}), \\
E[\tilde{\kappa}] &= \kappa - \frac{1}{n} \frac{A''}{2A'^2} + O(n^{-2}).
\end{aligned}$$

The expansion of $E[\hat{\kappa}]$ is coincident with Schou's result. Note that the term $(p-1)/2A'\kappa$ in the coefficient of $1/n$ is non negative and goes infinity as κ goes infinity, and $-A''/2A'^2$ is non-negative. This implies that the biases of them are quite large when κ is large. It is also shown that they have positive biases and the $1/n$ term of $E[\tilde{\kappa}]$ is smaller than that of $E[\hat{\kappa}]$. The variances of them are given as follows:

$$V[\hat{\kappa}] = \frac{1}{nA'} + O(n^{-2}) = V[\tilde{\kappa}].$$

The MSEs of them are expanded as

$$\begin{aligned}
E[(\hat{\kappa} - \kappa)^2] &= \frac{1}{nA'} + \frac{1}{n^2} \left\{ \frac{(p-1)(p-3)}{4A'^2\kappa^2} - \frac{3A''(p-1)}{2A'^3\kappa} + \frac{11A''^2}{4A'^4} + \frac{A^{(3)}}{4A'^3} \right\} \\
& \quad + O(n^{-3}), \\
E[(\tilde{\kappa} - \kappa)^2] &= \frac{1}{nA'} + \frac{1}{n^2} \left\{ \frac{p-1}{2A'^2\kappa^2} + \frac{11A''^2}{4A'^4} + \frac{A^{(3)}}{4A'^3} \right\} + O(n^{-3}).
\end{aligned}$$

Therefore it holds that

$$E[(\hat{\kappa} - \kappa)^2] - E[(\tilde{\kappa} - \kappa)^2] = \frac{1}{n^2} \left\{ \frac{(p-1)(p-5)}{2A'^2\kappa^2} - \frac{3A''(p-1)}{2A'^3\kappa} \right\} + O(n^{-3}).$$

The concentration probabilities of them are expanded for $b > 0$ as follows:

$$\begin{aligned} \Pr(-b < \sqrt{n}(\hat{\kappa} - \kappa) < b) &= \Phi(\sqrt{A'}b) - \Phi(-\sqrt{A'}b) \\ &+ \frac{1}{n} \left[- \left\{ \frac{(p-1)(p-3)}{4\sqrt{A'}\kappa^2} - \frac{A''(p-1)}{2A'\sqrt{A'}\kappa} + \frac{5A''^2}{12A'^2\sqrt{A'}} \right. \right. \\ &- \left. \left. \frac{A^{(3)}}{4A'\sqrt{A'}} \right\} b + \left\{ \frac{A''(p-1)}{3\sqrt{A'}\kappa} - \frac{2A''^2}{9A'\sqrt{A'}} + \frac{A^{(3)}}{4\sqrt{A'}} \right\} b^3 \right. \\ &- \left. \frac{A''^2}{9\sqrt{A'}} b^5 \right] \phi(\sqrt{A'}b) + O(n^{-\frac{3}{2}}), \end{aligned}$$

$$\begin{aligned} \Pr(-b < \sqrt{n}(\tilde{\kappa} - \kappa) < b) &= \Phi(\sqrt{A'}b) - \Phi(-\sqrt{A'}b) \\ &+ \frac{1}{n} \left[- \left\{ \frac{(p-1)}{2\sqrt{A'}\kappa^2} + \frac{5A''^2}{12A'^2\sqrt{A'}} - \frac{A^{(3)}}{4A'\sqrt{A'}} \right\} b \right. \\ &- \left. \left\{ \frac{2A''^2}{9A'\sqrt{A'}} - \frac{A^{(3)}}{4\sqrt{A'}} \right\} b^3 - \frac{A''^2}{9\sqrt{A'}} b^5 \right] \phi(\sqrt{A'}b) + O(n^{-\frac{3}{2}}), \end{aligned}$$

where $\Phi(x)$ and $\phi(x)$ are the distribution and the density functions of the standard normal distribution, respectively. These imply that

$$\begin{aligned} \Pr(-b < \sqrt{n}(\tilde{\kappa} - \kappa) < b) - \Pr(-b < \sqrt{n}(\hat{\kappa} - \kappa) < b) \\ &= \frac{1}{n} \left[\left\{ \frac{(p-1)(p-5)}{4\sqrt{A'}\kappa^2} - \frac{A''(p-1)}{2A'\sqrt{A'}\kappa} \right\} b \right. \\ &- \left. \frac{A''(p-1)}{3\sqrt{A'}\kappa} b^3 \right] \phi(\sqrt{A'}b) + O(n^{-\frac{3}{2}}). \end{aligned}$$

Note that each coefficient of the differences of the MSEs and concentration probabilities is non-negative when $p > 4$. For the cases $p = 2$ and $p = 3$, all the coefficients of the differences are not always positive and we cannot conclude which is smaller at a glance. The values of the coefficients of $1/n^2$ (for the MSEs) and $1/n$ (for concentration probabilities) in some special cases are given in Tables 1 and 2. The tables show that for small κ the MSE of $\hat{\kappa}$ is smaller but for other cases $\tilde{\kappa}$ is preferable and the differences between them become larger when κ goes larger.

From the results on the expectations of $\hat{\kappa}$ and $\tilde{\kappa}$ it is possible to modify the estimators so that they satisfy higher order asymptotic unbiasedness.

Table 1: The coefficients of $1/n^2$ of MSEs

κ	$p = 2$			$p = 3$		
	$\hat{\kappa}$	$\tilde{\kappa}$	$\hat{\kappa}_g$	$\hat{\kappa}$	$\tilde{\kappa}$	$\hat{\kappa}_g$
0.1	-96.910	200.820	201.516	9.973	902.748	903.612
0.5	1.434	10.678	9.919	11.792	39.944	39.916
1	14.288	12.699	5.578	18.184	17.244	13.971
5	1387.096	1194.199	221.954	451.227	327.846	73.861
10	5781.160	4943.107	940.131	1849.974	1349.978	299.996

Table 2: The coefficients of $1/n$ of concentration probabilities
coefficients of b

κ	$P = 2$			$p = 3$		
	$\hat{\kappa}$	$\tilde{\kappa}$	difference	$\hat{\kappa}$	$\tilde{\kappa}$	difference
1	-0.405	-1.133	0.728	-0.853	-2.077	1.224
5	-0.274	-0.178	-0.096	-0.436	-0.243	-0.193
10	-0.131	-0.092	-0.039	-0.217	-0.117	-0.100

coefficients of b^3

1	-0.183	-0.0575	-0.1256	-0.160	-0.0348	-0.125
5	0.00028	0.00481	-0.00453	-0.00617	0.00428	-0.0104
10	-0.0000213	0.000481	-0.000502	-0.000722	0.000611	-0.00133

coefficients of b^5

1	-0.00939	-0.00205
5	-0.0000780	-0.000136
10	-0.00000183	-0.00000444

Such estimators are given by

$$\hat{\kappa}_g = \hat{\kappa} - \frac{1}{n} \left(\frac{p-1}{2A'_p(\hat{\kappa})\hat{\kappa}} - \frac{A''_p(\hat{\kappa})}{2A'^2_p(\hat{\kappa})} \right),$$

$$\tilde{\kappa}_g = \tilde{\kappa} + \frac{1}{n} \frac{A''_p(\tilde{\kappa})}{2A'^2_p(\tilde{\kappa})}.$$

Then it is shown that the distributions of $\sqrt{n}(\hat{\kappa}_g - \kappa)$ and $\sqrt{n}(\tilde{\kappa}_g - \kappa)$ are

coincident up to $O(n^{-1})$. Further

$$\begin{aligned} E[(\hat{\kappa}_\varphi - \kappa)^2] &= \frac{1}{nA'} + \frac{1}{n^2} \left\{ \frac{(p-1)}{2A'^2\kappa^2} + \frac{A''^2}{2A'^4} \right\} + O(n^{-3}) \\ &= E[(\tilde{\kappa}_\varphi - \kappa)^2] + O(n^{-3}). \end{aligned}$$

Therefore $\hat{\kappa}_\varphi$ and $\tilde{\kappa}_\varphi$ are concluded to be quite similar. The values of the MSEs up to the order n^{-2} in some special cases are also given in the Table 1. It is noted that each value for moderate κ is smaller than the corresponding value of $\hat{\kappa}$ or $\tilde{\kappa}$. Note also that for the almost of all cases the MSEs of $\tilde{\kappa}$ are middle and similar to those of either $\hat{\kappa}$ or $\hat{\kappa}_\varphi$. When κ is quite small $\hat{\kappa}$ is most preferable but for the usual values of κ , $\hat{\kappa}_\varphi$ is best of these three estimators in the sense that the MSE is the smallest. It seems that it is better to use $\hat{\kappa}_\varphi$ instead of $\hat{\kappa}$ when κ is not so small.

We can express $\hat{\mu}$ as $\hat{\mu} = (A\mu + \Sigma^{1/2}\mathbf{y}/\sqrt{n}) / \|A\mu + \Sigma^{1/2}\mathbf{y}/\sqrt{n}\|$. Therefore the asymptotic expansions of the expectation and the MSE of $\hat{\mu}$ are given as follows:

$$\begin{aligned} E[\hat{\mu}] &= \mu - \frac{1}{n} \frac{p-1}{2A\kappa} \mu + \frac{1}{n^2} \left\{ \frac{(p-1)(15p-1)}{8A\kappa} \right. \\ &\quad \left. - \frac{A'(p-1)(p-4)}{A^2} + \frac{A''(p^2-1)}{3AA'} \right\} \mu + O(n^{-\frac{5}{2}}) \end{aligned}$$

and

$$E[(\hat{\mu} - \mu)(\hat{\mu} - \mu)'] = \frac{1}{n} \frac{1}{A\kappa} (I_p - \mu\mu') + O(n^{-2}).$$

Note that the term of the coefficient of $1/n$, $(p-1)/(2A\kappa)$ is quite large if κ is considered to be very small. It may be suggested to use

$$\hat{\mu}_\varphi = \left(1 + \frac{1}{n} \cdot \frac{p-1}{2A\hat{\kappa}} \right) \hat{\mu}$$

as a higher order asymptotic unbiased estimator.

3.3. Highly concentrated case

Next we consider the case when κ is large. In addition to the notations in section 3.2, we use the following notations:

$$Z = \sqrt{\kappa} B'_1 [x_1, \dots, x_n],$$

$\mathbf{1}_n$: an n -dimensional vector with all elements unity,

$$P_0 = \frac{1}{n} \mathbf{1}_n \mathbf{1}'_n \quad P_1 = I_n - P_0,$$

\mathbf{e}_j : an n -dimensional unit vector with j th element unity and others 0,

$$E_j = \mathbf{e}_j \mathbf{e}'_j.$$

It may be noted that Fujikoshi and Watamori [1992] have essentially obtained an asymptotic expansion of the distribution of Z . We can obtain stochastic expansions of $\hat{\kappa}$ and $\tilde{\kappa}$ in terms of Z . The stochastic expansions of $\hat{\kappa}$ and $\tilde{\kappa}$ are obtained as follows:

$$\begin{aligned} \frac{\hat{\kappa}}{\kappa} &= \frac{n(p-1)}{\text{tr } ZP_1Z'} + \frac{1}{\kappa} \left[\frac{p-1}{4(\text{tr } ZP_1Z')^2} \{(\text{tr } ZP_0Z')^2 + 2 \text{tr } ZP_0Z' \text{tr } ZP_1Z'\} \right. \\ &\quad \left. - n \sum_j (\text{tr } ZE_jZ')^2 \right] - \frac{p-3}{4} + O_p(\kappa^{-2}), \\ \frac{\tilde{\kappa}}{\kappa} &= \frac{(n-1)(p-1)}{\text{tr } ZP_1Z'} + \frac{1}{\kappa} \left[\frac{(n-1)(p-1)}{4(\text{tr } ZP_1Z')^2} \{(\text{tr } ZP_0Z')^2 + 2 \text{tr } ZP_0Z' \text{tr } ZP_1Z'\} \right. \\ &\quad \left. - n \sum_j (\text{tr } ZE_jZ')^2 \right] - \frac{p-3}{4} + O_p(\kappa^{-2}). \end{aligned}$$

Since the elements of Z are independently distributed as $N(0, 1)$ in the limiting case, we can see that $\text{tr } ZP_1Z'$ and $\text{tr } ZP_0Z'$ are asymptotically independent and

$$\begin{aligned} \text{tr } ZP_1Z' &\rightsquigarrow \chi_{(p-1)(n-1)}^2, \\ \text{tr } ZP_0Z' &\rightsquigarrow \chi_{(p-1)}^2. \end{aligned}$$

Thus the limiting distributions of $\hat{\kappa}$ and $\tilde{\kappa}$ are proportional to $1/\chi_{(p-1)(n-1)}^2$, which is coincident with Schou [1978] who also pointed out $\Pr(\tilde{\kappa} = 0) \rightarrow 0$. It is noted that neither $\hat{\kappa}$ nor $\tilde{\kappa}$ are asymptotically unbiased for large κ . These estimators may be modified as

$$\left(1 - \frac{2}{n(p-1)}\right) \hat{\kappa} \quad \text{and} \quad \left(1 - \frac{2}{(n-1)(p-1)}\right) \tilde{\kappa},$$

which are asymptotically unbiased for large κ . The above stochastic expansions will be useful in obtaining further refinements as in the large sample case. However, the investigation is left as a future work.

4. Multi-sample tests for mean directions (I)-large sample case

4.1. Primaries

As far as there seems not to be any confusion, we write $A_p(\kappa)$ as A and $A_p(\kappa_j)$ as A_j for simplicity. We consider the problems of testing

$$\begin{aligned} H_1: (\mu_1, \mu_2, \dots, \mu_k)C = 0 \quad \text{vs.} \quad K_1: (\mu_1, \mu_2, \dots, \mu_k)C \neq 0, \\ H_2: \mu_j \in V \quad \text{vs.} \quad K_2: \mu_j \notin V \text{ for some } j, \end{aligned}$$

where C is a given $k \times (k-1)$ matrix with rank $k-1$ and V is an s -dimensional subspace ($s \leq p-1$) based on k -random samples \mathbf{x}_{jl} of size n_j from $M_p(\mu_j, \kappa_j)$, $l = 1, \dots, n_j$, $j = 1, \dots, k$. Since all norms of μ_j 's are unit, we may take

$$C = \begin{pmatrix} I_{k-1} \\ c_1 & c_2 & \dots & c_{k-1} \end{pmatrix}, \quad c_j^2 = 1, \text{ for } \forall j,$$

without loss of generality. As in (2.17), we write V as $V = \{\mu \mid \mu = B_0 \zeta, \zeta' \zeta = 1\}$. So we can rewrite the hypotheses H_1 and H_2 as

$$\begin{aligned} H_1: \mu_j = b_j \mu, \quad b_j = c_1 c_j, \quad c_k = -1, \\ H_2: \mu_j = B_0 \zeta_j, \quad \|\zeta_j\| = 1. \end{aligned}$$

Consider a sequence of alternatives

$$\mu_j = (\mu_{0j} + n_j^{-\frac{1}{2}} \delta_j)(1 + 2n_j^{-1} \lambda_j)^{-\frac{1}{2}}, \quad (4.1)$$

where μ_{0j} is equal to $b_j \mu$ for H_1 or $B_0 \zeta_j$ for H_2 , $\lambda_j = \delta_j' \delta_j / 2$ and $\delta_j' \mu_{0j} = 0$. Choose a $p \times (p-s)$ matrix B_s such that $(B_0 \ B_s) \in \mathbf{O}(p)$. Let

$$\bar{\mathbf{x}}_j = \frac{1}{n_j} \sum_l \mathbf{x}_{jl}, \quad n = \sum_j n_j \quad \text{and} \quad \rho_j = \frac{n_j}{n}.$$

We assume that each ρ_j is $\mathbf{O}(1)$ as $n \rightarrow \infty$, i.e., $n_j \rightarrow \infty$ as $n \rightarrow \infty$. Under these notations and assumptions we consider the distribution of

$$\mathbf{y}_j = \sqrt{n_j} \sum_j^{-\frac{1}{2}} (\bar{\mathbf{x}}_j - A_j \mu_{0j} - A_j \delta_j / \sqrt{n_j}),$$

where $\Sigma_j = A_j' \mu_j \mu_j' + \frac{A_j}{\kappa_j} (I_p - \mu_j \mu_j')$. The characteristic function of \mathbf{y}_j is given by

$$\Psi_j(\mathbf{t}_j) = \left\{ \frac{a_p(\bar{\omega}_j)}{a_p(\kappa_j)} \right\}^{n_j} \exp \left\{ - \left(\sqrt{n_j} \frac{A_j}{\sqrt{A_j'}} \mu_{0j}' \mathbf{t}_j + \sqrt{A_j \kappa_j} \delta_j' \mathbf{t}_j \right) i \right\},$$

where

$$\begin{aligned}
\bar{\omega}_j &= \left\| \kappa_j \boldsymbol{\mu}_j + i \frac{1}{\sqrt{n_j}} \sum_j^{-\frac{1}{2}} \mathbf{t}_j \right\| \\
&= \left(\kappa_j^2 + \frac{2i\kappa_j}{\sqrt{n_j}} \mathbf{t}'_j \sum_j^{-\frac{1}{2}} \boldsymbol{\mu}_j + \frac{i^2}{n_j} \mathbf{t}'_j \sum_j^{-1} \mathbf{t}_j \right)^{\frac{1}{2}} \\
&= \kappa_j \left[1 + \frac{1}{\sqrt{n_j}} \frac{2}{\kappa_j \sqrt{A'_j}} \tilde{\alpha}_j + \frac{1}{n_j} \left\{ \left(\frac{1}{A'_j \kappa_j^2} - \frac{1}{A_j \kappa_j} \right) \tilde{\alpha}_j^2 + \frac{1}{A_j \kappa_j} \beta_j \right\} \right]^{\frac{1}{2}},
\end{aligned}$$

and

$$\tilde{\alpha}_j = i\boldsymbol{\mu}'_j \mathbf{t}_j, \quad \beta_j = i^2 \mathbf{t}'_j \mathbf{t}_j.$$

Noting that

$$\log a_p(\kappa + x) = \log a_p(\kappa) + \sum_{j=1}^{\infty} \frac{A^{(j-1)}}{j!} x^j, \quad (4.2)$$

where $A^{(j)} = (d^j/d\kappa^j)A$ denotes the j th differential of A about κ and $A^{(0)} = A$, we obtain

$$\log \Psi_j(\mathbf{t}_j) = \frac{1}{2} \beta_j + \frac{1}{\sqrt{n_j}} l_{1j}(\mathbf{t}_j) + \mathcal{O}(n_j^{-\frac{3}{2}}),$$

where

$$\begin{aligned}
l_{1j}(\mathbf{t}_j) &= e_{11j} \alpha_j^3 + e_{12j} \alpha_j \beta_j + e_{13j} \alpha_j \gamma_j + e_{14j} \alpha_j, \\
l_{2j}(\mathbf{t}_j) &= e_{21j} \alpha_j^4 + e_{22j} \alpha_j^2 \beta_j + e_{23j} \beta_j^2 + e_{24j} \alpha_j^2 \gamma_j + e_{25j} \beta_j \gamma_j + e_{26j} \alpha_j^2 \\
&\quad + e_{27j} \gamma_j^2 + e_{28j} \gamma_j, \\
\alpha_j &= i\boldsymbol{\mu}'_{0j} \mathbf{t}_j, \quad \gamma_j = i\boldsymbol{\delta}'_j \mathbf{t}_j,
\end{aligned}$$

and

$$\begin{aligned}
e_{11j} &= \frac{1}{2\sqrt{A'_j \kappa_j}} - \frac{\sqrt{A'_j}}{2A_j} + \frac{A''_j}{6A'_j \sqrt{A'_j}}, \\
e_{12j} &= -\frac{1}{2\sqrt{A'_j \kappa_j}} + \frac{\sqrt{A'_j}}{2A_j}, \\
e_{13j} &= -\sqrt{\frac{A_j}{A'_j \kappa_j}} + \sqrt{\frac{A'_j \kappa_j}{A_j}} = 2\sqrt{A_j \kappa_j} e_{12j}, \\
e_{14j} &= -\frac{A_j}{\sqrt{A'_j}} \lambda_j,
\end{aligned}$$

$$\begin{aligned}
e_{21j} &= -\frac{1}{2A_j'\kappa_j^2} + \frac{3}{8A_j\kappa_j} + \frac{A_j'}{8A_j^2} - \frac{A_j''}{4A_jA_j'} + \frac{A_j'''}{24A_j'^2}, \\
e_{22j} &= \frac{1}{2A_j'\kappa_j^2} - \frac{1}{4A_j\kappa_j} - \frac{A_j'}{4A_j^2} + \frac{A_j''}{4A_jA_j'}, \\
e_{23j} &= -\frac{1}{8A_j\kappa_j} + \frac{A_j'}{8A_j^2}, \\
e_{24j} &= \frac{\sqrt{A_j}}{4A_j'\kappa_j\sqrt{\kappa_j}} - \frac{1}{2\sqrt{A_j\kappa_j}} - \frac{A_j'\sqrt{\kappa_j}}{2A_j\sqrt{A_j}} + \frac{A_j''\sqrt{\kappa_j}}{4\sqrt{A_j}A_j'}, \\
e_{25j} &= -\frac{1}{2\sqrt{A_j\kappa_j}} + \frac{A_j'\sqrt{\kappa_j}}{2A_j\sqrt{A_j}}, \\
e_{26j} &= \frac{A_j}{A_j'\kappa_j} \lambda_j - \lambda_j, \\
e_{27j} &= \frac{A_j'\kappa_j}{2A_j} - \frac{1}{2}, \\
e_{28j} &= -\sqrt{A_j\kappa_j} \lambda_j.
\end{aligned}$$

Then

$$\Psi_j(\mathbf{t}_j) = \exp\left(\frac{1}{2}\beta_j\right) \left[1 + \frac{1}{\sqrt{n_j}} l_{1j}^*(\mathbf{t}_j) + \frac{1}{n_j} l_{2j}^*(\mathbf{t}_j) + O(n_j^{-\frac{3}{2}})\right], \quad (4.3)$$

where

$$\begin{aligned}
l_{1j}^*(\mathbf{t}_j) &= l_{1j}(\mathbf{t}_j), \\
l_{2j}^* &= \frac{1}{2} l_{1j}(\mathbf{t}_j)^2 + l_{2j}(\mathbf{t}_j) \\
&= \frac{1}{2} e_{11j}^2 \alpha_j^6 + e_{11j} e_{12j} \alpha_j^4 \beta_j + \frac{1}{2} e_{12j}^2 \alpha_j^2 \beta_j^2 + e_{11j} e_{13j} \alpha_j^4 \gamma_j + e_{12j} e_{13j} \alpha_j^2 \beta_j \gamma_j \\
&\quad + (e_{21j} + e_{11j} e_{14j}) \alpha_j^4 + (e_{22j} + e_{12j} e_{14j}) \alpha_j^2 \beta_j + e_{23j} \beta_j^2 + \frac{1}{2} e_{13j}^2 \alpha_j^2 \gamma_j^2 \\
&\quad + (e_{24j} + e_{13j} e_{14j}) \alpha_j^2 \gamma_j + e_{25j} \beta_j \gamma_j + \left(e_{26j} + \frac{1}{2} e_{14j}^2\right) \alpha_j^2 + e_{27j} \gamma_j^2 + e_{28j} \gamma_j.
\end{aligned}$$

Inverting (4.3), we have the following theorem.

THEOREM 4.1 *Under the assumption of (4.1) the joint probability density function of $Y = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$ can be asymptotically expanded as*

$$f(\mathbf{Y}) = \prod_m \phi_p(\mathbf{y}_m) \cdot \left[1 + \frac{1}{\sqrt{n}} \sum_j \frac{1}{\sqrt{\rho_j}} f_{1j}(\mathbf{y}_j) + \frac{1}{n} \left\{ \sum_j \frac{1}{\rho_j} f_{2j}(\mathbf{y}_j) + \sum_{j < l} \frac{1}{\sqrt{\rho_j \rho_l}} f_{1j}(\mathbf{y}_j) f_{1l}(\mathbf{y}_l) \right\} + O(n^{-\frac{3}{2}}) \right],$$

where $\phi_p(\mathbf{y}) = (2\pi)^{-p/2} \exp(-\mathbf{y}'\mathbf{y}/2)$,

$$\begin{aligned} f_{1j}(\mathbf{y}_j) &= e_{11j}(\mu'_{0j}\mathbf{y}_j)^3 + e_{12j}\mu'_{0j}\mathbf{y}_j \cdot \mathbf{y}'_j\mathbf{y}_j + e_{13j}\mu'_{0j}\mathbf{y}_j \cdot \boldsymbol{\delta}'_j\mathbf{y}_j + e_{14j}^*\mu'_{0j}\mathbf{y}_j, \\ f_{2j}(\mathbf{y}_j) &= \frac{1}{2} e_{11j}^2(\mu'_{0j}\mathbf{y}_j)^6 + e_{11j}e_{12j}(\mu'_{0j}\mathbf{y}_j)^4\mathbf{y}'_j\mathbf{y}_j + \frac{1}{2} e_{12j}^2(\mu'_{0j}\mathbf{y}_j)^2(\mathbf{y}'_j\mathbf{y}_j)^2 \\ &\quad + e_{11j}e_{13j}(\mu'_{0j}\mathbf{y}_j)^4\boldsymbol{\delta}'_j\mathbf{y}_j + e_{12j}e_{13j}(\mu'_{0j}\mathbf{y}_j)^2\mathbf{y}'_j\mathbf{y}_j \cdot \boldsymbol{\delta}'_j\mathbf{y}_j + e_{21j}^*(\mu'_{0j}\mathbf{y}_j)^4 \\ &\quad + e_{22j}^*(\mu'_{0j}\mathbf{y}_j)^2\mathbf{y}'_j\mathbf{y}_j + e_{23j}^*(\mathbf{y}'_j\mathbf{y}_j)^2 + \frac{1}{2} e_{13j}^2(\mu'_{0j}\mathbf{y}_j)^2(\boldsymbol{\delta}'_j\mathbf{y}_j)^2 \\ &\quad + e_{24j}^*(\mu'_{0j}\mathbf{y}_j)^2\boldsymbol{\delta}'_j\mathbf{y}_j + e_{25j}^*\mathbf{y}'_j\mathbf{y}_j \cdot \boldsymbol{\delta}'_j\mathbf{y}_j + e_{26j}^*(\mu'_{0j}\mathbf{y}_j)^2 \\ &\quad + e_{29j}^*\mathbf{y}'_j\mathbf{y}_j + e_{27j}^*(\boldsymbol{\delta}'_j\mathbf{y}_j)^2 + e_{28j}^*\boldsymbol{\delta}'_j\mathbf{y}_j + e_{20j}^*, \end{aligned}$$

and

$$\begin{aligned} e_{14j}^* &= -3e_{11j} - (p+2)e_{12j} + e_{14j}, \\ e_{21j}^* &= -\frac{15}{2}e_{11j}^2 - (p+8)e_{11j}e_{12j} + e_{21j} + e_{11j}e_{14j}, \\ e_{22j}^* &= -6e_{11j}e_{12j} - (p+6)e_{12j}^2 + e_{22j} + e_{12j}e_{14j}, \\ e_{23j}^* &= -\frac{1}{2}e_{12j}^2 + e_{23j}, \\ e_{24j}^* &= -6e_{11j}e_{13j} - (p+6)e_{12j}e_{13j} + e_{24j} + e_{13j}e_{14j}, \\ e_{25j}^* &= -e_{12j}e_{13j} + e_{25j}, \\ e_{26j}^* &= \frac{45}{2}e_{11j}^2 + 6(p+6)e_{11j}e_{12j} + \frac{1}{2}(p+4)(p+6)e_{12j}^2 - 6e_{21j} - 6e_{11j}e_{14j}, \\ &\quad - (p+4)(e_{22j} + e_{12j}e_{14j}) - \lambda_j e_{13j}^2 + e_{26j} + \frac{1}{2}e_{14j}^2, \\ e_{27j}^* &= -\frac{1}{2}e_{13j}^2 + e_{27j}, \\ e_{28j}^* &= 3e_{11j}e_{13j} + (p+4)e_{12j}e_{13j} - e_{24j} - e_{13j}e_{14j} - (p+2)e_{25} + e_{28j}, \\ e_{29j}^* &= 3e_{11j}e_{12j} + (p+4)e_{12j}^2 - e_{22j} - e_{12j}e_{14j} - 2(p+2)e_{23j}, \end{aligned}$$

$$\begin{aligned}
e_{20j}^* &= -\frac{15}{2}e_{11j}^2 - 3(p+4)e_{11j}e_{12j} - \frac{1}{2}(p+2)(p+4)e_{12j}^2 + 3e_{21j} + 3e_{11j}e_{14j}, \\
&+ (p+2)e_{22j} + (p+2)e_{12j}e_{14j} + p(p+2)e_{23j} + \lambda_j e_{13j}^2 - e_{26j} \\
&- \frac{1}{2}e_{14j}^2 - 2\lambda_j e_{27j}.
\end{aligned}$$

Hayakawa [1992] derived an asymptotic expansion of the joint probability density function of $\sqrt{n_j}(\bar{\mathbf{x}}_j - A_p(\kappa_j)\boldsymbol{\mu}_j)$'s. The above theorem can be also obtained from his result by considering an appropriate transformation. We note that our result has another merit since the limiting distribution is $N_{kp}(0, I_{kp})$, though our formula is slightly complicated.

4.2. LR Tests of H_1

First we consider testing hypothesis H_1 against K_1 when κ_j 's are known. Then the LR statistic T_{L11} is given as follows:

$$T_{L11} = 2n(\sum_j \rho_j \kappa_j \|\bar{\mathbf{x}}_j\| - \|\sum_j \rho_j \kappa_j b_j \bar{\mathbf{x}}_j\|).$$

We are interested in studying the distribution of T_{L11} under the null hypothesis H_1 and under a sequence of the alternatives (4.1). In general, T_{L11} is expanded in terms of $\mathbf{y} = (\mathbf{y}'_1 \cdots \mathbf{y}'_k)'$ as

$$T_{L11} = (\mathbf{y} + \boldsymbol{\delta})'(I_k - H) \otimes (I_p - M)(\mathbf{y} + \boldsymbol{\delta}) + \frac{1}{\sqrt{n}} q_{11}(\mathbf{y}) + \mathbf{O}_p(n^{-1}),$$

where

$$\begin{aligned}
H &= \boldsymbol{\beta}\boldsymbol{\beta}'/(\boldsymbol{\beta}'\boldsymbol{\beta}), \quad M = \boldsymbol{\mu}\boldsymbol{\mu}', \\
\boldsymbol{\delta} &= (\sqrt{A_1\kappa_1}\boldsymbol{\delta}'_1 \cdots \sqrt{A_k\kappa_k}\boldsymbol{\delta}'_k)', \\
\boldsymbol{\beta} &= (\sqrt{\rho_1 A_1 \kappa_1} b_1 \cdots \sqrt{\rho_k A_k \kappa_k} b_k)',
\end{aligned}$$

and \otimes denotes the Kronecker product. Here $q_{11}(\mathbf{y})$ is a polynomial of the first and the third degrees in the elements of \mathbf{y} . The nonnull distribution is given by

$$P(T_{L11} \leq x) = P(\chi_{(p-1)(k-1)}^2(\lambda) \leq x) + \mathbf{O}(n^{-1}),$$

where $\lambda = \{\sum_j A_j \kappa_j \lambda_j - \|\sum_j \sqrt{\rho_j} b_j A_j \kappa_j \boldsymbol{\delta}_j\|^2 / (2\sum_l \rho_l A_l \kappa_l)\}$.

Next we consider testing hypotheses H_1 when $\kappa_1 = \kappa_2 = \cdots = \kappa_k = \kappa$ but κ is unknown. The MLE $\hat{\kappa}$ of κ satisfies

$$A_p(\hat{\kappa}) = \sum_j \rho_j \|\bar{\mathbf{x}}_j\|.$$

Further let $\check{\kappa}$ satisfy

$$A_p(\check{\kappa}) = \|\sum \rho_j b_j \bar{x}_j\|$$

and this is the MLE of κ under the hypothesis H_1 . Then LR statistic T_{L12} is given as follows:

$$T_{L12} = 2n \{ \log a_p(\check{\kappa}) - \log a_p(\hat{\kappa}) - \check{\kappa} \|\sum_j \rho_j b_j \bar{x}_j\| + \hat{\kappa} \sum_j \rho_j \|\bar{x}_j\| \}.$$

It is noted that the Taylor expansion of $A_p^{-1}(\kappa)$ is given by

$$A_p^{-1}(A_p(\kappa) + x) = \kappa + \frac{1}{A'} x - \frac{A''}{2A'^3} x^2 + \frac{3A''^2 - A' A^{(3)}}{6A'^5} x^3 + \dots \quad (4.4)$$

Using this and the expansion (4.2) about $\log a_p(\kappa)$, T_{L12} is expanded in terms of \mathbf{y} under a sequence of alternatives (4.1) as

$$T_{L12} = (\mathbf{y} + \boldsymbol{\delta})'(I_k - H) \otimes (I_p - M)(\mathbf{y} + \boldsymbol{\delta}) + \frac{1}{\sqrt{n}} q_{11}^*(\mathbf{y}) + \mathbf{O}_p(n^{-1}),$$

where $q_{11}^*(\mathbf{y})$ is a polynomial with the same property with $q_{11}(\mathbf{y})$ and

$$\begin{aligned} H &= \boldsymbol{\beta}\boldsymbol{\beta}', & M &= \boldsymbol{\mu}\boldsymbol{\mu}', \\ \boldsymbol{\delta} &= (\sqrt{A\kappa}\boldsymbol{\delta}'_1 \cdots \sqrt{A\kappa}\boldsymbol{\delta}'_k)', \\ \boldsymbol{\beta} &= (\sqrt{\rho_1}b_1 \cdots \sqrt{\rho_k}b_k)', \end{aligned}$$

The nonnull distribution is given by

$$P(T_{L12} \leq x) = P(\chi_{(p-1)(k-1)}^2(\lambda) \leq x) + \mathbf{O}(n^{-1}),$$

where $\lambda = A\kappa \{ \sum_j \lambda_j - \|\sum_j \sqrt{\rho_j} b_j \boldsymbol{\delta}_j\|^2 / 2 \}$.

4.3. Test of H_2 when the concentration parameters are known

In this section we consider testing hypothesis H_2 against K_2 based on k -random samples of size n_j from $M_p(\boldsymbol{\mu}_j, \kappa_j)$ when κ_j 's are known. Then the two statistics, Watson statistic T_{W21} and LR statistic T_{L21} are given as follows:

$$\begin{aligned} \text{(i)} \quad T_{W21} &= n \sum_j \frac{\rho_j \kappa_j}{A_j} \|(I_p - B_0 B_0') \bar{x}_j\|^2, \\ \text{(ii)} \quad T_{L21} &= 2n \sum_j \rho_j \kappa_j (\|\bar{x}_j\| - \|B_0' \bar{x}_j\|). \end{aligned}$$

The limiting distributions of these statistics have been obtained by Watson [1983a]. Our purpose is to obtain asymptotic expansions of these distributions by extending the results due to Watamori [1992] in a special case

$k = 1$. Under a sequence of the alternatives (4.1), T_{W21} and T_{L21} are expanded in terms of Y as

$$\begin{aligned} T_{W21} &= q_0(Y) + \mathbf{O}_p(n^{-\frac{3}{2}}), \\ T_{L21} &= q_0(Y) + \frac{1}{\sqrt{n}} q_1(Y) + \frac{1}{n} q_2(Y) + \mathbf{O}_p(n^{-\frac{3}{2}}), \end{aligned}$$

where

$$q_0(Y) = \sum_j q_{00j}(\mathbf{y}_j), \quad (4.5)$$

$$q_1(Y) = - \sum_j \frac{\sqrt{A'_j}}{\sqrt{\rho_j A_j}} \boldsymbol{\mu}'_{0j} \mathbf{y}_j q_{00j}(\mathbf{y}_j),$$

$$q_2(Y) = \sum_j \frac{1}{\rho_j} \left\{ - \frac{1}{4A_j \kappa_j} q_{21j}(\mathbf{y}_j) + \frac{A'_j}{A_j^2} (\boldsymbol{\mu}'_{0j} \mathbf{y}_j)^2 \right\} q_{00j}(\mathbf{y}_j),$$

$$q_{00j}(\mathbf{y}_j) = \mathbf{y}'_j \mathbf{y}_j - \mathbf{y}'_j \mathbf{B}_0 \mathbf{B}'_0 \mathbf{y}_j + 2\sqrt{A_j \kappa_j} \boldsymbol{\delta}'_j \mathbf{y}_j + 2A_j \kappa_j \lambda_j,$$

$$q_{01j}(\mathbf{y}_j) = \mathbf{y}'_j \mathbf{y}_j - (\boldsymbol{\mu}'_{0j} \mathbf{y}_j)^2 + 2\sqrt{A_j \kappa_j} \boldsymbol{\delta}'_j \mathbf{y}_j + 2A_j \kappa_j \lambda_j,$$

$$q_{21j}(\mathbf{y}_j) = 2q_{01j}(\mathbf{y}_j) - q_{00j}(\mathbf{y}_j).$$

Thus the characteristic function of T_{W21} , $\Psi_{W21}(t)$ is given by

$$\Psi_{W21}(t) = \int \cdots \int \prod_j [\exp \{itq_{00j}(\mathbf{y}_j)\} \cdot \phi_p(\mathbf{y}_j)] \times g(\mathbf{y}_1, \dots, \mathbf{y}_k) \prod_j d\mathbf{y}_j,$$

where

$$\begin{aligned} g(\mathbf{y}_1, \dots, \mathbf{y}_k) &= 1 + \frac{1}{\sqrt{n}} \sum_j \frac{1}{\sqrt{\rho_j}} f_{1j}(\mathbf{y}_j) \\ &\quad + \frac{1}{n} \left\{ \sum_j \frac{1}{\rho_j} f_{2j}(\mathbf{y}_j) + \sum_{j < l} \frac{1}{\sqrt{\rho_j \rho_l}} f_{1j}(\mathbf{y}_j) f_{1l}(\mathbf{y}_l) \right\} + \mathbf{O}(n^{-\frac{3}{2}}). \end{aligned}$$

Consider the transformation

$$\mathbf{z}_j = A^{-\frac{1}{2}} \mathbf{y}_j - \frac{2it}{\sqrt{1-2it}} \sqrt{A_j \kappa_j} \boldsymbol{\delta}_j, \quad (4.6)$$

where

$$A = \frac{1}{1-2it} (I_p - \mathbf{B}_0 \mathbf{B}'_0) + \mathbf{B}_0 \mathbf{B}'_0.$$

Then we obtain

$$\Psi_{W_{21}}(t) = (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right) \times \quad (4.7)$$

$$E_{z_1, \dots, z_k} \left[g \left(A^{\frac{1}{2}} z_1 + \frac{2it}{1 - 2it} \sqrt{A_1 \kappa_1} \delta_1, \dots, A^{\frac{1}{2}} z_k + \frac{2it}{1 - 2it} \sqrt{A_k \kappa_k} \delta_k \right) \right],$$

where the expectation with respect to z_j is taken under $N_p(\mathbf{0}, I_p)$ and $\lambda = \sum_j A_j \kappa_j \lambda_j$. After calculating the expectation in (4.7), we obtain

$$\Psi_{W_{21}}(t) = (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right) \times \left[1 + \frac{1}{4n} \sum_{m=0}^4 d_{1m} (1 - 2it)^{-m} + O(n^{-\frac{3}{2}}) \right],$$

where

$$\begin{aligned} d_{10} &= 2\lambda^{(21)} + 6\lambda^{(22)} + 2(p-s)\lambda^{(11)} + (p-s)(p-s+2)\rho^{(1)}, \\ d_{11} &= -8\lambda^{(22)} - 2(p-s)\lambda^{(11)} - 2(p-s)(p-s+2)\rho^{(1)}, \\ d_{12} &= -4\lambda^{(21)} + 4\lambda^{(22)} - 2(p-s+2)\lambda^{(11)} + (p-s)(p-s+2)\rho^{(1)}, \\ d_{13} &= 2(p-s+2)\lambda^{(11)}, \quad d_{14} = 2\lambda^{(21)} - 2\lambda^{(22)}, \\ \lambda^{(21)} &= \sum_j A'_j \kappa_j^2 \lambda_j^2 / \rho_j, \quad \lambda^{(22)} = \sum_j A_j \kappa_j \lambda_j^2 / \rho_j, \\ \lambda^{(11)} &= \sum_j (A'_j \kappa_j / A_j - 1) \lambda_j / \rho_j, \quad \rho^{(1)} = \sum_j (A'_j / A_j^2 - 1 / A_j \kappa_j) / \rho_j. \end{aligned} \quad (4.8)$$

Inverting this characteristic function, we have the following theorem.

THEOREM 4.2 *Under a sequence of the alternatives (4.1) the distribution function of $T_{W_{21}}$ can be asymptotically expanded as*

$$P(T_{W_{21}} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) + \frac{1}{4n} \sum_{m=0}^4 d_{1m} P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + O(n^{-\frac{3}{2}}), \quad (4.9)$$

where d_{1m} 's are given by (4.8), and $\chi_f^2(\lambda)$ denotes a noncentral χ^2 -variate with f degrees of freedom and noncentrality parameter λ .

Letting $\delta_j = \mathbf{0}$ in (4.9), we obtain an asymptotic expansion of the null distribution of $T_{W_{21}}$,

$$P(T_{W_{21}} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} (p-s)(p-s+2)\rho^{(1)} \times \quad (4.10)$$

$$\{P(\chi_{k(p-s)}^2 \leq x) - 2P(\chi_{k(p-s)+2}^2 \leq x) + P(\chi_{k(p-s)+4}^2 \leq x)\} + \mathcal{O}(n^{-\frac{3}{2}}).$$

Next we consider the distribution of T_{L21} under (4.1). Similarly the characteristic function of T_{L21} , $\Psi_{L21}(t)$ can be expanded as

$$\begin{aligned} \Psi_{L21}(t) &= (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right) \\ &\times \left[1 + \frac{1}{4n} \sum_{m=0}^3 h_{1m} (1 - 2it)^{-m} + \mathcal{O}(n^{-\frac{3}{2}})\right], \end{aligned} \quad (4.11)$$

where

$$\begin{aligned} h_{10} &= 2\lambda^{(21)} + 6\lambda^{(22)} - 2(p-s)\lambda^* + (p-s)(p+s-4)\rho^*, \\ h_{11} &= -4\lambda^{(21)} - 8\lambda^{(22)} + 2(2p-s-1)\lambda^* - (p-s)(p+s-4)\rho^*, \\ h_{12} &= 2\lambda^{(21)} + 4\lambda^{(22)} - 2(p-1)\lambda^*, \quad h_{13} = -2\lambda^{(22)}, \\ \lambda^* &= \sum_j \lambda_j / \rho_j, \quad \rho^* = \sum_j 1 / (2A_j \kappa_j \rho_j), \end{aligned} \quad (4.12)$$

and $\lambda^{(21)}$ and $\lambda^{(22)}$ are given in (4.8). Inverting this characteristic function, we have the following theorem.

THEOREM 4.3 *Under a sequence of the alternatives (4.1) the distribution function of T_{L21} can be asymptotically expanded as*

$$\begin{aligned} P(T_{L21} \leq x) &= P(\chi_{k(p-s)}^2(\lambda) \leq x) \\ &+ \frac{1}{4n} \sum_{m=0}^3 h_{1m} P(\chi_{k(p-s)+2j}^2(\lambda) \leq x) + \mathcal{O}(n^{-\frac{3}{2}}), \end{aligned} \quad (4.13)$$

where h_{1m} 's are given by (4.12).

Letting $\delta_j = \mathbf{0}$ in (4.13), we obtain an asymptotic expansion of the null distribution of T_{L21} ,

$$\begin{aligned} P(T_{L21} \leq x) &= P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} (p-s)(p+s-4)\rho^* \\ &\{P(\chi_{k(p-s)}^2 \leq x) - P(\chi_{k(p-s)+2}^2 \leq x)\} + \mathcal{O}(n^{-\frac{3}{2}}). \end{aligned} \quad (4.14)$$

This result implies that $\tilde{T}_{L21} = \{1 + (p+s-4)\rho^*/(2nk)\} T_{L21}$ gives a better χ^2 -approximation, since

$$P(\tilde{T}_{L21} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \mathcal{O}(n^{-\frac{3}{2}}).$$

Let β_{W21} and β_{L21} be the powers of T_{W21} and T_{L21} with a level of

significance α . These α, β are different from those ones used in section 4.1. Then from (4.9) and (4.13) it is possible to obtain asymptotic expansions for β_{W21} and β_{L21} . A useful expression for such powers has been obtained by Fujikoshi [1988]. Applying his result to (4.9) and (4.13), we obtain the following theorem.

THEOREM 4.4. *Under a sequence of the alternatives (4.1) the powers β_{W21} and β_{L21} of T_{W21} and T_{L21} with a level of significance α are given by*

$$\begin{aligned} \beta_{W21} &= P(\chi_{k(p-s)}^2(\lambda) \geq x_\alpha) \\ &+ \frac{1}{n} \left[\{-\lambda^{(21)} - 3\lambda^{(22)} - (p-s)\lambda^{(1)}\} g_{k(p-s)+2}(x_\alpha; \lambda) \right. \\ &+ \left\{ -\lambda^{(21)} + \lambda^{(22)} + \frac{1}{2}(p-s+2)\lambda^{(1)} \right\} g_{k(p-s)+4}(x_\alpha; \lambda) \\ &\left. + \lambda^{(21)} - \lambda^{(22)} g_{k(p-s)+6}(x_\alpha; \lambda) \right] + O(n^{-\frac{3}{2}}), \end{aligned} \quad (4.15)$$

and

$$\begin{aligned} \beta_{L21} &= P(\chi_{k(p-s)}^2(\lambda) \geq x_\alpha) \\ &+ \frac{1}{n} \left[(-\lambda^{(21)} - 3\lambda^{(22)}) + (p-s)\lambda^* \right\} g_{k(p-s)+2}(x_\alpha; \lambda) \\ &+ \left\{ \lambda^{(21)} + \lambda^{(22)} - \frac{1}{2}(p-s+2)\lambda^* \right\} g_{k(p-s)+4}(x_\alpha; \lambda) \\ &\left. - \lambda^{(22)} g_{k(p-s)+6}(x_\alpha; \lambda) \right] + O(n^{-\frac{3}{2}}), \end{aligned} \quad (4.16)$$

where x_α is the upper α point of $\chi_{k(p-s)}^2$ and $g_f(x_\alpha; \lambda)$ is the probability density function of $\chi_f^2(\lambda)$.

Then taking the difference between (4.15) and (4.16),

$$\begin{aligned} \beta_{W21} - \beta_{L21} &= \frac{1}{n} \left[-(p-s) \left(\sum_j \frac{A'_j \kappa_j}{A_j \rho_j} \lambda_j \right) g_{k(p-s)+2}(x_\alpha; \lambda) \right. \\ &+ \left\{ -2 \sum_j \frac{A'_j \kappa_j^2}{\rho_j} \lambda_j^2 + \frac{1}{2}(p-s+2) \left(\sum_j \frac{A'_j \kappa_j}{A_j \rho_j} \lambda_j \right) \right\} g_{k(p-s)+4}(x_\alpha; \lambda) \\ &\left. + \left(\sum_j \frac{A'_j \kappa_j^2}{\rho_j} \lambda_j^2 \right) g_{k(p-s)+6}(x_\alpha; \lambda) \right] + O(n^{-\frac{3}{2}}). \end{aligned} \quad (4.17)$$

It is interesting to derive further reductions for (4.17), which clarify the difference of the powers.

4.4. Test of H_2 when the common concentration parameters are unknown

First we consider testing hypotheses H_2 against K_2 when $\kappa_1 = \kappa_2 = \dots = \kappa_k = \kappa$ but κ is unknown. The MLE $\hat{\kappa}$ of κ satisfies

$$A_p(\hat{\kappa}) = \sum_j \rho_j \|\bar{\mathbf{x}}_j\|. \tag{4.18}$$

Further let $\tilde{\kappa}$ satisfy

$$A_p(\tilde{\kappa}) = \sum_j \rho_j \|B'_0 \bar{\mathbf{x}}_j\|, \tag{4.19}$$

and this is the MLE of κ under the hypothesis H_2 . Then two statistics T_{W22} and T_{L22} are given as follows:

- (i) $T_{W22} = n \sum_j \frac{\rho_j \hat{\kappa}}{A_p(\hat{\kappa})} \|(I_p - B_0 B'_0) \bar{\mathbf{x}}_j\|^2,$
- (ii) $T_{L22} = 2n \{ \log a_p(\tilde{\kappa}) - \log a_p(\hat{\kappa}) - \tilde{\kappa} \sum_j \rho_j \|B'_0 \bar{\mathbf{x}}_j\| + \hat{\kappa} \sum_j \rho_j \|\bar{\mathbf{x}}_j\| \}.$

The statistic T_{W22} has been proposed by Watson [1983a] and T_{L22} is LR statistic. Watamori [1992] has obtained asymptotic expansions of the null and nonnull distributions for these statistics in a special case $k = 1$. The methods used there can be extended directly to multi-sample case, *i.e.*, $k > 1$. Using (4.2) and (4.4), T_{W22} and T_{L22} are expanded in terms of Y under a sequence of alternatives (4.1) as

$$T_{W22} = q_0(Y) + \frac{1}{\sqrt{n}} q'_1(Y) + \frac{1}{n} q'_2(Y) + O_p(n^{-\frac{3}{2}}), \tag{4.20}$$

$$T_{L22} = q_0(Y) + \frac{1}{\sqrt{n}} q^*_1(Y) + \frac{1}{n} q^*_2(Y) + O_p(n^{-\frac{3}{2}}), \tag{4.21}$$

where

$$q_0(Y) = \sum_j q_{00j}(y_j),$$

$$q^*_1(Y) = \frac{1}{\sqrt{A' \kappa}} \sum_j \sqrt{\rho_j} \mu'_{0j} \mathbf{y}_j q_0(Y) - \frac{\sqrt{A'}}{A} \sum_j \frac{1}{\rho_j} \mu'_{0j} \mathbf{y}_j q_{00j}(y_j),$$

$$q_1'(Y) = \left(\frac{1}{\sqrt{A'\kappa}} - \frac{\sqrt{A'}}{A} \right) \sum_j \sqrt{\rho_j} \mu'_{0j} y_j q_0(Y),$$

$$q_2^*(Y) = \frac{1}{4A'\kappa^2} q_0(Y) \sum_j q_{21j}(y_j) - \frac{1}{4A\kappa} \sum_j \frac{1}{\rho_j} q_{00j}(y_j) q_{21j}(y_j)$$

$$- \frac{A''}{2A'^2\kappa} q_0(Y) \left(\sum_j \sqrt{\rho_j} \mu'_{0j} y_j \right)^2$$

$$- \frac{1}{A\kappa} \sum_j \rho_j \mu'_{0j} y_j \sum_l \frac{1}{\sqrt{\rho_l}} \mu'_{0l} y_l q_{00l}(y_l) + \frac{A'}{A^2} \sum_j \frac{1}{\rho_j} (\mu'_{0j} y_j)^2 q_{00j}(y_j),$$

$$q_2'(Y) = \left(\frac{1}{2A'\kappa^2} - \frac{1}{2A\kappa} \right) q_0(Y) \sum_j q_{01j}(y_j)$$

$$- \left(\frac{A''}{2A'^2\kappa} + \frac{1}{A\kappa} - \frac{A'}{A^2} \right) \left\{ \sum_j \sqrt{\rho_j} \mu'_{0j} y_j \right\}^2 q_0(Y),$$

$$q_{00j}(y_j) = y_j' y_j - y_j' B_0 B_0' y_j + 2\sqrt{A\kappa} \delta_j' y_j + 2A\kappa \lambda_j,$$

$$q_{01j}(y_j) = y_j' y_j - (\mu'_{0j} y_j)^2 + 2\sqrt{A\kappa} \delta_j' y_j + 2A\kappa \lambda_j,$$

$$q_{21j}(y_j) = 2q_{01j}(y_j) - q_{00j}(y_j).$$

The characteristic functions of T_{W22} and T_{L22} , $\Psi_{W22}(t)$ and $\Psi_{L22}(t)$, respectively can be obtained in a similar way. Then we obtain

$$\Psi_{W22}(t) = (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right)$$

$$\times \left[1 + \frac{1}{4n} \sum_{m=0}^4 d_{2m} (1 - 2it)^{-m} + O(n^{-\frac{3}{2}}) \right],$$

and

$$\Psi_{L22}(t) = (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right)$$

$$\times \left[1 + \frac{1}{4n} \sum_{m=0}^4 h_{2m} (1 - 2it)^{-m} + O(n^{-\frac{3}{2}}) \right],$$

where $\lambda = A\kappa \sum_j \lambda_j$. The coefficients d_{2j} 's and h_{2j} 's are quite complicated and omitted here. Inverting these characteristic functions, we have the following theorem.

THEOREM 4.5 *Under a sequence of the alternatives (4.1) the distribution functions of T_{W22} and T_{L22} can be asymptotically expanded as*

$$P(T_{W22} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) \quad (4.22)$$

$$+ \frac{1}{4n} \sum_{m=0}^4 d_{2m} P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + \mathcal{O}(n^{-\frac{3}{2}}),$$

and

$$P(T_{L22} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) \quad (4.23)$$

$$+ \frac{1}{4n} \sum_{m=0}^4 h_{2m} P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + \mathcal{O}(n^{-\frac{3}{2}}).$$

Letting $\delta_j = \mathbf{0}$ in (4.22) and (4.23), we obtain asymptotic expansions of the null distributions of T_{W22} and T_{L22} ,

$$P(T_{W22} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} \sum_{m=0}^2 \tilde{d}_{2m} \rho P(\chi_{k(p-s)+2m}^2 \leq x) + \mathcal{O}(n^{-\frac{3}{2}}),$$

and

$$P(T_{L22} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} \sum_{m=0}^2 \tilde{h}_{2m} \rho P(\chi_{k(p-s)+2m}^2 \leq x) + \mathcal{O}(n^{-\frac{3}{2}}),$$

where

$$\begin{aligned} \tilde{d}_{20} &= (p-s)(p-s+2) \left(\frac{A'}{2A^2} - \frac{1}{2A\kappa} \right) \sum_j \frac{1}{\rho_j} + 2(p-s) \left(\frac{A'}{A^2} - \frac{2}{A\kappa} + \frac{1}{A'\kappa^2} \right) \\ &\quad + k(p-s)^2 \left(\frac{A'}{A^2} - \frac{2}{A\kappa} - \frac{1}{A'\kappa^2} \right) (1 - \sum_j \sqrt{\rho_j}) \\ &\quad + k(p-s) \left\{ \frac{A''}{A'^2\kappa} - \frac{A'}{A^2} + \frac{k(s-1)+2}{A\kappa} - \frac{k(s-1)}{A'\kappa^2} \right\}, \\ \tilde{d}_{21} &= -(p-s)(p-s+2) \left\{ \left(\frac{A'}{A^2} - \frac{1}{A\kappa} \right) \sum_j \frac{1}{\rho_j} + \left(-\frac{1}{A\kappa} + \frac{1}{A'\kappa^2} \right) \right\} \\ &\quad - k(p-s) \{ k(p-s) + 2 \} \left(\frac{A'}{A^2} - \frac{2}{A\kappa} + \frac{1}{A'\kappa^2} \right) \\ &\quad - k(p-s) \left\{ \frac{A''}{A'^2\kappa} - \frac{A'}{A^2} + \frac{k(s-1)+2}{A\kappa} - \frac{k(s-1)}{A'\kappa^2} \right\} \\ &\quad + 2(p-s) \{ k(p-s) + 1 \} \left(\frac{A'}{A^2} - \frac{2}{A\kappa} + \frac{1}{A'\kappa^2} \right) \sum_j \sqrt{\rho_j}, \\ \tilde{d}_{22} &= (p-s)(p-s+2) \left\{ \left(\frac{A'}{2A^2} - \frac{1}{2A\kappa} \right) \sum_j \frac{1}{\rho_j} + \left(-\frac{1}{A\kappa} + \frac{1}{A'\kappa^2} \right) \right\} \end{aligned}$$

$$+ (p-s)\{k(p-s)+2\}\left(\frac{A'}{2A^2}-\frac{1}{A\kappa}+\frac{1}{2A'\kappa^2}\right)(k-2\sum_j\sqrt{\rho_j}),$$

and

$$\tilde{h}_{20} = h_1^* - h_3^*,$$

$$\tilde{h}_{21} = -2h_1^* - h_2^* + h_3^*,$$

$$\tilde{h}_{22} = h_1^* + h_2^*,$$

$$h_1^* = (p-s)(p-s+2)\left(\frac{A'}{2A^2}-\frac{1}{2A\kappa}\right)\sum_j\frac{1}{\rho_j} + (p-s)(p-s+2)\frac{A'}{A^2}\sum_j\frac{1}{\rho_j^2} \\ - (p-s)\{k(p-s)+2\}\frac{1}{A\kappa}\sum_j\frac{1}{\sqrt{\rho_j}} + k(p-s)\{k(p-s)+2\}\frac{1}{2A'\kappa^2},$$

$$h_2^* = -(p-s)(p-s+2)\left(\frac{A'}{A^2}-\frac{3}{2A\kappa}\right)\sum_j\frac{1}{\rho_j} + k(p-s)\{k(p-s)+2\}\frac{1}{2A'\kappa^2} \\ + (p-s)\{k(p-s)+2\}\left(\frac{1}{A\kappa}-\frac{1}{A'\kappa^2}\right)\sum_j\sqrt{\rho_j},$$

$$h_3^* = k^2(p-s)(s-1)\frac{1}{A'\kappa^2} - k(p-s)\frac{A'}{A'^2\kappa} \\ + \left\{- (p-s)(s-1)\frac{1}{A\kappa} + 2(p-s)\frac{A'}{A^2} + k(p-s)^2\left(\frac{A'}{A^2}-\frac{1}{A\kappa}\right)\right\}\sum_j\frac{1}{\rho_j} \\ - \left\{2(p-s)\frac{1}{A\kappa} + k(p-s)^2\left(\frac{1}{A\kappa}-\frac{1}{A'\kappa^2}\right)\right\}\sum_j\sqrt{\rho_j}.$$

Now let β_{W22} and β_{L22} be the powers of T_{W22} and T_{L22} with a level of significance α . Then from (4.22) and (4.23) it is possible to obtain asymptotic expansions for β_{W22} and β_{L22} . However, it is difficult to obtain their difference in a simple form.

4.5. Test of H_2 when the concentration parameters are unknown

Here we consider testing hypothesis H_2 against K_2 when κ_j 's are unknown. The MLE $\hat{\kappa}_j$ of κ satisfies

$$A_p(\hat{\kappa}_j) = \|\bar{\mathbf{x}}_j\|. \quad (4.24)$$

Further let $\tilde{\kappa}_j$ satisfy

$$A_p(\tilde{\kappa}_j) = \|B'_0\bar{\mathbf{x}}_j\|, \quad (4.25)$$

and this is the MLE of κ_j under the hypothesis H_2 . Then two statistics T_{W23} and T_{L23} are given as follows:

$$(i) \quad T_{W23} = \sum_j \frac{n_j \hat{\kappa}_j}{A_p(\hat{\kappa}_j)} \|(I_p - B_0 B_0') \bar{x}\|^2,$$

$$(ii) \quad T_{L23} = 2n \sum_j \{ \rho_j \log a_p(\tilde{\kappa}_j) - \rho_j \log a_p(\hat{\kappa}_j) - \rho_j \tilde{\kappa}_j \|B_0' \bar{x}_j\| + \rho_j \hat{\kappa}_j \|\bar{x}_j\| \}.$$

The statistic T_{W23} has been proposed by Watson [1983a] and T_{L23} is LR statistic. T_{W23} and T_{L23} are expanded in terms of Y under the sequence of alternatives as

$$T_{W23} = q_0(Y) + \frac{1}{\sqrt{n}} q_1^{**}(Y) + \frac{1}{n} q_2''(Y) + O_p(n^{-\frac{3}{2}}), \tag{4.26}$$

$$T_{L23} = q_0(Y) + \frac{1}{\sqrt{n}} q_1^{**}(Y) + \frac{1}{n} q_2^{**}(Y) + O_p(n^{-\frac{3}{2}}), \tag{4.27}$$

where

$$\begin{aligned} q_1^{**}(Y) &= q_1(Y) + \sum_j \frac{1}{\sqrt{\rho_j A_j' \kappa_j}} \mu'_{0j} y_j q_{00j}(y_j) \\ &= \sum_j \frac{1}{\sqrt{\rho_j}} \left(\frac{1}{\sqrt{A_j' \kappa_j}} - \frac{\sqrt{A_j'}}{A_j} \right) \mu'_{0j} y_j q_{00j}(y_j), \end{aligned}$$

$$\begin{aligned} q_2^{**}(Y) &= q_2(Y) + \sum_j \frac{1}{\rho_j 4 A_j' \kappa_j^2} q_{21j}(y_j) q_{00j}(y_j) \\ &\quad - \sum_j \frac{1}{\rho_j} \left(\frac{1}{A_j \kappa_j} + \frac{1}{2 A_j'^2 \kappa_j} \right) (\mu'_{0j} y_j)^2 q_{00j}(y_j) \\ &= \sum_j \frac{1}{\rho_j} \left(\frac{1}{4 A_j' \kappa_j^2} - \frac{1}{4 A_j \kappa_j} \right) q_{21j}(y_j) q_{00j}(y_j) \\ &\quad + \sum_j \frac{1}{\rho_j} \left(\frac{A_j'}{A_j^2} - \frac{1}{A_j \kappa_j} - \frac{A_j''}{2 A_j'^2 \kappa_j} \right) (\mu'_{0j} y_j)^2 q_{00j}(y_j), \end{aligned}$$

$$q_2''(Y) = \sum_j \frac{1}{\rho_j} \left\{ \left(\frac{1}{2 A_j' \kappa_j^2} - \frac{1}{2 A_j \kappa_j} \right) q_{21j}(y_j) \right.$$

$$+ \left(\frac{A'_j}{A_j^2} - \frac{1}{A_j \kappa_j} - \frac{A''_j}{2A_j'^2 \kappa_j} \right) (\mu'_{0j} \mathbf{y}_j)^2 \left. \right\} q_{00j}(\mathbf{y}_j),$$

$q_0(Y)$, $q_{00j}(\mathbf{y}_j)$, $q_{01j}(\mathbf{y}_j)$ and $q_{21j}(\mathbf{y}_j)$ are the same as in (4.5). Similarly, the characteristic functions of T_{W23} and T_{L23} , $\Psi_{W23}(t)$ and $\Psi_{L23}(t)$ are given by

$$\begin{aligned} \Psi_{W23}(t) &= (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right) \\ &\times \left[1 + \frac{1}{4n} \sum_{m=0}^4 d_{3m} (1 - 2it)^{-m} + \mathcal{O}(n^{-\frac{3}{2}}) \right], \end{aligned}$$

and

$$\begin{aligned} \Psi_{L23}(t) &= (1 - 2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1 - 2it} \lambda\right) \\ &\times \left[1 + \frac{1}{4n} \sum_{m=0}^3 h_{3m} (1 - 2it)^{-m} + \mathcal{O}(n^{-\frac{3}{2}}) \right], \end{aligned}$$

respectively, where

$$\begin{aligned} d_{30} &= 2\lambda^{(21)} + 6\lambda^{(22)} + (p-s)(p+3s-6)\rho^{(2)} + (p-s)\rho^{(3)}, \\ d_{31} &= -4\lambda^{(21)} - 4\lambda^{(22)} + 4(s-2)\lambda^{(12)} + 2\lambda^{(13)} - 4(p-s)(s-2)\rho^{(2)} - (p-s)\rho^{(3)}, \\ d_{32} &= 2\lambda^{(21)} - 4\lambda^{(22)} + 2\lambda^{(23)} + 2(p-3s+6)\lambda^{(12)} - 2\lambda^{(13)} - (p-s)(p-s+2)\rho^{(2)}, \\ d_{33} &= 4\lambda^{(22)} - 4\lambda^{(23)} - 2(p-s+2)\lambda^{(12)}, \\ d_{34} &= -2\lambda^{(22)} + 2\lambda^{(23)}, \\ \lambda^{(23)} &= \sum_j A_j^2 / (A'_j \rho_j), \\ \lambda^{(12)} &= \sum_j \{1 - A_j / (A'_j \kappa_j)\} \lambda_j / \rho_j, \quad \lambda^{(13)} = \sum_j A_j A''_j \lambda_j / (A_j'^2 \rho_j), \\ \rho^{(2)} &= \sum_j \{1/A_j - 1/(A'_j \kappa_j)\} / (2\kappa_j \rho_j), \quad \rho^{(3)} = \sum_j A''_j / (A_j'^2 \kappa_j \rho_j), \end{aligned} \tag{4.28}$$

$\lambda^{(21)}$ and $\lambda^{(22)}$ are the same as in (4.8) and

$$\begin{aligned} h_{30} &= 2\lambda^{(21)} + 6\lambda^{(22)} + (p-s)(p+s-4)\rho^{(2)} + (p-s)\rho^{(3)}, \\ h_{31} &= -4\lambda^{(21)} - 4\lambda^{(22)} + 2(s-3)\lambda^{(12)} + 2\lambda^{(13)} - (p-s)(p+s-4)\rho^{(2)} - (p-s)\rho^{(3)}, \\ h_{32} &= 2\lambda^{(21)} - 4\lambda^{(22)} + 2\lambda^{(23)} - 2(s-3)\lambda^{(12)} - 2\lambda^{(13)}, \\ h_{33} &= 2\lambda^{(22)} - 2\lambda^{(23)}. \end{aligned} \tag{4.29}$$

Inverting these characteristic functions, we have the following theorem.

THEOREM 4.6 *Under a sequence of the alternatives (4.1) the distribution*

functions of T_{W23} and T_{L23} can be asymptotically expanded as

$$P(T_{W23} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) + \frac{1}{4n} \sum_{m=0}^4 d_{3m} P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + O(n^{-\frac{3}{2}}), \quad (4.30)$$

and

$$P(T_{L23} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) + \frac{1}{4n} \sum_{m=0}^3 h_{3m} P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + O(n^{-\frac{3}{2}}), \quad (4.31)$$

where d_{3m} 's and h_{3m} 's are given by (4.28) and (4.29), respectively.

Letting $\delta_j = \mathbf{0}$ in (4.30) and (4.31), we obtain asymptotic expansions of the null distributions of T_{W23} and T_{L23} ,

$$P(T_{W23} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} \sum_{m=0}^2 \tilde{d}_{3m} P(\chi_{k(p-s)+2m}^2 \leq x) + O(n^{-\frac{3}{2}}),$$

where

$$\begin{aligned} \tilde{d}_{30} &= (p-s)(p+3s-6)\rho^{(2)} + (p-s)\rho^{(3)}, \\ \tilde{d}_{31} &= -4(p-s)(s-2)\rho^{(2)} - (p-s)\rho^{(3)}, \\ \tilde{d}_{32} &= -(p-s)(p-s+2)\rho^{(2)}, \end{aligned}$$

and

$$\begin{aligned} P(T_{L23} \leq x) &= P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4n} \{(p-s)(p+s-4)\rho^{(2)} + (p-s)\rho^{(3)}\} \\ &\quad \times \{P(\chi_{k(p-s)}^2 \leq x) - P(\chi_{k(p-s)+2}^2 \leq x)\} + O(n^{-\frac{3}{2}}). \end{aligned}$$

This implies that the Bartlett correction factor for T_{L23} is given by

$$1 + \frac{1}{2nk} \{(p+s-4)\rho^{(2)} + \rho^{(3)}\}.$$

Now let β_{W23} and β_{L23} be the powers of T_{W23} and T_{L23} with a level of significance α . Then from (4.30) and (4.31) it is possible to obtain asymptotic expansions for β_{W23} and β_{L23} . By the same way as in Theorem 4.4, we obtain the following theorem.

THEOREM 4.7 *Under a sequence of the alternatives (4.1) the powers β_{W23} and β_{L23} of T_{W23} and T_{L23} with a level of significance α are coincident up to*

the order n^{-1} . Further, β_{W23} (or β_{L23}) is given by

$$\begin{aligned} \beta_{W23} &= P(\chi_{k(p-s)}^2(\lambda) \geq x_\alpha) \\ &+ \frac{1}{n} [(-\lambda^{(21)} - 3\lambda^{(22)})g_{k(p-s)+2}(x_\alpha; \lambda) \\ &+ \left\{ \lambda^{(21)} - \lambda^{(22)} + \frac{1}{2}(p-s+2)\lambda^{(12)} \right\} g_{k(p-s)+4}(x_\alpha; \lambda) \\ &+ (\lambda^{(22)} - \lambda^{(23)})g_{k(p-s)+6}(x_\alpha; \lambda)] + O(n^{-\frac{3}{2}}). \end{aligned}$$

Theorem 4.7 shows that the differences between the powers of T_{W23} and T_{L23} are very small when n is large and κ_j are unknown.

4.6. Explanatory notes

The differences between β_{W2j} and β_{L2j} can be numerically evaluated by using their asymptotic expansion formulas or doing simulation experiments. However, it will be difficult to state the differences as a simple rule. From the point of practical use, we can say that Watson statistic is calculated more easily while LR statistic has a better chi-square approximation. It may depend on our purpose which statistic we use.

As we mentioned before, H_1 is an extension of the hypothesis $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ where LR test has been studied by Hayakawa [1992]. If we take $c_j = -1$, then H_1 is the same as H_0 . On the other hand, even if we take $s = 1$ in H_2 , H_2 is different from H_0 , and this fact is also clear from our results. For one sample test, the procedures of testing $\mu = \mu_0$ with given μ_0 and $\mu = B_0\zeta$ with $p \times 1$ matrix B_0 are essentially the same.

5. Multi-sample tests for mean directions (II)-highly concentrated case

5.1. Primaries

In this section, we derive basic results on asymptotic expansions of various statistics when each concentration parameter κ_j is large and the sample size n_j is fixed. Most of notations are the same as corresponding ones in previous sections. However, our derivation process is slightly different, so the results are not always corresponding to one another. Moreover, we use similar notations to denote completely different things sometimes. Thus we will attempt to explain our notations as possible as we can even if they are the same in previous sections except in section 2.

Now, it is noted that most of methods to obtain asymptotic expansions under large samples in sections 3 and 4 are basically based on Taylor expansion. While, under a highly concentrated case, in addition to this, the

differential equations (2.12), (2.15) for $A_p(\kappa)$ and $a_p(\kappa)$ play key roles in our derivation process. In fact we often use the expansion for the logarithm of $a_p(\kappa)$, given in (2.14). Moreover we also note that this section heavily depends on Fujikoshi and Watamori [1992].

We consider the problems of testing H_j against $K_j(j = 1, 2)$ given in section 4.1 and the corresponding test statistics defined in sections 4.2 and 4.3. Since we are interesting in the situation where κ_j 's are large and n_j 's are fixed, instead of (4.1), we take a sequence of alternatives

$$\mu_j = (\mu_{0j} + \kappa_j^{-\frac{1}{2}} \delta_j)(1 + 2\kappa_j^{-1} \lambda_j)^{-\frac{1}{2}}, \tag{5.1}$$

where μ_{0j} is equal to $b_j \mu$ for H_1 or $B_0 \zeta_j$ for H_2 , $\lambda_j = \delta_j' \delta_j / 2$ and $\delta_j' \mu_{0j} = 0$.

Choose a $p \times (p - s)$ matrix B_s and an $s \times (s - 1)$ matrix Θ_{1j} such that $[B_0 B_s] \in \mathbf{O}(p)$ and $[\zeta_j \Theta_{1j}] \in \mathbf{O}(s)$. Let

$$\bar{\mathbf{x}}_j = \frac{1}{n_j} \sum_l \mathbf{x}_{jl}, \quad n = \sum_j n_j, \quad \kappa = \sum_j \kappa_j \quad \text{and} \quad \rho_j = \frac{\kappa_j}{\kappa}.$$

We assume that each ρ_j is $\mathbf{O}(1)$ as $\kappa \rightarrow \infty$, i.e., $\kappa_j \rightarrow \infty$ as $\kappa \rightarrow \infty$. Then the identity matrix I_p is decomposed as

$$I_p = B_s B_s' + B_0 \zeta_j \zeta_j' B_0' + B_0 \Theta_{1j} \Theta_{1j}' B_0'. \tag{5.2}$$

For $l = 1, \dots, n_j$ and $j = 1, \dots, k$, let

$$\mathbf{y}_{jl} = \sqrt{\kappa_j} B_s' \left(\mathbf{x}_{jl} - \frac{1}{\sqrt{\kappa_j}} \delta_j \right), \quad u_{jl} = 2\kappa_j (1 - \zeta_j' B_0' \mathbf{x}_{jl}), \quad \mathbf{w}_{jl} = \sqrt{\kappa_j} \Theta_{1j}' B_0' \mathbf{x}_{jl}.$$

These are the standardized statistics in a highly concentrated case. Our test statistics are functions of \mathbf{y}_{jl} and \mathbf{w}_{jl} , so we obtain asymptotic expansions of the distributions of them by extending Fujikoshi and Watamori [1992] to multi-sample case. The characteristic function of \mathbf{y}_{jl} and \mathbf{w}_{jl} is given by

$$\begin{aligned} \Psi(\mathbf{t}_{1j}, \mathbf{t}_{2j}) &= E[\exp(it'_{1j} \mathbf{y}_{jl} + it'_{2j} \mathbf{w}_{jl})] \\ &= \frac{a_p(\bar{\omega}_j)}{a_p(\kappa_j)} \exp(-it'_{1j} B_s' \delta_j), \end{aligned}$$

where

$$\begin{aligned} \bar{\omega}_j &= \kappa_j \left\| \mu_j + \frac{i}{\sqrt{\kappa_j}} (B_s \mathbf{t}_{1j} + B_0 \Theta_{1j} \mathbf{t}_{2j}) \right\| \\ &= \kappa_j \left\{ 1 + \frac{2i}{\sqrt{\kappa_j}} \mu_j' (B_s \mathbf{t}_{1j} + B_0 \Theta_{1j} \mathbf{t}_{2j}) + \frac{i^2}{\kappa_j} (\mathbf{t}'_{1j} \mathbf{t}_{1j} + \mathbf{t}'_{2j} \mathbf{t}_{2j}) \right\}^{\frac{1}{2}} \end{aligned}$$

$$= \kappa_j \left\{ 1 + \frac{1}{\kappa_j} (i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} + i^2 \mathbf{t}'_{2j} \mathbf{t}_{2j} + 2i \mathbf{t}'_{1j} B'_s \delta_j) - \frac{2i \lambda_j}{\kappa_j^2} \mathbf{t}'_{1j} B'_1 \delta_j + \mathcal{O}(\kappa_j^{-3}) \right\}.$$

Using the expansion (2.14) of the logarithm of $a_p(\kappa)$, we obtain

$$\begin{aligned} \log \Psi_j(\mathbf{t}_{1j}, \mathbf{t}_{2j}) &= \frac{1}{2} (i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} + i^2 \mathbf{t}'_{2j} \mathbf{t}_{2j}) \\ &\quad + \frac{1}{\kappa_j} \left\{ -\frac{p-1}{4} (i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} + i^2 \mathbf{t}'_{2j} \mathbf{t}_{2j} + 2i \mathbf{t}'_{1j} B'_s \delta_j) - i \lambda_j \mathbf{t}'_{1j} B'_s \delta_j \right. \\ &\quad \left. - \frac{1}{8} (i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} + i^2 \mathbf{t}'_{2j} \mathbf{t}_{2j})^2 \right\} + \mathcal{O}(\kappa_j^{-2}). \end{aligned}$$

Thus

$$\Psi_j(\mathbf{t}_{1j}, \mathbf{t}_{2j}) = \exp \left\{ \frac{1}{2} (i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} + i^2 \mathbf{t}'_{2j} \mathbf{t}_{2j}) \right\} + \mathcal{O}(\kappa_j^{-1}),$$

and in particular,

$$\begin{aligned} \Psi_j(\mathbf{t}_{1j}, \mathbf{0}) &= \exp \left(\frac{1}{2} i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j} \right) \left[1 - \frac{1}{8 \kappa_j} \{ \beta_j^2 + 4 \beta_j \gamma_j + 4 \gamma_j^2 \right. \\ &\quad \left. + 2(p-1) \beta_j + 4(p-1+2\lambda_j) \gamma_j \} + \mathcal{O}(\kappa_j^{-2}) \right], \end{aligned} \quad (5.3)$$

where

$$\beta_j = i^2 \mathbf{t}'_{1j} \mathbf{t}_{1j}, \quad \gamma_j = i \delta'_j B'_s \mathbf{t}_{1j}.$$

Using (5.3) we obtain the following theorem.

THEOREM 5.1 *Let $\mathbf{y}_j = \sqrt{\kappa_j n_j} B'_s \left(\bar{\mathbf{x}}_j - \frac{1}{\sqrt{\kappa_j}} \delta_j \right)$. Under the assumption of (5.1) the probability density function of $Y = [\mathbf{y}_1, \dots, \mathbf{y}_k]$ and $W = [\mathbf{w}_1, \dots, \mathbf{w}_k]$ are asymptotically independent. Further, the probability density functions of Y and W can be asymptotically expanded as follows, respectively.*

$$f_Y(Y) = \Phi_{n,p-s}(Y) \left[1 + \frac{1}{\kappa} b_1(Y) + \mathcal{O}(\kappa^{-2}) \right], \quad (5.4)$$

and

$$f_W(W) = \Phi_{n,s-1}(W) + \mathcal{O}(\kappa^{-1}), \quad (5.5)$$

where

$$\Phi_{n,p}(Y) = (2\pi)^{-\frac{np}{2}} \exp \left\{ -\frac{1}{2} \text{tr} YY' \right\},$$

$$b_1(Y) = \sum_j \frac{1}{\rho_j} b_{s_j}(\mathbf{y}_j)$$

$$\begin{aligned} b_{s_j}(\mathbf{y}_j) = & -\frac{1}{8n_j} [(\mathbf{y}'_j \mathbf{y}_j)^2 + 4\sqrt{n_j} \delta'_j B_s \mathbf{y}_j \cdot \mathbf{y}'_j \mathbf{y}_j + 2\{n(p-1) - (p-s+2)\} \mathbf{y}'_j \mathbf{y}_j \\ & + 4n_j (\delta'_j B_s \mathbf{y}_j)^2 + 4\{n_j(p-1) - (p-s+2) + 2n_j \lambda_j\} \sqrt{n_j} \delta'_j B_s \mathbf{y}_j \\ & - (p-s)(p+s-4) - 8\lambda_j]. \end{aligned}$$

In order to derive asymptotic distributions of test statistics, we introduce expansions of distributions of \mathbf{y}_{jl} 's and \mathbf{w}_{jl} 's in another form which are also obtained by (5.3).

THEOREM 5.2 *Let $Y_j = [\mathbf{y}_{j1}, \dots, \mathbf{y}_{jn_j}]$ and $W_j = [\mathbf{w}_{j1}, \dots, \mathbf{w}_{jn_j}]$. Under the assumption of (5.1) Y_j and W_j are asymptotically independent. Further, the probability density functions of Y_j and W_j can be asymptotically expanded as follows, respectively.*

$$f_{Y_j}(Y_j) = \Phi_{n_j, p-s}(Y_j) \left[1 + \frac{1}{\kappa} b_{1j}(Y_j) + \mathbf{O}(\kappa^{-2}) \right],$$

and

$$f_{W_j}(W_j) = \Phi_{n_j, s-1}(W_j) + \mathbf{O}(\kappa^{-1}),$$

where

$$\begin{aligned} b_{1j}(Y_j) = & \frac{1}{8\rho_j} \left[-\sum_l (\text{tr} \mathbf{e}_l \mathbf{e}'_l Y'_j Y_j)^2 - 4\sum_l (\text{tr} \mathbf{e}_l \delta'_j B_s Y_j) (\text{tr} \mathbf{e}_l \mathbf{e}'_l Y'_j Y_j) \right. \\ & - 2(s-3) \text{tr} Y_j Y'_j - 4\sum_l (\text{tr} \mathbf{e}_l \delta'_j B_s Y_j)^2 \\ & \left. - 4(s-3+2\lambda_j) \sum_l \text{tr} \mathbf{e}_l \delta'_j B_s Y_j + 8n_j \lambda_j + n_j(p-s)(p+s-4) \right] \end{aligned} \quad (5.6)$$

and \mathbf{e}_l is the n -dimensional column vector with l th element 1 and others 0.

Note that from the decomposition(5.2) of I_p , each \mathbf{x}_{jl} can be decomposed as

$$\begin{aligned} \mathbf{x}_{jl} = & B_s B'_s \mathbf{x}_{jl} + B_0 \Theta_{1j} \Theta'_{1j} B'_0 \mathbf{x}_{jl} + B_0 \zeta_j \zeta'_j B'_0 \mathbf{x}_{jl} \\ = & \frac{1}{\sqrt{\kappa \rho_j}} B_s (\mathbf{y}_{jl} + B'_s \delta_j) + \frac{1}{\sqrt{\kappa \rho_j}} B_0 \Theta_{1j} \mathbf{w}_{jl} + \frac{1}{2\kappa \rho_j} B_0 \zeta (1 - u_{jl}). \end{aligned}$$

Since $\|\mathbf{x}_{j\ell}\| = 1$, by taking the norms of both sides, we have

$$\begin{aligned} u_{j\ell} &= (\mathbf{y}_{j\ell} + B'_s \boldsymbol{\delta}_j)' (\mathbf{y}_{j\ell} + B'_s \boldsymbol{\delta}_j) + \mathbf{w}'_{j\ell} \mathbf{w}_{j\ell} \\ &+ \frac{1}{4\kappa\rho_j} \{(\mathbf{y}_{j\ell} + B'_s \boldsymbol{\delta}_j)' (\mathbf{y}_{j\ell} + B'_s \boldsymbol{\delta}_j) + \mathbf{w}'_{j\ell} \mathbf{w}_{j\ell}\}^2 + \mathcal{O}_p(\kappa_j^{-2}). \end{aligned} \quad (5.7)$$

5.2. Test of H_2 when the concentration parameters are known

Here we consider testing hypothesis H_2 against K_2 when κ_j 's are known. Then two statistics, Watson statistic T_{W_2} and LR statistic T_{L_2} are given by

$$\begin{aligned} \text{(i)} \quad T_{W_2} &= \kappa \sum_j \rho_j \|(I_p - B_0 B'_0) \bar{\mathbf{x}}_j\|^2, \\ \text{(ii)} \quad T_{L_2} &= 2\kappa \sum_j \rho_j n_j (\|\bar{\mathbf{x}}_j\| - \|B'_0 \bar{\mathbf{x}}_j\|). \end{aligned}$$

T_{W_2} was proposed by Watson [1983a]. T_{L_2} is the same as $T_{L_{21}}$ but the coefficient of T_{W_2} is slightly different from that of $T_{W_{21}}$. Under a sequence of alternatives (5.1), T_{W_2} and T_{L_2} are expanded in terms of Y_j and W_j as

$$\begin{aligned} T_{W_2} &= \sum_j q_{0j}(Y_j) + \mathcal{O}_p(\kappa^{-2}), \\ T_{L_2} &= \sum_j q_{0j}(Y_j) + \frac{1}{\kappa} \sum_j \tilde{q}_{1j}(Y_j^*, W_j) + \mathcal{O}(\kappa^{-2}), \end{aligned}$$

where

$$\begin{aligned} q_{0j}(Y_j) &= \text{tr}(Y_j + B'_s \boldsymbol{\delta}_j \mathbf{1}'_{n_j}) P_{2j} (Y_j + B'_s \boldsymbol{\delta}_j \mathbf{1}'_{n_j})' = \text{tr} Y_j^* P_{2j} Y_j^{*'}, \\ \tilde{q}_{1j}(Y_j^*, W_j) &= \frac{1}{\rho_j} \left\{ -\frac{1}{4n_j} (\text{tr} Y_j^* P_{2j} Y_j^{*'})^2 \right. \\ &\quad \left. + \frac{1}{2n_j} \text{tr} Y_j^* P_{2j} Y_j^{*'} (\text{tr} Y_j^* Y_j^{*'} + \text{tr} W_j P_{1j} W_j') \right\}, \\ P_{1j} &= I_{n_j} - \frac{1}{n_j} \mathbf{1}_{n_j} \mathbf{1}'_{n_j}, \quad P_{2j} = \frac{1}{n_j} \mathbf{1}_{n_j} \mathbf{1}'_{n_j}, \quad Y_j^* = Y_j + B'_s \boldsymbol{\delta}_j \mathbf{1}'_{n_j}, \end{aligned}$$

and $\mathbf{1}_n$ is the n -dimensional column vector with all elements 1. Thus the characteristic function of T_{W_2} , $\Psi_{W_2}(t)$ is given by

$$\begin{aligned} \Psi_{W_2}(t) &= \int \cdots \int \prod_j [\exp\{itq_{0j}(Y_j)\} \cdot \Phi_{n_j, p-s}(Y_j)] \\ &\quad \times \left\{ 1 + \frac{1}{\kappa} \sum_j b_{1j}(Y_j) + \mathcal{O}(\kappa^{-2}) \right\} \prod_j dY_j. \end{aligned}$$

Consider the transformation

$$Z_j = \left(Y_j - \frac{2it}{1-2it} B'_s \delta_j \mathbf{1}'_{n_j} \right) A_j^{-\frac{1}{2}}, \quad (5.8)$$

where

$$A_j = P_{1j} + \frac{1}{1-2it} P_{2j}.$$

Then we obtain

$$\begin{aligned} \Psi_{W_2}(t) &= (1-2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1-2it} \lambda\right) \\ &\times \left\{ 1 + \sum_j E_{Z_j} \left[\frac{1}{\kappa} b_{1j} \left(Z_j A_j^{\frac{1}{2}} + \frac{2it}{1-2it} B'_s \delta_j \mathbf{1}'_{n_j} \right) + \mathcal{O}(\kappa^{-2}) \right] \right\}, \end{aligned} \quad (5.9)$$

where

$$\lambda = \sum_j n_j \lambda_j.$$

The expectation with respect to Z_j in (5.9) is taken under the normal random matrix Z_j whose elements are independently distributed as $N(0, 1)$. After calculating the expectation in (5.9), we obtain

$$\begin{aligned} \Psi_{W_2}(t) &= (1-2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1-2it} \lambda\right) \\ &\times \left[1 + \frac{1}{4\kappa} \sum_{m=0}^4 d_m (1-2it)^{-m} + \mathcal{O}(\kappa^{-2}) \right], \end{aligned} \quad (5.10)$$

where

$$\begin{aligned} d_0 &= 6\lambda^{(2)} + 2(p-1)\lambda^{(1)} - 2(p-s)\lambda^* + (p-s)(p-1)\rho - (p-s)(p-s+2)\rho^*, \\ d_1 &= -8\lambda^{(2)} + 2(p-s)\lambda^* - (p-s)(p-1)\rho + 2(p-s)(p-s+2)\rho^*, \\ d_2 &= 4\lambda^{(2)} - 2(p-1)\lambda^{(1)} + 2(p-s+2)\lambda^* - (p-s)(p-s+2)\rho^*, \\ d_3 &= -2(p-s+2)\lambda^*, \quad d_4 = -2\lambda^{(2)}, \\ \lambda^* &= \sum_j (\lambda_j / \rho_j), \quad \lambda^{(1)} = \sum_j (n_j \lambda_j / \rho_j), \quad \lambda^{(2)} = \sum_j (n_j \lambda_j^2 / \rho_j), \\ \rho &= \sum_j (1 / \rho_j), \quad \rho^* = \sum_j 1 / (2n_j \rho_j). \end{aligned} \quad (5.11)$$

Inverting this characteristic function, we have the following theorem.

THEOREM 5.3 *Under a sequence of the alternatives (5.1) the distribution*

function of T_{W_2} can be asymptotically expanded as

$$P(T_{W_2} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) \quad (5.12)$$

$$+ \frac{1}{4\kappa} \sum_{m=0}^4 d_m P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + \mathbf{O}(\kappa^{-2}),$$

where d_m 's are given by (5.11), and $\chi_f^2(\lambda)$ denotes a noncentral χ^2 -variate with f degrees of freedom and noncentrality parameter λ .

Letting $\delta = \mathbf{0}$ in (5.12), we obtain an asymptotic expansion of the null distribution of T_{W_2} ,

$$P(T_{W_2} \leq x) = P(\chi_{k(p-s)}^2 \leq x) + \frac{1}{4\kappa} \sum_{m=0}^2 \tilde{d}_m P(\chi_{k(p-s)+2j}^2 \leq x) + \mathbf{O}(\kappa^{-2}),$$

where

$$\begin{aligned} \tilde{d}_0 &= (p-s)(p-1)\rho - (p-s)(p-s+2)\rho^*, \\ \tilde{d}_1 &= -(p-s)(p-1)\rho + (p-s)(p-s+2)\rho^*, \\ \tilde{d}_2 &= -(p-s)(p-s+2)\rho^*. \end{aligned}$$

Next we consider the distribution of T_{L_2} under (5.1). Then the characteristic function of T_{L_2} , $\Psi_{L_2}(t)$ is given by

$$\begin{aligned} \Psi_{L_2}(t) &= \int \cdots \int \prod_j [\exp\{itq_{0j}(Y_j)\} \cdot \Phi_{n_j, p-s}(Y_j)] \\ &\quad \times \left[1 + \frac{1}{\kappa} \sum_j \{itq_{1j}(Y_j^*) + b_{1j}(Y_j)\} + \mathbf{O}(\kappa^{-2}) \right] \prod_j dY_j. \end{aligned}$$

where

$$\begin{aligned} q_{1j}(Y^*) &= -\frac{1}{4n_j\rho_j} (\text{tr } Y_j^* P_{2j} Y_j^{*\prime})^2 \\ &\quad + \frac{1}{2n_j\rho_j} \text{tr } Y_j^* P_{2j} Y_j^{*\prime} \{ \text{tr } Y_j^* Y_j^{*\prime} + (s-1)(n-1) \}. \end{aligned}$$

Considering the transformation (5.8), we obtain

$$\begin{aligned} \Psi_{L_2}(t) &= (1-2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1-2it}\lambda\right) \quad (5.13) \\ &\quad \times \left\{ 1 + \sum_j E_{Z_j} \left[\frac{1}{\kappa} \left\{ itq_{1j} \left(Z_j A_j^{\frac{1}{2}} + \frac{1}{1-2it} B'_s \delta_j \mathbf{1}'_{n_j} \right) \right. \right. \right. \end{aligned}$$

$$+ b_1 \left(Z_j A_j^{\frac{1}{2}} + \frac{2it}{1-2it} B'_s \delta_j \mathbf{1}_{n_j} \right) \Big\} + \mathcal{O}(\kappa^{-2}) \Big\}.$$

After calculating the expectation in (5.13), we obtain

$$\begin{aligned} \Psi_{L_2}(t) &= (1-2it)^{-\frac{k(p-s)}{2}} \exp\left(\frac{2it}{1-2it} \lambda\right) \\ &\times \left[1 + \frac{1}{4\kappa} \sum_{m=0}^3 h_m (1-2it)^{-m} + \mathcal{O}(\kappa^{-2}) \right], \end{aligned} \tag{5.14}$$

where

$$\begin{aligned} h_0 &= 6\lambda^{(2)} + 2(p-1)\lambda^{(1)} - 2(p-s)\lambda^* + (p-s)(p+s-4)\rho^*, \\ h_1 &= -8\lambda^{(2)} - 2(p-1)\lambda^{(1)} + 2(2p-s-1)\lambda^* - (p-s)(p+s-4)\rho^*, \\ h_2 &= 4\lambda^{(2)} - 2(p-1)\lambda^*, \quad h_3 = -2\lambda^{(2)}. \end{aligned} \tag{5.15}$$

Inverting this characteristic function, we have the following theorem.

THEOREM 5.4 *Under a sequence of the alternatives (5.1) the distribution function of T_{L_2} can be asymptotically expanded as*

$$P(T_{L_2} \leq x) = P(\chi_{k(p-s)}^2(\lambda) \leq x) + \frac{1}{4\kappa} \sum_{m=0}^3 h_m P(\chi_{k(p-s)+2m}^2(\lambda) \leq x) + \mathcal{O}(\kappa^{-2}), \tag{5.16}$$

where h_m 's are given by (5.15).

Letting $\delta = \mathbf{0}$ in (5.16), we obtain an asymptotic expansion of the null distribution of T_{L_2} ,

$$\begin{aligned} P(T_{L_2} \leq x) &= P(\chi_{k(p-s)}^2 \leq x) \\ &+ \frac{1}{4\rho^*k\kappa} (p-s)(p+s-4) \{P(\chi_{k(p-s)}^2 \leq x) - P(\chi_{k(p-s)-2}^2 \leq x)\} \\ &+ \mathcal{O}(\kappa^{-2}). \end{aligned}$$

This implies that the Bartlett adjustment factor is given by

$$1 + \frac{1}{2\rho^*k\kappa} (p+s-4),$$

and

$$P\left(\left\{1 + \frac{1}{2\rho^*k\kappa} (p+s-4)\right\} T_{L_1} \leq x\right) = P(\chi_{k(p-s)}^2 \leq x) + \mathcal{O}(\kappa^{-2}).$$

Let β_{W_2} and β_{L_2} be the powers of T_{W_2} and T_{L_2} with a level of significance α . From (5.12) and (5.16) it is possible to obtain asymptotic expansions for β_{W_2} and β_{L_2} as the same way as before. Then we obtain the following theorem.

THEOREM 5.5 *Under a sequence of the alternatives (5.1) the powers β_{W_2} and β_{L_2} of T_{W_2} and T_{L_2} with a level of significance α are coincident up to the order κ^{-1} . Further, β_{W_2} (or β_{L_2}) is given by*

$$\begin{aligned} \beta_{W_2} = & P(\chi_{k(p-s)}^2(\lambda) \geq x_\alpha) \\ & - \frac{1}{\kappa} [\{3\lambda^{(2)} + (p-1)\lambda^{(1)} - (p-s)\lambda^*\} g_{k(p-s)+2}(x_\alpha; \lambda) \\ & - \{\lambda^{(2)} - (p-s+2)\lambda^*\} g_{k(p-s)+4}(x_\alpha; \lambda) \\ & + \lambda^{(2)} g_{k(p-s)+6}(x_\alpha; \lambda)] + O(\kappa^{-2}), \end{aligned}$$

where x_α is the upper α point of $\chi_{k(p-s)}^2$ and $g_f(x_\alpha; \lambda)$ is the probability density function of $\chi_f^2(\lambda)$.

Theorem 5.5 shows that the differences between the powers of T_{W_2} and T_{L_2} are very small when κ is large and n is fixed.

5.3. Explanatory notes

The LR statistic T_{L_1} for testing hypothesis H_1 against K_1 is given as follows:

$$T_{L_1} = 2\kappa(\sum_j \rho_j n_j \|\bar{\mathbf{x}}_j\| - \|\sum_j \rho_j n_j b_j \bar{\mathbf{x}}_j\|).$$

Although we have not obtained its asymptotic expansion yet, it will be derived by using Theorem 5.1. Since H_1 corresponds to H_2 when $s=1$, we need only the distributions of y_j and u_j . The detail is also left as a future study. Finally we note that the results in section 5 are multi-sample versions of Fujikoshi and Watamori [1992] which can be obtained in a parallel discussion.

6. Concluding remarks

In previous sections we derive asymptotic properties of some estimators and test statistics in large sample or highly concentrated case. Most of the results are obtained in the forms of asymptotic expansions. However, it is noted that in these discussions, the moment relationship (2.1) does not so influence the results, and the speciality that the sample space is a sphere does not appear explicitly.

As we mentioned before in first section, we regard a random vector \mathbf{x} on \mathcal{S}^{p-1} as a random vector in \mathcal{R}^p with unit norm $\|\mathbf{x}\| = 1$. By this embedding approach, we can apply many statistical theories in Euclidean space to statistical distributions on a sphere. However, the dimension of \mathbf{x} is not p but $p-1$, and this influences the structure of distributions on \mathcal{S}^{p-1} . Moreover, in \mathcal{R}^p , we do not have to distinguish a mean (or an expectation) from a barycenter. This means that we can use a mean as a representative point for a data set. On the other hand, when $\mathbf{x}_1, \dots, \mathbf{x}_n$ have norm 1, $\|\sum \mathbf{x}_j/n\| < 1$ unless all \mathbf{x}_j 's are the same point, *i.e.*, the sample mean is no longer a representative point on a sphere. Then we should think of what the notion of 'mean' is.

Anyway, we may not be able to take account of nature of the space entirely by the embedding approach. Therefore we need to consider another approach and introduce a new notion of 'mean point.' Watamori and Kakimizu [1994] have attempted to discuss these topics, though they consider only about a circle. Some other measures of location have also been proposed (see, *e.g.*, Lenth [1981] for robust one).

In practical situation, we often need quasi-random vectors on \mathcal{S}^{p-1} , and for computer generations of uniform or Langevin random vectors, there are some papers (see *e.g.*, Ulrich [1984] and Wood [1987]). By recent development of computers, some methods such as bootstrap have attracted many statisticians in various area. Fisher and Hall [1989] have applied bootstrap method for directional data and obtained confidence regions for mean direction.

A correlation coefficient is another example that we need new notions for directional data (or we need to extend notions in a Euclidean space to a sphere). Since $\mathcal{S}^p \times \mathcal{S}^q \sim \mathcal{S}^{p+q}$ in general, we can not treat $\mathbf{x}_1 \in \mathcal{S}^p$, $\mathbf{x}_2 \in \mathcal{S}^q$ in the same way as $(\mathbf{x}'_1, \mathbf{x}'_2)' \in \mathcal{S}^{p+q}$. Fisher and Lee [1983] have defined a correlation coefficient for circular data. Stephens [1979] considered it by embedding approach.

Some other statistical methods have been considered for directional data. These are nonparametric methods (see, *e.g.*, Rao [1984]), time series (see, *e.g.*, Breckling [1989]) and so on. Bagch and Kadane [1991] have applied Laplace approximations on the stance of Bayesian analysis. On the other hand, Mardia [1989] has shown that directional techniques are applicable to shape analysis. As we have seen above, it seems that there are many problems to be solved in this area.

Acknowledgement

The author wishes to thank Professor Y. Fujikoshi for his constant help

and useful comments.

References

- [1] P. Bagchi and J. B. Kadane, Laplace approximations to posterior moments and marginal distributions on circles, spheres, and cylinders, *Canadian J. of Statist.*, **19** (1991), 67–77.
- [2] R. Bartels, Estimation in a bidirectional mixture of von Mises distributions, *Biometrics*, **40** (1984), 777–784.
- [3] D. J. Best and N. I. Fisher, The bias of the maximum likelihood estimators of the von Mises-Fisher concentration parameters, *Commun. Statist.-Simula. Computa.*, **B10** (1981), 493–502.
- [4] M. S. Bingham and K. V. Mardia, Maximum likelihood characterization of the von Mises distribution, *Statistical Distribution in Scientific Work*, Vol. 3, (G. P. Patil *et al.* eds.), 1975, 387–398.
- [5] J. Breckling, *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, Springer-Verlag, *Lecture Note in Statistics*, **61** (1989).
- [6] R. Chou, Small sample theory of the Langevin distribution, *Austral. J. of Statist.*, **28** (1986), 335–344.
- [7] D. Collett and T. Lewis, Discriminating between the von Mises and wrapped normal distributions, *Austral. J. of Statist.*, **23** (1981), 73–79.
- [8] G. R. Ducharme and P. Milasevic, Spatial median and directional data, *Biometrika*, **74** (1987), 212–215.
- [9] G. R. Ducharme and P. Milasevic, Estimating the concentration of the Langevin distribution, *Canadian J. of Statist.*, **18** (1990), 163–169.
- [10] N. I. Fisher, *Statistical Analysis of Circular Data*, Cambridge University Press, 1993.
- [11] N. I. Fisher, T. Lewis and B. J. J. Embelton, *Statistical Analysis of Spherical Data*, Cambridge University Press, London, 1987.
- [12] N. I. Fisher and P. Hall, Bootstrap confidence regions for directional data, *J. Am. Statist. Assoc.*, **84** (1989), 996–1002.
- [13] N. I. Fisher and P. Hall, Bootstrap methods for directional data, *The Art of Statistical Science: A Tribute to G. S. Watson*, (K. V. Mardia eds.), Chichester: Wiley, 1992.
- [14] N. I. Fisher and A. J. Lee, A correlation coefficient for circular data, *Biometrika*, **70** (1983), 327–332.
- [15] N. I. Fisher and T. Lewis, Estimating the common mean direction of several circular or spherical distributions with differing dispersions, *Biometrika*, **70** (1983), 333–341.
- [16] R. A. Fisher, Dispersion on a sphere, *Proc. Roy. Soc.*, **A217** (1953), 295–305.
- [17] Y. Fujikoshi, Comparison of powers of a class of tests for multivariate linear hypothesis and independence, *J. Multivariate. Anal.*, **26** (1988), 48–58.
- [18] Y. Fujikoshi and Y. Watamori, Tests for the mean direction of the Langevin distribution with larg concentration parameter, *J. Multivariate Analysis*, **42** (1992), 210–225.
- [19] T. Hayakawa, On tests for the mean direction of the Langevin Distribution, *Ann. Inst. Statist. Math.*, **42** (1990), 359–373.
- [20] T. Hayakawa, Homogeneity of parameter of Langevin populations, *Research Memorandum*, **428** (1992).
- [21] P. E. Jupp and K. V. Mardia, A unified view of the theory of directional statistics, 1975–1988, *Internat. Statist. Rev.*, **57** (1989), 261–294.
- [22] R. V. Lenth, Robust measures of location for directional data, *Technometrics*, **23** (1981), 77–81.

- [23] K. V. Mardia, *Statistics of Directional Data*, Academic Press: London, 1972.
- [24] K. V. Mardia, *Statistics of directional data*, *J. Roy. Statist. Soc.*, **36** (1975a), 349–393.
- [25] K. V. Mardia, *Distribution theory for the von Mises-Fisher distribution and its application, Statistical Distribution in Scientific Work*, Vol. 1, (G. P. Patil *et al.* eds.), 1975b, 113–130.
- [26] K. V. Mardia, *Characterizations of directional distributions*, *ibid.*, Vol. 3, 1975c, 365–385.
- [27] K. V. Mardia, *Shape analysis of triangles through directional techniques*, *J. R. Statist. Soc. B*, **51** (1989), 449–458.
- [28] K. V. Mardia, J. T. Kent and J. M. Bibby, *Multivariate Analysis*, Academic Press: London, 1979.
- [29] J. S. Rao, *Nonparametric methods in directional data analysis*, *Handbook of Statistics*, Vol. 4, (P. R. Krishnaiah and P. K. Sen eds.), 1984, 755–770.
- [30] E. Ronchetti, *Optimal robust estimators for the concentration parameter of a von Mises-Fisher distribution*, *The Art of Statistical Science: A Tribute to G. S. Watson*, (K. V. Mardia eds.), Chichester: Wiley, 1992, 65–74.
- [31] B. D. Spurr and M. A. Koutbeiy, *A comparison of various methods for estimating the parameters in mixtures of von Mises distributions*, *Commun. Statist.-Simula.*, **20** (1991), 725–741.
- [32] M. A. Stephens, *Multi-sample tests for the Fisher distribution for directions*, *Biometrika*, **56** (1969), 169–181.
- [33] M. A. Stephens, *Multi-sample tests for the von Mises distribution*, *J. Am. Statist. Ass.*, **67** (1972), 456–461.
- [34] M. A. Stephens, *Vector correlation*, *Biometrika*, **66** (1979), 41–48.
- [35] M. A. Stephens, *Use of the von Mises distribution to analyse continuous proportions*, *Biometrika*, **69** (1982), 197–203.
- [36] G. Schou, *Estimation of the concentration parameter in von Mises-Fisher distributions*, *Biometrika*, **65** (1976), 369–377.
- [37] G. Ulrich, *Computer generation of distributions on the m -sphere*, *Appl. Statist.*, **33** (1984), 158–163.
- [38] R. Von Mises, *Über die “Ganzzahligkeit” der Atomgewicht und verwandte Fragen*, *Physikal. Z.*, **19** (1918), 490–500.
- [39] Y. Watamori, *Tests for a given linear structure of the mean direction of the Langevin distribution*, *Ann. Inst. Statist. Math.*, **44** (1992), 147–156.
- [40] Y. Watamori and O. Kakimizu, *On a geometric approach to distributions on a circle*, *Hiroshima Univ. Technical Report*, 1994, 94–14.
- [41] G. S. Watson and E. J. Williams, *On the construction of significance tests on the circle and the sphere*, *Biometrika*, **43** (1956), 344–352.
- [42] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, London, 1980.
- [43] G. S. Watson, *Statistics on Spheres*, Wiley: New York, 1983a.
- [44] G. S. Watson, *Large sample theory of the Langevin distribution*, *J. Statist. Plann. Inf.*, **8** (1983b), 245–256.
- [45] G. S. Watson, *Large sample theory for distributions on the hypersphere with rotational symmetries*, *Ann. Inst. Statist. Math.*, **35** (1983c), 303–319.
- [46] G. S. Watson, *The theory of concentrated Langevin distributions*, *J. Multi. Anal.*, **14** (1984), 74–82.
- [47] G. S. Watson, *Some estimation theory on the sphere*, *Ann. Inst. Statist. Math.*, **38** (1986), 263–275.
- [48] G. S. Watson, *The Langevin distribution on high-dimensional spheres*, *J. Appl. Statist.*,

15 (1988), 123–130.

- [49] A. Wood, The simulation of spherical distributions in the Fisher-Bingham family, *Comm. Statist.*, **B16** (1987), 885–898.

*Department of Mathematics
Faculty of Science
Hiroshima University
Higashi-Hiroshima, 739, Japan*