JOURNAL OF MULTIVARIATE ANALYSIS **59**, 141–152 (1996)
ARTICLE NO. 0057

Asymptotic Behavior of Sample Mean Direction for Spheres

Harrie Hendriks*

Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

Zinoviy Landsman

University of Haifa, Haifa, Israel

and

Texas Tech University

In this note we consider some asymptotic properties of empirical mean direction on spheres. We do not require any symmetry for the underlying density. Thus our results provide the framework for an asymptotic inference regarding mean direction under very weak model assumptions. Mean direction is a specialization of the more general concept of mean location applicable to arbitrary (compact) submanifolds of Euclidean space, to which the methods of this paper could be applied.

1. INTRODUCTION

In this paper we derive some asymptotic results for the empirical mean direction on spheres in an intrinsic manner that allow for generalization to more complicated manifolds. Moreover, we do not impose any symmetry condition on the underlying density. Thus our results provide the essential ingredients for estimating, hypothesis testing, and constructing confidence intervals without any symmetry restriction.

Received January 25, 1994; revised December 14, 1994.

AMS 1980 subject classifications: 62G05; 62H99.

Key words and phrases: mean direction, mean location, asymptotic behavior, sphere, sample, mean direction, sample mean location.

* To whom correspondence should be addressed at: Vakgroep Wiskunde, Katholieke Universiteit Nijmegen, Postbus 9010, 6500 GL Nijmegen, The Netherlands.

0047-259X/96 $18.00

The concept of mean direction has been studied before (Mardia, 1975; Watson, 1983). Consider the unit sphere $\mathbf{S}^{k-1}$ in $\mathbb{R}^k$, and let X be a random unit vector in $\mathbb{R}^k$. Let $\mathbf{E}X$ be the ordinary population mean or expectation vector; then we have the mean direction of X,

$$\mu = \frac{\mathbf{E}X}{\|\mathbf{E}X\|}, \tag{1}$$

provided that $\mathbf{E}X \neq 0$. Let be given n independent copies $X_1, ..., X_n$ of X. Let $\bar{X}_n = (1/n)(X_1 + \cdots + X_n)$ be the ordinary sample mean vector; then we have

$$S_n = \frac{\bar{X}_n}{\|\bar{X}_n\|}, \tag{2}$$

provided that $\bar{X}_n \neq 0$, so that indeed S_n is the sample mean direction.

We study the asymptotic behavior of the empirical mean direction under much weaker conditions than, e.g., those in Watson (1983). In particular, the assumption of rotational symmetry (around some axis) for the underlying density simplifies the mathematics considerably, but may often be quite unrealistic. In Watson (1983) this assumption leads to a limiting multivariate normal distribution with equal variances in all tangential directions. Without this symmetry assumption the variances may depend on the tangential direction.

From an intrinsic point of view it seems natural to define a limiting distribution for the empirical mean direction in the tangent space to the sphere at the point corresponding to the population mean direction. Approximations to the probabilities of actual interest, viz. the probability of finding the empirical mean direction in shrinking subsets of the sphere concentrated around the true mean direction, can be obtained from the limiting distribution on the tangent space via the exponential mapping (or even inverse projection for spheres). In a nonasymptotic setting the use of the exponential mapping to construct distributions on manifolds from given distributions on a tangent space was suggested in Jupp *et al.* (1989).

In Section 2 it will be shown that S_n is an asymptotically unbiased estimator with order of bias $O(n^{-1})$. Moreover, we will give an exact formula for the lowest order term of the bias and we will show the decomposition into a tangential component and an orthogonal component. A geometric interpretation of these components is given. The bias remainder term turns out to be of order $O(n^{-3/2})$.

In Section 3 we show that $\sqrt{n}(S_n - \mu)$ has a normal limiting distribution $\mathcal{N}(0, \tilde{\Sigma})$ with degenerate covariance matrix $\tilde{\Sigma}$. Indeed, this degeneracy is such that the limiting distribution is supported by the tangent space to the sphere at the point μ.

Next we write

$$S_n - \mu = T_n + N_n, \tag{3}$$

where N_n is the orthogonal component, i.e., N_n is a scalar multiple of μ, and T_n is the tangential component so that $\mu^T T_n = 0$. Theorem 2 entails that $\sqrt{n}\, T_n$ asymptotically has the $\mathcal{N}(0, \tilde{\Sigma})$-distribution. In Section 4 we further investigate the orthogonal part and show that $-2nN_n$ has a χ^2-type distribution in the limit which is the distribution of the sum of the squares of $(k-1)$ independent zero mean normal variables with possibly different variances. As an illustration of the main results the case of a Fisher–von Mises distribution is considered in Section 5.

Mean direction as a point on the sphere around which the density is imagined to be concentrated is generalized to the concept of mean location on manifolds (see Hendriks, 1991, and Hendriks *et al.*, 1992). Let us note that our approach can be extended to the study of sample mean location on compact submanifolds of $\mathbb{R}^k$ (see Hendriks *et al.*, 1996).

In this paper the transpose of a matrix or vector B will be denoted by B^T. The indicator function of a set A will be denoted by 1_A. The statement $Y_n = O_{L_1}(n^p)$ about a sequence of vector-valued random variables Y_n is taken to mean that $\mathbf{E}(\| Y_n \|) = O(n^p)$. This property implies that $Y_n = (O_{\mathbf{P}}(n^p)$, that is, $\mathbf{P}\{ \| Y_n \| > \varepsilon \} = O(n^p)$ for all $\varepsilon > 0$.

2. ASYMPTOTIC BEHAVIOR OF THE EXPECTATION VALUE OF THE SAMPLE MEAN LOCATION

Let X be the random variable X introduced in Section 1. Let $\mathbf{E}X = a \in \mathbb{R}^k$ and Σ be the covariance matrix of X; Here X is considered as a random element of the vector space $\mathbb{R}^k$, rather than $\mathbf{S}^{k-1}$. Consider the mapping $\Phi\colon \mathbb{R}^k \to \mathbf{S}^{k-1}$ defined by $\Phi(u) = u/\|u\|$ if $u \neq 0$, and $\Phi(0) = p$ for some $p \in \mathbf{S}^{k-1}$, say $p = (1, 0, ..., 0)$; then the mean direction $\mu = \Phi(a)$ and the sample mean direction $S_n = \Phi(\bar{X}_n)$.

Theorem 1. *Suppose that $a \neq 0$; then for $n \to \infty$,*

$$\mathbf{E}(S_n) = \mu - \frac{v}{n} + \frac{\tau}{n} + O(n^{-3/2}),$$

where

$$v = \frac{1}{2} \frac{1}{\|a\|^2} (\operatorname{Tr} \Sigma - \mu^T \Sigma \mu)\, \mu,$$

$$\tau = -\frac{1}{\|a\|^2} (\Sigma\mu - (\mu^T \Sigma \mu)\, \mu).$$

Thus $-v/n$ and τ/n give the behavior up to order n^{-1} of the orthogonal and tangential component of the bias term of S_n as an estimator of the mean location.

COROLLARY 1.

$$\|\mathbf{E}(S_n)\|^2 = 1 - \frac{2}{n}\mu^T v + O(n^{-3/2}) = 1 - \frac{1}{n\,\|a\|^2}(\mathrm{Tr}\,\Sigma - \mu^T \Sigma\mu) + O(n^{-3/2}).$$

COROLLARY 2. *The mean squared error of the estimator S_n of μ equals*

$$\mathbf{E}(\|S_n - \mu\|^2) = \frac{2}{n}\mu^T v + O(n^{-3/2}) = \frac{1}{n\,\|a\|^2}(\mathrm{Tr}\,\Sigma - \mu^T \Sigma\mu) + O(n^{-3/2}).$$

Proof of Theorem 1. Let $Z_n = \sqrt{n}(\bar{X}_n - a)$. Let R_n be defined by the following equation, as the remainder term of the Taylor series expansion of Φ at a of order 2,

$$S_n = \Phi(\bar{X}_n) = \Phi(a) + \frac{1}{\sqrt{n}}\Phi'_a(Z_n) + \frac{1}{2n}\Phi''_a(Z_n, Z_n) + R_n. \tag{4}$$

Denote

$$\Phi_a^{(l)}(v_1, ..., v_l) = D_{v_l} \cdots D_{v_1}\Phi(x)_{|x=a},$$

for $v_1, ..., v_l \in \mathbb{R}^k$. $\Phi'_a = \Phi_a^{(1)}$ is the Jacobian matrix of Φ at a and $\Phi''_a = \Phi_a^{(2)}$ is the vector of Hessian matrices of the coordinates of the vector valued map Φ at a.

Since $\|X\| = 1$, all moments of X are finite. From the well known inequality due to Dharmadhikari and Jogdeo (1969) (see also Petrov, 1975, p. 60) it follows that for $p \geqslant 2$, $\mathbf{E}(|Z_{ni}|^p) \leqslant C_p \mathbf{E}(|(X)_i - a_i|^p) \leqslant 2^p C_p$ and a fortiori that each moment of Z_n is finite and has a finite upper bound independent of n. Here Z_{ni} and $(X)_i$ denote the ith coordinates of the vectors Z_n and X. Moreover, $\|\Phi(\bar{X}_n)\| = 1$, so that all moments of $\Phi(\bar{X}_n)$ are finite. As a consequence, using Eq. (4), each moment of R_n has a finite upper bound, so that in particular the expectation value of R_n, $\mathbf{E}(R_n)$, does exist, and $\mathbf{E}(\|R_n\|^2)$ has a finite upper bound. Let R_{ni} denote the ith coordinate $(R_n)_i$ of R_n.

Let $A = \{u \in \mathbb{R}^k \mid \|u - a\| \leqslant \varepsilon\}$, then $A^c = \{u \in \mathbb{R}^k \mid \|u - a\| > \varepsilon\}$. Note that for $0 < \varepsilon < \|a\|$, Φ is three times continuously differentiable on A. In the

Taylor expansion, Eq. (4) for $\Phi(\bar{X}_n)$, conditional on the event that $\bar{X}_n \in A$, there are numbers $0 < \theta_i < 1$, depending on Z_n, such that

$$R_{ni} = \frac{1}{6} \frac{1}{n^{3/2}} (\Phi_i)^{(3)}_{a + \theta_i Z_n/\sqrt{n}} (Z_n, Z_n, Z_n).$$

From this it follows that

$$\mathbf{E}(\| R_n 1_A(\bar{X}_n) \|) \leqslant C n^{-3/2}$$

for some positive constant C, since for each combination of $j, l, m = 1, ..., k$,

$$\sup_{u \in A} \left| \frac{\partial^3 \Phi_i}{\partial x_j \, \partial x_l \, \partial x_m} \bigg|_u \right| < \infty$$

and, according to the Dharmadhikari and Jogdeo inequality, there are numbers $C_{jlm} > 0$ such that, for all n,

$$\mathbf{E}(|Z_{nj} Z_{nl} Z_{nm}|) < C_{jlm}.$$

On the other hand, for any $p \geqslant 2$,

$$\mathbf{E}(1_{A^c}(\bar{X}_n)) \leqslant \frac{1}{\varepsilon^p n^{p/2}} \mathbf{E}(\| Z_n \|^p) = O(n^{-p/2}), \tag{5}$$

because $\mathbf{E}(\| Z_n \|^p)$ is bounded from above by the Dharmadhikari and Jogdeo inequality.

As $\mathbf{E}(\| R_n \|^2)$ has a finite upper bound,

$$\begin{aligned} \mathbf{E}(\| R_n 1_{A^c}(\bar{X}_n) \|) &\leqslant \mathbf{E}(\| R_n \|^2)^{1/2} \, \mathbf{P}\{\bar{X}_n \in A^c\}^{1/2} \\ &= O(n^{-p/4}) \qquad \text{for any} \quad p \geqslant 2. \end{aligned}$$

We conclude that $\mathbf{E}(\| R_n \|) = O(n^{-3/2})$, or equivalently $R_n = O_{L_1}(n^{-3/2})$.

Note that Eq. (5) allows us to handle the remainder term for the Taylor expansion of any order, as far as we like.

As $\mathbf{E}(Z_n) = 0$, it is clear that $\mathbf{E}(\Phi'_a(Z_n)) = 0$. In order to finish the proof one has to calculate $\mathbf{E}((1/2n) \, \Phi''_a(Z_n, Z_n))$. Note that for $x, y \in \mathbb{R}^k$,

$$\begin{aligned} \Phi'_x(y) &= \frac{1}{\|x\|} (y - (\Phi(x)^T y) \, \Phi(x)) \\ &= \frac{1}{\|x\|} (E - \Phi(x) \, \Phi(x)^T)(y) \\ &= \frac{1}{\|x\|} \tan(y), \end{aligned} \tag{6}$$

where E denotes the identity matrix of rank k, and $\tan(\cdot)$ is the orthogonal projection onto the tangent space $T_{\Phi(x)}\mathbf{S}^{k-1} = \{v \in \mathbb{R}^k \mid \Phi(x)^T v = 0\}$. Note that Φ'_x can be expressed in terms of $\|x\|$ and $\Phi(x)$. Now for $z \in \mathbb{R}^k$,

$$\Phi''_x(y, z) = \frac{1}{\|x\|^2}\,(-y\Phi(x)^T z - z\Phi(x)^T y$$

$$+ 3\Phi(x)\,\Phi(x)^T y\Phi(x)^T z - \Phi(x)\,y^T z).$$

In particular, remembering that we are interested in Φ''_a, and noting that $\mu = \Phi(a)$, $\mu^T Z_n = Z_n^T \mu$, and $Z_n^T Z_n = \operatorname{Tr} Z_n^T Z_n = \operatorname{Tr} Z_n Z_n^T$, we have

$$\Phi''_a(Z_n, Z_n) = \frac{1}{\|a\|^2}\,(-2Z_n\mu^T Z_n + 3\mu(\mu^T Z_n)(\mu^T Z_n) - \mu(Z_n^T Z_n))$$

$$= \frac{1}{\|a\|^2}\,(-2Z_n Z_n^T \mu + 3\mu(\mu^T Z_n Z_n^T \mu) - \operatorname{Tr}(Z_n Z_n^T)\,\mu).$$

Note that we have a decomposition in an orthogonal vector and a tangent vector

$$Z_n Z_n^T \mu = \mu(\mu^T Z_n Z_n^T \mu) + [Z_n Z_n^T \mu - \mu(\mu^T Z_n Z_n^T \mu)],$$

because the inner product $\mu^T[Z_n Z_n^T \mu - \mu(\mu^T Z_n Z_n^T \mu)] = \mu^T Z_n Z_n^T \mu - \mu^T \mu(\mu^T Z_n Z_n^T \mu) = 0$ as $\mu^T \mu = 1$. Therefore

$$\Phi''_a(Z_n, Z_n) = -\frac{1}{\|a\|^2}\,(\operatorname{Tr}(Z_n Z_n^T) - (\mu^T Z_n Z_n^T \mu))\,\mu$$

$$-\frac{2}{\|a\|^2}\,(Z_n Z_n^T \mu - \mu(\mu^T Z_n Z_n^T \mu)). \tag{7}$$

As $\mathbf{E}(Z_n Z_n^T) = \Sigma$ we obtain $\mathbf{E}(\Phi''_a(Z_n, Z_n)) = 2(-\nu + \tau)$ for ν and τ as given in the statement of the theorem. ∎

Corollary 1 immediately follows from the orthogonality of μ and τ. Corollary 2 is a consequence of the fact that

$$\|S_n - \mu\|^2 = (S_n - \mu)^T (S_n - \mu)$$

$$= S_n^T S_n + \mu^T \mu - S_n^T \mu - \mu^T S_n$$

$$= 2 - 2\mu^T S_n,$$

as $S_n^T S_n = \mu^T \mu = 1$ and $\mu^T S_n = S_n^T \mu$. In particular, $\mathbf{E}(\|S_n - \mu\|^2) = 2 - 2\mu^T \mathbf{E}(S_n)$. Thus from the theorem, keeping in mind that $\mu^T \tau = 0$, it follows that

$$\mathbf{E}(\|S_n - \mu\|^2) = 2 - 2\mu^T \left(\mu - \frac{\nu}{n} + \frac{\tau}{n}\right) + O(n^{-3/2})$$

$$= 2\frac{\mu^T \nu}{n} + O(n^{-3/2}). \quad \blacksquare$$

Let us return to Eq. (4) and recall that in the proof of Theorem 1 it was shown that $R_n = O_{L_1}(n^{-3/2})$. Using Eq. (7) we get

$$S_n = \mu + \frac{1}{\sqrt{n}} \Phi'_a(Z_n) - \frac{1}{2n} \frac{1}{\|a\|^2} (\mathrm{Tr}\, Z_n Z_n^T - \mu^T Z_n Z_n^T \mu)\, \mu$$

$$- \frac{1}{n} \frac{1}{\|a\|^2} (Z_n Z_n^T \mu - \mu\mu^T Z_n Z_n^T \mu) + O_{L_1}(n^{-3/2}). \tag{8}$$

Note that the order $n^{-1/2}$ term is a random vector concentrated on the tangent space $T_\mu \mathbf{S}^{k-1}$ to $\mathbf{S}^{k-1}$ at μ. The term of order n^{-1} is decomposed into two components, namely into a vector proportional to μ and a vector in $T_\mu \mathbf{S}^{k-1}$. In the next section we will study the limit distribution of the vector $\sqrt{n}(S_n - \mu)$. We will show that its limit distribution is determined by $\Phi'_a(Z_n)$. That is why one can expect that the limit distribution will be degenerate on $\mathbb{R}^k$ and will be supported by the tangent plane $T_\mu \mathbf{S}^{k-1}$. The normal component, determined by $\mu^T \sqrt{n}(S_n - \mu)$, has a limit distribution concentrated in 0. In Section 4 we will consider the asymptotic behavior of the orthogonal component in more detail. In particular, it will be shown that $n\mu^T(S_n - \mu)$ has a limit distribution, determined by the first term of order n^{-1} in the above decomposition.

3. ASYMPTOTIC BEHAVIOR OF THE DISTRIBUTION OF SAMPLE MEAN DIRECTION

This section will be devoted to obtaining the asymptotical distribution of the random vector $\sqrt{n}(S_n - \mu)$. Our interest in this problem is twofold. First, it will be shown that the limit distribution is Gaussian with a degenerate covariance matrix and is supported by the tangent space $T_\mu \mathbf{S}^{k-1}$. Second, we indicate its use for testing the hypothesis H_0 that $\mu = \mu_0$. The critical area of the corresponding test lies in the space $T_\mu \mathbf{S}^{k-1}$ of dimension $(k-1)$. Note that Σ coincides with the covariance matrix of Z_n and that $\tilde{\Sigma} = \Phi'_a \Sigma (\Phi'_a)^T$ is the covariance matrix of $\Phi'_a(Z_n)$.

THEOREM 2. *Suppose* $a \neq 0$ *and* Σ *is nonsingular. For* $n \to \infty$, $\sqrt{n}(S_n - \mu)$ *converges in distribution to the degenerate normal distribution* $\mathcal{N}(0, \tilde{\Sigma})$, *with covariance matrix of rank* $k-1$.

Let $\xi = (\xi_1, ..., \xi_k)^T$ *be degenerately normally distributed with mean* 0 *and covariance matrix* $\tilde{\Sigma}$. *Then the support of* ξ *belongs to the tangent space* $T_\mu \mathbf{S}^{k-1} = \{x \in \mathbb{R}^k \mid \mu^T x = 0\}$.

Suppose without loss of generality that $a_k \neq 0$. Denote by $\tilde{\Sigma}_{(k-1)}$ the upper left submatrix of $\tilde{\Sigma}$ of dimension $k-1$ (cf. Anderson, 1966, p. 26) and $\mu^{(k-1)} = (\mu_1, ..., \mu_{k-1})^T$. Then $\tilde{\Sigma}_{(k-1)}$ is nonsingular and $\xi^{(k-1)} = (\xi_1, ..., \xi_{k-1})^T$ has a nondegenerate normal distribution with zero expectation vector and covariance matrix $\tilde{\Sigma}_{(k-1)}$. The degenerately normally distributed vector ξ can be reconstructed from $\xi^{(k-1)}$ as

$$\xi = \begin{pmatrix} E \\ -\dfrac{1}{\mu_k} (\mu^{(k-1)})^T \end{pmatrix} \cdot \xi^{(k-1)},$$

where E denotes the identity matrix of rank $k-1$.

The behavior of the degenerately normal vector ξ can be summarized by the following simple geometric consideration. The support of the distribution of $\xi = (\xi_1, ..., \xi_k)^T$ is concentrated in a hyperspace which can be parametrized by the parameters $\xi^{(k-1)} = (\xi_1, ..., \xi_{k-1})^T$ which are distributed $\mathcal{N}(0, \tilde{\Sigma}_{(k-1)})$.

Proof of Theorem 2. From Eq. (8) it follows immediately that

$$\sqrt{n}(S_n - \mu) = \Phi'_a(Z_n) + G_n = \frac{1}{\|a\|} \tan(Z_n) + G_n,$$

where $\mathbf{E}(\|G_n\|) = O(n^{-1/2})$ so that $G_n = O_{\mathbf{P}}(n^{-1/2})$. From the well known convergence theorem (Cramér, 1961, p. 254) the limit distribution of $\sqrt{n}(S_n - \mu)$ and that of $\Phi'_a(Z_n)$ are the same. Note that because Φ'_a is a degenerate matrix there is some peculiarity in applying the multivariate version of the Central Limit Theorem, where nondegeneracy of the first derivative matrix seems to be meant formally (see, for example, Rao (1973, 6a.2ii., 6a2.6, 6a.2iii, and 6a2.10). However, the characteristic function of the vector $\Phi'_a Z_n$ at the point t converges pointwise to the characteristic function of the vector Z_n at the point $(\Phi'_a)^T t$. This means that $\Phi'_a Z_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Phi'_a \Sigma \Phi'^T_a)$. The covariance matrix $\tilde{\Sigma} = \Phi'_a \Sigma (\Phi'_a)^T$ has rank $k-1$ and the normal distribution $\mathcal{N}(0, \tilde{\Sigma})$ as a matter of fact is a degenerate normal distribution (see Anderson, 1966, p. 25 ff) with support on $T_\mu \mathbf{S}^{k-1}$. ∎

Denote by $V=(1/n)\sum_{i=1}^{n}(X_i-\bar{X}_n)(X_i-\bar{X}_n)^T$ the sample covariance matrix, and set $\tilde{V}=\Phi'_a V(\Phi'_a)^T$. Let $\tilde{W}$ be the positive semidefinite symmetric square root of $\tilde{V}$, so that $\tilde{V}=\tilde{W}\cdot\tilde{W}$ (see Lancaster, 1969, Theorem 2.14.2). As $V\to\Sigma$ almost surely as $n\to\infty$, the linear transformation of $\mathbb{R}^k$ determined by V is nonsingular for large n with probability 1, so that $\tilde{V}$ and $\tilde{W}$ may be supposed to have their range equal to the range of Φ'_a, that is, equal to $T_\mu\mathbf{S}^{k-1}$. Let $\tilde{W}^+$ denote the Moore inverse of $\tilde{W}$ (see Rao, 1973, 1b.5 (viii)). Then it is known that $\tilde{W}\tilde{W}^+=\tan$, where $\tan\ =E-\mu\mu^T$ is the orthogonal projection from $\mathbb{R}^k$ onto $T_\mu\mathbf{S}^{k-1}$. Note that $\tilde{W}^+$ is symmetric. Moreover, $(\tilde{W}^+)^2$ is the Moore inverse $\tilde{V}^+$ of $\tilde{V}$. For testing the hypothesis $H_0\colon\mu=\mu_0$ the following Corollary to Theorem 2 and Cramér's convergence theorem is useful,

$$\sqrt{n}\,\tilde{W}^+(X_n-\mu)\xrightarrow{\mathscr{D}}\mathscr{N}(0,\tan\ ),$$

where $\mathscr{N}(0,\tan\ )$ is the degenerate normal distribution with support on $T_\mu\mathbf{S}^{k-1}$ and covariance matrix equal to the Euclidean inner product on $T_\mu\mathbf{S}^{k-1}$. As a consequence

$$n(S_n-\mu)^T\,\tilde{V}^+(S_n-\mu)\xrightarrow{\mathscr{D}}\chi^2_{k-1},\qquad n\to\infty.$$

4. ASYMPTOTIC BEHAVIOR OF THE ORTHOGONAL COMPONENT OF THE n^{-1} ORDER TERM

Let us recall Eq. (8) and give a simple geometric interpretation. It can be seen that $S_n-\mu$ decomposes into a tangent component T_n consisting of 2 terms, $(1/\sqrt{n})\,\Phi'_a(Z_n)$, of order $n^{-1/2}$ and a term T_n^1 of order n^{-1} and an orthogonal component N_n of order n^{-1}. By multiplying $S_n-\mu$ to the left by μ^T one gets the length of the orthogonal component N_n:

$$\mu^T(S_n-\mu)=-\|N_n\|=-\frac{1}{2n}\frac{1}{\|a\|^2}(\mathrm{Tr}\,Z_nZ_n^T-\mu^TZ_nZ_n^T\mu)+O_{L_1}(n^{-3/2}).$$

Note that the minus sign is appropriate since the angle between μ and $S_n-\mu$ is obtuse.

Recall that $\Phi'_a(Z_n)=(1/\|a\|)(Z_n-\mu\mu^TZ_n)$ from which it follows that

$$(\Phi'_aZ_n)^T(\Phi'_aZ_n)=\|\Phi'_aZ_n\|^2=\frac{1}{\|a\|^2}(\|Z_n\|^2-(\mu^TZ_n)^2)$$

$$=\frac{1}{\|a\|^2}(\mathrm{Tr}\,Z_nZ_n^T-\mu^TZ_nZ_n^T\mu).$$

So we get

$$K_n = -Zn\mu^T(S_n - \mu) = \|\Phi'_a Z_n\|^2 + O_{L_1}(n^{-1/2}) \tag{9}$$

and

THEOREM 3. *Suppose $a \neq 0$ and Σ is nonsingular. Then the random variable $K_n = -2n\mu^T(S_n - \mu)$ asymptotically has the same distribution as a weighted sum of $(k-1)$-independent χ^2_1-distributed random variables with weights equal to the nonzero eigenvalues (repeated according to their multiplicity) of $\tilde{\Sigma}$.*

Remark 1. If the nonzero eigenvalues of the matrix $\tilde{\Sigma}$ are all equal to λ, then

$$\frac{K_n}{\lambda} \xrightarrow{\mathscr{D}} \chi^2_{k-1}.$$

As a special case (cf. Watson, 1983, formula (4.2.9)), let $G = O(k)$ be the group of orthogonal matrices and $G_\mu = \{g \in G \mid g\mu = \mu\}$ be the isotropy group of μ. Suppose that for all $g \in G_\mu$, X and gX have the same distribution on $\mathbf{S}^{k-1}$. Then $\tilde{\Sigma} = \lambda(E - \mu\mu^T)$ with $\lambda = (1/(k-1)) \operatorname{Tr} \tilde{\Sigma}$.

Remark 2. If $k = 2$, the nontrivial eigenvalue of $\tilde{\Sigma}$ is $\operatorname{Tr} \tilde{\Sigma}$ and

$$\frac{K_n}{\operatorname{Tr} \tilde{\Sigma}} \xrightarrow{\mathscr{D}} \chi^2_1.$$

5. APPLICATION TO FISHER–VON MISES DISTRIBUTIONS

In this section we will illustrate the previous results with a von Mises distribution P on the circle of radius 1 and center 0. The density function with respect to the Lebesgue measure is as follows (see Mardia, 1975, p. 57):

$$g(x; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa\mu^T x).$$

Choose a unit vector γ such that $\mu^T\gamma = 0$. According to Mardia (1975, p. 62), the mean is $a = A(\kappa)\,\mu$, where $A(\kappa) = I_1(\kappa)/I_0(\kappa)$, therefore the mean location is μ. The variance matrix is given in Mardia (1975, p. 108, formulas (4.8.1–4.8.3)),

$$\Sigma = \tfrac{1}{2}\begin{pmatrix} 1 + \alpha_2 - 2\alpha^2 & \beta_2 - 2\alpha\beta \\ \beta_2 - 2\alpha\beta & 1 - \alpha_2 - 2\beta^2 \end{pmatrix}.$$

Using Mardia (1975, p. 62, formulas (3.4.43) and (3.4.51)), it follows that in the case $\mu=(1,0)$ we have $\beta=\beta_2=0$ and

$$\Sigma=\left(1-\frac{A(\kappa)}{\kappa}-A(\kappa)^2\right)\mu\mu^T+\frac{A(\kappa)}{\kappa}\gamma\gamma^T.$$

The generality of the last formula for Σ follows using the action of the orthogonal group $O(2)$. Note that $\Phi'_a=(1/A(\kappa))(E-\mu\mu^T)$, so that

$$\tilde{\Sigma}=\frac{1}{A(\kappa)^2}\frac{A(\kappa)}{\kappa}\gamma\gamma^T=\frac{1}{\kappa A(\kappa)}\gamma\gamma^T.$$

Now we conclude that

$$\sqrt{n}(S_n-\mu)\xrightarrow{\mathscr{D}}\mathcal{N}\left(0,\frac{1}{\kappa A(\kappa)}\gamma\gamma^T\right),$$

so that we get for example the limit

$$\gamma^T\sqrt{n}(S_n-\mu)\xrightarrow{\mathscr{D}}\mathcal{N}\left(0,\frac{1}{\kappa A(\kappa)}\right).$$

In applying Theorem 1 we note that the tangential component $\tau=0$ and that

$$\mathbf{E}(S_n)=\mu-\frac{1}{n}\frac{1}{2}\frac{1}{\|a\|^2}\left(1-A(\kappa)^2-1+\frac{A(\kappa)}{\kappa}+A(\kappa)^2\right)\mu+O(n^{-3/2})$$

$$=\mu-\frac{1}{n}\frac{1}{2}\frac{1}{\kappa A(\kappa)}\mu+O(n^{-3/2}),$$

$$\mathbf{E}(\|S_n-\mu\|^2)=\frac{1}{n}\frac{1}{\kappa A(\kappa)}+O(n^{-3/2}),$$

$$\|\mathbf{E}(S_n)\|^2=1-\frac{1}{n}\frac{1}{\kappa A(\kappa)}+O(n^{-3/2}).$$

Theorem 3 yields the following:

$$K_n=-2\mu^Tn(S_n-\mu)\xrightarrow{\mathscr{D}}\frac{1}{\kappa A(\kappa)}\chi_1^2.$$

REFERENCES

Anderson, T. W. (1966). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

Cramér, H. (1961). *Mathematical Methods of Statistics*. Princeton Univ. Press, Princeton, NJ.

Dharmadhikari, S. W., and Jogdeo, K. (1969). Bounds on moments of certain random variables. *Ann. Math. Statist.* **40** 1506–1508.

Hendriks, H. (1991). A Cramér-Rao type lower bound for estimators with values in a manifold. *J. Multivariate Anal.* **38** 245–261.

Hendriks, H., Janssen, J., and Ruymgaart, F. (1992). A Cramér-Rao type inequality for random variables in Euclidean manifolds. *Sankhyā-Ser. A* **54** 387–401.

Hendriks, H., Landsman, Z., and Ruymgaart, F. (1996). Asymptotic behavior of sample mean location for manifolds. *Stat. Prob. Letters* **26** 169–178.

Jupp, P., and Mardia, K. V. (1989). A unified view of the theory of directional statistics. *Internat. Statist. Rev.* **57** 261–294.

Lancaster, P. (1969). *Theory of Matrices*. Academic Press, New York/London.

Mardia, K. V. (1975). *Statistics of Directional Data*. Academic Press, London.

Petrov, V. V. (1975). *Sums of Independent Random Variables*. Springer-Verlag, Berlin.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Watson, G. S. (1983). *Statistics on Spheres*. Wiley, New York.