2012

Bayes Inference for a Tractable New Class of Non-symmetric Distributions for 3-Dimensional Rotations

Melissa Ann Bingham
University of Wisconsin - La Crosse

Danial J. Nordman
Iowa State University, dnordman@iastate.edu

Stephen B. Vardeman
Iowa State University, vardeman@iastate.edu

The final publication is available at Springer via http://dx.doi.org/10.1007/s13253-012-0107-9

Bayes Inference for a New Class of Nonsymmetric Distributions for 3-Dimensional Rotations

Melissa A. Bingham (bingham.meli@uwlax.edu)

University of Wisconsin-La Crosse

1725 State St.

La Crosse, WI 54601

Daniel J. Nordman (dnordman@iastate.edu)

Iowa State University

Ames, IA 50011

Stephen B. Vardeman (vardeman@iastate.edu)

Iowa State University

Ames, IA 50011

Abstract

Both existing models for nonsymmetric distributions on 3-dimensional rotations and their associated one-sample inference methods have serious limitations in terms of both interpretability and ease of use. Based on the intuitively appealing Uniform Axis-Random Spin (UARS) construction of Bingham, Nordman, and Vardeman (2009) for symmetric families of distributions, we propose new highly interpretable and tractable classes of nonsymmetric distributions that are derived from mixing UARS distributions. These have an appealing Preferred Axis-Random Spin (PARS) construction and (unlike existing models) directly interpretable parameters. Non-informative one-sample Bayes inference in these models is a direct generalization of UARS methods introduced in Bingham, Vardeman, and Nordman (2009), where credible levels were found to be essentially equivalent to frequentist coverage probabilities. We apply the new models and inference methods to a problem in biomechanics, where comparison of model parameters provides meaningful comparisons for the nature of movement about the calcaneocuboid joint of three different primate subjects.

Key Words

calcaneocuboid joint, circular distribution, credible sets, general von Mises-Fisher distribution, Haar meausure, Jeffreys prior, Markov chain Monte Carlo, mixture, preferred axis, UARS distribution

1 Introduction

Data in the form of 3-dimensional rotations are common in the study of biomechanics and human motion. As skeletal mammals move, their bones rotate around various joints. By connecting infra-red emitting diodes to various positions on limbs, the rotations about joints can be traced as movement occurs. In the application considered in this paper, we analyze movement about the calcaneocuboid joint, which lies between the calcaneous (heel) and cuboid bones of the foot (see Figure 1). Rotational movement about this joint while in a sitting position was measured over time as load was applied to the knee. Measurements were obtained from three primates; a human, a baboon, and a chimpanzee. All data treated in this paper were obtained by Prof. Thomas Greiner of the Physical Therapy Department at the University of Wisconsin-La Crosse. To compare and contrast the movement in the foot for the different subjects, we require methods for analyzing 3-dimensional orientation data.

Bingham, Nordman, and Vardeman (2009) considered the Uniform Axis-Random Spin (UARS) class of distributions on orientations in three dimensions. While the UARS class provided important new flexibility in the modeling of 3-dimensional orientation data, it is a symmetric class of distributions and may not provide an adequate fit to data showing some degree of asymmetry of distribution (as in the case of the calcaneocuboid data we will consider). Thus, it is of interest to extend the UARS class. While the general von Mises-Fisher distribution (referred to here as the matrix Fisher distribution) allows for nonsymmetric modeling and is the most widely studied and referenced distribution for rotations in the statistical literature (see Downs (1972), Khatri and Mardia (1977), Jupp and Mardia (1979), and Prentice(1986)), it suffers from practical limitations. Inference for the matrix Fisher distribution is not simple, and parameters are not easily interpreted. This is illustrated in the context of the calcaneocuboid data in Section 4. Further, simulation from the matrix Fisher distribution is not obvious. There is a need for a nonsymmetric class of distribu-

tions without the limitations of the matrix Fisher class, useful for applications like the one considered here. In the next section we develop such a statistical model. By providing a constructive definition of a new class of distributions, we arrive at physically meaningful parameters. In Section 3 we offer a direct and straightforward approach to inference for this new class of distributions using MCMC-based Bayes methods, and in Section 4 this methodology is applied to the analysis of calcaneocuboid data.

2 The Preferred Axis-Random Spin Distributions

Let Ω be the set of 3×3 orthogonal rotation matrices satisfying the right hand rule (with positive determinant). The UARS distributions of Bingham, Nordman, and Vardeman (2009) can be used to model a random orientation in Ω distributed about a central orientation of $\mathbf{S} \in \Omega$. $\mathbf{Q} \in \Omega$ from a UARS distribution with central orientation $\mathbf{S}$ can be represented as $\mathbf{Q} = \mathbf{S}\mathbf{T}(\mathbf{U}, r)$, where the random rotation matrix

$$\mathbf{T}(\mathbf{U}, r) = \mathbf{U}\mathbf{U}^T + (\mathbf{I}_{3\times3} - \mathbf{U}\mathbf{U}^T)\cos r + \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \sin r \in \Omega \qquad (1)$$

results from rotating the 3×3 identity matrix, $\mathbf{I}_{3\times3}$, about a random axis identified by unit vector $\mathbf{U} = (u_1, u_2, u_3)^T \in \mathbb{R}^3$ through an independently generated random angle $r \in (-\pi, \pi]$. For the UARS distributions, the random angle r is assumed to follow a circular distribution on $(-\pi, \pi]$ that is symmetric about 0 with spread depending on a concentration parameter $\kappa \geq 0$ (represented by the density $C(r|\kappa)$), and $\mathbf{U}$ is uniformly distributed on the unit sphere. Small $|r|$ indicates rotations from a UARS model deviating little from the central orientation $\mathbf{S}$, and since κ controls the spread of the circular distribution for r, it also controls variation in random rotations from the UARS distributions. Bingham, Nordman, and Vardeman (2009)

write $\mathbf{Q} \sim \text{UARS}(\mathbf{S}, \kappa)$, where $\mathbf{S}$ is the location parameter and κ is the spread parameter. A particular subclass of UARS distributions is specified by choosing a parametric circular distribution for r by specifying a density $C(r|\kappa)$.

Random orientations from a UARS distribution will scatter around the central orientation $\mathbf{S}$ in a *symmetric* fashion. Figure 2 (a) portrays 100 UARS observations plotted around $\mathbf{S}$ (represented by the rotated positions of the positive x, y, and z axes). Each observation (a rotation) corresponds to a set of three points on the sphere (representing the three columns of the rotation matrix), scattered around the unit vectors that are the columns of $\mathbf{S}$. We can see that the UARS observations form directionally symmetric patterns of rotated axes around the central position. There are situations, however, where data may scatter around a central orientation in patterns that are not directionally symmetric but have "preferred directions," such as is pictured in Figure 2 (b).

We next provide a constructive definition to develop a class of *nonsymmetric* distributions for 3-dimensional rotations. We also present some properties of this class and note here that while proofs of these properties are omitted, most follow fairly directly from results for the UARS class (see the Appendix of Bingham, Nordman, and Vardeman (2009)). Let $\mathbf{O} \sim \text{UARS}(\mathbf{I}_{3\times 3}, \kappa)$ and let $\mathbf{V}$ be a fixed vector. By multiplying a random symmetric pertubation (i.e. having random axis) $\mathbf{O}$ by a random rotation $\mathbf{T}(\mathbf{V}, p)$ around a fixed axis $\mathbf{V}$, where $\mathbf{T}$ is the function of a unit vector and angle defined in (1) and $p \in (-\pi, \pi]$ follows a circular distribution symmetric about 0 with concentration parameter τ (with density $D(p|\tau)$), we can "smear" distributions for orientations that are symmetric about the central orientation to create nonsymmetric orientations. We will say (for $\mathbf{O} \sim \text{UARS}(I_{3\times 3}, \kappa)$) that $\mathbf{P} = \mathbf{T}(\mathbf{V}, p)\mathbf{O}$ has Preferred Axis-Random Spin distribution with location $\mathbf{I}_{3\times 3}$, and parameters $\mathbf{V}$, κ, and τ and use the notation $\text{PARS}(\mathbf{I}_{3\times 3}, \kappa, \mathbf{V}, \tau)$ for the distribution of $\mathbf{P}$. We will refer to $\mathbf{V}$ as the "preferred axis of rotation." Property 1 below indicates that $\mathbf{P}$ and $\mathbf{P}'$ have the same distribution.

Property 1. *Both* $\mathbf{OT}(\mathbf{V}, p)$ *and* $\mathbf{T}(\mathbf{V}, p)\mathbf{O} \sim$ *PARS*$(\mathbf{I}_{3\times 3}, \kappa, \mathbf{V}, \tau)$.

Now, for $\mathbf{S} \in \Omega$, we'll say that $\mathbf{M} = \mathbf{SP} \sim \text{PARS}(\mathbf{S}, \kappa, \mathbf{V}, \tau)$. Due to the geometric nature of the construction of the model, we arrive at parameters that are easily interpreted. The parameter $\mathbf{S}$ serves as the central orientation for a $\text{PARS}(\mathbf{S}, \kappa, \mathbf{V}, \tau)$ distribution, κ controls the spread of orientations (with larger values indicating less spread), τ controls the degree of asymmetry in the orientations (with larger values indicating a more symmetric distribution), and $\mathbf{V}$ controls the direction of the asymmetry. The preferred axis of rotation, $\mathbf{V}$, is defined in terms of the original axes, so for $\text{PARS}(\mathbf{S}, \kappa, \mathbf{V}, \tau)$ observations (in the coordinate system that is established by rotation of the original axes by $\mathbf{S}$), the preferred axis of rotation is rotated to $\mathbf{SV}$.

Properties 2 and 3 below illustrate that while the order in which $\mathbf{S}$ and $\mathbf{O}$ appear in the PARS construction does not matter (Bingham, Nordman, and Vardeman (2009) give the result that both $\mathbf{OS}$ and $\mathbf{SO}$ are $\text{UARS}(\mathbf{S}, \kappa)$), it does matter (as far as establishing the preferred axis of rotation) whether rotation about the vector $\mathbf{V}$ or rotation by $\mathbf{S}$ is applied first. As the PARS class is defined, with rotation about $\mathbf{V}$ occurring *prior* to rotation by $\mathbf{S}$, $\mathbf{V}$ is the preferred axis of rotation in the original coordinate system and "smearing" symmetric distributions centered at $\mathbf{I}_{3\times 3}$ about $\mathbf{V}$ is equivalent to "smearing" symmetric distributions centered at $\mathbf{S}$ about $\mathbf{SV}$. Thus, if the rotation about $\mathbf{V}$ is applied *after* rotating the original axes by $\mathbf{S}$, we are "smearing" symmetric distributions centered at $\mathbf{S}$ about $\mathbf{V}$. This is equivalent to "smearing" symmetric distributions centered at $\mathbf{I}_{3\times 3}$ about $\mathbf{S}'\mathbf{V}$, giving preferred axis of rotation $\mathbf{S}'\mathbf{V}$ (Property 3).

Property 2. *Both* $\mathbf{SOT}(\mathbf{V}, p)$ *and* $\mathbf{OST}(\mathbf{V}, p) \sim$ *PARS*$(\mathbf{S}, \kappa, \mathbf{V}, \tau)$. *Consequently* $\mathbf{QT}(\mathbf{V}, p) \sim$ *PARS*$(\mathbf{S}, \kappa, \mathbf{V}, \tau)$ *where* $\mathbf{Q} \sim$ *UARS*$(\mathbf{S}, \kappa)$.

Property 3. *Both* $\mathbf{T}(\mathbf{V}, p)\mathbf{SO}$ *and* $\mathbf{T}(\mathbf{V}, p)\mathbf{OS} \sim$ *PARS*$(\mathbf{S}, \kappa, \mathbf{S}'\mathbf{V}, \tau)$. *Consequently* $\mathbf{T}(\mathbf{V}, p)\mathbf{Q} \sim$ *PARS*$(\mathbf{S}, \kappa, \mathbf{S}'\mathbf{V}, \tau)$ *and* $\mathbf{T}(\mathbf{SV}, p)\mathbf{Q} \sim$ *PARS*$(\mathbf{S}, \kappa, \mathbf{V}, \tau)$, *where* $\mathbf{Q} \sim$ *UARS*$(\mathbf{S}, \kappa)$.

Property 4 states that the resulting PARS distribution is the same whether we are considering preferred axis $\mathbf{V}$ or $-\mathbf{V}$. Thus, the preferred axis of rotation should be viewed as a line rather than a vector, as the direction here is non-identifiable.

Property 4. *The PARS*$(\mathbf{S}, \kappa, \mathbf{V}, \tau)$ *and PARS*$(\mathbf{S}, \kappa, -\mathbf{V}, \tau)$ *distributions are the same.*

By generating random PARS observations and plotting their effects on the directions of the three positive coordinate axes, we can visually examine how the different parameters affect the distribution. In Figure 3, 100 PARS observations are represented as point clouds on the sphere around the perpendicular axes x, y, and z, which represent the columns of central orientation $\mathbf{I}_{3\times3}$. (Each PARS observation is represented by 3 points, one near each axis.) The preferred axis of rotation, $\mathbf{V}$, is also displayed. Note that a specific subclass of the PARS class is chosen by placing particular distributions on both p and r, i.e. by specifying $D(p|\tau)$ and $C(r|\kappa)$. In this case, the von Mises circular distribution with mean direction 0 was used for both spins p and r (see Mardia and Jupp, 2000), and the figure shows that the parameters do control the aspects of the PARS distribution described above. (If one wanted to visualize a central orientation of $\mathbf{S}$, all aspects of the picture would simply be rotated by $\mathbf{S}$.)

To further examine the nature of the PARS distributions, 100,000 PARS$(\mathbf{I}_{3\times3}, 100, \mathbf{V}, 15)$ observations were generated for $\mathbf{V} = (1,1,1)/\sqrt{3}$ and $\mathbf{V} = (1,3,5)/\sqrt{35}$. The distributional density (with respect to the Haar measure; see Section 3) was computed for each of the observations and contours were located on the sphere such that the set of three ith contours, $i = 1, \ldots, 9$, simultaneously captures the directions of the rotated versions of three positive coordinate axes for the $(10\cdot i)\%$ of the data (or $10000\cdot i$ observations) with the largest density. Figure 4 displays the contours about $\mathbf{I}_{3\times3}$, represented by x, y and z unit vectors. The figure clearly shows how $\mathbf{V}$ affects the direction of the asymmetry in the distribution, with the extent of the "smearing" about a particular axis determined by the angle between $\mathbf{V}$ and that axis. In the case of $\mathbf{V} = (1,1,1)/\sqrt{3}$, each of the x, y and z unit vectors has the same

angular distance from the preferred axis of rotation and all three sets of contours exhibit the same degree of asymmetry. In the case of $\mathbf{V} = (1, 3, 5)/\sqrt{35}$, z makes the smallest absolute angle with the preferred axis of rotation, resulting in contours that are the most circular in nature, while x makes the largest absolute angle and shows the most asymmetry. Further, increasing the value of κ will result in contours that are further from the coordinate axes, and increasing the value of τ will result in contours that are more circular in nature. The constructive definition of the PARS model yields parameters that are easily interpreted and this will prove valuable in applications like that considered in Section 4.

3 Non-Informative One-Sample Bayes Inference for the PARS Distributions

In this section we develop a straightforward MCMC-based Bayes approach to inference for the PARS distributions. This approach, with the right prior choice, allows us to perform very tractable approximately frequentist inference in a situation where maximum likelihood would be far less tractable (and potentially even undefined). Since the model parameters include 3-dimensional rotations ($\mathbf{S}$) and unit vectors in $\mathbb{R}^3$ ($\mathbf{V}$), even where the likelihood is well-behaved its computation would require n calls to numerical integration per evaluation (as the density given below does not have an obvious closed form) and maximization would be difficult. The Bayes methodology developed here gives an inference framework for describing the PARS parameters belonging to special manifolds without the need for reparamaterization (e.g. using quaternions, Euler angles), while avoiding maximization of likelihood functions on these manifolds.

We begin by describing the density (with respect to the invariant Haar measure on rotations) of the PARS distributions. If $\mathbf{M} \sim \text{PARS}(\mathbf{S}, \kappa, \mathbf{V}, \tau)$, then $\mathbf{M} = \mathbf{ST}(\mathbf{V}, p)\mathbf{O}$

where $\mathbf{O} \sim \text{UARS}(\mathbf{I}_{3\times 3}, \kappa)$, and given the value of the spin p, we have

$$\mathbf{T}'(\mathbf{V}, p)\mathbf{S}'\mathbf{M} \sim \text{UARS}(\mathbf{I}_{3\times 3}, \kappa).$$

Therefore, the conditional density $f(\mathbf{M}|p)$ is obtained by using $\text{T}'(\mathbf{V}, p)\mathbf{S}'\mathbf{M}$ as the argument in the $\text{UARS}(\mathbf{I}_{3\times 3}, \kappa)$ density. The invariant Haar measure acts as a "uniform distribution" on Ω (Downs, 1972) and Bingham, Nordman, and Vardeman (2009, Sec. 2.2) give the $\text{UARS}(\mathbf{I}_{3\times 3}, \kappa)$ density with respect to the Haar measure. So, we have conditional density

$$f(\mathbf{M}|p, \mathbf{S}, \kappa, \mathbf{V}) = \frac{4\pi}{3 - tr(\mathbf{T}'(\mathbf{V}, p)\mathbf{S}'\mathbf{M})} C\left(\arccos\left[1/2(tr(\mathbf{T}'(\mathbf{V}, p)\mathbf{S}'\mathbf{M}) - 1)\right]|\kappa\right), \qquad (2)$$

with respect to the Haar measure, where $C(r|\kappa)$ is the density for the spin of the UARS distribution and tr stands for the trace of a matrix. Then, a joint density for $(\mathbf{M}, p)$ is given by

$$g(\mathbf{M}, p|\mathbf{S}, \kappa, \mathbf{V}, \tau) = f(\mathbf{M}|p, \mathbf{S}, \kappa, \mathbf{V})D(p|\tau), \qquad (3)$$

where $f(\mathbf{M}|r, \mathbf{S}, \kappa, \mathbf{V})$ is given in (2) and $D(p|\tau)$ is the density for the circular distribution placed on p. The PARS density for $\mathbf{M}$ does not have an obvious closed form, but for specific $\mathbf{M}$ can be evaluated by numerical integration of g in (3) with respect to $p \in (-\pi, \pi]$.

To perform Bayes inference for the PARS class, priors must first be placed on the parameters $\mathbf{S}$, κ, $\mathbf{V}$, and τ. For the location parameter $\mathbf{S}$ we use a prior uniform on Ω, so the prior distribution for $\mathbf{S}$ is specified by the density

$$h_1(\mathbf{S}) = 1 \qquad (4)$$

with respect to the Haar measure. For the parameters κ and τ, we will use Jeffreys priors.

So the prior densities for κ and τ (with respect to the Lebesgue measure) will be given by

$$h_2(\kappa) = \sqrt{-\mathrm{E}\left(\frac{d^2}{d^2\kappa}\log(C(r|\kappa))\right)} \tag{5}$$

and

$$h_3(\tau) = \sqrt{-\mathrm{E}\left(\frac{d^2}{d^2\tau}\log(D(p|\tau))\right)}. \tag{6}$$

Note that if r has von Mises circular distribution with density $C(r|\kappa) = [\sqrt{2\pi}I_0(\kappa)]^{-1}\exp(\kappa\cos r)$, the Jeffreys prior is given by the density

$$h_2(\kappa) = (I_0(\kappa))^{-1}\sqrt{I_0(\kappa)^2 - \frac{1}{\kappa}I_0(\kappa)I_1(\kappa) - I_1(\kappa)^2}, \quad \kappa \in [0,\infty),$$

where $I_i(\cdot)$ is the modified Bessel function of order i. Lastly, we place a uniform prior on the vector $\mathbf{V}$. If we write $\mathbf{V}$ in terms of polar coordinates $\theta \in [0,\pi]$ and $\phi \in [0,2\pi)$ so that $\mathbf{V} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, then the uniform distribution on the set of unit vectors has probability density (with respect to the Lebesgue measure) given by

$$h_4(\mathbf{V}) = h_4(\theta,\phi) = \sin\theta/4\pi. \tag{7}$$

Now suppose that $\mathbf{m}_1,\ldots,\mathbf{m}_n$ are PARS$(\mathbf{S},\kappa,\mathbf{V},\tau)$ observations with corresponding (unobservable) spins $p_1,\ldots,p_n$. Then we have a posterior density for the parameters $\mathbf{S}$, κ, $\mathbf{V}$, and τ, and for the unobservable $p_1,\ldots,p_n$ proportional to

$$G(\mathbf{S},\kappa,\mathbf{V},\tau,p_1,\ldots,p_n) = \prod_{i=1}^{n} g(\mathbf{m}_i,p_i|\mathbf{S},\kappa,\mathbf{V},\tau)h_1(\mathbf{S})h_2(\kappa)h_3(\tau)h_4(\mathbf{V}), \tag{8}$$

where g is given in (3), and h_j, $j = 1,2,3,4$, are given in (4)-(7).

To simulate values from the posterior (8), a Metropolis-Hastings within Gibbs algorithm can be used. We first describe the generation of proposals for updates of parameters and then

detail how these are used. If after $j-1$ iterations of the algorithm we have $\mathbf{S}^{j-1}$ as the current value of the parameter $\mathbf{S}$, then we obtain a candidate for $\mathbf{S}^j$ as $\mathbf{S}^{j*} \sim$ vM-UARS$(\mathbf{S}^{j-1}, \rho_1)$, where ρ_1 is a tuning parameter that can be adjusted to make the algorithm efficient and vM-UARS represents the member of the UARS class with von Mises circular distribution for the spin r (see Bingham, Nordman, and Vardeman (2009), Bingham, Vardeman, and Nordman (2009) for more on the vM-UARS distribution). As Bingham, Vardeman, and Nordman (2009) point out, this choice of proposal for $\mathbf{S}$ is symmetric in the sense that $F(\mathbf{S}'|\mathbf{S}, \rho_1) = F(\mathbf{S}|\mathbf{S}', \rho_1)$, for F the vM-UARS density. For a proposal for the parameter κ we will take $\log(\kappa^{j*}) \sim \mathrm{N}(\log(\kappa^{j-1}), \sigma_1^2)$. So, κ^{j*} is log-normal with parameters $\log(\kappa^{j-1})$ and σ_1^2 (a tuning parameter). If $t(x|\mu, \sigma^2)$ represents the log-normal density, then $\dfrac{t(\kappa'|\log(\kappa), \sigma_1^2)}{t(\kappa|\log(\kappa'), \sigma_1^2)} = \dfrac{\kappa}{\kappa'}$. The parameter τ is updated in a similar manner, with the tuning parameter σ_1^2 replaced by σ_2^2. For the parameter $\mathbf{V}$, we take as a candidate $\mathbf{V}^{j*} \sim \mathrm{FS}(\mathbf{V}^{j-1}, \rho_2)$, where $\mathrm{FS}(v, \beta)$ represents the Fisher distribution on the set of all unit vectors (i.e. the Fisher spherical distribution) with mean direction v and concentration parameter β (see Fisher, Lewis, and Embleton, 1987). Lastly, to update each p_i, $i = 1, \ldots, n$, we take as a proposal $p_i^{j*} \sim \mathrm{vM}(p_i^{j-1}, \eta)$, where $\mathrm{vM}(\alpha, \omega)$ is the von Mises circular distribution with mean α and concentration parameter ω (see Mardia and Jupp, 2000). We note that the choice of proposals for $\mathbf{V}$ and p_i are also symmetric.

Then, for observations $\mathbf{m}_i$, $i = 1, \ldots, n$, beginning with starting values $\mathbf{S}^0$, κ^0, $\mathbf{V}^0$, τ^0, and $\{\mathbf{p}_1^0, \ldots, \mathbf{p}_n^0\}$ we implement the algorithm for $j = 1, 2, \ldots$ as follows:

1. Generate $\mathbf{S}^{j*} \sim$ vM-UARS$(\mathbf{S}^{j-1}, \rho_1)$ as a proposal for $\mathbf{S}^j$.

2. Compute $a_j^1 = \dfrac{G(\mathbf{S}^{j*}, \kappa^{j-1}, \mathbf{V}^{j-1}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})}{G(\mathbf{S}^{j-1}, \kappa^{j-1}, \mathbf{V}^{j-1}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})}$ for G in (8) and generate $W_j^1 \sim$ Bernoulli$(\min(1, a_j^1))$.

3. Take $\mathbf{S}^j = W_j^1 \mathbf{S}^{j*} + (1 - W_j^1)\mathbf{S}^{j-1}$.

4. Generate $\log(\kappa^{j*}) \sim \mathrm{N}(\log(\kappa^{j-1}), \sigma_1^2)$, with κ^{j*} as a proposal for κ^j.

5. Compute $a_j^2 = \dfrac{G(\mathbf{S}^j, \kappa^{j*}, \mathbf{V}^{j-1}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})\kappa^{j*}}{G(\mathbf{S}^j, \kappa^{j-1}, \mathbf{V}^{j-1}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})\kappa^{j-1}}$ for G in (8) and generate $W_j^2 \sim$ Bernoulli$(\min(1, a_j^2))$.

6. Take $\kappa^j = W_j^2\kappa^{j*} + (1 - W_j^2)\kappa^{j-1}$.

7. Generate $\mathbf{V}^{j*} \sim \text{FS}(\mathbf{V}^{j-1}, \rho_2)$ as a proposal for $\mathbf{V}^j$, where FS represents the Fisher spherical distribution.

8. Compute $a_j^3 = \dfrac{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^{j*}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})}{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^{j-1}, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})}$ for G in (8) and generate $W_j^3 \sim$ Bernoulli$(\min(1, a_j^3))$.

9. Take $\mathbf{V}^j = W_j^3\mathbf{V}^{j*} + (1 - W_j^3)\mathbf{V}^{j-1}$.

10. Generate $\log(\tau^{j*}) \sim \text{N}(\log(\tau^{j-1}), \sigma_2^2)$, with τ^{j*} as a proposal for τ^j.

11. Compute $a_j^4 = \dfrac{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^j, \tau^{j*}, p_1^{j-1}, \ldots, p_n^{j-1})\tau^{j*}}{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^j, \tau^{j-1}, p_1^{j-1}, \ldots, p_n^{j-1})\tau^{j-1}}$ for G in (8) and generate $W_j^4 \sim$ Bernoulli$(\min(1, a_j^4))$.

12. Take $\tau^j = W_j^4\tau^{j*} + (1 - W_j^4)\tau^{j-1}$.

13. For $k = 1, \ldots, n$

    (a) Generate $p_k^{j*} \sim \text{vM}(p_k^{j-1}, \eta)$ as a proposal for p_k^j.

    (b) Compute $b_j^k = \dfrac{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^j, \tau^j, p_1^j, \ldots, p_{k-1}^j, p_k^{j*}, p_{k+1}^{j-1}, \ldots, p_n^{j-1})}{G(\mathbf{S}^j, \kappa^j, \mathbf{V}^j, \tau^j, p_1^j, \ldots, p_{k-1}^j, p_k^{j-1}, p_{k+1}^{j-1}, \ldots, p_n^{j-1})}$ for G in (8) and generate $U_j^k \sim$ Bernoulli$(\min(1, b_j^k))$.

    (c) Take $p_k^j = U_j^k p_k^{j*} + (1 - U_j^k)p_k^{j-1}$.

Once we have have used the above algorithm to obtain a large number of posterior draws (after an appropriate burn-in period), we can obtain point estimates and credible regions for each of the PARS parameters. For concreteness, we consider 95% credible sets. Suppose

we have a set of N posterior values for κ, $\Delta_\kappa = \{\kappa^1, \ldots, \kappa^N\}$. Then a point estimate for κ is taken to be $\hat{\kappa} = 1/N \sum_{i=1}^{N} \kappa^i$ and a 95% credible region for κ is taken to be $[\kappa_{.025}, \kappa_{.975}]$, where $\kappa_{.025}$ is the .025 quantile of Δ_κ and $\kappa_{.975}$ is the .975 quantile of Δ_κ. A point estimate and 95% credible interval for τ can be found from $\Delta_\tau = \{\tau^1, \ldots, \tau^N\}$ in a similar manner.

For the parameter $\mathbf{S}$ a point estimate is taken to be the matrix $\hat{\mathbf{S}}$ that minimizes $tr(\hat{\mathbf{S}}\bar{\mathbf{S}})$, where $\bar{\mathbf{S}} = 1/N \sum_{i=1}^{N} \mathbf{S}^i$ and $\Delta_S = \{\mathbf{S}^1, \ldots, \mathbf{S}^N\}$ is the set of posterior values for $\mathbf{S}$. The matrix $\hat{\mathbf{S}}$ is the mean direction of the matrices in Δ_S (Khatri and Mardia, 1977, p. 96). We use the 95% "set-of-cones" credible region of Bingham, Vardeman, and Nordman (2009) for $\mathbf{S}$. For $i = 1, \ldots, N$, we find the angle between each of the coordinate axes represented by $\hat{\mathbf{S}}$ with the corresponding axis represented by $\mathbf{S}^i$ and let δ^i represent the maximum of these three angles (i.e. δ_i is the maximal absolute diagonal element of $\hat{\mathbf{S}}^T\mathbf{S}^i$). Then, cones of constant angle $\delta_{.95}$ around the axes represented by $\hat{\mathbf{S}}$, where $\delta_{.95}$ = the .95 quantile of $\{\delta^1, \ldots, \delta^N\}$, create a region capturing 95% of sets of axes represented by matrices in Δ_S.

Lastly, let $\mathbf{V}^1, \ldots, \mathbf{V}^N$ be the posterior draws for the parameter $\mathbf{V}$. A point estimate for $\mathbf{V}$ is taken to be the line $\pm\hat{\mathbf{V}}$, where $\hat{\mathbf{V}}$ maximizes $\sum_{i=1}^{N} \left|\hat{\mathbf{V}}^T\mathbf{V}^i\right|$. Now, let γ^i be the angle between $\pm\hat{\mathbf{V}}$ and $\mathbf{V}^i$, $i = 1, \ldots, N$. If $\gamma_{.95}$ is the .95 quantile of $\{\gamma^1, \ldots, \gamma^N\}$, then cones of constant angle around $\pm\hat{\mathbf{V}}$ serve as a 95% credible region for the axis of asymmetry.

The Bayes methods outlined above provide inference for the (interpretable) parameters of PARS models. By employing the Metropolis-Hastings within Gibbs algorithm to simulate posterior values, there is no need to represent $\mathbf{S}$ using alternate parameterizations such as quaternions or Euler angles. The Bayes methodology also provides a direct approach to inference that is much more tractable than maximum likelihood would be (even in cases where it is well-defined), and in the next section we apply this Bayes methodology to the calcaneocuboid joint data.

4 Application to Movement of the Calcaneocuboid Joint

In the application considered here, we are interested in examining movement about the calcaneocuboid joint for a human, a chimpanzee, and a baboon. Each subject was placed in a sitting position with the foot flat on the floor. Load was then applied to the knee and movement was tracked by using infra-red emitting diode markers attached to the foot at the location of the calcaneous bone and the location of the cuboid bone. For each of the subjects, one application of load resulted in 125 observations giving the orientation of each of the markers at that time point. Although the data for each subject have been collected over time, for purposes of illustration we treat these observations as if they were independent and identically distributed. If the orientation of the calcaneous bone is represented as $\mathbf{B}$ and the orientation of the cuboid bone is represented as $\mathbf{M}$, then the resulting orientation that measures the relative orientation of the calcaneous bone with respect to the cuboid bone is given as $\mathbf{B}^T\mathbf{M}$. These relative orientations are the data that are analyzed here.

As mentioned in Section 1, although the matrix Fisher distribution might be used for nonsymmetric modeling of rotations, its parameters are not easily interpreted. To illustrate this, the matrix Fisher distribution was fit to each of the calcaneocuboid data sets using the derivations given by Khatri and Mardia (1977). For the case of 3×3 rotations, the matrix Fisher distribution has density $a(\mathbf{F}) \exp(tr(\mathbf{F}^T\mathbf{o}))$, $\mathbf{o} \in \Omega$ with respect to the Haar measure on Ω where $a(\mathbf{F})$ is a normalizing constant with $\mathbf{F}$ a 3×3 matrix of full rank. The parameter $\mathbf{F}$ can be decomposed as $\mathbf{F} = \mathbf{KM}$ where $\mathbf{M}$ is the "polar component" (sometimes called the "mean direction" as in Downs (1972)) and $\mathbf{K}$ is the "elliptic component" (Khatri and Mardia, 1977). For data $\mathbf{o}_1, \ldots, \mathbf{o}_n$, the maximum likelihood estimates of $\mathbf{M}$ and $\mathbf{K}$ can be obtained by considering the singular value decomposition of $\bar{\mathbf{o}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{o}_i$. This decomposition yields $\bar{\mathbf{o}} = \hat{\mathbf{\Delta}}\mathbf{D}_{\mathbf{g}}\hat{\mathbf{\Gamma}}$, where $\mathbf{D}_{\mathbf{g}}$ is a diagonal matrix with entries $\mathbf{g} = (g_1, g_2, g_3)$. The maximum likelihood estimate for $\mathbf{M}$ is $\hat{\mathbf{M}} = \hat{\mathbf{\Delta}}\hat{\mathbf{\Gamma}}$. Suppose $\mu_i(\boldsymbol{\phi}) = (\partial/\partial\phi_i)_0 F_1(3/2, \frac{1}{4}\mathbf{D}_{\boldsymbol{\phi}}^2)$ for $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3)$

where ${}_0F_1$ is the hypergeometric function with matrix argument. Then for $\hat{\boldsymbol{\phi}} = (\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3)$, solving $g_i = \mu_i(\hat{\boldsymbol{\phi}})$, $i = 1, 2, 3$, the maximum likelihood estimate for $\mathbf{K}$ is $\hat{\mathbf{K}} = \hat{\boldsymbol{\Delta}}\mathbf{D}_{\hat{\boldsymbol{\phi}}}\hat{\boldsymbol{\Delta}}^T$ (Khatri and Mardia, 1977). The estimates for $\mathbf{M}$ and $\mathbf{K}$ are given in Table 1 for each of the calcaneocuboid data sets.

While the polar components, $\mathbf{M}$, do provide an interpretable representation of where the data are centered (and are similar to the central orientations, $\mathbf{S}$, that will be obtained from fitting the PARS distributions), the elliptic components provide no clearly interpretable insight as to which of the three data sets exhibits the highest degree of asymmetry nor anything about "directions" of asymmetry. The matrix Fisher fit does not allow for subject matter comparison of the three calcaneocuboid data sets. Because the geometric construction of the PARS models gives easily interpretable parameters and the Bayes methodology of Section 3 provides a simple inference approach for estimating these parameters, we now turn our attention to fitting the PARS distributions to calcaneocuboid data.

The Bayes analyses of Section 3 were applied to each of the three data sets, where spins r and p were taken to have a von Mises distribution, so that $C(r|\kappa) = [\sqrt{2\pi}I_0(\kappa)]^{-1}\exp(\kappa\cos r)$ and $D(p|\tau) = [\sqrt{2\pi}I_0(\tau)]^{-1}\exp(\tau\cos p)$, for $I_0(\cdot)$ the modified Bessel function of order 0. A total of 10000 posterior simulations were obtained after a burn-in of 10000 iterations. Bayes point estimates for κ, τ, and $\pm\mathbf{V}$ are given in Table 3 and the Bayes estimates for $\mathbf{S}$ are given in Table 2. Because the markers may have not been placed comparably on all subjects, comparison of the estimated central orientations $\hat{\mathbf{S}}$ and the estimated preferred directions $\hat{\mathbf{V}}$ across subjects are not of subject matter interest. Though the estimated preferred axes of rotation may not be comparable across subjects, the preferred axis still has physical meaning for each subject. In this application, the preferred axis can be thought of as the primary rotational axis for the calcaneocuboid joint. The primary rotational axis is the anatomical axis about which observed movement occurs, as opposed to one of the axes of the reference frame (Ball and Greiner, 2011).

Figure 5 shows the data for each subject. Since the estimated central orientations $\hat{\mathbf{S}}$ are not comparable across subjects, each data set was rotated by its $\hat{\mathbf{S}}$ so all data sets approximately share central orientation at $\mathbf{I}_{3\times3}$ (represented by x, y, z in the figure). We can see that spread and amount of asymmetry differ across subjects. Figure 6 shows von Mises density curves for the fitted values of κ and τ given in Table 3. We see that the baboon data is the most concentrated around the central orientation (has the largest value of $\hat{\kappa}$) with the least amount of asymmetry (has the largest value of $\hat{\tau}$). While the human data also resulted in a fairly large value of $\hat{\kappa}$, it exhibits the highest degree of asymmetry (with the smallest $\hat{\tau}$ at 68.68). To illustrate the plausibility of the fitted PARS distributions, 125 PARS observations were generated using the estimated parameters from the chimp data. Figure 7 displays these observations in a similar manner to the actual chimp data from Figure 5 (c). We see that the simulated data closely resemble the actual data, providing some evidence of the adequacy of the PARS distribution with von Mises spins for describing these data. (Although not included here, plots similar to Figures 5 (a) and (b) were also obtained for simulated data based on the human and baboon parameter estimates.)

To more completely compare subjects, we consider credible regions for our parameters. Table 4 gives the 95% credible intervals for κ and τ along with the sizes of the cone angle for the credible regions for $\mathbf{V}$ and $\mathbf{S}$. We can see that none of the 95% credible intervals for τ overlap, so that each subject exhibits a different degree of asymmetry when considering movement around the calcaneocuboid joint, with the baboon showing much less asymmetry than the chimp and human. We also see that the baboon shows less variability in movement about the joint than the human and chimp, as evidenced by a credible region for κ that lies to the right of the (overlapping) regions for the human and chimp.

It might be expected that chimps and baboons would have similar movement about the calcaneocuboid joint, because the structure of their feet differs from that of a human foot. Human feet do not have the ability to effectively grasp and manipulate objects because toes

are shorter, and the arches are relatively higher in the human foot than those in chimps. However, chimps and humans are both part of the Hominidae family (sometimes referred to as Great Apes), while the baboon is part of the Cercopithecidae family (Old World Monkeys). Thus, despite the differences in the foot structure of humans and chimps, we see that they exhibit movement around the calcaneocuboid joint that is more alike than is that of the human and the baboon or the chimp and the baboon.

To provide evidence that the Bayes methodology presented here behaves as expected based on properties of similar methods for UARS models (with actual frequentist confidence levels approximating 95%), a small simulation study was done. Representing each subject, 100 samples of size $n = 125$ were generated from the PARS distribution using the estimated parameter values. For each of these 100 samples, a sample of size 8000 was generated from the posterior (8) (taken after a burn-in of 2000 iterations) using the algorithm outlined in Section 3 with starting values set at the true parameters. The 8000 posterior draws were used to find 95% credible regions for each parameter and it was determined whether the region captured the true parameter value (from which the data were generated). Empirical coverage rates are given in Table 5. The percentages in the table indicate that all credible regions are holding their nominal coverage rates and the Bayes methods are performing as desired for the parameter values estimated from the human, baboon, and chimp data. Median sizes of the credible regions (for the 100 samples) were also found for each case and are presented in Table 6. For κ and τ, the size of the credible region is expressed as the width of the credible interval. For $\mathbf{V}$, the size of the credible region is expressed as the angle between $\hat{\mathbf{V}}$ and the edge of the cones representing the credible region. For $\mathbf{S}$, the size of the credible region is expressed as the angle between any axis of $\hat{\mathbf{S}}$ and the edge of the cone around that axis.

5 Conclusion

The PARS distributions developed in this paper build upon the symmetric UARS distributions of Bingham, Nordman, and Vardeman (2009) by allowing for asymmetry. While the matrix Fisher distribution for orientations has been previously studied in the literature (Downs, 1972; Khatri and Mardia (1977); Jupp and Mardia (1979)), parameters of the distribution are not easily estimated or interpreted. Through the constructive definition of the PARS models, we have arrived at a class of distributions with directly interpretable parameters representing the location, spread, degree of asymmetry, and direction of asymmetry for a set of 3-dimensional orientations. The Bayes methodology developed for the PARS distributions provides a direct form of inference for estimating these parameters. This has proved valuable in our application, allowing us to compare rotations about the calcaneocuboid joint for a human, a baboon, and a chimp. Since rotation data naturally arise when considering movement around joints, the methods developed here have potential application in a variety of studies in the field of biomechanical motion. As such rotations are often collected over time, further work will involve using the UARS or PARS models in a time series context.

References

[1] Ball, K. A., and Greiner, T. M. (2011) "A Procedure to Refine Joint Kinematic Assessments: Functional Alignment," *Computer Methods in Biomechanics and Biomedical Engineering*, to appear.

[2] Bingham, M. A., Nordman, D. J., and Vardeman, S. B. (2009) "Modeling and Inference for Measured Crystal Orientations and a Tractable Class of Symmetric Distributions for Rotations in Three Dimensions," *Journal of the American Statistical Association*, 104, 1385–1397.

[3] Bingham, M. A., Vardeman, S. B., and Nordman, D. J. (2009) "Bayes One-Sample and One-Way Random Effects Analyses for 3-D Orientations with Application to Materials Science," *Bayesian Analysis*, 4, 607–630.

[4] Downs, T. D. (1972), "Orientation Statistics," *Biometrika*, 59, 665–676.

[5] Fisher, N. I., Lewis, T., and Embleton, B. J. J. (1987), *Statistical Analysis of Spherical Data*, Cambridge; New York: Cambridge University Press.

[6] Jupp, P. E., and Mardia, K. V. (1979), "Maximum Likelihood Estimators for the Matrix Von Mises-Fisher and Bingham Distributions," *The Annals of Statistics*, 7, 599–606.

[7] Khatri, C. G., and Mardia, K. V. (1977), "The Von Mises-Fisher Matrix Distribution in Orientation Statistics," *Journal of the Royal Statistical Society*, Ser. B, 39, 95–106.

[8] Mardia, K. V., and Jupp, P. E. (2000), *Directional Statistics*, Chichester; New York: John Wiley & Sons.

[9] Prentice, M. J. (1986), "Orientation Statistics without Parametric Assumptions," *Journal of the Royal Statistical Society*, Ser. B, 48, 214–222.

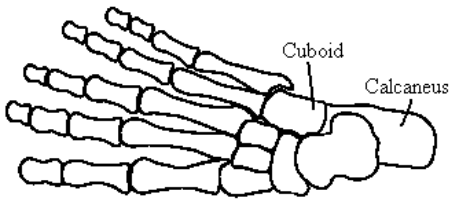
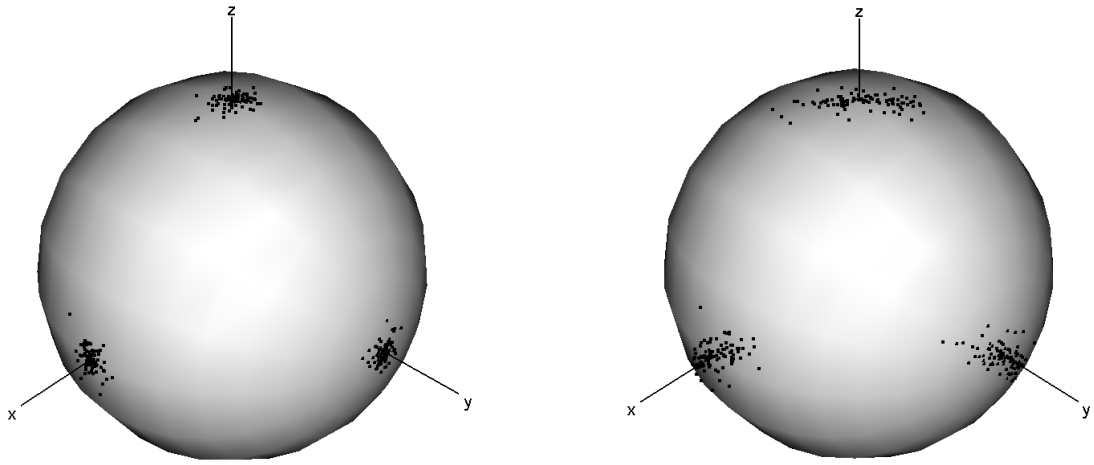
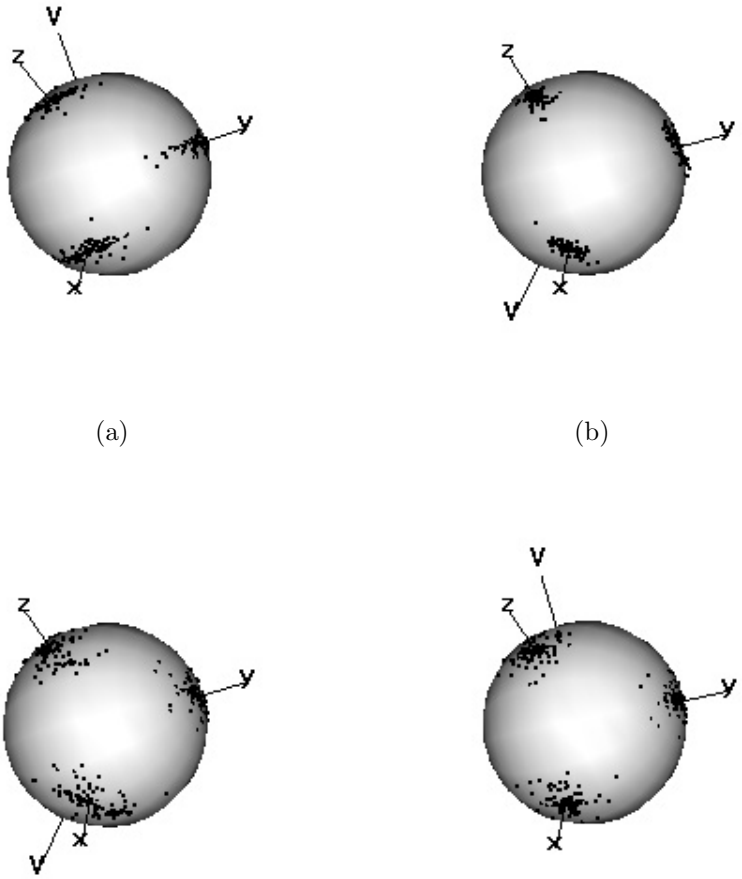


Figure 1: The calcaneous and cuboid bones of the foot (Image adapted from one at http://www.pt.ntu.edu.tw/hmchai/surfaceanatomy)



(a) (b)

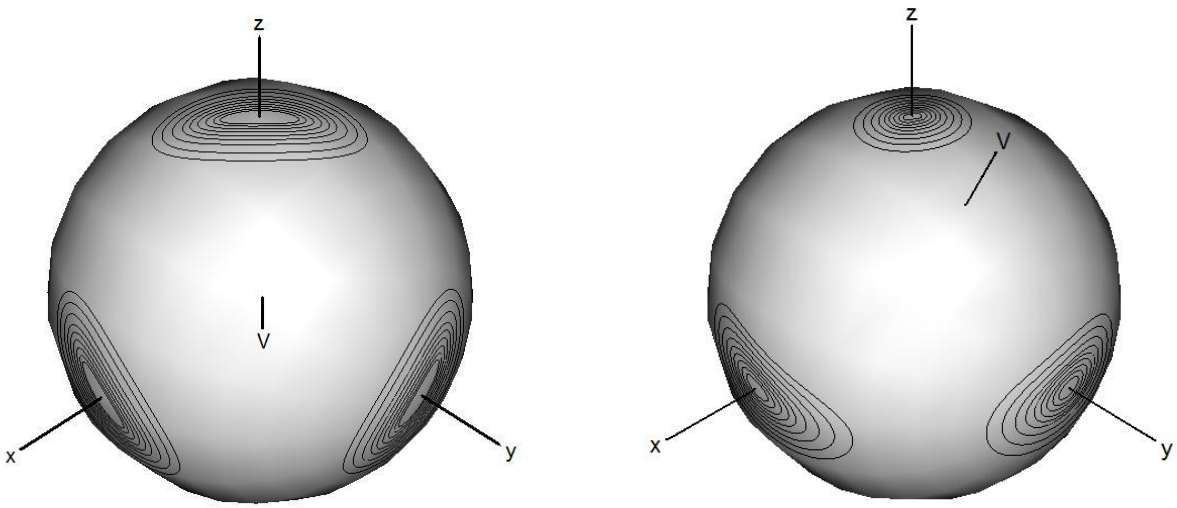
Figure 2: 100 random observations plotted around a central orientation represented by the perpendicular axes x, y, and z that are (a) symmetric and (b) nonsymmetric in nature



(a) (b)

(c) (d)

Figure 3: 100 PARS observations plotted around the principal direction at $\mathbf{I}_{3\times3}$ represented by the perpendicular x, y, and z axes using (a) $\kappa = 100, \tau = 20$, (b) $\kappa = 100, \tau = 100$, (c) $\kappa = 20, \tau = 20$, and (d) $\kappa = 20, \tau = 100$



(a) (b)

Figure 4: Contours capturing $10 \cdot i\%$ of the distribution, according to density, for the PARS($\mathbf{I}_{3\times 3}, 100, \mathbf{V}, 15$) distribution with (a) $\mathbf{V} = (1,1,1)/\sqrt{3}$ and (b) $\mathbf{V} = (1,3,5)/\sqrt{35}$

Table 1: Maximum likelihood estimates of the matrix Fisher components $\mathbf{M}$ and $\mathbf{K}$ for the calcaneocuboid data, by subject

	$\mathbf{M}$	$\mathbf{K}$
Human	$\begin{pmatrix} -0.104 & -0.919 & -0.381 \\ 0.332 & -0.393 & 0.858 \\ -0.938 & -0.037 & 0.346 \end{pmatrix}$	$\begin{pmatrix} 4113.9 & -891.5 & -387.7 \\ -891.5 & -1784.2 & 3579.7 \\ -387.7 & 3579.7 & 3281.1 \end{pmatrix}$
Baboon	$\begin{pmatrix} 0.906 & -0.369 & -0.209 \\ 0.388 & 0.523 & 0.759 \\ -0.171 & -0.768 & 0.617 \end{pmatrix}$	$\begin{pmatrix} 282.2 & -781.4 & -1506.6 \\ -718.4 & 1478.5 & -1218.7 \\ -1506.6 & -1218.7 & 1808.8 \end{pmatrix}$
Chimp	$\begin{pmatrix} -0.803 & 0.554 & -0.220 \\ 0.395 & 0.771 & 0.500 \\ 0.446 & 0.314 & -0.838 \end{pmatrix}$	$\begin{pmatrix} 50.9 & -138.1 & 287.7 \\ -138.1 & 850.4 & 141.3 \\ 287.7 & 141.3 & 120.9 \end{pmatrix}$

Table 2: Bayes estimates of the PARS central orientation $\mathbf{S}$ for the calcaneocuboid data, by subject

Human	$\begin{pmatrix} -0.109 & -0.916 & -0.385 \\ 0.324 & -0.399 & 0.858 \\ -0.940 & -0.031 & 0.340 \end{pmatrix}$
Baboon	$\begin{pmatrix} 0.906 & -0.366 & -0.211 \\ 0.388 & 0.521 & 0.761 \\ -0.168 & -0.771 & 0.614 \end{pmatrix}$
Chimp	$\begin{pmatrix} -0.798 & 0.557 & -0.230 \\ 0.396 & 0.771 & 0.498 \\ 0.455 & 0.307 & -0.836 \end{pmatrix}$

Table 3: Bayes estimates for κ, τ, and $\pm\mathbf{V}$ for the calcaneocuboid data, by subject

	$\hat{\kappa}$	$\hat{\tau}$	$\pm\hat{\mathbf{V}}$
Human	601.13	68.68	$\pm(-0.517,\ 0.249,\ 0.819)$
Baboon	1068.26	563.16	$\pm(0.581,\ 0.795,\ -0.175)$
Chimp	285.99	167.37	$\pm(0.601,\ 0.647,\ 0.469)$

Table 4: 95% Bayes credible regions for κ and τ, and cone angles for the 95% regions for $\mathbf{V}$ and $\mathbf{S}$

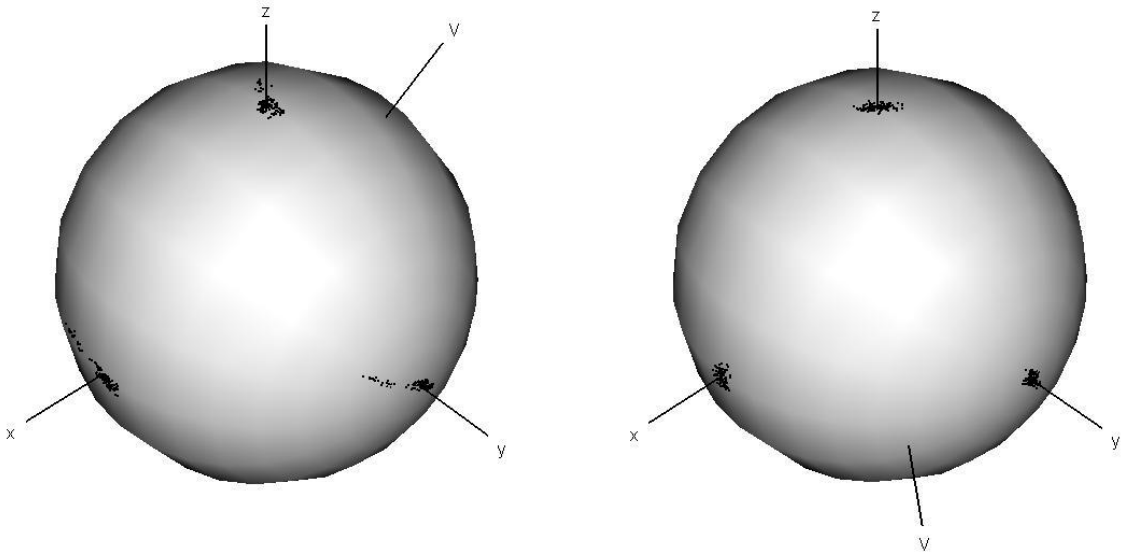
	κ	τ	$\mathbf{V}$	$\mathbf{S}$
Human	(346.40, 766.99)	(55.04, 85.57)	0.0080	0.0069
Baboon	(814.80, 1465.19)	(420.24, 691.92)	0.1070	0.0046
Chimp	(220.17, 361.41)	(119.20, 222.05)	0.1258	0.0280

Table 5: Coverage rates of 95% Bayes credible regions (based on 100 simulated samples) for parameters of the PARS distributions

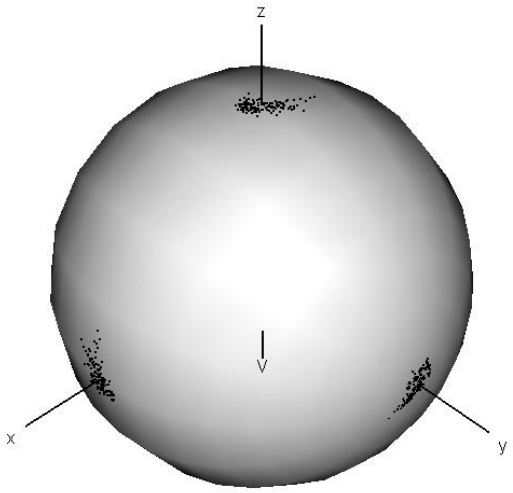
	κ	τ	$\mathbf{V}$	$\mathbf{S}$
Human	92%	94%	97%	96%
Baboon	94%	96%	93%	93%
Chimp	96%	96%	93%	95%

Table 6: Median size of Bayes credible sets for parameters of the PARS distributions

	κ	τ	$\mathbf{V}$	$\mathbf{S}$
Human	384.51	32.36	0.012	0.008
Baboon	662.85	251.15	0.084	0.004
Chimp	126.42	81.70	0.108	0.024

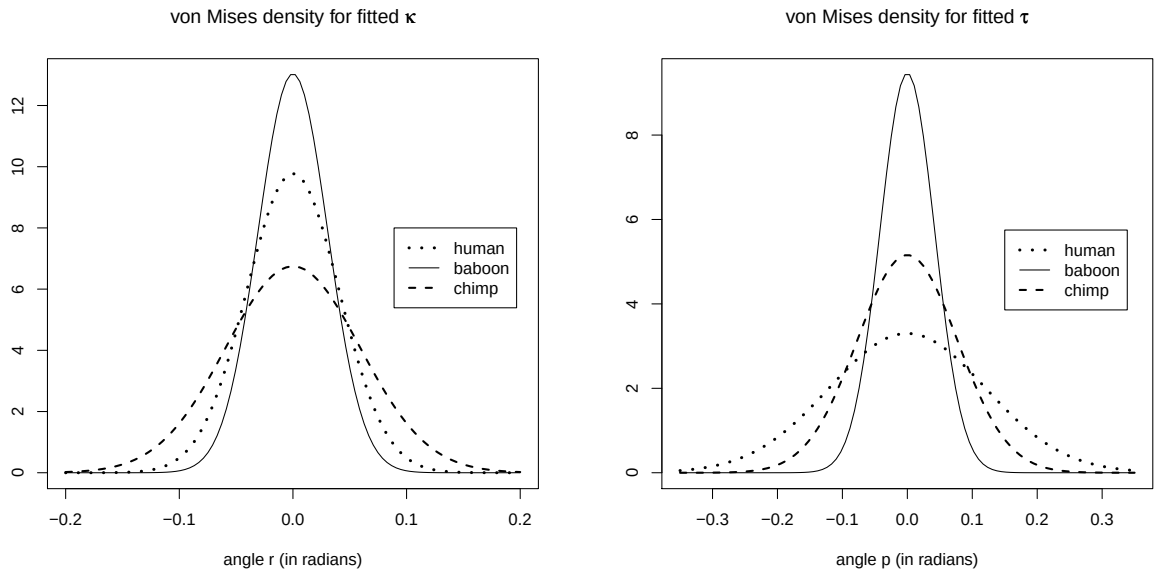


(a) (b)



(c)

Figure 5: 125 rotations about the calcaneocuboid joint, plotted around $\mathbf{I}_{3\times3}$ (represented by axes x, y, and z), for the (a) human, (b) baboon, and (c) chimp



(a) (b)

Figure 6: von Mises densities for the fitted values of (a) κ and (b) τ

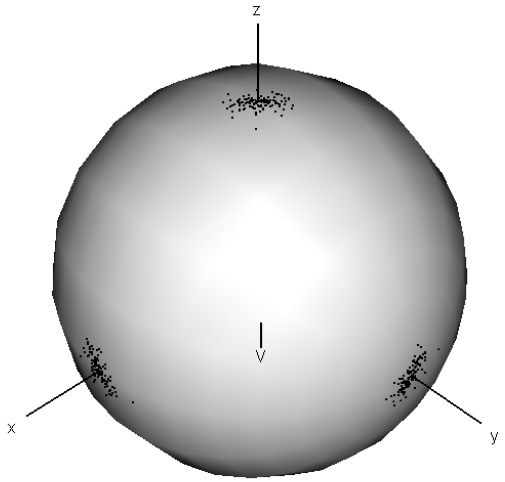


Figure 7: 125 rotations simulated using the parameter estimates for the chimp data