

Lectures in Paleomagnetism, 2005
by Lisa Tauxe

Citation:
<http://earthref.org/MAGIC/books/Tauxe/2005/>

June 8, 2005

Chapter 12

Beyond Fisher statistics

Suggested Reading

For background:

Chapter 1: Borradaile (2003)

Efron, B. and Tibshirani (1982)

Tauxe, L. et al. (1990)

To learn more:

Chapters 2-5: Fisher, et al. (1987)

12.1 Introduction

Paleomagnetists have depended since the 1950's on the special statistical framework developed by Fisher (1953) for the analysis of unit vector data. The power and flexibility of a variety of tools based on Fisher statistics enables quantification of parameters such as the degree of rotation of a crustal block, or whether the geomagnetic field really averages to a geocentric axial dipole independent of polarity. These tools, however, require that the paleomagnetic data belong to a particular parametric distribution - the Fisher distribution.

In many important situations, the Fisher distribution fails to represent paleomagnetic data adequately. To begin with, it is likely that the geomagnetic field itself produces directions that are far from Fisher distributed (Creer et al., 1959). Collections of paleomagnetic directions from the last five million years from around the globe (e.g., McElhinny and McFadden, 1997) suggest that data from equatorial regions (Figure 12.1a) are in fact significantly elongate while those from high latitudes are more symmetric (Figure 12.1c).

The failure of the Fisher model for directional data sets makes sense if the geomagnetic field is essentially controlled by a geocentric axial dipole with “noise” added by other terms, a topic we will cover in more detail in later lectures. Such models generate spherically symmetric distributions of the VGPs. When converted to equivalent directions are more elongate as the observation site approaches the equator (see Figure 12.1b,c). Because VGPs that are farther from the pole are associated with weaker field strengths in collections of paleomagnetic data and in many models of the field, the Fisher assumption of unit vector length over emphasizes the importance of the “outliers” and leads to mean inclinations that are shallower than the true mean (see e.g., Creer, 1983).

Another example of the inadequacy of the Fisher distribution is the fact that the magnetic field

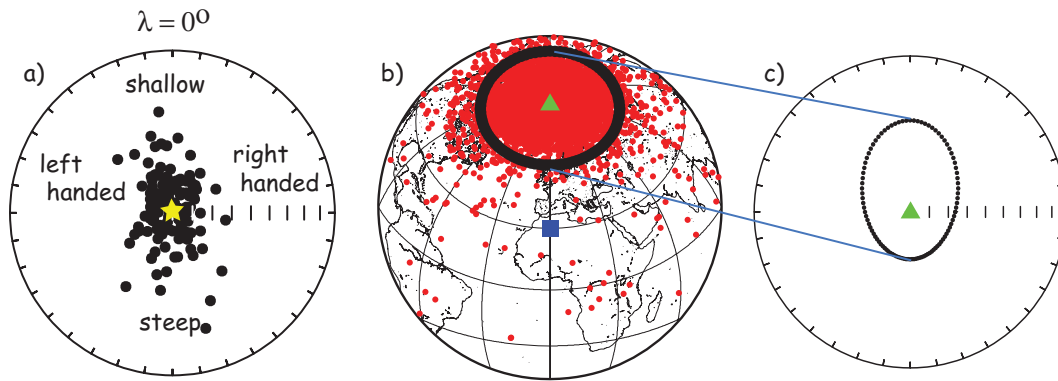


Figure 12.1: a) Paleomagnetic directions from the PSVRL database (see McElhinny and McFadden (1997)) compiled for latitude band 0-5° (N&S). Antipodes of reverse directions are used. Directions rotated to the expected direction (star: $D = 0^\circ, I = 0^\circ$) using the D', I' transformation in Lecture 2. Directions in the upper (lower) half of the diagram are shallower (steeper) than expected and those to the right (left) are right-handed (left-handed). b) VGPs from geomagnetic vectors evaluated from the statistical field model of Tauxe and Kent (2004) at 30°N (site of observation shown as square). The geographic pole is shown as a triangle. A set of VGP positions at 60° N are shown as the black ring. c) Directions observed at the site of observation [square in b)] converted from black ring of VGPs in a) which correspond to the VGP positions at 60°N. These directions have been projected along expected direction at site of observation (triangle). Note that a circularly symmetric ring about the geographic pole gives an asymmetric distribution of directions with a shallow bias. [Figures from Tauxe and Kent, 2004.]

exists in two stable polarity states. Because the Fisher distribution allows only uni-modal data, bipolar data must be separated into separate modes or one mode must be “flipped” to the antipode prior to calculating a mean. Furthermore, remanence vectors composed of several components tend to form streaked distributions. Similarly, structural complications (e.g., folding) can lead to streaked distributions of directional data.

Thus, non-Fisherian data are a fact of paleomagnetic life. The Fisher-based tests can frequently be inappropriate and could result in flawed interpretations. In Lecture 11 we learned the basics of Fisher statistics and how to test data sets against a Fisher model. In this lecture, we will discuss what to do when Fisher statistics fail. We will begin with essentially parametric approaches that treat certain types of non-fisherian data. We then turn to the use of non-parametric methods such as the bootstrap and jackknife in paleomagnetic applications.

12.2 Non-fisherian approaches to paleomagnetic vectors

12.2.1 The Kent Distribution

The D', I' data from equatorial sampling sites from the last five million years shown in Figure 12.1a have a more elliptical distribution than the symmetrical distribution required for a Fisherian data set. To treat such data, it is probably inappropriate to use a Fisher confidence ellipse and a distribution that allows data with elliptical directional dispersion would be better. The elliptical

12.2. NON-FISHERIAN APPROACHES TO PALEOMAGNETIC VECTORS

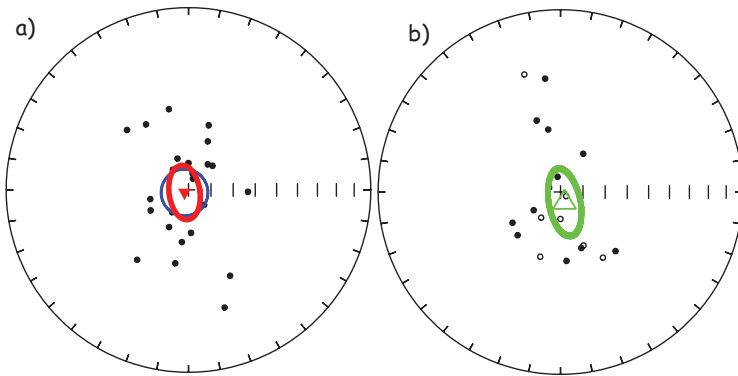


Figure 12.2: A subset of data from Figure 12.1a plotted with the Fisher circle of confidence (light blue line) and the Kent 95% confidence ellipse (heavy red line).

analogue of the Fisher distribution is the Kent distribution (Kent, 1982) and is given by:

$$F = c(\kappa, \beta)^{-1} \exp(\kappa \cos \alpha + \beta \sin^2 \alpha \cos 2\phi)$$

where α is the angle between a given direction and the true mean direction (estimated by the Fisher mean (see Lecture 11), and ϕ is an angle in the plane perpendicular to the true direction with $\phi = 0$ parallel to the major eigenvector $\mathbf{V}_2$ in that plane. κ is a concentration parameter similar to the Fisher κ and β is the “ovalness” parameter. $c(\kappa, \beta)$ is a complicated function of κ and β . When β is zero, the Kent distribution reduces to a Fisher distribution. Details of the calculation of Kent 95% confidence ellipse are given in the Appendix.

If we were to collect data from the equatorial region, we might well obtain a set of directions such as those shown in Figure 12.1b. The Fisher α_{95} circle of confidence for this data set is shown as the light blue line and the Kent ellipse as the heavy red line. The Kent ellipse clearly represents the distribution of data better than the Fisher α_{95} , being elongate in the same sense as the data themselves.

12.2.2 The Bingham distribution

The Kent distribution has the advantage that it can deal with elliptical data sets while the Fisher distribution cannot. However, many paleomagnetic data sets are also bi-modal and the Kent and Fisher distributions can only deal with data sets with a single polarity. It was precisely for treating bimodal, elliptical data that the Bingham distribution was developed (Bingham, 1974). The Bingham distribution is given by:

$$F = \frac{1}{4\pi d(k_1, k_2)} \exp(k_1 \cos^2 \phi + k_2 \sin^2 \phi) \sin^2 \alpha$$

where α and ϕ are as in the Kent distribution, k_1, k_2 are concentration parameters ($k_1 < k_2 < 0$) and $d(k_1, k_2)$ is a constant of normalization. Values for k_1, k_2 can be estimated by numerical integration and can be converted into 95% confidence ellipses, the details of which are given in the Appendix. In a nut shell, the $\mathbf{V}_1$ eigenvector of the orientation matrix (associated with the largest eigenvalue, see Lecture 9) is the mean direction and the 95% confidence ellipse semi-axes are proportional to

the intermediate and minimum eigenvalues. The Bingham mean therefore is not necessarily the same as the Fisher or Kent mean. If we take each vector end-point to be a mass, the Bingham mean is the axis about which the moment of inertia would be least. The Fisher mean is somewhat different, in that it is the vector sum of unit vectors. The Bingham mean is less effected by outliers than the Fisher mean, lying closer to the center of mass of data points.

The principle drawback of the Bingham distribution is that because the orientation matrix uses the entire data set (normal and reverse) the two modes are assumed to be antipodal and to share the same distribution parameters. The question of whether normal and reverse data sets are antipodal and have the same dispersion is in fact a question we may wish to ask! One could separate the two modes prior to calculation of the Bingham ellipse, but then the rationale for using the Bingham distribution is lost.

12.2.3 The Bingham-Le Goff approximation

Estimating the parameters for the Bingham ellipse exactly is computationally taxing and all of the available “canned” programs use the look up table of Mardia and Zemloch (1977; see Appendix). Le Goff et al. (1992) suggested some approximations which may be valid for concentrated distributions. They also introduced the concept of weighting results according to some reliability criteria. For the general case, however, it seems preferable to use the exact Kent (1982) ellipses on uni-modal data sets. These could of course be weighted if such weighting is desired.

12.2.4 The bi-Gaussian Distribution

Until now we have continued the Fisher assumption of unit vectors. As already mentioned, neglect of the vector strength can lead to bias. Love and Constable (2003) began the hard work of incorporating intensity information into the parameter estimation problem. Their method can handle bi-modal spherical Gaussian data such as those shown in Figure 12.3. Estimation of the Love parameters are beyond the scope of these lecture notes. Moreover, many data sets are not spherically symmetric as already noted and the Love and Constable (2003) approach must be generalized to elliptical, more “blade-like” data sets than the “cotton balls” currently treatable.

12.3 The simple bootstrap

As we have mentioned, “real” data may be pathological in several respects including bi-modal and elliptically distributed data. Neither the Kent nor the Bingham statistical methods have the test for common mean so critical to paleomagnetic studies. Moreover, none of the methods we have described so far can provide confidence ellipses for an off-center mean direction as is likely to occur in records of the geomagnetic field (see Figure 12.1c). Finally, data may be overprinted or contain the record of a paleomagnetic transition, resulting in “streaked” or non-antipodal mean directions, conditions that make the conventional methods inappropriate. In this section we will discuss alternative methods for estimating confidence bounds which is sufficiently flexible to accomodate all of these short comings, provided the data set is large enough.

In Figure 12.4a we show a fairly typical “not great” paleomagnetic data set. The data are elliptical, bi-modal and one has the suspicion that the normal and reverse modes may be neither antipodal nor share the same concentration or ovalness parameters. Clearly some non-parametric approach would be desirable. The approach for characterizing uncertainties for vectors we will take

12.3. THE SIMPLE BOOTSTRAP

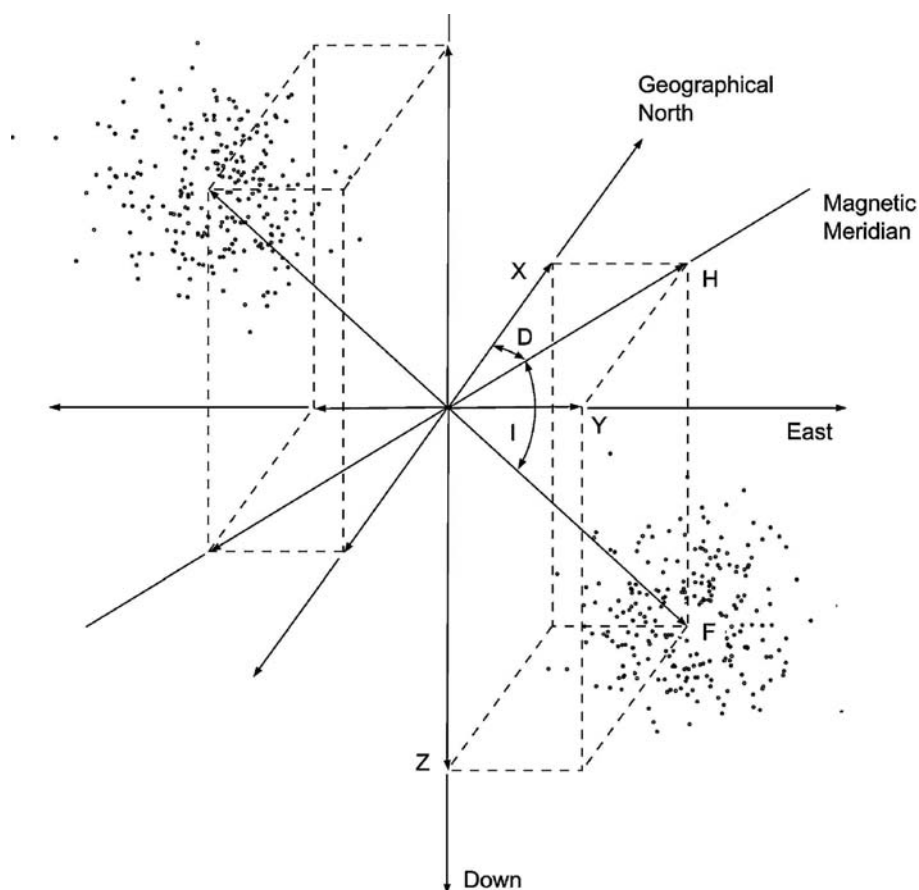


Figure 12.3: A bi-gaussian set of vectors suitable for treatment using the method of Love and Constable (2003). [Figure from Love and Constable, 2003.]

here is based on a technique known as the statistical *bootstrap*. As we shall see, the bootstrap has the flexibility to allow us to treat awkward data sets like that shown in Figure 12.4a.

The principles of the bootstrap are given in the Appendix. In essence, the parameter of interest (say, the mean vector) is calculated from many resampled data sets, whose data points are selected at random from the original data. The bootstrapped estimates “map out” the likely distribution of the parameter, allowing estimation of confidence regions. Before we extend the bootstrap from the scalar treatment in the appendix to vectors, it is important to point out that with the bootstrap, it is assumed that the underlying distribution is represented by the data, demanding that the data sets be rather large. Moreover, the bootstrap estimates are only asymptotically true, meaning that a large number of bootstrap calculations are required for the confidence intervals to be valid. It’s a good thing we have fast computers with huge hard-drives.

There are a variety of ways we can use the bootstrap to estimate confidence regions for paleomagnetic data. We will start with the most “Fisher” like approach of taking unit vectors of a single polarity. Then we will accommodate dual polarity data sets and develop analogous tests to those so useful for Fisher distributions.

To do a simple bootstrap on a data set with only one polarity (say the normal data in Figure 12.4a, we first randomly draw N data points from the data shown in Figure 12.4a. Each set

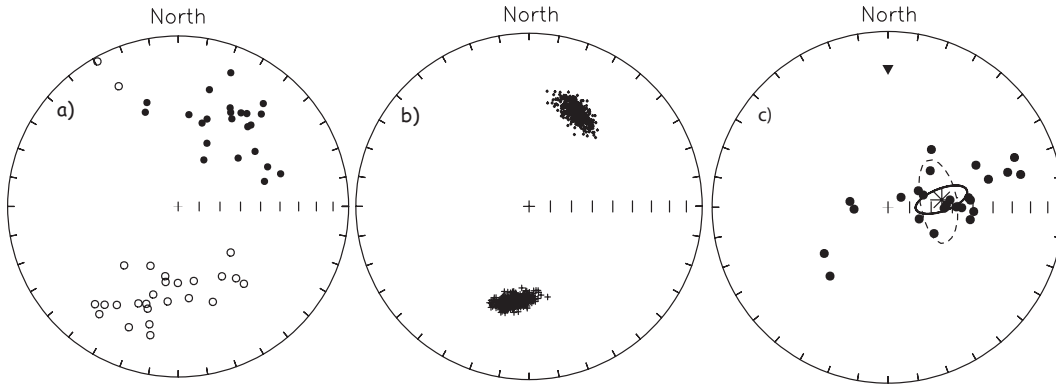


Figure 12.4: Hypothetical non-Fisherian data set. Normal and reversed polarity data that are not symmetrically distributed. Filled (open) circles plot on the lower (upper) hemisphere. b) Equal area projection of 500 bootstrapped means for para-data sets drawn from the data shown in a) Tiny circles (plusses) are lower (upper) hemisphere projections from the normal (reversed) polarity mode of each para-data set. c) Normal polarity data from a) transformed into VGPs and plotted in polar equal area projection. The dashed ellipse is obtained from VGP transformation of α_{95} , while the solid ellipse is determined by a bootstrap.

of N data points is a “para-data set” (also known as a “pseudo-sample”). We then calculate a Fisher mean of the para-data set (one little circle in Figure 12.4b). This resampling procedure can be repeated many times. We show 500 such bootstrapped means in Figure 12.4b.

Now we can estimate the region of 95% confidence for the bootstrapped means. A non-parametric approach would be to draw a contour enclosing 95% of the bootstrapped means. In many applications, paleomagnetists desire a more compact way of expressing confidence regions (for example, to list them in a table) and this necessitates some parametric assumptions about the distribution of the means. For this limited purpose, approximate 95% confidence regions can be estimated by noting the elliptical distribution of the bootstrapped means and by assuming that they are Kent (1982) distributed.

When paleomagnetic data are bimodal, as in Figure 12.4a, we can proceed in one of two ways. We could just calculate the principle eigenvector of the orientation matrix ($\mathbf{V}_1$) as in Bingham statistics of each bootstrapped para-data set or we can separate the data into two modes and calculate Fisher means for each mode separately (as in Figure 12.4b).

To separate the data into normal and reverse sub-sets, we first calculate the orientation matrix and its principal eigenvector $\mathbf{V}_1$. By transforming the data set into the coordinates of $\mathbf{V}_1$ (see appendix to Lecture 2), the transformed inclinations could be used to assign polarity in an automated and unbiased way; positive inclinations belong to one mode and negative inclinations to the other. After separation, Fisher means can be calculated for each mode separately. Alternatively, if a more robust estimate of the “average” direction is desired, one could calculate the principal eigenvector $\mathbf{V}_1$ of each mode, which is less sensitive to the presence of outliers.

12.4. THE PARAMETRIC BOOTSTRAP

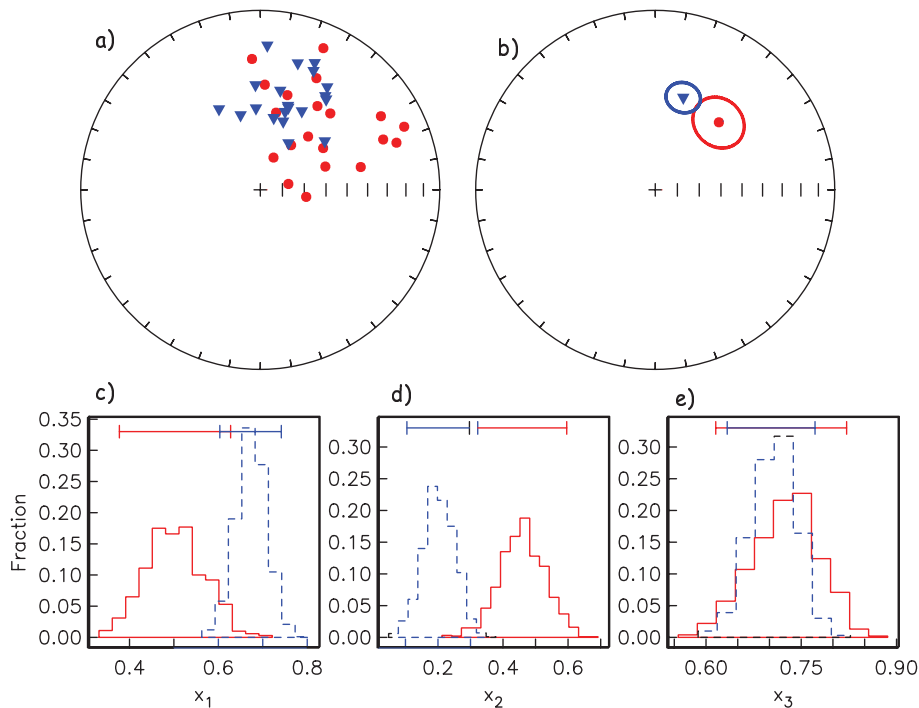


Figure 12.5: The bootstrap test for a common mean. a) Equal-area projections of two data sets two simulated Fisherian data sets (triangles and circles) each with κ of 20. b) Means and α_{95} of data sets shown in a). c) Histograms of x_1 components of the bootstrapped means from 500 para-data sets. Also shown are the bounds for each data set that include 95% of the components. The confidence intervals for the different data sets overlap. d) Same as c) but for the x_2 component. The confidence intervals do not overlap. e) same as d) but for the x_3 component.

12.4 The parametric bootstrap

The bootstrap just described is a “simple” or “naive” bootstrap in that no distribution is assumed for the data. We did assume, however, that all the uncertainty inherent in the data is reflected in the data distribution. If the data set is smaller than about $N = 20$, this leads to uncertainty ellipses that are too small (Tauxe et al., 1991). Many paleomagnetic data sets are smaller than this, yet they are demonstrably non-Fisherian. Luckily, if we are able to assume some parametric form for data from e.g., a given site, we can perform a superior technique which is known as the *parametric bootstrap*. As applied here, we assume that each site with N_s samples is Fisher distributed (in principle, a testable assumption). Then, after random selection of a particular site for inclusion in the para-data set, we draw N_s new directions from a Fisher distribution with the same mean direction, κ and N . From these, we then calculate a substitute mean direction, and use that in the para-data set. Otherwise, we follow the same procedure as before.

For large data sets ($N > 25$), the parametric and simple bootstraps yield very similar confidence ellipses. For smaller data sets, the parametric ellipses are larger, and are probably more realistic.

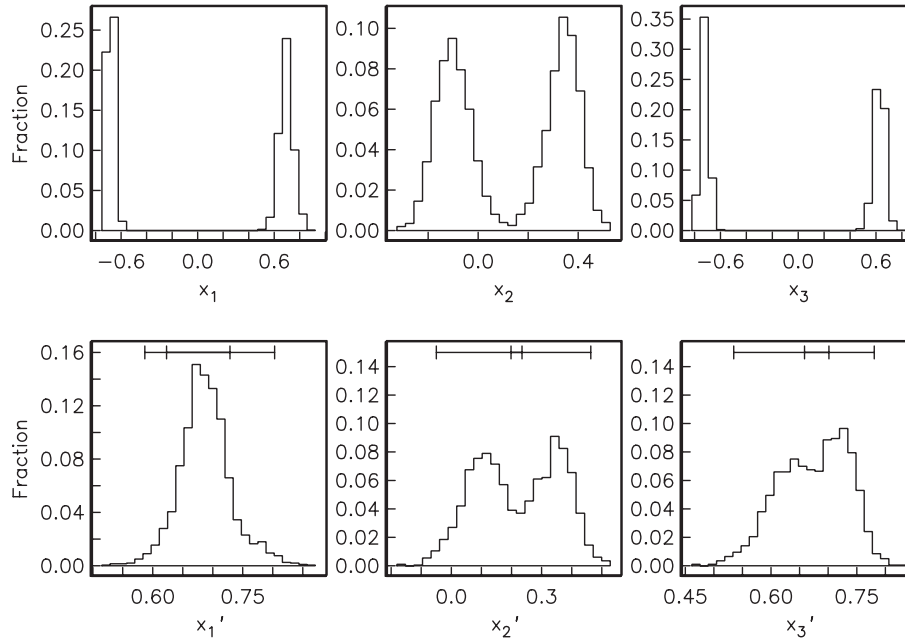


Figure 12.6: Histograms of cartesian coordinates of means of para-data sets drawn from the data shown in Figure 12.4a. In the lower plots, the reversed polarity mode has been flipped to the antipode (x'_i). The intervals containing 95% of each set of components are drawn (see last two figures). Because the confidence bounds from the two data sets overlap in all three components, the means of the reverse and normal modes cannot be distinguished at the 95% level of confidence; they pass the bootstrap reversals test.

12.5 Application to VGPs

We stated earlier that the transformation of Fisher distributed directional data to VGPs can cause a distortion from rotationally symmetric data to elliptically distributed data. To illustrate what happens if data are distributed as in our hypothetical normal data set, we calculate VGPs in Figure 12.4a, the ellipse described by dp, dm (Lecture 11), and our bootstrap confidence ellipses. The VGPs are plotted in Figure 12.4c and the dp, dm is shown as a dotted line. Because the declinations are smeared and the dp must point towards the site (shown as a triangle), the long axis of the so-called 95% confidence regions is perpendicular to the actual data distribution. The 95% confidence ellipse, calculated using the bootstrap, is shown as a solid line. The bootstrapped confidence ellipse gives a better sense of the uncertainty and follows the trend in the data.

12.6 When are two data sets distinct?

The test for a common mean addresses the question “can the means of two data sets be discriminated from one another?” Another way of putting it is, “If a set of bootstrap means is examined, are there two distinct groups or is there just one?” We explore these ideas by considering the same Fisherian data sets we used in Lecture 11 for the Watson’s V_w test. In Figure 12.5 we show two data sets (triangles and circles), each drawn from distributions with a κ of 20. The mean direction of each lies outside the confidence region of the other and the but the V_w test of Watson has a value

12.7. APPLICATION TO THE “REVERSALS TEST”

of 11.7 with a critical value of 6.3, hence they fail the test for a common mean.

In order to develop a bootstrap test analogous to the V_w test for use on non-Fisherian data sets, we first convert a set of bootstrapped mean directions to cartesian coordinates. Histograms of the cartesian coordinates of the bootstrap means (Figure 12.5c-e) are distinct in the x_2 component, confirming that the two means can be distinguished at the 95% confidence level.

12.7 Application to the “reversals test”

The so-called *reversals test* in paleomagnetism constitutes a test for a common mean for two modes, one of which has been “flipped” to its antipode. We apply our bootstrap test for common mean to the data shown in Figure 12.4. The histograms of the cartesian coordinates of the bootstrapped means are shown in Figure 12.6. There are two “humps” in the bootstrap test. However, the confidence intervals for the normal and reverse antipodes overlap, thereby suggesting that the two means cannot be distinguished at the 95% level of confidence. Thus, the data in Figure 12.4 pass the bootstrap reversals test.

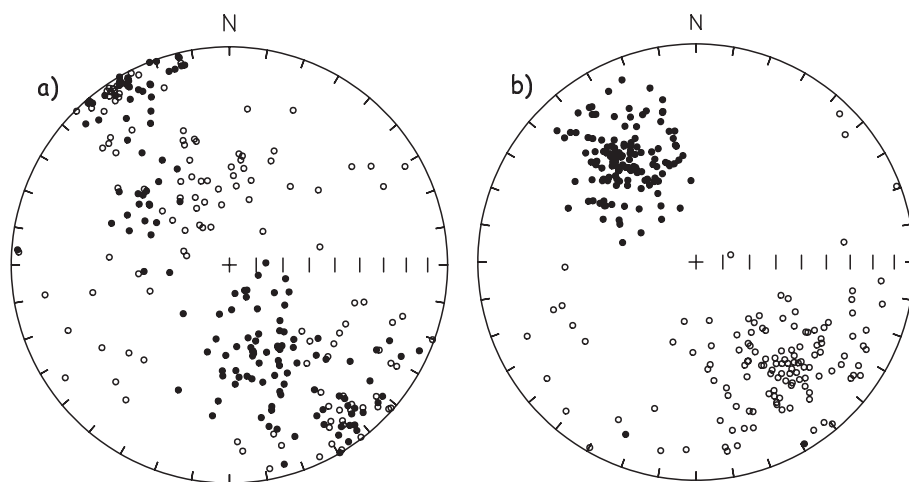


Figure 12.7: a) Equal area projection of a set of directions in geographic coordinates. The data are streaked in a girdle distribution and the polarity of many data points is ambiguous. b) Data from a) after 100% adjusting for tilt. Polarities are more readily identifiable. [Redrawn from Tauxe and Watson, 1994.]

12.8 Application to the fold test

A final test is extremely useful in paleomagnetism: the fold test (Lecture 9). One of the key components in paleomagnetic studies is to determine the coordinate system (geographic, tilt adjusted or somewhere in between) for which the directional data are most tightly clustered. If a rock has moved from its original position, was it magnetized in the original, in the present or in some other position? Moreover, is simple rotation about strike an appropriate method to restore the beds to their original positions? In the classic fold test as first proposed by Graham (1949), the directions of magnetization of a deformed rock unit are assumed to be most closely parallel in the

orientation in which the magnetization was acquired. Therefore, if a rock has retained an original magnetization through a subsequent folding or tilting event, the magnetic directions may cluster most tightly after they have been rotated back to their original positions. This of course is NOT true for elongate data such as those shown in Figure 12.1a.

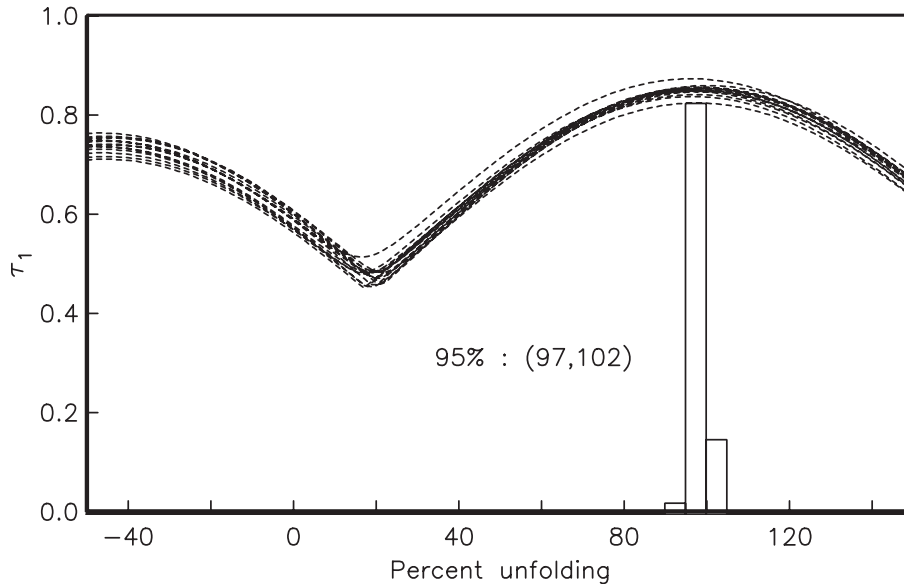


Figure 12.8: The largest eigenvalues τ_1 of the orientation matrices from representative para-data sets drawn from Figure 12.6. The directions are adjusted for tilt incrementally from -50% to 150%. The largest value of τ_1 occurs near 100% in all of the para-data sets. Histogram is of 500 maxima of τ_1 and the calculated 95% confidence interval. These data “pass” the bootstrap fold test. [Redrawn from Tauxe and Watson, 1994.]

The fold test appears at first glance to be simple, but it is not. The primary problem is that paleomagnetic vectors are never perfectly parallel. The scattered nature of the data means that a statistical test is necessary to determine whether clustering is “significantly” better in one orientation or another.

In Lecture 11 we suggested that variances could be compared using an F-test, so it was long the practice in paleomagnetism to compare estimated precisions before and after tilt adjustment (McElhinny, 1964). The ratio of the two estimates of κ were compared with those listed in statistical “F” tables (see Appendix to Lecture 11). Ratios higher than the “F” value for a given N were deemed to constitute a significant increase in concentration after adjusting for tilt, thus representing a positive fold test. This test can be done on the back of an envelope and is still in frequent use.

Although its simplicity is a great strength, there are several problems with the classical fold test. First, the geomagnetic field has two preferred states and is not perfectly dipolar. Directions observed in paleomagnetic samples are therefore not only scattered but are often of two polarities. Second, the magnetic directions may be most tightly clustered somewhere other than in “geographic” or 100% tilt adjusted coordinates. Finally, structural “corrections” are not perfectly known. Not only are the bedding orientations themselves often difficult to measure accurately, but detection of complications such as plunging folds, and multiple phases of tilting requires extensive field work.

12.8. APPLICATION TO THE FOLD TEST

It is nearly impossible to assess rotation about the vertical axis on the basis of field relations alone, as it results in no visible effect on the dip of the beds themselves. Because of this uncertainty, we might reasonably ask whether if the data are actually most tightly clustered at, say 90% tilt adjusted (as opposed to 100%), does this constitute a “failed” fold test.

We consider first the problem of dual polarity. We plot a data set in geographic coordinates in Figure 12.7a and in tilt adjusted coordinates in Figure 12.7b. The polarity is ambiguous and the calculation of κ necessitates using the tilt adjusted data to identify the polarity of the samples. The classic fold test requires calculation of κ which can only be done with data of a single polarity. Obviously, fold tests that rely on κ will not be straightforward with data such as these.

An alternative approach is based on the orientation matrix. In the orientation matrix, polarity does not play a role and the “tightness” of grouping is reflected in the relative magnitudes of the eigenvalues (τ). As the data become more tightly grouped, the variance along the principal axis grows and those along the other axes shrink. Thus, examination of the behavior of τ_1 during unfolding would reveal the point at which the tightest grouping is achieved, without knowledge of polarity.

Suppose we find that the degree of unfolding required to produce the maximum in τ_1 is 98%. Is this a positive fold test suggesting a pre-folding remanence or is the difference between 98% and 100% significant? For this we call on the familiar bootstrap. Numerous para-data sets can be drawn. We can then calculate the eigenparameters of the orientation matrix for a range of percent unfolding. Some examples of the behavior of τ_1 during tilt adjustment of representative para-data sets drawn from the data in Figure 12.7a are shown in Figure 12.8. Figure 12.8 is a histogram of maxima of τ_1 from 500 para-data sets. These are sorted and the 95% confidence interval for the degree of unfolding required to produce the tightest grouping (the highest τ_1) is thus constrained to lie between 97 and 102%.

The data from Figure 12.7a are shown after 100% tilt adjustment in Figure 12.7b. The tilt adjusted data are not only better grouped, but now the polarities of most samples can be readily determined. An advantage of the bootstrap approach is the fact that the data do not need prior editing to split them into normal and reversed polarity groups, which is a particularly onerous task for the data considered here.

For small data sets, we could of course employ a parametric bootstrap, whereby para-data sets are generated by first randomly selecting a site for inclusion, then by drawing a substitute direction from a Fisher distribution having the same D , I , N , and κ .

Appendix

A Estimation of Kent 95% confidence ellipse

Kent parameters are calculated by rotating unimodal directions x into the data coordinates x' by the transformation:

$$x' = \mathbf{\Gamma}^T x, \quad (\text{A1})$$

where $\mathbf{\Gamma} = (\gamma_1, \gamma_2, \gamma_3)$, and the columns of $\mathbf{\Gamma}$ are called the constrained eigenvectors of $\mathbf{T}$. The vector γ_1 is parallel to the Fisher mean of the data, whereas γ_2 and γ_3 (the major and minor axes) diagonalize $\mathbf{T}$ as much as possible subject to being constrained by γ_1 (see Kent (1982), but note that his x_1 corresponds to x_3 in conventional paleomagnetic notation). The following parameters may then be computed

$$\begin{aligned} \hat{\mu} &= N^{-1} \sum_k x_{k1}' \\ \hat{\sigma}_2^2 &= N^{-1} \sum_k (x_{k2}')^2 \\ \hat{\sigma}_3^2 &= N^{-1} \sum_k (x_{k3}')^2. \end{aligned} \quad (\text{A2})$$

As defined here, $\hat{\mu} = R/N$ (R is closely approximated by the equation for R in Lecture 11). Also to good approximation, $\hat{\sigma}_2^2 = \tau_2$, and $\hat{\sigma}_3^2 = \tau_3$, where τ_i are the eigenvalues of the orientation matrix. The semi-angles ζ_{95} and η_{95} subtended by the major and minor axes of the 95% confidence ellipse are given by:

$$\zeta_{95} = \sin^{-1}(\sigma_2 \sqrt{g}), \quad \eta_{95} = \sin^{-1}(\sigma_3 \sqrt{g}), \quad (\text{A3})$$

where $g = -2 \ln(0.05)/(N\hat{\mu}^2)$.

The tensor $\mathbf{\Gamma}$ is, to a good approximation, equivalent to $\mathbf{V}$, the eigenvectors of the orientation matrix. Therefore, the eigenvectors of the orientation matrix $\mathbf{V}$ give a good estimate for the directions of the semi-angles by:

$$\begin{aligned} D_\zeta &= \tan^{-1}(v_{22}/v_{12}), I_\zeta = \sin^{-1}v_{32}, \\ D_\eta &= \tan^{-1}(v_{23}/v_{13}), I_\eta = \sin^{-1}v_{33}, \end{aligned} \quad (\text{A4})$$

where for example the x_2 component of the smallest eigenvector $\mathbf{V}_3$ is denoted v_{23} .

B Estimation of Bingham 95% confidence parameters

The Bingham distribution is given by:

$$F = \frac{1}{4\pi d(k_1, k_2)} \exp(k_1 \cos^2 \phi + k_2 \sin^2 \phi) \sin^2 \alpha$$

where α and ϕ are as in the Kent distribution, k_1, k_2 are concentration parameters ($k_1 < k_2 < 0$) and $d(k_1, k_2)$ is a constant of normalization given by:

$$d(k_1, k_2) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \exp((k_1 \cos^2 \phi + k_2 \sin^2 \phi) \sin^2 \theta) \sin \theta d\theta d\phi.$$

To estimate the axes of the Bingham confidence ellipse, we first calculate the eigenparameters of the orientation matrix as for Kent parameters and described in the Appendix to Lecture 9.

C. THE STATISTICAL BOOTSTRAP

Note that through-out these lecture notes τ_i are the eigenvalues and $\mathbf{V}_i$ are the eigenvectors. The principle eigenvector $\mathbf{V}_1$ of the orientation matrix is associated with the largest eigenvalue τ_1 . In Bingham (1974), ω_1 is the τ_3 and ω_3 is τ_1 . In Bingham statistics, the $\mathbf{V}_1$ direction is taken as the mean - it is not always parallel to the Fisher mean of a unimodal set of directions.

The maximum likelihood estimates of k_1, k_2 , the concentration parameters are gotten by first maximizing the log likelihood function:

$$F = -N \log(4\pi) - N \log d(k_1, k_2) + k_1 \omega_1 + k_2 \omega_2.$$

These are listed for convenience in Table B1 as calculated by Mardia and Zemroch (1977). Once these are estimated, the semi-axes of the 95% confidence ellipse around the mean direction $\mathbf{V}_1$ are given by:

$$\epsilon_{ij}^2 = \chi_p^2(\nu) \sigma_{ij}^2,$$

where $\chi_p^2(\nu) = 5.99$ is the χ^2 value for significance ($p = .05$ for 95% confidence) with $\nu = 2$ degrees of freedom and

$$\sigma_{ij}^2 = \frac{1}{2N(\omega_i - \omega_j)(k_i - k_j)}$$

Bingham (1974) set $k_3 = 0$, so the semi axes of the confidence ellipse about the principle direction $\mathbf{V}_1$, associated with ω_3 , are therefore:

$$\epsilon_{32} = \frac{1.22}{-k_2 N(\omega_3 - \omega_2)}$$

and

$$\epsilon_{31} = \frac{1.22}{-k_1 N(\omega_3 - \omega_1)}.$$

Because $k_1 < k_2 < 0$, the semi-axes are positive numbers. Please note that we use the corrected version of Tanaka (1999) as opposed to the more oft quoted but erroneous treatment of Onstott (1980) in this appendix. Note also that the N is required for σ because we have normalized the ω s to sum to unity for consistency with other eigenvalue problems in this set of lecture notes. The N is missing in the treatment of Tanaka (1999) presumably because the eigenvalues sum to N . Finally, note that these values of ϵ are in radians and must be converted to degrees for most applications.

C The statistical bootstrap

In Figure B1, we illustrate the essentials of the statistical bootstrap. We will develop the technique using data drawn from a normal distribution. First, we generate a synthetic data set by drawing 500 data points from a normal distribution with a mean $\bar{x}$ of 10 and a standard deviation σ of 2. The synthetic data are plotted as a histogram in Figure B1a. In Figure B1b we plot the data as a Q-Q plot against the z_i expected for a normal distribution (see Abramowitz and Stegun (1970)).

In order to calculate the appropriate values for z_i assuming a normal distribution:

1. For $i = 1 \rightarrow N$, calculate $p = \frac{i}{N+1}$.

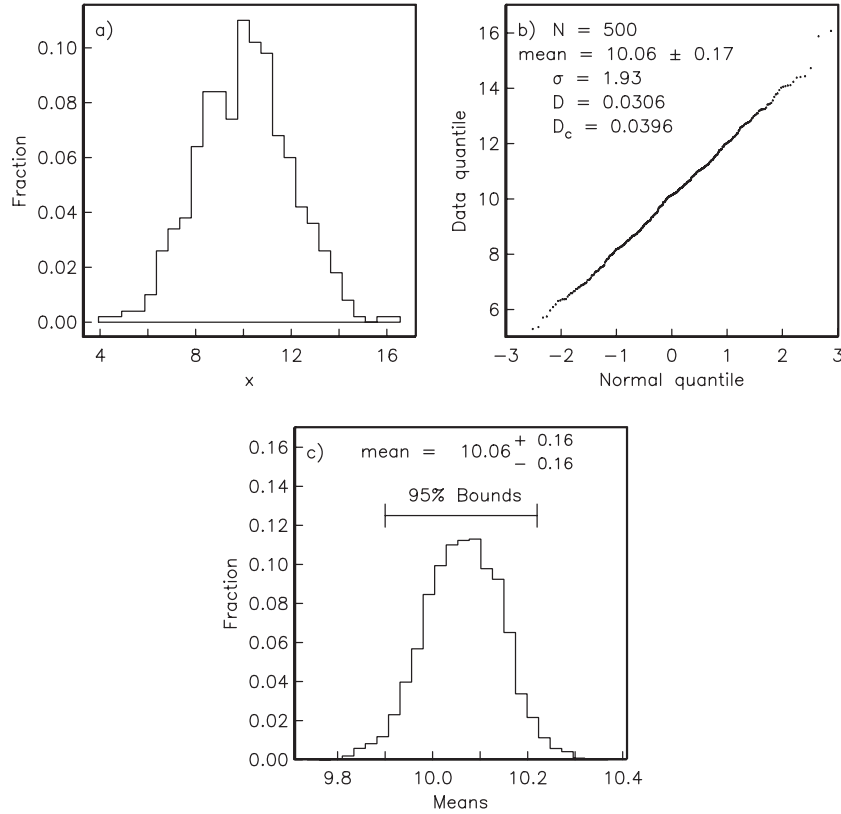


Figure B1: Bootstrapping applied to a normal distribution. a) 500 data points are drawn from a Gaussian distribution with mean of 10 and a standard deviation of 2. b) Q-Q plot of data in a). The 95% confidence interval for the mean is given by Gauss statistics as ± 0.17 . 10,000 new (para) data sets are generated by randomly drawing N data points from the original data set shown in a). c) A histogram of the means from all the para-data sets. 95% of the means fall within the interval $10.06^{+0.16}_{-0.16}$, hence the bootstrap confidence interval is similar to that calculated with Gaussian statistics.

2. If $p > 0.5$, then $q = 1 - p$; if $p < 0.5$, then $q = p$.
3. Calculate the following for all $p \neq 0.5$:

$$t = \sqrt{-2 \ln^{-1}(q)},$$

and

$$u = t - \frac{(a_1 + a_2 t + a_3 t^2)}{(1 + a_4 t + a_5 t^2 + a_6 t^3)},$$

where $a_1 = 2.515517, a_2 = 0.802853, a_3 = 0.010328, a_4 = 1.432788, a_5 = 0.189269, a_6 = 0.001388$.

4. If $p > 0.5$, then $z_i = u$; if $p < 0.5$, then $p = -u$.

C. THE STATISTICAL BOOTSTRAP

5. If $p = 0.5$, then $z_i = 0$.

The values of z_i calculated in this way for the simulated Gaussian distribution are plotted as the “normal quantile” data in Figure B1b.

The simulated data fall along a line in Figure B1b which suggests that they are normally distributed (as one would hope). To test this in a more quantitative way, we can calculate D_N^+ and D_N^- as follows:

1. Calculate the mean $\bar{x}$ and standard deviation σ for the data.

2. Then calculate:

$$p = \frac{x_i - \bar{x}}{\sqrt{2}\sigma},$$

and

$$q = \frac{1}{1 + 0.3275911|x|}.$$

3. Substitute q into the following expression (function 7.1.26 from Abramowitz and Stegun (1970):

$$\text{erf}(q) = 1 - e^{-p^2} [a_1 q + a_2 q^2 + a_3 q^3 + a_4 q^4 + a_5 q^5],$$

where $a_1 = 0.254829592$, $a_2 = -0.284496736$, $a_3 = 1.421413741$, and $a_5 = 1.061405429$.

4. Change the sign of $\text{erf}(q)$ such that it has the same sign as q .
5. Substitute $F(x) = 0.5(1 + \text{erf}(q))$ into Equations D3 and D4 in the Appendix to Lecture 11 for D_N^+ and D_N^- respectively. The Kolmogorov-Smirnov parameter D (e.g., Fisher et al., 1987) is the larger of D_N^+ or D_N^- .
6. The null hypothesis that a given data set is normally distributed can be rejected at the 95% level of confidence if D exceeds a critical value D_c given by $0.886/\sqrt{N}$.

Applying the foregoing to the data in Figure B1a yields a D value of 0.0306. Because $N = 500$, the critical value of D , D_c at the 95% confidence level is 0.0396. Happily, our program has produced a set of 500 numbers for which the null hypothesis of a normal distribution has not been rejected. The mean is about 10 and the standard deviation is 1.9. The usual Gaussian statistics allow us to estimate a 95% confidence interval for the mean as $\pm 1.96\sigma/\sqrt{N}$ or ± 0.17 .

The mean of the synthetic dataset is about 10 and the standard deviation is 1.9. The usual Gaussian statistics allow us to estimate a 95% confidence interval for the mean as $\pm 1.96\sigma/\sqrt{N}$ or ± 0.17 .

In order to estimate a confidence interval for the mean using the bootstrap, we first randomly draw a list of N data by selecting data points from the original data set. Some data points will be used more than once and others will not be used at all. We then calculate the mean of the “para-data set”. We repeat the procedure of drawing para-data sets and calculating the mean many times (say 10,000 times). A histogram of the “bootstrapped” means is plotted in Figure B1c. If these are sorted such that the first mean is the lowest and the last mean is the highest, the 95%

of the means are between the 250th and the 9,750th mean. These therefore are the 95% confidence bounds because we are approximately 95% confident that the true mean lies between these limits. The 95% confidence interval calculated for the data in Figure B1 by bootstrap is about ± 0.16 which is nearly the same as that calculated the Gaussian way. However, the bootstrap required orders of magnitude more calculations than the Gaussian method, hence it is ill-advised to perform a bootstrap calculation when a parametric one will do. Nonetheless, if the data are not Gaussian, the bootstrap provides a means of calculating confidence intervals when there is no quick and easy way. Furthermore, with a modern computer, the time required to calculate the bootstrap illustrated in Figure B1 was virtually imperceptible.

Bibliography

- Abramowitz, M. & Stegun, I. A. (1970), ‘Handbook of Mathematical Functions’.
- Bingham, C. (1974), ‘An antipodally symmetric distribution on the sphere’, *Ann. Statist.* **2**, 1201–1225.
- Borradaile, G. J. (2003), *Statistics of Earth Science Data: Their Distribution in Time, Space, and Orientation*, Springer, Berlin.
- Creer, K., Irving, E. & Nairn (1959), ‘Paleomagnetism of the Great Whin Sill’, *Geophys. J. Int.* **2**, 306–323.
- Creer, K. M. (1983), ‘Computer synthesis of geomagnetic paleosecular variations’, *Nature* **304**, 695–699.
- Efron, B. (1982), ‘The Jackknife, the Bootstrap and Other Resampling Plans’.
- Fisher, N. I., Lewis, T. & Embleton, B. J. J. (1987), ‘Statistical Analysis of Spherical Data’.
- Graham, J. W. (1949), ‘The stability and significance of magnetism in sedimentary rocks’, *Jour. Geophys. Res.* **54**, 131–167.
- Kent, J. T. (1982), ‘The Fisher-Bingham distribution on the sphere’, *J. R. Statist. Soc. B.* **44**, 71–80.
- Le Goff, M., Henry, B. & Daly, L. (1992), ‘Practical method for drawing a VGP path’, *Phys. Earth Planet. Inter.* **70**, 201–204.
- Love, J. & Constable, C. G. (2003), ‘Gaussian statistics for paleomagnetic vectors’, *Geophys. J. Int.* **152**, 515–565.
- Mardia, K. V. & Zemrock, P. J. (1977), ‘Table of maximum likelihood estimates for the Bingham distribution’, *J. Statist. Comput. Simul.* **6**, 29–34.
- McElhinny, M. W. (1964), ‘Statistical significance of the fold test in paleomagnetism’, *Geophys. Jour. R. astro. Soc.* **8**, 338–340.
- McElhinny, M. W. & McFadden, P. L. (1997), ‘Palaeosecular variation over the past 5 Myr based on a new generalized database’, *Geophys. J. Int.* **131**(2), 240–252.
- Onstott, T. C. (1980), ‘Application of the Bingham Distribution Function in paleomagnetic studies’, *J. Geophys. Res.* **85**, 1500–1510.

- Tanaka, H. (1999), 'Circular asymmetry of the paleomagnetic directions observed at low latitude volcanic sites', *Earth Planets Space* **51**, 1279–1286.
- Tauxe, L., Constable, C. G., Stokking, L. B. & Badgley, C. (1990), 'The use of anisotropy to determine the origin of characteristic remanence in the Siwalik red beds of northern Pakistan', *Jour. Geophys. Res.* **95**, 4391–4404.
- Tauxe, L. & Kent, D. V. (2004), A simplified statistical model for the geomagnetic field and the detection of shallow bias in paleomagnetic inclinations: Was the ancient magnetic field dipolar?, *in* e. a. Channell, J.E.T., ed., 'Timescales of the Paleomagnetic field', Vol. 145, American Geophysical Union, Washington, D.C., pp. 101–116.
- Tauxe, L., Kylstra, N. & Constable, C. (1991), 'Bootstrap statistics for paleomagnetic data', *Jour. Geophys. Res.* **96**, 11723–11740.
- Tauxe, L. & Watson, G. S. (1994), 'The fold test: an eigen analysis approach', *Earth Planet. Sci. Lett.* **122**, 331–341.