Geophys. J. R. astr. Soc. (1982) **69**, 307–319

Analysis of palaeomagnetic inclination data

P. L. McFadden* and A. B. Reid† *Department of Physics, University of Zimbabwe, PO Box MP 167, Mount Pleasant, Salisbury, Zimbabwe*

Received 1981 May 2; in original form 1980 December 17

Summary. The analysis of palaeomagnetic data where only inclinations are available is considered. Maximum likelihood estimates for the mean inclination I_0 and Fisher's precision parameter κ are derived. It is shown that they are in all cases biased although the bias is small for low inclinations. The case of steep inclinations and small values of κ is examined and it is shown that in this region I_0 and κ are not separable as distinct variables, because the lack of declination information in this region leads to fundamental ambiguities. Unbiased estimates for I_0 and $(1/\kappa)$ are derived for the case where the portion of the distribution folded about the vertical is insubstantial. A worked example of the method, with calculation of confidence limits, is appended.

1 Introduction

Palaeomagnetic investigations are sometimes performed on vertical borecores which have been recovered as fragmentary pieces. Under these conditions even 'relative' declinations are unknown and the only available data are the measured inclinations of magnetization. In a normal palaeomagnetic investigation both declinations and inclinations are available and neither of these constitute redundant information. Consequently it should be expected that a correct analysis of inclination data only will be extremely difficult and even impossible under some circumstances.

Initially it must be recognized that it is impossible to obtain an unconditional estimate of the true mean inclination which is unbiased, unless the true mean inclination happens to be zero. This is easily seen by considering the case of a true inclination of 90° (i.e. vertically down). Independent of the method of estimating the inclination of the true mean direction there will always be samples which will lead to an estimate less than 90°. Since it is impossible to obtain an estimate in excess of 90° this leads immediately to the conclusion that it is impossible to obtain an unbiased estimate. Evidently the bias in the estimate will decrease as the inclination of the true mean direction approaches zero and also as the precision of the observed population increases. However, it turns out that in many instances it is possible to obtain a conditional estimate for the inclination which is unbiased.

*Present address: Research School of Earth Sciences, Australian National University, PO Box 4, Canberra, ACT 2600, Australia.

†Present address: Geology Department, Tsumeb Corporation, PO Box 40, Tsumeb 9000, South West Africa.

12

Given that it is impossible to obtain an unbiased estimate for the inclination the estimate which probably has the least bias is the maximum likelihood estimate (mle). Therefore a general solution for the mles is derived with the assumption that the full data set of declinations and inclinations is Fisher distributed (Fisher 1953). An unbiased conditional estimate for the inclination is derived for those cases where it is possible, together with an unbiased estimate for the inverse of the precision parameter. The equations necessary for determining the errors associated with these estimates are also derived. The remaining cases are examined numerically.

2 The distribution

If the individual directions are Fisher distributed then the density is given by

$$P(\theta',\phi')\,d\theta'\,d\phi' = \frac{\kappa}{4\pi\sinh\kappa}\exp(\kappa\cos\theta')\sin\theta'\,d\theta'\,d\phi' \tag{1}$$

where κ is the precision parameter, θ' is the polar angle between an observation and the true mean direction and ϕ' is the uniformly distributed azimuthal angle about the true mean direction. The simplest method of analysing the problem is to transform to a set of variables related to the downward vertical. Defining the variables θ_0, θ and ϕ as they were defined by Briden & Ward (1966), θ_0 is the complement of the inclination of the true mean direction, θ is the complement of an observed inclination and ϕ is the azimuthal angle (about the downward vertical) of an observation. These variables are then related by the equations

$$\cos\theta' = \cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos\phi$$

$$\tan\phi' = \frac{\sin\theta\sin\phi}{\sin\theta_0\cos\theta - \cos\theta_0\sin\theta\cos\phi}$$

$$\cos\theta = \cos\theta_0\cos\theta' + \sin\theta_0\sin\theta'\cos\phi'$$

$$\tan\phi = \frac{\sin\theta'\sin\phi'}{\sin\theta_0\cos\theta' - \cos\theta_0\sin\theta'\cos\phi'} \tag{2}$$

and estimation of θ_0 is equivalent to estimation of I_0, the inclination of the true mean direction.

The Jacobian of the transformation is such that

$$\sin\theta'\,d\theta'\,d\phi' = \sin\theta\,d\theta\,d\phi \tag{3}$$

giving the density in terms of the variables θ and ϕ as

$$P(\theta,\phi)\,d\theta\,d\phi = \frac{\kappa}{4\pi\sinh\kappa}\exp\left[\kappa(\cos\theta_0\cos\theta + \sin\theta_0\sin\theta\cos\phi)\right]\sin\theta\,d\theta\,d\phi. \tag{4}$$

The marginal distribution of θ is then given by

$$P(\theta)\,d\theta = \frac{\kappa}{2\sinh\kappa}\exp(\kappa\cos\theta_0\cos\theta)\,G\sin\theta\,d\theta \tag{5}$$

where

$$G = \frac{1}{2\pi}\int_0^{2\pi}\exp(\kappa\sin\theta_0\sin\theta\cos\phi)\,d\phi$$

$$= \sum_{r=0}^{\infty}\frac{(\kappa\sin\theta_0\sin\theta)^{2r}}{2^{2r}(r!)^2}. \tag{6}$$

3 Previous analyses

The first serious attempt at an analysis of this problem was given by Briden & Ward (1966) who derived the distribution given above. They presented a fairly *ad hoc* analysis whereby they calculated numerically the values of certain parameters for a range of values of θ_0 and κ, producing from their results a nomogram and 28 tables for the estimation of θ_0 and κ from an observed sample, together with estimates for the errors. In obtaining their error estimates they assumed a normal distribution of errors as well as linear dependences of errors in θ_0 and κ. Unfortunately these assumptions are not well satisfied over a wide range of cases.

Kono (1980) has derived analytical expressions for estimates of $\theta_0(I_0)$ and κ. He does not state what type of estimates they are. He also asserts that 'the best estimate of precision parameter k, which is obtainable from (10) [his (10)], and the cosine of angular error (δ) are related through the equations derived by Fisher (1953);

$$1-\delta=\frac{N-R}{R}\left[\left(\frac{1}{P}\right)^{1/(N-1)}-1\right]$$

$$k=\frac{N-1}{N-R}$$

where p is the probability that the cosine of error angle is less than δ. ... Confidence limits for k can be obtained by the method of Cox (1969).' The accuracy of this assertion is doubtful since it requires that the estimates be unbiased and that

$$2\kappa R(1-\delta)\sim\chi_2^2$$

and

$$\frac{1}{k}\sim\frac{\chi^2_{2(N-1)}}{2\kappa(N-1)}$$

where '~' is to be read as 'is distributed as' and χ_b^2 is the chi-square distribution with b degrees of freedom (McFadden 1980a). There is no indication in his analysis that the estimates are unbiased nor that they have the above distributions. In a normal palaeomagnetic analysis N specimens give rise to $2N$ independent data (the N inclinations and the N declinations) from which the true declination and inclination have, in effect, to be estimated before κ is estimated, leading to the $2(N-1)$ degrees of freedom for the distribution of k. Given that only the inclinations are available there are only N independent data implying that whatever the distribution of the estimate k given by Kono (1980) it can have at most $(N-1)$ degrees of freedom. A similar argument shows that whatever the distribution of δ it can have at most one degree of freedom. Consequently confidence limits for κ cannot be obtained directly from the tables of Cox (1969, 1977).

Monte Carlo experiments show that the estimates given by Kono (1980) are in fact biased. For example, 200 random samples of size 5 drawn from a Fisher population with true mean inclination 85° and κ = 15 gave an average estimated inclination of 75.9°, which is quite heavily biased. The harmonic mean of the k values gave an estimate for κ of 32.3, a bias of just over two in the ratio of estimated value to true value. The arithmetic average of the k values gave 78.4 so k is not an unbiased estimate for κ either. At lower inclinations and higher precisions the bias in the inclination estimate becomes small and the bias in the estimate for $(1/\kappa)$ reduces. Again as an example, 200 random samples of size 5 from a Fisher population with true mean inclination of 60° and κ = 40 gave an average estimated inclination of 59.7°. The harmonic mean for the k values gave an estimate for κ of 53.3, a bias of 1.33 in the ratio of estimated value to true value. The arithmetic average of the k values gave 92.2.

Since the estimates given by Kono (1980) are biased (particularly that for κ) his error analysis is suspect. However, as will be shown in the Appendix, if the correct unbiased estimates can be obtained they are related in the manner he claimed.

4 General equations for maximum likelihood estimates

Given the density of equation (5) the likelihood function, $H(\boldsymbol{\theta})$, of an observed sample $(\theta_1, \theta_2, \dots, \theta_N) = \boldsymbol{\theta}$ is

$$H(\boldsymbol{\theta}) = \frac{\kappa^N}{2^N(\sinh\kappa)^N}\,\Pi\exp(\kappa\cos\theta_0\cos\theta_i)\,\Pi G_i \Pi\sin\theta_i$$

where G_i is the function G of equation (6) with $\sin\theta$ replaced by $\sin\theta_i$ and the products run over i from 1 to N. The log likelihood function, $h(\boldsymbol{\theta}) = \ln[H(\boldsymbol{\theta})]$ is then given by

$$h(\boldsymbol{\theta}) = N\ln(\kappa) - N\ln(\sinh\kappa) + \kappa\Sigma\cos\theta_0\cos\theta_i + \Sigma\ln(G_i) + \Sigma\ln(\sin\theta_i) - N\ln(2) \quad (7)$$

where the summation signs run over i from 1 to N. Differentiating $h(\boldsymbol{\theta})$ partially with respect to θ_0 and then κ, setting these partial differentials equal to zero and substituting $\hat{k}$ as the mle for κ and $\hat{\theta}_0$ as the mle for θ_0 gives

$$\hat{k}\sin\hat{\theta}_0\,\Sigma\cos\theta_i = \cot\hat{\theta}_0\,\Sigma(a_i/g_i) \quad (8a)$$

$$N\coth(\hat{k}) - \cos\hat{\theta}_0\,\Sigma\cos\theta_i - \frac{N}{\hat{k}} = \frac{1}{\hat{k}}\,\Sigma(a_i/g_i) \quad (8b)$$

$$a_i = \sum_{r=1}^{\infty}\frac{2r(\hat{k}\sin\hat{\theta}_0\sin\theta_i)^{2r}}{2^{2r}(r!)^2} \quad (8c)$$

$$g_i = \sum_{r=0}^{\infty}\frac{(\hat{k}\sin\hat{\theta}_0\sin\theta_i)^{2r}}{2^{2r}(r!)^2}. \quad (8d)$$

If $\hat{k}$ is large enough that $\coth(\hat{k}) = 1$, which will almost always be the case, then 8(a) and 8(b) give

$$\hat{k} = \frac{N}{N - \Sigma\cos\theta_i/\cos\hat{\theta}_0} \quad (9)$$

and this may be substituted in (8a) to solve for $\hat{\theta}_0$ by an iterative method. From the form of equation (9) it might be hoped that if k is defined as

$$k = \frac{N-1}{N - \Sigma\cos\theta_i/\cos\hat{\theta}_0} \quad (10)$$

then $(1/k)$ would be an unbiased estimate for $(1/\kappa)$. Unfortunately this turns out not to be the case. For example, 200 random samples of size 5 drawn from a Fisher population with the mean inclination 85° and $\kappa = 15$ (the same random samples were used here as were used for testing the estimates given by Kono 1980) gave an average estimated inclination of 77.0°. The harmonic mean of the k values gave an estimate for κ of 22.9. Quite evidently the biases are less than for the estimates given by Kono (1980) but this is of little value since the biases are still quite substantial. In performing the Monte Carlo experiments 20 terms were used in the calculation of the functions a_i and g_i.

It is perhaps disappointing that for steep inclinations and small values of κ the mles are so biased but, as was noted in Section 1, this is only to be expected. Under these conditions the vertical effectively constitutes a break point at which a substantial portion of the distribution is aliased or 'folded back' upon itself and it is because of this that the resulting distribution is intractable.

Examination of fig. 2 of Briden & Ward (1966) suggests that, for this region, the values of I_0 and κ are not separable as distinct variables. This arises because the bias is dependent on both I_0 and κ so that that they are inextricably coupled. This view is supported by the fact that some perfectly legitimate samples plot in the 'illegal' area of their nomogram. It would, however, appear possible to obtain a useful estimate of I_0 when κ is assumed, or *vice versa.*

Our own experiments with mles run into similar problems for the same reasons. Where a sample would plot in the 'illegal' region, we obtain an mle of 90° for I_0 accompanied by a biased estimate of κ. It is not surprising that these two independent approaches give similar results when it is remembered that half the data are effectively missing.

Examination of equation (10) makes it possible to predict what bias is likely to be encountered. Where an mle of 90° for I_0 is obtained it is likely to be an overestimate since 90° is the upper limit. The equivalent value of $\hat{\theta}_0$ is then 0° and it is apparent that κ will be underestimated. On the other hand, if some value less than 90° is obtained for $\hat{I}_0$, it is clear from the previous discussion that it will be biased towards too low a value, so that from (10) κ will be overestimated. This behaviour is illustrated in Table 1 which shows the results of some Monte Carlo experiments. Samples, all of size 50, were drawn randomly from a Fisher distribution with mean inclination I_0 and precision parameter κ. I_0 was allowed to vary between 90° and 75° while κ varied from 10 to 40 and the biases discussed are readily seen. We conclude that it is not possible to derive a satisfactory analysis for the case of steep inclinations with small values of κ when the declination data are missing. However, if the portion of the distribution 'folded back' at the vertical is insubstantial then the distribution becomes tractable via approximation and 'educated' guesses. This analysis will be performed in the next section.

Table 1. Results of Monte Carlo experiments on samples with steep inclinations and small values of κ. In all cases N = 50.

True values		Maximum likelihood estimates	
I_0	κ	$\hat{I}$	$\hat{k}$
90	10	73.2	16.2
85	10	90	8.7
80	10	90	6.9
75	10	90	8.9
90	20	90	14.9
85	20	75.8	44.1
80	20	90	15.5
75	20	71.1	26.6
90	40	83.0	62.1
85	40	83.0	62.2
80	40	78.2	58.8
75	40	72.6	51.4

5 Approximate analysis for the case where the probability that $\theta' > \theta_0$ is small

For this analysis it is simpler to rewrite the density of equation (4) as

$$P(\theta, \phi)\, d\theta\, d\phi = \frac{\kappa}{4\pi \sinh\kappa} \exp[\kappa \cos(\theta_0 - \theta)] \exp[-\kappa \sin\theta_0 \sin\theta\, (1 - \cos\phi)] \sin\theta\, d\theta\, d\phi \qquad (11)$$

which gives a better indication of the distortion of the distribution with respect to the Fisher distribution centred about the true mean direction. If the probability that θ' exceeds θ_0 is small then only an insubstantial portion of the distribution is 'folded back' at the vertical and any folding which does occur may be ignored. This situation is very common and may be judged from the data themselves. For example, consider the inclinations 70.0, 53.7, 59.7, 47.7, 78.1, 55.7, 55.5, 61.6, 54.5, and 45.3. The smallest value is 45.3° and the largest 78.1°, the range being 32.8° and a very rough estimate of the true mean inclination is about 58°. If the highest inclination, 78.1°, had been obtained from a part of the distribution which was 'folded back' then it would be an observation approximately 44° from the mean inclination. Given the range of the rest of the observations such a deviation from the mean is extremely unlikely and it may safely be concluded that the portion of the distribution 'folded back' is insubstantial.

If κ exceeds 5 (which is by far the most common situation) then the density of the distribution will be negligible for large values of θ' and if in addition that portion of the distribution folded back at the vertical is insubstantial then the density of equation (11) will be negligible for large values of ϕ and so the approximation

$$1 - \cos\phi = \tfrac{1}{2} \phi^2 \qquad (12)$$

will be acceptable. Consequently

$$\frac{1}{2\pi} \int_0^{2\pi} \exp[-\kappa \sin\theta_0 \sin\theta (1 - \cos\phi)]\, d\phi$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[-\tfrac{1}{2}\kappa\phi^2 \sin\theta_0 \sin\theta]\, d\phi$$

$$= (2\pi\kappa \sin\theta_0 \sin\theta)^{-1/2} \qquad (13)$$

giving the appropriate marginal distribution of θ as

$$P(\theta)\, d\theta = \frac{\sqrt{\kappa}}{2\sqrt{2\pi} \sinh\kappa} \exp[\kappa \cos(\theta_0 - \theta)] \left(\frac{\sin\theta}{\sin\theta_0}\right)^{1/2} d\theta, \qquad (14)$$

which is a more tractable distribution.

The likelihood function of this distribution may be written down immediately as

$$H(\boldsymbol{\theta}) = \frac{\kappa^{\frac{1}{2}N}}{2^N (2\pi)^{\frac{1}{2}N} (\sinh\kappa)^N} \Pi \exp[\kappa \cos(\theta_0 - \theta_i)] \frac{\Pi \sqrt{\sin\theta_i}}{(\sin\theta_0)^{\frac{1}{2}N}} \qquad (15)$$

giving the log likelihood function as

$$h(\boldsymbol{\theta}) = \tfrac{1}{2}N \ln(\kappa) - N \ln(\sinh\kappa) + \kappa\, \Sigma \cos(\theta_0 - \theta_i)$$
$$+ \Sigma \sqrt{\sin\theta_i} - \tfrac{1}{2}N \ln(\sin\theta_0) - N \ln(2) - \tfrac{1}{2}N \ln(2\pi). \qquad (16)$$

Differentiating $h(\boldsymbol{\theta})$ partially with respect to θ_0, setting this differential equal to zero and substituting the mles $\hat{k}$ and $\hat{\theta}_0$ for κ and θ_0 gives

$$\hat{k}\, \Sigma \sin(\theta_i - \theta_0) = \tfrac{1}{2}N \cot\hat{\theta}_0. \qquad (17)$$

Repeating the process but differentiating with respect to κ and assuming that $\hat{k}$ is large enough that $\coth\hat{k} = 1$, gives

$$\hat{k} = \frac{N}{2[N - \Sigma\cos(\hat{\theta}_0 - \hat{\theta}_i)]}. \tag{18}$$

Substitution of k from (18) into (17) gives $\hat{\theta}_0$ as a solution of

$$N\cos\hat{\theta}_0 + (\sin^2\hat{\theta}_0 - \cos^2\hat{\theta}_0)\,\Sigma\cos\theta_i - 2\sin\hat{\theta}_0\cos\hat{\theta}_0\,\Sigma\sin\theta_i = 0, \tag{19}$$

which may easily be solved by iteration. The equation actually has three solutions but it is always immediately obvious which is the correct solution. Alternatively, using the second derivative of $h(\boldsymbol{\theta})$ with respect to θ_0, if U is defined as

$$U = \tfrac{1}{2}N(\operatorname{cosec}^2\hat{\theta}_0 - \frac{\Sigma\cos(\hat{\theta}_0 - \theta_i)}{N - \Sigma\cos(\hat{\theta}_0 - \theta_i)} \tag{19a}$$

then with the correct solution for $\hat{\theta}_0$, U will be negative. With the two incorrect solutions for $\hat{\theta}_0$, U will be positive.

From the form of equation (18) and from the fact that the 'folding back' of the distribution is negligible it is reasonable to hope that if k is defined as

$$k = \frac{N-1}{2[N - \Sigma\cos(\hat{\theta}_0 - \theta_i)]} \tag{20}$$

then $(1/k)$ will be an unbiased estimate for $(1/\kappa)$. Monte Carlo experiments confirm that this is in fact the case. For example, 200 random samples of size 5 drawn from a Fisher population with true mean inclination 60° and $\kappa = 40$ (the same random samples were used here as were used for testing the estimates given by Kono 1980) gave an average estimated inclination of 59.7° and the harmonic mean of the k values from equation (20) gave an estimate for κ of 40.5.

Further, since $(1/k)$ does turn out to be an unbiased estimate for $(1/\kappa)$, it might also be hoped that

$$\frac{(N-1)\kappa}{k} \sim \chi^2_{(N-1)} \tag{21}$$

and again Monte Carlo experiments confirm that, to a very good approximation, this is so. Substituting for k from equation (20) gives

$$2\kappa\,[N - \Sigma\cos(\hat{\theta}_0 - \theta_i)] \sim \chi^2_{(N-1)}. \tag{22}$$

Returning to the marginal distribution of equation (14) it may be noted that if the density is negligible for large θ' then it is also negligible for large $(\theta_0 - \theta)$ and on substituting

$$z = (\theta_0 - \theta)\sqrt{\kappa} \tag{23}$$

with the approximation that

$$\cos(\theta_0 - \theta) = 1 - \tfrac{1}{2}(\theta_0 - \theta)^2 \tag{24}$$

the density reduces to

$$P(z)\,dz = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2}z^2)\sqrt{\frac{\sin\theta}{\sin\theta_0}}\,dz. \tag{25}$$

Thus except for the modifying term $(\sin\theta/\sin\theta_0)^{1/2}$ the variate z is normally distributed, mean zero and variance one, i.e. $z \sim n(0, 1)$. Assuming for the time being that in fact $z \sim n(0, 1)$ this would imply that

$$z^2 = \kappa(\theta_0 - \theta_i)^2 \sim \chi_1^2. \tag{26}$$

and on using the approximation of (24) again, that

$$2\kappa\,[1 - \cos(\theta_0 - \theta_i)] \sim \chi_1^2. \tag{27}$$

Summing over the N independent observations then gives

$$2\kappa\,[N - \Sigma\cos(\theta_0 - \theta_i)] \sim \chi_N^2. \tag{28}$$

Unless θ_0 is small (which is not the case here) the modifying term $(\sin\theta/\sin\theta_0)^{1/2}$ does not deviate much from unity except where the density is negligible. However, the approximation to normality (and therefore to a chi-square distribution) is not particularly good, but, on adding the distribution for several independent observations the modifying terms tend to cancel making the final chi-square distribution a much better approximation than any of the individual approximations. Again, Monte Carlo experiments confirm that for N greater than 3 the approximation of equation (28) is acceptable. From equations (22) and (28)

$$2\kappa\,[N - \Sigma\cos(\theta_0 - \theta_i)] = 2\kappa\,[N - \Sigma\cos(\hat{\theta}_0 - \theta_i)] + g \tag{29}$$

and

$$\chi_N^2 \sim \chi_{(N-1)}^2 + \chi_1^2$$

giving

$$g = 2\kappa\,[\Sigma\cos(\hat{\theta}_0 - \theta_i) - \Sigma\cos(\theta_0 - \theta_i)] \sim \chi_1^2. \tag{30}$$

If we define α_1 by the equation

$$\theta_0 = \hat{\theta}_0 - \alpha_1 \tag{31}$$

then α_1 is the angular error in the mle $\hat{\theta}_0$ on the high inclination side of θ_0. Substituting for θ_0 from (31a) into (30) and using the approximations that

$$\cos\alpha_1 = 1 - \tfrac{1}{2}\alpha_1^2; \qquad \sin\alpha_1 = \alpha_1, \tag{32}$$

equation (30) reduces to

$$\kappa(\alpha_1^2 C - 2\alpha_1 S) \sim \chi_1^2 \tag{33}$$

where

$$C = \Sigma\cos(\hat{\theta}_0 - \theta_i); \qquad S = \Sigma\sin(\hat{\theta}_0 - \theta_i). \tag{34}$$

Combining equations (33) and (22) gives the distribution independent of κ but conditional on observed C as

$$\frac{(N-1)(\alpha_1^2 C - 2\alpha_1 S)}{2(N-C)} \sim F[1, (N-1)] \tag{35}$$

where $F[a, b]$ is the F distribution with a and b degrees of freedom.

If f be the relevant critical value of $F[1, (N-1)]$, then the critical value of α_1 is given from

$$\frac{(N-1)(\alpha_1^2 C - 2\alpha_1 S)}{2(N-C)} = f \tag{36}$$

as

$$\alpha_1 = \frac{2S + \sqrt{4S^2 + 8fC(N-C)/(N-1)}}{2C}. \tag{37a}$$

If we further define α_2 by the equation

$$\theta_0 = \hat{\theta}_0 + \alpha_2 \tag{31b}$$

then α_2 is the angular error in the mle $\hat{\theta}_0$ on the low inclination side of θ_0. Following the same analysis gives the critical value of α_2 as

$$\alpha_2 = \frac{-2S + \sqrt{4S^2 + 8fC(N-C)/(N-1)}}{2C}. \tag{37b}$$

For one-tailed testing at the level of significance p, f must be the value of $F[1, (N-1)]$ which will be exceeded with probability p, giving the value of α_1 or α_2 which will be exceeded with probability p. For two-tailed testing (by far the more common case) at the level of significance p, f must be the value of $F[1, (N-1)]$ which will be exceeded with probability $\frac{1}{2}p$. This is explained further in the Appendix where a numerical example is worked.

Quite obviously a more compact notation is achieved by retaining only α_1 and allowing it to be both positive and negative giving

$$\alpha_1 = \frac{2S \pm \sqrt{4S^2 + 8fC(N-C)/(N-1)}}{2C}. \tag{37c}$$

The asymmetry in the error angle is an indication of the bias in $\hat{\theta}_0$, conditional on the observed value of C. Consequently θ_0 is given by

$$\theta_0 = \left(\hat{\theta}_0 - \frac{S}{C}\right) \pm \alpha \tag{38}$$

where

$$\alpha = \frac{\sqrt{4S^2 + 8fC(N-C)/(N-1)}}{2C}. \tag{39}$$

Working in degrees rather than radians, the inclination of the true mean direction, I_0, is then given by

$$I_0 = (90^\circ - \theta_0) = \left(90^\circ - \hat{\theta}_0 + \frac{180S}{\pi C}\right) \pm \alpha. \tag{40}$$

Usually it will be more convenient to use the approximation

$$\cos\alpha = 1 - \tfrac{1}{2}\alpha^2 \tag{41}$$

giving

$$\cos\alpha = 1 - \tfrac{1}{2}\left(\frac{S}{C}\right)^2 - \frac{f(N-C)}{C(N-1)}. \tag{42}$$

6 Very shallow inclinations

If the true inclination is very shallow it becomes difficult to distinguish normal from reverse polarity. Consequently it may be a simple matter to make the mistake of analysing a mixture of normal and reverse polarity data as if they were all observations from a single polarity population. The result of such an error would be one inferred inclination (instead of two) which would be biased towards zero, a precision estimate heavily biased towards too small a value and an apparent 'loss' of a reversal.

To avoid such an error care must be taken with shallow inclination data, particularly if the data are reasonably symmetrical about zero. A histogram of the inclinations should be drawn up to see if the distribution is bimodal and has a minimum at zero, since this would imply a mixing of reverse and normal polarity data. Additionally, if the sign of an observed inclination is correlated with the specimen position along the core then this is also evidence for the presence of mixed polarity data. If it becomes apparent that a mixed polarity sample is being analysed then the data should be separated according to the position of the specimen along the core. If the inclination data are plotted against specimen position along the core it should be possible to pick a point or region where the transition occurred and the data may be classified as normal or reverse depending on whether the specimen was above or below this point. If such a point or region cannot be located it is an indication that the true mean inclinations are too shallow for normal and reverse polarity to be distinguished with the available data.

7 Conclusions

It has been shown that the estimates given by Kono (1980) for the inclination of the true mean direction and for the precision are biased, in many instances quite seriously so. As a consequence of this bias in his estimates (particularly in the estimate for κ) his error analysis is suspect.

An examination of the case of steep inclinations and low values of κ gives good reason to believe that this situation is fundamentally intractable if declination data are not available. The problem arises because the indeterminacy of declination makes it impossible to distinguish observations on the near side of the vertical from those on the far side, leading to a 'folding back' or aliasing of the observed inclinations.

Given that the amount of 'folding back' of the distribution at the vertical is negligible a set of observed inclinations, $I_1, I_2, \ldots, I_N$, may be analysed in the following manner to obtain unbiased estimates for the inclination, I_0, of the true mean direction and the precision parameter of the distribution from which the inclinations were obtained, together with error estimates on each.

Define θ_0 as the complement of I_0 (i.e. $\theta_0 = 90^\circ - I_0$), θ_i as the complement of I_i, $\hat{\theta}_0$ as the maximum likelihood estimate for θ_0, C as

$$C = \Sigma \cos(\hat{\theta}_0 - \theta_i) = \cos\hat{\theta}_0 \, \Sigma \cos\theta_i + \sin\hat{\theta}_0 \, \Sigma \sin\theta_i \qquad (34)$$

and S as

$$S = \Sigma \sin(\theta_0 - \theta_i) = \sin\hat{\theta}_0 \, \Sigma \cos\theta_i - \cos\hat{\theta}_0 \, \Sigma \sin\theta_i$$

where the summations run over i from 1 to N. The value of $\hat{\theta}_0$ is then a solution of

$$N \cos\hat{\theta}_0 + (\sin^2\hat{\theta}_0 - \cos^2\hat{\theta}_0) \, \Sigma \cos\theta_i - 2 \sin\hat{\theta}_0 \cos\hat{\theta}_0 \, \Sigma \sin\theta_i = 0 \qquad (19)$$

which may be obtained by an iterative method. It is always obvious which solution is the correct one. However, if

$$U = \tfrac{1}{2}N\left(\operatorname{cosec}^2\hat{\theta}_0 - \frac{C}{N-C}\right) \tag{19a}$$

then U is negative with the correct solution for $\hat{\theta}_0$ and positive for the two incorrect solutions. The maximum likelihood estimate $\hat{k}$ for κ is then given by

$$\hat{k} = \frac{N}{2(N-C)} \tag{18}$$

and if k is defined as

$$k = \frac{N-1}{2(N-C)} \tag{20}$$

then $(1/k)$ is an unbiased estimated for $(1/\kappa)$. Furthermore

$$\frac{(N-1)\kappa}{k} \sim \chi^2_{(N-1)} \tag{21}$$

and so confidence limits for κ may be obtained by the method of Cox (1969, 1977); the observed value of k may be compared with the observed value from another study, either of inclinations only or the more normal palaeomagnetic study with declination data as well (see McFadden 1980a) by using the F distribution; and the observed value may be compared with a well-defined value for κ either by direct use of the chi-square distribution or via the F distribution (see McFadden 1980b). However, in each case it must be remembered that the number of degrees of freedom is half that for an observed k derived from a study with both declination and inclination data.

Finally, an unbiased estimate, I, for I_0, conditional on the observed value of C, is given by

$$I = \left(90° - \theta_0 + \frac{180S}{\pi C}\right). \tag{40}$$

The angle of confidence, α, in this estimate is given by

$$\cos\alpha = 1 - \tfrac{1}{2}\left(\frac{S}{C}\right)^2 - \frac{f(N-C)}{C(N-1)} \tag{42}$$

where f is the critical value of the F distribution with 1 and $(N-1)$ degrees of freedom. This error angle is conditional on the observed value of C and independent of the precision parameter κ.

In all instances of application it must be remembered that some of the distributions used to derive the above analysis were 'guessed at' and confirmed by Monte Carlo experiments, not theoretically derived. Consequently care should be exercised in cases of marginal significance.

Acknowledgment

The authors would like to thank Dr J. Kirschvink for having raised the problem of identifying a reversal when the inclinations are very shallow.

References

Briden, J. C. & Ward, M. A., 1966. Analysis of magnetic inclination in borecores, *Pure appl. Geophys.*, **63**, 133–152.

Cox, A., 1969. Confidence limits for the precision parameter κ, *Geophys. J. R. astr. Soc.*, **17**, 545–549.

Cox, A., 1977. Erratum: Confidence limits for the precision parameter κ, *Geophys. J. R. astr. Soc.*, **48**, 136.

Fisher, R. A., 1953. Dispersion on a sphere, *Proc. R. Soc. A*, **217**, 295–305.

Kono, M., 1980. Statistics of palaeomagnetic inclination data, *J. geophys. Res.*, **85**, 3878–3882.

McFadden, P. L., 1980a. The best estimate of Fisher's precision parameter κ, *Geophys. J. R. astr. Soc.*, **60**, 397–407.

McFadden, P. L., 1980b. Testing a palaeomagnetic study for the averaging of secular variation, *Geophys. J. R. astr. Soc.*, **61**, 183–192.

Appendix: numerical example

The inclination values 62.4, 61.6, 50.2, 65.2, 53.2, 61.4, 74.0, 60.0, 52.6 and 71.8 were obtained as a random sample from a Fisher population with a precision parameter of 40 and inclination of the true mean direction equal to 60°. These then give values of θ_i as 27.6, 28.4, 39.8, 24.8, 36.8, 28.6, 16.0, 29.4, 36.4 and 18.2 respectively. The analysis of these data is as follows.

$$\Sigma \cos\theta_i = 8.708 \qquad \Sigma \sin\theta_i = 4.748$$

and so from equation (19) $\hat{\theta}_0$ is a solution of

$$10\cos\hat{\theta}_0 + 8.708(\sin^2\hat{\theta}_0 - \cos^2\hat{\theta}_0) - 9.496\sin\hat{\theta}_0\cos\hat{\theta}_0 = 0.$$

This gives three solutions of 10.5°, 27.7° and 129.4°. Quite evidently the second solution is the correct one and therefore

$$\hat{\theta}_0 = 27.7°.$$

Substituting the above values into equation (34) gives

$$C = 9.917 \quad \text{and} \quad S = -0.1560$$

and therefore

$$\frac{180S}{\pi C} = -0.90°$$

giving the estimate I for I_0 as

$$I = (90° - \hat{\theta}_0 - 0.90) = 61.4°$$

from equation (40).

From equation (20) k is given as

$$k = \frac{9}{2(10 - 9.917)} = 54.2.$$

Using the distribution of equation (21)

$$\frac{9\kappa}{k} > 19.0 \quad \text{with probability } 0.025$$

and

$$\frac{9\kappa}{k} < 2.70 \quad \text{with probability } 0.025.$$

Consequently κ lies between 16.3 and 114.4 with 95 per cent confidence, compared with the known value (in this instance) of 40. Had this estimate for κ been obtained from a study with both declinations and inclinations then the 95 per cent confidence limits for κ would have been 24.8 and 94.9, indicating the loss of information in having only the inclinations.

To determine 95 per cent confidence limits for I a two-tailed test must be used. The F distribution with 1 and 9 degrees of freedom will exceed 7.21 with a probability of 0.025 giving α from equation (42) as 6.7°. Consequently the inclination of the true mean direction lies between 54.7° and 68.1° with 95 per cent confidence, compared with the known value (in this instance) of 60.0°.

If the estimate $k = 54.2$ had been obtained from an analysis of both declinations and inclinations this would imply a value of $R = 9.83$ giving the semi-angle of the cone of 95 per cent confidence about the estimated mean direction as 6.6°, almost exactly the same as the symmetrical 95 per cent confidence limit obtained for the estimate I. This is to be expected since an error value in the inclination gives half the information given by a cone of confidence about an estimated direction.

Given that one wishes to test, on the basis of the observed data, whether the inclination of the true mean direction is less than (for example) 65° a one-tailed test must be performed (a fairly rare situation). The F distribution with 1 and 9 degrees of freedom will exceed 5.12 with a probability of 0.05 giving α from equation (42) as 5.7°. The estimated inclination of 61.4° is only 3.6° less than the hypothesized 65° and so given the alternative hypothesis that the true mean inclination is less than 65° there is no statistical reason for rejecting (at the 95 per cent level of confidence) the null hypothesis that the inclination of the true mean direction is 65°.