Mon. Not. R. astr. Soc. (1984) **207**, 637–647

Indirectional statistics and the significance of an asymmetry discovered by Birch

David G. Kendall and G. Alastair Young* *Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB*

Received 1983 August 16; in original form 1983 July 25

Summary. Birch reported an apparent 'statistical asymmetry of the Universe'. The authors here develop 'indirectional analysis' as a technique for investigating statistical effects of this kind and conclude that the reported effect (*whatever may be its origin*) is strongly supported by the observations. The estimated pole of the asymmetry is at RA $13^h\,30^m$, Dec. $-37°$. The angular error in its estimation is unlikely to exceed 20–30°.

1 Introduction

Birch (1982) examined the position angles of elongation and polarization of 137 high luminosity classical double radio sources and reported a surprising topographic variation in the angular difference between the two position angles. He remarked that this could have an explanation on the cosmic scale and might even be an indication that 'the Universe is rotating'. The statistical significance of the Birch effect was queried by Phinney & Webster (1983), who moreover remarked that the effect, if it exists, might be an artefact associated with an error in the correction for galactic Faraday rotation. They invited us to examine the evidence for the effect, suggesting that we replace the first 45 measures in Birch's table 1 (PB Sample from 3CR) by more recent measures on 42 similar objects by Conway *et al.* (1983). This we have done, the number of objects in the total sample being now $N = 134$. We find the effect discovered by Birch to be supported to a high level of significance. We have nothing here to add to the debate concerning the interpretation of the effect, but feel that the statistical techniques used, which have some novel features, may be of interest to astronomers as well as statisticians and are applicable to other similar studies.

2 Indirectional statistics

We are concerned here with a problem in a novel field which might be called *indirectional statistics*. Directional statistics, now a much studied subject, began with von Mises (1918) who introduced the probability law

$$\frac{\exp(\kappa \cos\theta)}{2\pi I_0(\kappa)}\, d\theta \qquad (-\pi < \theta \leqslant \pi), \tag{1}$$

* G. A. Young was supported during the course of this work by a SERC research studentship and by a Benefactors' research studentship at St John's College, Cambridge.

which bears his name, as the simplest non-uniform distribution for the direction from the centre of a circle S^1 to a point labelled θ on its circumference. [Here $I_n(\kappa)$ denotes the modified Bessel function of the first kind and of the nth order.] Its importance was more clearly realized when Fisher (1953) introduced the corresponding probability law

$$\frac{\exp(\kappa \cos \psi)}{(4\pi/\kappa) \sinh \kappa} \sin \psi \, d\psi \, d\phi \tag{2}$$

as the simplest non-uniform distribution for the direction from the centre of a sphere S^2 to a point labelled (ψ, ϕ) on its surface, ψ being the polar angle. The present problem differs from those considered by von Mises and Fisher in being partly projective in character; the position angle of an axis of plane symmetry is only defined up to multiples of π (rather than 2π). This is why we choose to speak here of an indirection. The so-called elongation of the double radio source has this character, and so has the polarization. Thus the discrepancy angle Δ studied by Birch is actually the acute angle embraced by these two indirections in the tangent plane to the celestial sphere. The unit vector $\mathbf{p}$ locating the line-of-sight to the object is a direction in the ordinary sense. Birch counts Δ as positive if the rotation from the geometric axis to the nearest arm of the polarization axis is (say) clockwise about the line-of-sight and negative otherwise. Plainly Δ lies between $-\pi/2$ and $+\pi/2$ and furthermore these two extremes are to be identified. In fact Δ is also an indirection and when so measuring it we are making use of the familiar 'halving' map which converts the real projective space RP^1 into S^1.

An adaptation of the von Mises law,

$$\frac{\exp(\kappa \cos 2\Delta)}{\pi I_0(\kappa)} \, d\Delta \qquad (-\pi/2 < \Delta \leqslant \pi/2), \tag{3}$$

gives us a useful way of describing a non-uniform indirectional distribution symmetrical about $\Delta = 0$. To accommodate asymmetries of the kind in question here we add a multiple of $\sin 2\Delta$ to the exponent. If the asymmetry is to be controlled by a topographic regression, then to a first approximation it will be appropriate to replace the exponent at (3) by

$$\alpha \cos 2\Delta + (\boldsymbol{\lambda} \cdot \mathbf{p}) \sin 2\Delta. \tag{4}$$

Here $\mathbf{p}$ is the unit line-of-sight (directional) vector and $\boldsymbol{\lambda}$ is a (directional) vector parameter whose modulus $\beta = |\boldsymbol{\lambda}|$ measures the strength of the topographic effect. In $\boldsymbol{\mu} = \boldsymbol{\lambda}/\beta$ we have a unit vector – also a true direction – which locates the positive pole of the effect, the 'topographic pole'. To maintain total probability unity we must replace $I_0(\kappa)$ at (3) by

$$I_0\{\surd[\alpha^2 + (\boldsymbol{\lambda} \cdot \mathbf{p})^2]\}.$$

Finally, to complete the model, we must adjoin to this conditional Δ-distribution a multiplicative factor, say

$$f(\mathbf{p}) \, d\mathbf{p} \qquad (\mathbf{p} \in S^2) \tag{5}$$

which describes the distribution of source directions. This might, for instance, have the Fisher form (2), but in the present problem we leave it unspecified. It will not affect our calculations and is descriptive more of the social science of observatory building than of the natural science of the Universe.

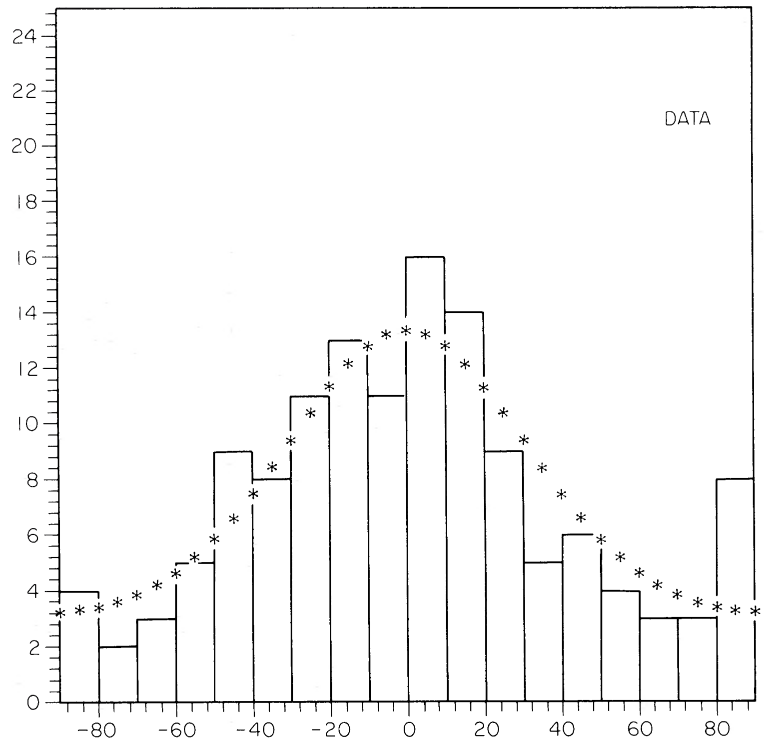

Figure 1. A fit (shown by ***) of the indirectional version of the von Mises distribution ($\alpha = 0.703$) to the whole ensemble of Δ-values, when the topographic relationship is ignored. This is the so-called marginal distribution of Δ.

3 Model for the Birch effect

We now have a model (model II) for the Birch effect expressed in the form of a joint probability law for $\mathbf{p}$ and Δ,

$$f(\mathbf{p})\,d\mathbf{p}\,\frac{\exp(\alpha\cos 2\Delta + (\boldsymbol{\lambda}\cdot\mathbf{p})\sin 2\Delta)}{\pi I_0\{\sqrt{[\alpha^2 + (\boldsymbol{\lambda}\cdot\mathbf{p})^2]}\}}\,d\Delta. \tag{6}$$

Apart from the f-factor it involves the four parameters α, λ_1, λ_2, λ_3. When $\beta = |\boldsymbol{\lambda}| = 0$, it reduces to the null model (model I) in the form

$$f(\mathbf{p})\,d(\mathbf{p})\,\frac{\exp(\alpha\cos 2\Delta)}{\pi I_0(\alpha)}\,d\Delta. \tag{7}$$

Fig. 1 shows that the marginal Δ-distribution is approximately of this type.

To test model I (the null hypothesis) against model II as alternative, we write down the log-likelihoods L_{I} and L_{II} for the data with respect to the two models, and maximize each of them separately with respect to the parameters; the appropriate Neyman–Pearson test statistic is then

$$T = \max(L_{\mathrm{II}}) - \max(L_{\mathrm{I}}).$$

From likelihood-ratio theory when N is large we shall expect T to have a sampling distribution of the form $\chi_3^2/2$ on the null hypothesis; it will indicate evidence for a genuine effect ($\beta > 0$) by a significantly large positive value provided that the noise associated with the small value of N is not too big. Notice that the density $f(\mathbf{p})$ contributes nothing to T.

The maximization of L_{I} is straightforward, but the maximization of

$$L_{\mathrm{II}} = \sum_{r=1}^{N} \{\log f[\mathbf{p}^{(r)}] + \alpha \cos 2\Delta^{(r)} + [\boldsymbol{\lambda} \cdot \mathbf{p}^{(r)}] \sin 2\Delta^{(r)} - \log \pi - \psi^{(r)}\}$$

where

$$\Psi^{(r)} = \log I_0 \langle \sqrt{\{\alpha^2 + [\boldsymbol{\lambda} \cdot \mathbf{p}^{(r)}]^2\}} \rangle,$$

is a much more delicate matter because the geometric disposition of the sources $\mathbf{p}^{(r)}$ then plays a role. The question is important and so we give a careful discussion; the reader who will accept the results on trust can jump at once to Section 4 below. The terms in f and π cancel out and can be ignored, and the remaining terms in α and $\boldsymbol{\lambda}$ are linear apart from the Ψ-term, so we start by examining the latter. Now

$$\Psi(\mathbf{p}; \alpha, \lambda_1, \lambda_2, \lambda_3) = \log I_0\{\sqrt{[\alpha^2 + (\lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3)^2]}\}$$

is a multiply real-analytic function of $(\alpha, \lambda_1, \lambda_2, \lambda_3)$ over all of R^4, because I_0 is an even entire function of its argument, and $I_0 \geqslant 1$ always.

To test $\Psi(\mathbf{p})$ for convexity, as a function of α and the λ_i's, for any given $\mathbf{p}$, we must inspect the Hessian quadratic form which with dummy variables $(a, \mathbf{L})$ reduces to

$$H(\mathbf{p}) = \frac{I_0'' I_0 - I_0'^2}{Q^2 I_0^2} \quad [a\alpha + (\mathbf{L} \cdot \mathbf{p})(\boldsymbol{\lambda} \cdot \mathbf{p})]^2 + \frac{I_0'}{Q^3 I_0} [a(\boldsymbol{\lambda} \cdot \mathbf{p}) - \alpha(\mathbf{L} \cdot \mathbf{p})]^2.$$

Here each modified Bessel function has argument $Q = \sqrt{[\alpha^2 + (\boldsymbol{\lambda} \cdot \mathbf{p})^2]}$. Now $I_0' > 0$ for $Q > 0$ and $I_0' = 0$ if $Q = 0$, while an application of the Schwarz inequality shows that $I_0'' I_0 \geqslant I_0'^2$ with equality if and only if $Q = 0$. Thus $H(\mathbf{p}) \geqslant 0$ always. In verifying this, notice that the singularity at $Q = 0$ is only apparent; in fact, near to $Q = 0$,

$$H(\mathbf{p}) = \frac{1}{2} [a^2 + (\mathbf{L} \cdot \mathbf{p})^2] - \frac{3}{16} [a\alpha + (\mathbf{L} \cdot \mathbf{p})(\boldsymbol{\lambda} \cdot \mathbf{p})]^2$$

$$- \frac{1}{16} [a(\lambda \cdot \mathbf{p}) - \alpha(\mathbf{L} \cdot \mathbf{p})]^2 + O(Q^2),$$

so that at $Q = 0$ we shall have $\alpha = (\boldsymbol{\lambda} \cdot \mathbf{p}) = 0$ and

$$H(\mathbf{p}) = \frac{1}{2} [a^2 + (\mathbf{L} \cdot \mathbf{p})^2] \geqslant 0.$$

It follows that $\Psi(\mathbf{p})$ is weakly convex over the whole of R^4 for every choice of $\mathbf{p}$.

Now L_{II} as a function of $(\alpha, \boldsymbol{\lambda})$ is the sum of two components; one is linear and the other is equal to

$$- \sum_{r=1}^{N} \Psi [\mathbf{p}^{(r)}]$$

Thus L_{II} is weakly concave everywhere in R^4, and it will be everywhere strongly concave if we can show that, for each choice of $(\alpha, \boldsymbol{\lambda})$,

$$\sum_{r=1}^{N} H[\mathbf{p}^{(r)}] > 0 \qquad \text{when} \qquad a^2 + (\mathbf{L} \cdot \mathbf{L}) > 0.$$

Let then α and $\boldsymbol{\lambda}$ be arbitrary and given, and let a and $\mathbf{L}$ be also fixed, with $a^2 + (\mathbf{L} \cdot \mathbf{L}) > 0$. We must show that $H[\mathbf{p}^{(r)}] > 0$ for some r.

(i) Let $Q(\mathbf{p}) = 0$, so that $\alpha = (\boldsymbol{\lambda} \cdot \mathbf{p}) = 0$. Then $H(\mathbf{p})$ vanishes if and only if $a = 0$ and $(\mathbf{L} \cdot \mathbf{p}) = 0$. (Notice that $a = 0$ implies $\mathbf{L} \neq \mathbf{0}$!)

(ii) Let $Q(\mathbf{p}) > 0$. Then $H(\mathbf{p})$ vanishes if and only if both

$$a\alpha + (\mathbf{L} \cdot \mathbf{p})(\boldsymbol{\lambda} \cdot \mathbf{p}) = 0,$$

and

$$-(\mathbf{L} \cdot \mathbf{p})\,\alpha + a(\boldsymbol{\lambda} \cdot \mathbf{p}) = 0.$$

Because $Q(\mathbf{p}) > 0$ we know that the determinant of coefficients in these linear equations must vanish, so $H(\mathbf{p})$ again equals zero if and only if $a = (\mathbf{L} \cdot \mathbf{p}) = 0$.

Putting (i) and (ii) together, we see that

$$\sum_{r=1}^{N} H[\mathbf{p}^{(r)}]$$

will fail to be everywhere strictly positive if and only if $[\mathbf{L} \cdot \mathbf{p}^{(r)}] = 0$ for some one non-vanishing vector $\mathbf{L}$ and every r.

Now the direction of this $\mathbf{L}$ will be the pole of a great circle, and so we have proved the

Theorem. L_{II} is everywhere strictly concave if and only if the sources do not all lie on a single great circle.

This makes clear what is the topographic situation in which our method breaks down. Of course in the present problem the condition is amply satisfied. In passing we note that we may expect large sampling errors in the estimates of the parameters whenever all the sources lie close to some one great circle, but in the present study the distribution of sources, patchy though it is, need not be expected to produce such disasters.

In order to assert that L_{II} has a global maximum at some finite point (necessarily a point of differentiability) we have to add to its strict concavity a guarantee that L_{II} tends to minus infinity in all directions in R^4. So choose $(\alpha^*, \boldsymbol{\lambda}^*)$ so that $\alpha^{*2} + \Sigma_1^3 \lambda_i^{*2} = 1$, put $\alpha = \alpha^* t$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}^* t$, and consider what happens when $t \to \infty$. We can of course assume that the condition of the theorem is satisfied. Now either (i) $\alpha^* \neq 0$, or (ii) $\alpha^* = 0$ and $\boldsymbol{\lambda}^* \neq \mathbf{0}$. In the case (i), each $Q^{(r)} \to \infty$, while in case (ii) $Q^{(r)}$ vanishes if $[\boldsymbol{\lambda}^* \cdot \mathbf{p}^{(r)}] = 0$ and $\to \infty$ otherwise. But for $\boldsymbol{\lambda}^* \neq 0$, we must have $[\boldsymbol{\lambda}^* \cdot \mathbf{p}^{(r)}] \neq 0$ *for at least one* r, so we can conclude that some at any rate of the $Q^{(r)}$s tend to infinity and that all the others vanish.

On looking back at the formula for L_{II} we see that its component terms can be sorted into two groups. Those associated with vanishing Qs all vanish, because for these terms $\alpha = [\boldsymbol{\lambda} \cdot \mathbf{p}^{(r)}] = 0$ and $\log I_0[Q^{(r)}] = 0$. The non-constant contribution of each one of the remaining terms is of the form

$$Q^{(r)} \cos[2\Delta^{(r)} + \epsilon^{(r)}] - Q^{(r)} + \frac{1}{2}\log[2\pi Q^{(r)}] + o(1)$$

as $Q^{(r)} \to \infty$, and so converges to $-\infty$ unless the cosine is +1. In practice it is clear that we can ignore accidents associated with exact Δ-values, and so we can confidently expect L_{II} to possess a finitely located global maximum which will be accessible to numerical analysis because of the global smoothness of L_{II}.

4 Method

The maximum of L_{II} can now be found by any of the standard packaged iterative routines as soon as we have suitable initial values. Our approach to obtaining starting values is to make a first-order approximation to the modified Bessel function I_0; with this approximation, suitable when α and β are not too large, the problem linearizes, giving initial estimates of $(\alpha, \boldsymbol{\lambda})$ in simple algebraic form and at the same time a linear approximation T_1 to the test statistic T itself.

Let A be the 3×3 matrix with components

$$a_{\mathrm{ij}} = \sum_{r=1}^{N} p_{\mathrm{i}}^{(r)} p_{\mathrm{j}}^{(r)}$$

and let B be the inverse of A (which certainly exists in view of the great circle assumption). We find that the first-order estimates of the parameters in model II are

$$\hat{\alpha} = \frac{2}{N} \sum_{r=1}^{N} \cos 2\Delta^{(r)},$$

and

$$\hat{\boldsymbol{\lambda}} = \mathsf{B}\mathbf{v}, \qquad \hat{\beta} = |\mathsf{B}\mathbf{v}|,$$

where $\mathbf{v}$ is the vector with components

$$v_{\mathrm{i}} = 2 \sum_{r=1}^{N} p_{\mathrm{i}}^{(r)} \sin 2\Delta^{(r)}.$$

The corresponding first-order test statistic is

$$T_1 = \frac{1}{4} \mathbf{v}^{\mathrm{T}} B \mathbf{v},$$

where $\mathbf{v}^{\mathrm{T}}$ is the transpose of $\mathbf{v}$.

5 Results

The results of the exact and the first-order calculations are as follows ($N = 134$). It is interesting that the first-order approximation should work so well for values of α and β which are far from small.

We have remarked that for large N the sampling distribution of T will be $\chi_3^2/2$ on the null hypothesis. The corresponding significance levels are $P = 0.005$ for $T = 6.42$ and $P = 0.001$ for $T = 8.13$. So the evidence for the Birch effect is significant at about the $P = 0.001$ level.

It is highly desirable to check this result in a manner which does not make use of asymptotic theory appropriate only for $N \to \infty$. Accordingly we used a data-based simulation test, set up as follows: for each simulation, hold the positions $\mathbf{p}^{(r)}$ of the sources fixed, and also keep the overall assemblage of observed $\Delta^{(r)}$s fixed, but assign the Δ-values randomly to the source positions. For each such scrambling of the Δ-values recompute the test statistic. Compare the observed value with the collection of simulated values.

Table 1. Estimates of parameters.

Basis	T	$\hat{\alpha}$	$\hat{\beta}$	Pole RA (h m)		Pole Dec.
First-order	7.32	0.663	0.823	13	30	−37°.2
Exact	8.02	0.727	0.933	13	30	−37°.4

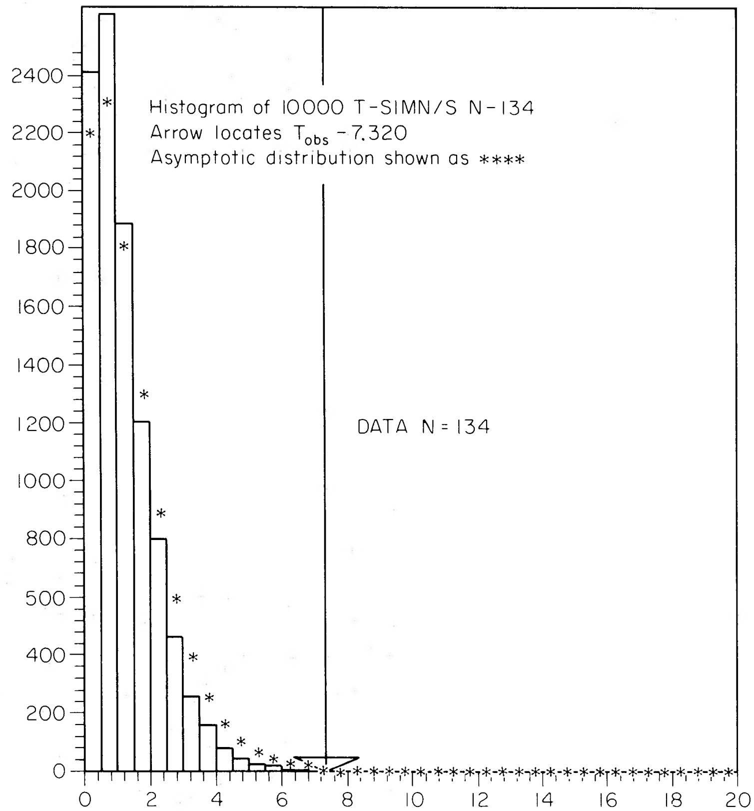

Figure 2. First-order analysis: the histogram of 10 000 simulated T_1-values. The T_1-value obtained from the data is shown by an arrow.

When the first-order statistic T_1 is being employed it is easy to perform this test with 10 000 simulations. When this is done a P-value of the order of 0.0005 is obtained, in agreement with the above asymptotic result. The histogram of simulated T_1-values is very striking (Fig. 2). It follows roughly, but far from exactly, the course of the $\chi_3^2/2$ curve, and the observed T_1-value stands well out in the 'tail'.

A data-based simulation test was also carried out using the exact test statistic T, but computation of that is slow and it is therefore very time-consuming to do more than a few hundred simulations. A few such experiments were carried out, these supporting the earlier conclusions.

6 Discussion

6.1 DISSECTING THE DATA

The 134 data values arise naturally in four sets: 134 = 42 (Conway) + 39 (Laing) + 28 (Ekers) + 25 (Conway). We repeated both types of analysis for each of the four subsets. The exact T-values were, respectively, 3.289, 4.627, 1.308 and 2.964. A useful, if rough, analysis of variance can now be constructed. In Table 2 we have doubled the T-values to get approximate χ_3^2 random variables.

The 'between groups' χ^2 is not significant, and indeed is below expectation. So there is no reason to suppose that the varying responses from the different groups reflect any more than the high sampling errors associated with the small sizes of the groups. What the

Table 2. Analysis of variance.

Data	$2T$-values	d.f.
Conway 1 ($N=42$)	6.578	3
Laing ($N=39$)	9.254	3
Ekers ($N=28$)	2.616	3
Conway 2 ($N=25$)	5.928	3
Between groups	8.335	9
Pooled	16.041	3
Total	24.376	12

individual subsets have to tell us is largely drowned by noise, but when they speak with one voice, in the 'pooled' line, the message is clear.

The two larger groups do in fact yield individual T-values which, compared with the asymptotic null distribution, are significant at the levels $P = 0.09$ and 0.025 respectively, but on account of the small sample sizes these should not be taken too seriously. For these two subsets simulation tests of 10 000 simulations (using T_1) gave P-values of 0.04 and 0.03 respectively.

We feel that it is worth reporting here the results of an analysis of variance test based on the contrast between the 'pooled' and 'between groups' mean squares. This is

$$F_{3,9} = \frac{16.041/3}{8.335/9} = 5.77.$$

The significance levels for this situation,

$$2.5 \text{ per cent}: F = 5.08,$$

$$1.0 \text{ per cent}: F = 6.99,$$

indicate that the overall Birch-effect is also significant when tested against the noise represented by the discrepancies between the four groups.

6.2 POWER OF THE TEST

It is desirable to check the power of the test, i.e. to see how it responds to a 'faked' effect. Accordingly we kept the source positions fixed, we set α and $\beta = |\boldsymbol{\lambda}|$ to the values estimated from the data by the exact method (see Table 1) and then assigned Δ-values to sources by sampling from the model II probability density with the pole $\boldsymbol{\mu}$ of the topographic regression set to each in turn of four widely different positions. The four resulting data-sets were then analysed, by the exact procedure, and yielded T-values as shown in Table 3.

In each case the faked effect is successfully detected (at $P = 2.5$ per cent or better; see the values of T) and the pole of the topographic regression is identified to within about 20–30° (angle between true and estimated directions of maximum effect). Experiment 2 indicates that reasonably large errors can occur purely as a result of sampling from the model II probability density.

A useful graphical illustration of our results makes use of the following measure of the

Table 3. Power of the test.

Experiment	1	2	3	4
Topographic pole:				
RA (h)	6	6	18	18
Dec. (°)	45	−45	45	−45
T-value	14.89	4.65	5.56	15.49
Estimated pole:				
RA (h m)	5 31	2 41	18 33	18 4
Dec. (°)	52.5	−62.6	39.4	−21.9
Angular error in pole (°)	8.9	33.0	8.3	23.1

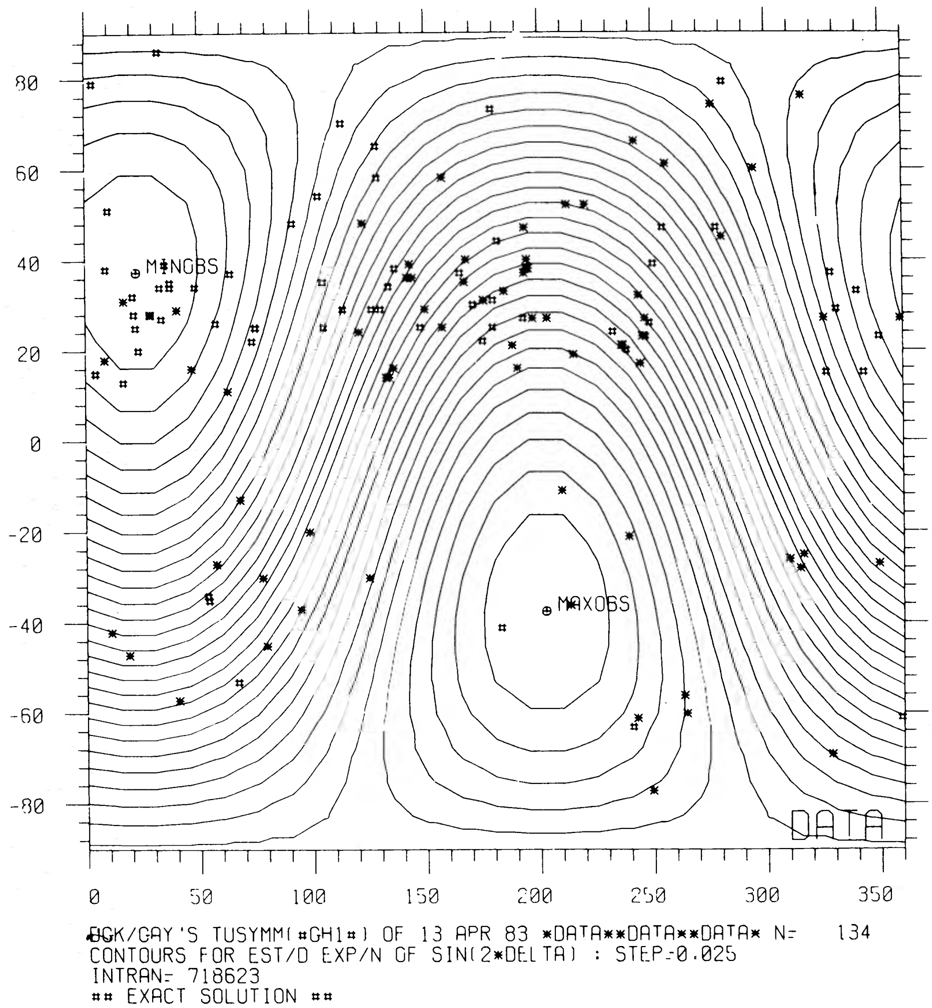

Figure 3. The contour plot of $Z = \mathbf{E} \sin 2\Delta$ as a function of position $\mathbf{p}$, based on the whole set of parameters estimated from the data by the exact procedure. The points indicate positions of observed sources (* for $\Delta > 0$, # for $\Delta < 0$). The estimated topographic pole is at MAXOBS. The vertical scale is of declination; the horizontal scale is of θ (degrees) where $\mathrm{RA} = 24\theta/360$. Note that RA is shown as increasing from left to right.

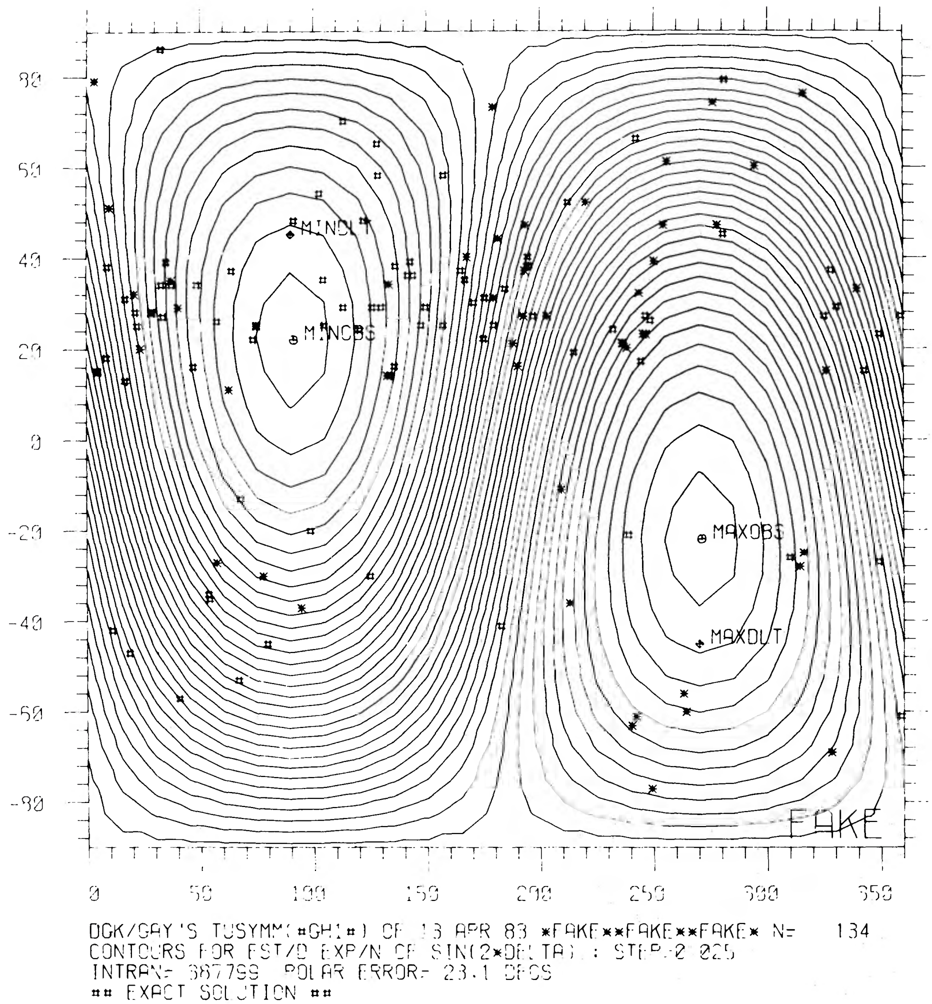

Figure 4. As for Fig. 3, but now, while the source positions are the same as before, the Δ-values are simulated using model II and the data-based values of α and β. The topographic pole of the distribution sampled from is at MAXDLT.

average degree of 'twist' in direction $\mathbf{p}$, when the model II Δ-distribution, conditional on $\mathbf{p}$, has parameters α and $\boldsymbol{\lambda} = \beta\boldsymbol{\mu}$:

$$Z(\mathbf{p}) = E(\sin 2\Delta \mid \mathbf{p}) = \frac{\beta(\boldsymbol{\mu}\cdot\mathbf{p})}{Q}\,\frac{I_1(Q)}{I_0(Q)}.$$

Here $E(\cdot)$ denotes mathematical expectation with respect to the conditional distribution and

$$Q = \sqrt{[\alpha^2 + \beta^2(\boldsymbol{\mu}\cdot\mathbf{p})^2]}.$$

Figures 3 and 4 show contour plots of the estimated Z (with contour interval 0.025) for (a) the actual observations and (b) the fourth of the four sets of 'faked' observations of Table 3. On these plots the observed source positions are shown as * when $\Delta > 0$ and as # when $\Delta < 0$.

In Fig. 4 the point marked MAXDLT shows the position of the pole of the faked effect, whilst in both Figs 3 and 4 the estimated pole is shown as the point marked MAXOBS. We can thus, in Fig. 4, examine the pole of the faked topographic regression, the observed source positions, the random Δ-values and the resulting estimated regression in the same diagram.

6.3 EFFECT OF OBSERVATIONAL ERRORS

We are informed that the observed Δ-values may be in error by as much as 10–20°. This does not invalidate our analysis, though it may have diluted the true effect. To check this we made a repeat analysis (by the exact procedure) of the observed data when each Δ-value was perturbed by a random and randomly signed fraction of 20° as a point in the projective space RP^1. The result of this analysis was: T-value = 6.49. Estimates: $\alpha = 0.666$, $\beta = 0.852$, pole: RA $12^{\mathrm{h}}\,40^{\mathrm{m}}$, Dec. 1 $-48°.2$. This result, and our study above the power of our test procedure, suggest some uncertainty in our estimates in Table 1, though the true pole of the topographic regression is unlikely to be more than 20–30° from the estimated pole. It is clear therefore that an acceptable explanation of the Birch effect must indicate a pole as close as this to our estimated pole at RA $13^{\mathrm{h}}\,30^{\mathrm{m}}$, Dec. 1 $-37°.4$.

Acknowledgments

We are most grateful to E. S. Phinney and R. L. Webster for drawing our attention to this problem and for much helpful advice in the course of our analysis. With regard to the new technical term 'indirection', which seems to us apt, the reader may like to note the first (and slightly different) use of that word by Shakespeare:

> And thus doe we of wisdome and of reach
> With windlesses, and with assaies of Bias,
> By indirections finde directions out.

References

Birch, P., 1982. *Nature,* **298,** 451.

Conway, R. G., Birch, P., Davis, R. J., Jones, L. R., Kerr, A. J. & Stannard, D., 1983. *Mon. Not. R. astr. Soc.,* **202,** 813.

Fisher, R. A., 1953. *Proc. R. Soc. Lond. (A),* **217,** 295.

Phinney, E. S. & Webster, R. L., 1983. *Nature,* **301,** 735.

von Mises, R., 1918. *Phys. Zeits.,* **19,** 490.