Internal Report No. 65
Frank Marshall
May 15, 1975

STATISTICAL PROPERTIES OF ANISOTROPY MEASUREMENTS

The smallest anisotropy that can be measured with confidence is limited by the accuracy of the directional intensity measurements. This internal report will examine how the statistical uncertainties in one's data propagate into uncertainties in derived anisotropies. It will be shown how to determine confidence intervals for a data sample or conversely how many counts are needed in order to measure a given anisotropy with a desired level of confidence.

The first part of this report will define the terms used. The next section will be a general discussion of the mathematical derivations with the gory details relegated to an Appendix. The results and examples of their application will be given in the final section.

INTRODUCTION

This report is concerned with the measurements of anisotropies by sampling directional intensities. The IMP instruments, for example, accumulate rates in eight different sectors corresponding to eight equally spaced directions in space. This report will assume ideal instruments for which corrections due to dead times or finite opening anglès can be neglected. It is also assumed that the sectors have an equal width in angle.

Given the accumulated rates for the different sectors it is possible to fit a cosine expansion to the data. Let Y_i be the number of counts from sector i and θ_i be the angle corresponding to sector i. Then the Y_i's can be approximated by the function

$$\Psi(\theta_i) = A + A_1 \cos(\theta_i - \Phi)$$

The anisotropy amplitude is associated with A_1/A and the direction of the anisotropy is Φ.

Generally a least squares fit is made to the data using this function and the resulting values are called the anisotropy amplitude and direction of the data sample. The remainder of this report will concern the relationship between these computed values and the corresponding values of the distribution sampled.

MATHEMATICS

Given a measured anisotropy amplitude, an experimenter still needs to determine the anisotropy amplitude of the distribution sampled. Appendix 1 gives a derivation of the probability of different anisotropy amplitudes given a measured amplitude. The method used is as follows : 1) a transformation is made to different parameters to describe the anisotropy 2) the probability of measuring an anisotropy given a true anisotropy distribution is determined 3) an assumption is made as to the apriori likelihood of observing an anisotropy 4) Bayes Theorem is applied and 5) a transformation back to the original parameters is made.

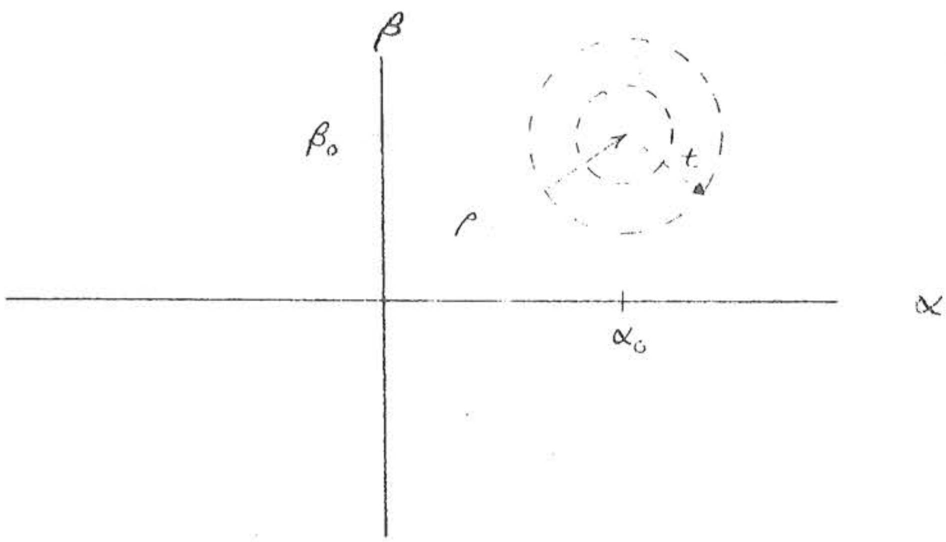
Although the variables A, A_1, and Φ give a convenient physical description of a distribution, they result in non-linear equations when one tries to make a least squares fit to a data sample. Consequently a change of variables is made to A, α and β where $\alpha = A_1 \cos\Phi$ and $\beta = A_1 \sin\Phi$

This is essentially just a transformation from polar to rectangular coordinates. With these new variables the least squares fitting results in linear equations which can readily be solved for A, α and β.

Uncertainties in the values of the Y_i's propagate into uncertainties in the values of A, α and β . The assumption is made that the uncertainties in the Y_i's are all equal. This assumption leads to an enormous simplification in the equations for A, α, and β and the corresponding uncertainties. It also makes subsequent probability calculations tractable. This assumption will be a good approximation if, for example, the anisotropy amplitude is small and the uncertainties in the Y_i's are due to Poisson statistics.

The next step is to determine the probability of observing a distribution that is characterized by A, α, and β given a true distribution that is exactly a first order cosine expansion. At this point another approximation is made. It is assumed that the relative uncertainties in A are small compared to the relative uncertainties in α or β . Again this will be a good approximation if the anisotropy amplitude is small.

The following diagram is a contour map of the relative probabilities of observing α and β given the corresponding parameters α_0 and β_0 of the true distribution.



The probability distribution for observing (α, β) is a two-dimensional Gaussian centered at (α_0, β_0) and is a function only of the distance from (α_0, β_0). This distance is labelled

t in the previous figure. From now on it will be assumed that the axes have been rotated so that $\beta_0 = 0$. Thus α_0 is proportional to the anisotropy amplitude.

The measured anisotropy amplitude is proportional to the distance to the origin (labelled ρ in the previous diagram) so the probability of observing an anisotropy amplitude is simply related to the probability of ρ. Thus the probability of observing any anisotropy amplitude given a true distribution completely characterized by A, A_1, and ϕ can be calculated.

The anisotropy amplitudes are divided by a scale factor z defined by:

$z = w\,(2/s)^{1/2}\,\sigma/A$ where $w = (\pi/s)/\sin(\pi/s)$

s = number of sectors

σ = uncertainties in Y_i's

A = average number of counts/sector

z is roughly the anisotropy amplitude expected if the true distribution is isotropic. w is a smoothing factor that corrects for the fact that sampling finite angle bins reduces the measured anisotropy amplitude. For 8 sectors and Poisson uncertainties

$z = 1.451\ N^{-1/2}$ where N is the total number of counts

Let $r = \delta_{observed}/z$

and $x = \delta_{true}/z$

Then

$$P(r \mid x) = e^{-r^2/2} \times e^{-x^2/2}\, I_0(xr)$$

Some of the properties of this distribution are displayed in Figure 1 . The solid line shows the most probable value of r as a function of x. Also shown are the 68.2% and 95.4% confidence intervals. There is a 68.2% probability that an observed r will be within the 68.2% confidence interval. Furthermore the interval is chosen to minimize its length.

Now making an assumption for the apriori likelihood of a given anisotropy amplitude and applying Bayes' Theorem, one can compute the probability that the true distribution is characterized by a given anisotropy amplitude if one measures a particular anisotropy amplutude. Unless otherwise stated all amplitudes are assumed equally likely apriori.

The results of this derivation are presented in the next section.

RESULTS

These results are only valid for anisotropy amplitudes much less than one and uncertainties in the count rates in all sectors roughly equal.

As discussed in the last section and shown in the Appendix the anisotropy amplitude and direction can be computed from the observed Y_i's by fitting the funcitonal form:

$$A + \alpha \cos \Theta_i + \beta \sin \Theta_i$$

and setting $A_1 = (\alpha^2 + \beta^2)^{1/2}$ and $\phi = \tan^{-1}(\beta / \alpha)$

Using the approximation that all the uncertainties in the Y_i's are equal, the best fit is made in the least squares sense when

$A = (1/s)\Sigma Y_i$ where s is the number of sectors

$$\alpha = (2/s)\Sigma Y_i \cos \Theta_i$$

$$\beta = (2/s)\Sigma Y_i \sin \Theta_i$$

The anisotropy amplitude is given by

$$\delta = w\ (A_1/A)$$

where w is the smoothing factor defined in the previous section. Again we define the scaling factor z

$$z = w\ (2/s)^{1/2}\ \sigma / A$$

and $r = \delta_{obs}/z$

$$x = \delta_{true}/z$$

z is approximately the anisotropy amplitude expected if the true distribution is isotropic.

Now the probability of x given the observation r is :

$$P(x \mid r) = (2/\pi)^{1/2} e^{-r^2/4}\, e^{-x^2/2}\, I_0(xr) \,/\, I_0(r^2/4)$$

where I_0 is the zero order modified Bessel function.
This probability distribution is plotted for r = 0, 2, and 3 in Figure 2 . Figure 3 is a plot of the mode and 68.2% and 95.4% confidence intervals for this distribution.

As mentioned above z is approximately the anisotropy due to the noise on the measurement. If the measured anisotropy amplitude if much larger than z (that is, when r is much larger than one), z becomes the sigma for the true anisotropy distribution. In the other limit, as the measured anisotropy becomes much smaller than z (r nearly zero) , the most likely true anisotrppy is zero and the distribution becomes a one-sided Gaussian. In fact, for all values of r less than the square root of 2 the mostly likely value of the true anisotropy is zero.

Figure 4 is a plot of the relative accuracy of a measurement as a function of the total number of counts (assuming Poisson statistics). The relative accuracy i~~f~~s defined as the length of the 68.2% confidence interval divided by the most probable anisotropy amplitude. The uppermost line shows how many counts are required so that the 68.2% confidence interval is only 10% of the most probable anisotropy. Also shown is the 60% accuracy line. The bottom line shows where the most probable value is zero, below which only upper limits can be placed on the value of the true anisotropy.

Thus far we have assumed that all anisotropy amplitudes are equally likely apriori. Other assumptions are possible. For example, one might assume that all values of α and β are apriori equally likely This corresponds to assuming that the apriori probability of an anisotropy amplitude is proportional to the amplitude. This assumption has the result that the probability distribution for an anisotropy amplitude given a true distribtuion is the same as that for the true distribution given an observation. Thus Figure 1 does double duty.

Calculations similar to those described above can be made for the propagation of errors in the determination of A and Φ. The standard deviation in A is $\sigma/\sqrt{\Sigma}$. The standard deviation in Φ has been determined in the limit of large r :

$$\sigma_\Phi = \sigma_\delta / \delta$$

Figure 1

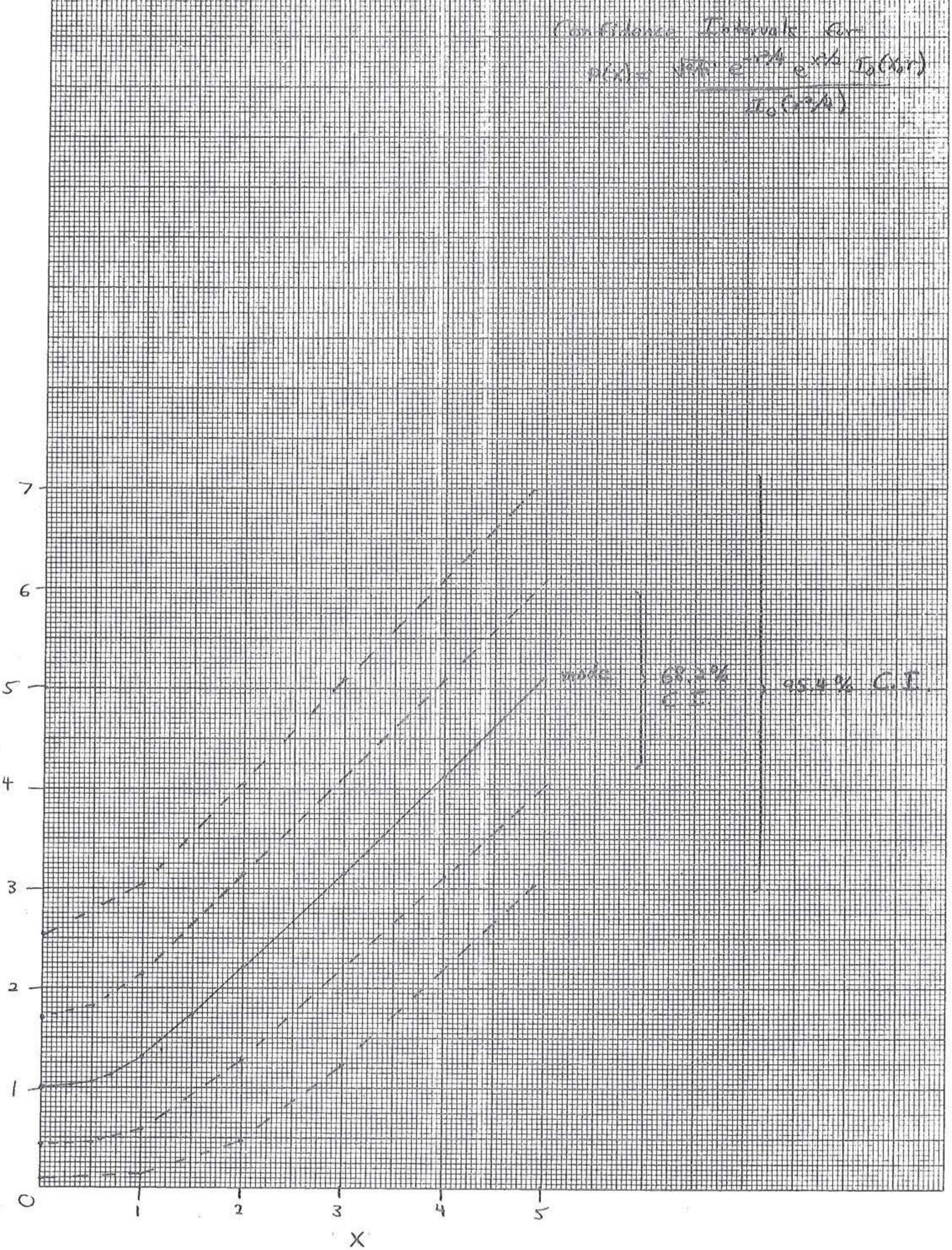
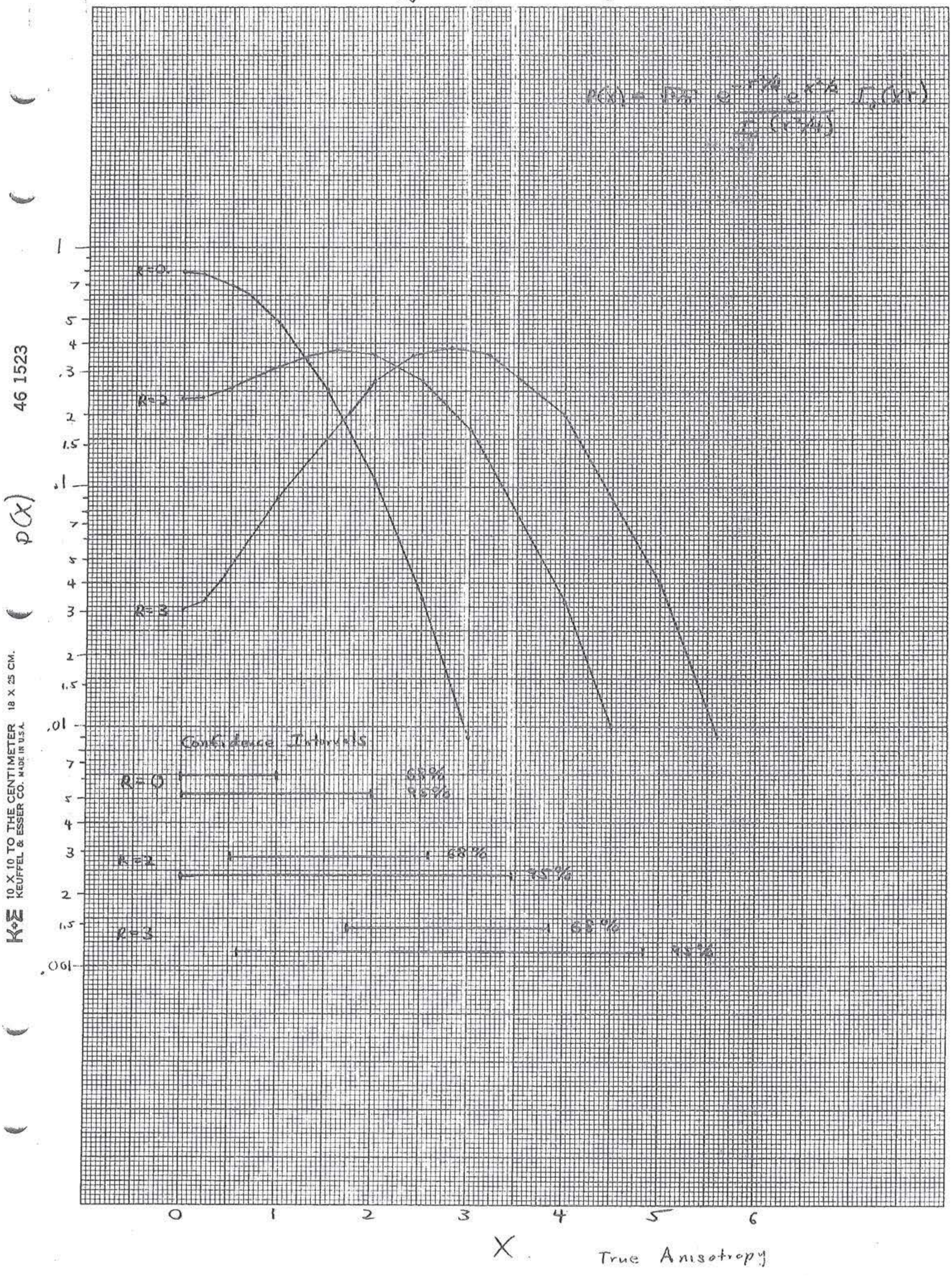

Confidence Intervals for
$p(x) = \frac{\sqrt{2/\pi}\, e^{-r^2/4}\, e^{-x^2/2}\, I_0(x r)}{I_0(r^2/4)}$
mode
68.2% C.I.
95.4% C.I.
r
x
0
1
2
3
4
5
6
7

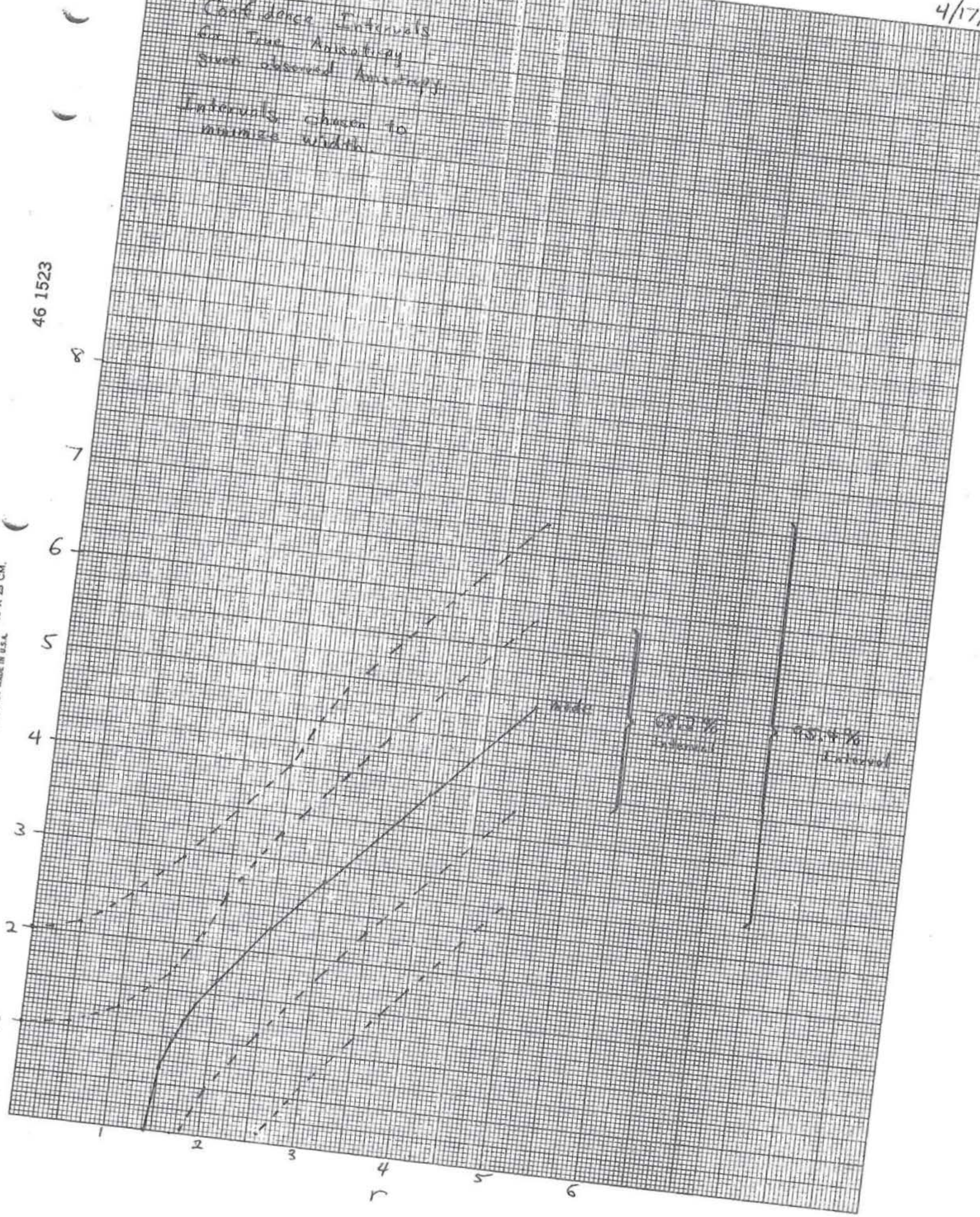
Figure 2



R=0
R=2
R=3
p(X)
Confidence Intervals
R=0
68%
95%
R=2
68%
95%
R=3
68%
95%
0
1
2
3
4
5
6
X
True Anisotropy
46 1523
10 X 10 TO THE CENTIMETER 18 X 25 CM.
KEUFFEL & ESSER CO. MADE IN U.S.A.

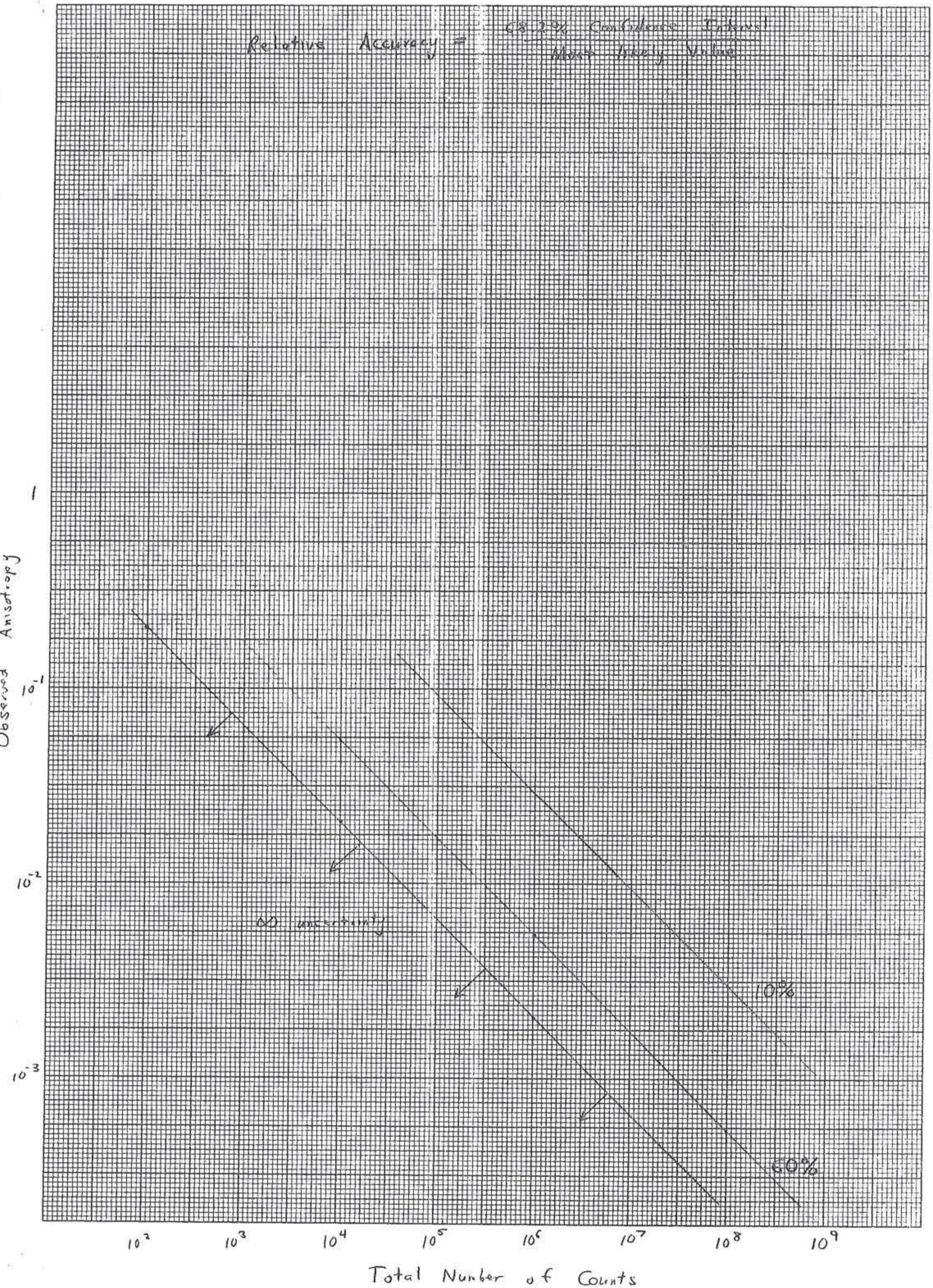
Figure 3

4/17/



Confidence Intervals
for True Anisotropy
given observed Anisotropy.
Intervals chosen to
minimize width.
8
7
6
5
4
3
2
1
1
2
3
4
5
6
r
Mode
68.2 %
Interval
95.4 %
Interval

Figure .


Relative Accuracy = 68.2% Confidence Interval / Most likely Value
1
10^{-1}
10^{-2}
10^{-3}
Observed Anisotropy
∞ uncertainty
10%
60%
10^2 10^3 10^4 10^5 10^6 10^7 10^8 10^9
Total Number of Counts

Anisotropy Calculations

After Zwickl[1]

Do general cosine expansion for 1 harmonic

$$\psi(\theta_i) = A + A_1 \cos(\theta_i - \phi_1)$$
$$= A + A_1 \cos\phi \cos\theta_i + A_1 \sin\phi_1 \sin\theta_i$$

Define
$$\alpha = A_1 \cos\Phi$$
$$\beta = A_1 \sin\Phi_1$$

So $A_1 = \sqrt{\alpha^2+\beta^2}$ $\Phi = \tan^{-1}(\beta/\alpha)$

Each sector observes Y_i counts
Assume each Y_i has σ_i associated with it.

Assume $(\forall i)$ $\sigma_i = \sigma$

Now minimize least squares

$$\mu^2 = \Sigma[Y_i - \psi(\theta_i)]^2$$

Easy to show that (see Appendix 2)

$A = \frac{1}{S}\Sigma Y_i$ where S is number of sectors

$\alpha = \frac{2}{S}\Sigma Y_i \cos\theta_i$ $\beta = \frac{2}{S}\Sigma Y_i \sin\theta_i$

From these

Anisotropy Amp $\xi = W A_1 / A$

Phase $\Phi = \tan^{-1}(\beta/\alpha)$

where $W = \frac{\pi}{S} / \sin(\frac{\pi}{S})$ is smoothing factor[2]

Would now like to determine an error estimate of these parameters

Standard Techinque[3] to find σ on a function of many varables

$$y = f(x_1, x_2, \dots)$$

is $\sigma_y^2 = \Sigma\, \sigma_{x_i}^2 \left(\frac{\partial y}{\partial x_i}\right)^2$

However this only works if X_i are uncorrellated and that the most probable value for y is given by

$$\bar{y} = f(\bar{X}_1, \bar{X}_2, \dots)$$

and the mean of a sample of y's tends to $\bar{y}$.

This technique will not always work for calculating σ_ξ since for a true anisotropy of 0

$$\bar{\xi}(\bar{Y}_i) = 0 \quad \text{but} \quad \bar{\xi} > 0 .$$

However σ_α and σ_β can be calculated using above technique.

Appendix 3 shows

$$\sigma_\alpha = \sigma_\beta = \sigma\sqrt{2/S}$$

$$\sigma_A = \sigma/\sqrt{S}$$

Assume we are sampling a distribution of form

$$\psi = A_0 + \delta_0 \cos\Theta$$

Let the number of counts in a sector for this distribution be $N_i = A_0 + \frac{\delta_0}{w}\cos\Theta_i$

A sample of this distribution yields n_i counts for sector i with uncertainty σ_i

Define

$$y_i \equiv n_i - N_i$$

so $\sigma_{y_i} = \sigma_i$ and $\langle y_i \rangle = 0$

Make approx $(\forall i)\ \sigma_i = \sigma$

Compute α and β for n's

$$\alpha_n = \tfrac{2}{5}\sum n_i \cos\Theta_i = \tfrac{2}{5}\sum (y_i \cos\Theta_i + N_i \cos\Theta_i)$$

$$= \alpha_y + \frac{\delta_0}{w}$$

$$\equiv \alpha_y + \alpha_0$$

similarly

$$\beta_n = \tfrac{2}{5}\sum y_i \sin\Theta_i + \tfrac{2}{5}N_i \sin\Theta_i$$

$$= \beta_y + 0$$

$$= \beta_y$$

Since n_i are isotropic

$$\langle \alpha_n \rangle = \alpha_0 \qquad \sigma_{\alpha_n} = \sigma\sqrt{2/5} = \sigma_{\beta_n}$$

$$\langle \beta_n \rangle = 0$$

Also $\langle A_n \rangle = A_0 \qquad \sigma_{A_n} = \sigma/\sqrt{5}$

So

$$P(\alpha_n) \cong \frac{1}{\sqrt{2\pi}\,\sigma_{\alpha_n}} e^{-(\alpha_n-\alpha_0)^2/2\sigma_{\alpha_n}^2}$$

similarly for β_n, A_n and since A_n, α_n, and β_n are indepentent linear combinations of gaussianly distributed variables[6]

$$P(\alpha_n, \beta_n, A_n) \cong \frac{1}{2\pi\,\sigma_{\alpha_n}^2} e^{-\left[\beta_n^2+\alpha_n^2-2\alpha_n\alpha_0+\alpha_0^2\right]/2\sigma_{\alpha_n}^2}$$

$$\times \frac{1}{\sqrt{2\pi}\,\sigma_{A_n}} e^{-(A_n-A_0)^2/2\sigma_{A_n}^2}$$

now define

$$x = \alpha_n/\sigma_{\alpha_n} \qquad y = \beta_n/\sigma_{\alpha_n} \qquad x_0 = \alpha_0/\sigma_{\alpha_n}$$

$$P(x, y, A_n) = \frac{1}{2\pi} e^{-(x^2+y^2+x_0^2-2xx_0)/2} \quad P(A_n)$$

define $r = \sqrt{x^2+y^2}$, $\theta = \tan^{-1}(y/x)$

$$P(r, \theta, A_n) = \frac{r}{2\pi} e^{-(r^2-2rx_0\cos\theta+x_0^2)/2} \quad P(A_n)$$

$$P(r, A_n) = \frac{r}{2\pi} e^{-x_0^2/2\,-r^2/2} \int_0^{2\pi} e^{+rx_0\cos\theta}\,d\theta \quad P(A_n)$$

$$P(r, A_n) = r\,e^{-x_0^2/2\,-r^2/2}\, I_0(x_0 r) \quad P(A_n) \qquad (4)$$

where I_0 is modified Bessel function

This is probability that a sample distribution can be characterized by r, A_n given that true distribution has characteristic x_0, A_0

That is, this is

$$P(r, A_n \mid x_0, A_0)$$

From Bayes' Theorem

$$\frac{P(X_{01}, A_{01} \mid r, A_n)}{P(X_{01}, A_{02} \mid r, A_n)} = \frac{P(X_{01}, A_{01})\, P(r, A_n \mid X_{01}, A_{01})}{P(X_{02}, A_{01})\, P(r, A_n \mid X_{01}, A_{01})}$$

Assuming all A_0's + δ_0's, equally likely
then all X_0's are qually likely

and

$$P(X_0, A_0) \equiv P(X_0, A_0 \mid r, A_n) = \frac{e^{-X_0^2/2}\, I_0(X_0 r)}{\int_0^\infty e^{-X_0^2/2}\, I_0(X_0 r)\, dX_0} \quad \frac{e^{-(A_0 - A_n)^2/2\sigma_{A_n}^2}}{\sqrt{2\pi}\,\sigma_{A_n}}$$

-5-

$$P(X_0, A_0) = \frac{\sqrt{2/\pi}\, e^{-r^2/4}}{I_0(r^2/4)}\, e^{-X_0^2/2}\, I_0(X_0 r) \quad P(A_0)$$

$$= P(X_0)\, P(A_0)$$

The characteristics of $P(X_0)$ are given in Figs 1 and 2

We will now relate r + X_0 to δ_{obs} and δ_{true}

$$X_0 = \alpha_0 / \sigma_{\alpha n} = \frac{\delta_0}{w} \Big/ \sigma\sqrt{2/s}$$

$$X_0 = \frac{\delta_0 \sqrt{s/2}}{w\,\sigma}$$

$$r = \sqrt{x^2 + y^2} = \sigma_{\alpha n}^{-1} \sqrt{\alpha_n^2 + \beta_n^2} = \frac{(A', \delta'/w)}{\sigma'} \sqrt{s/2}$$

where the ' indicates observed values

True anisotropy is $\xi_0 = \frac{\delta_0}{A_0}$

$$\xi_0 = \frac{X_0}{A_0} \frac{W\sigma}{\sqrt{S/2}}$$

$$P(\xi_0) = \int P\left(X_0 = \frac{\xi_0 A_0 \sqrt{S/2}}{W\sigma}\right) P(A_0)\, dA_0$$

$$\approx P\left(X_0 = \frac{\xi_0' A' \sqrt{S/2}}{W\sigma'}\right) \int P(A_0)\, dA_0$$

$$= P\left(X_0 = \frac{\xi_0' A' \sqrt{S/2}}{W\sigma'}\right)$$

have approximate $P(A_0)$ as delta function at A_n

The scale size for variations in $P(A_0) \sim \sigma_{A_0} = \sigma/\sqrt{S}$
scale size for variations in $P(X_0) \sim 1$ see Fig 1

A change in A_0 of $\sigma/\sqrt{S}$
leads to change in X_0 of $\frac{\xi_0 \sqrt{S/2}}{W\sigma} \frac{\sigma}{\sqrt{S}} = \frac{\xi_0}{W\sqrt{2}}$

$$\frac{\Delta X_0}{\lambda_{X_0}} = \frac{\xi_0 / W\sqrt{2}}{1} = \frac{\xi_0}{W\sqrt{2}}$$

So approximation valid for $\xi_0 << 1$

Thus to find $P(\xi_0)$ in this approximation

Define
$$Z = \frac{W\sigma'}{A'\sqrt{s/2}}$$

Now
$$r = \xi'/Z$$

and
$$P(\xi_0) = P\left(x_0 = \frac{\xi_0}{Z}\right)$$ using Fig 1 to characterize distribution

3/7/75

Average X_0

$$P(X_0) = \frac{\sqrt{2/\pi}}{I_0(r^2/4)} e^{-r^2/4} e^{-X_0^2/2} I_0(X_0 r)$$

$$\overline{X_0} = \int_0^\infty X_0 \, P(X_0) \, dX_0$$

$$= \frac{\sqrt{2/\pi}}{I_0(r^2/4)} e^{-r^2/4} e^{r^2/2}$$ G&R p717

$$= \frac{\sqrt{2/\pi} \; e^{r^2/4}}{I_0(r^2/4)}$$

$$\overline{X_0}_{\lim_{r\to\infty}} = r\left(1 - \frac{1}{2r^2}\right)$$

$$\overline{X_0}_{\lim_{r\to 0}} = \sqrt{\frac{2}{\pi}}\left(1 - \frac{r^2}{4}\right)$$

Remains uncertainty in ϕ

$$\phi = \tan^{-1}(\beta/\alpha)$$

have some problems about wrap around, $\tan^{-1}(0/0)$, etc

have $P(r,\theta) = \frac{r}{2\pi} e^{-(r^2+X_0^2-2rX_0\cos\theta)/2}$

$$P(\theta) = \frac{e^{-X_0^2/2}}{2\pi}\int_0^\infty r e^{-r^2/2 + r X_0\cos\theta}\, dr$$ $\mu = \frac{1}{2}$ $\nu = -X_0\frac{\cos\theta}{2}$

$$P(\theta) = \frac{e^{-X_0^2/2}}{2\pi}\left[1 + \frac{X_0\cos\theta}{2}\sqrt{2\pi}\, e^{X_0^2\cos^2\theta/2}\left[1+\Phi\left(+\frac{X_0\cos\theta}{\sqrt{2}}\right)\right]\right]$$ G+R p338

5.

Φ is probability integral

$$\Phi(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt \equiv erf(x) = -erf(-x)$$

as $X_0 \to 0$

$$P(\theta) = \frac{1}{2\pi}\left[1 + X_0\sqrt{\tfrac{2}{\pi}}\cos\theta\right]$$ so goes to constant centered @ $\theta = 0$

$X_0 \to \infty$

$$P(\theta) = \frac{e^{-X_0^2/2}}{2\pi}\left[\cancel{1} + X_0\cos\theta\sqrt{2\pi}\, e^{X_0^2\cos^2\theta/2}\right]$$

$$= \frac{X_0\cos\theta}{\sqrt{2\pi}} e^{-X_0^2(1-\cos^2\theta)/2}$$

$$P(\theta) = \frac{X_0\cos\theta}{\sqrt{2\pi}} e^{-\sin^2\theta X_0^2/2}$$

large X_0 well be centered near $\theta = 0$

$$= \frac{X_0(1-\cancel{\theta^2/2})}{\sqrt{2\pi}} e^{-\theta^2X_0^2/2}$$

so $\sigma_\theta = \frac{1}{X_0} = \frac{\sigma_n}{\alpha_0} = \frac{w\sigma\sqrt{2/s}}{\xi_0 A_0}$ ✓

$$= \frac{\sigma_{\xi_0}}{\xi}$$

Appendix 2.

Least Squares Fit

$$\Psi_i = A_0 + \alpha \cos\Theta_i + \beta \sin\Theta_i$$

$$\mu^2 = \sum\left[-Y_i + \Psi(\Theta_i)\right]^2$$
$$= \sum\left(-Y_i + A_0 + \alpha\cos\Theta_i + \beta\sin\Theta_i\right)^2$$

$$0 = \frac{\partial \mu^2}{\partial A_0} = +2\sum\left(-Y_i + A_0 + \alpha\cos\Theta_i + \beta\sin\Theta_i\right)$$
$$= -2\sum Y_i + 2SA_0 + 2\alpha\sum\cos\Theta_i + 2\beta\sum\overset{\sin}{\cancel{\cos}}\,\Theta_i$$
$$= -2\sum Y_i + 2SA_0$$

or
$$A_0 = \frac{1}{S}\sum Y_i$$

$$0 = \frac{\partial \mu^2}{\partial \alpha} = \sum 2\cos\Theta_i\left(-Y_i + A_0 + \alpha\cos\Theta_i + \beta\sin\Theta_i\right)$$
$$= -2\sum Y_i\cos\Theta_i + 2\alpha\sum\cos^2\Theta_i$$

or
$$\alpha = \frac{2}{S}\sum Y_i\cos\Theta_i$$

similarily

$$\beta = \frac{2}{S}\sum Y_i\sin\Theta_i$$

Appendix 3

$$\alpha = \frac{2}{S} \sum Y_i \cos\Theta_i$$

$$\sigma_\alpha^2 = \sum \sigma_i^2 \left(\frac{\partial \alpha}{\partial Y_i}\right)^2$$

$$= \sigma^2 \sum \frac{4}{S^2} \cos^2\Theta_i$$

$$= \frac{4\sigma^2}{S^2} \quad \frac{S}{2}$$

$$= 2\sigma^2/S$$

or

$$\sigma_\alpha = \sigma\sqrt{2/S}$$

σ_β works same way.

$$A_0 = \frac{1}{S} \sum Y_i$$

$$\sigma_{A_0}^2 = \sum \sigma_i^2 \left(\frac{\partial A_0}{\partial Y_i}\right)^2 = \sum \sigma^2 \frac{1}{S^2}$$

$$= \sigma^2/S$$

$$\sigma_{A_0} = \sigma/\sqrt{S}$$

Footnotes

1 Zwickl & Webber New Hampshire Conf

2 Chapman & Barkels *Geomagnetism* Vol II

3 Bevington, P. *Data Reduction and Error Analysis for the Physical Sciences*

4 Abramowitz, M. and Stegun, I *Handbook of Mathematical Functions*

5 Gradshteyn and Ryzhik p 710 6.618 1.

6 Cramér, H. *Mathematical Methods of Statistics* p 319