Vistas in Astronomy, Vol.29, pp.27-51, 1986
Printed in Great Britain.

0083-6656/86 $0.00 + .50

STATISTICAL METHODS OF STELLAR POLARIMETRY

D. Clarke* and B. G. Stewart**

*University Observatory, Acre Road/Maryhill Road, Glasgow G20 0TL, U.K.
**School of Electronic and Electrical Engineering, Robert Gordon's Institute of Technology, Aberdeen AB9 1FR, U.K.

1. INTRODUCTION

In principle, optical polarimetry can provide direct information about the geometries of stellar envelopes and binary systems, about stellar magnetic fields, about the properties of scattering particles (electrons, dust) within the stellar environments and about the dust in the interstellar medium. Generally stellar polarizations have low values and are consequently determined in experimental situations of low signal-to-noise.

Regular questions raised by the investigator in relation to discussions of the underlying astrophysics frequently include:

(i) Does the star exhibit polarization?
(ii) Has the polarization changed since the previous observation?
(iii) Is there a polarization difference for measurements made in two spectral intervals or do spectral measurements follow a particular predicted curve?

These points can only be tackled accurately with quantitative assessments if the characteristics of the various experimental noises are known and if the influences of these noises on the determined polarization parameters are fully understood.

Some twenty-five years ago, Serkowski (1958, 1962) demonstrated that noise introduces a bias to determinations of the degree of polarization, p, but this fact is still frequently unheeded in the current literature presenting new results. Indeed most papers treat the assessment of polarimetric uncertainties (errors) in naïve fashion, automatically assuming that repeated measurements follow normal distributions; they also apply standard formulae related to large sample statistics to situations in which only small samples are available. In some cases new data are incorrectly reduced and/or are presented with uncertainties or error bars which are incorrect (examples are given in the discussion below), this in turn possibly falsifying some of the claims that are made in relation to the underlying astrophysics. In other papers observational routines and reduction procedures are skimmed over and the discerning reader is left wondering if the overall experimental scheme was sufficiently reasonable to justify the claimed interpretations. By pointing to some of the shortcomings of previous work with cited examples, the aim is not to condemn but to advance the art of experimental polarimetry.

As stellar polarimetry improves in accuracy, with lower levels of polarization becoming detectable, it is important that the influences of the various noises are understood so that the investigated astrophysics can be discussed sensibly. Care must be taken that lack of appreciation of the effects of noise and misunderstanding of polarimetric statistics do not steer the interpretations to misleading conclusions. In this paper the statistics associated with polarization measurements will be presented in a deeper and more unified way than has been done previously, some of the material being presented for the first time.

Following the presentation in the next Section of the various definitions describing polarization, the statistical behaviour of the basic Normalized Stokes Parameters is discussed in Section 3. It is demonstrated that for the fainter stars, providing small photon counts per integration, and for the brighter stars whose measurement might be affected by scintillation noise, their repeated measurement may provide a non-normal (non-Gaussian) statistical behaviour. Attention is also drawn to the sensitivity of the determined values to background noise.

Section 4 deals with the statistical behaviour of p (the degree of polarization) and discusses the choice of estimator which is best for removing the effect of bias; the procedures for determining confidence intervals are also outlined. Section 5 discusses the determination of the position angle, ϕ, of polarization and the associated confidence intervals. An approach to the determination of confidence intervals for p when measurements are limited to small samples is presented in Section 6; tests for deciding whether polarization is present or not are also given. Section 7 presents the Welch test for comparing sets of data with different sample sizes and different variances. This test has general application but it is particularly useful in the assessment of polarimetric data.

Finally a scheme is summarised for the statistical investigation of any polarimetric data to check its validity. Such approaches are important in the establishment of new and more accurate polarimetric standard stars and in any experiments which advance stellar polarimetry to new levels of precision.

2. THE POLARIZATION PARAMETERS

The polarization state of a beam of light is generally expressed in the form of the familiar Stokes vector with the four elements {I, Q, U, V} having the dimensions of intensity (see, for example, Clarke and Grainger, 1971). The first Stokes parameter, I, represents the total intensity of the beam, the second and third parameters, Q and U, record the linear polarization intensities while the fourth parameter, V, describes the circular polarization intensity. If a non-zero value of V is simultaneously present with a non-zero value for either or both Q and U, the light is elliptically polarized. The value of Q corresponds to the intensity difference between orthogonal vibrations whose azimuths form the reference co-ordinate frame for the vector; for stellar polarimetry it is generally arranged that a positive Q value corresponds to the vibration azimuth being along the celestial north-south direction, while a negative Q value relates to the azimuth being aligned with the east-west direction. The value of U corresponds to the intensity difference between orthogonal vibrations whose azimuths are rotated by 45°, north through east, relative to the positive Q direction. The value of V is the intensity difference between the right-handed and left-handed circularly polarized components. In terms of the resolved electric field strengths, E, and their phases, δ, the Stokes vector is given by:

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} \langle E_N^2 + E_E^2 \rangle \\ \langle E_N^2 - E_E^2 \rangle \\ \langle 2E_N E_E \cos(\delta_E - \delta_N) \rangle \\ \langle 2E_N E_E \sin(\delta_E - \delta_N) \rangle \end{bmatrix} \tag{1}$$

where the subscripts N, E refer to the components with north/south and east/west azimuths respectively. The angular brackets denote that the enclosed quantities are expectation values with averaging performed over the frequency (wavelength) passband of the measuring equipment and over the experimental time used to make the measurements.

It is standard practice to normalize the parameters by dividing by the intensity so that we may write:

$$q = \frac{Q}{I}; \quad u = \frac{U}{I}; \quad v = \frac{V}{I} \qquad * \quad (2)$$

The degree of linear polarization, p, may now be defined as:

$$p = (q^2 + u^2)^{\frac{1}{2}} \qquad (3)$$

with the degree of circular polarization corresponding to v.

An angle θ may be defined by:

$$\theta = \arctan\left(\frac{U}{Q}\right) \quad \text{or} \quad \arctan\left(\frac{u}{q}\right) \qquad (4)$$

from which the position angle, ϕ, of the vibration of the linear polarization may be determined, i.e. $\phi = \theta/2$.

A less frequently used term is the degree of elliptical polarization, e, being defined as:

$$e = (q^2 + u^2 + v^2)^{\frac{1}{2}} \quad . \qquad (5)$$

3. THE STATISTICS OF NORMALIZED STOKES PARAMETERS

3.1 Effects of Photon Shot Noise

Polarimetric modulators take a variety of forms, the most favoured techniques employing either rotating waveplates or variable retardation waveplates prior to a fixed polarizer. The methods of recording the modulated intensity from the photodetector signal also vary from one instrument to another and, as a consequence, the most direct analytical treatment given below on the effects of noise on the determined Normalized Stokes Parameters (NSPs) may need to be modified to suit particular circumstances.

In essence, the parameters are determined from pairs of accumulated photon counts corresponding to orthogonal polarization modes with the difference being divided by the sum. Thus the first NSP may be written as:

$$q = \frac{n_1 - n_2}{n_1 + n_2} \qquad (6)$$

where n_1 is the number of photons recorded for the intensity associated with the direction of vibration parallel to the chosen reference axis and n_2 is the number of photons recorded for the orthogonal direction. Similar expressions also apply to u and v. Several instrumental designs provide signal outputs with records which can be inserted directly into equations of the form given by (6). Others require preliminary reduction and manipulation and the effects of these procedures may need to be taken into account in the assessment of the distribution of repeated determinations of the NSP.

In the determination of the uncertainties or confidence intervals for a set of repeated measurements it is generally assumed that the distribution is normal. However it has been shown by Clarke *et al.* (1983) that even with underlying statistically independent normal distributions of n_1 and n_2, the general probability distribution of q takes the form:

---

* The symbols used in the literature to represent Stokes parameters are various (see Clarke, 1974) but a preference has emerged in stellar polarimetry for I, Q, U, V. For the normalized parameters, the symbol q is also used to relate the degree of circular polarization rather than v in papers dealing solely with circular polarization. The definition of the sense of v in terms of being positive, negative, right-handed, left-handed, etc., has several alternatives - all of which are used in the literature.

$$P(q) = \frac{B \exp\left(\frac{B^2}{A} - C\right)}{(\pi A)^{\frac{1}{2}} A\sigma_1\sigma_2 (1 + q)^2} \tag{7}$$

where $$A = \frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\left(\frac{1-q}{1+q}\right)^2\right)$$

$$B = \frac{1}{2}\left(\frac{z_1}{\sigma_1^2} + \frac{z_2}{\sigma_2^2}\left(\frac{1-q}{1+q}\right)\right)$$

and $$C = \frac{1}{2}\left(\frac{z_1^2}{\sigma_1^2} + \frac{z_2^2}{\sigma_2^2}\right)$$

with z_1, z_2 being the true values of the signals and σ_1^2, σ_2^2 being the variances of the distributions of n_1, n_2 respectively. It has also been demonstrated that as the signal-to-noise of the experiment increases, Equation (7) approaches a Gaussian form but for moderate and low values (i.e. z_1/σ_1, z_2/σ_2 less than 50) the effects of kurtosis in the P(q) distribution need to be taken into account in the calculation of exact confidence intervals. Thus even under the best working conditions with the records of n_1 and n_2 exhibiting only the noise of photon counting statistics which can be considered as normal rather than Poissonian for z_1, z_2 both > 100, the probability distribution of q is non-normal.

The severity of the kurtosis as a function of the underlying true photon signal has been explored (Clarke, *et al.* - *loc. cit.*) by forming the ratio, R(q) of Equation (7) and a normal curve for P(q) given by

$$P_n(q) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left\{\frac{-(q - q_o)^2}{2\sigma_n}\right\} \tag{8}$$

where q_o is the true normalized Stokes parameter and σ_n^2 the formal variance based on photon shot noise and given by

$$\sigma_n = \left\{\frac{(1 - q_o)^2}{z}\right\}^{\frac{1}{2}} \quad * \tag{9}$$

with the total signal, z equal to $z_1 + z_2$. Figure 1 displays R(q) for the 68.2% (1σ), 95% (1.96σ) and 99% (2.575σ) confidence values as a function of z under the most favourable conditions of pure photon shot noise (i.e. $\sigma_1 = \sqrt{z_1}$ and $\sigma_2 = \sqrt{z_2}$). Although the central part of the P(q) curve as defined by Equation (7) is close to that of a normal distribution (Equation (8)), the effect of the tails of the distribution is strongly evident, particularly for z < 1000.

The form of P(q) can readily be confirmed by computer simulation studies and examples of determined values of R(q) generated by such an exercise are also displayed in Fig. 1. A value of q_o equal to zero was chosen. The program produced 25,000 values of n_1, n_2 from normal distributions with $\bar{n}_1 = z_1$ and $\bar{n}_2 = z_2$ and with variances given by $\sigma_1^2 = z_1$ and $\sigma_2^2 = z_2$; these simulated photon counts provided values of NSPs which were ordered into a histogram with a bin resolution of $\sigma_n/200$. From the histogram the probabilities associated with given NSP values corresponding to the 1σ, 1.96σ (95%) and 2.575σ (99%) points were determined so allowing R(q) values to be derived for these chosen confidence intervals. Repetitive runs of the program allowed the uncertainties of the simulations to be assessed.

---

* In various works by Serkowski (e.g. 1974a, 1974b) which discuss errors associated with NSPs at the photon counting limit, the quoted value with $q_o = 0$ is given as $\sigma_n = (2/z)^{\frac{1}{2}}$, a value which is $\sqrt{2}$ times larger than Equation (9). However this numerical difference merely comes from the definition of z. In this work z is taken as the total photon count associated with the determination of a single NSP while Serkowski defines z as the total photon count accumulated in the determination of both NSPs q and u.

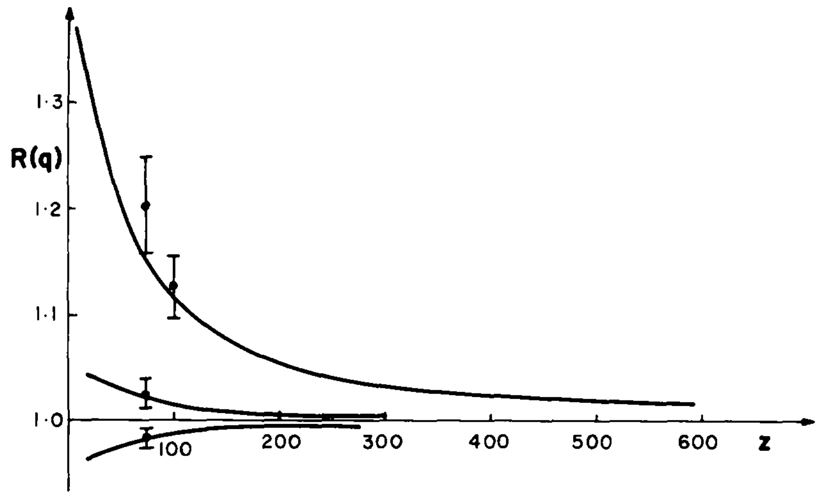


Fig. 1. The effect of positive kurtosis of the P(q) distribution for measurements limited by photon shot noise is indicated as a function of the total photon count z by displaying the behaviour of the distribution normalized to the perfect normal curve for values of (x) corresponding to 68.2% (1σ), 95% (1.96σ) and 99% (2.575σ) confidence levels. The veracity of the analytical expression for P(q) is confirmed by points which have been generated by computer simulation.

An alternative program allowed the values of the NSPs to be assessed at various percentage confidence intervals and those for 99% are compared in Table 1 with similar intervals for the normal curve based on the formal variance given by Equation (9). It can be seen that the real distribution is a few percent broader than the formal normal curve, the difference being larger at the lower z values.

TABLE 1

Normalized Stokes Parameters predicted by the normal curve (Equation (9)) are compared with real values of the 99% confidence level as determined by computer simulation for various photon counts with an underlying NSP value of zero.

z	Normal	Simulation
100	± 0.2575	± 0.2664
200	± 0.1821	± 0.1852
300	± 0.1486	± 0.1503
400	± 0.1287	± 0.1298
500	± 0.1152	± 0.1160

In the assembly of repeated measurements of NSPs, Clarke *et al.* (1983) have commented on the variety of numerical procedures that have been adopted. A fairly standard method involves collecting repeated values of n_1 and n_2 from which NSPs are calculated and a mean formed. This reduction scheme may be expressed as

$$\tilde{q} = \frac{1}{N}\sum_{i=1}^{N} q_i = \frac{1}{N}\sum_{i=1}^{N} \frac{n_{1i} - n_{2i}}{n_{1i} + n_{2i}} \tag{10}$$

from which the statistical behaviour of q_i may be investigated. For a given experimental time, the signals may be sampled frequently providing a large number of q_i values, each determined at low signal-to-noise, or the frequency may be reduced to provide a smaller sample of q_i values at a higher signal-to-noise. In the extreme, only one value might ensue, this being represented by

$$\bar{q} = \frac{\sum_{i=1}^{N} n_{1i} - \sum_{i=1}^{N} n_{2i}}{\sum_{i=1}^{N} n_{1i} + \sum_{i=1}^{N} n_{2i}} \quad . \tag{11}$$

Providing that the q_i values follow a normal distribution, the uncertainties associated with $\tilde{q}$ reduce according to $\sqrt{N}$ and in the limit the values of $\tilde{q}$ and $\bar{q}$ are identical and carry the same error. However, Fig. 1 indicates that this would only hold for values of z greater than a few thousand, this same number providing a critical minimum photon count for each NSP sample to achieve the best final result. Although for good statistical treatment a large number of samples is preferred, care should be taken not to oversample the signal so that NSP values are accumulated with too low a mean value of z. It may be mentioned here that the majority of data recording techniques are satisfactory in this regard but it may also be noted that most of the presented results are means from a small number of measurements and the quoted errors are underestimated, being based on the classical large sample calculation rather than the more appropriate Student t-distribution.

3.2 Effects of Non-normal Noises - Scintillation Noise

Comments have also been made by Clarke *et al.* (1983) on the behaviour of the P(q) distribution under conditions where perfect photon counting statistics do not pertain. For example, if the variances for the two photon counts are represented by $\sigma_1^2 = Kn_1$ and $\sigma_2^2 = Kn_2$ as is suggested by James and Sternberg (1969), then the kurtosis as might be depicted by the function, R(q), is influenced by the value of K. In the more general case where $\sigma_1^2 = K_1n_2$ and $\sigma_2^2 = K_2n_2$ with $K_1 \neq K_2$, then the P(q) distribution becomes skew and repeated measurements provide a mean value which in the limit would not equal q_o. In other words, a biased result would ensue.

According to the noise characteristics, the mean value, $\tilde{q}$, of the P(q) distribution can be obtained by numerically integrating Equation (7) and by differentiating this same expression, the most probable value, $\tilde{q}_{MP}$, can be determined. As an example of the biasing effect, Table 2 lists results for the case of q_o equal to zero, so that $z_1 = z_2$, but with $\sigma_1 = \sqrt{z_1}$ and $\sigma_2 = \sqrt{Kz_2}$. It can be seen that for photon count accumulations at low signal-to-noise, the biasing can be severe.

TABLE 2

The mean and most probable values of the NSP are given for an underlying value of q = 0 showing that biasing occurs when the variances of the component measurements are not equal, the effect being strongest under the lower signal-to-noise conditions.

$z_1 = z_2$	σ_1	σ_2	$\tilde{q}$	$\tilde{q}_{MP}$
50	7.0711	8	0.00155	-0.00290
		9	0.00333	-0.00615
100	10	11	0.00053	-0.00103
		12	0.00112	-0.00216
500	22.3607	23	0.00005	-0.00009
		24	0.00009	-0.00018

A similar situation arises if the noise associated with the photon counts can be represented as a constant fraction of the signal strengths (e.g. $\sigma_1 = \kappa n_1$ and $\sigma_2 = \kappa n_2$) as is the case for certain equipment generated noises or, to a first order, as might be considered to represent noise introduced by the earth's atmosphere (i.e.

transparency fluctuations, intensity scintillation). Analysis of the form of R(q) shows that the asymmetry of the P(q) distribution depends both on κ and q_o. If scintillation noise is present in the recorded data, as it may well be for bright stars measured with a mechanically rotating modulator (e.g. see Tinbergen, 1982), then the situation is made worse by the fact that repeated intensity measurements follow a log normal distribution (Young, 1974) which enhances the asymmetry of the P(q) distribution.

As can be seen from the above discussion, it is prudent to know the capabilities of any polarimeter in terms of the statistical behaviour of repeated measurements of NSPs over the wide range of experimental circumstances. The first undertaking is to check the normality of the results by calculating the skewness and kurtosis (see Wall, 1979) and comparing these determined parameters with the spread of values which is acceptable according to the number of NSPs in the sample. Useful tables which allow data to be tested for normality at the 95% and 99% confidence levels have recently been assembled by Brooks (1984). If normal statistics are deemed to apply, although the effects of systematic errors cannot be ruled out with certainty, the assignment of confidence intervals to the measurements is that much easier.

3.3 Effects of Background Subtraction

Although reports on observational schemes usually mention procedures for the removal of the sky and detector background signals, no reference is generally given to the estimation of the noise that is generated when the subtraction is performed. Indeed most experimental routines appear to be lax in providing guide lines as to how best to share out experimental time between recording the star signal with background and the background alone. Although measurements made by instruments with optical designs which allow direct compensation for sky background are more efficient in their use of telescope time, they still carry noise associated with the background level and the noise from the detector dark background is always present. However, no matter what procedure is employed for removing the effects of the background, it is invariably impossible to discern from the presented papers if the errors quoted for the corrected NSPs have allowed for the uncertainty of the background measurements.

The simplest procedure for removing the background is to record photon counts over several integrations with the telescope offset from the program star and to take means of the repeated measurements (M) so that

$$\bar{n}_{1B} = \frac{1}{M}\sum_{i=1}^{M} n_{1Bi} \quad \text{and} \quad \bar{n}_{2B} = \frac{1}{M}\sum_{i=1}^{M} n_{2Bi} \quad .$$

Subsequent measurements of the star are then adjusted by subtracting these mean background counts. However, the measurements carry noise and the experimentally determined value of the mean background may be offset from the true value. Its subsequent subtraction affects the determinations of the NSPs in a systematic fashion. If a sufficient number of integrations are made of the star, eventually the systematic error of the imprecise background subtraction will show through in the final result and the situation can only be improved by obtaining better estimates of the background.

The problem may be formulated in the following manner. By considering photon shot noise to apply, the noises of the "star plus background" integrations are $\pm\,(\overline{n_{1*} + n_{1B}})^{\frac{1}{2}}$ and $\pm\,(\overline{n_{2*} + n_{2B}})^{\frac{1}{2}}$. If the mean background signals are expressed as a fraction, f, of the star signal, the noises may be written as $\pm\,\bar{n}_{1*}^{\frac{1}{2}}\,(1 + f_1)^{\frac{1}{2}}$ and $\bar{n}_{2*}^{\frac{1}{2}}\,(1 + f_2)^{\frac{1}{2}}$. (For a star with small polarization with $\bar{n}_{1*} = \bar{n}_{2*}$ a large difference between f_1 and f_2 would reflect the presence of a sky background with significant polarization).

On recording the background, the noise on each integration may be expressed as $\pm\,(f_1\,\bar{n}_{1*})^{\frac{1}{2}}$ and $\pm\,(f_2\,\bar{n}_{2*})^{\frac{1}{2}}$. After M integrations the noises associated with the background subtraction are

$$\pm\left(\frac{f_1\,\bar{n}_{1*}}{M}\right)^{\frac{1}{2}} \quad \text{and} \quad \pm\left(\frac{f_2\,\bar{n}_{2*}}{M}\right)^{\frac{1}{2}} \quad .$$

By setting a criterion such that the noise in each of the "star plus background" integrations is X times that introduced by the background subtraction process, we may write:

$$\left\{\bar{n}_{1*}\,(1 + f_1)\right\}^{\frac{1}{2}} = X_1\left\{\frac{f_1\,\bar{n}_{1*}}{M}\right\}^{\frac{1}{2}}$$

and

$$\left\{\bar{n}_{2*}\,(1 + f_2)\right\}^{\frac{1}{2}} = X_2\left\{\frac{f_2\,\bar{n}_{2*}}{M}\right\}^{\frac{1}{2}} \quad .$$

Hence

$$M = \frac{X_1^2\,f_1}{(1 + f_1)} = \frac{X_2^2\,f_2}{(1 + f_2)} \quad .$$

If the criterion is related to the N integrations of "star plus background", we may write:

$$\frac{M}{N} = \frac{X_1^2\,f_1}{(1 + f_1)} = \frac{X_2^2\,f_2}{(1 + f_2)} \quad . \tag{12}$$

The ratio $\left(\frac{M}{N}\right)$ represents the fraction of observational time required to be devoted to measurement of the background relative to that used for measuring the star in order to achieve the set criterion.

As a rule of thumb, the smallest acceptable working value for X is about 5, from which the experiment, in terms of dividing observational time between N and M can be organised according to the f values. Table 3 displays M/N ratios for various values of X and f.

TABLE 3

The fraction of the time to be allocated to measurement of the background is given according to its strength (f) relative to the star signal and to the constant (X) set to describe its influence on the calculated results.

f	X=5	X=10	X=25
0.0001	0.002	0.010	0.062
0.0005	0.012	0.050	0.312
0.001	0.025	0.100	0.624
0.005	0.124	0.497	3.109
0.010	0.247	0.990	6.183
0.05	1.190	4.762	29.762
0.1	2.273	9.091	56.818
0.2	4.167	16.667	104.167

However, no matter what the value, the background subtraction noise should always be calculated and combined with noise of the star measurements to provide the figure for the overall error of the NSP.

For the situation with an unpolarized background (e.g. the noise is primarily from the detector) and the source is only weakly polarized so that $f_1 \doteqdot f_2 = f$, it may be readily shown that Equation (9) expressing the error of an NSP is modified to

$$\sigma_q = \sigma_n\left(1 + f(1 + \frac{M}{N})\right)^{\frac{1}{2}} \quad . \tag{13}$$

4. THE STATISTICS OF THE PARAMETER, p

4.1 Introduction

As can be seen from Equation (3), the degree of polarization, p, is a positive definite quantity and the noise associated with the component NSPs always enhances its determined value, so providing a biased result. The effect can be visualised by considering two sets of NSPs (q_i, u_i), both having a normal distribution with identical variances, giving rise to a set of polarization values, p_i (see Fig. 2). The mean value, defined by $\bar{p} = \frac{\Sigma p_i}{N}$ does not lie at the centre of the distribution. However an estimate of the centre is given by $p_c = (\bar{q}^2 + \bar{u}^2)^{\frac{1}{2}}$, this quantity frequently being chosen to represent the value of p (e.g. see Dyck and Jennings, 1971). When a line of length $\bar{p}$ or p_c, anchored at the origin is swept through the q, u plane, the generated arc does not divide the data into equal portions.

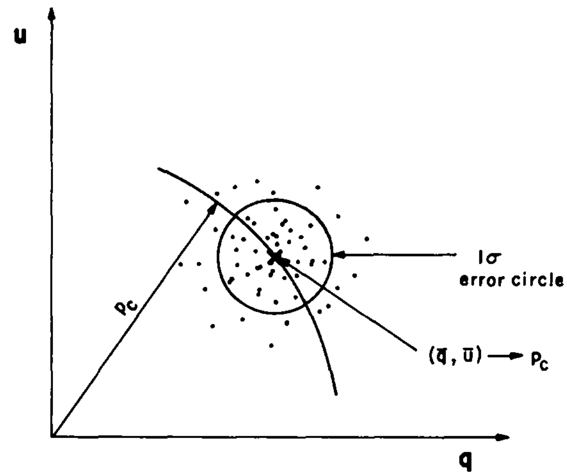


Fig. 2. A typical distribution of repeated NSP measurements is displayed with the centre of the distribution ($\bar{q}$, $\bar{u}$) providing a polarization value $p_c = (\bar{q}^2 + \bar{u}^2)^{\frac{1}{2}}$. In relation to the data distribution, it is obvious that if p_c is used as a radius to provide an arc through the data, its value is a biased estimate for p.

Many observers appear to be unaware of the statistical behaviour of p and assume that it follows a normal distribution. Without bias being taken into account, collations and comparisons of data are hazardous. Measurements made with different signal-to-noise ratios provide results with different amounts of bias. Mistaken polarimetric variability or polarization spectral dependence can all too easily be "detected" in data embracing a variety of signal-to-noise conditions. Again, because p follows a biased distribution, it is inappropriate merely to form simple means or weighted means from repeated measurements, but this is often done (e.g. see Dyck and Johnson, 1969).

In the establishment of polarimetric standard stars, special attention needs to be applied to the calculation of exact confidence intervals based on rigorous statistical principles. However, even the most recent observational studies of unpolarized standards (Tinbergen, 1982) and of polarized standard stars (Hsu and Breger, 1982) fall short in this most important area. More observations are certainly required here with deeper analysis.

It is sometimes boasted that polarimetry with a fast modulator can be performed in adverse conditions such as varying transparency caused by cirrus clouds. What is usually meant by the statement is that each result has errors within those given by the photon noise for the particular integration. However the usefulness of the measurements is limited as the data provide results with a range of bias according to the individual signal-to-noise values and calculation of the overall confidence limits is extremely difficult. Time dependent polarimetric changes are all too easy to "detect" in such data.

Some workers (e.g. Poeckert and Marlborough, 1976) do allow for bias in their data reductions according to a recipe proposed by Serkowski (1958, 1962) but it has recently been demonstrated by Simmons and Stewart (1985) [and see below] that this adopted procedure is not the best and also introduces a bias of its own.

Presentations of errors associated with p are also generally lax. In an early report on the wavelength dependence of interstellar polarization, Wolstencroft and Nandy (1971) present remarkable structure across the spectrum, but with an inadequate error assessment; the spectral values of p are given without reference to their associated ϕ values, the spread of the latter possibly being an indicator as to the validity of the interpretation on the variations of p.

Neither is it the case that all papers underestimate the errors associated with presented p values. It may be seen from Fig. 2 that for data centred away from the origin, a preliminary estimate for the variance of p is equal to that of the individual NSPs (see Equations (15)) and does not involve the combination of their variances as quoted by Klare *et al.* (1972) and Arsenijevic *et al.* (1979).

If the "best" estimate for p is to be presented from a series of observations and the associated errors prescribed, it is essential to have an understanding of the statistics of p. These problems are addressed below.

4.2 Estimation of the Degree of Polarization

Serkowski (1958, 1962) demonstrated that if repeated determinations of q and u come from normal distributions with equal variances, the mean value of p as determined from Equation (3) is a biased estimate of the true polarization p_o. Serkowski also suggested a method of correcting for the bias on the mean with the unbiased estimate being determined according to the uncertainty on the NSPs, their values being assumed to be known exactly.

In radio astronomy Wardle and Kronberg (1974) offered an alternative approach to correct for bias in p. The initial assumptions concerning the NSPs are the same as Serkowski's but the unbiased estimator is chosen as the value of p_o which makes the observed polarization p (biased) most probable.

Simmons and Stewart (1985) have introduced a further two methods of correcting for the bias in p, based on the maximum likelihood and the median estimators. In that paper they also perform an investigation and comparison of all four estimators.

Their initial assumption is that q and u are distributed normally and independently around the true values q_o, u_o and that the errors on the measured values q, u are known and both equal to σ. These assumptions are, however, invalidated in certain situations i.e. (a) when observations contain low signal-to-noise levels or scintillation dominated noise, in which case measurements of the NSPs come from non-normal distributions (Clarke *et al.*, 1983) and (b) in certain types of measurement techniques where a noise correlation is introduced between the NSPs (see, for example, Stewart, 1985).

Replacing q, u, q_o, u_o by q/σ, u/σ, q_o/σ, u_o/σ, the distribution function for p, known as the Rice distribution, is given by (Serkowski 1958, Vinokur 1965, Simmons and Stewart 1985)

$$F(p, p_o) = p \exp\left\{\frac{-\left[p^2 + p_o^2\right]}{2}\right\} J_o(ipp_o) \qquad (14)$$

where J_o = zeroth order Bessel function. The Rice distribution is shown in Fig. 3 for several values of p_o.

Adopting the nomenclature used by Simmons and Stewart for the unbiased estimators $\hat{p}_o$, *viz.*,

$\hat{p}_S$ for Serkowski's mean estimator, $\hat{p}_W$ for Wardle and Kronberg's most probable, $\hat{p}_{ML}$ for the maximum likelihood and $\hat{p}_M$ for the median, all four estimators are shown in Fig. 4, along with the uncorrected estimator $\hat{p}$.

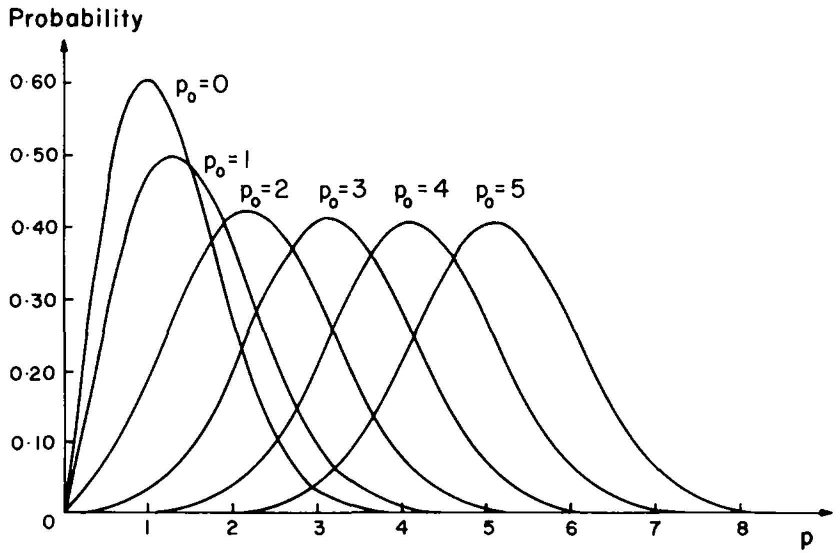


Fig. 3. The Rice distribution $F(p, p_0)$ as a function of p is displayed for values of p_0 = 1, 2, 3, 4 and 5.

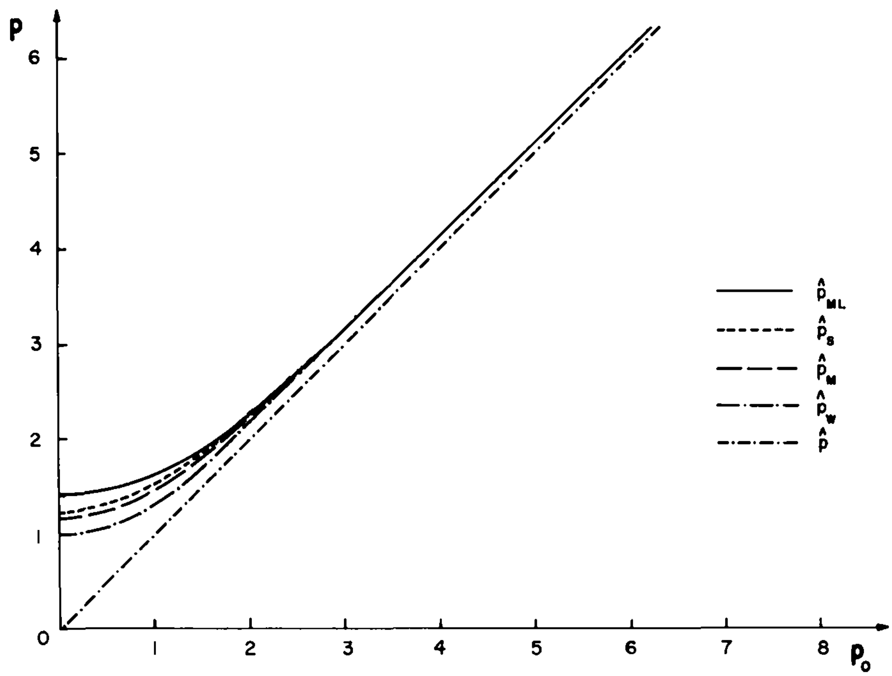
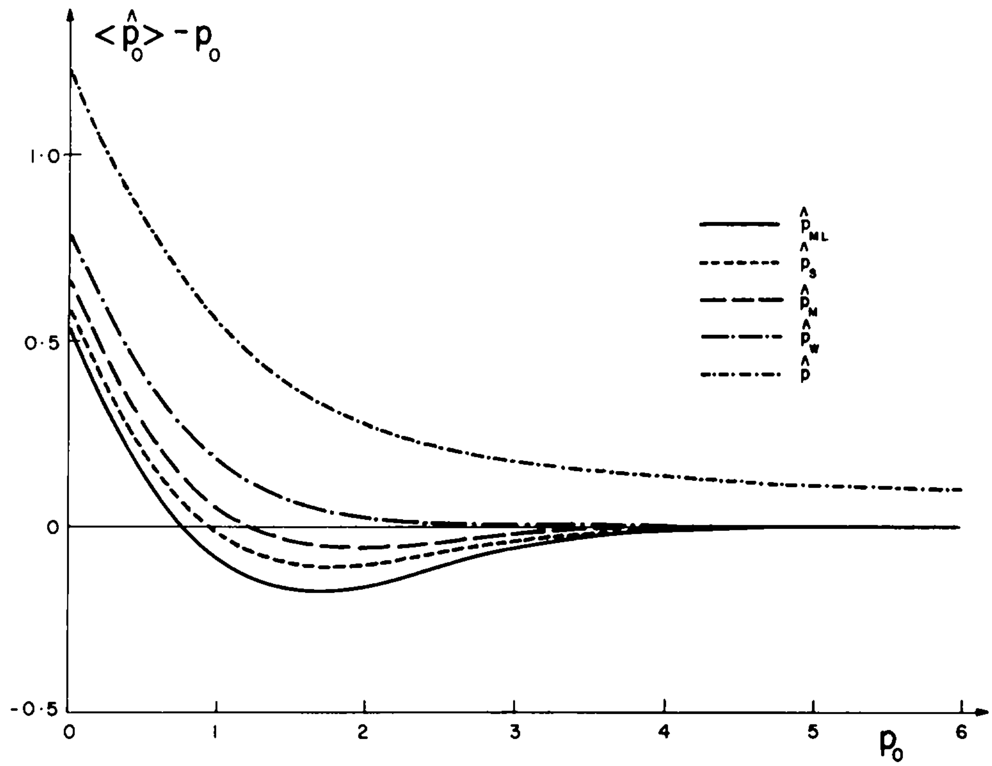


Fig. 4. Estimator curves for $\hat{p}_{ML}$, $\hat{p}_S$, $\hat{p}_M$, $\hat{p}_W$ and $\hat{p}$ are displayed and according to the appropriate choice may be used to obtain an estimated value of p_0 from the observed polarization.

Using angular brackets to signify expectation values, the bias $\langle\hat{p}_o\rangle - p_o$ and risk functions (or square error) $\langle(\hat{p}_o - p_o)^2\rangle$ for these estimators have also been evaluated by Simmons and Stewart. These are reproduced in Figures 5 (a) and (b).



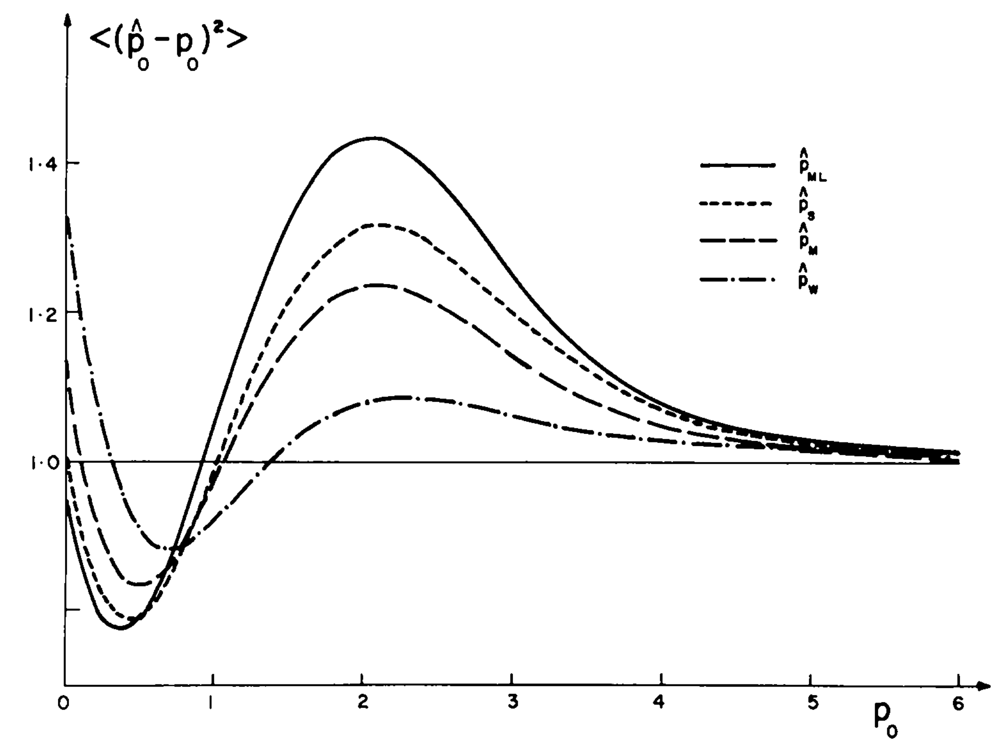


Fig. 5. (a) The bias $<\hat{p}_0> - p_0$ as a function of p_0 for the estimators p_{ML}, $\hat{p}_S$, $\hat{p}_M$, $\hat{p}_W$ and $\hat{p}$.
(b) The risk function (or square error) $<(\hat{p}_0 - p_0)^2>$ as a function of $\hat{p}_{ML}$, $\hat{p}_S$, $\hat{p}_M$ and $\hat{p}_W$.

From Figures 4, 5 (a) and (b) the following observations can be made.

(1) For values of $p > 4$ (i.e. signal-to-noise ratio $p/\sigma > 4$) all the correcting estimators agree, with the estimated value of p_o approximately given by $\hat{p}_o = (p^2 - 1)^{\frac{1}{2}}$.

(2) $\hat{p}_W$ and the naïve $\hat{p}$ are always positively biased, while the other estimators display positive and negative biasing depending on p_o.

(3) For small values of p_o, $p_o \lesssim 0.7$, $\hat{p}_{ML}$ is the best estimator, having least bias and associated risk.

(4) At larger values of p_o, $p_o \gtrsim 0.7$, $\hat{p}_W$ becomes the better estimator.

(5) The p_o region over which $\hat{p}_S$ and $\hat{p}_M$ are truly the best estimators is very small.

(6) $\hat{p}_{ML}$, $\hat{p}_S$, $\hat{p}_W$ and $\hat{p}_M$ all yield an estimated value of p_o which is less than $\hat{p}$.

Simmons and Stewart therefore conclude that Serkowski's $\hat{p}_S$ used by optical astronomers should not in general be applied. For small polarizations, $p_o \lesssim 0.7$, $\hat{p}_{ML}$ should be adopted and for $p_o \gtrsim 0.7$, $\hat{p}_W$ should then take over.

Other important features from their analysis can also be summarised in the following:-

(a) The continual application of $\hat{p}_W$ will always produce an overestimate of the true polarization because at all times it possesses positive bias. In other words, if the corrected values $\hat{p}_W$ are meaned, the resulting polarization will still contain a positive bias.

(b) As the biasing in $\hat{p}_{ML}$, $\hat{p}_M$ and $\hat{p}_S$ portrays positive and negative departures, their application will produce both overestimates or underestimates of polarization depending on the signal-to-noise ratio.

The noise on determinations of NSPs varies from observation to observation depending on many criteria, e.g. sample size, changes in seeing conditions, different telescopes and, perhaps of more astrophysical interest, changes in intrinsic intensity of stellar sources - for example binary or variable stars. In a set of measurements the ratio p_o/σ may be continually changing though p_o itself may remain constant. Even with debiased data, care must be taken not to allow variation of the systematic errors associated with the debiasing procedure to be misinterpreted as astrophysical effects.

(c) When the true polarization is zero, all estimators produce a positive bias. Hence, no matter which estimator is considered, in the limit a non-zero polarization will be determined.

Every area of investigation involving low polarizations will suffer from these effects. Some stellar sources reported to be displaying small values of polarization may therefore be intrinsically unpolarized.

These considerations are obviously important in establishing sets of stars to act as unpolarized standards. Even the most recent work done in this direction (e.g. Tinbergen, 1982) falls short in this respect on the discussion of the methods of data reduction and statistical procedures. That the effects are non-trivial shows up in the work of Serkowski, Mathewson and Ford (1975). In their study of the relationship between colour excess and polarization caused by interstellar dust, plots of $E_{(B-V)}$ against p reveal an upper boundary ($p \leq 9.0E_{(B-V)}$) under which all the stars lie. However, the same figure reveals a lower boundary ($p \geq 1.5E_{(B-V)}$) which was not commented on. It is suggested that this lower limit to p results from systematic errors which remain following the debiasing procedure. As there is a trend for the measured stars with high $E_{(B-V)}$ to be fainter and generally to provide polarimetry with greater noise, the systematic error grows with $E_{(B-V)}$, so giving a sloped lower boundary. The numerical value of $9.0E_{(B-V)}$ associated with the upper boundary must also be a slight overestimate.

Simmons and Stewart (1985) go further in the investigation of the estimators by deriving approximate estimator distributions for low polarization signal-to-noise levels. These are important in analysing a large number of independent polarization measures from one stellar source or a group of stellar sources to enable a decision as to their general polarimetric state.

In concluding this section, it is imperative that the observationalist is made aware of the consequences of bias produced for any adopted estimator. A major problem which now arises is in the direct comparison of small polarization measures taken from different instruments, telescopes, reduction procedures, etc., which have been corrected (or uncorrected) using different estimators. A better procedure would be for experimenters to present the original NSPs with associated errors so that their data might be made easily accessible to other workers. Any comparisons could then be made in a more unified way and be less susceptible to errors.

4.3 Confidence Intervals on the Degree of Polarization

A full statistical analysis of polarimetric measurement involves estimating the "best" value of p_0 and then prescribing a confidence interval or error bar on that estimated value. Often quoted formal errors on polarization are (e.g. Serkowski, 1962)

$$\sigma_p = \begin{cases} \sigma(2 - \frac{\pi}{2})^{\frac{1}{2}} & p_0 \gtrsim 0 \\ \sigma & p_0 \gg \sigma \end{cases} \qquad (15)$$

where again σ is the error on a measured NSP, it generally being taken that σ is symmetrical about p_0. However, the complete procedure involves correcting for bias and determining the asymmetric confidence intervals from the given distribution according to the particular signal-to-noise value.

Simmons and Stewart (1985) are the only authors to have given this problem any serious consideration, especially in the small p/σ regime. From the Rice distribution (Equation (14)) they construct precice percentage confidence intervals for p_0 using the uncorrected polarization p. Confidence intervals for 68%, 95% and 99% are reproduced in Figure 6, remembering that $p = p/\sigma$, $p_0 = p_0/\sigma$.

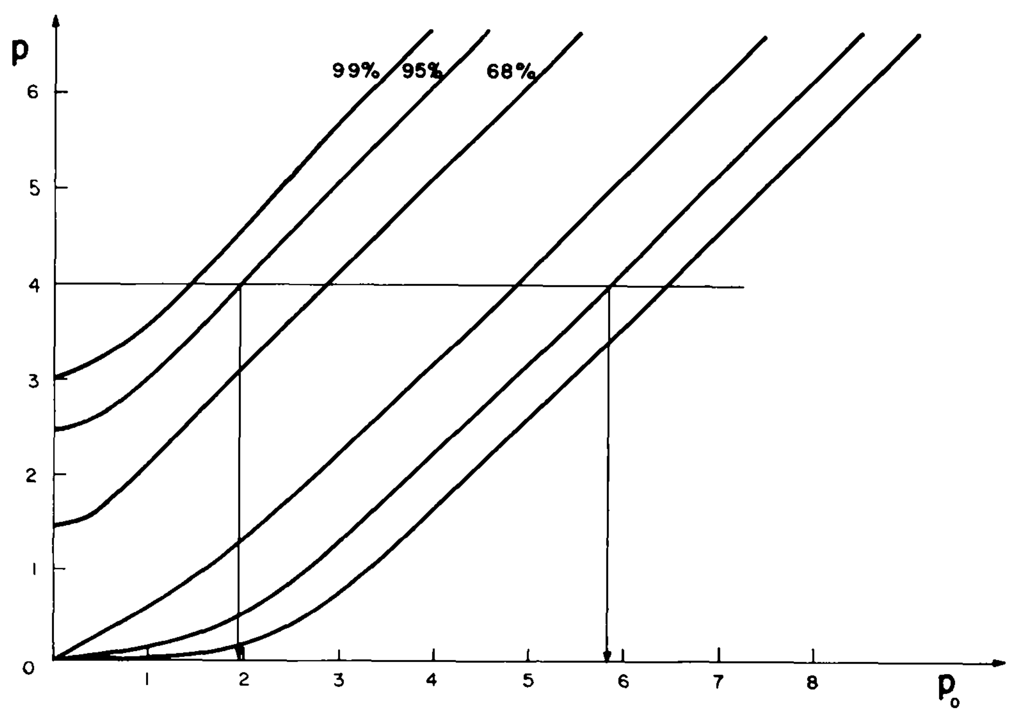


Fig. 6. Confidence intervals for the true polarization p_0 at the 68% (1σ), 95% (2σ) and 99% levels. Confidence intervals for an observed polarization p are found by drawing a line from p parallel to the p_0-axis. The upper and lower values of the intervals are then given by the p_0 co-ordinates of the points of intersection the parallel line makes with the curves. As an example, the 95% confidence interval is shown for $p_{observed} = 4$.

It can be seen that for $p \gtrsim 5$, the confidence intervals for p_o are those given by a normal distribution with mean value $\hat{p}_o = (p^2 - 1)^{\frac{1}{2}}$ and standard error = unity. However, for smaller polarizations or low signal-to-noise levels the confidence intervals for p_o are much wider than those associated with a normal distribution. In addition, they are asymmetric. Polarimetrists seem to be unaware of these important facts, most of them quoting normal error procedures which result in confidence intervals that are much smaller than the true ones. Furthermore, the formal method adopted by optical astronomers to estimate polarization error bars embraces values with $p_o < 0$ being contained in the confidence interval. Numerous examples of this fallacy exist in the literature. Clearly, by definition of p, all confidence intervals must have $p_o \geq 0$. It is interesting to note that in Markkanen's (1979) presentation of data for stars near the North Galactic Pole, the 1σ error bars are depicted as being symmetric except where they are cut short to prevent the embarrassment of crossing to the domain of $p_o < 0$.

5. ESTIMATION AND CONFIDENCE INTERVALS ON POSITION ANGLE

In a similar way to deriving the Rice distribution for p/σ, the distribution function for position angle φ is given by (Serkowski, 1962; Vinokur, 1965)

$$G(\theta,\theta_o,p_o) = \frac{1}{2\pi^{\frac{1}{2}}}\left\{\frac{1}{\pi^{\frac{1}{2}}} - \eta_o \exp(\eta_o^{\,2})\left[1 - \mathrm{erf}(\eta_o)\right]\right\}\exp\left(\frac{-p_o^{\,2}}{2}\right) \tag{16}$$

where $\eta_o = \frac{p_o}{2^{\frac{1}{2}}}\cos(\theta - \theta_o)$, $\theta = 2\phi$, and erf is the Gaussian error function. Unfortunately, $G(\theta,\theta_o,p_o)$ is a function of both θ_o and p making the problem of estimating and prescribing confidence intervals to θ very complicated. As yet no one has seriously treated this problem to the detail it deserves.

There are, however, a few remarks which can be made about the distribution. It is symmetric and, if p_o was unbiased, estimation of θ_o would produce an unbiased value of position angle. Conversely, substitution of a biased p_o value will result in a biased estimation of θ_o.

Quoted formal errors on estimated values of position angle are (Serkowski, 1958)

$$\sigma_\phi = \begin{cases} \dfrac{\pi}{12^{\frac{1}{2}}}\ \text{rad} = 51^{\circ}\!.96 & p_o \approx 0 \\[2ex] \dfrac{\sigma}{2p}\ \text{rad} = 28^{\circ}\!.65\ \dfrac{\sigma}{p} & p_o \gg \sigma \end{cases} \tag{17}$$

Whether these values for σ_ϕ are sufficiently good to calculate the various percentage confidence intervals for φ to any required accuracy at low polarizations or low signal-to-noise values has not yet been ascertained.

6. SMALL SAMPLE CONFIDENCE INTERVALS FOR p AND φ USING PROJECTION TECHNIQUES

6.1 The Calculation of Confidence Intervals

The previous sections have concentrated only on the estimation and assignment of p and θ (= 2φ) when the variances of the NSP distributions are assumed known, i.e. where repeated determinations of the NSPs provide large sample statistics. It has also been assumed that the uncertainties on the NSPs, i.e. σ_q, σ_u are the same and equal to σ. In many experimental conditions the above situation holds. However, when $\sigma_q \neq \sigma_u$, the resulting distribution functions for p and θ are more complicated and do not depend explicitly on p, p_o, θ and θ_o, making the point and interval estimation of p and θ very involved. As an added difficulty, experiments are often conducted with only a small sample of repeated measures of the NSPs being obtained, in which case the derived σ_q and σ_u represent estimates rather than true values. A mathematical derivation of the distributions of p and θ in these circumstances is indeed a formidable task and the resulting expressions will be extremely difficult to handle.

Determining the "best" estimator for the value of p_o and prescribing accurate confidence levels is a thorny problem. Without an analytic expression for the distribution function for p, small sample estimators and confidence intervals might be investigated using computer generated data. The accuracy of these determinations will depend on the sizes of the generated data sets, and, perhaps more critically, on the method of collecting polarization values in finite "bins" in order to produce a frequency distribution for any given value of p_o. This exercise has not yet been done.

When the determined values of p and σ have been obtained from a small set of repeated measures of the NSPs, with $p/\sigma \gtrsim 6$ say, it may be possible to assume that p comes from a normal distribution. A "best" estimate for p_o, would then perhaps be given by $\hat{p}_o = (p^2 - \sigma^2)^{\frac{1}{2}}$. Confidence intervals on $\hat{p}_o$ could therefore be assigned by using the sample dependent Student t modification on the value of σ. It must, however, be pointed out that for low values of p/σ, this method will produce only approximate estimates on the point and interval values of p.

Confidence intervals for p and for θ which take into consideration sample size and the condition $\sigma_q \neq \sigma_u$ have been developed by Stewart (1984). These intervals are conservative estimates and are constructed through the method of projection on the q-u plane. For example, in the large sample case, repeated measurements of NSPs would produce the mean values $(\bar{q}, \bar{u})$ and standard mean errors σ_q and σ_u. Confidence regions for the true NSPs (q_o, u_o) can be constructed by noting that

$$\frac{(\bar{q} - q_o)^2}{\sigma_q^{\,2}} + \frac{(\bar{u} - u_o)^2}{\sigma_u^{\,2}} \sim \chi_2^{\,2} \tag{18}$$

i.e. the left hand side is distributed in the form of a chi-square distribution with 2 degrees of freedom (see Simmons *et al.*, 1980; Simmons and Stewart, 1985). The values of $\chi_2^{\,2}$ at the 67%, 95% and 99% confidence levels are respectively 2.22, 5.99 and 9.21, therefore the appropriate elliptical confidence regions are obtained for (q_o, u_o) with semi-major and/or semi-minor axes given by $\sqrt{\sigma_q^{\,2}\,\chi_2^{\,2}}$ or $\sqrt{\sigma_u^{\,2}\,\chi_2^{\,2}}$. This region can be transformed into (p, θ) space and intervals constructed for p and θ as indicated in Figure 7 (a). The interval for p is obtained by drawing two perpendiculars from the origin to the ellipse, the intercepted values representing the upper and lower levels for p_o. The confidence intervals for θ are similar but are constructed by drawing two tangent lines from the origin to the ellipse.

For the circular case, $\sigma_q = \sigma_u = \sigma$, Figure 7 (b), the corresponding intervals are symmetrical around the estimated values p (biased) and θ.

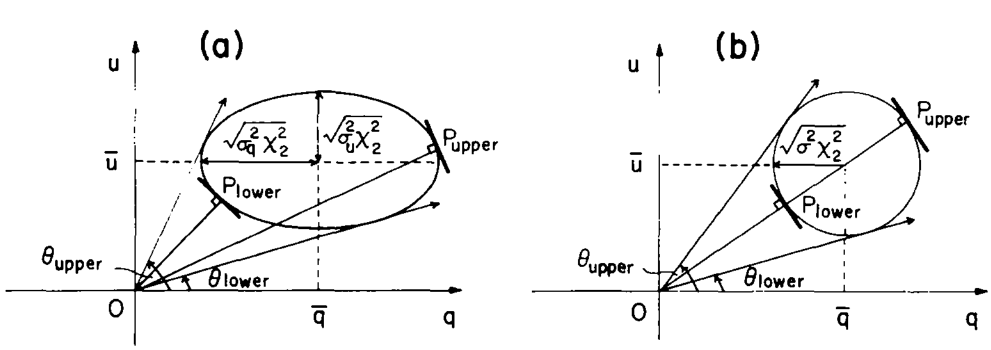


Fig. 7. Construction of confidence regions for the degree of polarization and position angle using projection techniques on the q, u plane for (a) the elliptical case with $\sigma_q^{\,2} \neq \sigma_u^{\,2}$ and for (b) the circular case with $\sigma_q^{\,2} = \sigma_u^{\,2} = \sigma^2$.

An interesting comparison can be made between the circular confidence intervals constructed by this method for the degree of polarization and the true intervals constructed by Simmons and Stewart (1985) (see Section 4.3). The projected intervals for p_o at the 67%, 95% and 99% levels are $p \pm 1.49\sigma$, $+2.45\sigma$ and $\pm 3.04\sigma$, respectively. (Notice that when the interval contains the origin the lower limit on polarization is zero). This merely states that the source can be considered polarized if p (biased) is greater than the predetermined value at the chosen confidence level. Comparing these values with the critical values of p derived by Simmons and Stewart (1985) shows that, if p is greater than the same three prescribed values, $p_o = 0$ is not contained in the confidence interval. This similarity occurs due to the method of constructing the interval from the Rice distribution, where at $p_o = 0$ ($q_o = u_o = 0$) the distribution is simply a chi-square with 2 degrees of freedom.

It can be seen that as the polarization becomes larger the true intervals in Figure 6 tend todards normality, becoming much smaller than those acquired through projection.

6.2 Testing for the Presence of Polarization

In the literature, various arbitrary criteria are used to decide whether or not a star exhibits polarization. For example, in a statistical study of supergiants, Coyne (1972) chose the polarization value corresponding to the 99% confidence interval from a distribution of measurements of unpolarized stars as representing the level at which a single observation providing that value or above deems the star to be "truly polarized".

The conclusions of the subsection immediately above related to very small polarizations suggests that sample dependent projection on the q-u plane is an ideal general method of testing for the existence of polarization.

Consider the following definitions:-

$$\bar{q} = \sum_{i=1}^{n_q} \frac{q_i}{n_q} \qquad \bar{u} = \sum_{i=1}^{n_u} \frac{u_i}{n_u}$$

$$\hat{S}_q^2 = \sum_{i=1}^{n_q} (q_i - \bar{q})^2/(n_q - 1) \qquad \hat{S}_u^2 = \sum_{i=1}^{n_u} (u_i - \bar{u})^2/(n_u - 1)$$

$$\hat{\sigma}_q^2 = S_q^2/n_q \qquad \hat{\sigma}_u^2 = S_u^2/n_u$$

where $(\bar{q}, \bar{u})$ are the mean values of the NPSs, n_q, n_u the respective sample numbers, $\hat{S}_q^2$, $\hat{S}_u^2$ estimates of the true sample variances S_q^2, S_u^2 and $\hat{\sigma}_q^2$, $\hat{\sigma}_u^2$ estimates of the true standard square mean errors σ_q^2, σ_u^2.

Stewart (1984) investigates three main cases

(1) $\sigma_q^2 = \sigma_u^2 = \sigma^2$ and $n_q = n_u = n$

(2) $\sigma_q^2 \neq \sigma_u^2$ and $n_q = n_u = n$

(3) $n_q \neq n_u$

(Before choosing the required case (1) or (2), the samples should be tested in consideration of $\sigma_q^2 = \sigma_u^2$. This is performed with an F-statistic test, the test being easily applicable and found in any good statistical text, *cf.* Snedecor and Cochran (1968)).

Case (1)

The projection statistic, Z, under consideration can be expressed in a similar way to the χ^2 statistic, i.e.

$$\frac{(\bar{q} - q_o)^2 + (\bar{u} - u_o)^2}{\hat{\sigma}^2} \sim Z \qquad (19)$$

where $\widehat{\sigma^2}$ is the *"pooled"* estimate of the true value σ^2. Since $n_q = n_u$, the $\widehat{\sigma^2}$ is defined by

$$\widehat{\sigma^2} = \frac{1}{2}\left(\widehat{\sigma_q^2} + \widehat{\sigma_u^2}\right) .$$

Stewart (1984) shows that after evaluating the distribution function for Z, a $(1 - \alpha)\times100\%$ confidence level for Z, i.e. $Z_{(1-\alpha)}$ is given by

$$Z_{(1-\alpha)} = 2(n-1)\left(\alpha^{-\frac{1}{(n-1)}} - 1\right)$$

The $(1-\alpha)$ 100% confidence regions for (q_o, u_o) are therefore derived from

$$\frac{(\bar{q} - q_o)^2 + (\bar{u} - u_o)^2}{\widehat{\sigma^2}} = Z_{(1-\alpha)} \tag{20}$$

representing a circle, centred on $(\bar{q}, \bar{u})$ with radius $\sqrt{\widehat{\sigma^2} Z_{(1-\alpha)}}$. Values of $Z_{(1-\alpha)}$ can be easily obtained from the above analytic expression. As expected, as $n \to \infty$ the percentage points tend towards those of a $\sqrt{\chi_2^2}$ distribution with 2 degrees of freedom.

<u>Cases (2) and (3)</u>

The formulation of these two cases is very similar. Under consideration is the distribution of a projection statistic Z, i.e.

$$\frac{(\bar{q} - q_o)^2}{\widehat{\sigma_q^2}} + \frac{(\bar{u} - u_o)^2}{\widehat{\sigma_u^2}} \sim Z . \tag{21}$$

It can be shown (Stewart, 1984) that the distribution function p(Z) for Z has the form

$$P(Z) = \frac{2 \cdot \nu_q^{\frac{\nu_q}{2}} \nu_u^{\frac{\nu_u}{2}}}{B\left(\frac{1}{2}, \frac{\nu_q}{2}\right) B\left(\frac{1}{2}, \frac{\nu_u}{2}\right)} \int_0^{\pi/2} \frac{d\theta}{\left[Z \sin^2\theta + \nu_q\right]^{\frac{1}{2}(\nu_q+1)} \left[Z \cos^2\theta + \nu_u\right]^{\frac{1}{2}(\nu_u+1)}} \tag{22}$$

where B (x, y) is the Beta function, and $\nu_q = (n_q - 1)$, $\nu_u = (n_u - 1)$.

$(1-\alpha)\times100\%$ confidence intervals can be constructed by numerically integrating P(Z) to the required $Z_{(1-\alpha)}$ confidence level, i.e.

$$\int_0^{Z_{(1-\alpha)}} P(Z)dZ = (1 - \alpha) .$$

The confidence region for (q_o, u_o) is therefore given by

$$\frac{(\bar{q} - q_o)^2}{\widehat{\sigma_q^2}} + \frac{(\bar{u} - u_o)^2}{\widehat{\sigma_u^2}} = Z_{(1-\alpha)} \tag{23}$$

which is an ellipse centred on $(\bar{q}, \bar{u})$ with semi-major, semi-minor axes given by $\sqrt{\widehat{\sigma_q^2}\, Z_{(1-\alpha)}}$, $\sqrt{\widehat{\sigma_u^2}\, Z_{(1-\alpha)}}$.

Tables 4 and 5 display $\sqrt{Z_{(1-\alpha)}}$ for cases (1) and (2). Table 4 shows $\sqrt{Z_{(1-\alpha)}}$ as a function of $n_q = n_u = n$, while Table 5 exhibits $\sqrt{Z_{(1-\alpha)}}$ as a function of ν_q and ν_u.

TABLE 4

Testing for the presence of polarization: $\sigma_q^2 \neq \sigma_u^2$; $n_q = n_u = n$

Values of $\sqrt{z_{(1-\alpha)}}$ at the 67%, 90%, 95% and 99% confidence levels

n	67%	90%	95%	99%
2	3.640	12.800	27.200	104.000
3	2.200	4.462	6.376	14.388
4	1.907	3.360	4.365	7.611
5	1.784	2.959	3.691	5.769
6	1.717	2.755	3.362	4.955
7	1.675	2.633	3.169	4.506
8	1.646	2.551	3.043	4.224
9	1.624	2.493	2.954	4.031
10	1.608	2.449	2.888	3.891
11	1.595	2.415	2.838	3.786
12	1.585	2.388	2.797	3.703
13	1.577	2.366	2.764	3.637
14	1.570	2.347	2.737	3.582
15	1.564	2.332	2.714	3.537
16	1.559	2.318	2.695	3.498
17	1.554	2.307	2.678	3.465
18	1.550	2.297	2.663	3.437
19	1.546	2.288	2.650	3.411
20	1.543	2.280	2.639	3.389
21	1.540	2.273	2.628	3.370
22	1.538	2.266	2.619	3.352
23	1.536	2.261	2.611	3.336
24	1.534	2.255	2.603	3.322
25	1.532	2.251	2.597	3.309
26	1.530	2.248	2.590	3.269
27	1.528	2.242	2.585	3.286
28	1.527	2.239	2.579	3.276
29	1.525	2.235	2.574	3.266
30	1.524	2.232	2.570	3.258
31	1.523	2.229	2.565	3.250
32	1.522	2.226	2.561	3.243
33	1.521	2.224	2.558	3.236
34	1.520	2.221	2.554	3.229
35	1.519	2.219	2.551	3.223
36	1.518	2.217	2.548	3.217
37	1.517	2.214	2.545	3.212
38	1.517	2.213	2.542	3.207
39	1.516	2.211	2.540	3.202
40	1.515	2.209	2.537	3.197
41	1.515	2.207	2.535	3.193
42	1.514	2.206	2.533	3.189
43	1.513	2.205	2.531	3.185
44	1.513	2.203	2.529	3.182
45	1.512	2.202	2.527	3.178
46	1.512	2.200	2.525	3.175
47	1.511	2.199	2.523	3.172
48	1.511	2.198	2.521	3.169
49	1.510	2.197	2.520	3.166
50	1.510	2.196	2.519	3.163
51	1.509	2.195	2.517	3.160
52	1.509	2.194	2.516	3.158
53	1.509	2.193	2.514	3.155
54	1.508	2.192	2.513	3.153
55	1.508	2.191	2.512	3.151
56	1.508	2.190	2.511	2.148
57	1.507	2.190	2.509	3.146
58	1.507	2.189	2.508	3.144
59	1.507	2.188	2.507	3.142
60	1.506	2.187	2.506	3.140
61	1.506	2.187	2.505	3.139
62	1.506	2.186	2.504	3.137
63	1.505	2.185	2.503	3.135
64	1.505	2.185	2.502	3.134
65	1.505	2.184	2.502	3.132
66	1.505	2.184	2.501	3.130
67	1.504	2.183	2.500	3.129
68	1.504	2.182	2.499	3.127
69	1.504	2.182	2.498	3.126
70	1.504	2.181	2.498	3.125
71	1.504	2.181	2.497	3.123
72	1.503	2.180	2.496	3.122
73	1.503	2.180	2.495	3.121
74	1.503	2.179	2.495	3.120
75	1.503	2.179	2.494	3.118
76	1.503	2.179	2.494	3.117
77	1.502	2.178	2.493	3.116
78	1.502	2.178	2.492	3.115
79	1.502	2.177	2.492	3.114
80	1.502	2.177	2.491	3.113
81	1.502	2.176	2.491	3.112
82	1.502	2.176	2.490	3.111
83	1.501	2.176	2.490	3.110
84	1.501	2.175	2.489	3.109
85	1.501	2.175	2.489	3.108
86	1.501	2.175	2.488	3.107
87	1.501	2.174	2.488	3.106
88	1.501	2.174	2.487	3.106
89	1.501	2.173	2.487	3.105
90	1.500	2.173	2.486	3.104
91	1.500	2.173	2.486	3.103
92	1.500	2.173	2.485	3.102
93	1.500	2.172	2.485	3.102
94	1.500	2.172	2.485	3.101
95	1.500	2.172	2.484	3.100
96	1.500	2.172	2.484	3.100
97	1.500	2.171	2.483	3.099
98	1.500	2.171	2.483	3.098
99	1.499	2.171	2.483	3.098
100	1.499	2.170	2.482	3.097
101	1.499	2.170	2.482	3.096
102	1.499	2.170	2.481	3.095
103	1.499	2.170	2.481	3.094
104	1.499	2.169	2.481	3.093
105	1.499	2.169	2.480	3.093
106	1.499	2.169	2.480	3.092
107	1.499	2.169	2.480	3.092
108	1.499	2.168	2.480	3.091
109	1.498	2.168	2.479	3.091
110	1.498	2.168	2.479	3.090
111	1.498	2.168	2.479	3.090
112	1.498	2.168	2.478	3.089
113	1.498	2.167	2.478	3.089
114	1.498	2.167	2.478	3.088
115	1.498	2.167	2.477	3.088
116	1.498	2.167	2.477	3.087
117	1.498	2.167	2.477	3.087
118	1.498	2.167	2.477	3.086
119	1.498	2.166	2.476	3.086
120	1.498	2.166	2.476	3.085
130	1.497	2.164	2.474	3.081
140	1.496	2.163	2.472	3.078
150	1.496	2.162	2.470	3.075
160	1.495	2.161	2.469	3.072
170	1.495	2.160	2.468	3.070
180	1.495	2.159	2.467	3.068
190	1.494	2.159	2.466	3.067
200	1.494	2.158	2.465	3.065
300	1.493	2.154	2.459	3.055
400	1.492	2.152	2.456	3.050
500	1.491	2.151	2.455	3.047
1000	1.490	2.148	2.451	3.041
5000	1.489	2.147	2.448	3.036
∞	1.489	2.146	2.448	3.035

TABLE 5

Testing for the presence of polarization: $n_q \neq n_u$ ($\nu_q \neq \nu_u$, where ν = n-1)
Values of $\sqrt{z_{(1-\alpha)}}$ at the 67%, 90%, 95% and 99% confidence levels

(1) Upper half - 67% level : Lower half - 90% level

ν_2 \ ν_1	5	10	15	20	25	30	40	50	60	70	80	90	100	120	150	∞
5	1.717 / 2.755	1.654	1.634	1.624	1.618	1.614	1.609	1.606	1.604	1.603	1.602	1.601	1.600	1.599	1.598	1.595
10	2.579	1.595 / 2.415	1.577	1.567	1.562	1.558	1.554	1.551	1.549	1.548	1.547	1.546	1.546	1.545	1.544	1.540
15	2.527	2.366	1.558 / 2.318	1.550	1.544	1.540	1.536	1.533	1.532	1.530	1.529	1.529	1.528	1.527	1.526	1.523
20	2.502	2.343	2.295	1.540 / 2.273	1.535	1.532	1.527	1.525	1.523	1.522	1.521	1.520	1.520	1.519	1.518	1.514
25	2.488	2.329	2.282	2.259	1.530 / 2.248	1.526	1.522	1.520	1.518	1.517	1.516	1.515	1.514	1.514	1.513	1.509
30	2.478	2.320	2.273	2.251	2.237	1.523 / 2.229	1.519	1.516	1.515	1.513	1.512	1.512	1.511	1.510	1.509	1.506
40	2.467	2.309	2.262	2.240	2.227	2.218	1.515 / 2.207	1.512	1.510	1.509	1.508	1.507	1.507	1.506	1.505	1.502
50	2.459	2.302	2.256	2.233	2.220	2.212	2.201	1.509 / 2.195	1.508	1.506	1.506	1.505	1.504	1.503	1.503	1.499
60	2.455	2.298	2.251	2.229	2.216	2.208	2.197	2.191	1.506 / 2.187	1.505	1.504	1.503	1.503	1.502	1.501	1.498
70	2.452	2.295	2.248	2.226	2.213	2.205	2.194	2.188	2.184	1.504 / 2.181	1.503	1.502	1.501	1.501	1.500	1.496
80	2.449	2.292	2.246	2.224	2.211	2.202	2.192	2.186	2.182	2.179	1.502 / 2.176	1.501	1.500	1.500	1.499	1.495
90	2.447	2.291	2.244	2.222	2.209	2.201	2.190	2.184	2.180	2.177	2.175	1.500 / 2.173	1.500	1.499	1.498	1.495
100	2.445	2.289	2.243	2.221	2.208	2.199	2.189	2.183	2.179	2.176	2.173	2.172	1.499 / 2.170	1.498	1.498	1.494
120	2.444	2.287	2.241	2.219	2.206	2.197	2.187	2.180	2.176	2.173	2.171	2.170	2.168	1.498 / 2.165	1.497	1.493
150	2.441	2.285	2.239	2.217	2.204	2.195	2.185	2.179	2.174	2.171	2.169	2.167	2.166	2.164	1.496 / 2.162	1.493
∞	2.432	2.276	2.230	2.208	2.195	2.187	2.177	2.170	2.166	2.163	2.161	2.159	2.158	2.156	2.154	1.489 / 2.145

TABLE 5, continued

$n_q \neq n_u$ ($\nu_q \neq \nu_u$, where $\nu = n-1$)

Values of $\sqrt{z_{(1-\alpha)}}$ at the 67%, 90%, 95% and 99% confidence levels

(2) Upper half - 95% level : Lower half - 99% level

ν_2 \ ν_1	5	10	15	20	25	30	40	50	60	70	80	90	100	120	150	∞
5	3.362 / 4.935	3.096	3.023	2.989	2.970	2.957	2.942	2.933	2.927	2.923	2.920	2.918	2.916	2.913	2.910	2.899
10	4.411	2.838 / 3.786	2.766	2.733	2.713	2.701	2.685	2.676	2.670	2.666	2.663	2.660	2.659	2.656	2.653	2.641
15	4.304	3.646	2.695 / 3.498	2.662	2.643	2.630	2.615	2.606	2.600	2.596	2.592	2.590	2.588	2.585	2.582	2.571
20	4.264	3.587	3.435	2.628 / 3.370	2.609	2.597	2.582	2.573	2.567	2.563	2.559	2.557	2.555	2.552	2.549	2.538
25	4.243	3.554	3.400	3.333	2.590 / 3.269	2.578	2.563	2.554	2.548	2.543	2.540	2.538	2.536	2.533	2.530	2.519
30	4.232	3.534	3.378	3.311	3.274	2.565 / 3.250	2.550	2.541	2.535	2.531	2.528	2.525	2.524	2.521	2.518	2.507
40	4.218	3.510	3.352	3.284	3.246	3.222	2.535 / 3.193	2.526	2.520	2.516	2.513	2.510	2.508	2.505	2.503	2.491
50	4.212	3.497	3.337	3.268	3.230	3.206	3.177	2.517 / 3.160	2.511	2.507	2.504	2.501	2.499	2.497	2.494	2.482
60	4.207	3.488	3.327	3.258	3.219	3.195	3.166	3.149	2.505 / 3.139	2.501	2.498	2.496	2.494	2.491	2.488	2.476
70	4.204	3.482	3.321	3.251	3.212	3.188	3.159	3.142	3.131	2.497 / 3.123	2.494	2.491	2.489	2.486	2.484	2.472
80	4.202	3.478	3.316	3.246	3.207	3.182	3.153	3.136	3.125	3.118	2.491 / 3.112	2.488	2.486	2.483	2.481	2.469
90	4.201	3.474	3.312	3.242	3.203	3.178	3.149	3.132	3.121	3.113	3.108	2.486	2.484	2.481	2.478	2.467
100	4.200	3.472	3.309	3.239	3.200	3.175	3.145	3.129	3.118	3.110	3.104	3.100	2.482 / 3.096	2.479	2.476	2.465
120	4.197	3.467	3.304	3.233	3.194	3.170	3.140	3.123	3.112	3.104	3.099	3.094	3.091	2.476 / 3.085	2.473	2.462
150	4.195	3.464	3.299	3.229	3.190	3.165	3.135	3.118	3.107	3.099	3.094	3.089	3.086	3.080	2.470 / 3.075	2.459
∞	4.190	3.449	3.283	3.211	3.171	3.146	3.116	3.099	3.088	3.080	3.074	3.070	3.066	3.060	3.055	2.448 / 3.035

7. THE DETECTION OF POLARIZATIONAL DIFFERENCES: THE WELCH TEST

A problem generally inadequately treated involves the quantitative assessment of detecting polarizational differences between two or more meaned values. Adopted procedures usually involve observing the error bars on the polarization points and then applying formal error combination formulae or even subjectively deciding whether the points are sufficiently displaced for differences to be present.

The astrophysical study might be in relation to deciding if a star has changed its polarization with time, or if two stars have differing polarizations, or if one star can be considered as being a member of a group representing a particular polarizational state, e.g. a set of unpolarized standard stars, or if the polarization of a star varies with wavelength. In all these cases, because of differing intensity levels and observing conditions the values for comparison are means based on data with different sample sizes (usually small) and different variances.

A suitable quantitative statistical test which takes into account sample size and variance is the Welch test, its application being appropriate to any of the situations outlined above (e.g. see Clarke and Brooks, 1985, and Clarke, Schwarz and Stewart, 1985).

When little or no correlation exists between the measurements of multivariate parameters (in this case the NSPs), testing for differences between the means of several data points should employ univariate rather than multivariate techniques (see Timm, 1975). In other words we should test for differences in q and u separately.

Given a set of g data points for q (and/or u) we usually have sample sizes $n_1 \neq n_2 \neq \ldots \neq n_g$ and standard mean errors $\hat{\sigma}_1^2 \neq \hat{\sigma}_2^2 \neq \ldots \neq \hat{\sigma}_g^2$. The problem of detecting a difference between the data points in these circumstances is generally referred to as the classical Behrens-Fisher problem. No exact analytical solution has been discovered but several statistics have been developed to cope with the situation.

Of the developed statistics, analysis by Brown and Forsythe (1974) and Kohr and Games (1974) conclude that in terms of size and power, (power being the probability of rejecting the hypothesis of equality of means when it is indeed false), the most recommended statistical test is the Welch test. The Welch statistic W is calculated from (see above references)

$$W = \frac{\sum_{i=1}^{g} w_i (q_i - \tilde{q})^2/(g-1)}{\left[1 + \frac{2(g-2)}{(g^2-1)} \sum_{i=1}^{g} (1 - w_i/u)^2/(n_i - 1)\right]} \tag{24}$$

where
$$w_i = \frac{1}{\hat{\sigma}_i^2}$$

$$u = \sum_{i=1}^{g} w_i$$

$$\tilde{q} = \sum_{i=1}^{g} w_i q_i / u \quad .$$

When all population means are equal (the standard mean errors being equal or unequal), W is approximately distributed as an F statistic with (g - 1) and f degrees of freedom where f is defined by

$$\frac{1}{f} = \left(3/(g^2 - 1)\right) \sum_{i=1}^{g} \left(1 - \frac{w_i}{u}\right)^2 / (n_i - 1)$$

The hypothesis of equality of population means is rejected at the usual 95% or 99% confidence levels.

Welch procedures may result in a polarimetric difference being detected between sets of points when in fact none of the individual points is polarized. This occurs because the Welch test weights observations with respect to mean values and standard mean errors. If for example two polarimetric measurements are close together and a third slightly displaced, more importance is attached to the concentrated pair lending support to the hypothesis that the outlying point does not come from the same population mean.

8. OBSERVATIONAL SCHEMES AND DATA REDUCTIONS

It is convenient to summarise the above discussions on the statistics of polarimetric parameters by considering an overall plan for making observations, for undertaking the data reductions and for combining or comparing measurements.

In the first place an estimate for the background signal strength should be obtained so that the observational time can be apportioned most efficiently, making sure that the background subtraction process does not impair the sought-for polarimetric accuracy. It must be remembered too that an imprecise background subtraction affects an assembly of repeated star measures in a systematic fashion. In the reporting of results, the noise of the background subtraction should be allowed for in the overall estimation of the uncertainties.

When organising the data collection, it is important to accumulate a reasonable sample of repeated measurements so that statistical tests can be applied. However, the signal should not be over-sampled to the degree that the individual photon count totals are less than a few thousand, otherwise a non-normal statistical behaviour will be induced into the distribution of NSPs. This would make comparisons between the observed behaviour and the expected that much more difficult.

Until the behaviour of an instrumental arrangement is firmly established over the wide range of observing conditions, the statistical behaviour of repeated determinations of the NSPs should be checked to see if their variance is consistent with photon counting statistics. It is also important to verify that they have a normal distribution and are not skew as might be the case if scintillation noise affects the measurements.

The effectiveness of operating an instrument claiming to provide results which are always limited only by photon counting statistics rather than by scintillation or variable atmospheric transparency should be judged in relation to the particular study. In this context, questions should be asked as to whether interpretations can be made from data with a wide range of signal-to-noise ratios, possibly producing different degrees of bias.

For data which is put on record for future statistical comparisons, it is better to tabulate and present values and uncertainties (allowing for sample dependence) for NSPs rather than p and ϕ. Values of p are generally biased and even debiasing procedures - of which the Serkowski recipe has been shown not to be the best - leave residual bias. For astrophysical studies involving comparisons of polarimetric data, it is preferable to perform any analysis on NSPs rather than repeated measures of p because of their simplicity of statistical behaviour. If claims are made for a particular model in respect of an analysis of p measurements, checks should be made as to the effect that the bias of p might have on the deduced model parameters.

Finally when comparing data sets with different variances with each set comprising small samples, procedures such as the Welch test should be applied rather than simplistic eyeball assessments using error bars.

REFERENCES

Arsenijevic, J., Vince, I., Kubicela, A. (1979) Proceedings of the III National Conference of Yugoslav Astronomers, Belgrade, 1977. (Belgrade 1979 - Union of Societies of Maths, Physics and Astronomers of Yugoslavia)

Brooks, A. (1984) Submitted to Biometrika. Also as an Internal Report - Dept. of Astronomy, Glasgow University.

Brown, M. B., Forsythe, A. B. (1974) Technometrics 16, 129

Clarke, D., Grainger, J. F. (1971) Polarized Light and Optical Measurement. Pergamon Press, Oxford.

Clarke, D. (1974) "Polarimetric Definitions" in Planets, Stars and Nebulae studied with photopolarimetry. Ed. T. Gehrels, University of Arizona Press

Clarke, D., Stewart, B. G., Schwarz, H. E., Brooks, A. (1983) Astron. Astrophys. 126, 260

Clarke, D., Brooks, A. (1985) Mon. Not. R. astr. Soc. 212, 211

Clarke, D., Schwarz, H. E., Stewart, B. G. (1985) Astron. Astrophys. 145, 232

Coyne, G. V. (1972) Specola Vaticana Commun. No: 53 - Taken from Proc. Third Coll. on Astrophys. Supergiant Stars. Trieste 1971 - pp. 93-107

Dyck, H. M., Johnson, H. R. (1969) Astrophys. J. 156, 389

Dyck, H. M., Jennings, M. C. (1971) Astron. Journal 76, 431

Hsu, J-C., Breger, M. (1982) Astrophys. J. 262, 732

James, J. F., Sternberg, R. S. (1969) The Design of Optical Spectrometers. Chapman and Hall, London

Klare, G., Neckel, Th., Schnur, G. (1972) Astron. Astrophys. Suppl. 5, 239

Kohr, R. L., Games, P. A. (1974) J. Exp. Educ. 43, 61

Markkanen, T. (1979) Astron. Astrophys. 74, 201

Poeckert, R., Marlborough, J. M. (1976) Astrophys. J. 206, 182

Serkowski, K. (1958) Acta. Astron. 8, 135

Serkowski, K. (1962) Advances in Astronomy and Astrophysics 1. Ed. Z. Kopal. Academic Press, New York, London

Serkowski, K. (1974a) "Polarimeters for Optical Astronomy" in Planets, Stars and Nebulae studied with photopolarimetry. Ed. T. Gehrels. University of Arizona Press.

Serkowski, K. (1974b) "Polarization Techniques" in Methods of Experimental Physics 12 - Part A. Ed. N. Carleton - Academic Press, New York, London

Serkowski, K., Mathewson, D. S., Ford, V. L. (1975) Astrophys. J. 196, 261

Simmons, J. F. L., Aspin, C., Brown, J. C. (1980) Astron. Astrophys. 91, 97

Simmons, J. F. L., Stewart, B. G. (1985) Astron. Astrophys. 142, 100

Snedecor, G. W., Cochran, W. G. (1968) Statistical Methods. 6th Edition. Iowa State Univ. Press

Stewart, B. G. (1984) Ph.D. Thesis, Univ. of Glasgow

Stewart, B. G. (1985) Astron. Astrophys. 143, 235

Timm, N. H. (1975) Multivariate Analysis with Application in Education and Psychology. Brooks-Cole, Monterey, Calif.

Tinbergen, J. (1982) Astron. Astrophys. 105, 53

Vinokur, M. (1965) Ann. Astrophysique 28, 412

Wall, J. V. (1979) Q. Jl. R. astr. Soc. 20, 138

Wardle, J. F. C., Kronberg, P. P. (1974) Astrophys. J. 194, 249

Wolstencroft, R. D., Nandy, K. (1971) Astrophys. Space Science 12, 158

Young, A. (1974) "Other Components in Photometric Systems" in Methods of Experimental Physics 12 - Part A. Ed. N. Carleton - Academic Press, New York, London