y, then the variances of the parameter estimates are minimum. Only for this choice of weights are the diagonal elements of $(A^TWA)^{-1}$ unbiased estimates of the variances of the parameter estimates. A nonlinear model like the structure factor formula may be replaced by a linear approximation. The estimates are then unbiased only to the extent that the linear approximation is valid, but biases can be reduced to arbitrarily small values by sufficiently precise observations. The "observations" may be raw data (as in the Rietveld method), net integrated intensities, or simple functions of the net integrated intensities, provided that the condition $\langle \underline{y}_i \rangle = M_i(\underline{x})$ is maintained. If the function is nonlinear, such as the extraction of a square root, some care must be taken to ensure this condition, but that care can be rewarded by improved precision in the parameter estimates.

If the model is correct, the properly weighted sum of squared residuals should equal $\underline{n} - \underline{p}$, where $\underline{n}$ is the number of observations and $\underline{p}$ is the number of parameters. Values larger than this are indicators of lack of fit. The common practice of assuming that these values are due to an incorrect scaling of weights that have the correct relative values is questionable, and inferences drawn from it should be viewed with caution. The positive square root of an estimated variance is an estimated standard deviation, or e. s. d. It is an indicator only of precision, which sets a lower limit to the uncertainty in the correspondence between estimated parameters and nature's values when the model is exactly correct. Statistical analysis can tell whether the model plausibly explains the observations. It cannot rule out the existence of systematic effects that bias the parameter estimates without contributing to lack of fit, nor can it rule out the existence of an entirely different model that would explain the observations as well or better. It therefore tells nothing about the actual accuracy of a measurement.

17.X-10
THE EFFECTS OF WEIGHTING SCHEMES ON ESTIMATED STANDARD DEVIATIONS AND ON ACCURACY.

By B.E. Robertson, Department of Physics and Astronomy, University of Regina, Regina, Sask., S4S 0A2, Canada.

Weights used in crystallographic least squares should be the reciprocals of the variances of the observations, but in a real experiment the variances of the observations are not known, other than their contribution from the Poisson counting statistics. Other errors which do not contribute to the difference between the estimate and the best value of a quantity used to describe a model (ie, they are not systematic errors) should also be used to determine the variances and the least-squares weights. How this should be done will depend on one's knowledge of the nature of other errors, and on the goals of the experiment. This is done by; (i) adding a contribution to the variance derived from the extent to which the sample variance of the intensities of the standard reflections exceeds the experimental variance (ii) replacing the experimentally determined variance with the sample variance obtained from the consistency of observations which should be identical according to the model, or adding to the experimentally determined variance a quantity such that the average modified experimental variance and sample variances are equal, or (iii) adding quantities to the variance such that the average value of Δ^2/σ^2 approaches (n-m)/m (where $\Delta = ||F_o| - |F_c||$ or $|I_o - I_c|$, σ^2 is the relevant variance, n is the number of observations and m is the number of least-square variables) for any groups of observations which may be averaged. All of these procedures may introduce systematic error into the weighting scheme. They normally lead to enhanced precision. However, if the values of the parameters which are to be estimated by the experiment are influenced by any of the systematic errors which are present, then the incorporation of systematic error into the least-squares weights is a "feed-back" process which may enhance or diminish the influence of systematic error. Furthermore, the enhanced precision obtained with modified weights may not imply enhanced accuracy.

In order to determine the effects of weight modification on the parameters of interest in routine structure determinations, we assume that an independent measure of the accuracy of the estimates of parameters from a refinement is their consistency; ie, more accurate refinements will lead to smaller sample variances among molecular parameters that are assumed to be exactly equal. This includes (I) chemically equivalent bonds which are not constrained to be equivalent by symmetry, (II) bonds in molecules in structures with more than one molecule per asymmetric unit, (III) multiple determinations of the same crystal structure. Consistency so determined has been used to investigate such questions as (a) the nature of the feedback from the contribution of systematic errors to the weights on accuracy, (b) the relation between accuracy and precision for various weighting schemes, (c) the effects of including weak reflections, (d) the merits of refinement on $|F|$ or I, etc.

17.X-11 THE BAYESIAN VIEWPOINT OF STATISTICAL PROCEDURES IN CRYSTALLOGRAPHY. by H. D. Flack, Laboratoire de Cristallographie, University of Geneva, 24 quai Ernest Ansermet, CH-1211 Genève 4, Switzerland.

From a Bayesian viewpoint, probability is subjective. A probability density function provides a measure of the crystallographer's degree of belief in the value of a random variable, a hypothesis, the estimate of physical parameters or a physical model used to interpret experimental results. Objectivity is regarded as being illusory - experimental observations are always interpreted through a model and one always has some prior idea of the numerical values entering into the model. Consider examples such as: weighting schemes to take account of systematic errors in the data or model, multiplication of e.s.d.'s by the goodness of fit value, averaging of symmetry-equivalent reflections, use of restraints (soft constraints, pseudo-observations), robust-resistant refinement. Such procedures are without any theoretical foundation in Statistics when viewed with the classical, Frequentist, notion of probability. On the other hand the Bayesian viewpoint does provide a clear framework within which to elaborate and criticize the above-mentionned procedures.

The Bayesian Three-Stage Model is a particularly fruitful

way to consider crystallographic refinement although to the present day it has not been used in practical applications to its full power. Stage One is associated with the observational error, typically quantum counting statistics for diffractometer measurements. Stage Two is concerned with physical modelling and its error - not only the relation between atomic parameters and calculated intensities but also the error associated with this modelling. Stage Three embodies prior knowledge of parameters - their possible values and dispersion. Stage One is very familiar and is the only error source appearing in a Frequentist analysis. Choice and adjustment of weighting schemes are closely related to stage Two. Use of restaints is clearly associated with stage Three.

Due to improvements in equipment and experimental technique, the extraction of physical and chemical information (and their error estimates) from many diffraction experiments, either single crystal or powder, is currently limited by model deficiences - systematic errors. Two courses of action are possible: Improve the models or explicitly quantify the poorness of the existing ones. The latter course must be undertaken using the Bayesian Philosophy.

17.X-12 REPORT ON THE WORK OF THE SUBCOMMITTEE ON STATISTICAL DESCRIPTORS. By D. Schwarzenbach, Institut de Cristallographie, Université de Lausanne, Bâtiment des Sciences Physiques, CH - 1015 Lausanne, Switzerland

The Subcommittee on Statistical Descriptors was established by the IUCr Commission on Crystallographic Nomenclature. The members are D. Schwarzenbach (chairman), H.D. Flack, W. Gonschorek, R.E. Marsh, E. Prince, B.E. Robertson and J.S. Rollett. Ex-officio members S.C. Abrahams and A.J.C. Wilson contributed actively to the work. The terms of reference are *to examine the validity of current statistical approaches used in estimating the variances in crystallographic quantities and to make recommendations for an improved methodology that rests securely on sound modern statistical theory and that can be widely adopted by the crystallographic community.* This ambitious programme may be surprising to some since there exists a vast literature on probability and statistics, and many of our undergraduate students are required to take courses in these disciplines. However, concern about the relevance of statistical theory as taught nowadays is widespread among experimentalists, be they crystallographers or particle physicists. Their problems are not confined to procedural issues, but have an important philosophical component originating from the need to justify the unavoidable lack of rigour with respect to mathematical theory in the treatment of experimental data. For this reason, one of the important discussions of the Subcommittee concerns the two alternate interpretations of probability, the *frequentist* and the *Bayesian* one. Another issue is the definition and use of the terms *model* and *observation*. Everybody agrees that the model, i.e. the conjecture about the physical reality used to interprete the data, comprises not only the crystal structure, but also data reduction procedures. The statement that *corrected observations and in particular structure amplitudes are not observed quantities* is of fundamental importance to some, and of minor consequence to others. In any case, recommendations of immediate value to practicing crystallographers cannot dispense with data treatment. Procedures to obtain mean values of symmetry-equivalent data and corresponding realistic variances become then important. Hotly debated were the kind of corrected observations which should be refined upon, in particular $|F|$ or $|F|^2$. The discussion of realistic weighting schemes to be used in least squares, and related issues such as the treatment of negative net intensities and the effects of neglecting weak reflections, makes it clear that variances of observations and corresponding weights are necessarily estimated from the observations themselves. Any weighting scheme based on the observed quantities results in bias of the refined parameters which might be reduced by weights computed by a combination of observed and calculated quantities. A preliminary report contains (a) a directory of statistical terms established for use by experimentalists; (b) a description of the statistical basis of refinement procedures and the Bayesian interpretation of statistics; (c) a section on the choice and significance of weighting schemes and (d) recommendations, some of which are easily implemented.

17.1-1 SIMULTANEOUS SEARCH FOR SYMMETRY-RELATED MOLECULES IN PATTERSON SPACE. By Christer E. Nordman, Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA.

In Patterson-space rotational search, the intramolecular vector set, or self-Patterson, of a known molecule or fragment is rotated to find the best fit to the Patterson, P. In translational search, an intermolecular vector set, or cross-Patterson, is correspondingly translated. New procedures, mainly for macromolecules, have been developed in which the search vector set is represented by a continuous vector density distribution of the search model. For rotation search, the self-Patterson is computed as the Patterson of a (triclinic) structure consisting of the known model placed in an orthogonal cell large enough to yield the self-Patterson free from intermolecular overlap. The vector density is transferred to a search grid of points lying on concentric spherical surfaces. This search grid is placed at the origin of P, and rotated through the three Euler angles to search for optimal fit. In non-triclinic crystals two (monoclinic), four (orthorhombic), or more self-Pattersons are embedded in P; these are related to each other by crystallographic symmetry axes. The method permits simultaneous search for all these self-Pattersons. At each search step the symmetry-related self-Patterson(s) are retrieved by two-dimensional interpolation on the spherical surfaces of the search grid. The sum of these two, four, or more copies of the self-Patterson is then compared to P. For translation search, the oriented model is placed in an extended crystal cell such that the computed Patterson yields separated self- and cross-Patterson regions, evaluated on a grid which matches that of P. The self-Patterson is subtracted from P, and the search done by comparing appropriate sums of overlapping model cross-Pattersons to P. Tests with myoglobin and two forms of superoxide dismutase will be described.