7.5. Statistical fluctuations

BY A. J. C. WILSON

7.5.1. Distributions of intensities of diffraction

Intensities of diffraction have two distinct probability distributions: (1) the *a priori* probability that an arbitrarily chosen reflection of a particular substance will have a particular 'true' intensity (R in the notation used below), and (2) the probability that a 'true' R will have an observed value R_o. The distributions of the first type depend on the symmetry and composition of the material, and are treated in Chapter 2.1 of Volume B of *International Tables for Crystallography* (Shmueli, 1993). The distributions of fluctuations of the second type, variations of the values R_o observed for a particular reflection, are treated here.

The crude counts or counting rates are rarely used directly in crystallographic calculations; they are subjected to some form of data processing to provide 'intensities of reflection' I_o in a form suitable for the determination of crystal structures, electron densities, line profiles for the study of defects, *etc*. The resulting intensity for the jth reflection, $I_{o,j}$, will be directly proportional to the corresponding $R_{o,j}$; when no confusion can arise, one of the subscripts will be omitted. The value of the proportionality factor c_j in

$$I_{o,j} = c_j R_{o,j} \qquad (7.5.1.1)$$

will be different for different reflections, since they will occur at different Bragg angles, have different absorption corrections, *etc*.

7.5.2. Counting modes

Whatever the radiation, in both single-crystal and powder diffractometry, the integrated intensity of a reflection is obtained as a difference between a counting rate averaged over a volume of reciprocal space intended to include the reflected intensity and a counting rate averaged over a neighbouring volume of reciprocal space intended to include only background. If these intentions are not effectively realized, there will be a systematic error in the measured intensity, but in any case there will be statistical fluctuations in the counting rates. The two basic modes (Parrish, 1956) are fixed-time counting and fixed-count timing. In the first, counts are accumulated for a pre-determined time interval, and the variance of the observed counting rate is proportional to the true (mean) counting rate. In the second, on the other hand, the counting is continued until a pre-determined number of counts is reached, and the variance of the observed counting rate is proportional to the square of the true counting rate. Put otherwise, the relative error in the intensities goes down inversely as the square root of the intensity for fixed-time counting, whereas it is independent of the intensity for fixed-count timing. Each mode has advantages, depending on the purpose of the measurements, and numerous modifications and compromises have been proposed in order to increase the efficiency of the use of the available time. References to some of the many papers are given in Section 7.5.7.

In principle, probability distributions can be determined for any postulated counting mode. In practice, they become complicated for all but the simplest modes; this is true even for the single measurement of the total counting rate or the background counting rate, but is even more pronounced for the distribution of their difference (the reflection-only rate). For most crystallographic purposes, however, it is only necessary to know the mean (to correct for bias, if present) and the variance (for the estimation of weights in refinement processes, see Part 8) of the distribution function.

7.5.3. Fixed-time counting

In the absence of disturbing influences (mains-voltage fluctuations, unrectified or unsmoothed high-tension supplies, 'dead time' of the counter or counter circuits, *etc*.), the number of counts recorded during the predetermined time interval used in the fixed-time mode will fluctuate in accordance with the Poisson probability distribution. If the 'true' number of counts to be expected in the interval is N, the probability that the observed number will be N_o is given by

$$p(N_o) = \exp(-N) N^{N_o} / N_o!, \qquad (7.5.3.1)$$

where all quantities appearing are necessarily non-negative. Both the mean and the variance of N_o are N. If the 'true' number of counts to be expected when the diffractometer is set to receive a reflection is T, and the 'true' number when it is set to receive the immediate background is B, the 'true' intensity of the reflection is

$$R = T - B, \qquad (7.5.3.2)$$

provided that the time interval used for the reflection is equal to the time interval used for the background. In practice, the 'observed' values of T_o and B_o fluctuate with probabilities given by (7.5.3.1) with T or B replacing N, so that the observed value,

$$R_o = T_o - B_o, \qquad (7.5.3.3)$$

will also fluctuate, and can, occasionally, take on negative values. The treatment of 'measured-as-negative' intensities is discussed in Section 7.5.6.

Although the sum of two Poisson-distributed variables is also Poisson, the difference is not, and the probability of R_o given by (7.5.3.3) has been shown to be (Skellam, 1946; Wilson, 1978)

$$p(R_o) = \exp\{-(B+T)\}(T/B)^{R_o/2} I_{|R_o|}[2(BT)^{1/2}], \qquad (7.5.3.4)$$

where I_n is the hyperbolic Bessel function of the first kind. The mean and variance of R_o are $T - B$ and $T + B$, respectively. If the times used for reflection and background are not equal, but are in the ratio of $k : 1$, the mean and variance of

$$R_o = T_o - kB_o \qquad (7.5.3.5)$$

are $T - kB$ and $T + k^2 B$, respectively. The distribution of R_o then involves a generalized Bessel function* depending on k:

$$p(R_o) = \exp\{-(B+T)\}(T^{2-k}/B)^{R_o/2} I^k_{|R_o|}[2(BT^k)^{1/2}]; \qquad (7.5.3.6)$$

for the properties of the generalized Bessel function, see Wright (1933) and Olkha & Rathie (1971). The probability distributions (7.5.3.4) and (7.5.3.6) seem to be very close to a normal distribution with the same mean and variance for R_o near R, but the probability of moderate deviations from R is less than normal and for large deviations is greater than normal.

In applications, it is usual to work with intensities expressed as counting rates, rather than as numbers of counts, even when fixed-time counting is used. The total intensity τ expressed as a counting rate is

* The term 'generalized Bessel function' seems to have no unique meaning in mathematics. The functions given that name by Paciorek & Chapuis (1994) belong to a different group.