Acta Cryst. (1981). A37, 227–228

Significance Testing of a Decrease in the *R* Factor. The Jackknife and Hamilton's *R*-Factor Ratio Test Applied to a Case of a False Minimum in Crystallography

BY DORIS M. THOMPSON,* MARY FRANCES RICHARDSON AND STUART M. ROTHSTEIN

Department of Chemistry, Brock University, St Catharines, Ontario, Canada L2S 3A1

(*Received* 9 *April* 1980; *accepted* 28 *October* 1980)

Abstract

The jackknife statistic and Hamilton's *R*-factor ratio test can both be applied to test the significance of a decrease in the *R* factor, and give similar results. In Hamilton's test, the form in which the hypothesis is cast may obscure the correct dimension of the hypothesis. This problem does not arise for the jackknife test.

Hamilton's well-known *R*-factor ratio test (Hamilton, 1965) addresses the problem of testing the significance of any decrease in *R* factor which occurs as refinement of a crystal structure continues. Describing successive models for which residuals are calculated usually involves releasing restraints on a number of parameters, which normally decreases $\mathscr{R}$. If the experimental *R*-factor ratio, $\mathscr{R} = R_I/R_{II}$, where R_{II} is the residual for the least restrained model (*i.e.* $R_I \geq R_{II}$), is greater than the tabulated value $\mathscr{R}_{b,N-m,\alpha}$ (Hamilton, 1965), then model II is significantly (at the α level) better than model I as a descriptor of the structure. The degrees of freedom, $N - m$, and the dimension of the hypothesis, b, are assumed known.

The two assumptions of linearity which Hamilton (1965) used to derive the distribution of $\mathscr{R}$ are almost never strictly satisfied in crystallographic problems, as he pointed out, which introduces uncertainty in the effective number of degrees of freedom. Furthermore, the dimension of the hypothesis, b, is usually (but not always) taken as the difference in the number of parameters refined in the two models, which sometimes leaves ambiguity in the definition of b (see below, however).

We have described non-parametric statistical tests based on the jackknife procedure (Rothstein, Richardson & Bell, 1978) and the Kendall τ coefficient (Richardson, Rothstein & Li, 1979) which are applicable specifically to cases where each of the structural models under consideration can be described by the same set of parameters (with the same restrictions). These tests seek to determine whether the distribution of weighted $(F_o - F_c)$ data for one model has a smaller variance than that distribution for another model at the α level of significance. Both tests gave similar results for the chemical problems treated so far, but further theoretical work (Rothstein, Bell & Richardson, 1981) showed that the Kendall procedure is not generally applicable. This theoretical work also invalidated a procedure based on Spearmin's rank correlation which has also been applied as an alternative to Hamilton's test (Li & Lee, 1980).

An important crystallographic situation which can be handled by the jackknife test is the case of the false minimum, an example of which has recently been described (Cotton & Rice, 1978). The structure of tris[diaqua-(tetra-μ-formato)-chromium(II)] decahydrate was refined to convergence at $R_1 = 0{\cdot}0551$, $R_2 = 0{\cdot}0805$, where $R_1 = \sum(||F_o| - |F_c||)/\sum|F_o|$ and $R_2 = [\sum w(||F_o| - |F_c||)^2/\sum w|F_o|^2]^{1/2}$. However, chemically unreasonable bond lengths caused a reassessment of the situation. After the positions of five atoms were set to produce reasonable bond lengths, refinement under the same conditions (number of parameters, number of observations) produced a new minimum ($R_1 = 0{\cdot}0504$, $R_2 = 0{\cdot}0684$). This structural model for the second minimum produced atom positions which were not greatly different from those of the model of the first minimum (average difference 0·023 Å for the 16 non-H atoms in the model, greatest difference 0·080 Å), but which gave more reasonable bond lengths in the suspect instances.

We applied the jackknife procedure (see references above) to the F_o and F_c data for model I, corresponding to the 'false' minimum, and for model II, which was assumed to be the better model. The jackknife statistic Q' was calculated with computer program *JACK* (Thompson, Richardson & Rothstein, 1980):

$$Q' \equiv \overline{\mathscr{L}}/V',$$

where

$$\overline{\mathscr{L}} = \sum \mathscr{L}_{-i}/N$$

$$\mathscr{L}_{-i} = N \ln(\mathscr{R}^2) - (N-1)\ln\mathscr{R}^2_{-i}$$

$$V'^2 = \sum_i (\mathscr{L}_{-i} - \overline{\mathscr{L}})^2/[N(N-1)]$$

$$\mathscr{R}^2 = R_I^2/R_{II}^2.$$

* Present address: Department of Chemistry, Austin College, Sherman, Texas 75090, USA.

0567-7394/81/020227-02$01.00

The quantity $\mathscr{R}^2$ is the square of the R-factor ratio, with $R_{\rm I}$ and $R_{\rm II}$ being the R factors for models I and II ($R_{\rm I} > R_{\rm II}$). The term $\mathscr{R}^2_{-i}$ is the squared R-factor ratio with the ith term deleted from both $R_{\rm I}$ and $R_{\rm II}$. The justification for this procedure is given by Rothstein, Richardson & Bell (1978).

Because of the large number of observations (1168 observed reflections were used), Q' was taken to approach the standard normal distribution; it may therefore be compared at tabulated upper-tail probabilities α with Z_α for a given significance level of α, Table 1. For an α level of 0·001, $Z_{0.001}$ is 3·1. Now Q' was found to be 3·91, which is greater than $Z_{0.001}$. The null hypothesis can therefore be rejected, and we can state that model II does describe the structure better than model I.

For comparison, Hamilton's R-factor ratio was calculated to be $\mathscr{R} = R$ (model I)/R (model II) = 1·39. The significance points of $\mathscr{R}$ (Hamilton, 1965) were obtained as

$$\mathscr{R}_{b,N-m,\alpha} = \left[\frac{b}{N-m} F_{b,N-m,\alpha} + 1\right]^{1/2}$$

with the subroutine *MDFI* (inverse of the F distribution function) from the IMSL Library (IMSL, 1979) for the determination of $F_{b,N-m,\alpha}$. Under the null hypothesis, H_0: model I also correctly describes the structure; alternative hypothesis, H_a: model II is the only correct description, the dimension of the hypothesis, b, is 133 (see below). From 1168 data points, $N - m = 1035$, and $\mathscr{R}_{133,1035,\alpha}$ was determined to be 1·073, 1·082, and 1·090 for significance levels α of 0·1, 0·01, and 0·001, respectively. Because the calculated ratio ($\mathscr{R} = 1{\cdot}39$) is not less than the critical value $\mathscr{R}_{133,1035,0.001}$, Hamilton's test also indicates a very small probability (less than 1 chance in 1000) that we would be wrong in rejecting the (null) hypothesis that the two models are equally good.

Hamilton's test is usually applied to cases in which the dimension of the hypothesis is clearly given by the difference in the number of parameters refined in the two models tested, and so we would like to call attention to the applicability of Hamilton's test, as well as the jackknife test, to the present situation. The dimension of the hypothesis is defined (Hamilton, 1965) as the rank of the matrix **Q** which specifies the set of linear equations (in m parameters X) which express the hypothesis: $\mathbf{Q}_{b,m}\mathbf{X}_{m,1} = \mathbf{Z}_{b,1}$. Therefore, the form in which the hypothesis is cast may obscure the correct value of b. For example, the null hypothesis is sometimes stated as 'H_0: the second structure fits the data better than the first structure'. This statement has frequently been assumed to be a yes/no hypothesis with $b = 1$ (Richardson, Rothstein & Li, 1979). However, the null hypothesis may also be stated as 'H_0: the data are also consistent with a structure describable by the *new* set of m parameters'. Now m equations are required for the hypothesis, and $b = m$. This is the case in Hamilton's first example (Hamilton, 1965) and in the present work.

Table 1. *Upper-tail probabilities* α *as a function of the statistic* Z_α

The values were obtained from Abramowitz & Stegun (1964).

α	Z_α	α	Z_α
0·0002	3·540	0·01	2·326
0·0005	3·291	0·02	2·054
0·001	3·090	0·5	1·645
0·002	2·878	0·1	1·282
0·005	2·576	0·2	0·842

Our results show that statistical tests such as these may be done on alternative structures in the case of a refined false minimum, independently of the 'chemical reasonableness' of either model of the structure. As pointed out previously (Richardson, Rothstein & Li, 1979), rejection of one model does not necessarily mean that the other is correct. The 'true' structure might lie at a minimum whose presence has not been discovered in the course of structure determination. Also, a particular α level means that one would be wrong to reject the null hypothesis $N\alpha$ times in N tests. One thus expects a statistical test to give the incorrect result a certain fraction of the time, even if the mathematical model upon which it is based holds rigorously for the physical data being tested. However, if both Hamilton's test and the jackknife test give the same indication for rejecting the null hypothesis, this should be strong evidence for rejecting one of the models over the other.

We thank Dr Gary Rice for helpful correspondence. This research was supported by a grant from the Natural Science and Engineering Research Council of Canada.

References

ABRAMOWITZ, M. & STEGUN, I. A. (1964). *Handbook of Mathematical Functions*. National Bureau of Standards (US), Washington.

COTTON, F. A. & RICE, G. W. (1978). *Inorg. Chem.* **17**, 688–692.

HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502–510.

IMSL Library (1979). International Mathematical and Statistical Library, Edition 7, IMSL LIB-0007, revised January 1979.

LI, W.-K. & LEE, S.-Y. (1980). *Org. Magn. Reson.* **13**, 97–99.

RICHARDSON, M. F., ROTHSTEIN, S. M. & LI, W.-K. (1979). *J. Magn. Reson.* **36**, 69–79.

ROTHSTEIN, S. M., BELL, W. D. & RICHARDSON, M. F. (1981). *J. Magn. Reson.* Submitted.

ROTHSTEIN, S. M., RICHARDSON, M. F. & BELL, W. D. (1978). *Acta Cryst.* A**34**, 969–974.

THOMPSON, D. M., RICHARDSON, M. F. & ROTHSTEIN, S. M. (1980). Submitted to the Quantum Chemistry Program Exchange (QCPE), Department of Chemistry, Indiana University, Bloomington, Indiana 47405, USA.