383. *The Computation of Velocities and Kinetic Constants of Reactions, with Particular Reference to Enzyme-catalysed Processes.*

By D. T. ELMORE, A. E. KINGSTON, and D. B. SHIELDS.

Curve-fitting procedures involving orthogonal polynomials are described which permit the computation of reaction velocities from kinetic data. The order of polynomial used is decided by applying Student's t test of significance to the coefficient of the orthogonal polynomial of highest order. Kinetic data at equal or unequal time intervals may be used. A weighted least-squares procedure is given for determining the Michaelis constant and maximum velocity of enzyme-catalysed reactions by computing the linear regression curves of $1/v$ on $1/[S]$ or $[S]/v$ on $[S]$, where v and $[S]$ are, respectively, the velocity and the substrate concentration. The weighted deviation of the worst point is tested for significance. A similar procedure is used for computing energies, enthalpies, and entropies of activation. All calculations have been programmed for the DEUCE digital computer.*

BOOMAN and NIEMANN[1] have described an empirical method of computing the velocity at zero time of enzyme-catalysed reactions from experimental data recording the progress of the reaction at equal intervals of time. In a subsequent brief communication,[2] this

---

* Copies of these programmes and operating instructions are available on application to the first author.

[1] Booman and Niemann, *J. Amer. Chem. Soc.*, 1956, **78**, 3642.

[2] Abrash, Kurtz, and Niemann, *Biochim. Biophys. Acta*, 1960, **45**, 378.

procedure was linked to the computation of Michaelis constants and was programmed for the DATATRON 220 computer. We have generalized this treatment by making it possible (*a*) to compute the velocity at any value of the independent variable which is included in the data, (*b*) to use any number of equidistant points up to 20, and (*c*) to compute initial velocities from experimental runs in which the value of the dependent variable at zero time is uncertain. In a later development, up to 32 unequally spaced points can be used in calculating up to 16 velocities at any value of the dependent variable. The two methods have been separately programmed for the DEUCE digital computer.

A function, $y = f(x)$, can be fitted to a set of $(n + 1)$ points at equidistant (Δx) values of the independent variable, the interval number being specified by $z(z = 0, 1 \cdots n)$, such that $x = \Delta x.z$, by using the orthogonal polynomials:[1,3]

$$y = \sum_{k=0}^{m} \gamma_{k,n} P_{k,n}(z); \tag{1}$$

$$\gamma_{k,n} = \sum_{z=0} y_z P_{k,n}(z) \Big/ \sum_{z=0}^{n} P^2_{k,n}(z); \tag{2}$$

$$P_{k,n}(z) = \sum_{j=0}^{k} (-1)^j \frac{[(k+j)(k+j-1)\cdot\cdots\cdot(j+1)][z(z-1)\cdot\cdots\cdot(z-j+1)]}{(k-j)!\,j!\,[n(n-1)\cdot\cdots\cdot(n-j+1)]}.$$

In practice, it was unnecessary to use polynomials higher than fifth order $(k \leqslant 5)$. For convenience, all the $P_{k,n}(z)$ values are multiplied by $P_{k,n}(0)$ in the subsequent treatment in order to give integral values, as suggested by Milne.[3] Hence we can write:

$$y = \sum_{k=0}^{m} a_k \gamma_{k,n}, \tag{3}$$

where

$$a_1 = P_{1,n}(0)\left[1 - \frac{2z}{n}\right],$$

$$a_2 = P_{2,n}(0)\left[1 - \frac{6z}{n} + \frac{6z(z-1)}{n(n-1)}\right],$$

$$a_3 = P_{3,n}(0)\left[1 - \frac{12z}{n} + \frac{30z(z-1)}{n(n-1)} - \frac{20z(z-1)(z-2)}{n(n-1)(n-2)}\right],$$

$$a_4 = P_{4,n}(0)\left[1 - \frac{20z}{n} + \frac{90z(z-1)}{n(n-1)} - \frac{140z(z-1)(z-2)}{n(n-1)(n-2)} + \frac{70z(z-1)(z-2)(z-3)}{n(n-1)(n-2)(n-3)}\right],$$

and

$$a_5 = P_{5,n}(0)\left[1 - \frac{30z}{n} + \frac{210z(z-1)}{n(n-1)} - \frac{560z(z-1)(z-2)}{n(n-1)(n-2)} + \frac{630z(z-1)(z-2)(z-3)}{n(n-1)(n-2)(n-3)} - \frac{252z(z-1)(z-2)(z-3)(z-4)}{n(n-1)(n-2)(n-3)(n-4)}\right].$$

The slope at any interval, z, can be obtained by differentiation of equation (3):

$$dy/dz = \sum_{k=1}^{m} a'_k \gamma_{k,n}, \tag{4}$$

---

[3] Milne, " Numerical Calculus," Princeton Univ. Press, Princeton, New Jersey, **1949**, p. 257.

where $a_1' = -2\mathrm{P}_{1,n}(0)/n$,

$$a_2' = \mathrm{P}_{2,n}(0)(12z - 6n)/n(n-1),$$

$$a_3' = \mathrm{P}_{3,n}(0)(-12n^2 + 6n + 60zn - 60z^2 - 4)/n(n-1)(n-2),$$

$$a_4' = \mathrm{P}_{4,n}(0)(-20n^3 + 30n^2 - 50n + 280z^3 - 420z^2n + 180zn^2 - 60zn + 100z)/n(n-1)(n-2)(n-3),$$

and

$$a_5' = \mathrm{P}_{5,n}(0)(-30n^4 + 90n^3 - 280n^2 + 100n - 1260z^4 - 1260z^2 + 420zn^3 + 2520z^3n - 1680z^2n^2 - 420zn^2 + z^2n + 1260zn - 48)/n(n-1)(n-2)(n-3)(n-4).$$

The gradient coefficients, a_k', can now be evaluated for sets of data comprising any number of points and at any interval. The coefficients $\gamma_{k,n}$ are determined for a particular set of data from equation (2) and the requisite values of $\mathrm{P}_{k,n}(z)$ quoted by Milne.[3] Substitution in equation (4) then affords the gradient, and the velocity is given by:

$$v = \frac{1}{\Delta x} \cdot \frac{\mathrm{d}y}{\mathrm{d}z}$$

Application of the method of least squares leads to the equation:

$$s_y^2 = \left\{ \sum_{z=0}^{n} y_z^2 - \sum_{k=0}^{m} \left[\sum_{z=0}^{n} \mathrm{P}^2_{k,n}(z) \right] \gamma^2_{k,n} \right\} \Big/ (n-m)$$

for the variance of the dependent variable, and the significance of the coefficient $\gamma_{k,n}$ can be tested by evaluating Student's t from the expression:

$$t_{n-k} = \gamma_{k,n} \left[\sum_{z=0}^{n} \mathrm{P}^2_{k,n}(z) \right]^{\frac{1}{2}} \Big/ s_y.$$

The standard deviation of the gradient can be calculated from the equation:

$$s_{\mathrm{d}y/\mathrm{d}z} = L_{m,n} s_y,$$

where
$$L_{m,n} = \sum_{k=1}^{m} \left[(a_k')^2 \Big/ \sum_{z=0}^{n} \mathrm{P}^2_{k,n}(z) \right].$$

The values of $L_{m,n}$, like those of the gradient coefficients, a_k', are independent of the experimental data. The standard deviation of the velocity is given then by:

$$s_v = L_{m,n} s_y / \Delta x.$$

In kinetic studies, y is some function which represents the concentration of starting materials or products, while x represents a time-scale. In our own case, y has usually taken the form of (*a*) moles of alkali or acid added by an autotitrator after a reaction at constant pH, or (*b*) optical density at a fixed wavelength. Experimental difficulties sometimes preclude a satisfactory start to a kinetic run, and there are two ways of dealing with this situation. Readings can be taken from the experimental trace starting at a point removed from zero time (Δx). This is made the interval in the calculation and z is set at -1 in order to calculate the coefficients a_k' and $L_{m,n}$. Because this procedure amounts to an extrapolation and because the interval is frequently larger than would normally be used, poor estimates of the gradient at zero time occasionally ensue. The alternative procedure involves using unequally spaced points so that poor regions in the trace can be omitted. The reading at zero time can be included or not as desired. The numerical and statistical treatments for this, second method are similar to those in the foregoing paragraphs. In equation (1), however, z is a linear transformation of the independent variable, x, such that $-1 \leqslant z \leqslant 1$. This transformation is effected by the introduction of two parameters, a and b, such that $a = (x_{\text{max.}} + x_{\text{min.}})/2$ and $b = (x_{\text{max.}} - x_{\text{min.}})/2$; then $z_i = (x_i - a)/b$. In this case, the relevant orthogonal polynomials are most easily derived [4] from the recurrence relationship:

$$\mathrm{P}_m(z_i) = (z_i - \alpha_m)\mathrm{P}_{m-1}(z_i) - \beta_m \mathrm{P}_{m-2}(z_i),$$

where

$$\alpha_m = \sum_i z_i \mathrm{P}^2{}_{m-1}(z_i) \Big/ \sum_i \mathrm{P}^2{}_{m-1}(z_i)$$

and

$$\beta_m = \sum \mathrm{P}^2{}_{m-1}(z_i) \Big/ \sum_i \mathrm{P}^2{}_{m-2}(z_i).$$

If we write $\mathrm{P}_m(z_i) = \sum_{k=0}^{m} \kappa_{m,k} z_i^k$, the coefficients $\kappa_{m,k}$ can be obtained from the analogous recurrence relationship:

$$\kappa_{m,k} = \kappa_{m-1,k-1} - \alpha_m \kappa_{m-1,k} - \beta_m \kappa_{m-2,k}$$

It is convenient to write the expansion in z in the form:

$$y_m(z) = \sum_{k=0}^{m} C_{m,k} z^k,$$

where

$$C_{m,k} = \sum_{j=0}^{m} \gamma_j \kappa_{j,k}$$

and γ_j follows from equation (2).

Evaluation of the variance of the dependent variable and the test of significance of the coefficients γ_j are carried out as in the case where equally spaced intervals are used. The gradient at any point z_i is then calculated from the equation:

$$\mathrm{d}y/\mathrm{d}z_i = \sum_{k=1}^{m} k C_{m,k} z_i^{k-1},$$

and the standard deviation of the gradient is given by:

$$S_{\mathrm{d}y/\mathrm{d}z_i} = s_y \left\{ \sum_{k=1}^{m} \left[\left(\sum_{j=1}^{k} j \kappa_{k,j} z_i^{j-1} \right)^2 \Big/ \sum_i \mathrm{P}^2{}_k(z_i) \right] \right\}^{\frac{1}{2}}$$

Then

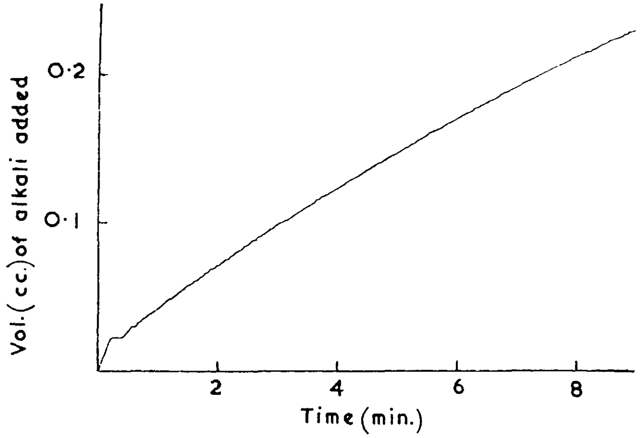
$$\frac{\mathrm{d}y}{\mathrm{d}x_i} = \frac{1}{b} \cdot \frac{\mathrm{d}y}{\mathrm{d}z_i} \text{ and } s_{\mathrm{d}y/\mathrm{d}x_i} = \frac{1}{b} \cdot s_{\mathrm{d}y/\mathrm{d}z_i}.$$

One set of data can be processed by either method on the DEUCE digital computer in 45—180 seconds, depending on the number of experimental points and the number of gradients required. The programmes are looped so that the order of the polynomial is increased without manual intervention. The t test is applied at each loop and, when this is not significant, the polynomial of unit lower order is evaluated.

Table 1 illustrates the use of both programmes to compute the initial velocity of a reaction (Figure) when a poor start required extrapolation. It will be seen that, in this example, variation of the number and spacing of points does not grossly affect the computed velocity or its standard deviation.

---

[4] (*a*) Forsythe, *J. Soc. Indust. Appl. Math.*, 1957, **5**, 74; (*b*) Guest, " Numerical Methods of Curve-fitting," Cambridge Univ. Press, 1961, p. 175.



Computation of initial velocity of the chymotrypsin-catalysed hydrolysis of methyl hippurate.

TABLE 1.

Computation of initial velocity of the chymotrypsin-catalysed hydrolysis of methyl hippurate by using different numbers and spacings of points.

Velocities ($10^4 v_0$; moles l.$^{-1}$ min.$^{-1}$) with standard deviations are given at the bottom of the Table. The figures in parentheses indicate the number of points, counted from the earliest observation in each column, used for each computation. (Sodium hydroxide = 0·1234N.)

x (min.) (Δx = 0·5)	$10^3 y$ (c.c./c.c.)	x (min.) (Δx = 0·75)	$10^3 y$ (c.c./c.c.)	x (min.) (Δx = 1·0)	$10^3 y$ (c.c./c.c.)	x (min.) (Δx irregular)	$10^3 y$ (c.c./c.c.)
0·5	2·72	0·75	3·52	1·0	4·18	0·75	3·52
1·0	4·18	1·50	5·74	2·0	7·06	0·95	4·02
1·5	5·74	2·25	7·84	3·0	9·84	1·25	4·98
2·0	7·06	3·00	9·84	4·0	12·28	1·60	6·00
2·5	8·46	3·75	11·72	5·0	14·66	2·00	7·06
3·0	9·84	4·50	13·52	6·0	16·86	2·45	8·36
3·5	11·06	5·25	15·14	7·0	18·90	2·95	9·70
4·0	12·28	6·00	16·86	8·0	20·82	3·50	11·06
4·5	13·52	6·75	18·34			4·10	12·54
5·0	14·66	7·50	19·90			4·75	14·02
5·5	15·78	8·25	21·24			5·45	15·66
6·0	16·86					6·20	17·20
6·5	17·90					7·00	18·90
7·0	18·90						
7·5	19·90						
3·921 ± 0·075 (8)		3·855 ± 0·049 (8)		3·840 ± 0·035 (8)		3·897 ± 0·115 (8)	
3·872 ± 0·038 (11)		3·963 ± 0·077 (11)				3·905 ± 0·077 (9)	
3·834 ± 0·021 (15)						3·927 ± 0·045 (10)	
						3·873 ± 0·048 (11)	
						3·873 ± 0·036 (12)	
						3·829 ± 0·034 (13)	

For an enzyme-catalysed reaction, the steady-state velocity, v, at a given substrate concentration, [S], is given by the Michaelis–Menten equation:

$$v = V[\mathrm{S}]/(K_{\mathrm{M}} + [\mathrm{S}]),$$

where V is the maximum velocity attained at infinite substrate concentration and K_{M} is the Michaelis constant. This equation can be rearranged in several ways in order to obtain linear relationships between functions of v and [S]:

$$\frac{1}{v} = \frac{1}{V} + \frac{K_{\mathrm{M}}}{V} \cdot \frac{1}{[\mathrm{S}]} \qquad (5)$$

$$\frac{[\mathrm{S}]}{v} = \frac{K_{\mathrm{M}}}{V} + \frac{[\mathrm{S}]}{V} \qquad (6)$$

$$\frac{v}{[\mathrm{S}]} = \frac{V}{K_{\mathrm{M}}} - \frac{v}{K_{\mathrm{M}}} \qquad (7)$$

There is no general agreement about which is the best form for computing K_{M} and V;[5-7] in addition, opinions differ about the use of weighted and unweighted observations in the least-squares method for computing the regression lines relating the variables in the above equations. The various arguments will not be reviewed here, but it may be noted[6] that equation (7) is inconvenient for statistical purposes, since both variables, $v/[\mathrm{S}]$ and v, are liable to experimental error. Application of the weighted least-squares method[6] to either equation (5) or (6), and the assumption that [S] is free from error, leads to the following expressions for V and K_{M}:

$$V = \frac{\sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 \sum\limits_i w_i v_i^4 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2}{\sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 \sum\limits_i w_i v_i^3 - \sum\limits_i w_i v_i^4/[\mathrm{S}_i] \sum\limits_i w_i v_i^3/[\mathrm{S}_i]}, \qquad (8)$$

and

$$K_{\mathrm{M}} = \frac{\sum\limits_i w_i v_i^3/[\mathrm{S}_i] \sum\limits_i w_i v_i^4 - \sum\limits_i w_i v_i^4/[\mathrm{S}_i] \sum\limits_i w_i v_i^3}{\sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 \sum\limits_i w_i v_i^3 - \sum\limits_i w_i v_i^4/[\mathrm{S}_i] \sum\limits_i w_i v_i^3/[\mathrm{S}_i]}, \qquad (9)$$

where w_i are the weights of the velocities v_i. The most efficient statistical weight is the reciprocal of the variance, *i.e.*, $w_i = 1/\sigma_v^2$. Application of the unweighted least-squares method to equation (5) (*i.e.*, $\sigma_{1/v} =$ constant) is numerically equivalent to using the weighted procedure in which $\sigma_v \propto v^2$, since $\sigma_{1/v} = \sigma_v/v^2$. Similarly, application of the unweighted procedure to equation (6) (*i.e.*, $\sigma_{[\mathrm{S}]/v} =$ constant) is numerically equivalent to using the weighted technique in which $\sigma_v \propto v^2/[\mathrm{S}]$, since $\sigma_{[\mathrm{S}]/v} = [\mathrm{S}]\sigma_v/v^2$. Wilkinson[7] has pointed out that, if experimental determinations of velocity at different substrate concentrations, such that $[\mathrm{S}]/K_{\mathrm{M}}$ varies from 1/3 to 3, have a constant variance, the relative weights of $1/v$ will vary by a factor of 81, while the relative weights of $[\mathrm{S}]/v$ will change by a factor of only 16/9. It should be pointed out, however, that if the measurements of velocity have a standard deviation which is proportional to the magnitude of the velocity, the ranges of the relative weights of $1/v$ and $[\mathrm{S}]/v$ are identical. If $[\mathrm{S}]/K_{\mathrm{M}}$ varies from 1/3 to 3, the relative weights of $1/v$ and $[\mathrm{S}]/v$ change by a factor of 9. The last supposition, that $\sigma_v \propto v$, leads to the equations:[6,8]

$$V = \frac{\sum\limits_i v_i^2 \sum\limits_i v_i^2/[\mathrm{S}_i]^2 - \left(\sum\limits_i v_i^2/[\mathrm{S}_i]\right)^2}{\sum\limits_i v_i \sum\limits_i v_i^2/[\mathrm{S}_i]^2 - \sum\limits_i v_i^2/[\mathrm{S}_i] \sum\limits_i v_i/[\mathrm{S}_i]},$$

and

$$K_{\mathrm{M}} = \frac{\sum\limits_i v_i^2 \sum\limits_i v_i/[\mathrm{S}_i] - \sum\limits_i v_i^2/[\mathrm{S}_i] \sum\limits_i v_i}{\sum\limits_i v_i \sum\limits_i v_i^2/[\mathrm{S}_i]^2 - \sum\limits_i v_i^2/[\mathrm{S}_i] \sum\limits_i v_i/[\mathrm{S}_i]}.$$

Rather than make any suppositions about the relative weights to be attached to experimental determinations of reaction velocity, we prefer to use the estimated value s_{v_i}, as

---

[5] Hofstee, *Nature*, 1959, **184**, 1296; Dixon and Webb, *ibid.*, p. 1298.
[6] Johansen and Lumry, *Compt. rend. Trav. Lab. Carlsberg*, 1961, **32**, 185.
[7] Wilkinson, *Biochem. J.*, 1961, **80**, 324.
[8] Stockell and Smith, *J. Biol. Chem.*, 1957, **227**, 1.

determined in the computation of velocity by the curve-fitting procedure outlined in the first part of this paper, in order to compute $1/s^2_{v_i}$. The latter is an estimate of w_i and is used in equations (8) and (9). In this way, reasonable account is taken of whether extrapolation has had to be employed to calculate the velocity and also of the order of the polynomial required to fit the experimental points. Undue disparity of weighting is avoided by using approximately the same number of experimental points for each velocity computation.

By using the method of weighted least squares to compute the regression line of $1/v$ on $1/[\mathrm{S}]$, estimates of the variance of the ordinal intercept, C_0, and slope, C_1, are given by:

$$s^2{}_{C_0} = \frac{s^2 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2}{\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2}$$

and $s^2{}_{C_1} = s^2 \sum\limits_i w_i v_i^4 \Big/ \left[\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2\right]$,

where

$$s^2 = \frac{1}{n-2}\left\{\sum_i w_i v_i^2 - \frac{\left(\sum\limits_i w_i v_i^3\right)^2}{\sum\limits_i w_i v_i^4} - \frac{\left(\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^3/[\mathrm{S}_i] - \sum\limits_i w_i v_i^3 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2}{\sum\limits_i w_i v_i^4\left[\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2\right]}\right\} \tag{10}$$

and n is the number of pairs of observations of v_i and $[\mathrm{S}_i]$. An estimate of the standard deviation of V, s_V, is given then by the expression:

$$s_V = V^2 s_{C_0}.$$

Since $K_{\mathrm{M}} = C_1/C_0$ and since C_1 and C_0 are not stochastically independent, an estimate of the variance of K_{M}, $s^2{}_{K_{\mathrm{M}}}$, is given by the equation:[9]

$$\begin{aligned} s^2{}_{K_{\mathrm{M}}} &= \left(\frac{\partial K_{\mathrm{M}}}{\partial C_0}\right)^2 s^2{}_{C_0} + \left(\frac{\partial K_{\mathrm{M}}}{\partial C_1}\right)^2 s^2{}_{C_1} + 2\,\frac{\partial K_{\mathrm{M}}}{\partial C_0}\cdot\frac{\partial K_{\mathrm{M}}}{\partial C_1}\operatorname{cov}(C_0, C_1) \\ &= \frac{C^2{}_1 s^2{}_{C_0}}{C^4{}_0} + \frac{s^2{}_{C_1}}{C^2{}_0} + \frac{2C_1 s^2 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]}{C^3{}_0\left[\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2\right]} \\ &= V^2 s^2\left[\frac{K^2{}_{\mathrm{M}} \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 + 2K_{\mathrm{M}} \sum\limits_i w_i v_i^4/[\mathrm{S}_i] + \sum\limits_i w_i v_i^4}{\sum\limits_i w_i v_i^4 \sum\limits_i w_i v_i^4/[\mathrm{S}_i]^2 - \left(\sum\limits_i w_i v_i^4/[\mathrm{S}_i]\right)^2}\right]. \end{aligned}$$

It is desirable to test the experimental data in order to see if any points can safely be rejected on the grounds that their deviations from the fitted line depart significantly from normal distributions. After the regression line has been obtained, and hence K_{M} and V from all the experimental data, the calculated value of the velocity, v'_i, is evaluated for each substrate concentration used. Weighted deviations, $\sqrt{w_i}(v_i - v_i^2/v_i')$, are calculated, and the point, for which $|\sqrt{w_i}(v_i - v_i^2/v_i')|$ is a maximum, is determined. The experimental observations corresponding to this point ($[\mathrm{S}_m]$, v_m, w_m) are removed from the remainder of the data, and a new regression line is computed. The calculated velocity, v_m', at the

---

[9] Ref. 4(*b*), p. 98.

substrate concentration [S_m] is computed from the new regression line; from this the corrected maximum weighted deviation, $\sqrt{w_m}(v_m - v_m^2/v_m')$ is evaluated. The standard deviation, s', of the dependent variable, and hence of the deviations of the dependent variable, for the second regression line is calculated according to equation (10). The significance of the corrected maximum weighted deviation from the second regression line is tested by evaluating Student's t from the expression:

$$t_{n-3} = |\sqrt{w_m}(v_m - v_m^2/v_m')/s'|.$$

Computation stops at this stage, and the values of K_M and V from the first regression line are punched out, unless the calculated value of t exceeds the value in the Tables. In the latter case, the data corresponding to the second worst point are removed, and the above procedure is repeated until the t test is passed or until only three points remain. This procedure for testing the homogeneity of residuals is based on that described for unweighted samples.[10]

TABLE 2.

Computation of K_M and V for the methyl hippurate–chymotrypsin system at pH 7·8 and 25°.

Initial substrate concn. (10^3[S]; mole l.$^{-1}$)	Velocity (10^4v; moles l.$^{-1}$ min.$^{-1}$)	Standard deviation (10^4v moles l.$^{-1}$ min.$^{-1}$)
20·0	4·8242	0·0217
12·0 •	4·0094	0·0336
8·0	4·1194	0·0273
6·4	3·7397	0·0133
4·8	3·2908	0·0185
3·2	2·8715	0·0600
1·6	1·8940	0·0268

	10^3K_M(moles l.$^{-1}$)	10^4V (moles l.$^{-1}$ min.$^{-1}$)	$t_{\text{calc.}}$	$t_{95\%}$ (Tables)
First computation (7 points)	3·105 ± 0·343	5·527 ± 0·187	3·80	2·78
Second computation (6 points) ...	3·231 ± 0·194	5·625 ± 0·108	2·17	3·18

• This point was rejected by the t test.

An example of the use of this method of computing K_M and V is given in Table 2; these particular data contain a point which was rejected as a result of applying the statistical test above. It will be noticed that the standard deviations of K_M and V are considerably reduced in the second computation from which this point was omitted.

We have also applied the weighted least-squares procedure to the problem of computing kinetic activation constants. The regression line of $\log_e k$ on $1/T$ is derived in order to calculate the parameters in the equations:

$$\log_e k = E_a/\boldsymbol{R}T + \log_e A \tag{11}$$

$$E_a = \Delta H^\ddagger + \boldsymbol{R}T_0 \tag{12}$$

$$\Delta S^\ddagger = \boldsymbol{R}(\log_e A - 1 - \log_e T_0 - \log_e \boldsymbol{k}/\boldsymbol{h}) \tag{13}$$

$$\Delta G^\ddagger = \Delta H^\ddagger - T_0\Delta S^\ddagger \tag{14}$$

where E_a is the energy of activation, $\Delta H^\ddagger$, $\Delta S^\ddagger$, and $\Delta G^\ddagger$ are the enthalpy, entropy, and free energy of activation at the absolute temperature T_0, A is the frequency factor, and $\boldsymbol{k}$ and $\boldsymbol{h}$ are Boltzmann's and Planck's constant, respectively. The data consist of rate constants, k_i, with standard deviations, s_{k_i} at n different temperatures T_i. The relevant

---

[10] Acton, " Analysis of Straight-Line Data," John Wiley & Sons, Inc., New York, 1959, p. 226.

equations, when T_i is assumed to be free from error, to compute the parameters of equation (11) are:

$$E_a = -\boldsymbol{R}\left[\frac{\sum_i w_i \sum_i w_i(\log_e k_i)/T_i - \sum_i w_i(\log_e k_i)\sum_i w_i/T_i}{\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2}\right]$$

and

$$\log_e A = \frac{\sum_i w_i/T_i^2 \sum_i w_i(\log_e k_i) - \sum_i w_i/T_i \sum_i w_i(\log_e k_i)/T_i}{\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2},$$

where

$$w_i = 1/\sigma^2_{\log_e k_i} \approx k_i^2/s^2_{k_i}.$$

It may be noted here that application of the unweighted least-squares procedure to equation (11) (*i.e.*, $\log_e k_i =$ constant) is equivalent to using the weighted method in which $s_{k_i} \propto k_i$.

The remaining parameters can be derived from equations (12—14). Estimates of the variances of the kinetic constants are given by the equations:

$$s^2_{E_a} = s^2_{\Delta H^\ddagger} = \frac{\boldsymbol{R}^2 s^2 \sum_i w_i}{\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2}$$

$$s^2_{\log_e A} = \frac{s^2 \sum_i w_i/T_i^2}{\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2},$$

$$s^2_{\Delta S^\ddagger} = \boldsymbol{R}^2 s^2_{\log_e A},$$

$$s^2_{\Delta G^\ddagger} = \frac{\boldsymbol{R}^2 s^2\left(T_0^2 \sum_i w_i/T_i^2 - 2T_0 \sum_i w_i/T_i + \sum_i w_i\right)}{\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2},$$

and

$$s^2 = \frac{1}{n-2}\left\{\sum_i w_i(\log_e k_i)^2 - \frac{\left(\sum_i w_i \log_e k_i\right)^2}{\sum_i w_i} - \frac{\left(\sum_i w_i \sum_i w_i \log_e k_i/T_i - \sum_i w_i \log_e k_i \sum_i w_i/T_i\right)^2}{\sum_i w_i\left[\sum_i w_i \sum_i w_i/T_i^2 - \left(\sum_i w_i/T_i\right)^2\right]}\right\}.$$

The experimental point having the greatest weighted deviation, $|\sqrt{w_i}(\log_e k_i - \log_e k_i')|$ from the regression line, where $\log_e k'_i$ is the fitted value of the dependent variable, is determined. The statistical test to decide if this point can be rejected then follows exactly the procedure outlined above in the computation of K_M and V. The programme is looped so that if a point is rejected a new set of kinetic constants is punched out.

The above programmes have been extensively and usefully used by us in kinetic studies of enzyme-catalysed and other reactions.

DEPARTMENTS OF BIOCHEMISTRY AND APPLIED MATHEMATICS,
QUEEN'S UNIVERSITY, BELFAST.

[*Received, August 1st,* 1962.]

---

[11] Curragh and Elmore, *J.*, 1962, 2948; Baines and Elmore, *Bull. Soc. Chim. biol.*, 1960, **42**, 1305 and unpublished work.