Biochem. J. (1974) **139**, 721–730
Printed in Great Britain

Statistical Considerations in the Estimation of Enzyme Kinetic Parameters by the Direct Linear Plot and Other Methods

By ATHEL CORNISH-BOWDEN*
Department of Biochemistry, University of Birmingham, P.O. Box 363, *Birmingham B15 2TT, U.K.*
and ROBERT EISENTHAL
Biochemistry Department, 4-West, University of Bath, Claverton Down, Bath BA2 7AY, U.K.

(*Received* 19 *December* 1973)

The statistical implications of the direct linear plot for enzyme kinetic data, described in the preceding paper (Eisenthal & Cornish-Bowden, 1974), are discussed for the case of the Michaelis–Menten equation. The plot is shown to lead directly to non-parametric confidence limits for the kinetic parameters, V and K_m, which depend on far less sweeping assumptions about the nature of experimental error than those implicit in the method of least squares. Median estimates of V and K_m can also be defined, which are shown to be more robust than the least-squares estimates in a wide variety of experimental situations.

The problem of how best to fit experimental observations to the Michaelis–Menten equation is an old one. Michaelis & Menten (1913) estimated the maximum velocity, V, and the Michaelis constant, K_m, from a plot of the observed velocity, v, against log s, the logarithm of the substrate concentration, but most workers have preferred to obtain a least-squares fit to a linear transformation of the Michaelis–Menten equation. The objections to such methods have been thoroughly discussed elsewhere (Johansen & Lumry, 1961; Wilkinson, 1961; Dowd & Riggs, 1965; Colquhoun, 1971), and we shall not rehearse them here. A much sounder procedure is to minimize the sum of squares of errors SSE, defined by

$$SSE = \sum_{i=1}^{n} w_i \left(v_i - \frac{V s_i}{K_m + s_i} \right)^2 \qquad (1)$$

where s_i and v_i are the ith of n substrate concentrations and velocities respectively and w_i is a weighting factor.

An important but often neglected aspect of eqn. (1) is the necessity to define w_i. It is usual to assume $w_i = 1$ for every observation, an assumption which implies that the errors in all of the velocities are of equal variance. Reich (1970) has cogently argued that this assumption is frequently unjustified in practice, and recommends that the statistical fluctuations in the data be carefully studied before any decision about weights is made. Although this is certainly sound advice, it is not always easy to follow, and there are additional difficulties inherent in the use of eqn. (1), and, indeed, all other least-squares criteria of closeness of fit. These stem from the fact that to demonstrate that a least-squares criterion is valid it is necessary to assume (i) that the errors are normally distributed; (ii) that the independent variable (s in the case of the Michaelis–Menten equation) is known exactly; and (iii) that the correct weights are known. In practice, it is unlikely that these assumptions can be justifiably maintained. Some of them may approximate to the truth, but in most experiments little information is available about any of them. Tests of normality exist, but they require many more observations than are generally made in enzyme kinetic experiments. Even when ample data are available it is most unusual for any tests of normality to be applied. Again, although there may be some experiments where it is reasonable to assume that one variable is largely free from error, there are others where such an assumption is manifestly absurd, as in the determination of dissociation constants from the Scatchard plot (Scatchard, 1949), in which both variables are calculated from the same observations. Finally, there is rarely sufficient information to permit a definitive assignment of weights, and any such assignment must depend on the system studied and the method used.

The main deviation from normality that seems to occur in enzyme kinetics in practice is a high incidence of outliers, i.e. observations with a much higher error than expected from the distribution of the majority of errors. An observation can sometimes be identified as an outlier and rejected from the analysis on independent evidence, e.g. because it was obtained under different experimental conditions from the other observations. More often there is no reason to reject it other than its failure to lie close to the fitted line. In such a case it may still be rejected after an appropriate statistical test is carried out (see, e.g., Bliss, 1967). However, many experimenters are reluctant to reject data on internal evidence alone, since this apparently introduces an unwelcome subjective

* To whom correspondence should be addressed.

element into the analysis. In any case, such a rejection requires a tacit admission that the experimental errors are not normally distributed, and so discredits the least-squares analysis. One may also note that rules for rejecting outliers have been received with very little enthusiasm by leading statisticians. For example, Anscombe (1960) states that 'all published proposals for rejection criteria, based on any kind of mathematical reasoning, from Pierce's (1852) onwards, have an unexplained starting point or objective, presented as though it were the only obvious one and in fact utterly obscure.' The problem is made particularly acute by the fact that the least-squares criterion is highly sensitive to outliers, and often produces a paradoxical situation where the observation recognized to be the worst makes the greatest contribution to the estimates of the parameters.

Distributions of errors with more numerous outliers than expected for a normal distribution are called leptokurtic or long-tailed distributions. They are by no means confined to enzyme kinetics, and Tukey & McLaughlin (1963) have suggested that the 'normal' distribution is actually so rare that it might more instructively be termed the 'pathological' distribution. They state that 'the typical distribution of errors and fluctuations has a shape whose tails are longer than that of a Gaussian distribution'.

Classical (or 'parametric') statistical methods rest heavily on assumptions about the distribution of errors, and recognition of the unreliable nature of these has led to the development of non-parametric statistics, in which as few assumptions as possible are made. The principal one that remains after discarding normality, uniform variance etc. is that the error in any observation is as likely to be positive as to be negative. Even this may of course be false in practice, but it is far less sweeping than the assumptions of parametric methods. If the classical assumptions are true, then parametric methods are appropriate, and provide more efficient tests than the corresponding non-parametric methods. However, this is scarcely a real disadvantage of non-parametric methods, since there is rarely any clear evidence of the validity of the classical assumptions.

Non-parametric methods have been most widely used in the analysis of multiple observations of a single variable, such as the effectiveness of a drug, where the distribution of results is unlikely to be normal, or where the quantification of results is somewhat arbitrary. In such cases the sample median is regarded as a more reliable estimate of the population (or 'average') value than the sample mean (Edgeworth, 1887). Many years ago, Edgeworth (1888) showed that the concept of a median could usefully be generalized into two or more dimensions, and that it could be applied to linear-regression problems. But his methods have not been widely used, and there appears to have been no previous attempt to apply non-parametric methods to the Michaelis–Menten equation, or indeed to any other problem in enzyme kinetics. The purpose of this paper is to show that the direct linear plot described in the preceding paper (Eisenthal & Cornish-Bowden, 1974) can be used to define rigorous non-parametric confidence limits for V and K_m in the Michaelis–Menten equation, and to obtain median estimates of V and K_m. These estimates prove to be almost as reliable as those derived from the least-squares criterion when the classical assumptions are true, and considerably more reliable when they are not.

Theoretical

For any observation (s_i, v_i), the Michaelis–Menten equation

$$v_i = \frac{Vs_i}{K_m + s_i} \tag{2}$$

can be rearranged to show that all values of V and K_m which agree exactly with the observation are related according to

$$V = v_i + \frac{v_i}{s_i} K_m \tag{3}$$

i.e. they lie on a straight line through the points $(-s_i, 0)$ and $(0, v_i)$. But if v_i and s_i are subject to error, eqn. (2) should be written

$$v_i = \frac{\mathscr{V} s_i}{\mathscr{K}_m + s_i} + \varepsilon_i \tag{4}$$

where $\mathscr{V}$ and $\mathscr{K}_m$ are the true (but unknown) values of V and K_m and ε_i is the difference between the observed and true velocities. Note that this equation does not imply that error is confined to v_i, since, if s_i is subject to error and v_i is exactly correct, v_i still differs from $\mathscr{V} s_i/(\mathscr{K}_m + s_i)$ by an amount which can be represented as ε_i. The same is true if both s_i and v_i are subject to error.

If ε_i is positive, the true parameters $\mathscr{V}$ and $\mathscr{K}_m$ are given by a point which lies below the line defined by eqn. (3). Similarly, if ε_i is negative, $(\mathscr{K}_m, \mathscr{V})$ lies above this line. In the absence of any other information, it is reasonable to suppose that ε_i is as likely to be positive as to be negative. (This assumption is of course included among the profusion of assumptions required by the least-squares treatment; it is unfortunately not possible to dispense with assumptions altogether.) So, if an infinite number of independent observations were made, half of the lines drawn according to eqn. (3) would pass below $(\mathscr{K}_m, \mathscr{V})$. In a finite experiment of n observations, n lines can be drawn according to eqn. (3), which intersect to divide the graph into $\frac{1}{2}(n^2+n+2)$ different regions, as illustrated in Fig. 1 for an experiment with $n = 5$. Each region corresponds to a different permutation of signs

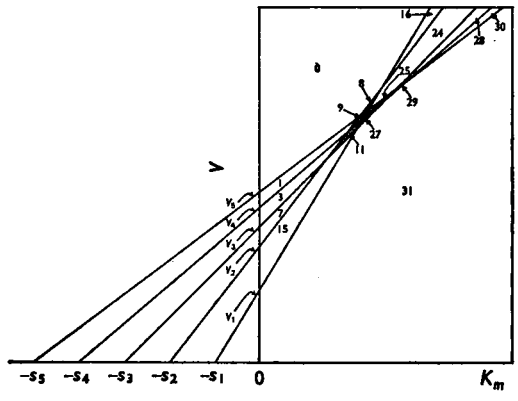

Fig. 1. *Direct linear plot of kinetic results with errors*

Observations of v at five different values of s are plotted as described by Eisenthal & Cornish-Bowden (1974). Each observation (s_i, v_i) is represented by a straight line through the points $(-s_i, 0)$ and $(0, v_i)$. The lines divide the graph into sixteen regions, each of which corresponds to a different permutation of signs among the calculated errors, as explained in the text.

among the ε_i values: for example, in Fig. 1 the region labelled 27 includes all values that $\mathscr{K}_m$ and $\mathscr{V}$ might have such that ε_1, ε_2, ε_4 and ε_5 were positive and ε_3 negative. This permutation of signs is conveniently expressed by the binary number 11011, which is 27 in decimal. The other labels in Fig. 1 can be understood similarly, after conversion into binary. For n observations, 2^n regions, labelled 0 to (2^n-1), can be defined, though only $\frac{1}{2}(n^2+n+2)$ appear in any single experiment. For example, Fig. 1 contains no region 14 (binary 01110); this indicates that for the data illustrated $\mathscr{K}_m$ and $\mathscr{V}$ cannot have values such that ε_1 and ε_5 are negative, and ε_2, ε_3 and ε_4 positive.

If every ε_i value has a median expectation of zero, and all ε_i values are independent, it follows that all possible permutations of signs among the ε_i values are equally likely. Since there are 2^n possible permutations of n signs, each permutation has a probability of 2^{-n}. For $n = 5$, there is only one permutation with all ε_i negative, and thus there is a 1/32 chance that all ε_i are negative. But there are $\binom{5}{2}$, using the common mathematical convention $\binom{n}{m}$ to represent $n!/[m!(n-m)!]$, i.e. 10, permutations containing two positive ε_i, and 10 more with three positive ε_i. Thus there is a 20/32 chance that there will be either two or three positive ε_i values in an experiment of five observations. This result provides a very simple way of defining a 20/32 (i.e. 62.5%) confidence region for $\mathscr{K}_m$ and $\mathscr{V}$ in such an experiment, comprising those regions of the graph that predict two or three positive ε_i values. In Fig. 1, this confidence region would consist of the regions labelled 3, 7, 9, 11, 24, 25 and 28.

Similarly, a 30/32 (i.e. 93.75%) confidence region can be defined by including regions for one or four positive ε_i values. In Fig. 1 this would include every region except 0 and 31.

Confidence regions as just defined are quite rigorous (much more so than those calculated from the sum of squares, since they depend on fewer assumptions) but they are not very convenient, because they extend to infinity, and include estimates of $\mathscr{K}_m$ and $\mathscr{V}$ that common sense would reject as absurd. This objection can be removed by considering the number of 'runs' of positive and negative signs in the series of errors, instead of the total numbers of positive and negative signs. For example, a sequence of signs ++−++ contains three runs, whereas −−+++ contains two. The numbers of runs in random sequences of binary digits obey the binomial distribution (Ising, 1925), and for n digits there are $2\binom{n-1}{m-1}$ permutations with m runs. So the probability that there are m runs in n digits is $2\binom{n-1}{m-1}/2^n$, or $\binom{n-1}{m-1}/2^{n-1}$. For example, with $n = 5$ there are two permutations with one run, eight with two, 12 with three, eight with four and two with five. So there is a 10/32 chance of at least four runs, 22/32 of at least three, etc. In Fig. 1 the regions with at least three runs are shaded, i.e. those labelled 8, 9, 11, 25, 27 and 29, to give a 22/32 (i.e. 68.75%) confidence region. This is just as rigorous as the 20/32 region defined previously, but is much more convenient and useful, since it is enclosed and relatively small in extent.

Confidence limits based on the numbers of runs of signs can be defined in this way for any value of n. If n is at least 9, it is possible to define an enclosed 95% confidence region, and if n is at least 12 it is also possible to define an enclosed 99% confidence region. Table 1 defines 95% and 99% confidence regions for up to 25 observations. If $n = 9$, 10 or 11, it is a simple matter to find the 95% confidence limits, since they enclose all finite regions. For n greater than 11, it is rather more laborious to locate the exact 95% limits, though it is a simple matter to determine whether any given point lies within them.

As well as determining joint confidence regions for $\mathscr{K}_m$ and $\mathscr{V}$, it is also desirable to define best-fit values. As mentioned in the introduction, the non-parametric analogue of the arithmetic mean is the median, which has the important advantage that its reliability is not seriously dependent on the form of the distribution curve. It has the second major advantage that estimation of a population median does not require accurate knowledge of the appropriate weights, i.e. an unweighted median is usually a good approximation to a correctly weighted median (Bowley, 1928). It is appropriate therefore to define median estimates of $\mathscr{K}_m$ and $\mathscr{V}$. But, as Haldane (1948) has pointed out, there are two quite different ways of defining a median, which are equivalent in the univariate case, but non-equivalent in multivariate

Table 1. *95% and 99% confidence regions*

For an experiment of n observations, the Table shows the probability p that there are at least m runs of positive and negative signs among the errors.

	$p \geqslant 95\%$		$p \geqslant 99\%$	
n	m	p	m	p
3	3	25.0*		
4	3	50.0*		
5	3	68.8*		
6	3	81.3*		
7	3	89.1*		
8	3	93.8*		
9	3	96.5		
10	3	98.1		
11	3	98.9		
12	4	96.7	3	99.4
13	4	98.1	3	99.7
14	5	95.4	3	99.8
15	5	97.4	4	99.4
16	5	98.2	4	99.6
17	6	96.2	4	99.8
18	6	97.5	5	99.4
19	7	95.2	5	99.6
20	7	96.8	6	99.0
21	7	97.9	6	99.4
22	8	96.1	6	99.6
23	8	97.4	7	99.2
24	9	95.3	7	99.5
25	9	96.8	7	99.7

* If n is less than 9, there is no enclosed confidence region with $p > 95\%$. In these cases the values of p for $m > 3$ are given.

cases. Thus the median of a univariate population is usually taken to be the estimate that divides the population into two equal portions, but it can equally well be defined by the fact that the sum of absolute (i.e. unsigned) deviations from any point is a minimum if the deviations are measured from the median (Laplace, 1798). Haldane (1948) refers to these as arithmetic and geometric medians respectively, and points out that both definitions provide difficulties when applied to multivariate populations. However, since the difficulties can be removed by precise definition, we shall not discuss them here.

The simplest way of defining median estimates of $\mathscr{V}$ and $\mathscr{K}_m$ for the Michaelis–Menten equation results from the fact that every pair of lines drawn according to eqn. (3) intersect to provide estimates of the parameters given by

$$V_{ij} = \frac{s_i - s_j}{\frac{s_i}{v_i} - \frac{s_j}{v_j}} \qquad (5)$$

$$K_{ij} = \frac{v_j - v_i}{\frac{v_i}{s_i} - \frac{v_j}{s_j}} \qquad (6)$$

for the ith and jth observations. Some of the intersections may be at infinity, if the lines are parallel, or indefinite, if they are coincident, but this is of little consequence in defining medians, provided that such intersections are few in number. For n observations, there are $\frac{1}{2}n(n-1)$ intersections altogether, which provide $\frac{1}{2}n(n-1)$ pairs of estimates. The estimates are not all independent, since there cannot be more than n independent functions of n observations, but there is no bias, and each observation is treated in the same manner. Consequently, the median estimate of $\mathscr{V}$ can be taken as the sample median of all the V_{ij}, and the median estimate of $\mathscr{K}_m$ as the sample median of all the K_{ij}. Since the V_{ij} and K_{ij} are treated as separate univariate samples in this definition the resulting estimates $\hat{V}$ and $\hat{K}_m$ are arithmetic medians in Haldane's terminology.

An alternative approach, corresponding to a special case of Haldane's geometric median, is to define the median estimates of $\mathscr{V}$ and $\mathscr{K}_m$ as those values which minimize the sum of absolute errors in v, *SAE*, defined by

$$SAE = \sum_{i=1}^{n} \left| v_i - \frac{V s_i}{K_m + s_i} \right| \qquad (7)$$

The minimization of the sum of absolute deviations has been recommended by other authors (Davies, 1967; Reich, 1970) as a criterion of closeness of fit, which greatly decreases the effect of bad observations, and is much less dependent on correct weighting than the least-squares criterion. But it has rarely if ever been used in practice because of practical difficulties. The partial derivatives of *SAE*, unlike those of *SSE* [defined by eqn. (1)] with respect to V and K_m, are not continuous functions, but contain saltuses, or jumps, at which the derivatives change abruptly from one value to another. The second partial derivatives are infinite at each saltus, and zero or very small at other points. Since most of the methods used for minimizing *SSE* depend on simple and continuous properties of the derivatives, they cannot be used for minimizing *SAE*. One exception is the robust and elegant method of Nelder & Mead (1965), but although this method can be used successfully it is normally very slow. For the Michaelis–Menten equation, the minimization of *SAE* is actually surprisingly simple, and requires somewhat less computational time than the minimization of *SSE* by the method of Wilkinson (1961). This simplicity results from the fact that the solution must lie on a line plotted according to eqn. (3), and is virtually certain to lie on a second line, i.e. at an intersection. Thus the infinite search in two dimensions that one might expect is decreased to a finite search of $\frac{1}{2}n(n-1)$ possibilities, and in fact far fewer than $\frac{1}{2}n(n-1)$ need be tested if the information obtained during the search is fully used. The validity of this approach depends on the fact, noted by Fisher (1961),

that for linear models of p parameters there must exist estimates of the parameters that minimize SAE and also satisfy p observations exactly. Now, although the Michaelis–Menten equation is non-linear in K_m, the method of Wilkinson (1961) for minimizing SSE depends on the fact that the equation can be accurately represented by a linear approximation over finite ranges of K_m. Thus the approach described for minimizing SAE is likely to be correct in all but a small proportion of experiments. Moreover, in these few experiments the difference in SAE between the true minimum and the putative minimum is likely to be insignificant.

It remains to be considered whether standard errors for the median estimates of $\mathscr{V}$ and $\mathscr{K}_m$ can usefully be defined. In view of the ease with which joint confidence regions can be found, there is very little point in considering the precision of each parameter separately. The very high correlation that always exists between estimates of $\mathscr{V}$ and $\mathscr{K}_m$ (Oliver, 1970) indicates that the error in either parameter is very heavily dependent on the error in the other, and that separation of the two is likely to be more misleading than helpful. Moreover, a standard error is strictly valid only if the statistic in question is normally distributed, and it is dangerous to ignore this point (Kendall & Stuart, 1969); but, as we shall show, estimates of $\mathscr{V}$ and $\mathscr{K}_m$ are generally skewed, even when the errors in v are normally distributed. Nonetheless, if it is felt that these arguments are without merit, and that no analytical method is complete without standard errors, the formulae given by Wilkinson (1961) for the standard errors of the least-squares estimates can be used. Under conditions where these are valid for the least-squares estimates they will also be valid for other unbiased estimates.

Methods

Simulated experiments were carried out on an International Computers Ltd. 1906A computer, by using programs written in FORTRAN.

In all cases the true parameters $\mathscr{V}$ and $\mathscr{K}_m$ were taken as 1.0 by definition. This involved no loss of generality because the numerical values are arbitrarily determined by the units of measurement.

Experimental errors were simulated by the use of a library routine which generated pseudo-random numbers uniformly distributed in the range 0 to 1. These were converted into normally distributed numbers by the method of Box & Muller (1958). The occurrence of outliers was simulated by selecting errors at random from two different normally distributed populations, in the following way: suppose that it is required that each number have a chance p of being drawn from a normal distribution with standard deviation σ_1, and a chance $(1-p)$ of being drawn from a normal distribution with standard deviation σ_2, where $\sigma_2/\sigma_1 = q$. Then, first a uniformly distributed random number r is generated and compared with p. If $r<p$, then a normally distributed number with standard deviation σ_1 is generated; but if $r \geqslant p$, then a normally distributed number with standard deviation σ_2 is generated. In the experiments described, p was usually 0.8, but other values were also tested to a limited extent, to confirm that the value of p was of no great importance in determining the results. q was in the range 1–8, and σ_1 and σ_2 were chosen such that the population standard deviation, given by $\sqrt{p\sigma_1^2+(1-p)\sigma_2^2}$, assumed a convenient value, such as 0.01 or 0.02.

The errors, ε, were introduced into the data in various ways, as follows: (i) simple errors in v, where $v = \mathscr{V}s/(\mathscr{K}_m+s)+\varepsilon$; (ii) relative errors in v, where $v = \mathscr{V}s(1+\varepsilon)/(\mathscr{K}_m+s)$; (iii) simple errors in s, where $v = \mathscr{V}(s-\varepsilon)/(\mathscr{K}_m+s-\varepsilon)$; and (iv) relative errors in s, where $v = \mathscr{V}s/[\mathscr{K}_m(1+\varepsilon)+s]$.

To compare the efficiencies of the various criteria of closeness of fit, 1000 experiments were simulated for each set of assumptions about the nature of experimental error. For each set of data, three pairs of estimates of $\mathscr{V}$ and $\mathscr{K}_m$ were found, and compared to determine which approximated to the true values best. The least-squares estimates were defined as those that minimized the sum of squares, SSE, as defined by eqn. (1) with every $w_i = 1$, and were found by the iterative procedure of Wilkinson (1961). The process was terminated after the fifth iteration, or after an iteration in which the estimate of $\mathscr{K}_m$ was altered by less than 0.01%, whichever was the sooner. Arithmetic-median estimates were defined as the sample medians of the $\frac{1}{2}n(n-1)$ estimates V_{ij} given by eqn. (5) and the $\frac{1}{2}n(n-1)$ estimates K_{ij} given by eqn. (6). An efficient routine for determining sample medians with only partial sorting was written. This routine used about 60% less time than the fastest published sorting routine (Singleton, 1969), and so permitted a considerable gain in speed. It is available on request to A. C.-B. Finally, geometric-median estimates were defined as those that minimized the sum of absolute errors, SAE, as defined by eqn. (7). These were found by searching the pairs of estimates defined by eqns. (5) and (6).

Results

One of the most striking characteristics of non-parametric statistical methods is their insensitivity to outliers, or highly aberrant observations. This is illustrated in Fig. 2, which shows the results of a series of determinations of $\hat{K}_m$ for sets of data in which the velocities at ten substrate concentrations from 0.2 to 2.0 were assigned errors from a normal population with standard deviation 0.005, except for the velocity at $s = 0.6$, which was assigned an error, ε_3, which varied from −0.03 to +0.03. The three types of esti-

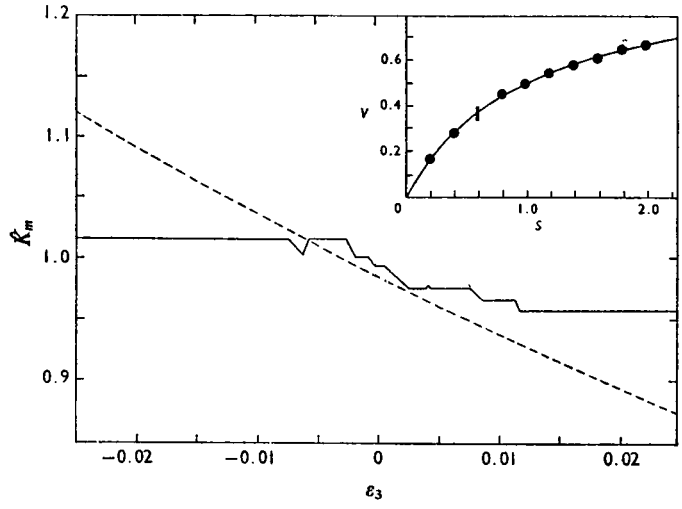

Fig. 2. *Variation of $\hat{K}_m$ with individual errors*

The plots show how the values of $\hat{K}_m$ found by two methods vary with ε_3, the error in the velocity for $s = 0.6$. The errors in the other nine velocities were drawn randomly from a normal population with standard deviation 0.005 and mean zero. The true parameters $\mathscr{V}$ and $\mathscr{K}_m$ were both set equal to 1.0. The broken line shows the value of $\hat{K}_m$ found by the least-squares method; the solid line shows the value of $\hat{K}_m$ found by the arithmetic-median method; the value for the geometric median varied in a similar way to the arithmetic median, but is omitted for the sake of clarity. Inset: A plot of v against s shows the data used in the test. The error bars for v_3 show the range of values used for ε_3. The line is drawn for the true values of the kinetic parameters.

mate behaved very differently in this experiment. The least-squares estimate was relatively insensitive to small errors in a single point, and so tended to be the most reliable estimate when outliers were absent. But in general the least-squares estimate of $\mathscr{K}_m$ deviates without limit from the true value as the error in any single point increases. By contrast, the arithmetic-median estimate responded in a rather irregular way to individual errors, and was very sensitive to small errors. The geometric-median estimate (omitted from Fig. 2, for clarity) was similar, but the variation with small errors was less irregular. In both cases the variation was strictly limited, and in general the median estimates of both parameters became more reliable than the least-squares estimates when outliers are present. The latitude of response of the median estimates to any one observation is determined by the precision of the other observations: in this example, if the other nine observations had been exactly correct, the median estimates of $\mathscr{K}_m$ would not have responded at all to variation of the aberrant observation; but if their standard deviation had been 0.01 instead of 0.05, the variation would have been about twice as great as it was.

The results shown in Fig. 2 are typical of a large number of similar tests. Estimates of $\mathscr{V}$ varied in a similar way, but the variation was less with all three methods, because $\mathscr{V}$ can generally be estimated more precisely than $\mathscr{K}_m$. With all three methods, the variability of both parameters was much greater for aberrant observations near the ends of the range than for observations in the middle. This is of considerable practical importance, because it is often more difficult to measure velocities accurately at the extreme values of s than in the middle of the range. So one may well expect to find outliers in that part of the experiment where they will do most damage to the estimates of $\mathscr{V}$ and $\mathscr{K}_m$. In contradiction of these results, Cleland (1967) has stated that 'it is the velocities obtained at substrate levels around K that are more important than either the higher or lower ones in determining K'. This is not correct in any of the examples that we have examined, and it seems unlikely to have any general validity.

The main series of simulated experiments was designed to compare the efficiencies of the three methods of estimating $\mathscr{V}$ and $\mathscr{K}_m$ over a wide range of assumptions about the nature of experimental error. Fortunately it is not necessary to describe the properties of all three methods in detail, because the two types of median proved to be so similar on average that it is possible to summarize the results for the geometric-median estimates very briefly. It was found that whatever type of error distribution was assumed they were slightly inferior to the arithmetic-median estimates, and in one case, when the standard deviation of v was assumed to be proportional to v, they were greatly inferior. So in the remainder of this

Table 2. *Results of* 56000 *simulated experiments*

The terminology for type of error is defined in the Methods section.

Type of error	n	$s/\mathscr{K}_m$ range	Standard deviation of ε*	Number of experiments (out of 1000) in which the least-squares estimate of $\mathscr{K}_m$ was closer to $\mathscr{K}_m$ than the arithmetic-median estimate, for various distributions of ε†			
				$q=1$	$q=2$	$q=4$	$q=8$
Simple errors in v	5	0.3–1.5	$0.01\mathscr{V}$	563	554	500	388
	10	0.1–1.0	$0.01\mathscr{V}$	609	571	469	308
	10	0.2–2.0	$0.005\mathscr{V}$	586	545	442	258
	10	0.2–2.0	$0.01\mathscr{V}$	612	502	426	302
	10	0.2–2.0	$0.02\mathscr{V}$	582	534	444	317
	10	1.0–10.0	$0.01\mathscr{V}$	549	527	462	330
	25	0.1–2.5	$0.01\mathscr{V}$	588	536	386	219
Relative errors in v	5	0.3–1.5	0.02	444	405	381	348
	10	0.1–1.0	0.02	371	354	322	239
	10	0.2–2.0	0.02	360	347	316	216
	10	1.0–10.0	0.02	613	562	444	287
	25	0.1–2.5	0.02	327	311	276	164
Simple errors in s	10	0.2–2.0	$0.02\mathscr{K}_m$	486	511	428	321
Relative errors in s	10	0.2–2.0	0.02	503	477	379	282

* Except in lines 3 and 5, the values were chosen so that a point with $s = \mathscr{K}_m$ would have the same standard deviation in each experiment.

† $q = \sigma_2/\sigma_1$, as defined in the Methods section. Each error had an 80% chance of being drawn from a normal population with standard deviation σ_1. The column for $q = 1$ corresponds to normally distributed errors.

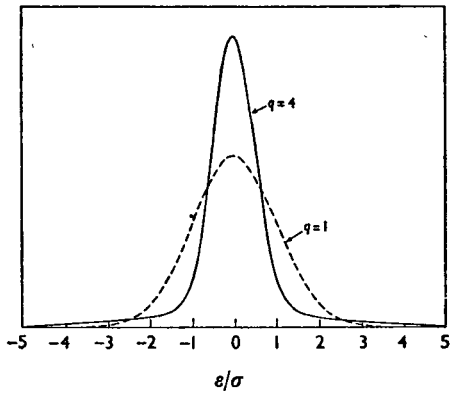

Fig. 3. *Leptokurtic distribution curve*

A distribution curve for $q = 4$, as defined in the Methods section (solid line) is shown with a normal curve (broken line) for comparison.

paper the least-squares estimates will be compared solely with the arithmetic-median estimates, which will be referred to simply as median estimates.

The results of the comparison under various different conditions are shown in Table 2. Although only the results for K_m are shown, the results for V and for V/K_m were almost identical. Two very striking facts are apparent: first, although the least-squares method is superior when all of its assumptions are true, as one would expect, the degree of superiority is remarkably slight. Even at best, it provides better estimates in only about 60% of experiments. Secondly, if any of the assumptions is incorrect, this slight superiority evaporates. Moreover, these conclusions are largely independent of the number of observations and the range of substrate concentrations. The distribution of errors does not have to deviate greatly from the normal curve for the least-squares method to be inferior. This may be judged from Fig. 3, which shows a normal curve of error and a curve with $q = 4$.

One objection may perhaps be made to the way the results are presented in Table 2. It might be argued that, although the median estimates of $\mathscr{K}_m$ might sometimes be better than the least-squares estimates, their distribution about $\mathscr{K}_m$ might be much more scattered. This can be checked by examining the actual distributions of estimates. Fig. 4(*a*) shows cumulative probability (probit) plots for the two estimates of $\mathscr{K}_m$ in a series of 1000 experiments in which all of the least-squares assumptions were true. Both plots show a positive curvature, indicating that in neither case are the estimates of $\mathscr{K}_m$ normally distributed, but that both are positively skewed. Both plots pass very close to the true value of $\mathscr{K}_m$, 1.0, at their midpoints, indicating that both types of estimate are unbiased. Finally, the median estimate shows a somewhat greater scatter, in accordance with

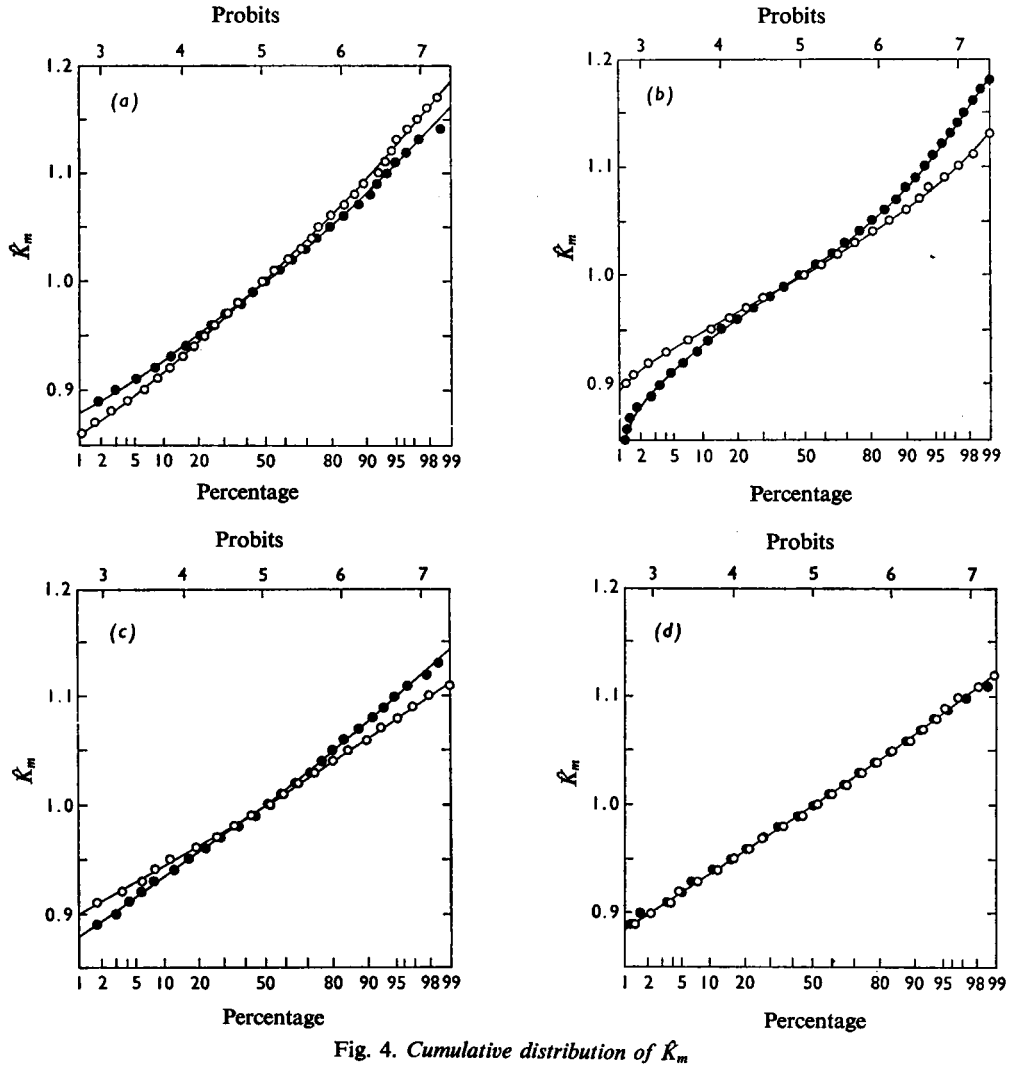

Fig. 4. *Cumulative distribution of* $\hat{K}_m$

Cumulative distribution (probit) plots are shown for the values of $\hat{K}_m$ found by two methods for simulated experiments in which the true value $\mathscr{K}_m$ was 1.0, with ten values of s in the range 0.2–2.0. The abscissa shows the percentage of experiments (out of 1000) in which the value of $\hat{K}_m$ was less than the value shown on the ordinate. The scale of the abscissa is designed so that the points would lie on a straight line if $\hat{K}_m$ were normally distributed. Four different assumptions were made about the nature of experimental error, i.e. (*a*) simple errors in v, normally distributed with standard deviation 0.01; (*b*) simple errors in v, 80% drawn from a normal population with standard deviation 0.005, 20% from a normal population with standard deviation 0.02; (*c*) relative errors in v, normally distributed with standard deviation 0.02; (*d*) simple errors in s, normally distributed with standard deviation 0.02. ●, Values estimated by the least-squares method; ○, values estimated by the arithmetic-median method.

the finding in Table 2 that it came closer to the true value in about 40% of experiments. Similar plots are shown in the remainder of Fig. 4 for some of the other situations considered in Table 2. If the proportion of outliers is increased beyond the normal expectation (Fig. 4*b*), the least-squares estimate deteriorates, but the median estimate improves, not only in relation to the least-squares estimate, but also absolutely. Both curves are leptokurtic, i.e. the relative frequency of very poor estimates is high, but this is much more pronounced for the least-squares estimates. If the wrong weights are used (Fig. 4*c*), the distribution of the median estimates is again much better than that of the least-squares.

Finally, if s rather than v is subject to error (Fig. 4*d*), the distributions are almost identical. All of these results are in excellent agreement with what would be expected from Table 2, and indicate that Table 2 does provide a realistic and informative picture.

The distributions of estimates of $\mathscr{V}$ have also been examined, with similar results, but less variability in all cases. The absolute variability of estimates shown in Fig. 4 is a function of the arbitrary value assumed for the population standard deviation of the errors, and so has no general applicability. However, the relative variability of the two types of estimate should be generally applicable.

Discussion

The results of the simulated experiments indicate that the method that we have described for finding median estimates of $\mathscr{V}$ and $\mathscr{K}_m$ is a very competitive alternative to any of the usual methods. In terms of reliability, it has some very important advantages, as it is far less dependent on assumptions about the nature of experimental error. Even if all of the usual assumptions are correct, the least-squares estimates are better than the median estimates in only a small majority of experiments. But it would be a bold experimenter who would claim that the distribution of errors was precisely normal, that the proper weights were precisely known, and that the values of s were precisely correct. So in practice even the slight advantage for the least-squares method under ideal conditions may be illusory. Certainly it is the experience of many workers in enzyme kinetics that outliers occur disconcertingly often, and it is precisely in such circumstances that the advantages of the median estimates are most pronounced. Similarly, the assumption that all observations should be weighted equally is one that is very difficult to justify in practice. Failure to use the correct weights gave the most striking results in experiments with a low range of s, a fairly common situation in practice, if the substrate is of limited solubility or availability.

It may be argued that the ill effects of outliers and incorrect weighting can be avoided by discarding outliers and using correct weights. Although this may be true, it is unrealistic because it presupposes reliable methods of recognizing outliers and of determining the correct weights. Now it is still common practice to display results in the form of a double-reciprocal plot, even when the best-fitting line is determined independently. On such a plot, a seriously aberrant point at high s may easily pass unnoticed, and so never even be considered for rejection, whereas a perfectly acceptable point at low s may appear to be so aberrant that it is wrongly rejected. This is illustrated in Fig. 5(*a*), which shows a double-reciprocal plot for a simulated experiment in which the velocities were assigned errors from a normal population of standard deviation 0.01, except for the point at $1/s = 0.5$, which was assigned an error of +0.05. The line is a least-squares line calculated by the method of Wilkinson (1961), i.e. it minimizes the sum of squares as defined by eqn. (1) with all $w_i = 1$. It is not at all evident from the plot that the point at $1/s = 0.5$ is highly aberrant, whereas the point at $1/s = 5.0$ appears to be aberrant, but in fact the error in v is 0.018, well within normal expectation. Contrast this with Fig. 5(*b*), which shows the same results plotted according

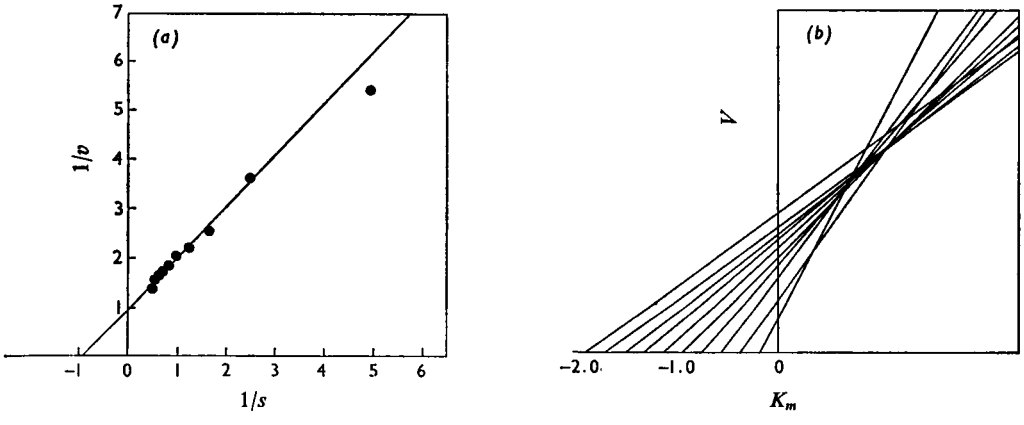

Fig. 5. *Ineffectiveness of the double-reciprocal plot for identifying outliers*

(*a*) Each velocity was assigned an error from a normal population with standard deviation 0.01, except for the one at $1/s = 0.5$, where the error was +0.05. The line minimizes the sum of squares of errors in v, and was calculated by the method of Wilkinson (1961). (*b*) The same results plotted according to the direct linear plot.

to the direct linear plot: here it is quite clear that the observation for $s = 0.2$ is in good agreement with the majority of observations, and that the observation for $s = 2.0$ is severely aberrant.

In addition to its greater reliability under non-ideal conditions, the method proposed in this paper has certain other advantages over the more usual methods. The most obvious of these is that it is the simplest method of estimating $\mathscr{V}$ and $\mathscr{K}_m$ that has ever been proposed, since it requires no calculation at all. The only other method that does not require calculation (apart from a single division by 2) is the estimation of $\mathscr{V}$ and $\mathscr{K}_m$ from a plot of v against s. But this requires the ability to draw a rectangular hyperbola accurately and to judge the location of its asymptotes, and is even less reliable than the double-reciprocal plot method. For small numbers of observations it is very easy to find the medians of the V_{ij} and K_{ij} on a graph, and, although the problem becomes more difficult with larger numbers of observations, it is still much easier to estimate the medians on a graph than it is to estimate the least-squares solutions on any plot. For routine analysis of large amounts of data, computation is more convenient, however, and we have incorporated the methods described in this paper into a computer program. This is written in FORTRAN, and is available on request from A. C.-B.

Another advantage of the method proposed here is that it focuses attention directly on V and K_m, rather than treating them as by-products of a plot whose main effect is to display the relationship between v and s. In practice, the purpose of measuring v at various values of s is almost always to obtain information about $\mathscr{V}$ and $\mathscr{K}_m$. It is rare for the values of v to have any interest except as means to this end. So a plot that displays more information about V and K_m, and their variability, is preferable to one that does not.

Finally it is worth remarking that the theory of the median method is extremely simple, and permits the setting up of joint confidence regions for the kinetic parameters with the use of no more statistical knowledge than is required for an understanding of coin-tossing experiments. This may be contrasted with the recondite nature of the theory of least squares. Indeed the uncritical acceptance of the least-squares method by many scientists owes far more to its mystique than to any real belief in its premises.

Although we have been concerned in this paper exclusively with the Michaelis–Menten equation, there does not seem *a priori* to be any reason why a similar approach should not profitably be made to other problems of interest to the biochemist.

References

Anscombe, F. J. (1960) *Technometrics* **2**, 123–147

Bliss, C. I. (1967) *Statistics in Biology*, vol. 1, pp. 442–444, McGraw–Hill, New York

Bowley, A. L. (1928) *F. Y. Edgeworth's Contributions to Mathematical Statistics*, p. 101, Royal Statistical Society, London

Box, G. E. P. & Muller, M. E. (1958) *Ann. Math. Statist.* **29**, 610–611

Cleland, W. W. (1967) *Advan. Enzymol.* **29**, 1–32

Colquhoun, D. (1971) *Lectures on Biostatistics*, pp. 257–272, Clarendon Press, Oxford

Davies, M. (1967) *J. Roy. Statist. Soc. B* **29**, 101–109

Dowd, J. E. & Riggs, D. S. (1965) *J. Biol. Chem.* **240**, 863–869

Edgeworth, F. Y. (1887) *Phil. Mag. (Ser. 5)* **24**, 268–271

Edgeworth, F. Y. (1888) *Hermathena* **6**, 279–285

Eisenthal, R. & Cornish-Bowden, A. (1974) *Biochem. J.* **139**, 715–720

Fisher, W. D. (1961) *J. Amer. Statist. Ass.* **56**, 359–362

Haldane, J. B. S. (1948) *Biometrika* **35**, 414–415

Ising, E. (1925) *Z. Phys.* **31**, 253–258

Johansen, G. & Lumry, R. (1961) *C. R. Trav. Lab. Carlsberg* **32**, 185–214

Kendall, M. G. & Stuart, A. (1969) *The Advanced Theory of Statistics*, vol. 1, pp. 228–245, Griffin, London

Laplace, P. S. (1798) *Mécanique Céleste*, book 3, chapter 5, section 40, Crapelet, Paris (*œvres Complètes de Laplace*, vol. 2, pp. 143–147, Gauthier-Villars Paris, 1878)

Michaelis, L. & Menten, M. L. (1913) *Biochem. Z.* **49**, 333–369

Nelder, J. A. & Mead, R. (1965) *Comput. J.* **7**, 308–313

Oliver, F. R. (1970) *J. Roy. Statist. Soc. C* **19**, 269–273

Pierce, B. (1852) *Astron. J.* **2**, 161–163

Reich, J. G. (1970) *FEBS Lett.* **9**, 245–251

Scatchard, G. (1949) *Ann. N.Y. Acad. Sci.* **51**, 660–672

Singleton, R. (1969) *Commun. ACM (Ass. Comput. Mach.)* **12**, 185–187

Tukey, J. W. & McLaughlin, D. H. (1963) *Sankhya Ser. A* **25**, 331–352

Wilkinson, G. N. (1961) *Biochem. J.* **80**, 324–332