Biophysical Chemistry 26 (1987) 1–8
Elsevier

BPC 01116

Conditions when statistical tests for model discrimination have high power

Some examples from pharmacokinetics, ligand binding, transient and steady-state enzyme kinetics

William G. Bardsley [a], Philip B. McGinlay [a] and Manuel G. Roig [b]

[a] *Department of Obstetrics and Gynaecology, University of Manchester, at St. Mary's Hospital, Whitworth Park, Manchester M13 0JH, U.K. and* [b] *Universidad de Salamanca, Departamento de Fisico Quimica, Aplicada y Tecnicas Instrumentales, Facultad de Farmacia, Salamanca, Spain*

Received 5 August 1986
Accepted 26 September 1986

Statistical test; Model discrimination; Nonlinear regression; Enzyme kinetics; Ligand binding; Pharmacokinetics

The situation where data pairs x_i, y_i are actually generated by a true model, $f(\vartheta, x)$, but erroneously fitted by a deficient model, $g(\varphi, x)$, is explored. A function, $Q(\vartheta)$, is described which is the average squared distance between $f(\vartheta, x)$ and the best-fit false model, $g(\varphi, x)$. For a range of x covering 5–95% 'saturation' of $f(\vartheta, x)$, $Q(\vartheta)$ is calculated numerically for sums of exponentials, binding functions and rational functions. In each case, the region of ϑ when the second model in the series can be reliably differentiated from the first by statistical tests is described.

1. Introduction

There are numerous occasions when experimental workers assume that an appropriate deterministic equation to fit data is one from a family $f_1, f_2, \ldots, f_n$. These models are then fitted sequentially and parameter redundancy estimated statistically. A good example would be the use of an F test [1] to estimate the contribution of possible predictor variables in multi-linear regression, or perhaps to fit a polynomial up to the highest statistically significant degree. The statistical power of such procedures could be calculated as a function of the data variance matrix, the design matrix and the true parameter vector for linear models. However, in life sciences, it is often appropriate to fit non-linear models to data and then it is not so clear how to proceed analytically.

The general theory of data fitting and model discrimination is well developed [2,3] and attention has been paid to design criteria for parameter estimates with the common non-linear models used in life sciences [4]. However, apart from Monte Carlo simulations [5] and some analytical results with exponential functions [6], the conditions imposed on parameters for correct model discrimination have not been explored for biochemical models.

In this paper we examine the first two models, f_1, f_2 say, of some series extensively used in protein denaturation, pharmacokinetics, binding isotherms, dose-response curves and enzyme kinetics and establish regions of parameter space where statistical tests can be expected to have high power.

2. General theory

Suppose the correct model is $f(\vartheta, x)$, but a deficient model, $g(\varphi, x)$, is fitted to data pairs

x_i, y_i to provide the best-fit parameter vector, $\hat{\varphi}$. Clearly

$$y_i = f(\vartheta, x_i) + \varepsilon_i$$
$$= g(\varphi, x_i) + \{f(\vartheta, x_i) - g(\varphi, x_i)\} + \epsilon_i$$

and so the estimated residuals would be

$$r_i = y_i - f(\hat{\vartheta}, x_i)$$

when fitting the correct model, or

$$r_i = y_i - g(\hat{\varphi}, x_i)$$
$$= f(\vartheta, x_i) - g(\hat{\varphi}, x_i) + \varepsilon_i$$

when fitting the wrong model. Evidently model discrimination is favoured when the model error, $f(\vartheta, x_i) - g(\hat{\varphi}, x_i)$, is large compared to the experimental error, ε_i. Proceeding to the limit of high-quality, dense data, we can claim that successful model discrimination requires a large value of the function $Q(\vartheta)$ defined as

$$F(\vartheta, \varphi) = \frac{1}{(x_2 - x_1)} \int_{x_1}^{x_2} \{f(\vartheta, x) - g(\varphi, x)\}^2 \, dx$$

$$Q(\vartheta) = \min F(\vartheta, \varphi) = F(\vartheta, \hat{\varphi})$$

where the average value of the squared model error has been minimised with respect to φ.

The problems are how to choose $g(\varphi, x)$ given $f(\vartheta, x)$, how to select sensible values for x_2 and x_1 and then how to calculate $Q(\vartheta)$.

The function $Q(\vartheta)$ just described is the average value of the square of the length of the model error between the limits x_1 and x_2. Since the value of $Q(\vartheta)$ depends on the choice of x_1 and x_2 we must adopt a convention. In some kinetic experiments, such as protein denaturation, it is possible to estimate the signal value at time $t = 0$ or $t = \infty$ by independent experiments. Then $x_1 = 0$, $x_2 = \infty$ would be appropriate. In ligand-binding experiments, the zero-saturation point can be estimated from the background and the infinite-saturation point can be estimated by extrapolation, although the actual data usually cover the range of say 10–90% saturation. With steady-state kinetics an additional complication is that substrate inhibition may be present. Sometimes experimental data can be obtained at such sufficiently high substrate values that the velocity is almost reduced to zero but on other occasions it is only possible to extend substrate levels to just beyond, or even just short of a turning point.

In this paper we chose x_1 so that $f(\vartheta, x_1) = 0.95\, f(\vartheta, 0)$ for monotonically decreasing profiles or $f(\vartheta, x_1) = 0.05\, f(\vartheta, \infty)$ for monotonically increasing profiles. The point x_2 was chosen so that $f(\vartheta, x_2) = 0.05 f(\vartheta, 0)$ for monotonically decreasing profiles or $f(\vartheta, x_2) = 0.95 f(\vartheta, \infty)$ for monotonically increasing profiles. For steady-state kinetic curves with a turning point at $x = x_t$, then x_1 was selected so that $f(\vartheta, x_1) = 0.05 f(\vartheta, x_t)$ and x_2 was chosen so that $f(\vartheta, x_2) = 0.5(f(\vartheta, x_t) + f(\vartheta, \infty))$.

Software was written in FORTRAN 77 using subroutines from the Numerical Algorithms Group (NAG) library. Polynomial equations were solved using CO2AEF, transcendental equations by CO5AZF and numerical integrations employed D01AJF. Where φ was 1-dimensional, $\partial F(\vartheta, \varphi)/\partial\varphi = 0$ was solved. Otherwise $\hat{\varphi}$ was obtained using the quasi-Newton optimisation algorithm E04JAF. Contour diagrams for $Q(\vartheta)$ were constructed by graphical interpolation of two-dimensional cross-sections of the three-dimensional surface.

3. A parallel first-order kinetic scheme

Suppose there are two independent chemical reactions taking place such as

$$X_1 \xrightarrow{k_1} Y_1$$

$$X_2 \xrightarrow{k_2} Y_2$$

as in the denaturation of two isoenzymes. Then the signal measured would take the form

$$F(t) = a_1 e^{-k_1 t} + a_2 e^{-k_2 t}.$$

After normalising, a suitable choice covering all possible parameter ranges for $f(\vartheta, x)$ and $g(\varphi, x)$

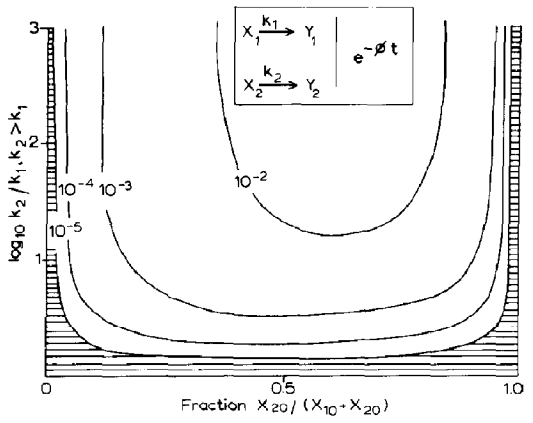

Fig. 1. Q for a parallel double exponential scheme fitted by a single exponential model with fixed amplitude. Model discrimination would fail for choices of parameters in the shaded region. Unless the fraction of either rapidly or slowly decaying component is very small, then rate constant ratios in excess of 2 lead to successful model identification.

would be

$$f(\vartheta, x) = (1-\vartheta_1)e^{-x} + \vartheta_1 e^{-\vartheta_2 x},$$

$$0 \leqslant \vartheta_1 \leqslant 1,\ \vartheta_2 = k_2/k_1 \geqslant 1.$$

$$g(\varphi, x) = e^{-\varphi x},\ \varphi \geqslant 0.$$

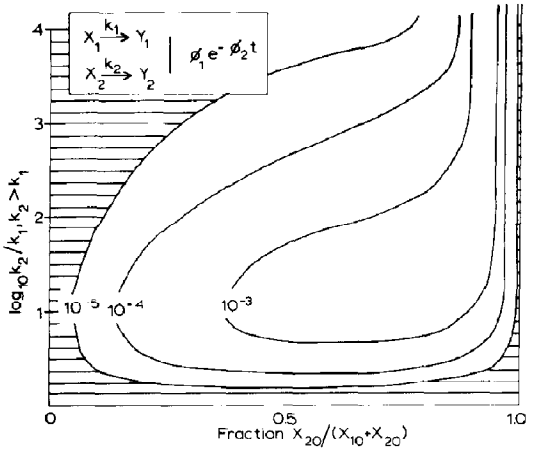

Fig. 2. Q for a parallel double exponential scheme fitted by a single exponential with variable amplitude. Model discrimination is now more difficult than with the case described in fig. 1. For most choices of the fractions of components there is an optimum ratio of k_2/k_1 of about 10 for successful model discrimination. This is because a small component of rapidly decaying exponential near the origin cannot be detected if the amplitude of the best-fit single exponential is allowed to vary.

This choice of $g(\varphi, x)$ presumes that $f(\varphi, 0) = 1$ is known with certainty and ϑ_1 can be thought of as the fraction of rapidly decaying species while ϑ_2 is the rate constant ratio. Normalisation consisted of choosing units so that $f(\vartheta, 0) = 1$ and selecting a time scale $x = k_1 t$.

The contour diagram for $Q(\vartheta)$ is shown in fig. 1. $Q(\vartheta)$ is zero for $\vartheta_2 = 1$ and all ϑ_1 and is also zero for $\vartheta_1 = 0$ and $\vartheta_1 = 1$. Elsewhere $Q(\vartheta)$ increases steadily with increasing ϑ_2 for fixed ϑ_1.

If the point $f(\vartheta, 0) = 1$ is not known with certainty then the experimentalist might be tempted to fit the data using

$$g(\varphi, x) = \varphi_1 e^{-\varphi_2 t},\ \varphi_1 \geqslant 0,\ \varphi_2 \geqslant 0.$$

The contours for $Q(\vartheta)$ in this circumstance are shown in fig. 2. The unsymmetrical nature of the surface and increased difficulty in detecting the correct model are clear.

The profiles in figs. 1 and 2 are rather dependent on the choice of x_1 and x_2 but the qualitative difference between the two results remains. If the single exponential $e^{-\varphi t}$ is fitted to the double exponential curve then model discrimination increases indefinitely as the two rate constants k_1 and k_2 separate. If the single exponential $\varphi_1 e^{-\varphi_2 t}$ with an amplitude factor is fitted, then, because φ_1 is allowed to vary, a small rapidly decaying portion of $f(\vartheta, x)$ near the origin can be overlooked. This is markedly influenced by the choice of x_1. A small slowly decaying component is not so easily overlooked unless a constant, φ_3 say, is added to the single exponential. The effect would then be dependent on the choice for x_2. However, for most reasonable choices for the proportion of fast to slow exponentials and x_1 and x_2, model discrimination reaches a maximum when the rate constants are separated by an order of magnitude as shown in fig. 2.

4. A consecutive compartmental model

Suppose three compartments A, B and C are linked as follows

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C.$$

Examples could be taken from radioactive decay, pharmacokinetics or transient enzyme kinetics with the solution

$$C(t) = (A_0 + B_0 + C_0) - \frac{k_2 A_0 e^{-k_1 t}}{(k_2 - k_1)} - \left\{ B_0 - \frac{k_1 A_0}{(k_2 - k_1)} \right\} e^{-k_2 t}, \qquad k_1 \neq k_2$$

$$= (A_0 + B_0 + C_0) - (kA_0 t + A_0 + B_0) e^{-kt}, \quad k_1 = k_2 = k.$$

Allowing the steps to be reversible leads to a more complicated expression of the same form with eigenvalues replacing the negative rate constants. Assuming $A(0) = I$, $B(0) = C(0) = 0$ and $\vartheta = k_2/k_1$, then normalising gives

$$f(\vartheta, x) = 1 - (\vartheta e^{-t} - e^{-\vartheta t})/(\vartheta - 1), \quad k_1 \neq k_2$$

$$= 1 - (t+1)e^{-t}, \qquad k_1 = k_2.$$

$$g(\varphi, x) = 1 - e^{-\varphi t}.$$

The graph of $Q(\vartheta)$ is in fig. 3 and shows how model discrimination becomes increasingly difficult as the two rate constants separate. An analytical treatment of the previous case with $\vartheta_1 = 0.5$ and in both this and the previous case with $x_1 = 0$, $x_2 = \infty$ has been given [6].

The behaviour shown in fig. 3 is in marked contrast to figs. 1 and 2. When $B_0 = C_0 = 0$ the equation for $C(t)$ becomes symmetrical with respect to k_1 and k_2. If either $k_1 \gg k_2$ or $k_2 \gg k_1$ then one of these steps becomes rate-limiting, the sigmoid character of the $C(t)$ profile is lost, the single exponential curve gives a good fit and the presence of the central compartment, B, cannot be detected.

5. Saturation functions of order 2 and 3

The general saturation function, $y(x)$, of order n can be written with n independent intrinsic binding constants, K_i, using the binding polynomial

$$p(x) = 1 + \binom{n}{1} K_1 x + \binom{n}{2} K_1 K_2 x^2 + \cdots + K_1 K_2 \cdots K_n x^n$$

and

$$y(x) = \frac{1}{n} \frac{\mathrm{dln}\ p(x)}{\mathrm{dln}\ x}.$$

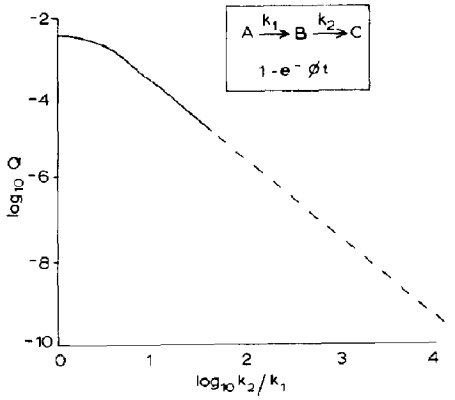

Fig. 3. Q for a sequential two compartment model fitted by a single exponential model. The correct model can be identified unless the rate constant ratio k_2/k_1 (or k_1/k_2) becomes too large, say greater than 100, as indicated by a dashed line. When the ratio is large, the sigmoidal character of $C(t)$ is lost, one step becomes rate-limiting and the $C(t)$ profile cannot be differentiated from a single exponential one.

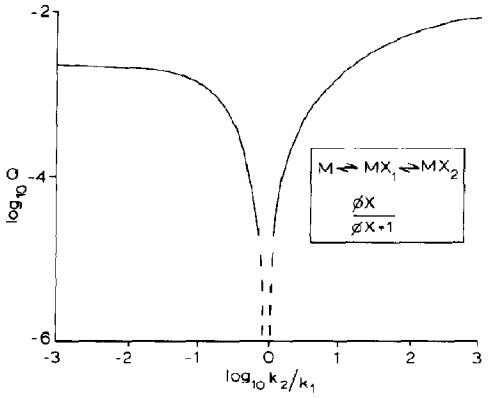

Fig. 4. Q for a two ligand binding site model fitted by hyperbolic binding model. The correct model can be identified unless the ratio K_2/K_1 is in the range of approx. 0.5–2 as indicated by the dashed line. When the problem is normalised as described in the text, the Q value is greater for positive cooperativity ($K_2 > K_1$) than for negative cooperativity ($K_2 < K_1$).

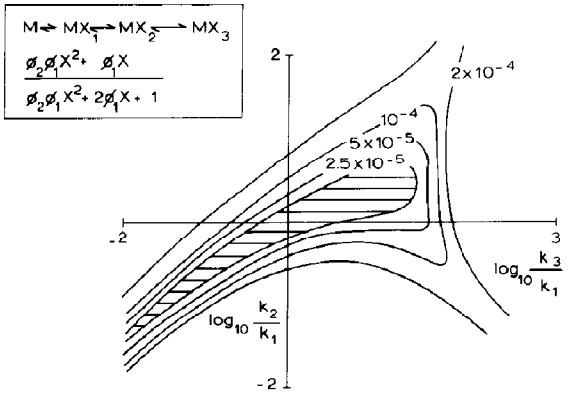

Fig. 5. Q for three ligand binding sites on a macromolecule fitted by a second-order binding model. Model discrimination is not possible for regions of parameter space indicated by the shaded section.

To consider the outcome of fitting order 2 by order 1, the appropriate pair after normalising are

$$f(\vartheta, x) = \frac{\vartheta x^2 + x}{\vartheta x^2 + 2x + 1}, \quad \vartheta = K_2/K_1$$

and

$$g(\varphi, x) = \frac{\varphi x}{\varphi x + 1}$$

leading to the $Q(\vartheta)$ function of fig. 4.

For fitting order 3 by order 2 then $Q(\vartheta)$ of fig. 5 results, using the pair

$$f(\vartheta, x) = \frac{\vartheta_2\vartheta_1 x^3 + 2\vartheta_1 x^2 + x}{\vartheta_2\vartheta_1 x^3 + 3\vartheta_1 x^2 + 3x + 1},$$

$$\vartheta_1 = K_2/K_1, \ \vartheta_2 = K_3/K_1$$

and

$$g(\varphi, x) = \frac{\varphi_2\varphi_1 x^2 + \varphi_1 x}{\varphi_2\varphi_1 x^2 + 2\varphi_1 x + 1}.$$

The precise shapes of some of these figures, especially figs. 4 and 5, depend on the sort of normalising procedure used. The one adopted here can be interpreted as follows: first K_1 is determined, then units of ligand concentration $x = K_1 X$ are chosen. This is equivalent to setting $K_1 = 1$ before fitting the deficient model.

6. High/low-affinity sites

The general n-th order saturation function, $y(x)$, for a mixture of m independent binding sites with association constant, K_i, is obtained from

$$p(x) = (1 + K_1 x)^{n_1}(1 + K_2 x)^{n_2} \ldots (1 + K_m x)^{n_m},$$

$$\sum n_i = n.$$

In fitting dose-response or receptor-binding curves it is usual to consider the special case

$$y(x) = \frac{f_1 K_1 x}{K_1 x + 1} + \frac{f_2 K_2 x}{K_2 x + 1}, \qquad f_1 + f_2 = 1$$

where the f_i are fractions of high/low-affinity sites.

After normalising, an appropriate choice to examine the fitting of order 2 by order 1 would be

$$f(\vartheta, x) = \frac{(1 - \vartheta_1)x}{x + 1} + \frac{\vartheta_1\vartheta_2 x}{\vartheta_2 x + 1},$$

$0 \leqslant \vartheta_1 \leqslant 1, \ \vartheta_2 = K_2/K_1 > 1.$

and

$$g(\varphi, x) = \frac{\varphi x}{\varphi x + 1}$$

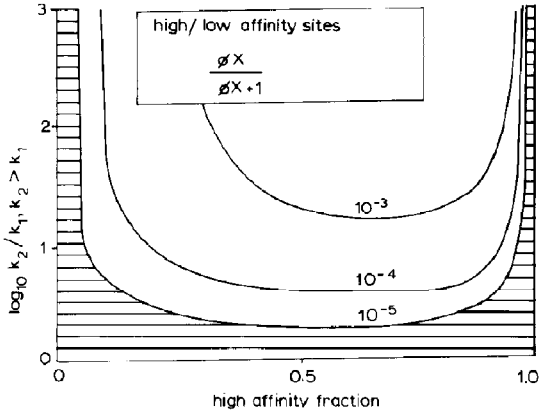

Fig. 6. Q for a mixture of high- and low-affinity sites fitted by a hyperbolic binding model. As long as the ratio of binding constants exceeds 2 and the fraction of either low- or high-affinity sites is not too small, the correct model will be identified. The shaded region illustrates regions of parameter space where model discrimination will fail.

and for this choice $Q(\vartheta)$ can be seen in fig. 6.

To understand the fitting of mixed low/high-affinity sites by the general order 2 saturation function requires careful consideration. The high/low-affinity sites model is a saturation function of order $n_1 + n_2$. The condition for it to be also a saturation function of order 2 is that

$$\frac{f_1 K_1 x}{K_1 x + 1} + \frac{f_2 K_2 x}{K_2 x + 1} = \frac{\vartheta_2 \vartheta_1 x^2 + \vartheta_1 x}{\vartheta_2 \vartheta_1 x^2 + 2\vartheta_1 x + 1},$$

for some ϑ_1, ϑ_2

or

$$K_1(2f_1 - 1) + K_2(2f_2 - 1) = 0$$

requiring $f_1 = f_2 = 0.5$. When $f_1 = 0$, $f_2 = 0$ or $K_1 = K_2$, $Q(\vartheta)$ is identically zero because it is then a saturation function of order 1.

$Q(\vartheta)$ of fig. 7 used the normalised forms

$$f(\vartheta, x) = \frac{(1 - \vartheta_1)x}{x + 1} + \frac{\vartheta_1 \vartheta_2 x}{\vartheta_2 x + 1}$$

and

$$g(\varphi, x) = \frac{\varphi_2 \varphi_1 x^2 + \varphi_1 x}{\varphi_2 \varphi_1 x^2 + 2\varphi_1 x + 1}.$$

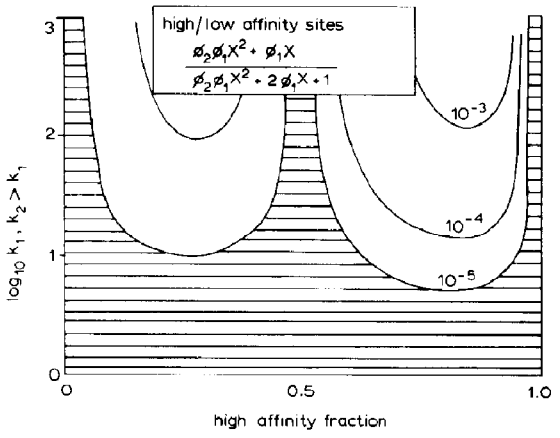

Fig. 7. Q for a mixture of high- and low-affinity sites fitted by a second-order binding model. The shaded area, where model discrimination will fail, is more extensive than in the previous figure. Fractions approaching 0.5 cannot be identified and a higher ratio of K_2/K_1, say 10, is required for successful model discrimination.

Fig. 6 is straightforward and should be compared to fig. 4. Fig. 7 is useful in a slightly different context. It is helpful in the situation where an experimental worker has obtained curved Scatchard plots and wishes to decide whether a general second-order model allowing for positive or negative cooperativity is required or whether a mixture of independent high/low-affinity sites in some proportion is justified. The second-order model has two degrees of freedom, the two binding constants. However, the mixture of high/low-affinity sites is not a second-order saturation function, since it has an additional parameter representing the fraction of high-affinity sites, i.e., the ratio of V_{max1}/V_{max2}, if a sum of two Michaelis-Menten equations is fitted.

7. A second-order steady-state kinetic model

All velocity vs. concentration curves, $v(S)$, resulting from second-order, one-substrate mechanisms can be modelled by a simple sequential mechanism [7] with rate equation

$$v = \frac{2E_0(k_1 K_1 S + k_2 K_1 K_2 S^2)}{1 + 2K_1 S + K_1 K_2 S^2}$$

where the k_i are product release constants and the K_i Michaelis constants, both corrected for statistical factors.

The intrinsic kinetic cooperativity is negative when $k_1 > k_2$ but substrate inhibition requires the stronger constraint $k_1 > 4k_2$. The intrinsic binding cooperativity is positive when $K_2 > K_1$ but a sigmoid $v(S)$ requires the stronger constraint $K_2/K_1 > 2k_1/k_2$. A concave-up double-reciprocal plot requires $(k_1/k_2)^2 + (K_2/K_1 - 1) > 0$ and so although $K_2 > K_1$ is a sufficient condition for this, it is not necessary.

The appropriate formulae are

$$f(\vartheta, x) = \frac{\vartheta_2 x^2 + \vartheta_1 x}{\vartheta_2 x^2 + x + 1}$$

and

$$g(\varphi, x) = \frac{\varphi_1 x}{\varphi_2 + 1}$$

where

$$f(\vartheta, x) = v/2k_2E_0$$

$$\vartheta_1 = k_1/2k_2$$

and

$$\vartheta_2 = K_2/4K_1.$$

The contour diagram for this choice is shown in fig. 8. $Q(\vartheta)$ is only zero when $\vartheta_2 - \vartheta_1(1 - \vartheta_1) = 0$.

We previously pointed out that detection of second-order terms in a 2:2 rational function is particularly easy [5].

In fact it is only for values of k_1, k_2, K_1 and K_2 near the locus $(k_1/k_2)^2 + (K_2/K_1 - 1) = 0$ that model discrimination will fail. Monte Carlo simulations have established that the F test statistics for 1:1 and 2:2 rational functions are approximately F-distributed, parameter estimates are relatively unbiased and the linear approximation covariance matrix gives useful confidence limits. However, the rational function

$$v/E_0 = \sum_{i=1}^{n} \alpha_i S^i / \sum_{i=0}^{n} \beta_i S^i$$

has $2n$ adjustable parameters and, in general, only two parameters, α_1/β_0 and α_n/β_n, can be estimated with any useful precision. So we have the paradox that statistical tests based upon parameter estimation could reliably indicate a kinetic order of 2, 3 or even 4 in extreme cases without yielding useful parameter estimates due to a dense correlation matrix, i.e., parameter redundancy.

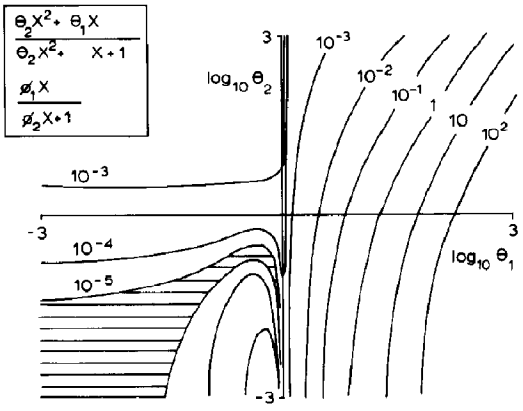

Fig. 8. Q for a second-order rational function fitted by a first-order rational function. Q is only zero when $\vartheta_1(1 - \vartheta_1) - \vartheta_2 = 0$. This limiting curve is surrounded by a strip of parameter space, shaded in the figure and almost entirely bounded by $\vartheta_1 < 1$, $\vartheta_2 < 1$, where the second-order character cannot be detected.

8. Conclusions

A reasonable question to ask is: What must be the magnitude of the parameters in $f(\vartheta, x)$ before it can be distinguished by statistical tests from $g(\varphi, x)$? Obviously the answer to this question will depend upon the number of experimental points, their spacing, the limit points x_1 and x_2, the nature of the error and the type of statistical test used. Often a qualitative answer can be given, usually in the form of: when the parameters are widely separated (or perhaps close together). In this paper we have attempted to give a quantitative answer by presenting calculated contours for the values of $Q(\vartheta)$. Where $Q(\vartheta)$ is large numerically then $f(\vartheta, x)$ can be distinguished from $g(\varphi, x)$ but this statement requires qualification.

First we note that $Q(\vartheta)$ depends upon the limits x_1 and x_2. We have found that the contour diagrams do not change very much for values of x_1 and x_2 that could be realised experimentally, e.g., 5–95%, 10–90% or even 15–85% saturation.

Secondly we have found that the diagrams do not change much in shape if logarithmic coordinates are used, corresponding to a geometrically spaced design. However, the numerical values of $Q(\vartheta)$ are different. In the case of monotonically increasing profiles like saturation functions, $Q(\vartheta)$ is much larger in logarithmic coordinates than in linear ones suggesting the superiority of geometrically spaced rather than arithmetically spaced design points. So calculating $Q(\vartheta)$ in alternative coordinates provides a method for studying alternative designs.

Thirdly we have generated data for these models, fitted $f(\vartheta, x)$ and $g(\varphi, x)$ by nonlinear regression and performed F tests and run tests to discover areas where $f(\vartheta, x)$ can be distinguished from $g(\varphi, x)$. Usually $f(\vartheta, x)$ cannot be differentiated from $g(\varphi, x)$ when values of $Q(\vartheta)$ are less

than 10^{-5}. However, with reasonable numbers of experimental points (10–15, say), sensible ranges of x_1 to x_2 (5–95 to 15–85%, say) and realistic errors (5–7.5% relative error, say), then $f(\vartheta, x)$ can usually be differentiated from $g(\varphi, x)$ in those regions of parameter space where $Q(\vartheta) > 10^{-4}$.

In conclusion we suggest that the shaded regions of figs. 1–8 represent regions of parameter space where the correct model, $f(\vartheta, x)$, cannot be reliably differentiated from the deficient model, $g(\varphi, x)$. Regions of parameter space inside the $Q(\vartheta) = 10^{-4}$ contour represent parameter values where the correct model, $f(\vartheta, x)$, can be distinguished from the deficient model, $g(\varphi, x)$, using for instance the F test with data of the extent and quality typically encountered in the biochemical laboratory.

References

1 N. Draper and H.S. Smith, Applied regression analysis, 2nd edn. (J. Wiley & Sons, New York, 1981).
2 Y. Bard, Nonlinear parameter estimation (Academic Press, New York, 1974).
3 V.V. Fedorov, Theory of optimal experiments (Academic Press, New York, 1972).
4 L. Endrenyi, Kinetic data analysis (Plenum Press, New York, 1981).
5 F.J. Burguillo, A.J. Wright and W.G. Bardsley, Biochem. J. 211 (1983) 23.
6 W.G. Bardsley, P.B. McGinlay and A.J. Wright, Biometrika 73 (1986) 501.
7 W.G. Bardsley, J. Theor. Biol. 104 (1983) 485