A COMPUTER PROGRAM FOR FITTING DATA TO THE MICHAELIS-MENTEN EQUATION [1]

Kenneth R. Hanson, Robert Ling, and Evelyn Havir

Department of Biochemistry
The Connecticut Agricultural Experiment Station
New Haven, Connecticut 06511
and the
Department of Statistics, Yale University
New Haven, Connecticut 06520

Received September 13, 1967

The statistical problems associated with fitting data to the Michaelis-Menten equation, $v = V_{max}[S]/(K_m + [S])$, have been analyzed by Johansen and Lumry (1961), Wilkinson (1961), and most recently by Bliss and James (1966). It is the purpose of this communication to indicate the advance in theory made by Bliss and James, and to describe the output of a computer program following their treatment. A computer program following the method of Wilkinson has been written by Cleland (1963, 1967).

Both Wilkinson, and Bliss and James assume: a) The data are a sample of a population whose mean values conform to the Michaelis-Menten equation. (Plots of initial velocity, v, against v/[S] provide the most sensitive method for recognizing deviations from the equation: Dowd and Riggs, 1965.) b) For all practical purposes [S] is free of error and the error in $\underline{v}$ for a given [S] is normally distributed about a true mean. c) The variance, σ^2 , in v is constant and

---

[1] This work was supported, in part, by N.S.F. Grant GB 2626.

independent of its mean values or of [S]. (Preliminary replicate experiments at several values of [S] may be used to establish that this is approximately true.)

From these assumptions maximum likelihood (ML) estimates for the constants of the equation are calculated by iterative procedures. Bliss and James start with a rough value of K_m, proceed by obtaining the least square fit to the bilinear regression $v = b_1 x + b_2 x_*$ where $x = [S]/(K_m + [S])$, and $x_* = [S]/(K_m + [S])^2$, and then iteratively correct K_m on the basis of the calculated ratio b_2/b_1 until this ratio is less than a pre-set very small value. The result is the linear regression $v = b_1 x$, where b_1 is the ML estimate for V_{max} and the final K_m is the ML estimate for this constant. They then calculate the sample variance, s^2 (an estimate of σ^2), and confidence limits for the ML estimates.[2] The expressions defining such bounds involve Students' _t_ at the desired level of probability for N-2 degrees of freedom (N = the number of observations). For large samples they approach a symmetrical distribution and may be calculated from asymptotic standard errors. These asymptotic standard errors are similar to the standard errors calculated by Wilkinson. Use of the Bliss and James equations thus give more reliable values for the confidence limits. The true bounds may be either over or under estimated by the Wilkinson procedure. If s is small enough or N large enough it will not matter which method is used, but such information is not available to the investigator prior to statistical analysis.

The above calculations are readily performed by digital computer.

---

[2] For example, at the 95% level of confidence there is a 19 in 20 chance that the true value for the constant of the population being sampled lies within the defined limits.

The output on running the program based on the method of Bliss and James, among other things, lists the following items: the progress of the iteration, the input data and the calculated ML values of x so that the scatter of points about the line $v = V_{max}\, x$ may be examined and outliers rejected, $\underline{s}^2$, $\underline{s}$, the ML estimates of V_{max}, $1/V$, K_m, $1/K_m$, V_{max}/K_m, K_m/V_{max}, and the asymptotic standard errors of these constants. Most significantly it lists the upper and lower confidence limit for each of the above forms of the ML constants at several levels of confidence, e.g. 50, 75, 95%. (The appropriate values of Students' $\underline{t}$ must be submitted as part of the data deck.)

The major practical use for this type of information is in examining the effect of some second controlled variate (a second substrate, an inhibitor, temperature, pH etc.) on the above constants with a view to distinguishing between plausible kinetic models. If, e.g., a theoretical curve relating K_m to inhibitor concentration cannot be constructed to lie within the 95% confidence limits for the various K_m points then it is probable that the particular theoretical model in question does not apply (see Wong and Hanes, 1962). In making such decisions it is very undesirable that the limits should be over or under estimated. If a particular model is accepted then ML estimates of desired constants may be computed using the full theoretical equation (Cleland, 1967; Kaiser, 1966).

Copies of the program written in FORTRAN IV are available on request. The symbols and computation follow closely the paper by Bliss and James and it should therefore be relatively easy to understand the program and make any required modifications. The program in its present form has been used with IBM 7094-7040 direct coupled system at the Yale Computer Center.

References

Bliss, C. I., and James, A. T., Biometrics, 22, 573 (1966).

Cleland, W. W., Nature, 198, 463 (1963); Adv. in Enzymology, 29, 1 (1967).

Dowd, J. E., and Riggs, D. S., J. Biol. Chem., 240, 863 (1965).

Johansen, G., and Lumry, R., Compt. Rend. Trav. Lab. Carlsburg, 32, 185 (1961).

Kaiser, A. Dissertation 3896, Eidge. Tec. Hoch., Zurich, Switzerland (1966).

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

Wong, J. T-F, and Hanes, C. S., Can. J. Biochem. Physiol. 40, 763 (1962).