Journal of Pharmacokinetics and Biopharmaceutics, Vol. 6, No. 1, 1978

SCIENTIFIC COMMENTARY

A Comparison of Numerical Integrating Algorithms by Trapezoidal, Lagrange, and Spline Approximation

K. C. Yeh[1] and K. C. Kwan[1]

Received June 9, 1977—Final August 23, 1977

In the trapezoidal method, linear interpolation between data points tends to overestimate or underestimate the area, depending on the concavity of the curve. In some instances, area estimates can be obtained by linear interpolation of logarithmically transformed data. Two alternative algorithms based on known interpolating functions have been implemented for area calculations. In the Lagrange method, the linear interpolations are replaced by cubic polynomial interpolations. In the spline method, the cubic functions are further modified so that the fitted curves are completely smooth. This report describes their computing procedures with numerical examples.

KEY WORDS: numerical integrating algorithms; trapezoidal approximation; Lagrange method; spline method.

INTRODUCTION

It is customary in biopharmaceutics to use a trapezoidal method to calculate areas under the concentration–time curve. The popularity of this method may be attributed to its simplicity both in concept and in execution (1,2). However, in cases where changes in curvature between data points are excessive or there are long intervals between data points, large algorithm errors are known to occur.

To circumvent the curvature problem, two alternative algorithms based on interpolating polynomials have been devised and implemented for area calculations in these laboratories. These polynomials are known as spline functions (3,4) and Lagrange interpolating functions (5). The purpose of this report is to describe computational procedures of these two methods, to compare their solutions along with those obtained by the

[1] Merck Sharp & Dohme Research Laboratories, West Point, Pennsylvania 19486.

0090-466X/78/0200-0079$05.00/0

linear and log trapezoidal methods, and to discuss the relative merit of the methods.

NUMERICAL METHODS

The purpose of a numerical method is to obtain practical solutions which otherwise would have been difficult or impossible to achieve. Because of two contributing factors, the solutions are seldom error free. First, experimental errors in the data are inevitable. These are called input errors. Second, additional errors are incurred when data are processed to produce numerical solutions. These are called algorithm errors. There are two types of algorithm errors (6). The *truncation* error is the difference between the true functional value and that calculated by numerical approximation. The *round off* error results from the fact that only a finite number of digits can actually be retained after each computational step, and any excess digits are lost.

In biopharmaceutics, experimental data such as plasma concentrations are usually recorded at discrete time points. The purpose of using an approximating function in the present case is to connect all data points so that reliable values of areas can be calculated by integration. Although the selection of a particular procedure is somewhat subjective, two basic factors are usually considered: speed and accuracy. When calculations are to be performed manually, easy and simple methods are clearly preferable. However, with the advent of high-speed electronic computers, accuracy becomes the major consideration since computational steps can be programmed and executed swiftly. Thus a method that increases the accuracy of solutions by minimizing algorithm errors should be attempted whenever the procedure is compatible with the limitations of available facilities.

Because of their convenient mathematical properties, polynomials are the most widely used among various curve-fitting approaches. The four procedures described below represent the application of polynomials to area calculations.

Linear Trapezoidal Method

The linear trapezoidal method is the best known numerical integrating method. The functional value y between two adjacent points (t_{i-1}, y_{i-1}) and (t_i, y_i) is approximated by a straight line:

$$y = a + bt \tag{1}$$

where

$$a = (t_i y_{i-1} - y_i t_{i-1})/h_i$$

$$b = (y_i - y_{i-1})/h_i$$

and

$$h_i = t_i - t_{i-1}$$

Integrating equation 1 from t_{i-1} to t_i gives the incremental area in that interval:

$$[\mathrm{AUC}]_{t_{i-1}}^{t_i} = \int_{t_{i-1}}^{t_i} y\, dt = \tfrac{1}{2} h_i (y_i + y_{i-1}) \tag{2}$$

To obtain the cumulative area over the interval $[t_1, t_n]$, where n is the number of data points, the above procedure is repeated for each i, where $i = 2, 3, \ldots, n$. The cumulative area then becomes

$$[\mathrm{AUC}]_{t_1}^{t_n} = [\mathrm{AUC}]_{t_1}^{t_2} + [\mathrm{AUC}]_{t_2}^{t_3} + \cdots + [\mathrm{AUC}]_{t_{n-1}}^{t_n} \tag{3}$$

It is apparent that linear interpolation between data points will tend to underestimate the area when data form a convex curve and to overestimate when the curve is concave. Further, the greater the h_i, the greater would be the error. The magnitude of error would also depend on the oscillatory nature of the curve, or the lack thereof, between data points.

Log Trapezoidal Method

A direct modification of the linear trapezoidal method is the so-called log trapezoidal method. In this modified version, the y values are assumed to vary linearly within each sampling interval on a semilogarithmic scale:

$$\ln(y) = \ln(y_{i-1}) + (t - t_{i-1}) \cdot \ln(y_i / y_{i-1}) / h_i \tag{4}$$

On integration, one obtains

$$[\mathrm{AUC}]_{t_{i-1}}^{t_i} = h_i (y_i - y_{i-1}) / \ln(y_i / y_{i-1}) \tag{5}$$

In pharmacokinetics, equation 5 is most appropriate when applied to data which appear to decline exponentially. Under such condition, the error produced is independent of h_i. However, the method may produce large errors when used in an ascending curve, near a peak, or in a steeply descending polyexponential curve. Furthermore, the method cannot be used if either concentration is zero or if the two values are equal. When they occur, equation 2 can be used as an alternative approach. Despite these limitations, the log trapezoidal method can be used advantageously in specific situations or in combination with a second method to yield optimal solutions.

Lagrange Method

In the Lagrange method, the linear function of equation 1 is replaced by a cubic polynomial:

$$y = a_i + b_i t + c_i t^2 + d_i t^3 \tag{6}$$

To interpolate between $t_{i-1} \le t \le t_i$, the equation is fitted to the nearest four data points (t_{i-2}, y_{i-2}), (t_{i-1}, y_{i-1}), (t_i, y_i), and (t_{i+1}, y_{i+1}). The function is thus forced to pass through all four points. The shape of the fitted curve in the middle interval may not always be linear, but may be parabolic, or sigmoidal with one inflection point. The four coefficients, a_i, b_i, c_i, and d_i, can be obtained by using the Lagrange multiplier formula (5), or by solving the following system of equations:

$$\begin{bmatrix} 1 & t_{i-2} & t_{i-2}^2 & t_{i-2}^3 \\ 1 & t_{i-1} & t_{i-1}^2 & t_{i-1}^3 \\ 1 & t_i & t_i^2 & t_i^3 \\ 1 & t_{i+1} & t_{i+1}^2 & t_{i+1}^3 \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \begin{bmatrix} y_{i-2} \\ y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix} \tag{7}$$

Once the coefficients are obtained, the incremental area on the interval $[t_{i-1}, t_i]$ is calculated by integrating equation 6:

$$[\mathrm{AUC}]_{t_{i-1}}^{t_i} = a_i h_i + \tfrac{1}{2} b_i (t_i^2 - t_{i-1}^2) + \tfrac{1}{3} c_i (t_i^3 - t_{i-1}^3) + \tfrac{1}{4} d_i (t_i^4 - t_{i-1}^4) \tag{8}$$

As an example, Fig. 1 shows the cubic polynomial

$$y = 1 + 7.5t - 5.5t^2 + t^3 \tag{9}$$

passing through (0,1), (1,4), (2,2), and (3,1). The area over the interval [1,2] by equation 8 is 3.17, whereas that by the trapezoidal method is 3.0

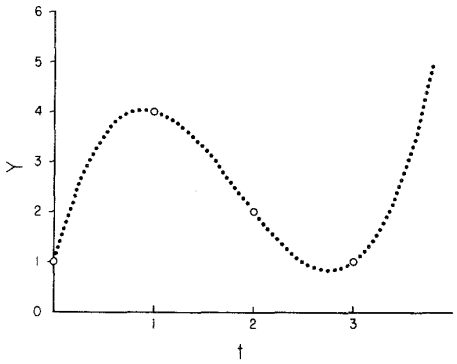

Fig. 1. A cubic polynomial fitted to four data points.

In order to use equation 6, the data must be theoretically smooth over the interval $[t_{i-2}, t_{i+1}]$. Under this condition, the cubic polynomial will usually give a better approximation than a single straight line in the middle interval $[t_{i-1}, t_i]$ because the two contiguous points also contribute to defining the functional behavior over this subinterval. Functional approximations over the intervals $[t_{i-2}, t_{i-1}]$ and $[t_i, t_{i+1}]$ tend to be less reliable in that shaping constraints are one-sided.

Equation 7 can be applied serially for each i, where $i = 3, 4, \ldots, n-1$, but not for the two end intervals, $[t_1, t_2]$ and $[t_{n-1}, t_n]$. To fit these two intervals, the nearest three points are used and fitted with a parabola:

$$y = a_i + b_i t + c_i t^2 \tag{10}$$

The three coefficients are calculated by solving a system of three simultaneous linear equations, analogous to equation 7, and the corresponding areas are obtained by integrating equation 10.

The cumulative area over the entire interval $[t_1, t_n]$ is computed by summation, using equation 3.

Areas calculated by the integration of parabolic equations (equation 10) are subject to greater error than those calculated by the integration of cubic equations (equation 6). In part, this is because approximations are usually better near the middle segment than at the two ends. Since equation 10 is quadratic, it may yield a minimum or a maximum that "should" not have been there. Unwanted oscillations may also occur with cubic equations. Therefore, it is desirable to monitor the interpolated curve over the entire interval. The monitoring can be accomplished by sampling a set of interpolated values within each interval $[t_{i-1}, t_i]$. In so doing, the suitability of the fitted curve can be evaluated intuitively in relation to the prevailing understanding or assumptions concerning the underlying kinetic mechanism.

Thus in the Lagrange method the n experimental points are linked by n-1 smooth curves, with each data point forming the knot of the chain. The fitted curve is piecewise smooth; i.e., it is differentiable within each interval $[t_{i-1}, t_i]$, but not at the data points. In other words, each knot forms a singular point and has no definitive first derivative because the two tangent lines of adjacent cubic functions at t_i may not coincide. This is similar to the trapezoidal method wherein the fitted curve is piecewise linear. However, the curve connected by serial cubic polynomials will look more curvilinear and natural than the polygonal curve formed by the linear trapezoidal method.

The use of serial low-degree (cubic) polynomials, each of which is fitted to a local region, is preferable to the use of a single high-degree polynomial. While a polynomial of degree n-1 or less can be expected to

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.