The Mathematical Background to the Analysis of Growth Curves

(Revised Text July 1971)

H. Goldstein

Institute of Child Health,
University of London,
England.

Let me begin by being provocative! I believe that an enormous amount of time has been spent in attempts to fit growth curves with very little to show for it, except thousands of graphs and a multitude of publications. In the past, I have myself spent time pursuing this new philosophers stone. I never created gold, and nor to my knowledge has anyone else.

It is very difficult to find any successful contributions to biological knowledge that have resulted from the study of theoretical growth curves. In one of the few studies which did succeed; Kidwell and Laird (1970), using the Gompertz equation in mice, discovered that graduation of the data by an ordinary quadratic curve would do the same job as well if not better, and with much less effort - so much for 'realistic' descriptions of growth.

Admittedly, one of the reasons for lack of success may be the difficulty in applying difficult curve fitting procedures to large amounts of data, and it could be argued that with computer programs these problems will disappear and we shall see much more use for these curves. I shall later on give detailed reasons for some doubts on this score, and these may be summarised simply by stating that, as yet, no simple curve has been found to completely satisfactorily describe human growth - even for such a 'well regulated' measurement as height and especially in the adolescent period. There nearly always appears to be some systematic error and one can anyway often fit an ordinary polynomial which will describe growth just as well as a theoretical curve such as the Gompertz. An interesting, if unintentional, admission of the inadequacy of the Gompertz in describing adolescent growth is given by Deming (1957) who fits curves to 48 children missing out those points which 'appeared to deviate markedly from the general trend'!

In direct constrast to the approach of the curve fitters is that of the event estimators; among the best known of whom are Tanner and Marshall and more recently Frisch and Revelle. If we can judge by results in terms of the increase in knowledge and understanding, these workers win hands down. Indeed, the work of Frisch and Revelle in particular I would argue, is a standard against which the models of the curve fitters has to be judged. Their quantification of the relationships between well defined events of growth imposes constraints on any theoretical models which are developed.

Let me now, having declared my prejudice, attempt to outline the background to the fitting of growth curves.

The fitting of growth curves is generally motivated in one of two ways. Either from biological considerations of the growth processes, leading to a mathematical equation, or as an attempt to 'graduate' (i.e. to follow) observed growth in a mathematically convenient way. In this paper I shall not deal with relative growth (allometry) but only with growth in a single measurement.

Examples of biologically motivated growth curves are rare, and often as in Kohn (1948) it is not possible to identify the parameters of the curve with meaningful biological entities. However, Weiss and Kavanu (1957) do attempt to make their parameters meaningful and derive a 14 parameter equation based in part on feedback mechanisms, to explain the growth of chickens, although there seems to have been little attempt to exploit their approach for human growth data. It is also a little difficult to see that it would have any real meaning, especially since the number of parameters is not very much less than the number of measurement points which are usually available.

The second motivation for growth curve fitting has followed two paths. One of these has attempted to incorporate mathematically identifiable events into the equation, such as an upper asymptote for final growth achieved or a parameter representing the age at which maximum growth rate is attained. The success of such equations must be judged both by their ability to fit the data adequately and by their usefulness in estimating important events during growth. The other path has been to fit simple curves which have an obvious interpretation and which can provide useful comparisons between groups. The most common example of this is the fitting of straight lines to estimate average growth rates, making only the assumption that a straight line is a relevant summary of growth in the period considered. Leech and Healy (1959) give an example of the use of this approach. Most of the other approaches by statisticians have also concentrated on graduating growth curves by low degree polynomials, and one of the most useful general models for such curves has been given by Rao (1965). The use of this model makes it possible to test, in a sample of individuals, whether a given degree polynomial fits the data adequately. We have used this model to fit curves of stature for children in the Harpenden growth study and find that between ages 5 and 10 years an adequate fit is obtained with a quadratic curve. After age 10 we find at least a cubic and possibly a higher order polynomial is necessary. This result implies that a different kind of growth curve may be needed. The reason for this is probably related to the well known fact that children enter the adolescent growth phase at different ages and, therefore, are effectively travelling along time scales measured from different ages. Most of the traditional growth curve equations for this period are designed to allow for this fact and some of these will now be considered in detail.

There are three basic requirements for a curve which has to describe the period of growth from the beginning of adolescence to maturity. First it should have an upper asymptote, which is approached as the child reaches maturity. Secondly it should have a single point of inflexion, namely a single time when the growth rate is a maximum. And thirdly there has to be a different time origin for each individual child, which for convenience we may think of as being the age at which peak velocity occurs. These requirements imply that in general we need a growth curve containing at least three unknown parameters. A study of the history of such curves indicates that this is indeed the case. In studies of rat and mice growth data, 3-parameter curves have often been claimed to give good fits (see e.g. Kidwell et al (1969)), and the Gompertz equation in particular has been discussed extensively (Laird et al (1965)).

Any attempt, however, to apply the same techniques to human growth is not necessarily bound to succeed. As an illustration, consider the relationship between size at time of peak velocity and final adult size. For the most commonly used 3-parameter curves, this is fixed by the nature of the curves at between 0.1 and 0.6, the Gompertz, for example, giving a value of 0.37. For human growth data, however, the actual value is about 0.9 (Frisch and Revelle (1969)) and hence we cannot expect such curves to estimate accurately <u>both</u> the point of peak velocity <u>and</u> the upper asymptote. This does not imply that we cannot find a 3-parameter curve which works. One method of searching for such a curve is to look at more general curves with 4 parameters and then to see whether it is possible to find a constant value for one of them so that conditions such as that above, are satisfied. One such type of curve is the general logistic sometimes known as Robertson's curve. The equation of this curve may be written

$$y = \frac{A}{(1 + ae^{-bt+c})^{1/a}}$$

The time of maximum velocity is given by $t = c/b$ and the value of the measurement y at this point is $A/(1+a)^{1/a}$ where A is the upper asymptote. If we use the value of 0.9 for the proportion of final growth attained at this time, then a has a value of about ·34. Setting a equal to this in equation (1) gives a 3-parameter curve which satisfies the condition on the growth at the time of peak velocity. There is still the question of whether it represents growth adequately in other respects, and this underlines a more general point, namely that if we have a special problem - for example, to ensure that a growth curve meets certain specified criteria - this will best be solved by the application of special techniques or, in other words, the derivation of a particular curve for the specified purpose.

It is still possible, however, that a more general curve such as equation (1) may be able to fulfil more than one purpose. As other authors have pointed out (Nelder 1961, Day 1966), estimating 4 parameters instead of 3 means that the estimates are less accurate and less easily identified with biologically meaningful events, particularly with the typically small number of points available in human growth curves. Furthermore since by simply increasing the number of parameters a closer fit is bound to occur, it is reasonable to ask whether a simple cubic polynomial, also a 4-parameter curve, might not also fit the data adequately. This curve does not have a strict upper asymptote, but it does have a point of inflexion, and if it can be made to fit the points closely enough we might well consider using it routinely, especially since it is very easy to fit by ordinary least squares. We have recently been comparing the fit of such a cubic with that of a four parameter logistic curve and a four parameter Gompertz (having an extra parameter for a lower asymptote and also an additional 'part parameter' since a selection of the points to be included must be made), and have not yet been able to choose between them for adequacy of fit. (slide) Whether, and for what, any of these curves is useful will depend on further analysis of our data. It is interesting that a very similar conclusion is reached by Kidwell and Howard (1970) for the growth of mice, where a quadratic and Gompertz were compared and no clear superiority of either was found.

In summary our experience - shared with other practitioners, is that there is as yet no single adequate graduation of growth for all events - a result expected on general theoretical grounds. The best advice is either to find a simple summary of the growth period e.g. by fitting a straight line to estimate growth rate or to estimate things by eye - after all the human brain is the best analogue computer we know of and is particularly good at this kind of problem.

Acknowledgements to J. M. Tanner.

SUMMARY

Two different approaches to the analysis of growth curves are considered. One is aimed at finding simple mathematical equations which describe a section of the growth curve with sufficient accuracy to allow, for example, valid comparisons between different groups of individuals. The other approach attempts to understand the biological processes affecting growth by setting up mathematical 'models' relating these processes to observed growth patterns.

Examples of the first approach are considered in detail.

REFERENCES

1. Day, N. E. (1966). Fitting curves to longitudinal data. Biometrics, 22, 276-291.

2. Frisch, R. E. and Revelle, R. (1969). The height and weight of adolescent boys and girls at the time of peak velocity of growth in height and weight: Longitudinal data. Human Biology, 41, 536-559.

3. Kidwell, J. F. and Chase, H. B. (1969). The inheritance of growth and form in the mouse; II. The Gompertz growth equation. Growth, 33, 339-352.

4. Kidwell, J. F. and Howard, A. (1970). The inheritance of growth and form in the mouse; III. Orthogonal Polynomials. Growth, 34, 87-97.

5. Kohn, P. (1948). Increase in weight and growth of children in the first year of life. Growth, 12, 150-159.

6. Laird, A. K., Sylvanus, A. T. and Barton, A. D. (1965). Dynamics of Normal Growth. Growth, 29, 233-248.

7. Leech, F. B. and Healy, M.J.R. (1959). The Analysis of Experiments on growth rate. Biometrics, 15, 98-106.

8. Marshall, W. A. and Tanner, J.M. (1970) Variations in the Pattern of Pubertal changes in Boys. Arch. Dis. Child. 45 13-23

9. Nelder, J. A. (1961). The fitting of a generalization of the logistic curve. Biometrics, 17, 98-110.

10. Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrics, 52, 447-461.

11. Weiss, P. and Kavanu, J. L. (1957). A model of growth and growth control in mathematical terms. J. Gen. Physiol., 41, 1-47.