1983

G L F R C   M O D E L I N G   G R O U P   S O F T W A R E

SEARCH

A USER'S MANUAL

BY

JACQUES REGNIERE AND D. WENDA BEILHARTZ

Non-Linear Regression Analysis: a Handbook to Commonly Used Equations, and Initial Parameter Estimation.

Jacques Régnière and D. Wenda Beilhartz

Department of the Environment
Canadian Forestry Service
Great Lakes Forest Research Centre
Sault Ste. Marie, Ontario

November 1983

Introduction

There are many reasons to use regression analysis in biology. Often times, the sole purpose is to demonstrate the significant effect of a controlled, or independent, variable (X) on some dependent variable (Y). In many cases, however, the analysis aims at describing the FORM of the relationship between X and Y, or at PREDICTING the values of Y from known values of X. In such cases, the possibility of non-linear responses must be considered. Quite frequently, biological relationships are non-linear; that is, the dependent variable Y is not affected in the same way by changes in variable X over the entire range of possible values of X. For example, consider the effect of fertilization on tree growth. A small amount of fertilizer may double the growth rate of a tree in a poor soil. However, there comes a point where further increases in fertilization will not produce additional growth response, simply because growth is also limited by other factors. Similar examples may be drawn from any biological field.

This paper is intended as a practical guide for those who wish to (or must) deal with non-linear regression. It is not intended as a treatise on this complex subject. Rather, after a brief discussion of the problems and techniques of non-linear regression, we discuss the use of various simple and commonly used equations. The discussion is centered around a computer program which was developed in conjunction with this handbook, and which is intended to simplify the task of performing non-linear regression analysis of data on our computer system. This program, called SEARCH (for grid-search), simplifies the task of obtaining initial estimates of equation parameters which are a prerequisite to the use of the BMDP non-linear regression program PAR. As an added feature, SEARCH also prepares, upon request, complete PAR control-language files. This further simplifies the task of non-linear regression analysis.

Rather than presenting equations and discussing what they can do, we have chosen to classify relationships into 8 frequently encountered SHAPES. We then discuss some of the simple equations which can be used to produce these shapes. While this makes the handbook rather long, it simplifies the task of consulting it. Within each shape, several equations are often proposed. For each equation, we discuss (1) the basic mathematical properties (e.g., parameter limits); (2) the role of parameters in establishing the exact form of a relationship; and (3) how to obtain initial values of the parameters with SEARCH.

Non-linear regression: some general principles

Whenever non-linear regression is to be used to describe a set of data, the analyst is faced with the problem of choosing an appropriate equation. In linear regression, the problem doesn't exist, since there is only one equation available. In non-linear regression, there may be an infinity of equations available to describe a set of data. How does one choose ? Several general principles (which apply to scientific inquiry in general) can be used in making the choice. Here are a few major ones.

A regression equation (when used to describe a relationship) is in fact a mathematically formulated hypothesis or theory (rarely a law). The parsimony principle should, therefore, be applied: it is best to find the simplest possible expression which accounts for a set of facts (data). Thus, one should look for the simplest equation possible to describe the major (non-trivial) features of a relationship. Simplicity, in non-linear regression, is not usually evaluated by the form of an equation, but rather by the number of parameters needed to describe a relationship. We all know that any set of n data points may be described PERFECTLY by a polynomial of order n-1. This, however, is a direct violation of statistical law, and of the parsimony principle.

Equations, like other scientific hypotheses or theories, should account for ALL known facts about a relationship, if they are to be generally applicable. In some cases, it may be possible to deduce some properties of a relationship from already established facts. In such cases, the form of a relationship may be described by a theoretical equation. When available, such equations should be used, since they make use of a much wider base of knowledge than just the data to be descibed. We recommend that pairs of variables be carefully examined, first, for trivial relationships (e.g., surface area vs radius).

Even when a theoretical relationship cannot be deduced, it is possible that some other authors have already developed or used some equation(s) to describe a relationship similar to the one at hand. This type of precedent may be a strong incentive to use one equation over another. In other instances, the same relationship may have been studied in a different way, or over a different range of values of X by someone else. Such information can and should be considered in choosing a regression model.

A little bit of logic may also be welcome. Many biological relationships have obvious logical constraints, which should be taken into account when a valid description is sought (e.g., zero intercept, limited growth rates). One should accomodate these logical arguments in the choice of a model.

Basically, one should look for an equation which will remain valid for any of the possible values of the independent variable, regardless of the range of value covered by the actual data to be analysed. Of course, this is only done whenever possible, and whenever the basic requirements of sound scientific inference and statistical analysis are not violated.

The techniques of non-linear regression

Non-linear regression is a complex analytical process best left to the computer. Nevertheless, it is useful to understand to a certain extent how the computer is programmed to obtain values of the parameters which "best" describe a set of data. First, one must realize that any given equation has its own limitations in terms of flexibility. One may think of an equation as some sort of rubber ruler, which can only be bent to a certain extent, and in a certain way. Thus, the "best" fitting parameter values may still not describe a set of data very well. This question of goodness of fit is of central importance in non-linear regression. Yet, it is not easy to devise rules as to how well an equation should fit data. This is all a question of degree and can usually be answered only in terms of a SPECIFIC example. Nevertheless, it may be useful to consult the chapter on parameter estimation in the easy-reading book by Gold entitled "Mathematical Modeling of Biological Systems" (1976).

Most non-linear regression programs are based on a technique first developed by D.W. Marquardt in 1963, called appropriately Marquardt's algorithm. This technique is based on minimization of the sum of squared residuals (observed minus predicted, squared), RSS for short. The computer is programmed to change parameter values in a logical fashion, compute the RSS, change the parameter values again, and repeat the process until any further reduction in this sum becomes negligible. This is called an iterative process. Several modifications of Marquardt's algorithm have been devised (e.g., Gaussian or Newtonian iteration). All are based on the same principle, but use a different kind of logic to modify parameter values before each iteration. Obviously, since the programs are built to MODIFY parameter estimates, initial values must be available. Much of the difficulty of non-linear regression lies in finding sufficiently accurate initial estimates of the parameters.

Parameter estimates should converge, eventually, to a set of values which minimizes the RSS. This is not necessarily true, however. In linear regression, the normal equations guarantee that a solution (parameter estimates) will be found. In non-linear regression, because the process is approximative, there is no guarantee that a suitable set of parameter values can be found. The first, most obnoxious, problem arises when convergence cannot be achieved. This usually means one of three things: (1) the initial values of the parameters were not good enough, in which case closer estimates must be found; (2) there is a gap in the data in a region of critical importance to the value of one or more of the parameters; or (3) there is no way that the equation chosen can describe the data (in other words, another equation should be used). A second, more intricate, problem has to do with how parameters affect the size of the residuals. In some cases (particularly whith complex equations) the program can get caught in what experts refer to as a local minimum in the RSS. One can visualise the RSS as a surface in the n-dimensional parameter space. In this surface there may be several hills and valleys, one of which is the lowest, and contains the true minimum RSS. However, if initial parameter estimates are not somewhere in the hills immediately around that valley, it is possible that the parameter estimates will converge in some other valley elsewhere, and that the true minimum RSS will not be found. This problem can sometimes be detected by examining the plot of residuals versus the X-variable. To correct this problem, better initial estimates of the parameters must be found.

A final word of caution concerns setting limits to the parameter values. The computer may, in the iterative process, try a wide range of parameter values, particularly if the initial estimates were not so good. In such cases, it is possible, even likely, that some "crazy" values will be tried. This can lead to problems, because there are some values which the computer cannot use. For example: division by zero, the root or logarithm of zero or of a negative number, the exponential of a number above about 87. Whenever possible, it is wise to set limits to the parameter values so that this type of irritating problem does not arise. The SEARCH program assigns upper and lower bounds (in the PAR control files) to the parameters which can lead to this type of problem.

Non-linear regression analysis is not quite as easy to perform as linear regression. Nevertheless, in most applications, problems of the type mentioned above do not occur. With a little practice, the computer programs should become easy to use.

The SEARCH Program

The SEARCH program was designed as a tool to obtain good initial parameter estimates for all the equations described in this handbook. Equations are referred to in the same fashion as in this handbook: by shape-type and equation number. As an additional, very useful feature, SEARCH prepares PAR control-language files upon request (PAR is the BMDP non-linear regression analysis program).

SEARCH obtains initial estimates of the parameters by finding the minimum residual sum of squares (RSS) for a range of values of each parameter of an equation. The equation is chosen by the user, after reference to this handbook. The ranges of each parameter are provided by the user. In this handbook, for each equation, we suggest simple methods to arrive at appropriate ranges. Here is, briefly, how to use SEARCH.

(1) What you need:

1. The data to be analysed must be stored in a computer file. Each observation [(X,Y) pair] should be entered on a separate line. The format is free, but values must be separated by at least one space, or a comma. SEARCH will read this file to the end, or up to 1000 lines, and store these data in active memory. Any valid IAS file name can be used. These files can be created in a number of ways, including the card reader, the PDS file-editor, or any of the data processing programs such as MINITAB, DATAENTRY, or DATATRIEVE.
2. A scatter diagram of the data to answer questions concerning parameter ranges and certain major features of the more complex relationships (such as maximum X, Y).
3. A copy of this handbook, for quick reference.

(2) Running the SEARCH program:

1. In response to a PDS prompt (PDS>), type
   Run DR0:[200,203]SEARCH
2. In response to prompts from SEARCH:
   a. enter the name of the file where the data are stored.
   b. specify an equation, by shape and equation number, as listed in this handbook.
   c. give parameter ranges or answer the simple questions concerning the data (this is where you need a graph of the data), and the number of steps (values) between the minimum and maximum of each parameter.

There are some aspects worth mentioning here. Parameters such as INTERCEPT, upper and lower ASYMPTOTES are frequently easy to estimate visually. Furthermore, final estimates of this type of parameter are very easily arrived at by non-linear regression algorithms. Therefore, a small number of steps (often only 1) can be used for such parameters in the SEARCH program. Other types of parameters, such as multipliers and exponents, are less easily arrived at, and require a larger number of steps (SEARCH will accept up to 25 for each parameter). Whenever a parameter is to be fixed (i.e., not re-estimated by PAR), enter 0 as the number of steps. SEARCH will then instruct PAR (via the control-language file) to FIX this parameter at the value provided by the user. We also point out that the number of calculations required to perform the grid-search in this program is the product of the number of steps requested for each parameter. This number, therefore, can become quite large when more than two or three parameters are being estimated. SEARCH has been programmed to refuse more than 10,000 loops through the data. This limits the number of steps which can be requested to a number which will not be prohibitive.

(3) When to stop SEARCH and get final estimates with PAR:

1. After obtaining parameter estimates for you, SEARCH will prompt for directions on what do no next. You may either try to obtain better estimates (F for further searching), stop (S for stop) or start over again with a different equation (R for restart). Whenever one or more of the initial parameter estimates arrived at by SEARCH lies against either limit of the range specified by the user, a new range encompassing this value should be tried. If this problem does not arise, we recommend that you stop, or perhaps perform a second run with narrower ranges in a complex equation.
2. If you opt to stop, SEARCH will ask you if you require a BMDP PAR file. If you don't, type N and the program will terminate. If you do, type Y and SEARCH will prompt you for the control file name and a title to be

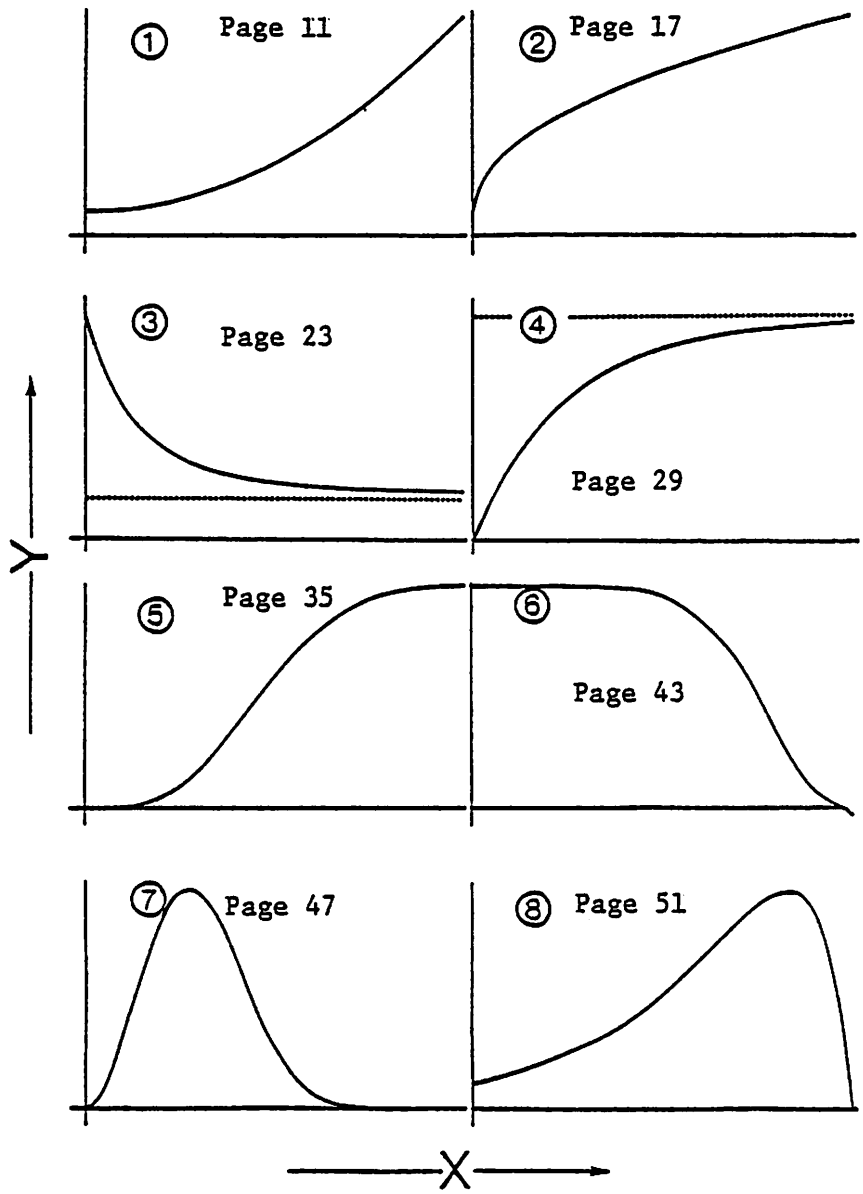
generated in the PAR output. This file contains all the instructions necessary to obtain final estimates of the parameters. There should rarely be any need to modify these control files. If this need occurs, however, consult Biometrics Services for assistance.

(4) Running the BMDP Non-linear regression program PAR:

1. In response to a PDS prompt (PDS>), type
   Run DR0:[200,203]PAR
2. In response to a PAR prompt (PAR>), type either
   a. TI:=control file name (for output directed to your terminal) or
   b. LP0:=control file name (for output directed to the line printer).

Classification of relationships by shape

In the figure below, we have classified commonly encountered types of non-linear relationships into 8 basic SHAPES. To use this handbook, first obtain a scatter diagram of the data to be analysed, consult this figure, and then go directly to the section where the selected shape is discussed.

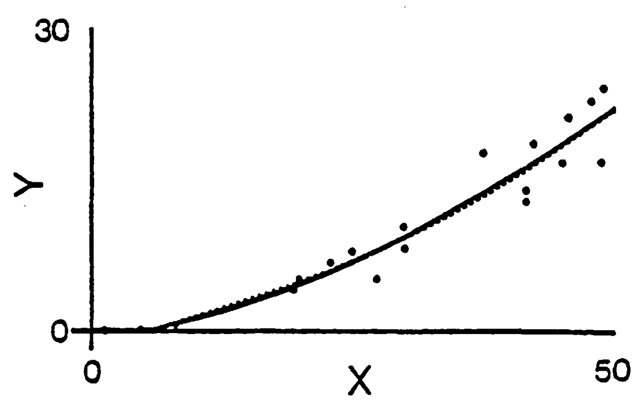


TYPE 1

Type 1 equations are characterized by a gradual increase in the slope of the X-Y relationship as X increases. This type of relationship is frequently encountered in biology. One must realize that Y increases very rapidly to infinity as X becomes larger. Few biological quantities actually do this. However, it may be that it is not Y which is limited to a realistic maximum, but rather X. In such cases, a Type 1 relationship is quite possible.

We propose 2 equations for this type of relationship. Both are simple (3 parameters). Equation [1-1] is nevertheless preferred, because of the intuitive simplicity of its parameters, and because the intercept can easily be fixed. This is not the case for equation [1-2]. An example of data which is suitably described by a Type 1 relationship is illustrated in Fig. 1-0. The two regression lines were obtained by fitting both equations to the data.

Fig. 1-0



TYPE 1 - Equation 1 (The Power Function)

$$Y = P1 + P2\ X^{P3} \quad \text{where } X \geq 0 \qquad [1\text{-}1]$$

1. Parameter limits

   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $1 < P3 < \infty$

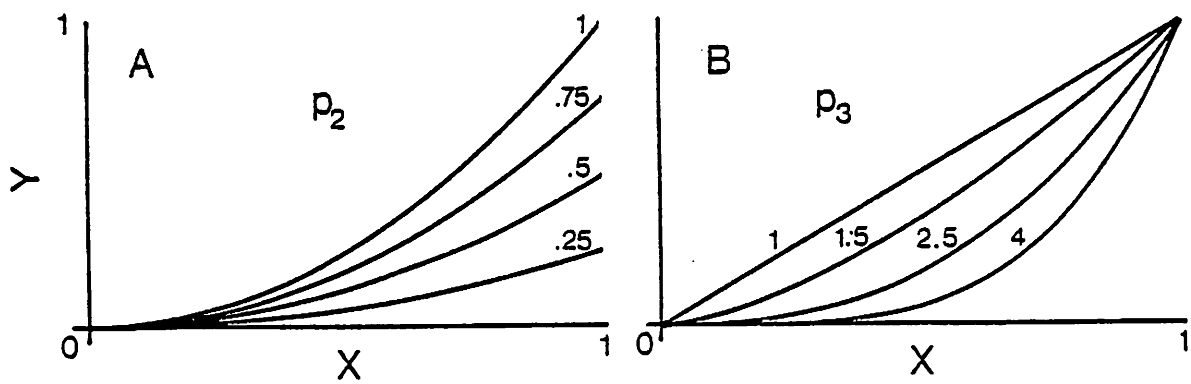
2. Role of the parameters

   P1: The intercept Yo (value of Y when X = 0).

   P2: Scales the values of Y with respect to X. Thus, changes the dimensions of the Y axis, without affecting either curvature or starting point (Fig. 1-1a).

   P3: Changes the curvature of the relationship (Fig. 1-1b). P3 also affects the size of the Y axis, and therefore P2 must be adjusted to compensate for changes in P3. When P3 = 1, the relationship is a straight line (linear regression model). When P3 is large (> 4) the relationship is virtually horizontal at low X values, and soars up as X increases.

Fig. 1-1



3. Initial parameter estimation (see page 7, item c)

A range of values for the parameters of equation [1-1] to be used in conjunction with SEARCH can be obtained in the following manner.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. Because this parameter is quite simple to estimate, we suggest a small number of steps (1 to 10). If an intercept can be deduced from logical arguments, P1 can be fixed by specifying # of steps = 0. In our example (Fig. 1-0), the intercept is obviously close to zero. However, we use the range [0, 0], # of steps = 1, because we want PAR to come up with a more precise estimate.

A range of values for P2 is not easily defined because it is dependent upon the value of P3. However, a rule of thumb can be applied. Since P3 > 1, we can assume that the maximum value of P2 will not exceed the ratio of the maximum Y-value to the maximum X-value (Ymax/Xmax). Obtain this ratio visually (i.e, is the ratio half? two-fold?), and try a range of [0, Ymax/Xmax], with a large number of steps (15 to 25). In our example (Fig. 1-0) the ratio Ymax/Xmax is approximately 1/2. Therefore, the range [0, .5] could be used.

Compare your data with Fig. 1-1b for curvature. Choose a realistic but sufficiently wide range of P3 values. FOLLOW THE RECOMMENDATION OF SEARCH concerning the maximum allowable value of P3. Failure to do so may lead to numbers too large for the computer to handle. In choosing the maximum value for P3, consider that P3 > 10 should rarely occur. In our example (Fig. 1-0), an appropriate range of values of P3 would be [1, 2.5]. We suggest a large number of steps (15 to 25) for this parameter as well.

For our example, SEARCH found a minimum RSS of 98.6 (S = .909), and PAR gave an RSS of 90.7 (S = .917) with final (converged) values of P1 = -.17, P2 = .037, and P3 = 1.64.

TYPE 1 - EQUATION 2 (The Exponential Function)

$$Y = P1 + P2\, e^{P3\, X} \qquad \text{where } X \geq 0 \qquad [1\text{-}2]$$

1. Parameter limits

   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $0 < P3 < \infty$

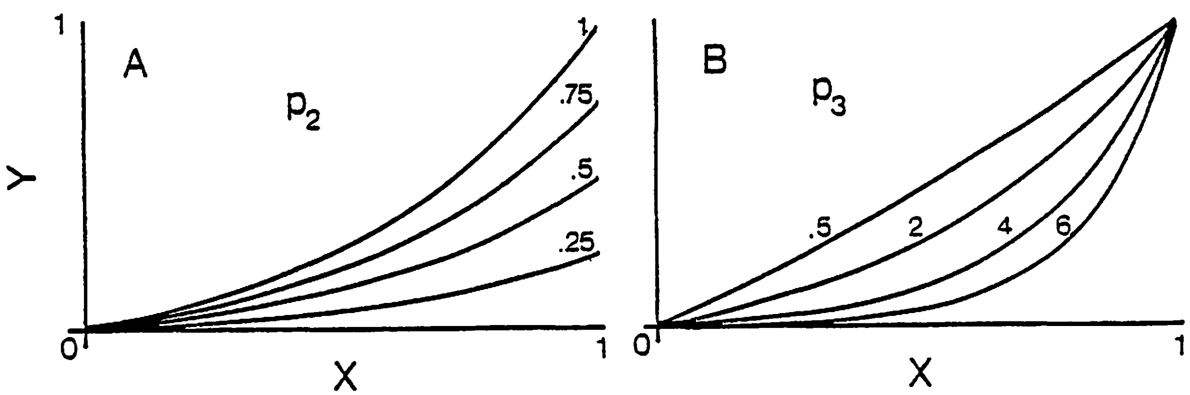
2. Role of the parameters

   P1: Part of the intercept Yo (value of Y when X = 0). We note that Yo = P1 + P2. P1 shifts the whole curve up and down on the Y axis. The intercept, Yo, cannot be readily forced through a fixed point (e.g., zero) with this equation in non-linear regression analysis.

   P2: Scales the values of Y with respect to X. Thus, changes the dimensions of the Y axis, without affecting the curvature (Fig. 1-2a). Also partly determines the intercept.

   P3: Changes the APPARENT curvature (Fig. 1-2b), by scaling the values of X. Determines the portion of an exponential curve which is seen over the range of values of X.

Fig. 1-2



3. Initial Parameter Estimation (see page 7, item c)

A range of values for the parameters of equation [1-2] to be used in conjunction with SEARCH can be obtained in the following manner.

First, compare your data with Fig. 1-2b, and choose a tentative range of values of P3. We note that the maximum X value in Fig. 1-2b is 1. Thus, since P3 SCALES X, a valid range of values for P3 can be obtained by dividing the values obtained from Fig. 1-2b by the maximum X value in your data (Xmax). A large number of steps (15 to 25) is recommended. In our example (Fig. 1-0), we obtained a range of [.5, 3] from Fig. 1-2b, and divided by Xmax = 50 to obtain a range of [.01, .06] for P3.

Parameter P2 depends on the value of P3. We suggest 0 as a convenient minimum. An appropriate maximum value of P2 should not exceed the maximum Y value (Ymax), since it would imply such a low P3 that the curvature of the relationship would be undetectable. Most likely, P2 will be much smaller than Ymax. A large number of steps (15 to 25) is recommended. In our example (Fig. 1-0), we chose a range of [0, 10], which implies a significant curvature.

A range of values of P1 is determined from the range of values of P2, and the approximate value of Yo. Because Yo = P1 + P2, we note that P1 = Yo - P2. Thus the general range [(Yo-maximum value of P2), (Yo-minimum value of P2)] should be used. A large number of steps (15 to 25) is recommended. In our example, since Yo is close to zero, we used the range [-10, 0] for P1.

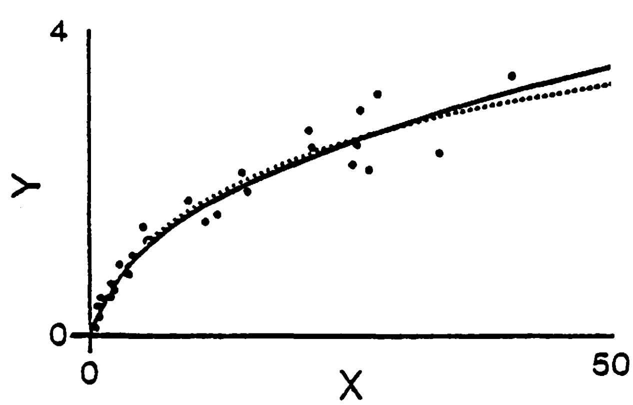
For our example, SEARCH found a minimum RSS of 91.6 (S = .916), and PAR gave an RSS of 91.2 (S = .916) with final (converged) values of P1 = -9.25, P2 = 8.42, and P3 = .027.

TYPE 2

Type 2 equations are characterized by a gradual decrease in the slope of the X-Y relationship as X increases. They resemble closely the "diminishing return" equations in all but one respect: the value of Y does not tend to a maximum, but continues increasing without bounds as X increases. In many cases, Type 2 equations are justified when values of X outside the observed range are unlikely.

We propose 2 equations for this type of relationship. Both are simple (3 parameters). Nevertheless, equation [2-1] is preferred, because it is more widely used in scientific literature. However, equation [2-2] has more intuitive appeal, since it is quite similar to a linear regression between Y and the logarithmic transform of X. An example of data which is suitably described by a Type 2 relationship is illustrated in Fig. 2-0. The two regression lines were obtained by fitting both equations to the data.

Fig. 2-0



TYPE 2 - EQUATION 1 (The Power Function)

$$Y = P1 + P2\ X^{P3} \qquad \text{where } X \geq 0 \qquad [2\text{-}1]$$

1. Parameter limits

   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $0 < P3 < 1$

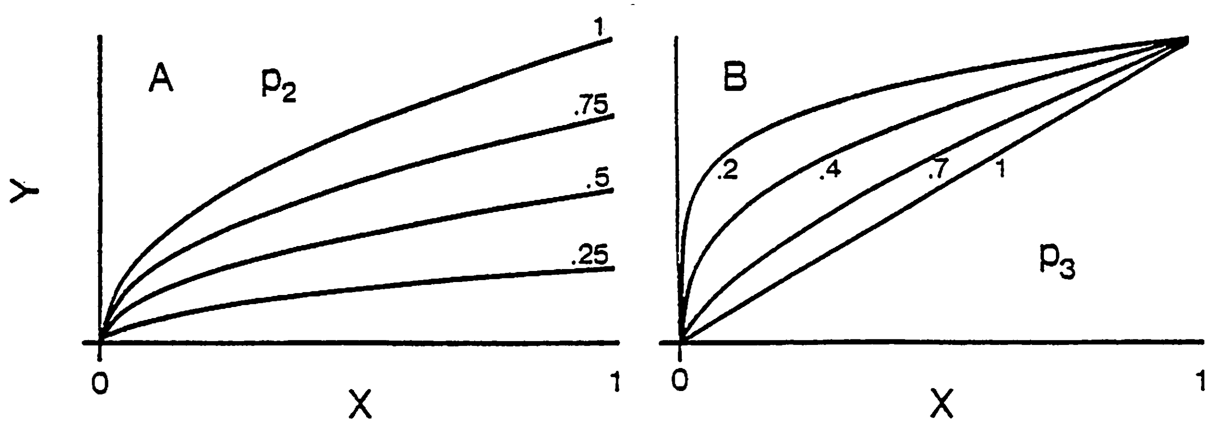
2. Role of the parameters

   P1: The intercept Yo (value of Y when X = 0).

   P2: Scales the values of Y with respect to X. Thus, changes the dimensions of the Y axis, without affecting either curvature or starting point. (Fig. 2-1a).

   P3: Changes the curvature of the relationship (Fig. 2-1b). P3 also affects the size of the Y axis. P2 must be adjusted to compensate for changes in P3. When P3 = 0, there is no relationship (horizontal line). When P3 is close to 1 the relationship is a straight line (linear regression model). With P3 small (but still larger than 0), the curvature is very strong.

Fig. 2-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [2-1] to be used in conjunction with SEARCH can be obtained in the following manner.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. If an intercept can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 2-0), the intercept is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

A range of values for P2 is not easily defined because it is dependent upon the value of P3. However, a rule of thumb can be applied. In the extreme, P3 = 1 (no curvature), and the minimum possible value of P2 is the ratio of maximum Y-value to maximum X-value (Ymax/Xmax). A maximum value for P2 is not as readily found, but one can assume that it should not exceed Ymax, when Xmax > 1. Obtain the Ymax/Xmax ratio visually (i.e, is the ratio half? Two-fold?). We suggest a range of [Ymax/Xmax, Ymax], with a large number of steps for this parameter (15 to 25). In our example (Fig. 2-0) the ratio Ymax/Xmax is approximately 1/10. Since the value of Ymax is about 4, the range [.1, 4] could be used.

Next, compare your data with Fig. 2-1b for curvature. Choose a realistic but sufficiently wide range of P3 values (often, the range [0,1] is convenient). In our example (Fig. 2-0), an appropriate range of values of P3 may be [.2, .7]. We suggest a large number of steps (15 to 25) for this parameter as well.

For our example, SEARCH found a minimum RSS of 1.86 (S = .948), and PAR gave an RSS of 1.55 (S = .956) with final (converged) values of P1 = -.354, P2 = .767, and P3 = .414.

TYPE 2 - EQUATION 2 (The Logarithmic Function)

$$Y = P1 + P2\ Ln(P3\ X + 1) \quad \text{where } X \geq 0 \qquad [2\text{-}2]$$

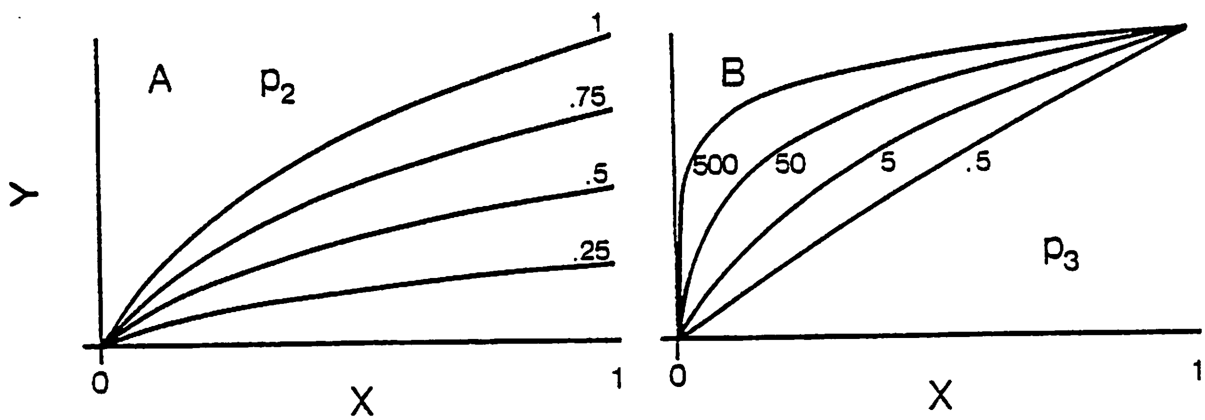
1. Parameter Limits
   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $0 < P3 < \infty$
2. Role of the Parameters

   P1: The intercept Yo (value of Y when X = 0).

   P2: Scales the values of Y with respect to X. Thus, changes the dimensions of the Y axis, without affecting either curvature or starting point (Fig. 2-2a).

   P3: Changes the APPARENT curvature (Fig. 2-2b), by scaling the values of X. Determines the portion of a logarithmic curve which is seen over the range of values of X.

Fig. 2-2



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [2-2] to be used in conjunction with SEARCH can be obtained in the following manner.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. If an intercept can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 2-0), the intercept is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

A range of values for P2 is not easily defined because it is dependent upon the value of P3. However, a rule of thumb can be applied. In the extreme, P3 is close to 0 (no curvature), and the maximum possible value of P2 can be infinite. In practical terms, however, the curvature would be undetectable if P3 < (.5/Ymax), in which case P2 would be roughly 2 x Ymax (where Ymax is the maximum value of Y). For strong curvatures, P2 would be much smaller. A practical range would therefore be [0, Ymax], with a large number of steps (15-25) for initial estimation. For our example (Fig. 2-0), an appropriate range of P2 would be [0, 4].

Next, compare your data with Fig. 2-2b, and choose a tentative range of values of P3. We note that the maximum X value in Fig. 2-2b is 1. Thus, since P3 SCALES X, a valid range of values for P3 can be obtained by dividing the values taken from Fig. 2-2b by the maximum X value in your data (Xmax). A large number of steps (15 to 25) is recommended. In our example (Fig. 2-0), we chose a range of [5, 50] from Fig. 2-2b, and divided by Xmax = 50 to obtain a range of [.1, 1] for P3.

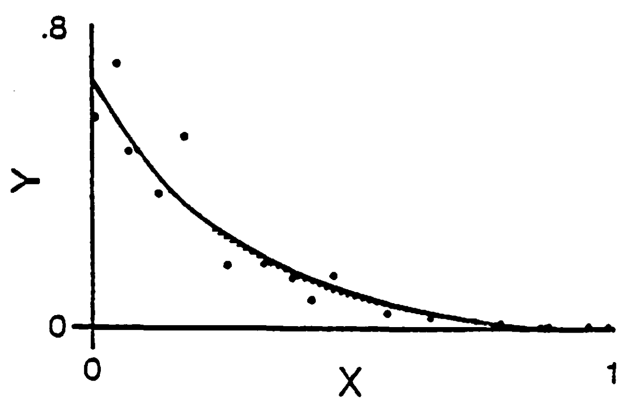
For our example, SEARCH found a minimum RSS of 1.65 (S = .954), and PAR gave an RSS of 1.59 (S = .955) with final (converged) values of P1 = .074, P2 = 1.171, and P3 = .297.

TYPE 3

Type 3 equations are characterized by a gradual decrease of Y as X increases. The rate of decrease of Y is rapid at first, but eventually becomes very small, as Y approaches a minimum value (e.g., Ymin = 0). This value is often called the lower ASYMPTOTE of Y, a value which is never quite reached by Y.

We propose 2 equations for this type of relationship. Although both are fairly simple, we prefer equation [3-1], because it has fewer parameters, and is much easier for PAR to handle. In equation [3-2], it is often necessary to FIX the value of parameter P3 in order to achieve convergence. We suggest trying equation [3-1] first. If the fit is unsatisfactory, then equation [3-2] can be tried. An example of data which is suitably described by a Type 3 equation is illustrated in Fig. 3-0. The two regression lines were obtained by fitting both equations to the data.

Fig. 3-0



TYPE 3 - EQUATION 1 (The Exponential Function)

$$Y = P1 + P2\, e^{P3\, X} \qquad \text{where } X \geq 0 \qquad [3\text{-}1]$$

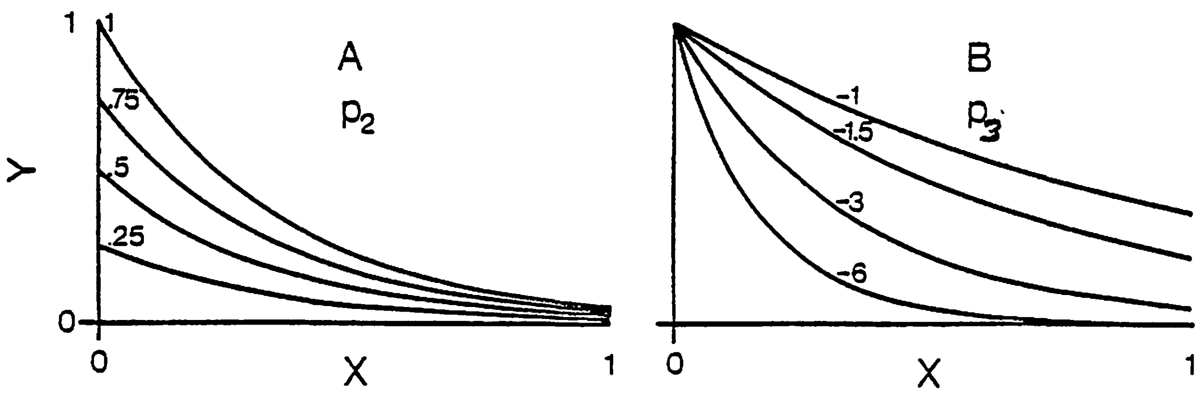
1. Parameter Limits
   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $-\infty < P3 < 0$
2. Role of the Parameters

   P1: Lower asymptote Ymin (value of Y when X->∞).

   P2: Changes the magnitude of Y (Fig. 3-1a). Therefore, determines, along with P1, the intercept Yo (value of Y when X = 0). We note that Yo = P1 + P2. Thus, P2 = Yo - P1.

   P3: Changes the APPARENT curvature of the relationship (Fig. 3-1b), by scaling the values of X. Determines the portion of an exponential curve which is seen over the range of values of X. When P3 = 0, there is no relationship (horizontal line). When P3 is very small (i.e., negatively large), the curvature is strong. With P3 close to zero, the curvature is weak.

Fig. 3-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [3-1] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the high range of X. If a lower asymptote can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 3-0), the lower asymptote is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

Second, estimate a range of values for the intercept (Yo). A suitable range of values of P2 is given by:

[P2min, P2max] = [Yomin-P1max, Yomax-P1min]

In our example (Fig. 3-0), Yo ranges between 0.6 and 0.8. Thus, since P1 is very close to 0, an appropriate range for P2 would be [.6-0, .8-0] = [.6, .8].

Third, compare your data with Fig. 3-1b, and choose a tentative range of values of P3. We note that the maximum X value in Fig. 3-1b is 1. Thus, since P3 SCALES X, a valid range of values for P3 can be obtained by dividing the values obtained from Fig. 3-1b by the maximum X value in your data (Xmax). A large number of steps (15 to 25) is recommended. In our example (Fig. 3-0), we obtained a range of [-6, -3] from Fig. 3-1b (since Xmax = 1 in this case, it was not necessary to alter this range.)

For our example, SEARCH found a minimum RSS of 0.08 (S = .920), and PAR gave an RSS of 0.076 (S = .925) with final (converged) values of P1 = -.034, P2 = .694, and P3 = -3.538.

TYPE 3 - EQUATION 2 (The Power Function)

$$Y = P1 + P2\,(P3\,X + 1)^{P4} \quad \text{where } X \geq 0 \qquad [3\text{-}2]$$

1. Parameter Limits
   - $-\infty < P1 < \infty$
   - $0 < P2 < \infty$
   - $0 < P3 < \infty$
   - $-\infty < P4 < 0$
2. Role of the Parameters

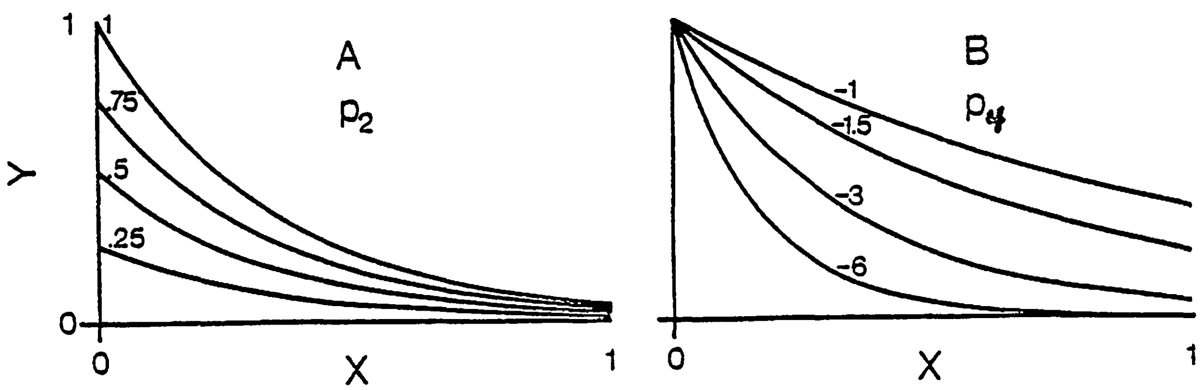
P1: Lower asymptote Ymin (value of Y when X-> ∞).

P2: Changes the magnitude of Y (Fig. 3-2a). Therefore, determines, along with P1, the intercept Yo (value of Y when X = 0). We note that Yo = P1 + P2. Thus, P2 = Yo - P1.

P3: Scales the values of X. This parameter gives additional flexibility to the equation, and is necessary because of the "+ 1" term in equation [3-2].

P4: Changes the curvature of the relationship (Fig. 3-2b). When P4 = 0, there is no relationship (horizontal line). When P4 is very small (i.e., negatively large), the curvature is strong. With P4 large (but still smaller than 0), the curvature is weak.

Fig. 3-2



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [3-2] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the high range of X. If a lower asymptote can be deduced from logical arguments, P1 can be fixed by specifying # of steps = 0. However, in our example (Fig. 3-0), although the lower asymptote is obviously close to zero, we still want PAR to estimate it and thus we use the range [0, 0] with # of steps = 1.

Second, estimate a range of values for the intercept (Yo). A suitable range of values of P2 is given by:

[P2min, P2max] = [Yomin-P1max, Yomax-P1min]

where Yomin is your estimate of the smallest value Yo could be, and Yomax is your estimate of the largest value Yo could be. We suggest a large number of steps (15 to 25) for this parameter. In our example (Fig. 3-0), Yo ranges between 0.6 and 0.8. Thus, since P1 is very close to 0, an appropriate range for P2 would be [.6-0, .8-0] = [.6, .8].

Parameter P3 is useful only to scale X so that a comparison can be made between data and the curves in Fig. 3-2b. Thus the value of P3 should be fixed (specify # of steps = 0) at 1/Xmax. In some cases, it may be useful to leave P3 as a parameter to be estimated, and then a range for Xmax should be obtained graphically. In our example, Xmax = 1, and therefore P3 = 1.

Finally, compare your data with Fig. 3-2b for curvature. Choose a realistic but sufficiently wide range of P4 values. In choosing the minimum value for P4, consider that P4 < -10 could rarely occur. In our example (Fig. 3-0), an appropriate range of values of P4 may be [-10, -4]. We suggest a large number of steps (15 to 25) for this parameter.

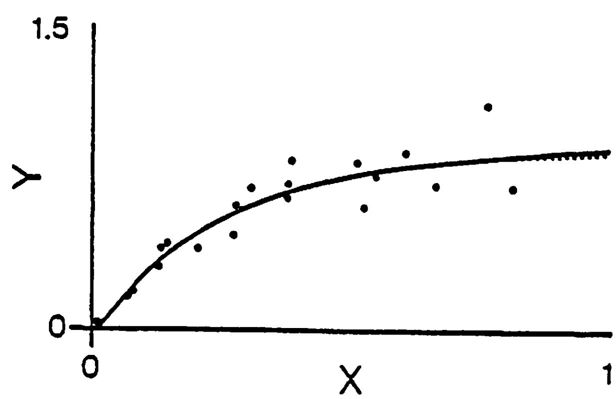
For our example, SEARCH found a minimum RSS of 0.09 (S = .908), and PAR gave an RSS of 0.08 (S = .921) with final (converged) values of P1 = -.08, P2 = .747, P3 = 1 (fixed), and P4 = -3.633.

TYPE 4

Type 4 equations are characterized by a gradual decrease in the slope of the X-Y relationship as X increases. The rate of increase of Y is very high at first, but gradually decreases to near zero, as Y approaches a maximum value (Ymax). This value is often called the upper ASYMPTOTE of Y, a value which is never quite reached by Y.

We propose 2 equations for this type of relationship. Equation [4-1] is preferred because it is simpler and more widely used in scientific literature (the typical diminishing return, or Poisson, function). Fitting equation [4-2] to data often requires that parameter P3 be fixed, so that convergence can be achieved. We suggest trying equation [4-1] first, and if a satisfactory fit is not obtained, then trying equation [4-2]. An example of data which is suitably described by a Type 4 equation is illustrated in Fig. 4-0. The two regression lines were obtained by fitting both equations to the data.

Fig. 4-0



TYPE 4 - EQUATION 1 (The Poisson Function)

$$Y = P1 + P2 \left(1 - e^{-P3\, X}\right) \quad \text{where } X \geq 0 \qquad [4\text{-}1]$$

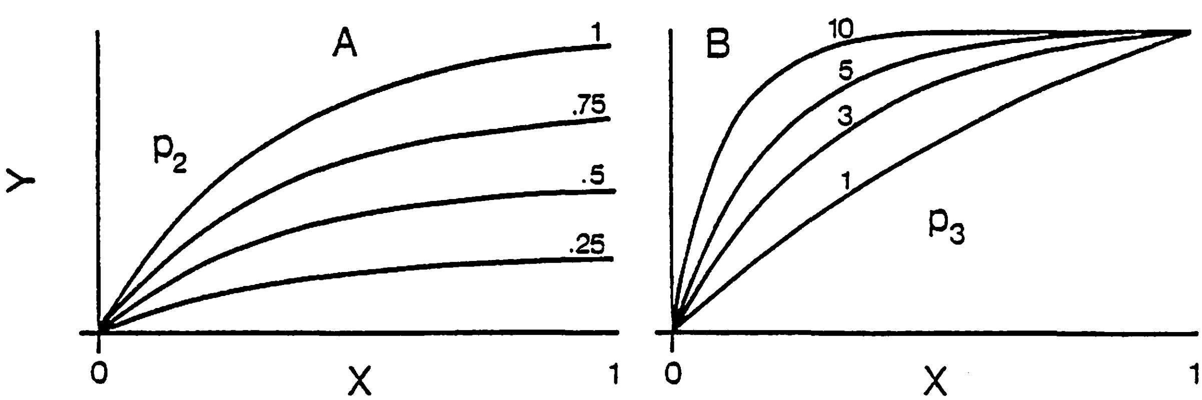
1. Parameter Limits
   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $0 < P3 < \infty$
2. Role of the Parameters

   P1: The intercept Yo (value of Y when X = 0).

   P2: Determines, along with P1, the upper asymptote (Fig. 4-1a). We note that Y = P1 + P2 when X-> ∞.

   P3: Changes the APPARENT curvature of the relationship (Fig. 4-1b), by scaling the values of X. Determines the portion of a Poisson curve which is seen over the range of values of X. When P3 = 0, there is no relationship (Y = P1). When P3 is large, the curvature is strong. With P3 close to zero, the curvature is weak.

Fig. 4-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [4-1] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. If an intercept can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 4-0), the intercept is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

Second, estimate a range of values for the upper asymptote (Amin, Amax). A suitable range of values of P2 is obtained by substracting the higher limit of P1 (P1max) from Amin, and the lower limit of P1 (P1min) from Amax:

[P2min, P2max] = [Amin-P1max, Amax-P1min]

In our example (Fig. 4-0), the upper asymptote lies between .75 and 1.25. Since the intercept is very close to 0, an appropriate range for P2 would be [.75-0, 1.25-0] = [.75, 1.25].

Third, compare your data with Fig. 4-1b, and choose a tentative range of values of P3. We note that the maximum X value in Fig. 4-1b is 1. Thus, since P3 SCALES X, a valid range of values for P3 can be obtained by dividing the values obtained from Fig. 4-1b by the maximum X value in your data (Xmax). A large number of steps (15 to 25) is recommended. In our example (Fig. 4-0), we obtained a range of [1, 5] from Fig. 4-1b (since Xmax = 1 in this case, it was not necessary to alter this range.)

For our example, SEARCH found a minimum RSS of 0.20 (S = .856), and PAR gave an RSS of 0.197 (S = .858) with final (converged) values of P1 = -.041, P2 = .929, and P3 = 4.09.

TYPE 4 - EQUATION 2 (The Power Function)

$$Y = P1 + P2\,(P3\,X + 1)^{P4} \quad \text{where } X \geq 0 \qquad [4\text{-}2]$$

1. Parameter Limits
   $-\infty < P1 < \infty$
   $-\infty < P2 < 0$
   $0 < P3 < \infty$
   $-\infty < P4 < 0$
2. Role of the Parameters

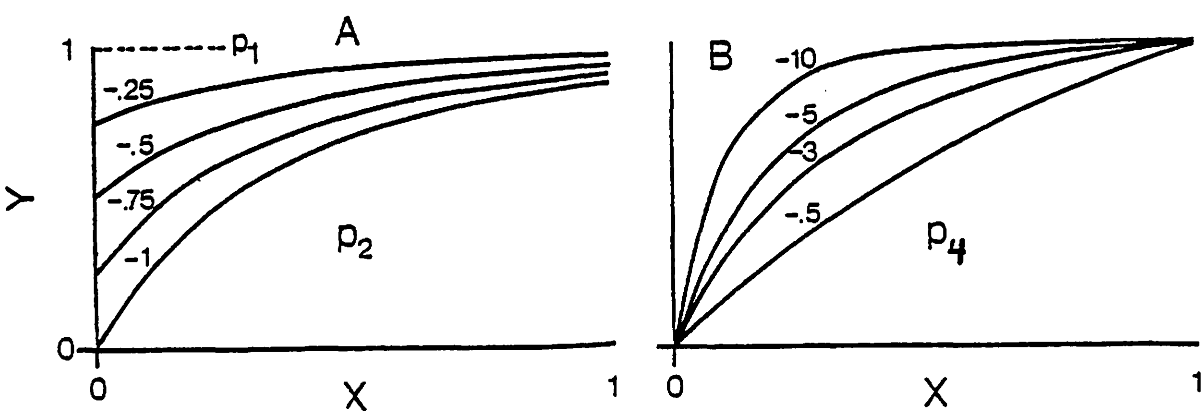
   P1: Upper asymptote (Y=Ymax when X->∞).

   P2: Changes the magnitude of Y (Fig. 4-2a). Therefore, determines, along with P1, the intercept (Yo when X = 0). We note that Yo = P1 + P2. Thus, P2 = Yo - P1. Rememoer that P2 < 0.

   P3: Scales the values of X. This parameter gives additional flexibility to the equation, and is necessary because of the "+ 1" term in equation 4-2.

   P4: Changes the curvature of the relationship (Fig. 4-2b). When P4 = 0, there is no relationship (horizontal line). When P4 is very small (i.e., negatively large), the curvature is strong. With P4 large (but still smaller than 0), the curvature is weak.

Fig. 4-2



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [4-2] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the high range of X. If an upper asymptote can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 4-0), the upper asymptote sould lie within the range [.7, 1.1].

Second, estimate a range of values for the intercept (Yo). A suitable range of values of P2 is given by:

[P2min, P2max] = [Yomin-P1max, Yomax-P1min]

We suggest a large number of steps (15 to 25) for this parameter. In our example (Fig. 4-0), Yo is very close to 0. Thus, since P1 ranges between .7 and 1.1, an appropriate range for P2 would be [0-.7, 0-1.1] = [-1.1, -.7]. We suggest a large number of iterations (15 to 25) for this parameter.

Parameter P3 is useful only to scale X so that a comparison can be made between data and the curves in Fig. 4-2b. Thus the value of P3 should be fixed (specify # of steps = 0) at 1/Xmax. In some cases, it may be useful to leave P3 as a parameter to be estimated, and then a range for Xmax should be obtained graphically. In our example, Xmax = 1, and therefore P3 = 1.

Finally, compare your data with Fig. 4-2b for curvature. Choose a realistic but sufficiently wide range of P4 values. In choosing the minimum value for P4, consider that P4 < -10 could rarely occur. In our example (Fig. 4-0), an appropriate range of values of P4 may be [-5, -1]. We suggest a large number of steps (15 to 25) for this parameter.

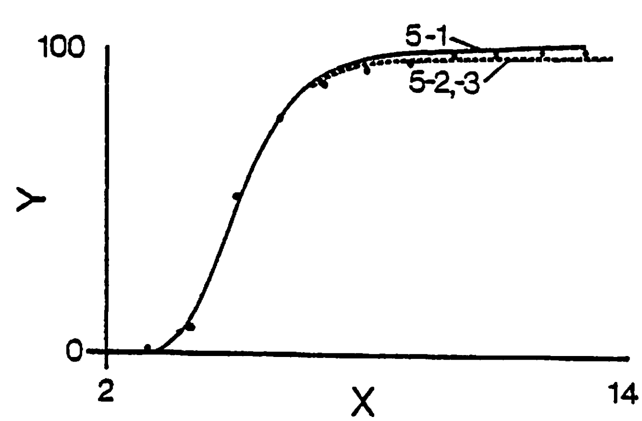
For our example, SEARCH found a minimum RSS of 0.199 (S = .857), and PAR gave an RSS of 0.198 (S = .857) with final (converged) values of P1 = .946, P2 =-1.0, P3 = 1 (fixed), and P4 = -4.202.

TYPE 5

Type 5 equations are the sigmoids. These relationships are characterized by the fact tnat Y is bounded by two asymptotes, upper and lower, between whicn it varies. We propose 3 different sigmoid functions. These 3 equations are not similar in complexity, and apply to slightly different sigmoid shapes or conditions. Equation [5-1] snould be applied to symetrical sigmoids or to relationships where the lower leg is much longer than the upper leg. Equation [5-2] should be applied to sigmoids where the lower leg is much shorter than the upper leg. Equation [5-3] applies to cases where X varies between two logically established limits (Xmin and Xmax) (e.g., 0 to 100%).

An example of data which is suitably described by a Type 5 equation is illustrated in Fig. 5-0. The three regression lines were obtained by fitting all three equations to the data.

Fig. 5-0



TYPE 5 - EQUATION 1 (The Logistic Equation)

$$Y = P1 + P2\,[1/(1 + e^{-P3(X-P4)})^{P5}] \quad \text{where } X \geq 0 \qquad [5\text{-}1]$$

1. Parameter Limits
   - $-\infty < P1 < \infty$
   - $-\infty < P2 < \infty$
   - $0 < P3 < \infty$
   - $-\infty < P4 < \infty$
   - $0 < P5 < \infty$
2. Role of the Parameters

   P1: The lower asymptote Yo (value of Y when X=0).

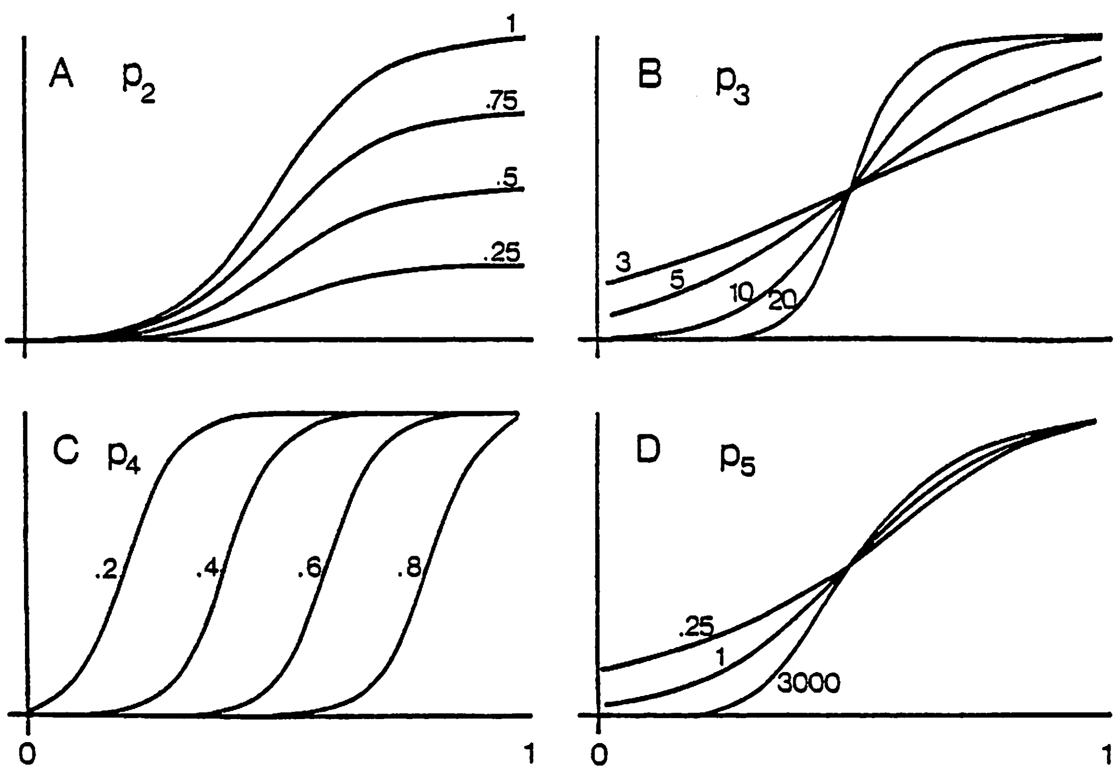
   P2: Determines, along with P1, the upper asymptote (Fig. 5-1a). We note that Y = P1 + P2 when X->∞ .

   P3: Changes the APPARENT curvature of the relationship (Fig. 5-1b), by scaling the values of X. When P3 is large, the curve occurs very rapidly, in a threshold-type of fashion. With P3 close to zero, the X-Y relationship is more gradual.

   P4: Displaces the curve along the X axis (Fig. 5-1c). When P5=1, P4 is the value of X when Y is at MIDPOINT between P1 and P2. Whenever P5 is not 1, this is no longer true.

   P5: Distorts the basic shape of the logistic equation (Fig. 5-1d). When 0 < P5 < 1, the lower leg is longer than the upper leg of the curve. When P5 >1, the upper leg is longer. Extreme skews of the latter type cannot be produced readily with equation [5-1], because they require very large values of P5.

Fig. 5-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [5-1] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. If a lower asymptote can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 5-0), the lower asymptote is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

Second, estimate a range of values for the upper asymptote (Amin, Amax). A suitable range of values of P2 is obtained by substracting the higher limit of P1 (P1max) from Amin, and the lower limit of P1 (P1min) from Amax:

[P2min, P2max] = [Amin-P1max, Amax-P1min]

In our example (Fig. 5-0), the upper asymptote is very close to 100. Since the lower asymptote is very close to 0, an appropriate range for P2 would be [100-0, 100-0] = [100, 100], with # of steps = 1.

Third, compare your data with Fig. 5-1b, and choose a tentative range of values of P3. We note that the maximum X value in Fig. 5-1b is 1. Thus, since P3 SCALES X, a valid range of values for P3 can be obtained by dividing the values obtained from Fig. 5-1b by the maximum X value in your data (Xmax). A large number of steps (15 to 25) is recommended. In our example (Fig. 5-0), we obtained a range of [10, 20] from Fig. 5-1b. Since Xmax = 13 in this case, the range [.7, 1.5] could be used.

You need not estimate a range of values for P4. This range depends on the value of P5. SEARCH has been programmed to provide this range itself, when given the range of values of X where Y is at about midpoint between P1 and P2. In our example, this range would be [4.5, 5.5]. We suggest a large number of steps (15 to 25) for this parameter.

Finally, a range of value of P5 can be obtained directly from Fig. 5-1d. Note that this range should be fairly wide, particularly when P5 >1. For our example, the range [1, 3000] would be necessary. Use a large number of steps (15 to 25) for this parameter.

For our example, SEARCH found a minimum RSS of 66.0 (S = .997), and PAR gave an RSS of 68.0 (S = .996) with final values of P1 = -.163, P2 = 100.3, P3 = 1.06, P4 = -2.834, and P5 = 2972.5. These values did not converge, however. This suggests that the amount of skew in our example was too large, and that equation [5-2] would be more suitable.

TYPE 5 - EQUATION 2 (The Poisson Function)

$$Y = P1 + P2\left(1 - e^{-P3\, X^{P4}}\right) \quad \text{where } X \geq 0 \qquad [5\text{-}2]$$

1. Parameter Limits
   $-\infty < P1 < \infty$
   $-\infty < P2 < \infty$
   $0 < P3 < \infty$
   $1 < P4 < \infty$
2. Role of the Parameters

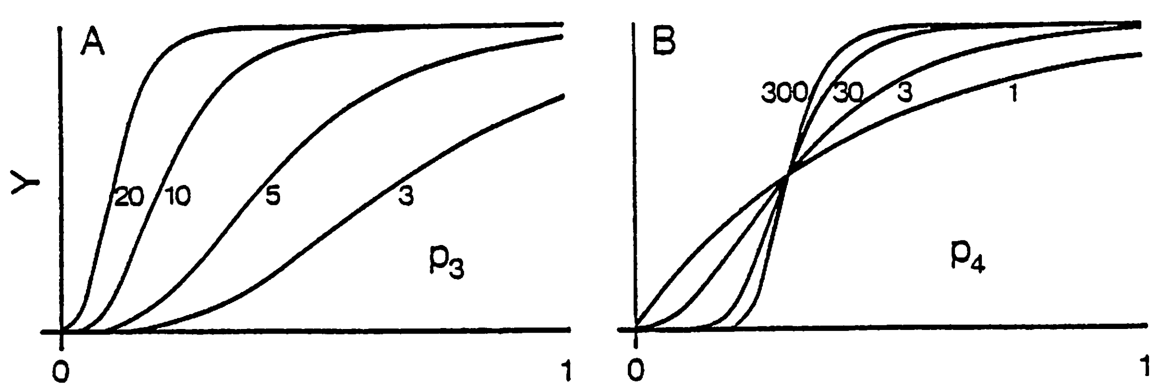
   P1: The lower asymptote Yo (value of Y when X=0).

   P2: Determines, along with P1, the upper asymptote. We note that Y = P1 + P2 when X->∞ .

   P3: Changes the APPARENT curvature of the relationship (Fig. 5-2a), by scaling the values of X. Determines the range of values of X over which the sigmoid curve occurs. The value of P3 is strongly affected by P4.

   P4: Distorts the basic shape of the Poisson equation (Fig. 5-2b). P4 works in the opposite manner as P5 in equation [5-1]. Here, the skew (lower leg always shorter than the upper) is most extreme when P4 = 1 (the typical Poisson). The amount of skew becomes smaller as P4 increases. At the limit, the curve is symetrical when P4 is infinite. Thus equation [5-2] should not be used to describe nearly symetrical sigmoids.

Fig. 5-2



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [5-2] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the low range of X. If a lower asymptote can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 5-0), the lower asymptote is obviously close to zero. We thus used the range [0, 0] (# of steps = 1).

Second, estimate a range of values for the upper asymptote (Amin, Amax). A suitable range of values of P2 is obtained by substracting the higher limit of P1 (P1max) from Amin, and the lower limit of P1 (P1min) from Amax:

[P2min, P2max] = [Amin-P1max, Amax-P1min]

In our example (Fig. 5-0), the upper asymptote is very close to 100. Since the lower asymptote is very close to 0, an appropriate range for P2 would be [100-0, 100-0] = [100, 100], with # of steps = 1.

SEARCH has been programmed to provide the range of values of P3. It will prompt for the range of values of X where Y is at about midpoint. In our example, this range would be [4.5, 5.5]. We suggest a large number of steps (15 to 25) for this parameter.

Finally, a range of value of P4 can be obtained directly from Fig. 5-2b. Note that this range should be fairly wide. For our example, the range [1, 300] would be needed. Use a large number of steps (15 to 25) for this parameter.

For our example, SEARCH found a minimum RSS of 65.0 (S = .997), and PAR gave an RSS of 40.2 (S = .998) with final (converged) values of P1 = -.206, P2 = 98.02, P3 = 1.14, and P4 = 205.1.

TYPE 5 - EQUATION 3

$$Y = P1\,(1 - Z)^{P2\,Z^{P3}} \qquad [5\text{-}3]$$

where $Z = (Xmax - X) / (Xmax - Xmin)$

(Xmin and Xmax user-supplied)

.sk 1
NOTE: THIS EQUATION SHOULD ONLY BE USED WHEN XMIN AND XMAX CAN BE OBJECTIVELY DEFINED.
.sk 1

1. Parameter Limits
   $-\infty < P1 < \infty$
   $1 < P2 < \infty$
   $1 < P3 < \infty$
2. Role of the Parameters

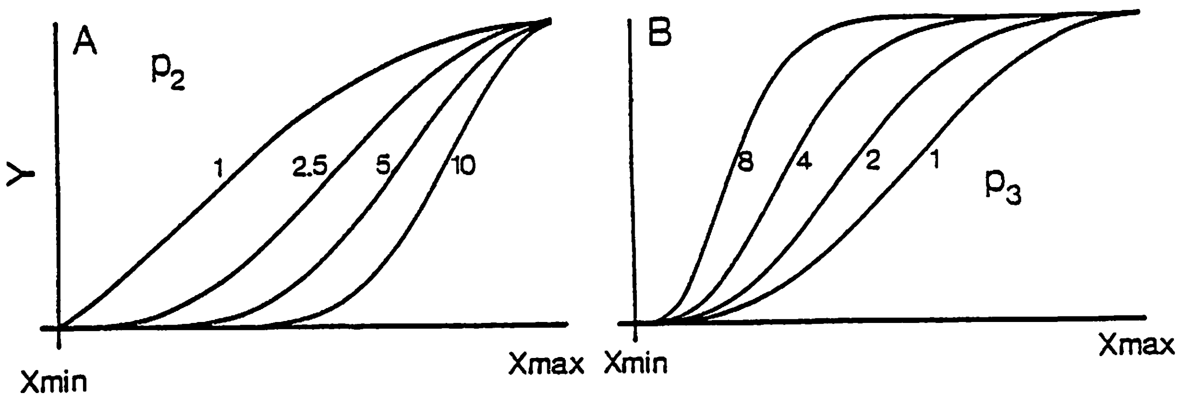
   P1: The upper asymptote Ymax (value of Y when $X \to \infty$).

   P2: Controls the steepness of the lower leg of the sigmoid (Fig. 5-3a). Also controls the position (relative to Xmin and Xmax) of the rise in Y. The higher P2, the later the rise.

   P3: Controls the steepness of the upper leg of the sigmoid (Fig. 5-3b). Also controls the position of the rise in Y. The higher P3, the earlier the rise.

   NOTE: P2 and P3 are in fact acting against each other, and when both are high, the curve is centrally located and very steep.

Fig. 5-3



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [5-3] to be used in conjunction with SEARCH can be obtained as follows.

First, obtain a range of values for P1. How wide a range can be determined from the amount of "noise" (error term) associated with Y in the high range of X. If an upper asymptote can be deduced from logical arguments, P1 can be fixed (specify # of steps = 0). In our example (Fig. 5-0), the upper asymptote is obviously close to 100. We thus used the range [100, 100] (# of steps = 1).

Second, a range of values of P2 and P3 can be obtained directly from Figs 5-3a and 5-3b. Because it is difficult to suggest how to choose them, we suggest that you use very wide ranges, with a large number of steps (15 to 25) for each parameter. In general, if the sigmoid is rather sharp, both P2 and P3 will be large, and vice-versa. For our example (Fig. 5-0), we used the ranges [1, 10] for both P2 and P3.

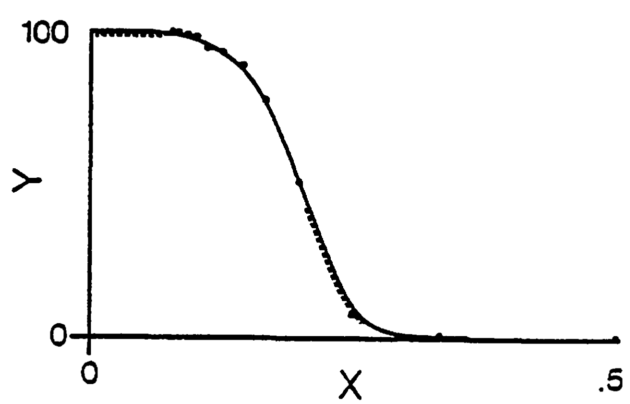
For our example, SEARCH found a minimum RSS of 106.0 (S = .994), and PAR gave an RSS of 62.9 (S = .997) with final (converged) values of P1 = 97.3, P2 = 5.233, P3 = 7.09.

TYPE 6

Type 6 relationships are the reversed sigmoids, where Y decreases from an upper limit at low X values, down to a lower asymptote at large X values. We suggest two equations for this type of relationship. Equation [6-1] is suitable for data where X can take any value. Equation [6-2] should be used only in cases where X has a limited scale (i.e., varies between strictly defined upper and lower limits, Xmin and Xmax).

In Fig. 6-0, we have illustrated an example of this type of relationship. The regression lines were obtained by fitting both equations to the data.

Fig. 6-0



TYPE 6 - EQUATION 1 (The Logistic Equation)

$$Y = P1 + P2\,[1/(1 + e^{-P3(X-P4)})^{P5}] \quad \text{where } X > 0 \quad [6-1]$$

1. Parameter Limits
   - $-\infty < P1 < \infty$
   - $-\infty < P2 < \infty$
   - $-\infty < P3 < 0$
   - $-\infty < P4 < \infty$
   - $0 < P5 < \infty$
2. Role of the Parameters

   We refer the reader to the discussion of the parameters of equation [5-1]. The only difference to be kept in mind, here, is that P3 is NEGATIVE. All statements applying to P3 remain valid as long as it is understood that a large P3 in equation [5-1] is equivalent to a small (negatively large) P3 in equation [6-1]. The effect of changing the sign of P3 here is to reverse the X axis.

3. Initial Parameter Estimation (see page 7, item c)

   Estimation of ranges for parameters of equation [6-1] is identical to that discussed in equation [5-1]. Keep in mind (1) that P1, the lower asymptote, now occurs at large values of X, and (2) that P3 is NEGATIVE. If you choose a range of, say, [3, 20] from Fig. 5-1b, you must change the sign to [-20, -3] before dividing by Xmax.

   For our example, SEARCH found a minimum RSS of 20.2 (S = .9989), and PAR gave an RSS of 6.6 (S = .9997) with final values of P1 = .492, P2 = 101.9, P3 = -26.6, P4 = .506, and P5 = 2248.4. The values did not converge. However, in view of the extremely high coefficient of determination, this is a minor setback.

TYPE 6 - EQUATION 2

$$Y = P1\,(1 - Z)^{P2\,Z^{P3}} \qquad [6\text{-}2]$$

where $Z = (X - Xmin) / (Xmax - Xmin)$

(Xmin and Xmax user-supplied)

NOTE: THIS EQUATION SHOULD ONLY BE USED WHEN XMIN AND XMAX CAN BE OBJECTIVELY DEFINED.

1. Parameter Limits
   $-\infty < P1 < \infty$
   $1 < P2 < \infty$
   $1 < P3 < \infty$
2. Role of the Parameters

   P1: The upper asymptote Ymax (value of Y when X->).

   We refer the reader to the discussion of the parameters of equation [5-3]. Note that the X axis is reversed in this case, by the transformation Z.

3. Initial Parameter Estimation (see page 7, item c)

   The reader is referred to the discussion of equation [5-3]. The parameter estimation process is the same in both cases.

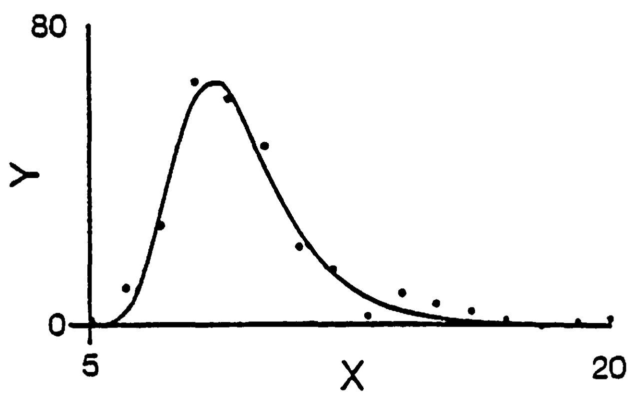
   For our example (Fig. 6-0), SEARCH found a minimum RSS of 233.7 (S = .987), and PAR gave an RSS of 8.0 (S = .9996) with final (converged) values of P1 = 99.4, P2 = 64.7, and P3 = 4.2.

TYPE 7

Type 7 relationships are characterized by an increase and then decrease of Y over the range of X ("bell-shaped" relationships), where the decreasing section is equal to or longer than the increasing section. We propose only one, very flexible equation for this type of relationship.

An example of data, and resulting regression line, are illustrated in Fig. 7-0.

Fig. 7-0



TYPE 7 - EQUATION 1

$$Y = P1 + [P2\ (X+1)]^{-P3}\ (1 - e^{-P4\ X})^{P5} \qquad [7\text{-}1]$$

where $X \geq 0$

1. Parameter Limits
   $-\infty < P1 < \infty$
   $0 < P2 < \infty$
   $1 < P3 < \infty$
   $0 < P4 < \infty$
   $1 < P5 < \infty$
2. Role of the Parameters

   P1: The lower limit Ymin (value of Y at X=0 or X->∞).

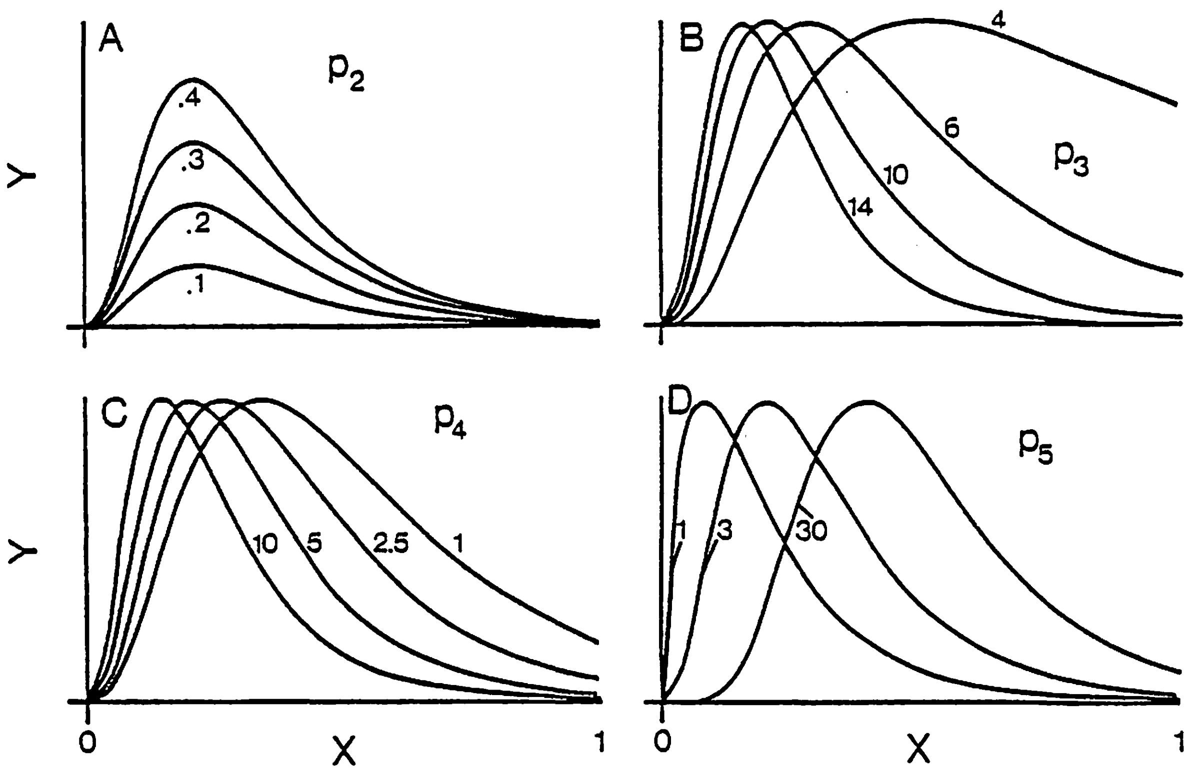
   P2: Scales the values of Y. Changes the size of the axes. Because of its position in the equation, however, the smaller P2 the larger Y (Fig. 7-1a).

   P3: Controls the rate of DECREASE of Y at larger X values (Fig. 7-1b). (It is the same as P4 in equation [3-1]). A large P3 makes Y decrease faster.

   P4: Scales X. Changes the rate of increase of Y (Fig. 7-1c). A large P4 makes Y increase more rapidly with X.

   P5: Changes the shape of the early portion of the increase of Y at lower values of X (Fig. 7-1d). This is the same as P4 in equation [5-2].

Fig. 7-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [7-1] to be used in conjunction with SEARCH can be obtained as follows.

First, estimate P1 visually. In most cases, P1 is zero, and this parameter can be fixed (specify # of steps = 0). In our example, this is the case.

P2 is very difficult to estimate, because it depends on the values of P3, P4, and P5. SEARCH has been programmed to compute P2 for you. Simply provide SEARCH with the maximum Y value, and the approximate value of X where this maximum occurs. In our example (Fig. 7-0), the maximum Y=65 occurs when X=8.5.

The range of P3 should be made as wide as possible. We suggest starting with [1, 20], with a large number of steps (10 to 20).

A tentative range of P4 is chosen from Fig. 7-1c. The, divide these values by Xmax. In our example (Fig. 7-0), we chose the range [1, 10], then divided by Xmax=20, to obtain [.05, 1]. A large number of steps (10 to 20) is recommended.

Parameter P5 is very difficult to estimate accurately. We suggest using a very wide range (e.g., [1, 1000]), with a large number of steps (10 to 20).

For our example, SEARCH found a minimum RSS of 192.6 (S = .973), and PAR gave an RSS of 174.2 (S = .976) with the final (converged) values P1 = 0 (fixed), P2 = .058, P3 = 9.96, P4 = .58, and P5 = 246.6.

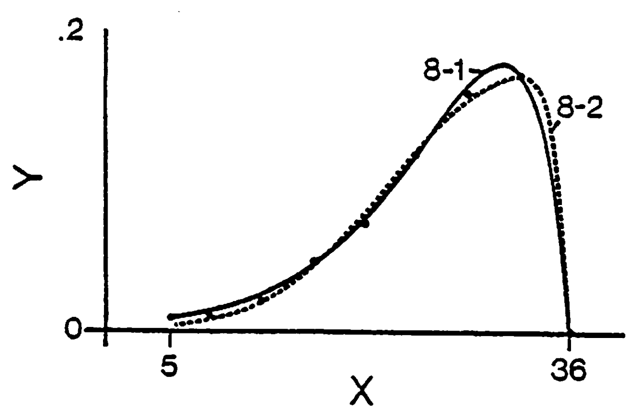
TYPE 8

Type 8 relationships are similar to type 7, in that Y rises and falls with increasing X. The difference between the two types is that the rise in Y here is more gradual than the collapse. Type 8 relationships are subject to THRESHOLD effects, particularly an UPPER threshold near which Y decays very rapidly. This form is frequently encountered in temperature-growth relationships. We suggest 2 equations for this type. Both are matched asymptote equations, developed by Logan et al. (Environ. Ent. 5: 1133-1140, 1976). To use these equations, there must be a basis to define a lower and an upper limit of X (Xmin and Xmax) below and above which Y=0.

Equation [8-1] is a combination of an exponential curve and a decay curve. Equation [8-2] is a combination of a logistic curve and a decay curve. Thus, the choice between them should be based on the "flatness" of the curve when y reaches its maximum value. A flat portion of curve there indicates that equation [8-2] should be used.

An example of data is illustrated in Fig. 8-0. The two regression lines were obtained by fitting both equations to the data.

Fig. 8-0



TYPE 8 - EQUATION 1

$$Y = P1 \{ e^{P2\,Z} - e^{[P2 - (1-Z)/P3]} \} \quad [8-1]$$

where $Z = (X - Xmin) / (Xmax - Xmin)$

(Xmin and Xmax user-supplied)

NOTE: THIS EQUATION SHOULD ONLY BE USED WHEN XMIN AND XMAX CAN BE OBJECTIVELY DEFINED.

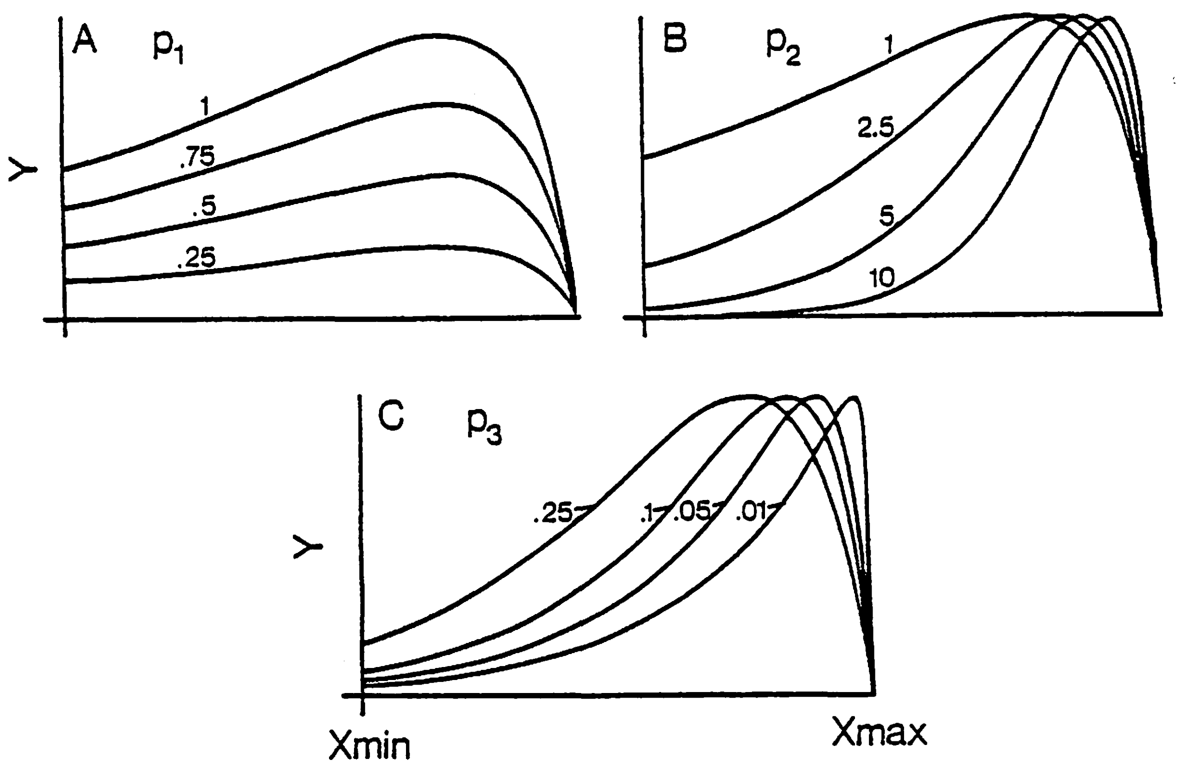
1. Parameter Limits
   - $0 < P1 < \infty$
   - $0 < P2 < \infty$
   - $0 < P3 < 1$
2. Role of the Parameters

   P1: The value of Y at X = Xmin (Fig. 8-1a).

   P2: Scales the transform Z, and determines the curvature of the curve in the lower range of X (exponential curve) (Fig. 8-1b).

   P3: Determines the width (relative to Z) of the "decay" portion of the curve (Fig. 8-1c).

Fig. 8-1



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [8-1] to be used in conjunction with SEARCH can be obtained as follows.

First, estimate a range of values for P1. This range depends on the amount of "noise" associated with Y in the low range of X. In our example, a range of [0, .02] could be used. We suggest a moderate number of steps for this parameter (i.e., 10).

Second, compare your data with Fig. 8-1b. Choose a range for P2. In our example, a range of [2.5, 7.5] would be suitable. Use a large number of steps for this parameter (15 to 25).

Finally, compare your data with Fig. 8-1c, and choose a range of values for P3. Note that P3 should rarely be smaller than 0.01, or larger than .5. In our example, the range [.01, .25] would be sufficient. Use a large number of steps (15 to 25) for this parameter.

For our example, SEARCH found a minimum RSS of .0001 (S = .9976), and PAR returned an RSS of .00007 (S = .998) with the values P1 = .0152, P2 = 5.23, and P3 = .16. These values did not converge. However, in view of the high coefficient of determination, this is not of much concern.

TYPE 8 - EQUATION 2

$$Y = P1 \{ 1/[1 + e^{(P2-P3\ Z)}] - e^{(Z-1)/P4} \} \qquad [8-2]$$

where $Z = (X - Xmin) / (Xmax - Xmin)$

(Xmin and Xmax user-supplied)

NOTE: THIS EQUATION SHOULD ONLY BE USED WHEN XMIN AND XMAX CAN BE OBJECTIVELY DEFINED.

1. Parameter Limits
   - $0 < P1 < \infty$
   - $0 < P2 < \infty$
   - $0 < P3 < \infty$
   - $0 < P4 < 1$
2. Role of the Parameters

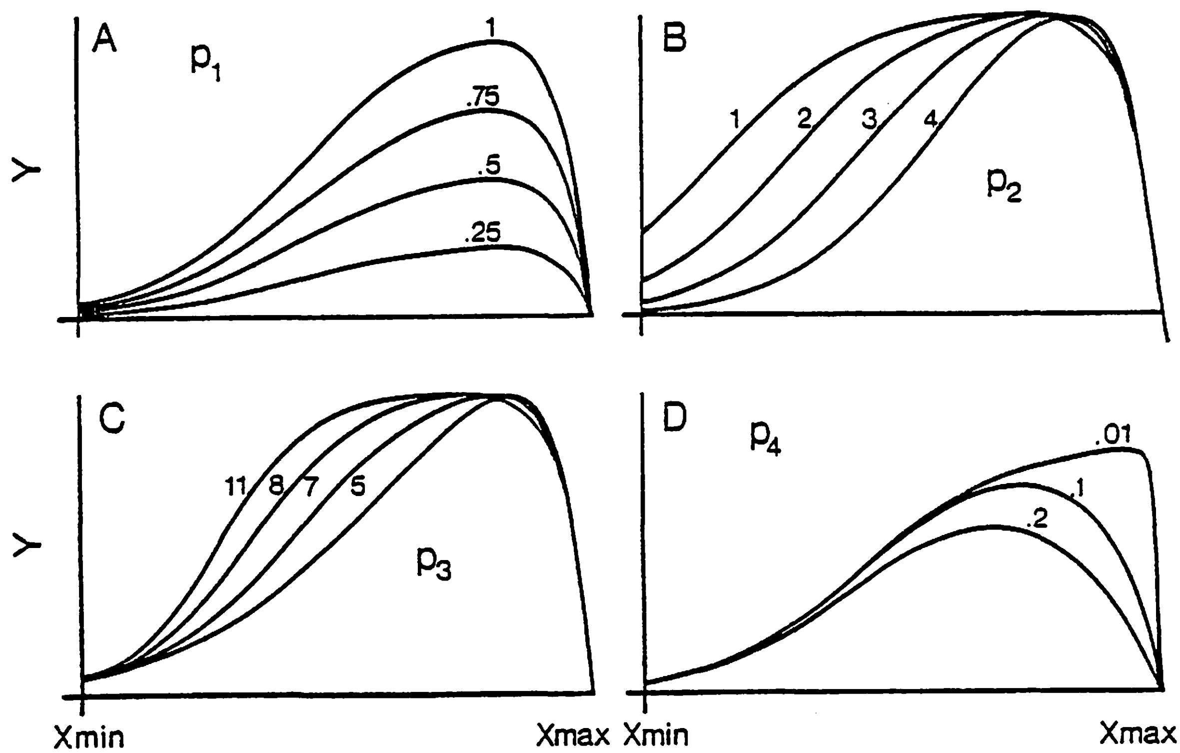
   P1: The "maximum" value of Y (Fig. 8-2a). This value, however, is only approximate, as it is affected by the value of P4.

   P2: Determines the initial value of Y at X = Xmin. (Fig. 8-2b).

   P3: Determines the rate at which the maximum Y is reached, and thus the "flatness" of the curve (Fig. 8-2c).

   P4: Determines the width (relative to Z) of the "decay" portion of the curve (Fig. 8-2d).

Fig. 8-2



3. Initial Parameter Estimation (see page 7, item c)

The range of values of the parameters of equation [8-2] to be used in conjunction with SEARCH can be obtained as follows.

First, estimate a range of values for P1. This range depends on the amount of "noise" associated with Y around its maximum. The value of P1 is affected by P4. When P4 is very small, P1 sould be very close to this actual maximum Y. But when P4 is large (e.g. .15) P1 must be larger than the maximum Y. Thus, be generous with the upper limit of the range of P1. In our example, a range of [.16, .21] could be used. We suggest a moderate number of steps for this parameter (i.e., 10).

Second, compare your data with Fig. 8-2b. Choose a range for P2. In our example, a range of [3, 10] would be suitable. Use a large number of steps for this parameter (10 to 20).

Third, compare your data with Fig. 8-2c, and determine a range of values for P3. In our example, a range of [5, 15] would be suitable. We suggest a large number of steps (10 to 20) for this parameter.

Finally, compare your data with Fig. 8-2d, and choose a range of values for P4. Note that P4 should rarely be smaller than 0.01, or larger than .5. In our example, the range [.01, .15] would be sufficient. Use a large number of steps (10 to 20) for this parameter.

For our example, SEARCH found a minimum RSS of .0002 (S = .995), and PAR returned an RSS of .00016 (S = .995) with the final (converged) values P1 = .194, P2 = 3.69, P3 = 6.89, and P4 = .035.